A Consistency-Gated Quaternion Error-State Kalman Filter for Multi-Sensor Inertial Navigation

Irfan Saric

December 2025

Abstract

We describe the design, derivation, implementation, and verification of an error-state Kalman filter (ESKF) that fuses a strapdown inertial measurement unit (IMU) with up to nine aiding modalities—GPS position and Doppler velocity, a barometric altimeter, a magnetometer, a downward laser altimeter (LiDAR), ultra-wideband (UWB) radio ranging, optical flow, a Doppler velocity log (DVL), and a direct attitude fix—to estimate the pose of an aerial vehicle. The filter follows the multiplicative, local error-state formulation of Solà [2] and Trawny and Roumeliotis [3]: a nominal state carries the full estimate and is integrated directly from the IMU, while a minimal 15-dimensional error state, reset to zero after every correction, carries the covariance. We derive the continuous-time error dynamics from first principles, discretize them, and give each of the nine measurement models with its exact error-state Jacobian; every Jacobian is validated against a finite difference. The central claim is not merely that the filter tracks well but that its *reported uncertainty is correct*: we verify this with the normalized estimation error squared (NEES) consistency test of Bar-Shalom et al. [8], run as a Monte-Carlo gate over many simulated flights and shown to hold throughout each flight, not only in the mean, and we reject outliers online with the companion normalized innovation squared (NIS) test. On the full nine-sensor fusion the filter is statistically consistent (position NEES 2.85 against an expected 3.0; full-state NEES 16.2 against 15.0) at 0.05 m position RMSE, while remaining faithful to inertial-navigation error and observability theory: with all aiding removed the position error grows without bound, and it is recovered to sub-metre accuracy by any single position-fixing modality, including GPS-denied UWB ranging. A recorded reference dataset and a replay tool accompany the filter, all reported figures are generated directly by the estimator, and the complete system—core, simulator, and a real-time 3D visualization—runs in the browser via WebAssembly and WebGL2.

1 Introduction

Aided inertial navigation estimates a vehicle's position, velocity, and orientation by integrating high-rate inertial measurements and correcting the accumulated drift with lower-rate aiding sensors [12, 13, 14]. Inertial integration alone is untenable over any useful horizon: accelerometer and gyroscope errors are integrated once into velocity and attitude and twice into position, so a bias of even a milli-g becomes metres of position error within seconds. Aiding sensors—GPS, a barometer, a camera, a radio ranging system—bound that drift, but only if the fusion that combines them is correct in a specific and often overlooked sense: it must report an honest uncertainty.

Orientation makes the estimation problem subtle. A unit quaternion has four components but only three degrees of freedom, so a filter that estimates it directly must either enforce the unit-norm constraint or carry a covariance for a quantity that is not free to move in four directions; either choice is awkward and, done naively, inconsistent. The *error-state* (equivalently indirect, or multiplicative) Kalman filter resolves this by estimating a minimal three-parameter rotation error in the tangent space of the current attitude estimate [5, 4, 2]. The nominal orientation absorbs the full, unconstrained rotation; the filter's covariance describes only a small perturbation about it, which is genuinely a three-vector and whose linearization is accurate because the perturbation is kept near zero by frequent resets.

The distinction this paper insists on is between an estimator that *tracks* well and one that is *consistent*. A tracker minimizes error; a consistent filter additionally reports a covariance that matches that error in distribution. The difference matters because a covariance is not decoration: downstream consumers—a path planner deciding whether a gap is safely traversable, a controller weighting a control action by confidence, a second estimator fusing this one's output—all trust the reported uncertainty. An overconfident filter (covariance too small) invites those consumers to act on noise as though it were signal; an underconfident one wastes information. A subtly wrong Jacobian or a mis-scaled process-noise term frequently leaves the RMSE almost unchanged while corrupting the covariance, so tracking error alone cannot detect it. The normalized estimation error squared (NEES) can, and it is the gate around which this work is organized.

Contributions. The contribution is one of rigor and reproducibility rather than novelty of formulation:

- a complete, self-contained derivation of the local error-state kinematics, their discretization, and the exact error-state Jacobian of each of **nine** measurement models, with every Jacobian checked against a central finite difference to 10^{-4};
- a discussion of the **observability** of the fused system that explains, rather than merely reports, why each sensor suite recovers the states it does;
- a verification methodology built on the **NEES consistency gate**, evaluated both pooled and as a function of time, complemented by an online **NIS** outlier test;
- full **reproducibility**: a recorded reference dataset, a replay harness, and figure data emitted directly by the filter, so every number and curve in this paper is real estimator output;
- an implementation whose numerical core is dependency-free and free of unsafe code, and which runs unchanged from a native test harness to a real-time browser visualization.

Source, the dataset, and an interactive demonstration are available at `https://github.com/Zarathos94/ekf-sim` and `https://irfansaric.com/showcase/ekf-sim`.

Outline. Section 2 fixes notation and frame conventions. Section 3 surveys related work and public datasets. Section 4 defines the nominal and error states. Sections 5 and 6 derive and discretize the error dynamics. Section 7 gives the nine measurement models, and Section 8 analyzes their observability. Section 9 covers injection, reset, and numerical care; Section 10 the consistency methodology. Sections 11–12 describe the simulation environment and present results, and Sections 13–15 cover implementation, limitations, and conclusions.

2 Notation and Frame Conventions

Two reference frames are used: a local-level *world* frame W (a fixed east-north-up tangent frame, with gravity along $-\mathbf{e}_3$) and a *body* frame B fixed to the vehicle. Vectors are columns in $\mathbb{R}^3$; $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ are the standard basis. The orientation of the body in the world is a *Hamilton* unit quaternion $\mathbf{q} \in \mathbb{S}^3$ with real part first, $\mathbf{q} = (q_w, q_x, q_y, q_z)$, and $\mathbf{R} = \mathbf{R}(\mathbf{q}) \in SO(3)$ is the corresponding rotation matrix mapping *body to world*, so a body-frame vector $\mathbf{u}_B$ appears in the world as $\mathbf{R}\mathbf{u}_B$ and a world vector in the body as $\mathbf{R}^\top \mathbf{u}_W$. We use the Hamilton convention throughout (as in [2]); it differs from the JPL convention of [3] by the handedness of the quaternion product and a transpose in $\mathbf{R}(\mathbf{q})$, a distinction that is a frequent source of sign errors and is spelled out in Appendix B.

The skew-symmetric (cross-product) operator $\lfloor\cdot\rfloor_\times : \mathbb{R}^3 \to \mathfrak{so}(3)$ is

$$\lfloor \mathbf{a} \rfloor_\times = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}, \qquad \lfloor \mathbf{a} \rfloor_\times \mathbf{v} = \mathbf{a} \times \mathbf{v}. \tag{1}$$

The maps $\mathrm{Exp} : \mathbb{R}^3 \to SO(3)$ and its inverse Log are the group exponential and logarithm (Appendix A). For a manifold state we write the generalized addition and subtraction $\boxplus$ and $\boxminus$: for orientation, $\mathbf{q} \boxplus \delta\boldsymbol{\theta} = \mathbf{q} \otimes \mathrm{Exp}(\frac{1}{2}\delta\boldsymbol{\theta})$ and $\mathbf{q}_2 \boxminus \mathbf{q}_1 = \mathrm{Log}(\mathbf{q}_1^{-1} \otimes \mathbf{q}_2)$; for vector states they reduce to ordinary $+$ and $-$. Expectation is $\mathbb{E}[\cdot]$, and $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes a Gaussian. Hats denote estimates and a leading δ denotes an error quantity.

