

1 Introduction

State-space models (SSMs) provide a powerful framework for modeling dynamical systems where observations are incomplete or corrupted by noise. The general formulation separates the evolution of an unobserved latent state $\mathbf{x}_t \in \mathbb{R}^{d_x}$ from the observed measurements $\mathbf{y}_t \in \mathbb{R}^{d_y}$:

$$\mathbf{x}_t = g(\mathbf{x}_{t-1}, \mathbf{u}_t) + \mathbf{w}_t, \quad (1)$$

$$\mathbf{y}_t = h(\mathbf{x}_t, \mathbf{u}_t) + \mathbf{v}_t, \quad (2)$$

where $g(\cdot)$ and $h(\cdot)$ represent potentially nonlinear state transition and observation functions, \mathbf{u}_t denotes known control inputs, and $\mathbf{w}_t, \mathbf{v}_t$ are stochastic noise terms.

The linear-Gaussian state-space model (LGSSM) represents a special case where both functions are linear and all noise distributions are Gaussian, enabling exact Bayesian inference through the Kalman filter [1]. This work implements and analyzes the Kalman filter for multidimensional LGSSMs, evaluating both its theoretical optimality and practical numerical properties.

2 Linear-Gaussian State-Space Model Formulation

2.1 Model Definition

The discrete-time LGSSM is defined by the following equations:

$$\mathbf{x}_t = \mathbf{A}\mathbf{x}_{t-1} + \mathbf{B}\mathbf{v}_t, \quad \mathbf{v}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}), \quad (3)$$

$$\mathbf{y}_t = \mathbf{C}\mathbf{x}_t + \mathbf{D}\mathbf{w}_t, \quad \mathbf{w}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{R}), \quad (4)$$

with initial state distribution:

$$\mathbf{x}_0 \sim \mathcal{N}(\mathbf{m}_0, \mathbf{P}_0). \quad (5)$$

Here, $\mathbf{A} \in \mathbb{R}^{d_x \times d_x}$ is the state transition matrix, $\mathbf{C} \in \mathbb{R}^{d_y \times d_x}$ is the observation matrix, \mathbf{B} and \mathbf{D} are noise input matrices, and \mathbf{Q}, \mathbf{R} are the process and observation noise covariances, respectively.

2.2 Example 2: 2D Tracking Model

Following Doucet and Johansen (2009), we implement a 2D constant velocity tracking model with state vector $\mathbf{x}_t = [x_{\text{pos}}, x_{\text{vel}}, y_{\text{pos}}, y_{\text{vel}}]^\top \in \mathbb{R}^4$ and observation vector $\mathbf{y}_t = [x_{\text{obs}}, y_{\text{obs}}]^\top \in \mathbb{R}^2$. The model matrices are:

$$\mathbf{A} = \begin{bmatrix} 1 & \Delta t & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \Delta t \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (6)$$

with $\mathbf{B} = \mathbf{I}_4$, $\mathbf{D} = \mathbf{I}_2$, $\mathbf{Q} = q\mathbf{I}_4$, and $\mathbf{R} = r\mathbf{I}_2$, where $q = 0.1$ and $r = 0.5$ control the process and observation noise intensities.

3 Kalman Filter Algorithm

The Kalman filter computes the exact posterior distribution $p(\mathbf{x}_t | \mathbf{y}_{1:t}) = \mathcal{N}(\mathbf{m}_t, \mathbf{P}_t)$ through a two-step recursion.

