

**Primal and dual mixed-integer least-squares
distributional statistics and global algorithm**

Teunissen, P. J.G.; Massarweh, L.

DOI

[10.1007/s00190-024-01862-1](https://doi.org/10.1007/s00190-024-01862-1)

Publication date

2024

Document Version

Final published version

Published in

Journal of Geodesy

Citation (APA)

Teunissen, P. J. G., & Massarweh, L. (2024). Primal and dual mixed-integer least-squares: distributional statistics and global algorithm. *Journal of Geodesy*, 98(7), Article 63. <https://doi.org/10.1007/s00190-024-01862-1>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.



Primal and dual mixed-integer least-squares: distributional statistics and global algorithm

P. J. G. Teunissen^{1,2,3} · L. Massarweh¹

Received: 25 February 2024 / Accepted: 5 May 2024
© The Author(s) 2024

Abstract

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 \hat{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.

Keywords Least-squares (LS) · Integer least-squares (ILS) · Primal mixed ILS · Dual mixed ILS · GNSS · Ambiguity resolution · Pull-in region · Ambiguity success-rate (SR) · Branch-and-bound (BB) · Convex-relaxation · Projected-gradient-descent (PGD)

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 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 Sect. 3, together with a representation of its objective function $\mathcal{D}(b)$. It is shown that it

✉ P. J. G. Teunissen
p.j.g.teunissen@tudelft.nl

¹ Department of Geoscience and Remote Sensing, Delft University of Technology, Delft, The Netherlands

² Department of Infrastructure Engineering, The University of Melbourne, Melbourne, Australia

³ GNSS Research Centre, Curtin University of Technology, Perth, Australia

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 Sects. 4 and 5, respectively. The dual approximation of Sect. 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 Sect. 5 consists of replacing the implicit ILS-estimator 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 $\mathcal{D}(b)$, Sect. 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 Sect. 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 Sect. 8. Finally, Sect. 9 contains the Summary and Conclusions.

The following notation is used: $E(\cdot)$ and $D(\cdot)$ 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 $P_M = M M^+$. $P_M^\perp = I - P_M$ is then the orthogonal projector that projects orthogonally on the orthogonal complement of $\mathcal{R}(M)$. The Q -weighted squared norm is denoted as $\|\cdot\|_Q^2 = (\cdot)^T Q^{-1}(\cdot)$, and $\lceil x \rceil$ 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 \leq denotes the all component-wise inequality \leq . $P[A]$ denotes the probability of event \mathcal{A} , $f_b(b)$ the probability density function (PDF) of the continuous random vector \hat{b} and $P[\check{a} = z]$ the probability mass function (PMF) of the integer random vector \check{a} . The non-

central 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 $\chi_\delta^2(p, 0)$.

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,

$$E(y) = Aa + Bb, \quad D(y) = Q_{yy} \quad (1)$$

with $y \sim \mathcal{N}_m(E(y), D(y))$ 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 ambiguity vector consisting of the integer carrier-phase ambiguities, $b \in \mathbb{R}^p$ the unknown baseline 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

$$\hat{a} = \bar{A}^+ y \quad \text{and} \quad \hat{b} = \bar{B}^+ y \quad (2)$$

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 (2) is used as input to obtain the *integer* estimate of the ambiguity vector a as

$$\check{a} = \arg \min_{a \in \mathbb{Z}^n} \|\hat{a} - a\|_{Q_{\hat{a}\hat{a}}}^2 \quad (3)$$

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 (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

$$\check{b} = \hat{b}(\check{a}) = \hat{b} - Q_{\hat{b}\hat{a}} Q_{\hat{a}\hat{a}}^{-1} (\hat{a} - \check{a}) \quad (4)$$

That (2), (3) and (4) are indeed the LS and ILS solutions of the GNSS model (1) follows readily from the *orthogonal decomposition* (Teunissen 1998a)

$$\|y - Aa - Bb\|_{Q_{yy}}^2 = \|P_{[A,B]}^\perp y\|_{Q_{yy}}^2 + F(a, b) \quad (5)$$

where

$$F(a, b) = \|\hat{a} - a\|_{Q_{\hat{a}\hat{a}}}^2 + \|\hat{b}(a) - b\|_{Q_{\hat{b}(a)\hat{b}(a)}}^2 \quad (6)$$

with $Q_{\hat{b}(a)\hat{b}(a)} = (B^T Q_{yy} B)^{-1}$ the variance matrix of $\hat{b}(a)$. As $\|P_{[A,B]}^\perp y\|_{Q_{yy}}^2$ is independent of a and b , the minimizers of $\|y - Aa - Bb\|_{Q_{yy}}^2$ are those of $F(a, b)$. It therefore follows from (6), recognizing $\hat{b}(\hat{a}) = \hat{b}$, that the real-valued minimizers of $F(a, b)$ are given by (2), while their mixed-integer counterparts are given by (3) and (4), respectively. We therefore have for the LS and the mixed ILS solutions,

$$\left\{ \begin{array}{l} \hat{a} \\ \hat{b} \end{array} \right\} = \arg \min_{a \in \mathbb{R}^n, b \in \mathbb{R}^p} F(a, b), \quad \left\{ \begin{array}{l} \check{a} \\ \check{b} \end{array} \right\} = \arg \min_{a \in \mathbb{Z}^n, b \in \mathbb{R}^p} F(a, b) \quad (7)$$

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

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

$$\left\{ \begin{array}{l} P[\check{a} = z] = \int_{\mathcal{P}_z} f_{\hat{a}}(\alpha) d\alpha \\ f_{\check{b}}(\beta) = \sum_{z \in \mathbb{Z}^n} f_{\hat{b}(z)}(\beta) P[\check{a} = z] \end{array} \right. \quad (8)$$

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 $\mathcal{P}_z = \{x \in \mathbb{R}^n \mid \|x - z\|_{Q_{\hat{a}\hat{a}}}^2 \leq \|x - u\|_{Q_{\hat{a}\hat{a}}}^2, \forall u \in \mathbb{Z}^n\}$. ■

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 $P[\check{a} = a]$. When the success-rate is high enough, one may neglect the uncertainty in \check{a} and describe the uncertainty in \check{b} by means of the PDF of $\hat{b}(a)$. This is made precise by the following bounds of Teunissen (1999b),

$$P[\hat{b}(a) \in \Omega] P[\check{a} = a] \leq P[\check{b} \in \Omega] \leq P[\hat{b}(a) \in \Omega] \quad (9)$$

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

$$P[\check{b} \in \Omega] \approx P[\hat{b}(a) \in \Omega] \quad (10)$$

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} , $P[\hat{b}(a) \in \Omega] \gg P[\hat{b} \in \Omega]$.

3 A dual mixed ILS formulation

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. 3) is computed and then the fixed baseline estimate as $\check{b} = \hat{b}(\check{a})$ (cf. 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

$$\min_{a \in \mathbb{Z}^n} \left[\min_{b \in \mathbb{R}^p} F(a, b) \right] = \min_{b \in \mathbb{R}^p} \left[\min_{a \in \mathbb{Z}^n} F(a, b) \right] \quad (11)$$

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

$$\left\{ \begin{array}{l} \hat{b}(a) = \arg \min_{b \in \mathbb{R}^p} F(a, b) \\ \check{a}(b) = \arg \min_{a \in \mathbb{Z}^n} F(a, b) \end{array} \right. \quad (12)$$

we have the following result.

Theorem 2 (Primal and Dual Mixed ILS) *Let the primal and dual objective functions be defined as*

$$\left\{ \begin{array}{l} \text{Primal : } \mathcal{P}(a) = \min_{b \in \mathbb{R}^p} F(a, b) = F(a, \hat{b}(a)) \\ \text{Dual : } \mathcal{D}(b) = \min_{a \in \mathbb{Z}^n} F(a, b) = F(\check{a}(b), b) \end{array} \right. \quad (13)$$

Then the mixed ILS solution is given as

$$\left\{ \begin{array}{l} \check{a} = \arg \min_{a \in \mathbb{Z}^n} \mathcal{P}(a) = \check{a}(\check{b}) \\ \check{b} = \arg \min_{b \in \mathbb{R}^p} \mathcal{D}(b) = \hat{b}(\check{a}) \end{array} \right. \quad (14)$$

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 $F(a, b)$,

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 multi-frequency 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 $P[\check{a}(b) = a]$.

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. 18) is given as*

$$P[\check{a}(b) = z] = P[\hat{a}(b) \in \mathcal{S}_z] = \int_{\mathcal{S}_z} f_{\hat{a}(b)}(x) dx, \forall z \in \mathbb{Z}^n \quad (20)$$

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

$$\hat{a}(b) \sim N_n(a + \Delta a, Q_{\hat{a}(b)\hat{a}(b)}) \quad (21)$$

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

The PMF of $\check{a}(b)$ (cf. 20) is driven by the PDF of $\hat{a}(b)$ (cf. 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 $\mathcal{S}_u, u \in \mathbb{Z}^n$, 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,

$$P[\check{a}(b) = z] \approx \delta_{uz} = \begin{cases} 1 & \text{if } z = u \\ 0 & \text{if } z \neq u \end{cases} \quad (22)$$

The ambiguity success-rate of $\check{a}(b)$ would then be large, i.e. $P[\check{a}(b) = a] \approx 1$, if $u = a$. For this to happen however,

we need $a + \Delta a \in \mathcal{S}_a$, 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, $\Delta a \in \mathcal{S}_0$. 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 $\Delta a = -Q_{\hat{a}\hat{b}}Q_{\hat{b}\hat{b}}^{-1}\Delta b$. Then*

$$\begin{aligned} \|\Delta a\|_{Q_{\hat{a}(b)\hat{a}(b)}}^2 &= \|P_A B \Delta b\|_{Q_{yy}}^2 \\ &\leq \|\Delta b\|_{Q_{\hat{b}(a)\hat{b}(a)}}^2 \end{aligned} \quad (23)$$

■

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 $\Delta a \in \mathcal{S}_0$. It would namely require knowledge of b such that $\Delta b = E(\hat{b}) - b$ 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 cannot 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

$$\begin{aligned} E \begin{bmatrix} p \\ \phi \end{bmatrix} &= \begin{bmatrix} 0 & D^T G \\ \lambda I_m & D^T G \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}, D \begin{bmatrix} p \\ \phi \end{bmatrix} \\ &= \begin{bmatrix} \sigma_p^2 D^T D & 0 \\ 0 & \sigma_\phi^2 D^T D \end{bmatrix} \end{aligned} \quad (24)$$

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

$$\begin{aligned} Q_{\hat{a}(b)\hat{a}(b)} &= \left(A^T Q_{yy}^{-1} A \right)^{-1} = \frac{\sigma_\phi^2}{\lambda^2} D^T D \\ \Delta a &= \left(A^T Q_{yy}^{-1} A \right)^{-1} A^T Q_{yy}^{-1} B \Delta b = \frac{1}{\lambda} D^T G \Delta b \end{aligned} \quad (25)$$

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 $|(\Delta a)_i| \leq \frac{2}{\lambda} \|\Delta b\|, i = 1, \dots, m$. This shows, as $\lambda \approx 20\text{cm}$ in case of GNSS, that $\Delta b = E(\hat{b}) - b$ 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 cannot expect the success-rate of $\check{a}(b)$ to

be large. That is, despite the high precision of the conditional estimate $\hat{a}(b)$, the influence of the unknown b is still too large to have $\check{a}(b) = a$ with high probability. This implies that one will have to evaluate $\mathcal{D}_2(b) = \|\hat{a}(b) - \check{a}(b)\|_{Q_{\hat{a}(b)\hat{a}(b)}}^2$ 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 $\lceil \hat{a} \rceil$, 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 $a + \Delta a \in \mathcal{S}_u \cap \mathcal{R}_u$, with \mathcal{R}_u denoting the integer-rounding pull-in region of $u \in \mathbb{Z}^n$. However, such assumption cannot 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 $\mathcal{S}_u \cap \mathcal{R}_u$. 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.

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,

$$\begin{aligned} (a) \quad \mathcal{D}^\circ(b) &= \|\hat{b} - b\|_{Q_{\hat{b}\hat{b}}}^2 + \min_{a \in \mathbb{Z}^n} \|\hat{a}(b) - a\|_{Q_{\hat{a}(b)\hat{a}(b)}}^2 \\ (b) \quad \mathcal{D}^\bullet(b) &= \|\hat{b} - b\|_{Q_{\hat{b}\hat{b}}}^2 + \|\hat{a}(b) - \mathcal{I}^\bullet(\hat{a}(b))\|_{Q_{\hat{a}(b)\hat{a}(b)}}^2 \end{aligned} \quad (26)$$

In the first type, we have replaced the conditional variance matrix $Q_{\hat{a}(b)\hat{a}(b)}$ by an approximation $Q_{\hat{a}(b)\hat{a}(b)}^\circ$, the idea being that the approximation will then allow for a simpler ambiguity minimization in (26a). For instance, when $Q_{\hat{a}(b)\hat{a}(b)}^\circ$ is chosen to be a diagonal matrix, the minimization in (26a) reduces to a straightforward componentwise integer rounding of $\hat{a}(b)$. In the second approximation type, we have replaced the integer ambiguity minimizer $\check{a}(b) = \arg \min_{a \in \mathbb{Z}^n} \|\hat{a}(b) - a\|_{Q_{\hat{a}\hat{a}}}$ of $\mathcal{D}_2(b)$ (cf. 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 (16) be replaced by $Q_{\hat{a}(b)\hat{a}(b)}^\circ$. Then*

$$\begin{aligned} &\|\hat{b} - b\|_{Q_{\hat{b}\hat{b}}}^2 + \|\hat{a}(b) - a\|_{Q_{\hat{a}(b)\hat{a}(b)}^\circ}^2 \\ &= \|\hat{a} - a\|_{Q_{\hat{a}\hat{a}}^\circ}^2 + \|\hat{b}^\circ(a) - b\|_{Q_{\hat{b}(a)\hat{b}(a)}^\circ}^2 \end{aligned} \quad (27)$$

with

$$\begin{aligned} Q_{\hat{a}\hat{a}}^\circ &= Q_{\hat{a}(b)\hat{a}(b)}^\circ + Q_{\hat{a}\hat{b}}Q_{\hat{b}\hat{b}}^{-1}Q_{\hat{b}\hat{a}} \\ \hat{b}^\circ(a) &= \hat{b} - Q_{\hat{b}\hat{a}}Q_{\hat{a}\hat{a}}^{\circ-1}(\hat{a} - a) \\ Q_{\hat{b}(a)\hat{b}(a)}^\circ &= Q_{\hat{b}\hat{b}} - Q_{\hat{b}\hat{a}}Q_{\hat{a}\hat{a}}^{\circ-1}Q_{\hat{a}\hat{b}} \end{aligned} \quad (28)$$

Proof see Appendix. \square

This result shows that replacing the variance matrix $Q_{\hat{a}(b)\hat{a}(b)}$ in $\mathcal{D}(b)$ (cf. 19) by $Q_{\hat{a}(b)\hat{a}(b)}^\circ$ provides an objective function of the type (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 $Q_{\hat{a}(b)\hat{a}(b)} \rightarrow Q_{\hat{a}(b)\hat{a}(b)}^\circ$ resulted in three changes of the primal formulation: $Q_{\hat{a}\hat{a}} \rightarrow Q_{\hat{a}\hat{a}}^\circ$, $\hat{b}(a) \rightarrow \hat{b}^\circ(a)$, and $Q_{\hat{b}(a)\hat{b}(a)} \rightarrow Q_{\hat{b}(a)\hat{b}(a)}^\circ$. These changes will therefore also drive the properties of the corresponding baseline- and integer ambiguity estimators. Using the quadratic identity (27), the following equivalence for the minimizer of $\mathcal{D}^\circ(b)$ can be established.

