Improved robustness of a robot pose filter using the direct product geometry on SE(3)

Jochen Trumpf * **Behzad Zamani** **

* *Australian National University (e-mail: Jochen.Trumpf@anu.edu.au)*
** *University of Melbourne (e-mail: Behzad.Zamani@unimelb.edu.au)*

Keywords: filtering theory, differential-geometric methods, non-linear systems theory

1. EXTENDED ABSTRACT

The current pose (position and orientation) of a mobile robot in 3D space $\mathbb{R}^3$ can be associated with an element of the Special Euclidean group SE(3), namely the rigid-body transformation that transforms space such that a chosen reference (space-fixed) coordinate frame is mapped to a chosen onboard (body-fixed) coordinate frame. The Special Euclidean group SE(3) is a semi-direct product of the Special Orthogonal group SO(3) with $\mathbb{R}^3$, accounting for the observation that spatial rotations and translations do not commute.

There is an extensive body of literature seeking to exploit this mathematical structure for the design of specialized nonlinear observers and filters that outperform general purpose nonlinear filters such as the Extended Kalman Filter (EKF), Unscented Kalman Filter (UKF), Minimum Energy Filter (MEF) or Particle Filter (PF) on the specific task of robot pose estimation. Among the notable approaches are; the Complementary Observer (CO), Baldwin et al. (2009); the Invariant Extended Kalman Filter (IEKF), Barrau and Bonnabel (2016), as a generalization of the Multiplicative Extended Kalman Filter (MEKF), Martin and Salaün (2010); the Equivariant Filter (EqF), van Goor et al. (2022); and the Geometric Approximate Minimum Energy filter (GAME), Saccon et al. (2016).

Any such filter that depends on second-order information (error covariance for stochastic filters, or the Hessian of the value function for deterministic filters) implicitly depends on a choice of local Riemannian metric on SE(3), even if this is often hidden by the formulation in specific coordinates and an implicit choice of the Euclidean metric in those coordinates. In the stochastic filtering literature, the most explicit articulation of this dependence can be found in the development of the EqF, specifically the reset step accounting for the nonlinear state space geometry in the update to the concentrated Gaussian distribution modeling the sensor noise. In the deterministic filtering literature this is more obvious since the Hessian of a function defined on a smooth manifold is not well-defined without reference to a metric structure.

While it seems natural to choose a left- or right-invariant Riemannian metric on SE(3) for this purpose, because it reflects the transformation of second-order information due to rigid-body motion either from a space-fixed or a body-fixed perspective, Ge et al. (2026), this is not the only relevant consideration in a nonlinear filtering context. The more salient question is which metric better approximates the full error information (probability distribution resp. value function) up to second order in-the-large. An alternative to invariant metrics are the *direct product* metrics on SE(3), endowing the group torsor of SE(3) with a direct product of an invariant metric on SO(3) and the Euclidean metric on $\mathbb{R}^3$.

In this presentation, we will show that — surprisingly — the standard direct product metric outperforms the standard left-invariant metric on SE(3) for a continuous-discrete GAME filter for robot pose from bearing measurements to *known* landmarks.[1] Specifically, we show through Monte Carlo simulations of a semi-realistic setting that performing the discrete measurement update step using the product metric retains acceptable performance up to an initial orientation estimation error of 7 degrees while the same filter using an invariant metric already deteriorates beyond usefulness at an initial orientation estimation error of 2-3 degrees. This is an important issue in practice where supplying a sufficiently accurate initial estimate requires complex and costly start-up procedures and/or launch equipment.

While the study presented in this work is quite limited, the general issue clearly deserves further exploration and research. In fact, the authors were caught completely off-guard by these results, where the initial plan had been to present recent new ideas on integrating the data fusion algorithm known as Covariance Intersection (CI) with a continuous-discrete GAME filter to enable filter augmentation based on bearing measurements to *unknown* landmarks, Zamani and Trumpf (2025). The conference presentation will still show some of these results, however, the authors would value the opportunity to discuss the completely unexpected strong metric dependence with other experts in the community.

