



Signal Processing
Systems
Mekelweg 4,
2628 CD Delft
The Netherlands
<https://sps.ewi.tudelft.nl/>

SPS-2026-00

M.Sc. Thesis

Targetless and Bias-Aware Extrinsic Calibration of Underwater Vehicles

Aybars Manav B.Sc.

Targetless and Bias-Aware Extrinsic Calibration of Underwater Vehicles

THESIS

submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE

in

ELECTRICAL ENGINEERING

by

Aybars Manav B.Sc.
born in Ankara, Turkey

This work was performed in:

Signal Processing Systems Group
Department of Microelectronics
Faculty of Electrical Engineering, Mathematics and Computer Science
Delft University of Technology



Delft University of Technology

Copyright © 2026 Signal Processing Systems Group
All rights reserved.

DELFT UNIVERSITY OF TECHNOLOGY
DEPARTMENT OF
MICROELECTRONICS

The undersigned hereby certify that they have read and recommend to the Faculty of Electrical Engineering, Mathematics and Computer Science for acceptance a thesis entitled “**Targetless and Bias-Aware Extrinsic Calibration of Underwater Vehicles**” by **Aybars Manav B.Sc.** in partial fulfillment of the requirements for the degree of **Master of Science**.

Dated: 29/06/2026

Chairman:

Dr. G. Joseph

Advisor:

Ir. J. Quist

Committee Members:

Dr. M. Kok

Dr. R.T. Rajan

Abstract

Inspection-class autonomous underwater vehicles often operate without external positioning. Instead, they use onboard sensors to estimate their state. Because no single sensor is sufficient in isolation, they typically combine a camera, an inertial measurement unit (IMU), a Doppler velocity log (DVL), and a pressure sensor, fused within a simultaneous localization and mapping (SLAM) estimator. Coherent fusion requires extrinsic calibration so that all measurements are expressed in a common frame. Existing calibration methods address only part of the problem: they often calibrate subsets of sensors, require dedicated calibration procedures or targets, adopt continuous-time formulations not directly compatible with a discrete-time estimator, or neglect the impact of uncertainty in the motion quantities used during calibration. This thesis develops a targetless joint extrinsic calibration that recovers the full set of extrinsic parameters between the camera, IMU, DVL, and pressure sensor from onboard data. The method is formulated within a discrete-time SLAM estimator. To the best of our knowledge, no prior method estimates the pressure-sensor translation jointly with the rest of the suite. First, we show that the established baseline carries a structural bias in the IMU–DVL translation, present regardless of data quality. Our corrected formulation removes this bias, establishing a precondition for any joint calibration built on the baseline. Second, we identify a subtle bias that remains, an errors-in-variables effect from conditioning on a noisy motion measurement. This bias cannot be eliminated through tuning. We show that its severity can be predicted from the sensor data before the run, without ground truth. A further reformulation reduces it. Finally, evaluated on real underwater data, the calibration runs on demand during a mission and improves the IMU–DVL translation estimate over the baseline, in the order predicted by the analysis.

Acknowledgments

I would like to thank my supervisor, Dr. Geethu Joseph, for her support and guidance throughout this work. Our discussions pushed me to look more deeply into the research, and her help was invaluable in the writing of this thesis. I am especially grateful for her encouragement, which gave me confidence when I needed it the most.

I would like to thank Dr. Manon Kok for her time and valuable feedback during this period, particularly on the survey paper we worked on together.

I thank my advisor, Ir. Joris Quist and everyone at Lobster Robotics for the opportunity to take on this topic and for the support that made the work possible.

I gratefully acknowledge the Justus van Effen Foundation for the scholarship that funded my studies and made this degree possible.

To those closest to me, thank you for the emotional support that carried me through the more difficult parts of this period.

Finally, I want to thank my family, who have always looked out for my health and stood behind me at every step.

Aybars Manav B.Sc.
Delft, The Netherlands
29/06/2026

Contents

Abstract	v
Acknowledgments	vii
1 Introduction	1
1.1 Motivation: Multi-Sensor Fusion and Extrinsic Calibration	1
1.2 State of the Art in Extrinsic Calibration	2
1.3 Research Questions and Contributions	4
2 The Calibration Problem	7
2.1 The Extrinsic Calibration Problem	7
2.2 Measurement Models and Residuals	9
2.2.1 Shared Foundation: Body Motion Model	10
2.2.2 Pressure Sensor Model	11
2.2.3 IMU Model	12
2.2.4 DVL Model	14
2.3 Calibration Cost and Its Minimization	16
2.4 Summary	18
3 Baseline AQUA-SLAM Analysis	19
3.1 AQUA-SLAM’s Calibration Formulation	19
3.2 IMU–DVL Lever Arm Bias in AQUA-SLAM	21
3.3 Experimental Setup	24
3.3.1 Simulation Environment	25
3.3.2 Trajectories	25
3.3.3 Noise Grades	27
3.3.4 Monte Carlo Protocol	28
3.4 Empirical Validation	29
3.4.1 Protocol	30
3.4.2 Lever Arm Collapse	30
3.4.3 Secondary Effects and Co-located Control	31
3.5 Summary	32
4 Formulation A: Lever-Arm-Corrected Velocity Model	33
4.1 Calibration State and Velocity Supply	33
4.2 The Residuals under Formulation A	33
4.2.1 IMU Velocity and Translation Residuals	34
4.2.2 Residual Covariances	34
4.3 The Coarse-to-Fine Pipeline	35
4.4 Empirical Validation	37
4.4.1 Protocol	37
4.4.2 AQUA-SLAM versus Formulation A	38

4.4.3	Pressure-Sensor Calibration	39
4.5	Summary	40
5	Errors-in-Variables Bias in Formulation A	41
5.1	Bias Under a Noisy Regressor	41
5.2	The Errors-in-Variables Bias in Formulation A	42
5.3	The Eigenvalue Diagnostic	45
5.4	The Information Degradation of the IMU Residuals	46
5.5	Empirical Validation	48
5.5.1	Protocol	48
5.5.2	The Diagnostic across Trajectories and Grades	48
5.5.3	Verifying the EIV Mechanism	49
5.5.4	Magnitude: A Conservative Indicator	49
5.5.5	Computing the Diagnostic in Practice	51
5.6	Summary	53
6	Formulation B: Velocity State Augmentation	55
6.1	Calibration State and Velocity Supply	55
6.2	The DVL Velocity Residual and Its Covariance	55
6.3	The EIV Diagnostic under Formulation B	56
6.3.1	Per-Factor Contributions	56
6.3.2	Comparison with Formulation A	58
6.4	Empirical Comparison: Formulation A versus Formulation B	58
6.4.1	Diagnostic Improvement	59
6.4.2	EIV-Coupled Parameters Improve	59
6.4.3	EIV-Uncoupled Parameters	61
6.4.4	A Non-EIV Bias Remains	61
6.5	Summary	63
7	Real-World Calibration	65
7.1	Calibration Pipeline	65
7.1.1	Platform, Sensors, and Sequences	65
7.1.2	Camera Trajectory	65
7.1.3	Protocol	66
7.2	Calibration Window Selection	66
7.2.1	The DVL Error Velocity	66
7.2.2	Two Gates and the Healthy Span	67
7.3	Three-Way Formulation Comparison	69
7.4	Convergence and Run Time	70
7.4.1	Convergence Against Window Length	70
7.4.2	Run Time	71
7.5	Downstream SLAM Validation	72
7.5.1	The Corrected Runtime Residual	72
7.5.2	Calibration Under the Corrected Model	74
7.6	Discussion	76

8	Conclusion	77
8.1	Limitations and Future Work	78
A	DVL noise standard deviation from the error velocity	83
A.1	Beam geometry and the error velocity	83
A.2	Per-beam standard deviation	83
A.3	Body-frame covariance	84
B	An Expectation Identity for the Cross-Product Matrix	85
C	Proof of Lemma 5.1: Asymptotic Bias of the WLS Estimator under EIV	87
D	DVL calibration-window selection	89
D.1	Window statistics	89
D.2	Gate thresholds	89
D.3	Acceptance rule	90

List of Figures

2.1	Two of the five elements of $\mathcal{E}$ on the vehicle: the DVL-IMU rotation ${}_b\mathbf{R}_d$ and camera-pressure lever arm ${}_c\mathbf{p}_{p/c}$	8
2.2	The navigation frame c_0 and the gravity direction ${}_{c_0}\mathbf{g}$ expressed in it. The world vertical (dashed) does not align with the axes of c_0 , so ${}_{c_0}\mathbf{g}$ is unknown a priori.	8
2.3	The camera trajectory $\{{}_{c_0}\mathbf{R}_{c_i}, {}_{c_0}\mathbf{p}_{c_i}\}_{i \in \mathcal{K}}$ expressed in the navigation frame c_0 , which is recovered beforehand by visual-only bundle adjustment, the trajectory enters the calibration as a fixed input.	9
2.4	Per-interval measurement sets between consecutive keyframes i and j . The IMU samples at high rate (orange ticks; gyroscope and accelerometer combined), the DVL provides body-velocity samples at a lower rate (green arrows), and the pressure sensor reports absolute depth as a step signal (blue).	10
3.1	Calibration Trajectory (Calib Medium variant): noiseless body velocity $\mathbf{v}$ (top), angular velocity $\boldsymbol{\omega}$ (middle), and linear acceleration $\mathbf{a}$ (bottom) in the relevant sensor frame. The sequenced per-axis excitation of $\mathbf{v}$ is visible in the top panel; the angular wiggle is persistent on all three axes.	26
3.2	The SE surrogate (left) alongside the real SE sequence (right). Each panel shows body velocity $\mathbf{v}$ (top), angular velocity $\boldsymbol{\omega}$ (middle), and linear acceleration $\mathbf{a}$ (bottom) in the relevant sensor frame. Both share the all-axis low-amplitude rotation character. The surrogate spans a larger translational range and depth of descent.	27
3.3	AQUA-SLAM IMU-DVL lever arm estimate magnitude $\ {}_b\hat{\mathbf{p}}_{d/b}\ $ and error magnitude $\ \bar{\mathbf{e}}_{b\hat{\mathbf{p}}_{d/b}}\ = \ {}_b\hat{\mathbf{p}}_{d/b} - {}_b\mathbf{p}_{d/b}^*\ $ across noise grades G1-G4 on the Calib Medium trajectory. Both quantities are means over $N = 200$ Monte Carlo trials, with error bars at ± 1 standard deviation. The dashed line marks the true magnitude $\ {}_b\mathbf{p}_{d/b}^*\ = 35.3$ cm.	30
4.1	The Formulation A coarse-to-fine calibration pipeline. Each stage activates a subset of parameters and residuals. Optimized parameters are shown in blue, fixed parameters in black, and parameters unobservable at that stage in red. The state is warm-started between passes. The final joint pass uses all parameters and all residuals.	36
4.2	IMU-DVL lever-arm error magnitude $\ \bar{\mathbf{e}}_{b\hat{\mathbf{p}}_{d/b}}\ = \ {}_b\hat{\mathbf{p}}_{d/b} - {}_b\mathbf{p}_{d/b}^*\ $ for AQUA-SLAM (red) and Formulation A (blue) across sensor grades G1-G4 on the Calib Medium trajectory. AQUA-SLAM's error is approximately equal to the true lever-arm magnitude $\ {}_b\mathbf{p}_{d/b}^*\ = 35.3$ cm at all grades, consistent with the AQUA-SLAM estimate collapsing to zero (Section 3.2). Formulation A's error grows with sensor noise grade and remains below AQUA-SLAM's at every grade. Error bars are ± 1 standard deviation across $N = 200$ Monte Carlo trials.	38

5.1	Predicted and realized IMU–DVL lever arm bias for the four trajectories. Green: oracle predicted bias magnitude swept over σ_g . Orange: realized total bias magnitude from $N = 200$ Monte Carlo trials. Blue dashed: realized component along $\hat{\mathbf{e}}_{\text{pred}}$. The upper axis marks the corresponding $\nu_{\min}^{(A)}$, and different regimes are shaded.	50
5.2	Robustness of $\nu_{\min}^{(A)}$ to extrinsic perturbation from the truth.	52
5.3	Operational $\hat{\nu}_{\min}^{(A)}$ against the oracle reference across the sixteen cells. . .	52
6.1	Oracle $\nu_{\min}^{(B)}$ versus $\nu_{\min}^{(A)}$ across the eight (trajectory, grade) cells, log-log axes. The dashed line marks $y = x$	59
6.2	Realized IMU–DVL lever arm bias magnitude under Formulation B versus Formulation A, log-log axes, across the eight (trajectory, grade) cells. Points below $y = x$ mark Formulation B’s improvement.	61
7.1	Window-quality statistics across the log, computed on 20 s windows at 2.5 s stride. The top two panels are gated: a window passes when its outlier fraction is below the dashed threshold (twice the inter-window median), and its median error velocity is below the per-window threshold, which scales as the inverse square root of the ping count. The lower two panels are reported but not used for selection. Points are colored by whether the window passes both gates, and the shaded band is the resulting healthy span after erosion.	68
7.2	Formulation B’s calibration errors against keyframe count, for windows anchored at 65 s and grown in five-keyframe increments to the full selected window. Rotation errors above, lever arm errors below; the final point at 184 keyframes is the selected window.	71
7.3	Wall-clock time of the calibration pipeline per stage against the keyframe count of the window; Stage 2 is inactive here (Section 7.1). The dashed line is the median time of a single AQUA-SLAM local bundle adjustment; half of the solves complete faster and half slower. The shaded band spans the middle half of the solve times, from the 25th to the 75th percentile. All timings are measured on an AMD Ryzen AI 7 350; both solvers run single-threaded, Ceres [1] for the calibration and g2o [2] for AQUA-SLAM’s bundle adjustment.	72
7.4	Per-run position ATE for the original baseline and the corrected residual, ten runs per configuration; the bar marks the median. The panels share no vertical scale: the Easy sequence is an order of magnitude more accurate than the other two.	73
7.5	Per-run position ATE for the four calibrations, ten runs per configuration; the bar marks the median. The panels share no vertical scale. . .	74

List of Tables

3.1	Canonical ground truth extrinsic set used throughout the synthetic experiments: IMU-DVL lever arm ${}_b\mathbf{p}_{d/b}^*$, camera-DVL lever arm ${}_c\mathbf{p}_{d/c}^*$, camera-pressure lever arm ${}_c\mathbf{p}_{p/c}^*$, IMU-camera rotation ${}_b\mathbf{R}_c^*$, and camera-DVL rotation ${}_c\mathbf{R}_d^*$. The two rotation matrices are dimensionless axis permutations; see the text.	25
3.2	Discrete-time per-sample standard deviations for the four noise grades on the GX5-AHRS + A50 + Bar30 sensor suite. G1 is datasheet-anchored, with the DVL entry taken from the error-velocity analysis of Appendix A; G3 is the AQUA-SLAM codebase configuration; G2 and G4 are interpolated and extrapolated around them. The G4 DVL white-noise row sits below G3 by construction; see the text.	28
3.3	Final-stage AQUA-SLAM Monte Carlo bias and standard deviation across noise grades G1–G4. Translation errors in cm; rotation and gravity errors in degrees. $N = 200$ trials per grade.	31
4.1	Monte Carlo bias and standard deviation for AQUA-SLAM and Formulation A across sensor grades G1–G4 on the Calib Medium trajectory ($N = 200$ trials). Translation in cm; rotation and gravity in degrees. . .	39
4.2	Pressure-sensor lever-arm calibration with Formulation A on the Calib Medium trajectory ($N = 200$ trials). The true value has magnitude $\ {}_c\mathbf{p}_{p/c}^*\ = 56.70$ cm.	39
5.1	Oracle $\nu_{\min}^{(A)}$ at the four sensor noise grades for the four trajectories. . .	48
5.2	IMU–DVL lever-arm bias on Calib Medium, realized with the gyroscope bias estimated and with it held fixed at its true value, against the oracle prediction (cm). The realized rows are mean $\pm$ per-trial standard deviation over $N = 200$ trials; the oracle prediction is deterministic. . .	51
6.1	Extrinsic-error comparison for Calib Medium: Formulation A (velocity substitution) versus Formulation B (velocity state). Each entry reports the bias magnitude $\ \bar{\mathbf{e}}\ $ and the per-trial 1σ scatter $\sigma_{\ \mathbf{e}\ }$ over the Monte Carlo trials.	60
6.2	Extrinsic-error comparison for the SE surrogate: Formulation A (velocity substitution) versus Formulation B (velocity state). Each entry reports the bias magnitude $\ \bar{\mathbf{e}}\ $ and the per-trial 1σ scatter $\sigma_{\ \mathbf{e}\ }$ over the Monte Carlo trials.	60
6.3	IMU–DVL lever arm bias $\ {}_b\bar{\mathbf{e}}_{d/b}\ $ under Formulations A and B, with the gyroscope bias estimated and with it held fixed at its true value (cm, mean $\pm$ per-trial 1σ over $N = 200$ trials).	62

7.1	Calibration errors on the selected window [65, 125] s (184 keyframes) against the reference calibration provided with the Tank dataset. Rotation errors are axis-angle magnitudes in degrees; lever arm errors are in centimeters.	69
7.2	Trajectory error of AQUA-SLAM with the baseline and the corrected velocity-transport residual, both with the reference extrinsics. Each cell is the median of the per-run RMSEs over ten runs, with the interquartile range (25th to 75th percentile) in parentheses. Loop closure disabled; errors after rigid-body alignment without scale correction, computed with the evo toolkit.	73
7.3	Trajectory error of AQUA-SLAM under the corrected residual with four calibrations: the dataset reference and the three estimates of Table 7.1. Each cell is the median of the per-run RMSEs over ten runs, with the interquartile range in parentheses. Loop closure disabled; errors after rigid-body alignment without scale correction, computed with the evo toolkit.	75

Introduction

Autonomous underwater vehicles (AUVs) have become central to subsea inspection. They examine pipelines and cables, assess offshore wind foundations, inspect hulls and harbors, and map the seabed and submerged archaeological sites [3]. These tasks reach depths and ranges where continuous human supervision is impractical. The vehicle must therefore operate on its own. Whether a mission succeeds and whether the data it returns are usable rests on the vehicle’s own ability to localize itself reliably [3].

On land and in the air, a vehicle can anchor its position to a global navigation satellite system and leave its onboard sensors to track only the short-term motion between satellite fixes. Underwater, that external anchor disappears, as satellite signals do not penetrate more than a few centimeters of water [3]. Acoustic positioning, using long-baseline or ultra-short-baseline arrays, can provide an alternative external reference, but it requires beacons to be deployed and surveyed in advance, or a vessel to track the vehicle from the surface [4]. Both options confine a mission to prepared areas and add cost that exploratory or long-range work cannot bear [4]. Consequently, underwater localization relies primarily on the vehicle’s onboard sensors.

One of the standard approaches to onboard localization is visual simultaneous localization and mapping (SLAM) [3], which recovers the vehicle’s trajectory from camera imagery. Underwater, however, vision alone is unreliable, and inspection vehicles therefore carry additional sensors that directly measure motion [5]. Fusing their readings requires knowing how each sensor is mounted on the vehicle: the extrinsic calibration. This thesis develops a targetless method for jointly estimating these extrinsics from the sensor suite’s own data.

1.1 Motivation: Multi-Sensor Fusion and Extrinsic Calibration

The camera is the natural starting point for an inspection vehicle: it directly observes the structures under survey. Visual SLAM depends on that imagery, but underwater, light attenuates rapidly, lowering contrast and shortening usable range. Turbidity and suspended particles obscure the scene [4, 6]. The image can drop out entirely for tens of seconds at a time [7]. Vision alone, therefore, cannot sustain the localization that an inspection mission requires.

This requirement naturally gives rise to a highly complementary sensor suite, in which each modality addresses a specific weakness of the others. An inertial measurement unit (IMU) reports rotation and acceleration at a high rate, bridging the visual dropouts that defeat the camera. Integrated on its own, though, it accumulates drift. A Doppler velocity log (DVL) bounds that drift by measuring the vehicle’s velocity directly against the seabed. Neither the camera, the IMU, nor the DVL observes

absolute depth, which a pressure sensor supplies. A magnetometer is the obvious further addition, but the ferrous structures in inspection AUV surveys, such as pipelines and ship hulls, heavily distort the local magnetic field and make its heading reference unreliable [8]. Consequently, modern inspection systems have converged on a visual-inertial-acoustic-depth sensing configuration [5, 9, 10].

Combining these sensors requires knowing how they are physically arranged on the vehicle. Each sensor reports in its own coordinate frame, fixed to its mounting location. Reconstructing a coherent motion means expressing every reading in a single common reference frame. This requires the rigid-body transformation between each sensor’s frame and the reference: a rotation together with a translation. These transformations together form the *extrinsics* of the suite. They are set by the physical mounting and remain stable during a mission, but they shift over weeks of operation due to drag, vibration, and wave loading [7]. Accurate knowledge of these extrinsics is critical: an error in them does not stay local. It propagates through the entire fusion pipeline [11], biasing any derived measurements.

Recovering this calibration accurately is therefore essential to the overall system performance, and is the focus of the thesis.

1.2 State of the Art in Extrinsic Calibration

Extrinsic calibration is an estimation problem. The extrinsics are stable within a mission but unknown a priori, and each sensor’s measurements depend on them through the vehicle’s motion. Calibration estimates the extrinsics by fitting that dependence to the recorded data. Existing methods that achieve this vary in five aspects: the calibration references they use, how they represent the trajectory, whether they calibrate the sensors in a cascade or jointly, how they ensure the motion is informative, and whether they account for noise in the measured motion. This section discusses them in turn.

The choice of calibration references differs across sensors, though every extrinsic is expressed in the same common frame, conventionally the IMU’s. The IMU thus defines that frame rather than being calibrated against anything, so it has no calibration tradition of its own. The camera and DVL must each be calibrated to it. Camera calibration is either target-based or targetless. Target-based methods recover the extrinsics by observing a known pattern, such as a checkerboard, from many viewpoints. The geometry they obtain is precise and repeatable, which is why such methods remain the laboratory standard [12, 13]. They require, however, that the pattern stay in view throughout, which constrains how the vehicle may move. Targetless methods drop the pattern and instead recover the extrinsics from the vehicle’s own motion; underwater, the camera’s vision-only trajectory serves as the reference [14].

DVL calibration follows a separate tradition. Established methods bootstrap from an external reference, aligning the DVL against a surface satellite fix [15, 16], a pre-calibrated high-grade inertial navigation system [17], or an acoustic positioning array [18]. In both traditions, the standard option depends on something external to the suite: a pattern held in the camera’s view, or a positioning reference for the DVL. Open water removes them all. A pattern cannot be held in view across a survey, and no external reference is available to a vehicle operating independently underwater [7, 14].

What remains for either sensor is the suite’s own data. Self-contained calibration is therefore not one option among several but the shared consequence of the deployment, and it unifies what are otherwise separate per-sensor traditions.

Recovering the extrinsics this way first requires recovering the vehicle’s trajectory from the suite’s own measurements, which admits different representations [19]. Continuous-time methods model the whole trajectory as a smooth curve, fixed by a set of control points and evaluated at any instant between them [20, 21]. This representation fuses sensors with different sampling rates without a separate state for each reading, and it allows small timing offsets between sensors to be estimated alongside the geometry. Frameworks built on this representation are conventionally posed as a single-batch optimization over the entire recording. This prioritizes accuracy but requires complete recording in advance, and recovering the continuous trajectory it conditions on is itself an offline computation [12]. Such a calibration cannot run during the mission it calibrates. In contrast, discrete-time methods maintain state only at the instants when the sensors report [7]. They give up the smooth interpolation between instants that the continuous-time curve provides. In return, the state they keep is the one that a running SLAM estimator already maintains, so a discrete-time calibration fits such a system without changing its representation. This makes it suitable for online operation during a mission [7].

Beyond how the motion is represented, calibration methods also differ in how they structure the estimation problem. The calibration can run as a cascade of steps or as a single joint estimate. A cascade calibrates the sensors in pairs or in sequence, each step resting on the one before. For the camera, IMU, and DVL, this typically means calibrating the camera and IMU first, then aligning the DVL against the resulting inertial trajectory. This keeps each step small, but an error early in the chain propagates to every subsequent step. A joint estimate solves for all the extrinsics at once, so every reading informs a single consistent result, with no error passed down the chain.

Even a joint formulation recovers the parameters only if the motion is informative enough to make them observable. Some trajectories fail to excite every degree of freedom, leaving part of the parameter space unobservable regardless of data quality. A constant-velocity pass along a plane is one such case, and it leaves some of the extrinsics unobservable [22]. An information-based metric scores how observable the parameters are under a given motion. Two approaches are commonly adopted to build on this metric. Active methods plan the motion in advance to score well, which secures observability but requires control over the trajectory [23, 24]. Passive methods, by contrast, take the motion as given and select the highest-scoring segments from the recording [25]. This suits opportunistic operation, as efficient approximations of these information scores can be computed alongside a running estimator. However, these metrics quantify whether a motion makes the parameters observable, not whether the noise in that motion biases the estimate; a motion can score well and still produce the bias.

Informative motion is therefore necessary but not sufficient. This leads to a final and often overlooked aspect of the estimation itself: noise can bias the estimate, not only scatter it. The calibration is driven by noisy sensor measurements. Standard estimators model noise as random error that scatters the estimate about the true value,

kept small by weighting each measurement by its noise [7]. But some of the noise displaces the estimate away from the true value rather than scattering it, a bias that these estimators leave unmodeled. This is the errors-in-variables problem [26, 27]. A calibration that ignores it can converge cleanly and report a confident estimate that is nonetheless wrong. Reliable calibration, therefore, requires methods aware of this bias, since convergence does not indicate accuracy.

Awareness of this bias completes the set of demands the inspection setting places on a calibration. An inspection AUV must calibrate itself in the field during deployments long enough for the extrinsics to drift, with no target to observe and no external reference available. A calibration fit for this setting must run during the mission, draw only on the suite’s own data, and estimate the camera, IMU, DVL, and pressure extrinsics as a single, consistent solution. It must also remain accurate when the motion it conditions on is itself measured and noisy. Each of these capabilities has been demonstrated, but only in isolation. While unified continuous-time frameworks achieve joint calibration across extensive terrestrial suites, they omit the DVL’s unique acoustic models [20]. The closest underwater method, AQUA-SLAM, calibrates the camera, IMU, and DVL jointly but does not include the pressure sensor, and even across those three, its formulation carries a structural bias in the IMU–DVL translation [7]. Other systems that fuse all four are localization frameworks, which take the full set of extrinsics as fixed input rather than calibrating any of them [5, 10]. The pressure sensor translation is thus the one extrinsic in the suite that no method estimates. The noisy-motion error, in turn, has been confronted only in concurrent and independent work, and only for a single DVL-sensor pair [28]. That work also adopts a continuous-time formulation rather than the discrete-time formulation used by most deployed SLAM estimators. The gap is therefore not a shortage of individual solutions but the absence of a single approach that achieves all four at once: full sensor coverage, self-contained in-mission operation, discrete-time compatibility, and robustness to noisy motion.

1.3 Research Questions and Contributions

To address this gap, this thesis develops a calibration framework tailored to the operational constraints of inspection-class AUVs. Specifically, it addresses the following main research question.

How can an inspection AUV jointly calibrate its camera, IMU, DVL, and pressure sensors in situ, from onboard data alone and within a discrete-time formulation, while accounting for its bias sources?

The central question breaks down into three subquestions, each addressed in later chapters.

1. In the established baseline formulation, what causes the bias in the IMU–DVL translation estimate, and how can it be corrected?
2. To what extent does the corrected formulation eliminate bias under realistic sensor noise, and can any remaining bias be diagnosed, without ground truth, before it is trusted, and then reduced?

3. Can the calibration operate efficiently enough to run during a mission, rather than only offline, and does it recover an accurate calibration from real underwater data that improves downstream SLAM?

This thesis makes the following contributions:

1. A targetless joint extrinsic calibration of the camera, IMU, DVL, and pressure sensor suite of an inspection-class AUV, validated on synthetic and real underwater data. The extrinsics are recovered together from the suite’s own measurements, with no target and no external aiding. To the best of our knowledge, it is the first method to estimate the pressure-sensor translation jointly with the rest of the suite, which no prior calibration recovers.
2. A closed-form analysis showing that the established baseline formulation estimates the IMU–DVL translation with a structural bias (Chapter 3), and a correction that removes it (Chapter 4). The bias comes from a term the baseline leaves out of its motion model, which drives the estimate toward zero, regardless of data quality. Restoring that term makes the translation estimate recoverable, making the correction necessary for any joint calibration built on the baseline.
3. The identification of a second, more subtle bias on the IMU–DVL translation estimate that the corrected formulation still carries (Chapter 5). This bias comes from noise in the rotation-rate measurement that the calibration relies on, and it grows with that noise. It is embedded in the solution, so calibration may consistently converge on an incorrect value, and better tuning alone is insufficient to mitigate it.
4. A diagnostic that flags this bias before a calibration is trusted (Chapter 5). Measuring the bias directly would need the ground truth value, yet its severity can be read in advance from the sensor data and a rough initial guess. The bias can then be identified before the calibration is used, rather than after.
5. A reformulation that reduces this bias by changing how the body’s velocity enters the estimate (Chapter 6). Treating the velocity as an independent latent variable, rather than reading it from the DVL, keeps the noisy rotation rate away from the part of the model that is sensitive to it. The bias is reduced but not fully removed.
6. A demonstration that the calibration can be run during a mission, and a principled way to choose which stretch of a recording to run it on (Chapter 7). The selection screens the data using the DVL’s own consistency, so it requires no calibration parameters or ground truth. On real underwater data [29], the calibration reproduces the accuracy ordering seen in simulation and holds up in downstream SLAM.

The Calibration Problem

This chapter formalizes the extrinsic calibration problem for the underwater suite introduced in Chapter 1. Section 2.1 defines the coordinate frames of the sensors and identifies the calibration state. Section 2.2 specifies the data from which the calibration is built and the sensor measurement models, forming a residual for each sensor from the mismatch between its measurement and the body motion. Section 2.3 assembles these residuals into the weighted-least-squares cost whose minimizer is the calibration estimate, together with the constraint-aware iteration that produces that minimizer.

2.1 The Extrinsic Calibration Problem

We consider an underwater vehicle that carries four sensors: an IMU, a stereo camera, a DVL, and a pressure sensor. The IMU's frame serves as the *body frame*, denoted b , which is the vehicle-fixed reference frame used for expressing inter-sensor geometry. The camera, DVL, and pressure-sensor frames are denoted c , d , and p , respectively.

The four frames are related by many possible inter-sensor transformations (rotations and positions or translations), but these are not independent. A minimal subset determines the rest by rigid-body composition, so only it needs to be estimated. One such minimal subset is

$$\mathcal{E} = \{ {}_b\mathbf{R}_c, {}_b\mathbf{p}_{c/b}, {}_b\mathbf{R}_d, {}_b\mathbf{p}_{d/b}, {}_c\mathbf{p}_{p/c} \}, \quad (2.1)$$

where ${}_b\mathbf{R}_c \in SO(3)$ is the rotation from the camera frame c to the IMU frame b , and ${}_b\mathbf{p}_{c/b} \in \mathbb{R}^3$ is the position of the camera frame c relative to frame b , expressed in frame b , also called IMU-camera *lever arm*. The remaining entries follow the same convention. Here, $SO(3)$ is the set of valid rotation matrices: the 3×3 orthogonal matrices with determinant one, which preserve lengths and angles. These five entries comprise two inter-sensor rotations and three lever arms (Figure 2.1). The camera and DVL each contribute a rotation; the pressure sensor does not. It enters $\mathcal{E}$ with only a lever arm because its scalar depth reading is invariant to its own orientation, which is therefore unobservable from the data and excluded from $\mathcal{E}$.

Calibration estimates $\mathcal{E}$, but also additional quantities on which the sensor measurements depend. The accelerometer measures *specific force*, for example, which is the body acceleration relative to gravitational acceleration. Recovering the body acceleration, therefore, requires knowing the direction of gravity. The gyroscope and accelerometer also carry a slowly drifting bias, which corrupts every sample and must therefore be estimated as well. Hence, $\mathcal{E}$, the gravity direction, and the biases together form the *calibration state*, which separates into time-invariant quantities ($\mathcal{E}$ and the gravity direction) and time-varying quantities (the biases).

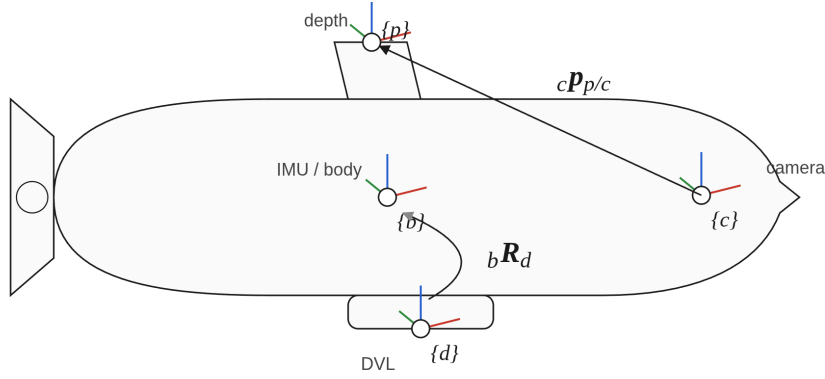


Figure 2.1: Two of the five elements of $\mathcal{E}$ on the vehicle: the DVL-IMU rotation ${}_b\mathbf{R}_d$ and camera-pressure lever arm ${}_c\mathbf{p}_{p/c}$.

We note that the gravity direction is constant over the calibration window. Representing it by a time-invariant set of coordinates, therefore, requires a frame that does not move with the vehicle. In a body-fixed frame, its coordinate representation changes over time as the vehicle turns. To avoid this, the first camera frame, denoted c_0 , serves as the *navigation frame*. It is the camera's frame at the initial time, frozen as a fixed reference. Gravity is then represented by its direction in c_0 , denoted as ${}_c\mathbf{g} \in S^2$, where S^2 denotes the unit sphere in $\mathbb{R}^3$. The magnitude $g \approx 9.81 \text{ m/s}^2$ is assumed known. Because the camera's initial orientation relative to vertical is not known, the direction of gravity in c_0 is unknown a priori and must be estimated (Figure 2.2).

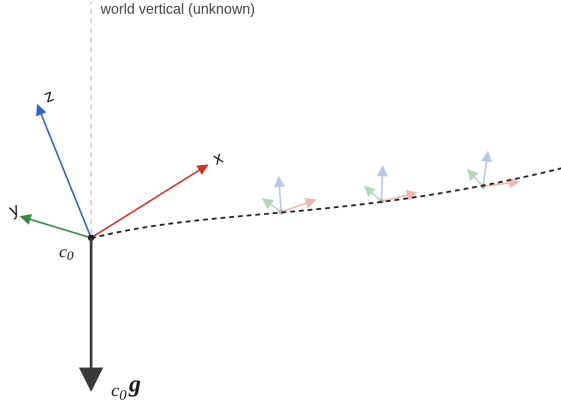


Figure 2.2: The navigation frame c_0 and the gravity direction ${}_c\mathbf{g}$ expressed in it. The world vertical (dashed) does not align with the axes of c_0 , so ${}_c\mathbf{g}$ is unknown a priori.

Unlike the gravity direction, the biases are time-varying, drifting slowly over the calibration window. Because the drift is slow, a value at every sample carries little more information than one at widely spaced times. Moreover, at the sensors' high sample rate, carrying one bias per sample would be impractical. Each bias is therefore carried only at a discrete set of times, the *keyframes* (set $\mathcal{K}$). Within each keyframe interval, the biases are treated as constant, so each is represented by a single value per keyframe.

The complete calibration state is then

$$\mathcal{X} = \left\{ \mathcal{E}, {}_{c_0}\mathbf{g}, \{ \mathbf{b}_i^g, \mathbf{b}_i^a \}_{i \in \mathcal{K}} \right\}. \quad (2.2)$$

Calibration is the set of values $\mathcal{X}^*$ that best explains the measurements. The next section provides what this requires: the data the estimate is built from, and a measurement model for each sensor. Each model predicts what its sensor should read, given the calibration state and the body motion over a keyframe interval. The difference between the actual reading and this prediction is the sensor’s residual. The models are built bottom-up: from a shared body motion model (Section 2.2.1) through the pressure sensor (Section 2.2.2), the IMU (Section 2.2.3), and the DVL (Section 2.2.4). The residuals they produce are assembled into the cost of Section 2.3.

2.2 Measurement Models and Residuals

The data collected over the calibration window comes in two forms: a fixed reference that the estimate is conditioned on, and noisy sensor measurements that the measurement models must explain. The fixed reference relies on the camera, which provides a reference trajectory consisting of poses (orientations and translations) $\{ {}_{c_0}\mathbf{R}_{c_i}, {}_{c_0}\mathbf{p}_{c_i} \}_{i \in \mathcal{K}}$ (Figure 2.3), where ${}_{c_0}\mathbf{p}_{c_i}$ abbreviates ${}_{c_0}\mathbf{p}_{c_i/c_0}$. The trajectory is recovered beforehand and treated as a fixed input rather than estimated. It comes from visual-only bundle adjustment [30] in a SLAM system, which refines camera poses from image features. This distinguishes the problem from SLAM, where the trajectory is itself part of the state and optimization problem.

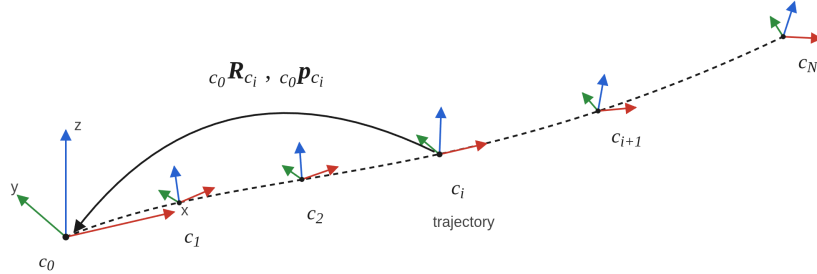


Figure 2.3: The camera trajectory $\{ {}_{c_0}\mathbf{R}_{c_i}, {}_{c_0}\mathbf{p}_{c_i} \}_{i \in \mathcal{K}}$ expressed in the navigation frame c_0 , which is recovered beforehand by visual-only bundle adjustment, the trajectory enters the calibration as a fixed input.

