Primal and dual mixed-integer models for Global Navigation Satellite Systems

Massarweh, L.

DOI
10.4233/uuid:54c03c6e-5a2e-448b-9f6a-869d673cd4dc

Publication date
2025

Document Version
Final published version

Citation (APA)
Massarweh, L. (2025). *Primal and dual mixed-integer models for Global Navigation Satellite Systems*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:54c03c6e-5a2e-448b-9f6a-869d673cd4dc

Primal and dual mixed-integer models for Global Navigation Satellite Systems

Primal and dual mixed-integer models for Global Navigation Satellite Systems

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus, Prof.dr.ir. T.H.J.J. van der Hagen,
chair of the Board for Doctorates
to be defended publicly on
Wednesday 29, January 2025 at 17:30 o'clock

by

Lotfi MASSARWEH

Master of Science in Aerospace Engineering
Delft University of Technology, the Netherlands
born in Matera, Italy

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus,	chairperson
Prof.dr.ir. P.J.G. Teunissen	Delft University of Technology, promotor
Dr.ir. A.A. Verhagen	Delft University of Technology, copromotor
Independent members:	
Prof.dr.ir P.N.A.M. Visser	Delft University of Technology
Prof.dr. B. Zhang	Chinese Academy of Sciences (CAS-APM)
Dr.ir. F. Lopez Dekker	Delft University of Technology
Dr. Z. Altamimi	Institut de Physique du Globe de Paris (IPGP)
Dr. G. Giorgi	German Aerospace Center (DLR)
Prof.dr.ir. S. Steele-Dunne	Delft University of Technology, reserve member

This research was initially funded by the EU's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Actions grant agreement No. 722023.

Keywords: Global Navigation Satellite System (GNSS), Integer Ambiguity Resolution, Mixed-Integer Least-Squares, Primal-Dual formulations, Least-squares AMBiguity Decorrelation Adjustment (LAMBDA).

An electronic version of this dissertation is available at http://repository.tudelft.nl/.

Considerate la vostra semenza:
fatti non foste a viver come bruti,
ma per seguir virtute e canoscenza.

Dante Alighieri, *La Divina Commedia*,
Inferno, Canto XXVI, versi 118–120.

Consider well your origins:
you were not made to live like brutes,
but to pursue virtue and knowledge.

Dante Alighieri, *The Divine Comedy*,
Inferno, Canto XXVI, lines 118–120.

Contents

List of Figures

List of Tables

Summary

Mixed-integer models arise in several geodetic problems, including precise positioning and remote sensing in Global Navigation Satellite Systems (GNSS), as well as deformation monitoring through Interferometric Synthetic Aperture Radar (InSAR) or fringe phase observations from Very Long Baseline Interferometry (VLBI). These problems generally involve two types of unknowns: integer ambiguities $a \in \mathbb{Z}^n$ and real-valued parameters $b \in \mathbb{R}^p$, whose accuracy can be significantly improved by correctly resolving the ambiguities. However, in some cases, a large number of ambiguity components are involved and need to be correctly resolved, therefore the ambiguity resolution process becomes a bottleneck for the computations.

One research question is therefore how to effectively tackle the challenges of high-dimensional ambiguity resolution and its computational complexity, while ensuring a successful resolution of the ambiguities. Additionally, a second question arises regarding whether it is possible to solve this challenging problem in the domain of real-valued parameters, e.g. positioning coordinates, given that those are usually the parameters of interest for the user. In response to such questions, this doctoral dissertation is structured in two main parts: the first one looks at the integer ambiguity domain, presenting new flexible estimators and algorithms; the second one focuses on the real-valued parameter domain where the integerness of ambiguities is still taken into account.

- In **Part I**, relative to the primal formulation for mixed-integer least-squares problems, we extend the principle of integer bootstrapping to the vectorial case, thus introducing the Vectorial Integer Bootstrapping (VIB) estimators. These VIB solutions are flexible since tailored estimators can be designed based on a user-chosen partitioning. A comprehensive statistical characterization of these integer estimators is provided, including properties such as probability bounds and success-rate approximations.

  At this point, by considering a global network estimation for orbit determination and time synchronization of multi-GNSS constellations, we evaluate VIB solutions with real world data. Hence, we demonstrate how it is possible to successfully fix thousands of ambiguity components in a fraction of a second, thus leading to centimeter-level orbit solutions based on a 14-station global network of well-distributed ground receivers tracking GPS and Galileo satellites. Results are then validated showing midnight orbital discontinuity errors at a level of few centimeters over a 3-month period in 2019.

  At last, a new Least-squares AMBiguity Decorrelation Adjustment (LAMBDA) v4.0 toolbox has been developed, now merging both *estimation* and *evaluation* capabilities, and can be considered an expansion and significant update of the

legacy LAMBDA 3.0 (2012) and Ps-LAMBDA 1.0 (2013) software. This free and open source toolbox aims to become a valuable resource for researchers and/or practitioners dealing with mixed-integer models in high dimensions. It introduces a new integer search algorithm and different (classes of) estimators, providing users with a larger set of tools and functions suitable for various applications.

- In **Part II**, we introduce a dual formulation, where the direct computation of real-valued parameters is performed, and an implicit integer least-squares problem is solved. For these dual mixed-integer least-squares problems, two methods of simplifying approximations are introduced, and their quality is described by a complete distributional analysis of their estimators. Due to the lack of convexity for dual functions, a globally convergent solution, which has finite termination with a guaranteed ϵ-tolerance, is proposed and constructed by combining a branch-and-bound principle with a special convex relaxation.

  The performance of dual solutions are therefore evaluated in the context of GNSS positioning, where a Branch-and-bound INteger-equivariant Global Optimizer (BINGO) method is described, following a normalization of dual problems, and each computational step is investigated. In the special case $p = 1$, an alternative P1 algorithm is introduced, which outperforms LAMBDA-based solutions and we demonstrate that its complexity (i.e. number of enumerated solutions) grows linearly with the ambiguity problem dimensionality. In the context of GNSS, where carrier-phase data is extremely precise, the computationally convenient dual approximations are shown to be quasi-optimal solutions to the original problem. Therefore, these approximations are statistically evaluated by studying their relation to their primal counterpart problems.

  Lastly, the Ambiguity Function (AF) method, developed in the 1980s for GPS baseline positioning, is studied and some new theory is introduced to provide the probability model by means of which the AF estimator becomes a maximum likelihood estimator. A similar branch-and-bound globally convergent approach is used to ensure optimal solutions, and their nonuniqueness is proved and discussed. At the end, a systematic comparison between the AF and the ILS principle is presented, showing how the two methods are fundamentally different, nonetheless can behave similarly under certain circumstances.

In summary, through the mathematical advancements, enhanced algorithms and estimators introduced in the new LAMBDA 4.0 toolbox, this dissertation provides theoretical and practical tools for the GNSS and non-GNSS communities tackling mixed-integer least-squares problems. These contributions are fundamental for addressing the high-dimensional challenges anticipated in future signals and systems, particularly in larger global networks with stricter timeliness requirements. Furthermore, the introduction of a dual formulation provides a new landscape of advanced solutions, which can also be extended to 'constrained' dual problems where additional constraints are applied to real-valued parameters. This will be an important subject of study, especially in safety-critical applications.

Samenvatting (in Dutch)

Mixed-integer modellen komen voor in verschillende geodetische problemen, waaronder nauwkeurige positiebepaling en remote sensing met Global Navigation Satellite Systems (GNSS), evenals deformatiemonitoring met behulp van Interferometric Synthetic Aperture Radar (InSAR) of fase-waarnemingen van Very Long Baseline Interferometry (VLBI). Deze problemen bevatten over het algemeen twee typen onbekenden: geheeltallige meerduidigheden $a \in \mathbb{Z}^n$ en reële parameters $b \in \mathbb{R}^p$, waarbij de nauwkeurigheid van de geschatte parameters aanzienlijk kan worden verbeterd door de geheeltallige meerduidigheden correct op te lossen. In sommige gevallen zijn echter veel meerduidigheidscomponenten betrokken die correct moeten worden opgelost, waardoor het proces van meerduidigheidsoplossing een knelpunt wordt in de berekeningen.

Een belangrijke onderzoeksvraag is daarom hoe de uitdagingen van hoogdimensionale meerduidigheidsoplossing en de bijbehorende rekencomplexiteit effectief kunnen worden aangepakt, terwijl tegelijkertijd een succesvolle oplossing van de meerduidigheden wordt gegarandeerd. Daarnaast rijst de vraag of het mogelijk is om dit uitdagende probleem op te lossen in het domein van de reële parameters, bijvoorbeeld de cooordinaten in geval van plaatsbepaling, aangezien dit meestal de parameters van belang zijn voor de gebruiker. Als antwoord op deze vragen is dit proefschrift gestructureerd in twee hoofdonderdelen: het eerste onderdeel richt zich op het domein van gehele meerduidigheden, waarbij nieuwe flexibele schatters en algoritmen worden gepresenteerd; het tweede onderdeel richt zich op het domein van reële parameters waarbij de geheeltalligheid van de meerduidigheden nog steeds in aanmerking wordt genomen.

- In **Deel I**, dat betrekking heeft op een primale formulering voor mixed-integer kleinste kwadraten problemen, breiden we het principe van geheltallige bootstrapping uit naar het vectoriële geval, waarmee we de Vectorial Integer Bootstrapping (VIB) schatters introduceren. Deze VIB-oplossingen zijn flexibel omdat op maat gemaakte schatters kunnen worden ontworpen op basis van een door de gebruiker gekozen opdeling. Een uitgebreide statistische karakterisering van deze geheeltallige schatters wordt verstrekt, inclusief eigenschappen zoals waarschijnlijkheidsgrenzen en benaderingen van het succespercentage.

  Vervolgens evalueren we VIB-oplossingen met echte data op basis van een globale netwerkschatting voor de baanbepaling en tijdsynchronisatie van multi-GNSS-constellaties. Hiermee laten we zien hoe het mogelijk is om duizenden meerduidigheden binnen een fractie van een seconde succesvol op te lossen, wat leidt tot baanoplossingen op centimeter-niveau op basis van een wereldwijd netwerk van 14 goed verdeelde grondontvangers die GPS- en Galileo-satellieten volgen. De resultaten worden vervolgens gevalideerd door het aantonen van

discontinuiiteitsfouten van enkele centimeters die om middernacht optreden, over een periode van drie maanden in 2019.

Ten slotte is er een nieuwe Least-squares AMBiguity Decorrelation Adjustment (LAMBDA) v4.0 toolbox ontwikkeld, die zowel *schattings-* als *evaluatiecapaciteiten* samenvoegt, en kan worden beschouwd als een uitbreiding en belangrijke update van de legacy LAMBDA 3.0 (2012) en Ps-LAMBDA 1.0 (2013) software. Deze gratis en open-source toolbox is bedoeld als waardevolle bron voor onderzoekers en/of praktijkmensen die te maken hebben met mixed-integer modellen in hoge dimensies. Het introduceert een nieuw geheel zoekalgoritme en verschillende (klassen van) schatters, waarmee gebruikers een grotere set tools en functies krijgen die geschikt zijn voor verschillende toepassingen.

- In **Deel II** introduceren we een duale formulering, waarbij de directe berekening van reële parameters wordt uitgevoerd en een impliciet geheeltallig kleinste kwadraten (integer least-squares, ILS) probleem wordt opgelost. Voor deze duale mixed-integer kleinste kwadraten problemen worden twee methoden voor vereenvoudigde benaderingen geïntroduceerd, en de kwaliteit wordt beschreven door een volledige analyse van de kansverdelingen van de schatters. Vanwege het gebrek aan convexiteit voor duale functies, wordt een globaal convergerende oplossing, met een eindige beëindiging en gegarandeerde ϵ-tolerantie, voorgesteld en geconstrueerd door een branch-and-bound-principe te combineren met een speciale convexe relaxatie.

  De prestaties van duale oplossingen worden vervolgens geëvalueerd in de context van GNSS plaatsbepaling, waarbij een Branch-and-bound INteger-equivariant Global Optimizer (BINGO) methode wordt beschreven, gebaseerd op een normalisatie van duale problemen, en waarbij elke rekenslag wordt onderzocht. In het speciale geval $p = 1$ wordt een alternatief P1-algoritme geïntroduceerd, dat beter presteert dan op LAMBDA gebaseerde oplossingen, en we tonen aan dat de complexiteit ervan (dat wil zeggen het aantal gesommeerde oplossingen) lineair groeit met de dimensie van het meerduidigheidsprobleem. In de context van GNSS, waar fasedata van de draaggolven uiterst nauwkeurig gemeten wordt, wordt de quasi-optimaliteit van de berekeningstechnisch handige duale benaderingen aangetoond. Vervolgens worden deze benaderingen statistisch geëvalueerd door de relatie met hun primale tegenhangers te bestuderen.

  Ten slotte wordt de Ambiguity Function (AF) methode, ontwikkeld in de jaren ‘80 voor relatieve GPS plaatsbepaling, bestudeerd en wordt een nieuwe theorie geïntroduceerd om het waarschijnlijkheidsmodel te bepalen waarmee de AF-schatter een maximum likelihood schatter wordt. Een vergelijkbare branch-and-bound benadering die globaal convergeert wordt gebruikt om optimale oplossingen te garanderen, en de niet-uniciteit wordt bewezen en besproken. Tot slot wordt een systematische vergelijking gepresenteerd tussen de AF- en de ILS-principes, waarbij wordt aangetoond hoe de twee methoden fundamenteel verschillend zijn, maar onder bepaalde omstandigheden vergelijkbaar kunnen optreden.

Samenvattend biedt dit proefschrift, door de wiskundige vooruitgangen, verbeterde algoritmen en schatters geïntroduceerd in de nieuwe LAMBDA 4.0-toolbox, theoretische en praktische tools voor de GNSS- en niet-GNSS-gemeenschappen die worstelen met mixed-integer least-squares problemen. Deze bijdragen zijn essentieel voor het aanpakken van de hoog-dimensionale uitdagingen die worden verwacht bij toekomstige signalen en systemen, met name in grotere wereldwijde netwerken met strengere tijdigheidseisen. Bovendien opent de introductie van een duale formulering een nieuw landschap van geavanceerde oplossingen, die ook kunnen worden uitgebreid naar 'geconstrueerde' duale problemen, waarbij extra beperkingen worden toegepast op reële parameters. Dit zal een belangrijk studieonderwerp zijn, vooral in veiligheidskritieke toepassingen.

1

Introduction

1.1 Background

In several geodetic applications, we often deal with a mixture of unknown parameters having integer and real values. The mixed-integer models are therefore of great interest for the geodetic community since they can be found in many different fields, such as Global Navigation Satellite Systems (GNSS), Very Long Baseline Interferometry (VLBI), Interferometric Synthetic Aperture Radar (InSAR), as well as other specific applications, e.g. acoustic waves for underwater navigation (Viegas and Cunha, 2007), or detection in Multiple-Input Multiple-Output (MIMO) communication systems (Damen et al., 2003).

In 1993, mixed-integer least-squares problems were examined by Teunissen (1993) and a primal orthogonal decomposition was introduced, which created the foundations for the Least-squares AMBiguity Decorrelation Adjustment (LAMBDA) method (Teunissen, 1995), in use since then. Although initially defined in the context of Global Positioning System (GPS) double-difference carrier-phase data, i.e. ultra precise but biased by an integer number of cycles, the method quickly gained popularity in different fields of geodesy, in particular the ones dealing with mixed-integer least-squares problems.

In the context of GNSS, these models are highly attractive, since after resolving for the carrier-phase integer ambiguities, it is possible to make use of the millimiter-level precision of phase data. Therefore, this starts acting as ultra-precise pseudo-range measurement, which can be used for precise positioning (Teunissen, 2020), time and frequency transfer (Mi et al., 2023), atmosphere remote sensing (Lu et al., 2018), and carrier-phase attitude determination (Giorgi, 2011). All these applications rely on the successful resolution of carrier-phase ambiguities; however, in some cases, the computational burden represents a major bottleneck to the process of Integer Ambiguity Resolution (IAR).

The high dimensionality issue is further exacerbated in the case of network processing, where multiple GNSS receivers might track up to hundreds of satellites; see De Jonge (1998), Ge et al. (2006), and Brack (2019). With the availability of multi-frequency systems and the deployment of additional constellations for Positioning, Navigation, and Timing (PNT), including the ones foreseen in low Earth orbit (Wang et al., 2024; Xu et al., 2024), new solutions for tackling High-Dimensional Ambiguity Resolution (HDAR) problems are required nowadays.

Some simple strategies were available in the last decades, such as well-known widelaning approaches (Blewitt, 1989; Cocard and Geiger, 1992), however limited to specific combinations of signals, or in other cases applicable only to differenced observations. For dense network processing, LAMBDA methods have already been shown to be highly efficient for problems with up to hundreds of ambiguity components; see Li and Teunissen (2011). With LAMBDA, a landscape of different solutions was introduced, along with different classes of estimators (Teunissen, 2003), thus supporting both the GNSS and non-GNSS communities dealing with mixed-integer problems.

However, today larger networks of globally distributed receivers are being processed, and IAR problems can easily involve thousands of components (Strasser, 2022). New solutions to high-dimensional problems are therefore required and with primary attention to real-valued parameters, given that those are usually the parameters of interest for the user. These important aspects will be discussed in this contribution, so providing GNSS examples, but starting here with a more general definition of mixed-integer models.

1.2 Brief review of mixed-integer models

Given $y \in \mathbb{R}^m$ as the vector of observables, with its variance-covariance matrix as $Q_{yy} \in \mathbb{R}^{m\times m}$, we consider the linear mixed-integer model (Leick et al., 2015):

$$\mathrm{E}\{y\} = Aa + Bb, \quad \mathrm{D}\{y\} = Q_{yy} \tag{1.1}$$

where $\mathrm{E}\{\cdot\}$ and $\mathrm{D}\{\cdot\}$ refer respectively to the expectation and the dispersion operator, while the integer and real-valued parameters are respectively given as $a \in \mathbb{Z}^n$ and $b \in \mathbb{R}^p$, with the full-rank design matrix given by $[A, B] \in \mathbb{R}^{m\times(n+p)}$. We assume $y \sim \mathcal{N}_m(\mathrm{E}\{y\}, \mathrm{D}\{y\})$, a m-dimensional normally distributed vector, and we seek a least-squares solution following the orthogonal decomposition (Teunissen, 1998):

$$\|y - Aa - Bb\|^2_{Q_{yy}} = \|\hat{e}\|^2_{Q_{yy}} + \mathcal{F}(a, b) \tag{1.2}$$

where $||x||^2_Q = x^T Q^{-1} x$, and with $\hat{e} = P^{\perp}_{[A,B]} y = y - (A\hat{a} + B\hat{b})$ being the least-squares residual vector, where $\|\hat{e}\|^2_{Q_{yy}}$ is not a function of a and b. In the context of GNSS, we generally refer to *a*mbiguity (integer) and *b*aseline (real-valued) parameters, see Verhagen (2005) for additional information. We refer to Table 1.1 for an overview of the relevant quantities considered in this dissertation.

Table 1.1: Overview of relevant quantities considered for mixed-integer models

Quantity	**Integer parameters**	**Real-valued parameters**
Vector of unknowns	$a \in \mathbb{Z}^n$	$b \in \mathbb{R}^p$
Design matrix	$A \in \mathbb{R}^{m\times n}$	$B \in \mathbb{R}^{m\times p}$
Left inverse matrix	$A^+ = \left(A^T Q_{yy}^{-1} A\right)^{-1} A^T Q_{yy}^{-1}$	$B^+ = \left(B^T Q_{yy}^{-1} B\right)^{-1} B^T Q_{yy}^{-1}$
Projection matrix onto $\mathcal{R}(\cdot)$	$P_A = AA^+$	$P_B = BB^+$
Projection matrix onto $\mathcal{R}(\cdot)^{\perp}$	$P_A^{\perp} = I_m - P_A$	$P_B^{\perp} = I_m - P_B$
Auxiliary matrix	$\bar{A} = P_B^{\perp} A$	$\bar{B} = P_A^{\perp} B$
Covariance matrix	$Q_{\hat{a}\hat{b}} = -A^+ B Q_{\hat{b}\hat{b}}$	$Q_{\hat{b}\hat{a}} = -B^+ A Q_{\hat{a}\hat{a}}$
Variance matrix	$Q_{\hat{a}\hat{a}} = \left(\bar{A}^T Q_{yy}^{-1} \bar{A}\right)^{-1}$	$Q_{\hat{b}\hat{b}} = \left(\bar{B}^T Q_{yy}^{-1} \bar{B}\right)^{-1}$
Variance matrix (conditioned)	$Q_{\hat{a}(b)} = \left(A^T Q_{yy}^{-1} A\right)^{-1}$	$Q_{\hat{b}(a)} = \left(B^T Q_{yy}^{-1} B\right)^{-1}$
Float estimator	$\hat{a} = \bar{A}^+ y$	$\hat{b} = \bar{B}^+ y$
Float estimator (conditioned)	$\hat{a}(\beta) = \hat{a} - Q_{\hat{a}\hat{b}} Q_{\hat{b}\hat{b}}^{-1} \left(\hat{b} - \beta\right)$ $\equiv A^+(y - B\beta), \forall \beta \in \mathbb{R}^p$	$\hat{b}(\alpha) = \hat{b} - Q_{\hat{b}\hat{a}} Q_{\hat{a}\hat{a}}^{-1} (\hat{a} - \alpha)$ $\equiv B^+(y - A\alpha), \forall \alpha \in \mathbb{Z}^n$

*$\mathcal{R}(\cdot)$ refers to *range space*.

From Eq.(1.2), we can observe that the minimizers of $\|y - Aa - Bb\|^2_{Q_{yy}}$ are those of the objective function $\mathcal{F} : \mathbb{Z}^n \times \mathbb{R}^p \to \mathbb{R}$, expressed here as

$$\mathcal{F}(a,b) = \begin{cases} \mathcal{F}_1(a,b) \overset{\text{def}}{=\!=} \|\hat{a} - a\|^2_{Q_{\hat{a}\hat{a}}} + \left\|\hat{b}(a) - b\right\|^2_{Q_{\hat{b}(a)}} & \text{(primal)} \\ \mathcal{F}_2(a,b) \overset{\text{def}}{=\!=} \left\|\hat{b} - b\right\|^2_{Q_{\hat{b}\hat{b}}} + \|\hat{a}(b) - a\|^2_{Q_{\hat{a}(b)}} & \text{(dual)} \end{cases} \tag{1.3}$$

with the 'primal' and 'dual' formulations introduced respectively by Teunissen (1993) and by Teunissen and Massarweh (2024). We will now briefly summarize these two formulations.

1.2.1 Primal mixed-ILS formulation

In the primal mixed-ILS formulation we consider $\mathcal{F}_1(a,b)$, and a three-step approach (Teunissen, 1995) can be adopted, i.e.

1) **Float solution**, where integerness of ambiguities is discarded and the float estimators $\hat{a} \in \mathbb{R}^n$ and $\hat{b} \in \mathbb{R}^p$ are computed via standard least-squares;

2) **Ambiguity resolution**, where we obtain a fixed ambiguity estimator $\check{a} = \mathcal{I}_a(\hat{a})$ given the mapping function $\mathcal{I}_a : \mathbb{R}^n \to \mathbb{Z}^n$ for the ILS estimator;

3) **Fixed solution**, where real-valued parameters $\hat{b} \in \mathbb{R}^p$ are conditionally updated onto the fixed ambiguities, thus leading to $\check{b} = \hat{b}(\check{a})$.

Notice that no constraints exist on the real-valued parameters, and the ambiguity-fixed solution will be given by $\hat{b}(a)$ for $a \in \mathbb{Z}^n$. Thus, the primal objective function $\mathcal{P} : \mathbb{Z}^n \to \mathbb{R}$ is defined here as

$$\mathcal{P}(a) = \min_{b \in \mathbb{R}^p} \mathcal{F}_1(a,b) = \mathcal{F}_1\left(a, \hat{b}(a)\right) = \|\hat{a} - a\|^2_{Q_{\hat{a}\hat{a}}} \tag{1.4}$$

and the (primal) mixed-ILS solution for the ambiguities is given by

$$\check{a} = \arg\min_{a \in \mathbb{Z}^n} \mathcal{P}(a) = \arg\min_{a \in \mathbb{Z}^n} \|\hat{a} - a\|^2_{Q_{\hat{a}\hat{a}}} \tag{1.5}$$

following the least-squares (float) estimation of a and b. It is possible to efficiently resolve the integer ambiguities using the recent **LAMBDA 4.0 toolbox** implementation, see Massarweh et al. (2025), given the $\hat{a}$-vector and $Q_{\hat{a}\hat{a}}$-matrix as inputs.

1.2.2 Dual mixed-ILS formulation

In the dual mixed-ILS formulation we consider $\mathcal{F}_2(a,b)$, and the dual objective function $\mathcal{D} : \mathbb{R}^p \to \mathbb{R}$ can be defined as

$$\mathcal{D}(b) = \min_{a \in \mathbb{Z}^n} \mathcal{F}_2(a,b) = \mathcal{F}_2\left(\check{a}(b), b\right) = \left\|\hat{b} - b\right\|^2_{Q_{\hat{b}\hat{b}}} + \|\hat{a}(b) - \check{a}(b)\|^2_{Q_{\hat{a}(b)}} \tag{1.6}$$

and the (dual) mixed-ILS solution for the parameters is given by

$$\check{b} = \arg\min_{b\in\mathbb{R}^p} \mathcal{D}(b) = \arg\min_{b\in\mathbb{R}^p} \left(\left\| \hat{b} - b \right\|^2_{Q_{\hat{b}\hat{b}}} + \|\check{\epsilon}(b)\|^2_{Q_{\hat{a}(b)}} \right) \tag{1.7}$$

with the (conditioned) ambiguity residuals $\check{\epsilon}(b) = \hat{a}(b) - \check{a}(b)$, for

$$\check{a}(b) = \arg\min_{a\in\mathbb{Z}^n} \|\hat{a}(b) - a\|^2_{Q_{\hat{a}(b)}} \tag{1.8}$$

that resembles the expression in Eq.(1.5). However, in the dual formulation, we make use of $Q_{\hat{a}(b)} \leq Q_{\hat{a}\hat{a}}$ since $Q_{\hat{a}\hat{a}} - Q_{\hat{a}(b)} \equiv Q_{\hat{a}\hat{b}} Q^{-1}_{\hat{b}\hat{b}} Q_{\hat{b}\hat{a}} \geq 0$, and the distribution of the conditioned float ambiguity estimator is given by

$$\hat{a}(\beta) \sim \mathcal{N}_n \left(a + \Delta a(\beta), Q_{\hat{a}(b)} \right) \tag{1.9}$$

so it is normally distributed around $a + \Delta a(\beta)$, with the ambiguity bias

$$\Delta a(\beta) = Q_{\hat{a}\hat{b}} Q^{-1}_{\hat{b}\hat{b}} \left(\beta - \mathrm{E}\{\hat{b}\} \right), \quad \beta \in \mathbb{R}^p \tag{1.10}$$

and given that $\hat{b}$ is an unbiased estimator, then $\mathrm{E}\{\hat{b}\} = b$ (true vector). This dual formulation will be extensively discussed and analyzed in Chapter 5.

1.2.3 A primal-dual relationship

In the primal formulation, we can consider the concept of pull-in regions, i.e. subsets of $\mathbb{R}^n$ where float vectors are mapped to the corresponding integer. For integer estimators, these regions cover the entire space $\mathbb{R}^n$ without gaps, while being translational invariant over the integers. For instance, the rounding pull-in regions are hyper-cubes where all float solutions are mapped to the same integer via a simple rounding operator.

In the dual formulation, we notice that potential integer candidates are those whose pull-in region is crossed by the conditioning hyperplane, i.e. $\hat{a}(\beta) = \hat{a} - Q_{\hat{a}\hat{b}} Q^{-1}_{\hat{b}\hat{b}} (\hat{b} - \beta)$, which can be illustrated in $\mathbb{R}^n$ as shown with Figure 1.1 for $p = 2, n = 3$. On the left panel, the conditioned hyperplane is illustrated in the ambiguity domain given $||\hat{b} - \beta||_{Q_{\hat{b}\hat{b}}} \leq R_0$, hence bounded by a magenta ellipse also depicted in the parameter domain, see right panel. Therefore, different real-valued parameters might belong to different pull-in regions. These regions depend on the estimators in use, e.g. simple rounding for Figure 1.1, and have been highlighted with colors in the illustration.

Note that the hyperplane's orientation in the ambiguity domain is mainly driven by the covariance matrix $Q_{\hat{a}\hat{b}} \in \mathbb{R}^{n\times p}$, while an equivalence can be seen in Eq.(1.3), since the equality $\mathcal{F}_1(a, b) \equiv \mathcal{F}_2(a, b)$ implies that

$$\|\hat{a} - a\|^2_{Q_{\hat{a}\hat{a}}} + \left\| \hat{b}(a) - b \right\|^2_{Q_{\hat{b}(a)\hat{b}(a)}} \equiv \left\| \hat{b} - b \right\|^2_{Q_{\hat{b}\hat{b}}} + \|\hat{a}(b) - a\|^2_{Q_{\hat{a}(b)\hat{a}(b)}} \tag{1.11}$$

where the left and the right side focus on the minimization problem in the ambiguity domain and in the real-valued parameter domain, respectively.

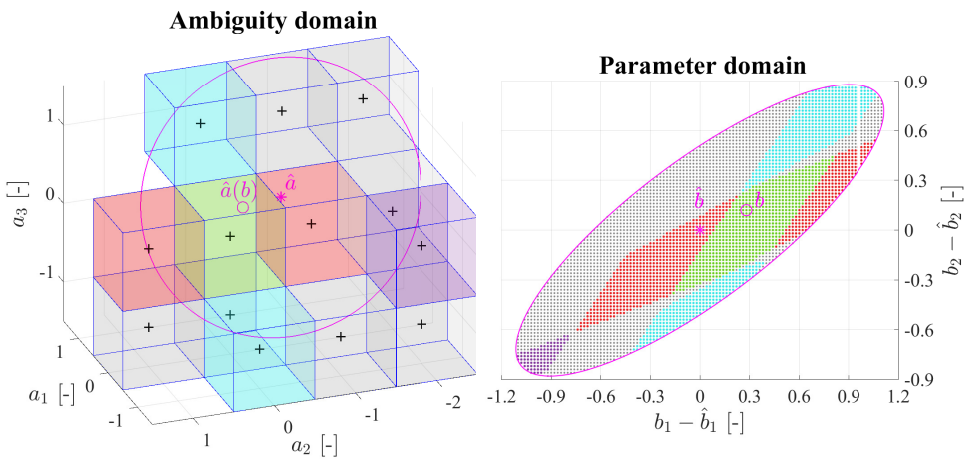


Figure 1.1: An illustrative example is shown for $p = 2, n = 3$, involving both the integer ambiguity domain (left) and the real-valued parameter domain (right). The conditioned hyperplane, centred in $\hat{b}$ and limited to a disk of a certain radius R_0, is represented by a magenta line, while a few selected pull-in regions have been highlighted using colors. The intersection of different pull-in regions with the magenta disk is then depicted in the real-valued parameter domain.

1.3 Research objectives and contribution

In this thesis work, we advance the theoretical and practical understanding of (linear) mixed-integer models through an investigation of ambiguity resolution strategies, focusing first on the domain of integer ambiguities $a \in \mathbb{Z}^n$, then on the domain of real-valued parameters $b \in \mathbb{R}^p$. The two approaches are, respectively, defined by a primal and a dual formulation, representing two complementary perspectives of the same mathematical problem, but some important differences exist. Therefore, in this two-part contribution, we first explore flexible solutions in the integer ambiguity domain, followed by new solutions in the real-valued parameter domain.

This study analyzes two cases (see Figure 1.2):

1. In the primal case, the high dimensionality n of the ambiguities represents a computational bottleneck for the integer ambiguity resolution process, especially when dealing with large GNSS network processing. With the advent of multi-GNSS/multi-frequency constellations, real-time providers of satellite corrections might deal with hundreds or even thousands of ambiguity components, which need to be successfully resolved within a short time. Meanwhile, for other GNSS applications, different requirements could be defined, where users might want to limit the probability of wrongly fixing the ambiguities, e.g. in safety-critical applications, or to compute minimum mean squared errors of user coordinates, e.g. in high-accuracy positioning.

In response to these needs, the current contribution formulates a new flexible Vectorial Integer Bootstrapping (VIB) estimator, as extension of the (scalar) bootstrapping principle. This VIB solution offers quasi-optimal performance, while it is suitable for tackling high-dimensional ambiguity resolution problems, such as the ones arising in the context of Orbit Determination and Time Synchronization (ODTS) with global networks (Montenbruck and Steigenberger, 2020). Some efficient implementations are investigated and provided to the community in a newly designed LAMBDA toolbox that introduces also different (classes of) estimators suitable for the different user applications. Both estimation and evaluation capabilities are integrated into a single, extensively documented software package, which is ready to be used in existing GNSS or non-GNSS software.

2. In the dual case, the focus shifts to real-valued parameters, which are often the parameters of interest for the users. Classical methods, such as the Ambiguity Function method were introduced in the 1980s, see Counselman and Gourevitch (1981), and were originally adopted, e.g., to tackle GPS baseline relative positioning. However, a rigorous study of solutions in the parameter domain has not been performed yet - to the best of our knowledge - and heuristic solutions are frequently preferred despite lacking a statistical foundation, so exposing theoretical gaps in their statistical characterization. Therefore, the ambiguity search in the parameters' domain represents still an open challenge.

   With this contribution, we introduce a dual formulation for mixed-integer models, which complements the primal formulation originally introduced by Teunissen (1993). This novel formulation focuses on the computation of real-valued parameters $b \in \mathbb{R}^p$, such that an implicit ILS problem is solved directly in the p-dimensional domain of b. A rigorous description of its distributional statistics is established, along with the introduction of globally convergent solutions based on easier-to-compute approximations. For the special case $p = 1$, a highly efficient algorithm is presented, with a computational complexity growing *linearly* with the dimensionality n. Lastly, the connection between dual ILS solutions and classical methods is analyzed in this work, thereby addressing long-standing theoretical research gaps in GNSS ambiguity resolution.

Ultimately, this dissertation takes a step further toward the establishment of a unified framework for solving mixed-integer least-squares problems. At the same time, we release a *free* and *open-source* new LAMBDA 4.0 toolbox[1] in support to many applications, e.g. for multi-GNSS/multi-frequency models arising in Precise Point Positining Real-Time Kinematics (PPP-RTK) network estimation (Teunissen and Khodabandeh, 2015), frequency-varying carrier phase signals (Khodabandeh and Teunissen, 2023), and in support of the future deployment of mega-constellations in low Earth orbit (LEO) for Positioning-Navigation-Timing (PNT). Although mainly focused on GNSS applications, many contributions of this thesis are applicable to non-GNSS problems described by (linear) mixed-integer models of Eq.(1.1).

[1]Available at http://pntlab.tudelft.nl/LAMBDA, see Massarweh et al. (2025).

1

1.4 Outline of the thesis

This doctoral thesis is organized in a two-part structure (see Figure 1.2). In Part I, we present Chapters 2, 3 and 4 pertaining to the primal formulation, thus looking at the n-dimensional ambiguity domain. In Part II, we introduce Chapters 5, 6, 7 and 8 related to the dual formulation proposed with this contribution, and we focus on the p-dimensional domain of real-valued parameters. Lastly, in Chapter 9, we summarize the conclusions and provide some recommendations for future work.

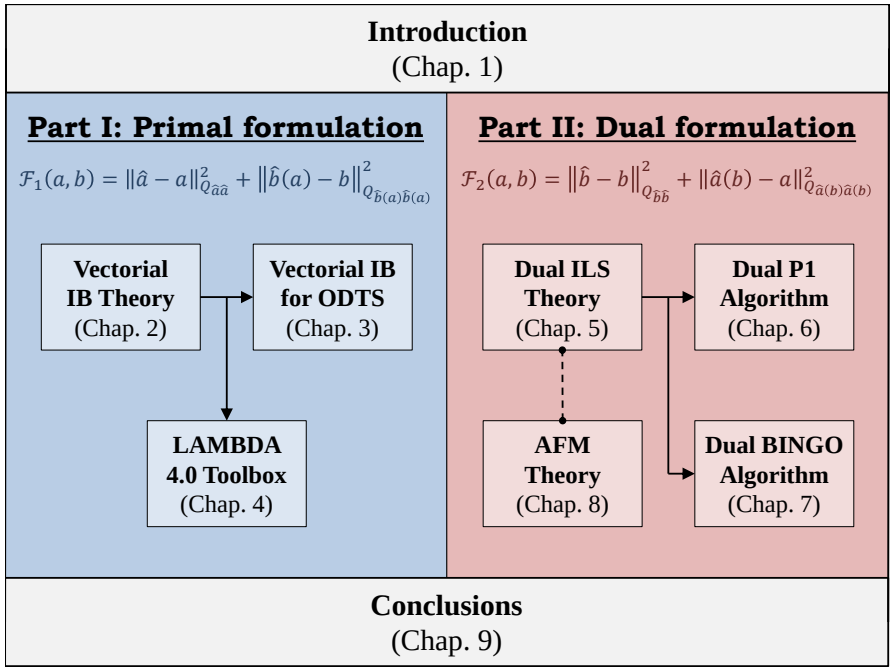


Figure 1.2: Graphical outline of this thesis.

1.4.1 Content per chapter

- **Chapter 2** introduces the theory of Vectorial Integer Boostrapping (VIB), including the mathematical definition of VIB pull-in regions, probability bounds and success rate. A performance ordering is presented with respect to legacy integer estimators, such as IR, IB and ILS, and a few examples are provided, e.g., in the context of GNSS regional network estimation.

  This chapter refers to the following publication:
  Teunissen PJG, Massarweh L, and Verhagen S (2021). Vectorial integer bootstrapping: flexible integer estimation with application to GNSS. In Journal of Geodesy 95, 99. https://doi.org/10.1007/s00190-021-01552-2

- **Chapter 3** considers the application of VIB estimators in Orbit Determination and Time Synchronization (ODTS) problems, based on global networks. First, the impact of ambiguity resolution is examined with respect to different network sizes, then a few different VIB-based implementations are presented and numerically assessed using GROOPS software, based on real-world data from 14 well-distributed ground stations. Finally, ODTS results for GPS and Galileo constellations are validated over a 3-month period in 2019.

  This chapter refers to the following publication:
  Massarweh L, Strasser S, and Mayer-Gürr T (2021). On vectorial integer bootstrapping implementations in the estimation of satellite orbits and clocks based on small global networks. In Advances in Space Research, 68(11), 4303-4320. https://doi.org/10.1016/j.asr.2021.09.023

- **Chapter 4** presents the new LAMBDA 4.0 toolbox implementation, i.e. free and open-source software for integer estimation, validation, and success rate evaluation. This major update introduces new algorithms, including an improved search strategy with a one-order reduction in the computational time, along with new (classes of) estimators: VIB, IAB and BIE. Moreover, it integrates *estimation* and *evaluation* capabilities, respectively based on LAMBDA 3.0 (2012) and Ps-LAMBDA 1.0 (2013) software, further enhanced here for tackling high-dimensional ambiguity resolution problems.

  This chapter refers to the following publication:
  Massarweh L, Verhagen S, and Teunissen PJG (2025). New LAMBDA toolbox for mixed-integer models: estimation and evaluation. In GPS Solutions 29, 14. https://doi.org/10.1007/s10291-024-01738-z

- **Chapter 5** introduces the dual mixed-integer least-squares theory, thus describing a dual formulation that deals with the direct computation of real-valued parameters in mixed-integer (linear) models. It establishes an equivalence with the primal formulation, which is used to determine the distributional properties of newly defined dual estimators. Two easy-to-compute approximations for dual problems are considered, based on an approximate weight matrix and an approximate integer map. Finally, a globally convergent method with finite termination is proposed and its three main constituents are outlined.

  This chapter refers to the following publication:
  Teunissen PJG, Massarweh L (2024). Primal and dual mixed-integer least-squares: distributional statistics and global algorithm. In Journal of Geodesy 98, 63. https://doi.org/10.1007/s00190-024-01862-1

- **Chapter 6** considers dual mixed-integer least-squares problems in the case of scalar real-valued parameters ($p = 1$). An ad hoc 'P1' algorithm is introduced, including geometrical insights, and it is demonstrated how the algorithm's complexity grows linearly with the ambiguity dimensionality n. A higher computational efficiency is demonstrated with respect to LAMBDA, and its potential quasi-optimal statistical performance is discussed.

  This chapter refers to the following publication:
  Massarweh L, Teunissen PJG (2024). An efficient 'P1' algorithm for dual mixed-integer least-squares problems with scalar real-valued parameters. In Journal of Applied Geodesy. https://doi.org/10.1515/jag-2024-0076

- **Chapter 7** investigates the performance of the dual formulation, focusing on the unconstrained dual problem, where real-valued parameters are freely defined in $\mathbb{R}^p$ for the general case $p \geq 1$. A globally convergent solution, namely the Branch-and-bound INteger-equivariant Global Optimizer (BINGO), is examined, relying on a branch-and-bound principle with lower bounds derived via a specific convex relaxation. Computational trade-offs are numerically evaluated, each step is extensively discussed, and finally, the ambiguity fixing performance is assessed in the context of GNSS models.

- **Chapter 8** introduces new theoretical developments for the classical GNSS Ambiguity Function (AF) method, providing a probabilistic model for the AF estimator, which is defined here as maximum likelihood estimator. The globally convergent algorithm used for the (approximate) dual formulation is adapted to the AF objective function by means of a dedicated convex lower bound, ensuring a unique global minimum with a guaranteed ϵ-tolerance. Lastly, a systematic comparison is made between the AF and ILS principles, showing that they are fundamentally different, even if they may still exhibit similarities under certain conditions.

  This chapter will be published as:
  Teunissen PJG, and Massarweh L (2024). Theory for the Ambiguity Function Method: Probability Model and Global Solution. In Journal of Geodesy (*submitted*).

- **Chapter 9** summarizes the conclusions of this research, and provides recommendations for future work.

References

Blewitt G (1989) Carrier phase ambiguity resolution for the Global Positioning System applied to geodetic baselines up to 2000 km. Journal of Geophysical Research: Solid Earth 94(B8):10187–10203, DOI 10.1029/JB094iB08p10187

Brack A (2019) Partial carrier-phase integer ambiguity resolution for high accuracy GNSS positioning. PhD thesis, Verlag Dr. Hut, Munich, Germany

Cocard M, Geiger A (1992) Systematic search for all possible widelanes. In: Proceedings of the 6th International Geodetic Symposium on Satellite Positioning, pp 312–318

Counselman CC, Gourevitch SA (1981) Miniature interferometric terminals for earth surveying: ambiguity and multipath with Global Positioning System. IEEE Transactions Geoscience and Remote Sensing GE-19(4):244–252

Damen M, El Gamal H, Caire G (2003) On maximum-likelihood detection and the search for the closest lattice point. IEEE Transactions on Information Theory 49(10):2389–2402

De Jonge PJ (1998) A processing strategy for the application of the GPS in networks. PhD thesis, Delft University of Technology, Delft, The Netherlands

Ge M, Gendt G, Dick G, Zhang F, Rothacher M (2006) A new data processing strategy for huge GNSS global networks. Journal of Geodesy 80(4):199–203, DOI 10.1007/s00190-006-0044-x

Giorgi G (2011) GNSS carrier phase-based attitude determination: Estimation and applications. PhD thesis, Delft University of Technology, Delft, The Netherlands

Khodabandeh A, Teunissen PJG (2023) Ambiguity-Fixing in Frequency-Varying Carrier Phase Measurements: Global Navigation Satellite System and Terrestrial Examples. NAVIGATION: Journal of the Institute of Navigation 70(2), DOI 10.33012/navi.580

Leick A, Rapoport L, Tatarnikov D (2015) GPS Satellite Surveying, 4th edn. John Wiley & Sons, Inc

Li B, Teunissen PJ (2011) High dimensional integer ambiguity resolution: A first comparison between lambda and bernese. Journal of Navigation 64(S1):192–210, DOI 10.1017/S037346331100035X

Lu C, Li X, Cheng J, Dick G, Ge M, Wickert J, Schuh H (2018) Real-time tropospheric delay retrieval from multi-GNSS PPP ambiguity resolution: Validation with final troposphere products and a numerical weather model. Remote Sensing 10(3):481

Massarweh L, Verhagen S, Teunissen PJG (2025) New LAMBDA toolbox for mixed-integer models: estimation and evaluation. GPS Solutions 29(1):14, DOI 10.1007/s10291-024-01738-z, available at http://pntlab.tudelft.nl/LAMBDA

Mi X, Zhang B, El-Mowafy A, Wang K, Yuan Y (2023) Undifferenced and uncombined GNSS time and frequency transfer with integer ambiguity resolution. Journal of Geodesy 97:13

Montenbruck O, Steigenberger P (2020) GNSS orbit determination and time synchronization. Position, Navigation, and Timing Technologies in the 21st Century: Integrated Satellite Navigation, Sensor Systems, and Civil Applications 1:233–258, DOI 10.1002/9781119458449.ch11

Strasser S (2022) Reprocessing multiple GNSS constellations and a global station network from 1994 to 2020 with the raw observation approach. PhD thesis, Graz University of Technology (90000), DOI 10.3217/978-3-85125-885-1

Teunissen PJG (1993) Least-squares estimation of the integer GPS ambiguities. In: IAG General Meeting. Invited Lecture. Section IV Theory and Methodology

Teunissen PJG (1995) The least-squares ambiguity decorrelation adjustment: A method for fast GPS integer ambiguity estimation. Journal of Geodesy 70(1–2):65–82, DOI 10.1007/BF00863419

Teunissen PJG (1998) GPS carrier phase ambiguity fixing concepts. In: Teunissen PJG, Kleusberg A (eds) GPS For Geodesy, 2nd edn, chap 8, pp 319–388

Teunissen PJG (2003) Towards a unified theory of GNSS ambiguity resolution. Journal of Global Positioning Systems 2(1):1–12

Teunissen PJG (2020) GNSS precise point positioning. In: Position, Navigation, and Timing Technologies in the 21st Century: integrated satellite navigation, sensor systems, and civil applications, vol 1, Wiley, pp 503–528

Teunissen PJG, Khodabandeh A (2015) Review and principles of PPP-RTK methods. Journal of Geodesy 89(3):217–240, DOI 10.1007/s00190-014-0771-3

Teunissen PJG, Massarweh L (2024) Primal and dual mixed-integer least-squares: Distributional statistics and global algorithm. Journal of Geodesy 98:63, DOI 10.1007/s00190-024-01862-1

Verhagen S (2005) The GNSS integer ambiguities: Estimation and validation. PhD thesis, Delft University of Technology, Delft, The Netherlands

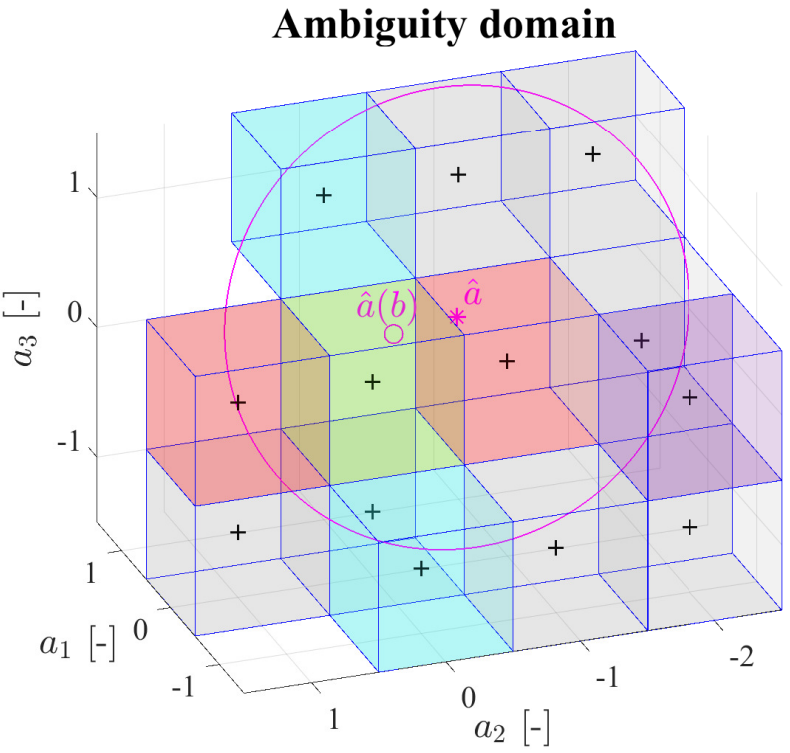
Viegas D, Cunha S (2007) Precise positioning by phase processing of sound waves. IEEE Transactions on Signal Processing 55(12):5731–5738

Wang K, El-Mowafy A, Cheng F, Yang X (2024) LEO augmentation in large-scale ionosphere-float PPP-RTK positioning. Journal of Surveying Engineering 150(2):04024001

Xu S, Yang Q, Du X, Xu X, Zhao Q, Yang L, Qin Y, Guo J (2024) Multi-GNSS precise point positioning enhanced by the real navigation signals from CENTISPACE LEO mission. Advances in Space Research 73(8):4175–4186, DOI 10.1016/j.asr.2024.01.017

I

Primal formulation



2

Vectorial Integer Bootstrapping: Flexible Integer Estimation with Application to GNSS

In this contribution we extend the principle of integer bootstrapping (IB) to a vectorial form (VIB). The mathematical definition of the class of VIB-estimators is introduced together with their pull-in regions and other properties such as probability bounds and success-rate approximations. The vectorial formulation allows sequential block-by-block processing of the ambiguities based on a user-chosen partitioning. In this way, flexibility is created, where for specific choices of partitioning, tailored VIB estimators can be designed. This wide range of possibilities is discussed, supported by numerical simulations and analytical examples. Further guidelines are provided, as well as the possible extension to other classes of estimators.

This chapter has been published as: Teunissen, P.J.G., Massarweh, L., and Verhagen, S. (2021). *Vectorial integer bootstrapping: flexible integer estimation with application to GNSS*. In Journal of Geodesy 95, 99. https://doi.org/10.1007/s00190-021-01552-2

2.1 Introduction

Global navigation satellite system (GNSS) ambiguity resolution is the process of resolving the unknown integer ambiguities of the observed carrier phases. Once they are successfully resolved, advantage can be taken of the mm-level precision of the carrier-phase measurements, thereby de facto turning them into ultra-precise pseudo-ranges. The practical importance of this becomes clear when considering the great variety of current and future GNSS models to which integer ambiguity resolution applies. A comprehensive overview of these GNSS models, together with their applications in surveying, navigation, geodesy and geophysics, can be found in textbooks such as (Strang and Borre, 1997; Teunissen and Kleusberg, 1998; Kaplan and Hegarty, 2006; Misra and Enge, 2006; Hofmann-Wellenhof et al., 2008; Borre and Strang, 2012; Leick et al., 2015; Teunissen and Montenbruck, 2017; Morton et al., 2021).

In the current theory of integer ambiguity resolution one can discriminate between three different classes of estimators: the class of integer (I) estimators (Teunissen, 1999), the class of integer-aperture (IA) estimators (Teunissen, 2003a) and the class of integer-equivariant (IE) estimators (Teunissen, 2003b). The classes are subsets of one another, with the I-class being the smallest and the IE-class the largest: $\mathrm{I} \subset \mathrm{IA} \subset \mathrm{IE}$. Members from all three classes have found their application in a wide range of different GNSS models, see e.g. (Brack et al., 2014; Hou et al., 2020; Odolinski and Teunissen, 2020; Psychas and Verhagen, 2020; Verhagen, 2005; Wang et al., 2018; Zaminpardaz et al., 2018). In this contribution, we restrict attention to the I-class.

Popular estimators in the I-class are integer rounding (IR), integer bootstrapping (IB) and integer least-squares (ILS). Integer rounding is the simplest, but has the poorest success-rate performance, while integer least-squares has the best performance, but is computationally the most complex (Teunissen, 1998, 1999). In terms of their success-rate performance, the three integer estimators can thus be ordered as follows

$$\mathrm{IR} \leq \mathrm{IB} \leq \mathrm{ILS} \tag{2.1}$$

where the equality holds if the ambiguities are perfectly decorrelated.

The IB-estimator has the attractive property that it is simple to compute and that it can have a close to optimal success-rate performance when used with a properly chosen ambiguity parametrization. The estimator is characterized by two operations that are alternately applied: a sequential conditional least-squares estimation and an integer mapping. With the IB these are both applied at the scalar-level. This is however not a necessity for the two principles to be applicable.

In this contribution we will generalize the scalar concept of integer bootstrapping to a vectorial form in which both the sequential conditional estimation and the integer mapping are vectorial. We will develop the associated theory and show that with the concept of vectorial integer bootstrapping (VIB) great flexibility is created for designing one's integer estimators when balancing computational simplicity against success-rate performance.

This contribution is organized as follows. In Section 2.2, we briefly review integer ambiguity resolution with a special emphasis on integer bootstrapping and its various properties. The concept of vectorial integer bootstrapping (VIB) is introduced in Section 2.3. Two different descriptions of the class of VIB-estimators are given, one that follows naturally from the definition of scalar integer bootstrapping and another that is more suitable to characterize its pull-in regions. In Section 2.4 we develop probabilistic properties of the VIB-estimators, with a special emphasis on their probability of correct integer estimation, the ambiguity success-rate, together with its easy-to-compute lower bounds and upper bounds. In this section we also provide a generalized version of the performance I-ordering defined in (2.1). It includes newly defined VIB-estimators and it shows the great flexibility one has in working with vectorial integer bootstrapping, which we also demonstrate through numerical examples. In Section 2.5, we provide further considerations for VIB-usage, including the choice of ambiguity parametrization, efficiency enhancing options when solving the normal equations and the different ways in which the VIB concept can be extended to other classes of estimators. Finally a summary with concluding remarks is provided in Section 2.6.

The following notation is used throughout: $\mathsf{E}(\cdot)$ denotes the expectation operator, $\mathsf{D}(\cdot)$ the dispersion operator, $\mathsf{P}(\cdot)$ probability of an event, $||\cdot||^2_Q = (\cdot)^T Q^{-1}(\cdot)$ the square-weighted-norm in the metric of Q, and $\mathcal{I} : \mathbb{R}^n \mapsto \mathbb{Z}^n$ an admissible integer map.

2.2 Mixed-integer model estimation

In this section a brief review is given of mixed-integer model estimation with an emphasis on the method of integer bootstrapping.

2.2.1 Mixed-integer model and ambiguity resolution

The basis of GNSS integer ambiguity resolution is formed by the mixed-integer model, given as

$$\mathsf{E}(y) = Aa + Bb, \qquad \mathsf{D}(y) = Q_{yy} \tag{2.2}$$

with $y \in \mathbb{R}^m$ the vector of observables containing the pseudo-ranges and carrier-phases, $(A, B) \in \mathbb{R}^{m\times(n+p)}$ the full-rank design matrix, $a \in \mathbb{Z}^n$ the vector of unknown integer ambiguities, $b \in \mathbb{R}^p$ the vector of real-valued unknown parameters (e.g. baseline components and atmospheric delays), and Q_{yy} the positive-definite variance-covariance (vc) matrix of the observables.

In the following we will refer to b simply as the baseline vector. The underlying distribution of the previous mixed-integer model will be assumed a multivariate normal distribution as common in many GNSS applications (Leick et al., 2015; Teunissen and Montenbruck, 2017; Morton et al., 2021). The solution to the (2.2) is obtained through the following three steps:

Step 1: the model is solved by means of least-squares (LS) estimation whereby the

integerness of the ambiguities is discarded. This gives the so-called *float* solution, together with its vc-matrix, expressed by

$$\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} \quad , \quad \begin{bmatrix} Q_{\hat{a}\hat{a}} & Q_{\hat{a}\hat{b}} \\ Q_{\hat{b}\hat{a}} & Q_{\hat{b}\hat{b}} \end{bmatrix} \tag{2.3}$$

Step 2: an admissible integer map $\mathcal{I} : \mathbb{R}^n \mapsto \mathbb{Z}^n$ is chosen to compute the integer ambiguity vector as

$$\check{a} = \mathcal{I}(\hat{a}) \tag{2.4}$$

The integer map is admissible when its *pull-in regions* $\mathcal{P}_z = \{x \in \mathbb{R}^n | \ \mathcal{I}(x) = z\}$, $z \in \mathbb{Z}^n$, cover $\mathbb{R}^n$, while being disjoint and integer translational invariant. It follows that those regions leave no gaps, have no overlaps and obey an integer remove–restore principle (Teunissen, 1999).

Some popular choices for $\mathcal{I}$ are integer rounding (IR), integer bootstrapping (IB), or integer least-squares (ILS). To enhance the probabilistic and/or numerical performance of ambiguity resolution, the integer map is often preceded by the decorrelating Z-transformation $\hat{z} = Z^T\hat{a}$ of the LAMBDA method (Teunissen, 1995), in which case the integer estimate of a is computed as

$$\check{a} = Z^{-T}\mathcal{I}(Z^T\hat{a}) \tag{2.5}$$

Step 3: after $\check{a} \in \mathbb{Z}^n$ has been validated, the ambiguity-resolved or *fixed* baseline solution can be given as

$$\check{b} = \hat{b} - Q_{\hat{b}\hat{a}}Q_{\hat{a}\hat{a}}^{-1}(\hat{a} - \check{a}) \tag{2.6}$$

Its vc-matrix, in case the uncertainty of $\check{a}$ may be neglected, is given as

$$Q_{\check{b}\check{b}} = Q_{\hat{b}\hat{b}} - Q_{\hat{b}\hat{a}}Q_{\hat{a}\hat{a}}^{-1}Q_{\hat{a}\hat{b}} \tag{2.7}$$

which shows by how much the precision of $\hat{b}$ will be improved as a consequence of imposing the integer ambiguity constraint $a \in \mathbb{Z}^n$.

2.2.2 Integer Bootstrapping

The IB-estimator is one of the most popular integer ambiguity estimators. Its popularity stems from the ease with which it can be computed and from its close to optimal performance once the ambiguities are sufficiently decorrelated. The IB-estimator, following (Teunissen, 1998) is here defined.

Definition 1 (*Scalar integer bootstrapping*). Let $\hat{a} = (\hat{a}_1, \ldots, \hat{a}_n)^T \in \mathbb{R}^n$ be the float solution and let $\check{a}_{\mathrm{IB}} = (\check{a}_1, \ldots, \check{a}_n)^T \in \mathbb{Z}^n$ denote the corresponding integer bootstrapped (IB) solution. Then

$$\begin{array}{lcll} \check{a}_1 & = & \lceil \hat{a}_1 \rfloor & \\ \check{a}_2 & = & \lceil \hat{a}_{2|1} \rfloor & = \lceil \hat{a}_2 - \sigma_{21}\sigma_1^{-2}(\hat{a}_1 - \check{a}_1) \rfloor \\ & \vdots & & \\ \check{a}_n & = & \lceil \hat{a}_{n|N} \rfloor & = \lceil \hat{a}_n - \sum_{j=1}^{n-1} \sigma_{n,j|J}\sigma_{j|J}^{-2}(\hat{a}_{j|J} - \check{a}_j) \rfloor \end{array} \tag{2.8}$$

where $\lceil \cdot \rfloor$ denotes integer rounding and $\hat{a}_{i|I}$ is the least-squares estimator of a_i conditioned on the values of the previous $I = \{1, \ldots, (i-1)\}$ sequentially rounded components, $\sigma_{i,j|J}$ is the covariance between $\hat{a}_i$ and $\hat{a}_{j|J}$, and $\sigma^2_{j|J}$ is the variance of $\hat{a}_{j|J}$. For $i = 1$, $\hat{a}_{i|I} = \hat{a}_1$. $\square$

This previous definition shows that the IB-estimator is driven by products of ambiguity conditional covariances and variances like $\sigma_{i,j|J} \cdot \sigma^{-2}_{j|J}$. As shown in (Teunissen, 2007), these are readily available when one works with the triangular decomposition of the ambiguity vc-matrix, $Q_{\hat{a}\hat{a}} = LDL^T$, thus making the computation of the IB-estimator particularly easy. The matrix L is lower unitriangular, so its entries are defined as

$$(L)_{ij} = \begin{cases} 0 & \text{for} \quad 1 \leq i < j \leq n \\ 1 & \text{for} \quad i = j \\ \sigma_{i,j|J} \cdot \sigma^{-2}_{j|J} & \text{for} \quad 1 \leq j < i \leq n \end{cases} \tag{2.9}$$

while we have

$$D = \text{diag}(\sigma^2_1, \sigma^2_{2|1}, \ldots, \sigma^2_{j|J}, \ldots) \tag{2.10}$$

Another attractive feature of the IB-estimator is that its probability of correct integer estimation, or success-rate, can be easily computed. Its analytical expression is given by Teunissen (1998) as follows

$$\mathsf{P}(\check{a}_{\text{IB}} = a) = \prod_{i=1}^{n} \left[2\Phi\left(\frac{1}{2\sigma_{i|I}} \right) - 1 \right] \tag{2.11}$$

with $\Phi(\cdot)$, i.e. cumulative normal distribution, being

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{1}{2} v^2 \right\} dv \tag{2.12}$$

Note that while the entries of L in $Q_{\hat{a}\hat{a}} = LDL^T$ drive the IB-estimator, entries of D are the ones that determine its success-rate. Moreover, the IB-estimator is not the best estimator within the class of I-estimators. Teunissen (1999) proved that of all admissible integer estimators, the ILS-estimator

$$\check{a}_{\text{ILS}} = \arg \min_{z \in \mathbb{Z}^n} ||\hat{a} - z||^2_{Q_{\hat{a}\hat{a}}} \tag{2.13}$$

is best in the sense that it has the largest possible success-rate. The price one pays for this optimality is that, in contrast to the easy-to-compute IB-estimator, the computation of (2.13) is based on a more elaborate integer search (Teunissen, 1995).

Although IB is not best in the class of integer estimators, it is best in a smaller class, namely the class of *sequential integer estimators*. This class was introduced in (Teunissen, 2007) as any $\mathcal{I} : \mathbb{R}^n \mapsto \mathbb{Z}^n$ for which

$$\mathcal{I}(x) = [x + (R - I_n)(x - \mathcal{I}(x))] \tag{2.14}$$

where $[\cdot]$ denotes component-wise integer rounding of its vectorial entry and R is an arbitrary unit lower triangular matrix.

Note that both IR and IB belong to this class. IR is obtained with the choice $R = I_n$ and IB with the choice $R = L^{-1}$. In (Teunissen, 2007) it is shown that of all sequential integer estimators, IB has the largest success-rate. Hence, for the success-rate of the three popular integer estimators, we have $\mathsf{P}(\breve{a}_{\mathrm{IR}} = a) \leq \mathsf{P}(\breve{a}_{\mathrm{IB}} = a) \leq \mathsf{P}(\breve{a}_{\mathrm{ILS}} = a)$. Thus integer rounding has poorest performance and integer-least-squares the best. In the following section we will extend this ordering by including the success-rate performance of a vectorial formulation for the integer bootstrapping.

2.3 Vectorial Integer Bootstrapping

In this section we introduce the concept of vectorial integer bootstrapping (VIB) together with a description of its pull-in regions that are illustrated by means of a few 3-dimensional examples.

2.3.1 The VIB estimator

The IB-estimator is characterized by two elements that are alternately applied, the sequential conditional estimation and the integer mapping. With the IB-estimator they are both applied at the scalar-level. This is however not a necessity for the two applied principles. Hence, we may generalize the scalar integer bootstrapping to a form in which both the sequential conditional estimation and the integer mapping are vectorial.

As a result we have the following definition of vectorial integer bootstrapping.

Definition 2 (*Vectorial integer bootstrapping*) Let $\hat{a} = (\hat{a}_1, \ldots, \hat{a}_v)^T \in \mathbb{R}^n$ be the float ambiguity solution, with $\hat{a}_i \in \mathbb{R}^{n_i}$, $i = 1, \ldots, v$ and $n = \sum_{i=1}^{v} n_i$, while let $\breve{a}_{\mathrm{VIB}} = (\breve{a}_1, \ldots, \breve{a}_v)^T \in \mathbb{Z}^n$ denote the corresponding vectorial integer bootstrapped solution. Then

$$\begin{array}{rclcl} \breve{a}_1 & = & \lceil \hat{a}_1 \rfloor_1 & & \\ \breve{a}_2 & = & \lceil \hat{a}_{2|1} \rfloor_2 & = & \lceil \hat{a}_2 - Q_{21} Q_{11}^{-1} (\hat{a}_1 - \breve{a}_1) \rfloor_2 \\ & \vdots & & & \\ \breve{a}_v & = & \lceil \hat{a}_{v|V} \rfloor_v & = & \lceil \hat{a}_v - \sum_{j=1}^{v-1} Q_{vj|J} Q_{jj|J}^{-1} (\hat{a}_{j|J} - \breve{a}_j) \rfloor_v \end{array} \tag{2.15}$$

where $\lceil \cdot \rfloor_i : \mathbb{R}^{n_i} \mapsto \mathbb{Z}^{n_i}$ is a still to be chosen admissible integer mapping, and $\hat{a}_{i|I}$ is the least-squares estimator of a_i conditioned on the values of the previous $I = \{1, \ldots, (i-1)\}$ sequentially integer estimated vectors, $Q_{ij|J}$ is the covariance matrix of $\hat{a}_i$ and $\hat{a}_{j|J}$, and $Q_{jj|J}$ is the variance matrix of $\hat{a}_{j|J}$. For $i = 1$, $\hat{a}_{i|I} = \hat{a}_1$. □

Note that each v integer map, $\lceil . \rfloor_i : \mathbb{R}^{n_i} \mapsto \mathbb{Z}^{n_i}$, in (2.15), can still be chosen freely. As an example consider the case where $\hat{a}$ is partitioned in two parts, $\hat{a} = (\hat{a}_1^T, \hat{a}_2^T)^T$.

Then $v = 2$ and the conditional estimator of a_2, when conditioned on a_1, is given as

$$\hat{a}_2(a_1) = \hat{a}_2 - Q_{\hat{a}_2\hat{a}_1} Q^{-1}_{\hat{a}_1\hat{a}_1} (\hat{a}_1 - a_1) \tag{2.16}$$

having as vc-matrix

$$Q_{\hat{a}_2\hat{a}_2|a_1} = Q_{\hat{a}_2\hat{a}_2} - Q_{\hat{a}_2\hat{a}_1} Q^{-1}_{\hat{a}_1\hat{a}_1} Q_{\hat{a}_1\hat{a}_2} \tag{2.17}$$

Would we now choose both $\lceil\cdot\rfloor_1$ and $\lceil\cdot\rfloor_2$ as ILS-maps, the corresponding VIB_{ILS} solution follows as

$$\begin{aligned} \breve{a}_{1,\text{VIB}_{\text{ILS}}} &= \arg \min_{z_1 \in \mathbb{Z}^{n_1}} ||\hat{a}_1 - z_1||^2_{Q_{\hat{a}_1\hat{a}_1}} \\ \breve{a}_{2,\text{VIB}_{\text{ILS}}} &= \arg \min_{z_2 \in \mathbb{Z}^{n_2}} ||\hat{a}_2(\breve{a}_{1,\text{VIB}_{\text{ILS}}}) - z_2||^2_{Q_{\hat{a}_2\hat{a}_2|a_1}} \end{aligned} \tag{2.18}$$

Thus now two ILS-problems need to be solved, but both at a lower dimension than that of the original full ILS-problem, which was defined as

$$\breve{a}_{\text{ILS}} = \arg \min_{z \in Z^n} ||\hat{a} - z||^2_{Q_{\hat{a}\hat{a}}} \tag{2.19}$$

The two solutions, (2.18) and (2.19), can be compared if we make use of the following orthogonal decomposition (Teunissen, 1995),

$$||\hat{a} - z||^2_{Q_{\hat{a}\hat{a}}} = ||\hat{a}_1 - z_1||^2_{Q_{\hat{a}_1\hat{a}_1}} + ||\hat{a}_2(z_1) - z_2||^2_{Q_{\hat{a}_2\hat{a}_2|a_1}} \tag{2.20}$$

This shows that instead of minimizing the sum of quadratic forms, as done with (2.19), the solution in (2.18) is obtained by minimizing the two quadratic forms separately. The integer-valued vector z_1 in (2.20) is obtained during the first minimization.

In general, the integer mapping in each block can be chosen as only IR, or only IB or only ILS, or combinations of them. Nonetheless, if for a certain n_i-dimensional block we consider IB as the estimator, this is equivalent to a further partitioning into n_i (scalar) blocks. It follows that adopting IB in each block, i.e. applying the VIB_{IB} estimator, will lead to the exactly same (scalar) IB formulation, thence this is not separately analyzed.

2.3.2 The VIB pull-in regions

In our VIB-definition given in (2.15) we have described the components of the VIB-estimator using the analogy with its scalar variant in (2.8). For the purpose of describing the pull-in regions of the VIB-estimator, we now provide its vectorial form thereby drawing on the analogy with (2.14).

Lemma 1 (VIB-*estimator*) Let $\mathcal{I}_{\text{VIB}} : \mathbb{R}^n \mapsto \mathbb{Z}^n$ be the VIB-defining admissible integer map. Then we have $\breve{a}_{\text{VIB}} = \mathcal{I}_{\text{VIB}}(\hat{a})$, with

$$\mathcal{I}_{\text{VIB}}(x) = \mathcal{I}\left(x + (\mathcal{L} - I_n)\left(x - \mathcal{I}_{\text{VIB}}(x)\right) \right) \tag{2.21}$$

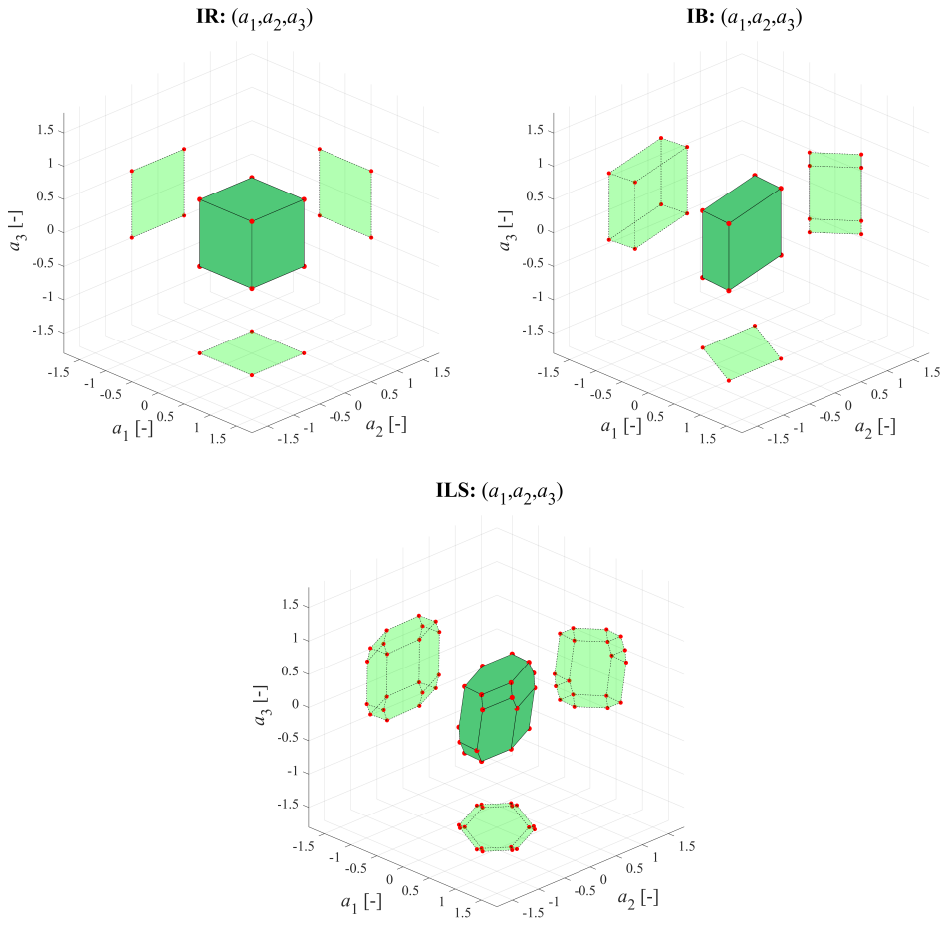


Figure 2.1: The 3D pull-in regions and their 2D projections are given for integer rounding (IR, cube), integer bootstrapping (IB, parallelopiped) and integer least-squares (ILS, parallelohedron). The vertexes of each region are given in red.

where $\mathcal{I} : \mathbb{R}^n \mapsto \mathbb{Z}^n$ is given as $\mathcal{I}(x) = (\lceil x_1 \rfloor_1^T, \ldots, \lceil x_v \rfloor_v^T)^T$ for $x = (x_1^T, \ldots, x_v^T)^T$. The $\mathcal{L} \in \mathbb{R}^{n \times n}$ is given as a lower block-triangular matrix, such that

$$\mathcal{L} = \begin{bmatrix} I_{n_1} & & & & \\ -Q_{\hat{a}_2\hat{a}_1} Q_{\hat{a}_1\hat{a}_1}^{-1} & I_{n_2} & & & \\ -Q_{\hat{a}_3\hat{a}_1} Q_{\hat{a}_1\hat{a}_1}^{-1} & -Q_{\hat{a}_3\hat{a}_2|a_1} Q_{\hat{a}_2\hat{a}_2|a_1}^{-1} & I_{n_3} & & \\ \vdots & \vdots & & \ddots & \\ -Q_{\hat{a}_v\hat{a}_1} Q_{\hat{a}_1\hat{a}_1}^{-1} & -Q_{\hat{a}_v\hat{a}_2|a_1} Q_{\hat{a}_2\hat{a}_2|a_1}^{-1} & & \ldots & I_{n_v} \end{bmatrix} \tag{2.22}$$

Proof. The proof follows directly by substitution thereby using the lower block-triangularity of $\mathcal{L}$ in (2.21). □

The pull-in region $\mathcal{P}_z$ of an integer estimator is defined as the region in which all float solutions are mapped to the same integer $z \in \mathbb{Z}^n$ by the integer estimator. Hence, for the VIB-estimator it is defined as

$$\mathcal{P}_{z,\mathrm{VIB}} = \{x \in \mathbb{R}^n |\, \mathcal{I}_{\mathrm{VIB}}(x) = z\}, \quad z \in \mathbb{Z}^n \tag{2.23}$$

The following Lemma shows how the pull-in regions are indeed driven by the VIB characterizing integer maps $\lceil . \rfloor_i : \mathbb{R}^{n_i} \mapsto \mathbb{Z}^{n_i}$ and sequential conditional estimation.

Lemma 2 (VIB *pull-in region*) The pull-in regions of $\breve{a}_{\mathrm{VIB}} = \mathcal{I}_{\mathrm{VIB}}(\hat{a})$ are given as

$$\mathcal{P}_{z,\mathrm{VIB}} = \{x \in \mathbb{R}^n |\, \mathcal{I}(\mathcal{L}(x - z)) = 0\}, \;\; z \in \mathbb{Z}^n \tag{2.24}$$

where $\mathcal{I}(x) = (\lceil x_1 \rfloor_1^T, \ldots, \lceil x_v \rfloor_v^T)^T$ and $\mathcal{L}$ is the lower block-triangular matrix given in (2.22).

Proof. Starting with (2.21), we have

$$\begin{aligned} \mathcal{I}_{\mathrm{VIB}}(x) &= \mathcal{I}(\; x + (\mathcal{L} - I_n)(x - z)\;) \\ &= \mathcal{I}(\; \mathcal{L}(x - z) + z\;) \\ &= \mathcal{I}(\mathcal{L}(x - z)) + z \end{aligned} \tag{2.25}$$

from which, using (2.23), the result follows. □

To gain further insights into the geometries of the VIB pull-in regions, in particular under different choices for the integer maps $\lceil . \rfloor_i : \mathbb{R}^{n_i} \mapsto \mathbb{Z}^{n_i}$, we provide a few graphical representations.

2.3.3 Graphics of VIB pull-in regions

We consider a three-dimensional float ambiguity vector $\hat{a} = (\hat{a}_1, \hat{a}_2, \hat{a}_3)^T \in \mathbb{R}^3$, having as vc-matrix

$$Q_{\hat{a}\hat{a}} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 \end{bmatrix} \cong \begin{bmatrix} 0.090 & -0.045 & 0.027 \\ -0.045 & 0.101 & 0.002 \\ 0.027 & 0.002 & 0.171 \end{bmatrix} \tag{2.26}$$

Its lower unitriangular matrix L and the diagonal matrix D from $Q_{\hat{a}\hat{a}} = LDL^T$ are given respectively as

$$L = \begin{bmatrix} 1.000 & 0 & 0 \\ -0.499 & 1.000 & 0 \\ 0.300 & 0.200 & 1.000 \end{bmatrix} \tag{2.27}$$

$$D = \begin{bmatrix} 0.30^2 & 0 & 0 \\ 0 & 0.28^2 & 0 \\ 0 & 0 & 0.40^2 \end{bmatrix} \tag{2.28}$$

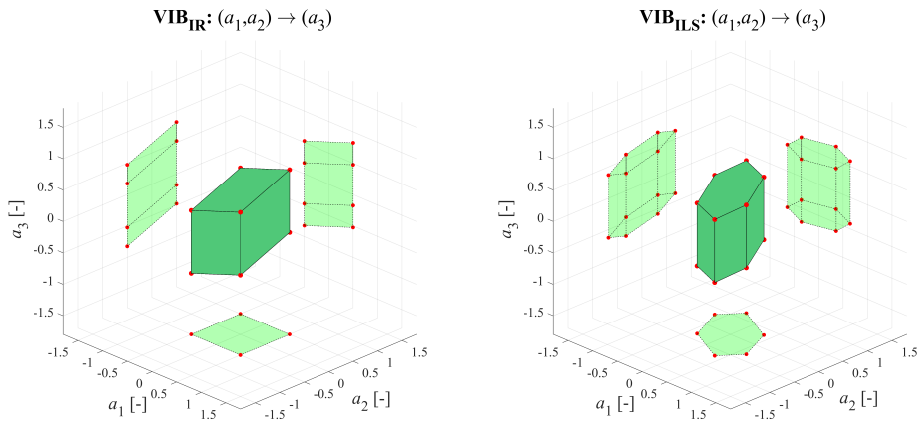


Figure 2.2: The 3D pull-in regions are given for the VIB_{IR} $[(a_1, a_2) \rightarrow (a_3)]$ estimator (left) and for the VIB_{ILS} $[(a_1, a_2) \rightarrow (a_3)]$ estimator (right). The arrow refers to a conditioning of the last component on the first two ambiguities.

The pull-in regions of the traditional integer estimators IR, IB and ILS are first considered. They are shown in Fig.2.1 as a cube, parallelepiped, and parallelohedron, respectively. Their projections on the three mutually orthogonal coordinate planes are also shown. Note that generally it is wise to start with the most precise ambiguities in the IB, in this case $\hat{a}_1$.

For the three-dimensional VIB-estimators, two different ambiguity-partitionings can be considered, depending on whether $\lceil \cdot \rfloor_1$ is a 1-dimensional or a 2-dimensional map. In Fig. 2.2, $\lceil \cdot \rfloor_1$ is 2-dimensional for which IR (Fig. 2.2, left) and ILS (Fig. 2.2, right) is chosen. The third ambiguity is then conditionally updated and rounded to its nearest integer. For the alternative case that $\lceil \cdot \rfloor_1$ is 1-dimensional, one starts with ordinary rounding, but then has different choices for the second step. In Fig. 2.3, the second step, where conditioning has taken place on the first ambiguity, is based on IR (Fig. 2.3, left) and ILS (Fig. 2.3, right).

As visible from these examples, different VIB-estimators have different geometries for their pull-in regions. Hence, their pull-in regions will have different fits to the float-ambiguity's confidence region and therefore also different success-rates.

2.4 VIB probability of correct integer estimation

In this section we study the success-rate of VIB estimators, along with bounds and possible approximations.

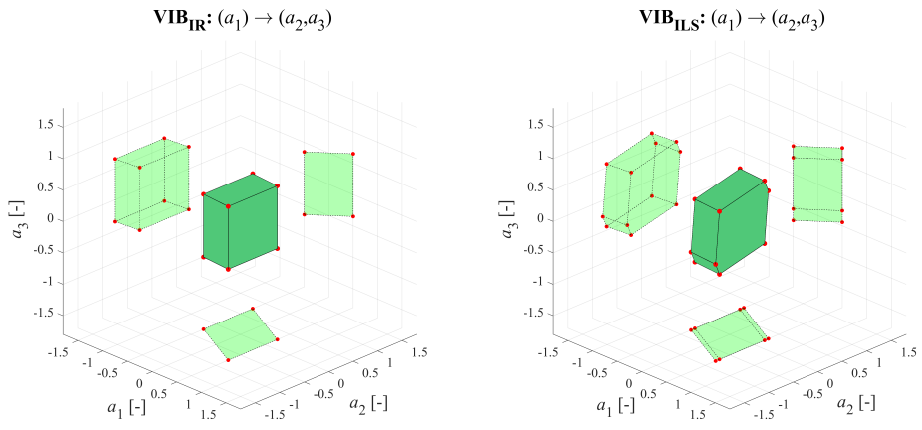


Figure 2.3: The 3D pull-in regions are given for the VIB_{IR} $[(a_1) \rightarrow (a_2, a_3)]$ estimator (left) and for the VIB_{ILS} $[(a_1) \rightarrow (a_2, a_3)]$ estimator (right). The arrow refers to a conditioning of the last two components on the first ambiguity.

2.4.1 The VIB success rate

As the success rate of an admissible integer estimator of $a \in \mathbb{Z}^n$ is given by the amount of probability mass its a-centred pull-in region covers of the probability density function (PDF) of $\hat{a}$, $f_{\hat{a}}(x)$, the VIB success rate is given by the integral

$$\mathsf{P}(\check{a}_{\text{VIB}} = a) = \int_{\mathcal{P}_{a,\text{VIB}}} f_{\hat{a}}(x)dx \tag{2.29}$$

The following Lemma shows how this can be computed from a product of probabilities.

Lemma 3 (VIB *success rate*) Let $\hat{a} \sim \mathcal{N}(a, Q_{\hat{a}\hat{a}})$ and let $\check{a}_{\text{VIB}}$ be the VIB-estimator of $a \in \mathbb{Z}^n$, as expressed in (2.15). Then

$$\mathsf{P}(\check{a}_{\text{VIB}} = a) = \prod_{i=1}^{v} \mathsf{P}(\hat{a}_{i|I} \in a + \mathcal{P}_i) \tag{2.30}$$

with

$$\mathsf{P}(\hat{a}_{i|I} \in a + \mathcal{P}_i) = \int_{\mathcal{P}_i} \frac{1}{\sqrt{|2\pi Q_{ii|I}|}} \exp\left\{-\tfrac{1}{2}||x_i||^2_{Q_{ii|I}}\right\} dx_i \tag{2.31}$$

and $\mathcal{P}_i = \{x \in \mathbb{R}^{n_i} \mid \lceil x \rfloor_i = 0\}$, $i = 1, \ldots, v$.

Proof. The proof is given in the Appendix. □

This general result has two familiar special cases. For $v = n$, then the above VIB success rate reduces to that of the (scalar) integer bootstrapping given in (2.11). For $v = 1$, it depends on the chosen integer mapping $\lceil \cdot \rfloor : \mathbb{R}^n \mapsto \mathbb{Z}^n$, for which IR, IB and ILS are the popular contenders.

2.4.2 Bounds and approximation of VIB success rate

We now discuss how existing bounds and approximations to the success rates of IR, IB and ILS can be used in order to obtain overall bounds and approximations for the VIB success rate as well.

Lemma 4 (VIB *success-rate bounds*) Let $\mathcal{P}_i = \{x \in \mathbb{R}^{n_i} \mid \lceil x \rfloor_i = 0\}$ be the origin-centred pull-in region of the integer map $\lceil . \rfloor_i : \mathbb{R}^{n_i} \mapsto \mathbb{Z}^{n_i}$ and let the bounds $\mathrm{LB}_i \leq \mathsf{P}(\hat{a}_{i|I} \in a + \mathcal{P}_i) \leq \mathrm{UB}_i$ be given. Then

$$\prod_{i=1}^{v} \mathrm{LB}_i \leq \mathsf{P}(\check{a}_{\mathrm{VIB}} = a) \leq \prod_{i=1}^{v} \mathrm{UB}_i \tag{2.32}$$

Proof. This follows from using the individual bounds LB_i and UB_i in (2.30). □

By making use of known bounds for IR and ILS, respectively, (2.32) can be easily applied to obtain the following two success-rate bounds.

Lemma 5 (*Success-rate bounds for* $\mathrm{VIB}_{\mathrm{IR}}$ *and* $\mathrm{VIB}_{\mathrm{ILS}}$) Let $\sigma^2_{j|I}$, for $j = q_{i-1}+1, \ldots, q_i$ ($q_0 = 0, q_i = \sum_{j=1}^{i} n_j$), and $i = 1, \ldots, v$, be the variance of the j-th ambiguity conditioned on a_I, with $I = \{1, \ldots, i-1\}$. Then with the Ambiguity Dilution of Precision (ADOP) of the ith subset $\hat{a}_{i|I} \in \mathbb{R}^{n_i}$,

$$\mathrm{ADOP}_i = |Q_{ii|I}|^{\frac{1}{2n_i}} \tag{2.33}$$

we have

$$\begin{aligned} \mathsf{P}(\check{a}_{\mathrm{VIB}_{\mathrm{IR}}} = a) &\geq \prod_{i=1}^{v} \left(\prod_{j=q_{i-1}+1}^{q_i} \left[2\Phi\left(\tfrac{1}{2\sigma_{j|I}}\right) - 1 \right] \right) \\ \mathsf{P}(\check{a}_{\mathrm{VIB}_{\mathrm{ILS}}} = a) &\leq \prod_{i=1}^{v} \mathsf{P}\left(\chi^2(n_i, 0) \leq \tfrac{c_{n_i}}{\mathrm{ADOP}_i^2} \right) \end{aligned} \tag{2.34}$$

where $\chi^2(n_i, 0)$ denotes a central chi-squared distributed random variable with n_i degrees of freedom, while

$$c_{n_i} = \frac{1}{\pi} \left[\frac{n_i}{2} \cdot \Gamma\left(\frac{n_i}{2}\right) \right]^{2/n_i} \tag{2.35}$$

with Γ being the gamma-function.

Proof. For the proof see the Appendix. □

These bounds can be used in the following sense:

- $\mathrm{VIB}_{\mathrm{IR}}$ can be considered good enough for ambiguity resolution if its lower bound is large enough.
- $\mathrm{VIB}_{\mathrm{ILS}}$ can be considered too poor if its upper bound is too small.

2.4.3 Performance ordering of VIB estimators

We now determine a performance ordering for the different integer estimators discussed. Such ordering can also be adopted to determine additional bounds on the success rates of well-known integer estimators.

First we can start with a generalization of (2.1) given a *fixed* partitioning $a = [a_1^T, \ldots, a_v^T]^T \in \mathbb{Z}^n$. In this case the success rate in (2.30) is respectively the smallest and the largest when all $\lceil . \rfloor_i$ correspond to IR and to ILS, i.e. $\lceil x_i \rfloor_i = \arg \min_{z_i \in \mathbb{Z}^{n_i}} ||x_i - z_i||^2_{Q_{ii|I}}$. It follows that the bounds of (2.34) are a lower and an upper bound for *all* VIB estimators given that same partitioning. Hence, for a fixed v, we have $\text{VIB}_{\text{IR}} \leq \text{VIB} \leq \text{VIB}_{\text{ILS}}$, whereas for $v = 1$ the relation simply reduces to (2.1).

Differently, we now allow v to vary, provided that it satisfies $n = \sum_{i=1}^{v} n_i$, so we are processing all ambiguity components. It follows that for $v = 1$, the entire n-dimensional subset is considered and the relation in (2.1) holds, while for $v = n$, each block is a scalar, so both VIB_{IR} and VIB_{ILS} are ultimately equivalent to IB. For the case $v \in (1, n)$, I-estimators adopted in each block will define overall VIB performance. Starting with IR, which is poorer than IB, we thus have $\text{VIB}_{\text{IR}} \leq \text{IB}$, but also $\text{VIB}_{\text{IR}} \geq \text{IR}$, since conditioning improves precision and an improved precision also improves the success rate of rounding (Teunissen, 2007). Similarly, since ILS is better than or equal to IB, we have $\text{VIB}_{\text{ILS}} \geq \text{IB}$, but always $\text{VIB}_{\text{ILS}} \leq \text{ILS}$ since ILS has the largest possible success rate in the class of I-estimators.

We summarize the above performance ordering results in the following Lemma.

Lemma 6 (VIB-*performance ordering*)

1. For fixed v, i.e. a fixed partitioning $a = [a_1^T, \ldots, a_v^T]^T$, with $a_i \in \mathbb{Z}^{n_i}$, we have

$$\text{VIB}_{\text{IR}} \leq \text{VIB} \leq \text{VIB}_{\text{ILS}} \tag{2.36}$$

2. For any $v > 0$, satisfying $n = \sum_{i=1}^{v} n_i$, we have

$$\text{IR} \leq \text{VIB}_{\text{IR}} \leq \text{IB} \leq \text{VIB}_{\text{ILS}} \leq \text{ILS} \tag{2.37}$$

□

This result shows that an easy way to improve IR is to already define blocks of ambiguities, and include some conditioning. Similarly, also the performance of IB can be still improved, while avoiding the computational effort required for a full ILS solution. A simple way to graphically summarize the previous lemma is given in Fig. 2.4, where the different estimators are ordered in terms of their success rate.

In (2.34) we have already made use of the ADOP quantity in order to formulate an upper bound. Given that the ADOP is a geometric average of the sequential conditional standard deviations of ambiguities (Teunissen, 1997a), it gives an approximation to the

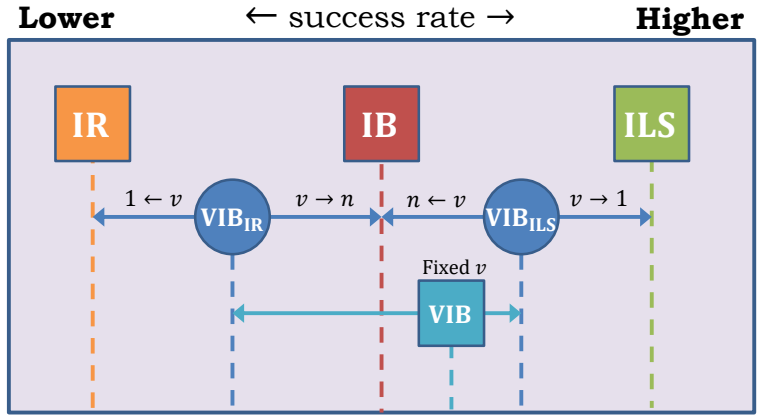


Figure 2.4: Graphical illustration of the success-rate ordering for well-known integer estimators, i.e. IR, IB and ILS. In addition, the VIB_{IR} and VIB_{ILS} are shown, when adopting only IR or ILS in each block, whose size is arbitrary. Lastly, with the cyan square a generic VIB is illustrated for a fixed block-size v, using arbitrarily selected I-estimators.

average precision of ambiguities, therefore it can be used for obtaining an approximation to the ILS success rate as

$$\mathsf{P}(\check{a}_{\text{ILS}} = a) \approx \left[2\Phi\left(\frac{1}{2\text{ADOP}}\right) - 1\right]^n \tag{2.38}$$

with $\text{ADOP} = |Q_{\hat{a}\hat{a}}|^{\frac{1}{2n}}$, similar to (2.33), but here referring to the full set of ambiguities. In a similar way, it can be used to provide an approximation to the success rate of the VIB_{ILS} estimator.