3 Related Work and Datasets

The indirect Kalman filter for attitude dates to spacecraft navigation; its canonical form is the multiplicative EKF (MEKF) of Lefferts, Markley, and Shuster [4], surveyed by Markley [5], Crassidis et al. [6], and, in quaternion form, by Trawny and Roumeliotis [3] and Solà [2]; it is the attitude workhorse of aided-INS texts [12, 13, 14] and of the modern tutorial treatment of Kok et al. [15]. For visual-inertial estimation the multi-state constraint Kalman filter (MSCKF) of Mourikis and Roumeliotis [16] keeps a sliding window of poses, and optimization-based systems such as VINS-Mono [17] and the factor-graph / incremental-smoothing approach of iSAM2 and preintegration [19, 20] trade recursion for re-linearization. The invariant EKF of Barrau and Bonnabel [18] improves consistency by exploiting the group structure of the state; the filter here uses the more common local-error parameterization but demonstrates consistency empirically and diagnoses it through observability. Radio ranging with UWB for micro-aerial vehicles follows Mueller et al. [25], and metric optical flow follows Honegger et al. [26].

Consistency and observability are tightly linked in this literature. Huang et al. [21] and Hesch et al. [22] showed that a standard EKF for visual-inertial navigation gains spurious information along the four truly unobservable directions of the problem (global position and heading), which manifests as an *inconsistent* filter; observability-constrained variants repair it. Closed-form observability analyses of inertial fusion are given by Martinelli [23]. We revisit these results in Section 8 to explain our scenario outcomes.

Public datasets are the standard proving ground for such estimators: EuRoC MAV [27] and TUM-VI [28] provide synchronized IMU, stereo imagery, and millimetre-accurate motion capture ground truth; UZH-FPV [29] targets aggressive first-person-view flight; and KITTI [30] provides automotive IMU, GNSS, and LiDAR. Because ground truth on such data is itself an estimate and the sensor suites differ from ours, we validate primarily against an *analytic* trajectory—where truth is exact and the NEES is therefore meaningful to the last digit—and additionally ship a recorded, replayable dataset (Section 11) so that our specific results are reproducible bit-for-bit.

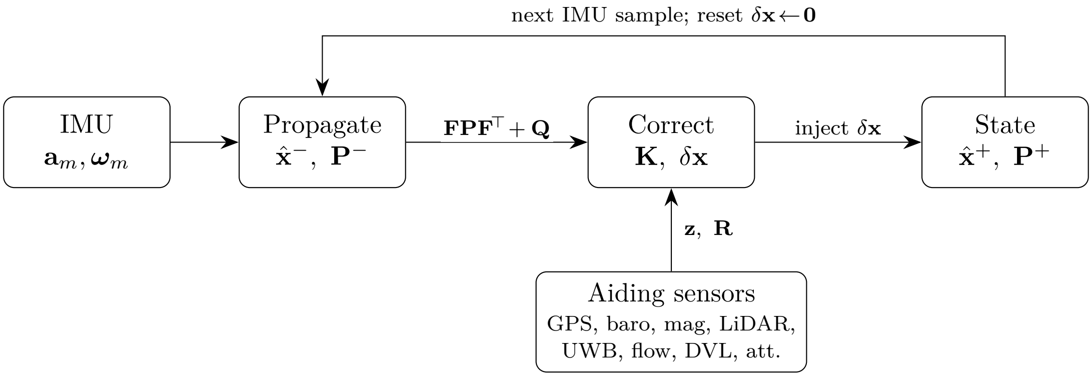

Figure 1: The error-state filter. The IMU drives the nominal state forward and inflates the error-state covariance ($\mathbf{P}^- = \mathbf{FPF}^\top + \mathbf{Q}$); each aiding measurement produces a correction $\delta\mathbf{x}$ that is injected into the nominal state, after which the error state is reset to zero and the cycle repeats on the next IMU sample.

Algorithm 1 Error-state Kalman filter, one cycle

```
Predict (each IMU sample): integrate p, v, q (Eq. 11); set P ← FPF^T + Q (Eqs. 12, 13)
for all available measurements z with model h, Jacobian H, noise R do
    y ← z − h(x̂);   S ← HPH^T + R                          ▷ innovation and its covariance
    if y^T S^{-1} y > γ then reject (NIS gate, Section 10)
    K ← PH^T S^{-1};   δx ← K y
    Inject: x̂ ← x̂ ⊞ δx;   P ← GPG^T;   reset δx ← 0
    P ← (I − KH)P(I − KH)^T + KRK^T                           ▷ Joseph form
end for
```

4 State Representation

The **nominal state** collects position $\mathbf{p} \in \mathbb{R}^3$, velocity $\mathbf{v} \in \mathbb{R}^3$, orientation $\mathbf{q} \in \mathbb{S}^3$ (body to world), and the accelerometer and gyroscope biases $\mathbf{a}_b, \boldsymbol{\omega}_b \in \mathbb{R}^3$:

$$\mathbf{x} = (\mathbf{p},\ \mathbf{v},\ \mathbf{q},\ \mathbf{a}_b,\ \boldsymbol{\omega}_b). \tag{2}$$

It carries the full, unconstrained estimate and is integrated directly from the IMU. The **error state** is the minimal 15-vector

$$\delta\mathbf{x} = (\delta\mathbf{p},\ \delta\mathbf{v},\ \delta\boldsymbol{\theta},\ \delta\mathbf{a}_b,\ \delta\boldsymbol{\omega}_b) \in \mathbb{R}^{15}, \tag{3}$$

in which the rotation error $\delta\boldsymbol{\theta}$ is defined *locally* (in the body frame) through the multiplicative relation

$$\mathbf{q}_{\text{true}} = \mathbf{q} \otimes \operatorname{Exp}\!\left(\tfrac{1}{2}\delta\boldsymbol{\theta}\right), \tag{4}$$

so that the corresponding rotation matrices satisfy $\mathbf{R}_{\text{true}} = \mathbf{R}\ \operatorname{Exp}(\delta\boldsymbol{\theta})$. The error state is reset to zero after every correction, so its 15×15 covariance $\mathbf{P} \in \mathbb{R}^{15\times15}$ stays well conditioned and its orientation block is an ordinary 3×3 matrix rather than a rank-deficient 4×4 one. This separation—an unconstrained nominal state plus a small, minimal, frequently reset error state—is the entire point of the error-state formulation: the extended Kalman filter's linearization is applied only to the small quantity, where it is accurate.

Figure 1 shows the overall dataflow and Algorithm 1 the per-cycle loop; the remainder of the paper fills in each block.

5 Continuous-Time Kinematics and Error Dynamics

We first write the true kinematics, then linearize them about the nominal state to obtain the dynamics of the error state, which the filter propagates.