1.1 *The simulation scenario*

We simulate a standard infrastructure inspection task carried out by a quadrotor drone in a 3D workspace of

[1] The standard left- and right-invariant metrics on SE(3) perform identically in this setting due to the specific invariance properties of the employed error cost function.

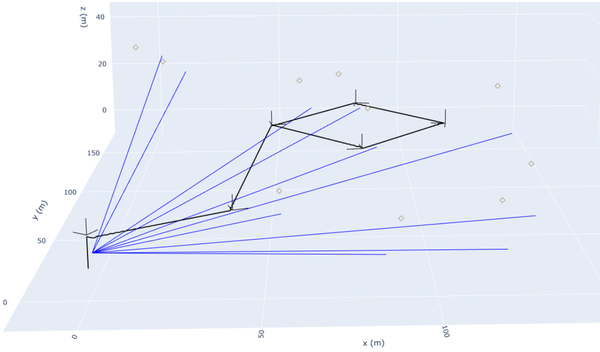

Fig. 1. Simulated scenario showing the ground truth robot trajectory and controller wayposes (in black), a randomly generated set of known landmarks (in brown) and simulated measurement rays to those landmarks (in blue).

size $150m \times 150m \times 60m$ (see Figure 1 on the next page). The drone (its trajectory shown in black in the figure) lifts off the ground from an initial position to the left, approaches the target zone at height 20m in two steps and then performs a square trajectory at level (for example, inspecting four sides of a building or other structure). The task takes approximately 140 seconds. Control of the drone is simulated with a standard waypose controller (wayposes are shown as black wireframes in the figure) that respects realistic constraints such as maximal velocity and acceleration constraints that would be typical for a robot of this type. We have deliberately decoupled the controller from the state estimator in this scenario in order not to obscure the differences between different filter implementations that are the actual target of this study.

Keeping with this philosophy, we have stripped out most of the complexity that would be found in an actual onboard state estimator for robot pose to condense the geometric essence of what is happening. In particular, we are not simulating actual Inertial Navigation System (INS) measurements but rather noisy measurements of the robot linear and angular velocities and noisy measurements of azimuth and elevation angles of bearings to randomly generated *known* landmarks in the environment (shown as brown diamonds in the figure, for example representing mapped features of surrounding structures).

The noise parameters are chosen to reflect the amount of noise that would be seen for these quantities in a more elaborate INS setup and are summed up in Table 1. The figure also shows an example of the measurement rays in blue as they would be perceived by the robot pose filter. Note how the rays emanate from a point that is not on the robot ground truth trajectory but where the filter estimates the robot to be at the time of measurement. The rays also reflect the orientation estimation error at the time of measurement as can be seen by the rays not actually pointing at the landmarks. The error shown in the figure corresponds to a heading estimation error of roughly 10 degrees.

The robot velocities are measured at 100Hz, a rate easily supported by modern INS systems, while the bearings to landmarks are measured only every 2 seconds, reflecting the complex computer vision pipeline needed to convert video from onboard camera systems to measured bearings. To further simplify the scenario we have assumed that all landmarks are visible at every measurement interval (this is the least realistic part of the setup).

While this is certainly a somewhat simplified and limited simulation scenario, it is still sufficiently complex to reflect the actual geometric challenges in the problem, such as the concentration of landmarks to one side of the robot through large parts of its initial trajectory, or the large distances to landmarks amplifying measurement errors in the filter. It is these practical challenges that have driven the development of specialized filters for this and related problems over the last 20 years.