The IMU, DVL, and pressure sensor comprise the second form of data: each sensor produces a sample stream that constrains the state. Their per-interval measurement sets are $\mathcal{Z}_{ij}^{\text{IMU}}$, $\mathcal{Z}_{ij}^{\text{DVL}}$, and $\mathcal{Z}_{ij}^P$, as illustrated in Figure 2.4, where i and $j > i$ denote consecutive keyframes. These are collected into

$$\mathcal{Z} = \left\{ \mathcal{Z}_{ij}^{\text{IMU}}, \mathcal{Z}_{ij}^{\text{DVL}}, \mathcal{Z}_{ij}^P \right\}_{(i,j)}, \quad (2.3)$$

indexed over consecutive keyframe intervals.

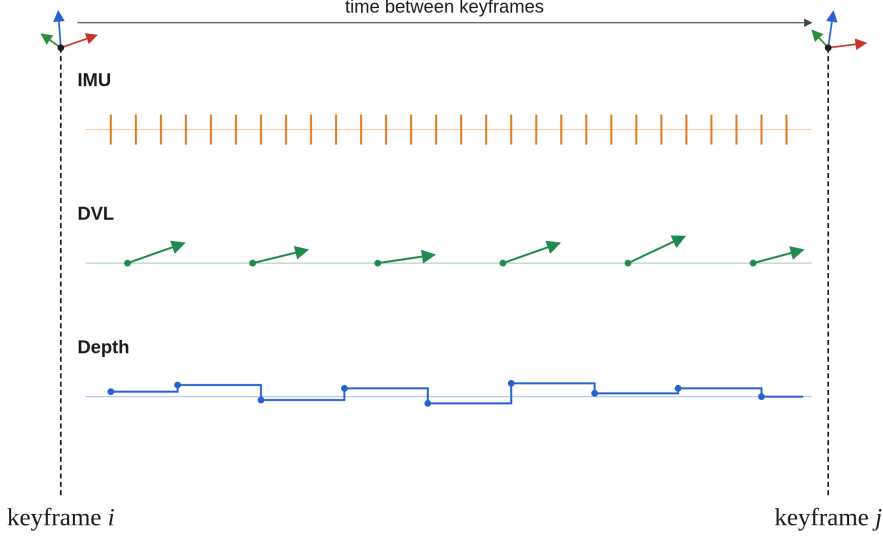


Figure 2.4: Per-interval measurement sets between consecutive keyframes i and j . The IMU samples at high rate (orange ticks; gyroscope and accelerometer combined), the DVL provides body-velocity samples at a lower rate (green arrows), and the pressure sensor reports absolute depth as a step signal (blue).

The keyframes and sample streams just introduced are referenced throughout the section by two index conventions and one subscript convention. The letters i and j index keyframes, and k indexes the higher-rate samples within a keyframe interval. A sensor-frame subscript carrying a keyframe index, such as c_i , b_i , d_i , or p_i , denotes that sensor’s frame at keyframe i ; for the IMU, DVL, and pressure streams, it identifies the first sample at or after keyframe i . Each interval (i, j) owns the half-open span $[t_i, t_j)$ of sample times, where t_i is the time of keyframe i and $\Delta t_{ij} = t_j - t_i$ is the interval duration, so a reading at a keyframe boundary belongs to a single interval.

2.2.1 Shared Foundation: Body Motion Model

The body motion model provides a unified kinematic description of the vehicle and serves as the backbone for all sensor models. It consists of the body pose, the body velocity in the navigation frame, and the body angular velocity in the body frame. The body pose is fixed by the camera trajectory; the body velocity and angular velocity are not.

The body pose is not part of the calibration state. It follows from the camera pose at keyframe i and the camera-IMU extrinsics by rigid-body transformation,

$${}_{c_0}\mathbf{R}_{b_i} = {}_{c_0}\mathbf{R}_{c_i} {}_b\mathbf{R}_c^T, \quad {}_{c_0}\mathbf{p}_{b_i} = {}_{c_0}\mathbf{p}_{c_i} - {}_{c_0}\mathbf{R}_{c_i} {}_b\mathbf{R}_c^T {}_b\mathbf{p}_{c/b}, \quad (2.4)$$

in which the camera pose is a known input from the trajectory, and the camera-IMU extrinsics are part of the calibration state. The body pose, therefore, varies with the extrinsics while the camera trajectory is held fixed. So residuals depend only on the extrinsics, not the body pose itself.

The body velocity and angular velocity are determined by neither the trajectory nor the state. The trajectory provides the pose at each keyframe but not its time derivatives, and the calibration state $\mathcal{X}$ of (2.2) does not include them either. Nevertheless, several of the sensor models introduced in the following sections depend explicitly on the body velocity ${}_{c_0}\mathbf{v}_{b_i} \in \mathbb{R}^3$ or the body angular velocity ${}_{b_i}\boldsymbol{\omega}_{b_i} \in \mathbb{R}^3$. Therefore, the calibration problem cannot be fully specified without introducing a representation for these latent motion variables and constraints that relate them to the available measurements. How this should be done is a central modeling choice, and the formulations developed in the following chapters differ precisely in how they resolve it.

2.2.2 Pressure Sensor Model

The pressure sensor is the simplest of the sensor suite and needs no integration. It measures hydrostatic pressure, which onboard electronics convert to depth below the water surface. The noisy reading $\tilde{p}_{p_i}$ is modeled as

$$\tilde{p}_{p_i} = {}_{c_0}\mathbf{g}^T {}_{c_0}\mathbf{p}_{p_i} + o_P + \eta_i^P, \quad (2.5)$$

where ${}_{c_0}\mathbf{p}_{p_i} \in \mathbb{R}^3$ is the pressure-sensor position in the navigation frame at keyframe i , $o_P \in \mathbb{R}$ is the datum offset between the navigation-frame origin and the water surface, and η_i^P is zero-mean Gaussian white noise with variance σ_P^2 . The first term is the projection of the sensor position onto the gravity direction ${}_{c_0}\mathbf{g}$, which is positive when the sensor lies below the navigation-frame origin.

The model carries the sensor position ${}_{c_0}\mathbf{p}_{p_i}$, which is not itself known. Because the sensor is rigidly mounted to the camera, this position decomposes into the known camera position and the unknown camera-pressure lever arm ${}_{c_0}\mathbf{p}_{p/c} \in \mathbb{R}^3$,

$${}_{c_0}\mathbf{p}_{p_i} = {}_{c_0}\mathbf{p}_{c_i} + {}_{c_0}\mathbf{R}_{c_i} {}_{c_0}\mathbf{p}_{p/c}. \quad (2.6)$$

Substituting this into the model leaves three unknowns: the gravity direction, the camera-pressure lever arm, and the datum offset. The first two are estimated in the calibration state. The datum offset is not: it is a single constant shared by every reading, which the residual removes.

The depth residual over the interval (i, j) differences the interval's first and last readings so that the shared datum offset cancels. The first reading is $\tilde{p}_{p_i}$, and the last is $\tilde{p}_{p_{j-1}}$, the last pressure sample before keyframe j ; using the last interior reading rather than the keyframe- j reading keeps consecutive intervals from sharing a sample. The residual is

$$r_{ij}^P = {}_{c_0}\mathbf{g}^T \left({}_{c_0}\mathbf{p}_{c_j} - {}_{c_0}\mathbf{p}_{c_i} + ({}_{c_0}\mathbf{R}_{c_j} - {}_{c_0}\mathbf{R}_{c_i}) {}_{c_0}\mathbf{p}_{p/c} \right) - (\tilde{p}_{p_{j-1}} - \tilde{p}_{p_i}). \quad (2.7)$$

The only random ingredients are the two depth readings, each of variance σ_P^2 , so the residual has variance $2\sigma_P^2$. Because consecutive intervals draw disjoint samples, adjacent depth residuals share no reading and are uncorrelated. The pressure residual depends only on the gravity direction and sensor position and does not involve body velocity terms.

2.2.3 IMU Model

The IMU model is developed in three parts, following [31]: a per-sample model that corrupts each raw gyroscope and accelerometer reading with bias and white noise; a scheme that compresses the high-rate samples over an interval into a single relative-motion constraint between two keyframes; and the residuals that follow from that constraint. Bias appears in each part: it enters the per-sample model, drifts across keyframes as a random walk, and is corrected within the compression scheme.

The gyroscope measures the body-frame angular velocity, and the accelerometer measures the body-frame specific force. At sample k , the noisy raw measurements decompose as

$${}_{b_k}\tilde{\boldsymbol{\omega}}_{b_k} = {}_{b_k}\boldsymbol{\omega}_{b_k} + \mathbf{b}_k^g + \boldsymbol{\eta}_k^g, \quad (2.8)$$

$${}_{b_k}\tilde{\mathbf{a}}_{b_k} = {}_{c_0}\mathbf{R}_{b_k}^T ({}_{c_0}\mathbf{a}_{b_k} - g {}_{c_0}\mathbf{g}) + \mathbf{b}_k^a + \boldsymbol{\eta}_k^a, \quad (2.9)$$

where ${}_{b_k}\tilde{\boldsymbol{\omega}}_{b_k}$ and ${}_{b_k}\tilde{\mathbf{a}}_{b_k}$ are the raw gyroscope and accelerometer readings at sample k , ${}_{b_k}\boldsymbol{\omega}_{b_k}$ is the true body angular velocity, ${}_{c_0}\mathbf{a}_{b_k}$ is the body acceleration in the navigation frame, $g {}_{c_0}\mathbf{g}$ is the gravity vector, and $\mathbf{b}_k^g$ and $\mathbf{b}_k^a$ are the true biases at sample k . The terms $\boldsymbol{\eta}_k^g, \boldsymbol{\eta}_k^a$ are zero-mean Gaussian white noise with covariances $\boldsymbol{\Sigma}_{\text{Gyro}}$ and $\boldsymbol{\Sigma}_{\text{Acc}}$.

The bias is assumed constant over the interval, so $\mathbf{b}_k^g$ and $\mathbf{b}_k^a$ take the keyframe values $\mathbf{b}_i^g$ and $\mathbf{b}_i^a$ carried in the state. Across keyframes, the bias is not constant. It is modeled as a random walk that links the values at successive keyframes,

$$\mathbf{b}_j^g = \mathbf{b}_i^g + \boldsymbol{\eta}_{ij}^{bg}, \quad \mathbf{b}_j^a = \mathbf{b}_i^a + \boldsymbol{\eta}_{ij}^{ba}, \quad (2.10)$$

where $\boldsymbol{\eta}_{ij}^{bg}$ and $\boldsymbol{\eta}_{ij}^{ba}$ are zero-mean Gaussian with covariances $\boldsymbol{\Sigma}_{bg} \Delta t_{ij}$ and $\boldsymbol{\Sigma}_{ba} \Delta t_{ij}$, scaling linearly with the interval, and $\boldsymbol{\Sigma}_{bg}$ and $\boldsymbol{\Sigma}_{ba}$ are the bias random-walk covariance densities. Linking the keyframes this way tracks the slow drift across the window rather than estimating it independently at each keyframe.

With the per-sample model and its bias in place, the high-rate samples over an interval are integrated into a per-interval measurement. If performed directly, this integration propagates the body motion forward, using the body pose at keyframe i as the initial condition. However, that pose depends on the camera-IMU rotation ${}_b\mathbf{R}_c$ through (2.4), so every change to ${}_b\mathbf{R}_c$ during optimization would force the whole integration to be redone. *Preintegration* [31] avoids this. It reparameterizes the integral into compound measurements independent of the pose at i , which can therefore be computed once and reused. Specifically, it defines three compound measurements: the preintegrated rotation $\Delta \tilde{\mathbf{R}}_{b_j/b_i}$, velocity $\Delta \tilde{\mathbf{v}}_{b_j/b_i}$, and position $\Delta \tilde{\mathbf{p}}_{b_j/b_i}$, all evaluated

from the raw samples at the keyframe- i bias values $\mathbf{b}_i^g, \mathbf{b}_i^a$:

$$\Delta \tilde{\mathbf{R}}_{b_j/b_i}(\mathbf{b}_i^g) = \prod_{k=i}^{j-1} \text{Exp}(({}_{b_k}\tilde{\boldsymbol{\omega}}_{b_k} - \mathbf{b}_i^g) \Delta t), \quad (2.11)$$

$$\Delta \tilde{\mathbf{v}}_{b_j/b_i}(\mathbf{b}_i^g, \mathbf{b}_i^a) = \sum_{k=i}^{j-1} \Delta \tilde{\mathbf{R}}_{b_k/b_i}(\mathbf{b}_i^g) ({}_{b_k}\tilde{\mathbf{a}}_{b_k} - \mathbf{b}_i^a) \Delta t, \quad (2.12)$$

$$\Delta \tilde{\mathbf{p}}_{b_j/b_i}(\mathbf{b}_i^g, \mathbf{b}_i^a) = \sum_{k=i}^{j-1} \left[\Delta \tilde{\mathbf{v}}_{b_k/b_i}(\mathbf{b}_i^g, \mathbf{b}_i^a) \Delta t + \frac{1}{2} \Delta \tilde{\mathbf{R}}_{b_k/b_i}(\mathbf{b}_i^g) ({}_{b_k}\tilde{\mathbf{a}}_{b_k} - \mathbf{b}_i^a) \Delta t^2 \right], \quad (2.13)$$

where Δt is the IMU sample period, and $\text{Exp} : \mathbb{R}^3 \rightarrow SO(3)$ is the exponential map, which maps an axis-angle vector, whose direction is the rotation axis and whose magnitude is the rotation angle, to the corresponding rotation matrix.

Each compound measurement is computed solely from the raw samples and is therefore independent of the extrinsics, the gravity direction, or the body pose at i . They do, however, still depend on bias. Therefore, a change in the bias estimate would require reintegration, negating the benefit of preintegration. To avoid this, the compound measurements are updated by a first-order Taylor expansion in the bias increments $\delta \mathbf{b}_i^g = \mathbf{b}_i^g - \bar{\mathbf{b}}_i^g$ and $\delta \mathbf{b}_i^a = \mathbf{b}_i^a - \bar{\mathbf{b}}_i^a$ around linearization points $\bar{\mathbf{b}}_i^g$ and $\bar{\mathbf{b}}_i^a$:

$$\Delta \tilde{\mathbf{R}}_{b_j/b_i}(\mathbf{b}_i^g) \approx \Delta \tilde{\mathbf{R}}_{b_j/b_i}(\bar{\mathbf{b}}_i^g) \text{Exp}\left(\frac{\partial \Delta \tilde{\mathbf{R}}_{b_j/b_i}}{\partial \mathbf{b}_i^g} \delta \mathbf{b}_i^g\right), \quad (2.14)$$

$$\Delta \tilde{\mathbf{v}}_{b_j/b_i}(\mathbf{b}_i^g, \mathbf{b}_i^a) \approx \Delta \tilde{\mathbf{v}}_{b_j/b_i}(\bar{\mathbf{b}}_i^g, \bar{\mathbf{b}}_i^a) + \frac{\partial \Delta \tilde{\mathbf{v}}_{b_j/b_i}}{\partial \mathbf{b}_i^g} \delta \mathbf{b}_i^g + \frac{\partial \Delta \tilde{\mathbf{v}}_{b_j/b_i}}{\partial \mathbf{b}_i^a} \delta \mathbf{b}_i^a, \quad (2.15)$$

$$\Delta \tilde{\mathbf{p}}_{b_j/b_i}(\mathbf{b}_i^g, \mathbf{b}_i^a) \approx \Delta \tilde{\mathbf{p}}_{b_j/b_i}(\bar{\mathbf{b}}_i^g, \bar{\mathbf{b}}_i^a) + \frac{\partial \Delta \tilde{\mathbf{p}}_{b_j/b_i}}{\partial \mathbf{b}_i^g} \delta \mathbf{b}_i^g + \frac{\partial \Delta \tilde{\mathbf{p}}_{b_j/b_i}}{\partial \mathbf{b}_i^a} \delta \mathbf{b}_i^a. \quad (2.16)$$

The expansion holds while the bias increment remains small; beyond a set threshold, the preintegration is recomputed and the Jacobians are re-evaluated at the new value.

Although free of the extrinsics and gravity, the preintegrations constrain both through their relation to the body motion. The preintegrated measurement model relates them to the body orientation, velocity, and position at keyframes i and j , each carrying preintegration noise:

$$\Delta \tilde{\mathbf{R}}_{b_j/b_i} = {}_b\mathbf{R}_{c \ c_0} \mathbf{R}_{c_i \ c_0}^T {}_b\mathbf{R}_{c_j}^T \text{Exp}(\delta \boldsymbol{\phi}_{b_j/b_i}), \quad (2.17)$$

$$\Delta \tilde{\mathbf{v}}_{b_j/b_i} = {}_b\mathbf{R}_{c \ c_0} \mathbf{R}_{c_i \ c_0}^T ({}_{c_0}\mathbf{v}_{b_j} - {}_{c_0}\mathbf{v}_{b_i} - g_{c_0} \mathbf{g} \Delta t_{ij}) + \delta \mathbf{v}_{b_j/b_i}, \quad (2.18)$$

$$\begin{aligned} \Delta \tilde{\mathbf{p}}_{b_j/b_i} = & {}_b\mathbf{R}_{c \ c_0} \mathbf{R}_{c_i \ c_0}^T ({}_{c_0}\mathbf{p}_{c_j} - {}_{c_0}\mathbf{p}_{c_i} - ({}_{c_0}\mathbf{R}_{c_j} - {}_{c_0}\mathbf{R}_{c_i}) {}_b\mathbf{R}_{c_i}^T {}_b\mathbf{p}_{c/b} \\ & - {}_{c_0}\mathbf{v}_{b_i} \Delta t_{ij} - \frac{1}{2} g_{c_0} \mathbf{g} \Delta t_{ij}^2) + \delta \mathbf{p}_{b_j/b_i}, \end{aligned} \quad (2.19)$$

where $\delta \boldsymbol{\phi}_{b_j/b_i}, \delta \mathbf{v}_{b_j/b_i}, \delta \mathbf{p}_{b_j/b_i}$ are zero-mean Gaussian preintegration noises whose joint covariance is derived in detail in [31].

Each model defines a residual: the discrepancy between the body-motion prediction and the compound measurement, with the bias corrections (2.14)–(2.16) applied at the current bias estimates $\mathbf{b}_i^g$ and $\mathbf{b}_i^a$. Stripping the noise terms, the three IMU residuals

are

$$\mathbf{r}_{ij}^{\text{IMU},R} = \text{Log}(\Delta \tilde{\mathbf{R}}_{b_j/b_i}(\mathbf{b}_i^g)^T {}_b\mathbf{R}_{c\ c_0} \mathbf{R}_{c_i\ c_0}^T \mathbf{R}_{c_j\ b} \mathbf{R}_c^T), \quad (2.20)$$

$$\mathbf{r}_{ij}^{\text{IMU},v} = {}_b\mathbf{R}_{c\ c_0} \mathbf{R}_{c_i}^T ({}_{c_0}\mathbf{v}_{b_j} - {}_{c_0}\mathbf{v}_{b_i} - g_{c_0} \mathbf{g} \Delta t_{ij}) - \Delta \tilde{\mathbf{v}}_{b_j/b_i}(\mathbf{b}_i^g, \mathbf{b}_i^a), \quad (2.21)$$

$$\begin{aligned} \mathbf{r}_{ij}^{\text{IMU},p} = & {}_b\mathbf{R}_{c\ c_0} \mathbf{R}_{c_i}^T ({}_{c_0}\mathbf{p}_{c_j} - {}_{c_0}\mathbf{p}_{c_i} - ({}_{c_0}\mathbf{R}_{c_j} - {}_{c_0}\mathbf{R}_{c_i}) {}_b\mathbf{R}_c^T {}_b\mathbf{p}_{c/b} \\ & - {}_{c_0}\mathbf{v}_{b_i} \Delta t_{ij} - \frac{1}{2} g_{c_0} \mathbf{g} \Delta t_{ij}^2) - \Delta \tilde{\mathbf{p}}_{b_j/b_i}(\mathbf{b}_i^g, \mathbf{b}_i^a), \end{aligned} \quad (2.22)$$

where $\text{Log} : SO(3) \rightarrow \mathbb{R}^3$ is the inverse of the exponential map Exp . It returns the axis-angle vector of a rotation matrix, so the rotation residual expresses the discrepancy between the predicted and measured rotation as a vector in $\mathbb{R}^3$. The rotation residual depends solely on the camera-IMU rotation. The velocity and position residuals additionally carry ${}_{c_0}\mathbf{g}$, the position residual also carries the camera-IMU lever arm, and both ${}_{c_0}\mathbf{v}_{b_i}$ and ${}_{c_0}\mathbf{v}_{b_j}$ enter the velocity and position residuals.

The bias residuals constrain the random-walk evolution of the IMU biases between keyframes:

$$\mathbf{r}_{ij}^{b^g} = \mathbf{b}_j^g - \mathbf{b}_i^g, \quad (2.23)$$

$$\mathbf{r}_{ij}^{b^a} = \mathbf{b}_j^a - \mathbf{b}_i^a. \quad (2.24)$$

The five components stack into the IMU residual $\mathbf{r}_{ij}^{\text{IMU}} \in \mathbb{R}^{15}$, defined as

$$\mathbf{r}_{ij}^{\text{IMU}} = \begin{bmatrix} \mathbf{r}_{ij}^{\text{IMU},R} \\ \mathbf{r}_{ij}^{\text{IMU},v} \\ \mathbf{r}_{ij}^{\text{IMU},p} \\ \mathbf{r}_{ij}^{b^g} \\ \mathbf{r}_{ij}^{b^a} \end{bmatrix}, \quad (2.25)$$

comprising the three preintegration residuals followed by the two bias residuals.

The IMU velocity and position residuals carry the body velocities ${}_{c_0}\mathbf{v}_{b_i}$ and ${}_{c_0}\mathbf{v}_{b_j}$ at the interval endpoints, as seen in Eqs. (2.21) and (2.22), one of the latent motion variables introduced in Section 2.2.1.

2.2.4 DVL Model

The DVL residual is modeled in two parts: a displacement residual accumulated over an interval, computed analogously to the IMU's compound measurements, and a velocity residual at a single keyframe. The displacement preintegration follows AQUA-SLAM [7], and the velocity residual uses the kinematically consistent velocity transport. Unlike the IMU preintegrations, both carry the extrinsics. This formulation again requires body angular velocity for velocity transport, highlighting the recurring dependence on latent motion variables. The two residuals are derived in turn after introducing the measurement model.

The DVL measures the velocity of the DVL frame, expressed in that frame. Unlike the IMU, it samples at a lower rate, sometimes only once or a few times per interval. At sample k ,

$${}_{d_k}\tilde{\mathbf{v}}_{d_k} = {}_{d_k}\mathbf{v}_{d_k} + \boldsymbol{\eta}_k^d, \quad (2.26)$$

where ${}_{d_k}\tilde{\mathbf{v}}_{d_k}$ is the raw DVL reading at sample k , ${}_{d_k}\mathbf{v}_{d_k}$ is the true DVL-frame velocity, and $\boldsymbol{\eta}_k^d$ is zero-mean Gaussian white noise with covariance $\boldsymbol{\Sigma}_{\text{DVL}}$.

The DVL displacement preintegration follows AQUA-SLAM [7]. Similar to the IMU, integrating the DVL velocity over an interval would depend on the body pose at keyframe i , which itself moves with the camera-IMU rotation. Preintegration again decouples the integral from that pose. Forming a body-frame displacement from the DVL velocity requires rotating each velocity measurement according to the vehicle's orientation change since keyframe i . This orientation evolves continuously over the interval and is available from the gyroscope at the higher IMU rate. The displacement is therefore accumulated at the gyroscope rate: each DVL reading is held constant between its samples (a zero-order hold, which neglects any change in velocity between samples) and rotated by the IMU's preintegrated rotation $\Delta\tilde{\mathbf{R}}_{b_k/b_i}$. The interval's DVL samples thus compress into a single compound measurement, the preintegrated DVL displacement,

$$\Delta\tilde{\mathbf{p}}_{d_j/d_i}(\mathbf{b}_i^g) = \sum_{k=i}^{j-1} \Delta\tilde{\mathbf{R}}_{b_k/b_i}(\mathbf{b}_i^g) {}_b\mathbf{R}_d {}_{d_k}\tilde{\mathbf{v}}_{d_k} \Delta t, \quad (2.27)$$

where Δt is the gyroscope sample period, ${}_{d_k}\tilde{\mathbf{v}}_{d_k}$ is the most recent DVL reading, and $\Delta\tilde{\mathbf{R}}_{b_k/b_i}$ is the preintegrated rotation of (2.11). The displacement $\Delta\tilde{\mathbf{p}}_{d_j/d_i}$ is the translation of the DVL origin between the two keyframes, expressed in the body frame at keyframe i .

Unlike the IMU preintegration, this displacement is not free of the extrinsics: it explicitly includes the DVL-IMU rotation ${}_b\mathbf{R}_d$. The displacement also depends on the gyroscope bias through the preintegrated rotation $\Delta\tilde{\mathbf{R}}_{b_k/b_i}$. A change in either DVL-IMU rotation ${}_b\mathbf{R}_d$ or $\mathbf{b}_i^g$ during optimization would, in principle, require reintegrating the sum. AQUA-SLAM [7] avoids this with a linear approximation, as for the IMU bias: the displacement is linearized about the current ${}_b\mathbf{R}_d$ and $\mathbf{b}_i^g$, then updated by a first-order correction for small increments. The sum is reintegrated only when an increment exceeds a set threshold, beyond which the first-order correction no longer holds. The correction in the gyroscope bias reuses the rotation bias-correction Jacobian of (2.14).

The preintegrated displacement relates the DVL measurement to the body's true motion and, thereby, constrains the extrinsics. Specifically, it relates the body orientation and position at keyframes i and j :

$$\Delta\tilde{\mathbf{p}}_{d_j/d_i} = {}_b\mathbf{R}_c {}_{c_0}\mathbf{R}_{c_i}^T ({}_{c_0}\mathbf{p}_{c_j} - {}_{c_0}\mathbf{p}_{c_i} + ({}_{c_0}\mathbf{R}_{c_j} - {}_{c_0}\mathbf{R}_{c_i}) {}_{c_0}\mathbf{p}_{d/c}) + \delta\mathbf{p}_{d_j/d_i}, \quad (2.28)$$

where $\delta\mathbf{p}_{d_j/d_i}$ is the zero-mean preintegration noise.

The corresponding DVL translation residual is the difference between this body-motion prediction and the preintegrated displacement over the interval (i, j) :

$$\mathbf{r}_{ij}^{\text{DVL},p} = {}_b\mathbf{R}_c {}_{c_0}\mathbf{R}_{c_i}^T ({}_{c_0}\mathbf{p}_{c_j} - {}_{c_0}\mathbf{p}_{c_i} + ({}_{c_0}\mathbf{R}_{c_j} - {}_{c_0}\mathbf{R}_{c_i}) {}_{c_0}\mathbf{p}_{d/c}) - \Delta\tilde{\mathbf{p}}_{d_j/d_i}(\mathbf{b}_i^g). \quad (2.29)$$

At the true extrinsics and true body motion, the prediction equals the noise-free displacement, so the residual reduces to the preintegration noise. Notably, this residual carries the extrinsics on both sides: ${}_b\mathbf{R}_c$ and ${}_{c_0}\mathbf{p}_{d/c}$ in the prediction, and ${}_b\mathbf{R}_d$ embedded within the preintegrated displacement.

The DVL also provides an instantaneous velocity constraint at a single keyframe using rigid-body velocity transport. At keyframe i , the body-origin velocity is carried to the DVL origin by the angular-velocity term and then rotated into the DVL frame, which gives the DVL velocity residual:

$$\mathbf{r}_i^{\text{DVL},v} = {}_b\mathbf{R}_{c\ c_0} \mathbf{R}_{c_i\ c_0}^T \mathbf{v}_{b_i} - ({}_b\mathbf{R}_{d\ d_i} \tilde{\mathbf{v}}_{d_i} - {}_{b_i}\boldsymbol{\omega}_{b_i} \times {}_b\mathbf{p}_{d/b}), \quad (2.30)$$

The velocity and displacement residuals stack into DVL residual $\mathbf{r}_{ij}^{\text{DVL}} \in \mathbb{R}^6$,

$$\mathbf{r}_{ij}^{\text{DVL}} = \begin{bmatrix} \mathbf{r}_{ij}^{\text{DVL},p} \\ \mathbf{r}_i^{\text{DVL},v} \end{bmatrix}. \quad (2.31)$$

The DVL velocity residual carries both the body velocity ${}_{c_0}\mathbf{v}_{b_i}$ and the body angular velocity ${}_{b_i}\boldsymbol{\omega}_{b_i}$, the two latent motion variables introduced in Section 2.2.1.

The pressure, IMU, and DVL residuals defined in this section are assembled into a single cost whose minimizer is the calibration estimate. Section 2.3 forms this cost and presents its minimization.

2.3 Calibration Cost and Its Minimization

The cost follows from a maximum-a-posteriori (MAP) treatment of the estimation: the calibration estimate is the value of $\mathcal{X}$ that maximizes the posterior $p(\mathcal{X} \mid \mathcal{Z}) \propto p(\mathcal{X}) p(\mathcal{Z} \mid \mathcal{X})$. The likelihood $p(\mathcal{Z} \mid \mathcal{X})$ factors under two independence assumptions: the measurements of different sensors are conditionally independent given the state, and the sample sets of different keyframe intervals are mutually independent. The factored posterior is

$$p(\mathcal{X} \mid \mathcal{Z}) \propto p(\mathcal{X}) \prod_{(i,j)} p(\mathcal{Z}_{ij}^{\text{IMU}} \mid \mathcal{X}) p(\mathcal{Z}_{ij}^{\text{DVL}} \mid \mathcal{X}) p(\mathcal{Z}_{ij}^P \mid \mathcal{X}), \quad (2.32)$$

with (i, j) ranging over consecutive keyframe intervals. The prior $p(\mathcal{X})$ encodes prior knowledge about the state variables, such as an initial guess for the extrinsics.

The half-open intervals make the sample sets of different intervals disjoint, so the second assumption, mutual independence across keyframe intervals, holds exactly. The first, conditional independence across sensors, needs to be qualified for the preintegrated measurements within an interval. The preintegrated rotation $\Delta \tilde{\mathbf{R}}_{b_k/b_i}$ appears in both the IMU preintegration and the DVL displacement preintegration. The gyroscope noise in that rotation, therefore, enters both factors, so they are not statistically independent. The sensor-wise conditional-independence assumption stated above is exact for the raw samples, whose sensor noise is physically independent, but only approximate for the preintegrated factors, which share the gyroscope-integrated rotation.

The thesis treats the factors as conditionally independent regardless. This makes the cost a composite-likelihood approximation rather than the exact MAP objective, a product of factors each correct on its own but not jointly exact. The practical effect on the estimate is benign. Under standard regularity conditions, the point estimate remains consistent, since each factor is correctly specified on its own. The effect on

the uncertainty is not benign: the reported posterior covariance is slightly optimistic, because the shared gyroscope noise is double-counted across the IMU and DVL factors. Previous works make the same approximation, without accounting for the correlation [5, 7]. A joint IMU-DVL preintegration factor that exactly captures the correlation is outside the scope of the thesis.

Under zero-mean Gaussian noise on each factor, the negative log of a factor is a squared Mahalanobis residual, up to constants that do not affect the minimizer. The prior is itself a Gaussian factor, so it contributes likewise, as a squared residual $\mathbf{r}^0$ penalizing departure from the initial guess. Each measurement residual $\mathbf{r}$ is the mismatch between a measured quantity and its predicted value from the state, with covariance $\Sigma_{\mathbf{r}}$. Maximizing the posterior is equivalent to minimizing the weighted-least-squares cost

$$\mathcal{X}^* = \arg \min_{\mathcal{X}} \|\mathbf{r}^0\|_{\Sigma_{\mathbf{r}^0}}^2 + \sum_{(i,j)} \left(\|\mathbf{r}_{ij}^{\text{IMU}}\|_{\Sigma_{\mathbf{r}_{ij}^{\text{IMU}}}}^2 + \|\mathbf{r}_{ij}^{\text{DVL}}\|_{\Sigma_{\mathbf{r}_{ij}^{\text{DVL}}}}^2 + \|\mathbf{r}_{ij}^P\|_{\Sigma_{\mathbf{r}_{ij}^P}}^2 \right), \quad (2.33)$$

where $\mathbf{r}^0$ is the prior residual and $\mathbf{r}_{ij}^{\text{IMU}}$, $\mathbf{r}_{ij}^{\text{DVL}}$, and $\mathbf{r}_{ij}^P$ are the per-interval IMU, DVL, and pressure residuals specified in Section 2.2.

This cost cannot be minimized directly with standard unconstrained optimization methods as $\mathcal{X}$ is constrained: the extrinsic rotations lie in $SO(3)$ and the gravity direction in S^2 , which are manifolds embedded in Euclidean spaces. To optimize the cost on a manifold, we linearize it around the current estimate. There, the manifold is approximated by a flat (Euclidean) space tangent to it, the tangent space, in which small changes around the current rotation or direction behave as ordinary vectors. Each update is computed in this tangent space and then mapped back onto the manifold, so the constraints hold at every step.

Each manifold needs its own concrete map. For the rotations, the map back is the exponential map Exp . The tangent space of $SO(3)$ is three-dimensional, so the update is an axis-angle vector $\boldsymbol{\phi} \in \mathbb{R}^3$. The update is applied by right multiplication, $\mathbf{R}^{(m+1)} = \mathbf{R}^{(m)} \text{Exp}(\boldsymbol{\phi})$. Since the product of two rotation matrices is also a rotation matrix, $\mathbf{R}^{(m+1)}$ remains in $SO(3)$. For the gravity direction, the tangent space is two-dimensional, and an analogous retraction [32] maps a two-dimensional vector to an updated unit vector.

With these update operations in place, the cost is minimized iteratively by Gauss-Newton optimization. Each iteration linearizes the residuals around the current estimate. The linearization yields a quadratic approximation of the cost, whose minimizer solves the Gauss-Newton normal equations

$$\mathbf{J}^T \Sigma^{-1} \mathbf{J} \delta \mathbf{x} = -\mathbf{J}^T \Sigma^{-1} \mathbf{r}, \quad (2.34)$$

where $\mathbf{r}$ is the stacked residual, $\mathbf{J}$ its Jacobian with respect to the update $\delta \mathbf{x}$, and Σ the block-diagonal residual covariance. The update is applied through the exponential map for the rotation components, the corresponding unit-vector operation for the gravity component, and standard addition for the lever arms and biases. The procedure repeats until convergence. In practice, the optimization is implemented using Ceres Solver [1], which assembles the linear system from user-supplied residuals and handles constraint-preserving operations internally.

2.4 Summary

This chapter formalized extrinsic calibration of the underwater suite as an approximate MAP estimation problem over the inter-sensor extrinsics, the gravity direction, and the per-keyframe IMU biases. The camera trajectory enters as a fixed input rather than a state variable. The body pose at each keyframe is derived from the trajectory and the camera-IMU extrinsics rather than being estimated.

Each sensor contributes a residual that compares a measurement to its prediction from body motion over a keyframe interval: a depth difference from the pressure sensor; preintegrated rotation, velocity, and position residuals from the IMU; and a preintegrated displacement and a velocity residual from the DVL. The residual-based formulation integrates measurements from the IMU, DVL, and pressure sensors within a unified optimization framework, solved via Gauss-Newton iteration with tangent-space updates for the rotation and gravity-direction components.

The residual models and optimization reveal a structural limitation: several residuals depend on the body velocity ${}_{c_0}\mathbf{v}_{b_i}$ and the body angular velocity ${}_{b_i}\boldsymbol{\omega}_{b_i}$, yet neither quantity is measured directly nor provided by the reference trajectory. The question, therefore, is how these velocities should be represented and constrained within the calibration framework, which motivates our calibration formulations. To position our formulation, the next chapter examines the formulation adopted by the state-of-the-art AQUA-SLAM, which serves as the baseline for camera-IMU-DVL calibration but does not incorporate pressure-sensor measurements.

Baseline AQUA-SLAM Analysis

3

Chapter 2 established the calibration state and posed the weighted-least-squares formulation underlying the calibration estimate. This chapter focuses on AQUA-SLAM [7], the closest existing method to the joint underwater calibration addressed in the thesis. Section 3.1 specifies AQUA-SLAM’s calibration state, the velocity it supplies for the open body velocity, and the residuals that result. Section 3.2 identifies a lever arm term that this velocity supply omits and derives the bias it induces on the IMU–DVL lever arm in closed form. Section 3.3 specifies the simulation environment, trajectories, noise grades, and Monte Carlo protocol shared by all the synthetic experiments in this thesis. Section 3.4 validates the analytically predicted bias across four noise grades.

3.1 AQUA-SLAM’s Calibration Formulation

AQUA-SLAM [7] is a tightly coupled acoustic-visual-inertial SLAM system that fuses a camera, an IMU, and a DVL. It self-calibrates the extrinsics among these sensors during operation, without a calibration target, by reusing the residuals of its SLAM estimator. That reuse is the starting point for this chapter.

AQUA-SLAM calibrates opportunistically. The extrinsics are observable only when the vehicle’s motion excites them, so the system calibrates on a keyframe window that provides both a well-tracked camera trajectory and sufficient excitation to make the geometry observable. It estimates the inter-sensor extrinsics and the gravity direction together with the per-keyframe IMU biases,

$$\mathcal{X}_{\text{AQUA}} = \left\{ {}_b\mathbf{R}_c, {}_d\mathbf{R}_c, {}_b\mathbf{p}_{d/b}, {}_d\mathbf{p}_{c/d}, {}_{c_0}\mathbf{g}, \{ \mathbf{b}_i^g, \mathbf{b}_i^a \}_{i \in \mathcal{K}} \right\}, \quad (3.1)$$

corresponding to fourteen time-invariant degrees of freedom, including twelve in the extrinsics and two in the gravity direction, along with six additional parameters per keyframe in the biases. Its extrinsics cover the same inter-sensor geometry as the subset $\mathcal{E}$ in (2.1), except for the pressure-sensor lever arm, which is absent because AQUA-SLAM does not use the pressure sensor.