Nominal kinematics. Let $\mathbf{a}_m$ and $\boldsymbol{\omega}_m$ be the accelerometer and gyroscope readings and define the bias-corrected specific force and body rate $\mathbf{a} = \mathbf{a}_m - \mathbf{a}_b$, $\boldsymbol{\omega} = \boldsymbol{\omega}_m - \boldsymbol{\omega}_b$. With gravity $\mathbf{g} = -g\,\mathbf{e}_3$ ($g = 9.806\,65\,\mathrm{m\,s^{-2}}$), the continuous nominal kinematics are

$$\dot{\mathbf{p}} = \mathbf{v}, \qquad \dot{\mathbf{v}} = \mathbf{R}\mathbf{a} + \mathbf{g}, \qquad \dot{\mathbf{q}} = \tfrac{1}{2}\,\mathbf{q} \otimes \boldsymbol{\omega}, \qquad \dot{\mathbf{a}}_b = \mathbf{0}, \qquad \dot{\boldsymbol{\omega}}_b = \mathbf{0}, \tag{5}$$

where $\boldsymbol{\omega}$ in the quaternion derivative is treated as a pure quaternion $(0, \boldsymbol{\omega})$. The biases are modelled as constant in the mean and driven only by noise (below).

Error kinematics. Write each true quantity as the nominal plus an error: $\mathbf{p}_t = \mathbf{p} + \delta\mathbf{p}$, $\mathbf{v}_t = \mathbf{v} + \delta\mathbf{v}$, $\mathbf{R}_t = \mathbf{R}\,\mathrm{Exp}(\delta\boldsymbol{\theta}) \approx \mathbf{R}(\mathbf{I} + \lfloor\delta\boldsymbol{\theta}\rfloor_\times)$, $\mathbf{a}_{b,t} = \mathbf{a}_b + \delta\mathbf{a}_b$, $\boldsymbol{\omega}_{b,t} = \boldsymbol{\omega}_b + \delta\boldsymbol{\omega}_b$. The IMU readings carry additive white noise $\mathbf{n}_a, \mathbf{n}_\omega$ and the biases random walks $\mathbf{n}_{ab}, \mathbf{n}_{\omega b}$. Substituting into (5) and retaining first-order terms yields the continuous error dynamics (derived in the manner of [2, 3]):

$$\dot{\delta\mathbf{p}} = \delta\mathbf{v}, \tag{6}$$

$$\dot{\delta\mathbf{v}} = -\mathbf{R}\lfloor\mathbf{a}\rfloor_\times\,\delta\boldsymbol{\theta} - \mathbf{R}\,\delta\mathbf{a}_b - \mathbf{R}\,\mathbf{n}_a, \tag{7}$$

$$\dot{\delta\boldsymbol{\theta}} = -\lfloor\boldsymbol{\omega}\rfloor_\times\,\delta\boldsymbol{\theta} - \delta\boldsymbol{\omega}_b - \mathbf{n}_\omega, \tag{8}$$

$$\dot{\delta\mathbf{a}}_b = \mathbf{n}_{ab}, \qquad \dot{\delta\boldsymbol{\omega}}_b = \mathbf{n}_{\omega b}. \tag{9}$$

The two coupling terms are the substance of the model. In (7), perturbing the rotation in $\mathbf{R}_t\mathbf{a}_t = \mathbf{R}(\mathbf{I} + \lfloor\delta\boldsymbol{\theta}\rfloor_\times)(\mathbf{a} - \delta\mathbf{a}_b - \mathbf{n}_a)$ and dropping second-order products gives the term $\mathbf{R}\lfloor\delta\boldsymbol{\theta}\rfloor_\times\mathbf{a} = -\mathbf{R}\lfloor\mathbf{a}\rfloor_\times\delta\boldsymbol{\theta}$ (using $\lfloor\boldsymbol{u}\rfloor_\times\boldsymbol{w} = -\lfloor\boldsymbol{w}\rfloor_\times\boldsymbol{u}$): an attitude error tilts the sensed specific force and thus corrupts the acceleration estimate, which is the dominant mechanism of inertial position drift. In (8), the local-error convention makes the orientation-error dynamics depend on the body rate through $-\lfloor\boldsymbol{\omega}\rfloor_\times$; this is the term whose *exact* discrete counterpart, rather than a first-order truncation, we retain in Section 6. Collecting (6)–(9) as $\dot{\delta\mathbf{x}} = \mathbf{A}\,\delta\mathbf{x} + \mathbf{L}\,\mathbf{n}$ gives the continuous system and noise matrices

$$\mathbf{A} = \begin{bmatrix} \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & -\mathbf{R}\lfloor\mathbf{a}\rfloor_\times & -\mathbf{R} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & -\lfloor\boldsymbol{\omega}\rfloor_\times & \mathbf{0} & -\mathbf{I} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}, \qquad \mathbf{L} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ -\mathbf{R} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix}, \tag{10}$$

with $\mathbf{n} = (\mathbf{n}_a, \mathbf{n}_\omega, \mathbf{n}_{ab}, \mathbf{n}_{\omega b})$ having power spectral density $\mathbf{Q}_c = \mathrm{diag}(\sigma_a^2\mathbf{I},\ \sigma_g^2\mathbf{I},\ \sigma_{ab}^2\mathbf{I},\ \sigma_{\omega b}^2\mathbf{I})$.

6 Discretization

Over one IMU interval Δt (rate $\approx 200\,\mathrm{Hz}$) the nominal state is integrated by

$$\mathbf{p} \leftarrow \mathbf{p} + \mathbf{v}\,\Delta t + \tfrac{1}{2}(\mathbf{Ra}+\mathbf{g})\,\Delta t^2, \quad \mathbf{v} \leftarrow \mathbf{v} + (\mathbf{Ra}+\mathbf{g})\,\Delta t, \quad \mathbf{q} \leftarrow \mathbf{q} \otimes \mathrm{Exp}\big(\tfrac{1}{2}\boldsymbol{\omega}\,\Delta t\big), \tag{11}$$

with the quaternion renormalized after each step. The error-state covariance is propagated by $\mathbf{P} \leftarrow \mathbf{FPF}^\top + \mathbf{Q}$, where $\mathbf{F} = \mathrm{Exp}_A(\mathbf{A}\,\Delta t)$ is the discrete state transition. Evaluating the matrix exponential blockwise, the nilpotent couplings contribute only their first-order terms while the orientation block is kept exact:

$$\mathbf{F} = \begin{bmatrix} \mathbf{I} & \mathbf{I}\,\Delta t & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & -\mathbf{R}\lfloor\mathbf{a}\rfloor_\times\,\Delta t & -\mathbf{R}\,\Delta t & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathrm{Exp}(-\boldsymbol{\omega}\,\Delta t) & \mathbf{0} & -\mathbf{I}\,\Delta t \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix}. \tag{12}$$

Here $\mathrm{Exp}(-\boldsymbol{\omega}\,\Delta t) = \exp(-\lfloor\boldsymbol{\omega}\rfloor_\times\Delta t)$ is the exact rotation integral (Appendix A) rather than $\mathbf{I} - \lfloor\boldsymbol{\omega}\rfloor_\times\Delta t$; the difference is third order in $\boldsymbol{\omega}\Delta t$ but accumulates at the high body rates of aggressive flight, and—being an orientation term—it is exactly the kind of small error the consistency test detects.