Theorem 3 *Let the approximate dual be given as*

$$\mathcal{D}^\circ(b) = \|\hat{b} - b\|_{Q_{\hat{b}\hat{b}}}^2 + \|\hat{a}(b) - \check{a}^\circ(b)\|_{Q_{\hat{a}(b)\hat{a}(b)}^\circ}^2 \quad (29)$$

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

$$\left. \begin{aligned} \check{a}^\circ &= \arg \min_{a \in \mathbb{Z}^n} \mathcal{P}^\circ(a) \\ \check{b}^\circ &= \hat{b}^\circ(\check{a}^\circ) \end{aligned} \right\} \Leftrightarrow \left\{ \begin{aligned} \check{b}^\circ &= \arg \min_{b \in \mathbb{R}^p} \mathcal{D}^\circ(b) \\ \check{a}^\circ &= \check{a}^\circ(\check{b}^\circ) \end{aligned} \right. \quad (30)$$

Proof see Appendix. \square

This equivalence can now be used to apply available theory for the primal formulation to determine the distributional properties of the estimators \check{a}° and \check{b}° . It should hereby be noted, however, although \check{a}° , like \check{a} , is still computed as the solution of an ILS-problem having \hat{a} as its input, the weight

matrix used is now different, $Q_{\hat{a}\hat{a}}^{\circ-1}$ instead of $Q_{\hat{a}\hat{a}}^{-1}$. 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 $Q_{\hat{b}(a)\hat{b}(a)}^\circ$ of (28) is therefore not the variance matrix of $\hat{b}^\circ(a)$. We have the following distributional result.

Theorem 4 (Distributions of \check{a}° and \check{b}°) *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 \check{a}° is given as*

$$\check{a}^\circ \sim P[\check{a}^\circ = z] = \int_{\mathcal{P}_z^\circ} f_{\hat{a}}(\alpha) d\alpha, \text{ with} \quad (31)$$

$$P[\check{a}^\circ = a] \leq P[\check{a} = a]$$

and the limiting PDF of the baseline estimator \check{b}° is given as

$$\lim_{P[\check{a}^\circ = a] \uparrow 1} \check{b}^\circ \sim \mathcal{N}_p(b, Q_{\hat{b}^\circ(a)\hat{b}^\circ(a)}), \text{ with} \quad (32)$$

$$Q_{\hat{b}^\circ(a)\hat{b}^\circ(a)} = Q_{\hat{b}(a)\hat{b}(a)} + T_{\hat{b}\hat{a}} Q_{\hat{a}\hat{a}} T_{\hat{b}\hat{a}}^T \geq Q_{\hat{b}(a)\hat{b}(a)}$$

where $\mathcal{P}_z^\circ = \{x \in \mathbb{R}^n \mid \|x - z\|_{Q_{\hat{a}\hat{a}}^\circ}^2 \leq \|x - u\|_{Q_{\hat{a}\hat{a}}^\circ}^2, \forall u \in \mathbb{Z}^n\}$ and $T_{\hat{b}\hat{a}} = Q_{\hat{b}\hat{a}} \left[Q_{\hat{a}\hat{a}}^{-1} - Q_{\hat{a}\hat{a}}^{\circ-1} \right]$. ■

Proof See Appendix. □

The above result shows that replacing $Q_{\hat{a}(b)\hat{a}(b)}$ by $Q_{\hat{a}(b)\hat{a}(b)}^\circ$ will always degrade the performance of the associated estimators. It will give a smaller ambiguity success-rate (cf. 31), as well as a poorer precision of the ambiguity-fixed baseline (cf. 32). Still, depending on the choice made for $Q_{\hat{a}(b)\hat{a}(b)}^\circ$, the degradation could be acceptably small, depending on the application.

To evaluate the success-rate $P[\check{a}^\circ = a]$, the multivariate integral of (31) needs to be computed. This is a nontrivial numerical task due to the geometric complexity of the pull-in region \mathcal{P}_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 $P[\check{a}^\circ = a]$ is driven by both $Q_{\hat{a}\hat{a}}$ and $Q_{\hat{a}\hat{a}}^\circ$, i.e. by the ambiguity variance matrix that determines $f_{\hat{a}}(\alpha)$ and by its approximation that determines the pull-in region \mathcal{P}_a° .

The following is such example where the two success-rates $P[\check{a} = a]$ and $P[\check{a}^\circ = a]$ are compared.

Example 2 Consider the GNSS model of Example 1 (cf. 24) and assume that in solving the dual problem we approximate the fully populated variance matrix $Q_{\hat{a}(b)\hat{a}(b)} = \frac{\sigma_\phi^2}{\lambda} D^T D$ with the diagonal matrix $Q_{\hat{a}(b)\hat{a}(b)}^\circ = 2 \frac{\sigma_\phi^2}{\lambda} I_m$. Then, with $Q_{\hat{a}\hat{b}} Q_{\hat{b}\hat{b}}^{-1} Q_{\hat{b}\hat{a}} = \frac{\sigma_\phi^2}{\lambda^2} D^T G [G^T P_D G]^{-1} G^T D$ and $P_D = D(D^T D)^{-1} D^T$, we get

$$Q_{\hat{a}\hat{a}}^\circ = \frac{1}{\lambda^2} \left(2\sigma_\phi^2 I_m + \sigma_p^2 D^T G [G^T P_D G]^{-1} G^T D \right) \quad (33)$$

With reference to Theorem 4 (cf. 31), Fig. 2 compares the two success-rates $P[\check{a} = a]$ and $P[\check{a}^\circ = a]$, based on (33), for a case of single-epoch, single-frequency L1 GPS, using a 10^{-4} phase-code variance ratio. It shows indeed that $P[\check{a}^\circ = a] \leq P[\check{a} = a]$, but also that $P[\check{a}^\circ = a]$ 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 $P[\check{a}^\circ = a]$, one may directly use the ILS success-rate $P[\check{a} = a]$ (cf. 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 \check{a}° .

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

(i) *If $Q_{\hat{a}\hat{a}}^\circ \geq Q_{\hat{a}\hat{a}}$, then*

$$\prod_{i=1}^n [2\Phi\left(\frac{1}{2\sqrt{d_i}}\right) - 1] \leq P[\check{a}^\circ = a] \quad (34)$$

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

(ii) *For any $Q_{\hat{a}\hat{a}}^\circ > 0$,*

$$P[\chi^2(0, n) \leq r^2] \leq P[\check{a}^\circ = a] \quad (35)$$

with

$$\begin{cases} r^2 &= \lambda_{\min} \times \frac{1}{4} \min_{z \in \mathbb{Z}^n / \{0\}} \|z\|_{Q_{\hat{a}\hat{a}}^\circ}^2 \\ \lambda_{\min} &= \min_{x \in \mathbb{R}^n} \frac{x^T Q_{\hat{a}\hat{a}}^\circ x}{x^T Q_{\hat{a}\hat{a}} x} \end{cases} \quad (36)$$

Proof see Appendix □

Note that lower-bound (34) is somewhat easier to compute than (35). It requires however that $Q_{\hat{a}\hat{a}}^\circ \geq Q_{\hat{a}\hat{a}}$, while no such restriction is placed on the lower-bound (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.

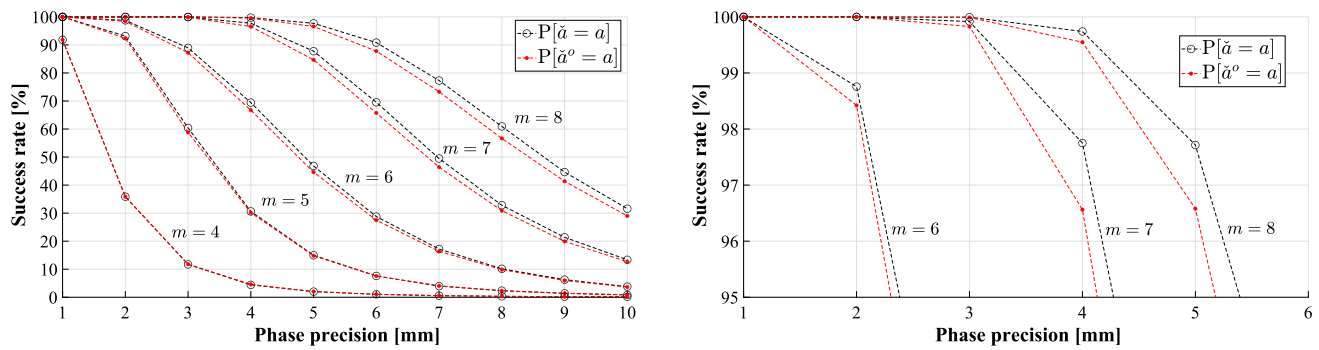


Fig. 2 Comparison of $P[\check{a}^o = a]$ with $P[\check{a} = a]$ as described in Example 2, based on the GNSS model of Example 1. The plot on the right side shows a zoom-in of the left one

5 Dual with approximate integer map

In this section we consider the second approximation of the dual function, $\mathcal{D}^\bullet(b)$ (cf. 26), and determine the statistical properties of its minimizer.

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

$$\mathcal{D}^\bullet(b) = \|\hat{b} - b\|_{Q_{\hat{b}\hat{b}}}^2 + \|\hat{a}(b) - \mathcal{I}^\bullet(\hat{a}(b))\|_{Q_{\hat{a}(b)\hat{a}(b)}}^2 \quad (37)$$

is given as

$$\check{b}^\bullet = \hat{b}(\check{a}^\bullet), \text{ with } \check{a}^\bullet = \arg \min_{z \in \Omega_a^\bullet} \|\hat{a} - z\|_{Q_{\hat{a}\hat{a}}}^2 \quad (38)$$

where

$$\Omega_a^\bullet = \{z \in \mathbb{Z}^n \mid z = \mathcal{I}^\bullet(\hat{a} + M\beta), \forall \beta \in \mathbb{R}^p\} \quad (39)$$

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

Proof see Appendix. \square

Note, in contrast to the baseline estimator \check{b}° (cf. 30), that the baseline estimator \check{b}^\bullet (cf. 38) is based, like the estimator \check{b} (cf. 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. 38) has the appearance of a standard ILS-estimator, it is generally not, unless of course $\mathcal{I}^\bullet(x)$ is chosen as $\mathcal{I}^\bullet(x) = \arg \min_{z \in \mathbb{Z}^n} \|x - z\|_{Q_{\hat{a}(b)\hat{a}(b)}}^2$, in which case $\check{a}^\bullet = \check{a}$, since then $\mathcal{D}^\bullet(b) = \mathcal{D}(b)$.

