Unified EKF Notation with Combined 8-DOF and 9-DOF Models

Aalto Internship – Sensor Fusion

Common EKF Notation

We adopt the textbook notation for Bayesian filtering and the Extended Kalman Filter (EKF):

$$\begin{aligned}
\text{State: } & x_k \in \mathbb{R}^n, \quad \text{control/input: } u_{k-1}, \quad \text{measurement: } y_k, \\
\text{Process model: } & x_k = f(x_{k-1}, u_{k-1}) + q_{k-1}, \quad q_{k-1} \sim \mathcal{N}(0, Q_{k-1}), && (1) \\
\text{Measurement model: } & y_k = h(x_k) + r_k, \quad r_k \sim \mathcal{N}(0, R_k). && (2)
\end{aligned}$$

At each step we maintain the Gaussian estimate (m_k, P_k) of x_k. We use the aliases

$$m_k^- \equiv m_{k|k-1}, \qquad P_k^- \equiv P_{k|k-1}.$$

EKF recursion The process and measurement models are

$$x_k = f(x_{k-1}, u_{k-1}) + q_{k-1}, \qquad q_{k-1} \sim \mathcal{N}(0, Q_{k-1}), \tag{3}$$

$$y_k = h(x_k) + r_k, \qquad r_k \sim \mathcal{N}(0, R_k). \tag{4}$$

We maintain a Gaussian estimate (m_k, P_k) of x_k, and use the aliases

$$m_k^- \equiv m_{k|k-1}, \qquad P_k^- \equiv P_{k|k-1}.$$

Prediction.

$$m_k^- = f\big(m_{k-1}, u_{k-1}\big), \tag{5}$$

$$P_k^- = F_{k-1}\, P_{k-1}\, F_{k-1}^\top + Q_{k-1}, \qquad F_{k-1} := \left.\frac{\partial f}{\partial x}\right|_{x=m_{k-1},\, u=u_{k-1}}. \tag{6}$$

Update (single sensor). Define the noise-free predicted measurement $\hat{y}_k := h(m_k^-)$ and the Jacobian $H_k := \left.\dfrac{\partial h}{\partial x}\right|_{x=m_k^-}$. Then

$$v_k = y_k - \hat{y}_k \;=\; y_k - h(m_k^-), \tag{7}$$

$$S_k = H_k\, P_k^-\, H_k^\top + R_k, \tag{8}$$

$$K_k = P_k^-\, H_k^\top\, S_k^{-1}, \tag{9}$$

$$m_k = m_k^- + K_k\, v_k, \tag{10}$$

$$P_k = (I - K_k H_k)\, P_k^- \quad (\text{or Joseph: } (I - K_k H_k) P_k^- (I - K_k H_k)^\top + K_k R_k K_k^\top). \tag{11}$$

Measurement notation (per sensor). For a specific sensor s at time k we write

$$y_k^{(s)} = h^{(s)}(x_k) + r_k^{(s)}, \qquad r_k^{(s)} \sim \mathcal{N}(0, R_k^{(s)}),$$

with

$$\hat{y}_k^{(s)} := h^{(s)}(m_{k|k-1}), \qquad H_k^{(s)} := \frac{\partial h^{(s)}}{\partial x}\bigg|_{x=m_{k|k-1}}.$$

Explicit computation at time k (sequential or joint). Starting from $(m_{k|k-1}, P_{k|k-1})$:

- **Sequential updates** over sensors $s = 1, \dots, S$:

$$v_k^{(s)} = y_k^{(s)} - \hat{y}_k^{(s)}, \quad S_k^{(s)} = H_k^{(s)} P_{k|k-1} (H_k^{(s)})^\top + R_k^{(s)}, \quad K_k^{(s)} = P_{k|k-1} (H_k^{(s)})^\top \big(S_k^{(s)}\big)^{-1},$$

$$m_{k|k} \leftarrow m_{k|k-1} + K_k^{(s)} v_k^{(s)}, \qquad P_{k|k} \leftarrow (I - K_k^{(s)} H_k^{(s)}) P_{k|k-1} \quad \text{(or Joseph with } R_k^{(s)}\text{)}.$$

  Use the output $(m_{k|k}, P_{k|k})$ of sensor s as the input $(m_{k|k-1}, P_{k|k-1})$ for sensor $s+1$.