The discrete process noise is $\mathbf{Q} = \int_0^{\Delta t}\mathbf{F}(\tau)\,\mathbf{LQ}_c\mathbf{L}^\top\mathbf{F}(\tau)^\top\,d\tau$, which for a step this short is accurately approximated by $\mathbf{LQ}_c\mathbf{L}^\top\Delta t$ evaluated with the block structure of (10), giving the diagonal form

$$\mathbf{Q} = \mathrm{diag}\big(\mathbf{0},\ \sigma_a^2\,\Delta t^2\,\mathbf{I},\ \sigma_g^2\,\Delta t^2\,\mathbf{I},\ \sigma_{ab}^2\,\Delta t\,\mathbf{I},\ \sigma_{\omega b}^2\,\Delta t\,\mathbf{I}\big). \tag{13}$$

The scaling deserves emphasis, because getting it wrong is the archetypal invisible bug. A per-sample accelerometer noise of standard deviation σ_a perturbs the integrated velocity by $\sigma_a\Delta t$ over the step and therefore injects a *variance* $\sigma_a^2\Delta t^2$; the same holds for the orientation term. This is also the only dimensionally consistent choice: σ_a has units of acceleration spectral density, and only $\sigma_a^2\Delta t^2$ produces a velocity variance. The bias random walks, by contrast, accumulate variance linearly and scale as Δt. A $\mathbf{Q}$ built with the wrong power of Δt can leave the RMSE almost unchanged while making the filter badly overconfident or conservative—a discrepancy that is glaring in the NEES of Section 10, and which this test caught during development.

7 Measurement Updates

For a measurement $\mathbf{z} = h(\mathbf{x}) + \boldsymbol{\nu}$, $\boldsymbol{\nu} \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$, the error-state Jacobian $\mathbf{H} = \partial h/\partial\delta\mathbf{x}$ is evaluated at the nominal state and applied through Algorithm 1. Because the error state is zero at the linearization point, the Jacobians take a common form: a measurement that depends on orientation does so through $\mathbf{R}^\top\boldsymbol{u}$ for some world vector $\boldsymbol{u}$, and the local perturbation $\mathbf{R}_t^\top = (\mathbf{I} - \lfloor\delta\boldsymbol{\theta}\rfloor_\times)\mathbf{R}^\top$ gives $\mathbf{R}_t^\top\boldsymbol{u} = \mathbf{R}^\top\boldsymbol{u} + \lfloor\mathbf{R}^\top\boldsymbol{u}\rfloor_\times\,\delta\boldsymbol{\theta}$ (using $-\lfloor\delta\boldsymbol{\theta}\rfloor_\times\boldsymbol{w} = \lfloor\boldsymbol{w}\rfloor_\times\delta\boldsymbol{\theta}$), hence

$$\left.\frac{\partial(\mathbf{R}^\top\boldsymbol{u})}{\partial\delta\boldsymbol{\theta}}\right|_{\delta\boldsymbol{\theta}=\mathbf{0}} = \lfloor\mathbf{R}^\top\boldsymbol{u}\rfloor_\times, \tag{14}$$

the pattern shared by the magnetometer, optical-flow, and DVL updates below. The nine models, with the non-zero blocks of $\mathbf{H}$, are as follows.

GPS position and Doppler velocity. Absolute position and world-frame velocity, $h = \mathbf{p}$ and $h = \mathbf{v}$, with $\partial h/\partial\delta\mathbf{p} = \mathbf{I}_3$ and $\partial h/\partial\delta\mathbf{v} = \mathbf{I}_3$ respectively. These are the only absolute references and, as Section 8 shows, the states that make the rest observable.

Barometer. Altitude only, $h = \mathbf{e}_3^\top\mathbf{p} = p_z$, with $\partial h/\partial\delta\mathbf{p} = \mathbf{e}_3^\top$.

Magnetometer. The world magnetic field $\mathbf{m}$ expressed in the body frame, $h = \mathbf{R}^\top\mathbf{m}$; by (14), $\partial h/\partial\delta\boldsymbol{\theta} = \lfloor\mathbf{R}^\top\mathbf{m}\rfloor_\times$. Treating the full field vector as the measurement, rather than collapsing it to a scalar heading, keeps the update linear in the error state and avoids the singularities of an explicit yaw extraction.

LiDAR altimeter. A downward beam ranging to flat ground senses the slant range $h = p_z/R_{33}$, where $R_{33} = \mathbf{e}_3^\top\mathbf{R}\,\mathbf{e}_3$ is the cosine between the body-down and world-down axes. Then $\partial h/\partial p_z = 1/R_{33}$, and differentiating R_{33} through $\mathbf{R}_t = \mathbf{R}\,\mathrm{Exp}(\delta\boldsymbol{\theta})$ gives $\partial h/\partial\delta\boldsymbol{\theta} = -\frac{p_z}{R_{33}^2}\,[-R_{32},\ R_{31},\ 0]$, so a single beam senses height *and* tilt.

UWB radio ranging. Scalar range to each fixed beacon $\mathbf{b}_k$, $h = \|\mathbf{p}-\mathbf{b}_k\|$, with $\partial h/\partial\delta\mathbf{p} = \hat{\mathbf{u}}^\top$, the unit line of sight from beacon to vehicle. A few beacons at surveyed positions make absolute position observable with no GPS at all [25].

Optical flow. The body-frame horizontal velocity, $h = \Pi\,\mathbf{R}^\top\mathbf{v}$ with $\Pi = [\mathbf{I}_2\ \ \mathbf{0}]$ selecting the first two components; then $\partial h/\partial\delta\mathbf{v} = \Pi\mathbf{R}^\top$ and, by (14), $\partial h/\partial\delta\boldsymbol{\theta} = \Pi\lfloor\mathbf{R}^\top\mathbf{v}\rfloor_\times$ [26].

Doppler velocity log (DVL). The full body-frame velocity, $h = \mathbf{R}^\top\mathbf{v}$, with $\partial h/\partial\delta\mathbf{v} = \mathbf{R}^\top$ and $\partial h/\partial\delta\boldsymbol{\theta} = \lfloor\mathbf{R}^\top\mathbf{v}\rfloor_\times$.

Attitude fix. A direct orientation measurement $\mathbf{z}_q$ (star tracker or vision): the innovation is the minimal rotation $\boldsymbol{y} = \mathrm{Log}(\hat{\mathbf{q}}^{-1}\otimes\mathbf{z}_q) = \mathbf{z}_q \boxminus \hat{\mathbf{q}}$ and $\partial h/\partial\delta\boldsymbol{\theta} = \mathbf{I}_3$—the standard multiplicative attitude update.

Every Jacobian is checked in the test suite against a central finite difference across all 15 error directions, agreeing to 10^{-4}; the attitude and velocity updates are additionally verified to converge from a seeded error, which fails if a sign is wrong. All updates use the Joseph-form covariance step of Algorithm 1, which remains symmetric positive-definite under finite precision. Table 1 summarizes the strapdown IMU and nine aiding modalities and the default noise levels used throughout.