The difference between the two integer estimators \check{a} and \check{a}^\bullet is driven by the characteristics of the integer set Ω_a^\bullet (cf. 39), which on its turn is driven by the $n \times p$ matrix M in

$$\hat{a}(\beta) = \hat{a} + M(\beta - \hat{b}) \quad (40)$$

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 (39) reduces then to the single integer vector $\Omega_a^\bullet = \{z = \mathcal{I}^\bullet(\hat{a})\}$, we have in that case $\check{a}^\bullet = \mathcal{I}^\bullet(\hat{a})$. In case (2), the subset Ω_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 (40). This is shown in Fig. 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 $\Omega_a^\bullet \neq \mathbb{Z}^n$, we have $\check{a}^\bullet \neq \check{a}$ 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 $\Omega_a^\bullet = \mathbb{Z}^n$. This shows that $\check{a}^\bullet = \check{a}$ 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, \check{a}^\bullet and \check{a} , is very unlikely in case of GNSS. Although the two estimators \check{a}^\bullet and \check{a} are

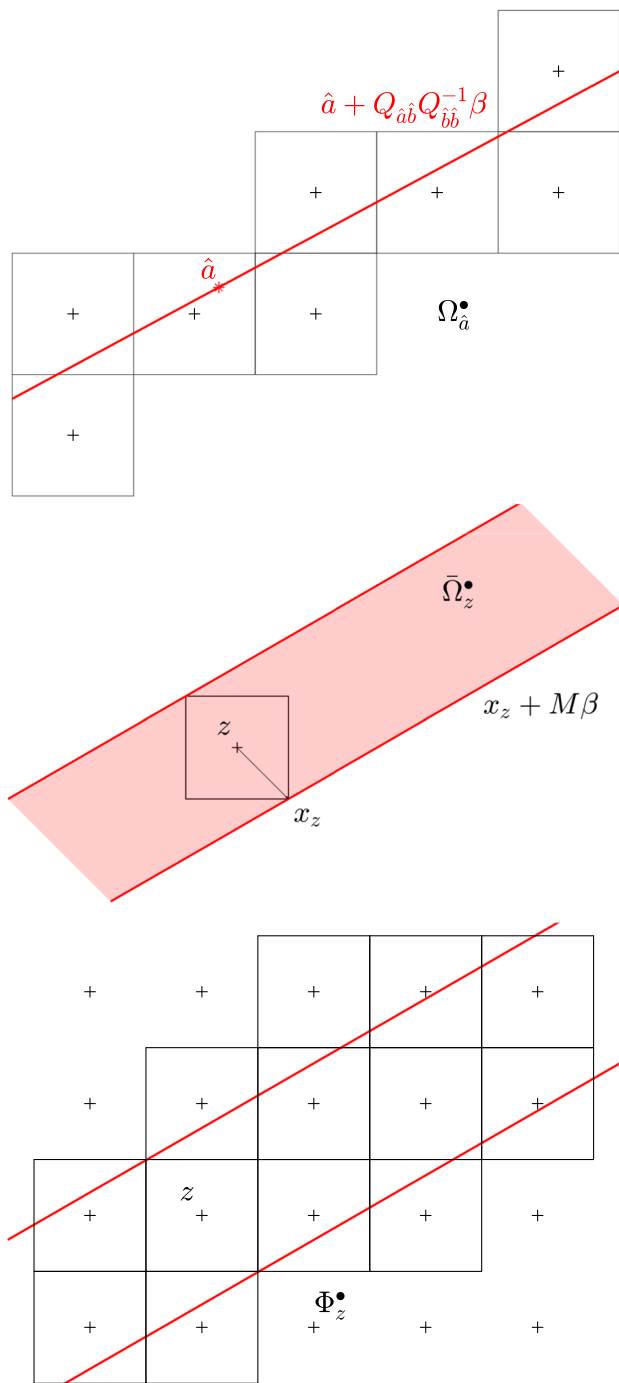


Fig. 3 (Top) The integer set $\Omega_{\hat{a}}^{\bullet} = \{z \in \mathbb{Z}^n \mid z = \mathcal{I}^{\bullet}(\hat{a} + M\beta), \forall \beta \in \mathbb{R}^p\}$ for $n = 2, p = 1$, and $\mathcal{I}(\cdot) = \lceil \cdot \rceil$; (Centre) The real-valued set $\bar{\Omega}_z^{\bullet} = \{x \in \mathbb{R}^n \mid z = \mathcal{I}^{\bullet}(x + M\beta), \exists \beta \in \mathbb{R}^p\}$ for $n = 2, p = 1$; (Bottom) The integer set $\Phi_z^{\bullet} = \{u \in \mathbb{Z}^n \mid u = \mathcal{I}^{\bullet}(x + M\beta), x \in \bar{\Omega}_z^{\bullet}, \forall \beta \in \mathbb{R}^p\}$ for $n = 2, p = 1$

then generally different, their integer sample outcomes can, of course, sometimes be the same. This happens when the outcome of \check{a} lies in $\Omega_{\hat{a}}^{\bullet}$.

5.2 A qualitative comparison of \check{a}^{\bullet} and \check{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 $\check{a}^{\bullet} \neq \check{a}$.

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

$$E_{\hat{a}} = \left\{ z \in \mathbb{Z}^n \mid \|\hat{a} - z\|_{Q_{\hat{a}\hat{a}}}^2 \leq \chi^2 \right\} \quad (41)$$

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

$$\begin{aligned} \check{a} &= \arg \min_{u \in E_{\hat{a}}} \|\hat{a} - u\|_{Q_{\hat{a}\hat{a}}}^2 \\ \check{a}^{\bullet} &= \arg \min_{u \in \Omega_{\hat{a}}^{\bullet}} \|\hat{a} - u\|_{Q_{\hat{a}\hat{a}}}^2 \end{aligned} \quad (42)$$

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

$$\Omega_{\hat{a}}^{\bullet} = \{z \in \mathbb{Z}^n \mid z = \mathcal{I}^{\bullet}(\hat{a} + M\beta), \forall \beta \in \mathbb{R}^p\} \quad (43)$$

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

$$E_{\hat{a}} = \left\{ z \in \mathbb{Z}^n \mid (\hat{a} - z)^T \left[Q_{\hat{a}(b)\hat{a}(b)} + MQ_{\hat{b}\hat{b}}M^T \right]^{-1} (\hat{a} - z) \leq \chi^2 \right\} \quad (44)$$

We can now compare the two integer sets (43) and (44). We will do so, for two extreme cases.

$$\text{Case (a) : } Q_{\hat{a}(b)\hat{a}(b)} = \text{small and } Q_{\hat{b}\hat{b}} = \text{large} \quad (45)$$

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 (45) is true, the ellipsoidal search space $E_{\hat{a}}$ (cf. 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, after all, 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 \hat{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).

Case (b) : $Q_{\hat{a}(b)\hat{a}(b)} = \text{large}$ and $Q_{\hat{b}\hat{b}} = \text{small}$ (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.

5.3 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 \mathcal{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

$$\mathcal{P}_z^{\bullet} = \left\{ x \in \mathbb{R}^n \mid z = \arg \min_{u \in \Omega_x^{\bullet}} \|x - u\|_{Q_{\hat{a}\hat{a}}}^2 \right\} \quad (47)$$

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 of $x \in \mathbb{R}^n$ to a subset. As z has to lie in Ω_x^{\bullet} , the choice of z implies the following subset for the values of x ,

$$\bar{\Omega}_z^{\bullet} = \{x \in \mathbb{R}^n \mid z = \mathcal{I}^{\bullet}(x + M\beta), \exists \beta \in \mathbb{R}^p\} \quad (48)$$

This is the set of x -values for which a $\beta \in \mathbb{R}^p$ exists, such that $x + M\beta$ gets mapped by $\mathcal{I}^{\bullet}(\cdot)$ to z , see Fig. 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 (47). As \mathcal{P}_z^{\bullet} is characterized by the minimization of $\|x - u\|_{Q_{\hat{a}\hat{a}}}^2$ over the integer subset Ω_x^{\bullet} , while, at the same time, this is constrained to all $x \in \bar{\Omega}_z^{\bullet}$, the integer set considered is actually

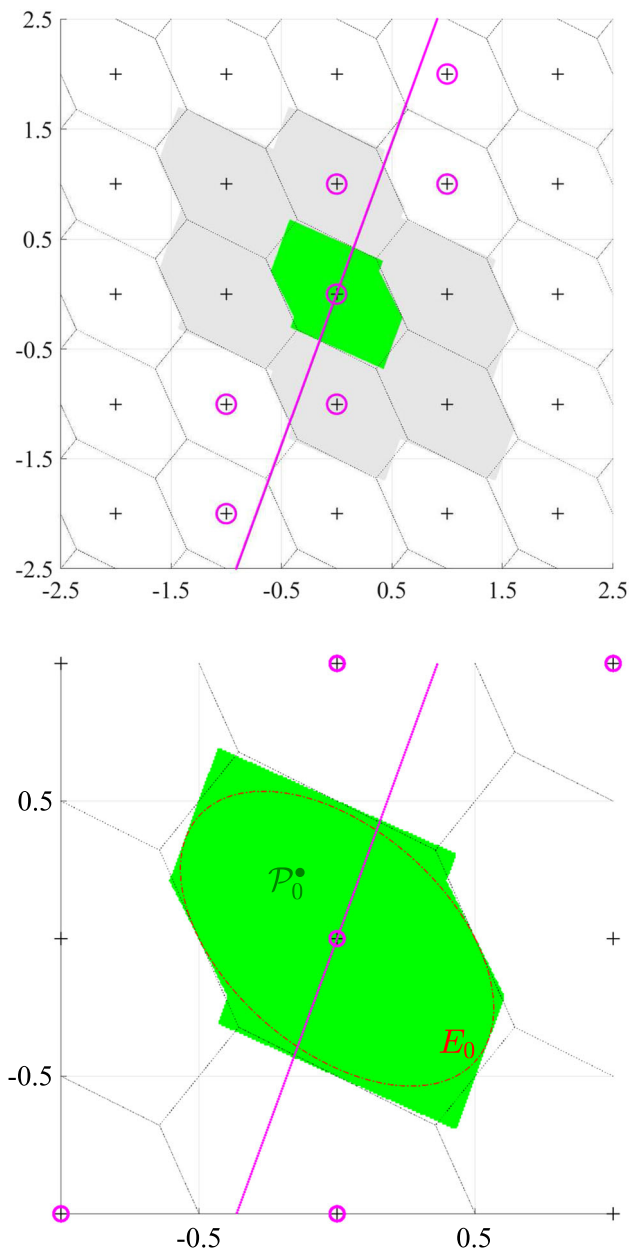


Fig. 4 For $p = 1, n = 2$, the pull-in regions \mathcal{P}_z^{\bullet} of the integer ambiguity estimator \check{a}^{\bullet} are shown, i.e. the regions where float ambiguity values are mapped to the same integer vector $z \in \mathbb{Z}^n$. In green colour, the \mathcal{P}_0^{\bullet} region is depicted for $z = 0$, surrounded by similar regions (in grey) for $z \neq 0$. The top plot shows a comparison with ILS pull-in regions, with the integer set Ω_0^{\bullet} depicted through encircled gridpoints; the bottom plot shows a zoom-in \mathcal{P}_0^{\bullet} of \check{a}^{\bullet} together with the (lower) bounding ellipse E_0 (cf. proof of Theorem 7)

$$\Phi_z^{\bullet} = \{u \in \mathbb{Z}^n \mid u = \mathcal{I}^{\bullet}(x), \forall x \in \bar{\Omega}_z^{\bullet}\} \quad (49)$$

It consists of all integer vectors to which the elements of $\bar{\Omega}_z^{\bullet}$ are mapped by $\mathcal{I}^{\bullet}(\cdot)$, see Fig. 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 \mathcal{P}_z^\bullet) *The pull-in regions \mathcal{P}_z^\bullet , $z \in \mathbb{Z}^n$, of \check{a}^\bullet can be represented, with*

$$\begin{aligned}\Omega_x^\bullet &= \{z \in \mathbb{Z}^n \mid z = \mathcal{I}^\bullet(x + M\beta), \forall \beta \in \mathbb{R}^p\} \\ \bar{\Omega}_z^\bullet &= \{x \in \mathbb{R}^n \mid z = \mathcal{I}^\bullet(x + M\beta), \exists \beta \in \mathbb{R}^p\} \\ \Phi_z^\bullet &= \{u \in \mathbb{Z}^n \mid u = \mathcal{I}^\bullet(x), \forall x \in \bar{\Omega}_z^\bullet\}\end{aligned}$$

as

$$\begin{aligned}\mathcal{P}_z^\bullet &= \left\{x \in \mathbb{R}^n \mid z = \arg \min_{u \in \Omega_x^\bullet} \|x - u\|_{Q_{\hat{a}\hat{a}}}^2\right\} \\ &= \left\{x \in \mathbb{R}^n \mid \|x - z\|_{Q_{\hat{a}\hat{a}}}^2 \leq \|x - \mathcal{I}^\bullet(y)\|_{Q_{\hat{a}\hat{a}}}^2, \forall y \in \bar{\Omega}_z^\bullet\right\} \\ &= \left\{x \in \mathbb{R}^n \mid \|x - z\|_{Q_{\hat{a}\hat{a}}}^2 \leq \|x - u\|_{Q_{\hat{a}\hat{a}}}^2, \forall u \in \Phi_z^\bullet\right\}\end{aligned}\quad (50)$$

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*

$$\begin{aligned}(i) \quad & \bigcup_{z \in \mathbb{Z}^n} \mathcal{P}_z^\bullet = \mathbb{R}^n \\ (ii) \quad & \mathcal{P}_{z_1}^\bullet \cap \mathcal{P}_{z_2}^\bullet = \emptyset, \forall z_1 \neq z_2 \\ (iii) \quad & \mathcal{P}_z^\bullet = \mathcal{P}_0^\bullet + z, \forall z \in \mathbb{Z}^n\end{aligned}\quad (51)$$

Proof see Appendix. \square

The admissibility property implies that if y is perturbed by Az to give $y' = y + Az$, the ambiguity float solution changes from $\hat{a} = \hat{A}^+ y$ to $\hat{a}' = \hat{A}^+ (y + Az) = \hat{a} + z$, 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 manageable 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 4 shows, for $n = 2$ and $p = 1$, an example of the pull-in regions \mathcal{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. $\mathcal{I}^\bullet(x) = \lceil x \rceil$. 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.

5.4 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 (Teunissen 1999b) as

$$\begin{cases} P[\check{a}^\bullet = z] = \int_{\mathcal{P}_z^\bullet} f_{\hat{a}}(\alpha) d\alpha \\ f_{\check{b}^\bullet}(\beta) = \sum_{z \in \mathbb{Z}^n} f_{\hat{b}(z)}(\beta) P[\check{a}^\bullet = z] \end{cases}\quad (52)$$

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 $P[\check{a}^\bullet = a]$ 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 $P[\check{a}^\bullet = a]$) *Let $\hat{a} \sim \mathcal{N}_n(a, Q_{\hat{a}\hat{a}})$ and $\check{a}^\bullet = \arg \min_{z \in \Omega_{\hat{a}}^\bullet} \|\hat{a} - z\|_{Q_{\hat{a}\hat{a}}}^2$, with $\Omega_{\hat{a}}^\bullet = \{u \in \mathbb{Z}^n \mid u = \mathcal{I}^\bullet(\hat{a} + Q_{\hat{a}\hat{b}} Q_{\hat{b}\hat{b}}^{-1} \beta), \forall \beta\}$. Then*

$$P[\chi^2(0, n) \leq r^2] \leq P[\check{a}^\bullet = a] \quad (53)$$

with $r^2 = \frac{1}{4} \min_{z \in \mathbb{Z}^n \setminus \{0\}} \|z\|_{Q_{\hat{a}\hat{a}}}^2$.