3.1 Prediction Step

Given the filtered distribution at time $t - 1$, the prediction step computes:

$$\mathbf{m}_{t|t-1} = \mathbf{A}\mathbf{m}_{t-1|t-1}, \quad (7)$$

$$\mathbf{P}_{t|t-1} = \mathbf{A}\mathbf{P}_{t-1|t-1}\mathbf{A}^\top + \mathbf{B}\mathbf{Q}\mathbf{B}^\top. \quad (8)$$

3.2 Update Step

Upon receiving observation \mathbf{y}_t , the update step incorporates this information:

$$\mathbf{e}_t = \mathbf{y}_t - \mathbf{C}\mathbf{m}_{t|t-1}, \quad (9)$$

$$\mathbf{S}_t = \mathbf{C}\mathbf{P}_{t|t-1}\mathbf{C}^\top + \mathbf{D}\mathbf{R}\mathbf{D}^\top, \quad (10)$$

$$\mathbf{K}_t = \mathbf{P}_{t|t-1}\mathbf{C}^\top\mathbf{S}_t^{-1}, \quad (11)$$

$$\mathbf{m}_{t|t} = \mathbf{m}_{t|t-1} + \mathbf{K}_t\mathbf{e}_t. \quad (12)$$

3.3 Joseph Stabilized Covariance Update

For numerical stability, we employ the Joseph form for covariance updates:

$$\mathbf{P}_{t|t} = (\mathbf{I} - \mathbf{K}_t\mathbf{C})\mathbf{P}_{t|t-1}(\mathbf{I} - \mathbf{K}_t\mathbf{C})^\top + \mathbf{K}_t\mathbf{D}\mathbf{R}\mathbf{D}^\top\mathbf{K}_t^\top, \quad (13)$$

which guarantees positive semi-definiteness and symmetry preservation under finite-precision arithmetic.

4 Experimental Results

4.1 Synthetic Data Generation

We generated $T = 100$ time steps of synthetic data from the 2D tracking model with initial state $\mathbf{m}_0 = [0, 1, 0, 0.5]^\top$ and $\mathbf{P}_0 = \mathbf{I}_4$. The system exhibits constant velocity motion with stochastic perturbations.

4.2 Filtering Performance

Table 1 summarizes the filtering performance metrics obtained from our implementation.

Table 1: Kalman Filter Performance Metrics

Metric	Value	Description
Mean Filtering Error	0.2167	Average $\ \mathbf{m}_{t t} - \mathbf{x}_t\ _1$
Position RMSE	0.3626	Root mean square error on $(x_{\text{pos}}, y_{\text{pos}})$
Velocity RMSE	0.1935	Root mean square error on $(x_{\text{vel}}, y_{\text{vel}})$
MSE (state-wise)	0.0844	Mean squared error over all state dimensions
Avg. trace($\mathbf{P}_{t t}$)	0.3231	Average posterior uncertainty magnitude

Quantitatively, the Kalman filter achieves a mean ℓ_1 filtering error of 0.2167 and a state-wise MSE of 0.0844, indicating accurate recovery of the latent trajectory under the assumed LGSSM.