Modality	Symbol dim.	Rate (Hz)	Noise σ	Frame	States aided
IMU (accel.)	3	200	$0.06\,\mathrm{m\,s^{-2}}$	body	propagation
IMU (gyro.)	3	200	$0.004\,\mathrm{rad\,s^{-1}}$	body	propagation
GPS position	3	5	$0.8\,\mathrm{m}$	world	$\mathbf{p}$
GPS Doppler vel.	3	5	$0.1\,\mathrm{m\,s^{-1}}$	world	$\mathbf{v}$
Barometer	1	20	$0.6\,\mathrm{m}$	world	p_z
Magnetometer	3	40	0.02 (unit)	body	$\delta\boldsymbol{\theta}$ (yaw)
LiDAR altimeter	1	25	$0.15\,\mathrm{m}$	body	$p_z, \delta\boldsymbol{\theta}$
UWB ranging	4	10	$0.35\,\mathrm{m}$	world	$\mathbf{p}$
Optical flow	2	30	$0.15\,\mathrm{m\,s^{-1}}$	body	$\mathbf{v}, \delta\boldsymbol{\theta}$
DVL / radar vel.	3	20	$0.05\,\mathrm{m\,s^{-1}}$	body	$\mathbf{v}, \delta\boldsymbol{\theta}$
Attitude fix	3	20	$0.01\,\mathrm{rad}$	–	$\delta\boldsymbol{\theta}$

Table 1: The strapdown IMU and nine aiding modalities, with default rates and noise standard deviations. UWB reports one range per beacon (four beacons); the magnetometer and DVL/flow aid orientation through the coupling of (14).

8 Observability

Which states a given sensor suite can actually pin down—its *observability*—explains the results of Section 12 far better than the RMSE numbers alone, and it is where an error-state filter can quietly become inconsistent if the analysis is ignored [21, 22].

With an absolute position reference (GPS or UWB), position and velocity are directly observable, and the attitude becomes observable through the inertial coupling of (7): an orientation error tilts the sensed specific force, so as the vehicle accelerates the resulting position/velocity residual reveals the tilt. Roll and pitch are additionally anchored by the gravity direction that the accelerometer senses whenever the specific force is not purely gravitational. Heading (yaw) is the delicate one: under constant-velocity motion with only inertial and position aiding it is *unobservable*, and the filter must not manufacture confidence about it—precisely the failure mode that observability-constrained EKFs were introduced to prevent [22, 23]. A magnetometer removes the ambiguity by measuring the world field directly (Section 7); so does sustained maneuvering, which couples yaw into the observable horizontal channels.

The gyroscope bias is observable whenever orientation is, and the accelerometer bias is observable when the vehicle rotates, since rotation modulates how a fixed body-frame bias projects into the world acceleration. This is why the analytic trajectory used here is a *banking* helix rather than a straight climb: the continuous change of attitude excites every bias direction and renders the full 15-state vector observable, which is a precondition for the NEES over all 15 dimensions to be meaningful.

These facts map directly onto the scenarios of Table 3. Removing all aiding leaves an unobservable, freely drifting integrator (dead reckoning). Restoring *any* absolute position sensor—GPS, or UWB ranging to surveyed beacons—restores observability of position and, through the inertial coupling, of the rest; this is why GPS-denied UWB ranging recovers sub-metre accuracy, and why the indoor suite (UWB, LiDAR, optical flow) needs no GPS. The vision-aided suite adds a direct attitude fix, which makes orientation observable outright and tightens the attitude RMSE by nearly an order of magnitude.

9 Injection, Reset, and Numerical Care

Once a correction $\delta\mathbf{x}$ is computed, it is folded into the nominal state and the error is reset to zero:

$$\mathbf{p} \leftarrow \mathbf{p} + \delta\mathbf{p}, \quad \mathbf{v} \leftarrow \mathbf{v} + \delta\mathbf{v}, \quad \mathbf{q} \leftarrow \mathbf{q} \otimes \mathrm{Exp}\big(\tfrac{1}{2}\delta\boldsymbol{\theta}\big), \quad \mathbf{a}_b \leftarrow \mathbf{a}_b + \delta\mathbf{a}_b, \quad \boldsymbol{\omega}_b \leftarrow \boldsymbol{\omega}_b + \delta\boldsymbol{\omega}_b. \tag{15}$$

Because the reset moves the linearization point of the orientation error onto the new nominal quaternion, the covariance must be transported by the reset Jacobian

$$\mathbf{G} = \mathrm{diag}\big(\mathbf{I},\ \mathbf{I},\ \mathbf{I} - \tfrac{1}{2}\lfloor\delta\boldsymbol{\theta}\rfloor_\times,\ \mathbf{I},\ \mathbf{I}\big), \qquad \mathbf{P} \leftarrow \mathbf{G}\,\mathbf{P}\,\mathbf{G}^\top. \tag{16}$$

Retaining the $\mathbf{I} - \frac{1}{2}\lfloor\delta\boldsymbol{\theta}\rfloor_\times$ term—frequently dropped, since it is within a whisker of the identity for small corrections—is one of the details the consistency test rewards. Three further numerical safeguards matter in finite precision: the covariance update uses the Joseph form $(\mathbf{I}-\mathbf{KH})\mathbf{P}(\mathbf{I}-\mathbf{KH})^\top + \mathbf{KRK}^\top$, which is positive-semidefinite by construction where the shorter $(\mathbf{I}-\mathbf{KH})\mathbf{P}$ form is not; the covariance is symmetrized as $\frac{1}{2}(\mathbf{P}+\mathbf{P}^\top)$ after each step to suppress asymmetry from round-off; and the quaternion is renormalized after every integration and injection so it never drifts off the unit sphere.

10 Consistency: NEES and NIS

The decisive test is *statistical consistency* [8]. Define the normalized estimation error squared

$$\varepsilon = \boldsymbol{e}^\top\mathbf{P}^{-1}\boldsymbol{e}, \qquad \boldsymbol{e} = \hat{\mathbf{x}} \boxminus \mathbf{x}_{\mathrm{true}} \in \mathbb{R}^{15}, \tag{17}$$

where the error is expressed in the filter's own tangent space: position and velocity differences, the body-frame logarithm $\mathrm{Log}(\hat{\mathbf{q}}^{-1}\otimes\mathbf{q}_{\mathrm{true}})$ for orientation, and bias differences. For a consistent filter whose error is zero-mean Gaussian with covariance $\mathbf{P}$, the quadratic form is chi-square distributed, so $\mathbb{E}[\varepsilon] = \dim$ (3 for position, 15 for the full state). Averaged over N independent Monte-Carlo runs, $N\bar{\varepsilon} \sim \chi^2(N\cdot\dim)$, giving the two-sided 95% acceptance interval

$$\bar{\varepsilon} \in \left[\ \frac{\chi^2_{0.025}(N\cdot\dim)}{N},\ \frac{\chi^2_{0.975}(N\cdot\dim)}{N}\ \right]. \tag{18}$$

Overconfidence lands the average above the band, conservatism below. Two methodological points make the test honest. First, the independent unit is the *flight*, not the sample: NEES values within a run are time-correlated, so the band is set by the number of runs, and pooling steady-state samples only sharpens the estimate inside that band. Second, the runs are seeded from a realistic initial error drawn from the initial covariance, so the test scores the filter rather than an omniscient initialization, and the NEES is reported per five-second window as well as pooled, because consistency must hold throughout the flight and not merely on average. The χ^2 quantiles are evaluated with the Wilson–Hilferty approximation, which is accurate to a fraction of a percent at the degrees of freedom used here.

