Gaussian Filtering summary and motivation for Particle Filtering

Particle Filtering in Dynamical Systems

Sai Sampath Kedari
sampath@umich.edu

Contents

1 Introduction

Bayesian state estimation provides a principled framework for inferring the latent state of a dynamical system from noisy measurements. At its core lies the Bayesian filtering recursion, which defines an exact update of probability distributions over time.

In practice, however, the filtering distributions produced by this recursion rapidly become intractable. A wide class of filtering methods has therefore been developed by introducing increasingly restrictive assumptions on system dynamics, noise models, or distributional structure.

This report summarizes that progression and identifies its fundamental limitations. The goal is to motivate the need for non-parametric representations of filtering distributions, setting the stage for particle filtering in subsequent work.

2 What Has Been Solved So Far

Before deriving particle filtering, it is essential to clearly position it within the broader landscape of Bayesian filtering methods developed in the previous reports. This section reconstructs the logical progression of ideas that lead from exact Bayesian inference to parametric filtering approximations, and exposes the structural limitations that ultimately motivate particle filtering.

2.1 The Core Inference Problem

We consider a general discrete-time state-space model

$$X_{n+1} = \Phi(X_n), \tag{1}$$
$$Y_n = h(X_n), \tag{2}$$

where X_n denotes the latent system state and Y_n the observation at time n. The functions $\Phi(\cdot)$ and $h(\cdot)$ may be linear or nonlinear, deterministic or stochastic, and no distributional assumptions are imposed at this level.

The central inference objective is to recursively compute the *filtering distribution*

$$p(X_n \mid Y_{1:n}),$$

that is, the posterior distribution of the current state given all measurements available up to time n.

This problem formulation is exact and fully general. However, it immediately raises a fundamental question: how can probability distributions be propagated through nonlinear dynamical systems over time?

2.2 Exact Bayesian Filtering

Starting from the state-space model, the filtering distribution satisfies the exact Bayesian filtering recursion, which decomposes naturally into two steps.

Prediction

$$p(X_n \mid Y_{1:n-1}) = \int p(X_n \mid X_{n-1})\, p(X_{n-1} \mid Y_{1:n-1})\, dX_{n-1}. \tag{3}$$

Update

$$p(X_n \mid Y_{1:n}) \propto p(Y_n \mid X_n)\, p(X_n \mid Y_{1:n-1}). \tag{4}$$

These equations are exact and impose no restrictions on the system dynamics, measurement model, or noise distributions. However, they operate in the space of probability measures. Under nonlinear dynamics, the filtering distribution generally becomes increasingly complex, high-dimensional, and analytically intractable.

Exact Bayesian filtering is therefore rarely computable in practice.

2.3 Why Exact Filtering Is Intractable

The intractability of Bayesian filtering is not merely computational; it is structural.

Even if the initial distribution is simple, nonlinear state transitions distort the shape of the distribution over time. Successive prediction and update steps introduce skewness, multimodality, and complex dependencies that cannot be captured by low-dimensional parametric families.

As a result, practical filtering methods must introduce approximations. The question is *where* and *how* to approximate.

2.4 Gaussian Filtering

To obtain tractable recursive estimators, a widely adopted approach is to restrict the filtering distribution to a parametric family that is closed under simple transformations. Gaussian filtering enforces the approximation

$$p(X_n \mid Y_{1:n}) \approx \mathcal{N}(m_n, C_n), \quad \text{for all } n,$$

thereby representing the belief over the system state using only its first two moments.

Under this assumption, the Bayesian filtering recursion no longer propagates arbitrary probability distributions. Instead, prediction and update are expressed as operations on the mean and covariance of a Gaussian random variable.

Importantly, the Gaussian filtering equations themselves are *exact consequences* of this distributional restriction. All Gaussian filters differ only in how the resulting integrals are evaluated or approximated.