Lemma 7 (VIB_{ILS} *success-rate approximation*) Let $\check{a}_i \in \mathbb{Z}^{n_i}$ be the i-th integer vector of the VIB estimator defined as

$$\check{a}_i = \lceil \hat{a}_i(\check{a}_I) \rfloor_i = \arg \min_{z \in \mathbb{Z}^{n_i}} ||\hat{a}_i(\check{a}_I) - z||^2_{Q_{ii|I}} \tag{2.39}$$

Then

$$\mathsf{P}(\check{a}_{\text{VIB}_{\text{ILS}}} = a) \approx \prod_{i=1}^{v} \left[2\Phi\left(\frac{1}{2\text{ADOP}_i}\right) - 1\right]^{n_i} \tag{2.40}$$

with ADOP_i given by (2.33). □

This approximation becomes better the more decorrelated the ambiguities are. The error of approximation vanishes in case of a full decorrelation. Note that the approximation in (2.40) becomes equal to the success rate of IB when $n_i = 1, \forall i$, and thus when $v = n$.

Finally we mention that if the bounds and approximations are not considered sharp enough, that one can still resort to Monte Carlo simulations of the required VIB-probabilities using the approaches as described in (Verhagen et al., 2013).

2.4.4 Numerical illustrations

We now present two numerical examples, one low- and one high-dimensional, to illustrate and exemplify the performance orderings of Lemma 6.

Example 1 (*Three-dimensional ambiguity space*) We consider the performance of different I-estimators when the triangular decomposition of the float ambiguity vc-matrix is given in (2.27) and (2.28). The success rate is numerically computed using 10^8 samples, where the float ambiguity vectors are synthetically generated from a normal distribution with zero-mean and vc-matrix given in (2.26). Table 2.1 shows results for VIB_{IR} and VIB_{ILS}, along with results for the well-known IR, IB and ILS estimators. Here the VIB-estimators have their third ambiguity conditioned on the other two, which are fixed by IR or ILS. The exact result for IB, together with lower bounds for IR and VIB_{IR}, and an ADOP approximation for ILS and VIB_{ILS}, are also given, together with a reference to their defining equation.

Table 2.1: Numerical simulations considering a 3-dimensional problem defined by the variance-covariance matrix in (2.26)

Integer estimator	Ambiguity success rate	Bounds/Approx. Value	Eqs.
IR	63.24 %	$\geq$ 61.86 %	(2.54)
VIB_{IR}	64.18 %	$\geq$ 63.11 %	(2.34)
IB	66.04 %	$=$ 66.04 %	(2.11)
VIB_{ILS}	66.82 %	$\approx$ 66.10 %	(2.40)
ILS	66.99 %	$\approx$ 67.85 %	(2.38)

The success-rate values are aligned with the VIB performance ordering of Lemma 6 (cf. 2.37). The VIB_{ILS} shows indeed a quasi-optimal performance, i.e. ILS, without the need of an integer search over the entire domain. Moreover, the VIB_{IR} estimator is better than IR, while being suboptimal to the IB method, whose success rate is available analytically. $\diamond$

Example 2 (*VIB in regional network*) A 9-station, dual-frequency GPS network for PPP-RTK processing with known station coordinates is considered (see Fig. 2.5). The data is processed based on a Kalman filter each 30s using precise orbits for DOY 293 in 2020, along with mathematical models and the software platform described in (Odijk et al., 2017).

For the purpose of illustrating the performance orderings, we focus attention on full-ambiguity resolution (FAR) and on a moment in time when the best FAR success rate is extremely low, which in the present example is the case at 7:15 (UTC), marked by epoch number 870, when a new satellite, PRN03, rises and is being tracked by all stations of the network. The dimension of the ambiguity space is $n = 146$. Fig. 2.6 shows the success rates of the different integer estimators employed, with all ambiguities decorrelated using the LAMBDA transformation (Teunissen, 1995).

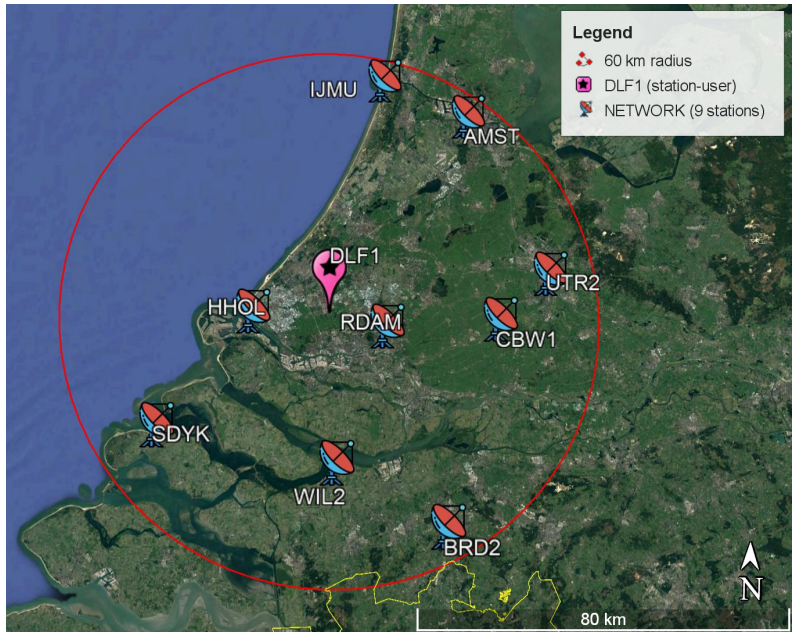


Figure 2.5: Example of regional network for GPS dual-frequency (L1-L2) data retrieved on DOY 293 (in 2020). The 9 stations are located within 60 km (circle) from a certain station-user.

In both VIB_{ILS} (top plot) and VIB_{IR} (bottom plot), we consider two different partitioning: $v = 2$ and $v = 20$. The VIB_{ILS} behaves optimally for $v = 2$, while smaller improvements are found for $v = 20$, i.e. when small subsets of 7-8 components are sequentially processed. We should observe that by decorrelating the ambiguities, an IB solution becomes quasi-optimal, thus a good lower bound to the ILS. The same holds also for the VIB_{ILS}.

In the bottom plot, the IR solution has a smaller success rate with respect to IB, but is extremely efficient. The VIB_{IR} solution for $v = 20$ approaches the IB success rate, with a large improvement over the IR solution. This second result is relevant since it empathizes how some conditioning operations can substantially improve robustness of the straightforward integer rounding. Its success rate will be always smaller than IB, but once the precision increases, this difference becomes negligible and a quasi-optimal integer solution can be obtained almost instantaneously. $\diamond$

2.5 Further VIB considerations

2.5.1 Ambiguity parametrization

In order to enhance the success-rate performance of VIB-estimation, considerations about the chosen ambiguity parametrization and their ordering are important. The

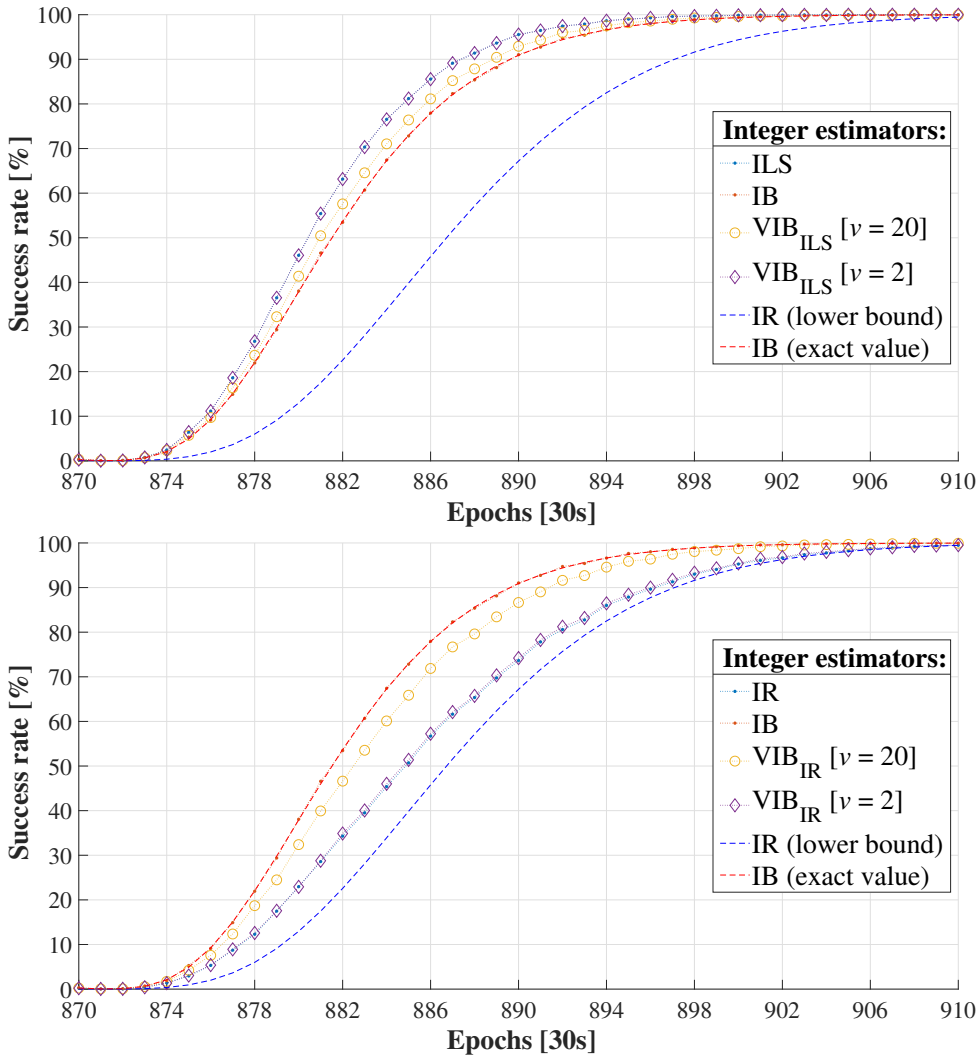


Figure 2.6: Network ambiguity success rates simulated for different I-estimators (DOY 293, 2020, 20,000 samples per epoch), with VIB_{ILS} referenced (top) and VIB_{IR} referenced (bottom)

general guide hereby is to aim forming blocks having the most precise ambiguities, followed by blocks that have the most precision gain from the conditioning, and this continues until the success rate drops below the required threshold. Although such can be achieved through the construction of the full-dimensional decorrelating Z-transformation (de Jonge and Tiberius, 1996), this process can in many GNSS applications be significantly aided by the *a priori* construction of proper ambiguity re-parametrizations. Here we show three such examples, with the first working on frequencies, the second on satellites and the third on antennas.

Example 3 (*Widelane-Narrowlane*) Widelaning is a popular multi-frequency technique of taking differences between ambiguities of different frequencies so as to obtain transformed ambiguities with a better precision (Hatch, 1989; Forsell et al., 1997; Teunissen, 1997c). When placed in the framework of VIB-estimation, the following steps and flexibility in the widelaning procedure can be recognized (here given for the dual-frequency case, but easily generalized to the multi-frequency case):

1. Apply the widelane transformation to get the float widelane and narrowlane ambiguity vectors, $\hat{a}_w$ and $\hat{a}_n$, respectively;
2. Integer estimate the widelane ambiguities as $\check{a}_w = \lceil \hat{a}_w \rfloor_1$, with $\lceil \cdot \rfloor_1$ being the integer map of IR or IB;
3. Float estimate the narrowlane ambiguities conditioned on the fixed widelane ambiguities as $\hat{a}_{n|w} = \hat{a}_n - Q_{\hat{a}_n \hat{a}_w} Q^{-1}_{\hat{a}_w \hat{a}_w} (\hat{a}_w - \check{a}_w)$;
4. Z-transform the narrowlane ambiguities to decorrelate, giving $\hat{z}_{n|w} = Z^T \hat{a}_{n|w}$ and its vc-matrix $Q_{\hat{z}_{n|w} \hat{z}_{n|w}}$;
5. Integer estimate the transformed narrowlane ambiguities as $\check{z}_n = \lceil \hat{z}_{n|w} \rfloor_2$, with $\lceil \cdot \rfloor_2$ being the integer map of IB or ILS.

The goal of the first step is to obtain ambiguities that are sufficiently precise so that simple estimators, like IR or IB, can achieve high-enough success rates in the second step. The goal of the third and fourth step is to benefit from the conditioning and decorrelation, before IB or ILS are applied. In case of ILS, the fourth step is aimed at improving the numerical efficiency of the ILS-computations, whereas for IB, it is aimed at improving the success rate. $\diamond$

Example 4 (*Multivariate geometry-free model*) Consider the dual-frequency, geometry-free model when tracking $v + 1$ satellites (Teunissen, 1997b). Due to its special structure, the $2v \times 2v$ vc-matrix of its DD float ambiguities is given as

$$Q_{\hat{a}\hat{a}} = (D_v^T D_v) \otimes Q \tag{2.41}$$

in which $D_v \in \mathbb{R}^{(v+1) \times v}$ represents the *between-satellite* differencing operator, $\otimes$ the Kronecker product and Q the vc-matrix of the single-differenced, dual-frequency ambiguities. It was shown in (Teunissen, 1997b), that by using the analytical LDL^T-decomposition of (2.41), the multivariate quadratic form of the ambiguities can be written in the form of a weighted sum of 2-dimensional quadratic forms,

$$||\hat{a} - a||^2_{Q_{\hat{a}\hat{a}}} = \sum_{i=1}^{v} \frac{i}{i+1} ||\hat{a}_i(a_I) - a_i||^2_Q \tag{2.42}$$

in which the $a_i \in \mathbb{Z}^2$, $i = 1, \ldots, v$, are, with respect to the reference satellite, the dual-frequency DD ambiguities of satellite i, $a_I = (a_1^T, \ldots, a_{i-1}^T)^T$, and $\hat{a}_i(a_I)$ is the conditional estimate

$$\hat{a}_i(a_I) = \hat{a}_i - \frac{1}{i} \cdot \sum_{j=1}^{i-1} (\hat{a}_j - a_j) \tag{2.43}$$

with $\hat{a}_i(a_I) = \hat{a}_1$ when $i = 1$. Instead of solving a high-dimensional ILS-problem by minimizing (2.42) over the $2v$-dimensional space of integers, the v two-dimensional quadratic forms of (2.42) are minimized in a sequential fashion. Hence, in this case the main VIB-estimation steps, for $i = 1, \ldots, v$, are:

1. Compute the conditional estimate $\hat{a}_{i|I} = \hat{a}_i(\check{a}_I)$;
2. ILS-estimate $\check{a}_i = \arg\min_{z \in \mathbb{Z}^2} ||\hat{a}_{i|I} - z||^2_Q$;
3. Update integer ambiguity vector: $\check{a}_{I+1} = (\check{a}_I^T, \check{a}_i^T)^T$.

Note, as each 2-dimensional ILS problem may still present highly correlated ambiguities, that a suitable 2×2 decorrelating Z-transform can be constructed, which then only has to be applied once to Q. $\diamond$

Example 5 (*Network array*) This example is taken from the concept of array-aided PPP introduced in (Teunissen, 2012). We assume to have an array with $r+1$ antennas, tracking $s+1$ GNSS satellites, on f frequencies. With $z_i \in \mathbb{Z}^{fs \times 1}$ being the integer vector of DD ambiguities of the ith baseline, the integer network ambiguity matrix $\mathcal{Z} = (z_1, \ldots, z_r) \in \mathbb{Z}^{fs \times r}$ can be formed of which the $fsr \times fsr$ vc-matrix of the float solution $\hat{\zeta} = \text{vec}(\hat{\mathcal{Z}})$ can be shown to read

$$Q_{\hat{\zeta}\hat{\zeta}} = D_r^T Q_r D_r \otimes N^{-1} \tag{2.44}$$

in which $D_r \in \mathbb{R}^{(r+1) \times r}$ represents the *between-antenna* differencing matrix, Q_r is a cofactor matrix by which the relative quality of the array-antennas can be modelled (i.e. $Q_r = I_{r+1}$ when all antennas have the same quality), and N is the $sf \times sf$ reduced normal matrix of the single-baseline ambiguities. Note, although (2.44) resembles the structure of (2.41), that (2.44) is a Kronecker product of two different types of matrices. As the receiver-antennae dependency is made explicit in matrix $D_r^T Q_r D_r$, differences in antenna-quality can be exploited in the VIB-conditioning. If we consider the case of 3 antennas, with $D_r = (e_2, -I_2)^T$, $e_2 = (1, 1)^T$, and $Q_r = \text{diag}(\sigma_1^2, \sigma_2^2, \sigma_3^2)$, then $\zeta = (z_1^T, z_2^T)^T \in \mathbb{Z}^{2fs \times 2fs}$ and (2.44) can be written as

$$Q_{\hat{\zeta}\hat{\zeta}} = \begin{bmatrix} (\sigma_1^2 + \sigma_2^2) \cdot N^{-1} & \sigma_1^2 \cdot N^{-1} \\ \sigma_1^2 \cdot N^{-1} & (\sigma_1^2 + \sigma_3^2) \cdot N^{-1} \end{bmatrix} \tag{2.45}$$

Would we now condition the ambiguities of the second baseline, $\hat{z}_2$, on those of the first baseline, $\hat{z}_1$, the resulting vc-matrix is obtained with the help of (2.45) as

$$Q_{\hat{z}_2\hat{z}_2|\hat{z}_1} = Q_{\hat{z}_2\hat{z}_2} \cdot \left(1 - \frac{1}{(1+\gamma_2)(1+\gamma_3)}\right) \tag{2.46}$$

where $\gamma_i = \sigma_i^2/\sigma_1^2$ is the variance-ratio between one of the two auxiliary antennas and the master. The precision of the ambiguities conditioned in a vectorial sense is improving by a factor 3/4 for antennas with the same precision. At the same time, the VIB improvement in the precision of the conditioned ambiguities $\hat{z}_{2|1}$ will actually be negligible if the second antenna (involved in the first baseline) has a very poor precision, a situation that thus should be avoided. $\diamond$

Next to frequencies, satellites and receivers, also other elements of the GNSS functional and stochastic model can in particular applications be exploited for the construction of a VIB-suitable ambiguity parametrization. Such can e.g. be driven by constellation, by satellite-elevation or by atmospheric impact. In a network, for instance, with very different baseline lengths, the ambiguities of the shorter baselines will generally be more precise and therefore candidates to be treated first (Blewitt, 1989). A similar consideration holds for ambiguities of high-elevation satellites, which are usually more precise than those of lower-elevation satellites.

2.5.2 Practical considerations

As VIB estimation is also aimed at reducing the computational complexities of integer estimation, it is important to recognize that in several of its computational steps a good use can be made of the, often readily available, Cholesky-decomposition of the system of normal equations. For instance, although in many of the expressions for ambiguity resolution the vc-matrix $Q_{\hat{a}\hat{a}}$ and/or its inverse $Q^{-1}_{\hat{a}\hat{a}}$ are needed, their explicit computation can often be avoided. Similarly, although the expressions of estimation often show the float ambiguity vector $\hat{a}$ explicitly, the computation of this full vector is not always needed, and this is particularly so in case of VIB-estimation.

To demonstrate this, let the partitioned system of normal equations, with the Cholesky decomposition for the normal matrix, be given as

$$\begin{bmatrix} C_{bb} & 0 \\ C_{ab} & C_{aa} \end{bmatrix} \begin{bmatrix} C_{bb} & 0 \\ C_{ab} & C_{aa} \end{bmatrix}^T \begin{bmatrix} \hat{b} \\ \hat{a} \end{bmatrix} = \begin{bmatrix} r_b \\ r_a \end{bmatrix} \tag{2.47}$$

Then the *reduced* system of normal equations for the ambiguities is given as $C_{aa}C_{aa}^T\hat{a} = \bar{r}_a$, from which it follows that the lower-triangular sub-matrix $C_{aa} \in \mathbb{R}^{n\times n}$ is directly related to the vc-matrix of the float ambiguity vector: $Q^{-1}_{\hat{a}\hat{a}} = C_{aa}C_{aa}^T$, and so $Q_{\hat{a}\hat{a}} = C_{aa}^{-T}C_{aa}^{-1}$. The matrix C_{aa}^{-1} (also lower triangular) can then be directly used in the decorrelation process prior to the actual integer estimation, thereby saving several matrix operations. We refer to de Jonge et al. (1996) for a more comprehensive description of these computational aspects, in particular describing further advantages relatively to the computation of the decorrelating Z-transformation. Here we also point to the VIB-flexibility as far as the Z-decorrelation is concerned. One can apply such decorrelation to all ambiguities, or one can restrict the Z-decorrelation to only when the mappings $\lceil . \rfloor_i : \mathbb{R}^{n_i} \mapsto \mathbb{Z}^{n_i}$ need to be computed.

In case of VIB, one can take another advantage of the triangular structure of C_{aa} by avoiding, through a proper ordering of the ambiguities, the explicit computation of their float values. With the triangular matrix C_{aa} partitioned as

$$C_{aa} = \begin{bmatrix} C_{11} & 0 \\ C_{12} & C_{22} \end{bmatrix} \tag{2.48}$$

it directly follows from the reduced normal equations $C_{aa}C_{aa}^T\hat{a} = \bar{r}_a$ that the conditional least-squares float solution of a_1, when conditioned on a_2, is given as

$\hat{a}_{1|2} = C_{11}^{-T}[C_{11}^{-1}\bar{r}_1 - C_{12}^T a_2]$. This shows that the computation of the conditional float solution $\hat{a}_{1|2}$ does not require the explicit computation of the float solution $\hat{a}_1$ and that it can be done efficiently by solving triangular systems of equations. Hence, when in case of VIB, the v-block partitioning is known, the ambiguities can be ordered accordingly to take advantage of this numerical gain.

As discussed in previous sections, a vectorial formulation enables also ad hoc parametrizations of the ambiguity components, where structure of a certain problem can be fully exploited. Given that the mathematical relations presented in this contribution are generalized for any different parametrization, it is indeed possible to also makes use of different decorrelation approaches (Jazaeri et al., 2014), which could further enhance efficiency of VIB-based strategies. The large variety of applications for such a vectorial formulation makes a comprehensive discussion about performances not practical within the scope of this work, therefore a subject of future researches. However, to further emphasize the available flexibility, we briefly highlight the possible extensions to other estimators.

2.5.3 Extensions to other classes of estimators

The flexibility of the VIB-formulation does not restrict it to the class of integer estimators only. It could include estimators from the IA-class (Teunissen, 2003a) or IE-class (Teunissen, 2002) as well. For each block, for instance, one can include IA-estimators having aperture pull-in regions that are particularly accommodated to the integer estimator $\lceil\cdot\rfloor_i : \mathbb{R}^{n_i} \mapsto \mathbb{Z}^{n_i}$ of that block. Such can then be used to include ambiguity-validation for each block, thereby providing flexibility and options to skip blocks when block-validation fails.

Also IE-estimators can be given a place in the VIB-framework, for instance when it turns out that the success rate drops below the required threshold when an additional block would be fixed. Instead of outputting the conditional float solution of the remaining ambiguities, one could then still apply, using their conditional vc-matrix, best integer equivariant estimation to these ambiguities to improve upon their mean squared errors. As such, and with the various options available, the concept of partial ambiguity resolution, introduced in (Teunissen et al., 1999), is generalized to the VIB-domain.

2.6 Summary and concluding remarks

In this contribution we introduced the concept of vectorial integer bootstrapping (VIB) as a generalization of the popular, but scalar, integer bootstrapping. As with integer bootstrapping, VIB-estimation is characterized by two alternating operations that are sequentially applied, conditioning and integer mapping. It is due to the vectorial formulation of these two principles, that the VIB-concept creates such flexibility in designing one's integer estimators. Many new integer estimators can be formulated, in particular when balancing computational simplicity against success-rate performance.

We presented the probabilistic properties of the VIB-estimators, with a special emphasis on their probability of correct integer estimation and the formulation of easy-to-compute lower bounds and upper bounds of their success rates. We provided a new ordering in the success-rate performance of various different VIB-estimators, together with corresponding numerical illustrations.

In order to enhance the success-rate performance of VIB-estimation, considerations about the chosen ambiguity parametrization and their ordering are important. The general guide hereby is to aim forming blocks having the most precise ambiguities, followed by blocks that have the most precision gain from the conditioning. Although such can be achieved through the construction of the full-dimensional decorrelating Z-transformation, it was shown by means of analytical examples that in many GNSS applications such can be significantly aided through the a-priori construction of proper ambiguity re-parametrizations.

We also discussed further considerations when implementing VIB. As it is aimed at reducing the computational complexities of integer estimation, we described how at several of its computational steps a good use can be made of the, often readily available, Cholesky-decomposition of the system of normal equations. This not only concerns the computation of the ambiguity vc-matrix, but also of the float solution itself. Finally, we indicated that the flexible VIB-concept lends itself to further extensions, in particular in combination with estimators from the IA- and IE-class.

2.7 Appendix

Proof of Lemma 3 (VIB *success rate*) We use the transformation of integral formula

$$\int_{\mathcal{R}} f(y)dy = \int_{T^{-1}(\mathcal{R})} f(T(x)) \cdot |\partial_x T(x)| \, dx \tag{2.49}$$

with $|\partial_x T(x)|$ being the determinant (in absolute value) of the Jacobian matrix of partial derivatives. Moreover, in our case we have $\mathcal{R} = \mathcal{P}_{a,\text{VIB}}$ and $f : \mathbb{R}^n \to \mathbb{R}$ as

$$f(y) = \frac{1}{\sqrt{|2\pi Q_{\hat{a}\hat{a}}|}} \exp\left\{-\tfrac{1}{2}||y-a||^2_{Q_{\hat{a}\hat{a}}}\right\} \tag{2.50}$$

As transformation $y = T(x)$ we choose

$$y_i = x_i + \sum_{j=1}^{i-1} Q_{ij|J} Q^{-1}_{jj|J}(x_j - a_j) \ , \ i = 1, \ldots, v \tag{2.51}$$

which leads for the transformed pull-in region to

$$T^{-1}(\mathcal{R}) = \{x \in \mathbb{R}^n \mid \mathcal{I}(x-a) = 0\} \tag{2.52}$$

with $\mathcal{I}(x) = (\lceil x_1 \rfloor_1^T, \ldots, \lceil x_v \rfloor_v^T)^T$, and for the transformed integrand we can write

$$f(T(x)) \cdot |\partial_x T(x)| = \prod_{i=1}^{v} \frac{\exp\{-\frac{1}{2}||x_i - a_i||^2_{Q_{ii|I}}\}}{\sqrt{|2\pi Q_{ii|I}|}} \tag{2.53}$$

since $|\partial_x T(x)| = 1$. Substitution of (2.52)(2.53) into (2.49) proves the result. □

Proof Lemma 5 (*Success-rate bounds for* $\mathrm{VIB_{IR}}$ *and* $\mathrm{VIB_{ILS}}$) The lower bound for $\mathrm{VIB_{IR}}$ follows from the IR lower bound given in (Teunissen, 1998), such that

$$\prod_{i=1}^{n} \left[2\Phi\left(\frac{1}{2\sigma_i}\right) - 1 \right] \leq \mathrm{P}(\check{a}_{\mathrm{IR}} = a) \tag{2.54}$$

Note that this lower bound is evaluated for the *unconditional* ambiguity standard deviations. The upper bound for $\mathrm{VIB_{ILS}}$ follows from the respective upper bound for an ILS estimator, also given as

$$\mathrm{P}(\check{a}_{\mathrm{ILS}} = a) \leq \mathrm{P}\left(\chi^2(n,0) \leq \frac{c_n}{\mathrm{ADOP}^2}\right) \tag{2.55}$$

with $\mathrm{ADOP} = |Q_{\hat{a}\hat{a}}|^{\frac{1}{2n}}$, and c_n given in (2.35). For a proof of this upper bound, see (Teunissen, 2000). □

References

Blewitt G (1989) Carrier phase ambiguity resolution for the global positioning system applied to geodetic baselines up to 2000 km. Journal of Geophysical Research: Solid Earth 94(B8):10187–10203

Borre K, Strang G (2012) Algorithms for Global Positioning. Wellesley-Cambridge Press

Brack A, Henkel P, Gunther C (2014) Sequential best integer equivariant estimation for GNSS. Navigation 61(2):149–158

Forsell B, Martin Neira M, Harris R (1997) Carrier phase ambiguity resolution in GNSS-2. Proc ION GPS 1997 pp 1727–1736

Hatch R (1989) Ambiguity resolution in the fast lane. Proc ION GPS89 pp 45–50

Hofmann-Wellenhof B, Lichtenegger H, Wasle E (2008) GNSS–global navigation satellite systems: GPS, GLONASS, Galileo, and more. Springer-Verlag Wien

Hou P, Zhang B, Liu T (2020) Integer-estimable GLONASS FDMA model as applied to Kalman-filter-based short- to long-baseline RTK positioning. GPS Solutions 24:93

Jazaeri S, Amiri-Simkooei A, Sharifi MA (2014) On lattice reduction algorithms for solving weighted integer least squares problems: comparative study. GPS Solutions 18(1):105–114, DOI 10.1007/s10291-013-0314-z

de Jonge P, Tiberius CCJM (1996) The LAMBDA method for integer ambiguity estimation: implementation aspects. LGR-Series Publications of the Delft Geodetic Computing Centre 12

de Jonge P, Tiberius CCJM, Teunissen PJG (1996) Computational aspects of the LAMBDA method for GPS ambiguity resolution. In: Proceedings of the 9th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 1996), Institute of Navigation, vol 9, pp 935–944

Kaplan ED, Hegarty C (2006) Understanding GPS, Principles and Applications, 2nd edition. Artech House

Leick A, Rapoport L, Tatarnikov D (2015) GPS Satellite Surveying, 4th edition. Wiley

Misra P, Enge P (2006) Global Positioning System : Signals, Measurements, and Performance, 2nd edition. Ganga-Jamuna

Morton J, van Diggelen F, Spilker JJ, Parkinson BP, Lo S, Gao G (2021) Position, Navigation, and Timing Technologies in the 21st Century: Integrated Satellite Navigation, Sensor Systems, and Civil Applications. Wiley IEEE Press

Odijk D, Khodabandeh A, Nadarajah N, Choudhury M, Zhang B, Li W, Teunissen PJG (2017) PPP-RTK by means of S-system theory: Australian network and user demonstration. Journal of Spatial Science 62(1):3–27

Odolinski R, Teunissen PJG (2020) Best integer equivariant estimation: performance analysis using real data collected by low-cost, single- and dual-frequency, multi-GNSS receivers for short- to long-baseline rtk positioning. Journal of Geodesy 94:91

Psychas D, Verhagen S (2020) Real-time PPP-RTK performance analysis using ionospheric corrections from multi-scale network configurations. Sensors 20(11:3012)

Strang G, Borre K (1997) Linear Algebra, Geodesy, and GPS. Wellesley-Cambridge Press

Teunissen PJG (1995) The least-squares ambiguity decorrelation adjustment: a method for fast GPS integer ambiguity estimation. Journal of Geodesy 70:65–82

Teunissen PJG (1997a) A canonical theory for short GPS baselines. Part III: the geometry of the ambiguity search space. Journal of Geodesy 71(8):486–501

Teunissen PJG (1997b) GPS double difference statistics: with and without using satellite geometry. Journal of Geodesy 71(3):137–148, DOI 10.1007/s001900050126

Teunissen PJG (1997c) On the GPS widelane and its decorrelating property. Journal of Geodesy 71(9):577–587, DOI 10.1007/s001900050126

Teunissen PJG (1998) Success probability of integer GPS ambiguity rounding and bootstrapping. Journal of Geodesy 72(10):606–612

Teunissen PJG (1999) An optimality property of the integer least-squares estimator. Journal of Geodesy 73(11):587–593

Teunissen PJG (2000) ADOP based upper bounds for the bootstrapped and the least squares ambiguity success. Artificial Satellites 35(4):171–179

Teunissen PJG (2002) A new class of GNSS ambiguity estimators. Artificial Satellites 37(4):111–120

Teunissen PJG (2003a) Integer aperture GNSS ambiguity resolution. Artificial Satellites 38(3):79–88

Teunissen PJG (2003b) Theory of integer equivariant estimation with application to GNSS. Journal of Geodesy 77:402–410

Teunissen PJG (2007) Influence of ambiguity precision on the success rate of GNSS integer ambiguity bootstrapping. Journal of Geodesy 81(5):351–358

Teunissen PJG (2012) A-PPP: Array-aided precise point positioning with global navigation satellite systems. IEEE Transactions on Signal Processing 60(6):2870–2881

Teunissen PJG, Kleusberg A (1998) GPS for Geodesy, 2nd edition. Springer, Berlin, Heidelberg

Teunissen PJG, Montenbruck O (2017) Handbook of Global Navigation Satellite Systems. Springer Verlag, Berlin

Teunissen PJG, Joosten P, Tiberius CCJM (1999) Geometry-free ambiguity success rates in case of partial fixing. Proc ION NTM 1999, San Diego pp 201–207

Verhagen S (2005) The GNSS integer ambiguities: Estimation and validation. PhD thesis, Delft University of Technology

Verhagen S, Li B, Teunissen PJG (2013) Ps-LAMBDA: ambiguity success rate evaluation software for interferometric applications. Computers & Geosciences 54:361–376

Wang K, Khodabandeh A, Teunissen PJG (2018) Five-frequency Galileo long-baseline ambiguity resolution with multipath mitigation. GPS Solutions 22(3):1–14

Zaminpardaz S, Wang K, Teunissen PJG (2018) Australia-first high-precision positioning results with new Japanese QZSS regional satellite system. GPS Solutions 22(4:101)

3

On VIB implementations in the estimation of satellite orbits/clocks based on small global networks

The estimation of satellite orbits and clocks plays a central role in different Global Navigation Satellite System (GNSS) domains, e.g. precise positioning or time transfer. Such products can be computed in the process of Orbit Determination and Time Synchronization (ODTS), which relies on a network of ground-based stations, well distributed around the globe. The mm-level precision of carrier-phase measurements is exploited in this network estimation following a correct resolution of their ambiguities. For several stations, satellites and/or signals, thousands of ambiguities have to be processed, which means having to deal with high-dimensional ambiguity resolution (HDAR) problems. In this research work, we firstly account for the impact of ambiguity resolution in a varying network size, based on GPS-only, Galileo-only and GPS+Galileo configurations. Using 25 or more stations, the accuracy (1D RMS orbital errors) of *fixed* solutions reaches a plateau at 1-2 cm. Hence, we focus on a small global network of 14 stations, where the model strength decreases, so does the reliability of the ambiguity fixing process and advantages over a *float* solution might become less evident. In order to allow reliable HDAR, two implementations of the Vectorial Integer Bootstrapping estimator are presented and evaluated with respect to their scalar counterpart. Finally, it is shown how the proposed fixing processes are more robust, still very efficient, and on certain days they provide a large improvement to satellite products. The orbital results are ultimately validated by considering the satellite midnight discontinuity errors over a 3-month period in 2019.

This chapter has been published as: Massarweh, L., Strasser, S., and Mayer-Gürr, T. (2021). *On vectorial integer bootstrapping implementations in the estimation of satellite orbits and clocks based on small global networks.* In Advances in Space Research, 68(11), 4303-4320. https://doi.org/10.1016/j.asr.2021.09.023

3.1 Introduction

The generation of satellite orbit and clock information represents an essential element for any Regional/Global Navigation Satellite System (RNSS/GNSS) and is generally based on the use of code and phase observations from a ground-based network of station receivers. The so-called process of Orbit Determination and Time Synchronization (ODTS) makes the best use of state-of-the-art knowledge on orbital dynamics for all tracked satellites (Montenbruck and Steigenberger, 2020). This a priori information can be consequently improved by means of an accurate functional and stochastic modeling of such measurements, ideally consistent with the models later adopted on the user side.

The contribution of carrier-phase measurements, very precise but ambiguous, is a key aspect in the network estimation, whether a global or a regional one. In fact, to exploit their mm-level precision, the integer ambiguities need to be correctly resolved, which is not an easy task when dealing with global networks (Ge et al., 2006). Given denser networks, many satellites from different GNSS constellations and/or multi-frequency signals, the number of ambiguity components grows rapidly, along with the complexity of the integer ambiguity resolution (IAR). This leads to the necessity for a balanced trade-off between *efficiency* and *robustness* of the IAR process, for example in support of a low-latency generation of precise satellite orbits and clocks products.

Simple and intuitive strategies can be used (Blewitt, 1989; Cocard and Geiger, 1992), e.g. well-known widelaning approaches, but sophisticated ones are available. For instance, the Least-squares AMBiguity Decorrelation Adjustment (LAMBDA) method, introduced in Teunissen (1995), which provides an effective solution to the IAR. The LAMBDA approach enables the adoption of a wide range of possible estimators, suitable for both low and high accuracy applications, and it has been proved to be relatively efficient up to a few hundred ambiguity components (Li and Teunissen, 2011). Expected performances for precise long-baseline positioning using both GPS and Galileo are presented in Odijk et al. (2014) based on a formal analysis, for instance in terms of predicted success rate. The advantages when considering GPS+Galileo have been further investigated (Nardo et al., 2016), where different long baseline scenarios (ground- and space-based) are simulated and examined.

However, nowadays, we might deal with problems of a much larger dimensionality (Chen et al., 2014), where thousands of ambiguity components are involved, and the overall complexity grows exponentially in such a “dimensional curse” (Verhagen et al., 2012b). This dimensionality aspect might lead to a computational bottleneck (Jazaeri et al., 2012), and the high-dimensional ambiguity resolution (hereinafter HDAR) problems are less likely solvable in a reliable way and in short computational times.

In this work, we consider the Vectorial Integer Bootstrapping (VIB) estimator described in Teunissen et al. (2021). The flexibility of such a VIB formulation allows for an arbitrary partitioning of the original n-dimensional ambiguity problem, along with a suitable choice of the integer estimator in use. Two (possible) implementations are described: firstly, a straightforward Cascade AR (CascAR) algorithm and secondly,

a blocked search approach in use by Graz University of Technology (TUG). Both the two algorithms are shown to be efficiently constructed starting from the same triangular decomposition, while following LAMBDA decorrelation of ambiguities.

Using the ODTS strategy and software developed at Graz University of Technology (Strasser et al., 2019), we analyze different network configurations in order to further define the impact of ambiguity resolution with respect to the network size. The latter is assessed by looking at the accuracy of satellite products. We focus on a small global network, where the correct ambiguity resolution shows to provide the most benefits, e.g. in terms of orbital errors. The two VIB implementations are validated over a 3-month period in 2019, using GPS+Galileo real data from a network of 14 well-distributed stations. Both these VIB-based approaches show the great convenience of this easy-to-implement and flexible formulation, while also leaving room for improvement. In fact, more advanced algorithms might be designed ad hoc in several GNSS domains, not limited to this network case.

In Section 3.2, the two algorithms are introduced as examples of implementation for the VIB estimator. In Section 3.3, the network estimation strategy for satellite orbits and clocks is delineated, along with configurations and a few scenarios considered in this work. The main results are presented and analyzed within Section 3.4, focusing on a small global network and also briefly comparing the two algorithms. In Section 3.5, we perform a validation over a 3-month period, followed by a short discussion on this vectorial formulation. Lastly, in Section 3.6, the overall conclusions are given.

3.2 Integer ambiguity resolution methodology

We start from a linear(ized) mixed-integer GNSS model estimation (Leick et al., 2015), given the vector of code and phase observations as $y \in \mathbb{R}^m$ with its positive-definite variance-covariance (vc-)matrix as $Q_{yy} \in \mathbb{R}^{m \times m}$, such that

$$\mathrm{E}\{y\} = Aa + Bb, \quad \mathrm{D}\{y\} = Q_{yy} \tag{3.1}$$

with $\mathrm{E}\{\cdot\}$ and $\mathrm{D}\{\cdot\}$ being the expectation and dispersion operators, respectively. The matrix $(A, B) \in \mathbb{R}^{m \times (n+p)}$ denotes a full-rank design matrix, with $a \in \mathbb{Z}^n$ as vector of carrier-phase integer ambiguity and $b \in \mathbb{R}^p$ as vector of real-valued parameters. The latter refer, without loss of generality, to estimable parameters in the ODTS process, for example satellite orbit and clock products, as later discussed in Section 3.3.

By means of an orthogonal decomposition (Teunissen, 1993), we can separate the quadratic objective function (with integerness constraints) into a sum of three squares as

$$||y - Aa - Bb||^2_{Q_{yy}} = ||\hat{e}||^2_{Q_{yy}} + ||\hat{a} - a||^2_{Q_{\hat{a}\hat{a}}} + ||\hat{b}(a) - b||^2_{Q_{\hat{b}(a)}} \tag{3.2}$$

where $||\cdot||^2_Q = (\cdot)^T Q^{-1} (\cdot)$ refers to a squared Mahalanobis distance (Mahalanobis, 1936), i.e. a weighted norm in the metric given by Q^{-1}. This previous decomposition leads to a subsequent three-step minimization of the original objective function.

In the first step, $\hat{e} \in \mathbb{R}^m$ is the residual vector of a least-squares solution that neglects the integerness constrains on ambiguities. These estimated (float) ambiguities $\hat{a} \in \mathbb{R}^n$, and their vc-matrix $Q_{\hat{a}\hat{a}} \in \mathbb{R}^{n\times n}$, are therefore used in a second step to solve an Integer Least Squares (ILS) problem, e.g. considering a many-to-one map $\mathcal{I} : \mathbb{R}^n \to \mathbb{Z}^n$, such that their integer-fixed estimate is $\check{a} = \mathcal{I}(\hat{a})$. Given that no constrains are taken into account for b, a fixed solution to the third step is easily found by considering $b = b(\check{a}) \overset{\text{def}}{=\!=} b - Q_{\hat{b}\hat{a}}Q_{\hat{a}\hat{a}}^{-1}(\hat{a} - \check{a})$ and, for $Q_{\check{a}\check{a}} \approx 0$ (Verhagen, 2005), we can safely assume $Q_{\check{b}\check{b}} \approx Q_{\hat{b}\hat{b}} - Q_{\hat{b}\hat{a}}Q_{\hat{a}\hat{a}}^{-1}Q_{\hat{a}\hat{b}}$.

3

A focal point in this three steps process becomes the minimization of $||\hat{a} - a||^2_{Q_{\hat{a}\hat{a}}}$, which can still be tackled by different classes of estimators, i.e. Integer (I) estimators (Teunissen, 1999), Integer Aperture (IA) estimators (Teunissen, 2003a) or Integer Equivariant (IE) estimators (Teunissen, 2002). These three classes are proper subsets of one another, thence $I \subset IA \subset IE$, with the I- and IE-class being the smallest and largest, respectively. Further mathematical details are summarized and extensively discussed in Teunissen (2003b).

3.2.1 Review of Vectorial Integer Bootstrapping estimators

The Vectorial Integer Bootstrapping (VIB) method, defined in Teunissen et al. (2021), hinges on a further orthogonal decomposition of the second term in Eq.(3.2). In fact, it is possible to define an *arbitrary* partitioning of the initial vector $a \in \mathbb{Z}^n$, i.e. $a_1 \in \mathbb{Z}^{n_1}, a_2 \in \mathbb{Z}^{n_2}$, where all ambiguity components are still involved, so $n = n_1 + n_2$. It follows that

$$\min_{a\in\mathbb{Z}^n} ||\hat{a} - a||^2_{Q_{\hat{a}\hat{a}}} = \min_{a_1\in\mathbb{Z}^{n_1}, a_2\in\mathbb{Z}^{n_2}} \left(||\hat{a}_1(a_2) - a_1||^2_{Q_{11|22}} + ||\hat{a}_2 - a_2||^2_{Q_{22}}\right) \tag{3.3a}$$

$$\approx \min_{a_1\in\mathbb{Z}^{n_1}} \left(||\hat{a}_1(a_2) - a_1||^2_{Q_{11|22}}\right) + \min_{a_2\in\mathbb{Z}^{n_2}} \left(||\hat{a}_2 - a_2||^2_{Q_{22}}\right) \tag{3.3b}$$

for $Q_{11|22} = Q_{11} - Q_{12}Q_{22}^{-1}Q_{21}$ as Schur complement (Zhang, 2006) of block Q_{22} in $Q_{\hat{a}\hat{a}}$. The latter follows from a conformable blocks' partitioning, i.e. $Q_{12} \in \mathbb{R}^{n_1\times n_2}$, such as

$$Q_{\hat{a}\hat{a}} = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{12}^T & Q_{22} \end{bmatrix} \tag{3.4}$$

with $\hat{a}_1(a_2) = \hat{a}_1 - Q_{12}Q_{22}^{-1}(\hat{a}_2 - a_2)$ here referring to the float ambiguity subset 1 conditioned on 2. In the VIB approach, the first term in parenthesis for Eq.(3.3a) is then minimized only accounting for $a_1 \in \mathbb{Z}^{n_1}$, thus assuming in Eq.(3.3b) that a_2 is given from the second minimization, i.e. $\check{a}_2 = \mathcal{I}_2(\hat{a}_2)$ for an admissible integer mapping $\mathcal{I}_2 : \mathbb{R}^{n_2} \to \mathbb{Z}^{n_2}$.

The VIB solution is consequently suboptimal, and its success rate depends upon the ambiguity parametrization, as it does for its scalar counterpart (Teunissen, 1998). Still, the integer search is now performed over two smaller subsets rather than over a large domain, whose complexity increases exponentially with the n dimensionality (Brack, 2019). A similar principle is, nonetheless, also known in the *scalar* Integer

Bootstrapping (IB) method, where the number of blocks m is equal to n, so having $n_i = 1, \forall i = 1, \ldots, m$. On the other hand, for $m = 1$ we have $n_1 = n$, and thence the overall ILS (optimal) solution is computed.

It is important to observe that each block is always conditioned on the previous ones, preferably going from the most to the least precise subset. This conditioning should not be neglected since it is a key aspect for improving the VIB success rate (Teunissen et al., 2021), and for the same parametrization it can lead to a solution that is closer to optimality than in the scalar IB case. In order to maximize the robustness of the fixing within each subset, all blocks are sequentially processed by means of an ILS estimator as discussed in the following sections.

3.2.2 Cascade Ambiguity Resolution (CascAR) algorithm

The CascAR algorithm is based on a quite general implementation of VIB estimators, and it allows having an efficient characterization of each subset, since it requires only a single $L^T DL$-decomposition of the (float) ambiguity vc-matrix. It follows as $Q_{\hat{a}\hat{a}} = L_{\hat{a}}^T D_{\hat{a}} L_{\hat{a}}$, where $L_{\hat{a}}, D_{\hat{a}} \in \mathbb{R}^{n \times n}$ are a lower unitriangular and a diagonal matrix, respectively. This triangular decomposition has a statistical interpretation, e.g. see Teunissen (1995), and it leads to the desired sum-of-blocks structure discussed in Section 3.2.1.

Given an arbitrary partitioning with m blocks, for $m \in [1, n]$, we can write

$$||\hat{a} - a||^2_{Q_{\hat{a}\hat{a}}} = \sum_{i=1}^{i=m} ||\hat{a}_{i|J} - a_i||^2_{Q_{ii|J}}, \quad J = \{i+1, \ldots, m\} \tag{3.5}$$

with the (conditioned) ambiguity subsets $\hat{a}_{i|J} \in \mathbb{R}^{n_i}$, for $n = n_1 + \ldots + n_m$. Furthermore, we set $\hat{a}_{m|J} = \hat{a}_m$, being the unconditioned block (i.e. firstly processed). In Teunissen et al. (2021), a block-decomposition is shown to define subsets, thus implying a certain selection for the partitioning, i.e. values n_i. Nonetheless, in a very similar way, the metric used in each sub-problem can be retrieved directly as $Q_{ii|J} = L_{ii}^T D_{ii} L_{ii}$, where a conformable (in size) partitioning of $L_{\hat{a}}, D_{\hat{a}}$ has been adopted. For the sake of clarity, additional mathematical details are given in Appendix.

We can describe the CascAR algorithm (see Fig. 3.1) by assuming that ambiguities are firstly decorrelated with the LAMBDA software (Verhagen et al., 2012a) by means of a Z-transformation, for $Z \in \mathbb{Z}^{n \times n}$ (unimodular), such that $\hat{z} = Z^T \hat{a}$ and $Q_{\hat{z}\hat{z}} = Z^T Q_{\hat{a}\hat{a}} Z$. This "pre-processing" step enhances the success rate for our VIB solution, improves the integer search process, and it also assures that ambiguity components are sorted based on their precision. The latter is an important element since it is convenient to firstly fix very precise blocks (i.e. with a high success rate), so to later condition the remaining ones on these more reliably fixed blocks.

In order to assure a full consistency with LAMBDA routines, the most precise components are set within the last block $i = m$, which is where we start. We then continue the cascade (conditioning) process to $i = m-1, m-2, \ldots$, till we reach $i = 1$.

The CascAR algorithm takes as inputs the triangular decomposition for a decorrelated ambiguity vc-matrix, i.e. $Q_{\hat{z}\hat{z}} = L_{\hat{z}}^T D_{\hat{z}} L_{\hat{z}}$, and a float vector $\hat{z}$ relative to the full set. Once the first partitioning has been defined by n_m, we initialize the aforementioned inputs with "*" and we divide the problem in two subsets I and II.

Hence, we get $\hat{z}_I \in \mathbb{R}^{n_I}$ and $\hat{z}_{II} \in \mathbb{R}^{n_{II}}$, where $n_I = n_1 + \ldots + n_{m-1}$ and $n_{II} = n_m$. This second block (presumed to be more precise) is processed in order to obtain an ILS solution $\check{z}_{II} \in \mathbb{Z}^{n_{II}}$, later adopted to condition the remaining ambiguities. We should observe that this conditioning operation takes place directly by means of conformable blocks in the matrix $L_{\hat{z}}$, so directly using $L_Q \in \mathbb{R}^{n_{II} \times n_I}$ and $L_{II} \in \mathbb{R}^{n_{II} \times n_{II}}$ (see details in Appendix).

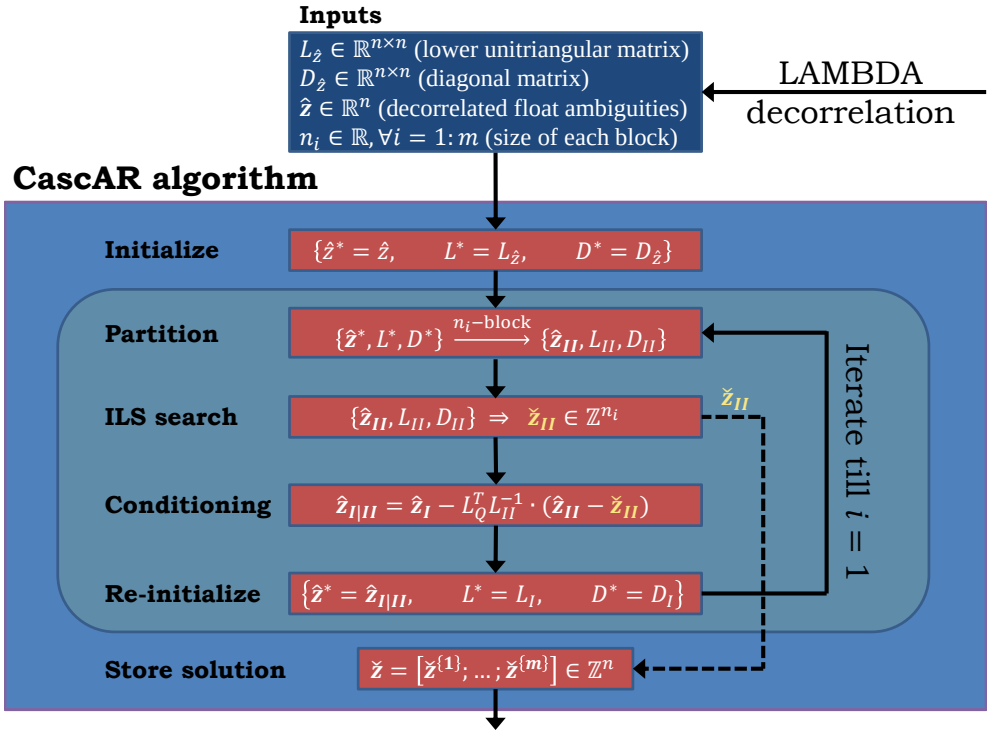


Figure 3.1: The CascAR algorithm is shown, where the initial ILS problem can be partitioned into m blocks, starting from ambiguity components that have been previously decorrelated in LAMBDA.

At the end of this first iteration, the block I is re-initialized with "*" and we re-iterate the process till $n_{II} = n_1$, while $n_I = 0$. During each i-th iteration, the integer-fixed solution $\check{z}^{\{i\}} \in \mathbb{Z}^{n_i}$ of the respective subset is stored in a single (column) vector $\check{z}_{\text{VIB}}$,

given as

$$\check{z}_{\mathrm{VIB}} = \begin{bmatrix} \check{z}^{\{1\}} \\ \vdots \\ \check{z}^{\{i\}} \\ \vdots \\ \check{z}^{\{m\}} \end{bmatrix} \in \mathbb{Z}^n \tag{3.6}$$

that still represents an integer solution to the problem as seen in Eq.(3.3b). The selection over number of blocks m and their respective dimensionality n_i is discussed in Section 3.3.2.1.

3.2.3 TUG's "blocked search" IAR method

When looking at vectorial approaches, there are several examples available in literature as highlighted in Teunissen et al. (2021). Another possible approach that can be placed within the family of VIB estimators is the so-called "blocked search" algorithm that has been firstly described in Strasser et al. (2019). It shows quite some similarities with the CascAR implementation, but with an alteration of the cascade processing.

This method considers overlapping blocks, de facto re-processing several components twice but as part of different subsets. The driving motivation is that the optimal solution for a certain ILS problem is unique, i.e. a definite n-dimensional set of integer ambiguities, and therefore fixed solutions for overlapping blocks should have the same integer components. If two adjacent blocks lead to different integer values for their overlapping components, then the integer search is performed over their joint subsets.

We should carefully consider that while this approach is not formally defined from a mathematical point of view, it still presents a quite interesting variation of the VIB approach and is therefore considered in this investigation. In Fig. 3.2 a schematic view of the TUG method is given for a simplified case, as a matter of example, based on a blocked partitioning of an ILS problem with $n = 400$, where each block involves 100 ambiguity components.

The TUG algorithm starts processing each 100-dimensional block in a main cascade and overlapping blocks (of same dimensionality) in a secondary one. When all integer-fixed components in common sub-blocks (i.e. here 50 ambiguities) are the same we obtain the respective fixed solution (in green color) for that portion of the block. When a different solution (in red color) is found between this main and secondary cascade, then the search is performed in the joint subset that contains 150 ambiguities. The solution of this larger integer search provides the 100 fixed components, including those 50 that differed originally, while the last 50 solution components are not yet accepted. In fact, these last 50 components overlap with a new block in the main cascade (i.e. from $\hat{z}_{200}$ till $\hat{z}_{101}$) that has not been processed yet. Once this block has been processed and the overlapping integer components match, we continue the process; otherwise the same "expansion and search" strategy is performed as before.

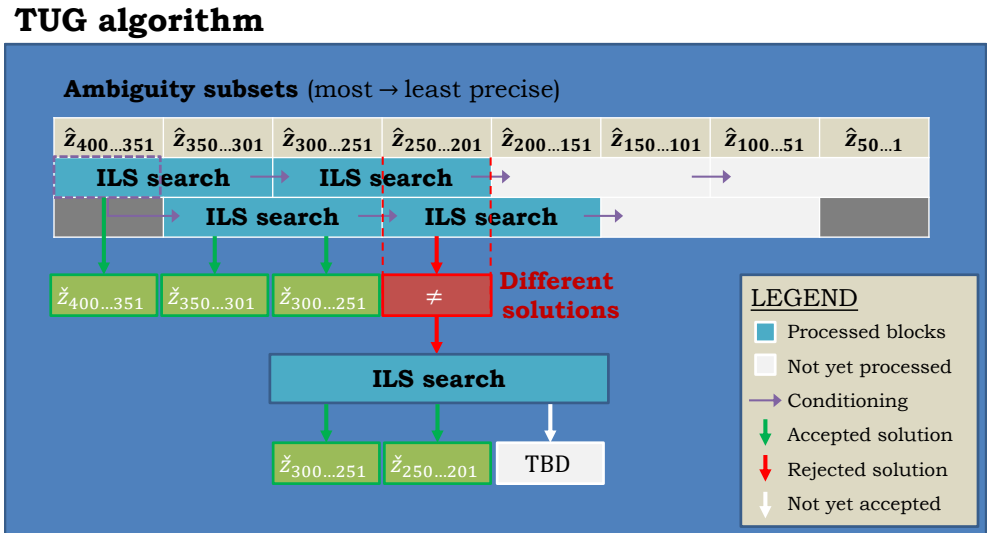


Figure 3.2: The TUG algorithm is shown based on decorrelated ambiguities. The overlapping subsets are fixed and used to conditioning the following ones. See text for more details.

Note that when working with such larger subsets, then we condition the successive ambiguities based on this new solution.

It is understandable that such an overlapping vectorial approach leads to higher computational times, since most components are processed twice. In addition to that, when rejections occur, we also deal with larger ILS problems. Increasing the dimensionality by 50% might lead (in some cases) to higher computational times, depending on the initially chosen size. Hence, it is also important to somehow limit the expansion strategy otherwise the entire process could become jammed in an interminable integer search process.

3.3 Processing strategy and scenarios selection

The experiments presented in this article are conducted using the open-source software GROOPS (Mayer-Gürr et al., 2021), which is developed and maintained at Graz University of Technology. The source code of GROOPS is available on GitHub (https://github.com/groops-devs/groops) together with documentation and an installation guide. The software features GNSS processing capabilities based on an implementation of the raw observation approach (Strasser et al., 2019).

It supports multi-GNSS and multi-signal environments and is optimized for the processing of large GNSS station networks. GROOPS was used to process up to 800 stations and 75 GPS, GLONASS and Galileo satellites per day for TUG's contribution (Strasser and Mayer-Gurr, 2021) to IGS repro3, the third reprocessing campaign of the

Table 3.1: Estimated parameters per component and their a priori constraints (σ)

Component	**Parametrization**
Satellite orbits	Initial state, 7 ECOM2 parameters, with stochastic pulse at midday ($\sigma = 0.1\,\mu$m/s per axis)
Satellite/receiver clocks	Epoch-wise
Satellite/receiver code biases	Constant per day, signal (e.g., C1C), and constellation
Satellite phase biases	Constant per day, frequency, and constellation
Receiver phase biases	Constant per day, signal (e.g., L1C) , and constellation
Phase ambiguities	Constant per continuous track and signal
Tropospheric zenith wet delay	Degree-1 spline with 2-hourly nodes per station ($\sigma = 5\,$m)
Tropospheric gradient delays	Constant-trend per day/station in the north/east ($\sigma = 5\,$m)
Ionospheric influence	Slant TEC per satellite-receiver line of sight and epoch

International GNSS Service (Johnston et al., 2017). GROOPS-based GNSS products are thus going to be incorporated into the next version of the International Terrestrial Reference Frame (i.e. ITRF2020).

The processing setup for the conducted experiments is similar to that of repro3 and generally follows what is documented in Strasser et al. (2019). All processing is performed in daily 24-hour batches. The parameters estimated per day are listed in Table 3.1. The seven ECOM2 solar radiation pressure parameters are $\{D_0, D_{2C}, D_{2S}, Y_0, B_0, B_{1C}, B_{1S}\}$ (cf. Arnold et al., 2015).

Strasser et al. (2019) provides more detailed information on the parametrizations used for the different components, for example how the ionospheric influence is separated from code biases and clock errors. The estimated code and phase biases account for satellite-specific and receiver-specific hardware biases (e.g., Håkansson et al., 2017). Combined satellite-and-receiver-specific biases, for example as observed by Hauschild et al. (2019), were not considered in the processing.

The system of equations is firstly solved in an iterative weighted batch least-squares adjustment. The ambiguities, initially treated as float-valued together with other real-valued parameters, are then decorrelated and fixed according to the methodology in Section 3.2. The experiments are limited to GPS and Galileo, either in a single- or dual-GNSS processing. To keep the setup as simple as possible, only observations with the RINEX 3 encoding C1C, C2W (GPS) and C1C/C1X, C5Q/C5X (Galileo) are used, along with their respective phase counterparts. The code and phase observations are processed at a 30-second sampling period. Satellite and receiver antenna calibrations are taken from "igsR3_2135.atx", i.e. the repro3 ANTEX file described in Villiger et al. (2020).

The station coordinates from a preliminary repro3 combined solution (IGSR03SNX) are introduced here as known, resulting in station-fixed solutions. This reduces the number of unknown parameters, allowing a more focused analysis of the impact of AR onto satellite products. In combination with the repro3 ANTEX file, it leads to products that are aligned to the IGSR3 reference frame.

Table 3.2: The list of a priori models considered for the dynamical accelerations of satellite orbital motion and corrections to GNSS code and phase observations

Dynamical accelerations	**Model adopted**	**Reference**
Earth's gravity field	GOCO06s	Kvas et al. (2021)
Astronomical tides	JPL DE432	Folkner et al. (2014)
Solid Earth tides	IERS 2010	Petit and Luzum (2010)
Ocean tides	FES2014b	Carrere et al. (2016)
Pole and ocean pole tides	IERS 2010	Petit and Luzum (2010)
Atmospheric tides	AOD1B RL06	Dobslaw et al. (2017)
General relativity	IERS 2010	Petit and Luzum (2010)
Solar radiation pressure	Box-wing	Rodriguez-Solano (2014)
Earth radiation pressure	Box-wing	Rodriguez-Solano (2009)
Antenna thrust	Narrow-beam	Steigenberger et al. (2018)
Observables corrections	**Model adopted**	**Reference**
Solid Earth tides	IERS 2010	Petit and Luzum (2010)
Ocean tides	FES2014b	Carrere et al. (2016)
Pole and ocean pole tides	IERS 2010	Petit and Luzum (2010)
Atmospheric tides	AOD1B RL06	Dobslaw et al. (2017)
Tropospheric delay	VMF3	Landskron and Böhm (2018)

Any potential reference frame and antenna model differences affecting satellite orbit comparisons are taken into account in our analyses by means of a 7-parameter Helmert transformation. For example, IGS Final/CODE MGEX products used for evaluation are both aligned to the IGb14 reference frame and are based on different satellite antenna Z-offsets, which results in a difference in scale that is accounted for by the transformation parameters. The small global network adopted in the experiments is detailed in Section 3.3.2. In case any of the stations are unavailable on a specific day, they are replaced by nearby IGS stations to keep the network geometry as stable as possible.

State-of-the-art force and principal models have been applied in this processing, as described in Strasser et al. (2019). These are summarized in Table 3.2, where we refer to both dynamical accelerations for the satellite orbital motion and corrections to the GNSS code and phase observations. Since these two elements refer to inertial and non-inertial reference systems, the Earth orientation is introduced here as known based on the IERS EOP 14 C04 (IAU2000A) time series, along with an additional model covering the high-frequency effects (Desai and Sibois, 2016).

An elevation-dependent stochastic model in the form $\sigma(el) = \sigma_z / \sin(el)$ is applied to define the a priori standard deviation of an observation based on its elevation angle "el", which is then used to weight the observations. Following repro3 configurations of GPS and Galileo, the standard deviation at zenith σ_z for raw code and phase observations is 22 cm and 1 mm, respectively.

3.3.1 Impact of ambiguity resolution with the network size

We start considering a global network of IGS stations, comprised of 60 well-distributed ground-based receivers, and a few sub-networks are selected starting from this largest one. The network and subsets are illustrated in Fig. 3.3, with a number of stations $M \in [10, 60]$.

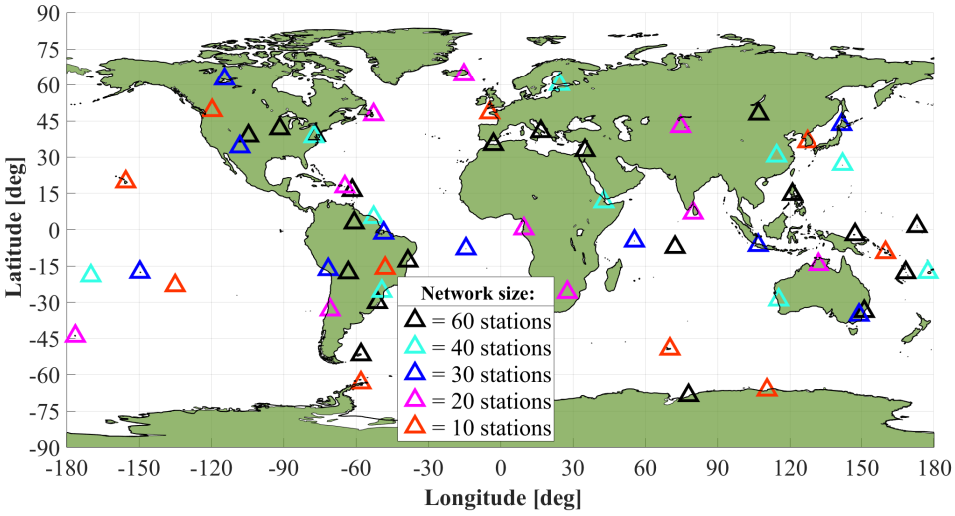


Figure 3.3: Illustration of networks with different sizes between 10 and 60, with ground-based station receivers well distributed around the globe. Each smaller network is a subset of larger ones.

For seven consecutive days (DOY 298-304 in 2019) we compute, on a daily basis, orbit and clock products considering GPS-only, Galileo-only and GPS+Galileo. Furthermore, in each configuration we compute both a float and a fixed solution. The latter is found based on an integer bootstrapping (IB) estimator and fixing for the most reliable subset (Verhagen et al., 2011) with a sufficient success rate, e.g. 99.9%. With IB, each ambiguity component is conditioned on the previous ones that are sequentially fixed by simply rounding. This conditioning introduces information that enhances the fixing success rate with respect to an Integer Rounding (IR) estimator where a component-wise rounding is performed. In this way, IB allows for a good computational efficiency with respect to integer search-based processes, even if it remains sub-optimal and largely dependent upon the ambiguity parametrization.

The resulting satellite orbits are compared with reference orbits referring to IGS Final and CODE MGEX (Prange et al., 2020) products for GPS and Galileo, respectively. The overall 1D root mean square (RMS) of the orbit differences is therefore computed for all GNSS satellites involved in a certain scenario. The RMS values (in log-scale) are shown in Fig. 3.4, where we observe some relevant features. Each curve is related to a single daily solution (seven in total), while increasing the network size. All these global networks are subsets of larger ones, while we have considered some different

sizes, e.g. $M = [10, 12, 14, 16, 18, 20, 25, 30, 40, 60]$.

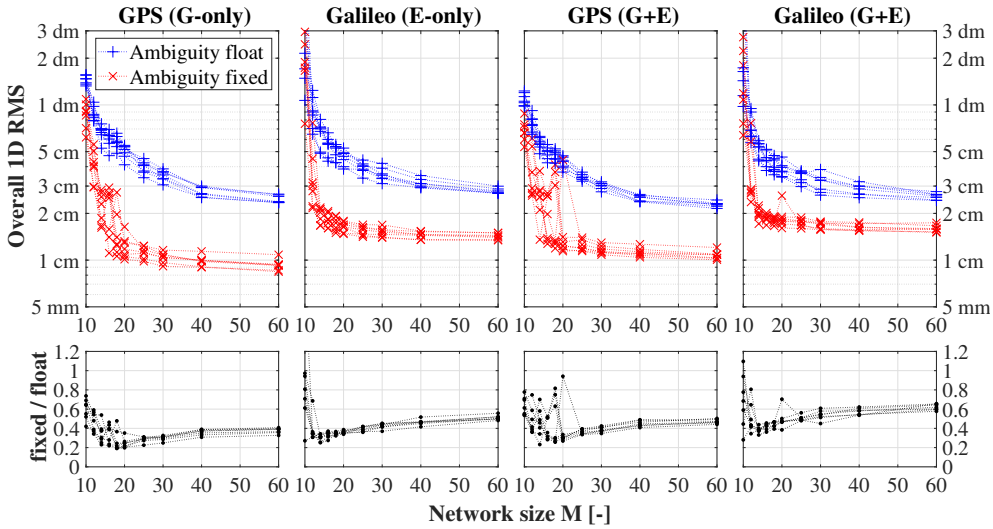


Figure 3.4: The results for orbital (1D RMS) errors given for different GNSS configurations over 7 consecutive days. The notation (G+E) refers to results based on a GPS+Galileo combined processing, rather than G-only or E-only. Both ambiguity-float (in blue) and ambiguity-fixed (in red) solutions are shown, while the fixed/float ratio is given (in black) in the bottom plots.

In the float solutions (blue), the underlying observational model strength increases with a larger network and the overall RMS substantially decreases down to a 2-3 cm level. This is visible for both constellations and each configuration. In the fixed solutions (red), the overall RMS decreases quickly, but it reaches a plateau with global networks of around 25-30 stations. Adding more stations only leads to very small improvements in terms of orbit differences, as they ultimately become limited by the consistency between our estimates and the reference orbits in terms of force modeling and/or processing strategy (Section 3.3). Nonetheless, for $M = 25$, the satellite orbit differences (1D RMS) for GPS and Galileo fixed solutions are at 1-2 cm level, whereas the respective float solutions are at around 3-5 cm.

When focusing only on fixed solutions for small-size networks, i.e. $M \leq 20$, we note that IAR benefits seem somehow limited, probably due to a lower model strength of such small networks. In fact, a crucial condition for improving the accuracy of estimates (conditioned on the fixed carrier-phase ambiguities) is related to the successfulness of the IAR process. After all, given stochastic inputs to an integer estimator, also the outputs are expected to manifest a certain stochastic nature (Teunissen, 1998) and the least precise components might not be correctly fixed to their integer values. These wrong fixes might deteriorate our solution (e.g. jumps), while it is visible that for $M > 20$ we can (generally) already reach the 1-2 cm level in terms of orbit difference (1D RMS).

It is also interesting to compare the single- and dual-constellation processing. In fact,

for the GPS+Galileo case, our fixed solutions (based on the same network) are showing slightly larger RMS values than for the respective GPS-only or Galileo-only scenario. The model inconsistencies between the two constellations might be compensated by real-valued ambiguities that are estimated (float solution), whereas they become more visible when applying the integer constraints (fixed solution). Moreover, in this dual-GNSS case, we are considering (integer-estimable) ambiguities mixed from both GPS and Galileo constellations, therefore processed as one full set. In this way, a between-constellation correlation is introduced. It seems interesting to further investigate such a GPS+Galileo scenario within small networks, e.g. looking at $M = 14$, whereas still thousands of ambiguities are involved.

3.3.2 Small-size network and configurations

We consider a small global network with 14 stations, and assuming a 5° elevation mask we obtain a visibility of 3 to 7 ground stations. For example, assuming a GPS satellite altitude, in Fig. 3.5 the number of ground-based stations visible at any point is illustrated by the depth-of-coverage (Blomenhofer et al., 2005). With Galileo satellites' higher altitude, and given the same elevation mask, station visibility is slightly increased. Hence, the depth-of-coverage is not exactly the same, nonetheless the patterns illustrated within Fig. 3.5 are found to be very similar.

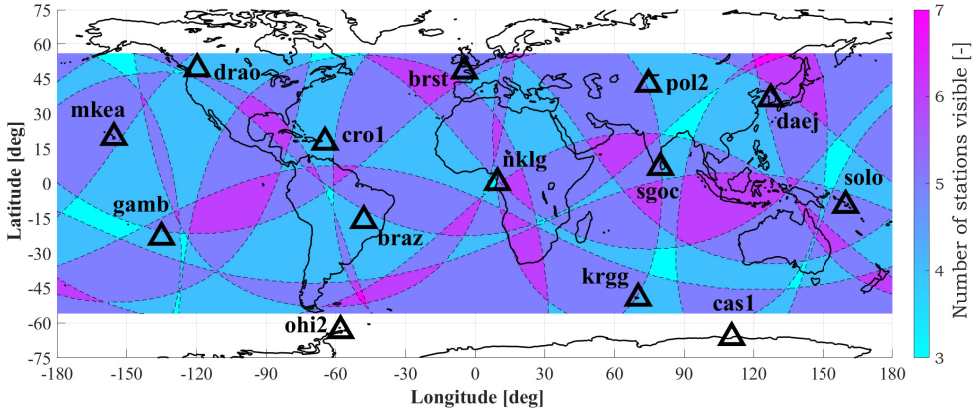


Figure 3.5: The depth-of-coverage is given for GPS satellite altitude and 5° elevation mask in this small global network, which consists of 14 well-distributed ground-based stations.

In all our analyses, we consider a total of around 30 GPS and 24 Galileo satellites. This number can however change if certain satellites are not available on specific days. In general, for such a dual-frequency processing (and specified number of satellites/stations), we can find an ambiguity dimensionality of $n \approx 830$, 1290 and 2120 for Galileo-only, GPS-only and GPS+Galileo, respectively. Hence, with only 14 stations, we are already dealing with quite large dimensionalities and with more Galileo satellites having been deployed in 2020, this dimensionality might further increase.

In some very rare cases, some stations are replaced since they are not available on specific days in the 3-month period considered for our validation, here DOY 244 till 334 in 2019. In that case, the station is replaced by a different close-by IGS station. As a matter of example, SGOC is replaced by DGAR on DOY 263, and KRGG by PERT on DOY 322.

3.3.2.1 Block size selection for partitioning

The selection of suitable block sizes for the partitioning follows a balanced trade-off between the computational complexity, i.e. the efficiency aspect, and the reliability of fixed solutions, i.e. the robustness aspect. In this research work we consider a fixed block size "q" for all subsets, and in both VIB implementations we adopt $q = 200$, while in the Section 3.2.3 an example was given based on $q = 100$. The former value is found to be suitable for solving a GPS+Galileo HDAR problem (i.e. $n > 2000$) within a few tenths of a second in most of the experiments. Given that generally q is not a perfect divisor of n, we first process a (most precise) block with n_m components where $n_m = n - q \cdot \lfloor n/q \rfloor$ with $\lfloor \cdot \rfloor$ as the floor function. It follows that the other ambiguity blocks are given by $n_i = q$, for $i = m - 1, \ldots, 1$ and the number of blocks is simply found as $m = 1 + \lfloor n/q \rfloor$.

It should be remarked that for very small block sizes, e.g. $q \ll 100$, the integer search process is almost instantaneous, but we also have many more conditioning operations to perform. The latter might contribute to increase the overall computational effort since it involves additional matrix operations. The cost of such matrix operations, as well as for the integer search in each block, depends upon the computational capabilities and/or hardware of the platform in use, but is not numerically discussed in this work. Nonetheless, by using larger blocks we can obtain more reliable solutions, since closer to the optimality given by an ILS solution. Still, it is important to avoid ending up into computational bottlenecks due to the exponential growth of number of candidates spanned in the search tree for too large subsets.

3.4 ODTS results and comparison

In this section we focus on the small global network described in Section 3.3.2. Here, the different IAR strategies are compared and their beneficial impact onto satellite products is evaluated. When comparing the fixed results, we will refer to the integer bootstrapping estimator as 'IB', while 'CascAR' and 'TUG' are the two algorithms presented in Section 3.2.2 and Section 3.2.3, respectively. For the sake of convenience, we might refer to 'VIB' when results from CascAR and TUG are identical, for example in comparison with the (scalar) IB.

3.4.1 Results for Galileo-only

We start with the Galileo-only processing, where up to 850 ambiguities need to be fixed for each day. In this scenario, over the same testing week (DOY 298-304 in 2019) used in Fig. 3.4, we would be able to actually compute an ILS solution in less

than a few seconds. The latter is possible due to the Z-transformation on ambiguity components, which improves the overall search time. Moreover, this decorrelation also makes quasi-optimal both the IB and VIB-based solutions, so leading here to the same integer solution as ILS. The latter is thus compared with respect to the float one.

The results are shown in Fig. 3.6, where the float and (ILS) fixed solutions are shown in black and magenta colors, respectively. For these seven days, the overall RMS of the entire constellation is computed epoch-wise. Looking at the radial component we see that a fixed solution allows RMS values below 3 cm for a large part of the day, whereas the float solution can exceed several centimeters. For the along-/cross-track components, large improvements are found and the overall RMS of fixed solutions are generally within a few centimeters. For what concerns the satellite clock component, smaller RMS errors are noted, but, as for the radial direction, the improvements are less pronounced.

These results are in agreement with the results of Li et al. (2019) and Laurichesse et al. (2013), since the ambiguity resolution in ODTS does not constrain much the radial direction due to its correlation with the clock error bias. Consequently, the largest improvements are generally found in the along-/cross-track error components. With respect to the float solution, the 1-week mean RMS value in the ILS solution is around 47% smaller for both radial and clock components. For what concerns the along- and cross-track components, the RMS value is 72% and 74% smaller, respectively.

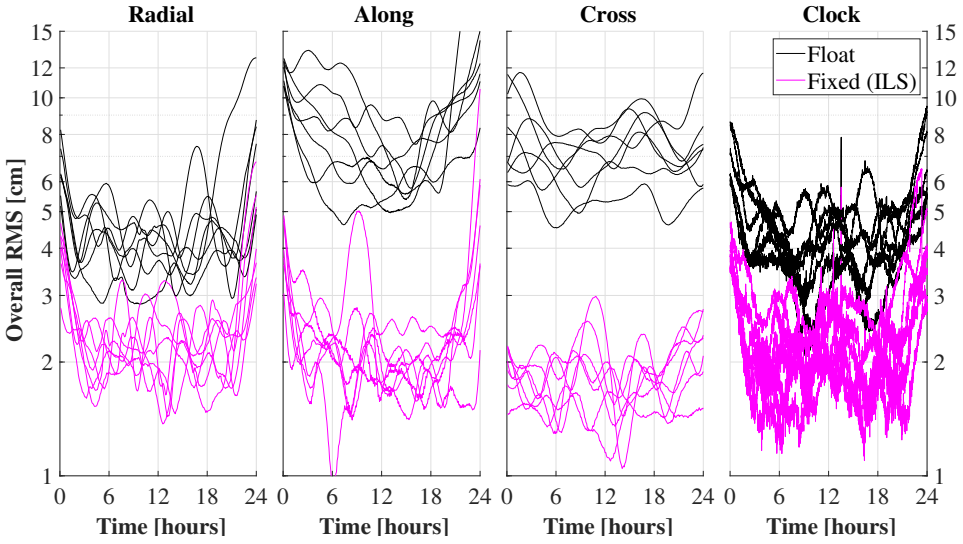


Figure 3.6: The comparison for seven days of the epoch-wise RMS error over the entire constellation, divided into radial/along/cross orbital components and satellite clock error.

For the sake of completeness, we might very briefly look at the error distribution, here based on all Galileo satellites over this 7-day analysis. In Fig. 3.7, the distribution is shown with four histograms (1 cm bins) for each component, while considering both a

float and a fixed solution. It is visible that in the fixed case, most of the errors are bounded within 3 cm, although the largest improvements mainly concern along and cross components.

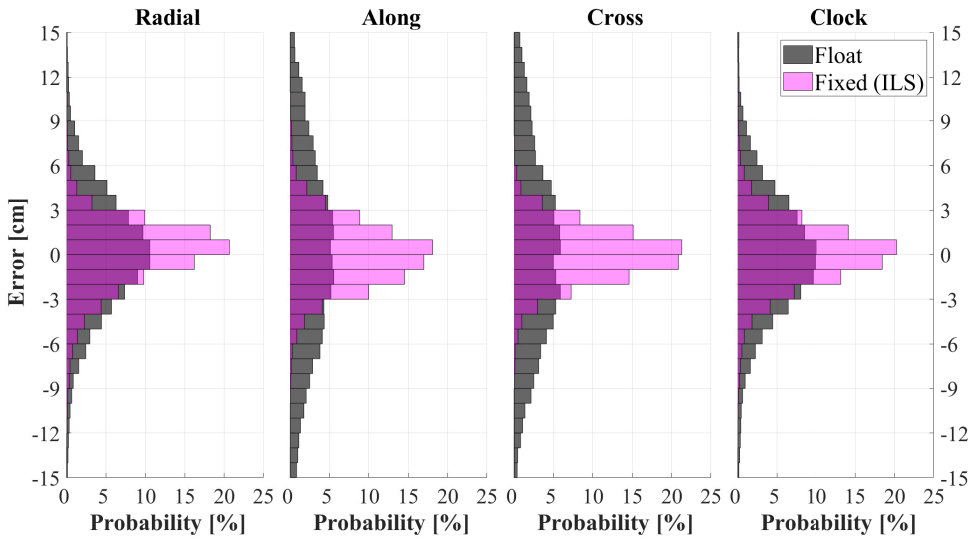


Figure 3.7: The error distribution is shown for all Galileo satellites considering DOY 298 till 304 in 2019. The results are overlapped for a float and a (ILS) fixed solution, separating each component.

3.4.2 Results for GPS-only

When considering the GPS-only process, we deal with a higher dimensionality due to more satellites available at that time. Moreover, large computations for an ILS solution make this optimal solution unfeasible, possibly also due to the poorer stochastic model that affects the precision of float ambiguity estimates (Amiri-Simkooei et al., 2016). Several elements might concur in the uncertainty of observables, as investigated in (Hou et al., 2021), such as the robustness of different signals tracking scheme or the receiver model in use. Still, such features have not been further investigated in the scope of this work.

In order to enable fixed solutions within seconds (or less), we consider a quite reliable (and very efficient) IB solution, as well as the TUG and CascAR algorithms. All three solutions can largely benefit from the decorrelating Z-transformation that has been previously applied with LAMBDA. In this second scenario, the fixed solutions found with IB, TUG and CascAR are now different, and differ in terms of satellite orbit and clock estimates. Here, we focus our analysis mostly on the orbits, with Fig. 3.8 showing the epoch-wise overall RMS values per orbital component.

In almost all cases, both VIB-based solutions lead to same results, meanwhile the IB solution shows some larger errors in particular for DOY 300-302. In fact, RMS values might even largely exceed the ones for the respective float solution, as a possible

consequence of wrong fixes. A different scale has been used for the radial component, where errors are within a few centimeters. It is interesting to observe, as expected, that the VIB-based results are generally equal to or better than the scalar IB ones (Teunissen et al., 2021).

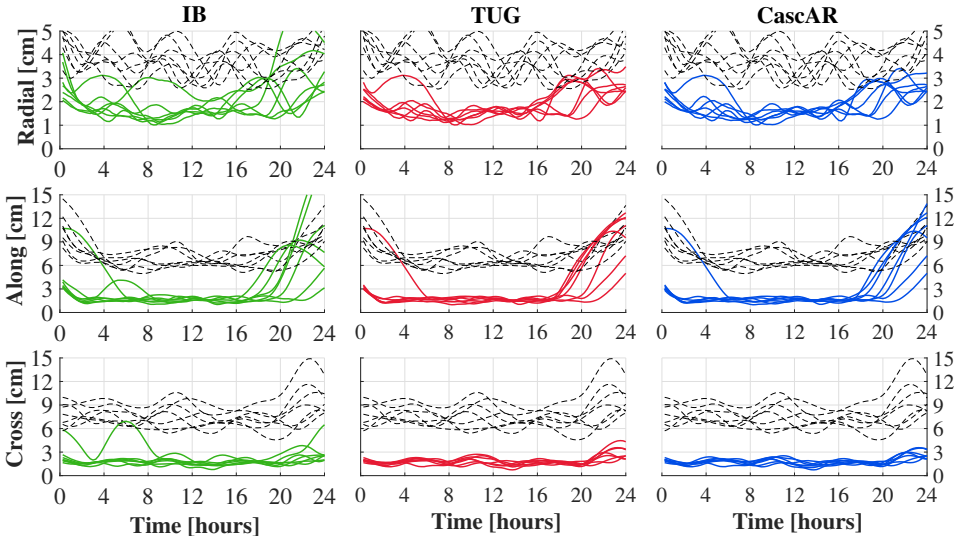


Figure 3.8: The epoch-wise RMS of orbital components is shown for the fixed solutions with respect to float (dashed lines). A total of 7 days (DOY 298-304 in 2019) is used in this GPS-only processing.

A more detailed analysis shows that in most cases those larger errors are related to a one or two satellites, for example on DOY 302, but on some other days might be affecting more satellites, e.g. on DOY 300. This last case is shown in Fig. 3.9, where specific GPS satellites presenting large orbital errors have been highlighted in color (same in all solutions). The remaining ones are still shown in grey dashed lines, but are generally bounded within 5 cm.

A clear improvement is observed with respect to the float solution, nonetheless, the IB solution leads for many satellites to large orbital errors, up to 15 cm in the radial component and several decimeters for the along-track. In the VIB-based solutions we get large errors only in the case of a single satellite (PRN16), most likely due to a wrong fix. These issues of specific satellites cause the noticeably increasing RMS, particularly in the along-track component, towards the end of a day, as visible for most days in Fig. 3.8. In most cases, the underlying issue is short observation arcs caused by the cut-off at day boundaries which lead to ambiguities that are not well determined and can cause issues in ambiguity resolution. Due to the GPS ground-track repeat period of 1 sidereal day, this effect is very similar for the 7 days shown in Fig. 3.8. However, this is coincidental and a slowly changing observation geometry leads to this effect being less pronounced for other parts in the 91-day period.

Still, it is visible (as also for other testing days) that VIB solutions generally offer

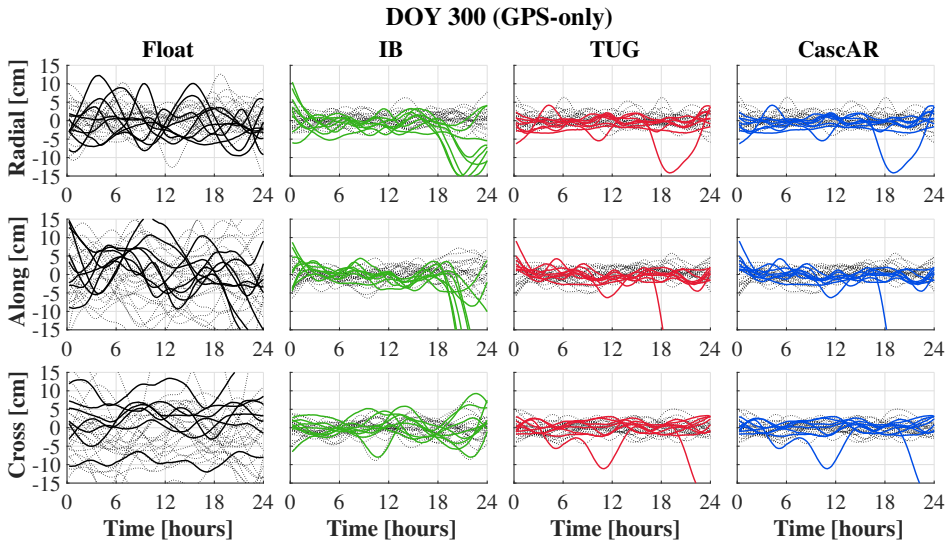


Figure 3.9: The orbital errors over 24 hours (DOY 300) for different solutions, considering GPS-only and highlighting (with colors) certain satellites where fixed-solution errors are quite large.

more reliable fixes without much impacting the computational effort, later discussed in Section 3.4.4. For instance, HDAR solutions of such a dimensionality (n≈1300) can be often found within hundredths of a second for standard desktop processors, e.g. Intel i7-4790 @3.6 GHz. As a matter of example, in Fig. 3.10 we show the RMS values (over these 24 hours on DOY 300) for all GPS satellites based on the four solutions previously discussed. The CascAR and TUG solutions are identical on this day, as is the case on almost all days of this scenario. With respect to the float solution, the R/A/C components are improved here by around 31/42/71% and 52/67/77% for the IB and both VIB solutions, respectively. As already seen for the Galileo-only ODTS results, larger benefits of fixed solutions are mostly observed in the along-/cross-track components.

3.4.3 Results for GPS+Galileo

At this point we can focus our investigation on the GPS+Galileo processing scenario. This is also the most interesting one, due to a larger ambiguity dimensionality involved, where in general for our small network we always have $n > 2000$. For a more extensive analysis of this HDAR problem, we consider a 3-month period in 2019, starting from DOY 244 till 334, both days included. We examine the solutions for these 91 days, so processing GPS+Galileo on a daily basis as described in Section 3.3.

Firstly, we start considering the RMS errors over all days while referring to each satellite for GPS and Galileo, and each orbital component. These RMS values are then divided with the ones relative to a float solution (dashed line), and ratios are provided, as percentages, in Fig. 3.11. Smaller % values imply larger improvements

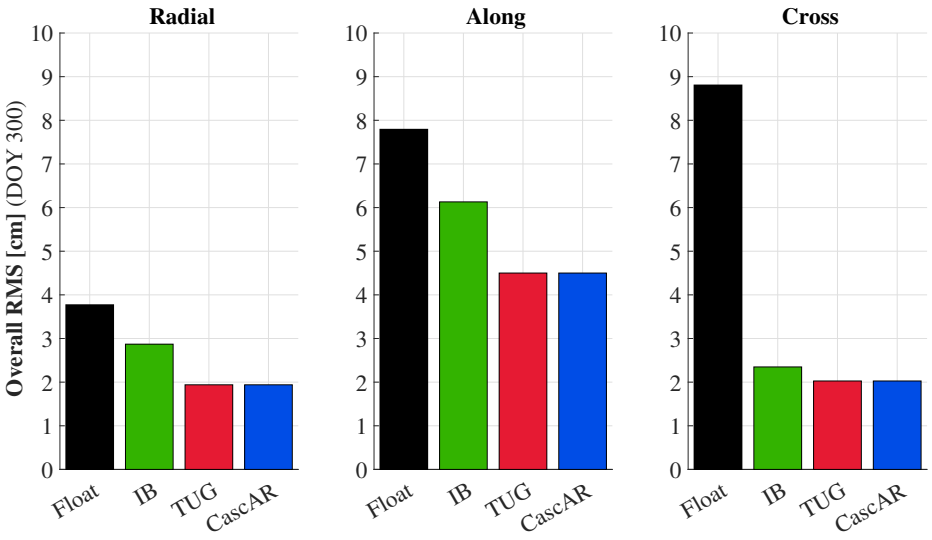


Figure 3.10: The overall RMS error is shown relative to a 24-hour period, i.e. DOY 300 in 2019. For each orbital component, we consider the float (black) and three fixed solutions.

with respect to the float solution, which is often the case. For GPS satellites, the mean improvement in terms of R/A/C orbital components is approximatively 28/44/56% and 33/50/63% for the IB and the VIB-based solutions, respectively. For Galileo satellites, this is approximatively 30/56/62% for the IB, while 32/60/64% again for both TUG and CascAR solutions. These two VIB solutions show a high consistency in the results, but not always the exact same solutions.