The same machinery, applied to the *innovation* rather than the estimation error, gives an online outlier gate. The normalized innovation squared $\boldsymbol{y}^\top\mathbf{S}^{-1}\boldsymbol{y}$ is χ^2 with dimension equal to that of the measurement, so a measurement whose NIS exceeds the upper quantile $\gamma = \chi^2_{0.99}(\dim\mathbf{z})$ is rejected before it corrupts the state (Algorithm 1, line 4). Unlike the NEES, the NIS needs no ground truth and runs live in the browser.

Quantity	NEES mean	Expected	95% band
Position (3-DOF)	2.85	3.0	[2.29, 3.81]
Full state (15-DOF)	16.23	15.0	[13.35, 16.74]

Table 2: NEES consistency over 40 Monte-Carlo flights, nine sensors active. Both statistics lie inside the band—the reported uncertainty is neither over- nor under-confident.

11 Simulation Environment and Test Data

Ground truth is an analytic banking helical trajectory—a horizontal orbit of radius 40 m and period 30 s with a superimposed vertical oscillation—so position, velocity, and acceleration are available in closed form and are exact. Orientation follows a coordinated turn (nose along the velocity, rolled into the turn by the bank angle that balances lateral acceleration against gravity), and the body angular rate is obtained by central-differencing the quaternion path. As Section 8 noted, the continual banking is deliberate: it excites every bias direction and makes the full state observable.

From this truth a strapdown IMU and the nine aiding sensors are synthesized, each with configurable noise, bias random walk, rate, and availability (Table 1). A `record` command writes the raw IMU stream, every aiding measurement, and ground truth to CSV; a `replay` command re-runs the filter on that CSV and re-derives the RMSE and NEES, so the dataset is an independently verifiable artifact rather than a live-only result. A `plotdata` command emits the exact series plotted in the next section. The filter is deliberately given a process noise that differs from the simulator's true noise—a filter never knows the truth—and it must remain consistent anyway.

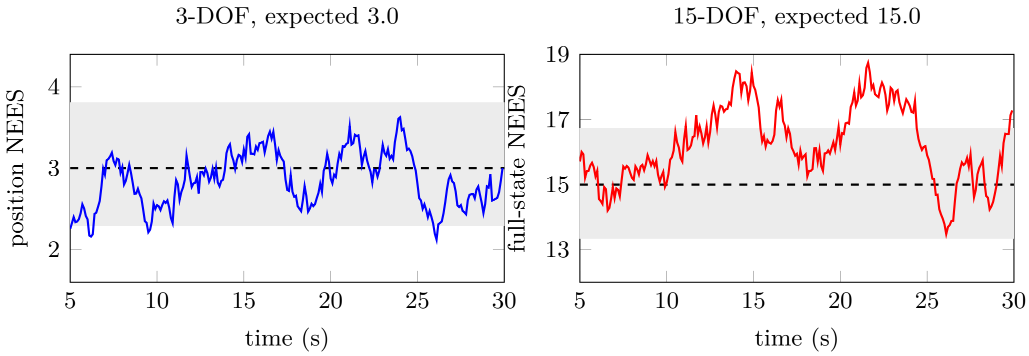

Figure 2: Ensemble-mean NEES over 40 flights as a function of time (blue/red), the expected value (dashed), and the 95% acceptance band (shaded). The mean tracks the expected dimension for the whole flight; the filter is consistent throughout, not merely on average.

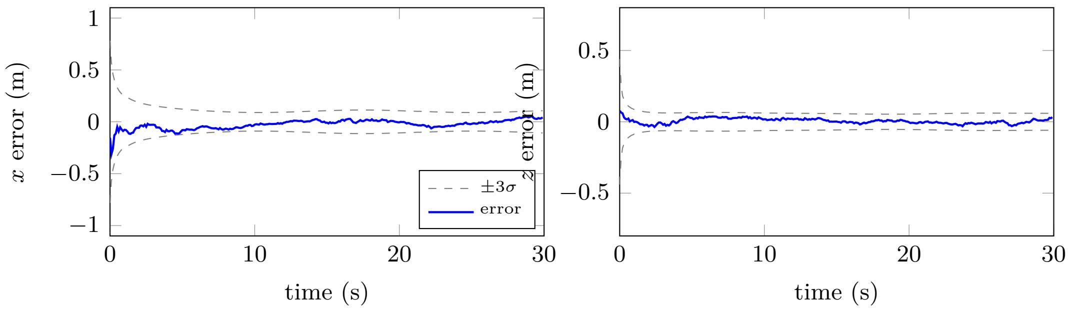

Figure 3: One flight's position error (blue) against its own $\pm 3\sigma$ envelope (dashed), for the horizontal x axis (left) and the vertical z axis (right). The envelope contracts as the filter converges and contains the error—consistency seen for a single run.

12 Results

All numbers and curves in this section are produced by the same `eskf` core that runs in the browser, via the command-line harness.

Consistency. Table 2 reports the Monte-Carlo consistency check over 40 flights of 30 s with all nine sensors active. Both the position and full-state NEES fall inside the band of (18). Figure 2 plots the ensemble-mean NEES as a function of time against the same band: it neither climbs (which would signal a filter growing overconfident, the classic inconsistency of an unconstrained EKF [21]) nor sags, but sits on the expected value throughout. The time-averaged RMSE is 0.049 m in position, 0.021 m s^{-1} in velocity, and 0.11° in attitude; replaying the recorded dataset independently reproduces 0.046 m.

Error against reported uncertainty. Figure 3 shows a single flight's position error against the $\pm 3\sigma$ envelope drawn from the filter's own covariance, for the horizontal (x) and vertical (z) axes. The envelope starts wide, reflecting the seeded initial uncertainty, and contracts as aiding information accumulates; the error stays within it. An envelope that the error escaped, or one far larger than the error, would both be visible here at a glance—this is the per-flight, per-axis view behind the aggregate NEES.

Scenario	Position (m)	Velocity (m/s)	Attitude (°)
Nominal (GPS + baro + mag)	0.47	0.24	0.99
IMU only (dead reckoning)	295	25.4	8.32
GPS dropout	16.6	1.87	2.05
GPS denied + UWB ranging	0.16	0.10	0.61
Indoor (UWB + LiDAR + flow)	0.08	0.04	0.23
Vision-aided (attitude + DVL)	0.14	0.03	0.12
Full fusion (all 9 sensors)	0.06	0.02	0.11
Noisy IMU (5×)	0.48	0.26	1.06
Bias drift (10×)	0.52	0.33	1.53

Table 3: Steady-state RMSE by scenario (mean of 12 flights of 30 s each). Dead reckoning diverges; any single absolute position fix, including UWB alone, recovers it.