The extrinsics are estimated as the minimizer of the cost (2.33), built from the SLAM residuals of Chapter 2. Those residuals leave two quantities open: the navigation-frame body velocity ${}_{c_0}\mathbf{v}_{b_i}$, on which the IMU velocity depends, and position residuals, and the body angular velocity ${}_{b_i}\boldsymbol{\omega}_{b_i}$, on which the DVL velocity residual depends. AQUA-SLAM supplies ${}_{c_0}\mathbf{v}_{b_i}$ from the DVL velocity ${}_{d_i}\mathbf{v}_{d_i}$. In SLAM mode, ${}_{d_i}\mathbf{v}_{d_i}$ is an estimated state, and the body-frame body velocity ${}_{b_i}\mathbf{v}_{b_i}$ follows from it by rotation alone,

$$({}_{b_i}\mathbf{v}_{b_i})_{\text{AQUA}} = {}_b\mathbf{R}_d {}_{d_i}\mathbf{v}_{d_i}, \quad (3.2)$$

where ${}_b\mathbf{R}_d = {}_b\mathbf{R}_c {}_c\mathbf{R}_d$ is the DVL-to-body rotation. This implicitly assumes that the DVL and IMU origins share a single velocity up to rigid-body rotation. This modeling choice is central to AQUA-SLAM whose implications are analyzed in Section 3.2.

In calibration mode, ${}_{d_i}\mathbf{v}_{d_i}$ is not estimated, but AQUA-SLAM fixes it to the raw measurement,

$${}_{d_i}\mathbf{v}_{d_i} \leftarrow {}_{d_i}\tilde{\mathbf{v}}_{d_i}, \quad (3.3)$$

so ${}_b\mathbf{v}_{b_i}$ is determined entirely by the DVL reading and the extrinsics. Carrying the rotation-only transport into the navigation frame is an exact frame change, so the approximation in (3.2) passes through unchanged:

$$\begin{aligned} ({}_{c_0}\mathbf{v}_{b_i})_{\text{AQUA}} &= {}_{c_0}\mathbf{R}_{c_i} {}_b\mathbf{R}_c^T ({}_b\mathbf{v}_{b_i})_{\text{AQUA}} = {}_{c_0}\mathbf{R}_{c_i} {}_b\mathbf{R}_c^T {}_b\mathbf{R}_d {}_{d_i}\tilde{\mathbf{v}}_{d_i} \\ &= {}_{c_0}\mathbf{R}_{c_i} {}_c\mathbf{R}_d {}_{d_i}\tilde{\mathbf{v}}_{d_i}, \end{aligned} \quad (3.4)$$

where the camera-to-body rotation cancels, ${}_b\mathbf{R}_c^T {}_b\mathbf{R}_d = {}_c\mathbf{R}_d$. This approximate $({}_{c_0}\mathbf{v}_{b_i})_{\text{AQUA}}$ replaces the true body velocity ${}_{c_0}\mathbf{v}_{b_i}$ in the IMU residuals of Chapter 2.

Substituting the approximate velocity (3.4) into the IMU velocity (2.21) and position (2.22) residuals gives AQUA-SLAM's IMU velocity and position residuals,

$$\begin{aligned} (\mathbf{r}_{ij}^{\text{IMU},v})_{\text{AQUA}} &= {}_b\mathbf{R}_c {}_{c_0}\mathbf{R}_{c_i}^T ({}_{c_0}\mathbf{R}_{c_j} {}_c\mathbf{R}_d {}_{d_j}\tilde{\mathbf{v}}_{d_j} - {}_{c_0}\mathbf{R}_{c_i} {}_c\mathbf{R}_d {}_{d_i}\tilde{\mathbf{v}}_{d_i} - g {}_{c_0}\mathbf{g} \Delta t_{ij}) \\ &\quad - \Delta \tilde{\mathbf{v}}_{b_j/b_i}(\mathbf{b}_i^g, \mathbf{b}_i^a), \end{aligned} \quad (3.5)$$

$$\begin{aligned} (\mathbf{r}_{ij}^{\text{IMU},p})_{\text{AQUA}} &= {}_b\mathbf{R}_c {}_{c_0}\mathbf{R}_{c_i}^T ({}_{c_0}\mathbf{p}_{c_j} - {}_{c_0}\mathbf{p}_{c_i} - ({}_{c_0}\mathbf{R}_{c_j} - {}_{c_0}\mathbf{R}_{c_i}) {}_b\mathbf{R}_c^T {}_b\mathbf{p}_{c/b} \\ &\quad - {}_{c_0}\mathbf{R}_{c_i} {}_c\mathbf{R}_d {}_{d_i}\tilde{\mathbf{v}}_{d_i} \Delta t_{ij} - \frac{1}{2} g {}_{c_0}\mathbf{g} \Delta t_{ij}^2) - \Delta \tilde{\mathbf{p}}_{b_j/b_i}(\mathbf{b}_i^g, \mathbf{b}_i^a). \end{aligned} \quad (3.6)$$

The gyroscope residual (2.20) and the DVL translation residual (2.29) carry neither ${}_{c_0}\mathbf{v}_{b_i}$ nor ${}_{b_i}\boldsymbol{\omega}_{b_i}$ and are unchanged. The DVL velocity residual (2.30) is dropped: with ${}_{d_i}\mathbf{v}_{d_i}$ fixed rather than estimated, it has no role, and the angular velocity ${}_{b_i}\boldsymbol{\omega}_{b_i}$ it depends on no longer enters the calibration.

The IMU-camera lever arm ${}_b\mathbf{p}_{c/b}$ appears only in the IMU translation residual (3.6). Since it is not among AQUA-SLAM's state variables, the rigid-body constraint decomposes it as

$${}_b\mathbf{p}_{c/b} = {}_b\mathbf{p}_{d/b} - {}_b\mathbf{R}_c {}_c\mathbf{p}_{d/c}, \quad (3.7)$$

through which the IMU-DVL lever arm ${}_b\mathbf{p}_{d/b}$ enters the calibration. No other residual provides an independent constraint on the IMU-DVL lever arm.

When that residual informs the lever arm depends on the order of AQUA-SLAM's initialization. AQUA-SLAM does not minimize these residuals jointly from the start. It proceeds in five stages. A vision-only bundle adjustment first fixes the camera trajectory. The next two stages use only the gyroscope and DVL translation residuals to initialize the rotations ${}_b\mathbf{R}_c$ and ${}_c\mathbf{R}_d$ and the camera-DVL lever arm ${}_c\mathbf{p}_{d/c}$, with the gyroscope biases added in the third stage. The fourth stage uses the IMU velocity and translation residuals to initialize the gravity direction. The IMU-DVL lever arm appears in the IMU translation residual at this stage, but it is held fixed while only the gravity direction is free, so it is not informed. The fifth stage refines all four residuals jointly over the full state. The IMU-DVL lever arm is informed only in this final stage,

after the rotations, the camera–DVL lever arm, and the gravity direction have already been initialized.

The estimated IMU–DVL lever arm, therefore, depends entirely on the IMU translation residual. Section 3.2 examines that dependence and derives the structural bias the residual imposes on the lever arm, arising from the implicit assumption in (3.2).

3.2 IMU–DVL Lever Arm Bias in AQUA-SLAM

The velocity transport (3.2) neglects the rigid-body coupling between translation and rotation by omitting the lever arm term. The body and DVL origins are two points on the same rigid body, separated by the IMU–DVL lever arm ${}_b\mathbf{p}_{d/b}$. Therefore, their velocities in the body frame are related by

$${}_b\mathbf{v}_{b_i} = {}_b\mathbf{R}_{d\ d_i}\mathbf{v}_{d_i} - {}_b\boldsymbol{\omega}_{b_i} \times {}_b\mathbf{p}_{d/b}, \quad (3.8)$$

with ${}_b\boldsymbol{\omega}_{b_i} \in \mathbb{R}^3$ the body angular velocity at keyframe i .

This omission is present in AQUA-SLAM in both SLAM and calibration modes, but its consequences differ between them. In SLAM mode, the DVL-frame velocity ${}_{d_i}\mathbf{v}_{d_i}$ is a free state, constrained by the DVL velocity residual, so the missing cross-product is absorbed into its estimate. In calibration mode, the DVL velocity residual is dropped, and the substitution (3.3) pins ${}_{d_i}\mathbf{v}_{d_i}$ to the raw measurement. The omitted term then has nowhere to be absorbed and propagates into the lever arm estimate. To see this clearly, we first consider an idealized setting, specified by the two assumptions stated next.

Assumption 3.1. *The IMU–DVL lever arm ${}_b\mathbf{p}_{d/b}$ is the only unknown. The rotations ${}_b\mathbf{R}_c$ and ${}_c\mathbf{R}_d$, the camera–DVL lever arm ${}_c\mathbf{p}_{d/c}$, the gravity direction ${}_{c_0}\mathbf{g}$, and the per-keyframe biases are fixed at their true values. The camera trajectory is given, and the gyroscope, accelerometer, and DVL measurements are noiseless.*

Assumption 3.2. *The body angular velocities $\{{}_b\boldsymbol{\omega}_{b_i}\}$ over the keyframe set span more than one dimension (the excitation condition). The angular velocity is twice differentiable in time.*

Under these assumptions, the following theorem states that the estimate does not merely shift away from the truth: it collapses to zero, regardless of how large the true lever arm is.

Theorem 3.1 (lever arm collapse). *Under Assumptions 3.1 and 3.2, the IMU–DVL lever arm enters the cost only through AQUA-SLAM’s IMU translation residual (3.6), and its least-squares estimate*

$${}_b\hat{\mathbf{p}}_{d/b} = \arg \min_{{}_b\mathbf{p}_{d/b}} \sum_{ij} \|(\mathbf{r}_{ij}^{\text{IMU},p})_{\text{AQUA}}\|^2 \quad (3.9)$$

satisfies

$${}_b\hat{\mathbf{p}}_{d/b} = \mathcal{O}(\theta) {}_b\mathbf{p}_{d/b}^*, \quad (3.10)$$

where ${}_b\mathbf{p}_{d/b}^\star$ is the true lever arm and $\theta = \max_{ij} \theta_{ij}$ is the largest inter-keyframe rotation angle, with $\theta_{ij} \in [0, \pi]$ the rotation angle of the relative body rotation ${}_b\mathbf{R}_{b_j}$ between keyframes i and j . In the small-angle regime $\theta \ll 1$, the estimate is a fraction $\mathcal{O}(\theta)$ of the truth, collapsed to near zero regardless of its true value.

Proof. Write $\mathbf{r}_{ij}({}_b\mathbf{p}_{d/b}) \equiv (\mathbf{r}_{ij}^{\text{IMU},p})_{\text{AQUA}}$, so the estimator (3.9) minimizes $J({}_b\mathbf{p}_{d/b}) = \sum_{ij} \|\mathbf{r}_{ij}({}_b\mathbf{p}_{d/b})\|^2$. The argument has two parts: an exact reduction of the minimizer to the residual at the true lever arm, and a small-angle expansion that sizes the result.

Exact reduction. The lever arm enters the residual only through the rigid-body link (3.7), which appears linearly in AQUA-SLAM's IMU translation residual (3.6). The residual is therefore affine in the lever arm, with a constant Jacobian

$$\bar{\mathbf{A}}_{ij} = -{}_b\mathbf{R}_{c_{c_0}} \mathbf{R}_{c_i}^T ({}_c\mathbf{R}_{c_j} - {}_c\mathbf{R}_{c_i}) {}_b\mathbf{R}_c^T = \mathbf{I} - {}_b\mathbf{R}_{b_j}, \quad (3.11)$$

using ${}_c\mathbf{R}_{c_i}^T {}_c\mathbf{R}_{c_j} = {}_c\mathbf{R}_{c_j}$ and ${}_b\mathbf{R}_{c_{c_i}} \mathbf{R}_{c_j} {}_b\mathbf{R}_c^T = {}_b\mathbf{R}_{b_j}$. Expanding the affine residual about the true lever arm gives

$$\mathbf{r}_{ij}({}_b\mathbf{p}_{d/b}) = \mathbf{r}_{ij}({}_b\mathbf{p}_{d/b}^\star) + \bar{\mathbf{A}}_{ij} ({}_b\mathbf{p}_{d/b} - {}_b\mathbf{p}_{d/b}^\star). \quad (3.12)$$

This residual does not vanish at the true lever arm. Subtracting the kinematically consistent residual $\mathbf{r}_{ij}^{\text{kin}}$ of (2.22), which vanishes at the truth for noiseless sensors, the two differ only through the omitted transport term, so

$$\begin{aligned} \mathbf{r}_{ij}({}_b\mathbf{p}_{d/b}^\star) &= \mathbf{r}_{ij}({}_b\mathbf{p}_{d/b}^\star) - \mathbf{r}_{ij}^{\text{kin}}({}_b\mathbf{p}_{d/b}^\star) \\ &= {}_b\mathbf{R}_{c_{c_0}} \mathbf{R}_{c_i}^T ({}_c\mathbf{v}_{b_i} - ({}_c\mathbf{v}_{b_i})_{\text{AQUA}}) \Delta t_{ij} \\ &= ({}_b\mathbf{v}_{b_i} - ({}_b\mathbf{v}_{b_i})_{\text{AQUA}}) \Delta t_{ij} = -[{}_b\boldsymbol{\omega}_{b_i}]_{\times} {}_b\mathbf{p}_{d/b}^\star \Delta t_{ij}. \end{aligned} \quad (3.13)$$

where $[\cdot]_{\times} : \mathbb{R}^3 \rightarrow \mathbb{R}^{3 \times 3}$ is the skew-symmetric matrix operator satisfying $[\mathbf{y}]_{\times} \mathbf{z} = \mathbf{y} \times \mathbf{z}$ for any $\mathbf{y}, \mathbf{z} \in \mathbb{R}^3$. This result is exact. The residual at the truth equals the omitted transport correction over the interval, and this nonzero offset is the source of the bias.

The cost is quadratic in the lever arm, so the minimizer sets its gradient to zero. With the affine form (3.12) this is the linear system

$$\left(\sum_{ij} \bar{\mathbf{A}}_{ij}^T \bar{\mathbf{A}}_{ij} \right) ({}_b\hat{\mathbf{p}}_{d/b} - {}_b\mathbf{p}_{d/b}^\star) = - \sum_{ij} \bar{\mathbf{A}}_{ij}^T \mathbf{r}_{ij}({}_b\mathbf{p}_{d/b}^\star).$$

The Hessian $\sum_{ij} \bar{\mathbf{A}}_{ij}^T \bar{\mathbf{A}}_{ij}$ is positive definite under the excitation condition, hence invertible, so

$${}_b\hat{\mathbf{p}}_{d/b} - {}_b\mathbf{p}_{d/b}^\star = - \left(\sum_{ij} \bar{\mathbf{A}}_{ij}^T \bar{\mathbf{A}}_{ij} \right)^{-1} \sum_{ij} \bar{\mathbf{A}}_{ij}^T \mathbf{r}_{ij}({}_b\mathbf{p}_{d/b}^\star). \quad (3.14)$$

To expose the collapse, refer this offset to the zero lever arm. Setting ${}_b\mathbf{p}_{d/b} = \mathbf{0}$ in the affine form (3.12) defines the residual at zero,

$$\mathbf{r}_{ij}(\mathbf{0}) = \mathbf{r}_{ij}({}_b\mathbf{p}_{d/b}^\star) - \bar{\mathbf{A}}_{ij} {}_b\mathbf{p}_{d/b}^\star, \quad \text{equivalently} \quad \mathbf{r}_{ij}({}_b\mathbf{p}_{d/b}^\star) = \bar{\mathbf{A}}_{ij} {}_b\mathbf{p}_{d/b}^\star + \mathbf{r}_{ij}(\mathbf{0}). \quad (3.15)$$

Substituting the right-hand form of (3.15) into the minimizer error (3.14),

$$\begin{aligned} {}_b\hat{\mathbf{p}}_{d/b} - {}_b\mathbf{p}_{d/b}^* &= -\left(\sum_{ij} \bar{\mathbf{A}}_{ij}^T \bar{\mathbf{A}}_{ij}\right)^{-1} \sum_{ij} \bar{\mathbf{A}}_{ij}^T (\bar{\mathbf{A}}_{ij} {}_b\mathbf{p}_{d/b}^* + \mathbf{r}_{ij}(\mathbf{0})) \\ &= -{}_b\mathbf{p}_{d/b}^* - \left(\sum_{ij} \bar{\mathbf{A}}_{ij}^T \bar{\mathbf{A}}_{ij}\right)^{-1} \sum_{ij} \bar{\mathbf{A}}_{ij}^T \mathbf{r}_{ij}(\mathbf{0}). \end{aligned}$$

The true lever arm cancels against the $-{}_b\mathbf{p}_{d/b}^*$ on the left, leaving

$${}_b\hat{\mathbf{p}}_{d/b} = -\left(\sum_{ij} \bar{\mathbf{A}}_{ij}^T \bar{\mathbf{A}}_{ij}\right)^{-1} \sum_{ij} \bar{\mathbf{A}}_{ij}^T \mathbf{r}_{ij}(\mathbf{0}). \quad (3.16)$$

The estimate itself is determined by the residual at the zero lever arm point. The reduction to this point is exact.

Small-angle sizing. For the smooth, slow underwater motion and keyframe spacing $\Delta t_{ij} \lesssim 0.15$ s used here, the inter-keyframe angle is small, $\theta_{ij} \ll 1$. Assumption 3.2 makes the angular velocity twice differentiable, so the relative rotation in (3.11) expands about $\Delta t_{ij} = 0$ with a bounded remainder,

$${}_b\mathbf{R}_{b_j} = \mathbf{I} + [{}_b\boldsymbol{\omega}_{b_i}]_{\times} \Delta t_{ij} + \frac{1}{2}([{}_b\boldsymbol{\omega}_{b_i}]_{\times}^2 + [{}_b\dot{\boldsymbol{\omega}}_{b_i}]_{\times}) \Delta t_{ij}^2 + \mathcal{O}(\theta_{ij}^3),$$

with ${}_b\dot{\boldsymbol{\omega}}_{b_i} \in \mathbb{R}^3$ the body angular acceleration at keyframe i . The Jacobian is then

$$\bar{\mathbf{A}}_{ij} = -[{}_b\boldsymbol{\omega}_{b_i}]_{\times} \Delta t_{ij} - \frac{1}{2}([{}_b\boldsymbol{\omega}_{b_i}]_{\times}^2 + [{}_b\dot{\boldsymbol{\omega}}_{b_i}]_{\times}) \Delta t_{ij}^2 + \mathcal{O}(\theta_{ij}^3), \quad (3.17)$$

with leading term $\mathcal{O}(\theta_{ij})$. Substituting the residual at the truth (3.13) and the Jacobian expansion (3.17) into the residual at zero (3.15), the first-order terms in Δt_{ij} cancel,

$$\mathbf{r}_{ij}(\mathbf{0}) = \frac{1}{2}([{}_b\boldsymbol{\omega}_{b_i}]_{\times}^2 + [{}_b\dot{\boldsymbol{\omega}}_{b_i}]_{\times}) {}_b\mathbf{p}_{d/b}^* \Delta t_{ij}^2 + \mathcal{O}(\theta_{ij}^3) {}_b\mathbf{p}_{d/b}^* = \mathcal{O}(\theta_{ij}^2) {}_b\mathbf{p}_{d/b}^*. \quad (3.18)$$

The residual at zero is one order below the residual at the truth.

The estimate (3.16) divides the accumulated numerator by the Hessian, and the two enter at different orders. The leading Jacobian gives the per-pair Hessian, using $[{}_b\boldsymbol{\omega}_{b_i}]_{\times}^T = -[{}_b\boldsymbol{\omega}_{b_i}]_{\times}$,

$$\bar{\mathbf{A}}_{ij}^T \bar{\mathbf{A}}_{ij} = -[{}_b\boldsymbol{\omega}_{b_i}]_{\times}^2 \Delta t_{ij}^2 + \mathcal{O}(\theta_{ij}^3) = \mathcal{O}(\theta_{ij}^2),$$

so the summed Hessian is $\mathcal{O}(\theta^2)$ and, under the excitation condition, its inverse is $\mathcal{O}(\theta^{-2})$. The numerator pairs the leading Jacobian transpose with the residual at zero (3.18),

$$\bar{\mathbf{A}}_{ij}^T \mathbf{r}_{ij}(\mathbf{0}) = \frac{1}{2} [{}_b\boldsymbol{\omega}_{b_i}]_{\times} ([{}_b\boldsymbol{\omega}_{b_i}]_{\times}^2 + [{}_b\dot{\boldsymbol{\omega}}_{b_i}]_{\times}) {}_b\mathbf{p}_{d/b}^* \Delta t_{ij}^3 + \mathcal{O}(\theta_{ij}^4) {}_b\mathbf{p}_{d/b}^* = \mathcal{O}(\theta_{ij}^3) {}_b\mathbf{p}_{d/b}^*.$$

Dividing the numerator by the Hessian,

$${}_b\hat{\mathbf{p}}_{d/b} = \mathcal{O}(\theta^{-2}) \cdot \mathcal{O}(\theta^3) {}_b\mathbf{p}_{d/b}^* = \mathcal{O}(\theta) {}_b\mathbf{p}_{d/b}^*,$$

which is (3.10) and with this the Theorem is established. $\square$

Discussion

In the noiseless setting, the estimate is exactly a fixed matrix times the true lever arm, ${}_b\hat{\mathbf{p}}_{d/b} = \mathbf{M} {}_b\mathbf{p}_{d/b}^*$. This is followed by composing two steps from the proof. The estimate (3.16) is a fixed linear combination of the per-pair residuals at zero $\mathbf{r}_{ij}(\mathbf{0})$. Each $\mathbf{r}_{ij}(\mathbf{0})$ is in turn linear in the true lever arm. The matrix $\mathbf{M}$ is constructed solely from the keyframe angular velocities and timing, so it does not depend on the lever arm it multiplies. The $\mathcal{O}(\theta)$ factor in (3.10) is the size of $\mathbf{M}$. A consistent estimator would have $\mathbf{M} = \mathbf{I}$ and return the true lever arm exactly. AQUA-SLAM’s $\mathbf{M}$ is instead near-zero, so the estimate returns essentially zero for any true lever arm. The error is multiplicative, not additive: it scales with the true lever arm, and no fixed offset removes it.

The collapse has a direct geometric interpretation. The omitted cross-product ${}_{b_i}\boldsymbol{\omega}_{b_i} \times {}_b\mathbf{p}_{d/b}$ is the velocity difference a nonzero lever arm produces between the IMU and DVL origins. Dropping it makes AQUA-SLAM’s velocity transport (3.2) behave as if the origins were co-located, ${}_b\mathbf{p}_{d/b} = \mathbf{0}$. This is what leaves the residual nonzero at the true lever arm in (3.13). The cost removes the inconsistency by driving the estimate to the value at which co-location holds, ${}_b\mathbf{p}_{d/b} = \mathbf{0}$. The collapse is the optimizer enforcing the model’s built-in co-location assumption, not a statement about the true geometry.

The bias is therefore structural, not statistical. It arises from a deterministic offset in the residual at the truth, persists in the noiseless limit, and does not scale with the sensor noise level. Under the excitation condition, its magnitude is approximately $\|{}_b\mathbf{p}_{d/b}^*\|$, since the estimate collapses to essentially zero and the error is then the full lever arm.

The collapse is conditional on the excitation that makes the lever arm observable. The per-pair Hessian contribution $-[{}_{b_i}\boldsymbol{\omega}_{b_i}]_{\times}^2 \Delta t_{ij}^2$ has range orthogonal to ${}_{b_i}\boldsymbol{\omega}_{b_i}$, so a single-axis rotation profile leaves the summed Hessian rank-deficient and the lever arm unobservable. This is a distinct failure mode from the excited collapse analyzed here. The order bound in (3.10) carries a constant set by the conditioning of this Hessian, so it describes the collapse only in the excited regime, away from the single-axis degeneracy.

One simplification in the proof does not affect the conclusion. The proof uses unweighted least squares, while the chapter cost (2.33) is weighted. Positive-definite weights that are $\mathcal{O}(1)$ in Δt_{ij} enter the Hessian and the gradient at the same order per pair, leaving the leading-order structure intact. The reduction is therefore a legitimate simplification, not a gap.

3.3 Experimental Setup

Section 3.2 produced a closed-form prediction of the lever arm collapse in the idealized noise-free setting. The result isolates the mechanism responsible for the bias, but it relies on simplifying assumptions, including exact measurements and a reduced parameter set. The remainder of the chapter examines how this behavior manifests under realistic conditions.

This section specifies the simulation environment used for that evaluation. The synthetic experiments in this thesis share a single setup, specified here once to avoid

Table 3.1: Canonical ground truth extrinsic set used throughout the synthetic experiments: IMU-DVL lever arm ${}_b\mathbf{p}_{d/b}^*$, camera-DVL lever arm ${}_c\mathbf{p}_{d/c}^*$, camera-pressure lever arm ${}_c\mathbf{p}_{p/c}^*$, IMU-camera rotation ${}_b\mathbf{R}_c^*$, and camera-DVL rotation ${}_c\mathbf{R}_d^*$. The two rotation matrices are dimensionless axis permutations; see the text.

Quantity	Value
${}_b\mathbf{p}_{d/b}^*$	$[-0.183, -0.065, +0.295]$ m
${}_c\mathbf{p}_{d/c}^*$	$[+0.115, +0.390, -0.600]$ m
${}_c\mathbf{p}_{p/c}^*$	$[+0.130, +0.305, -0.460]$ m
${}_b\mathbf{R}_c^*$	$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$
${}_c\mathbf{R}_d^*$	$\begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{bmatrix}$

repetition. All simulation data are generated using the HoloOcean simulator [33].

3.3.1 Simulation Environment

The simulator drives a vehicle through commanded linear and angular accelerations and emits IMU, DVL, pressure, and stereo-camera measurements consistent with the resulting motion. The vehicle trajectory is therefore an output of the simulation rather than a prescribed input. Sensor noise is added externally so that its magnitude can be controlled per experiment. The IMU runs at 200 Hz, the DVL at 5 Hz, the pressure sensor at 30 Hz, and keyframes are inserted at approximately 6 Hz.

The simulator provides ground truth at every timestep for the body pose, body velocity, sensor biases, and the full set of extrinsics. The ground truth extrinsic set used throughout the synthetic experiments is given in Table 3.1. The non-trivial entries of ${}_b\mathbf{R}_c^*$ and ${}_c\mathbf{R}_d^*$ reflect the differing native axis conventions of the IMU, camera, and DVL on the BlueROV (small, commercially available underwater robot); they are pure axis permutations rather than physical misalignments.

The camera trajectory produced by HoloOcean is used directly as the ground-truth visual estimate, matching the calibration setup in Chapter 2, where the camera trajectory is treated as a fixed input to the optimization.

3.3.2 Trajectories

The motion that drives the simulator comes from one of two trajectory families: a Calibration Trajectory built to excite the calibration, and a surrogate matched to the motion ranges of a real underwater sequence.

Calibration Trajectory. The Calibration Trajectory is a HoloOcean BlueROV run commanded through a sequenced per-axis acceleration profile. A full accelerate-

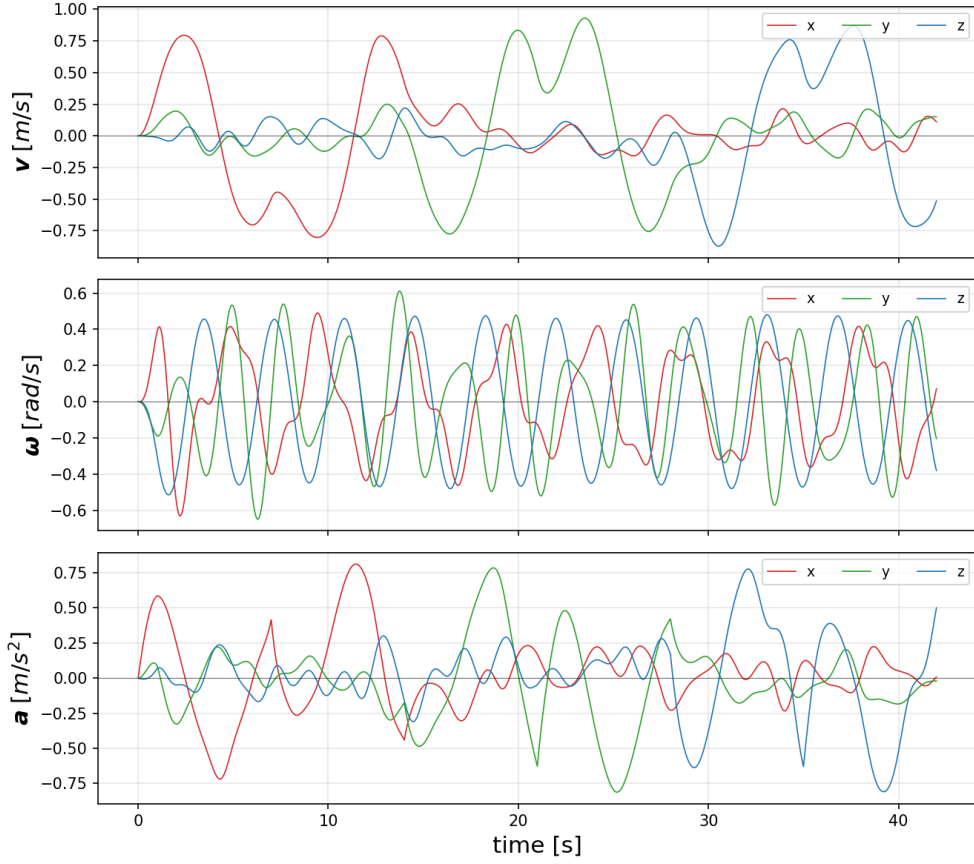


Figure 3.1: Calibration Trajectory (Calib Medium variant): noiseless body velocity $\mathbf{v}$ (top), angular velocity $\boldsymbol{\omega}$ (middle), and linear acceleration $\mathbf{a}$ (bottom) in the relevant sensor frame. The sequenced per-axis excitation of $\mathbf{v}$ is visible in the top panel; the angular wiggle is persistent on all three axes.

decelerate command is applied along the body x axis, then y , then z , while a low-frequency sinusoidal angular acceleration is commanded continuously on all three rotational axes. The full run lasts 42 seconds. Three rotational-excitation variants are used, denoted Calib Slow, Calib Medium, and Calib Fast, obtained by amplitude scaling the commanded angular acceleration at a fixed time base; the resulting peak angular velocity is $\omega_{\text{peak}} \in \{0.25, 0.6, 1.0\}$ rad/s, respectively. The translational command is identical across the three variants. Calib Medium is the canonical instance shown in Figure 3.1: the resulting body-velocity peaks reach approximately 0.85 m/s on each axis during its active phase, angular-velocity peaks reach approximately 0.6 rad/s persistently on all three axes, and linear-acceleration peaks reach approximately 0.8 m/s².

Structure Easy surrogate. The Structure Easy sequence (henceforth SE) is a real underwater data sequence from the Tank dataset [29], recorded on the sensor hardware described in Section 3.3.3. The SE surrogate is a HoloOcean run designed so that its resulting motion spans the same per-axis angular-velocity and linear-acceleration

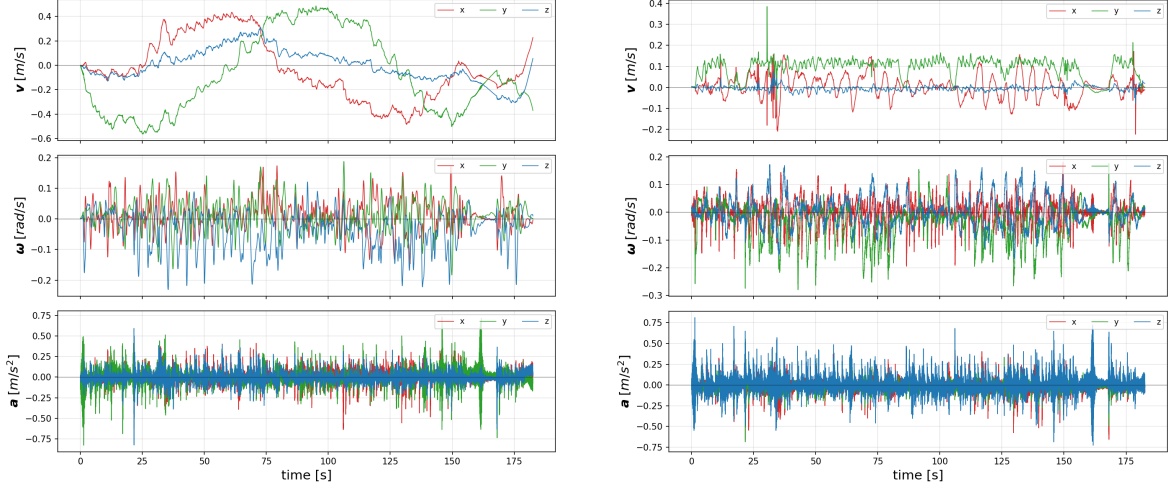


Figure 3.2: The SE surrogate (left) alongside the real SE sequence (right). Each panel shows body velocity $\mathbf{v}$ (top), angular velocity $\boldsymbol{\omega}$ (middle), and linear acceleration $\mathbf{a}$ (bottom) in the relevant sensor frame. Both share the all-axis low-amplitude rotation character. The surrogate spans a larger translational range and depth of descent.

ranges as SE: persistent low-amplitude angular velocity on all three axes with per-axis root-mean-square (RMS) around 0.05–0.07 rad/s, and linear acceleration with per-axis RMS of order 0.07–0.10 m/s². Figure 3.2 places the surrogate alongside SE for the same motion channels.

The surrogate matches SE on rotational and acceleration ranges, but is not a sample-by-sample replay and does not match SE on translational scale: it traverses a larger path length and descends through a substantially larger depth range than SE itself.

3.3.3 Noise Grades

Where a trajectory fixes the motion, a noise grade quantifies how strongly the sensors corrupt their measurements of that motion. A noise grade is a coupled configuration of the IMU, DVL, and depth white-noise levels, the IMU bias random-walk coefficients, and the turn-on bias priors. Four grades, G1 through G4, define a noise sweep on a fixed sensor suite: the MicroStrain 3DM-GX5-AHRS, the Water Linked A50, and the Blue Robotics Bar30. Discrete-time per-sample standard deviations follow from continuous-time noise densities σ_c and bias random-walk coefficients σ_c^b through $\sigma = \sigma_c \sqrt{f_s}$ and $\sigma^b = \sigma_c^b / \sqrt{f_s}$ at the rates of Section 3.3.1 where f_s indicates the rate of the sensor. The continuous-time units are rad/s/ $\sqrt{\text{Hz}}$ for the gyroscope, m/s²/ $\sqrt{\text{Hz}}$ for the accelerometer, and the corresponding 1/ $\sqrt{\text{Hz}}$ densities for the bias random walks.

G1 is the low-noise anchor, taken from the manufacturer’s datasheets. The gyroscope and accelerometer white-noise densities and the IMU bias random-walk coefficients are the GX5-AHRS datasheet figures, and the pressure standard deviation is the MS5837-30BA datasheet RMS at oversampling ratio 8192, namely 2 mm. The DVL white-noise level is the one G1 entry not given by a datasheet; it is obtained from the error-velocity analysis of the Structure Easy sequence in Appendix A. G3 is the upper

anchor, taken directly from the configuration files of the AQUA-SLAM’s open-source codebase [34] released alongside the published system [7], which deploys this hardware on the Tank dataset [29] in real underwater operation. These configuration values are from the AQUA-SLAM codebase which sets $\sigma_{\text{DVL}} = 0.1$ m/s, roughly $26\times$ the G1 value.

G2 and G4 fill in and extend the sweep around these anchors. G2 is the per-channel geometric midpoint between G1 and G3; the G1-to-G3 inflation is non-uniform across channels (approximately $14\times$ on the gyroscope, $4\times$ on the accelerometer, and $26\times$ on the DVL), so G2 inherits this asymmetry rather than applying a single factor. G4 represents shallow-water operation with beam-dropout-prone DVL conditions and vibration-coupled IMU readings. Its IMU white-noise rows are twice G3, while its DVL white-noise level is anchored at G1 and scaled by an operating-condition factor of approximately $24\times$, which places it below G3 by construction. The codebase σ_{DVL} absorbs other sources of error, so compounding G4 on top of G3 would conflate sensor noise with that slack.

The depth standard deviations at G2 through G4 reflect surface-wave-driven pressure fluctuations rather than sensor noise; they are stated estimates for typical AUV operating depths of 5 to 10 m. The turn-on bias rows give the at-power-on uncertainty, a fixed offset realized once when the sensor is powered on and held constant thereafter, distinct from the bias random walk that drifts during operation. Their dual role in the Monte Carlo protocol is described in Section 3.3.4.

Table 3.2: Discrete-time per-sample standard deviations for the four noise grades on the GX5-AHRS + A50 + Bar30 sensor suite. G1 is datasheet-anchored, with the DVL entry taken from the error-velocity analysis of Appendix A; G3 is the AQUA-SLAM codebase configuration; G2 and G4 are interpolated and extrapolated around them. The G4 DVL white-noise row sits below G3 by construction; see the text.