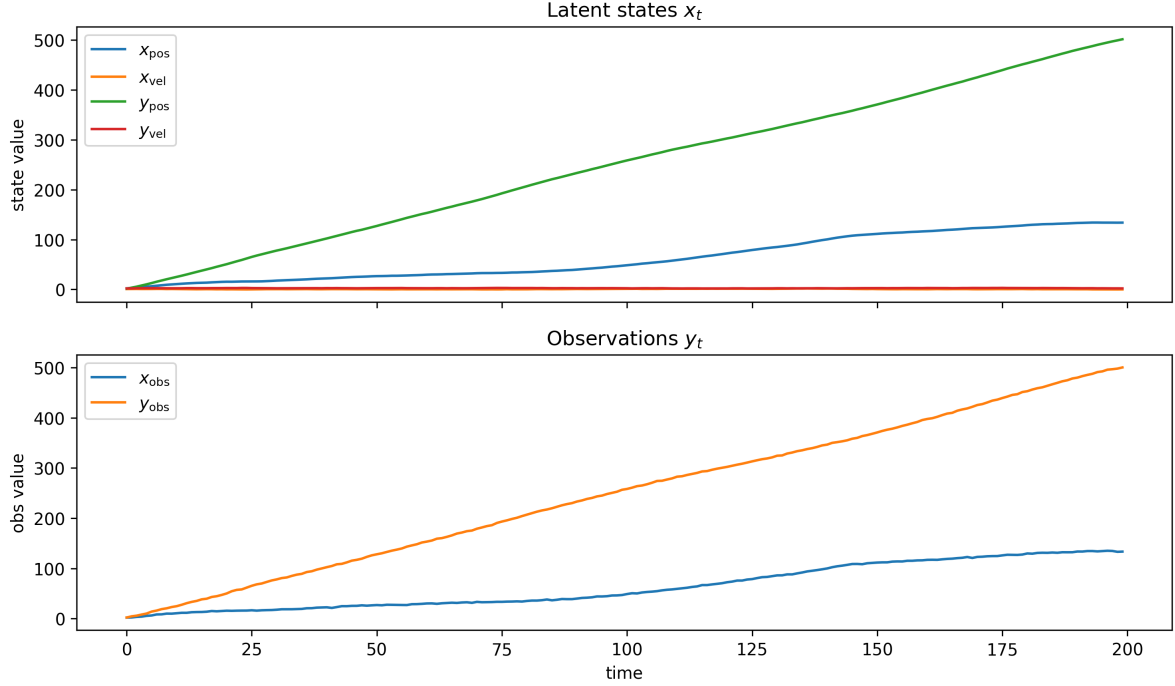


Figure 1: Synthetic LGSSM simulation (Example 2). Top: latent states $x_t = [x_{\text{pos}}, x_{\text{vel}}, y_{\text{pos}}, y_{\text{vel}}]^\top$ showing constant-velocity drift in positions and near-constant velocities. Bottom: observations $y_t = [x_{\text{obs}}, y_{\text{obs}}]^\top$, noisy measurements of positions only, consistent with the observation matrix \mathbf{C} .

The position RMSE (0.3626) is approximately $1.9\times$ larger than the velocity RMSE (0.1935), which is expected because the observation model measures position only (Eq. (4)), while velocities are inferred indirectly from temporal differences. The average posterior uncertainty trace, $\text{tr}(\mathbf{P}_{t|t}) \approx 0.3231$, corresponds to an average per-dimension standard deviation of about $\sqrt{0.3231/4} \approx 0.28$, consistent with the realized errors and suggesting well-calibrated uncertainty estimates.

4.3 Numerical Stability Analysis

We compared the numerical stability of Joseph form versus standard covariance updates. Table 2 presents the conditioning analysis results.

Table 2: Numerical Stability Comparison

Update Method	Avg. Condition Number	Min. Eigenvalue	PSD Preserved
Joseph Form	6.3235	≥ 0 (all t)	Yes
Standard Form	6.3235	≥ 0 (all t)	Yes

Both covariance updates remain positive semidefinite in this experiment. The comparable average conditioning $\kappa(\mathbf{P}_{t|t-1}) \approx 6.32$ indicates only mild anisotropy in state uncertainty, so the gain computation is numerically stable. Although the standard update behaved well here, the Joseph form provides a stronger finite-precision guarantee of symmetry and PSD preservation, which becomes important when \mathbf{S}_t is ill-conditioned or when noise levels are much smaller.