Proof see Appendix. \square

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. 19) simplify then to

$$\begin{aligned}\mathcal{D}_1(b) &= \|\hat{b} - b\|_{Q_{\hat{b}\hat{b}}}^2 = (\hat{b} - b)^2 / \sigma_b^2 \\ \mathcal{D}_2(b) &= \|\hat{a}(b) - \check{a}(b)\|_{Q_{\hat{a}(b)\hat{a}(b)}}^2 = (\hat{a}(b) - \lceil \hat{a}(b) \rceil)^2 / \sigma_{\hat{a}(b)}^2\end{aligned}\quad (54)$$

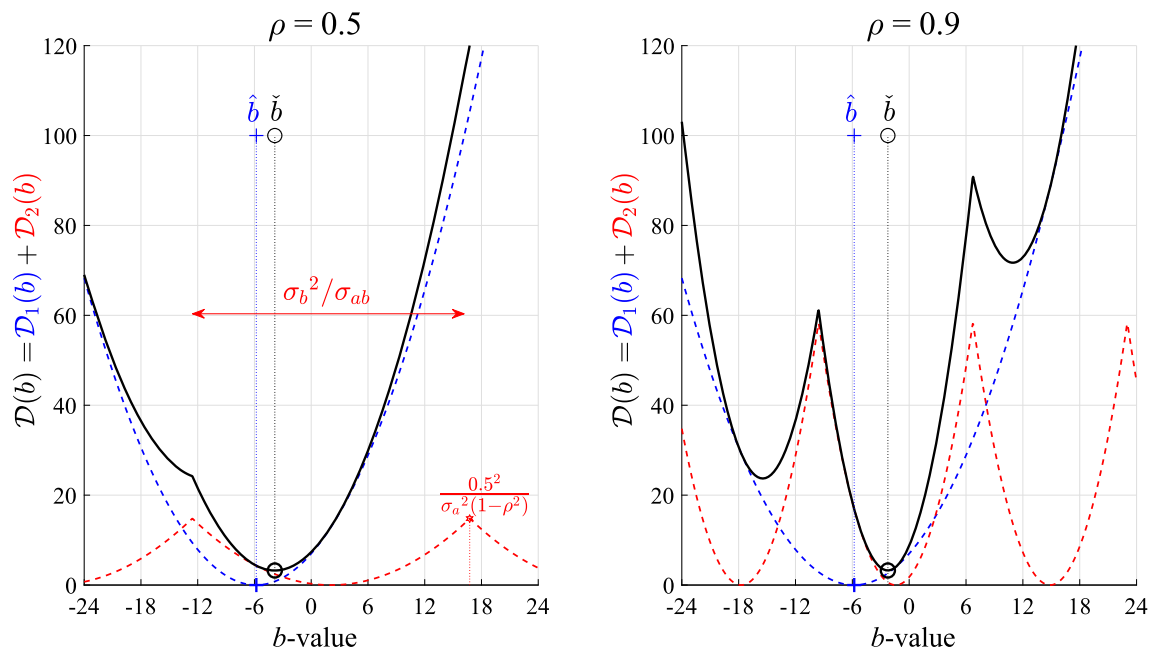


Fig. 5 The dual function $\mathcal{D}(b)$ for $n = p = 1$, $\sigma_b^2 = 2.2^2$, $\sigma_a^2 = 0.15^2$ and $\rho_{\hat{a}\hat{b}} = 0.5$ (left), $\rho_{\hat{a}\hat{b}} = 0.9$ (right)

For the second component, we may write $\mathcal{D}_2(b) = (\hat{a}(b) - z)^2 / \sigma_{\hat{a}(b)}^2$ if $\hat{a}(b) \in [z - \frac{1}{2}, z + \frac{1}{2}]$. Since $\hat{a}(b) = \hat{a} - \sigma_{\hat{a}\hat{b}} \sigma_{\hat{b}}^{-2} (\hat{b} - b)$, we have the equivalence

$$\hat{a}(b) \in [z - \frac{1}{2}, z + \frac{1}{2}] \Leftrightarrow b \in [b(z) - \frac{1}{2}\Delta, b(z) + \frac{1}{2}\Delta] \quad (55)$$

where $b(z) = \hat{b} + (z - \hat{a})\Delta$ and $\Delta = \sigma_b^2 / \sigma_{\hat{a}\hat{b}}$ (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

$$\mathcal{D}(b) = \sigma_b^{-2} \left\{ (\hat{b} - b)^2 + \frac{\rho_{\hat{a}\hat{b}}^2}{1 - \rho_{\hat{a}\hat{b}}^2} (b(z) - b)^2 \right\} \quad (56)$$

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 parabolas $\mathcal{D}_1(b) = (\hat{b} - b)^2 / \sigma_b^2$ and an infinite z -sequence of equally shaped parabola $\mathcal{D}_2(b) = \frac{\rho_{\hat{a}\hat{b}}^2}{1 - \rho_{\hat{a}\hat{b}}^2} (b(z) - b)^2 / \sigma_b^2$, centred at $b(z)$ and with domain $b \in [b(z) - \frac{1}{2}\Delta, b(z) + \frac{1}{2}\Delta]$.

Equation (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 Fig. 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 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

$$\begin{cases} \hat{b}(z) &= \arg \min_{\hat{a}(b) \in \mathcal{S}_z} \mathcal{D}(b) \\ \mathcal{D}(\hat{b}(z)) &= \|\hat{a} - z\|_{Q_{\hat{a}\hat{a}}}^2, \forall z \in \mathbb{Z}^n \end{cases} \quad (57)$$

(b) The global minimizer of $\mathcal{D}(b)$,

$$\check{b} = \arg \min_{b \in \mathbb{R}^n} \mathcal{D}(b) \in \mathcal{E}(r_z), \quad z \in \mathbb{Z}^n \quad (58)$$

resides in the ellipsoidal region $\mathcal{E}(r_z) = \{b \in \mathbb{R}^n \mid \|\hat{b} - b\|_{Q_{\hat{b}\hat{b}}}^2 \leq r_z^2\}$, for all $z \in \mathbb{Z}^n$, with $r_z^2 = \|\hat{a} - z\|_{Q_{\hat{a}\hat{a}}}^2$.

Proof First we prove (57). For $\hat{a}(b) \in \mathcal{S}_z$, we have $\check{a}(b) = z$, and therefore

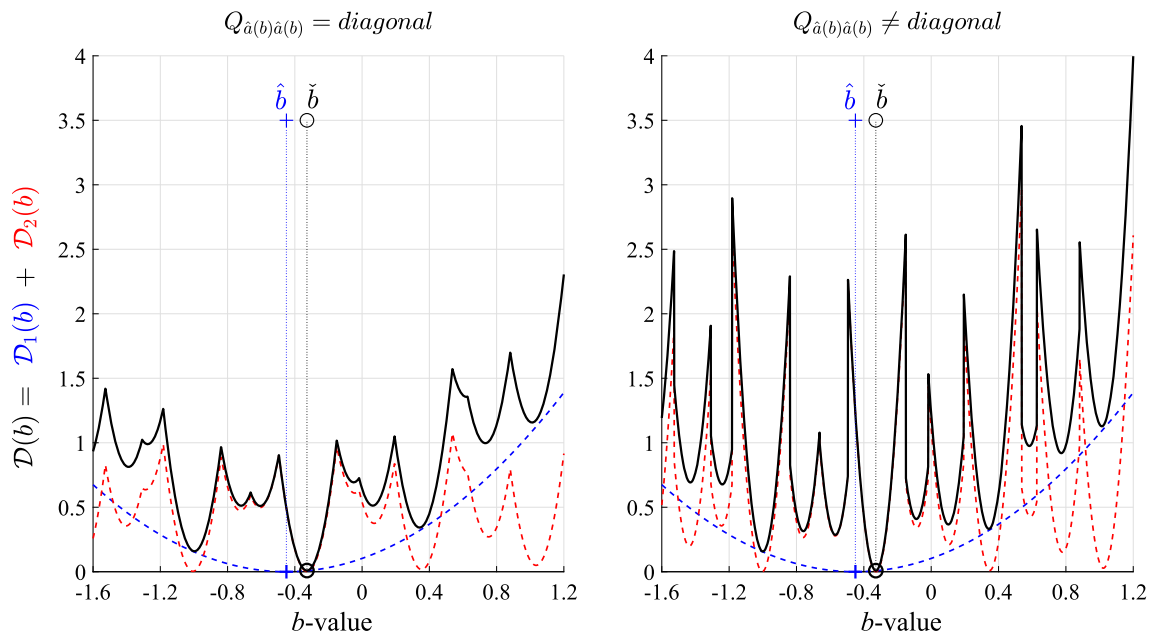


Fig. 6 The dual function $\mathcal{D}(b)$ for $p = 1$ and $n = 2$, with diagonal (left plot) and nondiagonal (right plot) conditional ambiguity variance matrix, having 0.8 correlation

$$\begin{aligned} \mathcal{D}(b) &= \|\hat{b} - b\|_{Q_{\hat{b}\hat{b}}}^2 + \|\hat{a}(b) - z\|_{Q_{\hat{a}(b)\hat{a}(b)}}^2 \\ &= \|\hat{a} - z\|_{Q_{\hat{a}\hat{a}}}^2 + \|\hat{b}(z) - b\|_{Q_{\hat{b}(a)\hat{b}(a)}}^2 \end{aligned} \quad (59)$$

showing that its local minimizer and minimum are given by (57). That no other minima exist of $\mathcal{D}(b)$ follows from the fact that the pull-in regions \mathcal{S}_z partition \mathbb{R}^n , $\forall z \in \mathbb{Z}^n$ (Teunissen 1999a). The proof of (58) follows by recognizing that as the global minimizer \check{b} is one of the local minimizers, \check{b} resides in the set $\{b \in \mathbb{R}^p \mid \mathcal{D}(b) \leq \|\hat{a} - z\|_{Q_{\hat{a}\hat{a}}}^2\}$, and thus also in the larger set E_z . \square

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 than 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 $\mathcal{D}(b)$.

7 Global minimization of dual

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 $\mathcal{E}(r) = \{b \in \mathbb{R}^p \mid \|\hat{b} - b\|_{Q_{\hat{b}\hat{b}}}^2 \leq r^2\}$ or any other of its circumscribing convex regions. The problem to be solved reads therefore

$$\check{b} = \arg \min_{b \in \mathcal{C} \subset \mathbb{R}^p} \mathcal{D}(b) \quad (60)$$

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 (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.

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 ,

$$\mathcal{C} = \cup_{i=1}^k \mathcal{B}_i \quad (61)$$

- to find *local* lower and upper bounds of $\mathcal{D}(b)$ for each box \mathcal{B}_i ,

$$L(\mathcal{B}_i) \leq \min_{b \in \mathcal{B}_i} \mathcal{D}(b) \leq U(\mathcal{B}_i), \quad (62)$$

- to form *global* bounds from the local bounds,

$$L_k \leq \min_{b \in \mathcal{C}} \mathcal{D}(b) \leq U_k \quad (63)$$

where

$$L_k = \min_{i=1, \dots, k} L(\mathcal{B}_i) \text{ and } U_k = \min_{i=1, \dots, k} U(\mathcal{B}_i) \quad (64)$$

- 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),

$$L_1 = L(\mathcal{C}) \leq \mathcal{D}(\check{b}) \leq U_1 = U(\mathcal{C}) \quad (65)$$

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,

$$\underbrace{\min(L(\mathcal{B}_1), L(\mathcal{B}_2))}_{L_2} \leq \mathcal{D}(\check{b}) \leq \underbrace{\min(U(\mathcal{B}_1), U(\mathcal{B}_2))}_{U_2} \quad (66)$$

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 (61), with associated global lower and upper bounds of $\mathcal{D}(\check{b})$ as given in (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.