Parameter	G1	G2	G3	G4
<i>IMU white noise</i>				
σ_g [rad/s]	1.02×10^{-3}	3.80×10^{-3}	1.41×10^{-2}	2.83×10^{-2}
σ_a [m/s ²]	3.22×10^{-2}	6.75×10^{-2}	1.41×10^{-1}	2.83×10^{-1}
<i>IMU bias random walk</i>				
σ_{bg} [rad/s]	2.52×10^{-7}	4.22×10^{-7}	7.07×10^{-7}	1.41×10^{-6}
σ_{ba} [m/s ²]	4.74×10^{-7}	1.83×10^{-6}	7.07×10^{-6}	1.41×10^{-5}
<i>DVL</i>				
σ_{DVL} [m/s]	3.82×10^{-3}	1.95×10^{-2}	1.00×10^{-1}	9.17×10^{-2}
<i>Depth (wave-driven)</i>				
σ_p [m]	2.00×10^{-3}	5.00×10^{-3}	1.20×10^{-2}	2.40×10^{-2}
<i>Turn-on bias priors</i>				
$\sigma_{bg}^{\text{prior}}$ [rad/s]	4.36×10^{-3}	1.53×10^{-2}	5.23×10^{-2}	1.05×10^{-1}
$\sigma_{ba}^{\text{prior}}$ [m/s ²]	1.96×10^{-2}	6.86×10^{-2}	2.35×10^{-1}	4.71×10^{-1}

3.3.4 Monte Carlo Protocol

Together, a trajectory and a noise grade define an experimental cell. Every reported result in the synthetic evaluation is the aggregate of a Monte Carlo set over one such

cell: the set holds the ground truth trajectory and ground truth extrinsics constant and repeats the calibration $N = 200$ times. Each repetition introduces new sensor noise at the cell’s grade, new IMU biases, and a new random initialization of the extrinsics. The random number generator is deterministically seeded by the trial index, ensuring reproducibility for each cell.

The optimizer is initialized far from the truth, so convergence cannot be attributed to a good starting guess. The initial rotation is a uniformly random rotation, carrying no information about the truth orientation, and each initial lever arm is drawn uniformly over a range well outside the vehicle’s physical envelope. These magnitudes are a stress-test design point rather than an empirical convergence boundary.

The IMU biases follow a random walk. The initial bias is drawn from a zero-mean Gaussian with the turn-on standard deviation in Table 3.2, and is then propagated over time by the per-sample random-walk standard deviation from the same table. The optimizer mirrors this generative model. A prior at the turn-on standard deviation anchors the initial bias state, and random-walk residuals at the per-sample standard deviation link successive bias states. No prior is placed on the extrinsics. This asymmetry is intentional: the turn-on uncertainty is genuinely known a priori on real hardware, whereas the extrinsic initialization is a stress test that should not inform the optimizer.

The calibration is solved as a batch nonlinear least-squares problem with the Ceres Solver [1], and the converged extrinsic estimate is recorded.

For each cell, we summarize each extrinsic by two scalars, reported for most experiments. Let $\hat{\mathbf{x}}_\ell \in \mathbb{R}^d$ be the estimate of a given extrinsic at trial ℓ , for $\ell = 1, \dots, N$, and let $\mathbf{x}^\star \in \mathbb{R}^d$ be its truth. The per-trial error is $\mathbf{e}_\ell = \hat{\mathbf{x}}_\ell - \mathbf{x}^\star$. The mean estimate over the cell is $\hat{\mathbf{x}} = \frac{1}{N} \sum_{\ell=1}^N \hat{\mathbf{x}}_\ell$, and the mean error vector is $\bar{\mathbf{e}} = \hat{\mathbf{x}} - \mathbf{x}^\star$. The two reported scalars are the bias and the across-run standard deviation,

$$\text{bias} = \|\bar{\mathbf{e}}\|, \quad \text{std} = \sqrt{\frac{1}{N-1} \sum_{\ell=1}^N \|\mathbf{e}_\ell - \bar{\mathbf{e}}\|^2}. \quad (3.19)$$

The bias is the systematic offset that survives averaging across trials. The standard deviation is the across-run scatter of the trial errors about that mean.

For the rotations and the gravity direction, the per-trial error $\mathbf{e}_\ell$ is a tangent-space error taken per axis. The construction is otherwise unchanged. The rotations use $d = 3$, and the gravity direction uses $d = 2$, matching its two degrees of freedom.

3.4 Empirical Validation

Section 3.2 predicted the lever arm collapse in closed form, under two idealizations: a single unknown lever arm and noiseless sensors. This section relaxes both and tests whether the collapse survives across the four noise grades of the sweep, using the simulation setting detailed in Section 3.3.

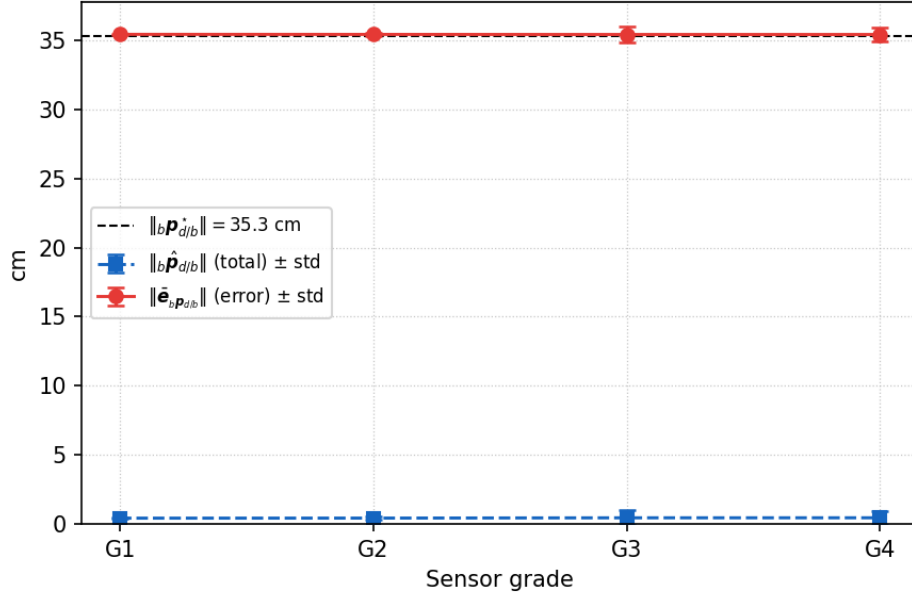


Figure 3.3: AQUA-SLAM IMU–DVL lever arm estimate magnitude $\|b \hat{\mathbf{p}}_{d/b}\|$ and error magnitude $\|\bar{\mathbf{e}}_{b \mathbf{p}_{d/b}}\| = \|b \hat{\mathbf{p}}_{d/b} - b \mathbf{p}_{d/b}^*\|$ across noise grades G1–G4 on the Calib Medium trajectory. Both quantities are means over $N = 200$ Monte Carlo trials, with error bars at ± 1 standard deviation. The dashed line marks the true magnitude $\|b \mathbf{p}_{d/b}^*\| = 35.3$ cm.

3.4.1 Protocol

The validation uses the shared setup of Section 3.3 over all four noise grades, with Calib Medium as the trajectory. Calib Medium is chosen because its angular excitation spans all three body axes, which the lever arm collapse requires to be observable. Each calibration runs through AQUA-SLAM’s coarse-to-fine pipeline in Section 3.1 and estimates the full state jointly, rather than a single lever arm of the theorem. A co-located configuration with $b \mathbf{p}_{d/b}^* = \mathbf{0}$ is run as a control, separating the bias of a nonzero lever arm from any implementation artifact.

3.4.2 Lever Arm Collapse

Figure 3.3 plots the IMU–DVL lever arm estimate magnitude and error magnitude across the four noise grades, against the true magnitude of 35.3 cm. The estimate collapses to near zero at every grade, with a mean magnitude of about 0.4 cm. The error magnitude is therefore approximately constant across grades, about 35.4 cm, marginally above the true magnitude. It changes by about 0.01 cm between G1 and G4, which is negligible against its magnitude. The two curves are complementary views of one fact: the estimate is driven to near zero, so the error settles at approximately the full lever arm magnitude, the collapse predicted by (3.10).

The Monte Carlo standard deviation, shown by the error bars on both curves, grows from sub-millimeter at G1 to about half a centimeter at G3. At G4, it does not increase further, because the DVL noise level sits below G3 by construction (Section 3.3). It

Table 3.3: Final-stage AQUA-SLAM Monte Carlo bias and standard deviation across noise grades G1–G4. Translation errors in cm; rotation and gravity errors in degrees. $N = 200$ trials per grade.

		Co-located				Realistic			
		G1	G2	G3	G4	G1	G2	G3	G4
${}_b\mathbf{p}_{d/b}$	bias [cm]	0.29	0.29	0.29	0.29	35.43	35.43	35.42	35.42
	std [cm]	0.03	0.11	0.52	0.49	0.03	0.11	0.55	0.52
${}_c\mathbf{p}_{d/c}$	bias [cm]	0.26	0.25	0.23	0.23	0.52	0.52	0.57	0.56
	std [cm]	0.11	0.58	2.98	2.73	0.11	0.58	2.98	2.73
${}_b\mathbf{R}_c$	bias [°]	0.036	0.036	0.036	0.035	0.037	0.037	0.036	0.036
	std [°]	0.002	0.009	0.034	0.068	0.002	0.009	0.034	0.068
${}_b\mathbf{R}_d$	bias [°]	0.047	0.059	0.173	0.161	0.185	0.174	0.134	0.135
	std [°]	0.055	0.278	1.424	1.306	0.055	0.278	1.423	1.309
${}_{c_0}\mathbf{g}$	bias [°]	0.019	0.024	0.045	0.042	0.713	0.732	0.688	0.692
	std [°]	0.022	0.112	0.569	0.526	0.022	0.109	0.553	0.511

remains more than an order of magnitude below the mean error across all grades. This scatter is the ordinary noise-driven variation about the collapsed estimate, and does not contradict the noise-independent bias prediction: the bias dominates the lever arm error throughout the sweep.

3.4.3 Secondary Effects and Co-located Control

Table 3.3 (realistic columns) reports the full extrinsic estimates. The IMU–DVL lever arm row is the same data as Figure 3.3 and is not restated here. The remaining extrinsics, which do not couple to the IMU–DVL lever arm through the velocity substitution, carry negligible bias. The camera–DVL lever arm ${}_c\mathbf{p}_{d/c}$ has bias below 0.6 cm at every grade, well over an order of magnitude below the IMU–DVL lever arm bias. The rotation extrinsics ${}_b\mathbf{R}_c$ and ${}_b\mathbf{R}_d$ have bias below 0.04° and 0.2° respectively. Their standard deviation grows with noise grade in the ordinary way, reaching about 3 cm for ${}_c\mathbf{p}_{d/c}$ and 1.4° for ${}_b\mathbf{R}_d$ at the highest grades, but this is noise-driven scatter and not a structural offset.

The gravity direction ${}_{c_0}\mathbf{g}$ exhibits a bias in the realistic configuration, unlike the other unaffected extrinsics. This bias arises from the coupling between gravity and the lever arm in AQUA-SLAM’s IMU translation residual. The predicted body translation contains a $\frac{1}{2}g {}_{c_0}\mathbf{g} \Delta t_{ij}^2$ gravity term alongside the velocity-substitution term that is structurally biased by the omitted lever arm correction. With the lever arm estimate driven toward zero, the residual offset at the truth is partially absorbed by adjusting ${}_{c_0}\mathbf{g}$. The observed gravity bias is therefore a secondary consequence of the primary IMU–DVL lever arm bias.

To confirm that the bias is caused specifically by the nonzero IMU–DVL lever arm, and not by an implementation artifact, the experiment is repeated with ${}_b\mathbf{p}_{d/b}^* = \mathbf{0}$, the

limit in which the IMU and DVL are co-located on the rigid body. The remaining setup is identical: the same trajectory, the same noise grades, and the same $N = 200$ trials per grade. Table 3.3 (co-located columns) shows the result. The IMU–DVL lever arm is recovered with sub-cm bias at every grade, about 0.29 cm across G1–G4, far below the realistic case’s ~ 35 cm bias. The gravity direction is recovered with bias below 0.05° across grades, well within the Monte Carlo scatter at G3–G4. This is consistent with zero structural bias, in contrast to the realistic case, where the gravity bias exceeds the noise at every grade. The other extrinsics are recovered with comparable accuracy to the realistic case. The implementation is therefore functionally correct: when there is no lever arm to bias, AQUA-SLAM’s calibration recovers the truth up to the noise-driven scatter common to both configurations. The bias documented here arises specifically from the velocity-substitution mechanism of Section 3.2 applied to a nonzero lever arm, which confirms its structural origin.

3.5 Summary

This chapter analyzed how AQUA-SLAM supplies the body velocity that its residuals leave open for calibration. AQUA-SLAM transports the DVL velocity to the body origin by rotation alone, omitting the cross-product ${}_{b_i}\boldsymbol{\omega}_{b_i} \times {}_b\boldsymbol{p}_{d/b}$ that a nonzero lever arm produces between the two origins.

This omission is benign in SLAM, where the DVL-frame velocity is a free state that absorbs the omitted term, but not in calibration, where the velocity is pinned to the raw measurement and the error has nowhere to go. Section 3.2 showed the consequence in closed form: AQUA-SLAM’s IMU translation residual does not vanish at the true lever arm. Since it is the sole constraint on the IMU–DVL lever arm, the estimate collapses toward zero regardless of the truth. The Monte Carlo study of Section 3.4 confirmed that the bias is structural rather than statistical: a near-constant error of the full lever arm magnitude across all four noise grades, independent of the noise level.

The remedy follows directly from the analysis. Restoring the omitted cross-product in the velocity supply yields a kinematically consistent model in which the IMU residuals are satisfied at the true state. Chapter 4 develops this correction as the kinematically-consistent formulation and tests whether it recovers the lever arm that the baseline drives to zero.

Formulation A: Lever-Arm-Corrected Velocity Model

4

This chapter develops Formulation A, an alternative kinematically consistent method to AQUA-SLAM for supplying the open body velocity, and evaluates whether it recovers the lever arm that the baseline drives to zero. Section 4.1 specifies Formulation A’s calibration state and the corrected velocity supply that distinguishes it from AQUA-SLAM. Section 4.2 carries that supply into the residuals, restating only the two that it changes and inheriting the rest from Chapter 2. Section 4.3 specifies the coarse-to-fine pipeline that solves the calibration. Section 4.4 evaluates Formulation A against AQUA-SLAM in simulation and reports the pressure-sensor lever-arm calibration it adds. Section 4.5 summarizes the chapter and isolates a residual noise-dependent bias that motivates our analysis in Chapter 5.

4.1 Calibration State and Velocity Supply

Formulation A estimates the calibration state

$$\mathcal{X} = \{ {}_b\mathbf{R}_c, {}_c\mathbf{R}_d, {}_c\mathbf{p}_{d/c}, {}_b\mathbf{p}_{d/b}, {}_c\mathbf{p}_{p/c}, {}_{c_0}\mathbf{g}, \{ \mathbf{b}_i^g, \mathbf{b}_i^a \}_{i \in \mathcal{K}} \}. \quad (4.1)$$

It represents the same inter-sensor geometry and gravity direction as AQUA-SLAM (Section 3.1), parameterizing the camera-DVL extrinsics in the camera frame rather than the DVL frame, and adds the pressure-sensor lever arm ${}_c\mathbf{p}_{p/c}$, which AQUA-SLAM does not estimate.

The body velocity ${}_{c_0}\mathbf{v}_{b_i}$ and angular velocity ${}_{b_i}\boldsymbol{\omega}_{b_i}$ are not part of the state. Formulation A supplies them from sensor samples. The angular velocity is the gyroscope sample with its bias removed, ${}_{b_i}\hat{\boldsymbol{\omega}}_{b_i} = {}_{b_i}\tilde{\boldsymbol{\omega}}_{b_i} - \mathbf{b}_i^g$; the body velocity is the DVL velocity transported to the body origin by the full rigid-body relation (3.8). Applying that relation to the measured samples and rotating into the navigation frame gives

$${}_{c_0}\mathbf{v}_{b_i} = {}_{c_0}\mathbf{R}_{c_i} {}_b\mathbf{R}_c^T ({}_b\mathbf{R}_{d_i} \tilde{\mathbf{v}}_{d_i} - {}_{b_i}\hat{\boldsymbol{\omega}}_{b_i} \times {}_b\mathbf{p}_{d/b}), \quad (4.2)$$

where ${}_{d_i}\tilde{\mathbf{v}}_{d_i}$ is the DVL sample at keyframe i and ${}_b\mathbf{R}_d = {}_b\mathbf{R}_c {}_c\mathbf{R}_d$.

This expression corresponds to AQUA-SLAM’s velocity supply equation (3.4) with the omitted cross-product correction restored. The correction contains the IMU–DVL lever arm ${}_b\mathbf{p}_{d/b}$, which therefore enters the IMU residuals through the constructed body velocity.

4.2 The Residuals under Formulation A

Using the corrected supply (4.2) affects only the IMU velocity and translation residuals. The rotation, DVL translation, and depth residuals carry no body velocity and are un-

changed from Chapter 2; the DVL velocity residual, which carried the angular velocity, is dropped. This section presents the two changed residuals and their covariances.

4.2.1 IMU Velocity and Translation Residuals

The velocity residual (2.21) depends on the body velocity at both keyframes. Substituting the velocity supply at each keyframe gives

$$\begin{aligned} \mathbf{r}_{ij}^{\text{IMU},v} = & {}_b\mathbf{R}_{c\ c_0}\mathbf{R}_{c_i}^T [{}_{c_0}\mathbf{R}_{c_j} {}_b\mathbf{R}_c^T ({}_b\mathbf{R}_{d\ d_j}\tilde{\mathbf{v}}_{d_j} - {}_{b_j}\hat{\boldsymbol{\omega}}_{b_j} \times {}_b\mathbf{p}_{d/b}) \\ & - {}_{c_0}\mathbf{R}_{c_i} {}_b\mathbf{R}_c^T ({}_b\mathbf{R}_{d\ d_i}\tilde{\mathbf{v}}_{d_i} - {}_{b_i}\hat{\boldsymbol{\omega}}_{b_i} \times {}_b\mathbf{p}_{d/b}) - g\ {}_{c_0}\mathbf{g}\ \Delta t_{ij}] - \Delta\tilde{\mathbf{v}}_{b_j/b_i}(\mathbf{b}_i^g, \mathbf{b}_i^a). \end{aligned} \quad (4.3)$$

The translation residual (2.22) carries the body velocity at keyframe i alone, integrated over the interval. Substituting the corrected supply gives

$$\begin{aligned} \mathbf{r}_{ij}^{\text{IMU},p} = & {}_b\mathbf{R}_{c\ c_0}\mathbf{R}_{c_i}^T [{}_{c_0}\mathbf{p}_{c_j} - {}_{c_0}\mathbf{p}_{c_i} - ({}_{c_0}\mathbf{R}_{c_j} - {}_{c_0}\mathbf{R}_{c_i}) {}_b\mathbf{R}_c^T {}_b\mathbf{p}_{c/b} \\ & - {}_{c_0}\mathbf{R}_{c_i} {}_b\mathbf{R}_{d\ d_i}\tilde{\mathbf{v}}_{d_i}\ \Delta t_{ij} + {}_{c_0}\mathbf{R}_{c_i} {}_b\mathbf{R}_c^T ({}_{b_i}\hat{\boldsymbol{\omega}}_{b_i} \times {}_b\mathbf{p}_{d/b})\ \Delta t_{ij} - \frac{1}{2} g\ {}_{c_0}\mathbf{g}\ \Delta t_{ij}^2] \\ & - \Delta\tilde{\mathbf{p}}_{b_j/b_i}(\mathbf{b}_i^g, \mathbf{b}_i^a). \end{aligned} \quad (4.4)$$

The velocity residual carries the cross-product correction at both keyframes, whereas the translation residual includes it only at keyframe i . The camera-IMU lever arm ${}_b\mathbf{p}_{c/b}$ is not a state variable; it expands through the rigid-body link (3.7) into ${}_{c_0}\mathbf{p}_{d/c}$ and ${}_b\mathbf{p}_{d/b}$.

4.2.2 Residual Covariances

The covariance models associated with the IMU residuals in Chapter 2 are presented in this section, as it is required to obtain statistically consistent weighting of heterogeneous sensor residuals in the cost function. The velocity and translation residuals are weighted separately, each by the inverse of its own covariance. Each covariance combines two parts: the marginal preintegration block, $\boldsymbol{\Sigma}_{\Delta\mathbf{v}_{b_j/b_i}}$ or $\boldsymbol{\Sigma}_{\Delta\mathbf{p}_{b_j/b_i}}$ (the velocity or position diagonal of the joint preintegration covariance [31]), and the noise the substitution injects.

Each residual is linear in its samples, so its noise propagates exactly through the residual Jacobian. The translation residual draws on the keyframe- i samples,

$$\frac{\partial \mathbf{r}_{ij}^{\text{IMU},p}}{\partial {}_{d_i}\tilde{\mathbf{v}}_{d_i}} = -{}_b\mathbf{R}_{c\ c_0}\mathbf{R}_{d\ d_i}\Delta t_{ij}, \quad \frac{\partial \mathbf{r}_{ij}^{\text{IMU},p}}{\partial {}_{b_i}\hat{\boldsymbol{\omega}}_{b_i}} = -[{}_b\mathbf{p}_{d/b}]_{\times}\Delta t_{ij}, \quad (4.5)$$

using $\mathbf{y} \times \mathbf{z} = -[\mathbf{z}]_{\times}\mathbf{y}$. Propagating the DVL and gyroscope sample covariances $\boldsymbol{\Sigma}_{\text{DVL}}$ and $\boldsymbol{\Sigma}_{\text{Gyro}}$ and adding the marginal preintegration block gives

$$\boldsymbol{\Sigma}_{\mathbf{r}_{ij}^{\text{IMU},p}}^{(\text{DVL})} = \Delta t_{ij}^2 {}_b\mathbf{R}_{c\ c_0}\mathbf{R}_{d\ d_i}\boldsymbol{\Sigma}_{\text{DVL}} {}_{c_0}\mathbf{R}_{d\ d_i}^T {}_b\mathbf{R}_c^T, \quad (4.6)$$

$$\boldsymbol{\Sigma}_{\mathbf{r}_{ij}^{\text{IMU},p}}^{(\text{Gyro})} = \Delta t_{ij}^2 [{}_b\mathbf{p}_{d/b}]_{\times}\boldsymbol{\Sigma}_{\text{Gyro}} [{}_b\mathbf{p}_{d/b}]_{\times}^T, \quad (4.7)$$

$$\boldsymbol{\Sigma}_{\mathbf{r}_{ij}^{\text{IMU},p}} = \boldsymbol{\Sigma}_{\Delta\mathbf{p}_{b_j/b_i}} + \boldsymbol{\Sigma}_{\mathbf{r}_{ij}^{\text{IMU},p}}^{(\text{DVL})} + \boldsymbol{\Sigma}_{\mathbf{r}_{ij}^{\text{IMU},p}}^{(\text{Gyro})}. \quad (4.8)$$

The velocity residual draws on both keyframes, the keyframe- j samples rotated through the relative body motion ${}_b\mathbf{R}_{b_j} = {}_b\mathbf{R}_{c\ c_0}\mathbf{R}_{c_i\ c_0}^T\mathbf{R}_{c_j\ b}\mathbf{R}_c^T$. Its Jacobians are

$$\begin{aligned}\frac{\partial \mathbf{r}_{ij}^{\text{IMU},v}}{\partial {}_{d_i}\tilde{\mathbf{v}}_{d_i}} &= -{}_b\mathbf{R}_{c\ c}\mathbf{R}_d, & \frac{\partial \mathbf{r}_{ij}^{\text{IMU},v}}{\partial {}_{d_j}\tilde{\mathbf{v}}_{d_j}} &= {}_{b_i}\mathbf{R}_{b_j\ b}\mathbf{R}_{c\ c}\mathbf{R}_d, \\ \frac{\partial \mathbf{r}_{ij}^{\text{IMU},v}}{\partial {}_{b_i}\hat{\boldsymbol{\omega}}_{b_i}} &= -[{}_b\mathbf{p}_{d/b}]_{\times}, & \frac{\partial \mathbf{r}_{ij}^{\text{IMU},v}}{\partial {}_{b_j}\hat{\boldsymbol{\omega}}_{b_j}} &= {}_{b_i}\mathbf{R}_{b_j} [{}_b\mathbf{p}_{d/b}]_{\times}.\end{aligned}\tag{4.9}$$

Propagation of the measurement noise gives

$$\begin{aligned}\boldsymbol{\Sigma}_{\mathbf{r}_{ij}}^{(\text{DVL})} &= {}_{b_i}\mathbf{R}_{b_j}\mathbf{R}_{c\ c}\mathbf{R}_d\boldsymbol{\Sigma}_{\text{DVL}}\mathbf{R}_d^T\mathbf{R}_{c\ b_i}\mathbf{R}_{b_j}^T \\ &\quad + {}_b\mathbf{R}_{c\ c}\mathbf{R}_d\boldsymbol{\Sigma}_{\text{DVL}}\mathbf{R}_d^T\mathbf{R}_c^T,\end{aligned}\tag{4.10}$$

$$\boldsymbol{\Sigma}_{\mathbf{r}_{ij}^{\text{IMU},v}}^{(\text{Gyro})} = {}_{b_i}\mathbf{R}_{b_j} [{}_b\mathbf{p}_{d/b}]_{\times}\boldsymbol{\Sigma}_{\text{Gyro}} [{}_b\mathbf{p}_{d/b}]_{\times}^T {}_{b_i}\mathbf{R}_{b_j}^T + [{}_b\mathbf{p}_{d/b}]_{\times}\boldsymbol{\Sigma}_{\text{Gyro}} [{}_b\mathbf{p}_{d/b}]_{\times}^T,\tag{4.11}$$

$$\boldsymbol{\Sigma}_{\mathbf{r}_{ij}^{\text{IMU},v}} = \boldsymbol{\Sigma}_{\Delta\mathbf{v}_{b_j/b_i}} + \boldsymbol{\Sigma}_{\mathbf{r}_{ij}^{\text{IMU},v}}^{(\text{DVL})} + \boldsymbol{\Sigma}_{\mathbf{r}_{ij}^{\text{IMU},v}}^{(\text{Gyro})}.\tag{4.12}$$

Weighting the two residuals separately is a block-diagonal approximation of their joint covariance: the diagonal blocks are kept, and the coupling between the velocity and translation residuals is discarded. That coupling is not negligible, since the two residuals share the keyframe- i samples and the preintegration ties velocity to position. Discarding it introduces a loss of statistical efficiency relative to a full joint weighting, but significantly simplifies implementation. Two additional correlations are treated analogously: (i) correlations between the substituted samples and the preintegration they also drive, and (ii) correlations between consecutive factors that share a common keyframe. The gyroscope bias also enters both the substitution and the preintegration, but it is a part of the estimated state, so its coupling is handled by the optimizer rather than through covariance coupling.

Finally, each weight depends on the IMU–DVL lever arm through the gyroscope contributions $\boldsymbol{\Sigma}^{(\text{Gyro})}$, which carry $[{}_b\mathbf{p}_{d/b}]_{\times}$. Re-evaluating the weight as the lever arm updates would lower the cost by inflating the covariance rather than reducing the residual. Therefore, the weights are frozen at a linearization point ${}_b\bar{\mathbf{p}}_{d/b}$, provided by the initialization procedure described in the next section.

4.3 The Coarse-to-Fine Pipeline

Formulation A is solved through a coarse-to-fine pipeline that activates parameters and residuals in stages, each warm-started from the previous (Figure 4.1). The camera trajectory is a fixed input from Chapter 2, so the pipeline begins after the vision-only bundle adjustment that AQUA-SLAM runs first, and proceeds in four stages rather than five.

Stage 1: gravity-independent initialization. The first stage initializes the rotations ${}_b\mathbf{R}_c$, ${}_c\mathbf{R}_d$ and the camera–DVL lever arm ${}_c\mathbf{p}_{d/c}$ from the gyroscope and DVL translation

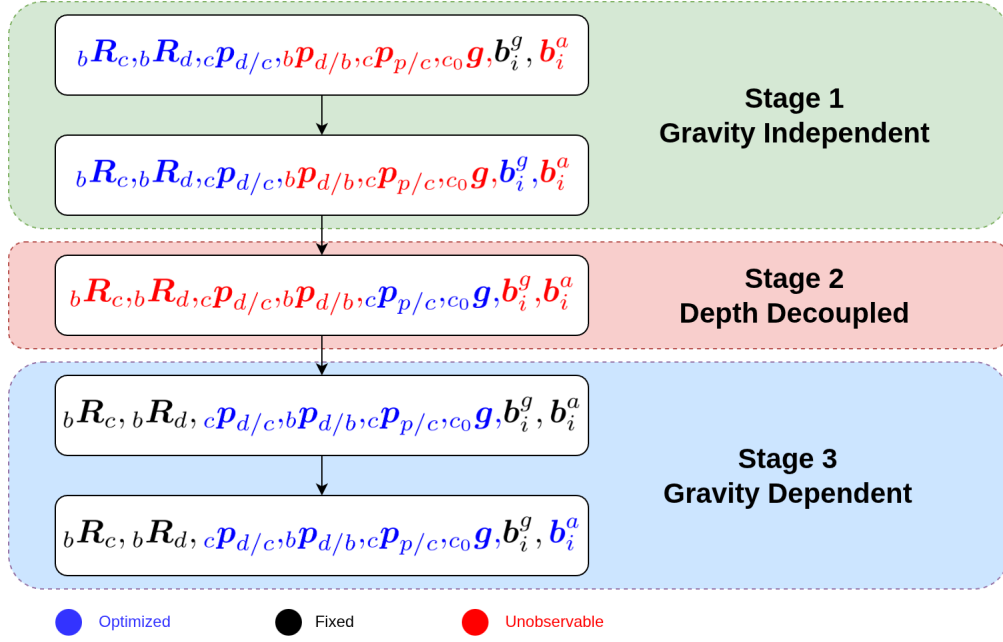


Figure 4.1: The Formulation A coarse-to-fine calibration pipeline. Each stage activates a subset of parameters and residuals. Optimized parameters are shown in blue, fixed parameters in black, and parameters unobservable at that stage in red. The state is warm-started between passes. The final joint pass uses all parameters and all residuals.

residuals alone, the two that involve neither gravity, the body velocity, nor the IMU–DVL lever arm. This stage is therefore decoupled from inertial effects and focuses purely on rotational alignment and extrinsic translation between the camera and DVL.

The optimization is performed in two passes: the first pass optimizes the rotations and ${}_c\mathbf{p}_{d/c}$ with the gyroscope bias held at zero; the second pass frees per-keyframe gyroscope biases $\{\mathbf{b}_i^g\}$, coupled across consecutive keyframes by a bias random-walk residual. All remaining parameters, including the gravity direction, the IMU–DVL lever arm, the pressure-sensor lever arm, and the accelerometer bias, are all unobservable here and held at their initial values.

Stage 2: depth-decoupled gravity initialization. The depth residual (2.7) depends only on the gravity direction ${}_{c_0}\mathbf{g}$ and the pressure-sensor lever arm ${}_c\mathbf{p}_{p/c}$, so the second stage uses it alone to initialize those two. All other parameters are fixed at their Stage 1 values. The gravity direction is initialized under an upright assumption, with gravity along c_0 ’s vertical axis of the reference camera frame.

Stage 3: gravity-dependent refinement. The third stage jointly refines the lever arms and gravity. It activates the IMU translation, IMU velocity, depth, and DVL translation residuals. Together these constrain the three lever arms ${}_c\mathbf{p}_{d/c}$, ${}_b\mathbf{p}_{d/b}$, ${}_c\mathbf{p}_{p/c}$, and the first three additionally constrain the gravity direction; the DVL translation residual, though gravity-independent, stays active to keep ${}_c\mathbf{p}_{d/c}$ constrained. The rotations and gyroscope bias are held at their Stage 1 values. A first pass optimizes

the lever arms and gravity together with the accelerometer bias at zero; a second pass frees the per-keyframe accelerometer biases $\{\mathbf{b}_i^a\}$. Stage 3 is the first to estimate the IMU–DVL lever arm, which enters only the IMU residuals activated here. Because the gyroscope-sample covariance $\Sigma^{(\text{Gyro})}$ depends on that lever arm and no estimate is yet available, it is omitted: ${}_b\mathbf{p}_{d/b}$ is set to zero for the IMU-residual covariance.

Stage 4: joint refinement. The final stage jointly releases all parameters and optimizes the full state (4.1) with respect to all residuals, together with the random-walk residuals for the gyroscope and accelerometer biases. All parameters are warm-started from the previous stages. The gyroscope-sample covariance $\Sigma^{(\text{Gyro})}$ omitted in Stage 3 is now included, evaluated at the warm-started lever arm ${}_b\bar{\mathbf{p}}_{d/b}$, the linearization point of Section 4.2.2, and held fixed through the stage. In practice, this joint stage yields only small refinements, suggesting that the staged estimates already lie close to the joint optimum.

The pipeline is implemented in C++ with the Ceres Solver [1]. Each stage is a separate nonlinear least squares problem solved by the Levenberg-Marquardt algorithm with sparse normal factorization, efficient for the banded sparsity that the per-keyframe biases and their random-walk residuals produce. The rotations are parameterized by local exponential maps on $SO(3)$ and the gravity direction on S^2 by a two-dimensional local chart, as in Section 2.3.

The pipeline retains AQUA-SLAM’s coarse-to-fine structure (Section 3.1) but differs in two respects. The first is the residual set. Formulation A’s IMU residuals carry the lever-arm correction (Section 4.1), which makes the IMU–DVL lever arm observable where AQUA-SLAM’s omitted-term bias (Section 3.2) collapses it. The second is the depth residual. It adds the pressure-sensor lever arm ${}_c\mathbf{p}_{p/c}$, which AQUA-SLAM does not estimate. It also changes how the gravity direction is initialized. AQUA-SLAM observes gravity only through the IMU residuals, and those residuals also carry ${}_b\mathbf{p}_{d/b}$. The lever arm is not constrained until the final joint stage, after gravity has already been initialized. Gravity is therefore initialized while the lever arm sits at its uninformed starting value, even though the two are strongly coupled. Formulation A’s depth residual constrains gravity without the lever arm. Stage 2 initializes gravity using only it, before the lever arm is estimated in Stage 3.

4.4 Empirical Validation

Formulation A makes the IMU–DVL lever arm observable by construction through the inclusion of the corrected IMU residuals. However, observability does not, on its own, guarantee recovery of the lever arm in the presence of sensor noise. This section uses a Monte Carlo simulation to assess how reliably the lever arm is recovered in practice and how that recovery degrades as sensor noise increases.

4.4.1 Protocol

The evaluation uses the Calib Medium trajectory of Section 3.3.2, whose angular excitation spans all three body axes, the condition that the IMU–DVL lever arm requires

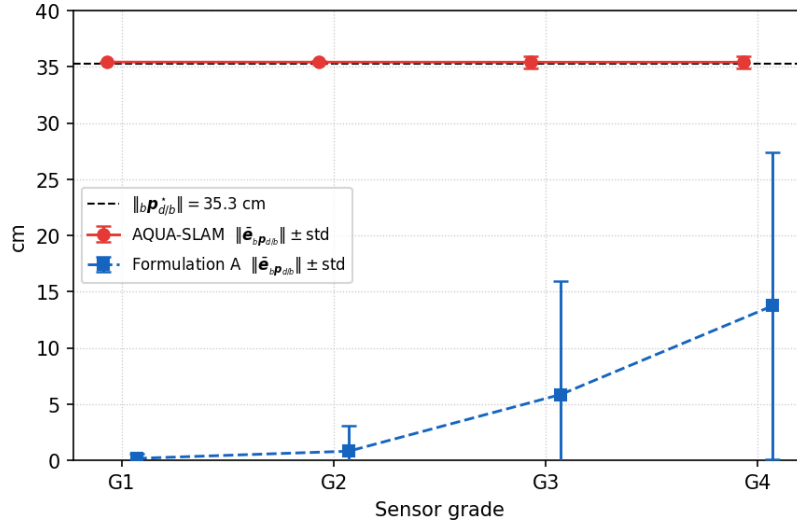


Figure 4.2: IMU–DVL lever-arm error magnitude $\|\bar{\mathbf{e}}_{b\hat{\mathbf{p}}_{d/b}}\| = \|\mathbf{b}\hat{\mathbf{p}}_{d/b} - \mathbf{b}\mathbf{p}_{d/b}^*\|$ for AQUA-SLAM (red) and Formulation A (blue) across sensor grades G1–G4 on the Calib Medium trajectory. AQUA-SLAM’s error is approximately equal to the true lever-arm magnitude $\|\mathbf{b}\mathbf{p}_{d/b}^*\| = 35.3$ cm at all grades, consistent with the AQUA-SLAM estimate collapsing to zero (Section 3.2). Formulation A’s error grows with sensor noise grade and remains below AQUA-SLAM’s at every grade. Error bars are ± 1 standard deviation across $N = 200$ Monte Carlo trials.

to be observable. It follows the Monte Carlo protocol of Section 3.3.4, where we set $N = 200$ trials per cell at the four sensor noise grades G1–G4 of Table 3.2, each cell summarized by the bias and across-run standard deviation of the estimates.

4.4.2 AQUA-SLAM versus Formulation A

To isolate the effect of the lever-arm correction, both methods are run on the same residual set, DVL translation and three IMU residuals, over the same set of estimated parameters, excluding the pressure-sensor lever arm. They then differ only in the velocity supply: Formulation A restores the correction that AQUA-SLAM omits (Section 4.1).

Figure 4.2 reports the IMU–DVL lever-arm error. AQUA-SLAM’s error remains at approximately the true lever-arm magnitude across all grades, the noise-independent collapse established in Section 3.2 and measured in Section 3.4. Formulation A’s error is far smaller and grows with sensor noise, from about 0.2 cm at G1 to about 14 cm at G4. The gap to AQUA-SLAM is about two orders of magnitude at G1 and narrows to a factor of two to three at G4 as the bias grows. At the operational grades G1–G2, the bias is at most about 0.8 cm and lies within the noise-driven scatter. At the high grades, the across-run scatter also grows large, reaching about 14 cm at G4, on the order of the bias (Table 4.1). A single estimate is therefore unreliable at these grades, even though the mean offset over the $N = 200$ trials remains well determined.