Prediction The prediction step propagates the prior Gaussian distribution through the system dynamics:

$$m_n^- = \int \Phi(x_{n-1})\, \mathcal{N}(x_{n-1}; m_{n-1}, C_{n-1})\, dx_{n-1}, \tag{5}$$

$$C_n^- = \int \big(\Phi(x_{n-1}) - m_n^-\big)\big(\Phi(x_{n-1}) - m_n^-\big)^\top \mathcal{N}(x_{n-1}; m_{n-1}, C_{n-1})\, dx_{n-1} + \Sigma. \tag{6}$$

Update The update step incorporates the measurement information by computing the predicted measurement moments and cross-covariances:

$$\mu_n = \int h(x_n)\,\mathcal{N}(x_n; m_n^-, C_n^-)\,dx_n, \tag{7}$$

$$S_n = \int \big(h(x_n) - \mu_n\big)\big(h(x_n) - \mu_n\big)^\top \mathcal{N}(x_n; m_n^-, C_n^-)\,dx_n + \Gamma, \tag{8}$$

$$U_n = \int (x_n - m_n^-)\big(h(x_n) - \mu_n\big)^\top \mathcal{N}(x_n; m_n^-, C_n^-)\,dx_n. \tag{9}$$

The filtering distribution is then updated as

$$m_n = m_n^- + U_n S_n^{-1}(y_n - \mu_n), \qquad C_n = C_n^- - U_n S_n^{-1} U_n^\top.$$

These equations define the *Gaussian filtering framework*. They are exact under the assumption that the filtering distribution is Gaussian. All subsequent Gaussian filtering methods arise from different strategies for approximating the above integrals when the system dynamics or measurement models are nonlinear.

2.5 Kalman Filtering: Exact Closure in the Linear Gaussian Case

When the system dynamics and measurement models are linear,

$$\Phi(X_n) = AX_n, \qquad h(X_n) = HX_n,$$

and the noise processes are Gaussian, the Gaussian family is preserved exactly under the Bayesian filtering recursion.

This yields the Kalman filter, which is an exact Bayesian estimator. No approximations are introduced; the restriction to Gaussian distributions is mathematically consistent with the system model.

This case represents a rare alignment between modeling assumptions and mathematical closure.

2.6 Extended Kalman Filter: Approximating the Model

When the system is nonlinear, Gaussianity is no longer preserved.

The Extended Kalman Filter addresses this by approximating the *system dynamics and measurement models themselves* via first-order Taylor expansions around the current mean:

$$\Phi(x) \approx \Phi(m) + A(x - m), \qquad h(x) \approx h(m) + H(x - m),$$

where A and H are Jacobians.

This transforms the nonlinear system into a locally linear one, enabling closed-form Gaussian updates. The approximation error arises from linearizing the model, not from numerical integration.

2.7 Gaussian–Hermite and Unscented Kalman Filters: Approximating the Integrals

Gaussian–Hermite and Unscented Kalman Filters take a different approach.

Rather than linearizing the system, they retain the original nonlinear dynamics and measurement models, and instead approximate the Gaussian filtering integrals directly using numerical quadrature.

The Gaussian–Hermite Kalman Filter employs Gaussian quadrature rules, while the Unscented Kalman Filter uses deterministic sigma points designed to match low-order moments.

In both cases, the approximation targets the *evaluation of expectations*, not the model structure. Nevertheless, the filtering distribution is still enforced to remain Gaussian at every time step.

2.8 The Fundamental Limitation of Gaussian Filtering

Despite their differences, all Gaussian filtering methods share a common structural assumption: the filtering distribution is constrained to a Gaussian family.

This restriction is severe. Many realistic dynamical systems produce filtering distributions that are multimodal, skewed, or heavy-tailed. Such distributions cannot be faithfully represented by any Gaussian, regardless of how accurately its moments are computed.

The failure is not numerical, but representational.

2.9 Motivation for Particle Filtering

Particle filtering arises when the parametric restriction is removed entirely.

Rather than approximating the filtering distribution by a fixed parametric family, particle filtering represents

$$p(X_n \mid Y_{1:n})$$

non-parametrically, using a weighted empirical measure of the form

$$p(X_n \mid Y_{1:n}) \approx \sum_{i=1}^{N} w_n^{(i)} \, \delta_{X_n^{(i)}}.$$

This representation allows arbitrary distributional shapes to emerge naturally under the Bayesian filtering recursion.

The central question addressed in this report is therefore:

> How does the exact Bayesian filtering recursion transform when probability distributions are approximated by weighted empirical measures and propagated sequentially in time?

Answering this question leads directly to *sequential self-normalized importance sampling*, which forms the probabilistic foundation of particle filtering.