Table 1. Standard deviations of initial estimates and measurement noises

Initial robot position	0.5m
Initial robot heading	1-7deg
Robot linear velocity	0.1m/s
Robot angular velocity	0.5deg/s
Azimuth and elevation angles	0.3deg

1.2 The system model and filtering algorithms

Let $X = \begin{bmatrix} R & p \\ 0_{1\times 3} & 1 \end{bmatrix}$ denote the pose matrix of a moving body in 3D space with $R \in \mathrm{SO}(3)$ denoting the attitude and $p \in \mathbb{R}^3$ the translation of the body-fixed frame with respect to the reference frame. The pose kinematics on SE(3) is then given by

$$\dot{X} = X \begin{bmatrix} \Omega \\ V \end{bmatrix}^{\wedge}, \; X(0) = X_0, \tag{1}$$

where the matrix $\begin{bmatrix} \Omega \\ V \end{bmatrix}^{\wedge} \in \mathfrak{se}(3)$ is the twist and embodies the angular velocity $\Omega \in \mathbb{R}^3$ and translational velocity $V \in \mathbb{R}^3$ of the body-fixed frame with respect to the reference frame.

The measured twist is comprised of angular and translational velocity measured in the body-fixed frame as represented by the following equations

$$\begin{aligned} &U = \begin{bmatrix} \Omega \\ V \end{bmatrix}^{\wedge} + (B\delta)^{\wedge}, \; B = \begin{bmatrix} B_\omega & 0_{3\times 3} \\ 0_{3\times 3} & B_v \end{bmatrix}, \; \delta = \begin{bmatrix} \delta_\omega \\ \delta_v \end{bmatrix}, \\ &U_\omega = \Omega + B_\omega \delta_\omega, \; U_v = V + B_v \delta_v, \end{aligned} \tag{2}$$

where $B_\omega, B_v \in \mathbb{R}^{3\times 3}$ are measurement error coefficient matrices determined from the sensor properties while $\delta_\omega, \delta_v \in \mathbb{R}^3$ are the unknown (here assumed zero mean Gaussian) angular and translational velocity errors incurred in the measurement process.

Denote by $\theta_i \in \mathcal{S}^1$ and $\psi_i \in \mathcal{S}^1$ the true azimuth and elevation bearing angles to a known landmark i, by $\theta_i^m \in \mathcal{S}^1$ and $\psi_i^m \in \mathcal{S}^1$ their measured versions and by $\delta_\theta, \delta_\psi \in \mathcal{S}^1$ their measurement errors. The errors are assumed to be drawn from normal distributions $\delta_\theta \sim \mathcal{N}(0, \sigma_\theta^2)$ and $\delta_\psi \sim \mathcal{N}(0, \sigma_\psi^2)$ with standard deviations $\sigma_\theta > 0$ and $\sigma_\psi > 0$, respectively.

$$\theta_i^m = \theta_i + \delta_\theta, \; \psi_i^m = \psi_i + \delta_\psi. \tag{3}$$

The following discrete update cost captures the uncertainty of the bearing measurement with respect to the robot pose X. Here, $\bar{p}_i$ denotes the position of landmark i in homogeneous reference frame coordinates.

$$
\begin{aligned}
J(X) &= \frac{1}{2}\|DX^{-1}\bar{p}_i\|^2, \\
D &\triangleq \begin{bmatrix} \dfrac{z_\theta^\perp (z_\theta^\perp)^\top}{\sigma_\theta} + \dfrac{z_\psi^\perp (z_\psi^\perp)^\top}{\sigma_\psi} & 0_{3\times 1} \\ 0_{1\times 3} & 0 \end{bmatrix}, \\
z_\theta^\perp &\triangleq \begin{bmatrix} -\sin(\theta_i^m) \\ \cos(\theta_i^m) \\ 0 \end{bmatrix}, \; z_\psi^\perp \triangleq \begin{bmatrix} -\cos(\theta_i^m)\sin(\psi_i^m) \\ -\sin(\theta_i^m)\sin(\psi_i^m) \\ \cos(\psi_i^m) \end{bmatrix}.
\end{aligned} \tag{4}
$$

This cost is as proposed in the target tracking literature, Lingren and Gong (1978), and captures the measurement error in the subspace perpendicular to the measured bearing.