At this point we can consider in Fig. 3.12 the results in terms of overall orbital errors (1D RMS) for GPS (left plot) and Galileo (right plot). For each day we compute values over all satellites (and components) of each constellation. By comparing the fixed solutions for Galileo, we see that they are usually quite similar, with the exception of specific days where the (scalar) IB solution shows much larger errors with respect to TUG or CascAR. This is for instance visible on DOY 268 where the 1D RMS value exceeds 12 cm. In these Galileo results, a slight discrepancy is also visible for the TUG and CascAR solutions, specifically on DOY 294, which is discussed in more details later in Section 3.4.4.

When looking at the GPS results, similarly to what was seen in Section 3.4.2, we observe several 1D RMS differences, and on several days we have high values for the IB. This is again most likely due to wrong fixes, which deteriorates the performance of the ODTS process in this small global network estimation. With respect to the Galileo results, float values are generally found at around 5 to 7 cm, whereas we have observed some much higher peaks in Galileo during the first four weeks, i.e. in September 2019. Nevertheless, this particular behavior has not been further investigated in the context of this research work.

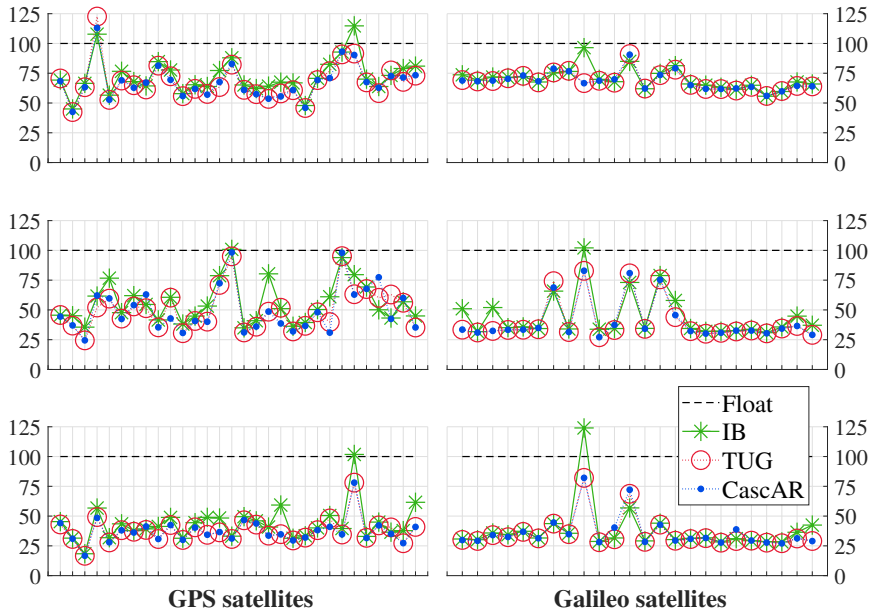


Figure 3.11: The results for all GPS and Galileo satellites are given in terms of RMS ratio w.r.t. float solution for each orbital component over the 3-month period considered (DOY 244-334 in 2019).

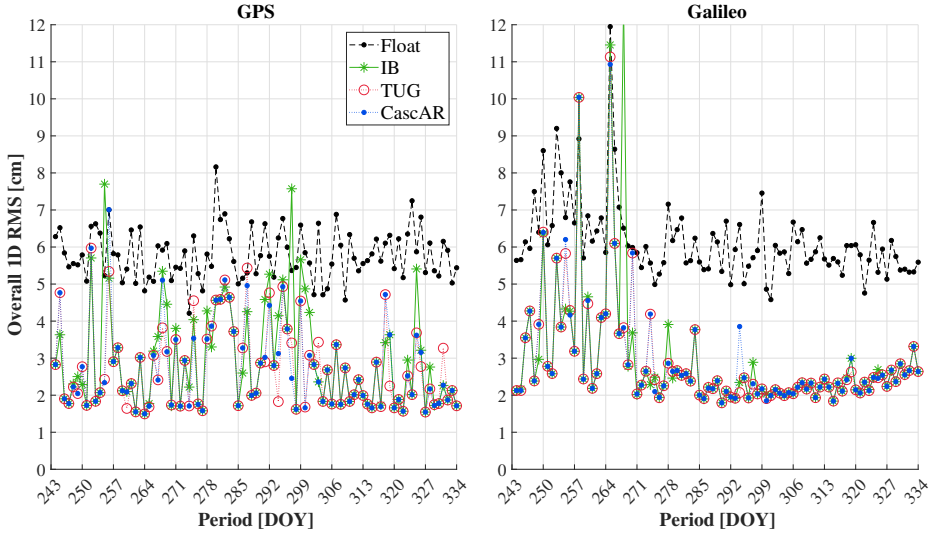


Figure 3.12: GPS+Galileo results in daily 1D RMS values over each constellation. The float solution is shown, along with three fixed ones: IB, TUG and CascAR.

Before proceeding with the validation of these GPS+Galileo results, we might firstly consider a small comparison between the two VIB methods.

3.4.4 Comparison between TUG and CascAR solutions

The high consistency shown between VIB-based solutions is mostly due to the similarity of the two algorithms, although in the TUG "blocked search" approach additional checks are performed on the integer solution by looking at overlapping blocks (see Section 3.2.3). If the solutions are always consistent in each overlap, then both TUG and CascAR are basically equivalent, i.e. return the same integer solution. This is the case for most of the days tested here, however, not always. By a detailed comparison in Fig. 3.12, only in a few occasions the two VIB-based results are quite different. For DOY 294 we clearly see CascAR being suboptimal to both IB and TUG. In Fig. 3.13, we present the orbital errors (as 3D scalar displacement) for all 54 satellites over the 24 hours on DOY 294. It is visible that in the CascAR solution, some wrong fixes occur and results are largely affected, with errors exceeding some decimeters. It should be mentioned that such large errors are not only affecting Galileo, but GPS as well, and in general only limited to very specific days.

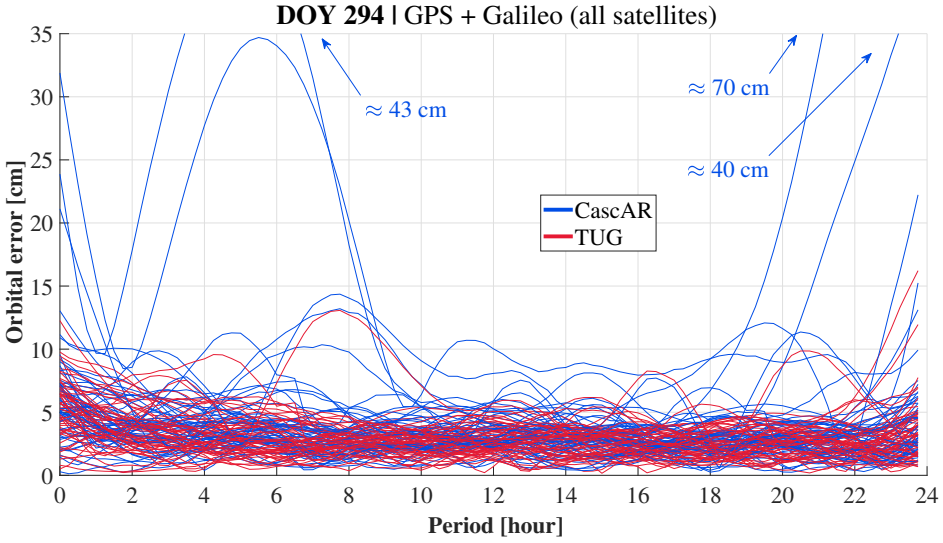


Figure 3.13: A detailed example for DOY 294 (in 2019) concerning the epoch-wise orbital errors of all satellites processed (different curves) with TUG and CascAR ambiguity resolution methods.

At this point we might wonder whether such robustness of the TUG algorithm comes at a higher computational cost, thence on a much lower efficiency. As mentioned, the VIB-based approaches allow for high flexibility in the design of most effective algorithms, but this also depends on the specific application, the considered scenarios and/or available computational capabilities. Hence, the daily computational time is very briefly compared for the different IAR methods considered in this test.

The IB solution is generally computed in a few milliseconds and it is only dependent upon the ambiguity problem dimensionality, which was quite constant over this 3-month period. For instance, in the GPS+Galileo processing we have always between 2050 and 2250 ambiguity components. In Fig. 3.14, we show the computation times for the two VIB-based solutions, based on an average over 100 runs. It is visible that both VIB-based solutions are computationally more expensive than IB, but they seem to provide a more reliable solution as shown in the previous section, and further validated in Section 3.5. Still, for most of the cases, a HDAR solution (i.e. $n > 2000$) can be found within a few tenths of a second, e.g. here based on the Intel Core i7-4790 @3.6 GHz central processor.

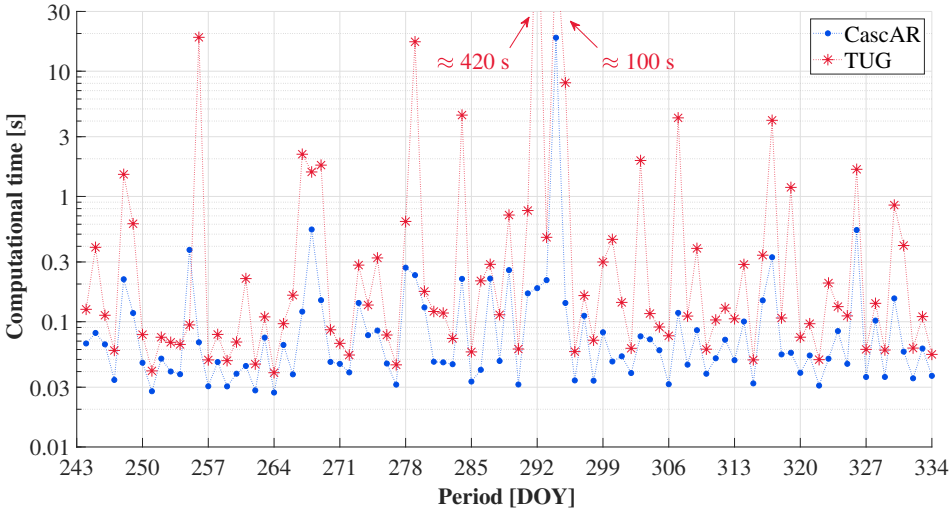


Figure 3.14: The computational time (log-scale), averaged over 100 runs, is given for the TUG and the CascAR algorithms with respect to the GPS+Galileo HDAR problem ($n > 2000$).

For what concerns the TUG algorithm, the computational time is indeed higher due to the overlapping strategy mentioned in Section 3.2.3. In only one case, on DOY 255, the search time is larger for the CascAR approach, while it only exceeds 0.6 seconds once, i.e. ∼20s on DOY 294. The 91-day averaged ratio of CascAR over TUG computational time is 47.8%, therefore being twice as efficient as the TUG algorithm. The much-increased search times observed on some days for the TUG method is caused by the expansion to larger block sizes in case the overlapping of fixed ambiguity components does not match. In fact, in many occasions we see that for the TUG approach it takes more than 0.6 seconds to compute an integer solution to the HDAR problem.

Finally, we should not forget that these examples of VIB algorithms are only two possible implementations, leaving further more advanced strategies still to be investigated. A few suggestions will be given in Section 3.5.2. We can now continue to focus on a validation of results for the GPS+Galileo scenario.

3.5 Validation and discussion

We can now present a further analysis of the results, where a different metric is adopted. In fact, as mentioned, the adoption of external products as 'ground truth' for the evaluation of orbit and clock errors has some limitations in terms of consistency. Given that we have processed network data on a daily basis, we might instead consider the orbital discontinuity at midnights. This evaluation strategy is independent from external sources and quantifies jumps in satellite orbits between consecutive days.

3.5.1 Analysis of orbital discontinuity at midnight

The GPS+Galileo scenario is considered, and for each satellite we compute the 1D RMS of orbital discontinuities at the midnight between two consecutive days. For the float solution we have always larger discontinuities, almost always exceeding 10 cm and it is therefore not illustrated here. Furthermore, TUG and CascAR solutions are matching basically in most of the results, except for 1-2 days. This difference has been already discussed in Section 3.4.4, so for this analysis we focus on IB versus VIB, the latter referring to the TUG solution. These results (over the 3-month period) are provided in Fig. 3.15 based on a color scale between 0 and 10 cm, while separating the satellite groups for each constellation.

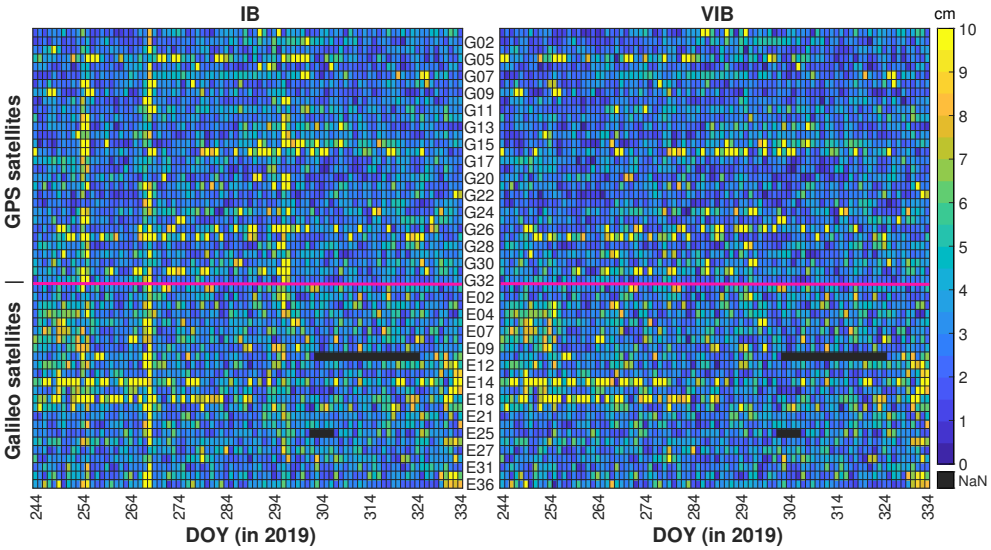


Figure 3.15: The satellite orbit discontinuity errors at midnight are shown for GPS+Galileo processing over a 3-month period in 2019. Both IB and VIB solutions are considered, as described in the text.

In the IB case, we see that many days present much larger values with respect to the vectorial counterpart. The VIB solution is sometimes exceeding the dm-level only for a few satellites, i.e. E14/E18 for Galileo and G05/G16/G26 for GPS. In the former case it should be noted that those two satellites are the ones placed into slightly elliptical

orbits, making it more challenging to properly model their orbit and attitude.

The most notable differences between IB and VIB clearly fully affect either one or both GNSS constellations, so we can examine the range of discontinuity values for both constellations. For each midnight epoch we compute a box plot over each system and results are shown in Fig. 3.16 for the three methods: the float solution, along with the fixed IB and VIB ones. A box plot visualizes the median value as a red line within a blue box that defines the interquartile range, i.e. the difference between 75th and 25th percentiles. The full range is given by grey whiskers, excluding outliers that are marked as red dots.

Many days show only small improvements when comparing IB to VIB, but on some days the discontinuities are much smaller in case of VIB-fixed solutions. However, improvements with respect to the float case are always evident. It is interesting to observe that for the (scalar) IB estimator we have, on specific days, errors exceeding the float ones. Once again, a wrong fixing can lead to an inconsistent satellite orbit, and this can impact the discontinuity values at the boundary/ies of that particular day. In this circumstance, many satellites might be affected and fixed results are visibly worse than the float solution.

At this point, for the sake of completeness, we define an overall empirical cumulative distribution of these orbit midnight discontinuity errors for each solution. Also in this case, both VIB-based solutions are represented by a single curve since no (visible) differences were found. In Fig. 3.17 we see that in terms of orbit midnight discontinuity errors, over the 91 days, the float solution has quite large errors, and in 50% of the cases those values exceed 8 cm. For the fixed solutions we have around 3 cm, with a negligible difference between the two methods. At 95% probability, float-solution errors are around 18.3 cm, roughly twice the respective value of the fixed solutions. This also confirms how substantial improvements can be achieved by robust (and ideally efficient) ambiguity resolution schemes in the estimation of satellite orbits and clocks, even when using a small global network.

We need to remark that such a good result is possible also thanks to the LAMBDA decorrelation of ambiguities. In fact, this can substantially improve the optimality of integer bootstrapping approaches. Still, the VIB solutions further improve the results with respect to the scalar IB, thus reducing the discontinuity errors from 10 to 8 cm (at 95% probability). As mentioned in Section 3.4.4, these more reliable solutions are possible without too much compromising the computational efficiency of the IAR process. Lastly, it should be remarked that enhancements at the decorrelation step might further benefit the here presented VIB formulations, where an example of comparison among different “reduction” methods can be found in Jazaeri et al. (2014).

3.5.2 Limitations and further possibilities

When adopting VIB, it is important not to generalize properties of a certain estimator to the entire ambiguity set. For example, the optimality of ILS within each block does not hold for the full set where VIB solutions are indeed suboptimal with respect to

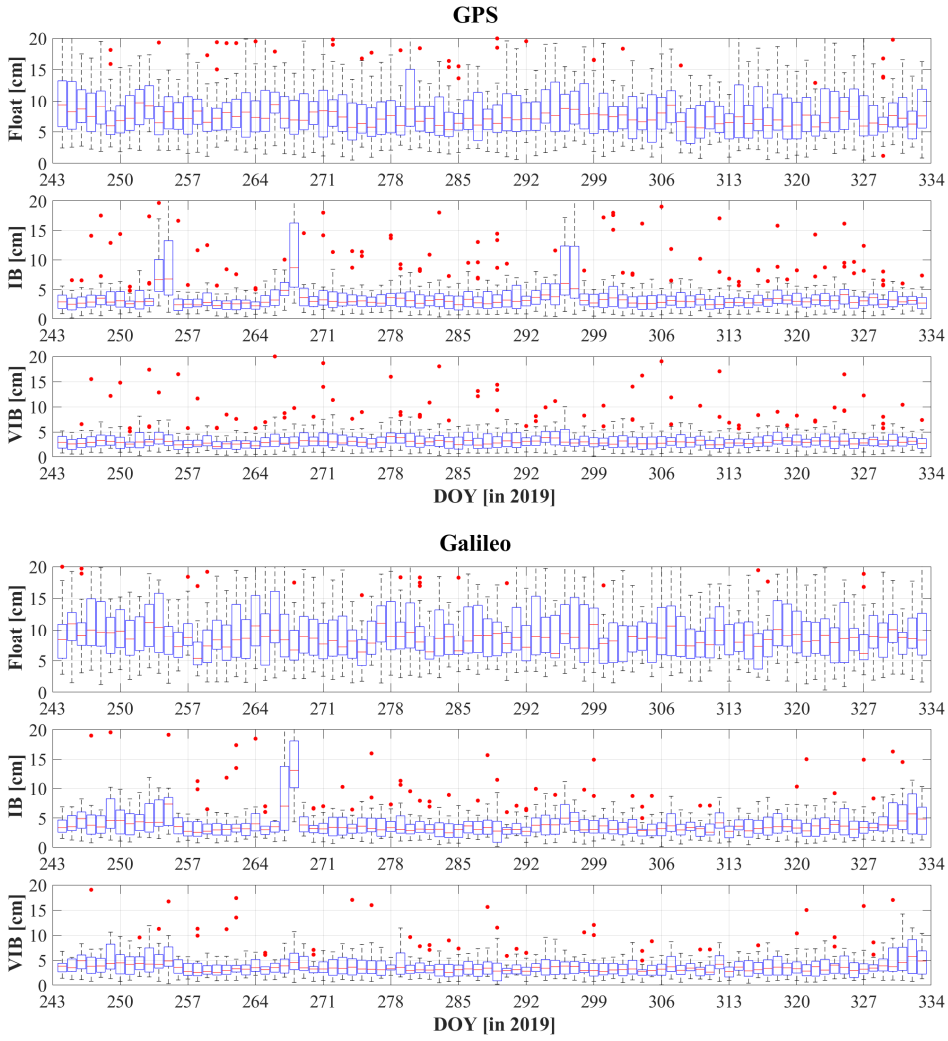


Figure 3.16: The box plot results for midnight discontinuities of GPS and Galileo considering three methods based on a float solution, then a scalar (IB) and a vectorial (VIB) integer bootstrapping.

an ILS solution over the full n-dimensional set. The use of larger ambiguity blocks can lead to higher robustness performance, but sometimes at a very high cost for the efficiency. With the VIB approaches, here focused on a small global network, the daily fixed solutions can still be computed in less than a second and generally lead to smaller errors. On some days, these errors might exceed a few decimeters when making use of the (still convenient) integer bootstrapping estimator.

The VIB formulation shall not be restricted to the use (in each subset) of I-estimators, e.g. ILS, but it can benefit from more advanced schemes (Teunissen et al., 2021). Hence,

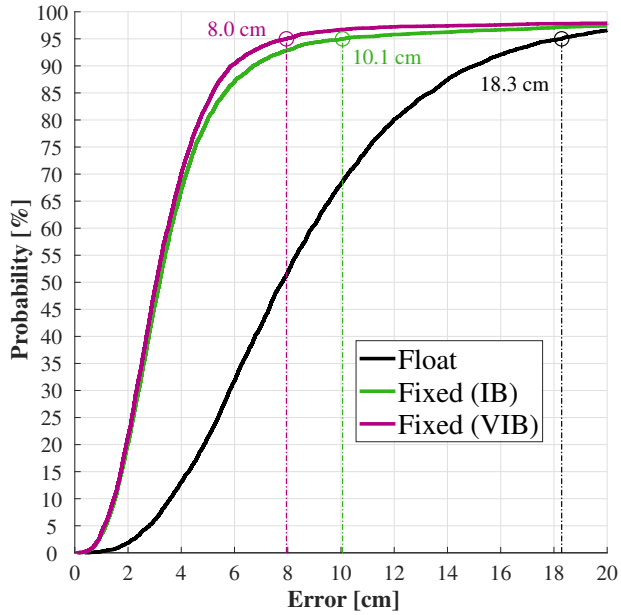


Figure 3.17: Empirical cumulative distribution for midnight discontinuity errors relative to the different solutions and all satellites processed over the 3-month period (DOY 244-334 in 2019). The error values at 95% probability are marked with circles.

by the use of a vectorial formulation, it is possible to combine different types and/or classes of estimators in this vectorial approach. As mentioned, a proper selection of the partitioning strategy can depend upon computational resources available, along with applications and latencies that users want to be addressed by the IAR process.

For instance, adoption of integer validation tests (i.e. a data-driven approach) is possible and it might further prevent the wrong fixing of some ambiguity subsets. In addition to that, more efficient estimators, for example rounding or bootstrapping, might be used only in the more precise ambiguity subsets, while adopting ILS in the remaining ones. Another notable example of VIB-like strategies is given by the Sequential Best Integer Equivariant (SBIE) approach, described in Brack et al. (2013), which enables a more efficient adoption of BIE-related solutions. Still, in their work, this was limited to a scalar case, therefore not taking full advantage from the vectorial formulation analyzed in this research contribution.

3.6 Conclusions

The high complexity involved in the search for an optimal integer least-squares solution to high-dimensional ambiguity resolution (HDAR) problems is still a challenge. In

order to find a trade-off between robustness and efficiency of a certain IAR process, in this work we present two implementations of a vectorial integer bootstrapping (VIB) estimator. Based on the VIB principle, introduced in Teunissen et al. (2021), it is possible to design more flexible IAR algorithms. This allows finding an effective solution to a given HDAR problem, thus balancing available computational resources and reliability of the fixed solutions.

The two algorithms are CascAR, as a straightforward implementation of VIB, and the "blocked search" method developed at Graz University of Technology. In both cases, we take advantage of the same initial triangular decomposition, along with LAMBDA Z-transformation to decorrelate the ambiguity components. These two VIB-based solutions are therefore compared with respect to their scalar counterpart, i.e. Integer Bootstrapping (IB) estimator, and investigated in the context of Orbit Determination and Time Synchronization (ODTS). This is performed based on TUG's open-source software GROOPS.

Firstly, we consider for different network sizes, between 10 and 60 stations, the 1D RMS orbital errors relative to GPS-only, Galileo-only and GPS+Galileo scenarios. In all cases, we adopt a 24-hour batch processing and 30-second sampling period for dual-frequency code and phase observations. The float and IB fixed solutions are considered, the latter showing a plateau at 1-2 cm, reached with 25 or more well-distributed stations. When looking at smaller networks the improvements over the float solutions are significant, but in several occasions wrong fixes might lead to quite large orbital errors. Hence, we focus our investigation on a small global network, i.e. 14 stations, where the three IAR strategies are evaluated and compared.

In the Galileo-only case, all fixed solutions seem to be optimal over the entire set and lead to large improvements over the float results, both evaluated with respect to the CODE MGEX products. On several days, the mean RMS (radial and clock) for a fixed solution can be around 50% smaller than the float one, and more than 70% for both along-and cross-track error components. These last two are in fact generally more affected by IAR, in agreement with literature. In the GPS-only case, with respect to IGS Final products, more visible differences are found among fixed solutions, where orbital errors in the radial/along/cross component can be improved (e.g. on DOY 300) by 31/42/71% and 52/67/77% for IB and VIB solutions, respectively.

The most interesting scenario considered is the GPS+Galileo processing, where more than 2000 ambiguities are fixed per day. This HDAR problem is not trivial, especially when looking for a computationally cheap and reliable fixed solution. Both GPS and Galileo are processed together, so an inter-system correlation of carrier-phase ambiguities might be introduced. The scenario is numerically assessed on a daily basis over a 3-month period (DOY 244 till DOY 334 in 2019), and then analyzed separately for each constellation. In some days, the IB solution shows much larger orbital errors (in terms of RMS over each single constellation), which can exceed several centimeters.

The validation of the GPS+Galileo results is performed by an intrinsic quality check, thus not relying on external orbital information. We look at the orbit discontinuity errors at midnights over the 3-month period. For multiple days, the discontinuities

based on an IB solution are quite large, whereas VIB-based methods provide better results. The two different VIB approaches present quite similar performances over all tests, as expected; nonetheless the TUG solution further improves the results on a few occasions. This comes at a higher computational cost, shown for each day over the testing period. Still, in most of the cases, this HDAR problem can be robustly solved within fractions of a second, while providing substantial benefit to the ODTS estimates. By means of VIB fixed solutions, with a global network of 14 stations, it is possible to compute orbits within a few centimeters error.

The vectorial formulation shows to be a suitable candidate, due to its flexibility, in the design of more effective IAR algorithms, in particular for HDAR problems. It enables a trade-off between efficiency and robustness, based on a very easy-to-implement strategy that has been extensively described in this work. Although referring here to the case of small global network estimation, its applicability extends over any mixed-integer GNSS model, and shall not be restricted to the network case. Additional works, also involving other classes of estimators or different strategies, shall be conducted to further exploit the VIB theory.

3.7 Appendix

The implementation of a generic VIB estimator can benefit from a conformed partitioning of the triangular decomposition for the ambiguity vc-matrix, i.e. $\tilde{Q} \simeq \tilde{L}^T \tilde{D} \tilde{L}$. Based on Eq.(3.4), we can consider $\tilde{L} \in \mathbb{R}^{n \times n}$ (lower unitriangular) and $\tilde{D} \in \mathbb{R}^{n \times n}$ (diagonal), such as

$$\tilde{L} = \begin{bmatrix} \tilde{L}_{11} & 0 \\ \tilde{L}_Q & \tilde{L}_{22} \end{bmatrix}, \quad \tilde{D} = \begin{bmatrix} \tilde{D}_1 & 0 \\ 0 & \tilde{D}_2 \end{bmatrix} \tag{3.7}$$

and for $k = 1, 2$, we have that $\tilde{L}_{kk} \in \mathbb{R}^{n_k \times n_k}$ is lower unitriangular, $\tilde{D}_{kk} \in \mathbb{R}^{n_k \times n_k}$ is diagonal, while $\tilde{L}_Q \in \mathbb{R}^{n_2 \times n_1}$ is a rectangular block matrix. The latter one represents the correlation between the two subsets, therefore involved in the conditioning, i.e. here $\{2\} \rightarrow \{1\}$.

It is straightforward to prove that each block of the vc-matrix $\tilde{Q}$ is retrieved from

$$\tilde{Q} \overset{\text{def}}{=\!=} \begin{bmatrix} \tilde{Q}_{11} & \tilde{Q}_{12} \\ \tilde{Q}_{21} & \tilde{Q}_{22} \end{bmatrix} = \begin{bmatrix} \tilde{L}_{11}^T \tilde{D}_1 \tilde{L}_{11} + \tilde{L}_Q^T \tilde{D}_2 \tilde{L}_Q & \tilde{L}_Q^T \tilde{D}_2 \tilde{L}_{22} \\ \tilde{L}_{22}^T \tilde{D}_2 \tilde{L}_Q & \tilde{L}_{22}^T \tilde{D}_2 \tilde{L}_{22} \end{bmatrix} \tag{3.8}$$

while the first ambiguity block, conditioned on the second one, is described by $\tilde{Q}_{11|22}$ as follows

$$\begin{aligned} \tilde{Q}_{11|22} &= \tilde{Q}_{11} - \tilde{Q}_{12} \tilde{Q}_{22}^{-1} \tilde{Q}_{21} \\ &= \tilde{Q}_{11} - (\tilde{L}_Q^T \tilde{D}_2 \tilde{L}_{22})(\tilde{L}_{22}^{-1} \tilde{D}_2^{-1} \tilde{L}_{22}^{-T})(\tilde{L}_{22}^T \tilde{D}_2 \tilde{L}_Q) \\ &= \tilde{L}_{11}^T \tilde{D}_1 \tilde{L}_{11} \end{aligned} \tag{3.9}$$

that represents its unique $L^T D L$-decomposition due to the form of matrices $\tilde{L}_{11}$ and $\tilde{D}_1$. These mathematical relationships also enable a straightforward computation

of the conditioning for the float ambiguity subset $\{1\}$, given $\hat{a}_k \in \mathbb{R}^{n_k}$ for $k = 1, 2$. Hence, we can write

$$\begin{aligned}\hat{a}_{1|2} &= \hat{a}_1 - \tilde{Q}_{12}\tilde{Q}_{22}^{-1} \cdot (\hat{a}_2 - \check{a}_2) \\ &= \hat{a}_1 - (\tilde{L}_Q^T \tilde{D}_2 \tilde{L}_{22})(\tilde{L}_{22}^{-1} \tilde{D}_2^{-1} \tilde{L}_{22}^{-T}) \cdot (\hat{a}_2 - \check{a}_2) \\ &= \hat{a}_1 - \tilde{L}_Q^T \tilde{L}_{22}^{-T} \cdot (\hat{a}_2 - \check{a}_2)\end{aligned} \tag{3.10}$$

with all matrix blocks already available from Eq.(3.7). In this case, for example, we have that $\check{a}_2 = \mathcal{I}_2(\hat{a}_2)$, where $\mathcal{I}_2$ can be any integer mapping function taken from the class of

- **I-estimators** (Teunissen, 1999), which always return $\check{a}_2 \in \mathbb{Z}^{n_2}$ (integer-valued);
- **IE-estimators** (Teunissen, 2002), which always return $\check{a}_2 \in \mathbb{R}^{n_2}$ (real-valued);
- **IA-estimators** (Teunissen, 2003a), which either return an integer- or a real-valued ambiguity vector, depending upon the result of the integer validation test.

With Eq.(3.9), we observe that the metric for the block $\{1\}$, being conditioned on $\{2\}$, is already 'extracted' from the same initial triangular decomposition. Furthermore, a selection of n_k values was arbitrary, and by setting $\tilde{L} \equiv \tilde{L}_{11}, \tilde{D} \equiv \tilde{D}_1$, we can further partition the sub-problem, thus equivalently obtaining an arbitrary number $m \in [1, n]$ of subsets. Each one can have an arbitrary dimensionality, as long as their sum is equal to n.

We should remark that such a partitioning strategy holds also if looking at a reversed conditioning, therefore $\{1\} \rightarrow \{2\} \rightarrow \ldots \rightarrow \{m\}$. In this case we would need to consider a different decomposition, e.g. see De Jonge and Tiberius (1996), so having $\tilde{Q} = \tilde{L}\tilde{D}\tilde{L}^T$, where $\tilde{L}$ and $\tilde{D}$ are once again a lower unitriangular and a diagonal matrix, respectively.

References

Amiri-Simkooei AR, Jazaeri S, Zangeneh-Nejad F, Asgari J (2016) Role of stochastic model on GPS integer ambiguity resolution success rate. GPS Solutions 20(1):51–61, DOI 10.1007/s10291-015-0445-5

Arnold D, Meindl M, Beutler G, Dach R, Schaer S, Lutz S, Prange L, Sośnica K, Mervart L, Jäggi A (2015) CODE's new solar radiation pressure model for GNSS orbit determination. Journal of Geodesy 89(8):775–791, DOI 10.1007/s00190-015-0814-4

Blewitt G (1989) Carrier phase ambiguity resolution for the Global Positioning System applied to geodetic baselines up to 2000 km. Journal of Geophysical Research: Solid Earth 94(B8):10187–10203, DOI 10.1029/JB094iB08p10187

Blomenhofer H, Ehret W, Blomenhofer E (2005) Investigation of the GNSS/Galileo integrity performance for safety of life applications. Proceedings of European GNSS

Brack A (2019) Partial carrier-phase integer ambiguity resolution for high accuracy GNSS positioning. PhD thesis, Technische Universität München

Brack A, Henkel P, Gunther C (2013) Sequential best integer-equivariant estimation for geodetic network solutions. In: Proceedings of the 2013 International Technical Meeting of The Institute of Navigation, pp 310–317

Carrere L, Lyard F, Cancet M, Guillot A, Picot N, et al. (2016) FES 2014, a new tidal model—validation results and perspectives for improvements. In: Proceedings of the ESA living planet symposium, pp 9–13

Chen H, Jiang W, Ge M, Wickert J, Schuh H (2014) An enhanced strategy for GNSS data processing of massive networks. Journal of Geodesy 88(9):857–867, DOI 10.1007/s00190-014-0727-7

Cocard M, Geiger A (1992) Systematic search for all possible widelanes. In: Proceedings of the 6th International Geodetic Symposium on Satellite Positioning, pp 312–318

De Jonge P, Tiberius C (1996) The LAMBDA method for integer ambiguity estimation: implementation aspects. Publications of the Delft Computing Centre, LGR-Series 12(12):1–47

Desai SD, Sibois AE (2016) Evaluating predicted diurnal and semidiurnal tidal variations in polar motion with GPS-based observations. Journal of Geophysical Research: Solid Earth 121(7):5237–5256, DOI 10.1002/2016JB013125

Dobslaw H, Bergmann-Wolf I, Dill R, Poropat L, Thomas M, Dahle C, Esselborn S, König R, Flechtner F (2017) A new high-resolution model of non-tidal atmosphere and ocean mass variability for de-aliasing of satellite gravity observations: AOD1B RL06. Geophysical Journal International 211(1):263–269, DOI 10.1093/gji/ggx302

Folkner WM, Williams JG, Boggs DH, Park RS, Kuchynka P (2014) The planetary and lunar ephemerides DE430 and DE431. Interplanetary Network Progress Report 196(1)

Ge M, Gendt G, Dick G, Zhang F, Rothacher M (2006) A new data processing strategy for huge GNSS global networks. Journal of Geodesy 80(4):199–203, DOI 10.1007/s00190-006-0044-x

Hauschild A, Steigenberger P, Montenbruck O (2019) Inter-receiver GNSS pseudorange biases and their effect on clock and DCB estimation. In: Proceedings of the 32nd International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2019), pp 3675–3685, DOI 10.33012/2019.16975

Hou P, Zhang B, Yuan Y (2021) Analysis of the stochastic characteristics of gps/bds/galileo multi-frequency observables with different types of receivers. Journal of Spatial Science 66(1):49–73, DOI 10.1080/14498596.2019.1578274

Håkansson M, Jensen ABO, Horemuz M, Hedling G (2017) Review of code and phase biases in multi-GNSS positioning. GPS Solutions 21(3):849–860, DOI 10.1007/s10291-016-0572-7

Jazaeri S, Amiri-Simkooei A, Sharifi M (2012) Fast integer least-squares estimation for GNSS high-dimensional ambiguity resolution using lattice theory. Journal of Geodesy 86(2):123–136, DOI 10.1007/s00190-011-0501-z

Jazaeri S, Amiri-Simkooei A, Sharifi MA (2014) On lattice reduction algorithms for solving weighted integer least squares problems: comparative study. GPS Solutions 18(1):105–114, DOI 10.1007/s10291-013-0314-z

Johnston G, Riddell A, Hausler G (2017) The International GNSS Service. In: Teunissen PJG, Montenbruck O (eds) Springer Handbook of Global Navigation Satellite Systems, Springer International Publishing, Cham, pp 967–982, DOI 10.1007/978-3-319-42928-1_33

Kvas A, Brockmann JM, Krauss S, Schubert T, Gruber T, Meyer U, Mayer-Gürr T, Schuh WD, Jäggi A, Pail R (2021) GOCO06s—a satellite-only global gravity field model. Earth System Science Data 13:99–118, DOI 10.5194/essd-13-99-2021

Landskron D, Böhm J (2018) VMF3/GPT3: refined discrete and empirical troposphere mapping functions. Journal of Geodesy 92(4):349–360, DOI 10.1007/s00190-017-1066-2

Laurichesse D, Cerri L, Berthias J, Mercier F (2013) Real time precise GPS constellation and clocks estimation by means of a Kalman filter. In: Proceedings of the 26th international technical meeting of the satellite division of the Institute of Navigation (ION GNSS+ 2013), pp 1155–1163

Leick A, Rapoport L, Tatarnikov D (2015) GPS satellite surveying, 4th edn. John Wiley & Sons

Li B, Teunissen PJG (2011) High dimensional integer ambiguity resolution: A first comparison between LAMBDA and Bernese. Journal of Navigation 64(S1):S192–S210, DOI 10.1017/S037346331100035X

Li Z, Li M, Shi C, Fan L, Liu Y, Song W, Tang W, Zou X (2019) Impact of ambiguity resolution with sequential constraints on real-time precise GPS satellite orbit determination. GPS Solutions 23(3):1–14, DOI 10.1007/s10291-019-0878-3

Mahalanobis PC (1936) On the generalized distance in statistics. In: Proceedings of the National Institute of Sciences of India, National Institute of Science of India

Mayer-Gürr T, Behzadpour S, Eicker A, Ellmer M, Koch B, Krauss S, Pock C, Rieser D, Strasser S, Süsser-Rechberger B, Zehentner N, Kvas A (2021) GROOPS: A software toolkit for gravity field recovery and GNSS processing. Computers & Geosciences p 104864, DOI 10.1016/j.cageo.2021.104864

Montenbruck O, Steigenberger P (2020) GNSS orbit determination and time synchronization. Position, Navigation, and Timing Technologies in the 21st Century: Integrated Satellite Navigation, Sensor Systems, and Civil Applications 1:233–258, DOI 10.1002/9781119458449.ch11

Nardo A, Li B, Teunissen PJG (2016) Partial ambiguity resolution for ground and space-based applications in a gps+ galileo scenario: a simulation study. Advances in Space Research 57(1):30–45

Odijk D, Arora BS, Teunissen PJG (2014) Predicting the success rate of long-baseline gps galileo (partial) ambiguity resolution. Journal of Navigation 67(3):385–401, DOI 10.1017/S037346331400006X

Petit G, Luzum B (2010) IERS conventions (2010). Tech. rep., Verlag des Bundesamts für Kartographie und Geodäsie

Prange L, Villiger A, Sidorov D, Schaer S, Beutler G, Dach R, Jaggi A (2020) Overview of CODE's MGEX solution with the focus on Galileo. Advances in Space Research 66(12):2786–2798, DOI https://doi.org/10.1016/j.asr.2020.04.038, scientific and Fundamental Aspects of GNSS - Part 1

Rodriguez-Solano CJ (2009) Impact of albedo modelling on GPS orbits. M.Sc. thesis, Technische Universität München, URL https://mediatum.ub.tum.de/doc/1368717/file.pdf

Rodriguez-Solano CJ (2014) Impact of non-conservative force modeling on GNSS satellite orbits and global solutions. PhD thesis, Technische Universität München, URL https://mediatum.ub.tum.de/doc/1188612/file.pdf

Steigenberger P, Thoelert S, Montenbruck O (2018) GNSS satellite transmit power and its impact on orbit determination. Journal of Geodesy 92(6):609–624, DOI 10.1007/s00190-017-1082-2

Strasser S, Mayer-Gurr T (2021) IGS repro3 products by Graz University of Technology (TUG). Data set available at https://doi.org/10.3217/dataset-4528-0723-0867

Strasser S, Mayer-Gurr T, Zehentner N (2019) Processing of GNSS constellations and ground station networks using the raw observation approach. Journal of Geodesy 93(7):1045–1057, DOI 10.1007/s00190-018-1223-2

Teunissen PJG (1993) Least-squares estimation of the integer GPS ambiguities. Invited Lecture, Section IV Theory and Methodology, IAG General Meeting, Beijing, China. Also in: Delft Geodetic Computing Centre LGR Series No. 6. Delft University of Technology

Teunissen PJG (1995) The least-squares ambiguity decorrelation adjustment: a method for fast GPS integer ambiguity estimation. Journal of Geodesy 70(1):65–82, DOI 10.1007/BF00863419

Teunissen PJG (1998) Success probability of integer GPS ambiguity rounding and bootstrapping. Journal of Geodesy 72(10):606–612, DOI 10.1007/s001900050199

Teunissen PJG (1999) An optimality property of the integer least-squares estimator. Journal of Geodesy 73(11):587–593, DOI 10.1007/s001900050269

Teunissen PJG (2002) A new class of GNSS ambiguity estimators. Artificial Satellites 37(4):111–120

Teunissen PJG (2003a) Integer aperture GNSS ambiguity resolution. Artificial Satellites 38(3):79–88

Teunissen PJG (2003b) Towards a unified theory of GNSS ambiguity resolution. Journal of Global Positioning Systems 2(1):1–12

Teunissen PJG, Massarweh L, Verhagen S (2021) Vectorial integer bootstrapping: Flexible integer estimation with application to GNSS. Journal of Geodesy (*accepted*)

Verhagen S, Teunissen PJG, van der Marel H, Li B (2011) GNSS ambiguity resolution: which subset to fix. In: IGNSS Symposium, pp 15–17

Verhagen S, Li B, Teunissen PJG (2012a) LAMBDA software package: Matlab implementation, version 3.0. Delft University of Technology and Curtin University, Perth, Australia

Verhagen S, Tiberius C, Li B, Teunissen PJG (2012b) Challenges in ambiguity resolution: biases, weak models, and dimensional curse. In: 2012 6th ESA Workshop on Satellite Navigation Technologies (Navitec 2012) & European Workshop on GNSS Signals and Signal Processing, IEEE, pp 1–8

Verhagen SA (2005) The GNSS integer ambiguities: Estimation and validation. PhD thesis, Delft University of Technology

Villiger A, Dach R, Schaer S, Prange L, Zimmermann F, Kuhlmann H, Wübbena G, Schmitz M, Beutler G, Jäggi A (2020) GNSS scale determination using calibrated receiver and Galileo satellite antenna patterns. Journal of Geodesy 94(9):1–13, DOI 10.1007/s00190-020-01417-0

Zhang F (2006) The Schur complement and its applications, vol 4. Springer Science & Business Media

4

New LAMBDA toolbox for mixed-integer models: Estimation and Evaluation

In this work we introduce the LAMBDA 4.0 toolbox, which provides an enhanced implementation for integer estimation, validation, and success rate evaluation. This free and open-source toolbox is a major update to LAMBDA 3.0 (2012), while it also integrates the functionalities from Ps-LAMBDA 1.0 (2013), thus respectively merging both *estimation* and *evaluation* capabilities. The new implementation provides redefined algorithms, such as an improved integer search strategy with a one-order reduction in the computational time, along with additional estimators: Vectorial Integer Bootstrapping (VIB), Integer Aperture Bootstrapping (IAB) and Best Integer Equivariant (BIE). This toolbox aims to become a valuable resource for researchers and/or practitioners dealing with mixed-integer models in high dimensions, e.g., terrestrial-based carrier-phase systems, multi-constellation Global Navigation Satellite Systems (GNSS), or other interferometric applications.

This chapter has been published as: Massarweh, L., Verhagen, S., and Teunissen, P.J.G. (2025). *New LAMBDA toolbox for mixed-integer models: Estimation and Evaluation.* In GPS Solutions 29, 14. https://doi.org/10.1007/s10291-024-01738-z

4.1 Introduction

The Least-squares AMBiguity Decorrelation Adjustment (LAMBDA) method was introduced in Teunissen (1993) as a numerically efficient and statistically optimal solution approach to Integer Ambiguity Resolution (IAR) problems. Such problems arise in many GNSS and non-GNSS mixed-integer model applications (see Table 4.1). For GNSS mixed-integer models, the millimeter-level precision of the ambiguous carrier-phase measurements can be exploited once the integer ambiguities are correctly resolved with high probability, also known as Success Rate (SR).

Table 4.1: Examples of applications dealing with IAR problems

	Application	Reference
GNSS	precise point positioning	(Teunissen, 2020)
	integer cycle-slip resolution	(Teunissen and de Bakker, 2015)
	time and frequency transfer	(Mi et al., 2023)
	atmosphere remote sensing	(Lu et al., 2018)
	carrier-phase attitude determination	(Giorgi, 2011)
	relative navigation for S/C formation flying	(Buist, 2013)
Non-GNSS	detection in MIMO communication systems	(Damen et al., 2003)
	InSAR permanent scatterer technique	(Kampes and Hanssen, 2004)
	phase data for InSAR deformation monitoring	(Teunissen, 2006)
	use of acoustic waves for underwater navigation	(Viegas and Cunha, 2007)
	fringe phase observations from VLBI	(Hobiger et al., 2009)

In this contribution we present the LAMBDA 4.0 toolbox, a new implementation dedicated to integer estimation, integer validation and SR computations. This new toolbox integrates all functionalities from LAMBDA 3.0 (Verhagen et al., 2012), and Ps-LAMBDA 1.0 (Verhagen et al., 2013), while incorporating newly redesigned algorithms that enhance the computational performance for high-dimensional problems. The novelties are not limited to computational aspects, since additional (classes of) estimators are implemented, including the Vectorial Integer Bootstrapping (Teunissen et al., 2021), Integer Aperture Bootstrapping (Teunissen, 2005) and Best Integer Equivariant (Teunissen, 2003b) solutions. In this way, we provide users with a larger collection of tools in support to several applications.

In Section 4.2, we discuss mixed-integer models, including three different classes of estimators. In Section 4.3, LAMBDA 4.0 is presented, where LAMBDA and Ps-LAMBDA capabilities are discussed, along with the software package provided to the user. In Section 4.4, we focus on 'estimation' and 'evaluation' capabilities for the new toolbox, providing some illustrative numerical examples. In Section 4.5, the main conclusions are summarized. A more detailed mathematical documentation, hereinafter referred to as 'LAMBDA Documentation', is given as supplementary material to this publication.

4.2 Mixed-integer models

We start with a linear(ized) mixed-integer model, given $y \in \mathbb{R}^m$ as vector of observables, and $Q_{yy} \in \mathbb{R}^{m \times m}$ its variance-covariance (vc-)matrix, such that

$$E\{y\} = Aa + Bb, \quad D\{y\} = Q_{yy} \tag{4.1}$$

where $E\{\cdot\}$ and $D\{\cdot\}$ are the expectation and dispersion operators, respectively. The vectors $a \in \mathbb{Z}^n$ and $b \in \mathbb{R}^p$ refer respectively to the integer ambiguities and real-valued parameters, while the full-rank design matrix is given by $[A, B] \in \mathbb{R}^{m \times (n+p)}$.

Usually, a three-step procedure is employed to solve such mixed-integer models, based on a weighted least-squares criterion, so considering the quadratic objective function

$$\|y - Aa - Bb\|^2_{Q_{yy}} = \|\hat{e}\|^2_{Q_{yy}} + \|\hat{a} - a\|^2_{Q_{\hat{a}\hat{a}}} + \left\|\hat{b}(a) - b\right\|^2_{Q_{\hat{b}(a)\hat{b}(a)}} \tag{4.2}$$

where $\|x\|^2_Q = x^T Q^{-1} x$, while $\hat{e} = y - A\hat{a} - B\hat{b}$ refers to the least-squares residuals' vector. Following this orthogonal decomposition (Teunissen, 1993), we compute

1. **Float solution**, where the integer constraints on the ambiguities are neglected, and a least-squares solution is found assuming $a \in \mathbb{R}^n$ and $b \in \mathbb{R}^p$, thus providing the float estimators $\hat{a} \in \mathbb{R}^n$ and $\hat{b} \in \mathbb{R}^p$.
2. **Ambiguity resolution**, where starting with the ambiguity float estimator $\hat{a} \in \mathbb{R}^n$ and its vc-matrix $Q_{\hat{a}\hat{a}} \in \mathbb{R}^{n \times n}$, we obtain the fixed ambiguity estimator $\breve{a} = \mathcal{J}_a(\hat{a})$ given a certain mapping function $\mathcal{J}_a$ defined for the selected estimator.
3. **Fixed solution**, where the real-valued (*float*) parameter $\hat{b} \in \mathbb{R}^p$ are now conditionally updated onto the (*fixed*) ambiguities, so leading to the fixed solution $\breve{b} = \hat{b}(\breve{a})$, while we assume no additional constraints exist onto the b-parameters.

The LAMBDA method (Teunissen, 1995) exploits the ambiguity re-parametrization by an admissible transformation matrix $Z \in \mathbb{Z}^{n \times n}$, where $\hat{z} = Z^T \hat{a}$ with $\hat{z} \in \mathbb{R}^n$, which serves to decorrelate the ambiguity components. This Z-transformation process increases SR for some estimators, while the SR is invariant for the Integer Least-Squares (ILS) estimator, whereas its computational performance can be enhanced.

4.2.1 Different classes of estimators

In Teunissen (2003c), three classes of estimators are introduced:

I. **Integer estimators**, i.e. I-class;

II. **Integer Aperture estimators**, i.e. IA-class;

III. **Integer Equivariant estimators**, i.e. IE-class;

with a set relationship I $\subset$ IA $\subset$ IE, as shown in Fig. 4.1 with one example per class.

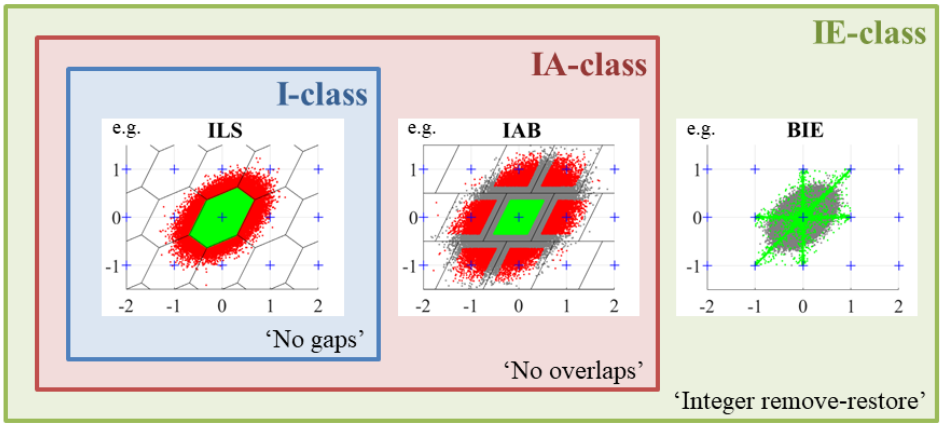


Figure 4.1: Overview of the three main classes and their set relationship, including an example of estimator per each class: Integer Least-Squares (ILS), Integer Aperture Bootstrapping (IAB) and Best Integer Equivariant (BIE).

Each class is characterized by different properties based on the family of maps $\mathcal{J}_a$ being defined. The I-class (Teunissen, 1999) is the most restrictive class, with many-to-one mapping functions defining *pull-in regions* that leave no gaps, have no overlaps and are invariant to integer translations. The IA-class (Teunissen, 2003a) is defined by dropping the 'no gaps' condition, and so-called *aperture regions* can be defined. The IE-class (Teunissen, 2002) is the largest class, where the 'no overlaps' condition is relaxed, while the 'integer remove-restore' property is still satisfied. We refer to Verhagen (2005) and 'LAMBDA Documentation' for more mathematical details.

4.3 LAMBDA 4.0 description

The new toolbox merges functionalities from LAMBDA 3.0 (Verhagen et al., 2012) and Ps-LAMBDA 1.0 (Verhagen et al., 2013), while providing advanced algorithms and additional estimators. A flowchart is shown in Fig. 4.2 with both LAMBDA and Ps-LAMBDA routines, respectively adopted for 'estimation' and 'evaluation'.

4.3.1 LAMBDA capabilities

The main LAMBDA script takes as inputs the $\hat{a}$-vector and $Q_{\hat{a}\hat{a}}$-matrix, along with relevant configuration parameters. The decorrelation starts with an $L^T DL$-decomposition, followed by a Z-transformation given an admissible matrix $Z \in \mathbb{Z}^{n\times n}$ (unimodular), so $Q_{\hat{z}\hat{z}} = Z^T Q_{\hat{a}\hat{a}} Z$. This matrix Z also sorts ambiguities based on their conditional variances, so enabling the selection of a suitable subset of most precise components needed for Partial Ambiguity Resolution (PAR). An overview of LAMBDA-related methods is given in Fig. 4.3.

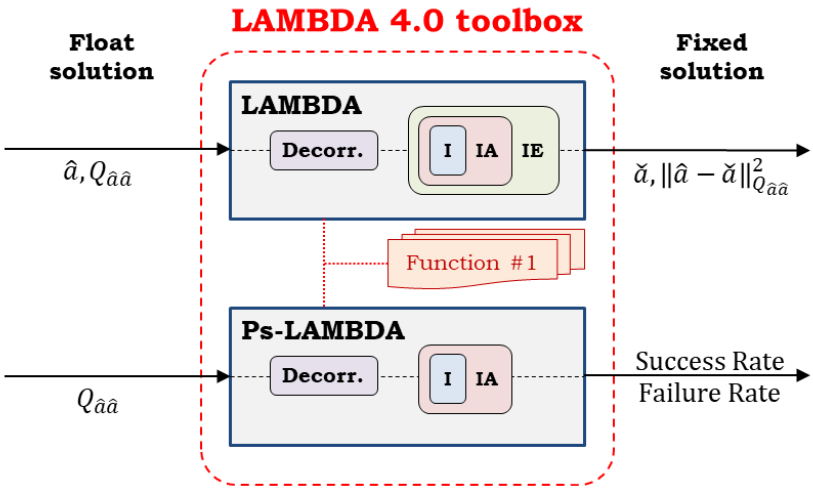


Figure 4.2: Flowchart of the LAMBDA 4.0 toolbox.

A new integer search strategy is adopted (Section 4.4.2.1), where a one-order reduction in computational time is possible, with large computational benefits for high-dimensional problems. Moreover, new coefficients (Hou et al., 2016) are adopted for the critical value of the Fixed Failure-rate Ratio Test (FFRT) solution, based on fitting functions. Lastly, new estimators are introduced:

- **I-class:** VIB (Teunissen et al., 2021), based on IR or ILS estimators, while supporting any arbitrary block partitioning of the whole ambiguity set;
- **IA-class:** IAB (Teunissen, 2005), for a user-defined aperture parameter $\beta_{IAB} \in (0, 1]$ or a maximum bootstrapping failure rate selected by the user;
- **IE-class:** BIE (Teunissen, 2003b), approximated here by a finite summation within a search ellipsoid for a user-defined significance level $\alpha_{BIE} \ll 1$.

4.3.2 Ps-LAMBDA capabilities

The main Ps-LAMBDA script takes as input only the $Q_{\hat{a}\hat{a}}$-matrix, along with a flag to enable the decorrelation of ambiguities, so users can still check the impact of different decorrelation strategies on each SR bound/approximation. The overview of Ps-LAMBDA-related methods is given in Fig. 4.4, while we refer to Verhagen et al. (2013) for more details.

The different methods serve as lower/upper bounds (LB/UB), or even approximation for the SR, while we recall the performance ordering (Teunissen et al., 2021):

$$\text{IR} \leq \text{VIB}_{\text{IR}} \leq \text{IB} \leq \text{VIB}_{\text{ILS}} \leq \text{ILS} \tag{4.3}$$

with Integer Rounding (IR) and Integer Least-Squares (ILS), which might also be

LAMBDA	Class	Description	Options
estimatorIR	I	**Integer Rounding**	-
estimatorIB	I	**Integer Bootstrapping**	-
estimatorILS	I	**Integer Least Squares** (search-and-shrink)	• # of candidates
estimatorILS_enum	I	**Integer Least Squares** (enumeration)	• # of candidates
estimatorPAR	I	**Partial Ambiguity Resolution** (with ILS estimator)	• # of candidates • Minimum SR
estimatorVIB	I	**Vectorial Integer Bootstrapping** (with IR or ILS estimator)	• Use IR or ILS • Partitioning of blocks
estimatorIA_RT	IA	**Integer Aperture w/ Ratio Test**	• Aperture parameter μ_{RT} • Maximum FR
estimatorIAB	IA	**Integer Aperture Bootstrapping**	• Aperture parameter β_{IAB} • Maximum FR
estimatorBIE	IE	**Best Integer Equivariant** (finite approximation)	• Significance level α_{BIE}

Figure 4.3: The LAMBDA functions implemented in LAMBDA 4.0 toolbox are listed, including their class, a short description and some relevant configuration options.

Ps-LAMBDA	Description	IR	VIB-IR	IB	VIB-ILS	ILS
IBexact	**Integer Bootstrapping** (exact SR, analytical)	UB	UB	exact	LB	LB
ADOPapprox	**ADOP method** (approximation)	UB	UB	UB	approx.	approx.
LB_Variance	**Variance method** (Lower Bound)	LB	LB	LB	LB	LB
UB_ADOP	**ADOP method** (Upper Bound)	UB	UB	UB	UB	UB
LB_Eigenvalue	**Eigenvalue method** (Lower Bound)	-	-	-	-	LB
UB_Eigenvalue	**Eigenvalue method** (Upper Bound)	UB	UB	UB	UB	UB
LB_Pullin	**Pull-in region method** (Lower Bound)	-	-	-	-	LB
UB_Pullin	**Pull-in region method** (Upper Bound)	UB	UB	UB	UB	UB
Numerical	**Numerical** (simulation-based)	approx.	approx.	approx.	approx.	approx.

Figure 4.4: The Ps-LAMBDA functions implemented in LAMBDA 4.0 toolbox are listed here, including a short description and their relation to popular Integer estimators.

applied vectorially as VIB_{IR} and VIB_{ILS} estimators, respectively. An exact SR can be computed only for Integer Bootstrapping (IB), while numerical simulations are available as approximation for all I- or IA-estimators of this toolbox, and Ps-LAMBDA performance can benefit from the enhanced algorithms implemented here.

4.3.3 Software package

The LAMBDA 4.0 software package contains different files/folders (see Table 4.2). Both 'LAMBDA' and 'Ps_LAMBDA' main scripts make use of the functionalities included in a dedicated "LAMBDA_toolbox" folder. However, more advanced users might benefit from directly using these functions in their software without the need to follow the logic defined by these main scripts. These can still easily be included into existing software applications, e.g., as for LAMBDA legacy versions already integrated in the "raPPPid" software package (Glaner and Weber, 2023).

Table 4.2: List of the principal files and/or folders in this software package

File/Folder	Description
README.txt	Text file with content overview
LAMBDA	Main script with **LAMBDA** functionalities
Ps_LAMBDA	Main script with **Ps-LAMBDA** functionalities
\ LAMBDA_toolbox	Folder with all toolbox functionalities
\ LAMBDA_examples	Folder with some illustrative examples
\ LAMBDA_papers	Folder with some relevant publications
LAMBDA – Documentation.pdf	PDF with main mathematical descriptions
LAMBDA – User Manual.pdf	PDF with the toolbox guide (per environment)

A detailed description of these new functionalities can be found in the supplementary material (see 'LAMBDA Documentation'). The current MATLAB implementation is based on v2024a, and it does not require any additional toolboxes. Moreover, this has been tested for backward compatibility starting with version 2018b, while a Python 3 version, including its dedicated user manual, is currently under development.

4.4 LAMBDA 4.0: estimation and evaluation

4.4.1 Estimation in the I-class, IA-class and IE-class

This LAMBDA 4.0 implementation offers a versatile toolbox with several estimators suitable for different problems, where users have access to three classes of estimators (see Fig. 4.5).

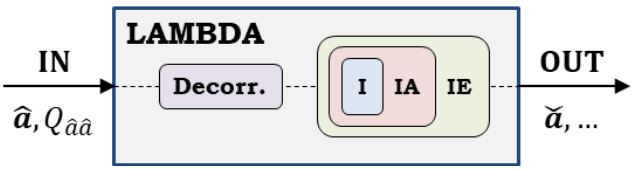


Figure 4.5: LAMBDA main script supporting different classes of estimators.

In some applications, it is important that users have control over the failure rate, and two different solutions could achieve this: a model-driven I-estimator and a data-driven IA-estimator. The former one uses PAR, where only part of the ambiguity vector is estimated, whereas the latter considers part of the pull-in region for the integer mapping. Fig. 4.6 shows a graphical example of IA-estimator based on ratio test, where a stronger and a weaker model are considered respectively in the left and right plots. The selection of the aperture parameter μ is fundamental for correctly 'controlling' the failure rate, where different aperture values are adopted for the two models, while still having a similar failure rate.

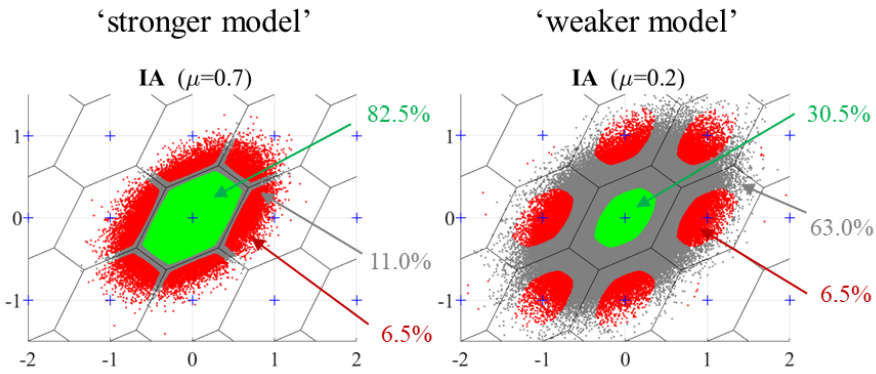


Figure 4.6: A comparison between a stronger and a weaker model is given using two different ratio tests, respectively for $\mu = 0.7$ and $\mu = 0.2$. Three regions can be observed, related to values of Success Rate (green), Failure Rate (red) and Undecided Rate (grey).

Another example concerns vectorial IB solutions, more suitable for high-dimensional problems. A VIB-based method is currently implemented in GROOPS software (Mayer-Gürr et al., 2021), where it is possible to resolve thousands of ambiguity components within a few tenths of a second in the context of small global networks for Orbit Determination and Time Synchronization (ODTS), see Massarweh et al. (2021).

Within the IE-class, the BIE solution is optimal in the minimum Mean Squared Error (MSE) sense. This property is beneficial for precise positioning applications, as shown by Odolinski and Teunissen (2020) in the context of multi-GNSS Real-Time Kinematics (RTK) using low-cost receivers, as well as by Yang et al. (2024) for PPP-AR applications. These BIE solutions will always outperform – in the MSE sense – both float and fixed solutions.

A numerical simulation is shown in Fig. 4.7, based on GPS (L1-only) single-baseline relative (kinematic) positioning with 6 satellites tracked (at 30s). In this illustrative example, we neglect the atmospheric delays (e.g., short baseline) and we adopt an elevation weighting $\propto 1/\sin(el)$, while having 30 cm and 3 mm standard deviation for the undifferenced code and phase observations (at zenith). Using 100,000 simulations, we compute the MSE ratio of different 'fixed' solutions with respect to the 'float' solution. A PAR solution is also shown in Fig. 4.7, based on an SR>99.0% criterion,

where only a subset of the ambiguities can be fixed, and improvements compared to the 'float' are negligible during the first epochs. At low SR values, the ILS estimator – optimal in terms of SR – might also lead to worse-than-float performance, whereas the BIE estimator is capable of outperforming all these solutions and it provides a minimum MSE value for the positioning coordinates computed here. We refer to Brack et al. (2023) for additional BIE-related results.

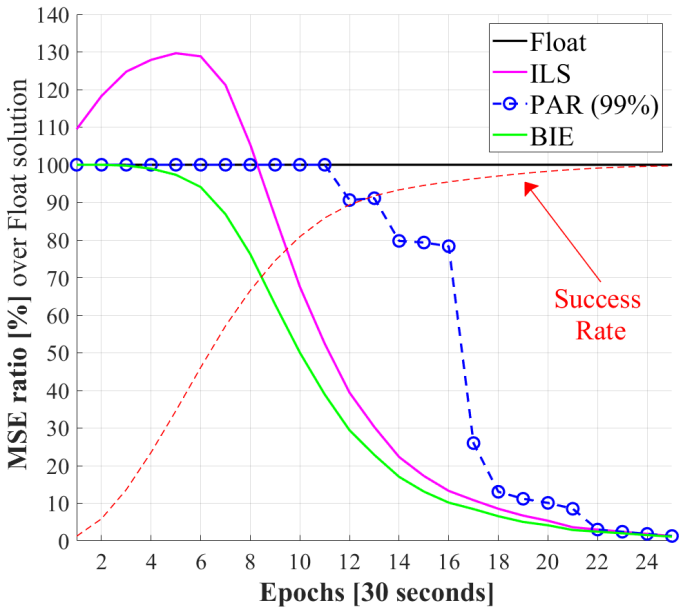


Figure 4.7: The ratio [%] of MSE of the baseline positioning coordinates with respect to the float solution is shown for different estimators based on GPS single-baseline single-frequency kinematic positioning, while using 30 cm/3 mm standard deviation for the undifferenced code/phase observations.

4.4.2 Numerical and statistical evaluations

For experimental and designing purposes, it is important to compute the relevant statistical measure of the different estimators, e.g. SR for the I-class, or also FR for the IA-class, as enabled in an efficient way by the new toolbox (see Fig. 4.8).

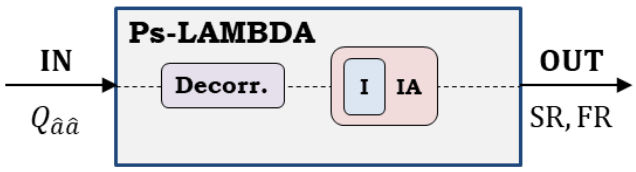


Figure 4.8: Ps-LAMBDA main script supporting different classes of estimators.

4.4.2.1 Numerical evaluations

In terms of numerical performance, the integer search strategy introduced by Ghasemmehdi and Agrell (2011) has been adopted, which is beneficial for both estimation and evaluation functionalities. The main advantage of this new search strategy is the removal of non-necessary conditioning operations when spanning (down) the search tree, see (ibid). An example is provided in Fig. 4.9, considering the old (in blue) and new (in red) search algorithms; see also Jazaeri et al. (2014).

The numerical test in Fig. 4.9 is based on a single-baseline geometry-free ionosphere-fixed model using GPS L1+L2 signals, with a standard deviation of 20cm/2mm respectively for the undifferenced code/phase observations. The number of satellites varies between 2 and 100 to increase the problem dimensionality, with larger improvements expected at higher dimensionalities.

The computational time is shown averaged over 1000 runs (top-left plot), along with the ratio between LAMBDA 3.0 and 4.0 (bottom-left plot). Moreover, the number of conditioning operations is illustrated (right plot), showing the results for each single run (dots), based on different float samples, along with their averages (circles). This new strategy does not only enhance I-estimators, such as ILS[1], PAR and VIB-ILS, but it has a positive impact also on Ps-LAMBDA functionalities, e.g., for SR/FR simulations.

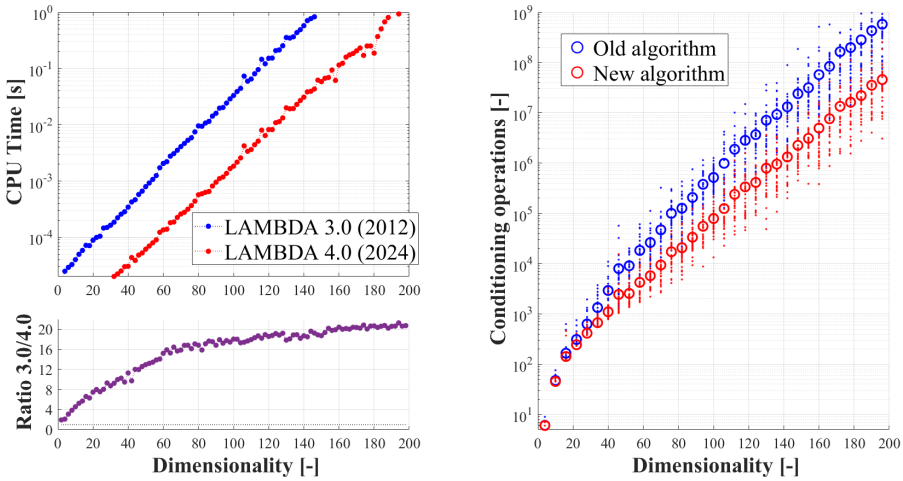


Figure 4.9: A numerical comparison is shown between LAMBDA 3.0 (OLD) and LAMBDA 4.0 (NEW) algorithm in terms of search computations for the optimal ILS solution.

[1]This 2011-algorithm has also been adopted in the MLAMBDA implementation starting from its 2016 version (https://www.cs.mcgill.ca/~chang/MLAMBDA.php), but limited only to the ILS estimator.

4.4.2.2 Statistical evaluations

In Fig. 4.10, we compare a few different SR bounds and approximations with respect to the ILS one, which has been numerically computed for the problem adopted in Fig. 4.7 In this example, the pull-in region method (in green) provides a close upper bound to ILS estimator (red line) for high SR values, while the variance method (in blue) is a close lower bound to IR estimator (orange square symbol). See Wang et al. (2016) for additional examples.

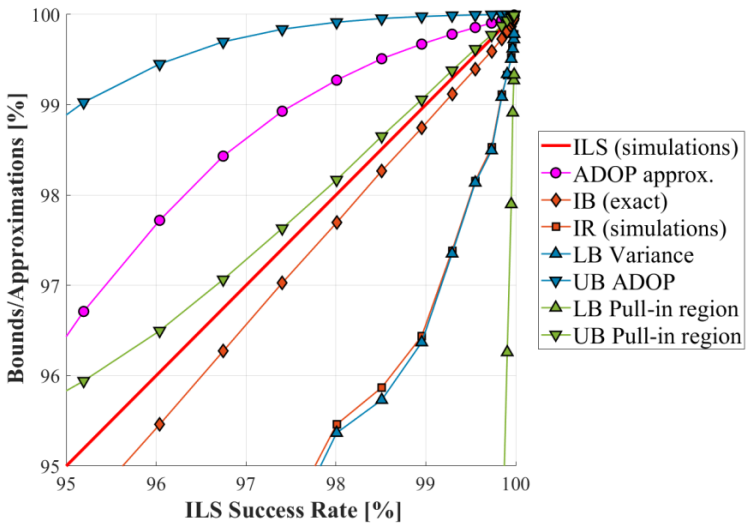


Figure 4.10: A few different bounds or approximations are compared with respect to the SR of ILS taken here as reference (in the range 95-100%), based on the example from Fig. 4.7, while IR and ILS numerical simulations are computed using 10^6 samples.

Moreover, two important cases could be considered (Teunissen, 2000):

1) SR sensitivity for changes in the stochastic model of observations;

2) SR sensitivity when using an incorrect stochastic model of observations;

In the first case, this analysis allows users to calculate how a precision improvement of their data would affect the SR. In the second case, it provides users with a way to understand how accurate they would need to know their observables' vc-matrix to maintain quasi-optimal performance in the ambiguity fixing. Hence, this provides a diagnostics to assess the impact on SR due to stochastic misspecifications.

We start with the first case, where we consider in Fig. 4.11 the model adopted in Fig. 4.7, and we vary the standard deviation for the undifferenced phase (left plot) or code (right plot) observations. Then, we compute the SR of ILS based on 10^6 samples, and we focus on the range 80-to-100%. A similar sensitivity analysis can be performed with other I- and IA- estimators, as well as considering bounds and/or approximations. See Nardo et al. (2016) for additional examples.

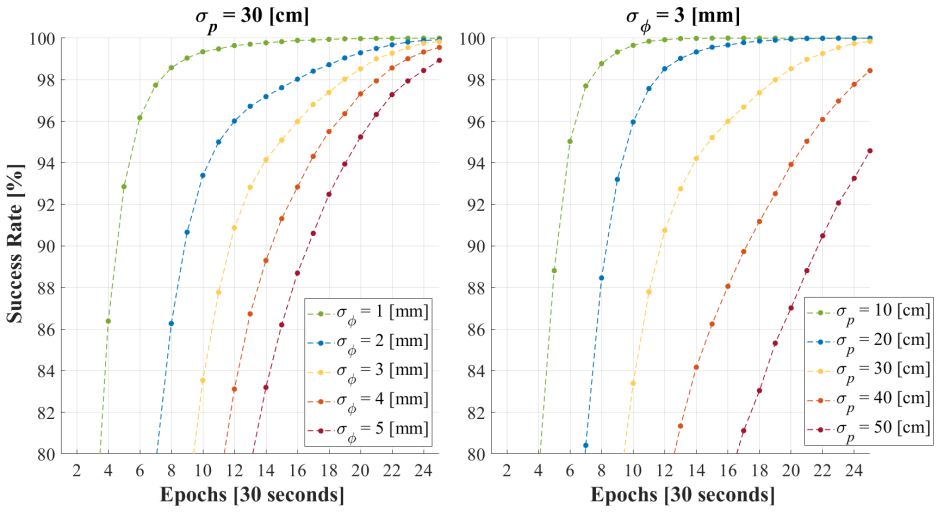


Figure 4.11: The SR of ILS is computed using 10^6 samples, based on the example of Fig. 4.7, where standard deviation for the undifferenced phase σ_ϕ (left plot) and code σ_p (right plot) observations has been varied. An elevation weighting scheme is adopted in all simulations.

For the second case, we present an example on the adoption of an incorrect stochastic model, which can negatively affect IAR performance. We consider a (short) single-baseline GPS (L1+L2) model, with 7 satellites tracked on a single epoch, having 30cm/3mm for the standard deviation of undifferenced code/phase data, respectively. The correct stochastic model is based on an elevation weighting scheme $\propto 1/\sin(el)$, while for the float solution computation we consider an incorrect stochastic model without any elevation weighting. The Ps-LAMBDA functionalities enable users to compare the correct and incorrect stochastic models, so computing the failure rate for an IA-estimator, e.g., with Ratio Test, while varying its aperture parameter $\mu \leq 1$.

The results are shown in Fig. 4.12, where SR of ILS is 92.9% or 77.3% when a correct or an incorrect model is used, respectively. Given different μ-values between 0.5 and 1.0, we can compute the FR, with a user-defined threshold (magenta dashed line) set here to 1%. Given $\mu = 0.7$, e.g., the FR is 0.9% or 3.0%, and the maximum FR threshold condition will be fulfilled only with the correct stochastic model.

4.5 Conclusions

In this contribution we introduce the LAMBDA 4.0 toolbox, a new implementation for the integer estimation, integer validation and success rate computation. This new toolbox builds its foundations over two legacy software implementations: LAMBDA 3.0 and Ps-LAMBDA 1.0, respectively introduced by Verhagen et al. (2012) and by Verhagen et al. (2013). Several modular functionalities are provided and can be used standalone to be easily integrated into existing GNSS or non-GNSS software.

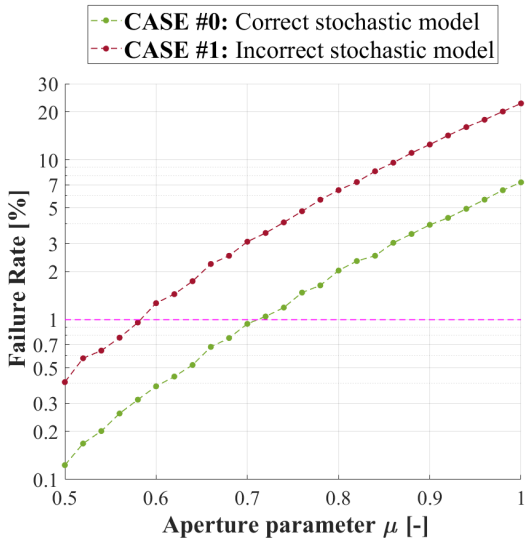


Figure 4.12: The failure rate is numerically computed for a (short) single-baseline GPS (L1+L2) model based on an Integer Aperture estimator with Ratio Test, where different values of the aperture parameter μ are used. An arbitrary user-defined threshold ($FR < 1\%$) is shown in magenta.

The LAMBDA 4.0 toolbox enhances the computational performances based on new routines to tackle any mixed-integer problem. For instance, an improved search strategy for optimal Integer Least-Squares (ILS) solutions has been implemented and becomes beneficial in many other Ps-LAMBDA functionalities. Moreover, new (classes of) estimators have been introduced, e.g., a flexible Vectorial Integer Bootstrapping (VIB) formulation for the high-dimensional problems, an Integer Aperture Bootstrapping (IAB) solution for the controlled failure rate applications, an optimal BIE solution that always outperforms – in the minimum Mean Squared Error sense – both float and fixed solutions.

The newly developed toolbox comes in response to the current need for advanced solutions to tackle high-dimensional problems in both an efficient and reliable way. Hence, this LAMBDA 4.0 toolbox can support many applications, e.g. for multi-GNSS and multi-frequency mixed-integer models arising in PPP-RTK network estimation (Teunissen and Khodabandeh, 2015), frequency-varying carrier phase signals (Khodabandeh and Teunissen, 2023) and in view to the future deployment of large satellite constellations in low-Earth orbit (LEO) tracked on ground. This landscape of possibilities is not limited to GNSS applications but it is expected to benefit all researchers dealing with mixed-integer models, thus providing a simple, versatile and effective tool to the research community.

References

Brack A, Männel B, Schuh H (2023) Two-epoch centimeter-level PPP-RTK without external atmospheric corrections using best integer-equivariant estimation. GPS Solutions 27:12

Buist P (2013) Multi-platform integrated positioning and attitude determination using GNSS. PhD thesis, Delft University of Technology, Delft, The Netherlands

Damen M, El Gamal H, Caire G (2003) On maximum-likelihood detection and the search for the closest lattice point. IEEE Transactions on Information Theory 49(10):2389–2402

Ghasemmehdi A, Agrell E (2011) Faster recursions in sphere decoding. IEEE Transactions on Information Theory 57(6):3530–3536

Giorgi G (2011) GNSS carrier phase-based attitude determination: Estimation and applications. PhD thesis, Delft University of Technology, Delft, The Netherlands

Glaner M, Weber R (2023) An open-source software package for precise point positioning: raPPPid. GPS Solutions 27:174

Hobiger T, Sekido M, Koyama Y, Kondo T (2009) Integer phase ambiguity estimation in next-generation geodetic very long baseline interferometry. Advances in Space Research 43(1):187–192

Hou Y, Verhagen S, Wu J (2016) An efficient implementation of fixed failure-rate ratio test for GNSS ambiguity resolution. Sensors 16(7):945

Jazaeri S, Amiri-Simkooei A, Sharifi M (2014) Modified weighted integer least squares estimations for GNSS IAR. Survey Review 46(335):112–121

Kampes B, Hanssen R (2004) Ambiguity resolution for permanent scatterer interferometry. IEEE Transactions on Geoscience and Remote Sensing 42(11):2446–2453

Khodabandeh A, Teunissen PJG (2023) Ambiguity-fixing in frequency-varying carrier phase measurements: Global navigation satellite system and terrestrial examples. NAVIGATION: Journal of the Institute of Navigation 70(2)

Lu C, Li X, Cheng J, Dick G, Ge M, Wickert J, Schuh H (2018) Real-time tropospheric delay retrieval from multi-GNSS PPP ambiguity resolution: Validation with final troposphere products and a numerical weather model. Remote Sensing 10(3):481

Massarweh L, Strasser S, Mayer-Gürr T (2021) On vectorial integer bootstrapping implementations in the estimation of satellite orbits and clocks based on small global networks. Advances in Space Research 68(11):4303–4320

Mayer-Gürr T, Behzadpour S, Eicker A, Ellmer M, Koch B, Krauss S, Pock C, Rieser D, Strasser S, Süsser-Rechberger B, et al. (2021) GROOPS: a software toolkit for gravity field recovery and GNSS processing. Computers and Geosciences 155:104864

Mi X, Zhang B, El-Mowafy ea A (2023) Undifferenced and uncombined GNSS time and frequency transfer with integer ambiguity resolution. Journal of Geodesy 97:13

Nardo A, Li B, Teunissen PJG (2016) Partial ambiguity resolution for ground and space-based applications in a gps+ galileo scenario: a simulation study. Advances in Space Research 57(1):30–45

Odolinski R, Teunissen PJG (2020) Best integer equivariant estimation: performance analysis using real data collected by low-cost, single- and dual-frequency, multi-GNSS receivers for short- to long-baseline RTK positioning. Journal of Geodesy 94:91

Teunissen PJG (1993) Least-squares estimation of the integer gps ambiguities. In: IAG General Meeting, Invited Lecture. Section IV Theory and Methodology

Teunissen PJG (1995) The least-squares ambiguity decorrelation adjustment: a method for fast gps integer ambiguity estimation. Journal of Geodesy 70(1):65–82

Teunissen PJG (1999) The probability distribution of the gps baseline for a class of integer ambiguity estimators. Journal of Geodesy 73:275–284

Teunissen PJG (2000) The success rate and precision of gps ambiguities. Journal of Geodesy 74:321–326

Teunissen PJG (2002) A new class of GNSS ambiguity estimators. Artificial Satellites 37(4):111–120

Teunissen PJG (2003a) Integer aperture GNSS ambiguity resolution. Artificial Satellites 38(3):79–88

Teunissen PJG (2003b) Theory of integer equivariant estimation with application to GNSS. Journal of Geodesy 77:402–410

Teunissen PJG (2003c) Towards a unified theory of GNSS ambiguity resolution. Journal of Global Positioning Systems 2(1):1–12

Teunissen PJG (2005) Integer aperture bootstrapping: a new GNSS ambiguity estimator with controllable fail-rate. Journal of Geodesy 79:389–397

Teunissen PJG (2006) On insar ambiguity resolution for deformation monitoring. Artificial Satellites 41(1):19–22

Teunissen PJG (2020) GNSS precise point positioning. In: Position, Navigation, and Timing Technologies in the 21st Century: integrated satellite navigation, sensor systems, and civil applications, vol 1, Wiley, pp 503–528

Teunissen PJG, de Bakker P (2015) Multivariate integer cycle-slip resolution: A single-channel analysis. In: Sneeuw N, Novák P, Crespi M, Sansò F (eds) VIII Hotine-Marussi Symposium on Mathematical Geodesy. International Association of Geodesy Symposia, vol 142, Springer, Cham

Teunissen PJG, Khodabandeh A (2015) Review and principles of PPP-RTK methods. Journal of Geodesy 89:217–240

Teunissen PJG, Massarweh L, Verhagen S (2021) Vectorial integer bootstrapping: flexible integer estimation with application to GNSS. Journal of Geodesy 95:99

Verhagen S (2005) The GNSS integer ambiguities: Estimation and validation. PhD thesis, Delft University of Technology, Delft, The Netherlands

Verhagen S, Li B, Teunissen PJG (2012) LAMBDA–Matlab implementation, version 3.0. Delft University of Technology and Curtin University

Verhagen S, Li B, Teunissen PJG (2013) Ps-lambda: Ambiguity success rate evaluation software for interferometric applications. Computers & Geosciences 54:361–376

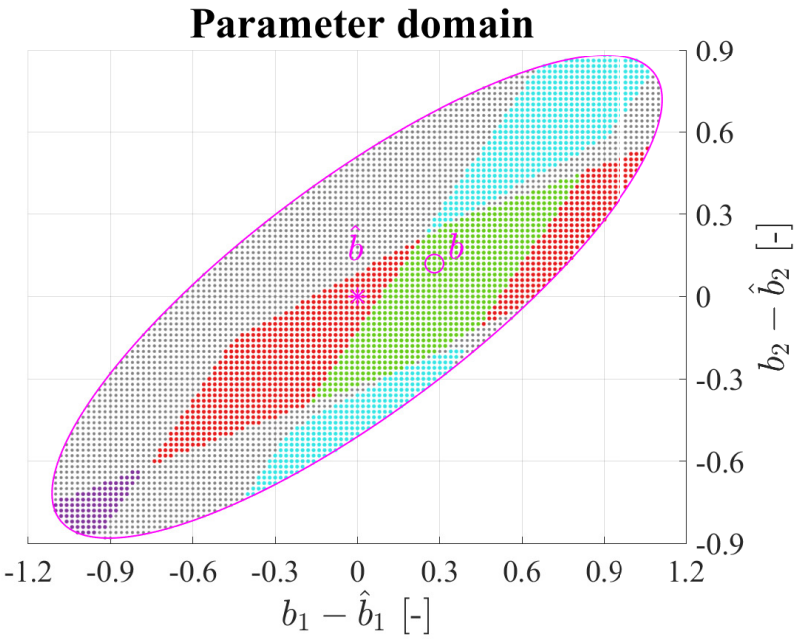
Viegas D, Cunha S (2007) Precise positioning by phase processing of sound waves. IEEE Transactions on Signal Processing 55(12):5731–5738

Wang L, Feng Y, Guo J, Wang C (2016) Impact of decorrelation on success rate bounds of ambiguity estimation. Journal of Navigation 69(5):1061–1081

Yang Y, Zhou F, Song S (2024) Improving precise point positioning (PPP) performance with best integer equivariant (BIE) estimator. GPS Solutions 28:50

II

Dual formulation



5

Primal and Dual Mixed-Integer Least-Squares: Distributional Statistics and Global Algorithm

In this contribution we introduce the dual mixed-integer least-squares problem and study it in relation to its primal counterpart. The dual differs from the primal formulation in the order in which the integer ambiguity vector $a \in \mathbb{Z}^n$ and baseline vector $b \in \mathbb{R}^p$ are estimated. As not the ambiguities, but rather the entries of b are usually the parameters of interest, the attractiveness of the dual formulation stems from its direct computation of b. It is shown that this potential advantage relies on the ease with which an implicit integer least-squares problem of the dual can be solved.

For the convoluted cases, we introduce two methods of simplifying approximations. To be able to describe their quality, we provide a complete distributional analysis of their estimators, thus allowing users to judge whether or not the approximations are acceptable for their application. It is shown that this approach implicitly introduces a new class of admissible integer estimators of which we also determine the pull-in regions.

As the dual function is shown to lack convexity, special care is required to be able to compute its global minimizer $\check{b}$. Our proposed method, which has finite termination with a guaranteed ϵ-tolerance, is constructed from combining the branch-and-bound principle, with a special convex-relaxation of the dual, to which the projected-gradient-descent method is applied to obtain the required bounds. Each of the method's three constituents are described, whereby special emphasis is given to the construction of the required continuously differentiable, convex lower bounding function of the dual.

This chapter has been published as: Teunissen, P.J.G., and Massarweh, L. (2024). *Primal and dual mixed-integer least-squares: distributional statistics and global algorithm.* In Journal of Geodesy 98, 63. https://doi.org/10.1007/s00190-024-01862-1

5.1 Introduction

The mixed-integer model forms the basis for ultraprecise GNSS parameter estimation (Hofmann-Wellenhof et al., 2008; Leick et al., 2015; Teunissen and Montenbruck, 2017). Characteristically the mixed-integer least-squares model parameters are usually computed in the order of first the integer ambiguities and then the ambiguity-resolved baseline parameters. There is in principle however no a-priori reason for this particular order. In this contribution we study the dual mixed-integer least-squares problem by reversing the computational order of the ambiguities and baseline parameters. This has the potential advantage of a direct computation of the baseline vector, without the need of an explicit computation of the resolved integer ambiguities. We study the opportunities and drawbacks of this approach, and show that certain approximations of the dual problem may have practical potential under specified conditions. We provide a complete distributional analysis of their estimators, thus allowing users to judge whether or not the approximations are acceptable for their application. We also develop the algorithmic details to ensure that the global minimizer of the dual function can be computed.

This contribution is organized as follows.

Section 5.2 provides a brief review of integer least-squares (ILS) ambiguity resolution, together with the distributional properties of the ambiguity- and baseline-estimators. The dual mixed-ILS formulation is introduced in Section 5.3, together with a representation of its objective function $\mathcal{D}(b)$. It is shown that it implicitly also relies on an ILS-problem, albeit with a metric driven by the more precise conditional ambiguity variance matrix. The potential advantage of the dual formulation in solving for b directly, relies therefore on the ease with which this implicit ILS problem can be solved. For the purpose of alleviating the computational demand on the implicit ILS problem, two approximations to the dual are introduced in Sections 5.4 and 5.5, respectively. The dual approximation of Section 5.4 consists of approximating the conditional ambiguity variance matrix. It is shown to which primal formulation this approximate dual belongs and a complete distributional description of its estimators, together with success-rate bounds, is provided. The approximation of Section 5.5 consists of replacing the implicit ILS-estimaor of the dual function by a simpler integer map. It is shown that as a result a new class of admissible integer estimators is found. Also for this class a distributional description of its estimators, together with success-rate bounds, is provided. With the purpose of providing insight in the challenge of minimizing $\mathcal{D}(b)$, Section 5.6 illustrates and describes the multimodality of the dual function. As the dual function lacks convexity, special algorithmic care is required to find its global minimizer. We present our proposed method in Section 5.7. It has finite termination with a guaranteed ϵ-tolerance and it is constructed from combining the branch-and-bound principle, with a special convex-relaxation of the dual, to which the projected-gradient-descent method is applied to obtain the required bounds. As the described approach of our proposed method is not restricted to the presented dual formulation, we provide an outlook for the constrained and partitioned dual problems in Section 5.8. Finally, Section 5.9 contains the Summary and Conclusions.