7.1.1 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

$$\mathcal{C} = \left\{ b \in \mathbb{R}^p \mid |b_\alpha - \hat{b}_\alpha| \leq r \sigma_{\hat{b}_\alpha}, \alpha = 1, \dots, p \right\} \quad (67)$$

It follows from the ellipsoidal *planes-of-support* lemma that the box \mathcal{C} is circumscribed by the ellipsoid $\mathcal{E}(r) = \{b \in \mathbb{R}^p \mid \|b - \hat{b}\|_{Q_{\hat{b}\hat{b}}}^2 \leq r^2\}$, 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, $r^2 = \chi_\alpha^2(p, 0)$ 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

$$U(\mathcal{B}) = \mathcal{D}(b_*) \geq \min_{b \in \mathcal{B}} \mathcal{D}(b) \quad (68)$$

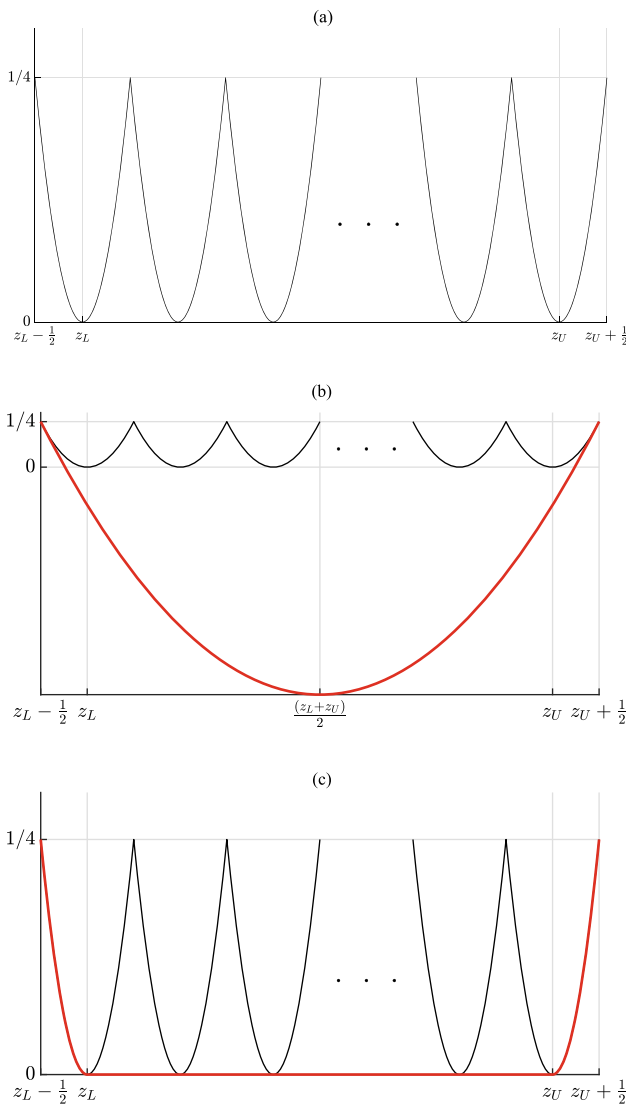


Fig. 7 **a** The function $g(x) = (x - \lceil x \rceil)^2$ as a sequence of cut-off parabolas on interval $[z_L - \frac{1}{2}, z_U + \frac{1}{2}]$; **b** with its parabolic lower bounding function; and **c** with its best possible, continuous differentiable lower bounding convex function

with b_* being the 'centre of gravity' of the box, i.e. if box \mathcal{B} is bounded as $b_L \leq b \leq b_U$, 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

$$L(\mathcal{B}) = \min_{b \in \mathcal{B}} \mathcal{D}_L(b) \leq \min_{b \in \mathcal{B}} \mathcal{D}(b) \quad (69)$$

We now show how this can be achieved.

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 $Q_{\hat{a}(b)\hat{a}(b)}^\circ = \text{diag}(\sigma_1^2, \dots, \sigma_n^2)$, the dual function (29) can be written as

$$\mathcal{D}^\circ(x) = \|b - \hat{b}\|_{Q_{\hat{b}\hat{b}}}^2 + \mathcal{G}(b) \quad (70)$$

with

$$\begin{aligned} \mathcal{G}(b) &= \sum_{i=1}^n g(x_i) / \sigma_i^2 \\ g(x) &= (x - \lceil x \rceil)^2 \\ x_i &= \hat{a}_i(b) = \hat{a}_i + m_i^T (b - \hat{b}) \\ m_i^T &= c_i^T Q_{\hat{a}\hat{b}} Q_{\hat{b}\hat{b}}^{-1} \\ c_i &= i\text{th column of unit matrix } I_n \end{aligned} \quad (71)$$

Since the first term of (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

$$\mathcal{G}_L(b) \leq \mathcal{G}(b), \quad \forall b \in \mathcal{B} \quad (72)$$

Once this function is found, we have found the dual convex lower bounding function as $\mathcal{D}_L^\circ(b) = \|b - \hat{b}\|_{Q_{\hat{b}\hat{b}}}^2 + \mathcal{G}_L(b) \leq \mathcal{D}(b), \forall b \in \mathcal{B}$.

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 parabolas, see Fig. 7a. 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 $\hat{a}_i(b)$, $i = 1, \dots, n$, that correspond with $b_L \leq b \leq b_U$. 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

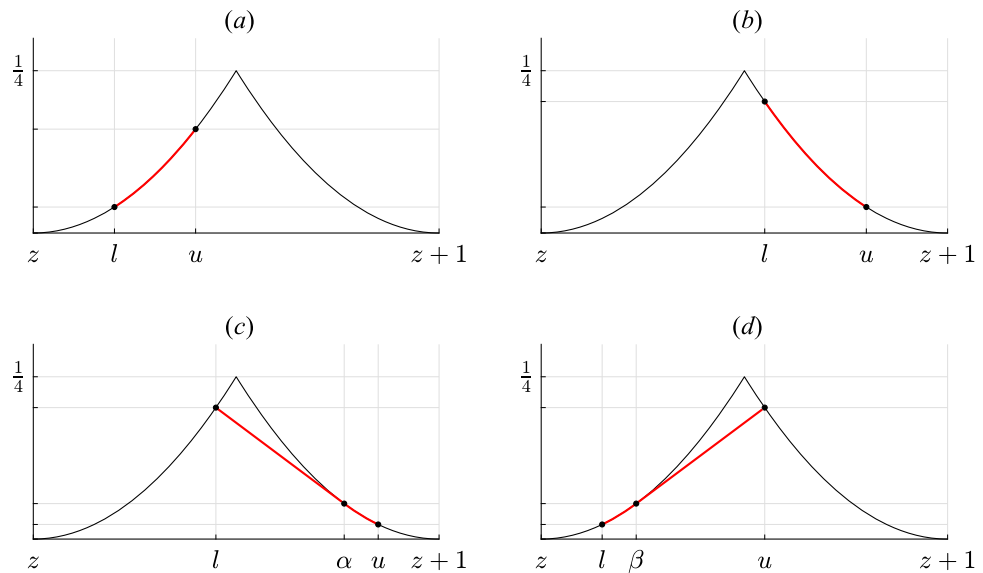
As the function $\mathcal{G}_L(b) = \sum_{i=1}^n g_{i,L}(\hat{a}_i(b)) / \sigma_i^2$ has to be convex for the box $b_{\alpha,L} \leq b_\alpha \leq b_{\alpha,U}$, $\alpha = 1, \dots, p$, the functions $g_{i,L}(x)$ need to be convex for the intervals $l_i \leq a_i(b) \leq u_i$ that correspond with this box. Application of the *projection-lemma* from the Appendix shows the relation between these intervals given as

$$b_{\alpha,L} \leq b_\alpha \leq b_{\alpha,U} \Rightarrow l_i \leq \hat{a}_i(b) \leq u_i \quad (73)$$

with

$$\begin{aligned} l_i &= \hat{a}_i(0) + \sum_{\alpha=1}^p \left(-\tilde{b}_\alpha |m_{i\alpha}| + \tilde{b}_\alpha m_{i\alpha} \right) \\ u_i &= \hat{a}_i(0) + \sum_{\alpha=1}^p \left(+\tilde{b}_\alpha |m_{i\alpha}| + \tilde{b}_\alpha m_{i\alpha} \right) \end{aligned} \quad (74)$$

Fig. 8 Four convex differentiable lower bounding functions (in red) of $g(x) = (x - \lceil x \rceil)^2$ on the interval $[l, u] \subset [z, z + 1]$, $z \in \mathbb{Z}$



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 $b_{\alpha,U} = b_{\alpha,L} = \tilde{b}_\alpha$, the result $\hat{a}(\tilde{b})$.

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 γ th interval $[b_{\gamma,L}, b_{\gamma,U}]$, with $\gamma \in \{1, \dots, p\}$ 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

$$\begin{aligned} l_i^1 &= l_i + \frac{1}{2} \tilde{b}_\gamma (|m_{i\gamma}| - m_{i\gamma}) \\ u_i^1 &= u_i - \frac{1}{2} \tilde{b}_\gamma (|m_{i\gamma}| + m_{i\gamma}) \end{aligned} \quad (75)$$

Hence, the length of the interval changes as $(u_i^1 - l_i^1) = (u_i - l_i) - b_\gamma |m_{i\gamma}|$, i.e. it gets shorter by $b_\gamma |m_{i\gamma}|$, 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 $g_{i,L}(x)$, $i = 1, \dots, n$, 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 $[l, u]$

Many different differentiable convex lower bounding functions $g_L(x)$ of $g(x) = (x - \lceil x \rceil)^2$ 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

$$g_L(x) = \alpha(x - \bar{z})^2 + \beta \quad (76)$$

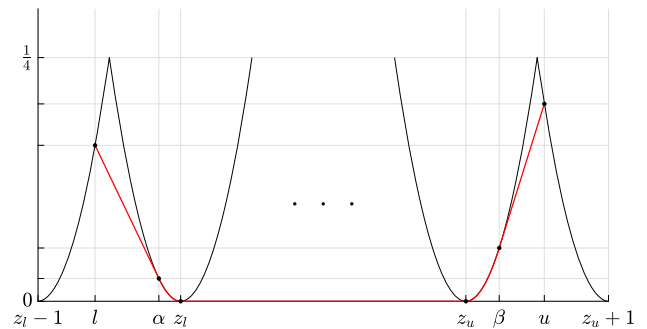


Fig. 9 The linear-parabolic differentiable lower bounding function (in red) of $g(x) = (x - \lceil x \rceil)^2$ when $l \in [z_l - 1, z_l - \frac{1}{2}]$, $u \in [z_u + \frac{1}{2}, z_u + 1]$ for $z_u \geq z_l$

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 Fig. 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 (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 Fig. 7c)

$$g_L(x) = \begin{cases} (x - z_l)^2 & \text{for } l = z_l - \frac{1}{2} \leq x \leq z_l \\ 0 & \text{for } z_l \leq x \leq z_u \\ (x - z_u)^2 & \text{for } z_u \leq x \leq z_u + \frac{1}{2} = u \end{cases} \quad (77)$$

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 (76) is a C^2 function, while (77) is only a C^1 function. The continuous differentiability of (77) is sufficient however for the

application of the projected gradient descent method (see Sect. 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 Fig. 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 Fig. 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 Fig. 8b.

For the third and the fourth case, we assume that l and u lie in different halves 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 Fig. 8c. It is given as:

$$g_L(x) = \begin{cases} \stackrel{u \leq \alpha}{=} a(x - l) + (l - z)^2 & \text{for } l \leq x \leq u \\ \stackrel{u \geq \alpha}{=} \begin{cases} = 2(\alpha - z - 1)(x - \alpha) + (\alpha - z - 1)^2 & \text{for } l \leq x \leq \alpha \\ = (x - z - 1)^2 & \text{for } \alpha \leq x \leq u \end{cases} \end{cases} \quad (78)$$

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 Fig. 8d. The convex lower bounding differentiable function is given as:

$$g_L(x) = \begin{cases} \stackrel{l \geq \beta}{=} a(x - l) + (l - z)^2 & \text{for } l \leq x \leq u \\ \stackrel{l \leq \beta}{=} \begin{cases} = (x - z)^2 & \text{for } l \leq x \leq \beta \\ = 2(\beta - z)(x - \beta) + (\beta - z)^2 & \text{for } \beta \leq x \leq u \end{cases} \end{cases} \quad (79)$$

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 Fig. 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

$$g_L(x) = \begin{cases} f_1(l, x) & \text{for } l \leq x \leq z_l \\ 0 & \text{for } z_l \leq x \leq z_u \\ f_2(x, u) & \text{for } z_u \leq x \leq u \end{cases} \quad (80)$$

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 (78), through replacement $z := z_l - 1$, and from (79), through replacement $z := z_u$, as

$$f_1(l, x) = \begin{cases} 2(\alpha - z_l)(x - \alpha) + (\alpha - z_l)^2 & \text{for } l \leq x \leq \alpha \\ (x - z_l)^2 & \text{for } \alpha \leq x \leq z_l \end{cases} \quad (81)$$

and

$$f_2(x, u) = \begin{cases} (x - z_u)^2 & \text{for } z_u \leq x \leq \beta \\ 2(\beta - z_u)(x - \beta) + (\beta - z_u)^2 & \text{for } \beta \leq x \leq u \end{cases} \quad (82)$$

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 (81) and (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 (77) to become

$$g_L(x) = \begin{cases} (x - z_l)^2 & \text{for } l \leq x \leq z_l \\ 0 & \text{for } z_l \leq x \leq z_u \\ (x - z_u)^2 & \text{for } z_u \leq x \leq u \end{cases} \quad (83)$$

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 $\mathcal{D}_L^\circ(b) = \|\hat{b} - b\|_{Q_{bb}}^2 + \sum_{i=1}^n g_{i,L}(\hat{a}_i(b))/\sigma_i^2$. It is continuous differentiable, convex and lower bounding $\mathcal{D}^\circ(b)$ on the interval $b_L \leq b \leq b_U$. Hence, it is now in the form that the PGD-method can be applied to it.

7.3 Projected gradient descent (PGD)

The PGD-method is designed to solve a constrained minimization problem

$$\min_{b \in \mathcal{B}} f(b) \quad (84)$$

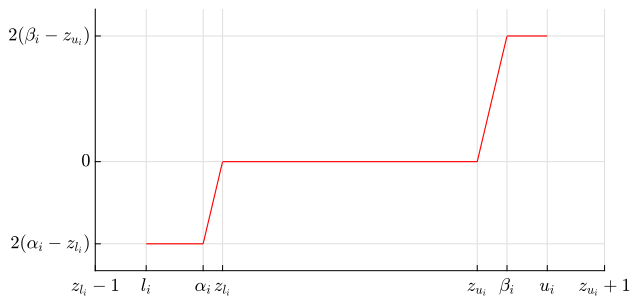


Fig. 10 The function $g'_{i,L}(x) = \frac{dg_{i,L}}{dx}(x)$ (cf. 91) on the interval $[l_i, u_i]$ for $l_i \in [z_{l_i} - 1, z_{l_i} - \frac{1}{2}]$, $u_i \in [z_{u_i} + \frac{1}{2}, z_{u_i} + 1]$

of which the objective function $f : \mathbb{R}^p \rightarrow \mathbb{R}$ 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 $\mathcal{B} = \{b \in \mathbb{R}^p \mid b_L \leq b \leq b_U\}$.