Table 4.1 reports the bias and standard deviation across all extrinsic parameters.

Table 4.1: Monte Carlo bias and standard deviation for AQUA-SLAM and Formulation A across sensor grades G1–G4 on the Calib Medium trajectory ($N = 200$ trials). Translation in cm; rotation and gravity in degrees.

		AQUA-SLAM				Formulation A			
		G1	G2	G3	G4	G1	G2	G3	G4
${}_b\mathbf{p}_{d/b}$	bias [cm]	35.43	35.43	35.42	35.42	0.17	0.81	5.86	13.75
	std [cm]	0.03	0.11	0.55	0.52	0.43	2.23	10.12	13.64
${}_c\mathbf{p}_{d/c}$	bias [cm]	0.52	0.52	0.57	0.56	0.21	0.21	0.20	0.23
	std [cm]	0.11	0.58	2.98	2.73	0.12	0.59	3.01	2.76
${}_b\mathbf{R}_c$	bias [°]	0.037	0.037	0.036	0.036	0.036	0.036	0.037	0.037
	std [°]	0.002	0.009	0.034	0.068	0.003	0.010	0.039	0.077
${}_b\mathbf{R}_d$	bias [°]	0.185	0.174	0.134	0.135	0.040	0.043	0.090	0.124
	std [°]	0.055	0.278	1.423	1.309	0.053	0.271	1.386	1.276
${}_c\mathbf{g}$	bias [°]	0.713	0.732	0.688	0.692	0.024	0.046	0.107	0.192
	std [°]	0.022	0.109	0.553	0.511	0.048	0.231	0.915	1.217

Table 4.2: Pressure-sensor lever-arm calibration with Formulation A on the Calib Medium trajectory ($N = 200$ trials). The true value has magnitude $\|{}_c\mathbf{p}_{p/c}^*\| = 56.70$ cm.

	G1	G2	G3	G4
bias [cm]	0.05	0.10	0.22	0.81
std [cm]	0.99	2.47	5.92	11.77

Beyond the lever arm, the gravity direction recovers from a $\sim 0.7^\circ$ bias under AQUA-SLAM to under 0.1° under Formulation A at G1–G2, which indicates that AQUA-SLAM’s gravity bias arose from coupling with the lever-arm bias rather than from an independent mechanism. The camera-DVL lever arm and the two rotations are recovered with sub-cm and sub-degree bias under both methods, because they are constrained substantially by the IMU rotation and DVL translation residuals, which do not depend on the velocity supply that distinguishes the two methods.

4.4.3 Pressure-Sensor Calibration

The comparison above held the residual set fixed, leaving out the depth residual. Restoring it adds the pressure-sensor lever arm ${}_c\mathbf{p}_{p/c}$, which AQUA-SLAM does not estimate. Table 4.2 reports its recovery under the full Formulation A pipeline, with all residuals active.

The pressure-sensor lever arm is recovered with bias below 1 cm at all grades. The standard deviation grows from about 1 cm at G1 to about 12 cm at G4 and exceeds the bias at every grade, reflecting weaker observability from the scalar depth residual than from the IMU and DVL vector residuals.

4.5 Summary

In summary, Formulation A restores the lever-arm correction that AQUA-SLAM omits, the single change that makes the IMU–DVL lever arm observable. The Monte Carlo evaluation confirms the effect. Where AQUA-SLAM’s estimate collapses to zero regardless of the true lever arm, Formulation A recovers it to within the noise-driven scatter at operational grades, and the coupled gravity bias resolves along with it. The pressure-sensor lever arm that AQUA-SLAM leaves uncalibrated is also recovered.

One feature of the evaluation remains open. Formulation A’s lever-arm bias is not constant across grades but grows with sensor noise (Table 4.1). This noise dependence sets the residual bias apart from AQUA-SLAM’s omitted-term bias, which is constant with respect to the noise level and persists even in the noiseless limit. Restoring the omitted correction has removed that noise-independent collapse, but left a smaller bias that grows with sensor noise. Chapter 5 identifies the mechanism behind this residual bias.

Errors-in-Variables Bias in Formulation A

5

This chapter identifies the mechanism underlying the noise-dependent bias left by Formulation A and characterizes it as an errors-in-variables (EIV) effect. It then provides a scalar diagnostic of the bias severity, computable without ground truth, to flag an exposed calibration. Section 5.1 states the bias for a general residual whose regressor carries noise. Section 5.2 shows that Formulation A’s IMU residuals are of this kind, so the lever arm is biased. Section 5.3 reduces the bias to a single scalar. Section 5.4 shows that the structure of the IMU residuals worsens the bias due to the slow, smooth motion of underwater operation. Section 5.5 validates the mechanism on four trajectories and four sensor noise grades and shows that the diagnostic is computable in practice.

5.1 Bias Under a Noisy Regressor

The mechanism underlying the noise-dependent bias can be interpreted via the EIV bias in regression: when the regressor multiplying the parameter carries noise, the least-squares estimate is biased even in the limit of unlimited data [27]. It is not specific to Formulation A, so it is first derived in a general setting before being specialized to the IMU residual structure of interest. A residual on a keyframe interval (i, j) carries a noisy-regressor structure when it takes the form

$$\mathbf{r}_{ij}(\mathbf{p}) = (\bar{\mathbf{A}}_{ij} + \delta\mathbf{A}_{ij})\mathbf{p} - \bar{\mathbf{A}}_{ij}\mathbf{p}^* + \boldsymbol{\eta}_{ij}, \quad (5.1)$$

where $\mathbf{p} \in \mathbb{R}^d$ is the parameter with true value $\mathbf{p}^*$, $\bar{\mathbf{A}}_{ij}$ is the deterministic regressor at the truth, $\delta\mathbf{A}_{ij}$ is a zero-mean random perturbation of the regressor caused by sensor noise, and $\boldsymbol{\eta}_{ij}$ is additive measurement noise. The full noisy regressor is written $\mathbf{A}_{ij} := \bar{\mathbf{A}}_{ij} + \delta\mathbf{A}_{ij}$. The perturbation $\delta\mathbf{A}_{ij}$ is independent of $\boldsymbol{\eta}_{ij}$ and independent across keyframe intervals, and the constant offset $-\bar{\mathbf{A}}_{ij}\mathbf{p}^*$ makes the residual vanish at $\mathbf{p} = \mathbf{p}^*$ in the noiseless limit.

The parameter is estimated by minimizing the WLS cost $J(\mathbf{p}) := \sum_{ij} \mathbf{r}_{ij}(\mathbf{p})^T \mathbf{W}_{ij} \mathbf{r}_{ij}(\mathbf{p})$, where $\mathbf{W}_{ij} \succeq 0$ is the weight, deterministic and independent of $\mathbf{p}$. We assume this cost sums over independent keyframe intervals with stationary motion and noise statistics. As the calibration window grows, the per-interval average of the cost concentrates around its expectation. The minimizer of J therefore converges to the minimizer of the expected cost $\mathbb{E}[J(\mathbf{p})]$ [27].

The expected cost is a quadratic in $\mathbf{p}$. Its curvature on each keyframe interval has two sources: the deterministic regressor and the perturbation. Three matrices per keyframe interval (i, j) carry this decomposition: the deterministic information $\bar{\mathbf{H}}_{ij}$,

the EIV penalty Λ_{ij} , and the full information $\mathbf{H}_{ij}$,

$$\bar{\mathbf{H}}_{ij} := \bar{\mathbf{A}}_{ij}^T \mathbf{W}_{ij} \bar{\mathbf{A}}_{ij}, \quad (5.2)$$

$$\Lambda_{ij} := \mathbb{E}[\delta \mathbf{A}_{ij}^T \mathbf{W}_{ij} \delta \mathbf{A}_{ij}], \quad (5.3)$$

$$\mathbf{H}_{ij} := \bar{\mathbf{H}}_{ij} + \Lambda_{ij}, \quad (5.4)$$

Both $\bar{\mathbf{H}}_{ij}$ and Λ_{ij} are positive semidefinite, and the cross terms in $\mathbb{E}[\mathbf{A}_{ij}^T \mathbf{W}_{ij} \mathbf{A}_{ij}]$ vanish because $\delta \mathbf{A}_{ij}$ is zero-mean, which gives (5.4).

Summing over the keyframe intervals of the calibration window gives the aggregates

$$\bar{\mathbf{H}} := \sum_{ij} \bar{\mathbf{H}}_{ij}, \quad \Lambda := \sum_{ij} \Lambda_{ij}, \quad \mathbf{H} := \bar{\mathbf{H}} + \Lambda. \quad (5.5)$$

The aggregate $\mathbf{H}$ is the expected Gauss-Newton information of the WLS problem at $\mathbf{p}^*$ under the sensor noise.

The expected cost combines these two aggregated curvatures into a single quadratic, and the two contributions are centered at different points. The deterministic information forms a bowl of curvature $\bar{\mathbf{H}}$ with its minimum at the truth. The EIV penalty forms a bowl of curvature Λ with its minimum at the origin. The centering follows from the perturbation term: it enters each residual as $\delta \mathbf{A}_{ij} \mathbf{p}$, and the expected weighted squares sum to $\mathbf{p}^T \Lambda \mathbf{p}$, smallest at $\mathbf{p} = \mathbf{0}$. The minimizer of the expected cost balances the two bowls and settles between their centers.

Lemma 5.1 (asymptotic bias under a noisy regressor). *For residuals of the form (5.1) with the independence assumptions above and $\mathbf{H} \succ 0$, the minimizer of the expected WLS cost is unique and equals*

$$\mathbf{p}^\infty := \arg \min_{\mathbf{p}} \mathbb{E}[J(\mathbf{p})] = \mathbf{p}^* - \mathbf{H}^{-1} \Lambda \mathbf{p}^*, \quad (5.6)$$

so the bias is

$$\mathbf{p}^\infty - \mathbf{p}^* = -\mathbf{H}^{-1} \Lambda \mathbf{p}^*. \quad (5.7)$$

Suppose the motion and noise statistics do not change as the window grows, so that each additional keyframe interval contributes information and a penalty of similar magnitude. Then, $\mathbf{H}$ and Λ scale proportionally with the number of keyframe intervals, the bias $\mathbf{p}^\infty - \mathbf{p}^*$ converges to a fixed nonzero value, and the estimator does not approach the ground truth as more data is collected.

The proof follows the standard treatment of errors-in-variables estimators [27] and is given in Appendix C.

5.2 The Errors-in-Variables Bias in Formulation A

Lemma 5.1 applies to any residual with a noisy-regressor structure. Now it remains to show that Formulation A's IMU residuals fit this EIV structure. To isolate the mechanism, two simplifying assumptions are introduced.

Assumption 5.1. *Only the IMU–DVL lever arm ${}_b\mathbf{p}_{d/b}$ is treated as unknown, with all other state parameters held at their true values.*

Assumption 5.2. *Gyroscope noise is isotropic, $\mathbb{E}[\boldsymbol{\eta}_i^g (\boldsymbol{\eta}_j^g)^T] = \sigma_g^2 \mathbf{I} \delta_{ij}$, where $\boldsymbol{\eta}_i^g \in \mathbb{R}^3$ is the gyroscope noise sample at keyframe i .*

Assumption 5.1 isolates the mechanism on the lever arm, which serves as the parameter $\mathbf{p}$ in Lemma 5.1. Assumption 5.2 adopts isotropy to obtain the penalties in closed form. Anisotropy changes those expressions but leaves the noisy-regressor structure, and with it the mechanism, intact.

Under these assumptions, Formulation A’s IMU translation and velocity residuals are instances of the noisy-regressor residual (5.1). Their aggregate information and EIV penalty specialize the general $\mathbf{H}$ and $\boldsymbol{\Lambda}$ of (5.5) to the Formulation A quantities $\mathbf{H}^{(A)}$ and $\boldsymbol{\Lambda}^{(A)}$. The theorem below states that bias caused by these residuals is characterized by these matrices.

Theorem 5.1 (EIV bias of the lever arm). *Under Assumptions 5.1 and 5.2, Formulation A’s IMU translation and velocity residuals carry a nonzero EIV penalty driven by gyroscope noise. Consequently, the lever arm estimator is asymptotically biased,*

$${}_b\mathbf{p}_{d/b}^\infty = {}_b\mathbf{p}_{d/b}^\star - (\mathbf{H}^{(A)})^{-1} \boldsymbol{\Lambda}^{(A)} {}_b\mathbf{p}_{d/b}^\star \neq {}_b\mathbf{p}_{d/b}^\star, \quad (5.8)$$

and the bias does not vanish as the calibration window grows.

Proof. Gyroscope noise enters each residual’s regressor through the lever arm correction term ${}_b\hat{\boldsymbol{\omega}}_{b_i} \times {}_b\mathbf{p}_{d/b}$. The proof first makes this explicit for the IMU translation residual, then the IMU velocity residual, and aggregates the two before applying Lemma 5.1.

The IMU translation residual (4.4) depends on the lever arm through two paths: the camera–body lever arm ${}_c\mathbf{p}_{c/b} = {}_b\mathbf{R}_c^T {}_b\mathbf{p}_{d/b} - {}_c\mathbf{p}_{d/c}$, and the lever arm correction term ${}_b\hat{\boldsymbol{\omega}}_{b_i} \times {}_b\mathbf{p}_{d/b} \Delta t_{ij}$. Both paths are linear in ${}_b\mathbf{p}_{d/b}$, so with the other parameters held at their truth, the residual is affine in the lever arm:

$$\mathbf{r}_{ij}^{\text{IMU},p}({}_b\mathbf{p}_{d/b}) = (\bar{\mathbf{A}}_{ij}^p + \delta \mathbf{A}_{ij}^p) {}_b\mathbf{p}_{d/b} - \bar{\mathbf{A}}_{ij}^p {}_b\mathbf{p}_{d/b}^\star + \boldsymbol{\eta}_{ij}^p, \quad (5.9)$$

with a deterministic regressor and gyroscope-noise perturbation

$$\bar{\mathbf{A}}_{ij}^p = (\mathbf{I} - \mathbf{R}_{ij}) + [{}_b\boldsymbol{\omega}_{b_i}^\star]_\times \Delta t_{ij}, \quad \mathbf{R}_{ij} := {}_b\mathbf{R}_{b_j}, \quad (5.10)$$

$$\delta \mathbf{A}_{ij}^p = [\boldsymbol{\eta}_i^g]_\times \Delta t_{ij}, \quad (5.11)$$

where ${}_b\boldsymbol{\omega}_{b_i}^\star$ is the true body angular velocity. The first term of $\bar{\mathbf{A}}_{ij}^p$ comes from the camera–body lever arm, and the second from the correction term. The perturbation enters through the correction term alone.

The translation residual is weighted by $\mathbf{W}_{ij}^p = \Sigma_{\mathbf{r}_{ij}^{\text{IMU},p}}^{-1}$, the inverse of its covariance (4.8), symmetric and positive definite. The deterministic information and EIV penalty specialize (5.2) and (5.3) to this residual,

$$\bar{\mathbf{H}}_{ij}^p = (\bar{\mathbf{A}}_{ij}^p)^T \mathbf{W}_{ij}^p \bar{\mathbf{A}}_{ij}^p, \quad \boldsymbol{\Lambda}_{ij}^p = \mathbb{E}[(\delta \mathbf{A}_{ij}^p)^T \mathbf{W}_{ij}^p \delta \mathbf{A}_{ij}^p]. \quad (5.12)$$

Substituting $\delta \mathbf{A}_{ij}^p = [\boldsymbol{\eta}_i^g]_{\times} \Delta t_{ij}$ and factoring the scalar Δt_{ij}^2 , the penalty depends on the gyroscope noise alone,

$$\boldsymbol{\Lambda}_{ij}^p = \Delta t_{ij}^2 \mathbb{E}[[\boldsymbol{\eta}_i^g]^T \mathbf{W}_{ij}^p [\boldsymbol{\eta}_i^g]_{\times}]. \quad (5.13)$$

The expectation is an instance of the cross-product expectation identity of Appendix B. Its transposed form (B.5), with the symmetric weight $\mathbf{W}_{ij}^p$ and the gyroscope covariance $\sigma_g^2 \mathbf{I}$ of Assumption 5.2, gives

$$\boldsymbol{\Lambda}_{ij}^p = \sigma_g^2 \Delta t_{ij}^2 (\text{tr}(\mathbf{W}_{ij}^p) \mathbf{I} - \mathbf{W}_{ij}^p). \quad (5.14)$$

Next, we look at the IMU velocity residual (4.3), which is affine in the lever arm in the same way, with correction terms at both keyframe endpoints. It takes the form

$$\mathbf{r}_{ij}^{\text{IMU},v}({}_b \mathbf{p}_{d/b}) = (\bar{\mathbf{A}}_{ij}^v + \delta \mathbf{A}_{ij}^v) {}_b \mathbf{p}_{d/b} - \bar{\mathbf{A}}_{ij}^v {}_b \mathbf{p}_{d/b}^* + \boldsymbol{\eta}_{ij}^v, \quad (5.15)$$

with

$$\bar{\mathbf{A}}_{ij}^v = [{}_b \boldsymbol{\omega}_{b_i}^*]_{\times} - \mathbf{R}_{ij} [{}_b \boldsymbol{\omega}_{b_j}^*]_{\times}, \quad (5.16)$$

$$\delta \mathbf{A}_{ij}^v = [\boldsymbol{\eta}_i^g]_{\times} - \mathbf{R}_{ij} [\boldsymbol{\eta}_j^g]_{\times}, \quad (5.17)$$

so both endpoints' gyroscope samples appear, coupled through the relative rotation $\mathbf{R}_{ij}$.

The velocity residual is weighted by $\mathbf{W}_{ij}^v = \boldsymbol{\Sigma}_{\mathbf{r}_{ij}^{\text{IMU},v}}^{-1}$, the inverse of its covariance (4.12), symmetric and positive definite. The deterministic information and EIV penalty specialize (5.2) and (5.3) to this residual,

$$\bar{\mathbf{H}}_{ij}^v = (\bar{\mathbf{A}}_{ij}^v)^T \mathbf{W}_{ij}^v \bar{\mathbf{A}}_{ij}^v, \quad \boldsymbol{\Lambda}_{ij}^v = \mathbb{E}[(\delta \mathbf{A}_{ij}^v)^T \mathbf{W}_{ij}^v \delta \mathbf{A}_{ij}^v]. \quad (5.18)$$

Since $\delta \mathbf{A}_{ij}^v$ carries both endpoint samples, the penalty integrand expands into four terms,

$$\begin{aligned} (\delta \mathbf{A}_{ij}^v)^T \mathbf{W}_{ij}^v \delta \mathbf{A}_{ij}^v &= [\boldsymbol{\eta}_i^g]^T_{\times} \mathbf{W}_{ij}^v [\boldsymbol{\eta}_i^g]_{\times} - [\boldsymbol{\eta}_i^g]^T_{\times} \mathbf{W}_{ij}^v \mathbf{R}_{ij} [\boldsymbol{\eta}_j^g]_{\times} \\ &\quad - [\boldsymbol{\eta}_j^g]^T_{\times} \mathbf{R}_{ij}^T \mathbf{W}_{ij}^v [\boldsymbol{\eta}_i^g]_{\times} + [\boldsymbol{\eta}_j^g]^T_{\times} \mathbf{R}_{ij}^T \mathbf{W}_{ij}^v \mathbf{R}_{ij} [\boldsymbol{\eta}_j^g]_{\times}. \end{aligned} \quad (5.19)$$

The two cross terms vanish in expectation because $\boldsymbol{\eta}_i^g$ and $\boldsymbol{\eta}_j^g$ are independent samples at distinct keyframes. Applying cross product identity (B.5) again to the two diagonal terms gives

$$\mathbb{E}[[\boldsymbol{\eta}_i^g]^T_{\times} \mathbf{W}_{ij}^v [\boldsymbol{\eta}_i^g]_{\times}] = \sigma_g^2 (\text{tr}(\mathbf{W}_{ij}^v) \mathbf{I} - \mathbf{W}_{ij}^v), \quad (5.20)$$

$$\mathbb{E}[[\boldsymbol{\eta}_j^g]^T_{\times} \mathbf{R}_{ij}^T \mathbf{W}_{ij}^v \mathbf{R}_{ij} [\boldsymbol{\eta}_j^g]_{\times}] = \sigma_g^2 (\text{tr}(\mathbf{R}_{ij}^T \mathbf{W}_{ij}^v \mathbf{R}_{ij}) \mathbf{I} - \mathbf{R}_{ij}^T \mathbf{W}_{ij}^v \mathbf{R}_{ij}). \quad (5.21)$$

The trace is rotation-invariant, $\text{tr}(\mathbf{R}_{ij}^T \mathbf{W}_{ij}^v \mathbf{R}_{ij}) = \text{tr}(\mathbf{W}_{ij}^v)$ and writing $\mathbf{I} = \mathbf{R}_{ij}^T \mathbf{R}_{ij}$ allows us to take rotations out as common terms in (5.21). Summing both equations,

$$\boldsymbol{\Lambda}_{ij}^v = \sigma_g^2 (\text{tr}(\mathbf{W}_{ij}^v) \mathbf{I} - \mathbf{W}_{ij}^v) + \sigma_g^2 \mathbf{R}_{ij}^T (\text{tr}(\mathbf{W}_{ij}^v) \mathbf{I} - \mathbf{W}_{ij}^v) \mathbf{R}_{ij}. \quad (5.22)$$

The penalty carries no Δt_{ij} factor, unlike the translation penalty (5.14).

The two residuals are weighted separately, following the block-diagonal approximation of the joint IMU covariance in Section 4.2.2, so the aggregate information and penalty sum the per-interval contributions,

$$\mathbf{H}^{(A)} = \sum_{ij} (\bar{\mathbf{H}}_{ij}^p + \mathbf{\Lambda}_{ij}^p + \bar{\mathbf{H}}_{ij}^v + \mathbf{\Lambda}_{ij}^v), \quad \mathbf{\Lambda}^{(A)} = \sum_{ij} (\mathbf{\Lambda}_{ij}^p + \mathbf{\Lambda}_{ij}^v). \quad (5.23)$$

Every term is positive semidefinite, and the translation penalty is positive definite, since $\text{tr}(\mathbf{W}_{ij}^p) \mathbf{I} - \mathbf{W}_{ij}^p \succ \mathbf{0}$ for the symmetric positive-definite weight. Hence $\mathbf{\Lambda}^{(A)} \succ \mathbf{0}$ and $\mathbf{H}^{(A)} \succ \mathbf{0}$, so Lemma 5.1 applies and yields the asymptotic minimizer (5.8), displaced from the truth because $\mathbf{\Lambda}^{(A)} \neq \mathbf{0}$. Under stationary motion and noise statistics $\mathbf{H}^{(A)}$ and $\mathbf{\Lambda}^{(A)}$ grow in proportion to the number of keyframe intervals, so the bias settles at a fixed nonzero value as the calibration window grows. $\square$

Theorem 5.1 establishes that the bias exists and is fixed by the matrix $(\mathbf{H}^{(A)})^{-1} \mathbf{\Lambda}^{(A)}$. The bias (5.8) depends on the true lever arm, the trajectory, and the noise, but whether the EIV mechanism dominates is a property of the matrix alone. The next section characterizes this matrix through its eigenvalue structure to obtain a scalar measure of its effect.

5.3 The Eigenvalue Diagnostic

For brevity, in this section, we drop the superscript (A) and write $\mathbf{H}$, $\mathbf{\Lambda}$, and $\mathbf{p}$ instead of $\mathbf{H}^{(A)}$, $\mathbf{\Lambda}^{(A)}$, and ${}_b\mathbf{p}_{d/b}$, respectively, where no ambiguity arises. The shrinkage matrix $\mathbf{H}^{-1} \mathbf{\Lambda}$ is summarized by the smallest generalized eigenvalue $\nu_{\min}$ of the pair $(\mathbf{H}, \mathbf{\Lambda})$. Consider the generalized eigenvalue problem

$$\mathbf{H} \mathbf{v} = \nu \mathbf{\Lambda} \mathbf{v}. \quad (5.24)$$

where $\mathbf{v}$ is a generalized eigen vector and ν is the corresponding generalized eigenvalue of the pair $(\mathbf{H}, \mathbf{\Lambda})$. Left-multiplying by $\mathbf{H}^{-1}$ gives $\mathbf{H}^{-1} \mathbf{\Lambda} \mathbf{v} = (1/\nu) \mathbf{v}$, so the eigenvectors of (5.24) are eigenvectors of the shrinkage matrix $\mathbf{H}^{-1} \mathbf{\Lambda}$ with eigenvalues $1/\nu$. When $\mathbf{\Lambda}$ is positive definite, the problem admits a complete set of eigenpairs $\{(\nu_k, \mathbf{v}_k)\}_{k=1}^d$ with $\nu_k \in [1, \infty)$. The lower bound follows from $\mathbf{H} = \bar{\mathbf{H}} + \mathbf{\Lambda} \succeq \mathbf{\Lambda}$. The aggregate $\mathbf{\Lambda}$ is a sum of positive semidefinite per-interval penalties (5.3), so it is positive definite whenever the perturbations excite every parameter direction somewhere in the window. Positive definiteness of $\mathbf{\Lambda}$ also gives $\mathbf{H} \succ \mathbf{0}$, so the inverse in the bias formula (5.7) exists. We order the eigenvalues $\nu_1 \leq \nu_2 \leq \dots \leq \nu_d$ with $\nu_{\min} := \nu_1$.

The eigenvectors span $\mathbb{R}^d$, so the true parameter expands as $\mathbf{p}^\star = \sum_k \alpha_k \mathbf{v}_k$. Substituting into (5.7) through the eigenvalue relation gives

$$\mathbf{p}^\infty - \mathbf{p}^\star = - \sum_k \frac{\alpha_k}{\nu_k} \mathbf{v}_k. \quad (5.25)$$

The bias decomposes along the eigenvectors, with each true component reduced by the fraction $1/\nu_k \in (0, 1]$. The estimate is therefore a shrunk version of the truth, each

direction pulled toward zero by its own factor. Among the eigenvectors, the attenuation is strongest along the eigenvector $\mathbf{v}_{\min}$ of the smallest eigenvalue. When $\mathbf{p}^*$ aligns with $\mathbf{v}_{\min}$ the bias is

$$\|\mathbf{p}^\infty - \mathbf{p}^*\|_{\mathbf{p}^* \propto \mathbf{v}_{\min}} = \frac{\|\mathbf{p}^*\|}{\nu_{\min}}. \quad (5.26)$$

Because the eigenvectors are $\mathbf{\Lambda}$ -orthonormal rather than Euclidean-orthogonal, (5.26) is not a Euclidean upper bound: a lever arm not aligned with any eigenvector can carry a larger Euclidean bias than $\|\mathbf{p}^*\|/\nu_{\min}$. Its role is different: the bias (5.7) is the shrinkage matrix applied to the truth, and $\nu_{\min}$ summarizes that matrix by its strongest contraction. The relative bias is therefore on the order of $1/\nu_{\min}$, concentrated along $\mathbf{v}_{\min}$.

Left-multiplying (5.24) by $\mathbf{v}_k^T$ expresses each eigenvalue as a ratio,

$$\nu_k = \frac{\mathbf{v}_k^T \mathbf{H} \mathbf{v}_k}{\mathbf{v}_k^T \mathbf{\Lambda} \mathbf{v}_k}, \quad (5.27)$$

the full information divided by the EIV penalty along $\mathbf{v}_k$. The eigenvalue is therefore a per-direction information-to-noise ratio. When $\nu_k \gg 1$, the information dominates the penalty, and the direction is effectively unbiased; when $\nu_k \rightarrow 1$, the penalty dominates, and the direction is unrecoverable.

5.4 The Information Degradation of the IMU Residuals

The bias along each direction is set by the ratio (5.27). Subtracting the penalty from the numerator through $\mathbf{H} = \bar{\mathbf{H}} + \mathbf{\Lambda}$ gives

$$\nu_k - 1 = \mathbf{v}_k^T \bar{\mathbf{H}} \mathbf{v}_k / \mathbf{v}_k^T \mathbf{\Lambda} \mathbf{v}_k, \quad (5.28)$$

so $\nu_{\min}$ is near its lower bound of one exactly when the deterministic information is small against the penalty. This section shows that the construction of Formulation A's IMU residuals systematically suppresses the deterministic information $\bar{\mathbf{H}}^{(A)}$ while leaving the gyroscope-noise penalty $\mathbf{\Lambda}^{(A)}$ unchanged. We begin by analyzing the sensitivity of the IMU residuals to the IMU-DVL lever arm, as outlined in the following theorem.

Theorem 5.2 (cancellation of the leading rotation content). *Let the body angular velocity be twice continuously differentiable within the keyframe interval, and let ${}_{b_i}\dot{\boldsymbol{\omega}}_{b_i}^*$ denote the body angular acceleration at keyframe i . Then the differencing that builds each IMU regressor removes its leading rotation content, the first-order content in the translation regressor (5.10) and the zeroth-order content in the velocity regressor (5.16), leaving the deterministic lever arm sensitivity in the centripetal and angular-acceleration effects alone,*

$$\bar{\mathbf{A}}_{ij}^p = -\frac{1}{2}([{}_{b_i}\boldsymbol{\omega}_{b_i}^*]_{\times}^2 + [{}_{b_i}\dot{\boldsymbol{\omega}}_{b_i}^*]_{\times}) \Delta t_{ij}^2 + \mathcal{O}(\Delta t_{ij}^3), \quad (5.29)$$

$$\bar{\mathbf{A}}_{ij}^v = -([{}_{b_i}\boldsymbol{\omega}_{b_i}^*]_{\times}^2 + [{}_{b_i}\dot{\boldsymbol{\omega}}_{b_i}^*]_{\times}) \Delta t_{ij} + \mathcal{O}(\Delta t_{ij}^2). \quad (5.30)$$

The gyroscope-noise perturbations (5.11) and (5.17) undergo no such cancellation.

Proof. By the regularity of the angular velocity, the body angular velocity at keyframe j expands as ${}_{b_j}\boldsymbol{\omega}_{b_j}^* = {}_{b_i}\boldsymbol{\omega}_{b_i}^* + {}_{b_i}\dot{\boldsymbol{\omega}}_{b_i}^* \Delta t_{ij} + \mathcal{O}(\Delta t_{ij}^2)$, with the remainder governed by the second derivative, and the relative rotation as

$$\mathbf{R}_{ij} = \mathbf{I} + [{}_{b_i}\boldsymbol{\omega}_{b_i}^*]_{\times} \Delta t_{ij} + \frac{1}{2}([{}_{b_i}\boldsymbol{\omega}_{b_i}^*]_{\times}^2 + [{}_{b_i}\dot{\boldsymbol{\omega}}_{b_i}^*]_{\times}) \Delta t_{ij}^2 + \mathcal{O}(\Delta t_{ij}^3),$$

where the correction from the rotation axis turning within the interval enters only at the third order.

In the translation regressor (5.10), the lever arm sensitivity is carried by the rigid-body link and the correction term. Substituting the expansion, the first-order content of the link is $-[{}_{b_i}\boldsymbol{\omega}_{b_i}^*]_{\times} \Delta t_{ij}$ and cancels exactly with the correction term, which gives (5.29). In contrast, the perturbation (5.11) enters through the correction term alone and finds no counterpart in the noiseless trajectory rotations of the link, so it does not cancel.

In the velocity regressor (5.16), the lever arm sensitivity is carried by the two endpoint correction terms. Substituting the endpoint expansion and the rotation expansion, their zeroth-order content cancels exactly, which gives (5.30). The perturbation (5.17) does not cancel, because the two endpoint samples are independent. $\square$

By Theorem 5.2, the differencing cancels the angular velocity that would have carried its leading sensitivity. So, the lever arm is observed through the centripetal and angular-acceleration effects alone, which constitute the deterministic information $\bar{\mathbf{H}}^{(A)}$. The residuals, therefore, carry information about the lever arm only to the extent that the motion excites these two effects.

Both surviving effects are weak when the angular rates are low, and the motion is smooth. The centripetal effect scales with the square of the body angular velocity $\|{}_{b_i}\boldsymbol{\omega}_{b_i}^*\|^2$, so low angular rates make it small. The angular-acceleration effect scales with the rate of change of the angular velocity $\|{}_{b_i}\dot{\boldsymbol{\omega}}_{b_i}^*\|$, so smooth motion makes it small as well. Low angular rates do not make the lever arm unobservable. Angular motion is what makes the lever arm observable, and the rates here are high enough to do so. Observability settles only whether the lever arm is constrained by the residuals, not how strongly. The surviving information is small, nonetheless, because the strong, first-order sensitivity is carried by the angular velocity, and the differencing cancels exactly that content.

In contrast, the penalty $\boldsymbol{\Lambda}^{(A)}$ is driven by the gyroscope noise σ_g^2 rather than by the angular rates, and the differencing that cancels the deterministic content leaves the perturbation that drives the penalty intact. As the angular rates fall, the deterministic information $\bar{\mathbf{H}}^{(A)}$ shrinks while the penalty $\boldsymbol{\Lambda}^{(A)}$ holds at full size, so the worst-direction ratio $\nu_{\min}^{(A)}$ is pinned toward its lower bound of one. Slow, smooth motion is the regime in which this is most severe, and it is precisely the operating condition for underwater calibration.

The keyframe interval further limits the magnitude of the surviving effects. The translation regressor carries a factor Δt_{ij}^2 , and the velocity regressor a factor Δt_{ij} , and the intervals of the trajectory are short. The interval is inherited from the camera keyframing of the trajectory input, so it cannot be lengthened to recover information; a longer interval would, in any case, prolong preintegration and accumulate more noise.

This analysis isolates the lever arm under Assumption 5.1. With all parameters free, the lever arm couples to the others, and the joint problem behaves differently. The regime identified here is therefore characterized empirically in the next section, which tests the realized bias against the diagnostic.

5.5 Empirical Validation

This section evaluates the diagnostic for Formulation A’s pair $(\mathbf{H}^{(A)}, \mathbf{\Lambda}^{(A)})$, given by $\nu_{\min}^{(A)}$. Monte Carlo trials verify that the realized bias aligns with the predicted direction when the EIV effect is significant or saturated, approaches the lever-arm magnitude as $\nu_{\min}^{(A)}$ approaches its lower bound, and increases with sensor noise.

5.5.1 Protocol

The validation follows the setup in Section 3.3. Four trajectories are used: Calib Slow, Calib Medium, and Calib Fast, which span a controlled sweep over angular excitation, and the SE surrogate, which checks the calibration performance in a low excitation survey-like sequence. The four sensor noise grades, G1 to G4, and the four trajectories define sixteen (trajectory, grade) cells. Each cell is evaluated under the Monte Carlo protocol of Section 3.3.4, with $N = 200$ trials. The trials drop the simplified-scenario assumption of Section 5.2 and jointly estimate all calibration parameters under realistic sensor noise.

5.5.2 The Diagnostic across Trajectories and Grades

Table 5.1 shows that the oracle $\nu_{\min}^{(A)}$, the diagnostic evaluated at the true extrinsics on noise-free gyroscope samples, spans nearly four orders of magnitude across the sixteen cells. The values decrease with the noise grade: on every trajectory, $\nu_{\min}^{(A)}$ drops by two to three orders of magnitude from G1 to G4. At a fixed grade, the values scale with the angular excitation: the ordering from Calib Fast down to the SE surrogate is preserved at every grade, with Calib Fast about an order of magnitude above Calib Slow at G1 and G2, and the gap narrowing at the higher grades. The diagnostic is therefore governed by the rotation excitation against the gyroscope noise.

Table 5.1: Oracle $\nu_{\min}^{(A)}$ at the four sensor noise grades for the four trajectories.

Trajectory	G1	G2	G3	G4
Calib Fast	10820	780	57.7	15.1
Calib Medium	3900	282	21.4	6.07
Calib Slow	835	61.1	5.37	2.08
SE surrogate	202	15.5	2.05	1.26

These values partition the cells into three regimes. In the saturated cell, where $\nu_{\min}^{(A)} < 2$ (the SE surrogate at G4, with Calib Slow at G4 and the SE surrogate at G3 just above the boundary), the diagnostic predicts that the lever arm component along

$\mathbf{v}_{\min}$ is largely lost. In the EIV-significant cells, with $\nu_{\min}^{(A)}$ between 2 and 100 (Calib Fast and Calib Medium at G3 and G4, Calib Slow at G2 to G4, and the SE surrogate at G2 and G3), the predicted bias is a measurable fraction of the lever arm. In the EIV-negligible cells, where $\nu_{\min}^{(A)} \geq 100$ (Calib Fast and Calib Medium at G1 and G2, Calib Slow and the SE surrogate at G1), the attenuation is below about one percent, and the predicted bias is at the millimeter scale. The remaining subsections test these predictions against the realized bias.

5.5.3 Verifying the EIV Mechanism

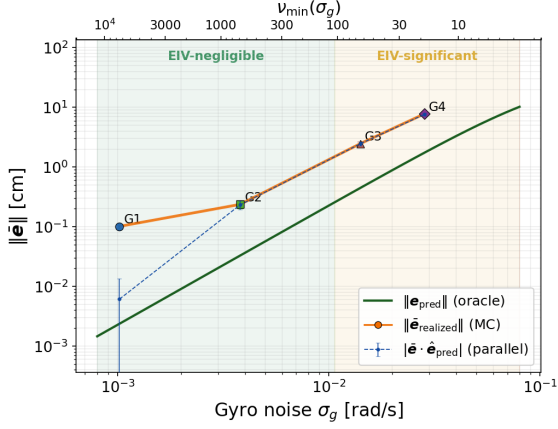
Figure 5.1 plots three quantities for each trajectory against the gyroscope noise σ_g , with the corresponding $\nu_{\min}^{(A)}$ on the upper axis: the predicted bias magnitude $\|(\mathbf{H}^{(A)})^{-1}\mathbf{\Lambda}^{(A)}{}_b\mathbf{p}_{d/b}^*\|$ (green); the realized total bias magnitude from the Monte Carlo trials (orange); and the realized component along the predicted EIV direction $\hat{\mathbf{e}}_{\text{pred}}$, the unit vector of the predicted bias $-(\mathbf{H}^{(A)})^{-1}\mathbf{\Lambda}^{(A)}{}_b\mathbf{p}_{d/b}^*$ (blue dashed). Both predicted quantities are evaluated at the true extrinsics. The different $\nu_{\min}^{(A)}$ regimes are shaded.