The following notation is used: $\mathsf{E}(\cdot)$ and $\mathsf{D}(\cdot)$ stand for the expectation and dispersion operator, respectively, and $\mathcal{N}_p(\mu, Q)$ denotes a p-dimensional, normally distributed random vector, with mean (expectation) μ and variance matrix (dispersion) Q. $\mathbb{R}^p$ and $\mathbb{Z}^p$ denote the p-dimensional spaces of real- and integer numbers, respectively, and the range space of a matrix M is denoted as $\mathcal{R}(M)$. The least-squares (LS) inverse of a full column rank matrix M is denoted as $M^+ = (M^T Q_{yy}^{-1} M)^{-1} M^T Q_{yy}^{-1}$ and the orthogonal projector onto $\mathcal{R}(M)$ as $P_M = MM^+$. $P_M^\perp = I - P_M$ is then the orthogonal projector that projects orthogonally on the orthogonal complement of $\mathcal{R}(M)$. The Q-weighted squared norm is denoted as $||.||_Q^2 = (.)^T Q^{-1}(.)$, and $\lceil x \rfloor$ denotes the rounding of x to the nearest integer. If applied to a vector, the rounding is understood to apply to each of its coordinates. $\cup$ and $\cap$ denote the union and intersection operators, and the vectorial inequality $\preceq$ denotes the all componentwise inequality $\leq$. $\mathsf{P}[\mathcal{A}]$ denotes the probability of event $\mathcal{A}$, $f_{\hat{b}}(b)$ the probability density function (PDF) of the continuous random vector $\hat{b}$ and $\mathsf{P}[\check{a} = z]$ the probability mass function (PMF) of the integer random vector $\check{a}$. The noncentral Chi-square distribution with p degrees of freedom and noncentrality parameter λ is denoted as $\chi^2(p, \lambda)$ and its δ-percentage critical value as $\chi^2_\delta(p, 0)$.

5.2 Brief review of ILS ambiguity resolution

We start from the mixed-integer model of (linearized) GNSS observation equations (Leick et al., 2015; Teunissen and Montenbruck, 2017), which in vector-matrix form reads,

$$\mathsf{E}(y) = Aa + Bb \ , \ \mathsf{D}(y) = Q_{yy} \tag{5.1}$$

with $y \sim \mathcal{N}_m(\mathsf{E}(y), \mathsf{D}(y))$ the m-vector of normally distributed pseudorange and carrier-phase observables, $[A, B] \in \mathbb{R}^{m \times (n+p)}$ the given design matrix of full rank $n+p$, $a \in \mathbb{Z}^n$ the unknown *a*mbiguity vector consisting of the integer carrier-phase ambiguities, $b \in \mathbb{R}^p$ the unknown *b*aseline vector consisting of the remaining real-valued parameters, such as e.g., position coordinates, atmosphere parameters, receiver/satellite clock parameters, and instrumental biases, and $Q_{yy} \in \mathbb{R}^{m \times m}$ the given positive-definite variance matrix of the observables. The above GNSS model may be given in undifferenced, single-differenced or double-differenced form. In any of these forms, the possible rank-defects in the design matrix are assumed eliminated through a careful reparametrization in clearly defined estimable parameters (Odijk et al., 2015; Teunissen, 2019).

The mixed integer least-squares (ILS) estimation of the integer ambiguity vector $a \in \mathbb{Z}^n$ and the real-valued baseline vector $b \in \mathbb{R}^p$ is executed in three steps (float-integer-fixed). In the first step, the integer constraint on a is discarded, giving the so-called *float*-solution of a and b as

$$\hat{a} = \bar{A}^+ y \text{ and } \hat{b} = \bar{B}^+ y \tag{5.2}$$

with $\bar{A} = P_B^\perp A$ and $\bar{B} = P_A^\perp B$. In the second step, the integer constraint $a \in \mathbb{Z}^n$ is invoked, and $\hat{a} \in \mathbb{R}^n$ of (5.2) is used as input to obtain the *integer* estimate of the

ambiguity vector a as

$$\breve{a} = \arg\min_{a \in \mathbb{Z}^n} ||\hat{a} - a||^2_{Q_{\hat{a}\hat{a}}} \tag{5.3}$$

with $Q_{\hat{a}\hat{a}} = (\bar{A}^T Q_{yy}^{-1} \bar{A})^{-1}$ being the variance matrix of $\hat{a}$. Once the integer solution (5.3) has been obtained, the expression of the *conditional* least-squares (LS) baseline estimator, $\hat{b}(a) = \hat{b} - Q_{\hat{b}\hat{a}} Q_{\hat{a}\hat{a}}^{-1}(\hat{a} - a)$ (i.e. conditioned on knowing a), is used in the third step to compute the ambiguity-*fixed* baseline solution as

$$\breve{b} = \hat{b}(\breve{a}) = \hat{b} - Q_{\hat{b}\hat{a}} Q_{\hat{a}\hat{a}}^{-1}(\hat{a} - \breve{a}) \tag{5.4}$$

That (5.2), (5.3) and (5.4) are indeed the LS and ILS solutions of the GNSS model (5.1) follows readily from the *orthogonal decomposition* (Teunissen, 1998a)

$$||y - Aa - Bb||^2_{Q_{yy}} = ||P^{\perp}_{[A,B]} y||^2_{Q_{yy}} + F(a,b) \tag{5.5}$$

where

$$F(a,b) = ||\hat{a} - a||^2_{Q_{\hat{a}\hat{a}}} + ||\hat{b}(a) - b||^2_{Q_{\hat{b}(a)\hat{b}(a)}} \tag{5.6}$$

with $Q_{\hat{b}(a)\hat{b}(a)} = (B^T Q_{yy} B)^{-1}$ the variance matrix of $\hat{b}(a)$. As $||P^{\perp}_{[A,B]} y||^2_{Q_{yy}}$ is independent of a and b, the minimizers of $||y - Aa - Bb||^2_{Q_{yy}}$ are those of $F(a,b)$. It therefore follows from (5.6), recognizing $\hat{b}(\hat{a}) = \hat{b}$, that the real-valued minimizers of $F(a,b)$ are given by (5.2), while their mixed-integer counterparts are given by (5.3) and (5.4), respectively. We therefore have for the LS and the mixed ILS solutions,

$$\left.\begin{matrix} \hat{a} \\ \hat{b} \end{matrix}\right\} = \underset{a \in \mathbb{R}^n, b \in \mathbb{R}^p}{\arg\min} F(a,b) \;, \quad \left.\begin{matrix} \breve{a} \\ \breve{b} \end{matrix}\right\} = \underset{a \in \mathbb{Z}^n, b \in \mathbb{R}^p}{\arg\min} F(a,b) \tag{5.7}$$

In order to judge the quality of the mixed ILS estimators $\breve{a}$ and $\breve{b}$, we need their probability distributions. They are given in the following theorem.

Theorem 1 (Teunissen, 1999b) The probability mass function (PMF) of $\breve{a}$ and the probability density function (PDF) of $\breve{b}$ are given as

$$\left\{ \begin{array}{lcl} \mathsf{P}[\breve{a} = z] & = & \int_{\mathscr{P}_z} f_{\hat{a}}(\alpha) d\alpha \\ f_{\breve{b}}(\beta) & = & \sum\limits_{z \in \mathbb{Z}^n} f_{\hat{b}(z)}(\beta) \mathsf{P}[\breve{a} = z] \end{array} \right. \tag{5.8}$$

with $\hat{a} \sim \mathcal{N}_n(a, Q_{\hat{a}\hat{a}})$, $\hat{b}(z) \sim \mathcal{N}_p(b - Q_{\hat{b}\hat{a}} Q_{\hat{a}\hat{a}}^{-1}(a - z), Q_{\hat{b}(z)\hat{b}(z)})$, and the pull-in region of $z \in \mathbb{Z}^n$ given as $\mathscr{P}_z = \{x \in \mathbb{R}^n | \; ||x - z||^2_{Q_{\hat{a}\hat{a}}} \leq ||x - u||^2_{Q_{\hat{a}\hat{a}}}, \forall u \in \mathbb{Z}^n\}$. ■

In the practice of GNSS ambiguity resolution one aims to resolve the ambiguities with a high *success-rate*, i.e. a high probability of correct integer estimation $\mathsf{P}[\breve{a} = a]$. When the success-rate is high enough, one may neglect the uncertainty in $\breve{a}$ and describe the uncertainty in $\breve{b}$ by means of the PDF of $\hat{b}(a)$. This is made precise by the following bounds of Teunissen (1999b),

$$\mathsf{P}[\hat{b}(a) \in \Omega] \mathsf{P}[\breve{a} = a] \leq \mathsf{P}[\breve{b} \in \Omega] \leq \mathsf{P}[\hat{b}(a) \in \Omega] \tag{5.9}$$

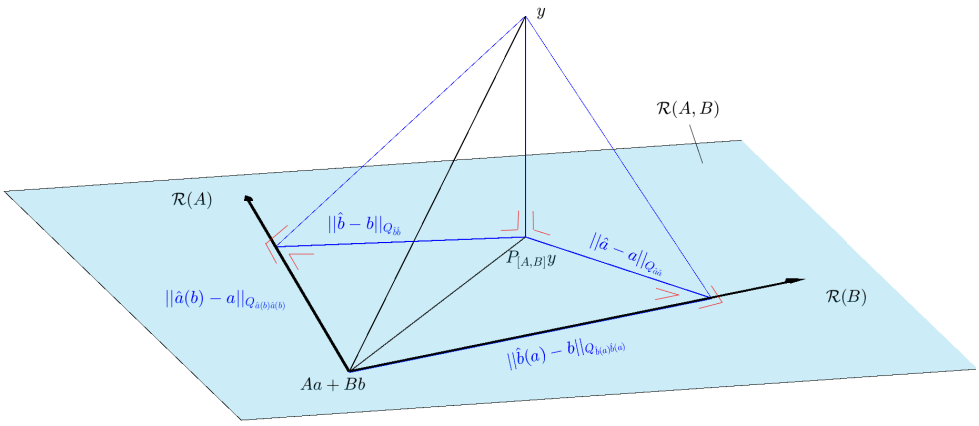


Figure 5.1: Primal and dual orthogonal decompositions, after (Teunissen, 1998a): $||y-Aa-Bb||^2_{Q_{yy}} = ||P^{\perp}_{[A,B]}y||^2_{Q_{yy}} + ||P_{[A,B]}(y-Aa-Bb)||^2_{Q_{yy}}$, with primal decomposition $||P_{[A,B]}(y-Aa-Bb)||^2_{Q_{yy}} = ||P_{\bar{A}}(y-Aa)||^2_{Q_{yy}} + ||P_B(y-Aa-Bb)||^2_{Q_{yy}} = ||\hat{a}-a||^2_{Q_{\hat{a}\hat{a}}} + ||\hat{b}(a)-b||^2_{Q_{\hat{b}(a)\hat{b}(a)}}$ and dual decomposition $||P_{[A,B]}(y-Aa-Bb)||^2_{Q_{yy}} = ||P_{\bar{B}}(y-Bb)||^2_{Q_{yy}} + ||P_A(y-Aa-Bb)||^2_{Q_{yy}} = ||\hat{b}-b||^2_{Q_{\hat{b}\hat{b}}} + ||\hat{a}(b)-a||^2_{Q_{\hat{a}(b)\hat{a}(b)}}$.

which hold true for any convex region $\Omega \subset \mathbb{R}^p$ centred at $\mathsf{E}(\hat{b})$. Thus when the success-rate $\mathsf{P}[\check{a}=a]$ is close enough to one, then

$$\mathsf{P}[\check{b} \in \Omega] \approx \mathsf{P}[\hat{b}(a) \in \Omega] \tag{5.10}$$

which in case of GNSS, due to the very precise carrier-phase data, is usually a much larger probability than that obtained from the float-solution $\hat{b}$, $\mathsf{P}[\hat{b}(a) \in \Omega] \gg \mathsf{P}[\hat{b} \in \Omega]$.

5.3 A dual mixed-ILS formulation

5.3.1 Primal and dual mixed-ILS

A characteristic of the 3-step solution approach is the order in which the mixed ILS solutions $\check{a}$ and $\check{b}$ are computed in the last two steps. First the integer ambiguity estimate $\check{a}$ (cf. 5.3) is computed and then the fixed baseline estimate as $\check{b} = \hat{b}(\check{a})$ (cf. 5.4). There is in principle however no a-priori reason for this particular order. The same solution will be obtained if one would interchange the order of the two minimization steps, since

$$\min_{a\in\mathbb{Z}^n}[\min_{b\in\mathbb{R}^p} F(a,b)] = \min_{b\in\mathbb{R}^p}[\min_{a\in\mathbb{Z}^n} F(a,b)] \tag{5.11}$$

With this equivalence, the solution to the mixed ILS problem can be formulated in two alternative ways, each working with a different objective function, namely a primal function $\mathcal{P}(a)$ that solely depends on the ambiguity vector and a dual function $\mathcal{D}(b)$

that solely depends on the baseline vector b. With the aid of the following short-hand notation

$$\begin{aligned} \hat{b}(a) &= \arg\min_{b\in\mathbb{R}^p} F(a,b) \\ \check{a}(b) &= \arg\min_{a\in\mathbb{Z}^n} F(a,b) \end{aligned} \tag{5.12}$$

we have the following result.

Theorem 2 (Primal and dual mixed-ILS) Let the primal and dual objective functions be defined as

$$\begin{aligned} \text{Primal}: \quad \mathcal{P}(a) &= \min_{b\in\mathbb{R}^p} F(a,b) = F(a,\hat{b}(a)) \\ \text{Dual}: \quad \mathcal{D}(b) &= \min_{a\in\mathbb{Z}^n} F(a,b) = F(\check{a}(b),b) \end{aligned} \tag{5.13}$$

Then the mixed ILS solution is given as

$$\begin{aligned} \check{a} &= \arg\min_{a\in\mathbb{Z}^n} \mathcal{P}(a) = \check{a}(\check{b}) \\ \check{b} &= \arg\min_{b\in\mathbb{R}^p} \mathcal{D}(b) = \hat{b}(\check{a}) \end{aligned} \tag{5.14}$$

■

This result shows that one has two routes available for computing the mixed ILS solution. Either one minimizes $\mathcal{P}(a)$ to obtain $\check{a}$ first and then $\check{b} = \hat{b}(\check{a})$, or one minimizes $\mathcal{D}(b)$ to obtain $\check{b}$ first and then $\check{a} = \check{a}(\check{b})$. Both routes determine the same minimum of $F(a,b)$,

$$\min_{a\in\mathbb{Z}^n, b\in\mathbb{R}^p} F(a,b) = \mathcal{P}(\check{a}) = \mathcal{D}(\check{b}) \tag{5.15}$$

The first route is the one described in the previous section. The second route is the object of study of the present contribution.

5.3.2 The dual objective function

If the parameters of interest are not the ambiguities, but rather the entries of b, it seems that working with the dual function $\mathcal{D}(b)$ is a natural way to go. To determine an explicit expression for $\mathcal{D}(b)$, it is useful to start from the orthogonal decomposition (5.6), but now with the roles of a and b interchanged, i.e.

$$F(a,b) = ||\hat{b} - b||^2_{Q_{\hat{b}\hat{b}}} + ||\hat{a}(b) - a||^2_{Q_{\hat{a}(b)\hat{a}(b)}} \tag{5.16}$$

where $\hat{a}(b) = \hat{a} - Q_{\hat{a}\hat{b}}Q^{-1}_{\hat{b}\hat{b}}(\hat{b} - b)$, $Q_{\hat{b}\hat{b}} = (\bar{B}^T Q^{-1}_{yy}\bar{B})^{-1}$, and $Q_{\hat{a}(b)\hat{a}(b)} = (A^T Q^{-1}_{yy} A)^{-1}$. For the geometry of the primal and dual orthogonal decompositions, see Figure 5.1. We can now obtain the following representations of the dual function $\mathcal{D}(b)$.

Lemma 1 (Dual objective function): Let $\mathscr{S}_z = \{x \in \mathbb{R}^n |\ ||x - z||^2_{Q_{\hat{a}(b)\hat{a}(b)}} \leq ||x - u||^2_{Q_{\hat{a}(b)\hat{a}(b)}}, \forall u \in \mathbb{Z}^n\}$ be the ILS pull-in regions of $Q_{\hat{a}(b)\hat{a}(b)}$, having $s_z(x)$ as its

indicator function, i.e. $s_z(x) = 1$ if $x \in \mathscr{S}_z$ and $s_z(x) = 0$ otherwise. Then

$$\begin{aligned} \mathcal{D}(b) &= \min_{a \in \mathbb{Z}^n} F(a,b) = \\ &= ||\hat{b} - b||^2_{Q_{\hat{b}\hat{b}}} + ||\hat{a}(b) - \check{a}(b)||^2_{Q_{\hat{a}(b)\hat{a}(b)}} \\ &= ||\hat{b} - b||^2_{Q_{\hat{b}\hat{b}}} + \sum_{z \in \mathbb{Z}^n} ||\hat{a}(b) - z||^2_{Q_{\hat{a}(b)\hat{a}(b)}} s_z(\hat{a}(b)) \end{aligned} \tag{5.17}$$

where

$$\check{a}(b) = \arg\min_{a \in \mathbb{Z}^n} ||\hat{a}(b) - a||^2_{Q_{\hat{a}(b)\hat{a}(b)}} \tag{5.18}$$

■

This shows that the dual function $\mathcal{D}(b)$ is a sum of two functions in b,

$$\mathcal{D}(b) = \underbrace{||\hat{b} - b||^2_{Q_{\hat{b}\hat{b}}}}_{\mathcal{D}_1(b)} + \underbrace{||\hat{a}(b) - \check{a}(b)||^2_{Q_{\hat{a}(b)\hat{a}(b)}}}_{\mathcal{D}_2(b)} \tag{5.19}$$

As the second function $\mathcal{D}_2(b)$ is formed from solving again an ILS problem, one may wonder whether anything would be gained by working with the dual $\mathcal{D}(b)$, in particular if we also note that the ILS problem of (5.18) needs to be re-evaluated for every different value of the unknown b. A comparison of the two ILS problems, (5.3) and (5.18), shows however that the second is formulated with respect to the *conditional* variance matrix $Q_{\hat{a}(b)\hat{a}(b)}$ and not with respect to $Q_{\hat{a}\hat{a}}$ as is the case with (5.3). Although both ILS problems can be solved efficiently by means of the LAMBDA method (Teunissen, 1995), we recall that herein the two dominant computational components are (1) the Z-decorrelation, and (2) the ellipsoidal integer search. Hence, if the structure of the conditional vc-matrix $Q_{\hat{a}(b)\hat{a}(b)}$ is such that one or both of these components can be skipped or simplified, then working with the dual $\mathcal{D}(b)$ could perhaps become attractive in some instances. For instance, if $Q_{\hat{a}(b)\hat{a}(b)}$ is diagonal, $\check{a}(b)$ equals the component-wise rounded version of $\hat{a}(b)$, and both components can be avoided. Diagonality of $Q_{\hat{a}(b)\hat{a}(b)}$ happens when the columns of A are mutually orthogonal in the metric of Q_{yy}. In the realm of GNSS, this is the case with the geometry-free GNSS model. Ease of computation would also be present if $\check{a}(b)$ would only be moderately dependent on b. To provide insight into this, we consider the probability mass function of $\check{a}(b)$ and in particular its success-rate $\mathsf{P}[\check{a}(b) = a]$.

5.3.3 Probability Mass Function of $\check{a}(b)$

For every b that we need to evaluate $\mathcal{D}(b)$, we need to compute the integer estimate $\check{a}(b)$. The performance of this integer estimator can be described by its probability mass function (PMF).

Lemma 2 (PMF of $\check{a}(b)$) The probability mass function of $\check{a}(b)$ (cf. 5.18) is given as

$$\begin{aligned} \mathsf{P}[\check{a}(b) = z] &= \mathsf{P}[\hat{a}(b) \in \mathscr{S}_z] \\ &= \textstyle\int_{\mathscr{S}_z} f_{\hat{a}(b)}(x)dx, \forall z \in \mathbb{Z}^n \end{aligned} \tag{5.20}$$

with pull-in regions $\mathscr{S}_z = \{x \in \mathbb{R}^n \mid ||x-z||^2_{Q_{\hat{a}(b)\hat{a}(b)}} \leq ||x-u||^2_{Q_{\hat{a}(b)\hat{a}(b)}}, \forall u \in \mathbb{Z}^n\}$ and the PDF $f_{\hat{a}(b)}(x)$ of $\hat{a}(b)$ given as

$$\hat{a}(b) \sim N_n(a + \Delta a, Q_{\hat{a}(b)\hat{a}(b)}) \tag{5.21}$$

where $\Delta a = -Q_{\hat{a}\hat{b}} Q^{-1}_{\hat{b}\hat{b}} \Delta b$, with $\Delta b = \mathsf{E}(\hat{b}) - b$. ■

The PMF of $\check{a}(b)$ (cf. 5.20) is driven by the PDF of $\hat{a}(b)$ (cf. 5.21). Its ambiguity success-rate can be evaluated with the bounds of Teunissen (2001) or with the simulation algorithms provided in Ps-LAMBDA (Verhagen et al., 2013). The PDF of $\hat{a}(b)$ is usually very peaked, especially in case of GNSS where we have $Q_{\hat{a}(b)\hat{a}(b)} \ll Q_{\hat{a}\hat{a}}$ due to the very precise phase data. Would this peakedness of the PDF be such that it is located over only a single pull-in region, say $\mathscr{S}_u$, $u \in \mathbb{Z}^n$, for a certain b, then the PMF of $\check{a}(b)$ could be well approximated for that value of b by a Kronecker delta function,

$$\mathsf{P}[\check{a}(b) = z] \approx \delta_{uz} = \begin{cases} 1 & \text{if } z = u \\ 0 & \text{if } z \neq u \end{cases} \tag{5.22}$$

The ambiguity success-rate of $\check{a}(b)$ would then be large, i.e. $\mathsf{P}[\check{a}(b) = a] \approx 1$, if $u = a$. For this to happen however, we need $a + \Delta a \in \mathscr{S}_a$, i.e. the bias in $\hat{a}(b)$ needs to be sufficiently small, with Δa residing in $\check{a}(b)$'s pull-in region of the origin, $\Delta a \in \mathscr{S}_0$. For the squared $Q_{\hat{a}(b)\hat{a}(b)}$-weighted norm of this ambiguity bias we have the following result.

Lemma 3 (Bias of $\hat{a}(b)$) Let $\Delta a = -Q_{\hat{a}\hat{b}} Q^{-1}_{\hat{b}\hat{b}} \Delta b$. Then

$$\begin{aligned} ||\Delta a||^2_{Q_{\hat{a}(b)\hat{a}(b)}} &= ||P_A B \Delta b||^2_{Q_{yy}} \\ &\leq ||\Delta b||^2_{Q_{\hat{b}(a)\hat{b}(a)}} \end{aligned} \tag{5.23}$$

■

Proof: see Appendix.

This result shows that for the to be accounted range of b-values, one can generally not expect the bias Δa to be small enough such that $\Delta a \in \mathscr{S}_0$. It would namely require knowledge of b such that $\Delta b = \mathsf{E}(\hat{b}) - b$ is sufficiently small with respect to the phase-driven, small standard deviations of $\hat{b}(a)$. Such can only be expected in a model having strong a-priori constraints on b. As the following example demonstrates, this can not be expected from a regular unconstrained GNSS model.

Example 1 Consider the single-frequency, single epoch, single baseline, double-differenced (DD), $m+1$ satellite GNSS model

$$\mathsf{E}\begin{bmatrix} p \\ \phi \end{bmatrix} = \begin{bmatrix} 0 & D^T G \\ \lambda I_m & D^T G \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}, \mathsf{D}\begin{bmatrix} p \\ \phi \end{bmatrix} = \begin{bmatrix} \sigma_p^2 D^T D & 0 \\ 0 & \sigma_\phi^2 D^T D \end{bmatrix} \tag{5.24}$$

with $p, \phi \in \mathbb{R}^m$ the DD pseudorange and carrier-phase data vectors, λ the wavelength, $D^T = [-e_m, I_m]$ the differencing matrix, $G \in \mathbb{R}^{(m+1)\times 3}$ the receiver-satellite geometry

matrix, and $\sigma_p^2, \sigma_\phi^2$ the variances of the single-differenced pseudoranges and carrier-phases. For this model the variance matrix of $\hat{a}(b)$ and its bias work out to be

$$\begin{array}{lcl} Q_{\hat{a}(b)\hat{a}(b)} & = & (A^T Q_{yy}^{-1} A)^{-1} = \frac{\sigma_\phi^2}{\lambda^2} D^T D \\ \Delta a & = & (A^T Q_{yy}^{-1} A)^{-1} A^T Q_{yy}^{-1} B \Delta b = \frac{1}{\lambda} D^T G \Delta b \end{array} \tag{5.25}$$

Recognizing that the rows of $D^T G$ consist of differences of the rows of G and that each row of G consists of a unit direction vector, the entries of Δa can be bounded from above as $|(\Delta a)_i| \leq \frac{2}{\lambda}||\Delta b||$, $i = 1, \ldots, m$. This shows, as $\lambda \approx 20$cm in case of GNSS, that $\Delta b = \mathsf{E}(\hat{b}) - b$ has to be very small indeed to ensure that $|(\Delta a)_i|$ stays below the subcycle level. □

As the above has demonstrated, without strong a-priori constraints on b, one can not expect the success-rate of $\check{a}(b)$ to be large. This implies that one will have to evaluate $\mathcal{D}_2(b) = ||\hat{a}(b) - \check{a}(b)||^2_{Q_{\hat{a}(b)\hat{a}(b)}}$ for a range of values of b and thus also solve as many ILS problems. It would therefore be beneficial, in case solving the ILS-problem is too time-consuming, if we could replace the evaluation of $\mathcal{D}_2(b)$ with a simpler one, without affecting the performance of the whole estimation process by much. One can ask oneself for instance, whether one can take advantage of the peakedness of the PDF of $\hat{a}(b)$ and replace the ILS estimator $\check{a}(b)$ by the integer-rounding (IR) estimator $\lfloor \hat{a} \rceil$, without a serious degradation in performance. Such would be possible if 'all' probability of the PDF of $\hat{a}(b)$ would be concentrated in the intersections of the ILS and IR pull-in regions, which would require $a + \Delta a \in \mathscr{S}_u \cap \mathscr{R}_u$, with $\mathscr{R}_u$ denoting the integer-rounding pull-in region of $u \in \mathbb{Z}^n$. However, such assumption can not be generally valid, as by changing b in $\Delta a = -Q_{\hat{a}\hat{b}} Q_{\hat{b}\hat{b}}^{-1} \Delta b$, one would also be able to pull $a + \Delta a$ out of such intersection $\mathscr{S}_u \cap \mathscr{R}_u$. It is therefore of importance, if one would replace the evaluation of $\mathcal{D}_2(b)$ by a simpler one, that one at the same time also has the ability to give a *rigorous* evaluation of the performance of such simplification. In the following we introduce two different simplifications of $\mathcal{D}(b)$ and study the probabilistic properties of their minimizers.

5.4 Dual with approximate weight matrix

In this and the next section we study the properties of the baseline- and ambiguity estimators when one works, instead with the dual $\mathcal{D}(b)$, with easier-to-compute approximations to it. The two types of approximation that we consider are,

$$\begin{array}{lcl} (a)\ \ \mathcal{D}^\circ(b) & = & ||\hat{b} - b||^2_{Q_{\hat{b}\hat{b}}} + \min\limits_{a \in \mathbb{Z}^n} ||\hat{a}(b) - a||^2_{Q^\circ_{\hat{a}(b)\hat{a}(b)}} \\ (b)\ \ \mathcal{D}^\bullet(b) & = & ||\hat{b} - b||^2_{Q_{\hat{b}\hat{b}}} + ||\hat{a}(b) - \mathcal{I}^\bullet(\hat{a}(b))||^2_{Q_{\hat{a}(b)\hat{a}(b)}} \end{array} \tag{5.26}$$

In the first type, we have replaced the conditional variance matrix $Q_{\hat{a}(b)\hat{a}(b)}$ by an approximation $Q^\circ_{\hat{a}(b)\hat{a}(b)}$, the idea being that the approximation will then allow for a simpler ambiguity minimization in (5.26a). For instance, when $Q^\circ_{\hat{a}(b)\hat{a}(b)}$ is chosen to be a diagonal matrix, the minimization in (5.26a) reduces to a straightforward

componentwise integer rounding of $\hat{a}(b)$. In the second approximation type, we have replaced the integer ambiguity minimizer $\check{a}(b) = \arg\min_{a\in\mathbb{Z}^n} ||\hat{a}(b) - a||^2_{Q_{\hat{a}\hat{a}}}$ of $\mathcal{D}_2(b)$ (cf. 5.19) by an arbitrary admissible integer estimator $\check{a}^\bullet(b) = \mathcal{I}^\bullet(\hat{a}(b))$, $\mathcal{I}^\bullet : \mathbb{R}^n \mapsto \mathbb{Z}^n$. This second type will be studied in the next section.

To determine the properties of the baseline estimator $\check{b}^\circ = \arg\min_{b\in\mathbb{R}^p} \mathcal{D}^\circ(b)$ and its corresponding integer ambiguity estimator, we again make use of the correspondence between the primal and dual formulations. The quadratic form identity, as provided by the following Lemma, forms the basis for establishing this correspondence.

Lemma 4 Let the conditional ambiguity variance matrix $Q_{\hat{a}(b)\hat{a}(b)}$ in (5.16) be replaced by $Q^\circ_{\hat{a}(b)\hat{a}(b)}$. Then

$$\begin{array}{c} ||\hat{b} - b||^2_{Q_{\hat{b}\hat{b}}} + ||\hat{a}(b) - a||^2_{Q^\circ_{\hat{a}(b)\hat{a}(b)}} = \\ ||\hat{a} - a||^2_{Q^\circ_{\hat{a}\hat{a}}} + ||\hat{b}^\circ(a) - b||^2_{Q^\circ_{\hat{b}(a)\hat{b}(a)}} \end{array} \tag{5.27}$$

with

$$\begin{array}{lcl} Q^\circ_{\hat{a}\hat{a}} & = & Q^\circ_{\hat{a}(b)\hat{a}(b)} + Q_{\hat{a}\hat{b}} Q^{-1}_{\hat{b}\hat{b}} Q_{\hat{b}\hat{a}} \\ \hat{b}^\circ(a) & = & \hat{b} - Q_{\hat{b}\hat{a}} Q^{\circ -1}_{\hat{a}\hat{a}} (\hat{a} - a) \\ Q^\circ_{\hat{b}(a)\hat{b}(a)} & = & Q_{\hat{b}\hat{b}} - Q_{\hat{b}\hat{a}} Q^{o-1}_{\hat{a}\hat{a}} Q_{\hat{a}\hat{b}} \end{array} \tag{5.28}$$

Proof: see Appendix. ■

This result shows that replacing the variance matrix $Q_{\hat{a}(b)\hat{a}(b)}$ in $\mathcal{D}(b)$ (cf. 5.19) by $Q^\circ_{\hat{a}(b)\hat{a}(b)}$ provides an objective function of the type (5.6). It therefore again establishes a primal-dual equivalence, but now one that is driven by the approximate dual function $\mathcal{D}^\circ(b)$. Note that the single replacement $Q_{\hat{a}(b)\hat{a}(b)} \to Q^\circ_{\hat{a}(b)\hat{a}(b)}$ resulted in three changes of the primal formulation: $Q_{\hat{a}\hat{a}} \to Q^\circ_{\hat{a}\hat{a}}$, $\hat{b}(a) \to \hat{b}^\circ(a)$, and $Q_{\hat{b}(a)\hat{b}(a)} \to Q^\circ_{\hat{b}(a)\hat{b}(a)}$. These changes will therefore also drive the properties of the corresponding baseline- and integer ambiguity estimators. Using the quadratic identity (5.27), the following equivalence for the minimizer of $\mathcal{D}^\circ(b)$ can be established.

Theorem 3 Let the approximate dual be given as

$$\mathcal{D}^\circ(b) = ||\hat{b} - b||^2_{Q_{\hat{b}\hat{b}}} + ||\hat{a}(b) - \check{a}^\circ(b)||^2_{Q^\circ_{\hat{a}(b)\hat{a}(b)}} \tag{5.29}$$

with $\check{a}^\circ(b) = \arg\min_{a\in\mathbb{Z}^n} ||\hat{a}(b) - a||^2_{Q^\circ_{\hat{a}(b)\hat{a}(b)}}$. Then the corresponding primal is $\mathcal{P}^\circ(a) = ||\hat{a} - a||^2_{Q^\circ_{\hat{a}\hat{a}}}$ and the minimizer $\check{b}^\circ$ of $\mathcal{D}^\circ(b)$, with corresponding integer ambiguity solution $\check{a}^\circ$, satisfies the primal-dual equivalence,

$$\left.\begin{array}{lcl} \check{a}^\circ & = & \arg\min_{a\in\mathbb{Z}^n} \mathcal{P}^\circ(a) \\ \check{b}^\circ & = & \hat{b}^\circ(\check{a}^\circ) \end{array}\right\} \Leftrightarrow \left\{\begin{array}{lcl} \check{b}^\circ & = & \arg\min_{b\in\mathbb{R}^p} \mathcal{D}^\circ(b) \\ \check{a}^\circ & = & \check{a}^\circ(\check{b}^\circ) \end{array}\right. \tag{5.30}$$

Proof: see Appendix. ■

This equivalence can now be used to apply available theory for the primal formulation to determine the distributional properties of the estimators $\breve{a}^\circ$ and $\breve{b}^\circ$. It should hereby be noted, however, although $\breve{a}^\circ$, like $\breve{a}$, is still computed as the solution of an ILS-problem having $\hat{a}$ as its input, the weight matrix used is now different, $Q^{\circ -1}_{\hat{a}\hat{a}}$ instead of $Q^{-1}_{\hat{a}\hat{a}}$. Also note, although $\hat{b}^\circ(a)$ has the same structure as $\hat{b}(a)$, that $\hat{a}$ is now *not* independent of $\hat{b}^\circ(a)$. The matrix $Q^\circ_{\hat{b}(a)\hat{b}(a)}$ of (5.28) is therefore not the variance matrix of $\hat{b}^\circ(a)$. We have the following distributional result.

Theorem 4 (Distributions of $\breve{a}^\circ$ and $\breve{b}^\circ$) With the PDF of $\hat{a} \sim \mathcal{N}_n(a, Q_{\hat{a}\hat{a}})$ denoted as $f_{\hat{a}}(\alpha)$, the PMF of the ambiguity estimator $\breve{a}^\circ$ is given as

$$\begin{array}{rcl} \breve{a}^\circ & \sim & \mathsf{P}[\breve{a}^\circ = z] = \int_{\mathscr{P}^\circ_z} f_{\hat{a}}(\alpha) d\alpha, \text{ with} \\ & & \mathsf{P}[\breve{a}^\circ = a] \leq \mathsf{P}[\breve{a} = a] \end{array} \tag{5.31}$$

and the limiting PDF of the baseline estimator $\breve{b}^\circ$ is given as

$$\begin{array}{rcl} \lim\limits_{\mathsf{P}[\breve{a}^\circ = a] \uparrow 1} \breve{b}^\circ & \sim & \mathcal{N}_p(b, Q_{\hat{b}^\circ(a)\hat{b}^\circ(a)}), \text{ with} \\ Q_{\hat{b}^\circ(a)\hat{b}^\circ(a)} & = & Q_{\hat{b}(a)\hat{b}(a)} + T_{\hat{b}\hat{a}} Q_{\hat{a}\hat{a}} T^T_{\hat{b}\hat{a}} \geq Q_{\hat{b}(a)\hat{b}(a)} \end{array} \tag{5.32}$$

where $\mathscr{P}^\circ_z = \{x \in \mathbb{R}^n | \; ||x-z||^2_{Q^\circ_{\hat{a}\hat{a}}} \leq ||x-u||^2_{Q^\circ_{\hat{a}\hat{a}}}, \forall u \in \mathbb{Z}^n\}$ and $T_{\hat{b}\hat{a}} = Q_{\hat{b}\hat{a}}[Q^{-1}_{\hat{a}\hat{a}} - Q^{\circ -1}_{\hat{a}\hat{a}}]$. ■

Proof: See Appendix. □

The above result shows that replacing $Q_{\hat{a}(b)\hat{a}(b)}$ by $Q^\circ_{\hat{a}(b)\hat{a}(b)}$ will always degrade the performance of the associated estimators. It will give a smaller ambiguity success-rate (cf. 5.31), as well as a poorer precision of the ambiguity-fixed baseline (cf. 5.32). Still, depending on the choice made for $Q^\circ_{\hat{a}(b)\hat{a}(b)}$, the degradation could be acceptably small, depending on the application.

To evaluate the success-rate $\mathsf{P}[\breve{a}^\circ = a]$, the multivariate integral of (5.31) needs to be computed. This is a nontrivial numerical task due to the geometric complexity of the pull-in region $\mathscr{P}^\circ_a$, over which the integration needs to be carried out. One approach is to rely on simulation, whereby the Ps-LAMBDA simulation tools of (Verhagen et al., 2013) can be used. Note hereby, that the success-rate $\mathsf{P}[\breve{a}^\circ = a]$ is driven by both $Q_{\hat{a}\hat{a}}$ and $Q^\circ_{\hat{a}\hat{a}}$, i.e. by the ambiguity variance matrix that determines $f_{\hat{a}}(\alpha)$ and by its approximation that determines the pull-in region $\mathscr{P}^\circ_a$.

The following example compares the two success-rates $\mathsf{P}[\breve{a} = a]$ and $\mathsf{P}[\breve{a}^\circ = a]$.

Example 2 Consider the GNSS model of Example 1 (cf. 5.24) and assume that in solving the dual problem we approximate the fully populated variance matrix $Q_{\hat{a}(b)\hat{a}(b)} = \frac{\sigma^2_\phi}{\lambda} D^T D$ with the diagonal matrix $Q^\circ_{\hat{a}(b)\hat{a}(b)} = 2\frac{\sigma^2_\phi}{\lambda} I_m$. Then, with $Q_{\hat{a}\dot{b}} Q^{-1}_{\dot{b}\dot{b}} Q_{\dot{b}\hat{a}} = \frac{\sigma^2_p}{\lambda^2} D^T G [G^T P_D G]^{-1} G^T D$ and $P_D = D(D^T D)^{-1} D^T$, we get

$$Q^\circ_{\hat{a}\hat{a}} = \tfrac{1}{\lambda^2} \left(2\sigma^2_\phi I_m + \sigma^2_p D^T G [G^T P_D G]^{-1} G^T D \right) \tag{5.33}$$

5

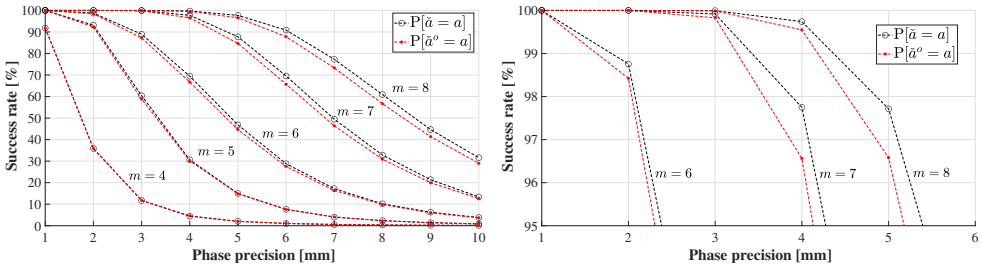


Figure 5.2: Comparison of $\mathsf{P}[\breve{a}^\circ = a]$ with $\mathsf{P}[\breve{a} = a]$ as described in Example 2, based on the GNSS model of Example 1.

With reference to Theorem 4 (cf. 5.31), Figure 5.2 compares the two success-rates $\mathsf{P}[\breve{a} = a]$ and $\mathsf{P}[\breve{a}^\circ = a]$, based on (5.33), for a case of single-epoch, single-frequency L1 GPS, using a 10^{-4} phase-code variance ratio. It shows that $\mathsf{P}[\breve{a}^\circ = a] \leq \mathsf{P}[\breve{a} = a]$, but also that $\mathsf{P}[\breve{a}^\circ = a]$ can still be acceptably large for some measurement scenarios. □

Instead of simulation, success-rate bounds may sometimes be used as an alternative. Upper-bounds are then useful to identify when successful ambiguity resolution would be problematic, while lower-bounds are useful to identify when to expect successful ambiguity resolution. As upper-bound of $\mathsf{P}[\breve{a}^\circ = a]$, one may directly use the ILS success-rate $\mathsf{P}[\breve{a} = a]$ (cf. 5.31), or alternatively, any of its simpler to compute upper-bounds given in (Teunissen, 2000a; Verhagen et al., 2013). The following Theorem provides two lower-bounds on the ambiguity success-rate of $\breve{a}^\circ$.

Theorem 5 (Success-rate lower bounds) Let $\hat{a} \sim \mathcal{N}_n(a, Q_{\hat{a}\hat{a}})$, $\breve{a}^\circ = \arg\min_{a \in \mathbb{Z}^n} ||\hat{a} - a||^2_{Q^\circ_{\hat{a}\hat{a}}}$. Then the success-rate of $\breve{a}^\circ$ can be lower-bounded as follows:

(i) If $Q^\circ_{\hat{a}\hat{a}} \geq Q_{\hat{a}\hat{a}}$, then

$$\prod_{i=1}^{n} [2\Phi\left(\tfrac{1}{2\sqrt{d_i}}\right) - 1] \leq \mathsf{P}[\breve{a}^\circ = a] \tag{5.34}$$

where $D = \mathrm{diag}(d_1, \ldots, d_n)$ is the diagonal matrix of the triangular decomposition $Q^\circ_{\hat{a}\hat{a}} = LDL^T$ and $\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}v^2)dv$.

(ii) For any $Q^\circ_{\hat{a}\hat{a}} > 0$,

$$\mathsf{P}[\chi^2(0, n) \leq r^2] \leq \mathsf{P}[\breve{a}^\circ = a] \tag{5.35}$$

with

$$\begin{cases} r^2 &= \lambda_{\min} \times \frac{1}{4} \min\limits_{z \in \mathbb{Z}^n / \{0\}} ||z||^2_{Q^\circ_{\hat{a}\hat{a}}} \\ \lambda_{\min} &= \min\limits_{x \in \mathbb{R}^n} \frac{x^T Q^\circ_{\hat{a}\hat{a}} x}{x^T Q_{\hat{a}\hat{a}} x} \end{cases} \tag{5.36}$$

Proof: see Appendix □

Note that lower-bound (5.34) is somewhat easier to compute than (5.35). It requires however that $Q^{\circ}_{\hat{a}\hat{a}} \geq Q_{\hat{a}\hat{a}}$, while no such restriction is placed on the lower-bound (5.35). Also note, although both lower-bounds are here presented in the context of the primal-dual formulations, that they are in fact success-rate lower-bounds of improperly weighted ILS-estimators, i.e. ILS-estimators that not use the inverse variance-matrix as their weight-matrix. They can therefore also be used more generally for studying the impact misspecifications in the stochastic model have on the success-rate.

5.5 Dual with approximate integer map

In this section we consider the second approximation of the dual function, $\mathcal{D}^{\bullet}(b)$ (cf. 5.26), and determine the statistical properties of its minimizer.

5.5.1 The minimizer of $\mathcal{D}^{\bullet}(b)$

The approximation $\mathcal{D}^{\bullet}(b)$ of $\mathcal{D}(b)$ is a result of replacing the integer vector $\check{a}(b) = \arg\min_{z\in\mathbb{Z}^n} ||\hat{a}(b) - z||^2_{Q_{\hat{a}(b)\hat{a}(b)}}$ in $\mathcal{D}(b)$ by the integer vector $\mathcal{I}^{\bullet}(\hat{a}(b))$, for which $\mathcal{I}^{\bullet} : \mathbb{R}^n \mapsto \mathbb{Z}^n$ may be chosen as any member from the class of admissible integer estimators, such as, for instance, integer rounding (IR), integer bootstrapping (IB), integer least-squares (ILS) or vectorial integer bootstrapping (VIB) (Teunissen et al., 2021). The following theorem provides the solution of minimizing $\mathcal{D}^{\bullet}(b)$.

Theorem 6 (Minimizer of $\mathcal{D}^{\bullet}(b)$) Let $\mathcal{I}^{\bullet} : \mathbb{R}^n \mapsto \mathbb{Z}^n$ be any admissible integer estimator. Then the minimizer $\check{b}^{\bullet} = \arg\min_{b\in\mathbb{R}^p} \mathcal{D}^{\bullet}(b)$ of the approximate dual

$$\mathcal{D}^{\bullet}(b) = ||\hat{b} - b||^2_{Q_{\hat{b}\hat{b}}} + ||\hat{a}(b) - \mathcal{I}^{\bullet}(\hat{a}(b))||^2_{Q_{\hat{a}(b)\hat{a}(b)}} \tag{5.37}$$

is given as

$$\check{b}^{\bullet} = \hat{b}(\check{a}^{\bullet}), \text{ with } \check{a}^{\bullet} = \arg\min_{z\in\Omega^{\bullet}_{\hat{a}}} ||\hat{a} - z||^2_{Q_{\hat{a}\hat{a}}} \tag{5.38}$$

where

$$\Omega^{\bullet}_{\hat{a}} = \{z \in \mathbb{Z}^n |\ z = \mathcal{I}^{\bullet}(\hat{a} + M\beta), \forall \beta \in \mathbb{R}^p\} \tag{5.39}$$

with $M = Q_{\hat{a}\hat{b}}Q^{-1}_{\hat{b}\hat{b}} \in \mathbb{R}^{n\times p}$.

Proof: see Appendix. □

Note, in contrast to the baseline estimator $\check{b}^{\circ}$ (cf. 5.30), that the baseline estimator $\check{b}^{\bullet}$ (cf. 5.38) is based, like the estimator $\check{b}$ (cf. 5.4), on the conditional LS baseline mapping $\hat{b}(a)$. However, $\check{b}$ and $\check{b}^{\bullet}$ make use of different integer ambiguity estimators in general. Although the integer estimator $\check{a}^{\bullet}$ (cf. 5.38) has the appearance of a standard ILS-estimator, it is generally not, unless of course $\mathcal{I}^{\bullet}(x)$ is chosen as $\mathcal{I}^{\bullet}(x) = \arg\min_{z\in\mathbb{Z}^n} ||x - z||^2_{Q_{\hat{a}(b)\hat{a}(b)}}$, in which case $\check{a}^{\bullet} = \check{a}$, since then $\mathcal{D}^{\bullet}(b) = \mathcal{D}(b)$.

5

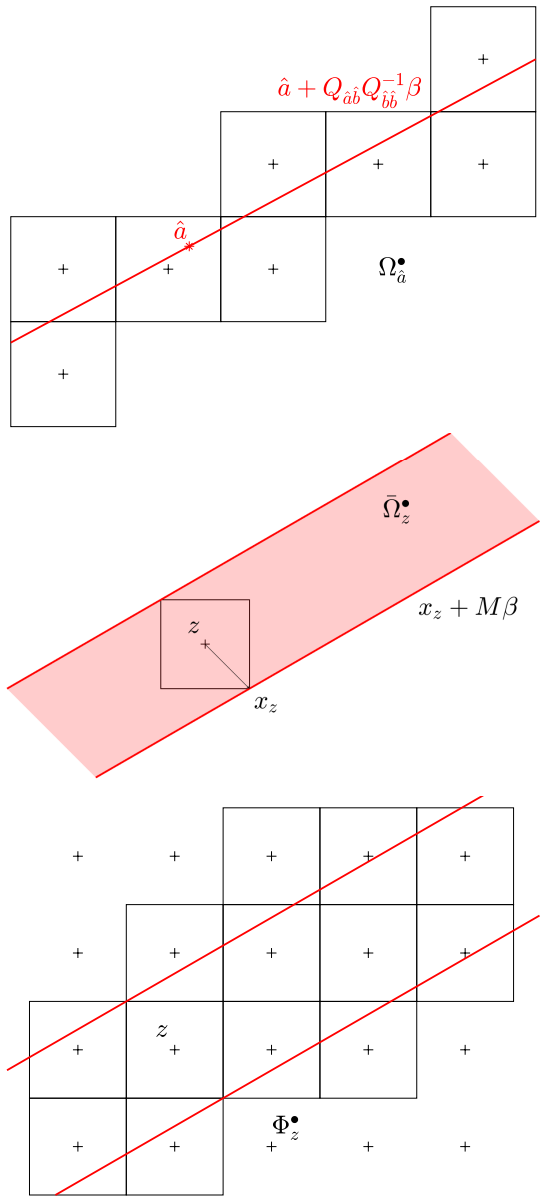


Figure 5.3: (Top) The integer set $\Omega^{\bullet}_{\hat{a}} = \{z \in \mathbb{Z}^n | \; z = \mathcal{I}^{\bullet}(\hat{a} + M\beta), \forall \beta \in \mathbb{R}^p\}$ for $n = 2, p = 1$, and $\mathcal{I}(.) = \lceil . \rfloor$; (Center) The real-valued set $\bar{\Omega}^{\bullet}_{z} = \{x \in \mathbb{R}^n | \; z = \mathcal{I}^{\bullet}(x + M\beta), \exists \beta \in \mathbb{R}^p\}$ for $n = 2, p = 1$; (Bottom) The integer set $\Phi^{\bullet}_{z} = \{u \in \mathbb{Z}^n | \; u = \mathcal{I}^{\bullet}(x + M\beta), x \in \bar{\Omega}^{\bullet}_{z}, \forall \beta \in \mathbb{R}^p\}$ for $n = 2, p = 1$.

The difference between the two integer estimators $\check{a}$ and $\check{a}^{\bullet}$ is driven by the characteristics of the integer set $\Omega^{\bullet}_{\hat{a}}$ (cf. 5.39), which on its turn is driven by the $n \times p$ matrix

M in

$$\hat{a}(\beta) = \hat{a} + M(\beta - \hat{b}) \tag{5.40}$$

The following three cases can be discriminated:

(1) $\hat{a}(\beta) \equiv \hat{a}$ if $M = 0$, i.e. $Q_{\hat{a}\hat{b}} = 0$ or $A^T Q_{yy}^{-1} B = 0$

(2) $\hat{a}(\beta)$ describes a *linear manifold* if $\text{rank}(M) < n$

(3) $\hat{a}(\beta)$ covers the whole of $\mathbb{R}^n$ if $\text{rank}(M) = n$

Case (1) happens if the estimators $\hat{a}$ and $\hat{b}$ are uncorrelated. As the integer set (5.39) reduces then to the single integer vector $\Omega_{\hat{a}}^{\bullet} = \{z = \mathcal{I}^{\bullet}(\hat{a})\}$, we have in that case $\breve{a}^{\bullet} = \mathcal{I}^{\bullet}(\hat{a})$. In case (2), the subset $\Omega_{\hat{a}}^{\bullet}$ contains all integer vectors to which $\mathcal{I}^{\bullet}(x)$ is mapped when x varies along the $\text{rank}(M)$-dimensional linear manifold (5.40). This is shown in Figure 5.3(Top) when $\mathcal{I}^{\bullet}(x)$ represents integer rounding and $n = 2, p = 1$, in which case the pull-in regions of $\mathcal{I}^{\bullet}(x)$ are unit-squares centred at integer grid points. As $\Omega_{\hat{a}}^{\bullet} \neq \mathbb{Z}^n$, we have $\breve{a}^{\bullet} \neq \breve{a}$ in case (2). In case (3), the invertibility of matrix M implies that the whole $\mathbb{R}^n$ is integer-mapped by $\mathcal{I}^{\bullet}(x)$, thus giving $\Omega_{\hat{a}}^{\bullet} = \mathbb{Z}^n$. This shows that $\breve{a}^{\bullet} = \breve{a}$ in case (3).

Note, as $\text{rank}(M) \leq \min(n, p)$, that case (3) can only happen if $p \geq n$. Hence, since $p < n$ in most GNSS models, the equality of the two estimators, $\breve{a}^{\bullet}$ and $\breve{a}$, is very unlikely in case of GNSS. Although the two estimators $\breve{a}^{\bullet}$ and $\breve{a}$ are then generally different, their integer sample outcomes can, of course, sometimes be the same. This happens when the outcome of $\breve{a}$ lies in $\Omega_{\hat{a}}^{\bullet}$.

5.5.2 A qualitative comparison of $\breve{a}^{\bullet}$ and $\breve{a}$

We now compare the two integer estimators for case (2), i.e. when $\hat{a}(\beta)$ describes a linear manifold of dimension $\text{rank}(M) < n$ and $\breve{a}^{\bullet} \neq \breve{a}$.

To aid the comparison between $\breve{a}^{\bullet}$ and $\breve{a}$, we first introduce the ambiguity search space

$$E_{\hat{a}} = \{z \in \mathbb{Z}^n | \; ||\hat{a} - z||^2_{Q_{\hat{a}\hat{a}}} \leq \chi^2\} \tag{5.41}$$

where χ^2 is assumed chosen such that $\breve{a} \in E_{\hat{a}}$ (note: for any integer $z_0 \in \mathbb{Z}^n$, e.g. $z_0 = \lceil \hat{a} \rfloor$, the value $\chi^2 = ||\hat{a} - z_0||^2_{Q_{\hat{a}\hat{a}}}$ satisfies this assumption). With the help of $E_{\hat{a}}$ we may write $\breve{a} = \arg \min_{u \in \mathbb{Z}^n} ||\hat{a} - u||^2_{Q_{\hat{a}\hat{a}}}$ in a similar form as that of $\breve{a}^{\bullet}$. We therefore have

$$\begin{aligned} \breve{a} &= \arg \min_{u \in E_{\hat{a}}} ||\hat{a} - u||^2_{Q_{\hat{a}\hat{a}}} \\ \breve{a}^{\bullet} &= \arg \min_{u \in \Omega_{\hat{a}}^{\bullet}} ||\hat{a} - u||^2_{Q_{\hat{a}\hat{a}}} \end{aligned} \tag{5.42}$$

which shows that the two estimators can be compared by comparing their respective search spaces, $E_{\hat{a}}$ vs $\Omega_{\hat{a}}^{\bullet}$. For $\Omega_{\hat{a}}^{\bullet}$ we have

$$\Omega_{\hat{a}}^{\bullet} = \{z \in \mathbb{Z}^n | \; z = \mathcal{I}^{\bullet}(\hat{a} + M\beta), \forall \beta \in \mathbb{R}^p\} \tag{5.43}$$

For $E_{\hat{a}}$ we may write, with the help of $M = Q_{\hat{a}\hat{b}}Q_{\hat{b}\hat{b}}^{-1}$ and $Q_{\hat{a}\hat{a}} = Q_{\hat{a}(b)\hat{a}(b)} + MQ_{\hat{b}\hat{b}}M^T$,

$$E_{\hat{a}} = \{z \in \mathbb{Z}^n |\ (\hat{a} - z)^T [Q_{\hat{a}(b)\hat{a}(b)} + MQ_{\hat{b}\hat{b}}M^T]^{-1}(\hat{a} - z) \leq \chi^2\} \tag{5.44}$$

We can now compare the two integer sets (5.43) and (5.44). We will do so, for two extreme cases.

$$\text{Case } (a) : Q_{\hat{a}(b)\hat{a}(b)} = \text{small and } Q_{\hat{b}\hat{b}} = \text{large} \tag{5.45}$$

This is the typical 'GNSS-case', in particular for instantaneous positioning. As the very precise carrier phase data do not contribute to the determination of b in case of a single epoch, the precision of $\hat{b}$ is solely driven by the noisy pseudorange data and $Q_{\hat{b}\hat{b}} = \text{large}$. Would b be known, then it are the very precise carrier phase data that predominantly determine the ambiguities and $Q_{\hat{a}(b)\hat{a}(b)} = \text{small}$.

When (5.45) is true, the ellipsoidal search space $E_{\hat{a}}$ (cf. 5.44) will have an extreme elongation in the directions of the range space of M and therefore closely resemble the integer set $\Omega_{\hat{a}}^{\bullet}$, which, afterall, is constructed from integer mapping the points of the linear manifold $\hat{a} + M\beta$. Under case (a) one can therefore expect the two estimators to be not too different, i.e. sample values of the ILS-estimator $\check{a} \in E_{\hat{a}}$ will not rarely be inside $\Omega_{\hat{a}}^{\bullet}$ as well. We hereby note that the GNSS-typical extreme elongation of $E_{\hat{a}}$ results in integer search-halting when solving for $\check{a}$. Resolving this bottleneck was the motivation for developing LAMBDA. By means of its decorrelating Z-transformation, the discontinuity in the spectrum of sequential conditional ambiguity variances is largely removed and search-halting avoided, see (Teunissen, 1995).

$$\text{Case } (b) : Q_{\hat{a}(b)\hat{a}(b)} = \text{large and } Q_{\hat{b}\hat{b}} = \text{small} \tag{5.46}$$

Now we have a different situation, which in GNSS-terminology could be described as having ultra-precise pseudorange data and very poor carrier-phase data. In such case the shape of the ellipsoidal search space $E_{\hat{a}}$ is primarily driven by $Q_{\hat{a}(b)\hat{a}(b)}$, implying that its shape will now generally *not* be aligned with $\Omega_{\hat{a}}^{\bullet}$. And this will even be more so if the ellipsoidal search space would have its elongation orthogonal to the range space of M. In this case one would expect the two estimators, $\check{a}^{\bullet}$ and $\check{a}$, to have different performances, i.e. sample values of the ILS-estimator $\check{a} \in E_{\hat{a}}$ are then not likely to reside inside $\Omega_{\hat{a}}^{\bullet}$ as well.

5.5.3 The pull-in regions of $\check{a}^{\bullet}$

In order to study the statistical properties of $\check{a}^{\bullet}$ and $\check{b}^{\bullet}$, it is useful to first determine the pull-in regions of $\check{a}^{\bullet}$. As the pull-in region $\mathscr{P}_z^{\bullet}$ of $\check{a}^{\bullet}$ is the region in which the float solution $\hat{a}$ gets mapped to $z \in \mathbb{Z}^n$, we have

$$\mathscr{P}_z^{\bullet} = \{x \in \mathbb{R}^n |\ z = \arg \min_{u \in \Omega_x^{\bullet}} ||x - u||^2_{Q_{\hat{a}\hat{a}}}\} \tag{5.47}$$

To further characterize this region, we recognize that the choice of $z \in \mathbb{Z}^n$, i.e. the integer-vector for which the pull-in region is described, already constrains the values

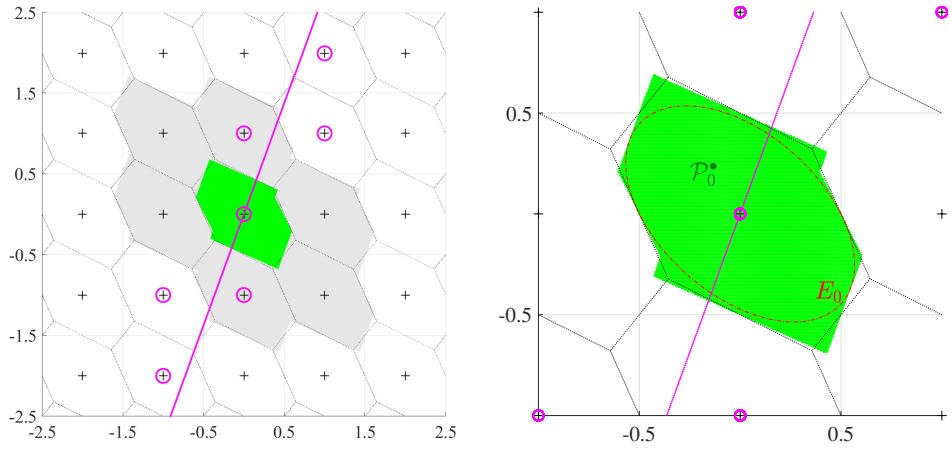


Figure 5.4: Pull-in regions $\mathscr{P}_z^\bullet$ of the integer ambiguity estimator $\check{a}^\bullet$ for $n = 2$ and $p = 1$. Top: Comparison of $\mathscr{P}_z^\bullet$ with ILS pull-in regions, showing integer set $\Omega_0^\bullet$ through encircled gridpoints; Bottom: A zoom-in showing $\mathscr{P}_0^\bullet$ of $\check{a}^\bullet$ together with the (lower) bounding ellipse E_0 (cf. proof of Theorem 7).

of $x \in \mathbb{R}^n$ to a subset. As z has to lie in $\Omega_x^\bullet$, the choice of z implies the following subset for the values of x,

$$\bar{\Omega}_z^\bullet = \{x \in \mathbb{R}^n | \; z = \mathcal{I}^\bullet(x + M\beta), \exists \beta \in \mathbb{R}^p\} \tag{5.48}$$

This is the set of x-values for which a $\beta \in \mathbb{R}^p$ exists, such that $x + M\beta$ gets mapped by $\mathcal{I}^\bullet(.)$ to z, see Figure 5.3(Middle).

With the help of the region $\bar{\Omega}_z^\bullet$ we can now characterize the whole integer set that is in play in the minimization of (5.47). As $\mathscr{P}_z^\bullet$ is characterized by the minimization of $||x - u||^2_{Q_{\hat{a}\hat{a}}}$ over the integer subset $\Omega_x^\bullet$, while, at the same time, this is constrained to all $x \in \bar{\Omega}_z^\bullet$, the integer set considered is actually

$$\Phi_z^\bullet = \{u \in \mathbb{Z}^n | \; u = \mathcal{I}^\bullet(x), \forall x \in \bar{\Omega}_z^\bullet\} \tag{5.49}$$

It consists of all integer vectors to which the elements of $\bar{\Omega}_z^\bullet$ are mapped by $\mathcal{I}^\bullet(.)$, see Figure 5.3(Bottom).

With the above three constructed sets, $\Omega_x^\bullet \subset \mathbb{Z}^n$, $\bar{\Omega}_z^\bullet \subset \mathbb{R}^n$, and $\Phi_z^\bullet \subset \mathbb{Z}^n$, we have the following three representations of the pull-in regions of $\check{a}^\bullet$.

Lemma 5 (Representations of pull-in region $\mathscr{P}_z^\bullet$) The pull-in regions $\mathscr{P}_z^\bullet$, $z \in \mathbb{Z}^n$, of $\check{a}^\bullet$ can be represented, with

$$\begin{aligned} \Omega_x^\bullet &= \{z \in \mathbb{Z}^n | \; z = \mathcal{I}^\bullet(x + M\beta), \forall \beta \in \mathbb{R}^p\} \\ \bar{\Omega}_z^\bullet &= \{x \in \mathbb{R}^n | \; z = \mathcal{I}^\bullet(x + M\beta), \exists \beta \in \mathbb{R}^p\} \\ \Phi_z^\bullet &= \{u \in \mathbb{Z}^n | \; u = \mathcal{I}^\bullet(x), \forall x \in \bar{\Omega}_z^\bullet\} \end{aligned}$$

as

$$\begin{aligned}\mathscr{P}_z^\bullet &= \{x \in \mathbb{R}^n |\ z = \arg\min_{u \in \Omega_x^\bullet} ||x-u||^2_{Q_{\hat{a}\hat{a}}}\} \\ &= \{x \in \mathbb{R}^n |\ ||x-z||^2_{Q_{\hat{a}\hat{a}}} \leq ||x-\mathcal{I}^\bullet(y)||^2_{Q_{\hat{a}\hat{a}}}, \forall y \in \bar{\Omega}_z^\bullet\} \\ &= \{x \in \mathbb{R}^n |\ ||x-z||^2_{Q_{\hat{a}\hat{a}}} \leq ||x-u||^2_{Q_{\hat{a}\hat{a}}}, \forall u \in \Phi_z^\bullet\}\end{aligned} \tag{5.50}$$

■

We can now use these pull-in representations to show that $\check{a}^\bullet$ is an *admissible* integer estimator. Recall that an integer estimator is said to be admissible if its pull-in regions are translational invariant and cover the whole space $\mathbb{R}^n$ without gaps and overlaps (Teunissen, 2002).

Lemma 6 (Admissible integer estimator $\check{a}^\bullet$) The integer estimator $\check{a}^\bullet$ is admissible as its pull-in regions satisfy

$$\begin{array}{ll}(i) & \bigcup\limits_{z \in \mathbb{Z}^n} \mathscr{P}_z^\bullet = \mathbb{R}^n \\ (ii) & \mathscr{P}_{z_1}^\bullet \cap \mathscr{P}_{z_2}^\bullet = \emptyset,\ \forall z_1 \neq z_2 \\ (iii) & \mathscr{P}_z^\bullet = \mathscr{P}_0^\bullet + z,\ \forall z \in \mathbb{Z}^n\end{array} \tag{5.51}$$

Proof: see Appendix. ■

The admissibility property implies that if y is perturbed by Az to give $y' = y + Az$, the ambiguity float solution changes from $\hat{a} = \bar{A}^+ y$ to $\hat{a}' = \bar{A}^+(y + Az) = \hat{a} + z$, and the integer ambiguity solution from $\check{a}^\bullet$ to $\check{a}'^\bullet = \check{a}^\bullet + z$. Hence, this provides the pleasant property, that if one wants to work with managable numbers, one can subtract arbitrary integers from the ambiguity float solution and still get the correct integer solution by restoring the subtracted integer at the end, i.e. if $\hat{a}' = \hat{a} - z$ then $\check{a}^\bullet = \check{a}'^\bullet + z$.

Figure 5.4 shows, for $n = 2$ and $p = 1$, an example of the pull-in regions $\mathscr{P}_z^\bullet$ of the integer ambiguity estimator $\check{a}^\bullet$. The choice made for the integer map $\mathcal{I}^\bullet : \mathbb{R}^n \mapsto \mathbb{Z}^n$ is in this case integer-rounding, i.e. $\mathcal{I}^\bullet(x) = \lceil x \rfloor$. The encircled integer gridpoints constitute the integer set $\Omega_{x=0}^\bullet$ (cf. Lemma 5) and the line through the origin has $M = Q_{\hat{a}\hat{b}} Q_{\hat{b}\hat{b}}^{-1}$ as its direction vector. For comparison also the hexagonian ILS pull-in regions of $\check{a}$ are shown, thus illustrating the close overlap between the two types of pull-in regions.

5.5.4 Distributions of $\check{a}^\bullet$ and $\check{b}^\bullet$

With the knowledge that the integer estimator $\check{a}^\bullet$ is admissible, we can now apply existing theory of Teunissen (1999b) to determine the distributions of $\check{a}^\bullet$ and $\check{b}^\bullet$.

Corollary (Distributions of $\check{a}^\bullet$ and $\check{b}^\bullet$): Let $f_{\hat{a}}(\alpha)$ be the PDF of $\hat{a} \sim \mathcal{N}_n(a, Q_{\hat{a}\hat{a}})$ and $f_{\hat{b}(z)}(\beta)$ be the PDF of $\hat{b}(z) \sim \mathcal{N}_p(b - Q_{\hat{b}\hat{a}} Q_{\hat{a}\hat{a}}^{-1}(a-z), Q_{\hat{b}(a)\hat{b}(a)})$. Then, as $\check{a}^\bullet$ is an admissible integer estimator and $\check{b}^\bullet = \hat{b}(\check{a}^\bullet)$, their PMF and PDF follow from

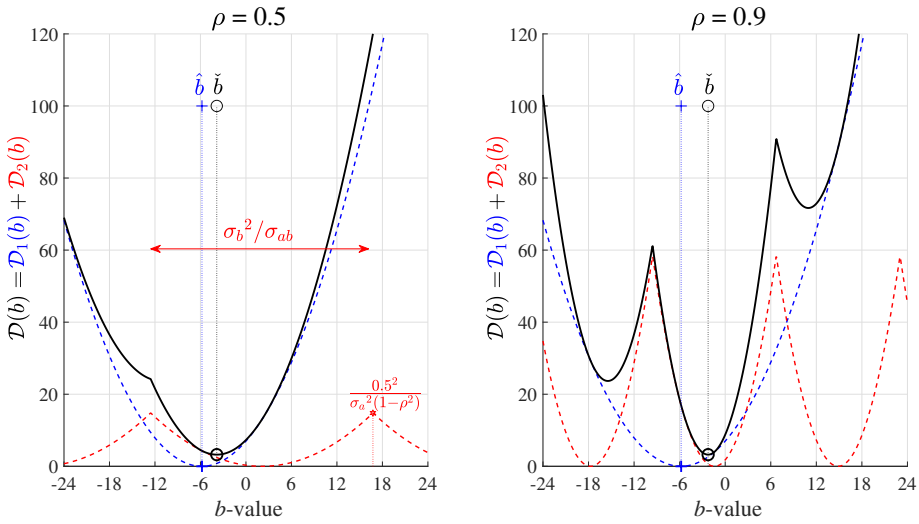


Figure 5.5: The dual function $\mathcal{D}(b)$ for $n = p = 1$, $\sigma_{\hat{b}}^2 = 2.2^2$, $\sigma_{\hat{a}}^2 = 0.15^2$ and $\rho_{\hat{a}\hat{b}} = 0.5$ (left), $\rho_{\hat{a}\hat{b}} = 0.9$ (right)

(Teunissen, 1999b) as

$$\left\{ \begin{array}{lcl} \mathsf{P}[\check{a}^\bullet = z] & = & \int_{\mathscr{P}_z^\bullet} f_{\hat{a}}(\alpha) d\alpha \\ f_{\check{b}^\bullet}(\beta) & = & \sum_{z \in \mathbb{Z}^n} f_{\hat{b}(z)}(\beta) \mathsf{P}[\check{a}^\bullet = z] \end{array} \right. \tag{5.52}$$

■

This result shows that the distribution of the ambiguity resolved baseline $\check{b}^\bullet$ can be approximated well by the peaked PDF $\mathcal{N}_p(b, Q_{\hat{b}(a)\hat{b}(a)})$ if the ambiguity success-rate $\mathsf{P}[\check{a}^\bullet = a]$ is sufficiently close to one. Verification whether or not the success-rate is large enough can be done by simulation or by using the following lower-bound.

Theorem 7 (Lower-bound of $\mathsf{P}[\check{a}^\bullet = a]$) Let $\hat{a} \sim \mathcal{N}_n(a, Q_{\hat{a}\hat{a}})$ and $\check{a}^\bullet = \arg \min_{z \in \Omega_{\hat{a}}^\bullet} ||\hat{a} - z||^2_{Q_{\hat{a}\hat{a}}}$, with $\Omega_{\hat{a}}^\bullet = \{u \in \mathbb{Z}^n |\ u = \mathcal{I}^\bullet(\hat{a} + Q_{\hat{a}\hat{b}} Q_{\hat{b}\hat{b}}^{-1} \beta), \forall \beta\}$. Then

$$\mathsf{P}[\chi^2(0, n) \leq r^2] \leq \mathsf{P}[\check{a}^\bullet = a] \tag{5.53}$$

with $r^2 = \frac{1}{4} \min_{z \in \mathbb{Z}^n / \{0\}} ||z||^2_{Q_{\hat{a}\hat{a}}}$.

Proof: see Appendix. ■

5.6 On the multimodality of $\mathcal{D}(b)$

So far we studied the distributional properties of the estimators that follow from the dual formulation. A convergent algorithm for actually computing these estimators has however not been developed yet. To be able to do so, it is useful to first illustrate some insightful characteristics of the dual function. To highlight some of its characteristics, we start with the simple one-dimensional case $n = p = 1$. The two components of $\mathcal{D}(b)$ (cf. 5.19) simplify then to

$$\begin{array}{rcl} \mathcal{D}_1(b) & = & ||\hat{b} - b||^2_{Q_{\hat{b}\hat{b}}} = (\hat{b} - b)^2/\sigma^2_{\hat{b}} \\ \mathcal{D}_2(b) & = & ||\hat{a}(b) - \check{a}(b)||^2_{Q_{\hat{a}(b)\hat{a}(b)}} = (\hat{a}(b) - \lceil \hat{a}(b) \rfloor)^2/\sigma^2_{\hat{a}(b)} \end{array} \tag{5.54}$$

For the second component, we may write $\mathcal{D}_2(b) = (\hat{a}(b)-z)^2/\sigma^2_{\hat{a}(b)}$ if $\hat{a}(b) \in [z-\frac{1}{2}, z+\frac{1}{2}]$. Since $\hat{a}(b) = \hat{a} - \sigma_{\hat{a}\hat{b}}\sigma^{-2}_{\hat{b}}(\hat{b} - b)$, we have the equivalence

$$\hat{a}(b) \in [z - \tfrac{1}{2}, z + \tfrac{1}{2}] \Leftrightarrow b \in [b(z) - \tfrac{1}{2}\Delta, b(z) + \tfrac{1}{2}\Delta] \tag{5.55}$$

where $b(z) = \hat{b} + (z - \hat{a})\Delta$ and $\Delta = \sigma^2_{\hat{b}}/\sigma_{\hat{a}\hat{b}}$ (note: $b(z)$ should here not be confused with $\hat{b}(z)$). Using this equivalence one can show that the sum of $\mathcal{D}_1(b)$ and $\mathcal{D}_2(b)$ can be written as

$$\mathcal{D}(b) = \sigma^{-2}_{\hat{b}} \left\{ (\hat{b} - b)^2 + \frac{\rho^2_{\hat{a}\hat{b}}}{1 - \rho^2_{\hat{a}\hat{b}}} (b(z) - b)^2 \right\} \tag{5.56}$$

for $b \in [b(z) - \frac{1}{2}\Delta, b(z) + \frac{1}{2}\Delta]$, $z \in \mathbb{Z}$, where $\rho_{\hat{a}\hat{b}}$ denotes the correlation coefficient of $\hat{a}$ and $\hat{b}$. Thus $\mathcal{D}(b)$ is the sum of a parabola $\mathcal{D}_1(b) = (\hat{b} - b)^2/\sigma^2_{\hat{b}}$ and an infinite z-sequence of equally shaped parabola $\mathcal{D}_2(b) = \frac{\rho^2_{\hat{a}\hat{b}}}{1-\rho^2_{\hat{a}\hat{b}}}(b(z) - b)^2/\sigma^2_b$, centred at $b(z)$ and with domain $b \in [b(z) - \frac{1}{2}\Delta, b(z) + \frac{1}{2}\Delta]$.

Equation (5.56) shows that the contribution of $\mathcal{D}_2(b)$ to $\mathcal{D}(b)$ is driven by the correlation coefficient $\rho_{\hat{a}\hat{b}}$; it is small if the correlation is small and it gets larger the closer the correlation coefficient gets to one. An illustration of $\mathcal{D}(b)$, together with its two components $\mathcal{D}_1(b)$ and $\mathcal{D}_2(b)$, is given in Figure 5.5 for two different values of the correlation coefficient, $\rho_{\hat{a}\hat{b}} = 0.5$ and $\rho_{\hat{a}\hat{b}} = 0.9$. It shows that $\mathcal{D}(b)$ is a *multimodal* function of which the multimodality, with its multiple local minima, gets more pronounced the larger the correlation coefficient gets, i.e. the more weight is given to $\mathcal{D}_2(b)$ in the sum of $\mathcal{D}(b)$. Figure 5.6 illustrates the multimodality of $\mathcal{D}(b)$ for $p = 1$ and $n = 2$, with diagonal (left) and nondiagonal (right) conditional ambiguity variance matrix. The multiple local minima of $\mathcal{D}(b)$ and also the domain in which its global minimizer is guaranteed to reside, are given by the following Lemma.

Lemma 7 (Local minimizers and global domain) (a) The local minimizers and corresponding minima of $\mathcal{D}(b)$ are

$$\left\{ \begin{array}{lcl} \hat{b}(z) & = & \arg \min_{\hat{a}(b) \in \mathscr{S}_z} \mathcal{D}(b) \\ \mathcal{D}(\hat{b}(z)) & = & ||\hat{a} - z||^2_{Q_{\hat{a}\hat{a}}} \ , \ \forall z \in \mathbb{Z}^n \end{array} \right. \tag{5.57}$$

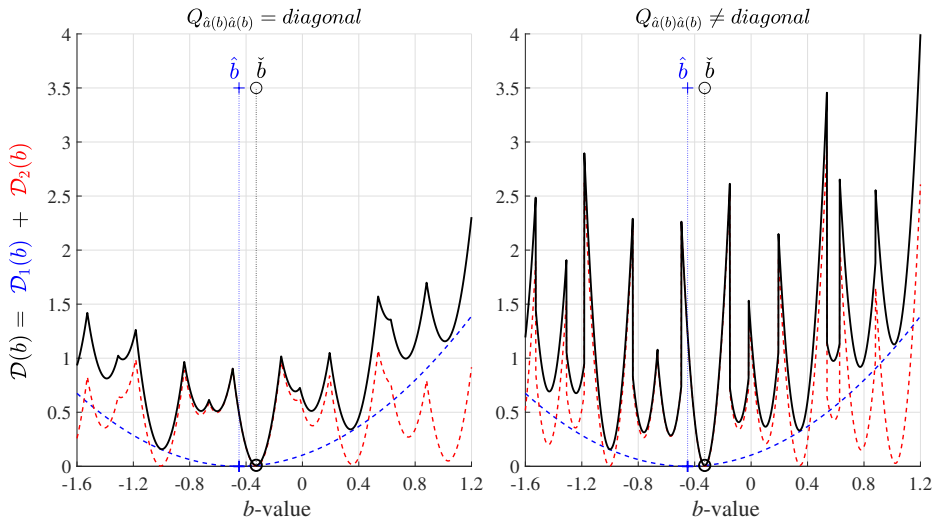


Figure 5.6: The dual function $\mathcal{D}(b)$ for $p = 1$ and $n = 2$, with diagonal (left plot) and nondiagonal (right plot) conditional ambiguity variance matrix.

5

(b) The global minimizer of $\mathcal{D}(b)$,

$$\check{b} = \arg \min_{b \in \mathbb{R}^n} \mathcal{D}(b) \in \mathcal{E}(r_z) \ , \ z \in \mathbb{Z}^n \tag{5.58}$$

resides in the ellipsoidal region $\mathcal{E}(r_z) = \{b \in \mathbb{R}^n \mid ||\hat{b} - b||^2_{Q_{\hat{b}\hat{b}}} \leq r_z^2\}$, for all $z \in \mathbb{Z}^n$, with $r_z^2 = ||\hat{a} - z||^2_{Q_{\hat{a}\hat{a}}}$. ■

Proof: First we prove (5.57). For $\hat{a}(b) \in \mathscr{S}_z$, we have $\check{a}(b) = z$, and therefore

$$\begin{aligned} \mathcal{D}(b) &= ||\hat{b} - b||^2_{Q_{\hat{b}\hat{b}}} + ||\hat{a}(b) - z||^2_{Q_{\hat{a}(b)\hat{a}(b)}} \\ &= ||\hat{a} - z||^2_{Q_{\hat{a}\hat{a}}} + ||\hat{b}(z) - b||^2_{Q_{\hat{b}(a)\hat{b}(a)}} \end{aligned} \tag{5.59}$$

showing that its local minimizer and minimum are given by (5.57). That no other minima exist of $\mathcal{D}(b)$ follows from the fact that the pull-in regions $\mathscr{S}_z$ partition $\mathbb{R}^n$, $\forall z \in \mathbb{Z}^n$ (Teunissen, 1999a). The proof of (5.58) follows by recognizing that as the global minimizer $\check{b}$ is one of the local minimizers, $\check{b}$ resides in the set $\{b \in \mathbb{R}^p \mid \mathcal{D}(b) \leq ||\hat{a} - z||^2_{Q_{\hat{a}\hat{a}}}\}$, and thus also in the larger set E_z. □

This result shows the size of the local minima and where they are located in b-space $\mathbb{R}^p$, but it does not show how their global minimum can be obtained, other then that it is confined to $\mathcal{E}(r_z)$. Due to the presence of this multimodality of $\mathcal{D}(b)$, one can therefore not expect standard iterative descent techniques (Teunissen, 1990) to be successful for finding its global minimum $\check{b}$. We will therefore have to develop a global algorithm dedicated to $\mathcal{D}(b)$.

5.7 Global minimization of dual function

In this section we present our proposed method for finding the global minimizer $\check{b}$ of the dual function $\mathcal{D}(b)$. According to Lemma 7, we can confine the search for $\check{b}$ to a convex set $\mathcal{C} \subset \mathbb{R}^p$, being either a suitably scaled ellipsoid $\mathcal{E}(r) = \{b \in \mathbb{R}^p | \; ||\hat{b} - b||^2_{Q_{\hat{b}\hat{b}}} \leq r^2\}$ or any other of its circumscribing convex regions. The problem to be solved reads therefore

$$\check{b} = \arg \min_{b \in \mathcal{C} \subset \mathbb{R}^p} \mathcal{D}(b) \tag{5.60}$$

The challenge in solving this problem is due to the multimodal dual function not being convex. Although various heuristic and stochastic methods for the *approximate* computation of nonconvex global minimizers exist (Zhigljavsky, 1991; Horst et al., 2000; Pardalos and Romeijn, 2002), we choose to present a method that has finite termination with a guaranteed ϵ-tolerance. Our method for solving (5.60) is constructed from the following three constituents:

1. *Branch and Bound (BB)*: Branch and bound algorithms (Lawler and Wood, 1966; Balakrishnan et al., 1991; Guida, 2015) are methods for *global* minimization of *nonconvex* problems. They are nonheuristic, in the sense that they maintain a provable upper and lower bound on the global minimum, i.e. they terminate with a guarantee that the computed solution has a prescribed accuracy.

2. *Convex relaxation*: To be able to compute the required lower bounds in the BB-algorithm, we construct differentiable, convex lower bounding functions of $\mathcal{D}(b)$ over convex sets. They are constructed such that the lower bounds converge to the nonconvex dual function as the convex sets shrink to a point.

3. *Projected gradient descent (PGD)*: As our convex lower bounding functions are only continuous differentiable (i.e. C^1-functions), the projected gradient descent method (Bertsekas, 1999; Nocedal and Wright, 2006) is used for the computation of their convex constrained minima.

We now describe each of these constituents and how they interrelate and integrate.

5.7.1 Branch and Bound (BB)

The basic idea of the BB-algorithm is

- to partition the initial box $\mathcal{C} \subset \mathbb{R}^p$ in k boxes $\mathcal{B}_i$,

$$\mathcal{C} = \cup_{i=1}^{k} \mathcal{B}_i \tag{5.61}$$

- to find *local* lower and upper bounds of $\mathcal{D}(b)$ for each box $\mathcal{B}_i$,

$$L(\mathcal{B}_i) \leq \min_{b \in \mathcal{B}_i} \mathcal{D}(b) \leq U(\mathcal{B}_i), \tag{5.62}$$

- to form *global* bounds from the local bounds,

$$L_k \leq \min_{b \in \mathcal{C}} \mathcal{D}(b) \leq U_k \tag{5.63}$$

where

$$L_k = \min_{i=1,\ldots,k} L(\mathcal{B}_i) \text{ and } U_k = \min_{i=1,\ldots,k} U(\mathcal{B}_i) \tag{5.64}$$

- to terminate if the difference of these bounds is small enough, $U_k - L_k \leq \epsilon$, else to refine the partition and repeat the process.

The efficacy of the BB-concept depends on the chosen method of partitioning, on the sharpness of the bounds and on the ease with which they can be computed. Importantly, for convergence, the bounds should become tight as the box shrinks to a point.

Although there exist a large variety of different BB-mechanizations, we shall here restrict ourselves to the simple approach where the partitioning of $\mathcal{C} \subset \mathbb{R}^p$ is sequentially constructed through a splitting in half of the boxes. So at the first level, we start with the trivial partitioning, which is $\mathcal{C}$ itself, and compute the lower and upper bounds $L_1 = L(\mathcal{C})$ and $U_1 = U(\mathcal{C})$ (hence, these bounds are local and global at the same time),

$$L_1 = L(\mathcal{C}) \leq \mathcal{D}(\check{b}) \leq U_1 = U(\mathcal{C}) \tag{5.65}$$

If $U_1 - L_1 \leq \epsilon$, the algorithm terminates. Otherwise we go to the second iteration level and partition $\mathcal{C}$ into two boxes $\mathcal{C} = \mathcal{B}_1 \cup \mathcal{B}_2$, and compute $L(\mathcal{B}_i)$ and $U(\mathcal{B}_i)$, $i = 1, 2$. The splitting of the box is usually done along its longest edge. Then we can construct new global lower and upper bounds,

$$\underbrace{\min(L(\mathcal{B}_1), L(\mathcal{B}_2))}_{L_2} \leq \mathcal{D}(\check{b}) \leq \underbrace{\min(U(\mathcal{B}_1), U(\mathcal{B}_2))}_{U_2} \tag{5.66}$$

As both $\mathcal{B}_1$ and $\mathcal{B}_2$ are 'smaller' than $\mathcal{C}$ (i.e. they are its partition), one can generally expect the local bounds for $\mathcal{B}_i$ to be sharper than the previous global bounds are for $\mathcal{B}_i$. One can therefore assume that the lower and upper bounds of the pair of boxes obtained by splitting are no worse than the lower and upper bounds of the box they were formed from.

If $U_2 - L_2 < \epsilon$, the algorithm terminates. Otherwise, we partition one of $\mathcal{B}_1$ and $\mathcal{B}_2$ into two boxes, to obtain a new partition of $\mathcal{C}$ into three boxes, and we compute the local lower and upper bounds for these new boxes. We then update the global lower bound L_3 as the minimum of the local lower bounds over the partition of $\mathcal{C}$, and similarly for the upper bound U_3. The choice which of the two boxes to split, $\mathcal{B}_1$ or $\mathcal{B}_2$, is based on the value of their local lower bound. The box to be split is the one of which the local lower bound equals the global lower bound, i.e. the one that has the smallest local lower bound. As at each iteration level a box is split into two, we have after k iterations a partitioning of the form (5.61), with associated global lower and upper bounds of $\mathcal{D}(\check{b})$ as given in (5.63), with L_k nondecreasing and U_k nonincreasing. Note, although the choice of which box to split may not be correct in the sense that it does not contain the solution $\check{b}$, at a certain following stage the BB-algorithm will revisit the nonselected box containing $\check{b}$ as its local lower bound will then have become the smallest.

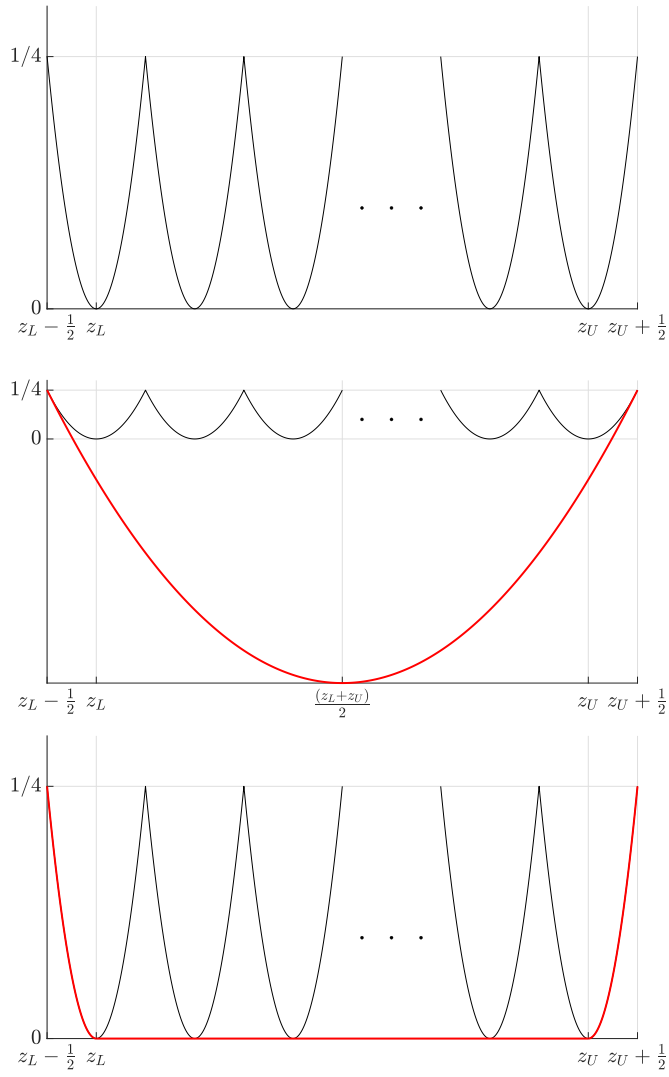


Figure 5.7: (*a*) The function $g(x) = (x - \lfloor x \rceil)^2$ as a sequence of cut-off parabola on interval $[z_L - \frac{1}{2}, z_U + \frac{1}{2}]$; (*b*) with its parabolic lower bounding function; and (c) with its best possible, continuous differentiable lower bounding convex function.

Initialization and Bounds

To start the BB-algorithm, the initial box $\mathcal{C}$ needs to be formed. We choose $\mathcal{C}$ to be the box

$$\mathcal{C} = \{b \in \mathbb{R}^p | \, |b_\alpha - \hat{b}_\alpha| \leq r\sigma_{\hat{b}_\alpha}, \alpha = 1, \ldots, p\} \tag{5.67}$$

It follows from the ellipsoidal *planes-of-support* lemma that the box $\mathcal{C}$ is circumscribed

by the ellipsoid $\mathcal{E}(r) = \{b \in \mathbb{R}^p |\ ||b - \hat{b}||^2_{Q_{\hat{b}\hat{b}}} \leq r^2\}$, see e.g. (Teunissen, 1995). The scalar $r > 0$ is a user-defined parameter. It can be set following Lemma 7, or by choosing a user-defined confidence-level. In the latter case, $r^2 = \chi^2_\alpha(p, 0)$ corresponds with a confidence-level $1 - \alpha$.

For the bounds we need to be able to compute upper and lower bounds of $\min_{b\in\mathcal{B}} \mathcal{D}(b)$ for any relevant box $\mathcal{B}$ that the BB-algorithm creates. The computation of local *upper* bounds $U(\mathcal{B})$ is rather straightforward, since any $b \in \mathcal{B}$ can be used for that purpose. We choose to compute the bound as

$$U(\mathcal{B}) = \mathcal{D}(b_*) \geq \min_{b\in\mathcal{B}} \mathcal{D}(b) \tag{5.68}$$

with b_* being the 'centre of gravity' of the box, i.e. if box $\mathcal{B}$ is bounded as $b_L \preceq b \preceq b_U$, then $b_* = \frac{1}{2}(b_L + b_U)$.

The computation of local *lower* bounds $L(\mathcal{B})$ is much more involved. We cannot use standard gradient-based methods for computing the minimizer, since $\mathcal{D}(b)$ is not convex and convergence is therefore not assured. The idea is therefore to find a *differentiable convex lower bounding* function $\mathcal{D}_L(b) \leq \mathcal{D}(b), \forall b \in \mathcal{B}$ such that the minimizer of $\mathcal{D}_L(b)$ over $\mathcal{B}$ can be computed with standard means and used as the local lower bound

$$L(\mathcal{B}) = \min_{b\in\mathcal{B}} \mathcal{D}_L(b) \leq \min_{b\in\mathcal{B}} \mathcal{D}(b) \tag{5.69}$$

We now show how this can be achieved.