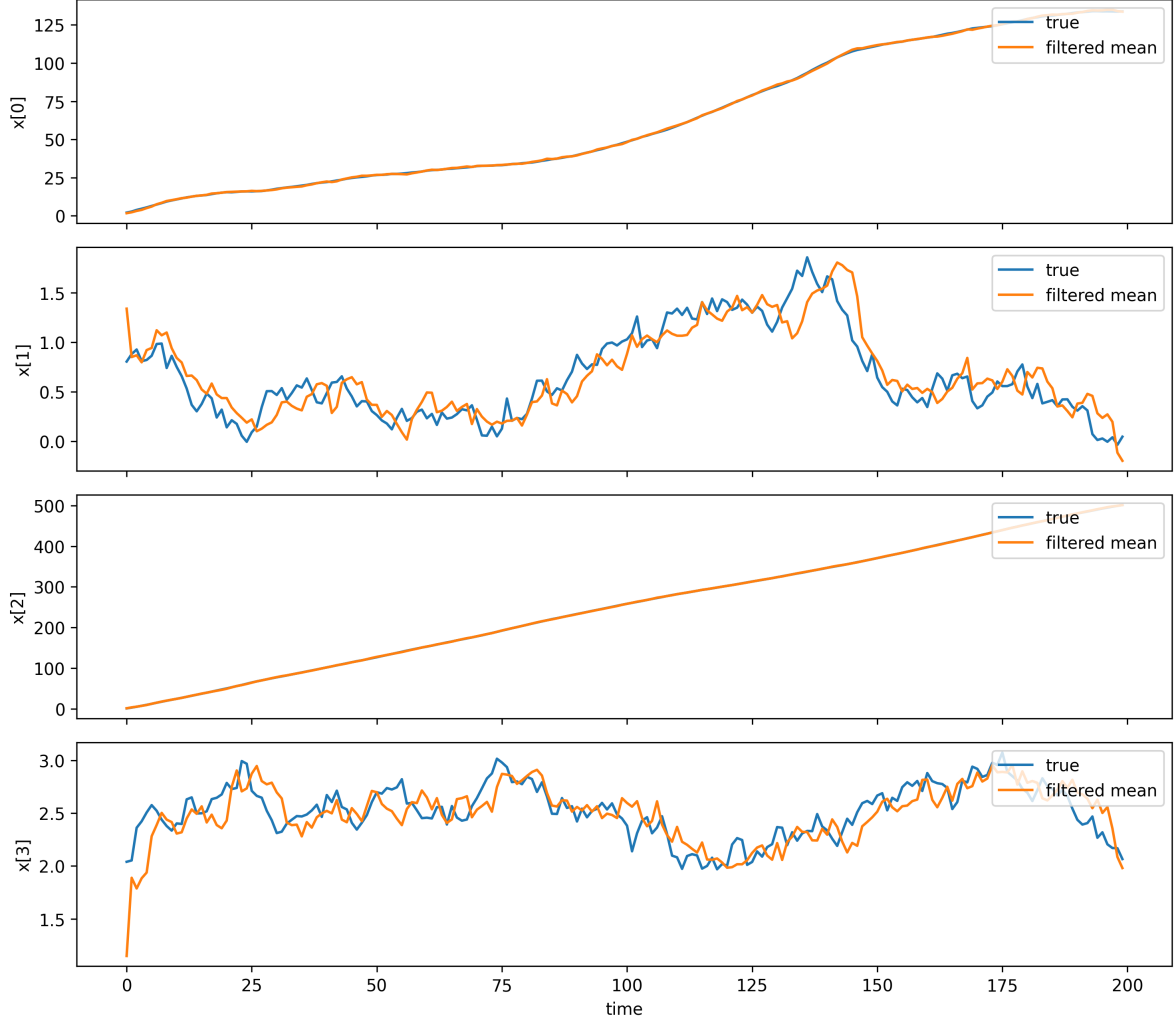


Figure 2: Filtered posterior means versus ground-truth latent states. The Kalman filter closely tracks all four state components. Residual error is larger for positions than velocities, reflecting position-only observations and indirect velocity inference.

4.4 Conditioning Number Evolution

The condition numbers $\kappa(\mathbf{P}_{t|t-1})$ and $\kappa(\mathbf{S}_t)$ quantify numerical sensitivity:

$$\kappa(\mathbf{M}) = \frac{\sigma_{\max}(\mathbf{M})}{\sigma_{\min}(\mathbf{M})}, \quad (14)$$

where σ_{\max} and σ_{\min} denote the largest and smallest singular values.