- **Joint (stacked) update** with

$$y_k^{\text{all}} = \begin{bmatrix} y_k^{(1)} \\ \vdots \\ y_k^{(S)} \end{bmatrix}, \quad \hat{y}_k^{\text{all}} = \begin{bmatrix} \hat{y}_k^{(1)} \\ \vdots \\ \hat{y}_k^{(S)} \end{bmatrix}, \quad H_k^{\text{all}} = \begin{bmatrix} H_k^{(1)} \\ \vdots \\ H_k^{(S)} \end{bmatrix}, \quad R_k^{\text{all}} = \operatorname{blkdiag}\big(R_k^{(1)}, \dots, R_k^{(S)}\big),$$

  and then apply (7)–(11) with the "all" quantities.

Multiple sensors at time k. Apply the above update *sequentially* over sensors $s = 1, \dots, S$ (order-independent for linear models), starting from $(m_{k|k-1}, P_{k|k-1})$ and ending at $(m_{k|k}, P_{k|k})$; or form a joint stacked update with

$$y_k^{\text{all}} = \begin{bmatrix} y_k^{(1)} \\ \vdots \\ y_k^{(S)} \end{bmatrix}, \quad H_k^{\text{all}} = \begin{bmatrix} H_k^{(1)} \\ \vdots \\ H_k^{(S)} \end{bmatrix}, \quad R_k^{\text{all}} = \operatorname{blkdiag}(R_k^{(1)}, \dots, R_k^{(S)}),$$

and use the same equations with "all".

The following two model families are unified under the above notation while preserving their full original equations: an 8-DOF planar ground robot model (RoboMaster S1) and a 9-DOF aerial model (drone). The 8-DOF content consolidates both formulations without loss; the 9-DOF section includes the full process/measurement expressions.

1 8-DOF Planar Ground Robot (RoboMaster S1)

This model targets 2D ground motion with IMU-driven prediction and mixed exteroceptive updates.

1.1 State and Inputs

$$\begin{aligned} x_k &= \begin{bmatrix} p_x & p_y & \theta & v_x & v_y & b_{a_x} & b_{a_y} & b_\omega \end{bmatrix}^\top \in \mathbb{R}^8, \\ u_{k-1} &= \begin{bmatrix} a_x^{\text{meas}} & a_y^{\text{meas}} & \omega^{\text{meas}} \end{bmatrix}^\top. \end{aligned} \tag{12}$$

Biases $b_{a_x}, b_{a_y}, b_\omega$ follow (random-walk) bias dynamics.

1.2 Process Model $f(\cdot)$ and Jacobian F_{k-1}

Define $R(\theta) = \left(\begin{smallmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{smallmatrix}\right)$ and the bias-compensated body acceleration

$$\begin{bmatrix} a'_x \\ a'_y \end{bmatrix} = \begin{bmatrix} a_x^{\text{meas}} - b_{a_x} \\ a_y^{\text{meas}} - b_{a_y} \end{bmatrix}.$$

Two equivalent forms of the same dynamics are retained:

Direct component form

$$f(x,u) = \begin{bmatrix} p_x + v_x\,\Delta t \\ p_y + v_y\,\Delta t \\ \theta + (\omega^{\text{meas}} - b_\omega)\,\Delta t \\ v_x + \big(a'_x\cos\theta - a'_y\sin\theta\big)\,\Delta t \\ v_y + \big(a'_x\sin\theta + a'_y\cos\theta\big)\,\Delta t \\ b_{a_x} \\ b_{a_y} \\ b_\omega \end{bmatrix}. \tag{13}$$

Rotation form Let $a^g = R(\theta)\,(a^b - b_a)$ denote world-frame specific force from body-frame IMU:

$$m_k^- = f(m_{k-1}, u_{k-1}) = \begin{bmatrix} x + v_x\Delta t \\ y + v_y\Delta t \\ \theta + (\omega_z^b - b_\omega)\Delta t \\ v_x + a_x^g\Delta t \\ v_y + a_y^g\Delta t \\ b_{a_x} \\ b_{a_y} \\ b_\omega \end{bmatrix}, \qquad a^g = R(\theta)\begin{bmatrix} a_x^b - b_{a_x} \\ a_y^b - b_{a_y} \end{bmatrix}. \tag{14}$$