5.7.2 Convex relaxation

We will develop the convex relaxation for the dual function $\mathcal{D}^\circ(x)$, as a similar approach can be developed for the other dual versions, like $\mathcal{D}^\bullet(x)$. Using the diagonal approximation $Q^\circ_{\hat{a}(b)\hat{a}(b)} = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_n^2)$, the dual function (5.29) can be written as

$$\mathcal{D}^\circ(x) = ||b - \hat{b}||^2_{Q_{\hat{b}\hat{b}}} + \mathcal{G}(b) \tag{5.70}$$

with

$$\begin{array}{lcl} \mathcal{G}(b) & = & \sum_{i=1}^n g(x_i)/\sigma_i^2 \\ g(x) & = & (x - \lfloor x \rceil)^2 \\ x_i & = & \hat{a}_i(b) = \hat{a}_i + m_i^T(b - \hat{b}) \\ m_i^T & = & c_i^T Q_{\hat{a}\hat{b}} Q^{-1}_{\hat{b}\hat{b}} \\ c_i & = & i\text{th column of unit matrix } I_n \end{array} \tag{5.71}$$

Since the first term of (5.70) is already convex for the whole space $\mathbb{R}^p$, we can concentrate on the second term and try to find a convex differentiable lower bounding function $\mathcal{G}_L(b)$ such that

$$\mathcal{G}_L(b) \leq \mathcal{G}(b),\ \forall b \in \mathcal{B} \tag{5.72}$$

Once this function is found, we have found the dual convex lower bounding function as $\mathcal{D}^\circ_L(b) = ||b - \hat{b}||^2_{Q_{\hat{b}\hat{b}}} + \mathcal{G}_L(b) \leq \mathcal{D}(b),\ \forall b \in \mathcal{B}$.

Note that the lack of convexity of $\mathcal{G}(b)$ is due to the single function $g(x)$, which itself is a sequence of cut-off parabolae, see Figure 5.7(a). Hence, if we can find a convex lower bounding function $g_L(x)$ of $g(x)$ on the required interval, then we automatically have constructed a convex $\mathcal{G}_L(b)$ on the required box $\mathcal{B}$. To do so, we first need to construct the intervals of $\hat{a}_i(b)$, $i = 1, \ldots, n$, that correspond with $b_L \preceq b \preceq b_U$. As these n intervals will differ, the lower bounding functions of $g(x)$ on these intervals will differ as well. They will be denoted as $g_{i,L}(x)$.

The interval $[l_i, u_i]$ for which $g_{i,L}(x)$ has to be convex

As the function $\mathcal{G}_L(b) = \sum_{i=1}^{n} g_{i,L}(\hat{a}_i(b))/\sigma_i^2$ has to be convex for the box $b_{\alpha,L} \leq b_\alpha \leq b_{\alpha,U}$, $\alpha = 1, \ldots, p$, the functions $g_{i,L}(x)$ need to be convex for the intervals $l_i \leq a_i(b) \leq u_i$ that correspond with this box. Application of the *projection-lemma* from the Appendix shows the relation between these intervals given as

$$b_{\alpha,L} \leq b_\alpha \leq b_{\alpha,U} \;\Rightarrow\; l_i \leq \hat{a}_i(b) \leq u_i \tag{5.73}$$

with

$$\begin{aligned} l_i &= \hat{a}_i(0) + \sum_{\alpha=1}^{p} (-\tilde{b}_\alpha |m_{i\alpha}| + \bar{b}_\alpha m_{i\alpha}) \\ u_i &= \hat{a}_i(0) + \sum_{\alpha=1}^{p} (+\tilde{b}_\alpha |m_{i\alpha}| + \bar{b}_\alpha m_{i\alpha}) \end{aligned} \tag{5.74}$$

where $\tilde{b}_\alpha = \frac{1}{2}(b_{\alpha,U} - b_{\alpha,L})$, $\bar{b}_\alpha = \frac{1}{2}(b_{\alpha,U} + b_{\alpha,L})$, and $m_{i\alpha} = c_i^T Q_{\hat{a}\hat{b}} Q_{\hat{b}\hat{b}}^{-1} c_\alpha$.

Note that the widths of the baseline intervals $b_{\alpha,U} - b_{\alpha,L}$ propagate into the widths of the ambiguity intervals as $u_i - l_i = \sum_{\alpha=1}^{p} (b_{\alpha,U} - b_{\alpha,L})|m_{i\alpha}|$. Thus as the baseline intervals get smaller due to the rectangular BB-splitting, the corresponding ambiguity intervals get smaller as well, producing in the limit, when $b_{\alpha,U} = b_{\alpha,L} = \check{b}_\alpha$, the result $\hat{a}(\check{b})$.

As the BB-splitting acts on only one interval at a time, we can now also show how this halfway splitting affects the intervals $[l_i, u_i]$. Let the γth interval $[b_{\gamma,L}, b_{\gamma,U}]$, with $\gamma \in \{1, \ldots, p\}$ be split halfway in $[b_{\gamma,L}, \bar{b}_\gamma]$ and $[\bar{b}_\gamma, b_{\gamma,U}]$. Then the new intervals $[l_i^1, u_i^1]$, corresponding with $[b_{\gamma,L}, \bar{b}_\gamma]$, can be expressed in the old as

$$\begin{aligned} l_i^1 &= l_i + \tfrac{1}{2}\tilde{b}_\gamma(|m_{i\gamma}| - m_{i\gamma}) \\ u_i^1 &= u_i - \tfrac{1}{2}\tilde{b}_\gamma(|m_{i\gamma}| + m_{i\gamma}) \end{aligned} \tag{5.75}$$

Hence, the length of the interval changes as $(u_i^1 - l_i^1) = (u_i - l_i) - \tilde{b}_\gamma |m_{i\gamma}|$, i.e. it gets shorther by $\tilde{b}_\gamma |m_{i\gamma}|$, where $\tilde{b}_\gamma = \frac{1}{2}(b_{\gamma,U} - b_{\gamma,L})$.

Now that we know the intervals $[l_i, u_i]$ over which the functions $g_{i,L}(x)$, $i = 1, \ldots, n$, need to provide a differentiable lower bounding of $g(x)$, we can start constructing these functions. As we will do so for an arbitrary interval $[l, u]$, we will dispense with the lower index i and write $g_L(x)$ instead of $g_{i,L}(x)$.

Convex lower bounding function of $g(x)$ on $[l, u]$

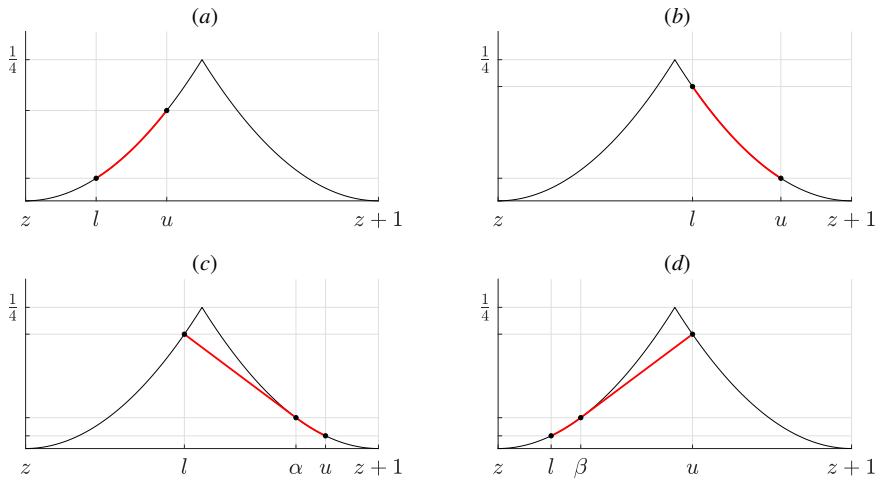


Figure 5.8: Four convex differentiable lower bounding functions (in red) of $g(x) = (x - \lceil x \rfloor)^2$ on the interval $[l, u] \subset [z, z+1]$, $z \in \mathbb{Z}$

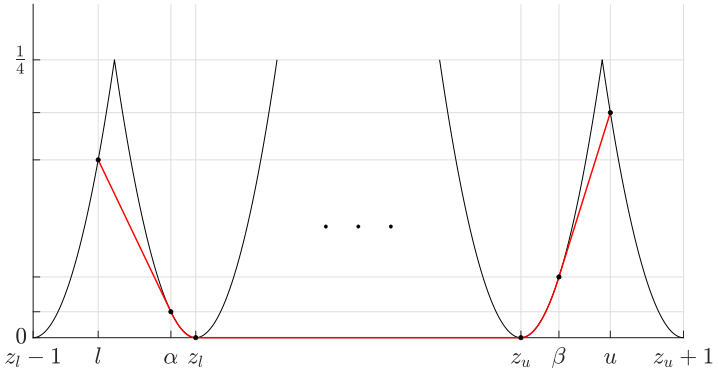


Figure 5.9: The linear-parabolic differentiable lower bounding function (in red) of $g(x) = (x - \lceil x \rfloor)^2$ when $l \in [z_l - 1, z_l - \frac{1}{2}]$, $u \in [z_u + \frac{1}{2}, z_u + 1]$ for $z_u \geq z_l$.

Many different differentiable convex lower bounding functions $g_L(x)$ of $g(x) = (x - \lceil x \rfloor)^2$ on $x \in [l, u]$ can be constructed. For example, if we assume for the moment that $l = z_l - \frac{1}{2}$, $z_l \in \mathbb{Z}$, and $u = z_u + \frac{1}{2}$, $z_u \in \mathbb{Z}$, then the parabola

$$g_L(x) = \alpha(x - \bar{z})^2 + \beta \tag{5.76}$$

with $\alpha = \frac{1}{2}(\frac{1}{2} + \tilde{z})^{-1}$, $\tilde{z} = \frac{z_u - z_l}{2}$, $\bar{z} = \frac{z_u + z_l}{2}$, and $\beta = -\frac{1}{2}\tilde{z}$, is such lower bounding function of $g(x)$, see Figure 5.7b. It gives a perfect fit to $g(x)$ if $u - l = 1$. However, the longer the interval $[l, u]$ becomes, i.e. the larger the difference $u - l$, the more negative the minimum β of the parabola becomes. With choice (5.76), the lower bounding fit to $g(x)$ gets thus poorer the more l and u differ.

A much better choice for the lower bounding function would be (see Figure 5.7c)

$$g_L(x) = \begin{cases} (x - z_l)^2 & \text{for} \quad l = z_l - \frac{1}{2} \leq x \leq z_l \\ 0 & \text{for} \quad z_l \leq x \leq z_u \\ (x - z_u)^2 & \text{for} \quad z_u \leq x \leq z_u + \frac{1}{2} = u \end{cases} \tag{5.77}$$

Also this choice gives a perfect fit when $u - l = 1$, while now its minimum is independent of the interval length. In fact, the minimum value of 0 provides the best possible convex lower bounding over the interval $[z_l, z_u]$. When we compare the smoothness of the above two choices, we note that (5.76) is a C^2 function, while (5.77) is only a C^1 function. The continuous differentiability of (5.77) is sufficient however for the application of the projected gradient descent method (see subsection 5.7.3).

So far we made the simplifying, but unrealistic assumption that $l = z_l - \frac{1}{2}$ and $u = z_u + \frac{1}{2}$. For the general situation in which $[l, u]$ can be any interval, we first assume that the interval lies inside the interval bordered by two consecutive integers: $[l, u] \subset [z, z + 1]$ for some $z \in \mathbb{Z}$. Then the following 4 different cases can be discriminated (see Figure 5.8):

Case 1: If $[l, u] \subset [z, z + \frac{1}{2}]$, then $g_L(x) = (x - z)^2$. Thus if both l and u lie in the first half of the interval, the parabola itself can be taken as the lower bounding function, see Figure 5.8a.

Case 2: If $[l, u] \subset [z + \frac{1}{2}, z + 1]$, then $g_L(x) = (x - z - 1)^2$. Thus if both l and u lie in the second half of the interval, the parabola centred at $z + 1$ can be taken as the lower bounding function, see Figure 5.8b.

For the third and the fourth case, we assume that l and u lie in different halfs of the interval $[z, z + 1]$, whereby then the discrimination has to be made whether or not the function value at l is larger than at u.

Case 3: Let $l \in [z, z + \frac{1}{2}]$, $u \in [z + \frac{1}{2}, z + 1]$, and $(l - z)^2 \geq (u - z - 1)^2$. Now the best convex lower bounding differentiable function is either a decreasing straight line, connecting the points $(l, (l-z)^2)$ and $(u, (u-z-1)^2)$, or a decreasing straight line that starts at the point $(l, (l - z)^2)$, is tangent of the parabola $y = (x - z - 1)^2$ at the point $(\alpha, (\alpha - z - 1)^2)$, and then continues along the parabola to the point $(u, (u - z - 1)^2)$, see Figure 5.8c. It is given as:

$$g_L(x) = \begin{cases} \overset{u \leq \alpha}{=\!=} a(x - l) + (l - z)^2 \text{ for } l \leq x \leq u \\ \overset{u \geq \alpha}{=\!=} \begin{cases} = & 2(\alpha - z - 1)(x - \alpha) + (\alpha - z - 1)^2 \\ & \text{for } l \leq x \leq \alpha \\ = & (x - z - 1)^2 \\ & \text{for } \alpha \leq x \leq u \end{cases} \end{cases} \tag{5.78}$$

with $\alpha = l + \sqrt{1 - 2(l - z)}$ and $a = \frac{(u-z-1)^2-(l-z)^2}{u-l} \leq 0$.

Case 4: Let $l \in [z, z+\frac{1}{2}]$, $u \in [z+\frac{1}{2}, z+1]$, and $(l-z)^2 \leq (u-z-1)^2$. This case is a 'mirror-image' of the previous one, see Figure 5.8d. The convex lower bounding differentiable function is given as:

$$g_L(x) = \begin{cases} \overset{l \geqq \beta}{\equiv} a(x-l) + (l-z)^2 \text{ for } l \leq x \leq u \\ \overset{l \leqq \beta}{\equiv} \begin{cases} = & (x-z)^2 \\ & \text{for } l \leq x \leq \beta \\ = & 2(\beta - z)(x-\beta) + (\beta - z)^2 \\ & \text{for } \beta \leq x \leq u \end{cases} \end{cases} \tag{5.79}$$

with $\beta = u - \sqrt{2(u-z)-1}$ and $a = \frac{(u-z-1)^2-(l-z)^2}{u-l} \geq 0$.

We can use the above insight also to construct lower bounding functions in case $[l,u] \subset [z, z+1]$ is not true, but instead $l \in [z_l - 1, z_l]$ and $u \in [z_u, z_u + 1]$ for some integers $z_l \leq z_u$, see Figure 5.9. When $l \in [z_l - 1, z_l - \frac{1}{2}]$ and $u \in [z_u + \frac{1}{2}, z_u + 1]$, the lower bounding function will be given as

$$g_L(x) = \begin{cases} f_1(l,x) & \text{for} \quad l \leq x \leq z_l \\ 0 & \text{for} \quad z_l \leq x \leq z_u \\ f_2(x,u) & \text{for} \quad z_u \leq x \leq u \end{cases} \tag{5.80}$$

with $f_1(l,x)$ and $f_2(x,u)$ constructed such that they provide the required lower bounds. They can be found directly from (5.78), through replacement $z := z_l - 1$, and from (5.79), through replacement $z := z_u$, as

$$f_1(l,x) = \begin{cases} 2(\alpha - z_l)(x-\alpha) + (\alpha - z_l)^2 & \text{for} \quad l \leq x \leq \alpha \\ (x-z_l)^2 & \text{for} \quad \alpha \leq x \leq z_l \end{cases} \tag{5.81}$$

and

$$f_2(x,u) = \begin{cases} (x-z_u)^2 & \text{for} \quad z_u \leq x \leq \beta \\ 2(\beta - z_u)(x-\beta) + (\beta - z_u)^2 & \text{for} \quad \beta \leq x \leq u \end{cases} \tag{5.82}$$

where $\alpha = l + \sqrt{1 - 2(l - z_l + 1)}$ and $\beta = u - \sqrt{2(u-z_u)-1}$.

Note that either one or both of the straight line components of (5.81) and (5.82) will be absent when $l \in [z_l - \frac{1}{2}, z_l]$ or/and $u \in [z_u, z_u + \frac{1}{2}]$. Both are absent when $l \in [z_l - \frac{1}{2}, z_l]$ and $u \in [z_u, z_u + \frac{1}{2}]$, in which case the lower bounding function generalizes (5.77) to become

$$g_L(x) = \begin{cases} (x-z_l)^2 & \text{for} \quad l \leq x \leq z_l \\ 0 & \text{for} \quad z_l \leq x \leq z_u \\ (x-z_u)^2 & \text{for} \quad z_u \leq x \leq u \end{cases} \tag{5.83}$$

With the above construction of $g_L(x)$ for the different scenarios we are now in the position to formulate our sought for lower bounding function as $\mathcal{D}_L^\circ(b) = ||\hat{b} - b||^2_{Q_{\hat{b}\hat{b}}} + \sum_{i=1}^n g_{i,L}(\hat{a}_i(b))/\sigma_i^2$. It is continuous differentiable, convex and lower bounding $\mathcal{D}^\circ(b)$ on the interval $b_L \preceq b \preceq b_U$. Hence, it is now in the form that the PGD-method can be applied to it.

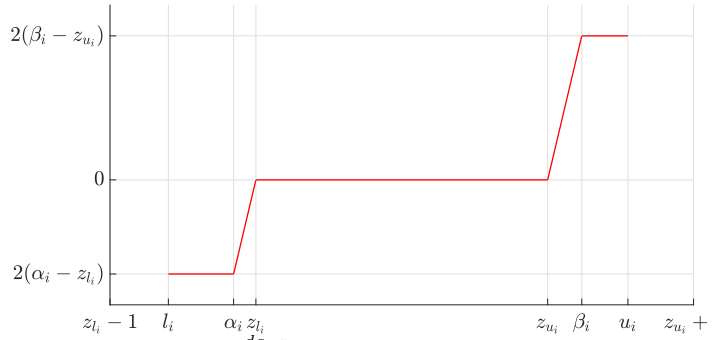


Figure 5.10: The function $g'_{i,L}(x) = \frac{dg_{i,L}}{dx}(x)$ (cf. 5.91) on the interval $[l_i, u_i]$ for $l_i \in [z_{l_i} - 1, z_{l_i} - \frac{1}{2}]$, $u_i \in [z_{u_i} + \frac{1}{2}, z_{u_i} + 1]$.

5.7.3 Projected Gradient Descent (PGD)

5

The PGD-method is designed to solve a constrained minimization problem

$$\min_{b \in \mathcal{B}} f(b) \tag{5.84}$$

of which the objective function $f : \mathbb{R}^p \to \mathbb{R}$ is only continuous differentiable, i.e. C^1-function (Bertsekas, 1999; Nocedal and Wright, 2006; Parikh and Boyd, 2013; Nesterov, 2018). This implies that only first order gradient information of the objective function can be used. If in addition both the objective function and the constraint are convex, then any local minimum is automatically a global minimum. This is the case before us when $f(b) := \mathcal{D}_L(b)$ and $\mathcal{B} = \{b \in \mathbb{R}^p | \; b_L \preceq b \preceq b_U\}$.

The PGD algorithmic steps for solving (5.84) are:

1. Initialize: Start with a feasible solution, $b^0 \in \mathcal{B}$ and then loop for $k = 0, \ldots$ until stop criterium:
2. Gradient descent step: Compute stepsize μ_k and gradient descent

$$\hat{b}^k = b^k - \mu_k \nabla f(b^k) \tag{5.85}$$

   such that $f(\hat{b}^k) \leq f(b^k)$.
3. Projection step: Project $\hat{b}^k$ onto $\mathcal{B}$ to ensure conformity with the constraints,

$$b^{k+1} = P_{\mathcal{B}}(\hat{b}^k) \text{ with } P_{\mathcal{B}}(y) = \arg\min_{b \in \mathcal{B}} ||y - b||^2 \tag{5.86}$$

This iterative scheme can be seen as repeatedly solving an approximate version of the original minimization problem, namely one in which the objective function $f(b)$ is approximated by a *quadratic* function $F_k(b) = f(b^k) + \nabla f(b^k)^T (b - b^k) + \frac{1}{2\mu_k} ||b - b^k||^2$ ($\mu_k > 0$). The above iterative scheme can then be summarized as repeatedly solving

$$b^{k+1} = \arg\min_{b \in \mathcal{B}} F_k(b) \tag{5.87}$$

To see this, we first rewrite the quadratic function $F_k(b)$ in the more convenient form $F_k(b) = \frac{1}{2\mu_k}||\hat{b}^k - b||^2 + c_k$, with $c_k = f(b^k) - \frac{1}{2}\mu_k||\nabla f(b^k)||^2$ and $\hat{b}^k = b^k - \mu_k \nabla f(b^k)$, showing that the gradient step (5.85) provides the unconstrained minimizer of $F_k(b)$. Substitution of $F_k(b) = \frac{1}{2\mu_k}||\hat{b}^k - b||^2 + c_k$ into (5.87) gives $b^{k+1} = \arg\min_{b\in\mathcal{B}} ||\hat{b}^k - b||^2$ and therefore $b^{k+1} = P_{\mathcal{B}}(\hat{b}^k)$, which is the projection step (5.86).

For computing the stepsize μ_k in each iteration, different linesearch strategies exist, from simple to advanced (Nesterov, 2018). One of the simplest starts with $\mu_k = 1$, followed by halving it, $\mu_k \leftarrow \mu_k/2$, until $f(b^k - \mu_k \nabla f(b^k)) < f(b^k)$. More involved accelerated strategies exist, where b^{k+1} is taken as a convex combination (weighted mean) of b^k and $P_{\mathcal{B}}(b^k - \mu_k \nabla f(b^k))$.

A potential complicating factor in applying the PGD-method lies in the projection onto the convex set $\mathcal{B}$, which, depending on the geometry of $\mathcal{B}$, can be quite involved. The PGD-method is only efficient if this projection can be done efficiently. Fortunately, in our case, with the convex set given as $\mathcal{B} = \{b \in \mathbb{R}^p |\ b_L \preceq b \preceq b_U\}$, the projection can be done very efficiently.

As $P_{\mathcal{B}}(y) = \arg\min_{b\in\mathcal{B}} ||y - b||^2$, the minimization problem to be solved is

$$\begin{aligned} \min_{b_L \preceq b \preceq b_U} ||y-b||^2 &= \min_{b_L \preceq b \preceq b_U} \sum_{\alpha=1}^{p} (y_\alpha - b_\alpha)^2 \\ &= \sum_{\alpha=1}^{p} \min_{b_{\alpha,L} \le b_\alpha \le b_{\alpha,U}} (y_\alpha - b_\alpha)^2 \\ &= \sum_{\alpha=1}^{p} (y_\alpha - \text{median}(b_{\alpha,L}, y_\alpha, b_{\alpha,U}))^2 \end{aligned} \tag{5.88}$$

in which the minimizer $\text{median}(b_{\alpha,L}, y_\alpha, b_{\alpha,U})$ denotes the *median* value of the triplet $b_{\alpha,L}$, y_α, and $b_{\alpha,U}$. We therefore have,

$$P_{\mathcal{B}}(y)_\alpha = \begin{cases} b_{\alpha,L} & \text{if} \quad y_\alpha \le b_{\alpha,L} \\ y_\alpha & \text{if} \quad b_{\alpha,L} \le y_\alpha \le b_{\alpha,U} \\ b_{\alpha,U} & \text{if} \quad b_{\alpha,U} \le y_\alpha \end{cases} \tag{5.89}$$

What now remains to be determined for the PGD-method to be applicable to solve $L(\mathcal{B}) = \min_{b\in\mathcal{B}} \mathcal{D}_L^\circ(b)$ is the gradient of the objective function.

The gradient of $\mathcal{D}_L^\circ(b)$

The gradient of $\mathcal{D}_L^\circ(b) = ||\hat{b} - b||^2_{Q_{\hat{b}\hat{b}}} + \mathcal{G}_L(b)$ is given as

$$\begin{aligned} \nabla\mathcal{D}_L(b) &= 2Q^{-1}_{\hat{b}\hat{b}}(b - \hat{b}) + \nabla\mathcal{G}_L(b) \\ &= 2Q^{-1}_{\hat{b}\hat{b}}(b - \hat{b}) + \sum_{i=1}^{n} \sigma_i^{-2} g'_{i,L}(\hat{a}_i(b)) m_i \\ &= 2Q^{-1}_{\hat{b}\hat{b}}[b - \hat{b} + Q_{\hat{b}\hat{a}} Q^{o-1}_{\hat{a}\hat{a}} s(b)] \end{aligned} \tag{5.90}$$

with $s(b) = \frac{1}{2}[g'_{1,L}(\hat{a}_1(b)), \ldots, g'_{n,L}(\hat{a}_n(b))]^T$ and $g'_{i,L}(x) = \frac{dg_{i,L}}{dx}(x)$. The entries of the vector $s(b)$ are driven by the intervals $[l_i, u_i]$, the derivatives $g'_{i,L}(x)$ of the functions

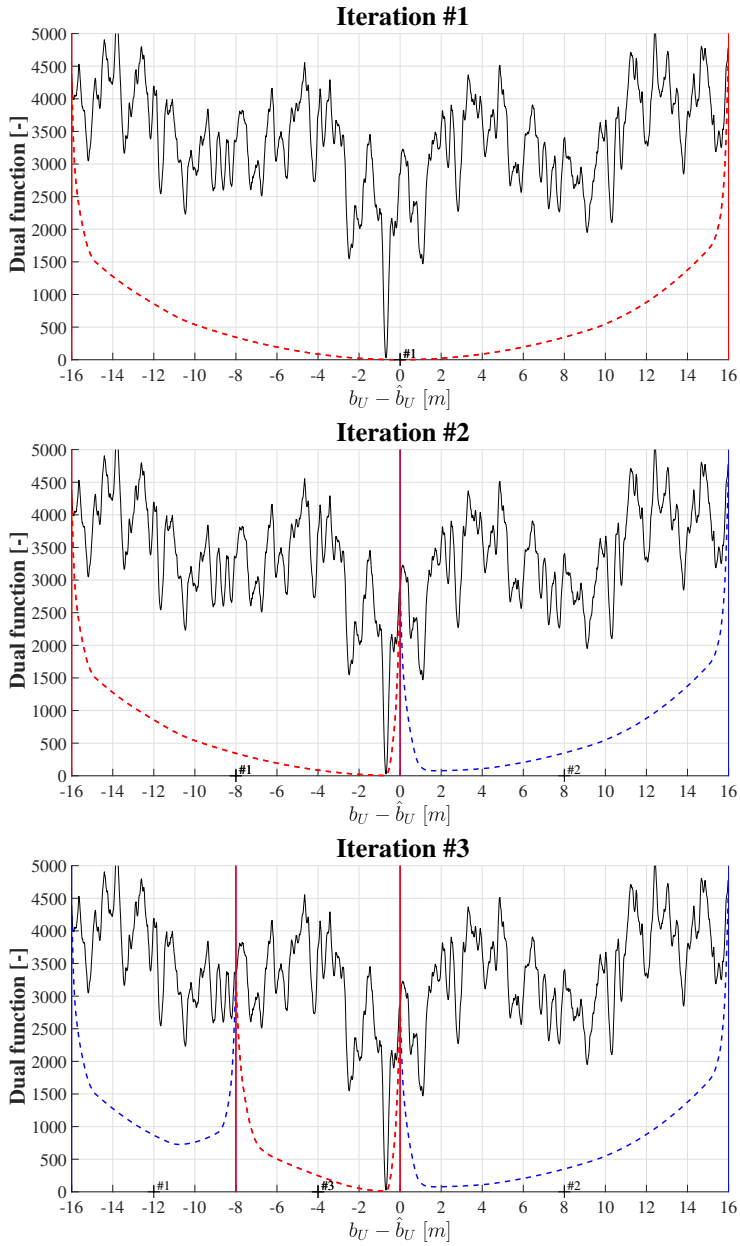


Figure 5.11: Multimodal dual-function of vertical positioning mixed-integer GNSS model, with its per iteration constructed convex lower bounding functions (red and blue) over intervals that get split for the red functions (i.e. intervals for which minimum of lower bounding function is lowest). Convergence was achieved in 7 iterations. Shown are the results of iterations #1, #2, #3, #6, and #7, with an additional zoom-in of #7. (Continued in Figure 5.12.)

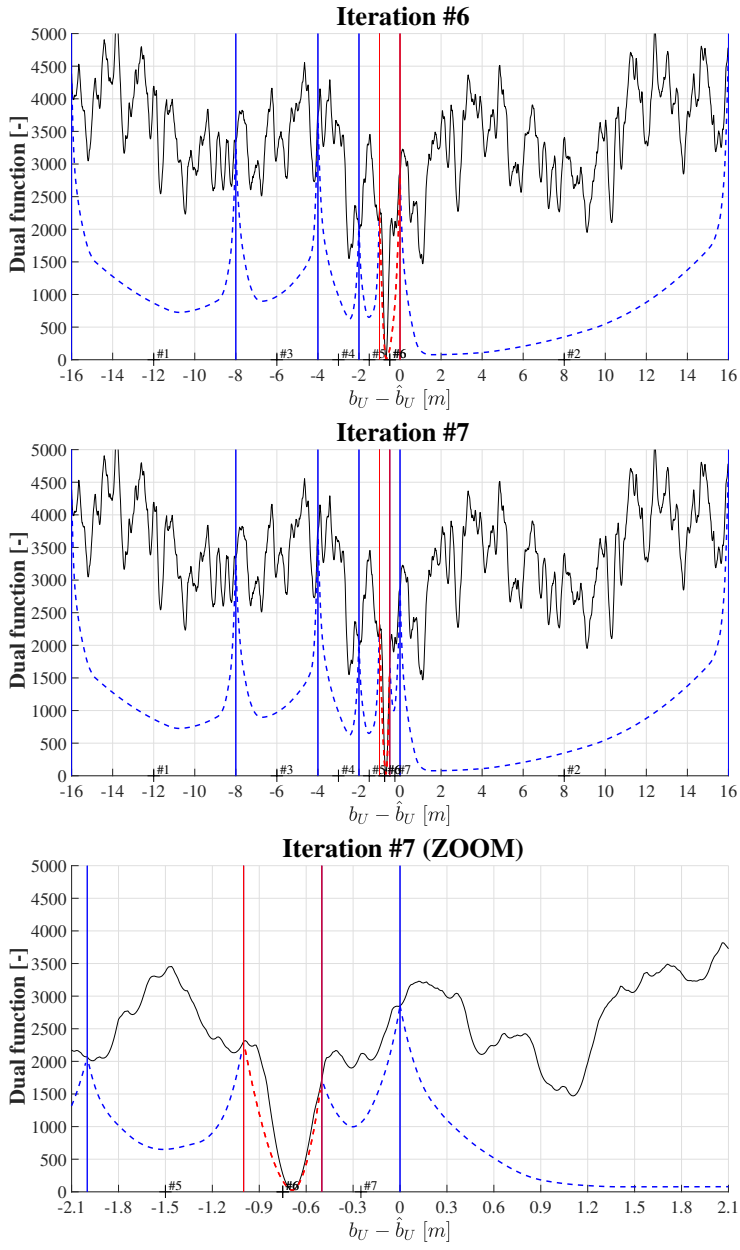


Figure 5.12: Continued from Figure 5.11

that are convex lower bounding on $[l_i, u_i]$, and the locations of the $\hat{a}_i(b)$ within the intervals $[l_i, u_i]$. For instance, for $l_i \in [z_{l_i} - 1, z_{l_i} - \frac{1}{2}]$ and $u_i \in [z_{u_i} + \frac{1}{2}, z_{u_i} + 1]$, with

$z_{u_i} \geq z_{l_i}$, the applicable derivative $g'_{i,L}(x)$ follows from (5.80), (5.81) and (5.82) as

$$g'_{i,L}(x) = \begin{cases} 2(\alpha_i - z_{l_i}) & \text{for} \quad l_i \leq x \leq \alpha_i \\ 2(x - z_{l_i}) & \text{for} \quad \alpha_i \leq x \leq z_{l_i} \\ 0 & \text{for} \quad z_{l_i} \leq x \leq z_{u_i} \\ 2(x - z_{u_i}) & \text{for} \quad z_{u_i} \leq x \leq \beta_i \\ 2(\beta_i - z_{u_i}) & \text{for} \quad \beta_i \leq x \leq u_i \end{cases} \tag{5.91}$$

with

$$\begin{aligned} \alpha_i &= l_i + \sqrt{1 - 2(l_i - z_{l_i} + 1)} \\ \beta_i &= u_i - \sqrt{1 - 2(z_{u_i} + 1 - u_i)} \end{aligned} \tag{5.92}$$

The behaviour of $g'_{i,L}(x)$ for $x \in [l_i, u_i]$ is illustrated in Figure 5.10. It shows that the entries of $s(b)$ are determined, in dependence of the location of $\hat{a}_i(b)$, through a mixed hard-soft thresholding, see Figure 5.10.

We now present two examples to illustrate the workings of our global algorithm. To provide an insightful graphical display of the box-splitting iterations, we show the results for the 1D and the 2D case, i.e. $b \in \mathbb{R}$ and $b \in \mathbb{R}^2$.

Example 3 (Horizontal position known) This example and the next one are based on the single-frequency, single-epoch, single-baseline linearized GNSS model of example 1 (cf. 5.24), using a $m + 1 = 31$ satellite configuration having $n = 30$ DD ambiguities, with signal wavelength $\lambda = 19.03$ cm, and pseudorange and carrier-phase standard deviations of $\sigma_p = 20$ cm and $\sigma_\phi = 0.2$ cm, respectively. As we use satellite-elevation weighting, the used variance matrices of the observables are: $\mathsf{D}(p) = \sigma_p^2 D^T Q D$ and $\mathsf{D}(\phi) = \sigma_\phi^2 D^T Q D$, with $Q = \text{diag}[\sin^{-1}(\text{el}_i), \ldots, \sin^{-1}(\text{el}_{m+1})]$, where el_i is the elevation angle of satellite i. In the current example we have $b \in \mathbb{R}$, as we assume the horizontal positions known, while in the next example, we have $b \in \mathbb{R}^2$, as the vertical positions are then assumed known. Note, as the model is in linearized form, that the parameters are increments with respect to the chosen approximate values, which in our case are taken as the correct values used for the simulation. Hence, an increment value of zero implies that the parameter value equals the correct value. In our case the float increment is computed to be $\hat{b} = 0.7042$ m.

Although the initialization in practice should follow our description of (5.67), we here use a larger, and thus more conservative, value for the initial box or interval size. Here we have chosen the initial interval to be $[-16\text{m}, +16\text{m}]$. The reason for this choice of larger interval length is that the corresponding increase in required number of iterations allows us to better show the various steps that are taken in the splitting process and construction of the convex lower bounding functions. The results of the various iteration steps are shown in Figure 5.11, starting at the top of the first column with iteration #1 and finishing at the bottom of the second column with iteration #7. The first interval $[-16\text{m}, +16\text{m}]$ is designated to be centred at 0, but which corresponds to the float solution $\hat{b}$. Thus the horizontal axis denotes the difference $b - \hat{b}$. The multimodal dual function is shown as a black curve, while the convex lower bounding functions are shown as either red or blue dashed curves. The lower bounding

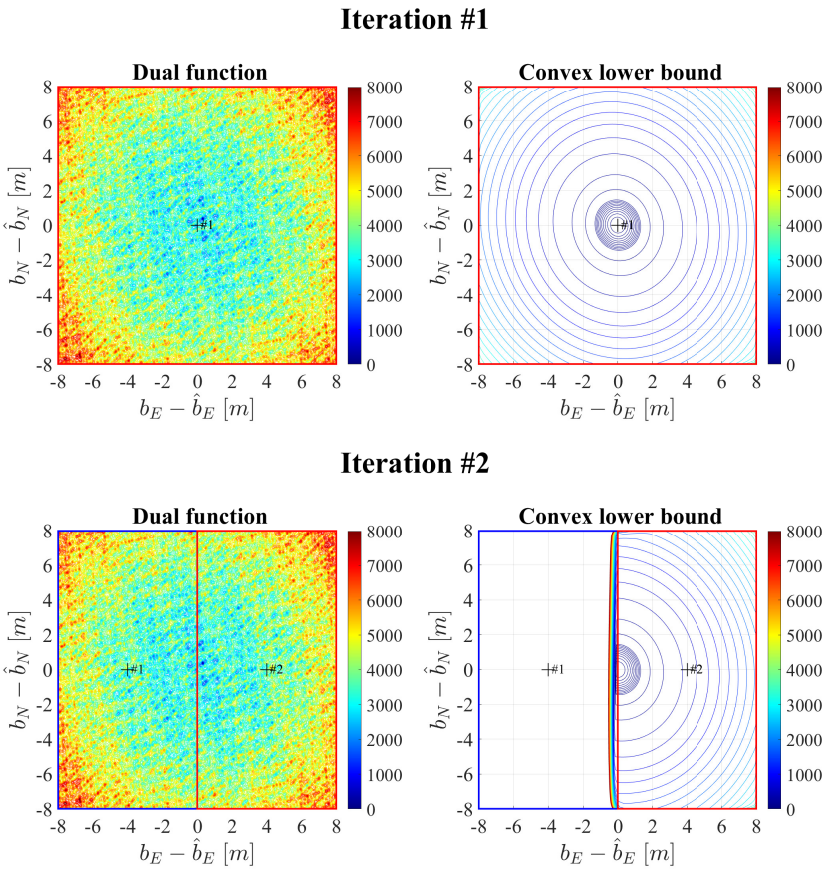


Figure 5.13: Multimodal dual-function of horizontal positioning mixed-integer GNSS model, with the contour lines of its per iteration constructed convex lower bounding functions over the rectangles that get split (i.e. rectangles for which minimum of lower bounding function is lowest). Convergence was achieved in 20 iterations. Shown are the results of iterations #1, #2, #10, #16, and #20. (Continued in Figure 5.14.)

function is shown as a red dashed curve if its minimum is the smallest, thus implying that it is its interval that will be split for the next iteration step. We thus start off with a red dashed lower bounding function, as is shown in Figure 5.11 at iteration #1.

With the initial interval of iteration #1 split in the middle, we obtain in iteration #2, two lower bounding functions. As the red curve has the smallest minimum, we continue with the left interval and split it in the middle, thus again giving us two lower bounding functions in iteration #3. This time it is the right interval that has the lower bounding function with smallest minimum. Splitting this interval in the middle and continuing in this way with the splitting process, we reach at the second last iteration #6 an interval that after splitting has in its left half the lower bounding

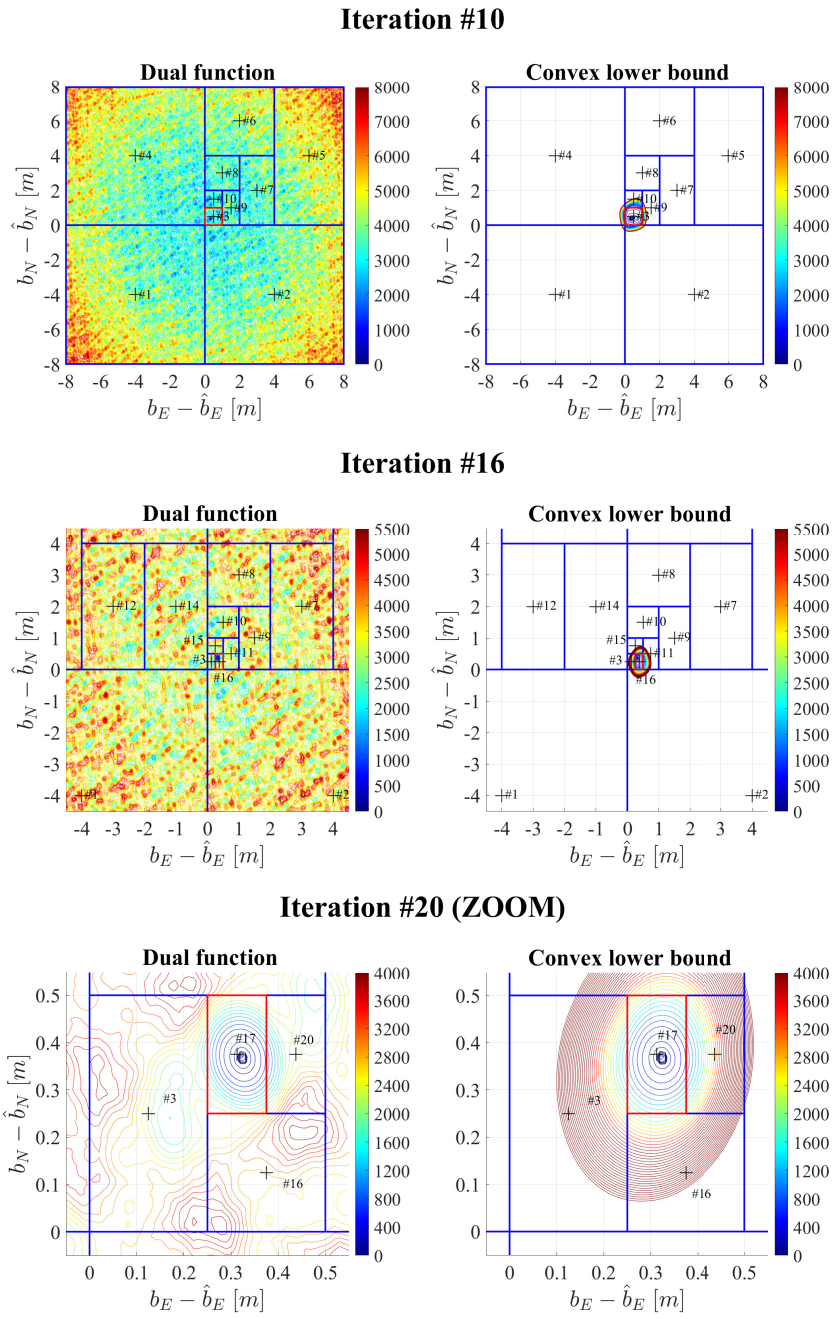


Figure 5.14: Continued from Figure 5.13.

function with smallest minimum. With this lower bounding function at iteration #7, the iteration stops, as then the fulfillment of our stop criterium, set at $\epsilon = 10^{-6}$, is realized. The obtained minimizer is then outputted to provide the asked for solution. In the current example, the iteration gives $\check{b} - \hat{b} = -0.7019$ m, which combined with the float increment $\hat{b} = 0.7042$ m, gives for the ambiguity-fixed increment $\check{b} = 0.0023$ m, thus showing that the fixed-solution differs by 2.3 mm from the correct value. □

Example 4 (Vertical position known) This example continues with the same model as used in the previous example, be it that now the vertical positions are assumed known and thus $b \in \mathbb{R}^2$. The results of the various iteration steps are shown in Figure 5.13, with iterations #1, #2, #10 and #16 columnwise in the first two rows, while the last iteration #20 is shown in the third row as a greater zoom-in. For each iteration, the function values of the dual function are shown in color coded form, as well, next to it, the contour lines of the convex lower bounding function for the box that is to be split. At each iteration step, its iteration number is located close to the centre of gravity of the red box, i.e. the box that will be split. The old iteration number is then moved to the centre of the box that remains unsplit. Iteration results #1, #2 and #10 are shown to the same scale, but for iteration #16 we use a zoom-in to show a greater detail and this is again done for the last iteration #20. Note, that even with this greater zoom-in, still a dense variability of the dual function is apparant, thus showing that it still has many local minima at this greater scale. In the red box containing the global minimum however, we now have a less pronounced multi-modality. At iteration #20 the stop criterium $\epsilon = 10^{-6}$ is fulfilled and as before the fixed solution is obtained at the mm-level. □

5.8 Constrained and partitioned dual problems

In this contribution we restricted our study of the dual problem to the mixed-integer model (5.1). It is possible however to generalize the developed methodology also to the constrained case $b \in \mathcal{C} \subset \mathbb{R}^p$ or to the case when only part of $b \in \mathbb{R}^p$ needs to be ambiguity-resolved. For these two cases we provide the following outlook.

Constrained Dual

For the constrained mixed-integer model

$$\mathsf{E}(y) = Aa + Bb \ , \ a \in \mathbb{Z}^n \ , \ b \in \mathcal{C} \subset \mathbb{R}^p \tag{5.93}$$

the constrained primal approach would be to first solve for the integer ambiguity a and then the baseline b. As we have, using $F(a, b)$ of (5.6),

$$\check{a} = \arg\min_{a \in \mathbb{Z}^n} F(a, \check{b}(a)) \ , \ \check{b}(a) = \arg\min_{b \in \mathcal{C}} F(a, b) \tag{5.94}$$

the primal function $F(a, \check{b}(a))$ would be used to set up a (nonellipsoidal) search space to determine $\check{a}$, which then on its turn is used to determine the ambiguity-resolved

constrained baseline as $\check{b} = \check{b}(\check{a})$. This is the approach used in the LAMBDA-based constrained solutions of GNSS attitude determination (Teunissen, 2010; Giorgi et al., 2010) and rank-defect bias estimation (Khodabandeh, 2022; Teunissen, 2006).

The dual formulation to (5.94) is

$$\check{b} = \arg\min_{b\in\mathcal{C}} F(\check{a}(b), b)\ ,\ \check{a}(b) = \arg\min_{a\in\mathbb{Z}^n} F(a, b) \tag{5.95}$$

In this case, first the constrained baseline is solved as $\check{b}$, from which the integer ambiguity follows as $\check{a} = \check{a}(\check{b})$. Earlier we demonstrated in Section 5.3.3, that for the typical unconstrained mixed-integer GNSS model, one can not expect the success-rate of $\check{a}(b)$ to be large. This may change however when constraints on b are included. If $\mathcal{C}$ is such that the variability in $\check{a}(b)$ is small in relation tot the size of the pull-in regions, then only a few integer candidates $\check{a}(b)$ may need to be considered, thus simplifying the computation of $\check{b}$ in (5.95) considerably. A typical example from the field of GNSS attitude determination is one where the distance between the antennae is sufficiently small in relation to the used signal wavelength.

In (5.95) we recognize the dual function as $\mathcal{D}(b) = F(\check{a}(b), b)$, thus giving

$$\check{b} = \arg\min_{b\in\mathcal{C}} \mathcal{D}(b) \tag{5.96}$$

This shows that if $\mathcal{C}$ would be easily and efficiently covered by a box $\mathcal{B}$, that our method of the previous section is directly applicable to the constrained mixed-integer model (5.93) as well. In the more general case, one can use the indicator function of $\mathcal{C}$, $p_{\mathcal{C}}(b) = \infty$ if $b \notin \mathcal{C}$, and $p_{\mathcal{C}}(b) = 0$ otherwise, to replace the constrained formulation (5.96) by the unconstrained formulation $\check{b} = \arg\min_{b\in\mathbb{R}^n} [\mathcal{D}(b) + p_{\mathcal{C}}(b)]$. Similar to our development in the previous section, the approach would then be to construct convex lower bounding functions using a continuous barrier function for $p_{\mathcal{C}}(b)$ (Nocedal and Wright, 2006; Nesterov, 2018).

Partitioned Dual

Earlier we mentioned that a potentially attractive feature of the dual formulation is that the minimization of the objective function takes place in $\mathbb{R}^p$ instead of $\mathbb{R}^n$. This can be attractive if p is much smaller than n and the dual function $\mathcal{D}(b)$ is easily formed or acceptably approximated. This changes however if the dimension of $b \in \mathbb{R}^p$ is large as well.

Still, even with p large, it could be that one is only interested in the ambiguity-resolved solution of some of the components of b, say b_1 of $b = [b_1^T, b_2^T]^T \in \mathbb{R}^{p_1+p_2}$. In that case one can combine the primal- and dual approach to obtain the decomposition

$$F(a, b_1, b_2) = ||\hat{b}_1 - b_1||^2_{Q_{\hat{b}_1\hat{b}_1}} + ||\hat{a}(b_1) - a||^2_{Q_{\hat{a}(b_1)\hat{a}(b_1)}} + ||\hat{b}_2(a, b_1) - b_2||^2_{Q_{\hat{b}_2(a,b_1)\hat{b}_2(a,b_1)}}$$

By now defining the partitioned dual $\mathcal{PD}(b_1)$ as

$$\mathcal{PD}(b_1) = ||\hat{b}_1 - b_1||^2_{Q_{\hat{b}_1\hat{b}_1}} + ||\hat{a}(b_1) - \check{a}(b_1)||^2_{Q_{\hat{a}(b_1)\hat{a}(b_1)}} \tag{5.97}$$

with $\breve{a}(b_1) = \arg\min_{a\in\mathbb{Z}^n} ||\hat{a}(b_1) - a||^2_{Q_{\hat{a}(b_1)\hat{a}(b_1)}}$, the solution for a and b becomes

$$\begin{cases} \breve{b}_1 &= \arg\min\limits_{b_1\in\mathbb{R}^{p_1}} \mathcal{PD}(b_1) \\ \breve{a} &= \breve{a}(\breve{b}_1) \\ \breve{b}_2 &= \hat{b}_2(\breve{a}, \breve{b}_1) \end{cases} \tag{5.98}$$

In this way one can still apply the dual approach to solve for b_1 in a lower dimensioned space. However, the price one pays for this reduction in dimension is that $\hat{a}(b_1)$ will be less precise than $\hat{a}(b)$, implying that, in dependence of the application, $\mathcal{PD}(b_1)$ may be more difficult to construct or to approximate.

5.9 Summary and conclusions

In this contribution we introduced and studied the dual mixed-integer least-squares formulation. The dual differs from its primal counterpart in the order in which the integer ambiguity vector $a \in \mathbb{Z}^n$ and baseline vector $b \in \mathbb{R}^p$ are estimated. In the primal formulation, the integer ambiguity estimator $\breve{a}$ is computed first, followed by the computation of the ambiguity-resolved baseline estimator $\breve{b} = \hat{b}(\breve{a})$. In the dual formulation, the order is reversed, giving first $\breve{b}$, followed by the ambiguity estimator $\breve{a} = \breve{a}(\breve{b})$. As not the ambiguities, but rather the entries of b are usually the parameters of interest, the dual formulation seems a natural way to go. And this seems even more so for applications where the baseline dimension p is considerable less than the ambiguity dimension n.

We show however that this optimism must somewhat be tempered, due to the fact that the formation of the dual objective function $\mathcal{D}(b)$ also requires the solution of an n-dimensional integer least-squares (ILS) problem, and even one that depends on the unknown baseline b. The potential advantage of the dual formulation lies therefore in the ease with which this implicit ILS problem can be solved. This problem is trivially solved in case the conditional ambiguity variance matrix is diagonal, which in case of GNSS happens for instance with the geometry-free model.

To make the dual formulation more computationally attractive, we introduced two approximations to it, both with the purpose of alleviating the computational demand on the implicit ILS problem. The first approximation is based on replacing the variance matrix of $\hat{a}(b)$ by a matrix of simpler structure (preferably diagonal), the second by replacing the integer estimator $\breve{a}(b)$ by a simpler estimator (preferably integer rounding).

To be able to describe the quality of their dual solutions, we provided a complete distributional analysis of the ambiguity- and baseline-estimators that follow from these approximations. This will allow users to rigorously judge whether or not the approximations made in the dual formulation are acceptable for their application. We hereby note that the second approximation of the dual formulation also resulted in the finding of a new class of admissible integer estimators, the pull-in regions of which were described as well.

To characterize the challenges that the computation of $\check{b}$, as a global minimizer of $\mathcal{D}(b)$, poses, we illustrated and emphasized the multimodality of the dual function $\mathcal{D}(b)$ with its many local minima. As a consequence, the dual function lacks convexity, thus requiring special care in developing an algorithm that is globally convergent.

Our proposed method, which has finite termination with a guaranteed ϵ-tolerance, is constructed from combining the branch-and-bound principle, with a special convex-relaxation of the dual, to which the projected-gradient-descent method is applied to obtain the required bounds. Each of the method's three constituents are described, whereby special emphasis is given to the construction of the required continuously differentiable, convex lower bounding function of the dual. Illustrative examples are given to provide insight into the workings of the method, while in future work its numerical performance for advanced models will be evaluated. Finally, we provided an outlook on solving a constrained and partitioned dual.

5

5.10 Appendix

Proof of Lemma 3: The equality of (5.23) follows from substituting $-Q_{\hat{a}\hat{b}}Q_{\hat{b}\hat{b}}^{-1} = (A^TQ_{yy}^{-1}A)^{-1}A^TQ_{yy}^{-1}B$ and $Q_{\hat{a}(b)\hat{a}(b)} = (A^TQ_{yy}^{-1}A)^{-1}$ in $||\Delta a||^2_{Q_{\hat{a}(b)\hat{a}(b)}}$. The inequality of (5.23) follows from $Q_{\hat{b}(a)\hat{b}(a)} = (B^TQ_{yy}^{-1}B)^{-1}$ and recognizing that the norm of a projected vector is never larger than the norm of the vector itself. □

Proof of Lemma 4: The quadratic identity follows as

$$\begin{aligned}
&||\hat{b}-b||^2_{Q_{\hat{b}\hat{b}}} + ||\hat{a}(b)-a||^2_{Q^\circ_{\hat{a}(b)\hat{a}(b)}} \\
&= \begin{bmatrix} \hat{a}(b)-a \\ \hat{b}-b \end{bmatrix}^T \begin{bmatrix} Q^\circ_{\hat{a}(b)\hat{a}(b)} & 0 \\ 0 & Q_{\hat{b}\hat{b}} \end{bmatrix}^{-1} \begin{bmatrix} \hat{a}(b)-a \\ \hat{b}-b \end{bmatrix} \\
&= \begin{bmatrix} \hat{a}-a \\ \hat{b}-b \end{bmatrix}^T \begin{bmatrix} Q^\circ_{\hat{a}(b)\hat{a}(b)} + Q_{\hat{a}\hat{b}}Q_{\hat{b}\hat{b}}^{-1}Q_{\hat{b}\hat{a}} & Q_{\hat{a}\hat{b}} \\ Q_{\hat{b}\hat{a}} & Q_{\hat{b}\hat{b}} \end{bmatrix}^{-1} \begin{bmatrix} \hat{a}-a \\ \hat{b}-b \end{bmatrix} \\
&= \begin{bmatrix} \hat{a}-a \\ \hat{b}^\circ(a)-b \end{bmatrix}^T \begin{bmatrix} Q^\circ_{\hat{a}\hat{a}} & 0 \\ 0 & Q_{\hat{b}\hat{b}} - Q_{\hat{b}\hat{a}}Q^{\circ-1}_{\hat{a}\hat{a}}Q_{\hat{a}\hat{b}} \end{bmatrix}^{-1} \begin{bmatrix} \hat{a}-a \\ \hat{b}^\circ(a)-b \end{bmatrix} \\
&= ||\hat{a}-a||^2_{Q^\circ_{\hat{a}\hat{a}}} + ||\hat{b}^\circ(a)-b||^2_{Q^\circ_{\hat{b}(a)\hat{b}(a)}}
\end{aligned} \tag{5.99}$$

□

Proof of Theorem 3: Using the quadratic identity (5.27), we may write $\mathcal{D}^\circ(b)$ as

$$\begin{aligned}
\mathcal{D}^\circ(b) &= ||\hat{b}-b||^2_{Q_{\hat{b}\hat{b}}} + \min_{a\in\mathbb{Z}^n} ||\hat{a}(b)-a||^2_{Q^\circ_{\hat{a}(b)\hat{a}(b)}} \\
&= \min_{a\in\mathbb{Z}^n} \left(||\hat{b}-b||^2_{Q_{\hat{b}\hat{b}}} + ||\hat{a}(b)-a||^2_{Q^\circ_{\hat{a}(b)\hat{a}(b)}} \right) \\
&= \min_{a\in\mathbb{Z}^n} \left(||\hat{a}-a||^2_{Q^\circ_{\hat{a}\hat{a}}} + ||\hat{b}^\circ(a)-b||^2_{Q^\circ_{\hat{b}(a)\hat{b}(a)}} \right)
\end{aligned} \tag{5.100}$$

from which it follows that $\check{b}^\circ = \arg\min_{b\in\mathbb{R}^p} \mathcal{D}^\circ(b)$ can be computed through the primal formulation as $\check{b}^\circ = \hat{b}^\circ(\check{a}^\circ)$, with $\check{a}^\circ = \arg\min_{a\in\mathbb{Z}^n} ||\hat{a}-a||^2_{Q^\circ_{\hat{a}\hat{a}}}$. What remains to be shown is that $\check{a}^\circ = \check{a}^\circ(\check{b}^\circ)$. Starting from the definition $\check{a}^\circ(b) = \arg\min_{a\in\mathbb{Z}^n} ||\hat{a}(b)-a||^2_{Q^\circ_{\hat{a}(b)\hat{a}(b)}}$, we may write with the aid of the quadratic identity (5.27),

$$\check{a}^\circ(b) = \arg\min_{a\in\mathbb{Z}^n} \left(||\hat{a}-a||^2_{Q^\circ_{\hat{a}\hat{a}}} + ||\hat{b}^\circ(a)-b||^2_{Q^\circ_{\hat{b}(a)\hat{b}(a)}} \right) \tag{5.101}$$

from which it follows, since $\check{b}^\circ = \hat{b}^\circ(\check{a}^\circ)$, that $\check{a}^\circ(\check{b}^\circ) = \arg\min_{a\in\mathbb{Z}^n} ||\hat{a}-a||^2_{Q^\circ_{\hat{a}\hat{a}}} = \check{a}^\circ$. □

Proof of Theorem 4: As $Q^\circ_{\hat{a}\hat{a}}$ is not the variance matrix of $\hat{a}$, inequality (5.31a) follows directly from Theorem 1 of (Teunissen, 2000b). From substituting the given matrices $R_{\hat{a}\hat{b}}$, $S_{\hat{a}\hat{a}}$ and $T_{\hat{a}\hat{b}}$, we obtain

$$R^T_{\hat{a}\hat{b}} R_{\hat{a}\hat{b}} = Q_{\hat{b}\hat{a}}[Q^{\circ-1}_{\hat{a}\hat{a}} Q_{\hat{a}\hat{a}} Q^{\circ-1}_{\hat{a}\hat{a}} - 2Q^{\circ-1}_{\hat{a}\hat{a}} + Q^{-1}_{\hat{a}\hat{a}}]Q_{\hat{a}\hat{b}} \tag{5.102}$$

which proves the equality of (5.32b). As (5.102) is positive semi-definite, the inequality of (5.32b) follows. □

Proof of Theorem 5 (Success-rate lower-bounds)

(*i*) The proof of (5.34) follows from combining an inequality-theorem of (Anderson, 1996) with the ILS-theorem of (Teunissen, 1999a). Let $\hat{a} \sim \mathcal{N}_n(a, Q_{\hat{a}\hat{a}})$, $\hat{a}^\circ \sim \mathcal{N}_n(a, Q^\circ_{\hat{a}\hat{a}})$ and $Q^\circ_{\hat{a}\hat{a}} \geq Q_{\hat{a}\hat{a}}$. Then it follows from Anderson's theorem that $\mathsf{P}[\hat{a}^\circ \in \mathscr{P}^\circ_a] \leq \mathsf{P}[\hat{a} \in \mathscr{P}^\circ_a]$, as the subset $\mathscr{P}^\circ_a \subset \mathbb{R}^n$ is convex and symmetric about a. As $\mathsf{P}[\hat{a}^\circ \in \mathscr{P}^\circ_a]$ can now be interpreted being the success-rate of a properly weighted ILS-estimator, it follows from Teunissen's optimality-theorem that any other integer estimator using $\hat{a}^\circ$ as input, will have a smaller success-rate. This therefore also holds true for integer bootstrapping, of which the success-rate is given by the left-hand side of (5.34), see (Teunissen, 1998b).

(*ii*) To prove the given lower-bound (5.35) of

$$\mathsf{P}[\check{a}^\circ = a] = \int_{\mathscr{P}^\circ_a} f_{\hat{a}}(\alpha) d\alpha \tag{5.103}$$

we will work with two ellipsoids, one being a subset of the other, while both are subsets of the pull-in region $\mathscr{P}^\circ_a$. First we describe the pull-in region in a way that facilitates the comparison with the planes-of-support formulation of ellipsoids:

$$\begin{aligned} \mathscr{P}^\circ_a &= a + \mathscr{P}^\circ_0 \\ &= a + \{x\in\mathbb{R}^n |\ x^T Q^{\circ-1}_{\hat{a}\hat{a}} x \leq ||x-z||^2_{Q^\circ_{\hat{a}\hat{a}}}, \forall z\in\mathbb{Z}^n\} \\ &= a + \{x\in\mathbb{R}^n |\ (z^T Q^{\circ-1}_{\hat{a}\hat{a}} x)^2 \leq \tfrac{1}{4}||z||^4_{Q^\circ_{\hat{a}\hat{a}}}, \forall z\in\mathbb{Z}^n\} \end{aligned} \tag{5.104}$$

Now we construct the first ellipsoid residing in $\mathscr{P}_0^\circ$. We have

$$\begin{array}{rcl} E_0^\circ & = & \{x \in \mathbb{R}^n |\ x^T Q_{\hat{a}\hat{a}}^{\circ -1} x \leq \chi^2\} \\ & \stackrel{a}{=} & \{x \in \mathbb{R}^n |\ (g^T x)^2 \leq \chi^2 g^T Q_{\hat{a}\hat{a}}^\circ g, \forall g \in \mathbb{R}^n\} \\ & \stackrel{b}{=} & \{x \in \mathbb{R}^n |\ (f^T Q_{\hat{a}\hat{a}}^{\circ -1} x)^2 \leq \chi^2 f^T Q_{\hat{a}\hat{a}}^{\circ -1} f, \forall f \in \mathbb{R}^n\} \\ & \stackrel{c}{\subset} & \{x \in \mathbb{R}^n |\ (z^T Q_{\hat{a}\hat{a}}^{\circ -1} x)^2 \leq \chi^2 ||z||^2_{Q^\circ_{\hat{a}\hat{a}}}, \forall z \in \mathbb{Z}^n\} \\ & \stackrel{d}{\subset} & \{x \in \mathbb{R}^n |\ (z^T Q_{\hat{a}\hat{a}}^{\circ -1} x)^2 \leq \chi^2 \frac{||z||^4_{Q^\circ_{\hat{a}\hat{a}}}}{\min\limits_{z \in \mathbb{Z}^n/\{0\}} ||z||^2_{Q^\circ_{\hat{a}\hat{a}}}}, \forall z \in \mathbb{Z}^n\} \end{array} \tag{5.105}$$

In (a) we used the planes-of-support formulation of the ellipsoid (Teunissen, 2001), while (b) follows from using the one-to-one transformation $g = Q_{\hat{a}\hat{a}}^{\circ -1} f$. With (c) we obtain a larger set as now only the integers are considered, while with (d) again a larger subset is created due to the larger right-hand side.

It follows from comparing (5.105) with (5.104) that

$$E_a^\circ = a + E_0^\circ \subset \mathscr{P}_a^\circ \ \text{ if } \ \chi^2 = \tfrac{1}{4} \min_{z \in \mathbb{Z}^n/\{0\}} ||z||^2_{Q^\circ_{\hat{a}\hat{a}}} \tag{5.106}$$

We now construct our second ellipsoid $E_a \subset E_a^\circ$ such that the probability $\mathsf{P}[\hat{a} \in E_a]$ is easy to compute. The ellipsoid is chosen as

$$\begin{array}{rcl} E_a & = & \{x \in \mathbb{R}^n |\ (x-a)^T Q_{\hat{a}\hat{a}}^{-1} (x-a) \leq r^2\} \\ & = & \{x = a + Q_{\hat{a}\hat{a}}^{1/2} v, v \in \mathbb{R}^n |\ v^T v \leq r^2\} \end{array} \tag{5.107}$$

If we use the same transformation also for E_a°, we may write

$$E_a^\circ = \{x = a + Q_{\hat{a}\hat{a}}^{1/2} v, v \in \mathbb{R}^n |\ v^T [Q_{\hat{a}\hat{a}}^{1/2} Q_{\hat{a}\hat{a}}^{\circ -1} Q_{\hat{a}\hat{a}}^{1/2}] v \leq \chi^2\} \tag{5.108}$$

By comparing (5.108) with (5.107) and noting that a sphere best fits inside an ellipsoid if its radius is equal to the ellipsoid's minor axis, it follows that

$$E_a \subset E_a^\circ \ \text{ if } \ r^2 = \lambda_{\min} \times \chi^2 \tag{5.109}$$

where $\lambda_{\min}$ is the smallest eigenvalue of $[Q_{\hat{a}\hat{a}}^{1/2} Q_{\hat{a}\hat{a}}^{\circ -1} Q_{\hat{a}\hat{a}}^{1/2}]^{-1}$ and thus the minimum of the Raileigh quotient

$$\lambda_{\min} = \min_{x \in \mathbb{R}^n} \frac{x^T Q_{\hat{a}\hat{a}}^\circ x}{x^T Q_{\hat{a}\hat{a}} x} \tag{5.110}$$

Since $E_a \subset E_a^\circ \subset \mathscr{P}_a^\circ$ (cf. 5.106 and 5.109), it follows that

$$\mathsf{P}[\check{a}^\circ = a] = \mathsf{P}[\hat{a} \in \mathscr{P}_a^\circ] \geq \mathsf{P}[\hat{a} \in E_a] = \mathsf{P}[\chi^2(0,n) \leq r^2]$$

which concludes the proof. □

Proof of Theorem 6: Recall that any admissible integer estimator $\mathcal{I}^\bullet(x)$ can be written as (Teunissen, 2000b),

$$\mathcal{I}^\bullet(x) = \sum_{z \in \mathbb{Z}^n} z \iota_z^\bullet(x) \tag{5.111}$$

where $\iota_z^\bullet(x)$ is the indicator function of its pull-in region $I_z^\bullet = \{x \in \mathbb{R}^n |\ z = \mathcal{I}^\bullet(x)\}$. With (5.111), we can rewrite the dual function as

$$\begin{aligned}
\mathcal{D}^\bullet(b) &= ||\hat{b} - b||^2_{Q_{\hat{b}\hat{b}}} + ||\hat{a}(b) - \mathcal{I}^\bullet(\hat{a}(b))||^2_{Q_{\hat{a}(b)\hat{a}(b)}} \\
&= ||\hat{b} - b||^2_{Q_{\hat{b}\hat{b}}} + \sum_{z\in\mathbb{Z}^n} ||\hat{a}(b) - z||^2_{Q_{\hat{a}(b)\hat{a}(b)}} \iota_z^\bullet(\hat{a}(b)) \\
&= \sum_{z\in\mathbb{Z}^n} \left(\hat{b} - b||^2_{Q_{\hat{b}\hat{b}}} + ||\hat{a}(b) - z||^2_{Q_{\hat{a}(b)\hat{a}(b)}} \right) \iota_z^\bullet(\hat{a}(b)) \\
&= \sum_{z\in\mathbb{Z}^n} \left(||\hat{a} - z||^2_{Q_{\hat{a}\hat{a}}} + ||\hat{b}(z) - b||^2_{Q_{\hat{b}(a)\hat{b}(a)}} \right) \iota_z^\bullet(\hat{a}(b))
\end{aligned}$$

This shows that

$$\mathcal{D}^\bullet(b) = ||\hat{a} - u||^2_{Q_{\hat{a}\hat{a}}} + ||\hat{b}(u) - b||^2_{Q_{\hat{b}(a)\hat{b}(a)}} \ \text{ if } \hat{a}(b) \in I_u^\bullet$$

Hence, if $\hat{a}(b) \in I_u^\bullet$ for some $u \in \mathbb{Z}^n$, then the minimizer and minimum of $\mathcal{D}^\bullet(b)$ are given as

$$\hat{b}(u) = \arg \min_{\iota_u^\bullet(\hat{a}(b))=1} \mathcal{D}^\bullet(b),\ \mathcal{D}^\bullet(\hat{b}(u)) = ||\hat{a} - u||^2_{Q_{\hat{a}\hat{a}}} \tag{5.112}$$

To find the *global* minimum of $\mathcal{D}^\bullet(b)$, we need to find the smallest value of $\mathcal{D}^\bullet(\hat{b}(u))$ for all $u \in \mathbb{Z}^n$ that satisfy $\hat{a}(b) \in I_u^\bullet$, $\forall b \in \mathbb{R}^p$. This integer set consists of the integer grid points of all pull-in regions $I_z^\bullet$ through which $\hat{a}(b)$ passes when b is varied,

$$\Omega_{\hat{a}}^\bullet = \{z \in \mathbb{Z}^n |\ z = \mathcal{I}^\bullet(\hat{a}(b)), \forall b \in \mathbb{R}^p\} \subset \mathbb{Z}^n \tag{5.113}$$

Hence, the global minimum of $\mathcal{D}^\bullet(b)$ is given as

$$\min_{b\in\mathbb{R}^p} \mathcal{D}^\bullet(b) = \min_{z\in\Omega_{\hat{a}}^\bullet} \mathcal{D}^\bullet(\hat{b}(z)) = \min_{z\in\Omega_{\hat{a}}^\bullet} ||\hat{a} - z||^2_{Q_{\hat{a}\hat{a}}} \tag{5.114}$$

and its corresponding minimizer as

$$\begin{aligned}
\check{b}^\bullet &= \arg \min_{b\in\mathbb{R}^p} \mathcal{D}^\bullet(b) = \hat{b}(\check{a}^\bullet), \text{with} \\
\check{a}^\bullet &= \arg \min_{z\in\Omega_{\hat{a}}^\bullet} ||\hat{a} - z||^2_{Q_{\hat{a}\hat{a}}}
\end{aligned} \tag{5.115}$$

□

Proof of Lemma 6: The first two conditions, (i) and (ii), for an integer estimator to be admissible, are not difficult to verify. The pull-in regions cover the whole space as $\arg \min_{u\in\Omega_x^\bullet} ||x - u||^2_{Q_{\hat{a}\hat{a}}}$ maps any $x \in \mathbb{R}^n$ to an integer vector. Furthermore, any $x \in \mathbb{R}^n$ that lies inside a pull-in region will be mapped uniquely to one integer vector. To verify the third condition, (iii), we need to show that $\mathscr{P}_z^\bullet = \mathscr{P}_0^\bullet + z$. We will prove this in steps. First we show that $\bar{\Omega}_z^\bullet = \bar{\Omega}_0^\bullet + z$ and $\Phi_z^\bullet = \Phi_0^\bullet + z$ (similarly one can show that $\Omega_x^\bullet + z = \Omega_{x+z}^\bullet$). We have

$$\begin{aligned}
\bar{\Omega}_0^\bullet + z &= \{x + z \in \mathbb{R}^n |\ 0 = \mathcal{I}^\bullet(x + M\beta), \exists \beta \in \mathbb{R}^p\} \\
&= \{y \in \mathbb{R}^n |\ 0 = \mathcal{I}^\bullet(y - z + M\beta), \exists \beta \in \mathbb{R}^p\} \\
&= \{y \in \mathbb{R}^n |\ z = \mathcal{I}^\bullet(y + M\beta), \exists \beta \in \mathbb{R}^p\} \\
&= \bar{\Omega}_z^\bullet,\ \forall z \in \mathbb{Z}^n
\end{aligned} \tag{5.116}$$

where use was made of the property $\mathcal{I}^\bullet(x+z) = \mathcal{I}^\bullet(x) + z$. Similarly, we have

$$\begin{array}{rcl} \Phi_0^\bullet + z & = & \{u+z \in \mathbb{Z}^n |\ u = \mathcal{I}^\bullet(x), \forall x \in \bar{\Omega}_0^\bullet\} \\ & = & \{v \in \mathbb{Z}^n |\ v = \mathcal{I}^\bullet(x) + z, \forall x \in \bar{\Omega}_0^\bullet\} \\ & = & \{v \in \mathbb{Z}^n |\ v = \mathcal{I}^\bullet(x+z), \forall x \in \bar{\Omega}_0^\bullet\} \\ & = & \{v \in \mathbb{Z}^n |\ v = \mathcal{I}^\bullet(y), \forall y \in \bar{\Omega}_0^\bullet + z\} \\ & = & \{v \in \mathbb{Z}^n |\ v = \mathcal{I}^\bullet(y), \forall y \in \bar{\Omega}_z^\bullet\} \\ & = & \Phi_z^\bullet,\ \forall z \in \mathbb{Z}^n \end{array} \tag{5.117}$$

where use was made of $\bar{\Omega}_z^\bullet = \bar{\Omega}_0^\bullet + z$. Finally, we have

$$\begin{array}{rcl} \mathscr{P}_0^\bullet + z & = & \{x+z \in \mathbb{R}^n |\ ||x||^2_{Q_{\hat{a}\hat{a}}} \leq ||x-u||^2_{Q_{\hat{a}\hat{a}}}, \forall u \in \Phi_0\} \\ & = & \{y \in \mathbb{R}^n |\ ||y-z||^2_{Q_{\hat{a}\hat{a}}} \leq ||y-(u+z)||^2_{Q_{\hat{a}\hat{a}}}, \forall u \in \Phi_0\} \\ & = & \{y \in \mathbb{R}^n |\ ||y-z||^2_{Q_{\hat{a}\hat{a}}} \leq ||y-v||^2_{Q_{\hat{a}\hat{a}}}, \forall v \in \Phi_0 + z\} \\ & = & \{y \in \mathbb{R}^n |\ ||y-z||^2_{Q_{\hat{a}\hat{a}}} \leq ||y-v||^2_{Q_{\hat{a}\hat{a}}}, \forall v \in \Phi_z\} \\ & = & \mathscr{P}_z^\bullet,\ \forall z \in \mathbb{Z}^n \end{array} \tag{5.118}$$

where use was made of $\Phi_z^\bullet = \Phi_0^\bullet + z$. □

Proof of Theorem 7: As $\mathsf{P}[\breve{a}^\bullet = a] = \mathsf{P}[\hat{a} \in \mathscr{P}_a^\bullet] = \int_{\mathscr{P}_a^\bullet} f_{\hat{a}}(\alpha) d\alpha$, the proof will be based on finding an ellipsoidal region $E_a = \{x \in \mathbb{R}^n |\ (x-a)^T Q_{\hat{a}\hat{a}}^{-1}(x-a) \leq r^2\}$ that resides in $\mathscr{P}_a^\bullet$. Once such region is found, we have $E_a \subset \mathscr{P}_a^\bullet$, and thus the lower-bound $\mathsf{P}[\hat{a} \in E_a] \leq \mathsf{P}[\hat{a} \in \mathscr{P}_a^\bullet]$, with $\mathsf{P}[\hat{a} \in E_a] = \mathsf{P}[\chi^2(0,n) \leq r^2]$. The challenge is therefore to find the proper value for r^2.

First we express the pull-in region $\mathscr{P}_a^\bullet = a + \mathscr{P}_0^\bullet$ in a more amendable form. We have

$$\begin{array}{rcl} \mathscr{P}_0^\bullet & = & \{x \in \mathbb{R}^n |\ ||x||^2_{Q_{\hat{a}\hat{a}}} \leq ||x - \mathcal{I}^\bullet(y)||^2_{Q_{\hat{a}\hat{a}}}, \forall y \in \bar{\Omega}_0^\bullet\} \\ & = & \{x \in \mathbb{R}^n |\ \mathcal{I}^\bullet(y)^T Q_{\hat{a}\hat{a}}^{-1} x \leq \tfrac{1}{2}||\mathcal{I}^\bullet(y)||^2_{Q_{\hat{a}\hat{a}}}, \forall y \in \bar{\Omega}_0^\bullet\} \end{array} \tag{5.119}$$

We now show how r^2 can be chosen such that $E_0 \subset \mathscr{P}_0^\bullet$. We have, with $q^2 = r^2/(\min_{z \in \mathbb{Z}^n/\{0\}} ||z||^2_{Q_{\hat{a}\hat{a}}})$,

$$\begin{array}{rcl} E_0 & & \{x \in \mathbb{R}^n |\ x^T Q_{\hat{a}\hat{a}}^{-1} x \leq r^2\} \\ & \overset{(i)}{=} & \{x \in \mathbb{R}^n |\ [g^T x]^2 \leq r^2 g^T Q_{\hat{a}\hat{a}} g, \forall g \in \mathbb{R}^n\} \\ & \overset{(ii)}{=} & \{x \in \mathbb{R}^n |\ [f^T Q_{\hat{a}\hat{a}}^{-1} x]^2 \leq r^2 f^T Q_{\hat{a}\hat{a}}^{-1} f, \forall f \in \mathbb{R}^n\} \\ & \overset{(iii)}{\subset} & \{x \in \mathbb{R}^n |\ [z^T Q_{\hat{a}\hat{a}}^{-1} x]^2 \leq r^2 ||z||^2_{Q_{\hat{a}\hat{a}}}, \forall z \in \mathbb{Z}^n\} \\ & \overset{(iv)}{\subset} & \{x \in \mathbb{R}^n |\ [z^T Q_{\hat{a}\hat{a}}^{-1} x]^2 \leq q^2 ||z||^4_{Q_{\hat{a}\hat{a}}}, \forall z \in \mathbb{Z}^n\} \\ & \overset{(v)}{\subset} & \{x \in \mathbb{R}^n |\ [\mathcal{I}^\bullet(y)^T Q_{\hat{a}\hat{a}}^{-1} x]^2 \leq q^2 ||\mathcal{I}^\bullet(y)||^4_{Q_{\hat{a}\hat{a}}}, \forall y \in \bar{\Omega}_0^\bullet\} \\ & \overset{(vi)}{\subset} & \{x \in \mathbb{R}^n |\ \mathcal{I}^\bullet(y)^T Q_{\hat{a}\hat{a}}^{-1} x \leq \left[q^2 ||\mathcal{I}^\bullet(y)||^4_{Q_{\hat{a}\hat{a}}}\right]^{\frac{1}{2}}, \forall y \in \bar{\Omega}_0^\bullet\} \\ & \overset{(5.119)}{=} & \mathscr{P}_0^\bullet \ \text{ if } \ r^2 = \tfrac{1}{4} \min_{z \in \mathbb{Z}^n/\{0\}} ||z||^2_{Q_{\hat{a}\hat{a}}} \end{array}$$

In (i) we used the *planes-of-support* formulation of the ellipsoid (Teunissen, 1995), while (ii) follows from using the one-to-one transformation $g = Q_{\hat{a}\hat{a}}^{-1} f$. With (iii) we

obtain a larger set as now only the integers are considered, while with both (iv) and (v) again larger subsets are created, once due to the larger right-hand side and once due to considering only a subset of integers rather than the whole integer space $\mathbb{Z}^n$. Finally, with (vi) we replaced the two-sided inequality by a one-sided one. Comparison with (5.119) shows that $\mathscr{P}_0^\bullet$ is obtained as stated. An example of the lower bounding ellipse E_0 is shown in Figure 5.4 for $n = 2$ and $p = 1$. □

Projection Lemma (projection of cube and rectangle):

(a) Largest and smallest projection of cube on a line:

$$\begin{aligned} \max_{-e \preceq x \preceq e} a^T x &= +\sum_{\alpha=1}^{p} |a_\alpha| \\ \min_{-e \preceq x \preceq e} a^T x &= -\sum_{\alpha=1}^{p} |a_\alpha| \end{aligned} \tag{5.120}$$

with $a = (a_1, \ldots, a_p)^T$ and $e = (1, \ldots, 1)^T$.

(b) Largest and smallest projection of rectangle on a line:

$$\begin{aligned} \max_{l \preceq x \preceq u} a^T x &= \sum_{\alpha=1}^{p} +\tfrac{1}{2}(u_\alpha - l_\alpha)|a_\alpha| + \tfrac{1}{2}(u_\alpha + l_\alpha)a_\alpha \\ \min_{l \preceq x \preceq u} a^T x &= \sum_{\alpha=1}^{p} -\tfrac{1}{2}(u_\alpha - l_\alpha)|a_\alpha| + \tfrac{1}{2}(u_\alpha + l_\alpha)a_\alpha \end{aligned} \tag{5.121}$$

■

Proof: (a) is easy to prove; (b) follows from (a) through the one-to-one transformation $x = \tilde{U}x' + \bar{x}$, where $\bar{x} = \frac{1}{2}(u + l)$, $\tilde{u} = \frac{1}{2}(u - l)$ and $\tilde{U} = \mathrm{diag}(\tilde{u}_1, \ldots, \tilde{u}_p)$. □

References

Anderson T (1996) Some inequalities for symmetric convex sets with applications. Ann Statist 24(2):753–762

Balakrishnan V, Boyd S, Balemi S (1991) Branch and bound algorithm for computing the minimum stability degree of parameter-dependent linear systems. Journal of Robust and Nonlinear Control 1(4):295–317

Bertsekas D (1999) Nonlinear programming. Athena Scientific, 2nd Ed

Giorgi G, Teunissen PJG, Verhagen S, Buist PJ (2010) Testing a New Multivariate GNSS Carrier Phase Attitude Determination Method for Remote Sensing Platforms. Advances in Space Research 46(2):118–129

Guida A (2015) A branch and bound algorithm for the global optimization and its improvements. PhD thesis, University of Florence pp 1–89

Hofmann-Wellenhof B, Lichtenegger H, Wasle E (2008) GNSS: Global Navigation Satellite Systems: GPS, Glonass, Galileo, and More. Springer, New York

Horst R, Pardalos P, Thoai N (2000) Introduction to Global Optimization, Second Edition. Kluwer Academic Publishers

Khodabandeh A (2022) Bias-bounded estimation of ambiguity: A method for radio interferometric positioning. IEEE Transactions on Signal Processing 70:3042–3057, DOI 10.1109/TSP.2022.3181344

Lawler EL, Wood DE (1966) Branch-And-Bound methods: A survey. Operations Research 14(4):699–719

Leick A, Rapoport L, Tatarnikov D (2015) GPS Satellite Surveying. 4th Edition Wiley

Nesterov Y (2018) Lectures on convex optimization. Springer Optimization and Its Applications 137

Nocedal J, Wright S (2006) Numerical optimization. 2nd ed. Springer-Verlag

Odijk D, Zhang B, Khodabandeh A, Odolinski R, Teunissen PJG (2015) On the estimability of parameters in undifferenced, uncombined GNSS network and PPP-RTK user models by means of S-system theory. J Geod 90(1):15–44

Pardalos P, Romeijn H (2002) Handbook of Global Optimization. Prentice-hall Englewood Cliffs, New Jersey

Parikh N, Boyd S (2013) Proximal algorithms. Foundations and Trends in Optimization 1(3):123–231

Teunissen PJG (1990) Nonlinear least squares. Manuscripta Geodaetica 15(3):137–150

Teunissen PJG (1995) The least-squares ambiguity decorrelation adjustment: a method for fast GPS integer ambiguity estimation. J Geod 70(1-2):65–82

Teunissen PJG (1998a) GPS carrier phase ambiguity fixing concepts. Chapter 8 in GPS For Geodesy, 2nd Ed PJG Teunissen and A Kleusberg (Eds) pp 319–388

Teunissen PJG (1998b) Success probability of integer GPS ambiguity rounding and bootstrapping. J Geod 72(10):606–612

Teunissen PJG (1999a) An optimality property of the integer least-squares estimator. J Geod 73(11):587–593

Teunissen PJG (1999b) The probability distribution of the GPS baseline for a class of integer ambiguity estimators. J Geod 73:275–284

Teunissen PJG (2000a) ADOP based upperbounds for the bootstrapped and least-squares ambiguity success rates. Artificial Satellites 30(4):171–179

Teunissen PJG (2000b) The success rate and precision of GPS ambiguties. J Geod 74:321–326

Teunissen PJG (2001) Integer estimation in the presence of biases. J Geod 75(7):399–407

Teunissen PJG (2002) The parameter distributions of the integer GPS model. J Geod 76(1):41–48

Teunissen PJG (2006) On InSAR ambiguity resolution for deformation monitoring. Artificial Satellites 41(1):19–22

Teunissen PJG (2010) Integer Least Squares Theory for the GNSS Compass. Journal of Geodesy, Springer, DOI 101007/s00190-010-0380-8

Teunissen PJG (2019) A new GLONASS FDMA model. GPS Solutions, 23, 100, https://doiorg/101007/s10291-019-0889-0

Teunissen PJG, Montenbruck O (eds) (2017) Springer Handbook of Global Navigation Satellite Systems. Springer

Teunissen PJG, Massarweh L, Verhagen S (2021) Vectorial integer bootstrapping: flexible integer estimation with application to GNSS. Journal of Geodesy 95(99):1–14

Verhagen S, Li B, Teunissen PJG (2013) Ps-LAMBDA: Ambiguity success rate evaluation software for interferometric applications. Computers and Geosciences 54:361–376

Zhigljavsky A (1991) Theory of Global Random Search. Mathematics and its applications. Kluwer Academic Publishers

6

An efficient 'P1' algorithm for dual mixed-integer least-squares problems with scalar real-valued parameters

In this contribution we consider mixed-integer least-squares problems, where the integer ambiguities $a \in \mathbb{Z}^n$ and real-valued parameters $b \in \mathbb{R}^p$ are estimated. Both a primal and a dual formulation can be considered, with the latter concerning the ambiguity resolution process taking place into the parameters' domain. We study the $p = 1$ case, where an ad hoc 'P1' algorithm is introduced, and some geometrical insights are provided. It is demonstrated how the algorithm's complexity (i.e., number of candidate integer solutions to be evaluated) grows *linearly* with the ambiguity dimensionality n, differently from the primal formulation where an *exponential* growth is observed. By means of numerical simulations, here based on Global Navigation Satellite System (GNSS) models, we show the efficiency of this proposed 'P1' algorithm, meanwhile also demonstrating its quasi-optimal statistical performance.

This chapter has been published as: Massarweh, L., and Teunissen, P.J.G. (2024). *An efficient P1 algorithm for dual mixed-integer least-squares problems with scalar real-valued parameters.* In Journal of Applied Geodesy. https://doi.org/10.1515/jag-2024-0076

6.1 Introduction

The Integer Ambiguity Resolution (IAR) process concerns the successful resolution of the *unknown* integer ambiguities present in mixed-integer models. For instance, in the context of Global Navigation Satellite Systems (GNSS), carrier-phase IAR is the key to fast and high-precision baseline estimation (Teunissen, 2017). Once these ambiguities have been correctly resolved, the carrier-phase data starts acting as very precise pseudo-range data, so enabling users' precise positioning and navigation, see (Blewitt, 1989; Leick et al., 2015).

When considering mixed-integer least-squares problems, two equivalent formulations are possible, denoted as *primal* and *dual*, respectively introduced by Teunissen (1993) and by Teunissen and Massarweh (2024). In the primal formulation, integer ambiguities are firstly resolved followed by a conditional least-squares baseline estimator, so computing ambiguity-fixed baseline solutions. Efficient algorithms exist for tackling such IAR problems within the ambiguity domain, for instance with the Least-Squares AMBiguity Decorrelation Adjustment (LAMBDA) method, see (Teunissen, 1995). On the other hand, in the dual formulation, the IAR process takes place directly in the parameters' domain and globally convergent solutions could be defined, as presented by Teunissen and Massarweh (2024).

In this contribution we further study dual mixed-ILS problems, focusing on the case $p = 1$, i.e. scalar real-valued parameter $b \in \mathbb{R}$, meanwhile assuming an arbitrary number of integer ambiguity components in $a \in \mathbb{Z}^n$, $n \geq 1$. A deterministic P1 algorithm is introduced here as an efficient implementation for the ambiguity search process, now taking place in the parameters' domain (Jazaeri et al., 2012). By defining the algorithm complexity as 'number of candidate solutions evaluated', it is demonstrated that, in the dual formulation (for $p = 1$), the complexity grows linearly with the dimensionality n. Besides the numerical performance, compared here against LAMBDA method, we also investigate statistical performances, thus showing that quasi-optimal solutions can be obtained by the P1 algorithm.

In Section 6.2, a brief review of dual mixed-integer least-squares models is given, then focusing on the case $n \geq p = 1$. In Section 6.3, the P1 algorithm is presented, along with some geometrical insights, and the linear growth of complexity with respect to the dimensionality n is demonstrated. The performance is numerically investigated in Section 6.4, i.e. considering GNSS models, while the main conclusions are summarized in Section 6.5.

6.2 Review of dual mixed ILS models

The dual formulation for mixed-integer least-squares models was introduced by Teunissen and Massarweh (2024). We start here with the observables' vector $y \in \mathbb{R}^m$, so

$$y \sim \mathcal{N}_m \left(Aa + Bb, Q_{yy} \right), \quad a \in \mathbb{Z}^n, \quad b \in \mathbb{R}^p \tag{6.1}$$

where $\sim$ refers to *distributed as*, given a m-dimensional normal distribution with expectation $E\{y\} = Aa + Bb$ and dispersion $D\{y\} = Q_{yy}$ for $Q_{yy} \in \mathbb{R}^{m \times m}$ being the variance-covariance of y. The full-rank design matrix is given by $[A, B] \in \mathbb{R}^{m \times (n+p)}$, while integer ambiguities and real-valued parameters are respectively denoted as a and b.

The dual formulation considers a dual objective function $\mathcal{D} : \mathbb{R}^p \to \mathbb{R}$ given by

$$\mathcal{D}(b) = \left\| b - \hat{b} \right\|^2_{Q_{\hat{b}\hat{b}}} + \| \hat{a}(b) - \breve{a}(b) \|^2_{Q_{\hat{a}(b)}} \tag{6.2}$$

where the (dual) mixed ILS solution for the real-valued parameters follows as

$$\breve{b} = \arg \min_{b \in \mathbb{R}^p} \mathcal{D}(b) \tag{6.3}$$

given the (conditioned) ambiguity vectors

$$\hat{a}(b) = \hat{a} - Q_{\hat{a}\hat{b}} Q_{\hat{b}\hat{b}} \left(\hat{b} - b \right), \quad \breve{a}(b) = \arg \min_{a \in \mathbb{Z}^n} \| \hat{a}(b) - a \|^2_{Q_{\hat{a}(b)}} \tag{6.4}$$

with $Q_{\hat{a}(b)} \in \mathbb{R}^{n \times n}$ as conditional variance-covariance matrix of ambiguities. The latter ones are conditioned onto the current b-value that could freely be defined in $\mathbb{R}^p$.

Two approximations of Eq.(6.2) are discussed in (ibid), such as

I) **Approximate weighting**, where we replace the conditional variance matrix $Q_{\hat{a}(b)}$ by an approximation $Q^{\circ}_{\hat{a}(b)}$, e.g. diagonal matrix, such that

$$\mathcal{D}^{\circ}(b) = \left\| b - \hat{b} \right\|^2_{Q_{\hat{b}\hat{b}}} + \| \hat{a}(b) - \breve{a}^{\circ}(b) \|^2_{Q^{\circ}_{\hat{a}(b)}} \tag{6.5}$$

for $\breve{a}^{\circ}(b) = \arg \min_{a \in \mathbb{Z}^n} \| \hat{a}(b) - a \|^2_{Q^{\circ}_{\hat{a}(b)}}$.