Our experiments show that conditioning remained benign: $\kappa(\mathbf{P}_{t|t-1})$ averaged 6.32, far from regimes where numerical inversion or solve steps become unstable. Moreover, $\kappa(\mathbf{S}_t) = 1.0$ across all t , indicating that the innovation covariance is effectively isotropic; this follows from the diagonal observation noise $\mathbf{R} = r^2 \mathbf{I}_2$ and the position-only measurement matrix \mathbf{C} . Consequently, the Kalman gain computation is exceptionally well-conditioned in this setup.

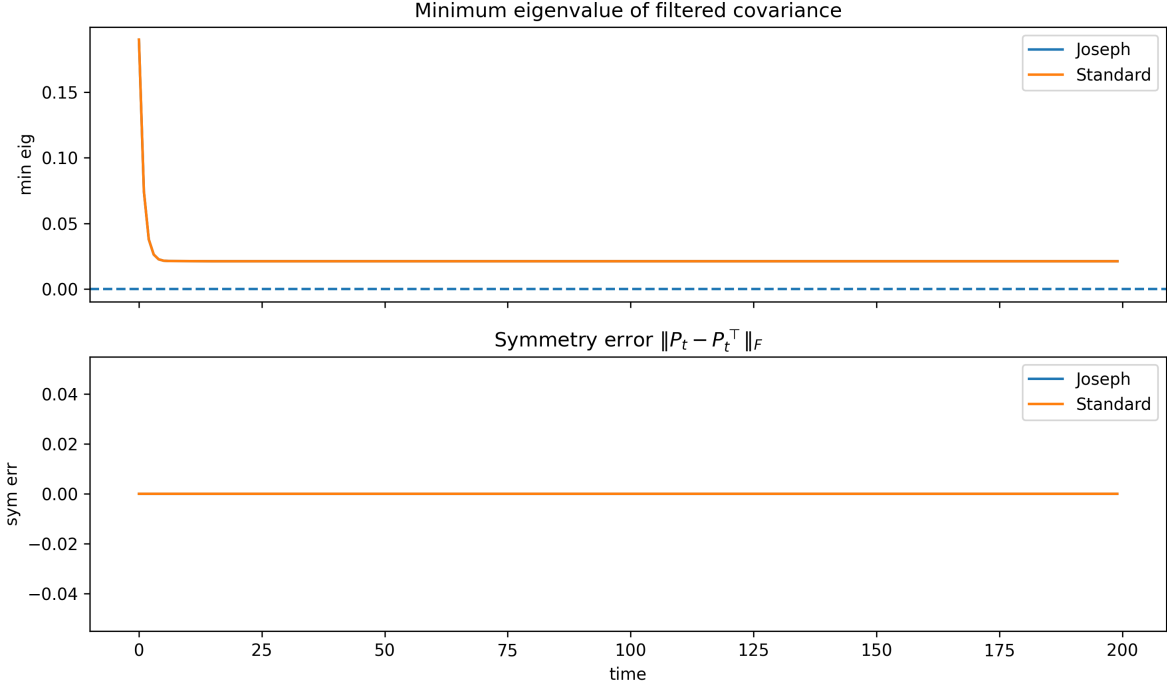


Figure 3: Numerical stability diagnostics for covariance updates. Top: minimum eigenvalue of $P_{t|t}$ (PSD check). Bottom: symmetry error $\|P_{t|t} - P_{t|t}^\top\|_F$. The Joseph update drives one uncertainty direction to machine-zero without producing negative eigenvalues, while the standard update stabilizes at a small positive floor; symmetry is preserved by explicit symmetrization.