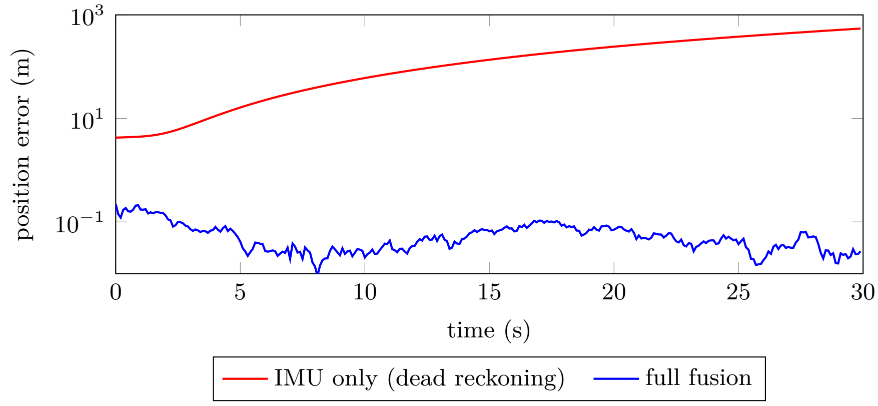

Figure 4: Position error versus time (log scale). Unaided inertial integration diverges to hundreds of metres as bias is integrated twice into position; the full nine-sensor fusion holds a few centimetres. The gap is five orders of magnitude by 30 s.

Behaviour across sensor suites. Table 3 shows steady-state RMSE as the sensor suite is changed, and Figure 4 contrasts the two extremes over time. With *all* aiding removed the filter dead-reckons on the IMU alone and the position error reaches hundreds of metres over 30 s—the quadratic error growth that accelerometer and gyroscope bias predict for an unaided strapdown system [13, 14], and exactly what observability theory requires of an unobservable integrator. Any single position-fixing modality restores sub-metre accuracy: GPS-denied UWB ranging recovers from 16.6 m to 0.16 m, and the indoor and vision-aided suites—using no GPS at all—stay well under a metre, consistent with the observability argument of Section 8.

13 Implementation and Reproducibility

The system is a small Rust workspace. The numerical core (`eskf`) has no dependencies, forbids unsafe code, and carries its own fixed-size linear algebra, quaternion algebra, simulator, and NEES metric; it is exercised by a native test suite that includes the finite-difference Jacobian checks and the exp/log round-trips of Appendix A. A command-line harness (`eskf-cli`) is the correctness gate: `check` runs the Monte-Carlo NEES test of Section 10, `scenarios` produces Table 3, `record/replay` manage the reference dataset, and `plotdata` emits the series behind Figures 2–4. The identical core is compiled to WebAssembly (`eskf-wasm`) and driven by a hand-written WebGL2 front end that renders the estimate, ground truth, and live covariance ellipsoid in real time, with per-sensor toggles and noise controls. Because the browser and the gate share one implementation, the consistency proven at the command line is the consistency running in the demonstration. The reference dataset and the figure data are committed alongside the source, so every table and figure in this paper can be regenerated from a clean checkout with a single command.

14 Limitations and Future Work

The filter is a first-order (extended) estimator with additive Gaussian noise, and it inherits the usual limitations. Linearization error grows with the update interval and the motion aggressiveness; an iterated update or a sigma-point (unscented) propagation [24] would reduce it at some cost. The local-error parameterization is consistent here but is known to accrue spurious information along unobservable directions in the harder visual-inertial case; the invariant/equivariant EKF [18] addresses this at the level of the state's group structure and is the natural next step. Measurements are assumed time-synchronized and in-sequence; real systems need delayed-state or out-of-sequence handling. Finally, validation here is against a simulator with exact ground truth, which is what makes the NEES meaningful; the natural complement is a run against a public benchmark with motion capture, such as EuRoC [27] or TUM-VI [28], where the sensor models would have to be matched to real hardware.

15 Conclusion

We have presented a quaternion error-state Kalman filter fusing up to nine aiding sensors for inertial navigation, in the standard local-error formulation, with the error kinematics derived from first principles, discretized with care, and each of the nine measurement Jacobians given and finite-difference-verified. The emphasis throughout is that correctness is *demonstrated*, not asserted: the full estimation loop passes a Monte-Carlo NEES consistency gate in aggregate and over time, rejects outliers with a companion NIS test, and reproduces the inertial-navigation and observability behaviour theory predicts—unbounded dead-reckoning drift, and graceful recovery from any single absolute position aid, including GPS-denied ranging. A recorded dataset, a replay tool, and figure data emitted by the filter itself make every result reproducible, and the complete system runs unchanged from the command-line gate to a real-time browser visualization.

A The $SO(3)$ Exponential and Logarithm

For a rotation vector $\boldsymbol{\phi} \in \mathbb{R}^3$ with $\theta = \|\boldsymbol{\phi}\|$ and $\hat{\boldsymbol{\phi}} = \boldsymbol{\phi}/\theta$, the group exponential is Rodrigues' formula

$$\mathrm{Exp}(\boldsymbol{\phi}) = \exp(\lfloor \boldsymbol{\phi} \rfloor_\times) = \mathbf{I} + \frac{\sin\theta}{\theta} \lfloor \boldsymbol{\phi} \rfloor_\times + \frac{1-\cos\theta}{\theta^2} \lfloor \boldsymbol{\phi} \rfloor_\times^2, \tag{19}$$

with the small-angle limit $\mathbf{I} + \lfloor \boldsymbol{\phi} \rfloor_\times$ (used with a series near $\theta = 0$ to avoid the removable singularity); the corresponding unit quaternion is $\mathrm{Exp}(\frac{1}{2}\boldsymbol{\phi}) = \big(\cos\frac{\theta}{2},\ \sin\frac{\theta}{2}\,\hat{\boldsymbol{\phi}}\big)$. The orientation block of (12) is $\mathrm{Exp}(-\boldsymbol{\omega}\,\Delta t)$ evaluated with this formula rather than truncated. The inverse (the logarithmic map) recovers the rotation vector from a unit quaternion $(q_w, \boldsymbol{q}_v)$ as

$$\mathrm{Log}(q_w, \boldsymbol{q}_v) = 2\ \mathrm{atan2}\big(\|\boldsymbol{q}_v\|,\ q_w\big)\,\frac{\boldsymbol{q}_v}{\|\boldsymbol{q}_v\|}, \tag{20}$$

and defines the $\boxminus$ used to form the orientation component of the estimation error in Section 10. Both maps are unit-tested for round-trip accuracy and against known rotations.