II) **Approximate mapping**, where we replace the integer minimizer of Eq.(6.4) by an arbitrary admissible estimator $\mathcal{I}^{\bullet} : \mathbb{R}^n \to \mathbb{Z}^n$, e.g. integer rounding, such that

$$\mathcal{D}^{\bullet}(b) = \left\| b - \hat{b} \right\|^2_{Q_{\hat{b}\hat{b}}} + \| \hat{a}(b) - \mathcal{I}^{\bullet}(\hat{a}(b)) \|^2_{Q_{\hat{a}(b)}} \tag{6.6}$$

and the two approximations differ since in the 'approximate weighting' case we neglect off-diagonal terms in $Q_{\hat{a}(b)}$, whereas in the 'approximate mapping' case we might adopt simpler rounding estimators for the many-to-one mapping function $\mathcal{I}^{\bullet}$.

6.2.1 Particular case for $n \geq p = 1$

We focus on the case $p = 1$, thus defining $\beta \in \mathbb{R}$ and $\mathcal{D}\colon \mathbb{R} \to \mathbb{R}$, where

$$\mathcal{D}(\beta) = \mathcal{D}_1(\beta) + \mathcal{D}_2(\beta) = \frac{\left(\beta - \hat{b}\right)^2}{\sigma^2_{\hat{b}}} + \|\hat{a}(\beta) - \check{a}(\beta)\|^2_{Q_{\hat{a}(b)}} \tag{6.7}$$

for $\hat{a}(\beta) = \hat{a} - q_{\hat{a}\hat{b}}\sigma^{-2}_{\hat{b}}\left(\hat{b} - \beta\right)$ given $q_{\hat{a}\hat{b}} \in \mathbb{R}^n$, while $\check{a}(\beta)$ follows Eq.(6.4). Note that in the parameter domain the dual objective function is composed by a parabolic term $\mathcal{D}_1(\beta)$ and a periodic-like term $\mathcal{D}_2(\beta)$, as shown in Fig. 6.5 by Teunissen and Massarweh (2024). At the same time, in the ambiguity domain we are able to define 'pull-in regions', i.e. subsets of $\mathbb{R}^n$ where float vectors are mapped to the corresponding integer.

For integer estimators, the pull-in regions are translational invariant over the integers and cover the entire space $\mathbb{R}^n$ without gaps and overlaps [8, 9]. Moreover, from Eq.(6.7), we observe that potential integer candidates are the ones belonging to pull-in regions crossed by the conditioned line $\hat{a}(\beta)$ for $\beta \in \mathbb{R}$. In Fig. 6.1, we provide an illustrative example for $p = 1, n = 2$ given $\beta \in \left[\beta_{MIN} = \hat{b} - 1.7, \beta_{MAX} = \hat{b} + 1.7\right]$, with these bounds for β further specified later, where

$$\begin{pmatrix} \hat{a}_1 \\ \hat{a}_2 \\ \hat{b} \end{pmatrix} = \begin{pmatrix} +0.4 \\ -0.6 \\ +0.2 \end{pmatrix}, \quad \begin{bmatrix} \sigma^2_{\hat{a}_1} & \sigma_{\hat{a}_1\hat{a}_2} & \sigma_{\hat{a}_1\hat{b}} \\ \sigma_{\hat{a}_2\hat{a}_1} & \sigma^2_{\hat{a}_2} & \sigma_{\hat{a}_2\hat{b}} \\ \sigma_{\hat{b}\hat{a}_1} & \sigma_{\hat{b}\hat{a}_2} & \sigma^2_{\hat{b}} \end{bmatrix} \cong \begin{bmatrix} +0.733 & -0.666 & +0.294 \\ -0.666 & +1.031 & -0.637 \\ +0.294 & -0.637 & +0.490 \end{bmatrix}, \tag{6.8}$$

such that $q_{\hat{a}\hat{b}} = \left(\sigma_{\hat{b}\hat{a}_1}, \sigma_{\hat{b}\hat{a}_2}\right)^T$, and therefore

$$Q_{\hat{a}(b)} = Q_{\hat{a}\hat{a}} - q_{\hat{a}\hat{b}}q^T_{\hat{a}\hat{b}}\sigma^{-2}_{\hat{b}} \cong \begin{bmatrix} +0.557 & -0.284 \\ -0.284 & +0.203 \end{bmatrix}, \quad Q^{\circ}_{\hat{a}(b)} \cong \begin{bmatrix} 0.557 & 0 \\ 0 & 0.203 \end{bmatrix} \tag{6.9}$$

with two pull-in regions represented in red and blue respectively when using $Q_{\hat{a}(b)}$ and its diagonal approximation $Q^{\circ}_{\hat{a}(b)}$ as inverse-weighting matrix. In the first case (in red), we are looking at Integer Least-Squares pull-in regions (i.e. hexagons), while in the second case (in blue) we have more simple Integer Rounding pull-in regions (i.e. unit squares).

The weighting approximation leads to an approximate dual objective function

$$\mathcal{D}^{\circ}(\beta) = \mathcal{D}_1(\beta) + \mathcal{D}^{\circ}_2(\beta) = \frac{\left(\beta - \hat{b}\right)^2}{\sigma^2_{\hat{b}}} + \|\hat{a}(\beta) - \check{a}^{\circ}(\beta)\|^2_{Q^{\circ}_{\hat{a}(b)}} \tag{6.10}$$

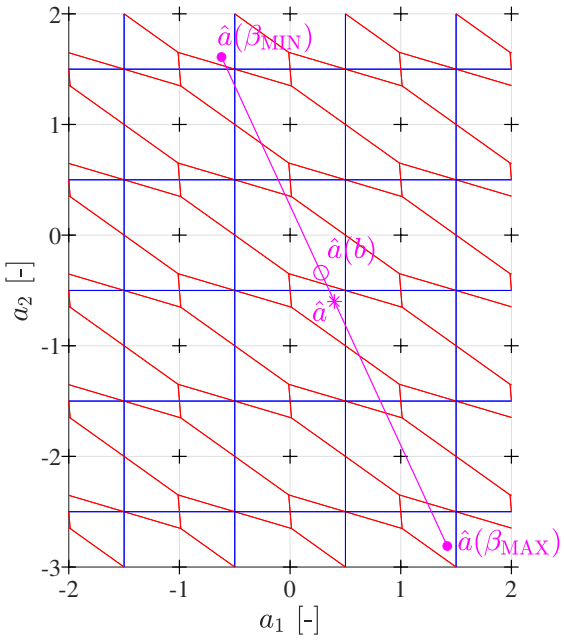


Figure 6.1: The conditioned line $\hat{a}(\beta)$ is shown in magenta given $\beta \in [\beta_{MIN}, \beta_{MAX}]$, i.e. defined between the extreme points $\hat{a}\,(\beta_{MIN})$ and $\hat{a}\,(\beta_{MAX})$. The magenta circle refers to $\hat{a}(b)$ for the true $b \in \mathbb{R}$, while the asterisk refers to $\hat{a}\left(\hat{b}\right) \equiv \hat{a}$, i.e. the float ambiguity solution. Moreover, two pull-in regions are defined as hexagons (in red) or unit squares (in blue) when making use of $Q_{\hat{a}(b)}$ or $Q^{\circ}_{\hat{a}(b)}$ as inverse-weighting matrix, respectively.

so using $\mathcal{D}^{\circ}(\beta)$ instead of $\mathcal{D}(\beta)$, both illustrated in Fig. 6.2, with a common parabolic term $\mathcal{D}_1$ highlighted in green color. The periodic-like terms $\mathcal{D}_2$ and $\mathcal{D}^{\circ}_2$ are depicted in red and blue colors, as for their pull-in regions, respectively on the left and right side. Note that when $Q_{\hat{a}(b)}$ is diagonal, no approximation takes place at all, and the original dual problem is being considered. Even if $Q_{\hat{a}(b)}$ is diagonal, it does not imply that also $Q_{\hat{a}\hat{a}}$ is diagonal, and in the primal formulation we might deal with highly correlated (unconditioned) ambiguities.

The required interval $[\beta_{\text{MIN}}, \beta_{\text{MAX}}]$ can be found starting with an initial guess $\beta_0 \overset{\text{def}}{=\!=} \hat{b}$, so leading to $\hat{a}\,(\beta_0) = \hat{a}$, and then seeking new solutions β_j for $j > 0$, such that

$$\mathcal{D}\,(\beta_j) \leq \mathcal{D}\,(\beta_0) = \|\hat{a} - \check{a}\,(\beta_0)\|^2_{Q_{\hat{a}(b)}} \tag{6.11}$$

and therefore

$$\frac{\left(\hat{b} - \beta_j\right)^2}{\sigma^2_{\hat{b}}} \equiv \mathcal{D}_1\,(\beta_j) \leq \mathcal{D}\,(\beta_j) \quad \overset{\text{Eq.(6.11)}}{\rightarrow} \quad \frac{\left(\hat{b} - \beta_j\right)^2}{\sigma^2_{\hat{b}}} \leq \mathcal{D}\,(\beta_0) \tag{6.12}$$

where the first inequality follows from the definition of $\mathcal{D}(\beta)$, see Eq.(6.7).

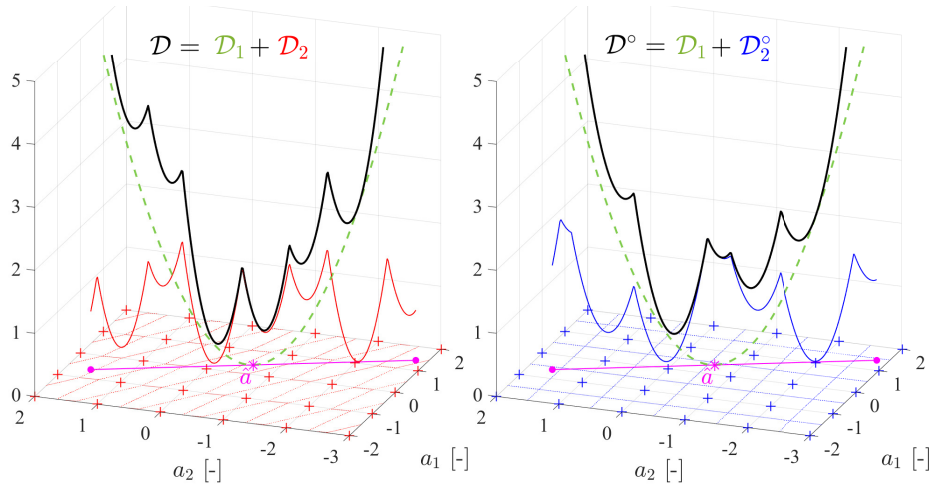


Figure 6.2: The dual objective function $\mathcal{D}(\beta)$ and its approximation $\mathcal{D}^\circ(\beta)$ are shown in the left and right plots, respectively, with the parabolic term $\mathcal{D}_1(\beta)$ shown as green dashed line. The periodic-like contributions $\mathcal{D}_2(\beta)$ and $\mathcal{D}_2^\circ(\beta)$ are respectively given in red and blue, as well as their pull-in regions, while the conditioning line $\hat{a}(\beta)$ is depicted in magenta, centered in $\hat{a}(\hat{b}) \equiv \hat{a}$ (magenta asterisk), where we have $\mathcal{D}_1(\hat{b}) = 0$.

Hence, we are able to define an initial search radius $R_0 = \sigma_{\hat{b}}\sqrt{\mathcal{D}(\beta_0)}$ such that $-R_0 \leq \beta_j - \hat{b} \leq R_0$ for any j-th candidate solution, i.e. interval $\left[\hat{b} - R_0, \hat{b} + R_0\right]$ is found. A grid search could be performed over this interval, but in this work we will consider an alternative efficient approach for the enumeration of all potential solutions, as discussed next.

6.3 The P1 algorithm

In the earlier Fig. 6.1 we notice how several β values might belong to the same pull-in regions, implying that a grid search might be inefficient. At the same time, pull-in regions are convex regions for any $n > 0$, and the conditioning line $\hat{a}(\beta)$ will at most intersect them twice. By identifying any β value associated with each different integer vector $\check{a}(\beta)$, we can obtain a finite set of integer candidates where we expect to find a global minimizer. This is the main idea of the 'P1' algorithm described in this section.

The intersection between $\hat{a}(\beta)$ and ILS pull-in regions is not trivial, therefore we will restrict our discussion to the approximated dual objective function $\mathcal{D}^\circ(\beta)$. Hence, we make use of $Q^\circ_{\hat{a}(b)}$ after neglecting the off-diagonal terms of $Q_{\hat{a}(b)}$, but notice that in some mixed-integer models the latter one might already be a diagonal matrix, as discussed later. Based on the example of Fig. 6.1, we can consider again an illustration of the conditioning line $\hat{a}(\beta)$ for $\beta \in [\beta_{MIN}, \beta_{MAX}]$ given in magenta color in the Fig. 6.3 (note a rotated plot).

The intersections with each interface (of unit-square regions) are shown by magenta squares, while the middle points β_j^{MP} are computed between two consecutive intersections, here shown as (filled) blue hexagrams. Each candidate β_j^{MP} will belong to an individual rounding pull-in region.

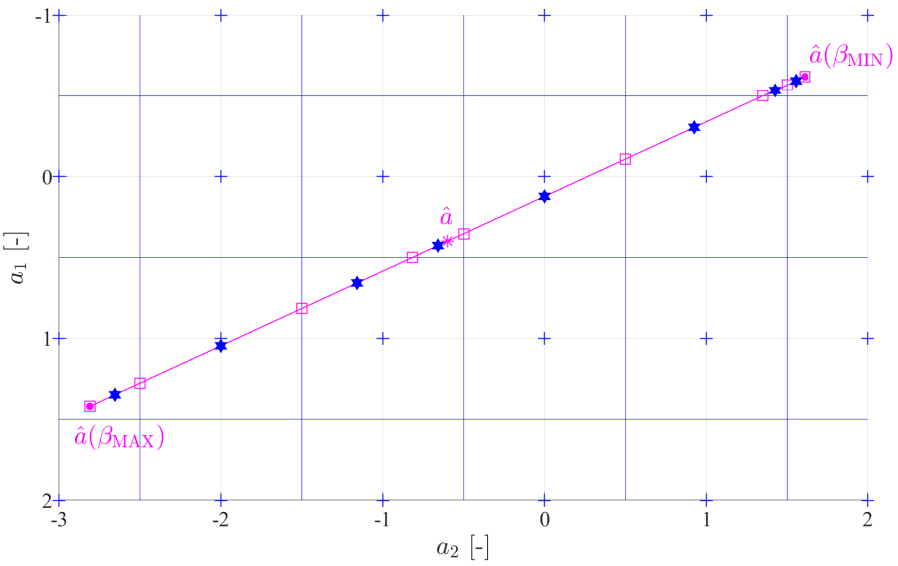


Figure 6.3: Illustration of the example $p = 1, n = 2$, from Fig. 6.1, with the conditioning line $\hat{a}(\beta)$ in magenta color given $\beta \in [\beta_{MIN}, \beta_{MAX}]$. The intersections with all interfaces are shown as magenta squares, while the middle points β_j^{MP} are given as (filled) blue hexagrams.

6.3.1 Algorithm description

The P1 algorithm starts with an initial search radius R_0 and we consider individually each i-th component of the conditioned ambiguity vector. By making use of the expression of the conditioned line $\hat{a}(\beta)$ projected along each component, i.e. $\hat{a}_i(\beta) = \hat{a}_i + q_{\hat{a}_i\hat{b}}\sigma_{\hat{b}}^{-2}\left(\beta - \hat{b}\right)$, it is then possible to bound each i-th term since $\beta - \hat{b} \in [-R_0, +R_0]$, so having

$$\hat{a}_i(\beta) \in [\hat{a}_i - \Delta_i, \hat{a}_i + \Delta_i], \quad \Delta_i \overset{\text{def}}{=\!=} \frac{\left|q_{\hat{a}_i\hat{b}}\right|}{\sigma_{\hat{b}}^2} R_0 \tag{6.13}$$

where $q_{\hat{a}_i\hat{b}}$ refers to the i-th component of the vector $q_{\hat{a}\hat{b}} \in \mathbb{R}^n$. In this way, we can compute all the integer components $\nu_i \in \mathbb{Z}$ between $\hat{a}_i - \Delta_i$ and $\hat{a}_i + \Delta_i$, and define values β_{ν_i} that are related to the interfaces with rounding pull-in regions. These values are then collected from all i-th ambiguity components in one *sorted* list, i.e. $\{\beta_{MIN}, \ldots, \beta_{j-1}, \beta_j, \beta_{j+1}, \ldots, \beta_{MAX}\}$, while also including both two extrema $\beta_{MIN} = \hat{b} - R_0$ and $\beta_{MAX} = \hat{b} + R_0$.

The middle point can be found by $\beta_j^{MP} = (b_j + b_{j+1})/2$, and will belong to a single integer pull-in region, hence refers to an individual integer candidate $u_j \overset{\text{def}}{=\!=} \check{a}^{\circ}\left(\beta_j^{MP}\right) \in \mathbb{Z}^n$. In the solutions' evaluation process, we can start with values β_j^{MP} that are closer to $\hat{b}$, since they will be associated with a smaller parabolic term $\mathcal{D}_1\left(\beta_j^{MP}\right)$. Moreover, during the evaluations, we can make use of results from Eq.(6.12) in order to reduce the search radius R_0 after each evaluation. This leads to discarding several outer values of β that could not further minimize the dual objective function.

This search-and-shrink strategy is similar to the primal counterpart already adopted in LAMBDA, but now we 'shrink' the real-valued parameters' domain (i.e. on the conditioning line) where the enumeration process takes place. Once all potential candidates in the current interval are evaluated, the process stops and the global minimum β^* is obtained.

A summary of 'P1' algorithm is given in Algorithm 1, which consists of three parts: INITIALIZATION where an initial guess allows computing the initial search radius R_0, ENUMERATION of the integer candidates that are found inside the initial interval, and MAIN SEARCH, where such potential solutions are evaluated, including the aforementioned 'shrinking' strategy.

Algorithm 1 Summary of the 'P1' algorithm described in Section 6.3.1

INPUTS:

$$\hat{a} \in \mathbb{R}^n,\ \hat{b} \in \mathbb{R}, Q_{\hat{a}\hat{a}} \in \mathbb{R}^{n \times n}, q_{\hat{a}\hat{b}} \in \mathbb{R}^n,\ \sigma_{\hat{b}} \in \mathbb{R}$$

INITIALIZATION:

Given $\beta_0 = \hat{b}$ (initial guess), compute $\hat{a}\,(\beta_0) = \hat{a}$ based on Eq.(6.4), which is used to find $\breve{a}_0^{\circ} = \lceil \hat{a} \rfloor \in \mathbb{Z}^n$, along with an initial search radius $R_0 = \sigma_{\hat{b}} \sqrt{\mathcal{D}^{\circ}\left(\hat{b}\,(\breve{a}_0^{\circ})\right)}$.

ENUMERATION: % Find all potential integer candidates

For each i-th ambiguity component $\hat{a}_i$, with $i = 1, \ldots, n$, find the intersections with rounding pull-in regions given $\nu_i \in \mathbb{Z} \cap [\hat{a}_i - \Delta_i, \hat{a}_i + \Delta_i]$, see Eq.(6.13).

Collect all β-values in one *sorted* list: $\{\beta_{MIN}, \ldots, \beta_j, \ldots, \beta_{MAX}\}$, so including the extrema $\beta_{MIN} = \hat{b} - R_0$ and $\beta_{MAX} = \hat{b} + R_0$.

Compute the middle points $\beta_j^{MP} = (b_j + b_{j+1})/2$, each associated to a single integer candidate, and sorted now starting from smaller values of $\left|\beta_j^{MP} - \hat{b}\right|$.

MAIN SEARCH: % Evaluate each potential integer candidate

Set $\mathcal{D}^*$ as the current best function value from the step INITIALIZATION, with the current best solution defined by $\beta^* = \hat{b}\,(\breve{a}_0^{\circ})$.

% Iterate over each j-th potential solution, sorted by $\left|\beta_j^{MP} - \hat{b}\right|$.
for $j = 1, \ldots, N$
    % Outside the interval, search is over.
    if $\left|\beta_j^{MP} - \hat{b}\right| > R_0$
        Break loop;
    end

    % Evaluation of the current integer candidate.
    Compute the integer vector $u_j \overset{\text{def}}{=\!=} \breve{a}^{\circ}\left(\beta_j^{MP}\right) = \lceil \hat{a}\left(\beta_j^{MP}\right) \rfloor$, which it is used to evaluate a new objective function $\mathcal{D}_j^{NEW} = \mathcal{D}^{\circ}\left(\hat{b}\,(u_j)\right)$, see Eq.(6.15).

    % Update the current best solution
    if $\mathcal{D}_j^{NEW} < \mathcal{D}^*$
        Save $\mathcal{D}^* = \mathcal{D}_j^{NEW}$ and $\beta^* = \beta_j^{NEW} = \hat{b}\,(u_j)$;

        % Shrinking step, see Eq.(6.12)
        Update current R_0 (search radius);
    end
end

OUTPUTS:

$$\beta^* \in \mathbb{R},\quad \breve{a}^{\circ}\,(\beta^*) \in \mathbb{Z}^n,\quad \mathcal{D}^* \in \mathbb{R}$$

NOTE: the symbol $\lceil \cdot \rfloor$ defines the integer rounding operator.

In Algorithm 1, the main search takes place starting with smaller values $\left|\beta_j^{MP} - \hat{b}\right|$, then considering that $\left|\beta_j^{MP} - \hat{b}\right| \leq \left|\beta_{j+1}^{MP} - \hat{b}\right|$ for any $j > 0$. Therefore, once the j-th solution is found outside the interval $\left[\hat{b} - R_0, \hat{b} + R_0\right]$, given the current radius R_0 (updated after each evaluation), we can stop the for-loop iterations. For the evaluation of each integer candidate $u_j \overset{\text{def}}{=\!=} \check{a}^{\circ}\left(\beta_j^{MP}\right)$, we might directly exploit a primal formulation, thus evaluating

$$\mathcal{P}^{\circ}(u_j) = \|\hat{a} - u_j\|^2_{Q^{\circ}_{\hat{a}\hat{a}}}, \quad \check{b} \overset{\text{def}}{=\!=} \hat{b}(u_j) = \hat{b} - q^T_{\hat{a}\hat{b}} {Q^{\circ}_{\hat{a}\hat{a}}}^{-1}(\hat{a} - u_j) \tag{6.14}$$

given the approximate primal objective function $\mathcal{P}^{\circ} : \mathbb{Z}^n \to \mathbb{R}$ for $Q^{\circ}_{\hat{a}\hat{a}} = Q^{\circ}_{\hat{a}(b)} + q_{\hat{a}\hat{b}} q^T_{\hat{a}\hat{b}} / \sigma^2_{\hat{b}}$, so following Lemma 4 by Teunissen and Massarweh (2024).

On the other hand, the candidates' evaluation could also be performed directly in the dual formulation, starting from a selected integer value $u \in \mathbb{Z}^n$. Using Lemma 7 from (ibid), we know that the local minimizer for $\hat{a}(\beta) \in \mathcal{S}_u$, i.e. pull-in region of u such that $\check{a}^{\circ}(\beta) = u$, can be computed based on a primal conditioning $\hat{b}(u) = \hat{b} - q^T_{\hat{a}\hat{b}} {Q^{\circ}_{\hat{a}\hat{a}}}^{-1}(\hat{a} - u)$, where

$$\mathcal{P}^{\circ}(u) = \mathcal{D}^{\circ}\left(\hat{b}(u)\right), \quad \forall u \in \mathbb{Z}^n \tag{6.15}$$

so now evaluating $\mathcal{D}^{\circ}$ and no other minima exist given $\hat{a}(\beta) \in \mathcal{S}_u$, see (ibid). In this way it is possible to initialize using $Q^{\circ}_{\hat{a}\hat{a}}$ rather than $Q^{\circ}_{\hat{a}(b)}$, so leading to fewer integer candidates.

6.3.2 Algorithm complexity

Before presenting a numerical analysis of the performance, we can briefly discuss the *linear* growth of complexity, here defined by the number of potential candidates that are evaluated during the enumeration process. In fact, in the dual formulation with $\mathcal{D}^{\circ}(\beta)$, the algorithm's complexity can be shown to be linearly increasing with the dimensionality $n > 0$.

Starting with Eq.(6.13), we can approximate a maximum number of candidates of each i-th component directly using $N_i^{MAX} \cong 2\Delta_i$, such that the total number of candidates (i.e. the middle points computed between successive intersections) follows as

$$N^{MAX} = \sum_{i=1}^{i=n} N_i^{MAX} \cong \sum_{i=1}^{i=n} 2R_0 \frac{\left|q_{\hat{a}_i\hat{b}}\right|}{\sigma^2_{\hat{b}}} = \frac{2R_0}{\sigma^2_{\hat{b}}} \sum_{i=1}^{i=n} \left|q_{\hat{a}_i\hat{b}}\right| \tag{6.16}$$

where the approximation $\cong$ gets more accurate for larger values of $N_i^{MAX}, i = 1, \ldots, n$. Given that $\left|q_{\hat{a}_i\hat{b}}\right| = \left|\rho_{\hat{a}_i\hat{b}}\right| \sigma_{\hat{a}_i} \sigma_{\hat{b}}$ for the correlation term $\rho_{\hat{a}_i\hat{b}} \in [-1, +1]$, we notice

that each N_i^{MAX} grows with the (unconditioned) standard deviation $\sigma_{\hat{a}_i}$ of the i-th ambiguity component $\hat{a}_i$, which is dependent on the underlying model strength. Note that the search of a global minimum will take around N^{MAX} evaluations of all the middle points β_j^{MP} currently in the list, meanwhile we continue here by defining

$$\overline{\left|q_{\hat{a}_i\hat{b}}\right|} = \frac{1}{n}\sum_{i=1}^{i=n}\left|q_{\hat{a}_i\hat{b}}\right| \tag{6.17}$$

and we further reformulate the approximation in Eq.(6.16) as

$$N^{MAX} \cong \frac{2R_0}{\sigma_{\hat{b}}^2} n\overline{\left|q_{\hat{a}_i\hat{b}}\right|} \tag{6.18}$$

where $\overline{\left|q_{\hat{a}_i\hat{b}}\right|}$ is smaller than the largest entry (in absolute value) of $q_{\hat{a}\hat{b}} \in \mathbb{R}^n$. This shows how N^{MAX} grows linearly with the dimensionality n given an initial search radius R_0, which can be computed based on Eq.(6.11), e.g., using $R_0 = \sigma_{\hat{b}}\sqrt{\mathcal{D}(\beta_0)}$, therefore we obtain

$$N^{MAX} \cong 2n \cdot \sqrt{\mathcal{D}(\beta_0)} \cdot \frac{\overline{\left|q_{\hat{a}_i\hat{b}}\right|}}{\sigma_{\hat{b}}} \tag{6.19}$$

and the complexity is indeed dependent upon three elements: the problem dimensionality, the initial guess β_0 and the variance/covariance terms. Note that given $\overline{\left|q_{\hat{a}_i\hat{b}}\right|} = 0$, no correlation exists at all between the float ambiguities $\hat{a}_i$ and the parameter $\hat{b}$, so the latter is actually the global minimizer for our dual problem.

6.4 Numerical assessments

We start by a simple numerical example for a multi-frequency geometry-free model, where the conditional variance matrix $Q_{\hat{a}(b)}$ is diagonal, whereas the unconditional matrix $Q_{\hat{a}\hat{a}}$ is not. Given a single-epoch single-baseline ionosphere-fixed scenario, two receivers track two Galileo satellites based on a standard deviation of $\sigma_p = 30\ cm$ and $\sigma_\phi = 3\ mm$ respectively for the undifferenced code and phase observations.

The mixed-integer model of Eq.(6.1), given J frequencies, is based on

$$A = \begin{bmatrix} 0 \\ \Lambda_J \end{bmatrix}, \quad B = \begin{bmatrix} e_J \\ e_J \end{bmatrix}, \quad Q_{yy} = 4\begin{bmatrix} \sigma_p^2 I_J & \\ & \sigma_\phi^2 I_J \end{bmatrix} \tag{6.20}$$

where $I_J \in \mathbb{Z}^{J\times J}$ refers to the identity matrix, $e_J \in \mathbb{Z}^J$ is a vector of 1s, while $\Lambda_J \in \mathbb{R}^{J\times J}$ is a diagonal matrix with its entries as the signal wavelengths, where

$n \equiv J$. The factor '4' arises from the covariance propagation law following the double-differencing operator for code and phase observables.

We consider four scenarios based on Galileo signal, i.e. E1+E6 ($n = 2$), E1+E6+E5a ($n = 3$) E1+E6+E5a+E5b ($n = 4$) and E1+E6+E5a+E5b+E5 ($n = 5$), where computational time over 2000 different samples has been presented in Fig. 6.4. In all these simulations, the results for $p = 1$ using the P1 algorithm perfectly match with the ILS solutions computed by LAMBDA 4.0 toolbox (Massarweh et al., 2025), since $Q_{\hat{a}(b)}$ is a diagonal matrix, where

$$Q^{-1}_{\hat{a}(b)} = A^T Q^{-1}_{yy} A = \frac{\Lambda_J^T \Lambda_J}{4\sigma^2_\phi} = \frac{1}{4\sigma^2_\phi} diag\left(\lambda_1^2, \ldots, \lambda_J^2\right) \tag{6.21}$$

therefore no approximation takes place in our dual formulation, and

$$\mathcal{D}(\beta) \equiv \mathcal{D}^\circ(\beta) = n\frac{\left(\beta - \hat{b}\right)^2}{4\sigma_p^2} + \frac{1}{4\sigma^2_\phi}\sum_{i=1}^{i=n} \lambda_i^2 \check{\epsilon}_i^2(\beta), \quad \forall \beta \in \mathbb{R} \tag{6.22}$$

for the ambiguity residuals' defined by $\check{\epsilon}_i(\beta) = \hat{a}_i(\beta) - \lceil\ \hat{a}_i(\beta)\rfloor$ for $i = 1, \ldots, n$ and $\lceil\cdot\rfloor$ being the integer rounding operation. Notice how the parabolic term $\mathcal{D}_1$ and periodic-like term $\mathcal{D}_2$ are mainly driven by the precision of code and phase measurements, respectively. Moreover, in this illustrative example, we observe a smaller computational time for the P1 algorithm with respect to LAMBDA method, and its efficiency becomes more visible when increasing the ambiguity problem dimensionality.

Based on 2000 samples ($n = 3$), we show in Fig. 6.5 the maximum number of integer candidates (in blue) as defined by Eq.(6.16), which seems to well approximate the potential number of candidates (in orange) iterated in the MAIN SEARCH step (see Algorithm 1), i.e. number of middle points previously found in ENUMERATION. However, by means of the search-and-shrink approach, we note that the actual number of integer candidates evaluated (in yellow) is substantially reduced. This demonstrates substantial improvements in terms of efficiency once accounting for a search-and-shrink strategy in the 'P1' algorithm, as adopted in the numerical results of Fig. 6.4.

At this point, we continue with a different numerical example, where the matrix $Q_{\hat{a}(b)}$ is not diagonal and suboptimal performance (with respect to a primal ILS estimator) might be expected, so we focus on statistical performances rather than computational ones.

6.4.1 Statistical performance for $Q_{\hat{a}(b)}$ not diagonal

We consider a single-epoch single-baseline geometry-based ionosphere-fixed model, with m satellites tracked on GPS L1 frequency. We assume the horizontal position known, leading to $p = 1$ estimation of the vertical (UP) coordinate $b_{UP} \in \mathbb{R}$, where

$$A = \begin{bmatrix} 0 \\ \lambda_1 I_{m-1} \end{bmatrix}, \quad B = \begin{bmatrix} D_m^T G \\ D_m^T G \end{bmatrix}, \quad Q_{yy} = 2\begin{bmatrix} D_m^T Q_{pp} D_m & \\ & D_m^T Q_{\phi\phi} D_m \end{bmatrix} \tag{6.23}$$

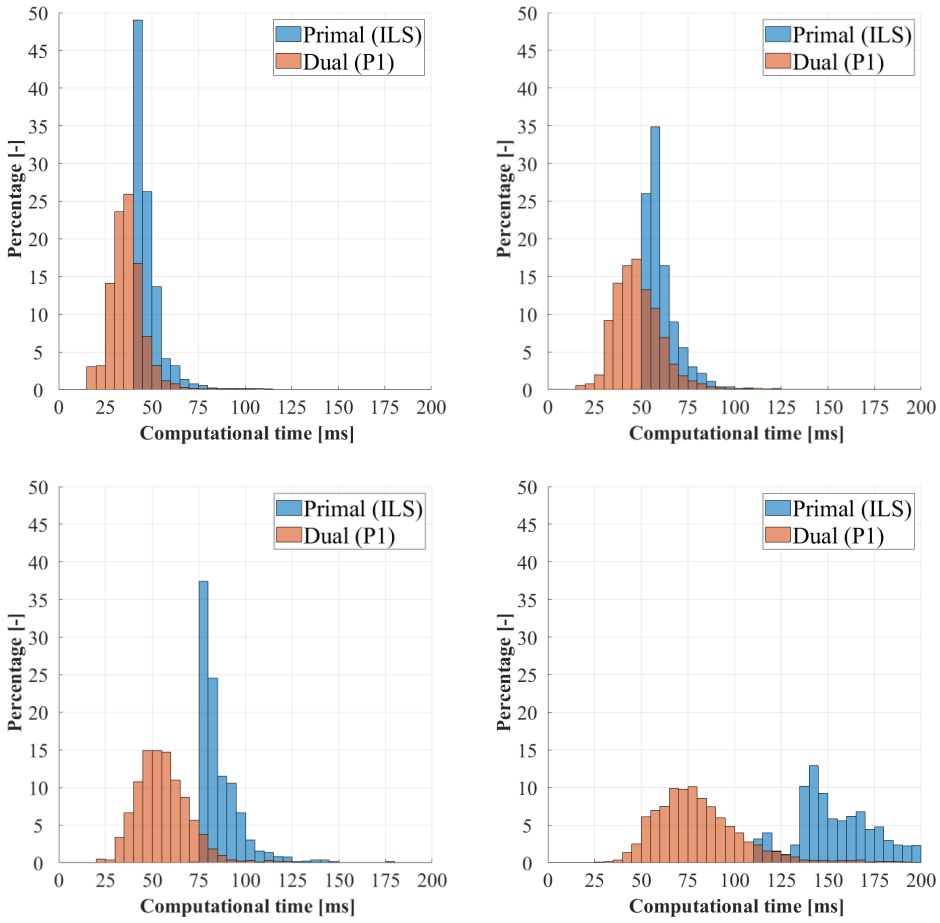


Figure 6.4: The computational times for primal (ILS) and dual (P1) solutions are shown given four different Galileo scenarios, i.e. E1+E6 ($n = 2$, top left), E1+E6+E5a ($n = 3$, top right) E1+E6+E5a+E5b ($n = 4$, bottom left) and E1+E6+E5a+E5b+E5 ($n = 5$, bottom right). For each scenario, we use 2000 different float samples, see text for more information.

with $I_n \in \mathbb{Z}^{n\times n}$ refers to the identity matrix for $n = m-1$, while $D_m^T = [-e_{m-1}, I_{m-1}]$ is the between-satellite single differencing with respect to the first (pivot) satellite, and the vector $G \in \mathbb{R}^m$ consists of $\sin(el^\circ)$ terms for each GPS satellite, see (Odijk and Teunissen, 2008).

The stochastic model follows by a covariance propagation law for the undifferenced code and phase standard deviation (at zenith), respectively given as 30 cm and 3 mm, while an elevation weighting $\propto 1/\sin(el^\circ)$ is adopted. Therefore, both matrices $Q_{\phi\phi}$ and Q_{pp} are diagonal with entries being the elevation-dependent standard deviations. Note that a term '2' in the expression for Q_{yy} arises from the between-receiver differencing.

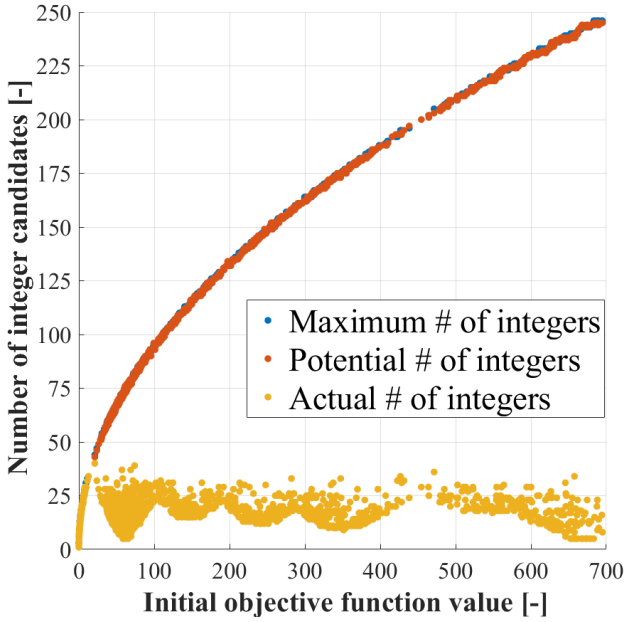


Figure 6.5: The maximum number of expected integer candidates is shown (in blue) based on the Eq.(6.16), while the potential number of integers refers to middle points (in orange) actually computed in this analysis over 2000 different samples for $n = 3$. Lastly, we also provide the actual total number of integers being evaluated (in yellow), largely reduced thanks to the search-and-shrink strategy adopted by the 'P1' algorithm.

At this point, given $Q^{\circ}_{\hat{a}\hat{a}}$, we aim to analyze such dual approximation of the variance-covariance matrix $Q_{\hat{a}\hat{a}}$ for the estimated (unconditioned) float ambiguities. The objective is to assess how accurate the dual approximation is, where as a matter of comparison we will consider also a primal approximation given $Q^{\odot}_{\hat{a}\hat{a}}$. Three different scenarios are compared:

- **Case #1**: We use $Q_{\hat{a}\hat{a}} = Q_{\hat{a}(b)} + Q_{\hat{a}\hat{b}}Q^{-1}_{\hat{b}\hat{b}}Q_{\hat{b}\hat{a}}$;
- **Case #2**: We use $Q^{\circ}_{\hat{a}\hat{a}} = Q^{\circ}_{\hat{a}(b)} + Q_{\hat{a}\hat{b}}Q^{-1}_{\hat{b}\hat{b}}Q_{\hat{b}\hat{a}}$, with $Q^{\circ}_{\hat{a}(b)}$ diagonal;
- **Case #3**: We use $Q^{\odot}_{\hat{a}\hat{a}} = \left(Q_{\hat{a}(b)} + Q_{\hat{a}\hat{b}}Q^{-1}_{\hat{b}\hat{b}}Q_{\hat{b}\hat{a}}\right)^{\odot}$, with $Q^{\odot}_{\hat{a}\hat{a}}$ diagonal;

where in the **Case #1**, we consider a (primal) Integer Least-Squares solution based on the full variance-covariance matrix $Q_{\hat{a}\hat{a}}$, in the **Case #2** we adopt a dual approximation based on $Q^{\circ}_{\hat{a}\hat{a}}$ where $Q^{\circ}_{\hat{a}(b)}$ accounts only for the diagonal entries of $Q_{\hat{a}(b)}$ (that is not diagonal), while in the **Case #3** we consider a (primal) Integer Rounding solution based on $Q^{\odot}_{\hat{a}\hat{a}}$, being the diagonal matrix extracted from $Q_{\hat{a}\hat{a}}$. In LAMBDA, a decorrelation process (Teunissen, 1995) is used to reduce the correlation between (unconditioned) float ambiguities, based on an admissible transformation matrix

$Z \in \mathbb{Z}^{n\times n}$ (unimodular), so $\hat{z} = Z^T\hat{a}$. Still, in order to directly compare statistical performances of these three cases, no ambiguity decorrelation has been used here, while in **Case #1** such a re-parametrization would only affect the computational time.

In Fig. 6.6, the errors for the 'UP' component are shown using 6000 samples, where a total of 8 satellites has been tracked, i.e. $n = 7$. In grey color, the float solution is illustrated, having a standard deviation $\sigma_{\hat{b}} \cong 1.612$ [m], while the fixed results are presented in green and red color referring to correctly fixed and incorrectly fixed ambiguities, respectively. The success rate (SR) is computed for the three scenarios, with $SR \cong 97.9\%, \cong 97.0\%$, $\cong 6.3\%$ for each case. The poor success rate of the primal rounding is visible in the large UP errors, with a root mean squares (RMS) value of around 1.6 [m] in contrast to the 1.6 [cm] found for the fixed solutions when ambiguities are correctly resolved, i.e. $a = 0$ (in green).

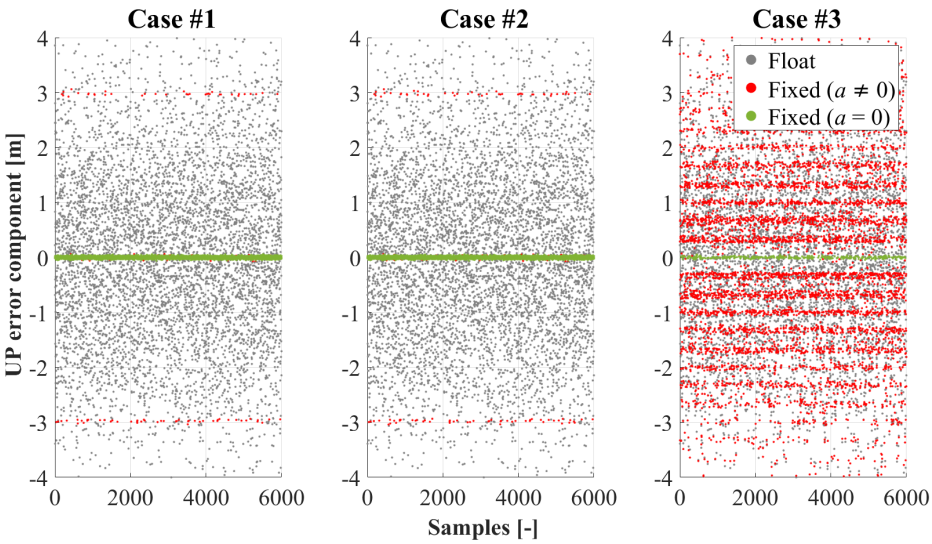


Figure 6.6: The 'UP' error component [m] is shown for different samples generated based on a single-epoch single-baseline geometry-based ionosphere-fixed model for L1 signal tracked by 8 GPS satellites. In grey color, we show the float solution, along with fixed solutions in red or green respectively for incorrectly resolved or correctly resolved ambiguities. See text for more details on the three cases considered for this example.

We can further investigate the good performance of the dual formulation by looking at the two individual terms of $Q_{\hat{a}\hat{a}}$, i.e. $Q_{\hat{a}(b)}$ and $Q_{\hat{a}\hat{b}}Q_{\hat{b}\hat{b}}^{-1}Q_{\hat{b}\hat{a}}$, and we re-write those as

$$Q_{\hat{a}\hat{a}} \triangleq Q_{\hat{a}(b)} + Q_{\hat{a}\hat{b}}Q_{\hat{b}\hat{b}}^{-1}Q_{\hat{b}\hat{a}} = \left(A^T Q_{yy}^{-1} A\right)^{-1} + A^+ B Q_{\hat{b}\hat{b}} B^T \left(A^+\right)^T \tag{6.24}$$

with $A^+ = \left(A^T Q_{yy}^{-1} A\right)^{-1} A^T Q_{yy}^{-1}$ as the left inverse matrix of $A \in \mathbb{R}^{m\times n}$. In this last example, the conditional variance-covariance matrix of the ambiguities $Q_{\hat{a}(b)}$ was given by

$$Q_{\hat{a}(b)} = \left(A^T Q_{yy}^{-1} A\right)^{-1} = 2\left(D_m^T Q_{\phi\phi} D_m\right)/\lambda_1^2 \tag{6.25}$$

where the (conditioned) L1 ambiguities are correlated through the between-satellite single-differencing operator. In several GNSS models, like the one adopted for our last numerical example, most of the correlation in the unconditional variance matrix $Q_{\hat{a}\hat{a}}$ arise due to the presence of real-valued parameters, meanwhile $Q_{\hat{a}(b)}$ is generally also small due to the much higher precision of phase measurements, e.g. $\sigma_p/\sigma_\phi \cong 100$.

For sake of completeness, in Fig. 6.7 we can show the numerical values of $Q_{\hat{a}\hat{a}}$ and its individual matrix terms, where it is visible how the approximation $Q^{\circ}_{\hat{a}(b)}$ would actually have little impact, so $Q^{\circ}_{\hat{a}\hat{a}}$ is very similar to the original matrix $Q_{\hat{a}\hat{a}}$. In the **Case #3**, where large off-diagonal components have been neglected, the approximation is poor, and therefore the correlation among ambiguities is not taken properly into account.

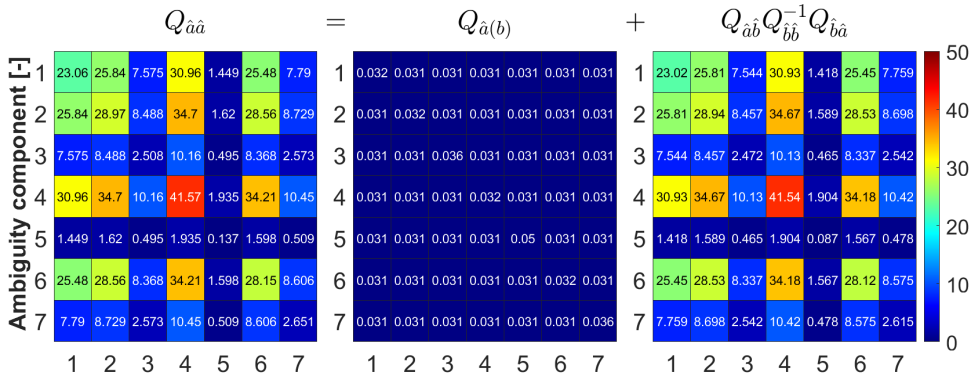


Figure 6.7: The entries of matrix $Q_{\hat{a}\hat{a}}$, as well as $Q_{\hat{a}(b)}$ and $Q_{\hat{a}\hat{b}}Q_{\hat{b}\hat{b}}^{-1}Q_{\hat{b}\hat{a}}$, are shown based on the numerical example used in Fig. 6.6. Note that the elevation for the eight satellites was 62.6°, 49.6°, 48.8°, 43.9°, 18.5°, 18.2°, 9.3° and 7.3°, using a weighting scheme $\propto 1/\sin(el°)$.

6.4.2 Additional remarks

The proposed P1 algorithm is limited to the rounding pull-in regions (i.e. unit hypercubes in $\mathbb{R}^n$) since they allow efficiently computing intersections and middle points, later used for the enumeration of potential integer candidates (see Section 6.3.1). In this way, in the dual case it was possible to demonstrate a *linear* growth of the complexity with respect to the ambiguity dimensionality n, differently from the *exponential* growth found in the primal problem, e.g. see (Fincke and Pohst, 1985). At the same time, we have restricted the algorithm to the case $p = 1$ since the computation of these intersections in a more general case $p > 1$ is not trivial, even while working with the rounding pull-in regions. In fact, in this case, the enumeration of candidates belonging to each pull-in region would require a more expensive search for all intersections, i.e. most likely very inefficient. On the other hand, the potential repeated use of P1 algorithm for searching over different directions in $\mathbb{R}^p$ (parameters' domain) is also possible, and it will be subject of future works.

6.5 Conclusions

In this contribution we have considered mixed-integer models, and the dual mixed-Integer Least-Squares (ILS) formulation introduced by Teunissen and Massarweh (2024). In the dual problems, the resolution of integer ambiguities takes place in the domain of real-valued parameters, which are freely defined in $\mathbb{R}^p$. We focus here on the $p = 1$ case, where a scalar parameter is estimated together with an arbitrary number of ambiguities $n \geq 1$.

As an alternative to a grid search approach, we present the 'P1' algorithm, i.e. a deterministic global solution for the minimization of dual problems ($p = 1$). This algorithm is based on the intuition that potential integer candidates are the ones whose pull-in region is crossed by the conditioning line $\hat{a}(\beta) = \hat{a}(0) + q_{\hat{a}\hat{b}}\sigma_{\hat{b}}^{-2}\beta$, for $\beta \in \mathbb{R}$. Therefore, we show how algorithm's complexity grows linearly with respect to the dimensionality n. Moreover, the proposed 'P1' algorithm makes use of a search-and-shrink strategy that further enhances its computational performance, here evaluated numerically with respect to an implementation of the LAMBDA method, see (Massarweh et al., 2025). As the ambiguity dimensionality n increases, the P1 algorithm computationally outperforms LAMBDA search process that takes place in the primal formulation.

Given that the algorithm assumes an approximate conditional matrix $Q^{\circ}_{\hat{a}(b)}$, i.e. after neglecting the off-diagonal terms of $Q_{\hat{a}(b)}$, this work shows how quasi-optimal performance could still be achieved when $Q_{\hat{a}(b)}$ provides a small contribution to the unconditional matrix $Q_{\hat{a}\hat{a}}$. This is often the case in GNSS models, where the term $Q_{\hat{a}(b)}$ is generally small since driven by the high precision of phase data, while most of correlation among (unconditioned) ambiguities is due to the presence of the estimated real-valued parameters. Lastly, as this efficient algorithm is limited to the case $p = 1$, a further extension to the case $p > 1$ will be investigated in the future. This is possible, e.g., by considering a partitioned dual formulation (Teunissen and Massarweh, 2024), thus solving the problem in a lower dimensioned space, or by exploiting heuristic search strategies in $\mathbb{R}^p$, thus looking along multiple 'search lines' that intersect with pull-in regions, where the P1 algorithm could be effectively employed.

References

Blewitt G (1989) Carrier phase ambiguity resolution for the global positioning system applied to geodetic baselines up to 2000 km. Journal of Geophysical Research 94(B8):10187–10203, DOI 10.1029/JB094iB08p10187

Fincke U, Pohst M (1985) Improved methods for calculating vectors of short length in a lattice, including a complexity analysis. Mathematics of Computation 44(170):463–471, DOI 10.2307/2007966

Jazaeri S, Amiri-Simkooei A, Sharifi M (2012) Fast integer least-squares estimation for GNSS high-dimensional ambiguity resolution using lattice theory. Journal of Geodesy 86:123–136, DOI 10.1007/s00190-011-0501-z

Leick A, Rapoport L, Tatarnikov D (2015) GPS Satellite Surveying. John Wiley & Sons, DOI 10.1002/9781119018612

Massarweh L, Verhagen S, Teunissen PJG (2025) New LAMBDA toolbox for mixed-integer models: estimation and evaluation. GPS Solutions 29(1):14, DOI 10.1007/s10291-024-01738-z, available at http://pntlab.tudelft.nl/LAMBDA

Odijk D, Teunissen PJG (2008) ADOP in closed form for a hierarchy of multi-frequency single-baseline GNSS models. Journal of Geodesy 82:473–492, DOI 10.1007/s00190-007-0197-2

Teunissen PJG (1993) Least-squares estimation of the integer GPS ambiguities. In: IAG General Meeting. Invited Lecture. Section IV Theory and Methodology

Teunissen PJG (1995) The least-squares ambiguity decorrelation adjustment: a method for fast GPS integer ambiguity estimation. Journal of Geodesy 70(1-2):65–82, DOI 10.1007/BF00863419

Teunissen PJG (2017) Carrier phase integer ambiguity resolution. In: Teunissen PJG, Montenbruck O (eds) Springer Handbook of Global Navigation Satellite Systems, Springer Handbooks, Springer, Cham, DOI 10.1007/978-3-319-42928-1_23

Teunissen PJG, Massarweh L (2024) Primal and dual mixed-integer least-squares: distributional statistics and global algorithm. Journal of Geodesy 98(63), DOI 10.1007/s00190-024-01862-1

7

Branch-and-bound INteger-equivariant Global Optimizer: implementation and performance in dual mixed-ILS problems

In this contribution we investigate the performance in *dual* mixed-Integer Least-Squares (ILS) problems introduced by Teunissen and Massarweh (2024). We focus on the unconstrained *dual* problem, where the integerness of ambiguities is taken into account, while the real-valued parameters are freely defined in $\mathbb{R}^p$. We describe in detail the main implementation steps for the Branch-and-bound INteger-equivariant Global Optimizer (BINGO), a globally convergent solution to *dual* problems. This implementation employs a normalization of the *dual* objective function, and it is based on the branch-and-bound (BB) approach, with BB lower bounds constructed with a convex relaxation, then minimized via projected gradient descent. In this work, we first evaluate the computational trade-offs of the proposed BINGO implementation, and later we assess the ambiguity fixing performance with respect to the *primal* formulation in the context of Global Navigation Satellite System (GNSS).

7.1 Introduction

Mixed-integer models arise in several geodetic problems, e.g. precise point positioning (PPP) and atmosphere remote sensing with multi-constellation Global Navigation Satellite Systems (GNSS), or deformation monitoring by Interferometric Synthetic Aperture Radar (InSAR), or fringe phase observations from Very Long Baseline Interferometry (VLBI). These problems generally involve two types of unknowns: integer ambiguities and real-valued parameters, where the accuracy of these estimated parameters can significantly be improved following a correct resolution of the ambiguities; see Hofmann-Wellenhof et al. (2008).

In Teunissen and Massarweh (2024), mixed-integer least-squares problems have been studied, and a *dual* formulation has been introduced in addition to the *primal* one originally presented by Teunissen (1993). The primal formulation created the foundations for the Least-squares AMBiguity Decorrelation Adjustment (LAMBDA) method, i.e. statistically optimal and numerically efficient approach to the ambiguity resolution (Teunissen, 1995). Moreover, a few different (classes of) estimators were introduced (Teunissen, 2003), with the resolution process taking place in the ambiguity domain. However, the new dual problem formulation introduces a framework that focuses the ambiguity resolution process in the domain of real-valued parameters, given that they are typically the parameters of interest for the user.

In this research work, we numerically investigate the performance in 'unconstrained dual' problems, where we account for the integerness of ambiguities without any additional constraint onto the real-valued parameters, therefore freely defined in $\mathbb{R}^p$. In Section 7.2, we discuss a globally convergent solution to dual problems, as introduced by Teunissen and Massarweh (2024), thus presenting an implementation of the Branch-and-bound INteger-equivariant Global Optimizer (BINGO), and investigating main computational trade-offs using illustrative examples. In Section 7.3, a numerical analysis of the ambiguity-fixing performance is carried out based on GNSS models, e.g., by comparing the success rate (SR) between *primal* and *dual* formulations. Lastly, in Section 7.4, the main conclusions are summarized.

7.2 Branch-and-bound INteger-equivariant Global Optimizer (BINGO)

The Branch-and-bound INteger-equivariant Global Optimizer (BINGO) represents a globally convergent solution to dual problems, see (ibid), and makes use of three principles:

- **Branch and Bound (BB) algorithm**, for achieving the global minimization of non-convex problems given an initial search domain, see Lawler and Wood (1966);
- **Convex relaxation**, for computing the required lower bound in the branch-and-bound iterations, here based on a convex continuous differentiable C^1-function;
- **Projected Gradient Descent (PGD),** used for minimizing this convex lower bound over a constrained set (e.g. p-dimensional box), see Parikh and Boyd (2013);

with the proposed implementation following a normalization step, as discussed below.

7.2.1 Dual problem normalization

We start from a Cholesky decomposition of the covariance matrix for the ambiguities and the real-valued parameters, with the Cholesky factor $C \in \mathbb{R}^{(n+p)\times(n+p)}$ given as

$$C = \begin{bmatrix} C_{\hat{a}} & 0 \\ C_Q & C_{\hat{b}} \end{bmatrix} \quad \overset{Q \triangleq C^T C}{\rightarrow} \quad \begin{bmatrix} Q_{\hat{a}\hat{a}} & Q_{\hat{a}\hat{b}} \\ Q_{\hat{b}\hat{a}} & Q_{\hat{b}\hat{b}} \end{bmatrix} \triangleq \begin{bmatrix} C_{\hat{a}}^T C_{\hat{a}} + C_Q^T C_Q & C_Q^T C_{\hat{b}} \\ C_{\hat{b}}^T C_Q & C_{\hat{b}}^T C_{\hat{b}} \end{bmatrix} \tag{7.1}$$

where all sub-matrices are conformable in size, with $C_{\hat{a}} \in \mathbb{R}^{n\times n}$ and $C_{\hat{b}} \in \mathbb{R}^{p\times p}$ being lower triangular, while $C_Q \in \mathbb{R}^{p\times n}$ is a rectangular matrix. The Cholesky decomposition of the conditional ambiguity covariance matrix follows simply as $Q_{\hat{a}(b)} = C_{\hat{a}}^T C_{\hat{a}}$, and it can be noticed that $Q_{\hat{a}\hat{b}} Q_{\hat{b}\hat{b}}^{-1} \equiv C_Q^T C_{\hat{b}}^{-T}$.

Differently from the ambiguities, where an admissible transformation is needed to preserve their integerness (Teunissen, 1994), for the real-valued parameters $b \in \mathbb{R}^p$ without additional constraints, we can consider any linear transformation $\mathcal{T}: \mathbb{R}^p \rightarrow \mathbb{R}^p$ as long as $\mathcal{T}$ is invertible. Therefore, we make use of $\delta = C_{\hat{b}}^{-T}\left(b - \hat{b}\right) \in \mathbb{R}^p$, so that $\hat{\delta} = C_{\hat{b}}^{-T}\left(\hat{b} - \hat{b}\right) = 0$, and the transformed parameters will have a *standard* multivariate normal distribution, i.e. $\mathcal{N}_p\left(E\left\{\delta\right\}, D\left\{\delta\right\}\right)$ given $E\left\{\delta\right\} = 0$ and $D\left\{\delta\right\} = I_p$. Furthermore, we have $Q_{\hat{\delta}\hat{a}} = C_Q$, while the conditional ambiguity covariance matrix $Q_{\hat{a}(\delta)} = C_{\hat{a}}^T C_{\hat{a}}$ will not be affected by this linear transformation.

The 'normalized' function $\bar{\mathcal{F}}_2 : \mathbb{Z}^n \times \mathbb{R}^p \rightarrow \mathbb{R}$ is expressed here as

$$\bar{\mathcal{F}}_2(a, \delta) \overset{\text{def}}{=\!=} \|\delta\|_2^2 + \|\hat{a}(\delta) - a\|_{Q_{\hat{a}(\delta)}}^2, \quad \hat{a}(\delta) = \hat{a} + C_Q^T \delta \tag{7.2}$$

where $a \in \mathbb{Z}^n$ and $\delta \in \mathbb{R}^p$, given the 'normalized' dual objective function

$$\bar{\mathcal{D}}(\delta) = \min_{a \in \mathbb{Z}^n} \bar{\mathcal{F}}_2(a, \delta) = \bar{\mathcal{F}}_2\left(\check{a}(\delta), \delta\right) = \|\delta\|_2^2 + \|\hat{a}(\delta) - \check{a}(\delta)\|_{Q_{\hat{a}(\delta)}}^2 \tag{7.3}$$

with the dual mixed ILS solution found as

$$\check{\delta} = \arg\min_{\delta \in \mathbb{R}^p} \bar{\mathcal{D}}(\delta) = \arg\min_{\delta \in \mathbb{R}^p} \left(\|\delta\|_2^2 + \|\check{\epsilon}(\delta)\|_{Q_{\hat{a}(\delta)}}^2 \right) \tag{7.4}$$

given the (conditioned) ambiguity residuals $\check{\epsilon}(\delta) = \hat{a}(\delta) - \check{a}(\delta)$, for

$$\check{a}(\delta) = \arg\min_{a \in \mathbb{Z}^n} \|\hat{a}(\delta) - a\|_{Q_{\hat{a}(\delta)}}^2 \tag{7.5}$$

and once the global solution $\check{\delta} \in \mathbb{R}^p$ has been found, we compute the real-valued ambiguity-fixed vector in its original parametrization simply as $\check{b} = \hat{b} + C_{\hat{b}}^T \check{\delta}$.

7.2.2 Algorithm implementation aspects

We consider the dual weighting approximation $Q_{\hat{a}(\delta)}^{\circ}$ described by Teunissen and Massarweh (2024), so we neglect the off-diagonal terms of $Q_{\hat{a}(\delta)}$, such that

$$Q_{\hat{a}(\delta)}^{\circ} = \operatorname{diag}\left(\sigma_{\hat{a}_1(\delta)}^2, \ldots, \sigma_{\hat{a}_n(\delta)}^2 \right) \tag{7.6}$$

and the approximate (normalized) dual objective function is

$$\bar{\mathcal{D}}^{\circ}(\delta) = \|\delta\|_2^2 + \bar{\mathcal{G}}(\delta) \tag{7.7}$$

with the (non-negative) second periodic-like term $\bar{\mathcal{G}} : \mathbb{R}^p \to \mathbb{R}$ given as

$$\bar{\mathcal{G}}(\delta) = \sum_{i=1}^{i=n} \frac{g\left(x_i(\delta)\right)}{\sigma_{\hat{a}_i(\delta)}^2}, \quad g(x) = \left(x - \lceil x \rfloor\ \right)^2, \quad x_i(\delta) \equiv \hat{a}_i(\delta) = \hat{a}_i + c_{Q,i}^T \delta \tag{7.8}$$

where $c_{Q,i} \in \mathbb{R}^p$ refers to the i-th column for $i = 1, \ldots, n$, of matrix $C_Q \equiv Q_{\hat{\delta}\hat{a}} \in \mathbb{R}^{p \times n}$, while $\lceil \cdot \rfloor$ is the rounding operator. At the problem initialization, we might consider a guess solution $\delta_0 \overset{\text{def}}{=\!=} \hat{\delta} \equiv 0$ (float solution) to bound the dual objective function as

$$\bar{\mathcal{D}}^{\circ}(0) = \bar{\mathcal{G}}(0) = \sum_{i=1}^{i=n} \frac{g\left(\hat{a}_i\right)}{\sigma_{\hat{a}_i(\delta)}^2} \leq \sum_{i=1}^{i=n} \frac{0.25}{\sigma_{\hat{a}_i(\delta)}^2} = \frac{\operatorname{tr}\left({Q_{\hat{a}(\delta)}^{\circ}}^{-1} \right)}{4} \tag{7.9}$$

where $g(x) \leq 0.25, \forall x \in \mathbb{R}$, while $tr(Q)$ refers to the trace of matrix Q.

However, this upper bound will be too large for small (conditional) variances $\sigma^2_{\hat{a}_i(\delta)}$ and for large n values. Instead, in the dual initialization we consider a local minimizer given $\hat{a}\,(\delta_0) = \hat{a} \in \mathcal{S}_u$, where $u = \lceil \hat{a} \rfloor \ \in \mathbb{Z}^n$ is the integer rounding solution. Using Lemma 7 from (ibid), we can therefore define an initial search radius R_0, such as

$$R_0 \overset{\text{def}}{=\!=} \sqrt{\mathcal{P}^\circ\left(\lceil \hat{a} \rfloor\right)} = \|\hat{a} - \lceil \hat{a} \rfloor\|_{Q^\circ_{\hat{a}\hat{a}}} \tag{7.10}$$

so making use of the approximate inverse-weighting matrix $Q^\circ_{\hat{a}\hat{a}}$. Note that R_0 is defined here as 'radius' since we seek new solutions δ_j for $j > 0$ that further minimize $\bar{\mathcal{D}}^\circ$, i.e.

$$\|\delta_j\|_2^2 \leq \bar{\mathcal{D}}^\circ\left(\delta_j\right) \leq \bar{\mathcal{D}}^\circ\left(\delta_0\right) \tag{7.11}$$

where the first inequality follows by the definition given in Eq. 7.7. Therefore, starting with an initial guess $\delta_0 = 0$, we are able to constrain the feasible domain in $\Theta_{R_0} \subset \mathbb{R}^p$, with

$$\Theta_{R_0} = \{\delta \in \mathbb{R}^p \mid \|\delta\|_2 \leq R_0\} \tag{7.12}$$

that represents a p-ball of radius R_0. In BINGO, as well as other branch-and-bound methods, it is common to set the initial search domain using a p-dimensional box, namely

$$\mathcal{C} = \{\delta \in \mathbb{R}^p | -R_0 \leq \delta_j \leq R_0,\ j = 1, \ldots, p\} \tag{7.13}$$

that will later be partitioned during the BB k-iterations into smaller boxes, e.g.

$$\mathcal{B}_k = \{\delta \in \mathbb{R}^p | \delta_{j,L} \leq \delta_j \leq \delta_{j,U}, j = 1, \ldots, p\}, \quad k > 0 \tag{7.14}$$

where $\delta_{j,L}$ and $\delta_{j,U}$ are the extrema of the current partitioned box along each j-th component, and they relate to the ambiguity intervals $l_i \leq \hat{a}_i(\delta) \leq u_i$ given $i = 1, \ldots, n$. For additional information, we refer to Eq.(74) in (ibid), where $m_{i\alpha}$ is equivalent here to entries $\left(Q_{\hat{a}\hat{\delta}}\right)_{i\alpha}$ based on the indexing notation $i \in [1, n]$ and $\alpha \in [1, p]$.

Given any box $\mathcal{B}$, the convex lower bounding function $\bar{\mathcal{G}}_L(\delta) \leq \bar{\mathcal{G}}(\delta), \forall \delta \in \mathcal{B}$ can be found considering $g_{i,L} \overset{\text{def}}{=\!=} g_L\left(x_i\right) \leq g\left(x_i\right)$ for $i = 1, \ldots, n$. A definition of these functions $g_{i,L}$ is summarized in Table 7.1, with the two auxiliary functions 'funL' and 'funU' respectively described in Table 7.2 and Table 7.3. The derivatives $g'_{i,L} = \frac{dg_{i,L}(x_i)}{dx_i}$ are not shown since they can easily be computed for these polynomial functions. Note that the implementation in Table 7.1 has been simplified here after re-defining the followings: $l_i := l_i - z_L$, $x_i := x_i - z_L$, $u_i := u_i - z_L$ and $z_U := \lfloor u_i \rfloor - z_L$, given a common integer offset $z_L := \lfloor l_i \rfloor$ with $\lfloor \cdot \rfloor$ referring to the floor function. Hence, we have $z_U = 0$ if $u_i \in (l_i, 1)$, otherwise $z_U > 0$.

The branch-and-bound iterations involve computing both upper bound and convex lower bound for each box, thus partitioning the initial box $\mathcal{C} \subset \mathbb{R}^p$ in smaller boxes. Given a box $\mathcal{B}$, the computation of upper bounds $U\left(\mathcal{B}\right)$ is possible by evaluating

Table 7.1: An overview is given of the Convex Lower Bound (CLB) term $g_{i,L}$, computed after re-defining $l_i := l_i - z_L, x_i := x_i - z_L, u_i := u_i - z_L$ and $z_U := \lfloor u_i \rfloor - z_L$, for $z_L := \lfloor l_i \rfloor$.

CASE	**Condition**	**Sub-Condition**	**CLB function**
$z_U = 0$	$l_i < 0.5, u_i \leq 0.5$	-	$g_{i,L} = x_i^2$
	$l_i \geq 0.5, u_i > 0.5$	-	$g_{i,L} = (x_i - 1)^2$
	$l_i < 0.5, u_i > 0.5$	$\lvert l_i \rvert > \lvert u_i - 1 \rvert$	$g_{i,L} = \text{funL}\,(l_i, u_i, x_i)$
		$\lvert l_i \rvert < \lvert u_i - 1 \rvert$	$g_{i,L} = \text{funU}\,(l_i, u_i, x_i)$
		$\lvert l_i \rvert = \lvert u_i - 1 \rvert$	$g_{i,L} = l_i^2$
$z_U \neq 0$	$x_i \geq 1, x_i \leq z_U$	-	$g_{i,L} = 0$
	$x_i < 1$	$l_i < 0.5$	$g_{i,L} = \text{funL}\,(l_i, 1, x_i)$
		$l_i \geq 0.5$	$g_{i,L} = (x_i - 1)^2$
	$x_i > z_U$	$u_i > z_U + 0.5$	$g_{i,L} = \text{funU}\,(0, u_i - z_U, x_i - z_U)$
		$u_i \leq z_U + 0.5$	$g_{i,L} = (x_i - z_U)^2$

Table 7.2: Auxiliary function 'funL' adopted in Table 7.1

$\text{funL}(l_i, u_i, x_i)$	
▶ Compute $\alpha_i = l_i + \sqrt{1 - 2l_i}$, while note that $x_i \in [l_i, u_i]$	
Condition	**Function**
$u_i < \alpha_i$	$g_{iL} = l_i^2 + (x_i - l_i)\frac{(u_i-1)^2 - l_i^2}{u_i - l_i}$
$x_i \leq \alpha_i$	$g_{iL} = 2(\alpha_i - 1)(x_i - \alpha_i) + (\alpha_i - 1)^2$
$x_i \geq \alpha_i$	$g_{iL} = (x_i - 1)^2$

Table 7.3: Auxiliary function 'funU' adopted in Table 7.1

$\text{funU}(l_i, u_i, x_i)$	
▶ Compute $\beta_i = u_i - \sqrt{2u_i - 1}$, while note that $x_i \in [l_i, u_i]$	
Condition	**Function**
$l_i > \beta_i$	$g_{i,L} = l_i^2 + (x_i - l_i)\,\frac{(u_i-1)^2 - l_i^2}{u_i - l_i}$
$x_i \leq \beta_i$	$g_{i,L} = x_i^2$
$x_i \geq \beta_i$	$g_{i,L} = 2\beta_i\,(x_i - \beta_i) + \beta_i^2$

any feasible solution within this box domain, while for the lower bounds $L(\mathcal{B})$ the minimization takes place via a convex relaxation and projected gradient descent, as discussed later. At the k-th iteration, the global upper bound U_k and the global lower bound L_k are eventually found by

$$L_k = \min_{q=1,\ldots,k} L(\mathcal{B}_q), \quad U_k = \min_{q=1,\ldots,k} U(\mathcal{B}_q) \tag{7.15}$$

and the BINGO algorithm converges if $U_k - L_k < \epsilon_{BB}$.

Note that a small tolerance $\epsilon_{BB} \ll 1$ is provided by the user as convergence criterion, but in addition we can also set a maximum number of iterations K_{MAX} to limit excessively long computational times. Still, in this case, the global optimal solution for the approximate dual problem is not guaranteed.

7.2.3 Analysis of computational trade-offs

We can investigate now the computational performance, thus looking at different trade-offs for the proposed implementation. Hence, we examine the following aspects:

I. Problem normalization (Section 7.2.3.1)

II. Problem initialization (Section 7.2.3.2)

III. Boxes' splitting strategy (Section 7.2.3.3)

IV. Convex lower bound minimization (Section 7.2.3.4)

V. The ϵ_{BB}-termination criterion (Section 7.2.3.5)

7.2.3.1 Benefits of normalization

The problem normalization in Section 7.2.1 is convenient since with each function evaluation we compute $\|\delta\|_2^2 \equiv \delta^T\delta$ instead of $\left\|b-\hat{b}\right\|^2_{Q_{\hat{b}\hat{b}}} \equiv \left(b-\hat{b}\right)^T Q^{-1}_{\hat{b}\hat{b}}\left(b-\hat{b}\right)$, so avoiding several matrix operations, while using $\hat{a}(\delta)-\hat{a} = C_Q^T\delta$ for the conditioned ambiguities rather than $\hat{a}(b)-\hat{a} = Q_{\hat{a}\hat{b}}Q^{-1}_{\hat{b}\hat{b}}\left(b-\hat{b}\right)$. The global minimum is invariant to a re-parametrization of the b parameters, however as boxes' definition and splitting change, then the computational steps might slightly differ following a linear transformation $\mathcal{T}:\mathbb{R}^p \to \mathbb{R}^p$.

In Fig. 7.1, we compare the original (left) and normalized dual problem (right) for the Example 3 in Fig. 12 from (ibid), based $p = 2$ and $n = 30$, while using a different float sample. The initial feasible domain, where we expect to find the global minimum, is bounded by a p-dimensional hypersphere or hyperellipsoid, respectively when using the normalized or the original formulation (see magenta dashed line). Hence, the BINGO initial search domain for the normalized case is always reduced to a p-dimensional hypercube (right plot).

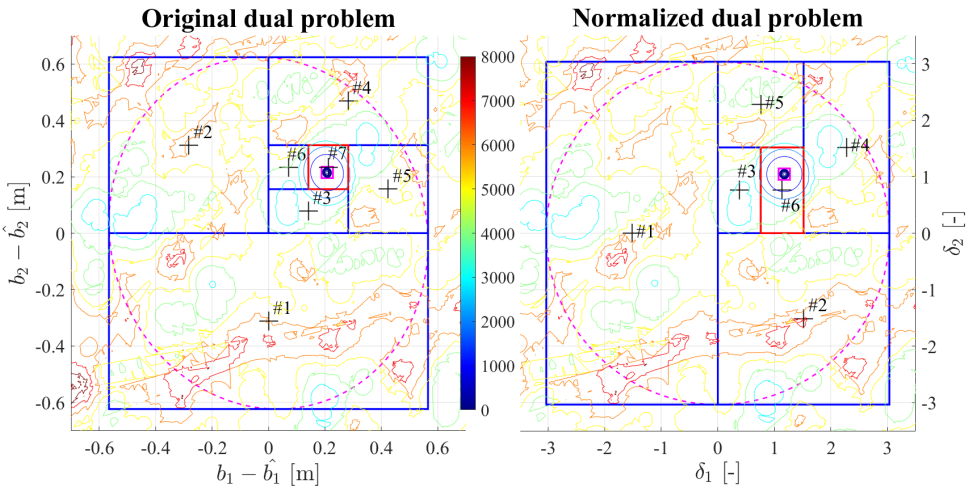


Figure 7.1: The comparison between original (left) and normalized (right) formulations of the dual problem is shown based on Example 3 in Fig. 12 from (ibid). In red, the box relative to the optimal solution found with BINGO method, along with previously partitioned boxes in blue and the initial ellipsoid in magenta dashed line. See text for more information, and note that the BB convergence criterion is based here on a small tolerance $\epsilon_{BB} = 10^{-6}$.

In this example, the initial ellipse (or sphere) is based on a statistical criterion, where $\left\|b - \hat{b}\right\|^2_{Q_{\hat{b}\hat{b}}} \sim \chi^2(p)$, following a Chi-squared central distribution with $p = 2$ degrees of freedom. Hence, we can define a 99.0% confidence level, leading to $R_0 \simeq 3.035$ and the smallest axis-aligned initial bounding box in the original or the normalized case is given as $\left\{\left|b_1 - \hat{b}_1\right| \leq R_0\ \sigma_{\hat{b}_1} \cong 0.565, \left|b_2 - \hat{b}_2\right| \leq R_0\ \sigma_{\hat{b}_2} \cong 0.624\right\}$ or as $\{|\delta_1| \leq R_0\ , |\delta_2| \leq R_0\}$, for $\sigma_{\hat{b}_1} \cong 0.186$ and $\sigma_{\hat{b}_2} \cong 0.206$.

7.2.3.2 Initialization strategies

The initialization is a key aspect for the performance in the BINGO solution, where we could consider four different strategies:

a) **Statistical criterion**, where $R_0 = \chi_\alpha(p)$ for a given confidence level, e.g., $\alpha = 99\%$, as adopted in the previous example of Fig. 7.1;

b) **Dual initial guess**, where $R_0 = \|\hat{a} - \lceil\ \hat{a}\rfloor\ \|_{Q_{\hat{a}(b)}}$, using the weighting matrix $Q^{-1}_{\hat{a}(b)}$ and given $\hat{a}\,(b_0) \equiv \hat{a}$, since we have assumed $b_0 = \hat{b}$;

c) **Primal initial guess**, where $R_0 = \|\hat{a} - \lceil\ \hat{a}\rfloor\ \|_{Q_{\hat{a}\hat{a}}}$, thus using the matrix $Q^{-1}_{\hat{a}\hat{a}}$ instead;

d) **Constrained set**, where each j-th component of real-valued parameters is constrained between some known *a priori* values b_j^{MIN} and b_j^{MAX}.

In (a), we expect the true parameter vector b to be found in a neighborhood of $\hat{b}$, which is distributed as $\mathcal{N}\left(b, Q_{\hat{b}\hat{b}}\right)$. However, given the true integer vector $a \in \mathbb{Z}^n$, we have $\hat{b}(a) \sim \mathcal{N}\left(b, Q_{\hat{b}(a)}\right)$, which is also expected to be found in a neighborhood of b. However, notice that $\hat{b}(a)$ might not belong to the search domain initially set by $R_0 = \chi_\alpha(p)$. On the other hand, (b) and (c) are similar, starting with the integer ambiguity vector $\breve{a} = \lceil \hat{a} \rfloor$, which ensures that the global solution is found. In this case, a primal initial guess is preferred since $Q^{-1}_{\hat{a}\hat{a}} \leq Q^{-1}_{\hat{a}(b)}$, and therefore we start with a much smaller radius R_0. Lastly, with strategy (d), we are making use of additional information, e.g. given a specific application in use, de facto constraining $b \in \Omega \subset \mathbb{R}^p$, but this is not further discussed here since we will mainly focus on 'unconstrained dual' problems.

We consider now the example in Fig. 7.1 for the normalized dual problem, and we assume different R_0 values. The number of boxes partitioned in BINGO for a different initial search radius $R_0 \in [3, 30]$ is shown on the left plot in Fig. 7.2, while on the right we provide an illustration of the search domain for the case $R_0 = 30$, along with the contour values of the (normalized) dual cost function, with a global solution achieved in 23 iterations. We observe that a larger initial search radius leads to a higher number of partitioned boxes, but this relation is however not a monotonically increasing function and it also depends upon other aspects, such as the boxes' splitting strategy, as discussed next.

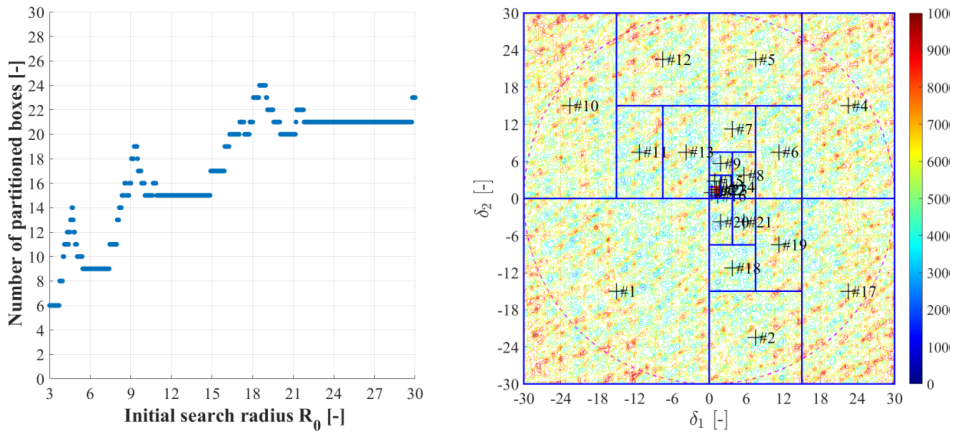


Figure 7.2: On the left plot, we show the number of partitioned boxes (or BB iterations) for the normalized dual problem of Fig. 7.1, while varying the initial search radius R_0 between 3 and 30. On the right plot, we show the dual cost function starting with $R_0 = 30$ (see magenta dashed line), along with the 23 iterations needed before the algorithm converged.

7.2.3.3 Boxes' splitting strategy

So far, the splitting of boxes was based on halving the longest edge for the most promising box, i.e. the one having the lowest global CLB value. In this way, we were

able to control the 'condition number' of the boxes during their partitioning (leading to 'children' boxes or sub-regions), which is defined by Boyd and Mattingley (2007) as follows

$$cond\,(\mathcal{B}) = \frac{\nu_{\max}}{\nu_{\min}} \overset{\text{def}}{=\!=} \frac{\max_{j=1\ldots p}\,(\delta_{j,U} - \delta_{j,L})}{\min_{j=1\ldots p}\,(\delta_{j,U} - \delta_{j,L})} \leq \max\{cond\,(\mathcal{C})\,, 2\} \tag{7.16}$$

for any box $\mathcal{B} \in \mathcal{L}_k$ (in the current list), where $\mathcal{C}$ is the initial search box. Hence, during each branching step, condition number of children boxes is not larger than the maximum between 2 and the condition number of $\mathcal{C}$. In this way, (ibid) is able to prove convergence assuming that bounds become sharper as boxes shrink, while upper and lower bounds should be cheap to compute for making the BB approach computationally efficient.

We investigate now if an uneven splitting of the boxes is convenient, thus taking the longest edge $\nu_{\max}$ of a certain box, thus dividing it into $\mu_{BB}\nu_{\max}$ and $(1-\mu_{BB})\,\nu_{\max}$ for the splitting ratio $\mu_{BB} \in (0,1)$. In Fig. 7.3, we take the numerical example from Fig. 7.1 (right) starting with $R_0 = 10$ and varying $\mu_{BB} = 0.001, \ldots, 0.999$, so computing the total number of partitioned boxes needed for each splitting ratio value (see left plot).

For the boundary cases, i.e. $\mu_{BB} \to 0$ and $\mu_{BB} \to 1$, the number of splits increases up to a maximum set here to $K_{MAX} = 100$, while the nominal case $\mu_{BB} = 0.5$ is illustrated on the right side with all iterations. The use of splitting ratios μ_{BB} close to 0 or 1 is certainly not suggested since too elongated boxes will be created during the BB iterations, and the search for a global solution will require more iterations given the same initial search domain. The algorithm partitioning steps are not perfectly symmetric for μ_{BB} and $1-\mu_{BB}$, and in general, the exact number of iterations will depend upon several factors, e.g. the initial search radius or dual function topology varying based on the float ambiguity solution. In the reminder of this work we will keep using $\mu_{BB} = 0.5$, while we briefly discuss now some potential alternatives to bisection.

Alternatives to Bisection

The previous results are derived from splitting the most promising box in two parts along its longest edge. However, several BB alternatives exist and different partitioning strategies can be considered, e.g. adopted in the 'DIRECT method' for black-box global optimization, as introduced by Jones et al. (1993). A discussion of the 'DIRECT method' is outside the scope of this work since it aims at minimizing black-box objective functions, with little information available, see Jones and Martins (2021), i.e. differently from our dual mixed-ILS problems.

At this point, it is possible to consider trisections instead of bisections, splitting the longest edge into three parts instead of two. An illustrative sketch is presented for $p = 2$ in Fig. 7.4, where the bisection (left) and trisection (right) approaches are shown starting with an initial radius R_0 (in magenta), so that the initial domain is defined as $\mathcal{C} = \{x \in \mathbb{R}^p : |x_i| \leq R_0, \;\; i = 1, \ldots, p\}$, with each box in $\mathbb{R}^p$ given by

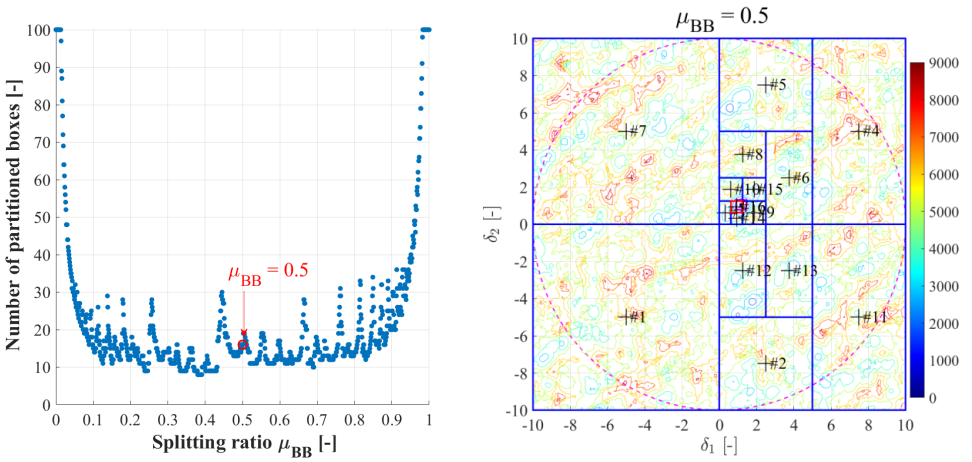


Figure 7.3: The number of partitioned boxes is shown on the left with respect to the splitting ratio $\mu_{BB} \in [0, 1]$ for the problem of Fig. 7.1 while using an initial search radius $R_0 = 10$. On the right side, we show the BB iterations for the specific case $\mu_{BB} = 0.5$, given $\epsilon_{BB} = 10^{-6}$.

$\mathcal{B}_{q_1 q_2 \dots q_k}$ for indexes q_s referring to the q_s-th partition (or children box) at the s-th iteration. A different number of iterations is considered for the bisection and trisection cases, respectively leading to a volume for the optimal box (in green color) given by $vol\left(\mathcal{B}_{21111}\right) = vol\left(\mathcal{C}\right)/32$ and $vol\left(\mathcal{B}_{213}\right) = vol\left(\mathcal{C}\right)/27$.

Both bisection and trisection strategies can be understood as *binary* and *ternary* search trees, as illustrated in Fig. 7.5, with the most promising box (marked by a red asterisk) being split at each BB iteration. This interpretation allows extending the splitting strategy to an arbitrary number of children boxes starting from the most promising one, but one should carefully consider that a higher number of evaluations (per generation) will take place.

The branching strategy can be generalized to $\mathcal{B}_{q_1 q_2 \dots q_k}$ with the index q_s defining the q_s-th partition (or children box) at the s-th iteration. This can be varied during iterations and it opens a vast landscape of possibilities for BB methods, where also the parallelization of the children boxes' evaluation is possible. Still, from our experience, higher-order splitting approaches are not suggested as several evaluations will then occur in regions of the search domain most likely less promising for finding the global optimum.

Overall, the splitting strategy can be regarded as a trade-off between *width* and *depth* of the search tree associated to branching iterations (Morrison et al., 2016). However, the algorithm's complexity is not solely related to the number of boxes evaluated, but also to the computational effort required for computing the bounds (Boyd and Mattingley, 2007), as discussed in the next section. For sake of simplicity, in the reminder of this work we proceed by considering only the bisection strategy.

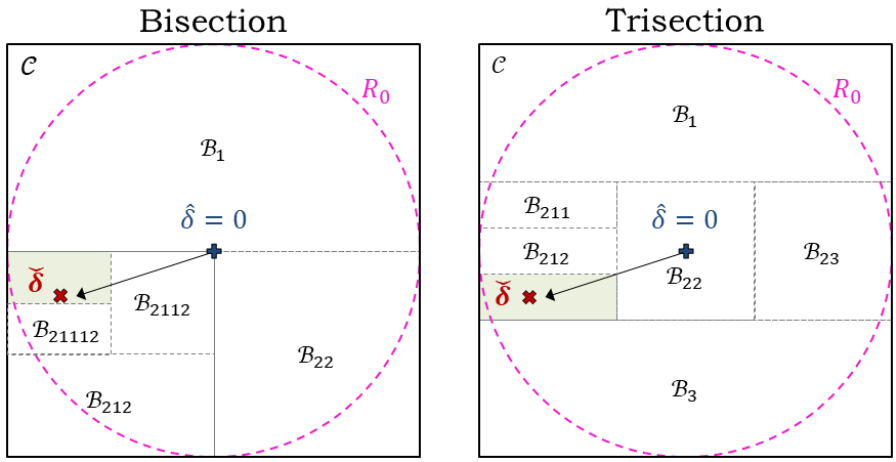


Figure 7.4: An illustrative comparison between bisection and trisection splitting approaches is presented for a 2-dimensional case in normalized coordinates, where $\hat{\delta} = 0$ (float solution, in blue) and $\check{\delta} \neq 0$ (fixed solution, in red). The partitioned boxes $\mathcal{B}_{q_1 q_2 \dots q_k}$ are defined based on indexes q_s referring to the q_s-th partition (or children box) at the s-th iteration.

7.2.3.4 Convex lower bound minimization

At each box evaluation, given $\mathcal{B} = \{\delta \in \mathbb{R}^p \mid \delta_{j,L} \leq \delta_j \leq \delta_{j,U}, j = 1, \dots, p\}$, the upper bound can be computed starting with the initial guess, e.g., $\delta_C = (\delta_{1,C}, \dots, \delta_{p,C})^T \in \mathbb{R}^p$, which refers to the center of the box given $\delta_{j,C} = 0.5\,(\delta_{j,U} + \delta_{j,L})$ for $j = 1, \dots, p$. This initial guess leads to a conditioned (float) ambiguity $\hat{a}_C = \hat{a}\,(\delta_C)$, thence rounded to $\check{a}_C = \lceil \hat{a}_C \rfloor \in \mathbb{Z}^n$, so we obtain a *local* minimizer $\delta^{UB} = \hat{\delta}\,(\check{a}_C)$ of $\bar{\mathcal{D}}^\circ$. This local minimum will represent an upper bound of the dual function inside $\mathcal{B}$. As the box shrinks, it might become the only minimum in $\mathcal{B}$, at which point there is no reason to further split this box.

On the other hand, the computation for the CLB minimization is less trivial and represents a major processing burden for the BINGO algorithm. The CLB definition has been provided in Table 7.1, see description in Section 7.2.2. In this section, we will investigate a few different computational aspects related to the CLB minimization, therefore briefly discussing

(1) the projection operation;

(2) the selection of step length;

(3) the selection of step direction;

based on the projected gradient descent method, hereinafter PGD; see Schmidt (2020) and Andersen (2024) for additional information.

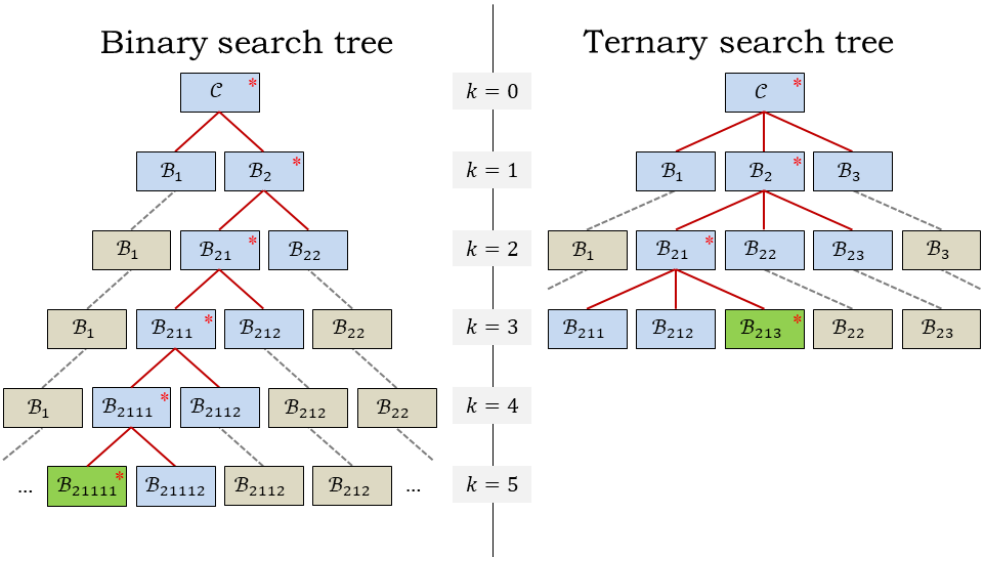


Figure 7.5: The search tree presentation of Fig. 7.4 is shown here for the bisection and trisection strategies, respectively defined by a binary (left side) and a ternary (right side) search tree. In green color, the box where the global minimum lies, while the red asterisk refers to the most promising box (at each k-th iteration) for the generation of children boxes (in blue), which will be evaluated. In this illustration, we require 11 (left) versus 10 (right) evaluations.