The Jacobian $F = \frac{\partial f}{\partial x}$ evaluated at (m_{k-1}, u_{k-1}) is (explicit 8×8):

$$F = \begin{pmatrix} 1 & 0 & 0 & \Delta t & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & \Delta t & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & -\Delta t \\ 0 & 0 & \Delta t(-a'_x\sin\theta - a'_y\cos\theta) & 1 & 0 & -\Delta t\cos\theta & \Delta t\sin\theta & 0 \\ 0 & 0 & \Delta t(a'_x\cos\theta - a'_y\sin\theta) & 0 & 1 & -\Delta t\sin\theta & -\Delta t\cos\theta & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \tag{15}$$

Process noise

$$Q = \operatorname{diag}\big(\sigma_{p_x}^2, \sigma_{p_y}^2, \sigma_\theta^2, \sigma_{v_x}^2, \sigma_{v_y}^2, \sigma_{b_{a_x}}^2, \sigma_{b_{a_y}}^2, \sigma_{b_\omega}^2\big). \tag{16}$$

or (exact discretization using continuous PSDs $(q_a, q_\omega, q_{b_a}, q_{b_\omega})$):

$$Q_k = \operatorname{blkdiag}\left(\begin{bmatrix} \frac{\Delta t^4}{4} q_a\, I_2 & \frac{\Delta t^3}{2} q_a\, I_2 \\ \frac{\Delta t^3}{2} q_a\, I_2 & \Delta t^2 q_a\, I_2 \end{bmatrix},\ \Delta t\, q_\omega,\ \Delta t\, q_{b_a}\, I_2,\ \Delta t\, q_{b_\omega} \right). \tag{17}$$

1.3 Measurement Models $h(\cdot)$, $\hat{y}_k$, H_k, R

Wheel odometry (body-frame velocity)

$$y_k^{\text{odom}} = h_{\text{odom}}(x_k) + r_k^{\text{odom}} = R(\theta_k)^\top \begin{bmatrix} v_x \\ v_y \end{bmatrix}_k + r_k^{\text{odom}}, \tag{18}$$

$$\hat{y}_k^{\text{odom}} = R(\theta_{k|k-1})^\top \begin{bmatrix} v_x \\ v_y \end{bmatrix}_{k|k-1}, \tag{19}$$

$$H_k^{\text{odom}} = \begin{bmatrix} 0 & 0 & -v_x \sin\theta + v_y \cos\theta & \cos\theta & \sin\theta & 0 & 0 & 0 \\ 0 & 0 & -v_x \cos\theta - v_y \sin\theta & -\sin\theta & \cos\theta & 0 & 0 & 0 \end{bmatrix} \Bigg|_{x=m_{k|k-1}}, \tag{20}$$

$$R^{\text{odom}} = \operatorname{diag}(\sigma^2_{v_{b,x}},\, \sigma^2_{v_{b,y}}). \tag{21}$$

Magnetometer (heading)

$$y_k^{\text{mag}} = h_{\text{mag}}(x_k) + r_k^{\text{mag}} = \theta_k + r_k^{\text{mag}}, \tag{22}$$

$$\hat{y}_k^{\text{mag}} = \theta_{k|k-1}, \qquad H_k^{\text{mag}} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \tag{23}$$

$$R^{\text{mag}} = \sigma_\psi^2. \tag{24}$$

GPS position (2D)

$$y_k^{\text{pos}} = \begin{bmatrix} p_x \\ p_y \end{bmatrix} + r_k^{\text{pos}}, \tag{25}$$

$$\hat{y}_k^{\text{pos}} = \begin{bmatrix} p_x \\ p_y \end{bmatrix}_{k|k-1}, \qquad H_{\text{pos}} = \begin{bmatrix} I_2 & 0_{2\times 6} \end{bmatrix}, \tag{26}$$

$$R^{\text{pos}} = \operatorname{diag}(\sigma^2_{\text{GPS},x},\, \sigma^2_{\text{GPS},y}). \tag{27}$$

GPS velocity (2D)

$$y_k^{\text{vel}} = \begin{bmatrix} v_x \\ v_y \end{bmatrix} + r_k^{\text{vel}}, \tag{28}$$

$$\hat{y}_k^{\text{vel}} = \begin{bmatrix} v_x \\ v_y \end{bmatrix}_{k|k-1}, \qquad H_{\text{vel}} = \begin{bmatrix} 0_{2\times 3} & I_2 & 0_{2\times 3} \end{bmatrix}, \tag{29}$$

$$R^{\text{vel}} = \operatorname{diag}(\sigma^2_{\text{GPS},v_x},\, \sigma^2_{\text{GPS},v_y}). \tag{30}$$