B Quaternion Algebra and Conventions

We use Hamilton quaternions with the real part first, $\mathbf{q} = (q_w, \boldsymbol{q}_v)$, and the product

$$\mathbf{q}_1 \otimes \mathbf{q}_2 = \begin{pmatrix} q_{1w}q_{2w} - \boldsymbol{q}_{1v}\cdot\boldsymbol{q}_{2v} \\ q_{1w}\boldsymbol{q}_{2v} + q_{2w}\boldsymbol{q}_{1v} + \boldsymbol{q}_{1v} \times \boldsymbol{q}_{2v} \end{pmatrix}, \tag{21}$$

which is right-handed ($\boldsymbol{ij} = \boldsymbol{k}$). The rotation matrix mapping body to world is

$$\mathbf{R}(\mathbf{q}) = (q_w^2 - \boldsymbol{q}_v^\top \boldsymbol{q}_v)\mathbf{I} + 2\,\boldsymbol{q}_v\boldsymbol{q}_v^\top + 2\,q_w \lfloor \boldsymbol{q}_v \rfloor_\times, \tag{22}$$

and satisfies $\mathbf{R}(\mathbf{q}_1 \otimes \mathbf{q}_2) = \mathbf{R}(\mathbf{q}_1)\mathbf{R}(\mathbf{q}_2)$. The local (body-frame) perturbation $\mathbf{q} \otimes \mathrm{Exp}(\frac{1}{2}\delta\boldsymbol{\theta})$ of (4) corresponds to a right multiplication of $\mathbf{R}$, $\mathbf{R}\ \mathrm{Exp}(\delta\boldsymbol{\theta})$, which is why every orientation-dependent measurement Jacobian in Section 7 enters through $\mathbf{R}^\top$ and the pattern of (14). The JPL convention of [3] instead orders the product left-handedly and transposes $\mathbf{R}(\mathbf{q})$; the two conventions describe the same physics but differ by transposes that, if mixed, produce sign errors in exactly the coupling terms of (10). Fixing one convention—here, Hamilton—and testing the rotation composition against direction-cosine matrices is the practical guard against that class of bug.

References

[1] R. E. Kalman, "A new approach to linear filtering and prediction problems," *J. Basic Eng.*, vol. 82, no. 1, pp. 35–45, 1960.

[2] J. Solà, "Quaternion kinematics for the error-state Kalman filter," arXiv:1711.02508, 2017.

[3] N. Trawny and S. I. Roumeliotis, "Indirect Kalman filter for 3D attitude estimation," Univ. of Minnesota, Tech. Rep. 2005-002, 2005.

[4] E. J. Lefferts, F. L. Markley, and M. D. Shuster, "Kalman filtering for spacecraft attitude estimation," *J. Guidance, Control, and Dynamics*, vol. 5, no. 5, pp. 417–429, 1982.

[5] F. L. Markley, "Attitude error representations for Kalman filtering," *J. Guidance, Control, and Dynamics*, vol. 26, no. 2, pp. 311–317, 2003.

[6] J. L. Crassidis, F. L. Markley, and Y. Cheng, "Survey of nonlinear attitude estimation methods," *J. Guidance, Control, and Dynamics*, vol. 30, no. 1, pp. 12–28, 2007.

[7] M. D. Shuster, "A survey of attitude representations," *J. Astronautical Sciences*, vol. 41, no. 4, pp. 439–517, 1993.

[8] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation*. New York: Wiley, 2001.

[9] P. S. Maybeck, *Stochastic Models, Estimation, and Control, Vol. 1*. New York: Academic Press, 1979.

[10] A. Gelb (ed.), *Applied Optimal Estimation*. Cambridge, MA: MIT Press, 1974.

[11] C. F. Van Loan, "Computing integrals involving the matrix exponential," *IEEE Trans. Automatic Control*, vol. 23, no. 3, pp. 395–404, 1978.

[12] J. A. Farrell, *Aided Navigation: GPS with High Rate Sensors*. New York: McGraw-Hill, 2008.

[13] P. D. Groves, *Principles of GNSS, Inertial, and Multisensor Integrated Navigation Systems*, 2nd ed. Boston: Artech House, 2013.

[14] D. H. Titterton and J. L. Weston, *Strapdown Inertial Navigation Technology*, 2nd ed. IET, 2004.

[15] M. Kok, J. D. Hol, and T. B. Schön, "Using inertial sensors for position and orientation estimation," *Foundations and Trends in Signal Processing*, vol. 11, no. 1–2, pp. 1–153, 2017.

[16] A. I. Mourikis and S. I. Roumeliotis, "A multi-state constraint Kalman filter for vision-aided inertial navigation," in *IEEE ICRA*, 2007, pp. 3565–3572.

[17] T. Qin, P. Li, and S. Shen, "VINS-Mono: A robust and versatile monocular visual-inertial state estimator," *IEEE Trans. Robotics*, vol. 34, no. 4, pp. 1004–1020, 2018.

[18] A. Barrau and S. Bonnabel, "The invariant extended Kalman filter as a stable observer," *IEEE Trans. Automatic Control*, vol. 62, no. 4, pp. 1797–1812, 2017.

[19] M. Kaess, H. Johannsson, R. Roberts, V. Ila, J. J. Leonard, and F. Dellaert, "iSAM2: Incremental smoothing and mapping using the Bayes tree," *Int. J. Robotics Research*, vol. 31, no. 2, pp. 216–235, 2012.

[20] C. Forster, L. Carlone, F. Dellaert, and D. Scaramuzza, "On-manifold preintegration for real-time visual–inertial odometry," *IEEE Trans. Robotics*, vol. 33, no. 1, pp. 1–21, 2017.

[21] G. P. Huang, A. I. Mourikis, and S. I. Roumeliotis, "Observability-based rules for designing consistent EKF SLAM estimators," *Int. J. Robotics Research*, vol. 29, no. 5, pp. 502–528, 2010.

[22] J. A. Hesch, D. G. Kottas, S. L. Bowman, and S. I. Roumeliotis, "Consistency analysis and improvement of vision-aided inertial navigation," *IEEE Trans. Robotics*, vol. 30, no. 1, pp. 158–176, 2014.

[23] A. Martinelli, "Vision and IMU data fusion: Closed-form solutions for attitude, speed, absolute scale, and bias determination," *IEEE Trans. Robotics*, vol. 28, no. 1, pp. 44–60, 2012.

[24] S. J. Julier and J. K. Uhlmann, “Unscented filtering and nonlinear estimation,” *Proc. IEEE*, vol. 92, no. 3, pp. 401–422, 2004.

[25] M. W. Mueller, M. Hamer, and R. D’Andrea, “Fusing ultra-wideband range measurements with accelerometers and rate gyroscopes for quadrocopter state estimation,” in *IEEE ICRA*, 2015, pp. 1730–1736.

[26] D. Honegger, L. Meier, P. Tanskanen, and M. Pollefeys, “An open source and open hardware embedded metric optical flow CMOS camera for indoor and outdoor applications,” in *IEEE ICRA*, 2013, pp. 1736–1741.

[27] M. Burri et al., “The EuRoC micro aerial vehicle datasets,” *Int. J. Robotics Research*, vol. 35, no. 10, pp. 1157–1163, 2016.

[28] D. Schubert, T. Goll, N. Demmel, V. Usenko, J. Stückler, and D. Cremers, “The TUM VI benchmark for evaluating visual-inertial odometry,” in *IEEE/RSJ IROS*, 2018, pp. 1680–1687.

[29] J. Delmerico, T. Cieslewski, H. Rebecq, M. Faessler, and D. Scaramuzza, “Are we ready for autonomous drone racing? The UZH-FPV drone racing dataset,” in *IEEE ICRA*, 2019, pp. 6713–6719.

[30] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, “Vision meets robotics: The KITTI dataset,” *Int. J. Robotics Research*, vol. 32, no. 11, pp. 1231–1237, 2013.