(1) *Projection inside the loop*

The PGD algorithmic steps described in Teunissen and Massarweh (2024) start by a feasible solution $\delta^{q=0} \in \mathcal{B}$, therefore looping q with an alternation between the *gradient descent* and *projection* steps. Firstly, we compute $\hat{\delta}^k = \delta^k - \mu_k \nabla \bar{\mathcal{D}}_L^\circ \left(\delta^k\right)$ given a step-size $\mu_k > 0$ such that $\bar{\mathcal{D}}_L^\circ \left(\hat{\delta}^k\right) \leq \bar{\mathcal{D}}_L^\circ \left(\delta^k\right)$, then we ensure conformity with the box constraints by a projection $\delta^{k+1} = P_{\mathcal{B}} \left(\hat{\delta}^k\right)$ where for each j-th component the projection reduces to the median value of each triplet $\left\{\delta_{j,L}, \hat{\delta}_j^k, \ \delta_{j,U}\right\}$ due to our box geometry. This operation is computationally cheap, whereas a large number of PGD iterations $k \gg 0$ might still be needed before we converge to the minimum of the CLB function (i.e. needed at each new box evaluation).

However, given the step-size μ_k for $\bar{\mathcal{D}}_L^\circ \left(\hat{\delta}^k\right) \leq \bar{\mathcal{D}}_L^\circ \left(\delta^k\right)$, after the projection step we cannot assure $\bar{\mathcal{D}}_L^\circ \left(\delta^{k+1}\right) \leq \bar{\mathcal{D}}_L^\circ \left(\delta^k\right)$, and the PGD might take a long time to converge to the CLB minimizer $\delta^{LB} \in \mathcal{B}$ with $\bar{\mathcal{D}}_L^* = \bar{\mathcal{D}}_L^\circ \left(\delta^{LB}\right)$. Therefore, we can make use of a different approach, as suggested in Section 11.2.3.2 by Schmidt et al. (2011), known as "backtracking along the projection arc", see Bertsekas (1999). Given that our projection operator $P_{\mathcal{B}}$ is very cost efficient, we can conduct a line search for μ_k, such that

$$\bar{\mathcal{D}}_L^{\circ}\left(\delta^{k+1}\right) \leq \bar{\mathcal{D}}_L^{\circ}\left(\delta^{k}\right), \quad \delta^{k+1} = P_{\mathcal{B}}\left(\hat{\delta}^{k} \overset{\text{def}}{=\!=} \delta^{k} - \mu_k \nabla \bar{\mathcal{D}}_L^{\circ}\left(\delta^{k}\right)\right) \tag{7.17}$$

so directly searching along the projection arc rather than in a descent direction. Ultimately, the PGD algorithm stops when $\left\|\delta^{k+1} - \delta^{k}\right\|_2 < \gamma_{PGD}$ for a small positive value γ_{PGD}.

An example is shown in Fig. 7.6, where we compare the backtracking along the descent direction (left) and along the projection arc (right). In the first approach, we descent outside the blue box with $\hat{\delta}^k$ (red circles), then projected back onto the constraints as δ^{k+1} (black dots). Still, after five iterations we see that $\delta^5 \equiv \delta^1$ and the PGD is therefore *trapped* in a configuration that will never converge, except if changing adaptively the step-size (i.e. rather not trivial). In the proposed implementation, the backtracking occurs along projection arcs, so when reaching an edge of the box, the search continues directly along this direction, and convergence is assured. For instance, the point δ^4 in Fig. 7.6 diverges for the 'projection outside the loop' approach, while converges towards the CLB minimum, $\delta^{LB} \in \mathcal{B}$ shown as magenta square, when making use of the 'projection inside the loop'.

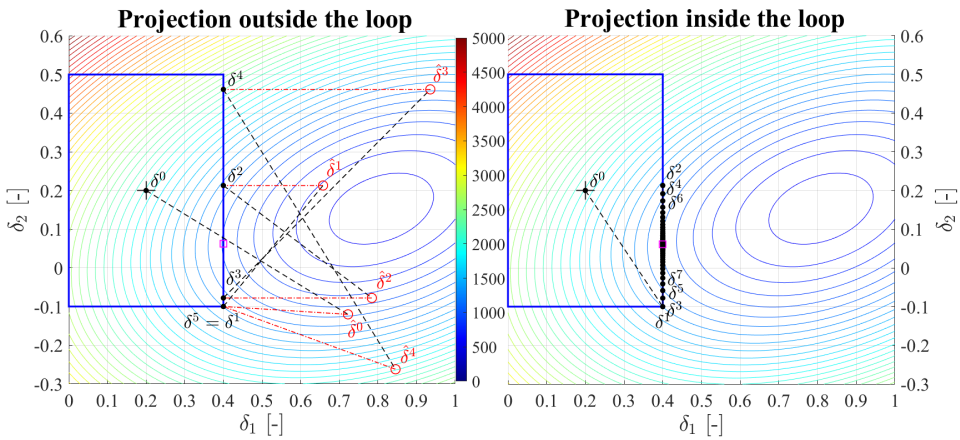


Figure 7.6: The iterations of Project Gradient Descent are shown for two cases: with projection outside (left) or inside (right) the loop for the line search. See text for more details.

As initial value in the box $\mathcal{B}$, we can consider $\delta^0 \overset{\text{def}}{=\!=} P_{\mathcal{B}}\left(\delta^{UB}\right)$, so starting with the solution from the upper bound computations, then projected into $\mathcal{B}$. The PGD computations described in this section are summarized in Algorithm 2, while we continue by investigating the PGD in relation to both 'step length' and 'descent direction'.

Algorithm 2 Summary of the 'PGD' algorithm

INPUTS:

$$\mathcal{B} \in \mathbb{R}^p, \quad \gamma_{PGD} \ll 1, \quad k_{MAX} \in \mathbb{Z}$$

INITIALIZATION:

Initialize $\delta^0 \in \mathcal{B}$, e.g., using the vector δ^{UB} obtained in the upper bound computations, then projected into $\mathcal{B}$.

PROJECTED GRADIENT DESCENT:

Set $\delta^* = \delta^0$ and $\bar{\mathcal{D}}^* = \bar{\mathcal{D}}_L^\circ\left(\delta^0\right)$;

% PGD iterations
for $k = 1, \ldots, k_{MAX}$
    Define $\delta^k = \delta^{k-1}$;

    % Gradient descent with projection step
    Compute $\bar{\mathcal{D}}_L^\circ\left(\delta^k\right)$ and $\nabla\bar{\mathcal{D}}_L^\circ\left(\delta^k\right)$, being respectively the approximate dual function and its gradient;

    % Line search iterations
    for $\mu_k = 2^0, 2^{-1}, 2^{-2}, 2^{-3}, \ldots$
        $\delta^{k+1} = P_{\mathcal{B}}\left(\delta^k - \mu_k \nabla\bar{\mathcal{D}}_L^\circ\left(\delta^k\right)\right)$

        % Check if solution improved
        Compute $\bar{\mathcal{D}}_L^\circ\left(\delta^{k+1}\right)$;
        if $\bar{\mathcal{D}}_L^\circ\left(\delta^{k+1}\right) \leq \bar{\mathcal{D}}_L^\circ\left(\delta^k\right)$
            $\delta^* = \delta^{k+1}$;
            $\bar{\mathcal{D}}^* = \bar{\mathcal{D}}_L^\circ\left(\delta^{k+1}\right)$;
            Break loop; % Inner loop
        end
    end

    % Check if PGD converged
    if $\left\|\delta^{k+1} - \delta^k\right\|_2 < \gamma_{PGD}$
        Break loop; % Outer loop
    end
end

OUTPUTS:

$$\delta^* \in \mathbb{R}^p, \quad \bar{\mathcal{D}}^* \in \mathbb{R}$$

(2) *Selection of the step length*

The step-size $\mu_k > 0$ scales the gradient $\nabla\bar{\mathcal{D}}_L^\circ$ evaluated at δ^k, so leading to the displacement $\delta^k - \hat{\delta}^k = \mu_k \nabla\bar{\mathcal{D}}_L^\circ\left(\delta^k\right)$, as illustrated in Fig. 7.6, where the black dashed

line represents the descent trajectory with $\bar{\mathcal{D}}_L^{\circ}\left(\delta^{k+1}\right) \leq \bar{\mathcal{D}}_L^{\circ}\left(\delta^{k}\right)$ for a suitable value μ_k. We can introduce now the step-length $\eta_k = \mu_k \left\|\nabla\bar{\mathcal{D}}_L^{\circ}\left(\delta^{k}\right)\right\|_2$, such that

$$\delta^{k+1} = P_{\mathcal{B}}\left(\hat{\delta}^k \stackrel{\text{def}}{=\!=} \delta^k - \eta_k \frac{\nabla\bar{\mathcal{D}}_L^{\circ}\left(\delta^{k}\right)}{\left\|\nabla\bar{\mathcal{D}}_L^{\circ}\left(\delta^{k}\right)\right\|_2}\right) \tag{7.18}$$

meaning that $\left\|\hat{\delta}^k - \delta^k\right\|_2 = \eta_k \geq 0$. As shown in Algorithm 2, the CLB evaluation with $\bar{\mathcal{D}}_L^{\circ}$ is needed at each k-th iteration, as well as for each candidate step-size μ_k (or step-length η_k) in the inner loop. That backtracking strategy concerns a search line for a suitable step that is halved till a smaller CLB value is found, where a too small initial step might lead to many PGD iterations (i.e. outer loop), while a too large step might lead to unnecessary evaluations of $\bar{\mathcal{D}}_L^{\circ}$ during the line-search iterations (i.e. inner loop).

We take the example of Fig. 7.6 (right) and examine how the backtracking search operates along the box edge when using the step-size μ_k or the step-length η_k, respectively shown in the left or in the right plot of Fig. 7.7. In the latter case, convergence is achieved in fewer operations since the δ-displacement is not subject to the local gradient value during the descent step that might become too large or too small. Moreover, the backtracking search for $\eta_k = \left\{\eta_k^0, \eta_k^1 = \frac{\eta_k^0}{2^1}, \ldots, \eta_k^q = \frac{\eta_k^0}{2^q}\right\}$ is optimized here since the step-length will decrease once we reach an edge of the box, so we can avoid evaluations for larger values of η_k^0 in the successive PGD iteration.

For instance, we start with $\eta_k^0 \approx \eta_{k-1}^{q_{last}}$, so making use of the previous step length found during the q_{last}-th iteration of the inner loop (see Algorithm 2). We opt here for using $\eta_k^0 = 2\eta_{k-1}^{q_{last}}$ to avoid η_k^0 monotonically decreasing with k, while the length-halving strategy adopted in our algorithms can be regarded as a special case of

$$\eta_k = \left\{\eta_k^0, \eta_k^1 = \tau\eta_k^0, \ldots, \eta_k^q = \tau\eta_k^{q-1}\right\} \tag{7.19}$$

for $\tau = 0.5$, see Armijo (1966). By way of example, considering the left panel of Fig. 7.7 for the step-size μ_k, then we are able to reduce the total number of CLB evaluations (of the inner loop) from 1428 to 300, therefore approximatively from 15 to 3 dual function evaluations per each PGD iteration of the outer loop (here equal to 96).

(3) *Selection of the step direction*

Besides step size or step length, the descent direction is likewise important when minimizing the CLB. A class of 'iterative descent' methods is discussed by Teunissen (1990), where the necessary and sufficient conditions for convergence are given. By *iterative*, we refer here to algorithms that generate a sequence of vectors δ^k for $k \geq 0$, i.e. dependent on previous ones, while by *descent* we refer to the fact that new candidates correspond to decreasing values of the convex lower bound $\bar{\mathcal{D}}_L^{\circ}$.

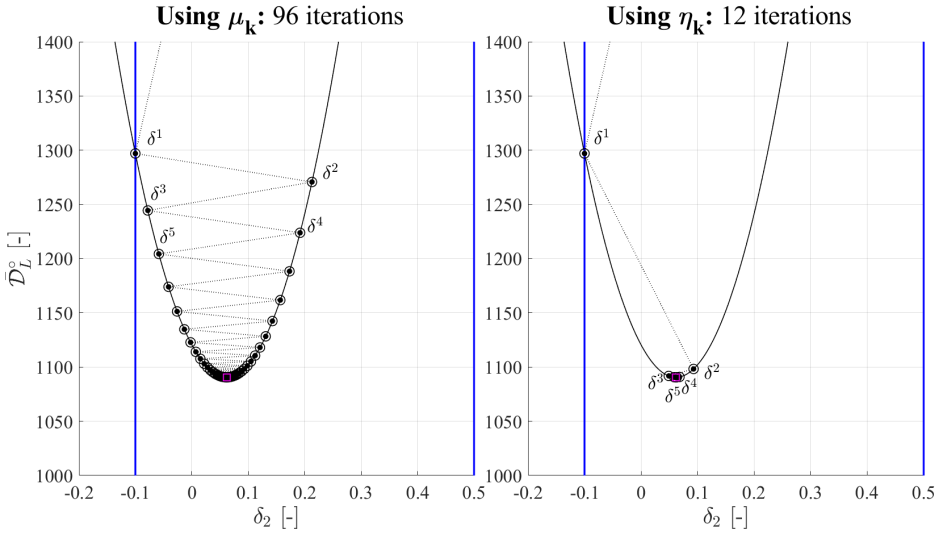


Figure 7.7: The comparison of PGD iterations is shown along the variable δ_2 for $\delta_1 = 0.4$, given the projection-in-the-loop case in the example of Fig. 7.6. On the left and right plots, we show the adoption respectively of step-size μ_k or step-length η_k for minimizing the CLB function $\bar{\mathcal{D}}_L^\circ$, with the minimizer shown as magenta square. See text for more detail.

For instance, in this class, we can consider

- **Steepest Descent method**, see Cauchy (1847), where in a *linear* approximation the direction follows the steepest descent of $\bar{\mathcal{D}}_L^\circ$ at δ^k;
- **Newton's method**, which computes the minimum of a *quadratic approximation* for $\bar{\mathcal{D}}_L^\circ$ at each iteration step, and it has a local quadratic rate of convergence;
- **Trust Region method**, introduced by Levenberg (1944), and modified by Marquardt (1963), i.e. a compromise between the previous two methods.

while – in our experience – we observe that many PGD iterations often occur at the edges of the box currently evaluated. Therefore, the projection-in-the-loop strategy will anyhow alter the descent direction, as already discussed in Fig. 7.6. Notice that if the convex lower bound minimizer is found at one edge of the box, then the gradient might not be zero, nevertheless it is expected to be orthogonal to the constraint contour lines going through that point.

Concerning the PGD convergence criterion, note that this depends upon the user-selected threshold $\gamma_{PGD} > 0$, where we recall $\delta = C_{\hat{b}}^{-T}\left(b - \hat{b}\right)$, such that

$$\left\|\delta^{k+1} - \delta^k\right\|_2 \equiv \left\|b^{k+1} - b^k\right\|_{Q_{\hat{b}\hat{b}}} < \gamma_{PGD} \ll 1 \tag{7.20}$$

and a direct relation follows between the threshold γ_{PGD} and $\Delta b^{k+1} = b^{k+1} - b^k$. In

general, a small value $\gamma_{PGD} = 10^{-4}$ seems to work well in practice, while in the next section we will lastly also discuss a convergence criterion for the BB iterations.

7.2.3.5 The ϵ_{BB}-termination criterion

For the algorithm convergence, we have considered the following stopping criterion

$$U_k - L_k \leq \epsilon_{BB}, \tag{7.21}$$

with the *global* bounds computed in Eq. 7.15, while $U(\mathcal{B})$ and $L(\mathcal{B})$ are the *local* upper bound and lower bound for a certain box $\mathcal{B}$, thus starting with the box $\mathcal{C} = \bigcup_{q=1}^{q=k} \mathcal{B}_q$ at each k-th BB iteration. After defining $g_i(\delta) \equiv g\left(\hat{a}_i(\delta)\right)$ and $g_{i,L}(\delta) \equiv g_L\left(\hat{a}_i(\delta)\right)$ for $i = 1, \ldots, n$, we are able to re-write the algorithm convergence condition such as

$$\left[\left\|\delta^{UB}\right\|_2^2 + \sum_{i=1}^{i=n} \frac{g_i\left(\delta^{UB}\right)}{\sigma^2_{\hat{a}_i(\delta)}}\right] - \left[\left\|\delta^{LB}\right\|_2^2 + \sum_{i=1}^{i=n} \frac{g_{i,L}\left(\delta^{LB}\right)}{\sigma^2_{\hat{a}_i(\delta)}}\right] \leq \epsilon_{BB} \tag{7.22}$$

or equivalently

$$\left(\left\|\delta^{UB}\right\|_2^2 - \left\|\delta^{LB}\right\|_2^2\right) + \sum_{i=1}^{i=n} \frac{g_i\left(\delta^{UB}\right) - g_{i,L}\left(\delta^{LB}\right)}{\sigma^2_{\hat{a}_i(\delta)}} \leq \epsilon_{BB} \tag{7.23}$$

where δ^{UB} and δ^{LB} refer to the solutions respectively obtained from the upper and the lower *global* bound computations. Once the global minimum is found, we assume $\delta^{UB} \equiv \delta^{LB}$.

Following the CLB definition, we will always have $g_i(\delta) - g_{i,L}(\delta) \geq 0, \forall \delta \in \mathcal{B}$, so the second summation term cannot be negative. However, even for the same input $\delta \in \mathcal{B}$, it is not guaranteed that $g_i(\delta) = g_{i,L}(\delta)$ since it will depend on the intervals $[l_i, u_i]$ used to define each convex lower term $g_{i,L}$ for $i = 1, \ldots, n$. Moreover, from our experience, the CLB minimizer δ^{LB} might be affected by small numerical errors, since it is computed via PGD iterations using a threshold condition, e.g., $\gamma_{PGD} = 10^{-4}$. At the same time, a direct relation $\epsilon_{BB} \leftrightarrow \|\Delta\delta\|_2$ for the displacement vector $\Delta\delta \stackrel{\text{def}}{=\!=} \delta^{UB} - \delta^{LB}$ if not straightforward here and, unfortunately, literature on branch-and-bound optimization primarily provides convergence proofs, see Balakrishnan et al. (1991). Still, we know that bounds become sharper as the box shrinks to a point, and in our numerical examples we generally make use of $\epsilon_{BB} = 10^{-6}$, i.e. in most of the cases achieving $U_k - L_k \approx 10^{-12}$ after convergence.

These numerical considerations are certainly dependent on the specific problem at hand, and they should therefore be assessed by BINGO's users based on their individual applications. As discussed in Section 7.2.2, a maximum number of iterations might be set to limit large computational times, but in this case the algorithm's optimality is not guaranteed, and a sub-optimal ambiguity fixing performance may occur. The statistical performance will be further studied in the next section.

7.3 Analysis of dual ambiguity-fixing performance

So far, we have discussed different implementation aspects related to the BINGO algorithm, thus showing important computational trade-offs and describing a few different possible mechanizations. Still, as presented by Teunissen and Massarweh (2024), we make use of the dual 'weighting approximation', where the periodic-like term in $\mathcal{D}^\circ(b)$ is weighted by using the conditional matrix $Q^\circ_{\hat{a}(b)}$ instead of $Q_{\hat{a}(b)}$. Hence, we neglect all off-diagonal terms, i.e. correlation among (conditioned) ambiguities, and this approximation leads to a much simpler definition of the convex lower bound function (see Table 7.1), which is required for the BB iterations.

We consider now the equivalent approximate primal problem, used here to study the statistical ambiguity-fixing performance and to determine the properties of ambiguity-*fixed* dual estimators. Following (ibid), we can transform $Q_{\hat{a}\hat{a}} \to Q^\circ_{\hat{a}\hat{a}}$, with

$$Q^\circ_{\hat{a}\hat{a}} = Q^\circ_{\hat{a}(b)} + Q_{\hat{a}\hat{b}} Q^{-1}_{\hat{b}\hat{b}} Q_{\hat{b}\hat{a}} \tag{7.24}$$

while given $\alpha \in \mathbb{Z}^n$ we also have

$$\hat{b}^\circ(\alpha) = \hat{b} - Q_{\hat{b}\hat{a}} {Q^\circ_{\hat{a}\hat{a}}}^{-1} (\hat{a} - \alpha), \quad Q^\circ_{\hat{b}(a)} = Q_{\hat{b}\hat{b}} - Q_{\hat{b}\hat{a}} {Q^\circ_{\hat{a}\hat{a}}}^{-1} Q_{\hat{a}\hat{b}} \tag{7.25}$$

where $Q^\circ_{\hat{b}(a)}$ is not the variance matrix of $\hat{b}^\circ(a)$, since the latter is not independent of $\hat{a}$. Still, it is possible to conclude[1] that $Q^\circ_{\hat{b}(a)} \geq Q_{\hat{b}(a)}$, where the equality holds if $Q_{\hat{a}\hat{a}} = Q^\circ_{\hat{a}\hat{a}}$.

In this section we will therefore discuss and evaluate two cases: $Q_{\hat{a}(b)}$ diagonal, i.e. no approximation takes place, and then $Q_{\hat{a}(b)}$ nondiagonal. In the latter case we numerically compare the performance with respect to the original primal formulation, i.e. based on GNSS user positioning model.

7.3.1 Case with $Q_{\hat{a}(b)}$ diagonal

We start with the case where the conditional ambiguities are perfectly uncorrelated, and the matrix $Q_{\hat{a}(b)}$ is diagonal, so we look for models where the matrix $A^T Q^{-1}_{yy} A$ is diagonal. One GNSS example where this condition occurs is the multivariate Integer Cycle-Slips (I-CS) Resolution approach introduced by Teunissen and de Bakker (2015), and later discussed in Li and Melachroinos (2019).

Given a single-receiver user 'u' tracking each satellite 's' on J signal frequencies with wavelengths λ_j for $j = 1, \ldots, J$, we consider the time-differenced code and phase observables respectively as $\Delta p^s_u \in \mathbb{R}^J$ and $\Delta \phi^s_u \in \mathbb{R}^J$. Assuming we have no time

[1] In Theorem 4 by (ibid), it is proved that

$$Q^\circ_{\hat{b}(a)} = Q_{\hat{b}(a)} + Q_{\hat{b}\hat{a}} \left[Q^{-1}_{\hat{a}\hat{a}} - {Q^\circ_{\hat{a}\hat{a}}}^{-1} \right] Q_{\hat{a}\hat{a}} \left[Q^{-1}_{\hat{a}\hat{a}} - {Q^\circ_{\hat{a}\hat{a}}}^{-1} \right] Q_{\hat{a}\hat{b}} \geq Q_{\hat{b}(a)} \tag{7.26}$$

where, for a sufficiently high success rate, we have $\check{b}^\circ \sim \mathcal{N}_p \left(b, Q^\circ_{\hat{b}(a)} \right)$ given $\check{b}^\circ = \arg\min_{b \in \mathbb{R}^p} \mathcal{D}^\circ(b)$.

correlation (between epochs) and no correlation among different frequencies, the covariance matrices are diagonal and given by $2Q_{pp}$ and $2Q_{\phi\phi}$. The functional and stochastic models follow as

$$\begin{aligned} \mathrm{E}\left\{\begin{pmatrix} \Delta p_u^s \\ \Delta \phi_u^s \\ \Delta i_0 \end{pmatrix}\right\} &= \begin{bmatrix} e_J & +\mu & 0 \\ e_J & -\mu & \Lambda \\ 0 & 1 & 0 \end{bmatrix} \begin{pmatrix} \Delta \rho_u^s \\ \Delta i_{u,1}^s \\ c_r^s \end{pmatrix}, \\ \mathrm{D}\left\{\begin{pmatrix} \Delta p_u^s \\ \Delta \phi_u^s \\ \Delta i_0 \end{pmatrix}\right\} &= \begin{bmatrix} 2Q_{pp} & 0 & 0 \\ 0 & 2Q_{\phi\phi} & 0 \\ 0 & 0 & \sigma^2_{\Delta i_{u,1}^s} \end{bmatrix} \end{aligned} \tag{7.27}$$

where $e_J = (1, \ldots, 1)^T$ and $\Lambda = \mathrm{diag}\,(\lambda_1, \ldots, \lambda_J)$. The ionosphere pseudo-observation $\Delta i_0 \in \mathbb{R}$ is weighted by $\sigma^2_{\Delta i_{u,1}^s}$ and set to zero, so it constrains the time-differenced slant ionospheric delay on the first frequency given by $\Delta i_{u,1}^s$, scaled by the coefficients $\mu = (\mu_1, \ldots, \mu_J)^T$ given $\mu_j = f_1^2/f_j^2$. The time difference of the lumped parameter including frequency-independent terms, e.g. the receiver-satellite range, is given by $\Delta \rho_u^s \in \mathbb{R}$, while the integer cycle-slip vector is $c_r^s \in \mathbb{Z}^J$.

Hence, this is a mixed-integer model with $p = 2$ and $n = J$, where the conditional covariance ambiguity matrix is diagonal and its inverse is defined as

$$Q^{-1}_{\check{a}(b)} = \frac{\Lambda^T Q^{-1}_{\phi\phi} \Lambda}{2} = \mathrm{diag}\left(\frac{\lambda_1^2}{2\sigma^2_{\phi_1}}, \ldots, \frac{\lambda_J^2}{2\sigma^2_{\phi_J}}\right) \in \mathbb{R}^{J \times J} \tag{7.28}$$

so each conditional ambiguity variance is ultimately given by

$$\sigma^2_{\check{a}_j(b)} = \frac{2\sigma^2_{\phi_j}}{\lambda_j^2}, \quad \forall j = 1, \ldots, J \tag{7.29}$$

where generally we have $\sigma_{\phi_j}/\lambda_j \approx 0.01$ in the GNSS scenarios.

An illustrative example is given in Fig. 7.8 where we show the dual objective function $\mathcal{D}(b) : \mathbb{R}^{p=2} \to \mathbb{R}$ given $b = \left(\Delta \rho_u^s, \Delta i_{u,1}^s\right)^T$, thus considering GPS L1, L1+L2 and L1+L2+L5 signals based on 30cm and 3mm as standard deviation of code and phase (undifferenced) observations, respectively. We set $\sigma^2_{\Delta i_{u,1}^s} = 0.01^2$, which leads to a SR of 17.7%, 92.9% and 99.8% for the three configurations, and we show the float solution (black asterisk), the fixed solution (magenta diamond) and the region (red line) in the parameters' domain relative to true integer vector $\check{a}(b) = a \in \mathbb{Z}^J$.

As mentioned in Teunissen and Massarweh (2024), when columns of A are mutually orthogonal in the metric of Q_{yy}, then $Q_{\hat{a}(b)}$ is diagonal and no approximation is used. Hence, optimal solutions are achieved and this is the case for other GNSS models, such as the multi-frequency geometry-free models. In these ones, the conditional variance matrix is diagonal, whereas the unconditional vc-matrix $Q_{\hat{a}\hat{a}}$ is not.

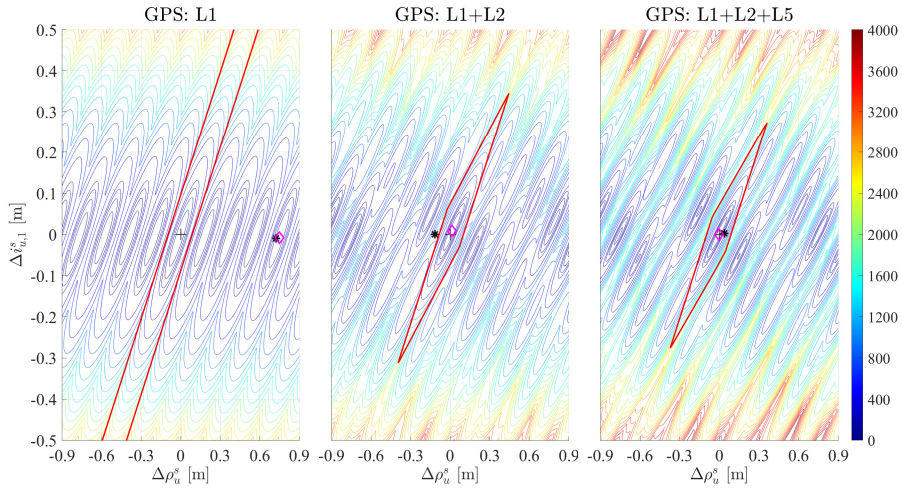


Figure 7.8: The illustration of the dual objective function for the multivariate integer cycle-slip model described in Eq. 7.27 is given based on three different configurations for GPS L1, L1+L2 and L1+L2+L5 signals. The float and fixed solutions are represented respectively as a black asterisk and a magenta diamond, while the region associated to the true integer vector $\check{a}(b) = a \in \mathbb{Z}^J$ is marked with a red line. See text for additional information.

Moreover, the unconditionally estimated ambiguities are highly correlated, with the correlation coefficient between two ambiguities $\hat{a}_i$ and $\hat{a}_j$ given as $\rho_{\hat{a}_i\hat{a}_j} \approx 1$ due to the very small variance ratio between phase and code.

7

7.3.2 Case with $Q_{\hat{a}(b)}$ not diagonal

We consider the case where conditional ambiguities are correlated, and the matrix $Q_{\hat{a}(b)}$ is not diagonal. Still, as said in Massarweh and Teunissen (2024), this conditional ambiguity covariance matrix is generally small given the high precision of carrier-phase measurements in GNSS models. By looking at Eq. 7.24, assuming $Q^{\circ}_{\hat{a}(b)} \ll Q_{\hat{a}\hat{b}}Q^{-1}_{\hat{b}\hat{b}}Q_{\hat{b}\hat{a}}$, we notice that the dual approximation of $Q^{\circ}_{\hat{a}(b)}$ might not have a large influence on $Q^{\circ}_{\hat{a}\hat{a}}$. Following (ibid), we can therefore compare three scenarios: adoption of $Q_{\hat{a}\hat{a}}$ for a primal ILS solution, $Q^{\circ}_{\hat{a}\hat{a}}$ for a dual approximation (with $Q^{\circ}_{\hat{a}(b)}$ diagonal), and $Q^{\odot}_{\hat{a}\hat{a}}$ (diagonal) for a primal IR solution where no decorrelation is performed.

We take a nominal 24-satellite constellation for GPS and Galileo tracked on L1+L5 and E1+E5a by a single (short) baseline at middle latitude, e.g., near REDU station, every 5 minutes over a 24-hour period. An illustration is given in Fig. 7.9, where the number of visible satellites (i.e. 7° elevation mask) is shown per constellation on the left side, along with a skyplot on the right side. Notice that around 13-to-17 satellites are considered for this multi-GNSS dual-frequency constellation, given $\lambda_1 = 19.03\ cm$ and $\lambda_2 = 25.48\ cm$.

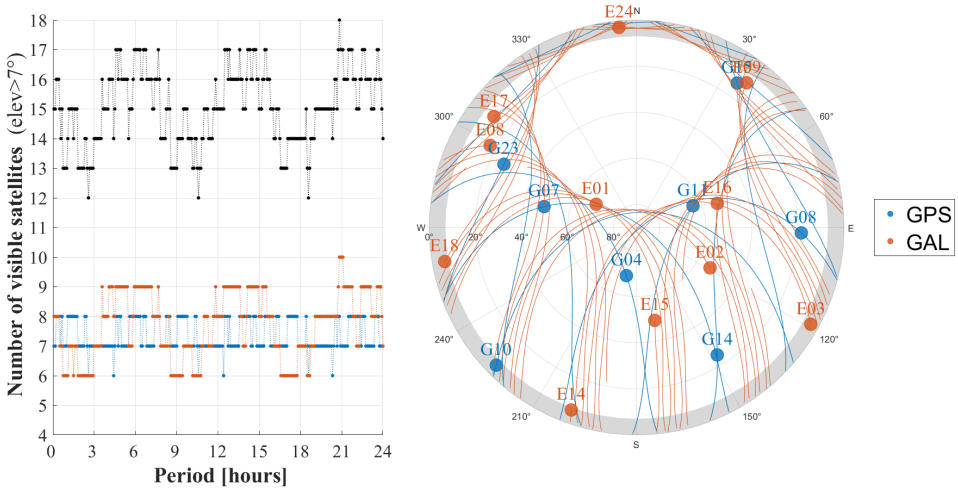


Figure 7.9: The number of visible satellites with a 7° elevation mask is shown on the left panel for nominal GPS and Galileo constellations tracked by one single (short) baseline at middle latitude (e.g. near REDU) over a 24-hour period, along with a sky plot on the right panel.

At each epoch, the functional and stochastic models (Odijk and Teunissen, 2008) are:

$$\begin{aligned} \mathrm{E}\left\{\begin{pmatrix}\phi_1\\ \phi_2\\ p_1\\ p_2\end{pmatrix}\right\} &= \begin{bmatrix}\lambda_1 I_{m-2} & & D_m^T G\\ & \lambda_2 I_{m-2} & D_m^T G\\ & & D_m^T G\\ & & D_m^T G\end{bmatrix}\begin{pmatrix}a_1\\ a_2\\ b\end{pmatrix}, \qquad b = \begin{pmatrix}b_E\\ b_N\\ b_U\end{pmatrix} \in \mathbb{R}^{p=3}\\ \mathrm{D}\left\{\begin{pmatrix}\phi_1\\ \phi_2\\ p_1\\ p_2\end{pmatrix}\right\} &= 2\begin{bmatrix}D_m^T Q_{\phi_1\phi_1} D_m & & & \\ & D_m^T Q_{\phi_2\phi_2} D_m & & \\ & & D_m^T Q_{p_1p_1} D_m & \\ & & & D_m^T Q_{p_2p_2} D_m\end{bmatrix}\end{aligned} \tag{7.30}$$

with a_j referring to DD ambiguities for each j-th signal, while

$$G = \begin{bmatrix}\vdots & \vdots & \vdots\\ \cos(El^\circ)\sin(Az^\circ) & \cos(El^\circ)\cos(Az^\circ) & \sin(El^\circ)\\ \vdots & \vdots & \vdots\end{bmatrix} \tag{7.31}$$

and $D_m^T = \mathrm{blkdiag}\left(D_{m_G}^T, D_{m_E}^T\right)$ being a block diagonal matrix for the between-satellite single differencing $D_{m_S}^T \in \mathbb{R}^{(m_S-1)\times m_S}$ for each 'S' constellation, so $m = m_G + m_E$. The choice of a pivot satellite per constellation is here necessary due to the existence of receiver differential GPS-Galileo inter-system biases on the overlapping frequencies, see Odijk and Teunissen (2013). The variance matrices for code $Q_{p_jp_j}$ and phase $Q_{\phi_j\phi_j}$ are both diagonal, where the variance terms $\sigma_{p_j}^2$ and $\sigma_{\phi_j}^2$ are then multiplied by an elevation weight $\sin^{-2}(el^\circ)$. Lastly, we make use of $\sigma_{p_1} = 30\ cm$ and $\sigma_{p_2} = 20\ cm$, along with $\sigma_{\phi_1} = 3\ mm$ and $\sigma_{\phi_2} = 2\ mm$.

Recalling the block diagonal structure of the variance-covariance matrix Q_{yy} of observables, we note that $Q_{\hat{a}(b)}$ is a block diagonal matrix with each frequency block equal to $2\left(D_m^T Q_{\phi_j\phi_j} D_m\right)/\lambda_j^2$, therefore off-diagonal entries still exist due to the between-satellite differencing of phase observations, nonetheless these terms are generally small given a ratio $\sigma_{\phi_j}/\lambda_j \ll 1$. Furthermore, the block diagonal structure of D_m^T (due to the choice of one pivot satellite per constellation) leads further defining constellation-wise blocks, so

$$\frac{2}{\lambda_j^2}\left(D_m^T Q_{\phi_j\phi_j} D_m\right) = \frac{2}{\lambda_j^2}\begin{bmatrix} D_{m_G}^T Q_{\phi_j\phi_j}^G D_{m_G} & \\ & D_{m_E}^T Q_{\phi_j\phi_j}^E D_{m_E} \end{bmatrix} \tag{7.32}$$

after separating measurements of GPS (G) and Galileo (E) constellations.

For the statistical performance analysis, we generate observation data distributed as $\mathcal{N}(0, Q_{yy})$, then we calculate the 'empirical SR' (based on 100,000 samples) to evaluate fixing performance of the three cases, where in the approximations of Case #2 and Case #3 we are neglecting the off-diagonal terms of matrix $Q_{\hat{a}(b)}$ and $Q_{\hat{a}\hat{a}}$, respectively.

As visible in the left panel of Fig. 7.10, the dual approximation achieves similar SR performance as the primal ILS solution, i.e. expected by the close approximation made in Eq. 7.24. On the right panel, we evaluate the Root Mean Squared Error (RMSE) of the ambiguity-fixed parameters, with values for the 3D positioning RMSE at centimeter level for Case #1 and Case #2. Still, the dual approximation conditioning $\hat{b}^\circ(a)$ is based on the approximated $Q_{\hat{a}\hat{a}}^\circ$ instead of $Q_{\hat{a}\hat{a}}$, so some wrong fixes are visible, while for Case #3 we find RMSE at around 35-40 cm, i.e. consequence of the poorer fixing performance when using the approximation $Q_{\hat{a}\hat{a}}^{\odot}$.

We focus now on a single epoch, i.e. at 20:50, with 18 tracked satellites and we illustrate the horizontal errors and vertical errors (for all samples) respectively in Fig. 7.11 and Fig. 7.12. As visible, centimeter-level solutions are still possible, nonetheless the performance is slightly deteriorated due to potential wrong fixes (e.g. when SR is lower), but also due to the fact that $\hat{b}^\circ(a) \in \mathbb{R}^p$, and so we expect $Q_{\hat{b}(a)}^\circ \geq Q_{\hat{b}(a)}$.

7.3.3 Summary remarks

As mentioned, the matrix $Q_{\hat{a}(b)}$ is generally less correlated, and in the previous example the correlation existed only due to a between-satellite differencing of phase observations. Thus, a block diagonal structure as $Q_{\hat{a}(b)} = blkdiag\left(Q_{\hat{a}_1(b)}, Q_{\hat{a}_2(b)}\right)$ allows further simplifying a dual formulation, given $\mathcal{H}_2 : \mathbb{Z}^{n_1} \times \mathbb{Z}^{n_2} \times \mathbb{R}^p \to \mathbb{R}$, where

$$\mathcal{H}_2(a_1, a_2, b) \overset{\text{def}}{=\!=} \left\|\hat{b} - b\right\|_{Q_{\hat{b}\hat{b}}}^2 + \|\hat{a}_1(b) - a_1\|_{Q_{\hat{a}_1(b)}}^2 + \|\hat{a}_2(b) - a_2\|_{Q_{\hat{a}_2(b)}}^2 \tag{7.33}$$

with $n = n_1 + n_2$, while assuming $\hat{a}_q(b) = \hat{a}_q - Q_{\hat{a}_q\hat{b}} Q_{\hat{b}\hat{b}}^{-1}\left(\hat{b} - b\right) \in \mathbb{R}^{n_q}$ for each subset q, and given $Q_{\hat{a}_1(b)\hat{a}_2(b)} = 0$. Notice that the expression above represents no approximation of the original problem.

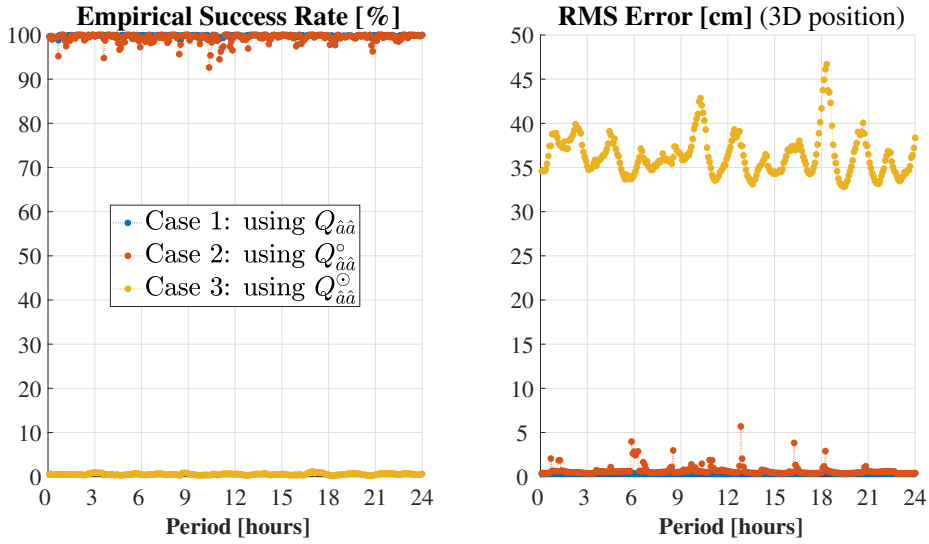


Figure 7.10: The Empirical SR (left) and Root Mean Square Errors (right) are computed based on three cases described in this section, using 10^5 samples at each epoch, see text for details.

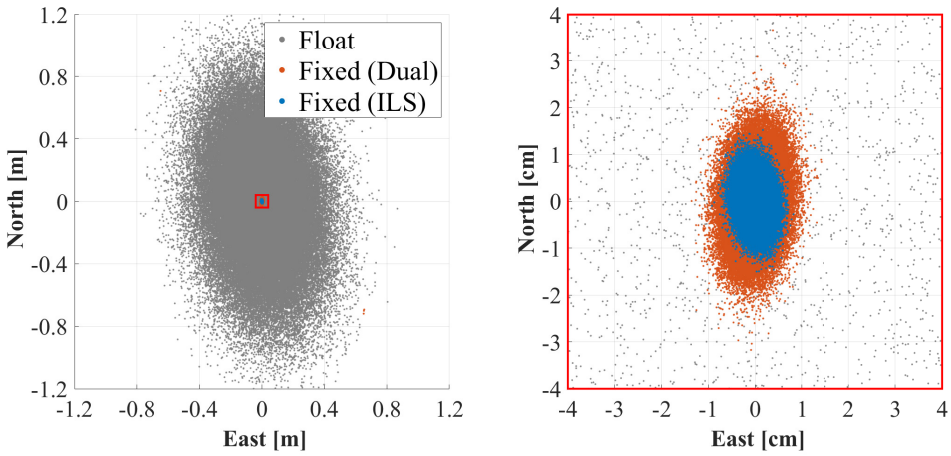


Figure 7.11: The horizontal positioning errors of the east and north coordinates are given for a single epoch based on 10^5 samples, thus showing float solutions (in grey), approximate dual solutions (in orange) and primal ILS solutions (in blue), with a zoom-in on the right panel.

This last consideration leads also to a possibility of performing partial ambiguity resolution, where only a subset of the ambiguities is resolved, e.g. $\check{a}_2 = \hat{a}_2(b) \in \mathbb{R}^{n_2}$. This comes in addition to the partitioned dual formulation suggested in Section 8 by Teunissen and Massarweh (2024), i.e. splitting of real-valued parameters. For this dual PAR approach, different criteria might be considered, for example neglecting specific

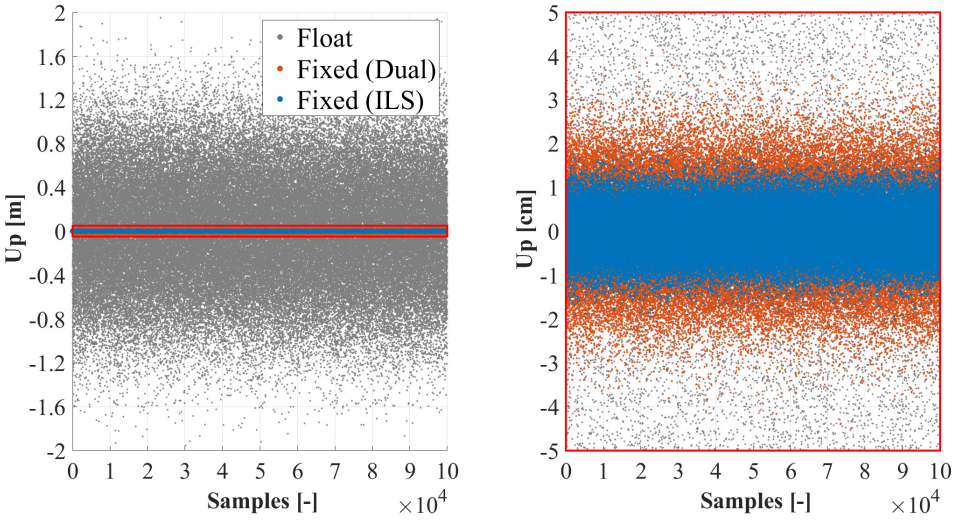


Figure 7.12: The vertical errors of the up coordinate is given for a single epoch based on 10^5 samples, thus showing float solutions (in grey), approximate dual solutions (in orange) and primal ILS solutions (in blue), with a zoom-in on the right panel.

satellites/signals or large $\sigma_{\hat{a}_i(b)}$ values, and will be investigated in future works.

For sake of clarity, we highlight that current BINGO computations are not over performing (yet) the very recent LAMBDA 4.0 implementation (Massarweh et al., 2025), as the latter has benefitted from decades of research works. Still, the landscape of possibilities introduced by this dual formulation is certainly vast and yet to be explored. As an example, the possible generalization to 'constrained dual' problems is yet to be exploited when some additional constraints on the real-valued parameters exist and could be taken into account rigorously.

7.4 Conclusions

In this contribution we have investigated the performance in dual mixed-integer least-squares problems, introduced by Teunissen and Massarweh (2024). We focus on the 'unconstrained dual' problems, thus accounting for the integerness of ambiguities and freely defining the real-valued parameters in $\mathbb{R}^p$. The globally convergent algorithm discussed in (ibid) is here extensively examined, therefore discussing a Branch-and-bound INteger-equivariant Global Optimizer (BINGO) implementation for solving dual problems. Following a normalization of the dual problem, we perform an in-depth analysis of the computational performance and we scrutinize individual computational steps in order to evaluate the impact of several different elements: problem normalization and initialization, boxes' splitting strategy, convex lower bound minimization, and ϵ_{BB}-termination criterion.

In addition to the computational performance, we evaluate the statistical performance for the dual approximation used in the BINGO approach, and we show how this might still provide an accurate representation of the original problem. By means of numerical analysis for a multi-GNSS single-baseline positioning example, we demonstrate how quasi-optimal solutions are possible, with quasi-optimal results comparable (for the selected scenarios) to the primal integer least-squares estimator.

Overall, the landscape of possibilities and potential applications for the dual mixed-ILS formulation is vast and still to be explored, where additional aspects shall be investigated, e.g. decorrelation, parameters' splitting or partial ambiguity resolution. Lastly, we remark that we have focused here on 'unconstrained dual' problems, while some additional constraints on real-valued parameters might still be accounted for in a rigorous way with the dual formulation, and this is subject of future research.

References

Andersen A (2024) Projected gradient algorithm. ECS at University Southampton, UK. Available at https://angms.science/doc/CVX/CVX_PGD.pdf (Accessed: 7 September 2024)

Armijo L (1966) Minimization of functions having lipschitz continuous first partial derivatives. Pacific Journal of Mathematics 16(1):1–3

Balakrishnan V, Boyd S, Balemi S (1991) Branch and bound algorithm for computing the minimum stability degree of parameter-dependent linear systems. International Journal of Robust and Nonlinear Control 1(4):295–317

Bertsekas D (1999) Nonlinear Programming, 2nd edn. Athena Scientific

Boyd S, Mattingley J (2007) Branch and bound methods. Notes for EE364b, Stanford University, 2006-07

Hofmann-Wellenhof B, Lichtenegger H, Wasle E (2008) GNSS: Global Navigation Satellite Systems: GPS, GLONASS, Galileo, and More. Springer, New York

Jones DR, Martins JRRA (2021) The DIRECT algorithm: 25 years later. Journal of Global Optimization 79:521–566, DOI 10.1007/s10898-020-00952-6

Jones DR, Perttunen CD, Stuckman BE (1993) Lipschitzian optimization without the lipschitz constant. Journal of Optimization Theory and Applications 79(1):157–181, DOI 10.1007/BF00941892

Lawler EL, Wood DE (1966) Branch-and-bound methods: A survey. Operations Research 14(4):699–719

Li T, Melachroinos S (2019) An enhanced cycle slip repair algorithm for real-time multi-GNSS, multi-frequency data processing. GPS Solutions 23:1, DOI 10.1007/s10291-018-0792-0

Massarweh L, Teunissen PJG (2024) An efficient 'P1' algorithm for dual mixed-integer least-squares problems with scalar real-valued parameters. Journal of Applied Geodesy DOI 10.1515/jag-2024-0076

Massarweh L, Verhagen S, Teunissen PJG (2025) New LAMBDA toolbox for mixed-integer models: estimation and evaluation. GPS Solutions 29(1):14, DOI 10.1007/s10291-024-01738-z, available at http://pntlab.tudelft.nl/LAMBDA

Morrison DR, Jacobson SH, Sauppe JJ, Sewell EC (2016) Branch-and-bound algorithms: A survey of recent advances in searching, branching, and pruning. Discrete Optimization 19:79–102

Odijk D, Teunissen PJG (2008) ADOP in closed form for a hierarchy of multi-frequency single-baseline GNSS models. Journal of Geodesy 82:473–492, DOI 10.1007/s00190-007-0197-2

Odijk D, Teunissen PJG (2013) Characterization of between-receiver gps-galileo inter-system biases and their effect on mixed ambiguity resolution. GPS Solutions 17:521–533, DOI 10.1007/s10291-012-0298-0

Parikh N, Boyd S (2013) Proximal algorithms. Foundations and Trends in Optimization 1(3):123–231

Schmidt M (2020) First-order optimization algorithms for machine learning - projected-gradient methods. The University of British Columbia, Canada. Available at https://www.cs.ubc.ca/~schmidtm/Courses/5XX-S20/S5.pdf (Accessed: 24 July 2024)

Schmidt M, Kim D, Sra S (2011) Projected newton-type methods in machine learning. In: Optimization for Machine Learning

Teunissen PJG (1990) Nonlinear least-squares. Manuscripta Geodaetica 15(3):137–150

Teunissen PJG (1993) Least-squares estimation of the integer gps ambiguities. In: IAG General Meeting. Invited Lecture. Section IV Theory and Methodology

Teunissen PJG (1994) A new method for fast carrier phase ambiguity estimation. In: Proceedings of 1994 IEEE Position, Location and Navigation Symposium - PLANS'94, Las Vegas, NV, USA, pp 562–573, DOI 10.1109/PLANS.1994.303362

Teunissen PJG (1995) The least-squares ambiguity decorrelation adjustment: A method for fast gps integer ambiguity estimation. Journal of Geodesy 70(1–2):65–82

Teunissen PJG (2003) Towards a unified theory of GNSS ambiguity resolution. Journal of Global Positioning Systems 2(1):1–12

Teunissen PJG, de Bakker PF (2015) Multivariate integer cycle-slip resolution: A single-channel analysis. In: N Sneeuw MC P Novák, Sansò F (eds) VIII Hotine-Marussi Symposium on Mathematical Geodesy, International Association of Geodesy Symposia, vol 142, pp Springer, Cham, DOI 10.1007/1345_2015_69

Teunissen PJG, Massarweh L (2024) Primal and dual mixed-integer least-squares: Distributional statistics and global algorithm. Journal of Geodesy 98:63, DOI 10.1007/s00190-024-01862-1

8

Theory for the Ambiguity Function Method: Probability Model and Global Solution

In this contribution we introduce some new theory for the classical GNSS ambiguity function (AF) method. We provide the probability model by means of which the AF-estimator becomes a maximum likelihood estimator and we provide a globally convergent algorithm for computing the AF-estimate. The algorithm is constructed from combining the branch-and-bound principle, with a special convex-relaxation of the multimodal ambiguity function, to which the projected-gradient-descent method is applied to obtain the required bounds. We also provide a systematic comparison between the AF-principle and that of integer least-squares (ILS). From this comparison the conclusion is reached that the two principles are fundamentally different, although there are identified circumstances under which one can expect AF- and ILS-solutions to behave similarly.

This chapter will be published as: Teunissen, P.J.G., and Massarweh, L. (2024). *Theory for the Ambiguity Function Method: Probability Model and Global Solution.* In Journal of Geodesy (*submitted*).

8.1 Introduction

The ambiguity function (AF) method is one of the earliest methods for estimating baselines from integer ambiguous GNSS carrier-phase data. The method was introduced and popularized by Counselman and Gourevitch (1981) and Remondi (1984, 1991), while its original idea of eliminating dependence on the '2π' ambiguities goes back to (Rogers et al., 1978). In fact, it was this property of invariance that contributed to the initial popularity of the method. It promised the capability of determining precise baselines, without the explicit need of having to resolve the values of integer carrier-phase ambiguities.

Although the AF-method is one of the oldest methods, its statistical and numerical evolution did not keep pace with the theoretical developments of other methods of mixed-integer inference (Teunissen, 2003b, 2017; Hartman, 2021). Currently we have different classes of mixed-integer estimators, with identified optimal estimators within each class, together with numerically efficient computational algorithms (Teunissen, 1995). For instance, the best integer-equivariant (BIE) estimator (Teunissen, 2003a) is minimum-mean-squared-error optimal in the largest class, while the integer least-squares (ILS) estimator (Teunissen, 1999) is best in maximizing the probability of correct integer estimation within the smaller integer-class.

As the maturity of the AF-method is not on par with the current methods of mixed-integer estimation, it is the goal of the present contribution to help fill in some of the theoretical gaps. In doing so, the two main innovations of this contribution are: (1) the provision of a probability model by means of which the AF-solution is given a statistical basis, and (2) the provision of a global optimizer of the AF likelihood function, having finite termination with a guaranteed epsilon-tolerance.

This contribution is organized as follows. In Section 8.2 we provide a brief review of the ambiguity-function method, together with examples of its use. Then in Section 8.3 we draw attention to the possible nonuniqueness of the AF-solution. This is new, as this problem has not been addressed before in the AF-method's literature. We prove under which conditions the AF-solution is nonunique and how one can verify whether or not this nonuniqueness is problematic for the specific problem or application under consideration.

In Section 8.4, we introduce our probability model for the AF-method. It shows what distributional assumptions need to be made in order for the AF-estimator to become a maximum likelihood estimator. It therefore provides, for the first time, a statistical basis for the AF-method of GNSS baseline determination. In doing so, we also generalize the classical expression of the AF-estimator by showing how a varying precision of the carrier-phase observables can be incorporated into the estimation scheme.

In order to describe the characteristics of the ambiguity objective function qualitatively, we provide its multivariate gradient and Hessian in Section 8.5. Their analysis indeed shows that the ambiguity objective function is severely multimodal and that, in the

absence of very accurate baseline initializations, iterative gradient descent methods will not be able to locate the sought for maximizer of the likelihood function. A global optimization method is therefore asked for, which we introduce in Section 8.6. Our proposed method, which has finite termination with a guaranteed epsilon-tolerance, is constructed from combining the branch-and-bound principle, with a special convex-relaxation of the ambiguity function, to which the projected-gradient-descent method is applied to obtain the required bounds. Each of the method's three constituents is described in this chapter, whereby special emphasis is given to the construction of the required continuously differentiable, convex lower bounding function of the multimodal ambiguity function.

In Section 8.7 we compare the AF estimation principle with that of integer least-squares (ILS). Although we exemplify the various marked differences between the two principles, we also show under which identified circumstances one can expect AF- and ILS-solutions to be close. We do this by making use of the primal-dual equivalence of mixed-ILS theory as introduced in (Teunissen and Massarweh, 2024). The presented theory is supported by means of several examples in which the workings and performance of the AF-method are numerically and graphically illustrated. Finally, Section 8.8 contains the Summary and Conclusions.

The following notation is used: $\mathsf{E}(.)$ and $\mathsf{D}(.)$ stand for the expectation and dispersion operators, respectively, and $\mathcal{N}_p(\mu, Q)$ denotes a p-dimensional, normally distributed random vector, with mean (expectation) μ and variance matrix (dispersion) Q. $\mathbb{R}^p$ and $\mathbb{Z}^p$ denote the p-dimensional spaces of real- and integer numbers, respectively. The Q-weighted squared norm is denoted as $||.||^2_Q = (.)^T Q^{-1}(.)$, and $\lceil x \rfloor$ denotes the rounding of x to the nearest integer. If applied to a vector, the rounding is understood to apply to each of its coordinates. $\cup$ and $\cap$ denote the union and intersection operators, and the vectorial inequality $\preceq$ denotes the all componentwise inequality $\leq$. The gradient of a function $F(b)$ is denoted as $\partial_b F(b)$ and the central Chi-square distribution with p degrees of freedom is denoted as $\chi^2(p, 0)$, with $\chi^2_\delta(p, 0)$ being its δ-percentage critical value.

8

8.2 The AF-method: a brief review

The single-baseline, k-epochs, f-frequencies, and s-satellites GNSS ambiguity function (AF) is generally defined as (Mader, 1992; Lachapelle et al., 1992; Leick et al., 2015),

$$\mathrm{AF}(b) = \sum_{t=1}^{k} \sum_{j=1}^{f} \sum_{i=1}^{s} \cos[\tfrac{2\pi}{\lambda_j}(\phi^{ri}_{12,j}(t) - \rho^{ri}_{12}(t, b))] \tag{8.1}$$

in which λ_j is the wavelength of the jth frequency, $\phi^{ri}_{12,j}(t) = [\phi^i_{2,j}(t) - \phi^i_{1,j}(t)] - [\phi^r_{2,j}(t) - \phi^r_{1,j}(t)]$ the double-differenced (DD) phase-observable, in units of range, on frequency j at epoch t of receivers $1, 2$ and satellites r, i, and $\rho^{ri}_{12}(t, b)$ is its corresponding DD-range, which depends on b, the unknown baseline vector between receivers 1 and 2.

Note that AF(b) is *invariant* for any pertubations of $\phi^{ri}_{12,j}(t)$ that are integer multiples of the wavelength λ_j. Hence, it is invariant for integer cycle slips in the phase data, as well as for the presence of the DD integer ambiguities $a^{ri}_{12,j} \in \mathbb{Z}$ in the observation equations $\phi^{ri}_{12,j}(t) = \lambda_j a^{ri}_{12,j} + \rho^{ri}_{12}(t,b) + \epsilon^{ri}_{12,j}(t)$. This invariance has in fact been the overarching motivation for introducing the AF-concept (Counselman and Gourevitch, 1981; Remondi, 1984; Mader, 1992; Hofmann-Wellenhof et al., 2008; Leick et al., 2015). It promises namely of being able to resolve the unknown baseline b, without the explicit need of having to resolve the values of the integer ambiguities $a^{ri}_{12,j}$.

As AF(b) reaches its maximum value when all DD phase errors $\epsilon^{ri}_{12,j}(t)$ are identically zero, the chosen AF-approach for resolving the unknown baseline is to *aim* for a solution that satisfies $\check{b} = \arg\max_b \mathrm{AF}(b)$. The usual approach for doing so is by direct evaluation of the ambiguity function on the vertices of a three-dimensional rectangular grid, centred at an approximate baseline solution. The vertex that provides the largest ambiguity function value is then selected as the solution sought (Rogers et al., 1978; Remondi, 1984; Hofmann-Wellenhof et al., 2008). It will be clear that the numerical and statistical efficacy depends on the chosen grid spacing, grid size, and grid location.

Although the AF-principle, of working with an integer-ambiguity invariant objective function that is maximized when the errors are zero, forms the basis of all publications in the GNSS AF-literature, it is important to realize that different authors applied the principle to different objective functions. As a consequence, different baseline results will be obtained even when these authors would be using the same original data.

Some authors work on single-differenced data, while others apply the principle to double-differenced data, and some work directly with the cosine function, as in (8.1), while other authors work with the complex phasor function, being the analytical representation of a cosine function. Remondi (1984, 1991), for instance, works with phasor norms, as a result of which his ambiguity objective function contains, in contrast to (8.1), both cosine and sine functions. The same phasor norms are also used by Han and Rizos (1996), but they use it on double-differenced data, instead of on single-differenced data as is done in (Remondi, 1984, 1991; Remondi and Hilla, 1993; Hofmann-Wellenhof et al., 2008).

In this contribution we will use, as in (8.1), a 'sum of cosines' as our basis for constructing the ambiguity-invariant objective function. This is consistent with the original formulation of Rogers et al. (1978), but more importantly, it will allow us to formulate a probabilistic model for the ambiguity function method.

8.3 On the nonuniqueness of the AF-solution

As mentioned earlier, the attractiveness of the AF-method is the integer-ambiguity invariance of its objective function. This invariance, however, holds for any integer pertubations of its argument and thus also for those that may be generated by changes in the baseline. If such case happens, one can not expect the baseline solution of the AF-method to be unique.

To study the possible nonuniqueness of the AF-solution, we first introduce a useful compact notation for the ambiguity function. Let $\epsilon_\phi = [\epsilon_{\phi_1}, \ldots, \epsilon_{\phi_m}]^T$ be an m-vector with its entries expressed in cycles and let $e_m = [1, \ldots, 1]^T$ be the m-vector of ones. Then we introduce for the AF-function the compact notation

$$\mathrm{AF} = e_m^T \cos[2\pi(\epsilon_\phi)] := \sum_{i=1}^{m} \cos[2\pi(\epsilon_{\phi_i})] \tag{8.2}$$

Thus $\cos[2\pi(\epsilon_\phi)]$ is the vector that consists of the componentwise cosine values of $2\pi\epsilon_{\phi_i}$. As the general system of GNSS carrier-phase observation equations can be written in vector-matrix cycle-form as

$$\phi = A_\phi a + B_\phi b + \epsilon_\phi,\ a \in \mathbb{Z}^n, b \in \mathbb{R}^p, A_\phi \in \mathbb{Z}^{m\times n} \tag{8.3}$$

the to-be-maximized objective function of the AF-method follows upon substitution of $\epsilon_\phi = \phi - A_\phi a - B_\phi b$ into (8.2) as

$$\mathrm{AF}(b) = e_m^T \cos[2\pi(\phi - B_\phi b)] \tag{8.4}$$

Note that this formulation generalizes that of (8.1) in the sense that ϕ need now not be retricted to a DD-form and that b need not be restricted to a single baseline. As (8.4) applies to any carrier-phase system of the form (8.3), it holds in principle for undifferenced data and networks as well.

Also note, due to the property of the cosine-function, that the ambiguity-part $A_\phi a \in \mathbb{Z}^m$ of system (8.3) disappeared from the objective function $\mathrm{AF}(b)$. This is also the principal attractiveness of the method as it implies that no explicit 'integer ambiguity resolution' is required when maximizing (8.4). We now show what this 'invariance' does to the uniqueness of the AF-maximizer.

Theorem 1 (Nonuniqueness of AF-solution): Let $m \times (n + p)$ design matrix $[A_\phi, B_\phi]$ (cf. 8.3) be of full column rank and let $Z = [Z_1, Z_2]$, $Z_1 \in \mathbb{Z}^{m\times(m-p)}$, $Z_2 \in \mathbb{Z}^{m\times p}$, be an admissible ambiguity transformation (i.e. Z and Z^{-1} have integer entries), satisfying $B_\phi^{\perp T}[Z_1, Z_2] = [L, 0]$, where $B_\phi^\perp$ is a basis matrix of the null space of B_ϕ^T. Then the ambiguity function (8.4) satisfies

$$\mathrm{AF}(b + B_\phi^+ Z_2 \tilde{z}_2) = \mathrm{AF}(b)\ ,\ \forall \tilde{z}_2 \in \mathbb{Z}^p \tag{8.5}$$

in which B_ϕ^+ is a left-inverse of B_ϕ (i.e. $B_\phi^+ B_\phi = I_p$). ■

8

Proof: From $B_\phi^{\perp T}[Z_1, Z_2] = [L, 0]$, it follows that Z_2 is an integer basis matrix of the range space of B_ϕ, i.e. $Z_2 = B_\phi X$ for some invertible $p \times p$ matrix $X = B_\phi^+ Z_2$. Therefore

$$\begin{aligned} \mathrm{AF}(b) &= e_m^T \cos[2\pi(\phi - B_\phi b)] \\ &= e_m^T \cos[2\pi(\phi - B_\phi b - Z_2 \tilde{z}_2)] \\ &= e_m^T \cos[2\pi(\phi - B_\phi (b + X\tilde{z}_2))] \\ &= e_m^T \cos[2\pi(\phi - B_\phi (b + B_\phi^+ Z_2 \tilde{z}_2))] \\ &= \mathrm{AF}(b + B_\phi^+ Z_2 \tilde{z}_2) \end{aligned} \tag{8.6}$$

□

The important consequence of this result is that the AF-solution may not be unique even if the design matrix $[A_\phi, B_\phi]$ (cf. 8.3) is of full column rank. Thus even if the phase-only system of observation equations produces a unique float-solution, with corresponding integer least-squares (ILS) solution, the solution produced by the AF-method may not be unique. This possible lack of baseline uniqueness is here identified for the first time as it is not part of the deliberations in the classical AF-literature (Counselman and Gourevitch, 1981; Remondi, 1984; Mader, 1992; Hofmann-Wellenhof et al., 2008; Leick et al., 2015).

The condition under which the above lack of uniqueness occurs is when an admissible integer matrix $Z = [Z_1, Z_2]$ can be constructed such that $B^{\perp T}[Z_1, Z_2] = [L, 0]$. This is always possible when the entries of matrix $B^\perp$ are *rational*, see Theorem 2 in (Teunissen and Khodabandeh, 2022). We hereby note, even if the actual entries of $B^\perp$ are not rational, that in the context of numerical computing they could be, which then still introduces numerically the above-mentioned lack of uniqueness.

Whether or not the above-identified lack of uniqueness is problematic from a practical point of view may depend on how large the smallest pertubation $B^+ Z_2 \tilde{z}_2$ will be. If the smallest such pertubation is sufficiently large, then a local maximizer $\check{b}$ of $\mathrm{AF}(b)$ could still be acceptable, since the next nearest maximizer will then be far away. The smallest distance between the local maximizers is given by

$$D_{\min} = \min_{z \in \mathbb{Z}^p \setminus \{0\}} ||B_\phi^+ Z_2 z||_Q \tag{8.7}$$

with Q as user-chosen positive definite matrix, e.g. $Q = (B_\phi^T B_\phi)^{-1}$. This can be computed efficiently with LAMBDA method (Teunissen, 1995; Massarweh et al., 2025).

8.4 A probability model for the Ambiguity Function method

In this section we will develop our probability model for the AF method. The starting idea is to find a probability density function (PDF) that has the solution of the AF method, $\check{b} = \arg\max_b e_m^T \cos[2\pi(\phi - B_\phi b)]$ (cf. 8.4), as its *maximum likelihood* estimate.

We ignore for the moment that the solution may be nonunique, then

$$\begin{aligned}
\check{b} &= \arg\max_b e_m^T \cos[2\pi(\phi - B_\phi b)] \\
&\overset{(a)}{=} \arg\max_b \exp\{e_m^T \cos[2\pi(\phi - B_\phi b)]\} \\
&\overset{(b)}{=} \arg\max_b \exp\{e_m^T \cos[z - \bar{z}]\}, z = 2\pi\phi, \bar{z} = 2\pi B_\phi b \\
&\overset{(c)}{=} \arg\max_b \frac{\exp\{e_m^T \cos[z-\bar{z}]\}}{\int_\Omega \exp\{e_m^T \cos[z-\bar{z}]\} dz}, \ \Omega = [-\pi, +\pi]^m \\
&\overset{(d)}{=} \arg\max_b \prod_{i=1}^m \frac{\exp\{\cos[z_i-\bar{z}_i]\}}{\int_{-\pi}^{+\pi} \exp\{\cos[z_i-\bar{z}_i]\} dz_i}, z_i = c_i^T z, \bar{z}_i = c_i^T \bar{z} \\
&\overset{(e)}{=} \arg\max_b \prod_{i=1}^m f(z_i|\bar{z}_i, 1) \\
\text{with} \quad & f(x|\mu,\kappa) = \frac{\exp\{\kappa \cos[x-\mu]\}}{\int_{-\pi}^{+\pi} \exp\{\kappa \cos[x-\mu]\} dx}
\end{aligned} \tag{8.8}$$

This result can be explained as follows. By taking the exponential exp in step (a), we obtain a nonnegative objective function which has the same maximizer as the original objective function. In step (b), we simplify the argument by setting $z = 2\pi\phi$ and $\bar{z} = 2\pi B_\phi b$. In step (c), we normalize the objective function such that it now can be interpreted as being a PDF. As function of z, it integrates to 1 over Ω, and as function of b, it still has $\check{b}$ as its maximizer. It is thus the multivariate PDF of the random vector z, having $\bar{z} = 2\pi B_\phi b$ as its parameter vector. As the PDF is symmetric about $\bar{z}$, $\bar{z} = 2\pi B_\phi b$ is also the mean of z. In step (d), we applied the property that the exponential of a sum can be written as a product of exponentials. As a result the multivariate PDF is written as an m-product of univariate PDFs. Here, c_i denotes the canonical unit vector having its only nonzero entry of 1 as its ith entry; Step (e) follows by recognizing that all m PDFs are the same, except for their means $\bar{z}_i$, $i = 1, \ldots, m$. Here we also recognize that the PDFs $f(z_i|\bar{z}_i, 1)$, $i = 1, \ldots, m$, are all special cases of the well-known *circular normal* distribution $\mathcal{CN}(\mu, \kappa)$, having $f(x|\mu, \kappa)$ as its PDF (Gumbel et al., 1953). The circular normal distribution $\mathcal{CN}(\mu, \kappa)$, with mean μ and *concentration* parameter κ, is also known as the *von Mises* distribution. Its denominator is given as $2\pi I_0(\kappa) = \int_{-\pi}^{+\pi} \exp\{\kappa \cos[x - \mu]\} dx$, where $I_0(\kappa)$ is the modified Bessel function of the first kind of order 0.

The above has shown that the solution $\check{b}$ of the ambiguity function method can now be interpreted as being the maximum of the likelihood function of $2\pi\phi$, if the m random variables $2\pi\phi_i$, $i = 1, \ldots, m$, are independent and distributed as $2\pi\phi_i \sim \mathcal{CN}(2\pi c_i^T B_\phi b, 1)$. Noting that this probabilistic result is obtained with circular normal distributions having unity concentration parameters, our above derivation now also shows how to introduce a weighting scheme into the ambiguity function method. If we replace the sum of cosines, $e_m^T \cos[2\pi(\phi - B_\phi b)]$, by the weighted sum $w^T \cos[2\pi(\phi - B_\phi b)]$, $w = [w_1, \ldots, w_m]^T$, a similar derivation as above shows, with $z_i = 2\pi c_i^T \phi$ and $\bar{z}_i = 2\pi c_i^T B_\phi b$, that

$$\arg\max_b w^T \cos[2\pi(\phi - B_\phi b)] = \arg\max_b \prod_{i=1}^m f(z_i|\bar{z}_i, w_i) \tag{8.9}$$

This shows how the classical objective function (8.4) needs to be changed in order to incorporate a varying precision of the GNSS carrier-phase measurements, thus allowing to include, for instance, frequency and/or elevation dependency in the precision description.

We are now in the position to summarize the above findings in a theorem. In order to do so, we will include the pseudorange data, and work, instead of with the phase-only system (8.3), with the extended partitioned system

$$\begin{bmatrix} p \\ \phi \end{bmatrix} = \begin{bmatrix} 0 & B_p \\ A_\phi & B_\phi \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} + \begin{bmatrix} \epsilon_p \\ \epsilon_\phi \end{bmatrix}, \; a \in \mathbb{Z}^n, \; b \in \mathbb{R}^p, \; A_\phi \in \mathbb{Z}^{m \times n} \tag{8.10}$$

in which the pseudoranges in p are expressed in units of range, while the carrier-phases in ϕ are still expressed in cycles.

By now assuming the pseudorange observables to be normally distributed and independent of the circular-normal distributed carrier-phase observables, we obtain a complete probabilistic model for the ambiguity function method and one that also eliminates the lack of uniqueness discussed in the previous section. We have the following result.

Theorem 2 (AF-ML estimator) Let $p \sim \mathcal{N}_m(B_p b, Q_{pp})$, with $\text{rank} B_p = p$, $Q_{pp} > 0$, be independent of $2\pi\phi \sim \prod_{i=1}^m \mathcal{CN}(2\pi c_i^T B_\phi b, w_i)$. Then the *likelihood function* of b is given as

$$L(b) = \frac{\exp\{-\frac{1}{2}||p - B_p b||^2_{Q_{pp}}\}}{|2\pi Q_{pp}|^{1/2}} \frac{\exp\{w^T \cos[2\pi(\phi - B_\phi b)]\}}{\prod_{i=1}^m 2\pi I_0(w_i)} \tag{8.11}$$

and its maximizer as

$$\check{b}_{\text{AF}} = \arg\max_{b \in \mathbb{R}^p} \left(-\tfrac{1}{2}||p - B_p b||^2_{Q_{pp}} + w^T \cos[2\pi(\phi - B_\phi b)] \right) \tag{8.12}$$

■

This result shows how the solution of the ambiguity function method can be interpreted as a maximum likelihood (ML) estimator, through which it also automatically inherits all the known properties of ML-estimators. Hence, by means of the above identified probabilistic model one can now apply known likelihood estimation and testing results.

Note, by making use of the trigonometric identity $\cos 2x = 1 - 2\sin^2 x$, that we may write the maximization-problem (8.12) also as a minimization-problem,

$$\check{b}_{\text{AF}} = \arg\min_{b \in \mathbb{R}^p} \left(||p - B_p b||^2_{Q_{pp}} + 4w^T \sin^2[\pi(\phi - B_\phi b)] \right) \tag{8.13}$$

If we assume B_p to be of full column rank, we may use the orthogonal decomposition $||p - B_p b||^2_{Q_{pp}} = ||P^\perp_{B_p} p||^2_{Q_{pp}} + ||\hat{b} - b||^2_{Q_{\hat{b}\hat{b}}}$, with $P^\perp_{B_p} = I_m - B_p B_p^+$, $\hat{b} = B_p^+ p$, $B_p^+ = Q_{\hat{b}\hat{b}} B_p^T Q_{pp}^{-1}$, $Q_{\hat{b}\hat{b}} = (B_p^T Q_{pp}^{-1} B_p)^{-1}$, and write (8.13) also as

$$\check{b}_{\text{AF}} = \arg\min_{b \in \mathbb{R}^p} \left(||\hat{b} - b||^2_{Q_{\hat{b}\hat{b}}} + 4w^T \sin^2[\pi(\phi - B_\phi b)] \right) \tag{8.14}$$

This formulation shows how the quadratic term on the right-hand side, and therefore the inclusion of pseudorange data through $\hat{b}$, acts as a *regularizer* on the phase-based part of the objective function. It is through this pseudorange-based regularization that the earlier mentioned nonuniqueness of the classical ambiguity function (8.4) is eliminated. Note in this regard that although a full rank of (8.10) guarantees uniqueness of the float LS-solution, that this not necessarily implies uniqueness of $\check{b}_{\mathrm{AF}}$. This is only the case if also B_p is of full rank. With single-differenced data for instance, B_p will be rank defect as the phase-clock cannot be determined from code data alone. In that case, the integerness of the phase-clock coefficients implies that the phase-clock can only be AF-determined up to an integer multiple of the wavelength.

To be able to work with (8.12) or (8.14), we still need to show how the weighting vector w can be chosen. For this we rely on the properties of the circular normal distribution and the fact that GNSS carrier-phase measurements are ultra-precise. The circular normal distribution $\mathcal{CN}(\mu, \kappa)$ is a symmetric unimodular distribution having its mode at μ. The parameter κ drives its peakedness or concentration. For $\kappa \to 0$, the circular normal distribution converges to the uniform distribution, while for $\kappa \to \infty$, it converges to the point distribution $\delta(x - \mu)$. As for large, but finite values of κ, the normal distribution $\mathcal{N}(\mu, \sigma^2 = \frac{1}{\kappa})$ provides a good approximation to $\mathcal{CN}(\mu, \kappa)$ (Gumbel et al., 1953), we will use the precision of the carrier-phase observables to set the entry values of the weight vector w. For example, if σ_i denotes the phase standard deviation when expressed in units of range, i.e. $\sigma_i = \sigma_{\lambda_i \phi_i}$ with λ_i the wavelength, then the variance of $2\pi\phi_i$ can be taken as the reciprocal value of the corresponding concentration parameter, i.e. $\frac{4\pi^2\sigma_i^2}{\lambda_i^2} = \frac{1}{w_i}$.

Example 1 *(Geometry-free model)* Consider the single-epoch, multi-frequency, double-differenced (DD) geometry-free model

$$\mathsf{E}\begin{bmatrix} p \\ \phi \end{bmatrix} = \begin{bmatrix} 0 & e_f \\ I_f & \Lambda_f^{-1} e_f \end{bmatrix} \begin{bmatrix} a \\ \rho \end{bmatrix} \tag{8.15}$$

with $p \in \mathbb{R}^f$ the DD pseudorange vector expressed in units of range, $\phi \in \mathbb{R}^f$ the DD carrier-phase vector expressed in cycles, $e_f = (1, \ldots, 1)^T$ the f-vector of ones, $\Lambda_f = \mathrm{diag}(\lambda_1, \ldots, \lambda_f)$ the diagonal matrix of f wavelengths, $a \in \mathbb{Z}^f$ the DD integer ambiguity vector and $\rho \in \mathbb{R}$ the scalar DD range. Then, with $\mathsf{D}(p) = \sigma_p^2 I_f$, the two parts of the objective function of (8.14) work out as

$$\begin{aligned} ||\hat{b} - b||^2_{Q_{\hat{b}\hat{b}}} &:= \left(\frac{\hat{\rho} - \rho}{\sigma_p / \sqrt{f}}\right)^2, \ \hat{\rho} = \tfrac{1}{f} \textstyle\sum_{i=1}^f p_i \\ 4w^T \sin^2[\pi(\phi - B_\phi b)] &:= 4 \textstyle\sum_{i=1}^f w_i \sin^2[\frac{\pi(\tilde{\phi}_i - \rho)}{\lambda_i}] \\ w_i &:= \frac{\lambda_i^2}{4\pi^2 \sigma^2_{\tilde{\phi}_i}} \end{aligned} \tag{8.16}$$

with $\sigma^2_{\tilde{\phi}_i}$ being the variance of $\tilde{\phi}_i = \lambda_i \phi_i$. □

8.5 Multimodal Ambiguity Function and its optimality domain

In this section we describe and illustrate some of the defining characteristics of the ambiguity function. In order to do so, we work from now on, for convenience sake, instead of with the objective function (8.12), with that of (8.13), i.e.

$$F(b) = ||p - B_p b||^2_{Q_{pp}} + 4w^T \sin^2[\pi(\phi - B_\phi b)] \tag{8.17}$$

Hence, since $F(b) = 2w^T e_m - 2\mathrm{AF}(b)$, the maximum likelihood solution of the AF-method is then computed, not as a maximizer, but as the minimizer of $F(b)$.

8.5.1 Minimization by Iterative Gradient Descent

As $F(b)$ is a smooth function (in fact, it is a C^∞ function, having continuous derivatives of all orders), one may think of applying methods of *iterative gradient descent* to obtains its minimum. Such methods adhere to the following scheme (Teunissen, 1990; Nocedal and Wright, 2006):

$$b_{k+1} = b_k - t_k Q(b_k) \partial_b F(b_k),\ k = 0, 1, \ldots \tag{8.18}$$

in which $t_k > 0$ is a to be chosen step size and $Q(b_k) > 0$ a to be chosen positive-definite matrix. The iteration is started by an externally provided initial approximation b_0 of the minimizer.

Through the choice of $Q(b_k)$, one can choose the direction of descent, and through the choice of t_k, one can enforce that $F(b_{k+1}) \leq F(b_k)$. For computing the stepsize t_k in each iteration, different linesearch strategies exist, from simple to advanced (Nesterov, 2018). One of the simplest starts with $t_k = 1$, followed by halving it, $t_k \leftarrow t_k/2$, until $F(b^k - t_k Q(b_k)\partial_b F(b^k)) < F(b^k)$. The simplest choice for $Q(b_k)$ would be to choose it as an identity matrix, $Q(b_k) = I_k$. As the resulting direction $-\partial_b F(b_k)$ points in the direction of steepest descent of $F(b)$ at b_k, this method is known as the *steepest descent method*. It has a local linear rate of convergence, but the iterations have the potential to zig-zag when the contours of $F(b)$ are elongated at the minimizer. This is avoided when $Q(b_k)$ is chosen as the inverse-Hessian of $F(b)$, $Q(b_k) = [\partial^2_{bb} F(b_k)]^{-1}$. This gives *Newton's method*, which is known to have a local quadratic convergence. Contrary to the steepest descent method, Newton's method does not need a linesearch strategy to enforce local convergence. That is, when the Hessian is positive-definite, the method has a quaranteed convergence for points sufficiently close to the solution. This is a consequence of the method being based on a linear approximation of the vanishing gradient of $F(b)$ at the minimizer: $0 = \partial_b F(\check{b}_{\mathrm{AF}}) \approx \partial_b F(b_k) + \partial^2_{bb} F(b_k)(\check{b}_{\mathrm{AF}} - b_k)$.

To apply the above iterative descent methods to (8.17) and verify whether or not a minimizer is obtained, the gradient and Hessian of $F(b)$ are needed. They are given as follows.

Lemma 1 (Ambiguity Function Gradient and Hessian): The gradient and Hessian of the objective function $F(b) = ||p - B_p b||^2_{Q_{pp}} + 4w^T \sin^2[\pi(\phi - B_\phi b)]$ are

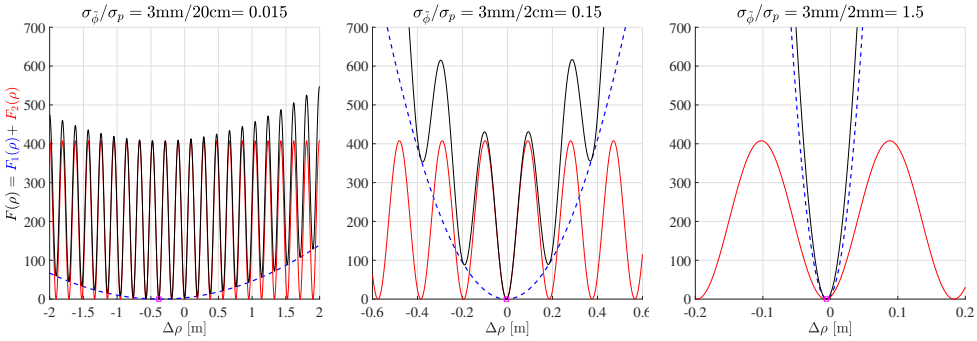


Figure 8.1: The $L1$ single-frequency, geometry-free ambiguity function $F(\rho) = F_1(\rho) + F_2(\rho)$ (cf. 8.21) is illustrated for different standard deviation values of the code and phase data. As the parabolic term $F_1(\rho)$ (in blue) rapidly increases for larger values $\epsilon = \sigma^2_{\tilde{\phi}}/\sigma^2_p$, the ambiguity function $F(\rho)$ becomes strictly convex for $\epsilon > 1$ (note: $\Delta\rho$ is the difference between the variable ρ and the range model value).

given as

$$\begin{aligned} \partial_b F(b) &= -2\left(B_p^T Q_{pp}^{-1}(p - B_p b) + B_\phi^T \varphi\right) \\ \partial^2_{bb} F(b) &= 2\left(B_p^T Q_{pp}^{-1} B_p + B_\phi^T D B_\phi\right), \end{aligned} \tag{8.19}$$

with

$$\begin{aligned} \varphi &= \sum_{i=1}^{m} (2\pi w_i \sin[2\pi c_i^T[\phi - B_\phi b)])c_i \\ D &= \mathrm{diag}(d_1(b), \ldots, d_m(b)) \\ d_i(b) &= 4\pi^2 w_i \cos[2\pi c_i^T(\phi - B_\phi b)] \end{aligned}$$

and where c_i is the ith canonical unit vector. ■

Note, due to the presence of the diagonal matrix D, that the Hessian matrix (cf. 8.19) of the ambiguity function is not necessarily positive-definite. Hence, it may not be invertible, or, when it is, it may not provide a descent direction. To avoid this from happening, one may think of regularizing the Hessian as $\partial^2_{bb} F(b) - 2B_\phi^T D B_\phi$, when $\partial^2_{bb} F(b)$ fails to be positive-definite. The so obtained regularized iteration can then again be interpreted as an iterative descent method, but now corresponding with the choice $Q(b_k) = [2(B_p^T Q_{pp}^{-1} B_p)]^{-1}$. Note that, with $t_k = 1$, this iteration boils down to

$$b_{k+1} = (B_p^T Q_{pp}^{-1} B_p)^{-1}[B_p^T Q_{pp}^{-1} p + B_\phi^T \varphi] \tag{8.20}$$

which can be seen to be the *fixed point iteration* of the system $\partial_b F(b) = 0$.

8.5.2 On the Ambiguity Function multimodality

Although the above descent methods converge to a minimum of $F(b)$, it depends on the initial approximation b_0 whether the minimum is a *local* or a *global* minimum of $F(b)$. One can have some confidence in having computed the global minimum, if b_0

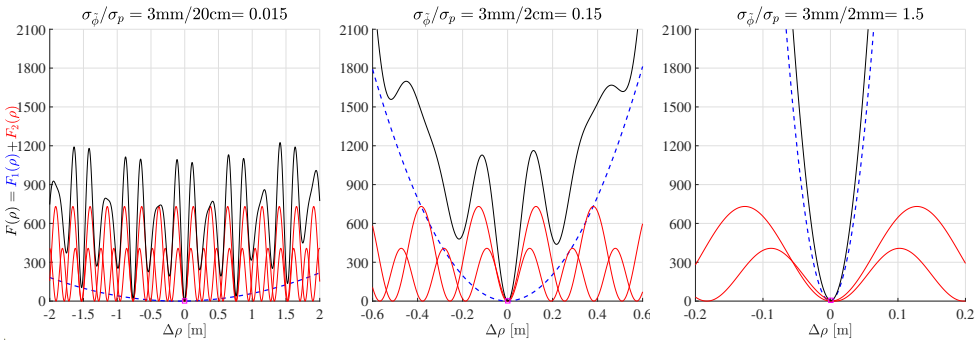


Figure 8.2: The $L1$-$L2$ dual-frequency, geometry-free ambiguity function $F(\rho) = F_1(\rho) + F_2(\rho)$ is illustrated for different standard deviation values of the code and phase data as presented in Fig.8.1. The term $F_2(\rho)$ is a sum of two sine squared functions, illustrated individually in the plots (red curves), each one having different amplitudes and periods.

would already be close enough to the global minimizer $\check{b}_{\mathrm{AF}}$. To see whether or not one can reasonably expect this to be the case, we take the ambiguity function of the single-frequency geometry-free model as an example. For the single-frequency case, the geometry-free ambiguity function follows from (8.16) as

$$F(\rho) = \tfrac{(p-\rho)^2}{\sigma_p^2} + 4w \sin^2\left(\tfrac{\pi}{\lambda}[\tilde{\phi} - \rho]\right), \; w = \tfrac{\lambda^2}{4\pi^2\sigma_{\tilde{\phi}}^2} \tag{8.21}$$

thus having first and second derivatives

$$\begin{aligned} d_\rho F(\rho) &= \tfrac{2}{\sigma_{\tilde{\phi}}^2}\left[\epsilon(\rho - p) - \tfrac{\lambda}{2\pi}\sin\left(\tfrac{2\pi}{\lambda}[\tilde{\phi} - \rho]\right)\right] \\ d_{\rho\rho}^2 F(\rho) &= \tfrac{2}{\sigma_{\tilde{\phi}}^2}\left[\epsilon + \cos\left(\tfrac{2\pi}{\lambda}[\tilde{\phi} - \rho]\right)\right] \end{aligned} \tag{8.22}$$

with ϵ being the phase-code variance ratio $\epsilon = \sigma_{\tilde{\phi}}^2/\sigma_p^2$. This shows that all those points where the straight line $y = \epsilon(\rho - p)$ intersects the sine-function $y = \frac{\lambda}{2\pi}\sin(\frac{2\pi}{\lambda}(\tilde{\phi} - \rho))$, are points for which $d_\rho F(\rho) = 0$. These are therefore the points where the local minima and maxima of the function $F(\rho)$ are located. Note that their number increases, when ϵ gets smaller, i.e. when the descending straight line $y = \epsilon(\rho - p)$ gets less tilted.

From the second derivative we learn that of these points, only those are minima for which $\cos(\frac{2\pi}{\lambda}(\tilde{\phi} - \rho)) > -\epsilon$ holds. This shows, since $\cos x \geq -1$ for all $x \in \mathbb{R}$, that $d_{\rho\rho}^2 F(\rho) > 0$ for all $\rho \in \mathbb{R}$, if $\epsilon > 1$. Hence, this is the condition for which $F(\rho)$ is convex. Thus if $\epsilon > 1$, then $F(\rho)$ has only a single minimum. As ϵ equals the phase-code variance ratio, this would require the precision of the pseudorange (code) observables to be better than that of the phase observables, which clearly is not the case with GNSS.

In fact, in case of GNSS it is the reciprocal value of ϵ that is large, i.e. the phase-code variance ratio is very small, $\epsilon \approx 10^{-4}$. This implies that the almost horizontal line $y = \epsilon(\rho - p)$ will have a large number of intersections with $y = \frac{\lambda}{2\pi}\sin(\frac{2\pi}{\lambda}(\tilde{\phi} - \rho))$. Hence, in the typical GNSS case, there will be a large number of minima and maxima

from which one would then need to select the global minimum. The ρ-values for the minima will have to satisfy $\cos(\frac{2\pi}{\lambda}(\tilde{\phi}-\rho)) > -\epsilon \approx 0$. If we define the small value δ_ϵ as satisfying $\epsilon + \cos(\frac{1}{2}\pi + 2\pi\delta_\epsilon) = 0$, then $d^2_{\rho\rho}F(\rho) > 0$ for all DD range values satisfying

$$\rho/\lambda \in \phi + z + (-\tfrac{1}{4} - \delta_\epsilon, +\tfrac{1}{4} + \delta_\epsilon), \forall z \in \mathbb{Z} \tag{8.23}$$

Thus of all the solutions satisfying $d_\rho F(\rho) = 0$, those ρ/λ that lie in one of the integer translated intervals (8.23) will be local minima and thus candidates for a global minimum.

To see how this set of 'integer translated intervals' generalizes to the multidimensional case, we need to study the region for which the $p \times p$ Hessian-matrix $\partial^2_{bb}F(b)$ is positive-definite. We have the following result.