The realized bias direction tracks the prediction wherever the EIV effect is significant or saturated. At grades G2 to G4, on every trajectory, the parallel component (blue) coincides with the realized total (orange), so the bias lies almost entirely along the predicted direction, with the angle between them below about twelve degrees. The agreement weakens only at the EIV-negligible cells at G1, and it does so in step with $\nu_{\min}^{(A)}$: the angle grows from about thirty-five degrees at the SE surrogate and Calib Slow, to about fifty-seven degrees at Calib Medium and about eighty-seven degrees at Calib Fast, where the realized bias is nearly orthogonal to the prediction. This breakdown is the expected signature, not a discrepancy. When the predicted EIV bias is on the millimeter scale, it no longer determines the realized bias; unrelated contributions instead determine the direction. The directional agreement is therefore present exactly in the cells where the EIV bias is large, and the magnitude behavior at the lowest- $\nu_{\min}^{(A)}$ cells, examined next, makes the identification decisive.

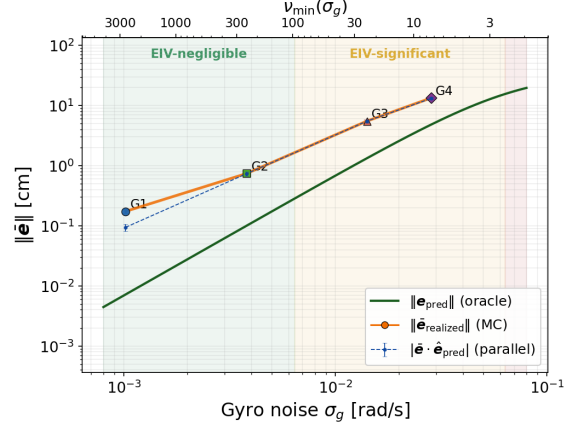
Across all four panels, the realized bias grows monotonically with the noise grade. The behavior at the lowest- $\nu_{\min}^{(A)}$ cells is the decisive evidence. At the SE surrogate at G3 and G4, and Calib Slow at G4, both the predicted and the realized bias climb toward $\|{}_b\mathbf{p}_{d/b}^*\| = 0.35$ m as $\nu_{\min}^{(A)}$ approaches its lower bound. A bias from another source has no reason to approach the lever arm magnitude at exactly the cells where $\nu_{\min}^{(A)}$ collapses. Together with the monotonic growth and the directional agreement in the EIV-significant and saturated cells, this establishes the residual bias as the errors-in-variables effect of Section 5.2.

5.5.4 Magnitude: A Conservative Indicator

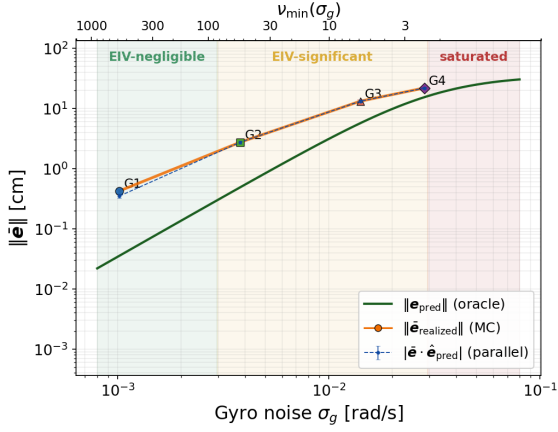
Across the four panels of Figure 5.1, the realized bias exceeds the closed-form predicted magnitude at every cell, and the excess shrinks monotonically with $\nu_{\min}^{(A)}$ on every trajectory. The realized bias exceeds the prediction by more than fortyfold in the highest- $\nu_{\min}^{(A)}$ cell (Calib Fast at G1) and by only about five percent in the most saturated cell (the



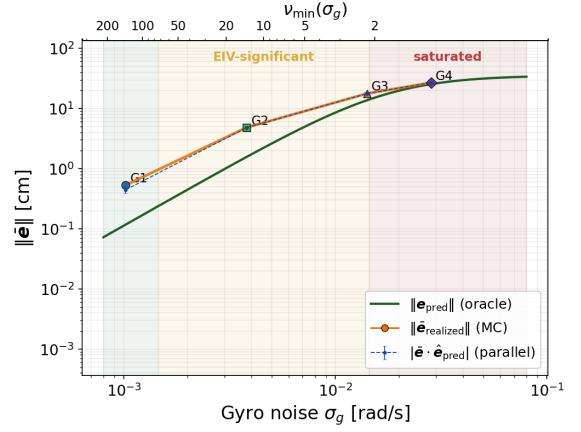
(a) Calib Fast ($\omega_{\text{peak}} = 1.0$ rad/s)



(b) Calib Medium ($\omega_{\text{peak}} = 0.6$ rad/s)



(c) Calib Slow ($\omega_{\text{peak}} = 0.25$ rad/s)



(d) SE surrogate ($\omega_{\text{peak}} \approx 0.2$ rad/s)

Figure 5.1: Predicted and realized IMU–DVL lever arm bias for the four trajectories. Green: oracle predicted bias magnitude swept over σ_g . Orange: realized total bias magnitude from $N = 200$ Monte Carlo trials. Blue dashed: realized component along $\hat{\mathbf{e}}_{\text{pred}}$. The upper axis marks the corresponding $\nu_{\text{min}}^{(A)}$, and different regimes are shaded.

SE surrogate at G4). The gap is large where $\nu_{\text{min}}^{(A)}$ is high, because the white-noise bias, i.e., the EIV bias driven by the white gyroscope noise that $\mathbf{\Lambda}^{(A)}$ models, is then only a small part of the realized bias. It closes in the saturation band, where that white-noise bias dominates. The closed form is therefore a conservative underestimate, and its tightness tracks how strongly the white-noise bias dominates.

The excess of the realized bias over the prediction has multiple sources. The dominant one, where the EIV effect is significant, is the gyroscope bias estimation error, which the penalty does not model. The correction term removes the estimated gyroscope bias from the raw sample, so the bias-state estimation error enters on the same path as the white gyroscope noise and perturbs the same regressor. This error does not vanish, because the gyroscope bias is co-estimated rather than known, yet $\mathbf{\Lambda}^{(A)}$ is built from the white gyroscope noise alone and excludes it. A control run isolates it:

Table 5.2: IMU–DVL lever-arm bias on Calib Medium, realized with the gyroscope bias estimated and with it held fixed at its true value, against the oracle prediction (cm). The realized rows are mean $\pm$ per-trial standard deviation over $N = 200$ trials; the oracle prediction is deterministic.

	G1	G2	G3	G4
Realized	0.17 ± 0.01	0.74 ± 0.07	5.49 ± 0.34	13.29 ± 0.49
Realized, gyroscope bias fixed	0.31 ± 0.01	0.28 ± 0.05	0.38 ± 0.26	2.22 ± 0.24
Oracle prediction	0.01	0.10	1.33	4.79

with the gyroscope bias fixed to its true value rather than estimated, the realized bias on Calib Medium falls by about fourteenfold at G3 and sixfold at G4 (Table 5.2, note that results slightly differ from Table 4.1 due to included depth residual). With the gyroscope bias removed, the realized bias drops below even the white-noise prediction at these grades. This is consistent with that prediction being a simplified-scenario indicator in a joint estimation problem rather than a point predictor. The prediction thus overstates the white-noise bias on its own and understates the full realized bias. Its conservatism is a property of the operational case, where the estimated gyroscope bias introduces an error that the closed-form solution does not model.

The remaining sources are not separated by the data. The simplified-scenario analysis of Section 5.2 treats the lever arm in isolation, whereas the trials estimate all parameters jointly, so coupling to the co-estimated parameters affects the realized bias. The DVL noise, the preintegration noise, and the preintegration discretization error enter additively rather than through the regressor, none of them multiplying the lever arm, so none adds to the EIV penalty, and their contribution to the realized bias is not characterized here.

The underestimate precludes post-hoc correction. A tight magnitude prediction would let the calibration be corrected by inverting the predicted attenuation. Since the prediction is conservative rather than tight, the diagnostic’s quantitative role is order-of-magnitude flagging, not point prediction of the post-calibration error.

5.5.5 Computing the Diagnostic in Practice

The flagging role requires a diagnostic that can be computed during a real calibration. The validation above used the oracle $\nu_{\min}^{(A)}$, which requires the true extrinsics and noise-free gyroscope samples, neither of which is available. The operational diagnostic $\hat{\nu}_{\min}^{(A)}$ is computed instead from the current extrinsic estimate and the actual noisy gyroscope samples, with no ground truth and no optimizer covariance.

The dependence on the estimate is structural. The lever arm correction term $b_i \hat{\omega}_{b_i} \times {}_b\mathbf{p}_{d/b}$ makes the residual’s gyroscope-noise sensitivity depend on the lever arm, so the residual covariance, the weight, and both $\mathbf{H}^{(A)}$ and $\mathbf{\Lambda}^{(A)}$ are evaluated at the current estimate ${}_b\hat{\mathbf{p}}_{d/b}$. The deployed quantity is the diagnostic at the estimate rather than at the truth, and its usefulness rests on $\nu_{\min}^{(A)}$ being insensitive to the estimate over the error range that arises in practice.

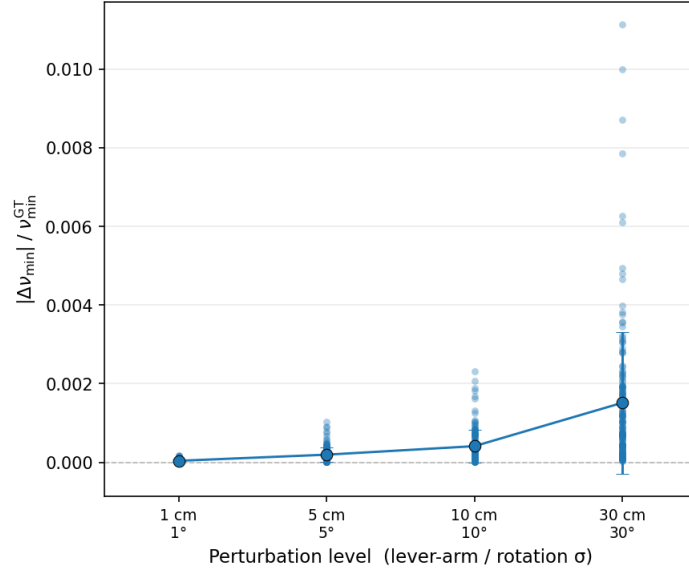


Figure 5.2: Robustness of $\nu_{\min}^{(A)}$ to extrinsic perturbation from the truth.

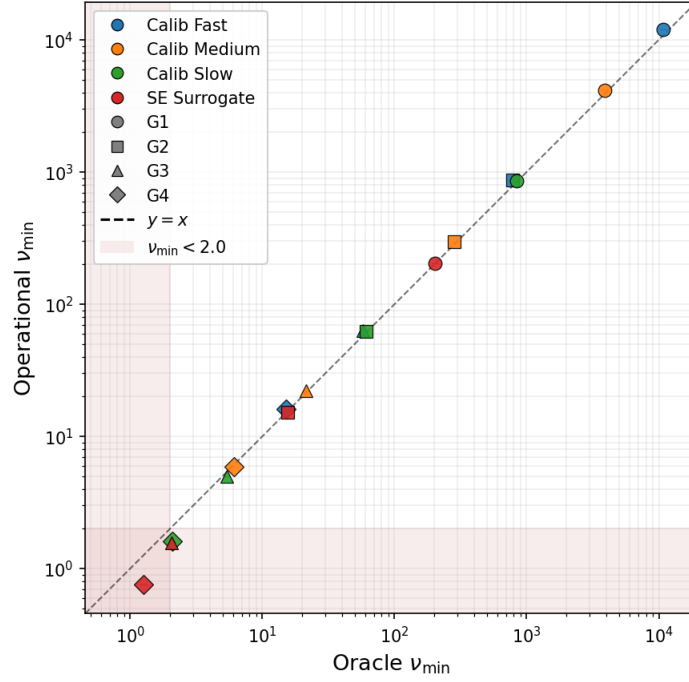


Figure 5.3: Operational $\hat{\nu}_{\min}^{(A)}$ against the oracle reference across the sixteen cells.

Figure 5.2 reports the relative change in $\nu_{\min}^{(A)}$ under a sweep of extrinsic perturbations from the truth. At a 10 cm and 10° perturbation, the change is below 0.5%, and at the 30 cm and 30° extreme, it reaches only about 1%. An initial estimate within this range, therefore, changes the diagnostic by at most 1%.

Figure 5.3 plots the operational $\hat{\nu}_{\min}^{(A)}$ from noisy gyroscope samples against the oracle

reference across the sixteen cells. The two agree within about 10% for $\nu_{\min}^{(A)} \gtrsim 5$. At the most saturated cells, the operational diagnostic falls below one, which the oracle cannot. The bound $\nu_{\min}^{(A)} \geq 1$ holds for the expected information; the realized $\hat{\mathbf{H}}^{(A)}$ fluctuates about that expectation, and at the saturated cells, where $\nu_{\min}^{(A)}$ sits near one, the fluctuation carries the realized value below it. Those are the cells the oracle already flags as saturated, so the dip below one loses no information.

A small $\nu_{\min}^{(A)}$ flags a calibration that is structurally exposed to EIV bias, and the optimizer does not reveal this on its own. The optimizer’s reported covariance describes the scatter of the estimate, not the systematic offset, so a well-converged calibration can still carry a large EIV bias. The operational diagnostic is computable from the calibration inputs alone, and it is stable enough to serve this screening role.

5.6 Summary

Formulation A removed AQUA-SLAM’s omitted-term collapse, a noise-independent bias that had made the IMU–DVL lever arm unrecoverable. A smaller residual bias remained, and it grew with the sensor noise. This chapter identified that residual bias as an errors-in-variables effect, where gyroscope noise enters the lever arm correction term as a noisy regressor and shrinks the least-squares estimate toward zero, direction by direction. The severity was reduced to a single scalar, the smallest generalized eigenvalue $\nu_{\min}^{(A)}$, which measures the information-to-noise ratio in the most-affected direction. Monte Carlo trials across four trajectories and four noise grades confirmed the mechanism: where the effect is significant, the realized bias aligns with the predicted direction, and at the lowest- $\nu_{\min}^{(A)}$ cells it climbs toward the lever arm magnitude as $\nu_{\min}^{(A)}$ approaches one. A control with the gyroscope bias held fixed identified the gyroscope bias estimation error as the dominant source of the excess over the white-noise prediction at the higher grades. The diagnostic is a conservative indicator that underpredicts the realized bias and is stable and computable at deployment.

The bias is structural, a property of where the expected cost is minimized, so no choice of weighting or solver removes it. By Theorem 5.2, the construction of each IMU residual cancels its leading rotation content while the gyroscope-noise perturbation does not cancel. The lever arm sensitivity survives only through the centripetal and angular-acceleration effects, scaled by the keyframe interval. These effects are weak in the smooth, slowly varying motion of underwater operation, so each residual contributes little deterministic information about the lever arm yet incurs the gyroscope-noise penalty $\mathbf{\Lambda}^{(A)}$ in full. This holds the deterministic part of $\mathbf{H}^{(A)}$ down against $\mathbf{\Lambda}^{(A)}$, which pins $\nu_{\min}^{(A)}$ toward its lower bound. The extent of the reduction depends on the rotation excitation and the gyroscope noise level. Following this principle, we next develop an alternative formulation, called Formulation B, that modifies the residual structure to preserve the leading rotation content. This construction directly counteracts the cancellation mechanism that drives $\nu_{\min}^{(A)}$ toward its lower bound, and is presented in the next chapter.

Formulation B: Velocity State Augmentation

6

This chapter develops Formulation B, which supplies the open body velocity as a per-keyframe state rather than by substitution. Keeping the velocity in the state removes the lever arm from the information-poor IMU residuals. The chapter evaluates whether this reduces the EIV bias that Chapter 5 diagnosed in Formulation A. Section 6.1 specifies Formulation B’s calibration state and the velocity supply that distinguishes it from Formulation A. Section 6.2 develops the DVL velocity residual that Formulation B reinstates. Section 6.3 applies the diagnostic of Chapter 5 to Formulation B and compares it with Formulation A. Section 6.4 evaluates the two formulations in simulation. Section 6.5 summarizes the comparison.

6.1 Calibration State and Velocity Supply

Formulation B augments Formulation A’s calibration state with the body velocity at every keyframe, expressed in the navigation frame c_0 ,

$$\mathcal{X}^{(B)} = \mathcal{X}^{(A)} \cup \{ {}_{c_0}\mathbf{v}_{b_i} \}_{i \in \mathcal{K}}, \quad (6.1)$$

where $\mathcal{X}^{(A)}$ is Formulation A’s state of (4.1), collecting the inter-sensor extrinsics, the gravity direction ${}_{c_0}\mathbf{g}$, and the per-keyframe biases. Where Formulation A supplies the body velocity by substituting the DVL measurement through the rigid-body relation (4.2), Formulation B keeps it in the state and lets the residuals constrain it.

The body-velocity state must be initialized before the residuals that depend on it become active. Formulation B reuses Formulation A’s coarse-to-fine pipeline (Section 4.3), where those residuals enter in the third stage. The velocities are initialized at the end of the second stage by rotating the DVL-measured velocities into the navigation frame,

$${}_{c_0}\mathbf{v}_{b_i} \leftarrow {}_{c_0}\mathbf{R}_{c_i b} \mathbf{R}_{c b}^T \mathbf{R}_{d b} \tilde{\mathbf{v}}_{d_i}, \quad (6.2)$$

through the zero lever arm transport that AQUA-SLAM uses. The DVL velocity residual enters the residual set in the same third stage, so the lever arm and the velocities are refined together from these initial values.

6.2 The DVL Velocity Residual and Its Covariance

The DVL velocity residual is the only residual Formulation B adds to the set. The IMU velocity and translation residuals are inherited from Chapter 2, given by (2.21) and (2.22). They now depend on the body velocity in the state rather than on a substituted measurement. They are not restated. Neither depends on the DVL measurement

or the keyframe gyroscope sample, so their covariances are the marginal preintegration blocks of [31].

The DVL velocity residual is the residual of (2.30), with the body velocity taken from the state and the angular velocity supplied by the bias-corrected gyroscope sample,

$$\mathbf{r}_i^{\text{DVL},v} = {}_b\mathbf{R}_{c\ c_0}\mathbf{R}_{c_i\ c_0}^T\mathbf{v}_{b_i} - ({}_b\mathbf{R}_{d\ d_i}\tilde{\mathbf{v}}_{d_i} - {}_{b_i}\hat{\boldsymbol{\omega}}_{b_i} \times {}_b\mathbf{p}_{d/b}). \quad (6.3)$$

The prediction transports the body-velocity state into the body frame at keyframe i . The measurement side rotates the DVL reading into the body frame and applies the rigid-body correction for the lever arm. The residual is expressed in the body frame, matching the DVL translation residual (2.29).

The residual depends on two independent raw measurements at keyframe i . The DVL reading enters through the additive term, and the gyroscope sample enters through the cross-product term. The residual is linear in each, so the noise propagates exactly through the Jacobians

$$\frac{\partial \mathbf{r}_i^{\text{DVL},v}}{\partial {}_{d_i}\tilde{\mathbf{v}}_{d_i}} = -{}_b\mathbf{R}_d, \quad \frac{\partial \mathbf{r}_i^{\text{DVL},v}}{\partial {}_{b_i}\hat{\boldsymbol{\omega}}_{b_i}} = -[{}_b\mathbf{p}_{d/b}]_{\times}, \quad (6.4)$$

using $\mathbf{y} \times \mathbf{z} = -[\mathbf{z}]_{\times}\mathbf{y}$. Propagating the DVL and gyroscope sample covariances $\boldsymbol{\Sigma}_{\text{DVL}}$ and $\boldsymbol{\Sigma}_{\text{Gyro}}$ gives the residual covariance

$$\boldsymbol{\Sigma}_{\mathbf{r}_i^{\text{DVL},v}} = {}_b\mathbf{R}_d \boldsymbol{\Sigma}_{\text{DVL}} {}_b\mathbf{R}_d^T + [{}_b\mathbf{p}_{d/b}]_{\times} \boldsymbol{\Sigma}_{\text{Gyro}} [{}_b\mathbf{p}_{d/b}]_{\times}^T. \quad (6.5)$$

The gyroscope contribution scales with the lever arm magnitude, so the lever arm appears in the covariance as well as in the residual. The covariance is evaluated at the linearization point ${}_b\tilde{\mathbf{p}}_{d/b}$ and held there, following the fixed-weight convention of Section 4.2.2, so that updates to the lever arm cannot lower the cost by inflating its own covariance.

6.3 The EIV Diagnostic under Formulation B

With the residual set specified, the EIV diagnostic of Chapter 5 can be applied to Formulation B. It is built from the aggregate information $\mathbf{H}$, the aggregate penalty $\boldsymbol{\Lambda}$, and the smallest generalized eigenvalue $\nu_{\min}$ of the pair $(\mathbf{H}, \boldsymbol{\Lambda})$. This section assembles those aggregates for Formulation B under the simplified scenario of Section 5.2, in which only the IMU–DVL lever arm is unknown, and compares them with Formulation A’s.

6.3.1 Per-Factor Contributions

Under Formulation B, the lever arm is observed by two residuals, the IMU translation residual and the DVL velocity residual. The IMU velocity residual observed it under Formulation A’s velocity substitution but no longer does, so it leaves the analysis. The two contributions differ in one respect: only the DVL velocity residual has a noisy regressor.

IMU translation residual. The IMU translation residual depends on the IMU-DVL lever arm through the IMU-camera lever arm. Its regressor is the rotation difference $\mathbf{I} - \mathbf{R}_{ij}$ and its perturbation is zero,

$$\bar{\mathbf{A}}_{ij}^{p,B} = \mathbf{I} - \mathbf{R}_{ij}, \quad \mathbf{R}_{ij} := {}_{b_i}\mathbf{R}_{b_j}, \quad \delta \mathbf{A}_{ij}^{p,B} = \mathbf{0}, \quad (6.6)$$

where $\bar{\mathbf{A}}_{ij}^{p,B}$ and $\delta \mathbf{A}_{ij}^{p,B}$ refer to the deterministic and perturbed parts of the regressor of the IMU-DVL lever arm over interval (i, j) . The corresponding penalty $\mathbf{\Lambda}_{ij}^{p,B}$, is then zero, and the deterministic information $\bar{\mathbf{H}}_{ij}^{p,B}$ is

$$\mathbf{\Lambda}_{ij}^{p,B} = \mathbf{0}, \quad \bar{\mathbf{H}}_{ij}^{p,B} = (\mathbf{I} - \mathbf{R}_{ij})^T \mathbf{W}_{ij}^{p,B} (\mathbf{I} - \mathbf{R}_{ij}), \quad (6.7)$$

with $\mathbf{W}_{ij}^{p,B} = \Sigma_{\Delta \mathbf{p}_{b_j/b_i}}^{-1}$ the marginal IMU translation preintegration weight. The information is nonzero whenever $\mathbf{R}_{ij} \neq \mathbf{I}$.

DVL velocity residual. With the bias-corrected sample in (6.3) decomposed into truth and noise, ${}_{b_i}\hat{\boldsymbol{\omega}}_{b_i} = {}_{b_i}\boldsymbol{\omega}_{b_i}^* + \boldsymbol{\eta}_i^g$, and the cross product written as $[{}_{b_i}\hat{\boldsymbol{\omega}}_{b_i}]_{\times} {}_b\mathbf{p}_{d/b}$. Then the regressor of IMU-DVL lever arm in DVL velocity residual, denoted by $\mathbf{A}_i^{\text{DVL},v}$, splits into a deterministic part and a perturbation,

$$\bar{\mathbf{A}}_i^{\text{DVL},v} = [{}_{b_i}\boldsymbol{\omega}_{b_i}^*]_{\times}, \quad \delta \mathbf{A}_i^{\text{DVL},v} = [\boldsymbol{\eta}_i^g]_{\times}. \quad (6.8)$$

The deterministic part scales with the keyframe angular velocity, and the perturbation is the per-sample gyroscope noise in skew-symmetric form. The per-keyframe inverse covariance then weights the deterministic regressor, giving the deterministic information

$$\bar{\mathbf{H}}_i^{\text{DVL},v} = [{}_{b_i}\boldsymbol{\omega}_{b_i}^*]_{\times}^T \mathbf{W}_i^{\text{DVL},v} [{}_{b_i}\boldsymbol{\omega}_{b_i}^*]_{\times}, \quad \mathbf{W}_i^{\text{DVL},v} = \Sigma_{\mathbf{r}_i^{\text{DVL},v}}^{-1}. \quad (6.9)$$

The DVL velocity residual carries a single gyroscope sample, so its penalty is the expectation

$$\mathbf{\Lambda}_i^{\text{DVL},v} = \mathbb{E}[[\boldsymbol{\eta}_i^g]_{\times}^T \mathbf{W}_i^{\text{DVL},v} [\boldsymbol{\eta}_i^g]_{\times}]. \quad (6.10)$$

This expectation is an instance of the cross-product expectation identity of Appendix B. Under isotropic gyroscope noise $\boldsymbol{\eta}_i^g \sim \mathcal{N}(\mathbf{0}, \sigma_g^2 \mathbf{I})$, its transposed form (B.5) applies with the symmetric weight $\mathbf{W}_i^{\text{DVL},v}$ and gives

$$\mathbf{\Lambda}_i^{\text{DVL},v} = \sigma_g^2 (\text{tr}(\mathbf{W}_i^{\text{DVL},v}) \mathbf{I} - \mathbf{W}_i^{\text{DVL},v}). \quad (6.11)$$

Summing over the calibration window gives the aggregates

$$\mathbf{H}^{(B)} = \sum_{ij} \bar{\mathbf{H}}_{ij}^{p,B} + \sum_i (\bar{\mathbf{H}}_i^{\text{DVL},v} + \mathbf{\Lambda}_i^{\text{DVL},v}), \quad \mathbf{\Lambda}^{(B)} = \sum_i \mathbf{\Lambda}_i^{\text{DVL},v}. \quad (6.12)$$

The penalty $\mathbf{\Lambda}^{(B)}$ collects the DVL velocity contributions alone, since the IMU translation residual adds none. Whether Formulation B's bias is smaller than Formulation A's depends on the information relative to the penalty, not on the penalty alone. The next subsection makes the comparison through $\nu_{\min}$.

6.3.2 Comparison with Formulation A

Two structural changes take Formulation A’s aggregates (5.23) to Formulation B’s (6.12), and both raise $\nu_{\min}$. Both follow from holding the body velocity in the state rather than substituting the DVL measurement. The first removes a penalty term, and the second restores information that the differencing had canceled.

The first change removes the gyroscope-noise penalty from the IMU translation residual. In Formulation A, the gyroscope sample entered this residual through the lever arm correction term, which supplied the perturbation $\delta \mathbf{A}_{ij}^p$ of (5.11) and with it the penalty Λ_{ij}^p . With the body velocity held in the state, the correction term no longer appears. The lever arm then enters the residual only through the camera–body lever arm, so its regressor is the rotation difference $\mathbf{I} - \mathbf{R}_{ij}$ of (6.6). This regressor is free of gyroscope noise, so the perturbation and the penalty $\Lambda_{ij}^{p,B}$ both vanish.

The second change moves the velocity constraint from the differenced IMU velocity residual to the per-keyframe DVL velocity residual. In Formulation A, the IMU velocity residual differenced the lever arm correction terms at the two endpoint keyframes, and by Theorem 5.2, this differencing canceled the zeroth-order rotation content and left the lever arm sensitivity in the weak centripetal and angular-acceleration effects. The DVL observes absolute velocity, so the DVL velocity residual constrains the lever arm at a single keyframe and needs no differencing. Its regressor is the angular velocity $[_{b_i}\boldsymbol{\omega}_{b_i}^\star]_\times$ of (6.8), exactly the zeroth-order content that the differencing in Formulation A’s IMU residuals had canceled. The information (6.9) then accrues at every keyframe and scales with the keyframe angular velocity.

Together, the two changes raise the ratio of information to penalty. The IMU translation residual now contributes information free of any penalty, and the DVL velocity residual contributes information at the same order as its penalty rather than below it. This predicts $\nu_{\min}^{(B)} > \nu_{\min}^{(A)}$ for the same trajectory and sensor grade, and with it a weaker attenuation along the most-affected direction (5.26). The bias is reduced but not removed, because the DVL velocity residual still depends on the gyroscope sample as a noisy regressor, so the penalty $\Lambda^{(B)}$ is nonzero.

6.4 Empirical Comparison: Formulation A versus Formulation B

This section evaluates Formulation B against Formulation A by Monte Carlo simulation on the Calib Medium and SE surrogate trajectories of Section 3.3.2. Four results are reported: the structural improvement in the diagnostic (Section 6.4.1), the calibration improvement on the parameters coupled to the EIV mechanism in Formulation A (Section 6.4.2), the EIV-uncoupled parameters (Section 6.4.3), and the non-EIV bias that Formulation B exposes, traced to its sources by a gyroscope bias control (Section 6.4.4).

The four sensor noise grades, G1 to G4, and the two trajectories define eight (trajectory, grade) cells. The simulation follows the Monte Carlo protocol of Section 3.3.4, with $N = 200$ trials per cell. Both formulations jointly estimate all calibration parameters under realistic noise on each sensor. Two further control runs repeat the comparison on both trajectories with the gyroscope bias held fixed at its true value

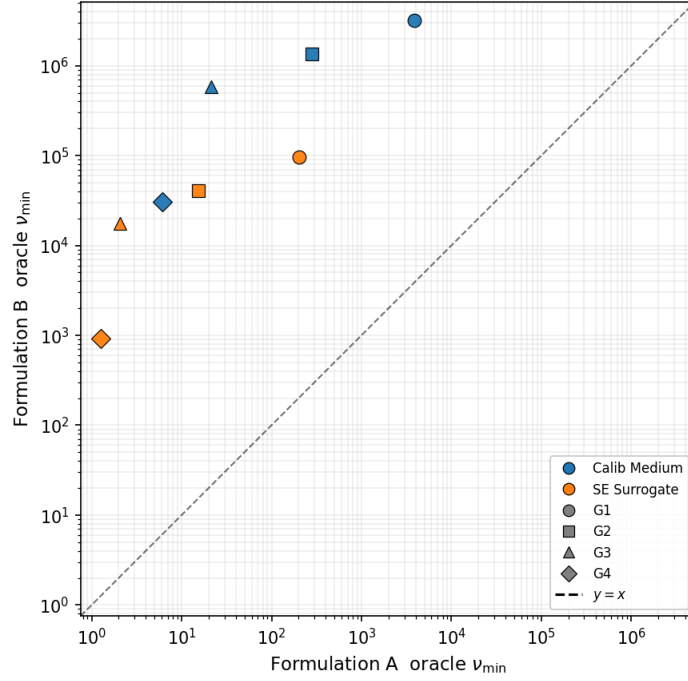


Figure 6.1: Oracle $\nu_{\min}^{(B)}$ versus $\nu_{\min}^{(A)}$ across the eight (trajectory, grade) cells, log-log axes. The dashed line marks $y = x$.

rather than estimated, isolating the gyroscope bias estimation error.

Tables 6.1 and 6.2 report the bias and per-trial scatter for the extrinsic parameters and the gravity direction of the state (4.1) under both formulations.

6.4.1 Diagnostic Improvement

Figure 6.1 plots the oracle $\nu_{\min}^{(B)}$ against $\nu_{\min}^{(A)}$ across the eight (trajectory, grade) cells. Every cell sits well above the identity line ($y = x$ reference). The oracle $\nu_{\min}^{(B)}$ exceeds $\nu_{\min}^{(A)}$ by between roughly three and four orders of magnitude at every cell. The cells where Formulation A is most saturated, the SE surrogate at G3 and G4 with $\nu_{\min}^{(A)}$ near its lower bound, move into the EIV-negligible regime under Formulation B, with $\nu_{\min}^{(B)}$ above 100, the negligible threshold of Chapter 5. The two structural changes of Section 6.3.2 both raise $\nu_{\min}$, and together they account for a gain of several orders of magnitude.

6.4.2 EIV-Coupled Parameters Improve

Two parameters are coupled to the EIV mechanism in Formulation A: the IMU–DVL lever arm ${}_b\mathbf{p}_{d/b}$, on which the penalty acts directly, and the gravity direction ${}_{c_0}\mathbf{g}$, which couples to the lever arm through the joint calibration. The same coupling appeared between AQUA-SLAM and Formulation A in Section 4.4.2. Both improve under Formulation B.

Table 6.1: Extrinsic-error comparison for Calib Medium: Formulation A (velocity substitution) versus Formulation B (velocity state). Each entry reports the bias magnitude $\|\bar{\mathbf{e}}\|$ and the per-trial 1σ scatter $\sigma_{\|\mathbf{e}\|}$ over the Monte Carlo trials.

Grade	Form	${}_b\mathbf{R}_c$ [deg]	${}_c\mathbf{R}_d$ [deg]	${}_c\mathbf{p}_{d/c}$ [cm]	${}_b\mathbf{p}_{d/b}$ [cm]	${}_c\mathbf{p}_{p/c}$ [cm]	${}_{c_0}\mathbf{g}$ [deg]
G1	A	0.036 ± 0.003	0.041 ± 0.053	0.211 ± 0.115	0.173 ± 0.401	0.054 ± 0.988	0.031 ± 0.060
	B	0.035 ± 0.003	0.062 ± 0.053	0.259 ± 0.114	0.711 ± 0.166	0.052 ± 0.989	0.036 ± 0.024
G2	A	0.036 ± 0.010	0.051 ± 0.270	0.206 ± 0.588	0.737 ± 1.919	0.099 ± 2.469	0.083 ± 0.236
	B	0.033 ± 0.010	0.068 ± 0.269	0.250 ± 0.581	0.717 ± 0.627	0.065 ± 2.472	0.033 ± 0.051
G3	A	0.037 ± 0.039	0.108 ± 1.384	0.202 ± 3.012	5.495 ± 8.579	0.223 ± 5.924	0.251 ± 0.643
	B	0.029 ± 0.036	0.102 ± 1.380	0.217 ± 2.982	0.691 ± 3.007	0.102 ± 5.933	0.037 ± 0.107
G4	A	0.037 ± 0.077	0.139 ± 1.270	0.234 ± 2.763	13.286 ± 12.553	0.806 ± 11.772	1.206 ± 1.443
	B	0.029 ± 0.072	0.103 ± 1.265	0.270 ± 2.732	0.792 ± 2.904	0.173 ± 11.865	0.050 ± 0.215

Table 6.2: Extrinsic-error comparison for the SE surrogate: Formulation A (velocity substitution) versus Formulation B (velocity state). Each entry reports the bias magnitude $\|\bar{\mathbf{e}}\|$ and the per-trial 1σ scatter $\sigma_{\|\mathbf{e}\|}$ over the Monte Carlo trials.

Grade	Form	${}_b\mathbf{R}_c$ [deg]	${}_c\mathbf{R}_d$ [deg]	${}_c\mathbf{p}_{d/c}$ [cm]	${}_b\mathbf{p}_{d/b}$ [cm]	${}_c\mathbf{p}_{p/c}$ [cm]	${}_{c_0}\mathbf{g}$ [deg]
G1	A	0.033 ± 0.012	0.014 ± 0.087	0.517 ± 0.593	0.530 ± 1.386	0.167 ± 3.198	0.023 ± 0.013
	B	0.027 ± 0.011	0.009 ± 0.071	0.519 ± 0.485	0.972 ± 0.556	0.168 ± 3.197	0.023 ± 0.004
G2	A	0.031 ± 0.046	0.038 ± 0.384	0.410 ± 2.619	4.816 ± 7.530	0.361 ± 8.000	0.026 ± 0.052
	B	0.018 ± 0.035	0.022 ± 0.359	0.361 ± 2.463	0.873 ± 2.534	0.353 ± 7.993	0.021 ± 0.010
G3	A	0.027 ± 0.171	0.130 ± 1.873	0.864 ± 12.770	17.674 ± 23.206	0.841 ± 19.248	0.038 ± 0.221
	B	0.008 ± 0.102	0.097 ± 1.842	0.615 ± 12.470	0.757 ± 12.545	0.824 ± 19.184	0.020 ± 0.025
G4	A	0.032 ± 0.337	0.114 ± 1.718	0.710 ± 11.700	26.835 ± 16.675	1.663 ± 38.330	0.048 ± 0.227
	B	0.018 ± 0.205	0.092 ± 1.686	2.034 ± 11.026	2.796 ± 11.246	1.655 ± 38.307	0.021 ± 0.050

Figure 6.2 compares the realized IMU–DVL lever arm bias magnitude $\|{}_b\hat{\mathbf{p}}_{d/b} - {}_b\mathbf{p}_{d/b}^*\|$ under the two formulations across the eight cells. The reduction follows the EIV regimes that Chapter 5 assigned to these trajectories. At the cells where the EIV effect is significant or saturated, Calib Medium at G3 and G4 and the SE surrogate at G2 to G4, Formulation A’s bias ranges from about 5 to 27 cm, and Formulation B reduces it by factors of about five to twenty-three: Calib Medium G4 drops from 13.3 to 0.8 cm and the SE surrogate G4 from 26.8 to 2.8 cm. At the EIV-negligible cells, Calib Medium at G1 and G2 and the SE surrogate at G1, Formulation A’s bias is already at the sub-centimeter non-EIV floor of Section 6.4.4, and Formulation B does not reduce it; at the two G1 cells, it is marginally higher, both below one centimeter.