Due to space constraints, we are unable to provide the full continuous discrete GAME filter derivation for this system and measurement model here. The reader is referred to our previous paper, Zamani and Trumpf (2019), for a related setup with full filter derivation. The resulting GAME filter equations are as follows.

$$
\begin{aligned}
&\dot{\hat{X}} = \hat{X}U, \quad \hat{X}(0) = \hat{X}_0, \\
&\dot{P} = -\bar{U}^\top P - P\bar{U} + BB^\top, \qquad (5) \\
&\bar{U} \triangleq \begin{bmatrix} (U_\omega)^\times & 0 \\ (U_v)^\times & (U_\omega)^\times \end{bmatrix}, \quad \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}^\times \triangleq \begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix},
\end{aligned}
$$

for the propagation step and

$$P^+ = (P^{-1} + H^{-1})^{-1}, \tag{6}$$

$$\hat{X}^+ = \hat{X}\exp_{\hat{X}}(-P^+ g), \tag{7}$$

for the discrete update step, where H is the Hessian and g is the gradient of J at $\hat{X}$ computed with respect to the chosen Riemannian metric. Note that also the exponential map exp depends on the choice of metric.

1.3 *Simulation results*

Figures 2 and 4 sum up the main result of this study. They show that the filter using the direct product metric is significantly more robust to initial heading error levels, both in terms of outlier performance and worst case performance. A fully randomized Monte Carlo simulation of 10,000 Monte Carlo runs was performed separately for each initial error setting, making these results statistically significant. A variation of the initial heading error distribution (rather than the position error distribution) was chosen since those errors are amplified via the long distance to landmarks when propagating to the position variables in the filter.

Figure 5 shows a typical inlier run where both filters converge despite a number of large initial update steps. A typical outlier run is shown in Figure 3, demonstrating the typical failure mode of the filter using the invariant metric. No such failure was observed for the filter using the direct product metric in any of the settings, however, the slow deterioration of the average heading estimation error observed in the top plot in Figure 4 suggests that this would eventually occur at sufficiently large initial heading errors. From a practical perspective, the filter using the invariant metric is already unusable at the quite moderate initial heading error level of 2-3 degrees.

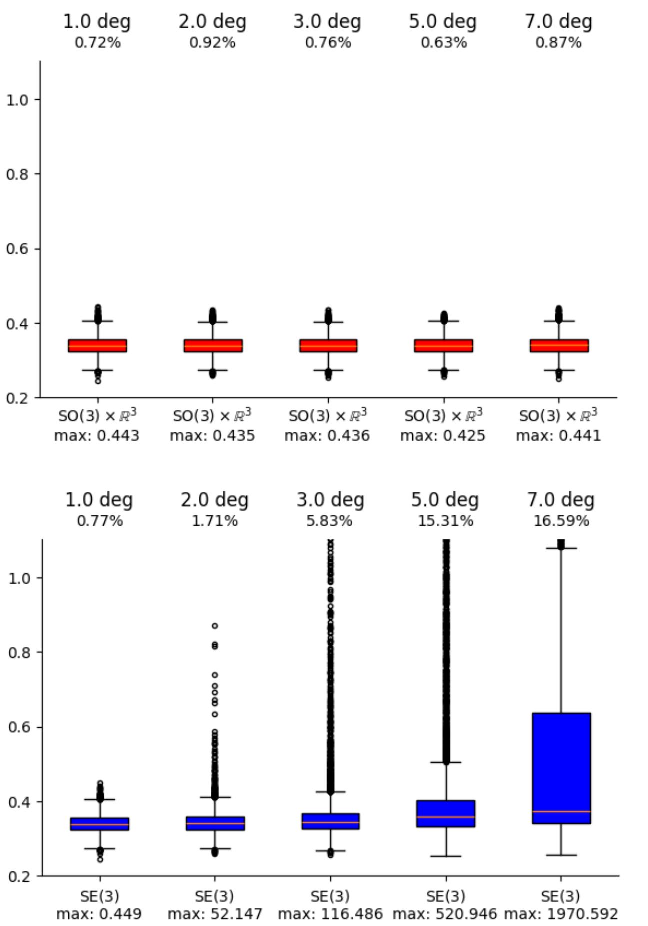

Fig. 2. Distribution of average robot position estimation error in metres along the entire trajectory for the filters using the direct product metric (above) or the invariant metric (below). The different levels of initial heading error are shown above the plots together with the corresponding outlier percentages (the outliers beyond 1.5 times Interquartile Range are represented by small circles in the plots). The maximal observed error is shown below the plots.