Lemma 2 (On ambiguity function's convexity) Let $\Omega_{\mathrm{PD}} = \{b \in \mathbb{R}^p | \partial^2_{bb}F(b) > 0\}$ be the region in which the ambiguity function is convex and define the convex polytope

$$\Omega^\circ_{\mathrm{PD}}(z) = \{b \in \mathbb{R}^p |\ l(z) \preceq B_\phi b \preceq r(z)\}, z \in \mathbb{Z}^n \tag{8.24}$$

with $l(z) = (\phi - \delta_\epsilon - \frac{1}{4}e_m) + z$, $r(z) = (\phi + \delta_\epsilon + \frac{1}{4}e_m) + z$. Then

$$\Omega^\circ_{\mathrm{PD}} = \underset{z\in\mathbb{Z}^m}{\cup} \Omega^\circ_{\mathrm{PD}}(z) \subset \Omega_{\mathrm{PD}} \tag{8.25}$$

for $\delta_\epsilon = 0$, while if $B_\phi = \Lambda^{-1}B_p$, with Λ the diagonal wavelength matrix, and Q_{pp} is diagonal, relation (8.25) even holds with the entries of $\delta_\epsilon = (\delta_{\epsilon_1}, \ldots, \delta_{\epsilon_m})^T$ satisfying $\epsilon_i + \cos(\frac{1}{2}\pi + 2\pi\delta_{\epsilon_i}) = 0$, $\epsilon_i = \sigma^2_{\hat{\phi}_i}/\sigma^2_{p_i} = \lambda_i^2/(4\pi^2 w_i \sigma^2_{p_i})$, $i = 1, \ldots, m$. If then also B_p is invertible, we have $\Omega^\circ_{\mathrm{PD}} = \Omega_{\mathrm{PD}}$. ■

Proof: Relation (8.25), for $\delta_\epsilon = 0$, follows from the fact that $D > 0$ (cf. 8.19) implies $2(B_p^T Q_{pp}^{-1} B_p + B_\phi^T D B_\phi) = \partial^2_{bb}F(b) > 0$. Similarly, if $B_\phi = \Lambda^{-1}B_p$, then $Q_{pp}^{-1} + \Lambda^{-1}D\Lambda^{-1} > 0$ implies $2(B_p^T Q_{pp}^{-1} B_p + B_\phi^T D B_\phi) = \partial^2_{bb}F(b) > 0$, which gives (8.25) for the given δ_ϵ when Q_{pp} is diagonal. If also B_p is invertible, then $Q_{pp}^{-1} + \Lambda^{-1}D\Lambda^{-1} > 0$ is equivalent to $2(B_p^T Q_{pp}^{-1} B_p + B_\phi^T D B_\phi) = \partial^2_{bb}F(b) > 0$, which gives the equality $\Omega^\circ_{\mathrm{PD}} = \Omega_{\mathrm{PD}}$. □

This result shows how the region over which the ambiguity function is convex can be *inscribed* by a set of translated convex polytopes. This set is thus not guaranteed to contain the global minimizer. This is only true in the special case when B_p is invertible.

It will be clear from the above that for the typical GNSS case, i.e. when the phase-code variance ratio is small, the ambiguity function will show many local minima and maxima, and thus exhibit a pronounced multimodal variability. This is illustrated in Figure 8.1 for the single-frequency, geometry-free model and in Figure 8.2 for the dual-frequency, geometry-free model.

Due to the multimodality of $F(b)$ and the difficulty of knowing a-priori whether our initial approximation b_0 resides in the same convexity region as the global minimizer,

it will generally not be possible to solve for $\check{b}_{\text{AF}}$ by using one of the iterative gradient descent methods from the start. Fortunately, we do know how we can provide a convex bound to the region in which the search for the global minimizer can be conducted. We have the following result.

Lemma 3 (Optimality domain) Let $\hat{b} = \arg\min_{b\in\mathbb{R}^p} ||p - B_p b||^2_{Q_{pp}}$. Then

$$\begin{aligned} &\check{b}_{\text{AF}} = \arg\min_{b\in\mathbb{R}^p} F(b) \in \Omega, \text{ with} \\ &\Omega = \{b \in \mathbb{R}^p |\ ||p - B_p b||^2_{Q_{pp}} \leq F(\hat{b})\} \end{aligned} \tag{8.26}$$

■

Proof: Clearly $\check{b}_{\text{AF}} \in \Omega_{\text{F}} = \{b \in \mathbb{R}^p |\ F(b) \leq F(b_0)\}$ for any $b_0 \in \mathbb{R}^p$ and thus also for $b_0 = \hat{b}$. Then the result follows by recognizing that $||p - B_p b||^2_{Q_{pp}} \leq F(b)$ for all $b \in \mathbb{R}^p$. □

Geometrically, the region Ω of (8.26) captures all those $b \in \mathbb{R}^p$ of which the vectors $B_p b$ have a weighted squared distance to p that is not greater than $F(\hat{b})$. With B_p being of full column rank and by using the orthogonal decomposition $||p - B_p b||^2_{Q_{pp}} = ||P^{\perp}_{B_p} p||^2_{Q_{pp}} + ||\hat{b} - b||^2_{Q_{\hat{b}\hat{b}}}$, in which $Q_{\hat{b}\hat{b}} = (B_p^T Q_{pp}^{-1} B_p)^{-1}$, the region can also be written in ellipsoidal form,

$$\Omega = \{b \in \mathbb{R}^p |\ ||\hat{b} - b||^2_{Q_{\hat{b}\hat{b}}} \leq r^2\} \tag{8.27}$$

with $r^2 = F(\hat{b}) - ||P^{\perp}_{B_p} p||^2_{Q_{pp}}$. Note, since $\hat{b} \sim \mathcal{N}_p(b, Q_{\hat{b}\hat{b}})$, that Ω would become a confidence region with confidence level $1 - \alpha$, if r^2 would be chosen as $r^2 = \chi^2_\alpha(p, 0)$.

8

8.6 Global minimization of the Ambiguity Function

In this section we present our proposed method for finding the global minimizer $\check{b}_{\text{AF}}$ of the ambiguity function $F(b)$. The method is based on that of Teunissen and Massarweh (2024), be it that a different convex relaxing lower bounding function needs to be constructed. We therefore first provide a brief review of the characteristic components of the algorithm, followed by our construction of the required convex relaxation of the ambiguity function.

8.6.1 Branch-and-bound based minimization

As $\check{b}_{\text{AF}} \in \Omega$ (cf. 8.27), we can reformulate our original minimization problem $\min_{b\in\mathbb{R}^p} F(b)$ as the minimization of $F(b)$ over a bounded convex region, $\min_{b\in\Omega} F(b)$. This can be further simplified if we replace Ω by the ellipsoid-circumscribing box $\mathcal{C} = \{b \in \mathbb{R}^p |\ |b_\alpha - \hat{b}_\alpha| \leq r\sigma_{\hat{b}_\alpha}, \alpha = 1, \ldots, p\} \supset \Omega$. Hence, this brings our task to solving

$$\check{b}_{\text{AF}} = \arg\min_{b\in\mathcal{C}\subset\mathbb{R}^p} F(b) \tag{8.28}$$

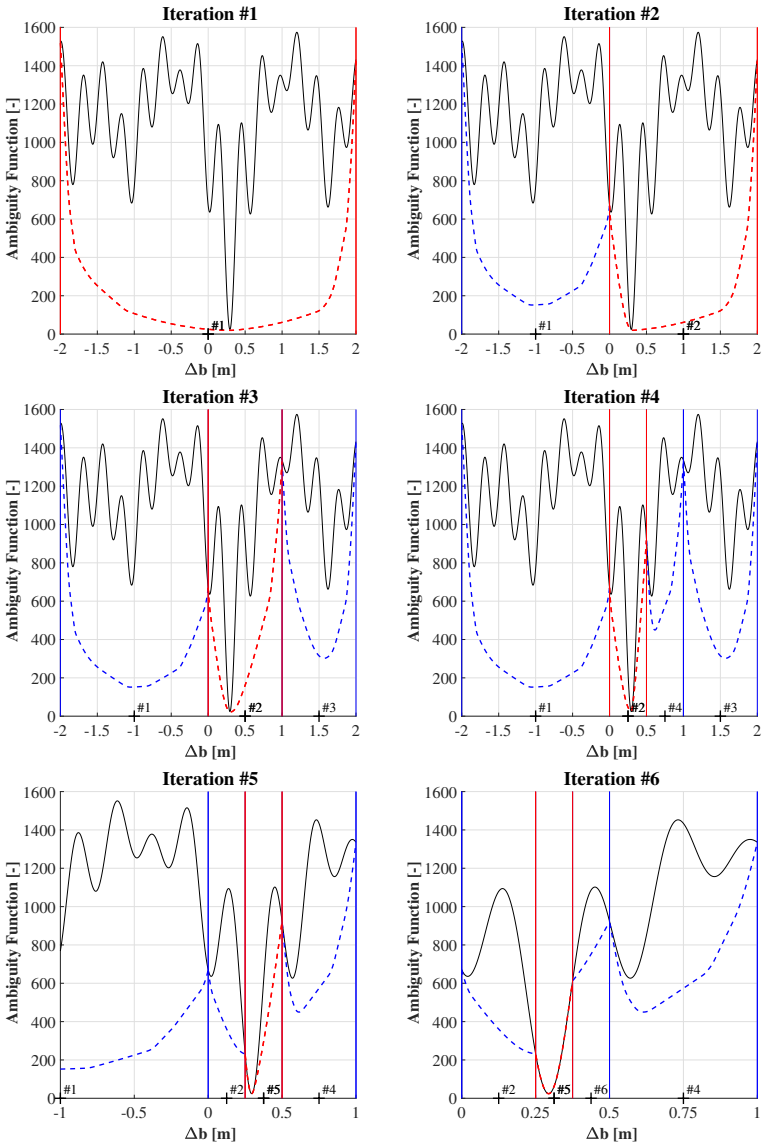


Figure 8.3: An illustrative example for $p = 1$, showing the branch-and-bound iterations for computing the global minimum of the ambiguity function (in black). In the first iteration, the interval $[-2, +2]$ is defined as search domain for the global optimum, where Δb refers to the initial value. The convex lower bound is represented by dashed lines for each interval, where at each iteration the most promising one (in red color) is halved, thus further isolating the global optimum.

The challenge in solving (8.28) is due to the multimodality of the ambiguity function. It is not convex and it has a multitude of local minima over $\mathcal{C}$. As this challenge is similar to the ones of dual mixed-integer least-squares computations, the same

algorithmic components will be used as in the method of Teunissen and Massarweh (2024), i.e. a branch-and-bound driven minimization, for which the required lower bounds are constructed from projected gradient descent solutions of a convex relaxed objective function.

Table 8.1: Overview of different cases and conditions for the definition of convex lower bound terms $g_{i,L}$.

Case	Condition	Function	Plot
$z_U = 0$	$[l_i, u_i] \in \left[0, \frac{1}{4}\right]$	$g_{i,L} = \sin^2(\pi x_i)$	Figure 8.4a
	$[l_i, u_i] \in \left[\frac{3}{4}, 1\right]$		
	$[l_i, u_i] \in \left[\frac{1}{4}, \frac{3}{4}\right]$	$g_{i,L} = \sin^2(\pi l_i) + (x_i - l_i)\frac{\sin^2(\pi u_i) - \sin^2(\pi l_i)}{u_i - l_i}$	Figure 8.4b
	$l_i \in \left[0, \frac{1}{4}\right], u_i \in \left[\frac{1}{4}, 1\right]$	$g_{i,L} = \mathrm{funL}(l_i, u_i, x_i)$	Figure 8.4c
	$l_i \in \left[0, \frac{3}{4}\right], u_i \in \left[\frac{3}{4}, 1\right]$	$g_{i,L} = \mathrm{funU}(l_i, u_i, x_i)$	Figure 8.4d
$z_U > 0$	$x_i < 1$	$g_{i,L} = \begin{cases} \mathrm{funU}(l_i, 1, x_i), & l_i < \frac{3}{4} \\ \sin^2(\pi x_i), & \text{otherwise} \end{cases}$	Figure 8.5
	$x_i \in [1, z_U]$	$g_{i,L} = 0$	
	$x_i > z_U$	$g_{i,L} = \begin{cases} \mathrm{funL}(0, u_i - z_U, x_i - z_U), & u_i > z_U + \frac{1}{4} \\ \sin^2(\pi x_i), & \text{otherwise} \end{cases}$	

Branch and Bound (BB)

Branch and bound algorithms (Lawler and Wood, 1966; Balakrishnan et al., 1991; Guida, 2015) are methods for global minimization of nonconvex problems. They are nonheuristic, in the sense that they maintain a provable upper and lower bound on the global minimum, i.e. they terminate with a guarantee that the computed solution has a prescribed accuracy. BB-algorithms involve partitioning a problem into subproblems (branching) and solving these subproblems to the optimal level, using bounds to eliminate the need to consider suboptimal solutions (bounding). Although there exist a large variety of different BB-mechanizations, we shall here restrict ourselves to the simple approach where the partitioning of the initial box $\mathcal{C} \subset \mathbb{R}^p$ is sequentially constructed through a splitting in half of the boxes. So at the first level, we start with the trivial partitioning, which is $\mathcal{C}$ itself, and compute the lower and upper bounds $L_1 = L(\mathcal{C})$ and $U_1 = U(\mathcal{C})$ (hence, these bounds are local and global at the same time),

$$L_1 = L(\mathcal{C}) \leq F(\check{b}_{\mathrm{AF}}) \leq U_1 = U(\mathcal{C}) \tag{8.29}$$

If $U_1 - L_1 \leq \epsilon$, the algorithm terminates (note: here the user-defined stop-criterion ϵ should not be confused with the phase-code variance ratio). Otherwise we go to the second iteration level and partition $\mathcal{C}$ into two boxes $\mathcal{C} = \mathcal{B}_1 \cup \mathcal{B}_2$, and compute $L(\mathcal{B}_i)$ and $U(\mathcal{B}_i)$, $i = 1, 2$. The splitting of the box is usually done along its longest edge.

Then we can construct new global lower and upper bounds,

$$\min(L(\mathcal{B}_1), L(\mathcal{B}_2)) \leq F(\check{b}_{\mathrm{AF}}) \leq \min(U(\mathcal{B}_1), U(\mathcal{B}_2)) \tag{8.30}$$

with $L_2 = \min(L(\mathcal{B}_1), L(\mathcal{B}_2))$ and $U_2 = \min(U(\mathcal{B}_1), U(\mathcal{B}_2))$. As both $\mathcal{B}_1$ and $\mathcal{B}_2$ are 'smaller' than $\mathcal{C}$ (i.e. they are its partition), one can generally expect the local bounds for $\mathcal{B}_i$ to get sharper in subsequent BB-iterations and the difference between upper and lower bounds to converge uniformly to zero, see (Balakrishnan et al., 1991).

If $U_2 - L_2 < \epsilon$, the algorithm terminates. Otherwise, we partition one of $\mathcal{B}_1$ and $\mathcal{B}_2$ into two boxes, to obtain a new partition of $\mathcal{C}$ into three boxes, and we compute the local lower and upper bounds for these new boxes. We then update the global lower bound L_3 as the minimum of the local lower bounds over the partition of $\mathcal{C}$, and similarly for the upper bound U_3. The choice which of the two boxes to split, $\mathcal{B}_1$ or $\mathcal{B}_2$, is based on the value of their local lower bound. The box to be split is the one of which the local lower bound equals the global lower bound, i.e. the one that has the smallest local lower bound. As at each iteration level a box is split into two, we have after k iterations a partitioning of the form $\mathcal{C} = \cup_{i=1}^{k} \mathcal{B}_i$, with associated global lower and upper bounds of $F(\check{b}_{\mathrm{AF}})$ as

$$L_k \leq \min_{b \in \mathcal{C}} F(b) \leq U_k \tag{8.31}$$

where $L_k = \min_{i=1,\ldots,k} L(\mathcal{B}_i)$ is nondecreasing, while $U_k = \min_{i=1,\ldots,k} U(\mathcal{B}_i)$ replaces U_{k-1} only if $U_k < U_{k-1}$, thus assuring that the global upper bound is nonincreasing (i.e. possibly speeding up the BB algorithm). Partitioning terminates if the difference of these bounds is small enough, $U_k - L_k \leq \epsilon$.

With reference to Lemma 3 (cf. 8.26), we note that the above procedure can be aided by applying box-shrinking, i.e. any b^* that has a function value $F(b^*)$ smaller than the previously used, can be used to shrink the set Ω.

Lower and Upper Bounds

For the bounds we need to be able to compute lower and upper bounds of $\min_{b \in \mathcal{B}} F(b)$ for any relevant box $\mathcal{B}$ that the BB-algorithm creates. The computation of local *lower* bounds $L(\mathcal{B})$ is the most challenging. We cannot use standard gradient-based methods for computing the minimizer over $\mathcal{B}$, since $F(b)$ is not necessarily convex over $\mathcal{B}$ and convergence is therefore not assured. The idea is therefore to find a differentiable convex lower bounding function $F_L(b) \leq F(b), \forall b \in \mathcal{B}$, such that the minimizer of $F_L(b)$ over $\mathcal{B}$ can be easily computed and used as the local lower bound

$$L(\mathcal{B}) = \min_{b \in \mathcal{B}} F_L(b) \leq \min_{b \in \mathcal{B}} F(b) \tag{8.32}$$

where $\mathcal{B} = \{b \in \mathbb{R}^p |\ b_L \preceq b \preceq b_U\}$. Our construction of $F_L(b)$ is described in Sect. 8.6.2.

In contrast to the lower bound, the computation of local *upper* bounds $U(\mathcal{B})$ is straightforward, since any $b_* \in \mathcal{B}$ can in principle be used for that purpose,

$$U(\mathcal{B}) = F(b_*) \geq \min_{b \in \mathcal{B}} F(b) \tag{8.33}$$

A simple choice is to take b_* as the 'centre of gravity' of the box, i.e. if box $\mathcal{B}$ is bounded as $b_L \preceq b \preceq b_U$, then $b_* = \frac{1}{2}(b_L + b_U)$. Another simple choice would be to take b_* as the minimizer of the lower bounding function, $b_* = \arg\min_{b\in\mathcal{B}} F_L(b)$ (cf. 8.32), the idea being that if $F_L(b)$ approximates $F(b)$ well in $\mathcal{B}$, then $F_L(b_*)$ should not differ too much from $F(b_*)$. A third option is to exploit the smoothness of the AF-function and compute b_* as a local minimizer of $F(b)$ over $\mathcal{B}$, thereby using, for instance, the projected gradient descent method (see below). This local minimizer becomes then automatically the global minimizer over $\mathcal{B}$, once the branching has reached the stage that $F(b)$ is truely convex over $\mathcal{B}$.

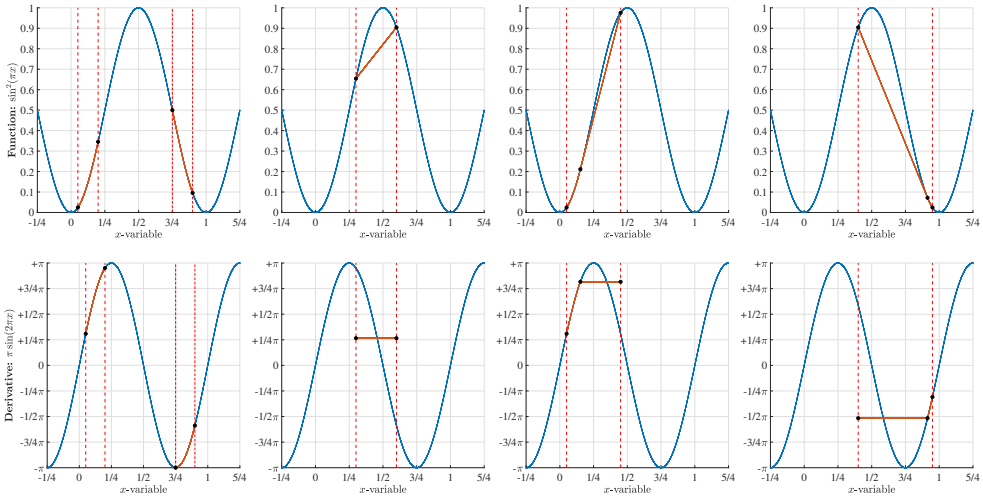


Figure 8.4: Illustration of the different convex lower bound terms $g_{i,L}$ (top) and their derivatives $g'_{i,L}$ (bottom) for the case $z_U = 0$ in Table 8.1. We consider $\{l = 0.05, u = 0.20\}$ and $\{l = 0.75, u = 0.90\}$ in the first column, $\{l = 0.30, u = 0.60\}$ in the second column, $\{l = 0.05, u = 0.45\}$ in the third column, and $\{l = 0.40, u = 0.95\}$ in the fourth column.

8

Projected Gradient Descent (PGD)

As our constructed lower bounding convex function $F_L(b)$ is only continuously differentiable, we use the PGD-method to solve for the lower bound (8.32). The algorithmic steps for doing so are as follows (Bertsekas, 1999; Nocedal and Wright, 2006):

1. Initialize: Start with a feasible solution, $b^0 \in \mathcal{B}$ and then loop for $k = 0, \ldots$ until the stop criterium $||b^{k+1} - b^k|| \leq \delta_{\text{PGD}}$ is satisfied, given a user-selected small threshold $\delta_{\text{PGD}} \ll 1$.

2. PGD step: Compute stepsize $t_k > 0$ and projected gradient descent

$$b^{k+1} = \text{P}_{\mathcal{B}}(b^k - t_k \partial_b F_L(b^k)) \tag{8.34}$$

such that $F_L(b^{k+1}) \leq F_L(b^k)$, where

$$\text{P}_{\mathcal{B}}(y) = \arg\min_{b\in\mathcal{B}} ||y - b||^2 \tag{8.35}$$

As $\mathcal{B}$ is a box, the orthogonal projection $\mathrm{P}_{\mathcal{B}} : \mathbb{R}^p \to \mathcal{B}$ can be computed very efficiently. We have for $\alpha = 1, \ldots, p$,

$$\mathrm{P}_{\mathcal{B}}(y)_\alpha = \begin{cases} b_{\alpha,L} & \text{if} \quad y_\alpha \leq b_{\alpha,L} \\ y_\alpha & \text{if} \quad b_{\alpha,L} \leq y_\alpha \leq b_{\alpha,U} \\ b_{\alpha,U} & \text{if} \quad b_{\alpha,U} \leq y_\alpha \end{cases} \tag{8.36}$$

in which $b_{\alpha,L}$, y_α, and $b_{\alpha,U}$ denote the α-coordinates of b_L, y, and b_U, respectively.

Example 2 *(BB height-determination)* Figure 8.3, first panel, shows, over an initial interval $[-2, +2]$ meters, the ambiguity function (8.17) of an $L1$ single-frequency, DD short-baseline model for which only the height difference is assumed unknown, i.e. $p = 1$. The true, simulated, height-difference is 30.0cm, while the AF-minimizer is found to be 30.3cm, using a pseudo range (code) and phase measurement precision of 20cm and 2mm, respectively (note: Δb is shown with respect to initial value).

In the first iteration, the initial search interval $[-2, +2]$ is taken and the convex lower bounding function (red dashed line) is evaluated. This interval represents our initial box $C \in \mathbb{R}^p$ for $p = 1$, which is therefore split in two for the second iteration. For both intervals, we compute the respective convex lower bounding functions (blue and red dashed line) and their minimum value, i.e. convex lower bound (CLB). The one with lowest CLB value is taken as most promising interval, here marked in red color. For the upper bound computations, we simply consider the center of current intervals, i.e. $b_* = (b_L + b_U)/2$, and compute $F(b_*)$. As the intervals get shorter, the difference between upper and lower local bounds get smaller as well.

The process is repeated and with further iterations the global minimum is isolated in smaller intervals where the difference between $F(b)$ and $F_L(b)$ also gets smaller, until convergence. Note that in the last iteration, we have isolated the global minimum in an interval where the ambiguity function is convex and iterative gradient descent could have been adopted, see Section 8.5.1. □

8.6.2 Convex relaxation of the Ambiguity Function

We now determine our differentiable convex lower bounding (CLB) function of

$$\begin{aligned} F(b) &= ||p - B_p b||^2_{Q_{pp}} + 4w^T \sin^2[\pi(\phi - B_\phi b)] \\ &= ||p - B_p b||^2_{Q_{pp}} + \underbrace{4 \sum_{i=1}^{m} w_i \sin^2[\pi c_i^T (\phi - B_\phi b)]}_{G(b)} \end{aligned} \tag{8.37}$$

As the first term on the right-hand side is already convex, we concentrate on finding a convex relaxation of the second term $G(b)$. We thus aim to find a differentiable convex function $G_L : \mathcal{B} \to \mathbb{R}$, such that

$$G_L(b) \leq G(b), \ \forall b \in \mathcal{B} \tag{8.38}$$

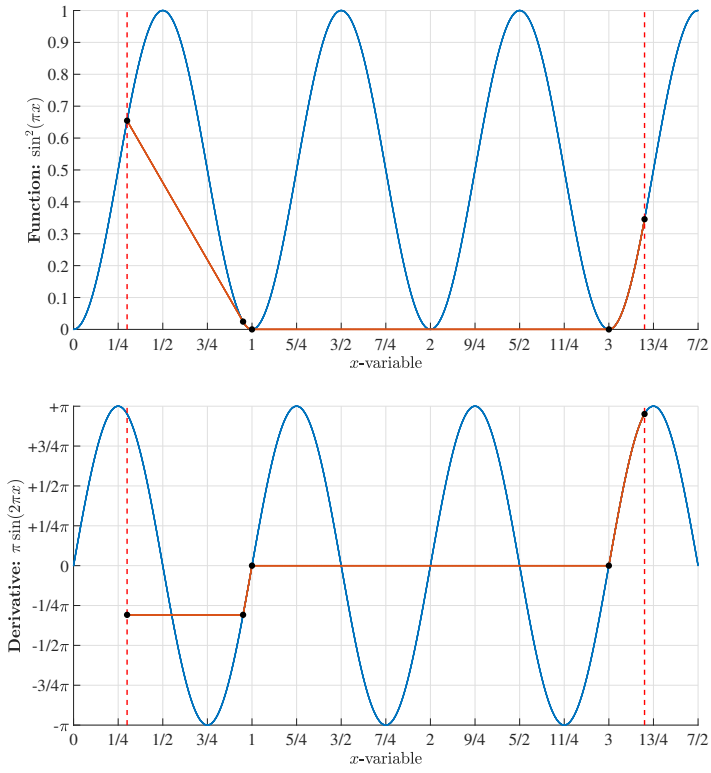


Figure 8.5: Illustration of the convex lower bound term $g_{i,L}$ given the case $z_U > 0$ described in Table 8.1, along with the associated derivative terms $g'_{i,L}$, where we consider for this example a particular interval $\{l = 0.3, u = 3.2\}$.

8

so that we obtain the convex lower bounding (CLB) function

$$F_L(b) = ||p - B_p b||^2_{Q_{pp}} + G_L(b) \leq F(b), \ \forall b \in \mathcal{B} \tag{8.39}$$

with its gradient given by

$$\partial_b F_L(b) = -2B_p^T Q_{pp}^{-1}(p - B_p b) + \partial_b G_L(b) \tag{8.40}$$

Given that $G(b)$ is a summation of scalar terms,

$$G(b) = 4\sum_{i=1}^{m} w_i g_i(x_i(b)), \ g_i(x_i(b)) = \sin^2(\pi x_i(b)) \tag{8.41}$$

where $x_i(b) = c_i^T(\phi - B_\phi b) \in \mathbb{R}$, we may seek a CLB function of similar structure,

$$G_L(b) = 4\sum_{i=1}^{m} w_i g_{i,L}(x_i(b)) \tag{8.42}$$

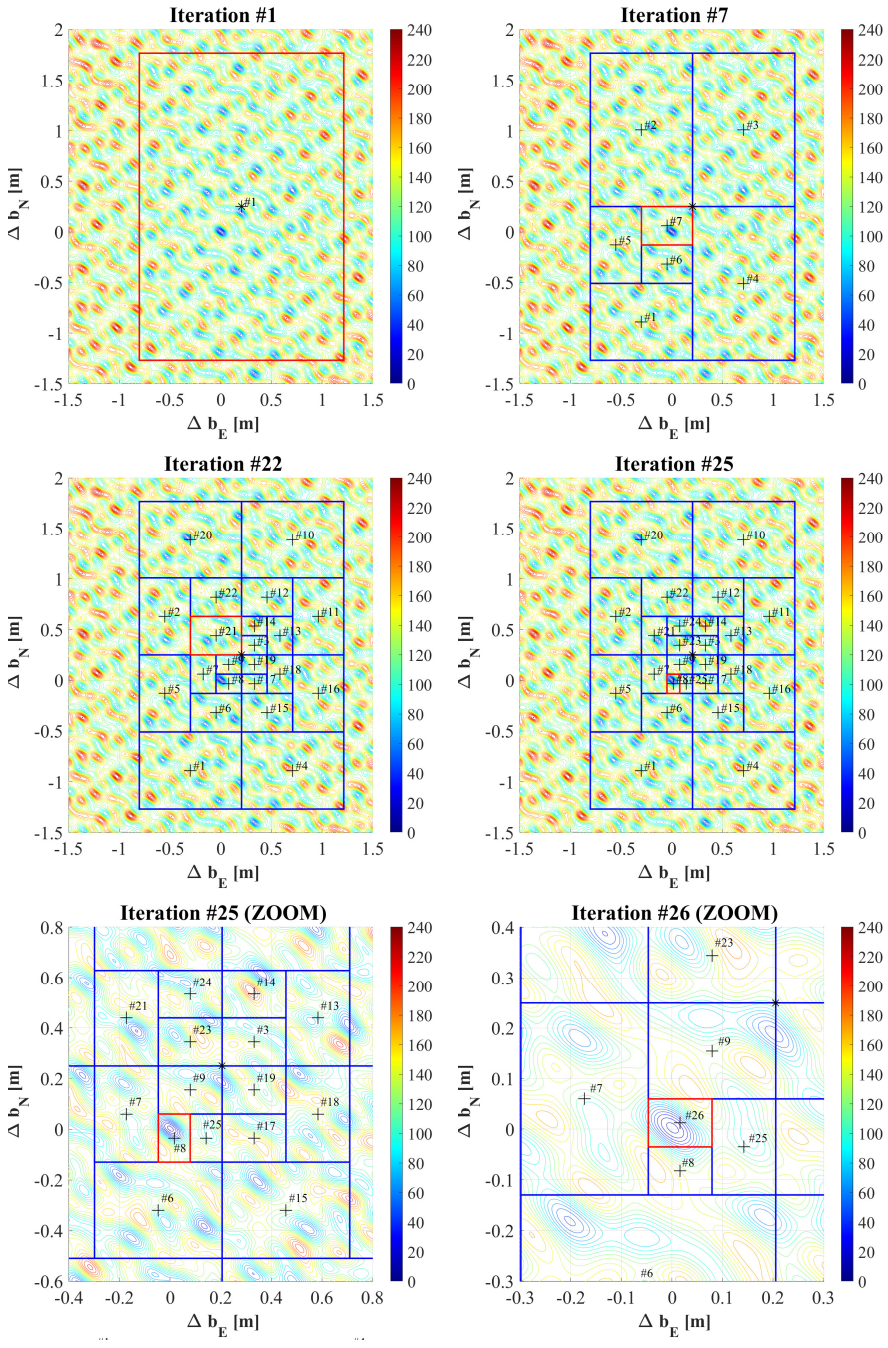


Figure 8.6: A branch-and-bound iteration towards the global minimum of the ambiguity function $F(b)$ (cf. 8.17) for an $L1$ single-frequency, DD short-baseline model with the horizontal position coordinates assumed unknown, where at each iteration the most promising box (in red color) is halved.

such that $g_{i,L}(x_i(b)) \leq g_i(x_i(b)), \forall b \in \mathcal{B}$, thus satisfying (8.38). The box constraint implies that also $x_i(b)$ is bounded in an interval $[l_i, u_i]$ for which $g_{i,L}$ is required to be convex. These intervals can be found based on the projection-lemma described in Teunissen and Massarweh (2024).

At this point we provide the definition of $g_{i,L}$ for $x_i \in [l_i, u_i]$, and different cases should be distinguished as summarized in Table 8.1. Notice that we define $z_L = \lfloor l_i \rfloor$ and $z_U = \lfloor u_i \rfloor$, where $\lfloor \cdot \rfloor$ refers to the floor function. Hence, we subtract z_L from the aforementioned quantities, so they are re-defined as $x_i := x_i - z_L$, $l_i := l_i - z_L$, $u_i := u_i - z_L$, and $z_U := z_U - z_L$. Thus, we always have $l_i \in [0, 1]$ with $u_i \in (l_i, 1)$ if $z_U = 0$ or $u_i > 1$ if $z_U > 0$. These two cases are separated in Table 8.1, and are respectively illustrated in Figure 8.4 and Figure 8.5.

When constructing these CLB terms, we make use of two auxiliary functions 'funL' and 'funU' that are defined as:

Auxiliary Function #1:

$$\text{funL}(l, u, x) = \begin{cases} \sin^2(\pi l) + (x - l)\frac{\sin^2(\pi u) - \sin^2(\pi l)}{u - l}, & l \geq x_T \\ \sin^2(\pi x), & l < x_T, x \leq x_T \\ \sin^2(\pi x_T) + \pi(x - x_T)\sin(2\pi x_T), & l < x_T, x > x_T \end{cases} \tag{8.43}$$

where $x_T \in [0, \frac{1}{4}]$ is obtained from

$$\sin^2(\pi u) = \sin^2(\pi x_T) + \pi(u - x_T)\sin(2\pi x_T) \tag{8.44}$$

Auxiliary Function #2:

$$\text{funU}(l, u, x) = \begin{cases} \sin^2(\pi l) + (x - l)\frac{\sin^2(\pi u) - \sin^2(\pi l)}{u - l}, & u \leq x_T \\ \sin^2(\pi x), & u > x_T, x \geq x_T \\ \sin^2(\pi x_T) + \pi(x - x_T)\sin(2\pi x_T), & u > x_T, x < x_T \end{cases} \tag{8.45}$$

where $x_T \in \left[\frac{3}{4}, 1\right]$ is obtained from

$$\sin^2(\pi l) = \sin^2(\pi x_T) + \pi(l - x_T)\sin(2\pi x_T) \tag{8.46}$$

At the same time, the derivative in $b \in \mathbb{R}^p$ of each CLB term $g_{i,L}$ follows as

$$\partial_b g_{i,L}(x_i(b)) = g'_{i,L}(x_i)\partial_b x_i(b) = -g'_{i,L}(x_i)B_\phi^T c_i \tag{8.47}$$

given $x_i \in [l_i, u_i]$, while $g'_{i,L}(x_i) = \partial_{x_i} g_{i,L}(x_i)$ can be easily computed based on the elementary expressions provided in Table 8.1, and has been illustrated for the two cases $z_U = 0$ and $z_U > 0$, respectively in Figure 8.4 and Figure 8.5. The gradient $\partial_b G_L(b)$ is ultimately computed as

$$\partial_b G_L(b) = -4\sum_{i=1}^{m} w_i g'_{i,L}(x_i)B_\phi^T c_i \tag{8.48}$$

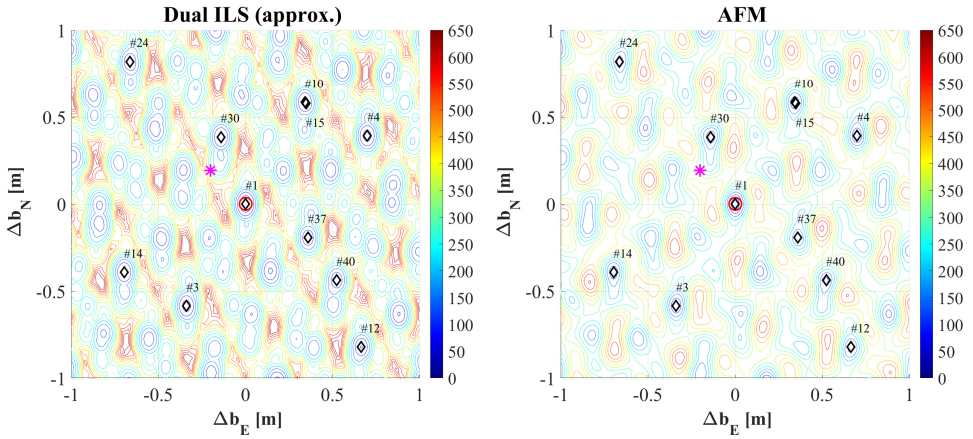


Figure 8.7: Contourlines zoom-in of the dual-ILS function $\mathcal{D}^\circ(b)$ (left) and ambiguity function $F(b)$ (right) for the single-epoch model of example 3. The colourbar scale of $\mathcal{D}^\circ(b)$ is half that of $F(b)$. The float solution $\hat{b}$ is shown as a magenta asterisk and the AF-solution $\breve{b}_{\mathrm{AF}} = \arg\min_b F(b)$ as a red circle. The ten black diamonds show the locations of 10 out of the 40 smallest local minima of $\mathcal{D}^\circ(b)$, with diamond #1 showing the location of the ILS-solution $\breve{b}^\circ = \arg\min_b \mathcal{D}^\circ(b) = \hat{b}^\circ(\breve{a}^\circ)$, with $\breve{a}^\circ = \arg\min_{a\in\mathbb{Z}^n} ||\hat{a}-a||^2_{Q^\circ_{\hat{a}\hat{a}}}$.

and $\partial_b F_L(b)$ will resemble the expression shown in (8.19), after substituting $\pi\sin(2\pi x_i)$ with $g'_{i,L}(x_i)$ since we are considering now the convex lower bound of $F(b)$.

Example 3 *(BB position determination)* In analogy with example 2, this example illustrates the BB-iterations towards the global minimum of the ambiguity function $F(b)$ (cf. 8.17) in case of an $L1$ single-frequency, DD short-baseline model for which only the horizontal position coordinates are assumed unknown. The panels of Figure 8.6 show the contourlines of $F(b)$, together with the per iteration-step increasing box-densification. Shown are the results for iterations #1, #7, #22, #25, together with a zoom-in of the last two iteration steps, #25 and #26. The red box is every time the most promising box to be split. It is the box with lowest CLB value, computed from the convex lower bounding functions that we introduced in Section 8.6.2. □

8.7 The AF- and LS-principle compared

The fact that the ambiguity objective function (8.17) has a 'sum-of-squares' structure and that the approximation $\sin^2(x) \approx x^2$ holds for small x, has led some authors to link the AF-principle to that of least-squares, e.g. (Rogers et al., 1978; Lachapelle et al., 1992; Leick et al., 2015), with some even stating that the two are 'fundamentally equivalent'. Although we show in this section that this statement is incorrect, we also show under which identified circumstances one can expect AF- and ILS-solutions to have similar behaviour.

8.7.1 A least-squares relation

To make a strict comparison between the AF- and ILS-objective functions possible, we rely on the primal-dual equivalence of mixed-ILS theory as introduced in (Teunissen and Massarweh, 2024). We therefore first summarize the for current purposes relevant material in the following theorem (see Section 4 of (ibid)).

Theorem 3 (Primal-Dual mixed ILS): Let the dispersion of the float ambiguity and baseline estimators, $\hat{a} \in \mathbb{R}^n$ and $\hat{b} \in \mathbb{R}^p$, be given as

$$\mathrm{D}\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = \begin{bmatrix} Q_{\hat{a}\hat{a}} & Q_{\hat{a}\hat{b}} \\ Q_{\hat{b}\hat{a}} & Q_{\hat{b}\hat{b}} \end{bmatrix} \tag{8.49}$$

and define the baseline objective function as

$$\mathcal{D}^\circ(b) = ||\hat{b} - b||^2_{Q_{\hat{b}\hat{b}}} + ||\hat{a}(b) - \breve{a}^\circ(b)||^2_{Q^\circ_{\hat{a}(b)\hat{a}(b)}} \tag{8.50}$$

with $\hat{a}(b) = \hat{a} - Q_{\hat{a}\hat{b}}Q^{-1}_{\hat{b}\hat{b}}(\hat{b} - b)$, $\breve{a}^\circ(b) = \arg\min_{z\in\mathbb{Z}^n} ||\hat{a}(b) - z||^2_{Q^\circ_{\hat{a}(b)\hat{a}(b)}}$ and $Q^\circ_{\hat{a}(b)\hat{a}(b)}$ being an approximation of the actual baseline-conditioned ambiguity variance matrix $Q_{\hat{a}(b)\hat{a}(b)} = Q_{\hat{a}\hat{a}} - Q_{\hat{a}\hat{b}}Q^{-1}_{\hat{b}\hat{b}}Q_{\hat{b}\hat{a}}$. Then the minimizer of $\mathcal{D}^\circ(b)$,

$$\breve{b}^\circ = \arg\min_{b\in\mathbb{R}^p} \mathcal{D}^\circ(b) \tag{8.51}$$

is the solution of the mixed ILS-problem

$$\left.\begin{matrix} \breve{a}^\circ \\ \breve{b}^\circ \end{matrix}\right\} = \arg\min_{a\in\mathbb{Z}^n, b\in\mathbb{R}^p} \left(||\hat{a} - a||^2_{Q^\circ_{\hat{a}\hat{a}}} + ||\hat{b}^\circ(a) - b||^2_{Q^\circ_{\hat{b}(a)\hat{b}(a)}} \right)$$

where

$$\begin{array}{lcl} Q^\circ_{\hat{a}\hat{a}} & = & Q^\circ_{\hat{a}(b)\hat{a}(b)} + Q_{\hat{a}\hat{b}}Q^{-1}_{\hat{b}\hat{b}}Q_{\hat{a}\hat{b}} \\ \hat{b}^\circ(a) & = & \hat{b} - Q_{\hat{b}\hat{a}}Q^{\circ-1}_{\hat{a}\hat{a}}(\hat{a} - a) \\ Q^\circ_{\hat{b}(a)\hat{b}(a)} & = & Q_{\hat{b}\hat{b}} - Q_{\hat{b}\hat{a}}Q^{\circ-1}_{\hat{a}\hat{a}}Q_{\hat{a}\hat{b}} \end{array}$$

■

Proof: See Section 4 of (Teunissen and Massarweh, 2024). □

From the above theorem two important conclusions can be drawn. First, if $Q^\circ_{\hat{a}(b)\hat{a}(b)}$ is chosen to be equal to $Q_{\hat{a}(b)\hat{a}(b)}$, then also $Q^\circ_{\hat{a}\hat{a}} = Q_{\hat{a}\hat{a}}$ and $\hat{b}^\circ(a) = \hat{b}(a)$, from which follows that $\breve{b}^\circ = \arg\min_{b\in\mathbb{R}^p} \mathcal{D}^\circ(b)$ will be identical to

$$\breve{b} = \hat{b}(\breve{a}) \text{ with } \breve{a} = \arg\min_{a\in\mathbb{Z}^n} ||\hat{a} - a||^2_{Q_{\hat{a}\hat{a}}} \tag{8.52}$$

Thus with the choice $Q^\circ_{\hat{a}(b)\hat{a}(b)} = Q_{\hat{a}(b)\hat{a}(b)}$, the baseline solution $\breve{b}^\circ = \arg\min_{b\in\mathbb{R}^p} \mathcal{D}^\circ(b)$ is identical to the ILS-baseline estimator $\breve{b}$, as a consequence of which it will also share

its statistical optimality properties such as having a maximum ambiguity success-rate (Teunissen, 1999).

The second conclusion that can be drawn from the above theorem is that even if $Q^{\circ}_{\hat{a}(b)\hat{a}(b)} \neq Q_{\hat{a}(b)\hat{a}(b)}$, the minimizer $\check{b}^{\circ} = \arg\min_{b\in\mathbb{R}^p} \mathcal{D}^{\circ}(b)$ is still an ILS-baseline estimator, namely

$$\check{b}^{\circ} = \hat{b}^{\circ}(\check{a}^{\circ}) \text{ with } \check{a}^{\circ} = \arg\min_{a\in\mathbb{Z}^n} ||\hat{a}-a||^2_{Q^{\circ}_{\hat{a}\hat{a}}} \tag{8.53}$$

but now one which uses an incorrect ambiguity-weighting through $Q^{\circ}_{\hat{a}\hat{a}}$ and an incorrect baseline-mapping through $\hat{b}^{\circ}(a)$; compare (8.52) with (8.53). Hence, (8.53) will not have the optimality properties of (8.52). Using the statistical and distributional properties of $\check{a}^{\circ}$ and $\check{b}^{\circ}$ as given in (Teunissen and Massarweh, 2024), one can study by how much these properties differ from those of the optimal estimators in (8.52).

As the above makes clear, this departure from optimality is driven by the difference between $Q^{\circ}_{\hat{a}(b)\hat{a}(b)}$ and $Q_{\hat{a}(b)\hat{a}(b)}$. In (ibid) we discussed some cases for which the estimators $\check{b}^{\circ}$ and $\check{b}$ performed similarly due to their small difference between the matrices $Q^{\circ}_{\hat{a}(b)\hat{a}(b)}$ and $Q_{\hat{a}(b)\hat{a}(b)}$.

For our current AF-analysis, the relevance of the above Theorem 3 lies now in the fact that with a special choice for $Q^{\circ}_{\hat{a}(b)\hat{a}(b)}$, the ILS-baseline producing objective function $\mathcal{D}^{\circ}(b)$ can be given a structure that closely resembles that of the ambiguity objective function $F(b)$. For the ambiguity objective function given in (8.17), assuming $w_i = \frac{1}{4\pi^2\sigma^2_{\phi_i}}$, we have

$$F(b) = ||\hat{b}-b||^2_{Q_{\hat{b}\hat{b}}} + \sum_{i=1}^{m} \frac{\sin^2\big(\pi c_i^T(\phi - B_\phi b)\big)}{\pi^2\sigma^2_{\phi_i}} \tag{8.54}$$

A very similar structure is obtained for $\mathcal{D}^{\circ}(b)$ if $Q^{\circ}_{\hat{a}(b)\hat{a}(b)} = \text{diag}(\sigma^2_{\hat{a}_1(b)}, \ldots, \sigma^2_{\hat{a}_n(b)})$. With this choice we have $\check{a}^{\circ} = \arg\min_{a\in\mathbb{Z}^n} ||\hat{a}-a||^2_{Q^{\circ}_{\hat{a}\hat{a}}} = (\lceil \hat{a}_1(b)\rfloor, \ldots, \lceil \hat{a}_n(b)\rfloor)^T$, and therefore $\mathcal{D}^{\circ}(b) = ||\hat{b}-b||^2_{Q_{\hat{b}\hat{b}}} + \sum_{i=1}^{n} \frac{1}{\sigma^2_{\hat{a}_i(b)}} \big(\hat{a}_i(b) - \lceil \hat{a}_i(b)\rfloor\big)^2$, which can be worked out as

$$\begin{aligned} \mathcal{D}^{\circ}(b) \quad &= \quad ||\hat{b}-b||^2_{Q_{\hat{b}\hat{b}}} + \sum_{i=1}^{n} \frac{d^2\big(c_i^T A^+_\phi(\phi - B_\phi b)\big)}{\sigma^2_{(A^+_\phi\phi)_i}} \\ &\text{with} \\ d(x) \quad &= \quad (x - \lceil x\rfloor) \\ A^+_\phi \quad &= \quad (A_\phi^T Q^{-1}_{\phi\phi} A_\phi)^{-1} A_\phi^T Q^{-1}_{\phi\phi} \\ \sigma^2_{(A^+_\phi\phi)_i} \quad &= \quad c_i^T (A_\phi^T Q^{-1}_{\phi\phi} A_\phi)^{-1} c_i \end{aligned} \tag{8.55}$$

The two expressions, (5.6) and (8.55), look very similar, but with the following marked differences,

1. The functions used in the two respective sums are $\frac{\sin^2(\pi x)}{\pi^2}$ in (5.6) and $(x-\lceil x \rfloor)^2$ in (8.55).
2. The arguments used in the two functions are $c_i^T(\phi - B_\phi b)$ in (5.6) and $c_i^T A_\phi^+(\phi - B_\phi b)$ in (8.55), while normalized with the variances $\sigma^2_{\phi_i}$ in (5.6) and $\sigma^2_{(A_\phi^+ \phi)_i}$ in (8.55).
3. Furthermore, as a consequence, the sum in (5.6) is over m, the dimension of the phase vector ϕ, while the sum in (8.55) is over n, the dimension of the ambiguity vector a.

In the absence of these differences, the two functions $F(b)$ and $\mathcal{D}^\circ(b)$ would be identical and the ambiguity function determined baseline $\breve{b}_{\text{AF}} = \arg\min_{b\in\mathbb{R}^p} F(b)$ would be a true ILS-baseline, albeit one determined from an incorrect ambiguity variance matrix.

Note that the second of the above differences (and implictly the third as well) is due to the fact that $\mathcal{D}^\circ(b)$ is based on a conditional least-squares ambiguity estimation, while this is not true for $F(b)$. The ambiguity function $F(b)$ works namely directly on the phase data and it thus therefore not exploit any time-constancy in the ambiguities if such would be present. Hence, if we assume to work with DD phase data in either a single-epoch model or a multi-epoch model in which all ambiguities are disconnected in time, then $A_\phi = I_m$, from which follows that the last two of the above differences disappear, since then $A_\phi^+ = I_m$ and $m = n$. As immunity for cycle slips is considered one of the attractive features of AF methods, we summarize the properties for the case $A_\phi = I_m$ separately in the following corollary.

Corollary 1 (AF as ILS approximation) If $A_\phi = I_m$, replacement of $\frac{\sin^2 \pi x}{\pi^2}$ in (5.6) by $(x - \lceil x \rfloor)^2$ turns the ambiguity function $F(b)$ into $\mathcal{D}^\circ(b)$, the minimizer of which is the ILS-baseline $\breve{b}^\circ$ (cf. 8.53) with $Q^\circ_{\hat{a}(b)\hat{a}(b)} = \text{diag}(\sigma^2_{\hat{a}_1(b)}, \ldots, \sigma^2_{\hat{a}_n(b)})$. ■

8

This result shows that the extend to which the AF-solution $\breve{b}_{\text{AF}} = \arg\min_{b\in\mathbb{R}^p} F(b)$ can be considered an approximate ILS-solution hinges on the approximation of $(x-\lceil x \rfloor)^2$ by $\frac{\sin^2 \pi x}{\pi^2}$. Although the maxima of these two functions are quite different, their minima are identical and their local behaviour around these minima is also very similar. For $x = z + \delta$, with $z \in \mathbb{Z}$ and $\delta =$ small, we have namely $\frac{\sin^2 \pi x}{\pi^2} \approx \delta^2 = (x - \lceil x \rfloor)^2$. This shows that under the condition of Corollary 1, one can indeed expect the minimizers of $F(b)$ and $\mathcal{D}^\circ(b)$ to be quite close.

This is illustrated, for the model of example 3, in Figure 8.7. To the left we have the contourlines of $\mathcal{D}^\circ(b)$ and to the right those of $F(b)$. Apart from their difference in scale, the two contour plots show a very similar topography, with their local minima at almost identical locations. The ten black diamonds, for instance, show the locations of 10 out of the 40 smallest local minima of $\mathcal{D}^\circ(b)$, but they are at the same time also very close to the corresponding local minima of $F(b)$. And this also holds true for their global minimum, with the black diamond #1 identifying the ILS-solution $\breve{b}^\circ = \arg\min_b \mathcal{D}^\circ(b)$ and the red circle identifying $\breve{b}_{\text{AF}} = \arg\min_b F(b)$.

As the strength of the underlying model is not in play in the properties captured by Corollary 1, the close-to-ILS behaviour of the AF-solution $\check{b}_{\mathrm{AF}}$ will not change when varying the ambiguity success-rate of the model. In Table 8.2 we show the formal standard deviation (σ_{UP}) and the simulated RMS-values of $\hat{b}$ (float), $\check{b}^{\circ}$ (ILS), and $\check{b}_{\mathrm{AF}}$, for three different success-rates (SR) of the model used in example 2. The success-rates were reduced by reducing the number of satellites from 8 via 6 to 5. These results show the consistency between σ_{UP} and the float-RMS values and how the AF-RMS values follow those of ILS. The RMS-values of AF and ILS get poorer for smaller success-rates and for very small success-rates even poorer than the RMS of the float solution. This shows just as for ILS, it is the ambiguity success-rate that plays a decisive role in the quality of the AF-solution. Hence, although the ambiguity function is a function of b only, the statistical and probabilistic quality of its minimizer is still, like the ILS estimator of b, in a large part driven by the ambiguity success-rate.

Table 8.2: Formal standard deviation (σ_{UP}) and simulated RMS-values of $\hat{b}$ (float), $\check{b}^{\circ}$ (ILS), and $\check{b}_{\mathrm{AF}}$, for three different success-rates (SR) of the model used in Example 2. For $\check{b}^{\circ}$ (cf. 8.53), matrix $Q^{\circ}_{\hat{a}(b)\hat{a}(b)}$ was chosen as $Q^{\circ}_{\hat{a}(b)\hat{a}(b)} = \mathrm{diag}(\sigma^2_{\hat{a}_1(b)}, \ldots, \sigma^2_{\hat{a}_n(b)})$.

SR [%]	σ_{UP} [cm]	Float [cm]	ILS [cm]	AFM [cm]
99.99	101.9	101.8	1.0	1.1
86.84	194.7	194.5	113.6	113.9
45.06	418.1	418.0	372.5	372.9

As an additional remark to our comparison of the two objective functions (5.6) and (8.55), we note that the difference in their phase-arguments can easily be eliminated by *re-defining* the ambiguity function such that $c_i^T(\phi - B_\phi b)$ and $\sigma^2_{\phi_i}$ in (5.6) are replaced by $c_i^T A^+_\phi(\phi - B_\phi b)$ and $\sigma^2_{(A^+_\phi \phi)_i}$, respectively. This however, would eliminate the immunity-to-cycle-slip property of the ambiguity function.

8

8.7.2 What about differencing?

The above considerations have shown that the extend to which the AF-solution can be expected to be close to the statistically optimal ILS-solution $\check{b}$ depends on the differences listed for (8.52) and (8.53), and on how well $Q^{\circ}_{\hat{a}(b)\hat{a}(b)}$ approximates the actual ambiguity variance matrix $Q_{\hat{a}(b)\hat{a}(b)}$. However, in addition to this, there are also two other aspects that one should keep in mind when comparing AF with ILS. The first is their difference in solution-uniqueness as discussed in Section 8.3 and the second is their ability to handle data transformations, like e.g. phase differencing.

In our review Section 8.2 we already alluded to the fact that the AF baseline solution lacks invariance against different forms of data differencing. To understand the essence of this better, consider data differencing as a linear transformation and then first recall the property which the least-squares principle has with respect to invertible linear transformations of the data. In case of least-squares the parameter solution remains invariant provided the weighting accomodates the effect of the data transformation. Such accomodation however, is not generally available with the

AF-method. The w-vector in (8.12) can take care of variations in precision, but not accomodate any correlations that a linear transformation of the data may introduce. Another way of describing this difference between the two principles is to consider the PDF transformation rule for the normal distribution and the circular normal distribution. The normal distribution is closed under linear transformations. That is, a linear transformation of a normally distributed random vector is again normally distributed. Such is, however, not the case with the circular normal distribution. A linear combination of circular normally distributed random variables is *not* circular normally distributed anymore.

This lack of being closed under linear transformations implies that one has to be careful when formulating the carrier-phase part of the AF likelihood function. Under the assumption that the considered carrier-phase observables are circular normally distributed, usage of (8.12), as maximum likelihood estimator, requires the carrier-phase observables to be independent. This implies, in the context of GNSS, that only undifferenced (UD) or (between-receiver) single-differenced (SD) carrier-phase observables would qualify as potential entries of the m-vector ϕ. Double-differenced (DD) carrier-phase observables, being correlated amongst themselves, would then not qualify in principle.

But usage of the AF-method with DD carrier-phase observables is of course not forbidden. As our numerical examples have shown, their AF-results can become quite close to the ILS-solutions and similarly to the SD-based AF-results, as a consequence of the typically high-precision of the GNSS phase-observables. What one should keep in mind however is that the solution would then not be invariant for the in principle arbitrary way in which DD observables can be defined. Hence, from the same UD data set, one would then obtain different AF solutions when using different definitions of the DD observables. Although these solution differences can be small, it is important that this lack of invariance is understood when applying the AF-method.

8.8 Summary and conclusions

In this contribution we introduced new theory for the ambiguity function method. Its two main components are (*i*) the provision of a probability model by means of which the AF-estimator can be identified as a maximum likelihood estimator, and (*ii*) the provision of a global optimizer of the AF-likelihood function, having finite termination with a guaranteed epsilon-tolerance.

It was shown that for the AF-estimator to be a maximum likelihood estimator, the multivariate distribution of the phase data must consist of independent circular normal distributions. Although the suggestion of the circular normal distribution for phase data is not new, see (Cai et al., 2007), our linkage of the circular normal distribution to the ambiguity function method and its requirement for the AF-estimator to be a maximum likelihood estimator, are new. In this context we also showed how the ambiguity function method can be generalized so as to enable the inclusion of varying weights for the phase data.

Although the attractiveness of the ambiguity function method is the integer-ambiguity invariance of its objective function, we showed that this same invariance may also cause nonuniqueness in its solution. As this possible nonuniqueness appears to be overlooked in the AF-literature, we determined in Theorem 1 the explicit conditions that need to be satisfied for such nonuniqueness to occur. This nonuniqueness is then removed by code-regularization in Theorem 2.

To better understand the challenges of computing the AF-solution, we first characterized the multimodality of the ambiguity function, provided its gradient and Hessian matrix in Theorem 2 and determined a convex region in which its minimizer is guaranteed to reside. As iterative gradient descent methods alone will not be able to ensure the determination of the global minimizer of the multimodal ambiguity function, we introduced our globally convergent algorithm which is constructed from combining the branch-and-bound principle, with a special convex-relaxation of the ambiguity function, to which the projected-gradient-descent method is applied. Each of the method's three constituents was described, with special emphasis to the construction of the required continuously differentiable, convex lower bounding function of the multimodal ambiguity function. Several examples were provided in which the workings and performance of our AF-algorithm were numerically and graphically illustrated.

Finally, a further comparison between the AF-principle and that of ILS-estimation was made using the primal-dual equivalence of mixed-ILS theory as introduced in (Teunissen and Massarweh, 2024). Based on this equivalence, as summarized in Theorem 3, it was shown that the differences are driven by those listed for (8.52) and (8.53), as well as by the impact of neglecting the correlation between the baseline conditioned, float ambiguities. From this comparison, as well as from the identified differences in nonuniqueness and dependence on linear transformations, the conclusion is reached that the two principles are fundamentally different, although there are identified circumstances, as was shown, under which one can expect AF- and ILS-solutions to behave similarly.

References

Balakrishnan V, Boyd S, Balemi S (1991) Branch and bound algorithm for computing the minimum stability degree of parameter-dependent linear systems. Journal of Robust and Nonlinear Control 1(4):295–317

Bertsekas D (1999) Nonlinear programming. Athena Scientific, 2nd Ed

Cai J, Grafarend EW, Hu C (2007) The statistical property of the GNSS carrier phase observations and its effects on the hypothesis testing of the related estimators. In: Proceedings of ION GNSS 20th International Technical Meeting of the Satellite Division, pp 331–338

Counselman CC, Gourevitch SA (1981) Miniature interferometric terminals for earth surveying: ambiguity and multipath with Global Positioning System. IEEE Transactions Geoscience and Remote Sensing GE-19(4):244–252

Guida A (2015) A branch and bound algorithm for the global optimization and its improvements. PhD thesis, University of Florence pp 1–89

Gumbel EJ, Greenwood JA, Durand D (1953) The circular normal distribution: theory and tables. J Amer Statist Assoc 48(261):131–152

Han S, Rizos C (1996) Improving the computational efficiency of the ambiguity function algorithm. Journal of Geodesy (70):330–341

Hartman S (2021) The History of RTK - Part 3: Everybody Wanted It. The American Surveyor pp 1–9

Hofmann-Wellenhof B, Lichtenegger H, Wasle E (2008) GNSS: Global Navigation Satellite Systems: GPS, Glonass, Galileo, and More. Springer, New York

Lachapelle G, Canon ME, Erickson C, Falkenberg W (1992) High-precision C/A code technology for rapid-static DGPS surveys. Proc 6th Int Geodetic Symposium on Satellite Positioning, DMA/OSU pp 1–6

Lawler EL, Wood DE (1966) Branch-And-Bound methods: A survey. Operations Research 14(4):699–719

Leick A, Rapoport L, Tatarnikov D (2015) GPS Satellite Surveying. 4th Edition Wiley

Mader GL (1992) Rapid static and kinematic Global Positioning System solutions using the ambiguity function technique. Journal of Geophysical Research 97(B3):2371–2383

Massarweh L, Verhagen S, Teunissen PJG (2025) New LAMBDA toolbox for mixed-integer models: estimation and evaluation. GPS Solutions 29(1):14, DOI 10.1007/s10291-024-01738-z, available at http://pntlab.tudelft.nl/LAMBDA

Nesterov Y (2018) Lectures on convex optimization. Springer Optimization and Its Applications 137

Nocedal J, Wright S (2006) Numerical optimization. 2nd ed. Springer-Verlag

Remondi BW (1984) Using the Global Positioning System phase observable for relative geodesy: modelling, processing and results. PhD Dissertation, University of Texas

Remondi BW (1991) Pseudo-kinematic GPS results using the ambiguity function method. Navigation, Journal of the Institute of Navigation 38(1):17–36

Remondi BW, Hilla SA (1993) Pseudo-kinematic surveying based on full-wavelength dual-frequency GPS observations. NOAA Technical Memorandum NOS NGS-56, 24p

Rogers AEE, Knight CA, Hintegregger HF, Witney AR (1978) Geodesy by radio interferometry: Determination of a 1.24-km base line vector with ∼5-mm repeatability. Journal of Geophyscial Research 83(B1):325–333

Teunissen PJG (1990) Nonlinear least squares. Manuscripta Geodaetica 15(3):137–150

Teunissen PJG (1995) The least-squares ambiguity decorrelation adjustment: a method for fast GPS integer ambiguity estimation. J Geod 70(1-2):65–82

Teunissen PJG (1999) An optimality property of the integer least-squares estimator. J Geod 73(11):587–593

Teunissen PJG (2003a) Theory of integer equivariant estimation with application to GNSS. Journal of Geodesy 77(7-8):402–410

Teunissen PJG (2003b) Towards a unified theory of gnss ambiguity resolution. Journal of Global Positioning Systems 2(1):1–12

Teunissen PJG (2017) Carrier-phase integer ambiguity resolution. Chapter 23 in Handbook of GNSS, Teunissen and Montenbruck (Eds) pp 661–685

Teunissen PJG, Khodabandeh A (2022) PPP–RTK theory for varying transmitter frequencies with satellite and terrestrial positioning applications. Journal of Geodesy 96:84

Teunissen PJG, Massarweh L (2024) Primal and dual mixed-integer least-squares: distributional statistics and global algorithm. Journal of Geodesy 98(63):1–26

9

Conclusions

9.1 Conclusions

In this dissertation, we considered mixed-integer models in the context of Global Navigation Satellite Systems (GNSS), but most of the theoretical contributions of this work can be generalized also to non-GNSS problems dealing with a mixture of integer ambiguities $a \in \mathbb{Z}^n$ and real-valued parameters $b \in \mathbb{R}^p$. This thesis tackled mixed-integer least-squares problems, and the search for candidate solutions, from two different perspectives. First, we looked into the discrete domain of ambiguities, used to conditionally update real-valued parameters, e.g. positioning coordinates. Then, we formulated a dual problem mainly concerned with these parameters of interest for the user, and therefore implicitly solving for the unknown ambiguities.

The main conclusions can be summarized as follows:

- In **Part I (Primal formulation)**, we started by introducing in Chapter 2 new integer estimators, so extending the bootstrapping principle to the vectorial case. The Vectorial Integer Bootstrapping (VIB) estimators were compared against legacy estimators such as integer rounding, integer bootstrapping and the (success-rate optimal) integer least-squares. A performance ordering was therefore established, showing that by means of larger blocksets of ambiguities it is possible to obtain quasi-optimal performance, since ambiguity correlation is more extensively considered. It was also demonstrated how other techniques, e.g. widelane-narrowlane (Teunissen, 1997), belong to this family of vectorial solutions. The vectorial solutions were proved to be flexible since balancing computational efficiency and success rate performance by means of a user-selected partitioning of the ambiguity resolution problem. Lastly, the possible extension of these vectorial solutions to other classes of estimators was discussed.

Overall, VIB-based solutions are highly effective in high-dimensional problems, as demonstrated in Chapter 3 where we compared and analyzed different ambiguity fixing solutions for Orbit Determination and Time Synchronization (ODTS) processes. The ODTS aims at estimating accurate satellite products, and it relies on a global network of ground station receivers. First, the impact of ambiguity resolution was studied with respect to the network size, showing that a plateau is reached with global networks of around 25–30 stations, whereas a factor '2' improvement is maintained when comparing *float* and *fixed* solutions. The VIB solutions proved to be suitable for such applications, where thousands of ambiguities were fixed in a fraction of a second, based on a 14-station network of well-distributed receivers. Daily RMS error values of few centimeters were obtained, then validated by a 3-month GPS+Galileo campaign where we looked at midnight orbit discontinuity errors. At 95% probability, errors of 8.0 cm, 10.1 cm and 18.3 cm were obtained for VIB, (scalar) IB, and float solutions.

After these flexible vectorial solutions were established, we developed a unified software framework for integer estimation and evaluation, built upon the legacy of LAMBDA (2012) and Ps-LAMBDA (2013) software. Thus, a new LAMBDA 4.0 toolbox was implemented (in MATLAB), extensively documented, then released, free and open source, to the community. This implementation not only included a versatile implementation of VIB estimators, but it also accounted for additional (classed of) estimators, such as Integer Aperture Bootstrapping (IAB) and Best Integer Equivariant (BIE). In this way, it was possible to provide users with a large set of functionalities for different applications. Moreover, this implementation took advantage of recently enhanced algorithms, including a search strategy with one-order reduction in the computational time, enabling numerical success-rate simulations for different classes of estimators.

- In **Part II (Dual formulation)**, we shifted our attention back to real-valued parameters, and we introduced in Chapter 5 a dual formulation for mixed-integer least-squares problems, based on a dual orthogonal decomposition. The latter is complementary to the primal decomposition introduced by Teunissen (1993), which created the foundations for the LAMBDA method and current state-of-the-art theory in GNSS ambiguity resolution. A close primal-dual relation was proved, then exploited in order to study the statistical properties of these new dual ILS estimators. Two easy-to-compute approximations were introduced, based on an approximate weighting and an approximate integer mapping, with dual objective functions respectively denoted as $\mathcal{D}^{\circ}$ and $\mathcal{D}^{\bullet}$. While the weighting approximation implies non-optimal performance, it was shown how quasi-optimal solutions are still possible for GNSS models, mainly due to the high precision of carrier-phase measurements. With the second approximation, new pull-in regions were defined and the distributional statistics was discussed. Given the multimodality of the dual objective function, a globally convergent solution was proposed based on a branch-and-bound principle combined with a well-defined convex relaxation. In this way, optimal solutions were guaranteed (given an ϵ-tolerance) for the approximate dual problems.

Our investigation continued by focusing on the special case $p = 1$, where scalar real-valued parameters were considered given an arbitrary number of ambiguities n. Although this is likely not representative of most GNSS models, this simple analysis provided important geometrical insights and demonstrated that global solutions can be found based on an enumeration of solutions taking place in the domain of real-valued parameters. Hence, the P1 algorithm was presented and numerically evaluated against LAMBDA 4.0 implementation (see Chapter 4), showing superior computational performance for such particular problems. In addition, it was also mathematically proved that this algorithm has a complexity (i.e. number of candidate solutions to be enumerated) growing *linearly* with respect to the ambiguity dimensionality n.

When looking at the more general case $p \geq 1$, we investigated the branch-and-bound strategy from Chapter 5, and we provided an extensive analysis of the computational performance, including different implementation aspects for a Branch-and-bound INteger-equivariant Global Optimizer (BINGO) solution. A dual problem normalization was introduced, i.e. convenient for the computations, along with the evaluation of several different elements: search space initialization, boxes' splitting strategy, convex lower bound minimization approach, and so on. By analyzing GNSS models, it was demonstrated that quasi-optimal solutions are feasible and comparable with the primal ILS estimator. Moreover, this highlighted a landscape of potential applications, where additional constraints on the real-valued parameters can be rigorously taken into account.

Ultimately, the methodology introduced for the dual formulation was exploited for an in-depth analysis of Ambiguity Function (AF) methods. These were originally developed in the 1980s for GPS baseline positioning based on different formulations, but unfortunately often confused in literature. In this final part of the dissertation, we addressed and bridged some existing theoretical gaps related to the AF principle. Hence, we demonstrated the nonuniqueness of the AF solution, later solved via a code-based regularization, and we introduced a probabilistic model that justifies the AF methods as maximum likelihood estimator. This is possible when the multivariate distribution of phase data consists of independent circular normal distributions. Therefore, a similar branch-and-bound strategy was applied to provide a global solution (with guaranteed ϵ-tolerance), based on a different convex lower bound function. To conclude, a comparison between the Ambiguity Function principle and the one of Integer Least-Squares estimators was presented, based on the primal-dual equivalence introduced in Chapter 5. In this way, we proved that the two principles are fundamentally different, although there are identified circumstances under which the two solutions might behave similarly.

9.2 Recommendations for future work

Based on this thesis's contributions, there are some aspects that we consider to be of interest for future research works, summarized by the following recommendations:

- Vectorial bootstrapping estimators defined in the primal formulation (Part I) proved to be flexible and highly efficient methods. However, they were mainly analyzed in the context of integer estimators, so their extension to the class of Integer Aperture or Integer Equivariant estimators is still an important subject of future investigation. See example in Miao et al. (2024).
- As VIB solutions are still dependent upon the ambiguity parametrization, it is important to study how admissible Z-transformations could be generalized to the vectorial case. An example was given in Chapter 5 with the widelane-narrowlane transformation, but more general methods should be developed in view of future navigation systems in low Earth orbit (Wang et al., 2024), where the ambiguity problem dimensionality will play a central role.
- Although LAMBDA 4.0 toolbox presented many new algorithms and introduced different (classes of) estimators, it is fundamental to further study the uncertainty description for all these ambiguity-fixed solutions. For instance, the Best Integer Equivariant (BIE) estimator is optimal in terms of minimum mean squared errors, but this solution is not normally distributed and its probabilistic characterization is therefore necessary.
- The introduction of a dual mixed-integer least-squares problem opened up a wide range of possible analyses. In fact, this research work mainly focused on the 'unconstrained' case, where no additional constraints existed for the real-valued parameters. However, several GNSS or non-GNSS models do have constraints that can be exploited to enhance the success rate, i.e. probability of correctly fixing the ambiguities, and they could be taken into account more rigorously in this dual formulation.
- In the geometric interpretation of the P1 algorithm, the link between primal and dual formulations should be further explored, as also noted in Chapter 5, since an enumeration in the parameter domain is possible (and proved for the $p = 1$ case). This could potentially lead to new search algorithms (for $p > 1$) that make better use of the information from real-valued parameters.
- A global solution in the dual formulation was made possible by a branch-and-bound approach, which used a dedicated lower convex bound function constructed differently for the ILS and AF principles. However, it represents only one potential approach to guarantee optimal solutions in dual mixed-integer least-squares problems, and other strategies might be investigated. For instance, the adoption of simple heuristic methods (e.g. genetic algorithms, particle swarm optimization, etc.) or alternative global optimization methods can be explored, thus leading to new solutions and more advanced applications.

References

Miao W, Li B, Gao Y, et al. (2024) Vectorial integer bootstrapping of best integer equivariant estimation (VIB-BIE) for efficient and reliable gnss ambiguity resolution. Journal of Geodesy 98:30, DOI 10.1007/s00190-024-01836-3

Teunissen PJG (1993) Least-squares estimation of the integer gps ambiguities. In: Invited lecture, section IV theory and methodology, IAG general meeting, Beijing, China, pp 1–16

Teunissen PJG (1997) On the GPS widelane and its decorrelating property. Journal of Geodesy 71(9):577–587, DOI 10.1007/s001900050126

Wang K, El-Mowafy A, Cheng F, Yang X (2024) LEO augmentation in large-scale ionosphere-float PPP-RTK positioning. Journal of Surveying Engineering 150(2):04024001

Acknowledgements

This has certainly been a long journey. However, here we are, in the last few sentences, summarizing the countless adventures, experiences, and remarkable people I met along the way. Brevity is not my strong suit; nonetheless, I will try to be concise.

I shall begin with my wonderful years at Deimos Engenharia in Lisbon, Portugal. I am deeply thankful to my fantastic office mates, **Samuele Salvi** and **David Cordeiro**, along with **João Malés** (a big hug to Anita), as well as **Maria**, **Rui**, **Ricardo(s)**, **Vasco**, **Gonzalo**, **Luis**, and many others, for their friendship and camaraderie. I will always cherish the endless hours of board games with them, as well as the support of my great managers, **Pedro Tavares** and **Pedro Freire da Silva**, who always encouraged my initiatives and ideas. In particular, I would like to thank **Manuel Marcote** and **José Fernandes Vasconcelos**, both kind and talented individuals. I am also grateful to **Catarina Bastos** for her enthusiasm and encouragement over my years at Deimos. I will surely remember our Friday football games, the cervejas, as well as the many travels and experiences we shared.

During my research period at Técnico Lisboa, I had the privilege to be supervised by **Rodrigo Ventura**, a brilliant professor and scientist with an incredible dedication to work. Even if I could not always fully express my gratitude, I will always be indebted for the incredible opportunities you provided. The amazing people in the Institute for Systems and Robotics (Lisboa) will always hold a special place in my heart. Likewise, I will never forget **Marco Zecchetto**, a true friend who stood by my side in the hardest times. Your humility and genuine friendship have been invaluable to me.

I was more fortunate than I could have ever hoped for. Through the Marie-Curie scholarship under the Horizon 2020 TREASURE program, I had the privilege of meeting many great colleagues/friends and sharing numerous adventures. I deeply thank **Juliana Damaceno**, **Karl Bolmgren**, **Caner Savas**, **Wenjian Qin**, **Kai Guo**, **Brian Weaver**, **Hossein Ghobadi**, and **Paola Testa** - what an amazing group! Above all, I cannot forget the kindness and friendship of **Hongyang Ma**, along with the forever-lasting GADIP3 Team: **Jon Bruno**, **Francesco Darugna**, and **Dimitrios Psychas**. You are incredible researchers, immensely talented and full of passion. We shared memories and stories I will never forget. A great future lies ahead for you all, but ours was not a *goodbye* - only an *arrivederci*.

These first three years were filled with adventures, and listing everyone would be a challenging task. I lived in a beautiful house of 10 people, met extraordinary friends and shared too many (crazy) experiences. I especially thank **Gabriel** for keeping that big house sane, along with the 'Fortaleza Paranormal' girls: **Béatrice**, **Alice**, **Catherine**, and **Paula** [**aka Broccoli**]. Surviving that house was surely a significant achievement and a life-long memory of those wonderful years.

During that time, I also had the pleasure of meeting and working with **Sebastian Strasser**, one of the most talented researchers I have ever met, along with several friends from the Caltech Space Challenge 2019: **Paolo Cappuccio**, **Marc Rovira Navarro**, **Alessia Gloder**, **Paolo Panicucci**, and many individuals with bright futures ahead. Remember, one day humanity will explore Saturn's moon Enceladus and uncover whether life hides within its "tiger stripes".

I moved back to Delft in 2020. I will forever be indebted to **Dr. Sandra Verhagen** and **Prof. Peter Teunissen**, whom I can jokingly consider my 'parents-in-mathematics'. You not only taught me far more than I could have ever imagined, but you also truly changed my life by believing in me when I could not. Still, what got me even more impressed was your patience and endurance in listening to my hours-long presentations and thousands of slides; how you survived that is truly remarkable. I cannot forget to thank **Christiaan Tiberius**, **Hans van der Marel** and **Alireza Amiri-Simkooei**, whose dedication and expertise have been an inspiration for me. Also thanks to my colleagues and friends of the PNTLab group: **Sebastian Ciuban**, **Chengyu Yin**, **Bob van Noort**, **Serge Kaplev**, **Carlos Fortuny Lombraña**, and **Lucas Alvarez Navarro**.

Many more I have met during these last years at TU Delft, including my old office mates **José Dias** and **Andrés Yarce Botero**, as well as **Sophie de Roda Husman** (with your Monday morning smile), **Wietske** (whose name I still cannot pronounce), **Ann-Sofie**, **Weilun**, **Dai**, **Yan**, **Max**, **Felix**, **Alex** and many other truly remarkable people. Without any doubt, this last year has been challenging for me, but I must deeply thank my office mate **Annelot Broerse** for her joyful spirit and her radiant "Buongiorno!" - even on days when cycling to work felt like swimming. Having been involved on several projects, I have additional people I should thank. To **Aram Vroom**, **Tom van den Oever**, and **Javier Navarro Juan**: I will never forget those years debugging MATLAB code, as well as the DLR Team (**Gabriele Giorgi**, **Cécile Deprez**, **Manuele Dassié** and **Christian Trainotti**) for the countless hours realizing how nasty C++ programming is. I shall thank many others I had the pleasure to collaborate with, like **Marco Fortunato**, **Ciro Gioia**, ..., and last but not least, **Jose van den IJssel**, a brilliant researcher and such an amazing person.

Of course, my time at TU Delft has not been only about work. I have important people to thank for their friendship and for introducing me to new sports. Thanks to **Reka**, who created our lovely 'VolleyBolley' group, along with **Kiran**, **Girish**, **Tao**, **Suzan** and others too numerous to name. I cannot forget the **van de Laar** brothers, i.e. **Thomas** and **Matheus**, for challenging me in sport pursuits I never thought possible. I also owe thanks to **Anna** (or "Susskina"), a truly remarkable person with a big heart (and a loud voice), as well as **Fatih**, **Fabra**, **Shreyas** and others who tolerated my alcohol-free beers. Finally, to **Flora Guarnotta**, a wonderful person and a dear friend of mine; I promise to train hard and one day to finally conquer the *full Ironman*! "Anything is Possible".

Though I was living in the Netherlands, I have always carried the many friends and family around the world deep in my heart.

To the 'Bobba La Gobba' crew: **Marco**, **Roberto**, **Lorenzo**, **Rosalia**, **Ilaria** and **Flavia**, along with the new little members *soon to appear*; as well as **Elena** and **Francesco**, **Lanzi** and **Laura**, **Ferd**, **Marta**, **Sof**, **Ilaria**, **Edo**, and many more. Our friendship transcends space and time; when we meet it always feels like home. To both my **Italian** and **Palestinian families**, your love has always been my strength. Seeing the turmoil of recent years has broken my heart, but I hope to visit you soon. After all: "*Non nobis solum nati sumus ortusque nostri partem patria vindicat, partem amici*" (Not for us alone are we born; our country, our friends, have a share in us).

Ultimately, I dedicate this thesis to the most important people in my life. To one of my best friends **Niccolò Gastaldello**, who unfortunately left us in May 2021. I will always remember our adventures. You were a true inspiration and I regard you as the best aerospace engineer I have ever known. To my brother **Omar** and my **father**, your unconditional support and understanding mean everything. To the memories of **mamma**, who left us the very last day before I first moved to TU Delft, back in 2014. In this long journey, I have never forgotten you, you gave me strength for all my achievements. And to **Elena De Lazzari**, my *Cuoricino* (little heart), who supported me with love, who tolerated my 3:00 am "brilliant" ideas (often wrong), and the very one who has made the end of this long journey possible. Love you. ♡

Curriculum Vitæ

Lotfi Massarweh

Born in Matera (Italy), 14/05/1992.

Education

2024– 2020 **Doctor of Philosophy**
Delft University of Technology

Thesis: *"Primal and dual mixed-integer models for Global Navigation Satellite Systems"*

Supervisor: Dr.ir. Sandra Verhagen
Promotor: Prof.dr.ir. Peter J.G. Teunissen

2016– 2014 **Master of Science in Aerospace Engineering**
Delft University of Technology

Thesis: *"Linear stability and bifurcations of periodic Lagrange orbits in the Elliptic Restricted 3-Body Problem. An investigation at L1/L2 in the Earth-Moon system."*

Supervisor: Dr.ir. Ernst J.O. Schrama
Advisor: Dr.ir. Wim T. van Horssen

2014– 2011 **Bachelor of Science in Aerospace Engineering**
Politecnico di Milano

Work experience

Current– Jan 2021	**Researcher** Delft University of Technology (in Delft/NL) Geoscience and Remote Sensing (GRS) department
Aug 2020– Sep 2017	**Aerospace Engineer** Deimos Engenharia S.A. (in Lisbon/PT) Research fellow of Horizon 2020 TREASURE project, funded by Marie Skłodowska-Curie Actions (MSCA) programme
Aug 2019– Sep 2017	**Researcher** Instituto Superior Técnico (IST, in Lisbon/PT) Intelligent Robots and Systems group (IRSg), part of the Institute for Systems and Robotics (ISR|Lisboa)
Jul 2017– Jan 2017	**Researcher** Delft University of Technology (in Delft/NL) Astrodynamics and Space missions (AS) group
Dec 2016– Nov 2016	**Researcher** Delft Space Institute (DSI, in Delft/NL)
Nov 2015– Aug 2015	**Researcher** Centre National de la Recherche Scientifique (CNRS, in Toulouse/FR) Internship, collaboration with Institut Supérieur de l'Aéronautique et de l'Espace (ISAE) and Centre National D'Etudes Spatiales (CNES)

Grants

1. Awarded in May 2024 as prime contractor by ESA for the **Contract No. 4000144530/24/NL/AK/mk**, "NAVISP-EL1-071: Technologies for Reliable Ambiguity Resolution and Integrity in High Accuracy Positioning".
2. Awarded in March 2022 as prime contractor by ESA for the **Contract No. 4000137444/22/D/SR**, "Kinematic Precise Orbit Determination for Satellites Based on the Galileo High Accuracy Service (HAS)" – T710-601GN.
3. Awarded with a **Marie Skłodowska-Curie scholarship** (project ID 722023) as Early Stage Researcher (ESR) part of Innovative Training Network (ITN) funded through EU H2020 Research and Innovation Programme.

Awards

2019 **1st Prize – Galileo Masters 2019 'University Challenge'**
Organized by AZO Space of Innovation, presented at the European Space Week 2019 (EUSW2019).

2019 **1st Prize – International Summer School on GNSS 2019**
Organized by the European Space Agency (ESA) and the EC Joint Research Centre (JRC), Vila Nova de Cerveira, Portugal.

2019 **1st Prize – Caltech Space Challenge 2019**
Organized by the California Institute of Technology (Caltech), along with the NASA's Jet Propulsion Laboratory (JPL).

2019 **1st Prize – Galileo App Competition 2018/2019**
Organized by the European Space Agency, in collaboration with EC, Google and the European GNSS Agency. Co-founder of the GNSS Android-based Dual-frequency Iono-estimated PPP (GADIP3) app.

2018 **Finalist – Farming by Satellite Prize 2018**
Organized by the European GNSS Agency, in collaboration with CLAAS and the European Environment Agency (EEA).

List of Publications

The list of the doctoral candidate's publications (in light blue, non-GNSS related).

Peer-reviewed publications

1. Teunissen, P.J.G., and **Massarweh, L.** (2024). *Theory for the Ambiguity Function Method: Probability Model and Global Solution.* In Journal of Geodesy (submitted).

2. **Massarweh, L.**, and Teunissen, P.J.G. (2024). *An efficient P1 algorithm for dual mixed-integer least-squares problems with scalar real-valued parameters.* In Journal of Applied Geodesy. https://doi.org/10.1515/jag-2024-0076

3. **Massarweh, L.**, Verhagen, S., and Teunissen, P.J.G. (2025). *New LAMBDA toolbox for mixed-integer models: Estimation and Evaluation.* In GPS Solutions 29, 14. https://doi.org/10.1007/s10291-024-01738-z

4. Teunissen, P.J.G., and **Massarweh, L.** (2024). *Primal and dual mixed-integer least-squares: distributional statistics and global algorithm*, J Geod 98, 63. https://doi.org/10.1007/s00190-024-01862-1

5. Nathan, E., Balachandran, K., Cappuccio, P., Di, J., Doerksen, K., Gloder, A., Li, M., **Massarweh, L.**, Peev, T., Santra, S., Rovira-Navarro, M., and Limonchik, D. (2022). *A multilander New Frontiers mission concept study for Enceladus: SILENUS.* Front. Astron. Space Sci. 9:995941. https://doi.org/10.3389/fspas.2022.995941

6. **Massarweh, L.**, Strasser, S., and Mayer-Gürr, T. (2021). *On vectorial integer bootstrapping implementations in the estimation of satellite orbits and clocks based on small global networks*, Advances in Space Research, 68(11), 4303-4320. https://doi.org/10.1016/j.asr.2021.09.023

7. Teunissen, P.J.G., **Massarweh, L.**, and Verhagen, S. (2021). *Vectorial integer bootstrapping: flexible integer estimation with application to GNSS.* J Geod 95, 99. https://doi.org/10.1007/s00190-021-01552-2

8. **Massarweh, L.**, Fortunato, M., and Gioia, C. (2020). *Assessment of Real-time Multipath Detection with Android Raw GNSS Measurements by Using a Xiaomi Mi 8 Smartphone*, at 2020 IEEE/ION Position, Location and Navigation Symposium (PLANS), Portland, OR, USA, 2020, pp. 1111-1122. https://doi.org/10.1109/PLANS46316.2020.9110169

9. **Massarweh, L.**, and Cappuccio, P. (2020). *On the Restricted 3-Body Problem for the Saturn-Enceladus system: mission geometry & orbit design for plume sampling missions.* In AIAA Scitech 2020 Forum (p. 0467). https://doi.org/10.2514/6.2020-0467

10. Psychas, D., Bruno, J., **Massarweh, L.** and Darugna, F. (2019). *Towards Submeter Positioning using Android Raw GNSS Measurements*, Proceedings of the 32nd

International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2019), Miami, Florida, September 2019, pp. 3917-3931. https://doi.org/10.33012/2019.17077

11. **Massarweh, L.**, Darugna, F., Psychas, D. and Bruno, J. (2019). *Statistical Investigation of Android GNSS Data: Case Study Using Xiaomi Mi 8 Dual-Frequency Raw Measurements*, Proceedings of the 32nd International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2019), Miami, Florida, September 2019, pp. 3847-3861. https://doi.org/10.33012/2019.17072

12. Doerksen, K., Balachandran, K., Cappuccio, P., Di, J., Fuchs, J., Gloder, A., Jolitz, R., Li, M., Limonchik, D., **Massarweh, L.**, Meszaros, A., Naftalovich, D., Nathan, E., Peev, T., Rovira-Navarro, M., and Santra, S. (2019). *Rapid Mission Concept Development at the 2019 Caltech Space Challenge: A Small Lander Network Studying the Habitability of Enceladus*, at the 3rd Symposium on Space Educational Activities, in Leicester, United Kingdom, September 16-18, 2019. https://doi.org/10.29311/2020.52

13. Garcia, R.F., Bruinsma, S., **Massarweh, L.**, and Doornbos, E. (2016). *Medium-scale gravity wave activity in the thermosphere inferred from GOCE data.* Journal of Geophysical Research: Space Physics, 121(8), 8089-8102. https://doi.org/10.1002/2016JA022797

Oral and poster presentations

1. *Impact of Modelling Errors on LEO-PNT Performance: A Brief Numerical Investigation on Ambiguity-Fixed Positioning*, presented during NAVITEC 2024 at ESA/ESTEC in Noordwijk, 11-13 December, The Netherlands.

2. *Kinematic POD for real- and near-real-time using Galileo HAS - GHASP3 project*, presented at the 9th International Colloquium on Scientific and Fundamental Aspects of GNSS, 25-27 September 2024, in Wroclaw, Poland.

3. *Numerical assessment of LEO-PNT performance: a comparison between ambiguity-float and ambiguity-fixed precise point positioning solutions*, presented at the European Navigation Conference (ENC) 2024 at ESA/ESTEC in Noordwijk, 22-24 May, The Netherlands.

4. *LAMBDA 4.0: an advanced tool for integer estimation, validation and success rate simulation*, presented at the 28th General Assembly of the International Union of Geodesy and Geophysics (IUGG), 11-20 July 2023, in Berlin, Germany.

5. **Massarweh, L.**, and Teunissen, P.J.G. (2024). *A novel globally convergent maximizer for the multivariate carrier-phase integer ambiguity function*, presented at EGU General Assembly 2024, Vienna, Austria, 14–19 Apr 2024, EGU24-19367. https://doi.org/10.5194/egusphere-egu24-19367

6. *LAMBDA 4.0: an enhanced toolbox for high dimensional ambiguity resolution*, presented at the European Navigation Conference (ENC) 2023 at ESA/ESTEC in Noordwijk, 31 May-2 June, The Netherlands.

7. *LEO-PNT constellation: the next revolution for GNSS - challenges, benefits and geoscience opportunities*, seminar held in hybrid format at TU Delft on March 20th,

2023. [Webinar link] Password: 1tgWb?zt

8. *SILENUS: A Mission Concept Investigating the Habitability of Enceladus*, presented in the session MIT2 at the EPSC-DPS Joint Meeting 2019, in Geneva, Switzerland, September 15-20 2019.

9. *GADIP3 app: GNSS Android-based Dual-frequency Iono-estimated PPP*, poster at the 2nd TREASURE Workshop, in Toulouse, France, May 21-23, 2019.

    - Presented at the 3rd GNSS Raw Measurements Taskforce Workshop, in Prague, GSA Headquarter, June 26, 2019.

10. *Real-time provision of high accuracy orbits and clocks: beyond the EKF*, presented at the 2nd Workshop of TREASURE, in Toulouse, France, May 21-23, 2019.

11. *A low-priced augmenting support for UAV's precise positioning based on dual-frequency mass-market receivers*, presented at the European Space Week 2018, in Marseille, France, December 5, 2018.

12. *Insights on the possible use of Unscented Kalman Filters for real-time OD*, presented at the 1st Workshop of TREASURE, in Rome, Italy, April 17, 2018.

13. *Using platform magnetometers to observe and detect Space Weather events*, poster presentation at the 14th European Space Weather Week, Nov 27 – Dec 01, 2017, in Ostend, Belgium.

14. *Gravity wave activity in the thermosphere inferred from GOCE data, and its dependence on solar flux conditions*, presented at the European Geophysical Union General Assembly 2016 in Vienna, Austria, April 17-22, 2016.

Toolboxes

1. **Massarweh, L.**, Verhagen, S. and Teunissen, P.J.G. (2024). LAMBDA toolbox: Version 4.0 - Documentation (MATLAB), Delft University of Technology.