Feedback Particle Filter: Design, Approximation, and Error Analysis
Presented at the University of California Irvine

Amirhossein Taghvaei
Ph.D. advisor: Prashant G. Mehta

Coordinated Science Laboratory
University of Illinois at Urbana-Champaign

May 9, 2018
Nonlinear filtering  (this talk)


Sampling

- A. Taghvaei, P. G. Mehta. Accelerated gradient flow for probability distributions, ICML, 2019

Deep learning


Common theme

- Control problem on probability distribution
- Interacting particle system as algorithm
Research Overview

Nonlinear filtering (this talk)

Sampling
- A. Taghvaei, P. G. Mehta. Accelerated gradient flow for probability distributions, ICML, 2019

Deep learning

Common theme
- Control problem on probability distribution
- Interacting particle system as algorithm
Research Overview

Nonlinear filtering (this talk)

Sampling
- A. Taghvaei, P. G. Mehta. Accelerated gradient flow for probability distributions, ICML, 2019

Deep learning

Common theme
- Control problem on probability distribution
- Interacting particle system as algorithm
Outline

- Part I: Background
  - Filtering problem
  - Kalman filter (1960’s)
  - Ensemble Kalman filter & particle filter (1990s)

- Part II: Feedback Particle Filter (2013-)
  - Design
  - Approximation
  - Error analysis

Message of part I:
Kalman filter → Ensemble Kalman filter → Feedback Particle filter
and importance controlled interacting particle systems

Message of part II: How to analyze these systems
Outline

- Part I: Background
  - Filtering problem
  - Kalman filter (1960's)
  - Ensemble Kalman filter & particle filter (1990s)

- Part II: Feedback Particle Filter (2013-)
  - Design
  - Approximation
  - Error analysis

Message of part I:
Kalman filter $\rightarrow$ Ensemble Kalman filter $\rightarrow$ Feedback Particle filter
and importance controlled interacting particle systems

Message of part II: How to analyze these systems
Outline

- Part I: Background
  - Filtering problem
  - Kalman filter (1960’s)
  - Ensemble Kalman filter & particle filter (1990s)

- Part II: Feedback Particle Filter (2013–)
  - Design
  - Approximation
  - Error analysis

Message of part I:

Kalman filter $\rightarrow$ Ensemble Kalman filter $\rightarrow$ Feedback Particle filter

and importance controlled interacting particle systems

Message of part II: How to analyze these systems
Filtering Problem
Example: Navigation

State

Hidden state: Position and orientation of quadrotor

Observation: Camera, GPS, and motion sensor

Problem: Estimate the state based on observation

Observation

Filtering approach: Compute the conditional probability distribution
Filtering Problem
Example: Navigation

**State**

**Hidden state:** Position and orientation of quadrotor

**Observation:** Camera, GPS, and motion sensor

**Problem:** Estimate the state based on observation

**Filtering approach:** Compute the conditional probability distribution
Filtering problem
Mathematical formulation in continuous-time

Dynamical system:

**State process:** \( \text{d}X_t = (\text{dynamical model}), \quad X_0 \sim p_0(\cdot) \)

**Observation process:** \( \text{d}Z_t = h(X_t) \text{d}t + \text{d}W_t \)

Filtering objective: Compute the posterior distribution \( P(X_t|Z_{[0,t]}) \)

Solution:
- In principle, Bayes rule. In practice, impossible to implement it
- Linear Gaussian setting: Kalman filter
- General setting: Approximate solutions

Filtering problem
Mathematical formulation in continuous-time

Dynamical system:

State process: \( \text{d}X_t = (\text{dynamical model}), \quad X_0 \sim p_0(\cdot) \)

Observation process: \( \text{d}Z_t = h(X_t) \text{d}t + \text{d}W_t \)

Filtering objective: Compute the posterior distribution \( P(X_t | Z_{[0,t]}) \)

Solution:

- In principle, Bayes rule. In practice, impossible to implement it
- Linear Gaussian setting: Kalman filter
- General setting: Approximate solutions

J. Xiong, An introduction to stochastic filtering theory, 2008
Kalman-Bucy Filter

**Linear Gaussian setting:**
- linear dynamics: \( dX_t = AX_t \, dt + \sigma_B \, dB_t \)
- linear observation model: \( h(x) = Hx \)

**Kalman-Bucy filter:** \( P(X_t|Z_t) \) is Gaussian \( \mathcal{N}(m_t, \Sigma_t) \)

Update for mean:
\[
\begin{align*}
    dm_t &= Am_t \, dt + K_t \left( dZ_t - Hm_t \, dt \right) \\
&= \text{dynamics} + \text{correction}
\end{align*}
\]

Update for covariance:
\[
\frac{d\Sigma_t}{dt} = (\text{Ricatti equation})
\]

Kalman gain:
\[
K_t := \Sigma_t H^T
\]