Optional nonholonomic constraint (NHC)

$$y_k^{\text{nhc}} = 0 + r_k^{\text{nhc}}, \quad \hat{y}_k^{\text{nhc}} = \begin{bmatrix} -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} v_x \\ v_y \end{bmatrix}_{k|k-1}, \tag{31}$$

$$H_{\text{nhc}} = \begin{bmatrix} 0 & 0 & -(\cos\theta)\, v_x - (\sin\theta)\, v_y & -\sin\theta & \cos\theta & 0 & 0 & 0 \end{bmatrix} \Bigg|_{x=m_{k|k-1}}, \tag{32}$$

$$R^{\text{nhc}} = \sigma^2_{\text{NHC}}. \tag{33}$$

Optional zero-velocity update (ZUPT)

$$y_k^{\text{zupt}} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + r_k^{\text{zupt}}, \tag{34}$$

$$\hat{y}_k^{\text{zupt}} = \begin{bmatrix} v_x \\ v_y \end{bmatrix}_{k|k-1}, \qquad H_{\text{zupt}} = \begin{bmatrix} 0_{2\times3} & I_2 & 0_{2\times3} \end{bmatrix}, \tag{35}$$

$$R^{\text{zupt}} = \sigma_{\text{ZUPT}}^2 I_2. \tag{36}$$

1.4 EKF Steps (Using Common Notation)

Apply (5)–(11) with f in (13) or (14), F in (15), and the relevant $(\hat{y}, H, R)$ from (20)–(36).

2 9-DOF Aerial Vehicle (Drone)

This model estimates 3D position, velocity, and attitude (Euler ZYX), driven by IMU prediction and GPS/barometer updates.

2.1 State, Inputs, and Process Model

$$x_k = \begin{bmatrix} p_x & p_y & p_z & v_x & v_y & v_z & \phi & \theta & \psi \end{bmatrix}^\top \in \mathbb{R}^9, \tag{37}$$

$$u_{k-1} = \begin{bmatrix} a_{k-1} \\ \omega_{k-1} \end{bmatrix}, \qquad g = \begin{bmatrix} 0 \\ 0 \\ -9.81 \end{bmatrix}.$$

With Δt the sampling interval,

$$f(x,u) = \begin{bmatrix} p + v\,\Delta t \\ v + \big(R(\theta)\,a + g\big)\,\Delta t \\ \theta + J(\theta)\,\omega\,\Delta t \end{bmatrix}. \tag{38}$$

The ZYX (yaw–pitch–roll) body-to-world rotation and Euler-rate mapping are:

$$R(\phi,\theta,\psi) = \begin{pmatrix} c_\psi c_\theta & c_\psi s_\theta s_\phi - s_\psi c_\phi & c_\psi s_\theta c_\phi + s_\psi s_\phi \\ s_\psi c_\theta & s_\psi s_\theta s_\phi + c_\psi c_\phi & s_\psi s_\theta c_\phi - c_\psi s_\phi \\ -s_\theta & c_\theta s_\phi & c_\theta c_\phi \end{pmatrix}, \tag{39}$$

$$J(\phi,\theta) = \begin{pmatrix} 1 & s_\phi \dfrac{s_\theta}{c_\theta} & c_\phi \dfrac{s_\theta}{c_\theta} \\ 0 & c_\phi & -s_\phi \\ 0 & s_\phi / c_\theta & c_\phi / c_\theta \end{pmatrix}. \tag{40}$$

Process noise

$$Q = \operatorname{diag}\!\big(\sigma_p^2 I_3,\ \sigma_v^2 I_3,\ \sigma_\theta^2 I_3\big). \tag{41}$$

2.2 Measurement Models ($\hat{y}_k$, H_k, R)

GPS position (3D)

$$y_k^G = h_{\text{GPS}}(x_k) + r_k^G = \begin{bmatrix} p_x \\ p_y \\ p_z \end{bmatrix} + r_k^G, \tag{42}$$

$$\hat{y}_k^G = \begin{bmatrix} p_x \\ p_y \\ p_z \end{bmatrix}_{k|k-1}, \qquad H_G = \begin{bmatrix} I_3 & 0_{3\times 6} \end{bmatrix}, \tag{43}$$

$$R^G = \operatorname{diag}(\sigma^2_{\text{GPS},x}, \sigma^2_{\text{GPS},y}, \sigma^2_{\text{GPS},z}). \tag{44}$$