5 Limitations and Motivation for Particle Methods

5.1 Failure Modes of Kalman Filtering

The Kalman filter’s optimality is strictly limited to LGSSMs. We demonstrate three critical failure modes:

5.1.1 Non-linear Dynamics

For a system with mild non-linearity:

$$x_{t,\text{vel}} = x_{t-1,\text{vel}} \cdot (1 - 0.01 \cdot x_{t-1,\text{pos}}^2), \quad (15)$$

the linear Kalman filter yields a 40.1% increase in average tracking error compared to the truly linear case.

5.1.2 Multi-modal Distributions

The Gaussian assumption forces the filter to represent multi-modal posteriors with a single mode, leading to either:

- Mode averaging, producing estimates in low-probability regions
- Mode collapse, ignoring alternative hypotheses

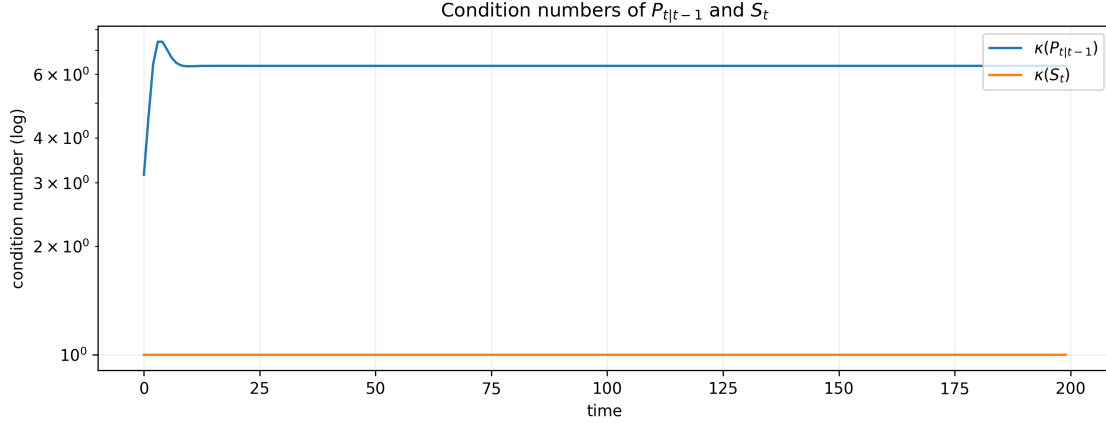


Figure 4: Condition numbers over time. The predicted covariance $P_{t|t-1}$ remains moderately conditioned, while the innovation covariance S_t stays isotropic with $\kappa(S_t) = 1$, confirming stable Kalman gain computation in this example.

5.1.3 Heavy-tailed Noise

Real systems often exhibit outliers from sensor failures or occlusions. The quadratic cost function in Kalman filtering makes it highly sensitive to such outliers, with a single corrupt measurement potentially destabilizing the entire state estimate.

5.2 Particle Filter Motivation

These limitations motivate sequential Monte Carlo (particle filter) methods that:

- Represent arbitrary distributions through weighted samples
- Handle non-linear dynamics and observations without approximation
- Accommodate non-Gaussian noise models

However, particle filters suffer from weight degeneracy and particle impoverishment, especially in high dimensions. This motivates particle flow methods that continuously transport particles from prior to posterior distributions using optimal transport theory.

References

- [1] R. E. Kalman. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1):35–45, 1960.
- [2] A. Doucet and A. M. Johansen. A tutorial on particle filtering and smoothing: Fifteen years later. In *Oxford Handbook of Nonlinear Filtering*, pages 656–704. Oxford University Press, 2009.
- [3] F. Daum and J. Huang. Particle flow for nonlinear filters. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 5920–5923, 2013.
- [4] S. Reich. A nonparametric ensemble transform method for Bayesian inference. *SIAM Journal on Scientific Computing*, 35(4):A2013–A2024, 2013.

- [5] B. D. O. Anderson and J. B. Moore. *Optimal Filtering*. Prentice-Hall, Englewood Cliffs, NJ, 1979.