The PGD algorithmic steps for solving (84) are:

1. Initialize: Start with a feasible solution, $b^0 \in \mathcal{B}$ and then loop for $k = 0, \dots$ until stop criterium:
2. Gradient descent step: Compute stepsize μ_k and gradient descent

$$\hat{b}^k = b^k - \mu_k \nabla f(b^k) \quad (85)$$

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,

$$b^{k+1} = P_{\mathcal{B}}(\hat{b}^k) \text{ with } P_{\mathcal{B}}(y) = \arg \min_{b \in \mathcal{B}} \|y - b\|^2 \quad (86)$$

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

$$b^{k+1} = \arg \min_{b \in \mathcal{B}} F_k(b) \quad (87)$$

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 (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 (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 (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 $\mathcal{B} = \{b \in \mathbb{R}^p \mid b_L \leq b \leq b_U\}$, 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

$$\begin{aligned} \min_{b_L \leq b \leq b_U} \|y - b\|^2 &= \min_{b_L \leq b \leq b_U} \sum_{\alpha=1}^p (y_{\alpha} - b_{\alpha})^2 \\ &= \sum_{\alpha=1}^p \min_{b_{\alpha,L} \leq b_{\alpha} \leq b_{\alpha,U}} (y_{\alpha} - b_{\alpha})^2 \\ &= \sum_{\alpha=1}^p (y_{\alpha} - \text{median}(b_{\alpha,L}, y_{\alpha}, b_{\alpha,U}))^2 \end{aligned} \quad (88)$$

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,

$$P_{\mathcal{B}}(y)_{\alpha} = \begin{cases} b_{\alpha,L} & \text{if } y_{\alpha} \leq b_{\alpha,L} \\ y_{\alpha} & \text{if } b_{\alpha,L} \leq y_{\alpha} \leq b_{\alpha,U} \\ b_{\alpha,U} & \text{if } b_{\alpha,U} \leq y_{\alpha} \end{cases} \quad (89)$$

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 $\mathcal{D}_L^{\circ}(b) = \|\hat{b} - b\|_{Q_{\hat{b}\hat{b}}}^2 + \mathcal{G}_L(b)$ is given as

$$\begin{aligned} \nabla \mathcal{D}_L(b) &= 2Q_{\hat{b}\hat{b}}^{-1}(b - \hat{b}) + \nabla \mathcal{G}_L(b) \\ &= 2Q_{\hat{b}\hat{b}}^{-1}(b - \hat{b}) + \sum_{i=1}^n \sigma_i^{-2} g'_{i,L}(\hat{a}_i(b)) m_i \\ &= 2Q_{\hat{b}\hat{b}}^{-1} \left[b - \hat{b} + Q_{\hat{a}\hat{a}} Q_{\hat{a}\hat{a}}^{-1} s(b) \right] \end{aligned} \quad (90)$$

with $s(b) = \frac{1}{2} [g'_{1,L}(\hat{a}_1(b)), \dots, g'_{n,L}(\hat{a}_n(b))]^T$ 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 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 $z_{u_i} \geq z_{l_i}$,

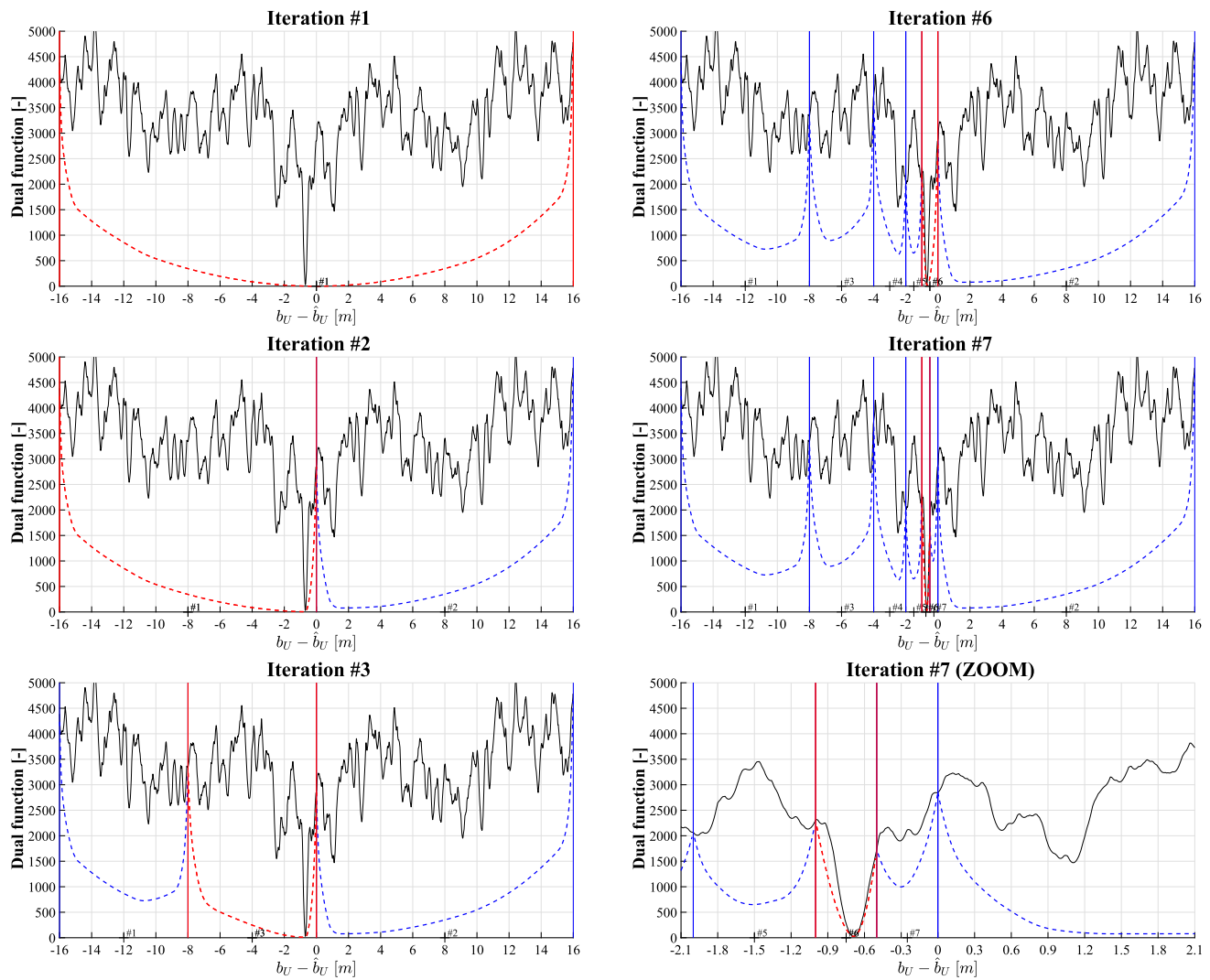


Fig. 11 Multimodal dual-function of vertical positioning mixed-integer GNSS model, with its per iteration constructed convex lower bounding functions (red and blue) over intervals that get split for the red functions

(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

the applicable derivative $g'_{i,L}(x)$ follows from (80), (81) and (82) as

$$g'_{i,L}(x) = \begin{cases} 2(\alpha_i - z_{l_i}) & \text{for } l_i \leq x \leq \alpha_i \\ 2(x - z_{l_i}) & \text{for } \alpha_i \leq x \leq z_{l_i} \\ 0 & \text{for } z_{l_i} \leq x \leq z_{u_i} \\ 2(x - z_{u_i}) & \text{for } z_{u_i} \leq x \leq \beta_i \\ 2(\beta_i - z_{u_i}) & \text{for } \beta_i \leq x \leq u_i \end{cases} \quad (91)$$

with

$$\begin{aligned} \alpha_i &= l_i + \sqrt{1 - 2(l_i - z_{l_i} + 1)} \\ \beta_i &= u_i - \sqrt{1 - 2(z_{u_i} + 1 - u_i)} \end{aligned} \quad (92)$$

The behaviour of $g'_{i,L}(x)$ for $x \in [l_i, u_i]$ is illustrated in Fig. 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 Fig. 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. 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: $D(p) = \sigma_p^2 D^T Q D$ and $D(\phi) = \sigma_\phi^2 D^T Q D$, with

$Q = \text{diag}[\sin^{-1}(e_{l_1}), \dots, \sin^{-1}(e_{l_{m+1}})]$, where e_{l_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 (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 Fig. 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 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 Fig. 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 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. \square

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 Fig. 12, 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 colour 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 apparent, 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. \square

8 Constrained and partitioned dual

In this contribution we restricted our study of the dual problem to the mixed-integer model (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

$$E(y) = Aa + Bb, \quad a \in \mathbb{Z}^n, \quad b \in \mathcal{C} \subset \mathbb{R}^p \quad (93)$$

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 (6),

$$\check{a} = \arg \min_{a \in \mathbb{Z}^n} F(a, \check{b}(a)), \quad \check{b}(a) = \arg \min_{b \in \mathcal{C}} F(a, b) \quad (94)$$

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).

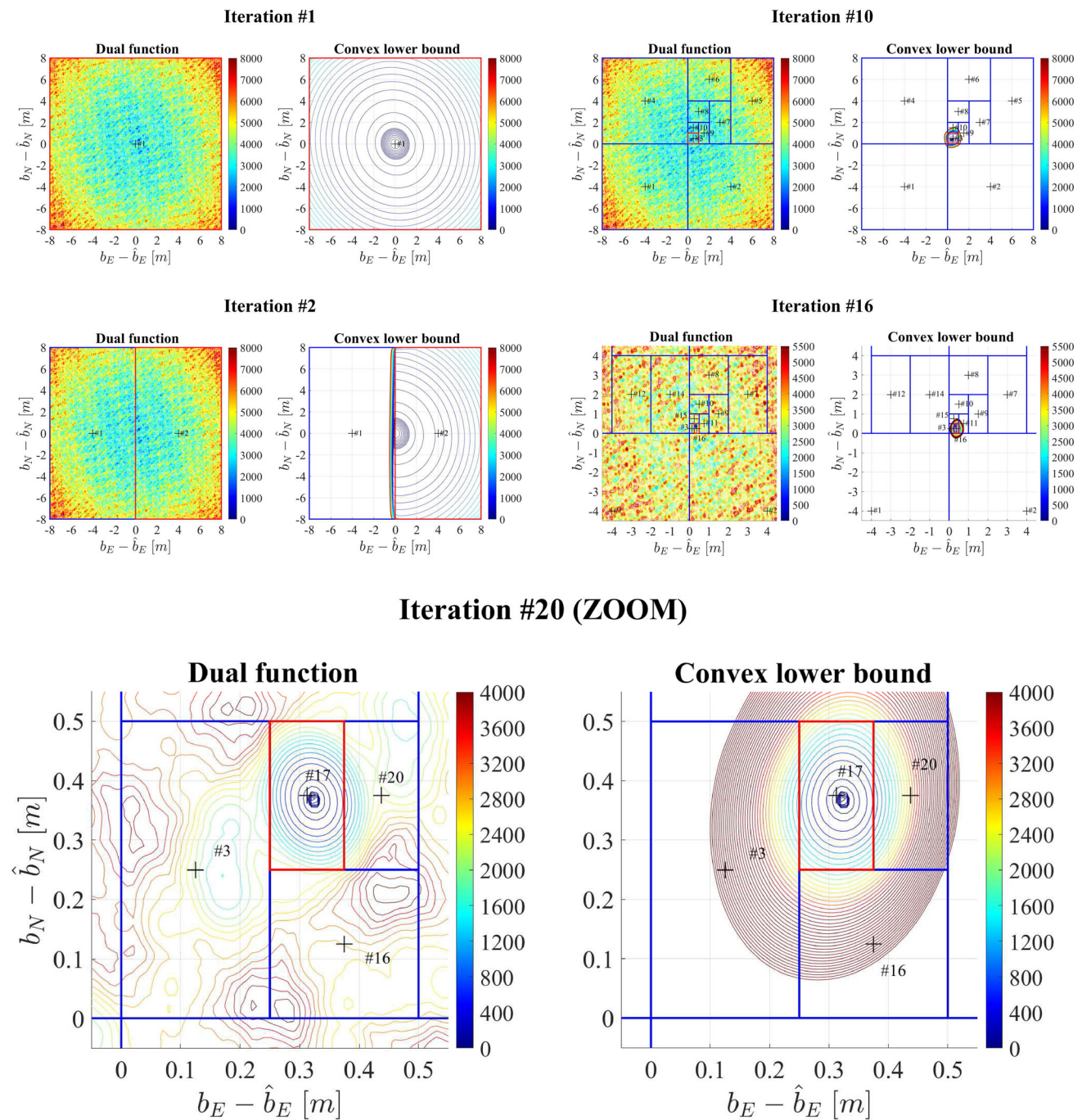


Fig. 12 Multimodal dual-function of horizontal positioning mixed-integer GNSS model, with the contour lines of its per iteration constructed convex lower bounding functions over the rectangles that

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

The dual formulation to (94) is

$$\check{b} = \arg \min_{b \in \mathcal{C}} F(\check{a}(b), b), \quad \check{a}(b) = \arg \min_{a \in \mathbb{Z}^n} F(a, b) \quad (95)$$

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 Sect. 3.3, that for the typical unconstrained mixed-integer GNSS model, one cannot 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 to 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 (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 (95) we recognize the dual function as $\mathcal{D}(b) = F(\check{a}(b), b)$, thus giving

$$\check{b} = \arg \min_{b \in \mathcal{C}} \mathcal{D}(b) \quad (96)$$

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 (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 (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

$$F(a, b_1, b_2) = \|\hat{b}_1 - b_1\|_{Q_{\hat{b}_1 \hat{b}_1}}^2 + \|\hat{a}(b_1) - a\|_{Q_{\hat{a}(b_1) \hat{a}(b_1)}}^2 + \|\hat{b}_2(a, b_1) - b_2\|_{Q_{\hat{b}_2(a, b_1) \hat{b}_2(a, b_1)}}^2$$

By now defining the partitioned dual $\mathcal{PD}(b_1)$ as

$$\mathcal{PD}(b_1) = \|\hat{b}_1 - b_1\|_{Q_{\hat{b}_1 \hat{b}_1}}^2 + \|\hat{a}(b_1) - \check{a}(b_1)\|_{Q_{\hat{a}(b_1) \hat{a}(b_1)}}^2 \quad (97)$$

with $\check{a}(b_1) = \arg \min_{a \in \mathbb{Z}^n} \|\hat{a}(b_1) - a\|_{Q_{\hat{a}(b_1) \hat{a}(b_1)}}^2$, the solution for a and b becomes

$$\begin{cases} \check{b}_1 = \arg \min_{b_1 \in \mathbb{R}^{p_1}} \mathcal{PD}(b_1) \\ \check{a} = \check{a}(\check{b}_1) \\ \check{b}_2 = \hat{b}_2(\check{a}, \check{b}_1) \end{cases} \quad (98)$$

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.