Barometer altitude

$$y_k^B = h_{\text{baro}}(x_k) + r_k^B = p_z + r_k^B, \tag{45}$$

$$\hat{y}_k^B = p_{z,\,k|k-1}, \qquad H_B = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \tag{46}$$

$$R^B = \sigma^2_{\text{baro}}. \tag{47}$$

Magnetometer (heading)

$$y_k^M = h_{\text{mag}}(x_k) + r_k^M = \psi_k + r_k^M, \tag{48}$$

$$\hat{y}_k^M = \psi_{k|k-1}, \qquad H_M = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \tag{49}$$

$$R^M = \sigma^2_\psi. \tag{50}$$

2.3 EKF Jacobian and Recursion (Fully Expanded)

The process Jacobian at (m_{k-1}, u_{k-1}) is

$$F_{k-1} = \begin{pmatrix} I_3 & \Delta t\, I_3 & 0_{3\times 3} \\ 0_{3\times 3} & I_3 & \dfrac{\partial\big(R(\phi,\theta,\psi)a\big)}{\partial[\phi,\theta,\psi]} \\ 0_{3\times 3} & 0_{3\times 3} & I_3 + \Delta t\,\dfrac{\partial\big(J(\phi,\theta)\omega\big)}{\partial[\phi,\theta]} \end{pmatrix}_{\theta=\theta_{k-1},\, a=a_{k-1},\, \omega=\omega_{k-1}}. \tag{51}$$

Explicit partials for $R(\phi,\theta,\psi)a$ Let $a = \begin{bmatrix} a_x & a_y & a_z \end{bmatrix}^\top$. Then

$$\frac{\partial(Ra)}{\partial[\phi,\theta,\psi]} = \Big[\, (\partial_\phi R)\, a \quad (\partial_\theta R)\, a \quad (\partial_\psi R)\, a \,\Big],$$

where (with $c\cdot = \cos(\cdot)$, $s\cdot = \sin(\cdot)$):

$$\partial_\phi R = \begin{pmatrix} 0 & c_\psi s_\theta c_\phi + s_\psi s_\phi & -c_\psi s_\theta s_\phi + s_\psi c_\phi \\ 0 & s_\psi s_\theta c_\phi - c_\psi s_\phi & -s_\psi s_\theta s_\phi - c_\psi c_\phi \\ 0 & c_\theta c_\phi & -c_\theta s_\phi \end{pmatrix},$$

$$\partial_\theta R = \begin{pmatrix} -c_\psi s_\theta & c_\psi c_\theta s_\phi & c_\psi c_\theta c_\phi \\ -s_\psi s_\theta & s_\psi c_\theta s_\phi & s_\psi c_\theta c_\phi \\ -c_\theta & -s_\theta s_\phi & -s_\theta c_\phi \end{pmatrix},$$

$$\partial_\psi R = \begin{pmatrix} -s_\psi c_\theta & -s_\psi s_\theta s_\phi - c_\psi c_\phi & -s_\psi s_\theta c_\phi + c_\psi s_\phi \\ c_\psi c_\theta & c_\psi s_\theta s_\phi - s_\psi c_\phi & c_\psi s_\theta c_\phi - s_\psi s_\phi \\ 0 & 0 & 0 \end{pmatrix}.$$

Explicit partials for $J(\phi,\theta)\omega$ Let $\omega = \begin{bmatrix}\omega_x & \omega_y & \omega_z\end{bmatrix}^\top$. Then

$$\frac{\partial(J\omega)}{\partial[\phi,\theta]} = \Big[\,(\partial_\phi J)\,\omega \quad (\partial_\theta J)\,\omega\,\Big],$$

with

$$\partial_\phi J = \begin{pmatrix} 0 & c_\phi \dfrac{s_\theta}{c_\theta} & -\,s_\phi \dfrac{s_\theta}{c_\theta} \\ 0 & -\,s_\phi & -\,c_\phi \\ 0 & \dfrac{c_\phi}{c_\theta} & -\,\dfrac{s_\phi}{c_\theta} \end{pmatrix},$$

$$\partial_\theta J = \begin{pmatrix} 0 & \dfrac{s_\phi}{c_\theta^2} & \dfrac{c_\phi}{c_\theta^2} \\ 0 & 0 & 0 \\ 0 & s_\phi \dfrac{s_\theta}{c_\theta^2} & c_\phi \dfrac{s_\theta}{c_\theta^2} \end{pmatrix}.$$

EKF recursion Prediction (5)–(6) uses (38), (51), (41). Updates (7)–(11) use $(\hat{y}, H, R)$ from (43)–(50) as appropriate.