Formulation B also lowers the per-trial scatter on the lever arm in every cell, for example, from 12.6 to 2.9 cm for Calib Medium G4. The restored sensitivity that removes the structural cancellation of Theorem 5.2 also sharpens the cost along the lever arm direction, so the gain appears in the spread of single-run estimates as well as in the bias.

The gravity direction shows the same coupling on Calib Medium. Formulation A’s

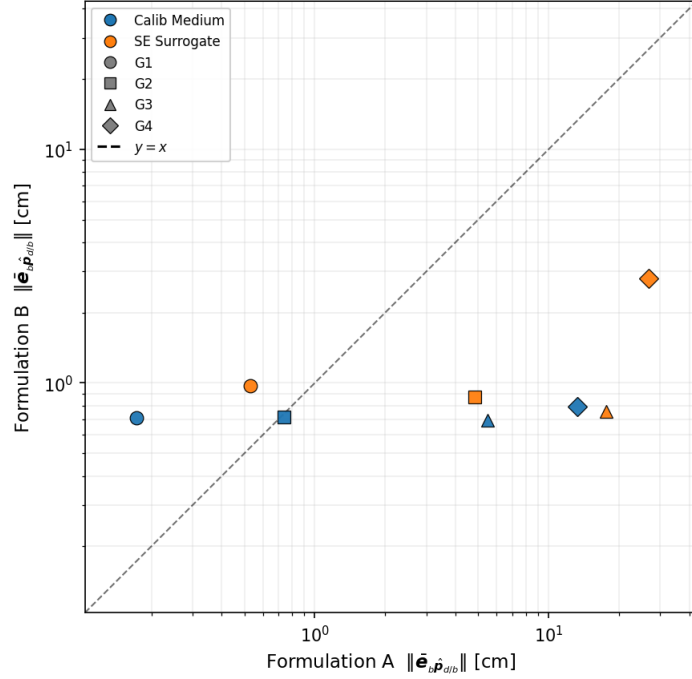


Figure 6.2: Realized IMU–DVL lever arm bias magnitude under Formulation B versus Formulation A, log-log axes, across the eight (trajectory, grade) cells. Points below $y = x$ mark Formulation B’s improvement.

gravity bias there grows from 0.03° at G1 to 1.2° at G4, tracking the growth of its lever arm bias, while Formulation B’s stays between 0.03° and 0.05° across all grades (Tables 6.1 and 6.2). On the SE surrogate, Formulation A’s gravity bias stays below 0.05° at every grade despite the much larger lever arm bias, so the coupling is visible only on Calib Medium. When the EIV bias drives the lever arm away from its true value, part of the error is absorbed into the gravity direction, and when Formulation B suppresses that bias, the gravity direction recovers with the lever arm.

6.4.3 EIV-Uncoupled Parameters

The parameters not coupled to the EIV mechanism are recovered in the same way under both formulations. The two rotation extrinsics ${}_b\mathbf{R}_c$ and ${}_c\mathbf{R}_d$, the camera–DVL lever arm ${}_c\mathbf{p}_{d/c}$, and the camera–pressure lever arm ${}_c\mathbf{p}_{p/c}$ agree in both bias and per-trial scatter across all cells (Tables 6.1 and 6.2). None is subject to the noisy regressor that drives the EIV bias, so the formulation change does not affect them measurably. The one visible bias difference, the camera–DVL lever arm at SE surrogate G4, lies well within that cell’s per-trial scatter.

6.4.4 A Non-EIV Bias Remains

Figure 6.2 shows that Formulation B does not drive the IMU–DVL lever arm bias to zero. A residual remains, near 0.7 cm on Calib Medium across all grades and at a

Table 6.3: IMU–DVL lever arm bias $\|_b \bar{e}_{d/b}\|$ under Formulations A and B, with the gyroscope bias estimated and with it held fixed at its true value (cm, mean $\pm$ per-trial 1σ over $N = 200$ trials).

	Gyro bias	G1	G2	G3	G4
<i>Calib Medium</i>					
A	estimated	0.17 ± 0.40	0.74 ± 1.92	5.50 ± 8.58	13.29 ± 12.55
	fixed	0.31 ± 0.24	0.28 ± 1.20	0.38 ± 6.25	2.22 ± 6.12
B	estimated	0.71 ± 0.17	0.72 ± 0.63	0.69 ± 3.01	0.79 ± 2.90
	fixed	0.42 ± 0.13	0.43 ± 0.59	0.41 ± 2.98	0.48 ± 2.77
<i>SE surrogate</i>					
A	estimated	0.53 ± 1.39	4.82 ± 7.53	17.67 ± 23.21	26.84 ± 16.68
	fixed	0.32 ± 1.10	0.45 ± 5.49	5.93 ± 28.44	18.29 ± 22.82
B	estimated	0.97 ± 0.56	0.87 ± 2.53	0.76 ± 12.55	2.80 ± 11.25
	fixed	0.83 ± 0.56	0.72 ± 2.52	0.59 ± 12.51	2.57 ± 11.19

comparable level on the SE surrogate at G1 to G3, rising to about 2.8 cm at the SE surrogate G4. With $\nu_{\min}^{(B)}$ in the hundreds or above at every cell, the worst-direction attenuation $1/\nu_{\min}^{(B)}$ is at the 10^{-3} level or below, which places any EIV-attributable bias orders of magnitude under this residual. The residual is therefore non-EIV in origin.

This residual belongs to the class of non-EIV effects identified for Formulation A in Section 5.5.4: contributions beyond the white-noise bias that the closed-form EIV prediction does not cover. Chapter 5 identified gyroscope bias estimation error as the dominant such contribution under Formulation A. Table 6.3 isolates it under both formulations by holding the gyroscope bias at its true value instead of estimating it. The change between the estimated and fixed rows is the gyroscope bias contribution to the lever arm bias. Under Formulation A, this contribution grows with the grade, reaching 11.1 cm at Calib Medium G4 and 8.5 cm at SE surrogate G4. Under Formulation B, it is flat across all grades, near 0.3 cm on Calib Medium and below 0.25 cm on the SE surrogate. The gyroscope bias error enters through the same gyroscope sample as the white noise that the penalty models. The restored sensitivity of Section 6.3.2 therefore de-amplifies it for the same reason it suppresses the white-noise EIV bias: the lever arm is now read from the strong angular-velocity regressor rather than the weak post-cancellation one.

Fixing the gyroscope bias under Formulation A collapses the Calib Medium bias to below the white-noise prediction. On the saturated SE surrogate, it instead leaves a large residual, because the white-noise EIV bias dominates that trajectory, so the gyroscope bias contribution is a smaller part of Formulation A’s bias there than on Calib Medium. Even with the gyroscope bias fixed, Formulation B retains a sub-centimeter residual in most cells and about 2.8 cm in the SE surrogate G4. This non-gyroscope bias residual, including the SE surrogate G4 elevation, lies within the large per-trial scatter at the higher grades, and its source is not separated by these results.

6.5 Summary

Formulation B keeps the body velocity in the calibration state rather than substituting it from the DVL. The lever arm then leaves the insensitive IMU residuals and is observed via the reinstated DVL velocity residual, which constrains it at each keyframe using the keyframe angular velocity rather than a velocity difference. By the diagnostic of Chapter 5 this predicts a $\nu_{\min}$ above Formulation A's.

The simulation confirms the effect. Across the tested cells $\nu_{\min}^{(B)}$ exceeds $\nu_{\min}^{(A)}$ by between three and four orders of magnitude. The IMU–DVL lever arm bias falls by factors of about five to twenty-three, where Formulation A's bias was EIV-dominated, its per-trial scatter falls with it, and the coupled gravity direction recovers. The parameters not coupled to the EIV mechanism are recovered with equivalent precision, within per-trial scatter. A centimeter-level non-EIV residual remains; a gyroscope bias control shows that the velocity-state augmentation de-amplifies the gyroscope bias estimation error that the EIV mechanism amplified under Formulation A, so this residual is not the EIV bias.

The comparison in this chapter rests on simulation, where the per-sensor noise follows the assumed model. Whether Formulation B's reduction of the IMU–DVL lever arm bias survives on real data, where that model holds only approximately, and the camera trajectory is itself imperfect, is the question Chapter 7 takes up.

Real-World Calibration

This chapter applies the three calibration formulations of the preceding chapters to real data: AQUA-SLAM’s baseline formulation (Chapter 3), Formulation A (Chapter 4), and Formulation B (Chapter 6). The application to real data requires one additional step beyond the simulation protocol: a calibration window on which the assumed noise model holds. Section 7.1 describes the platform, the camera trajectory input, and the protocol shared by the three formulations. Section 7.2 presents the selection of the calibration window and validates it. Section 7.3 compares the three formulations on the selected window. Section 7.4 reports the convergence of the calibration as keyframes accumulate and its run time. Section 7.5 validates the corrected velocity transport and the recovered calibration within the SLAM estimator. Section 7.6 discusses the findings.

7.1 Calibration Pipeline

This section specifies what is held constant across the three formulations: the platform and sequence, the camera trajectory input, and the solver protocol.

7.1.1 Platform, Sensors, and Sequences

The Tank dataset [29] provides underwater sequences recorded on a BlueROV2 equipped with the GX5-AHRS IMU, the WaterLinked A50 DVL, the Bar30 depth sensor, and a custom underwater stereo camera. Sensor specifications, rates, and noise characteristics correspond to the G3 grade of Section 3.3, the configuration used in AQUA-SLAM’s released implementation [34].

Calibration uses the Structure Easy sequence. Its survey-type motion, sustained low-amplitude rotation on all three body axes with bursty linear acceleration and no targeted per-axis excitation, is characterized in Section 3.3. The vehicle maneuvers around an underwater structure covered with AprilTag markers, from which the trajectory ground truth is computed.

7.1.2 Camera Trajectory

The camera trajectory used as the fixed input for the calibration is produced by an AQUA-SLAM run in stereo-only mode: the IMU and DVL measurements are not fused into the trajectory estimate, the depth sensor is not used, and loop closure is disabled. In this mode, AQUA-SLAM estimates the trajectory from the stereo images using local bundle adjustment and outputs a per-keyframe camera trajectory that the three calibration formulations consume. The AprilTag detections are not used at this step, so the trajectory input remains independent of the ground truth. The trajectories differ

from those in the simulation due to noise. Their estimation error is not modeled, and it propagates into the recovered extrinsics as an error source that the formulations cannot separate from sensor noise.

7.1.3 Protocol

The three formulations are applied as specified in Sections 3.1, 4.1, and 6.1, without modification. We exclude the depth residual from all three formulations to ensure a fair comparison; consequently, the pressure-sensor lever arm is not estimated. The Tank dataset does not provide a reference against which its recovery could be evaluated, and the sequences contain almost no vertical motion, so it is unclear whether the depth residual could constrain this lever arm at all. Each formulation uses its own coarse-to-fine initialization: AQUA-SLAM follows its baseline initialization, Formulation A the pipeline of Section 4.3, and Formulation B the same pipeline extended with the velocity-state initialization of Section 6.1. Random initialization of the calibration state, Ceres with Levenberg-Marquardt, and default tolerances are employed unchanged from the Monte Carlo protocol of Section 3.3.

7.2 Calibration Window Selection

The calibration formulations assume independent and identically distributed zero-mean Gaussian per-beam DVL noise. However, in the recorded log, this assumption holds only for part of the sequence, so the calibration window must be selected. AQUA-SLAM likewise calibrates on a window of this sequence, but the window and its selection are not reported [7], so selection is an unaddressed prerequisite for reproducing the experiment. This section selects the window by gating on two statistics of a motion-independent DVL residual.

7.2.1 The DVL Error Velocity

A four-beam Janus DVL measures four projections of the body-frame velocity onto known beam directions. These geometric relationships are captured by the beam matrix $\mathbf{E} \in \mathbb{R}^{4 \times 3}$, where each row corresponds to the unit direction vector of one acoustic beam expressed in the body frame. Since there are four measurements but only three velocity degrees of freedom, it has rank three, so its left null space is one-dimensional; let $\mathbf{n} \in \mathbb{R}^4$ be its unit-norm generator, determined by the beam geometry alone. The *error velocity* $\hat{\zeta}_k := \mathbf{n}^T \tilde{\mathbf{b}}_k$, the projection of the beam-velocity vector $\tilde{\mathbf{b}}_k \in \mathbb{R}^4$ at ping k onto $\mathbf{n}$, annihilates the body-velocity contribution entirely. So, under the standard noise model, it is a zero-mean Gaussian with per-beam standard deviation σ_b , independent of vehicle motion. Any deviation from this prediction, therefore, reveals a violation of the noise model. Importantly, this construction requires only the beam velocities and the ping timestamps, without any dependence on calibration parameters or external ground truth. Appendix A develops the error velocity in detail.

Each error-velocity sample decomposes into a systematic offset and random scatter, and the selection tests the two separately. The scatter is estimated once, globally: the

per-beam noise standard deviation $\hat{\sigma}_b$ is the median absolute deviation of the error velocity over the full log, scaled to the Gaussian standard deviation by $1/\Phi^{-1}(0.75) \approx 1.4826$, where Φ^{-1} is the standard normal quantile function [35]. On Structure Easy this gives $\hat{\sigma}_b = 2.07 \times 10^{-3}$ m/s. The estimate is global rather than per-window because a contaminated window inflates its own estimate, thereby hiding its own outliers, so the reference standard deviation must not adapt to the window under test. The median absolute deviation tolerates up to half the samples being contaminated, and the contaminated portions of this log are well below that, so the global estimate remains anchored by the clean majority.

7.2.2 Two Gates and the Healthy Span

The log is scanned with sliding windows of 20 s duration at a stride of 2.5 s, and two statistics are computed for each window. The first is the outlier fraction: the fraction of samples whose deviation from the window median exceeds $3\hat{\sigma}_b$, which tests the Gaussian tails after centering. A window passes if its fraction does not exceed twice the median fraction across all scanned windows.

The second statistic is the magnitude of the window median itself, which tests the centering that the first statistic removes: the outlier fraction is blind by construction to a window whose samples scatter narrowly about a nonzero offset. Under the noise model, the median of the P pings in a window is asymptotically Gaussian with standard deviation $\sqrt{\pi/2} \hat{\sigma}_b / \sqrt{P}$ [36], so a window passes if its median magnitude does not exceed three of these; the threshold follows from the model rather than from a choice, and adapts to the window’s sample count and the sensor’s noise level. A bias common to all four beams is annihilated by the null-space projection just as the velocity is, so a failing median reflects per-beam asymmetry, a deterministic violation of the assumed beam geometry within the window. This is also the statistic through which sustained single-beam anomalies surface: a beam that persistently returns a degraded velocity, from a weak or corrupted acoustic return, shifts the window median, whereas a transient anomaly on any beam lands in the outlier fraction instead.

Two further statistics are computed but not gated. The lag-one autocorrelation of the error velocity tests independence across pings. Because correlation enters the DVL residuals additively rather than through a regressor, it leaves the calibration estimate consistent and only renders the reported covariance optimistic. It is therefore reported as a caveat on the estimated uncertainty rather than used for selection. The timing jitter of the ping schedule is computed and found roughly constant across the log, so it provides no discrimination on this log and is not used. Formal definitions of all computed statistics appear in Appendix D.

Figure 7.1 shows the two gated statistics, the autocorrelation, and the jitter across the log. Both gates flag the start and the end of the log and remain low through the middle, with transitions on the order of a few seconds. On this log, the two gates flag largely the same intervals, although they remain structurally distinct. Since the median gate detects uniform offsets, the outlier fraction is blind. A sample is accepted if every scan window covering it passes both gates, which erodes the passing run of window centers by half a window duration on each side; the resulting healthy span is [60, 130] s.

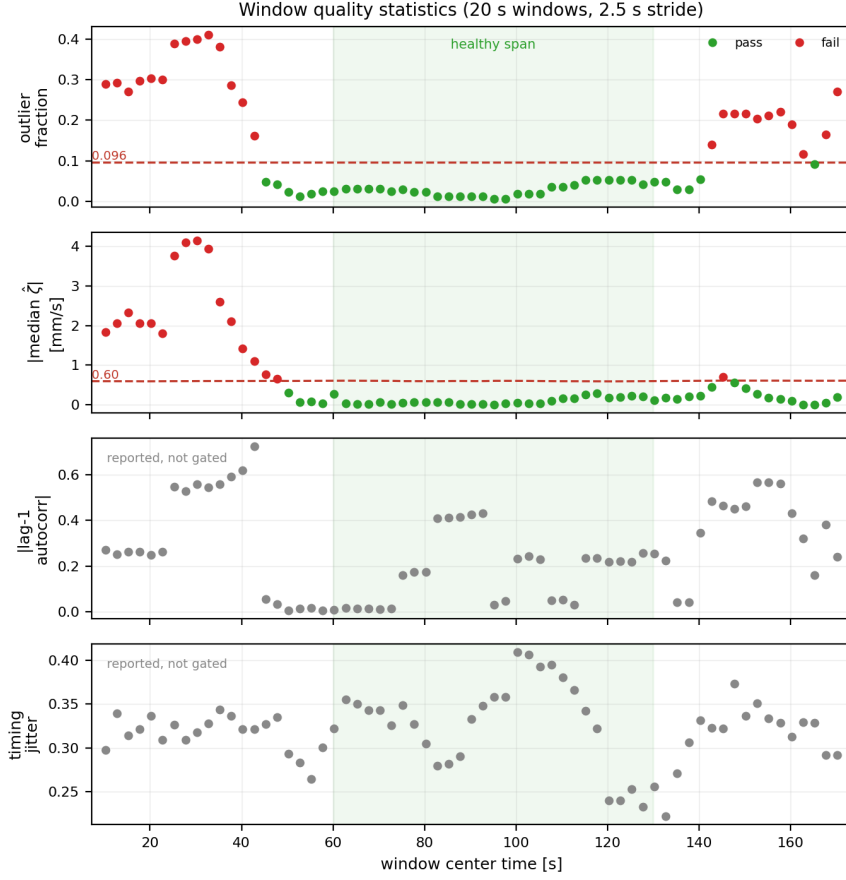


Figure 7.1: Window-quality statistics across the log, computed on 20 s windows at 2.5 s stride. The top two panels are gated: a window passes when its outlier fraction is below the dashed threshold (twice the inter-window median), and its median error velocity is below the per-window threshold, which scales as the inverse square root of the ping count. The lower two panels are reported but not used for selection. Points are colored by whether the window passes both gates, and the shaded band is the resulting healthy span after erosion.

The healthy span covers 70 s against a 60 s calibration-window target, so placement is nearly determined; we take the 60 s window centered in the span, $[65, 125]$ s, which covers 184 keyframes. The 60 s target reflects the keyframe count: a 20 s window covers roughly 60 keyframes, too few for a well-constrained calibration on a survey-type sequence. This window is used for the three-way comparison of Section 7.3.

The selection depends only on the beam geometry and the ping stream: it consults neither the extrinsic parameters nor the calibration residuals, so it cannot bias the calibration toward any solution. The error velocity and $\hat{\sigma}_b$ likewise serve selection only. The same construction bounds what the selection can see: it tests only the DVL’s own noise model, and a fault confined to the camera trajectory or the IMU is invisible to it.

Table 7.1: Calibration errors on the selected window [65, 125] s (184 keyframes) against the reference calibration provided with the Tank dataset. Rotation errors are axis-angle magnitudes in degrees; lever arm errors are in centimeters.

		AQUA-SLAM	Formulation A	Formulation B
${}_b\mathbf{R}_c$	[°]	0.30	0.29	0.15
${}_b\mathbf{R}_d$	[°]	1.05	0.87	0.53
${}_c\mathbf{R}_d$	[°]	0.78	0.60	0.49
${}_b\mathbf{p}_{c/b}$	[cm]	13.33	5.25	1.70
${}_b\mathbf{p}_{d/b}$	[cm]	14.83	8.13	4.24
${}_c\mathbf{p}_{d/c}$	[cm]	2.32	3.34	2.97

7.3 Three-Way Formulation Comparison

The three formulations are applied to the selected window under the protocol of Section 7.1, and the recovered extrinsics are compared against the reference calibration provided with the dataset. Table 7.1 reports the errors: rotations as the axis-angle magnitude in degrees, lever arms in centimeters.

The IMU–DVL lever arm ${}_b\mathbf{p}_{d/b}$ reproduces the cumulative pattern predicted by the synthetic experiments. The reference value provided with the dataset has a magnitude of 18.4 cm, and AQUA-SLAM’s error of 14.83 cm likely represents the collapsed lever arm error plus estimation scatter, consistent with the collapse analysis in Chapter 3: with the correction term omitted, the data do not constrain this lever arm. Restoring the omitted term, Formulation A reduces the error to 8.13 cm, and Formulation B to 4.24 cm. The IMU–camera lever arm ${}_b\mathbf{p}_{c/b}$ follows the same chain, 13.33, 5.25, and 1.70 cm, as the rigid-body identity ties its recovery to the other two extrinsic pairs.

The camera–DVL lever arm behaves differently, and as predicted, all three formulations recover it to between 2.3 and 3.4 cm. This parameter is observed through the DVL translation residual, which all three formulations share, and it does not enter the residuals that the corrections of Formulation A and Formulation B modify, so the formulations have no mechanism by which to differ on it.

The rotation extrinsics recover to within a degree under all three formulations, with differences of a few tenths of a degree between them. This is consistent with the synthetic finding of Chapter 6 that the corrections leave rotation recovery unchanged, so the comparison supports no ranking on the rotations.

The recovered errors can be read against the SE surrogate of Chapter 6 (Table 6.2), which matches the excitation of the real Structure Easy sequence and uses the platform’s G3 noise grade. The surrogate carries a larger IMU–DVL lever arm than the platform, 35 cm in Table 3.1 against the 18.4 cm reference. Since both EIV bias and the per-trial scatter scale with that magnitude, the lever-arm comparison is made in relative terms. For rotations, there is no such scaling, so they are compared in degrees. Under Formulation B, the surrogate gives each parameter a near-zero bias and a wide per-trial scatter, the latter being the imprint of the weak survey excitation. A single real calibration is therefore one draw from that distribution, and an error comparable

to the per-trial scatter reflects that scatter rather than a systematic offset.

On the real Structure Easy sequence, the recovered errors bear this out. The IMU–DVL error is about a quarter of the reference, smaller than the surrogate’s per-trial scatter of about a third, so it lies within one scatter of the near-zero bias. The camera–DVL rotation error of 0.49° is well within its per-trial scatter of close to two degrees. The camera–IMU rotation error of 0.15° lies furthest out, about 1.5 times its per-trial scatter of 0.10° , still inside the spread a single draw covers. The camera–DVL lever arm error of 2.97 cm is not coupled to the EIV mechanism and is recovered alike by all three formulations. The IMU–DVL lever arm is the only parameter the surrogate shows as biased under Formulation A, and Formulation B reduces that bias to the near-zero level, so the value recovered on real Structure Easy reflects per-trial scatter about a near-zero bias rather than a residual systematic offset.

The calibration results are conditioned on the selected window. Applied instead to a 60-second window that overlaps the contaminated start of the log, [40, 100] s, Formulation B’s error on the IMU–DVL lever arm rises from 4.24 to 12.81 cm. The structural improvements of the corrected formulations do not survive a window on which the DVL noise model fails.

7.4 Convergence and Run Time

With the window settled, two practical questions remain: whether the calibration improves predictably as the window grows, and how long it takes to run.

7.4.1 Convergence Against Window Length

Formulation B is applied to windows anchored at the start of the selected window, $t = 65$ s, and grown forward in increments of five keyframes, from ten keyframes until the window reaches its full extent at 125 s. Every window in the sequence lies inside the healthy span, so the experiment isolates the effect of accumulating data that satisfies the noise model. Figure 7.2 reports the error magnitudes against the keyframe count, rotations in degrees, and lever arms in centimeters.

Every window in the sequence converges from a random initialization, including the smallest at 10 keyframes, so the curves show how the error of the converged estimate decreases as the window grows. The parameters reach their final error at different points along the sweep. The camera–DVL lever arm settles within roughly twenty keyframes and varies between 2.5 and 5 cm thereafter. The rotation extrinsics reach the sub-degree regime by roughly eighty keyframes. The IMU–camera and IMU–DVL lever arms settle last: their errors level off between roughly 7 and 10 cm at the midpoint of the sweep and fall below 5 cm only after roughly 125 keyframes. Each increment grows two things at once, the sample count and the motion content of the window. The curve, therefore, does not separate the averaging of noise from the arrival of new excitation, and the late improvement of the IMU lever arms may reflect either or both. The sequence is also nested, so adjacent points share their data, and the fluctuations along the curves reflect the entering keyframes rather than independent repeats; the trends are read at the few-centimeter scale of single-window scatter. The two IMU

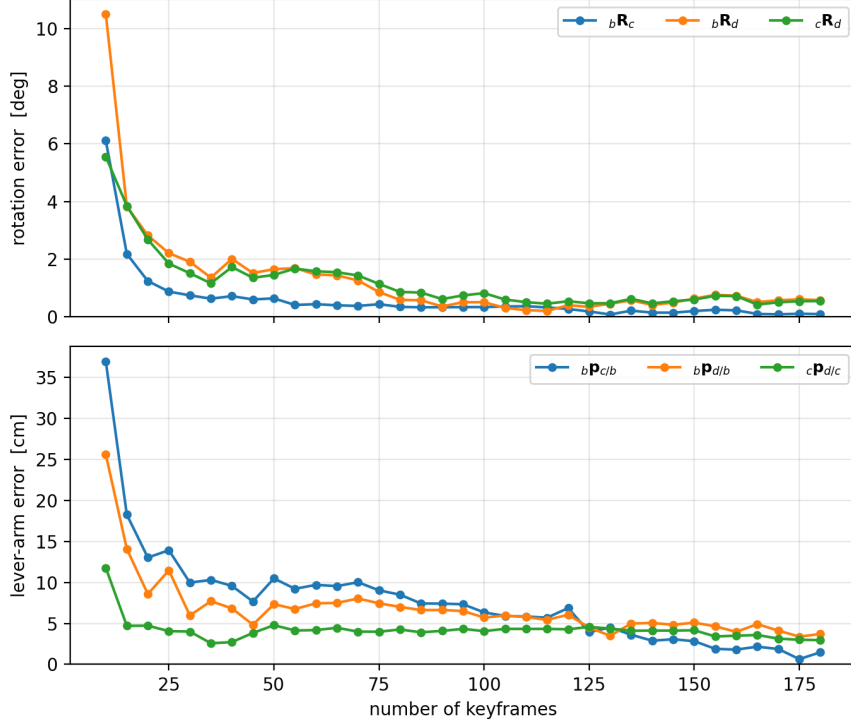


Figure 7.2: Formulation B’s calibration errors against keyframe count, for windows anchored at 65 s and grown in five-keyframe increments to the full selected window. Rotation errors above, lever arm errors below; the final point at 184 keyframes is the selected window.

lever arms reach their final error only with most of the window, so the 60 s target of Section 7.2 provides little margin beyond what the slowest parameters require.

7.4.2 Run Time

The calibration runs once per deployment, so its run time is a one-off overhead rather than a per-frame budget. To contextualize this cost, Figure 7.3 reports two complementary quantities on a shared scale. The first is the wall-clock run time of the coarse-to-fine pipeline of Section 4.3, broken down by stage as a function of the keyframe count of the window. The second is the distribution of single AQUA-SLAM local bundle adjustment times on the same machine. The bundle adjustment serves as a natural reference because it is continuously executed online during operation. The two quantities are not directly comparable; however, the bundle adjustment runs repeatedly during operation, typically at keyframe insertion, whereas the calibration runs once over the window. The comparison, therefore, indicates only relative order of magnitude.

The full calibration on the selected window completes in 0.47 s, roughly twice the median time for a single local bundle adjustment (0.25 s), with the middle half of the solves between 0.20 and 0.30 s. The run time grows roughly linearly with the keyframe count, and a 100-keyframe window calibrates in about the time required for a typical local bundle adjustment. Stage 2 of the pipeline incorporates the depth residual and is

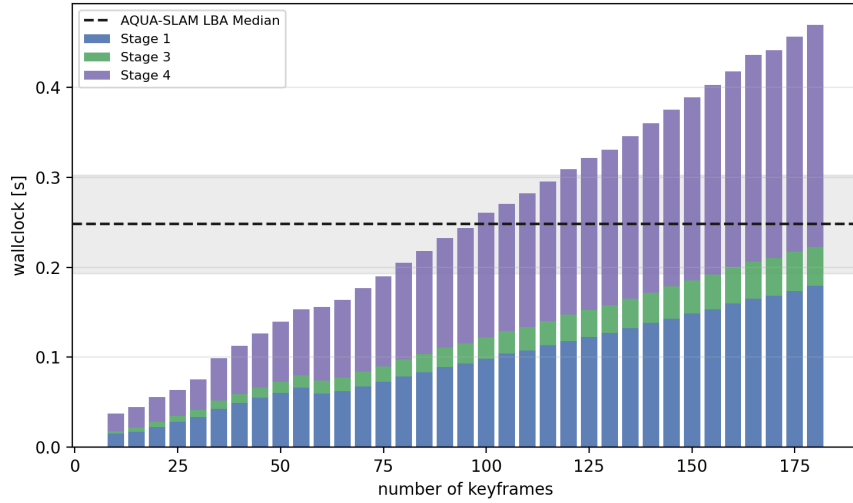


Figure 7.3: Wall-clock time of the calibration pipeline per stage against the keyframe count of the window; Stage 2 is inactive here (Section 7.1). The dashed line is the median time of a single AQUA-SLAM local bundle adjustment; half of the solves complete faster and half slower. The shaded band spans the middle half of the solve times, from the 25th to the 75th percentile. All timings are measured on an AMD Ryzen AI 7 350; both solvers run single-threaded, Ceres [1] for the calibration and g2o [2] for AQUA-SLAM’s bundle adjustment.

inactive here, since the depth residual is excluded (Section 7.1), so the reported time comprises the three active stages.

7.5 Downstream SLAM Validation

The calibration serves the SLAM system, so the final validation runs AQUA-SLAM end-to-end. The validation proceeds in two steps. The first corrects the runtime velocity-transport residual, and the second evaluates the recovered calibrations under that corrected residual. The order matters, because the omitted cross-product term of Chapter 3 is the only lever arm-dependent term in the velocity channel. The original runtime residual is therefore partially blind to the IMU–DVL lever arm, and a comparison of calibrations under it would understate the effect of calibration quality. Correcting the residual first removes that blindness, so the calibration comparison runs on a residual that can register the differences between calibrations.

7.5.1 The Corrected Runtime Residual

AQUA-SLAM is run end-to-end on each evaluation sequence in two configurations, one with its original baseline velocity-transport residual and one with the corrected residual of Formulation A in Chapter 4. Both configurations use the reference extrinsics, so the residual form is the only difference between them. The evaluation sequences are the three Structure sequences of the dataset: Easy, Medium, and Hard. Their ground truth is derived from the AprilTag [37] structure and covers the portions of each trajectory on

Table 7.2: Trajectory error of AQUA-SLAM with the baseline and the corrected velocity-transport residual, both with the reference extrinsics. Each cell is the median of the per-run RMSEs over ten runs, with the interquartile range (25th to 75th percentile) in parentheses. Loop closure disabled; errors after rigid-body alignment without scale correction, computed with the evo toolkit.

Sequence	Residual	ATE position [cm] median (25–75%)	ATE rotation [deg] median (25–75%)
Structure Easy	baseline	3.02 (2.58–3.50)	0.86 (0.81–0.98)
	corrected	2.52 (2.38–2.60)	0.88 (0.85–0.92)
Structure Medium	baseline	34.31 (26.92–37.03)	8.33 (6.62–9.07)
	corrected	25.12 (19.75–36.04)	6.44 (4.84–9.08)
Structure Hard	baseline	27.82 (21.56–34.57)	6.12 (5.61–7.12)
	corrected	23.26 (20.09–28.00)	4.92 (4.63–5.18)

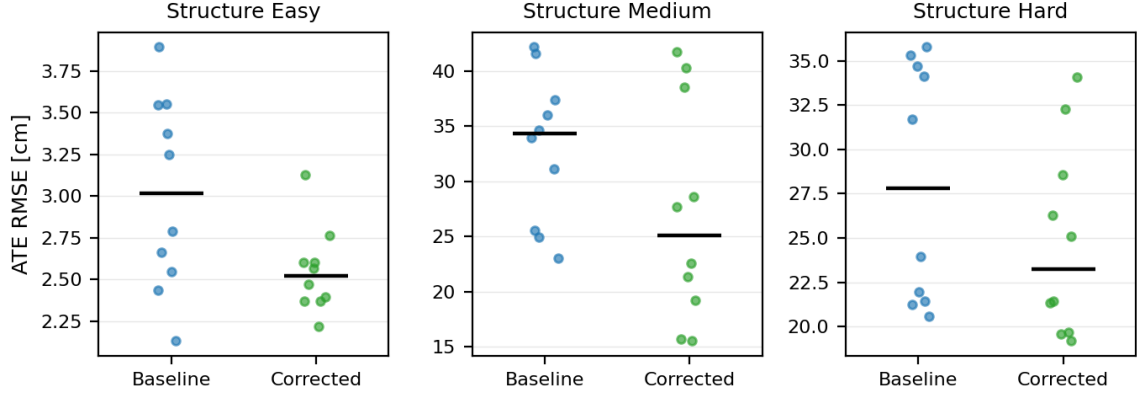


Figure 7.4: Per-run position ATE for the original baseline and the corrected residual, ten runs per configuration; the bar marks the median. The panels share no vertical scale: the Easy sequence is an order of magnitude more accurate than the other two.

which the structure is visible [29], and the error is evaluated over the covered portion. Loop closure is disabled throughout. Trajectory accuracy is measured with the evo toolkit [38] as the absolute trajectory error (ATE) in position and rotation, after a rigid-body alignment to the ground truth, which fits a rotation and translation but no scale. Each run is summarized by its root-mean-squared error (RMSE) over the trajectory. Each configuration is run ten times. The reported error is the median of the per-run RMSEs, the spread is their interquartile range.

Table 7.2 and Figure 7.4 report the comparison. The corrected residual improves the median position error on all three sequences by between 16 and 27 percent, and on no sequence does any metric degrade beyond the run-to-run scatter. The separation from the scatter differs by sequence. On Structure Easy, the position improvement is clear: the corrected configuration’s interquartile range lies almost entirely below that of the baseline, while the rotation error is unchanged. On Structure Hard, the clearest

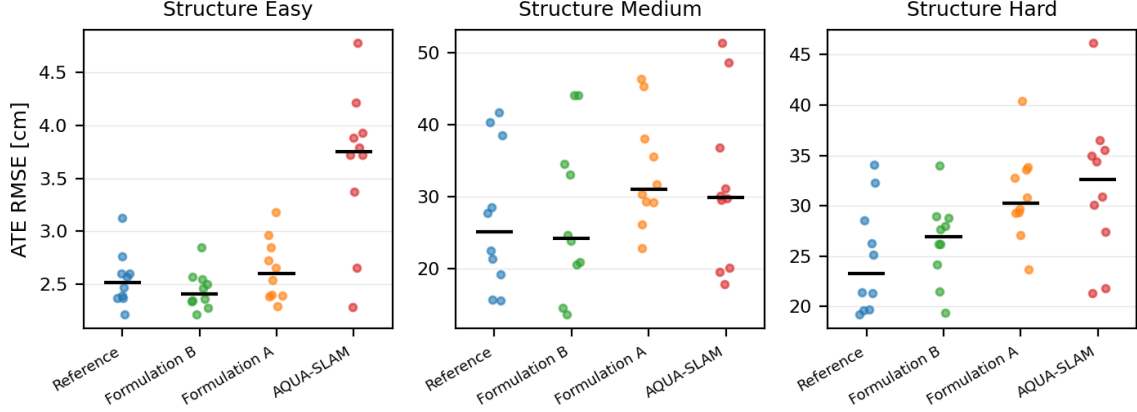


Figure 7.5: Per-run position ATE for the four calibrations, ten runs per configuration; the bar marks the median. The panels share no vertical scale.

effect appears in the rotation error, where the two interquartile ranges do not overlap, 4.63–5.18 against 5.61–7.12 degrees. On Structure Medium, the median improvement is the largest, roughly a quarter in both metrics. However, the run-to-run spread is also the largest, and the interquartile ranges overlap, so the improvement in this sequence does not stand out from the scatter. The omitted velocity-transport term therefore carries a measurable, if modest, cost in the live estimator, and degrades no metric on any sequence: the analytical finding of Chapter 3 holds at the system level.

The size of the effect is tied to the platform. The omitted term is the cross product of the angular velocity and the lever arm, and both are small here: the vehicle moves slowly with sustained low-amplitude rotation, and the IMU–DVL separation is 18 cm. On platforms with larger sensor separations or faster rotational dynamics, the omitted term scales with both, so a larger system-level benefit is expected there, although the present data do not measure it.

7.5.2 Calibration Under the Corrected Model

With the residuals fixed to the corrected form, the extrinsics become the variable of interest: AQUA-SLAM is run with four calibrations on the three Structure sequences, following the protocol of the previous subsection. The four are the reference calibration provided with the dataset and the three estimates in Table 7.1, recovered over the selected window using AQUA-SLAM’s baseline formulation, Formulation A, and Formulation B. The design completes the three-way comparison at the system level: it tests whether the recovered calibrations transfer across sequences, and what the calibration errors of Table 7.1 cost in trajectory accuracy. Structure Easy is the calibration sequence and is reported as such; the transfer claim rests on Medium and Hard. The reference is itself a calibration, so agreement with it is the strongest claim the data supports.

Table 7.3 and Figure 7.5 report the four calibrations. On Structure Easy, the recovered calibrations of Formulation B and Formulation A match the reference, all three

Table 7.3: Trajectory error of AQUA-SLAM under the corrected residual with four calibrations: the dataset reference and the three estimates of Table 7.1. Each cell is the median of the per-run RMSEs over ten runs, with the interquartile range in parentheses. Loop closure disabled; errors after rigid-body alignment without scale correction, computed with the evo toolkit.