The specific nature of the failure mode of the filter, with the estimated orientation of the robot essentially flipping by 180 degrees, is due to the compact topology of SO(3) and the resulting presence of critical points of the error vector fields at antipodal points. This is a well-known issue in this branch of the control literature, Bhat and Bernstein (2000), and one of the reasons why this topic remains interesting and mathematically challenging.

REFERENCES

Baldwin, G., Mahony, R., and Trumpf, J. (2009). A nonlinear observer for 6 DOF pose estimation from inertial and bearing measurements. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2237–2242.

Barrau, A. and Bonnabel, S. (2016). The invariant Extended Kalman Filter as a stable observer. *IEEE Transactions on Automatic Control*, 62(4), 1797–1812.

Bhat, S. and Bernstein, D. (2000). A topological obstruction to continuous global stabilization of rotational motion and the unwinding phenomenon. *Systems&Control Letters*, 39, 63–70.

Ge, Y., Delama, G., Scheiber, M., Fornasier, A., van Goor, P., Weiss, S., and Mahony, R. (2026). The difference between the left and

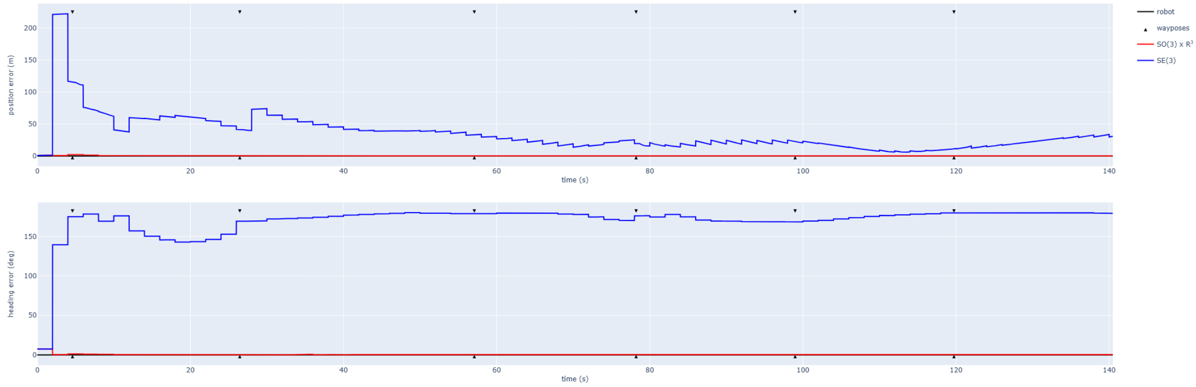

Fig. 3. Typical failure mode for the filter using the invariant metric (shown in blue). The heading estimation error (below) quickly escalates to almost 180deg (the theoretical maximum error) while the filter using the direct product metric (shown in red) stays stable in both position error (above) and heading error (below).