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 \check{a} is computed first, followed by the computation of the ambiguity-resolved baseline estimator $\check{b} = \hat{b}(\check{a})$. In the dual formulation, the order is reversed, giving first \check{b} , followed by the ambiguity estimator $\check{a} = \check{a}(\check{b})$. 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 multi-frequency 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 $\check{a}(b)$ 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.

10 Appendix

Proof of Lemma 3: The equality of (23) follows from substituting $-Q_{\hat{a}\hat{b}}Q_{\hat{b}\hat{b}}^{-1} = (A^T Q_{yy}^{-1} A)^{-1} A^T Q_{yy}^{-1} B$ and $Q_{\hat{a}(b)\hat{a}(b)} = (A^T Q_{yy}^{-1} A)^{-1}$ in $\|\Delta a\|_{Q_{\hat{a}(b)\hat{a}(b)}}^2$. The inequality of (23) follows from $Q_{\hat{b}(a)\hat{b}(a)} = (B^T Q_{yy}^{-1} B)^{-1}$ and recognizing that the norm of a projected vector is never larger than the norm of the vector itself. \square

Proof of Lemma 4: The quadratic identity follows as

$$\begin{aligned} & \|\hat{b} - b\|_{Q_{\hat{b}\hat{b}}}^2 + \|\hat{a}(b) - a\|_{Q_{\hat{a}(b)\hat{a}(b)}}^2 \\ &= \begin{bmatrix} \hat{a}(b) - a \\ \hat{b} - b \end{bmatrix}^T \begin{bmatrix} Q_{\hat{a}(b)\hat{a}(b)} & 0 \\ 0 & Q_{\hat{b}\hat{b}} \end{bmatrix}^{-1} \begin{bmatrix} \hat{a}(b) - a \\ \hat{b} - b \end{bmatrix} \\ &= \begin{bmatrix} \hat{a} - a \\ \hat{b} - b \end{bmatrix}^T \begin{bmatrix} Q_{\hat{a}(b)\hat{a}(b)} + Q_{\hat{a}\hat{b}}Q_{\hat{b}\hat{b}}^{-1}Q_{\hat{b}\hat{a}} & Q_{\hat{a}\hat{b}} \\ Q_{\hat{b}\hat{a}} & Q_{\hat{b}\hat{b}} \end{bmatrix}^{-1} \begin{bmatrix} \hat{a} - a \\ \hat{b} - b \end{bmatrix} \\ &= \begin{bmatrix} \hat{a} - a \\ \hat{b}^\circ(a) - b \end{bmatrix}^T \begin{bmatrix} Q_{\hat{a}\hat{a}} & 0 \\ 0 & Q_{\hat{b}\hat{b}} - Q_{\hat{b}\hat{a}}Q_{\hat{a}\hat{a}}^{-1}Q_{\hat{a}\hat{b}} \end{bmatrix}^{-1} \begin{bmatrix} \hat{a} - a \\ \hat{b}^\circ(a) - b \end{bmatrix} \\ &= \|\hat{a} - a\|_{Q_{\hat{a}\hat{a}}}^2 + \|\hat{b}^\circ(a) - b\|_{Q_{\hat{b}(a)\hat{b}(a)}}^2 \end{aligned} \quad (99)$$

\square

Proof of Theorem 3: Using the quadratic identity (27), we may write $\mathcal{D}^\circ(b)$ as

$$\begin{aligned} \mathcal{D}^\circ(b) &= \|\hat{b} - b\|_{Q_{\hat{b}\hat{b}}}^2 + \min_{a \in \mathbb{Z}^n} \|\hat{a}(b) - a\|_{Q_{\hat{a}(b)\hat{a}(b)}}^2 \\ &= \min_{a \in \mathbb{Z}^n} \left(\|\hat{b} - b\|_{Q_{\hat{b}\hat{b}}}^2 + \|\hat{a}(b) - a\|_{Q_{\hat{a}(b)\hat{a}(b)}}^2 \right) \\ &= \min_{a \in \mathbb{Z}^n} \left(\|\hat{a} - a\|_{Q_{\hat{a}\hat{a}}}^2 + \|\hat{b}^\circ(a) - b\|_{Q_{\hat{b}(a)\hat{b}(a)}}^2 \right) \end{aligned} \quad (100)$$

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 $\check{a}^\circ = \arg \min_{a \in \mathbb{Z}^n} \|\hat{a} - a\|_{Q_{\hat{a}\hat{a}}}^2$. What remains to be shown is that $\check{a}^\circ = \check{a}^\circ(\check{b}^\circ)$. Starting from the definition $\check{a}^\circ(b) =$

$\arg \min_{a \in \mathbb{Z}^n} \|\hat{a}(b) - a\|_{Q_{\hat{a}(b)\hat{a}(b)}}^2$, we may write with the aid of the quadratic identity (27),

$$\check{a}^\circ(b) = \arg \min_{a \in \mathbb{Z}^n} \left(\|\hat{a} - a\|_{Q_{\hat{a}\hat{a}}}^2 + \|\hat{b}^\circ(a) - b\|_{Q_{\hat{b}(a)\hat{b}(a)}}^2 \right) \quad (101)$$

from which it follows, since $\check{b}^\circ = \hat{b}^\circ(\check{a}^\circ)$, that $\check{a}^\circ(\check{b}^\circ) = \arg \min_{a \in \mathbb{Z}^n} \|\hat{a} - a\|_{Q_{\hat{a}\hat{a}}}^2 = \check{a}^\circ$. \square

Proof of Theorem 4: As $Q_{\hat{a}\hat{a}}^\circ$ is not the variance matrix of \hat{a} , inequality (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

$$R_{\hat{a}\hat{b}}^T R_{\hat{a}\hat{b}} = Q_{\hat{b}\hat{a}} \left[Q_{\hat{a}\hat{a}}^{\circ-1} Q_{\hat{a}\hat{a}} Q_{\hat{a}\hat{a}}^{\circ-1} - 2Q_{\hat{a}\hat{a}}^{\circ-1} + Q_{\hat{a}\hat{a}} \right] Q_{\hat{a}\hat{b}} \quad (102)$$

which proves the equality of (32b). As (102) is positive semi-definite, the inequality of (32b) follows. \square

Proof of Theorem 5 (Success-rate lower-bounds) (i) The proof of (34) follows from combining an inequality-theorem of (Anderson 1996) with the ILS-theorem of (Teunissen 1999a). Let $\hat{a} \sim \mathcal{N}_n(a, Q_{\hat{a}\hat{a}})$, $\hat{a}^\circ \sim \mathcal{N}_n(a, Q_{\hat{a}\hat{a}}^\circ)$ and $Q_{\hat{a}\hat{a}}^\circ \geq Q_{\hat{a}\hat{a}}$. Then it follows from Anderson's theorem that $P[\hat{a}^\circ \in \mathcal{P}_a^\circ] \leq P[\hat{a} \in \mathcal{P}_a^\circ]$, as the subset $\mathcal{P}_a^\circ \subset \mathbb{R}^n$ is convex and symmetric about a . As $P[\hat{a}^\circ \in \mathcal{P}_a^\circ]$ 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}° 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 (34), see (Teunissen 1998b).

(ii) To prove the given lower-bound (35) of

$$P[\check{a}^\circ = a] = \int_{\mathcal{P}_a^\circ} f_{\hat{a}}(\alpha) d\alpha \quad (103)$$

we will work with two ellipsoids, one being a subset of the other, while both are subsets of the pull-in region \mathcal{P}_a° . First we describe the pull-in region in a way that facilitates the comparison with the planes-of-support formulation of ellipsoids:

$$\begin{aligned} \mathcal{P}_a^\circ &= a + \mathcal{P}_0^\circ \\ &= a + \left\{ x \in \mathbb{R}^n \mid x^T Q_{\hat{a}\hat{a}}^{\circ-1} x \leq \|x - z\|_{Q_{\hat{a}\hat{a}}^\circ}^2, \forall z \in \mathbb{Z}^n \right\} \\ &= a + \left\{ x \in \mathbb{R}^n \mid (z^T Q_{\hat{a}\hat{a}}^{\circ-1} x)^2 \leq \frac{1}{4} \|z\|_{Q_{\hat{a}\hat{a}}^\circ}^4, \forall z \in \mathbb{Z}^n \right\} \end{aligned} \quad (104)$$

Now we construct the first ellipsoid residing in \mathcal{P}_0° . We have

$$\begin{aligned} E_0^\circ &= \left\{ x \in \mathbb{R}^n \mid x^T Q_{\hat{a}\hat{a}}^{\circ-1} x \leq \chi^2 \right\} \\ &\stackrel{a}{=} \left\{ x \in \mathbb{R}^n \mid (g^T x)^2 \leq \chi^2 g^T Q_{\hat{a}\hat{a}}^\circ g, \forall g \in \mathbb{R}^n \right\} \\ &\stackrel{b}{=} \left\{ x \in \mathbb{R}^n \mid (f^T Q_{\hat{a}\hat{a}}^{\circ-1} x)^2 \leq \chi^2 f^T Q_{\hat{a}\hat{a}}^{\circ-1} f, \forall f \in \mathbb{R}^n \right\} \\ &\stackrel{c}{\subset} \left\{ x \in \mathbb{R}^n \mid (z^T Q_{\hat{a}\hat{a}}^{\circ-1} x)^2 \leq \chi^2 \|z\|_{Q_{\hat{a}\hat{a}}^\circ}^2, \forall z \in \mathbb{Z}^n \right\} \\ &\stackrel{d}{\subset} \left\{ x \in \mathbb{R}^n \mid (z^T Q_{\hat{a}\hat{a}}^{\circ-1} x)^2 \leq \chi^2 \frac{\|z\|_{Q_{\hat{a}\hat{a}}^\circ}^4}{\min_{z \in \mathbb{Z}^n \setminus \{0\}} \|z\|_{Q_{\hat{a}\hat{a}}^\circ}^2}, \forall z \in \mathbb{Z}^n \right\} \end{aligned} \quad (105)$$

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 (105) with (104) that

$$E_a^\circ = a + E_0^\circ \subset \mathcal{P}_a^\circ \text{ if } \chi^2 = \frac{1}{4} \min_{z \in \mathbb{Z}^n \setminus \{0\}} \|z\|_{Q_{\hat{a}\hat{a}}^\circ}^2 \quad (106)$$

We now construct our second ellipsoid $E_a \subset E_a^\circ$ such that the probability $P[\hat{a} \in E_a]$ is easy to compute. The ellipsoid is chosen as

$$\begin{aligned} E_a &= \left\{ x \in \mathbb{R}^n \mid (x - a)^T Q_{\hat{a}\hat{a}}^{-1} (x - a) \leq r^2 \right\} \\ &= \left\{ x = a + Q_{\hat{a}\hat{a}}^{1/2} v, v \in \mathbb{R}^n \mid v^T v \leq r^2 \right\} \end{aligned} \quad (107)$$

If we use the same transformation also for E_a° , we may write

$$E_a^\circ = \left\{ x = a + Q_{\hat{a}\hat{a}}^{1/2} v, v \in \mathbb{R}^n \mid v^T [Q_{\hat{a}\hat{a}}^{1/2} Q_{\hat{a}\hat{a}}^{\circ-1} Q_{\hat{a}\hat{a}}^{1/2}] v \leq \chi^2 \right\} \quad (108)$$

By comparing (108) with (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

$$E_a \subset E_a^\circ \text{ if } r^2 = \lambda_{\min} \times \chi^2 \quad (109)$$

where λ_{\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

$$\lambda_{\min} = \min_{x \in \mathbb{R}^n} \frac{x^T Q_{\hat{a}\hat{a}}^\circ x}{x^T Q_{\hat{a}\hat{a}} x} \quad (110)$$

Since $E_a \subset E_a^\circ \subset \mathcal{P}_a^\circ$ (cf. 106 and 109), it follows that

$$P[\check{a}^\circ = a] = P[\hat{a} \in \mathcal{P}_a^\circ] \geq P[\hat{a} \in E_a] = P[\chi^2(0, n) \leq r^2]$$

which concludes the proof. \square

Proof of Theorem 6: Recall that any admissible integer estimator $\mathcal{I}^\bullet(x)$ can be written as (Teunissen 2000b),

$$\mathcal{I}^\bullet(x) = \sum_{z \in \mathbb{Z}^n} z \iota_z^\bullet(x) \quad (111)$$

where $\iota_z^\bullet(x)$ is the indicator function of its pull-in region $I_z^\bullet = \{x \in \mathbb{R}^n \mid z = \mathcal{I}^\bullet(x)\}$. With (111), we can rewrite the dual function as

$$\begin{aligned} \mathcal{D}^\bullet(b) &= \|\hat{b} - b\|_{Q_{\hat{b}\hat{b}}}^2 + \|\hat{a}(b) - \mathcal{I}^\bullet(\hat{a}(b))\|_{Q_{\hat{a}(b)\hat{a}(b)}}^2 \\ &= \|\hat{b} - b\|_{Q_{\hat{b}\hat{b}}}^2 + \sum_{z \in \mathbb{Z}^n} \|\hat{a}(b) - z\|_{Q_{\hat{a}(b)\hat{a}(b)}}^2 \iota_z^\bullet(\hat{a}(b)) \\ &= \sum_{z \in \mathbb{Z}^n} \left(\|\hat{b} - b\|_{Q_{\hat{b}\hat{b}}}^2 + \|\hat{a}(b) - z\|_{Q_{\hat{a}(b)\hat{a}(b)}}^2 \right) \iota_z^\bullet(\hat{a}(b)) \\ &= \sum_{z \in \mathbb{Z}^n} \left(\|\hat{a} - z\|_{Q_{\hat{a}\hat{a}}}^2 + \|\hat{b}(z) - b\|_{Q_{\hat{b}(a)\hat{b}(a)}}^2 \right) \iota_z^\bullet(\hat{a}(b)) \end{aligned}$$