Angle wrapping. After any update that touches angles (θ, ϕ, ψ), wrap them to $(-\pi, \pi]$.

Appendix: Matrix Properties and Joseph Form

Matrix properties (assumptions and consequences)

Let $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$. At time k:

- $Q \in \mathbb{R}^{n\times n}$: process-noise covariance (covariance of q_{k-1} in (3)), symmetric PSD ($Q \succeq 0$).
- $P^- \in \mathbb{R}^{n\times n}$: prior covariance, symmetric PSD. If $P_{k-1} \succeq 0$ and $Q \succeq 0$, then with $P^- = FP_{k-1}F^\top + Q$ we have $P^- \succeq 0$ and P^- is symmetric.
- $H \in \mathbb{R}^{m\times n}$: measurement Jacobian (linearized h); m is measurement dimension.
- $R \in \mathbb{R}^{m\times m}$: measurement-noise covariance, symmetric PD ($R \succ 0$).
- $S \in \mathbb{R}^{m\times m}$: innovation covariance, $S = HP^-H^\top + R$. Then $S = S^\top$ and $S \succ 0$ (since $R \succ 0$), so S^{-1} exists and $(S^{-1})^\top = S^{-1}$.
- $K \in \mathbb{R}^{n\times m}$: Kalman gain, $K = P^-H^\top S^{-1}$.
- Posterior covariance P: symmetric in exact arithmetic; Joseph update preserves symmetry/PSD under round-off.

Two handy equalities

$$\begin{aligned} &\text{(i)} \quad S\,(S^{-1})^\top = I \quad \text{since } S = S^\top \;\Rightarrow\; (S^{-1})^\top = S^{-1}. \\ &\text{(ii)} \quad S\,K^\top \;=\; H\,P^- \quad \text{because } K^\top = (S^{-1})^\top HP^- = S^{-1}HP^-. \end{aligned}$$

Joseph form from the general covariance update

Start from the "general" form

$$P \;=\; P^- - K\,S\,K^\top, \qquad S = HP^-H^\top + R, \qquad K = P^-H^\top S^{-1}.$$

Consider the Joseph (stabilized) form

$$P \;=\; (I - KH)\,P^-\,(I - KH)^\top \;+\; KRK^\top.$$

Expand the RHS:

$$\begin{aligned}
(I-KH)P^-(I-KH)^\top + KRK^\top &= P^- - KHP^- - P^-H^\top K^\top + KHP^-H^\top K^\top + KRK^\top \\
&= P^- - KHP^- - P^-H^\top K^\top + K(HP^-H^\top + R)K^\top \\
&= P^- - KHP^- - P^-H^\top K^\top + KSK^\top.
\end{aligned}$$

Now use $KHP^- = P^-H^\top S^{-1}HP^-$ and its transpose $P^-H^\top K^\top = P^-H^\top S^{-1}HP^-$. Hence

$$(I-KH)P^-(I-KH)^\top + KRK^\top = P^- - 2\,P^-H^\top S^{-1}HP^- + KSK^\top.$$

But $KSK^\top = P^-H^\top S^{-1}SS^{-1}HP^- = P^-H^\top S^{-1}HP^-$, so the middle terms cancel to yield

$$(I-KH)P^-(I-KH)^\top + KRK^\top \;=\; P^- - KSK^\top,$$

which proves the Joseph form is algebraically equivalent to the general form.

Notes on Tuning and Initialization

For the 8-DOF model, either diagonal Q (Eq. (16)) or exact-discretized Q_k (Eq. (17)) may be used; R values reflect sensor trust (odom/mag or GPS/NHC/ZUPT) and are given in the measurement models above. For the 9-DOF model, Q and R are diagonal as stated (Eqs. (41), (44)–(50)); GPS and barometer arrive at lower rates than IMU and are incorporated on arrival (magnetometer yaw can be fused similarly). Asynchronous, IMU-rate prediction with opportunistic updates is assumed in both.