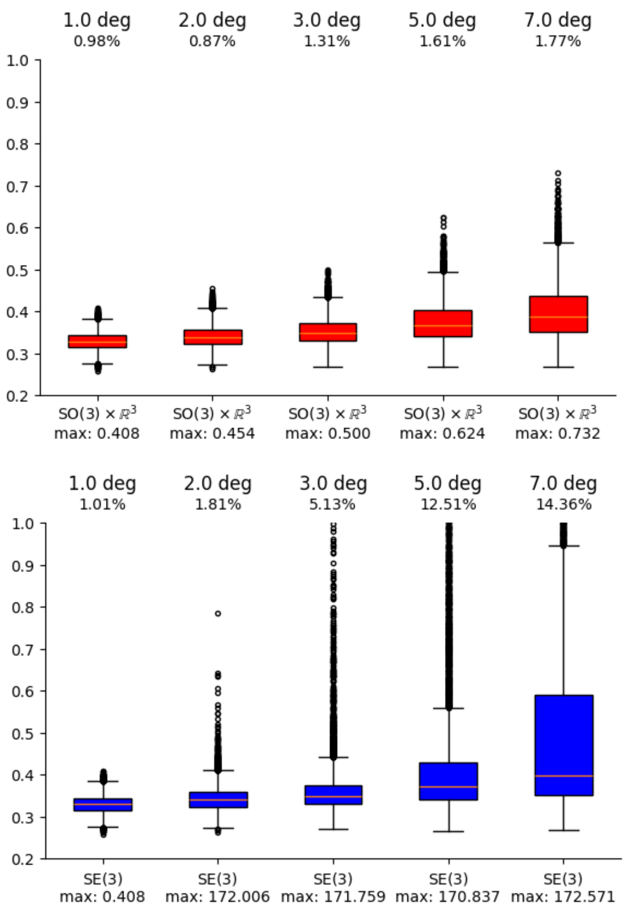

Fig. 4. Distribution of average robot heading estimation error in degrees along the entire trajectory for the filters using the direct product metric (above) or the invariant metric (below). The different levels of initial heading error are shown above the plots together with the corresponding outlier percentages (the outliers beyond 1.5 times Interquartile Range are represented by small circles in the plots). The maximal observed error is shown below the plots.

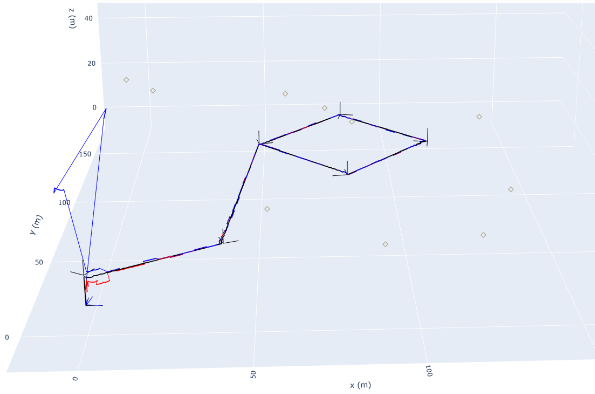

Fig. 5. A typical simulation run showing fast convergence of both filter trajectories following a small number of initial larger update steps. The filter using the invariant metric (shown in blue) typically performs much worse in the initial phase.

right invariant Extended Kalman Filter. *Control Engineering Practice*, 167, 106656.

Lingren, A.G. and Gong, K.F. (1978). Position and velocity estimation via bearing observations. *IEEE Transactions on Aerospace and Electronic systems*, AES-14(4), 564–577.

Martin, P. and Salaün, E. (2010). Generalized Multiplicative Extended Kalman Filter for aided attitude and heading reference system. In *AIAA Guidance, Navigation, and Control Conference*, 8300.

Saccon, A., Trumpf, J., Mahony, R., and Aguiar, A. (2016). Second-order-optimal minimum-energy filters on Lie groups. *IEEE Transactions on Automatic Control*, 61, 2906–2919.

van Goor, P., Hamel, T., and Mahony, R. (2022). Equivariant Filter (EqF). *IEEE Transactions on Automatic Control*, 68(6), 3501–3512.

Zamani, B. and Trumpf, J. (2025). A modular nonlinear continuous-discrete filter for concurrent 3D robot pose and landmark position estimation. In *Proceedings of the 29th International Conference on Methods and Models in Automation and Robotics (MMAR)*, 467–472.

Zamani, M. and Trumpf, J. (2019). Discrete update pose filter on the Special Euclidean group SE(3). In *Proceedings of the 58th IEEE Conference on Decision and Control (CDC)*, 635–641.