This shows that

$$\mathcal{D}^\bullet(b) = \|\hat{a} - u\|_{Q_{\hat{a}\hat{a}}}^2 + \|\hat{b}(u) - b\|_{Q_{\hat{b}(a)\hat{b}(a)}}^2 \text{ if } \hat{a}(b) \in I_u^\bullet$$

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

$$\hat{b}(u) = \arg \min_{\iota_u^\bullet(\hat{a}(b))=1} \mathcal{D}^\bullet(b), \quad \mathcal{D}^\bullet(\hat{b}(u)) = \|\hat{a} - u\|_{Q_{\hat{a}\hat{a}}}^2 \quad (112)$$

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,

$$\Omega_{\hat{a}}^\bullet = \{z \in \mathbb{Z}^n \mid z = \mathcal{I}^\bullet(\hat{a}(b)), \forall b \in \mathbb{R}^p\} \subset \mathbb{Z}^n \quad (113)$$

Hence, the global minimum of $\mathcal{D}^\bullet(b)$ is given as

$$\min_{b \in \mathbb{R}^p} \mathcal{D}^\bullet(b) = \min_{z \in \Omega_{\hat{a}}^\bullet} \mathcal{D}^\bullet(\hat{b}(z)) = \min_{z \in \Omega_{\hat{a}}^\bullet} \|\hat{a} - z\|_{Q_{\hat{a}\hat{a}}}^2 \quad (114)$$

and its corresponding minimizer as

$$\begin{aligned} \check{b}^\bullet &= \arg \min_{b \in \mathbb{R}^p} \mathcal{D}^\bullet(b) = \hat{b}(\check{a}^\bullet), \text{ with} \\ \check{a}^\bullet &= \arg \min_{z \in \Omega_{\hat{a}}^\bullet} \|\hat{a} - z\|_{Q_{\hat{a}\hat{a}}}^2 \end{aligned} \quad (115)$$

\square

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 $\arg \min_{u \in \Omega_x^\bullet} \|x - u\|_{Q_{\hat{a}\hat{a}}}^2$ 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 $\mathcal{P}_z^\bullet = \mathcal{P}_0^\bullet + z$. 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

$$\begin{aligned} \bar{\Omega}_0^\bullet + z &= \{x + z \in \mathbb{R}^n \mid 0 = \mathcal{I}^\bullet(x + M\beta), \exists \beta \in \mathbb{R}^p\} \\ &= \{y \in \mathbb{R}^n \mid 0 = \mathcal{I}^\bullet(y - z + M\beta), \exists \beta \in \mathbb{R}^p\} \\ &= \{y \in \mathbb{R}^n \mid z = \mathcal{I}^\bullet(y + M\beta), \exists \beta \in \mathbb{R}^p\} \\ &= \bar{\Omega}_z^\bullet, \quad \forall z \in \mathbb{Z}^n \end{aligned} \quad (116)$$

where use was made of the property $\mathcal{I}^\bullet(x + z) = \mathcal{I}^\bullet(x) + z$. Similarly, we have

$$\begin{aligned} \Phi_0^\bullet + z &= \{u + z \in \mathbb{Z}^n \mid u = \mathcal{I}^\bullet(x), \forall x \in \bar{\Omega}_0^\bullet\} \\ &= \{v \in \mathbb{Z}^n \mid v = \mathcal{I}^\bullet(x) + z, \forall x \in \bar{\Omega}_0^\bullet\} \\ &= \{v \in \mathbb{Z}^n \mid v = \mathcal{I}^\bullet(x + z), \forall x \in \bar{\Omega}_0^\bullet\} \\ &= \{v \in \mathbb{Z}^n \mid v = \mathcal{I}^\bullet(y), \forall y \in \bar{\Omega}_0^\bullet + z\} \\ &= \{v \in \mathbb{Z}^n \mid v = \mathcal{I}^\bullet(y), \forall y \in \bar{\Omega}_z^\bullet\} \\ &= \Phi_z^\bullet, \quad \forall z \in \mathbb{Z}^n \end{aligned} \quad (117)$$

where use was made of $\bar{\Omega}_z^\bullet = \bar{\Omega}_0^\bullet + z$. Finally, we have

$$\begin{aligned} \mathcal{P}_0^\bullet + z &= \{x + z \in \mathbb{R}^n \mid \|x\|_{Q_{aa}}^2 \leq \|x - u\|_{Q_{aa}}^2, \forall u \in \Phi_0\} \\ &= \{y \in \mathbb{R}^n \mid \|y - z\|_{Q_{aa}}^2 \leq \|y - (u + z)\|_{Q_{aa}}^2, \forall u \in \Phi_0\} \\ &= \{y \in \mathbb{R}^n \mid \|y - z\|_{Q_{aa}}^2 \leq \|y - v\|_{Q_{aa}}^2, \forall v \in \Phi_0 + z\} \\ &= \{y \in \mathbb{R}^n \mid \|y - z\|_{Q_{aa}}^2 \leq \|y - v\|_{Q_{aa}}^2, \forall v \in \Phi_z\} \\ &= \mathcal{P}_z^\bullet, \quad \forall z \in \mathbb{Z}^n \end{aligned} \quad (118)$$

where use was made of $\Phi_z^\bullet = \Phi_0^\bullet + z$. \square

Proof of Theorem 7: As $P[\hat{a}^\bullet = a] = P[\hat{a} \in \mathcal{P}_a^\bullet] = \int_{\mathcal{P}_a^\bullet} f_{\hat{a}}(\alpha) d\alpha$, the proof will be based on finding an ellipsoidal region $E_a = \{x \in \mathbb{R}^n \mid (x - a)^T Q_{aa}^{-1} (x - a) \leq r^2\}$ that resides in \mathcal{P}_a^\bullet . Once such region is found, we have $E_a \subset \mathcal{P}_a^\bullet$, and thus the lower-bound $P[\hat{a} \in E_a] \leq P[\hat{a} \in \mathcal{P}_a^\bullet]$, with $P[\hat{a} \in E_a] = P[\chi^2(0, n) \leq r^2]$. The challenge is therefore to find the proper value for r^2 .

First we express the pull-in region $\mathcal{P}_a^\bullet = a + \mathcal{P}_0^\bullet$ in a more amenable form. We have

$$\begin{aligned} \mathcal{P}_0^\bullet &= \{x \in \mathbb{R}^n \mid \|x\|_{Q_{aa}}^2 \leq \|x - \mathcal{I}^\bullet(y)\|_{Q_{aa}}^2, \forall y \in \bar{\Omega}_0^\bullet\} \\ &= \{x \in \mathbb{R}^n \mid \mathcal{I}^\bullet(y)^T Q_{aa}^{-1} x \leq \frac{1}{2} \|\mathcal{I}^\bullet(y)\|_{Q_{aa}}^2, \forall y \in \bar{\Omega}_0^\bullet\} \end{aligned} \quad (119)$$

We now show how r^2 can be chosen such that $E_0 \subset \mathcal{P}_0^\bullet$. We have, with $q^2 = r^2 / (\min_{z \in \mathbb{Z}^n \setminus \{0\}} \|z\|_{Q_{aa}}^2)$,

$$\begin{aligned} E_0 &= \{x \in \mathbb{R}^n \mid x^T Q_{aa}^{-1} x \leq r^2\} \\ &\stackrel{(i)}{=} \{x \in \mathbb{R}^n \mid [g^T x]^2 \leq r^2 g^T Q_{aa} g, \forall g \in \mathbb{R}^n\} \\ &\stackrel{(ii)}{=} \{x \in \mathbb{R}^n \mid [f^T Q_{aa}^{-1} x]^2 \leq r^2 f^T Q_{aa}^{-1} f, \forall f \in \mathbb{R}^n\} \\ &\stackrel{(iii)}{\subset} \{x \in \mathbb{R}^n \mid [z^T Q_{aa}^{-1} x]^2 \leq r^2 \|z\|_{Q_{aa}}^2, \forall z \in \mathbb{Z}^n\} \\ &\stackrel{(iv)}{\subset} \{x \in \mathbb{R}^n \mid [z^T Q_{aa}^{-1} x]^2 \leq q^2 \|z\|_{Q_{aa}}^4, \forall z \in \mathbb{Z}^n\} \\ &\stackrel{(v)}{\subset} \{x \in \mathbb{R}^n \mid [\mathcal{I}^\bullet(y)^T Q_{aa}^{-1} x]^2 \leq q^2 \|\mathcal{I}^\bullet(y)\|_{Q_{aa}}^4, \forall y \in \bar{\Omega}_0^\bullet\} \\ &\stackrel{(vi)}{\subset} \left\{x \in \mathbb{R}^n \mid \mathcal{I}^\bullet(y)^T Q_{aa}^{-1} x \leq \left[q^2 \|\mathcal{I}^\bullet(y)\|_{Q_{aa}}^4\right]^{\frac{1}{2}}, \forall y \in \bar{\Omega}_0^\bullet\right\} \\ &\stackrel{(119)}{=} \mathcal{P}_0^\bullet \text{ if } r^2 = \frac{1}{4} \min_{z \in \mathbb{Z}^n \setminus \{0\}} \|z\|_{Q_{aa}}^2 \end{aligned}$$

In (i) we used the *planes-of-support* formulation of the ellipsoid (Teunissen 1995), while (ii) follows from using the one-to-one transformation $g = Q_{aa}^{-1} f$. 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 (119) shows that \mathcal{P}_0^\bullet is obtained as stated. An example of the lower bounding ellipse E_0 is shown in Figure 4 for $n = 2$ and $p = 1$. \square

Projection Lemma (projection of cube and rectangle):

(a) Largest and smallest projection of cube on a line:

$$\begin{aligned} \max_{-e \leq x \leq e} a^T x &= + \sum_{\alpha=1}^p |a_\alpha| \\ \min_{-e \leq x \leq e} a^T x &= - \sum_{\alpha=1}^p |a_\alpha| \end{aligned} \quad (120)$$

with $a = (a_1, \dots, a_p)^T$ and $e = (1, \dots, 1)^T$.

(b) Largest and smallest projection of rectangle on a line:

$$\begin{aligned} \max_{l \leq x \leq u} a^T x &= \sum_{\alpha=1}^p + \frac{1}{2} (u_\alpha - l_\alpha) |a_\alpha| + \frac{1}{2} (u_\alpha + l_\alpha) a_\alpha \\ \min_{l \leq x \leq u} a^T x &= \sum_{\alpha=1}^p - \frac{1}{2} (u_\alpha - l_\alpha) |a_\alpha| + \frac{1}{2} (u_\alpha + l_\alpha) a_\alpha \end{aligned} \quad (121)$$

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 $\tilde{U} = \text{diag}(\tilde{u}_1, \dots, \tilde{u}_p)$. \square

Author contributions PT devised the statistical and algorithmic concepts of the dual mixed-integer problem, developed the theory and wrote the paper, LM checked the theory and text, developed the integrated algorithm and computed the numerical/graphical results.

Data availability All data generated or analysed during this study are included in this contribution.

Declarations

Conflict of interest The authors declare that they have no Conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

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. *J Robust Nonlin Control* 1(4):295–317
- Bertsekas D (1999) *Nonlinear programming*. Athena Scientific, 2nd Edn
- Giorgi G, Teunissen PJG, Verhagen S, Buist PJ (2010) Testing a new multivariate GNSS carrier phase attitude determination method for remote sensing platforms. *Adv Space Res* 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*, 2nd edn. Kluwer Academic Publishers, Dordrecht
- Khodabandeh A (2022) Bias-bounded estimation of ambiguity: a method for radio interferometric positioning. *IEEE Trans Signal Process* 70:3042–3057. <https://doi.org/10.1109/TSP.2022.3181344>
- Lawler EL, Wood DE (1966) Branch-And-Bound methods: a survey. *Oper Res* 14(4):699–719
- Leick A, Rapoport L, Tatarnikov D (2015) *GPS satellite surveying*, 4th edn. Wiley, Hoboken
- Nesterov Y (2018) *Lectures on convex optimization*. Springer optimization and its applications 137
- Nocedal J, Wright S (2006) *Numerical optimization*, 2nd edn. Springer, Berlin
- 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. *Found Trends Optimiz* 1(3):123–231
- Teunissen PJG (1990) Nonlinear least squares. *Manuscr Geodaet* 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. *Artif Satell* 30(4):171–179
- Teunissen PJG (2000b) The success rate and precision of GPS ambiguities. *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. *Artif Satell* 41(1):19–22
- Teunissen PJG (2010) Integer least squares theory for the GNSS compass. *J Geodesy*. <https://doi.org/10.1007/s00190-010-0380-8>
- Teunissen PJG (2019) A new GLONASS FDMA model. *GPS Solut* 23:100. <https://doi.org/10.1007/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. *J Geodesy* 95(99):1–14
- Verhagen S, Li B, Teunissen PJG (2013) Ps-LAMBDA: ambiguity success rate evaluation software for interferometric applications. *Comput Geosci* 54:361–376
- Zhigljavsky A (1991) *Theory of global random search*. Mathematics and its applications. Kluwer Academic Publishers, Dordrecht