Sequence	Calibration	ATE position [cm] median (25–75%)	ATE rotation [deg] median (25–75%)
Structure Easy	reference	2.52 (2.38–2.60)	0.88 (0.85–0.92)
	Formulation B	2.41 (2.34–2.54)	0.83 (0.80–0.87)
	Formulation A	2.60 (2.40–2.82)	0.86 (0.84–0.97)
	AQUA-SLAM	3.75 (3.46–3.91)	1.18 (1.12–1.42)
Structure Medium	reference	25.12 (19.75–36.04)	6.44 (4.84–9.08)
	Formulation B	24.23 (20.60–34.20)	6.19 (5.19–8.52)
	Formulation A	31.06 (29.21–37.48)	7.71 (7.38–9.49)
	AQUA-SLAM	29.93 (22.45–35.45)	7.54 (5.74–8.78)
Structure Hard	reference	23.26 (20.09–28.00)	4.92 (4.63–5.18)
	Formulation B	26.96 (24.69–28.60)	5.86 (5.60–6.19)
	Formulation A	30.26 (29.28–33.39)	6.49 (6.04–6.94)
	AQUA-SLAM	32.64 (28.10–35.39)	6.46 (6.12–7.03)

within 2.4 to 2.6 cm, while AQUA-SLAM’s recovered calibration is measurably worse at 3.75 cm and carries a heavier tail in both metrics. On Structure Medium, the reference and Formulation B are indistinguishable, 25.1 and 24.2 cm with closely matching interquartile ranges, and the other two sit slightly higher within the run-to-run spread. On Structure Hard, the three recovered calibrations are ordered as Formulation B, Formulation A, AQUA-SLAM, with Formulation B the closest to the reference; the reference retains a small advantage, its rotation interquartile range lying below all three.

The per-run spread on Structure Medium and Structure Hard is large and is shared across all four calibrations, so on these two sequences, the comparison is distributional rather than pointwise: configuration medians separated by a few centimeters fall within the run-to-run variation and are not ordered. The spread originates in the SLAM estimator rather than in the calibration, since it is present at the same magnitude under the reference as well.

Two conclusions remain robust despite this spread. The recovered calibration of Formulation B matches the reference on the two sequences where the comparison is decisive, Easy and Medium, and is the closest of the three recovered calibrations to it on Hard. So, a targetlessly recovered calibration transfers across sequences without loss when compared to a reference that is itself a calibration. AQUA-SLAM’s recovered calibration, the only one carrying the structural collapse of Chapter 3, consistently degrades the trajectory on Easy, the sequence where the run-to-run spread is smallest and the comparison most sensitive; the effect is small in absolute terms, about 1.2 cm of median position error above the other three, and on the harder sequences its deficit is absorbed into the shared spread.

7.6 Discussion

The real data reproduce the pattern predicted by the simulation. AQUA-SLAM’s formulation leaves the IMU–DVL lever arm essentially unrecovered. Formulation A recovers it, and Formulation B is consistent with the further reduction that the analysis predicts. The cm-level lever arms and degree-level rotations are the normal yield of this excitation rather than a degraded outcome. The structural finding of Chapter 3 and the bias analysis of Chapters 5 and 6 therefore hold on a real platform and are not artifacts of the simulation.

The calibration can be run in the field, subject to one data requirement. Its cost is about two local bundle adjustments, paid once per deployment (Section 7.4.2), and the window it uses is selected from the DVL’s own measurements without ground truth (Section 7.2). The requirement is that the window both supports the assumed noise model and provides sufficient excitation for the slowest parameters to settle. The first part is not optional, since on a window that overlaps the contaminated start of the log, the corrected formulations lose their advantage entirely. The second part depends on the motion rather than on a fixed budget. The IMU lever arms settle last, and on this survey-type sequence, they need most of the sixty-second window, whereas richer excitation would settle them sooner. The calibration is therefore within reach during a mission, given a clean window that the vehicle’s motion makes long enough.

Calibration quality reaches the trajectory where the SLAM estimator is sensitive enough to register it. Correcting the velocity-transport residual improves trajectory accuracy on all three sequences and degrades none (Section 7.5.1), and under that residual, the calibration recovered by Formulation B transfers across sequences to match the dataset reference (Section 7.5.2). The structural collapse, as carried by AQUA-SLAM’s recovered calibration, is not free, degrading the trajectory on Structure Easy, where the run-to-run spread is smallest, and the comparison is most sensitive. On the harder sequences, the spread is shared across all calibrations and originates in the SLAM estimator, so it, rather than the calibration, bounds what the comparison can resolve.

The improvement the corrections produce is genuine but small on this platform, for the reason given in Section 7.5.1: the slow motion and short sensor separation keep the omitted velocity-transport term small throughout. The gains demonstrated here are therefore a floor rather than a representative measure. The results are also drawn from a single platform, a single tank, and a single route type, so they characterize the formulations in this regime rather than across the range of platforms to which the corrections apply.

This discussion has shown that the corrected formulations carry the simulation results onto a real platform, run in the field at a cost the platform can absorb during a mission, and improve the downstream trajectory where the SLAM estimator can resolve the difference. These results fall within the low-excitation regime of this survey-type deployment, where the effect of the corrections is smallest. Chapter 8 draws these real-data findings together with the analysis and simulation of the preceding chapters, and sets out the limitations that remain and the future work that the combined results motivate.

Conclusion

This thesis addressed the main research question: *How can an inspection AUV jointly calibrate its camera, IMU, DVL, and pressure sensors in situ, from onboard data alone and within a discrete-time formulation, while accounting for its bias sources?* It develops a targetless joint calibration that recovers all four extrinsics from the suite’s own measurements in a discrete-time formulation. To the best of our knowledge, it is the first method to estimate the pressure-sensor translation jointly with the rest of the suite. The following discussion reflects on these findings through the research subquestions.

In the established baseline formulation, what causes the bias in the IMU–DVL translation estimate, and how can it be corrected? The baseline’s velocity transport omits the angular-velocity cross-product term that a nonzero translation produces, leaving a noise-independent structural bias that drives the estimate toward zero. Formulation A restores that term, which makes the translation observable and recovers it to within the noise-driven scatter.

To what extent does the corrected formulation eliminate bias under realistic sensor noise, and can any remaining bias be diagnosed, without ground truth, before it is trusted, and then reduced? It does not fully eliminate the bias: a smaller errors-in-variables bias remains, because the gyroscope sample enters the correction term as a noisy regressor, and this bias grows with sensor noise. Its severity collapses to a single scalar that flags an exposed calibration before it is trusted, without ground truth, and Formulation B, which carries the body velocity as a per-keyframe state, reduces but does not remove it.

Can the calibration operate efficiently enough to run during a mission, rather than only offline, and does it recover an accurate calibration from real underwater data that improves downstream SLAM? It runs on a finite window at a cost on the order of two local bundle adjustments, so it is runnable on demand during a mission rather than only offline. On a window selected by the DVL’s own consistency, the real data reproduces the accuracy ordering seen in simulation, and the calibration recovered by Formulation B matches the dataset reference downstream, up to the resolution a shared per-run estimator variance allows.

Taken together, these findings show that an inspection AUV can calibrate its full extrinsic suite jointly and without a target. It can recover even the parameter that resists calibration the most and judge the result to be trustworthy without ground truth; all from its own measurements during a mission, at a cost no greater than the estimation it already runs. Recovering that parameter, the IMU–DVL lever arm, came down to a corrected model of how the sensors’ motion couples, not to more data. Once the velocity transport has carried its full rigid-body term, the lever arm becomes observable, and the surviving bias is noise-driven rather than structural and silent.

8.1 Limitations and Future Work

Formulation B reduces but does not remove the EIV bias on the IMU–DVL translation, leaving a centimeter-scale residual of non-EIV origin whose source the present results do not separate. The diagnostic that flags the EIV bias models only the white gyroscope noise, so it underpredicts the realized bias and serves as a screen rather than a calibrated error estimate. It treats the translation in isolation rather than as coupled to the other extrinsics. Modeling the further sources of EIV-like bias that the diagnostic omits, the gyroscope bias error, and the camera trajectory treated as exact, would extend the analysis. Separately, a joint IMU–DVL preintegration [39] factor would capture the gyroscope correlation that the present cost neglects and relax the zero-order hold in the DVL translation preintegration.

The real-data evaluation is based on a single platform, tank, and route type, with slow motion and a short sensor separation, which keeps the omitted velocity-transport term small. The formulations, therefore, need evaluation in higher-excitation regimes, on larger or faster vehicles with greater sensor separation, where the omitted term is larger, and the benefit is expected to grow. The pressure-sensor translation, the one extrinsic no prior method recovers, is also untested on real data, because the dataset provides no reference for it and contains almost no vertical motion, so a deployment with vertical motion and a depth reference would close that gap.

The window selection is not autonomous, as the window is selected manually within the healthy span identified by the gates. It also screens only the DVL noise model, so a fault confined to the camera trajectory or the gyroscope passes undetected. Packaging the selection as an automated, excitation-aware pipeline stage would remove the manual step, and complementary screens on the gyroscope and camera rotations would catch trajectory-side faults that the DVL screen cannot detect.

Bibliography

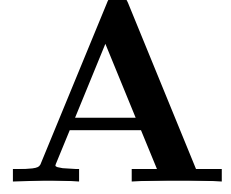
- [1] Agarwal, Sameer, Mierle, Keir, and The Ceres Solver Team. Ceres Solver, 2023.
- [2] Rainer Kümmerle, Giorgio Grisetti, Hauke Strasdat, Kurt Konolige, and Wolfram Burgard. G2o: A general framework for graph optimization. In *2011 IEEE International Conference on Robotics and Automation*, pages 3607–3613, May 2011. ISSN: 1050-4729.
- [3] Liam Paull, Sajad Saeedi, Mae Seto, and Howard Li. AUV Navigation and Localization: A Review. *IEEE Journal of Oceanic Engineering*, 39(1):131–149, January 2014.
- [4] Xiaotian Wang, Xinnan Fan, Pengfei Shi, Jianjun Ni, and Zhongkai Zhou. An overview of key SLAM technologies for underwater scenes. *Remote Sensing*, 15(10):2496, 2023.
- [5] Shuoshuo Ding, Tiedong Zhang, Dapeng Jiang, and Ming Lei. Underwater Visual-Inertial-Acoustic-Depth SLAM with DVL Preintegration for Degraded Environments. *arXiv preprint arXiv:2510.21215*, 2025.
- [6] Mohamed Heshmat, Lyes Saad Saoud, Muayad Abujabal, Atif Sultan, Mahmoud Elmezain, Lakmal Seneviratne, and Irfan Hussain. Underwater SLAM meets deep learning: challenges, multi-sensor integration, and future directions. *Sensors*, 25(11):3258, 2025.
- [7] Shida Xu, Kaicheng Zhang, and Sen Wang. Aqua-slam: Tightly-coupled underwater acoustic-visual-inertial slam with sensor calibration. *IEEE Transactions on Robotics*, 2025.
- [8] Franz S Hover, Ryan M Eustice, Ayoung Kim, Brendan Englot, Hordur Johannsson, Michael Kaess, and John J Leonard. Advanced perception, navigation and planning for autonomous in-water ship hull inspection. *The International Journal of Robotics Research*, 31(12):1445–1464, October 2012.
- [9] Lin Zhao, Mingxi Zhou, and Brice Loose. Tightly-coupled visual-DVL-inertial odometry for robot-based ice-water boundary exploration. In *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 7127–7134. IEEE, 2023.
- [10] Yuanbo Xue, Yang Hu, Dejin Zhang, Chih-Yung Wen, and Bing Wang. AVIP: Acoustic-Visual-Inertial-Pressure Fusion-based Underwater Localization System with Multi-Centric Calibration. In *2025 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1518–1525, October 2025. ISSN: 2153-0866.
- [11] Yulin Yang, Patrick Geneva, and Guoquan Huang. Multi-Visual-Inertial System: Analysis, Calibration and Estimation, September 2024. arXiv:2308.05303 [cs].

- [12] Paul Furgale, Joern Rehder, and Roland Siegwart. Unified temporal and spatial calibration for multi-sensor systems. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1280–1286, November 2013. ISSN: 2153-0866.
- [13] Joern Rehder, Roland Siegwart, and Paul Furgale. A General Approach to Spatiotemporal Calibration in Multisensor Systems. *IEEE Transactions on Robotics*, 32(2):383–398, April 2016.
- [14] Shida Xu, Tomasz Luczynski, Jonatan Scharff Willners, Ziyang Hong, Kaicheng Zhang, Yvan R. Petillot, and Sen Wang. Underwater Visual Acoustic SLAM with Extrinsic Calibration. In *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 7647–7652, September 2021. ISSN: 2153-0866.
- [15] Bo Xu, Lianzhao Wang, Shengxin Li, and Jiao Zhang. A Novel Calibration Method of SINS/DVL Integration Navigation System Based on Quaternion. *IEEE Sensors Journal*, 20(16):9567–9580, August 2020.
- [16] Shede Liu, Tao Zhang, and Yongyun Zhu. A GNSS Aided Calibration Method for DVL Error Based on the Optimal-REQUEST. *IEEE Sensors Journal*, 22(22):21899–21910, November 2022.
- [17] Bo Xu and Yu Guo. A Novel DVL Calibration Method Based on Robust Invariant Extended Kalman Filter. *IEEE Transactions on Vehicular Technology*, 71(9):9422–9434, September 2022.
- [18] J.C. Kinsey and L.L. Whitcomb. Towards in-situ calibration of gyro and Doppler navigation sensors for precision underwater vehicle navigation. In *Proceedings 2002 IEEE International Conference on Robotics and Automation (Cat. No.02CH37292)*, volume 4, pages 4016–4023 vol.4, May 2002.
- [19] Giovanni Cioffi, Titus Cieslewski, and Davide Scaramuzza. Continuous-Time Vs. Discrete-Time Vision-Based SLAM: A Comparative Study. *IEEE Robotics and Automation Letters*, 7(2):2399–2406, April 2022.
- [20] Shuolong Chen, Xingxing Li, Shengyu Li, Yuxuan Zhou, and Xiaoteng Yang. iKalibr: Unified Targetless Spatiotemporal Calibration for Resilient Integrated Inertial Systems. *IEEE Transactions on Robotics*, 41:1618–1638, 2025.
- [21] Steven Lovegrove, Alonso Patron-Perez, and Gabe Sibley. Spline Fusion: A continuous-time representation for visual-inertial fusion with application to rolling shutter cameras. In *BMVC*, volume 2, page 8, 2013.
- [22] Yulin Yang, Patrick Geneva, Kevin Eickenhoff, and Guoquan Huang. Degenerate Motion Analysis for Aided INS With Online Spatial and Temporal Sensor Calibration. *IEEE Robotics and Automation Letters*, 4(2):2070–2077, April 2019.
- [23] Karol Hausman, James Preiss, Gaurav S. Sukhatme, and Stephan Weiss. Observability-Aware Trajectory Optimization for Self-Calibration With Application to UAVs. *IEEE Robotics and Automation Letters*, 2(3):1770–1777, July 2017.

- [24] Jiang Wang, Yaozhong Kang, Linya Fu, Kazuhiro Nakadai, and He Kong. Observability-Aware Active Calibration of Multisensor Extrinsic for Ground Robots via Online Trajectory Optimization. *IEEE Sensors Journal*, 25(17):33022–33036, September 2025.
- [25] Thomas Schneider, Mingyang Li, Michael Burri, Juan Nieto, Roland Siegwart, and Igor Gilitschenski. Visual-inertial self-calibration on informative motion segments. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pages 6487–6494, May 2017.
- [26] Olav Reiersøl. Identifiability of a Linear Relation between Variables Which Are Subject to Error. *Econometrica*, 18(4):375–389, 1950.
- [27] Wayne A. Fuller. *Measurement Error Models*. John Wiley & Sons, September 2009. Google-Books-ID: Nalc0DkAJRYC.
- [28] Hongxu Zhao, Guangyang Zeng, Yunling Shao, Tengfei Zhang, and Junfeng Wu. Spatiotemporal Calibration of Doppler Velocity Logs for Underwater Robots. *arXiv preprint arXiv:2510.24571*, 2025.
- [29] Shida Xu, Jonatan Scharff Willners, Joshua Roe, Sean Katagiri, Tomasz Luczynski, Yvan Petillot, and Sen Wang. Tank dataset: An underwater multi-sensor dataset for SLAM evaluation. *The International Journal of Robotics Research*, 45(4):541–551, April 2026.
- [30] Bill Triggs, Philip F. McLauchlan, Richard I. Hartley, and Andrew W. Fitzgibbon. Bundle Adjustment — A Modern Synthesis. In Bill Triggs, Andrew Zisserman, and Richard Szeliski, editors, *Vision Algorithms: Theory and Practice*, pages 298–372, Berlin, Heidelberg, 2000. Springer.
- [31] Christian Forster, Luca Carlone, Frank Dellaert, and Davide Scaramuzza. On-manifold preintegration for real-time visual-inertial odometry. *IEEE Transactions on Robotics*, 33(1):1–21, 2016.
- [32] Christoph Hertzberg, René Wagner, Udo Frese, and Lutz Schröder. Integrating generic sensor fusion algorithms with sound state representations through encapsulation of manifolds. *Information Fusion*, 14(1):57–77, January 2013.
- [33] Easton Potokar, Spencer Ashford, Michael Kaess, and Joshua G. Mangelson. Holocean: An underwater robotics simulator. In *2022 International Conference on Robotics and Automation (ICRA)*, pages 3040–3046. IEEE, 2022.
- [34] Shida Xu and Sense Robotics Lab. AQUA-SLAM Repository, 2026.
- [35] Peter J. Rousseeuw and Christophe Croux. Alternatives to the Median Absolute Deviation. *Journal of the American Statistical Association*, 88(424):1273–1283, 1993.
- [36] A. W. van der Vaart. *Asymptotic Statistics*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 1998.

- [37] Edwin Olson. AprilTag: A robust and flexible visual fiducial system. In *2011 IEEE international conference on robotics and automation*, pages 3400–3407. IEEE, 2011.
- [38] Michael Grupp. evo: Python package for the evaluation of odometry and slam, 2017.
- [39] Pau Vial, Joan Solà, Narcís Palomeras, and Marc Carreras. On Lie Group IMU and Linear Velocity Preintegration for Autonomous Navigation Considering the Earth Rotation Compensation. *IEEE Transactions on Robotics*, 41:1346–1364, 2025.
- [40] C. N. Flagg, G. Schwartz, E. Gottlieb, and T. Rossby. Operating an Acoustic Doppler Current Profiler aboard a Container Vessel. *Journal of Atmospheric and Oceanic Technology*, 15(1):257–271, February 1998.
- [41] Joan Solà, Jeremie Deray, and Dinesh Atchuthan. A micro Lie theory for state estimation in robotics, December 2021. arXiv:1812.01537 [cs.RO].

DVL noise standard deviation from the error velocity



This appendix derives the DVL error velocity for the implemented beam geometry and describes the estimation of the per-beam noise standard deviation σ_b from the raw ping stream. The estimate supplies the one G1 entry of Table 3.2 not given by the manufacturer’s datasheet, and the reference standard deviation for the calibration-window selection of Section 7.2. The procedure is a coarse, motion-independent estimate rather than a precise sensor calibration.

A.1 Beam geometry and the error velocity

The WaterLinked A50 is a four-beam Janus DVL in the X-pattern. Each beam direction is parameterized by an azimuth θ and an elevation ϕ measured from the sensor’s horizontal plane, with the implemented values $\theta = 45^\circ$ and $\phi = 67.5^\circ$, that is, beams tilted 22.5° from the vertical. The beam matrix $\mathbf{E} \in \mathbb{R}^{4 \times 3}$ stacks the unit beam directions as rows,

$$\mathbf{E} = \begin{bmatrix} -\cos \theta \cos \phi & \sin \theta \cos \phi & \sin \phi \\ -\cos \theta \cos \phi & -\sin \theta \cos \phi & \sin \phi \\ \cos \theta \cos \phi & -\sin \theta \cos \phi & \sin \phi \\ \cos \theta \cos \phi & \sin \theta \cos \phi & \sin \phi \end{bmatrix}, \quad (\text{A.1})$$

and the per-ping measurement model is $\tilde{\mathbf{b}} = \mathbf{E}\mathbf{v} + \boldsymbol{\eta}$ with body velocity $\mathbf{v} \in \mathbb{R}^3$ and per-beam noise $\boldsymbol{\eta}$, $\mathbb{E}[\boldsymbol{\eta}\boldsymbol{\eta}^T] = \sigma_b^2 \mathbf{I}_4$.

The beam matrix has rank three, so its left null space is one-dimensional, and for the sign pattern above the unit generator is $\mathbf{n} = \frac{1}{2}(1, -1, 1, -1)^T$ exactly, for any azimuth and elevation: the alternating signs cancel the x and y columns pairwise and the equal z entries in alternating sum. Projecting a ping onto $\mathbf{n}$ therefore annihilates the motion contribution,

$$\hat{\zeta} = \mathbf{n}^T \tilde{\mathbf{b}} = \mathbf{n}^T \boldsymbol{\eta}, \quad \mathbb{E}[\hat{\zeta}] = 0, \quad \mathbb{E}[\hat{\zeta}^2] = \sigma_b^2, \quad (\text{A.2})$$

where the variance follows from $\|\mathbf{n}\| = 1$. The combination is known as the DVL error velocity and is widely used as a quality indicator [40]. Because the projection removes the velocity entirely, every statistic of $\hat{\zeta}$ is independent of the vehicle motion.

A.2 Per-beam standard deviation

The standard deviation is estimated robustly over the full log: $\hat{\sigma}_b$ is the median absolute deviation of the error-velocity samples about their median, divided by $0.6745 = \Phi^{-1}(0.75)$, which calibrates the median absolute deviation of a Gaussian to its standard deviation. The robust estimator is insensitive to acoustic outliers, which would inflate

a sample standard deviation. On the Structure Easy sequence of the Tank dataset [29] this gives $\hat{\sigma}_b = 2.07 \times 10^{-3}$ m/s.

A.3 Body-frame covariance

The body-frame velocity covariance follows by standard propagation from the least-squares beam-to-velocity solution, $\Sigma_{\text{DVL}} = \sigma_b^2 (\mathbf{E}^T \mathbf{E})^{-1}$. For the symmetric geometry above with $\theta = 45^\circ$ this is diagonal, with horizontal variances $\sigma_b^2 / (2 \cos^2 \phi)$ and vertical variance $\sigma_b^2 / (4 \sin^2 \phi)$; at $\phi = 67.5^\circ$ the horizontal standard deviation is $1.85 \sigma_b$ and the vertical $0.54 \sigma_b$. The G1 table entry takes the largest diagonal entry as an isotropic scalar, treating the DVL as isotropic at the worst axis: $1.85 \times 2.07 \times 10^{-3} = 3.82 \times 10^{-3}$ m/s, the value in Table 3.2.

An Expectation Identity for the Cross-Product Matrix

B

For a vector $\mathbf{y} \in \mathbb{R}^3$, the cross-product matrix $[\mathbf{y}]_{\times} \in \mathbb{R}^{3 \times 3}$ is the skew-symmetric matrix satisfying $[\mathbf{y}]_{\times} \mathbf{z} = \mathbf{y} \times \mathbf{z}$ for every $\mathbf{z} \in \mathbb{R}^3$, so that $[\mathbf{y}]_{\times}^T = -[\mathbf{y}]_{\times}$.

Lemma B.1 (cross-product expectation identity). *Let $\boldsymbol{\eta} \in \mathbb{R}^3$ be a zero-mean random vector with isotropic covariance $\mathbb{E}[\boldsymbol{\eta} \boldsymbol{\eta}^T] = \sigma^2 \mathbf{I}$, and let $\mathbf{N} \in \mathbb{R}^{3 \times 3}$ be deterministic. Then*

$$\mathbb{E}[[\boldsymbol{\eta}]_{\times} \mathbf{N} [\boldsymbol{\eta}]_{\times}] = \sigma^2 (\mathbf{N}^T - \text{tr}(\mathbf{N}) \mathbf{I}). \quad (\text{B.1})$$

Proof. The map $\mathbf{N} \mapsto [\mathbf{y}]_{\times} \mathbf{N} [\mathbf{y}]_{\times}$ is linear in $\mathbf{N}$, so it is enough to evaluate it on a rank-one matrix and sum over a decomposition of $\mathbf{N}$. For any $\mathbf{u}, \mathbf{w} \in \mathbb{R}^3$ the rank-one product rearranges as

$$[\mathbf{y}]_{\times} \mathbf{u} \mathbf{w}^T [\mathbf{y}]_{\times} = [\mathbf{u}]_{\times} \mathbf{y} \mathbf{y}^T [\mathbf{w}]_{\times}, \quad (\text{B.2})$$

which moves the fixed vector out of the two cross-product matrices and into the central dyad. The rearrangement uses the antisymmetry $[\mathbf{y}]_{\times} \mathbf{u} = \mathbf{y} \times \mathbf{u} = -\mathbf{u} \times \mathbf{y} = -[\mathbf{u}]_{\times} \mathbf{y}$ on the left factor and its transpose $\mathbf{w}^T [\mathbf{y}]_{\times} = -\mathbf{y}^T [\mathbf{w}]_{\times}$ on the right factor.

Write $\mathbf{N} = \sum_r \mathbf{u}_r \mathbf{w}_r^T$ as a sum of rank-one terms. Applying (B.2) to each term gives

$$[\mathbf{y}]_{\times} \mathbf{N} [\mathbf{y}]_{\times} = \sum_r [\mathbf{u}_r]_{\times} \mathbf{y} \mathbf{y}^T [\mathbf{w}_r]_{\times}, \quad (\text{B.3})$$

valid for every fixed $\mathbf{y}$.

Setting $\mathbf{y} = \boldsymbol{\eta}$ and taking the expectation introduces the probabilistic assumption at a single point, the isotropic covariance $\mathbb{E}[\boldsymbol{\eta} \boldsymbol{\eta}^T] = \sigma^2 \mathbf{I}$ in the central dyad,

$$\mathbb{E}[[\boldsymbol{\eta}]_{\times} \mathbf{N} [\boldsymbol{\eta}]_{\times}] = \sigma^2 \sum_r [\mathbf{u}_r]_{\times} [\mathbf{w}_r]_{\times}. \quad (\text{B.4})$$

The standard product identity $[\mathbf{u}]_{\times} [\mathbf{w}]_{\times} = \mathbf{w} \mathbf{u}^T - (\mathbf{u}^T \mathbf{w}) \mathbf{I}$ turns each summand into $\mathbf{w}_r \mathbf{u}_r^T - (\mathbf{u}_r^T \mathbf{w}_r) \mathbf{I}$ [41]. Summing over r , the first part gives $\sum_r \mathbf{w}_r \mathbf{u}_r^T = \mathbf{N}^T$ by transposition of the decomposition, and the second gives $\sum_r \mathbf{u}_r^T \mathbf{w}_r = \text{tr}(\mathbf{N})$ since $\text{tr}(\mathbf{u}_r \mathbf{w}_r^T) = \mathbf{u}_r^T \mathbf{w}_r$. This leaves $\sigma^2 (\mathbf{N}^T - \text{tr}(\mathbf{N}) \mathbf{I})$, which is (B.1). $\square$

By the skew-symmetry $[\boldsymbol{\eta}]_{\times}^T = -[\boldsymbol{\eta}]_{\times}$, the transposed product follows from (B.1) as $\mathbb{E}[[\boldsymbol{\eta}]_{\times}^T \mathbf{N} [\boldsymbol{\eta}]_{\times}] = \sigma^2 (\text{tr}(\mathbf{N}) \mathbf{I} - \mathbf{N}^T)$. When $\mathbf{N}$ is symmetric, this becomes

$$\mathbb{E}[[\boldsymbol{\eta}]_{\times}^T \mathbf{N} [\boldsymbol{\eta}]_{\times}] = \sigma^2 (\text{tr}(\mathbf{N}) \mathbf{I} - \mathbf{N}). \quad (\text{B.5})$$

Proof of Lemma 5.1: Asymptotic Bias of the WLS Estimator under EIV

C

The proof establishes the minimizer of the expected weighted least-squares cost and the asymptotic bias that follows. The argument follows the standard treatment of errors-in-variables estimators [27].

The weighted least-squares cost is

$$J(\mathbf{p}) := \sum_{ij} \mathbf{r}_{ij}(\mathbf{p})^T \mathbf{W}_{ij} \mathbf{r}_{ij}(\mathbf{p}), \quad (\text{C.1})$$

with $\mathbf{W}_{ij} \succeq 0$ being the weight, deterministic and independent of $\mathbf{p}$.

Take the expectation of the cost (C.1), holding $\mathbf{p}$ fixed. Write $\mathbf{A}_{ij} = \bar{\mathbf{A}}_{ij} + \delta \mathbf{A}_{ij}$ for the noisy regressor, so the residual becomes $\mathbf{r}_{ij}(\mathbf{p}) = \mathbf{A}_{ij}\mathbf{p} - \bar{\mathbf{A}}_{ij}\mathbf{p}^* + \boldsymbol{\eta}_{ij}$. Expanding $\mathbf{A}_{ij} = \bar{\mathbf{A}}_{ij} + \delta \mathbf{A}_{ij}$ and using $\mathbb{E}[\delta \mathbf{A}_{ij}] = \mathbf{0}$,

$$\begin{aligned} \mathbb{E}[\mathbf{A}_{ij}^T \mathbf{W}_{ij} \mathbf{A}_{ij}] &= \bar{\mathbf{A}}_{ij}^T \mathbf{W}_{ij} \bar{\mathbf{A}}_{ij} + \mathbb{E}[\delta \mathbf{A}_{ij}^T \mathbf{W}_{ij} \delta \mathbf{A}_{ij}] = \bar{\mathbf{H}}_{ij} + \boldsymbol{\Lambda}_{ij}, \\ \mathbb{E}[\mathbf{A}_{ij}^T \mathbf{W}_{ij} \bar{\mathbf{A}}_{ij}] &= \bar{\mathbf{A}}_{ij}^T \mathbf{W}_{ij} \bar{\mathbf{A}}_{ij} = \bar{\mathbf{H}}_{ij}, \end{aligned}$$

the cross terms in $\delta \mathbf{A}_{ij}$ dropping because it is zero-mean.

Expanding the per-residual quadratic and substituting these,

$$\mathbb{E}[\mathbf{r}_{ij}(\mathbf{p})^T \mathbf{W}_{ij} \mathbf{r}_{ij}(\mathbf{p})] = \mathbf{p}^T (\bar{\mathbf{H}}_{ij} + \boldsymbol{\Lambda}_{ij}) \mathbf{p} - 2 \mathbf{p}^T \bar{\mathbf{H}}_{ij} \mathbf{p}^* + c_{ij},$$

where c_{ij} collects the $\mathbf{p}$ -independent terms. The $\boldsymbol{\eta}_{ij}$ cross term vanishes because $\boldsymbol{\eta}_{ij}$ is zero-mean and independent of the regressor. Summing over the keyframe intervals and using $\mathbf{H} = \bar{\mathbf{H}} + \boldsymbol{\Lambda}$,

$$\mathbb{E}[J(\mathbf{p})] = \mathbf{p}^T \mathbf{H} \mathbf{p} - 2 \mathbf{p}^T \bar{\mathbf{H}} \mathbf{p}^* + C, \quad (\text{C.2})$$

with C independent of $\mathbf{p}$, a positive-definite quadratic with a unique minimizer.

The gradient of (C.2) vanishes where

$$\nabla_{\mathbf{p}} \mathbb{E}[J(\mathbf{p})] = 2 \mathbf{H} \mathbf{p} - 2 \bar{\mathbf{H}} \mathbf{p}^* = \mathbf{0},$$

that is, $\mathbf{H} \mathbf{p} = \bar{\mathbf{H}} \mathbf{p}^*$. With $\mathbf{H}$ invertible the minimizer is

$$\mathbf{p}^\infty = \mathbf{H}^{-1} \bar{\mathbf{H}} \mathbf{p}^*,$$

and substituting $\bar{\mathbf{H}} = \mathbf{H} - \boldsymbol{\Lambda}$,

$$\mathbf{p}^\infty - \mathbf{p}^* = \mathbf{H}^{-1} \bar{\mathbf{H}} \mathbf{p}^* - \mathbf{p}^* = \mathbf{H}^{-1} (\bar{\mathbf{H}} - \mathbf{H}) \mathbf{p}^* = -\mathbf{H}^{-1} \boldsymbol{\Lambda} \mathbf{p}^*, \quad (\text{C.3})$$

deriving the desired result in (5.7).

In the standard errors-in-variables treatment, the minimizer of the expected cost is the value to which the weighted least-squares estimate converges as the data grow. Here, the minimizer is $\mathbf{p}^\infty$. The bias (C.3) is its displacement from the truth, and it does not shrink with more data. Under stationary motion and noise statistics, each added interval (i, j) contributes $\bar{\mathbf{H}}_{ij}$ and $\mathbf{\Lambda}_{ij}$ of similar magnitude, so $\bar{\mathbf{H}}$, $\mathbf{\Lambda}$, and $\mathbf{H}$ grow in proportion to the number of intervals. The ratio $\mathbf{H}^{-1}\mathbf{\Lambda}$ then approaches a fixed limit, and the bias stays nonzero whenever $\mathbf{\Lambda}\mathbf{p}^\star \neq \mathbf{0}$.

DVL calibration-window selection

D

This appendix gives the formal definitions of the window statistics and the derivations of the gate thresholds used by the calibration-window selection of Section 7.2. The error velocity itself, its null-space construction, and the estimation of the per-beam noise standard deviation are developed in Appendix A; the selection procedure and its results are presented in the chapter and are not repeated here.

D.1 Window statistics

The log is scanned with windows of duration 20 s at a stride of 2.5 s, which yields 65 windows with a median of 167 pings per window on Structure Easy. For a window $\mathcal{W}$ containing P pings, the statistics below are computed from the error-velocity samples $\hat{\zeta}_k$, $k \in \mathcal{W}$, with $\hat{\sigma}_b$ the global per-beam standard deviation of Appendix A; it is the median absolute deviation calibrated to the Gaussian standard deviation, so it carries the units of the samples and the factor 3 below means three standard deviations. No sensor validity flags are consulted at any point.

The window median and the two gated statistics are

$$m_{\mathcal{W}} = \text{med}_{k \in \mathcal{W}} \hat{\zeta}_k, \quad f_{\mathcal{W}} = \frac{1}{P} \#\{k \in \mathcal{W} : |\hat{\zeta}_k - m_{\mathcal{W}}| > 3\hat{\sigma}_b\}, \quad b_{\mathcal{W}} = |m_{\mathcal{W}}|, \quad (\text{D.1})$$

where the outlier fraction $f_{\mathcal{W}}$ tests the Gaussian tails after centering on the window median, and the median bias $b_{\mathcal{W}}$ tests the centering itself. Centering on the median rather than the mean targets the same center, since the two coincide for the assumed zero-mean Gaussian, but remains stable in the presence of the outliers that the fraction counts, which would displace a window mean. The two reported, ungated statistics are the lag-one autocorrelation and the timing jitter,

$$\hat{\rho}_1(\mathcal{W}) = \frac{\sum_k (\hat{\zeta}_k - \bar{\zeta})(\hat{\zeta}_{k+1} - \bar{\zeta})}{\sum_k (\hat{\zeta}_k - \bar{\zeta})^2}, \quad j_{\mathcal{W}} = \frac{\text{std}_{k \in \mathcal{W}}(\Delta t_k)}{\text{mean}_{k \in \mathcal{W}}(\Delta t_k)}, \quad (\text{D.2})$$

with $\bar{\zeta}$ the window mean and Δt_k the inter-ping intervals.

D.2 Gate thresholds

The outlier-fraction gate is relative to the log. A window passes if its outlier fraction $f_{\mathcal{W}}$ does not exceed $c = 2$ times the median outlier fraction over all scanned windows. The threshold cannot be taken from the model, because the Gaussian prediction for the fraction beyond 3σ is 0.3%, whereas windows on which the model is otherwise usable carry 1 to 5%. An absolute bound at the Gaussian value would therefore reject the

entire log. The cross-window median is anchored by the clean majority of the log, for the same robustness reason as the global standard deviation. The gate, therefore, tests whether a window's tails are atypically heavy relative to the log. On Structure Easy the threshold evaluates to $f_W \leq 0.096$.

The median-bias gate is derived from the noise model. For P iid samples from a distribution with density f , the sample median is asymptotically Gaussian about the true median with variance

$$\frac{1}{4Pf(0)^2},$$

a standard large-sample result [36]. For a zero-mean Gaussian with standard deviation σ_b , the density at zero is $f(0) = 1/(\sigma_b\sqrt{2\pi})$. Substituting gives the variance $\pi\sigma_b^2/(2P)$, so the window median has standard deviation $\sqrt{\pi/2}\sigma_b/\sqrt{P}$ under the model. A window passes if

$$b_W \leq 3\sqrt{\pi/2}\hat{\sigma}_b/\sqrt{P}, \quad (\text{D.3})$$

which under the noise model excludes a clean window with probability 0.3%. At the median ping count of this log, the threshold evaluates to $b_W \leq 0.60$ mm/s.

Two Gaussian constants appear in the procedure and serve distinct roles. The first, $0.6745 = \Phi^{-1}(0.75)$, calibrates the median absolute deviation to the standard deviation [35], and is applied once in the estimation of $\hat{\sigma}_b$ (Appendix A). The second, $\sqrt{\pi/2} \approx 1.2533$, is the sampling penalty of the median as a location estimator, and is applied here. Both follow from the assumed Gaussian noise model.

D.3 Acceptance rule

A sample at time t is accepted if every scan window containing t passes both gates. Equivalently, the run of passing window centers is eroded by half a window duration at each end, which is the conservative reading: a window statistic is a property of everything inside the window, so a sample adjacent to a contaminated interval sits inside tainted windows and is excluded even if the sample itself is clean. The healthy span is the longest contiguous run of accepted samples, and the calibration window is placed within it as described in the chapter.