**Properties**
- Close relation to optimal control theory
- Strong results about the stability of the filter

**Question:** What is the generalization to the nonlinear and non-Gaussian setting?

R. E Kalman and R. S Bucy. New results in linear filtering and prediction theory, 1961
Kalman-Bucy Filter

**Linear Gaussian setting:**
- Linear dynamics: \( \text{d}X_t = AX_t \text{d}t + \sigma_B \text{d}B_t \)
- Linear observation model: \( h(x) = Hx \)

**Kalman-Bucy filter:** \( P(X_t|Z_t) \) is Gaussian \( \mathcal{N}(m_t, \Sigma_t) \)

**Update for mean:**
\[
\text{d}m_t = Am_t \text{d}t + K_t (\text{d}Z_t - Hm dt)
\]

**Update for covariance:**
\[
\frac{\text{d} \Sigma_t}{\text{d}t} = (\text{Ricatti equation})
\]

**Kalman gain:**
\[
K_t := \Sigma_t H^T
\]

**Properties**
- Close relation to optimal control theory
- Strong results about the stability of the filter

**Question:** What is the generalization to the nonlinear and non-Gaussian setting?

R. E Kalman and R. S Bucy. New results in linear filtering and prediction theory, 1961
Kalman-Bucy Filter

**Linear Gaussian setting:**
- linear dynamics: \( dX_t = AX_t \, dt + \sigma_B \, dB_t \)
- linear observation model: \( h(x) = Hx \)

**Kalman-Bucy filter:** \( P(X_t | Z_t) \) is Gaussian \( \mathcal{N}(m_t, \Sigma_t) \)

\[
\begin{align*}
\text{Update for mean:} & \quad dm_t = Am_t \, dt + K_t \left( dZ_t - Hm_t \, dt \right) \\
\text{Update for covariance:} & \quad \frac{d\Sigma_t}{dt} = (\text{Ricatti equation}) \\
\text{Kalman gain:} & \quad K_t := \Sigma_t H^T
\end{align*}
\]

**Properties**
- Close relation to optimal control theory
- Strong results about the stability of the filter

**Question:** What is the generalization to the nonlinear and non-Gaussian setting?

R. E Kalman and R. S Bucy. New results in linear filtering and prediction theory, 1961
Kalman-Bucy Filter

Linear Gaussian setting:
- linear dynamics: \( dX_t = AX_t \, dt + \sigma_B \, dB_t \)
- linear observation model: \( h(x) = Hx \)

Kalman-Bucy filter: \( P(X_t|Z_t) \) is Gaussian \( \mathcal{N}(m_t, \Sigma_t) \)

Update for mean:
\[
dm_t = Am_t \, dt + K_t \left( dZ_t - Hm_t \, dt \right)
\]

Update for covariance:
\[
d\Sigma_t \, dt = \text{(Ricatti equation)}
\]

Kalman gain:
\[
K_t := \Sigma_t H^T
\]

Properties
- Close relation to optimal control theory
- Strong results about the stability of the filter

Question: What is the generalization to the nonlinear and non-Gaussian setting?

R. E. Kalman and R. S. Bucy. New results in linear filtering and prediction theory, 1961
Monte-Carlo Approximation

- Filtering problem has no finite-dim. solution in general → approximations
- Monte-Carlo method: Approximate with empirical distribution of $N$ particles

$$P(X_t) \approx \frac{1}{N} \sum_{i=1}^{N} \delta_{X^i_t}$$

Example:

**State process:** \(dX_t = a(X_t) \, dt + \sigma(X_t) \, dB_t\)

**Objective:** compute \(P(X_t)\)

Monte-Carlo method: Simulate \(N\) independent samples

$$dX^i_t = a(X^i_t) \, dt + \sigma(X^i_t) \, dB^i_t, \quad \text{for} \quad i = 1, \ldots, N$$

Question: Can we generalize this idea to the filtering problem?
Monte-Carlo Approximation

- Filtering problem has no finite-dim. solution in general $\rightarrow$ approximations
- Monte-Carlo method: Approximate with empirical distribution of $N$ particles

$$P(X_t) \approx \frac{1}{N} \sum_{i=1}^{N} \delta_{X_t^i}$$

Example:

**State process:** \( dX_t = a(X_t) \, dt + \sigma(X_t) \, dB_t \)

**Objective:** compute \( P(X_t) \)

Monte-Carlo method: Simulate $N$ independent samples

\[ dX_t^i = a(X_t^i) \, dt + \sigma(X_t^i) \, dB_t^i, \quad \text{for} \quad i = 1, \ldots, N \]

Question: Can we generalize this idea to the filtering problem?
Monte-Carlo Approximation

- Filtering problem has no finite-dim. solution in general $\rightarrow$ approximations
- Monte-Carlo method: Approximate with empirical distribution of $N$ particles

$$P(X_t) \approx \frac{1}{N} \sum_{i=1}^{N} \delta_{X^i_t}$$

Example:

**State process:** \(dX_t = a(X_t) \, dt + \sigma(X_t) \, dB_t\)

**Objective:** compute \(P(X_t)\)

**Monte-Carlo method:** Simulate $N$ independent samples

\[dX^i_t = a(X^i_t) \, dt + \sigma(X^i_t) \, dB^i_t, \quad \text{for } i = 1, \ldots, N\]

Question: Can we generalize this idea to the filtering problem?
Monte-Carlo Approximation

- Filtering problem has no finite-dim. solution in general → approximations
- Monte-Carlo method: Approximate with empirical distribution of $N$ particles

$$P(X_t) \approx \frac{1}{N} \sum_{i=1}^{N} \delta_{X_t^i}$$

Example:

**State process:**  
$$dX_t = a(X_t) \, dt + \sigma(X_t) \, dB_t$$

**Objective:** compute $P(X_t)$

**Monte-Carlo method:** Simulate $N$ independent samples

$$dX^i_t = a(X^i_t) \, dt + \sigma(X^i_t) \, dB^i_t, \quad \text{for} \quad i = 1, \ldots, N$$

**Question:** Can we generalize this idea to the filtering problem?
Ensemble Kalman filter
Monte-Carlo approximation of Kalman filter

**Idea:** Propagate particles \( \{X_t^i\}_{i=1}^N \sim P(X_t|Z_t) \) instead of mean and covariance

\[
dX_t^i = AX_t^i \, dt + \sigma_B \, dB_t^i + K_t^{(N)} \left( dZ_t - \frac{HX_t^i + N^{-1} \sum_{j=1}^N HX_t^j}{2} \, dt \right) \quad X_0^i \sim p_0
\]

- \( K_t^{(N)} = \Sigma_t^{(N)} H^T \)
- \( \Sigma_t^{(N)} \) is empirical covariance

**Properties**
- In the limit \( (N = \infty) \), the mean and variance evolve according to the Kalman filter
- Computational complexity is \( O(Nd) \), efficient when \( d \gg N \)
- It is not exact in a general setting

**Question:** What is the generalization of EnKF in a general setting?

---

K. Bergemann and S. Reich. *An ensemble Kalman-Bucy filter for continuous data assimilation*, 2012
Ensemble Kalman filter
Monte-Carlo approximation of Kalman filter

Idea: Propagate particles \( \{ X_t^i \}_{i=1}^N \) \( \sim P(X_t|Z_t) \) instead of mean and covariance

\[
dX_t^i = A X_t^i \, dt + \sigma_B \, dB_t^i + K_t^{(N)} \left( dZ_t - \frac{H X_t^i + N^{-1} \sum_{j=1}^{N} H X_t^j}{2} \, dt \right)
\]

- \( K_t^{(N)} = \Sigma_t^{(N)} H^T \)
- \( \Sigma_t^{(N)} \) is empirical covariance

Properties

- In the limit (\( N = \infty \)), the mean and variance evolve according to the Kalman filter
- Computational complexity is \( O(Nd) \), efficient when \( d \gg N \)
- It is not exact in a general setting

Question: What is the generalization of EnKF in a general setting?

K. Bergemann and S. Reich. An ensemble Kalman-Bucy filter for continuous data assimilation, 2012
Ensemble Kalman filter
Monte-Carlo approximation of Kalman filter

**Idea:** Propagate particles \( \{X_t^i\}_{i=1}^N \sim P(X_t|Z_t) \) instead of mean and covariance

\[
\begin{align*}
\text{d}X_t^i &= AX_t^i \text{d}t + \sigma_B \text{d}B_t^i + K_t^{(N)} \left( \text{d}Z_t - \frac{HX_t^i + N^{-1} \sum_{j=1}^N HX_t^j}{2} \text{d}t \right) \\
X_0^i &\sim p_0
\end{align*}
\]

- \( K_t^{(N)} = \Sigma_t^{(N)} H^T \)
- \( \Sigma_t^{(N)} \) is empirical covariance

**Properties**

- In the limit (\( N = \infty \)), the mean and variance evolve according to the Kalman filter
- Computational complexity is \( O(Nd) \), efficient when \( d \gg N \)
- It is not exact in a general setting

**Question:** What is the generalization of EnKF in a general setting?

K. Bergemann and S. Reich. An ensemble Kalman-Bucy filter for continuous data assimilation, 2012
Idea: Approximate the posterior $P(X_t|Z_t)$ using weighted dist. of particles $\{X^i_t, W^i_t\}_{i=1}^N$

$$P(X_t|Z_t) \approx \sum_{i=1}^N W^i_t \delta_{X^i_t}$$

Update the weights based on the likelihood model (importance sampling)
**Idea:** Approximate the posterior $P(X_t|Z_t)$ using weighted dist. of particles $\{X^i_t, W^i_t\}_{i=1}^N$

$$P(X_t|Z_t) \approx \sum_{i=1}^N W^i_t \delta_{X^i_t}$$

Update the weights based on the likelihood model (importance sampling)

---


**Particle filter**
Sequential Monte-Carlo method

**Idea:** Approximate the posterior $P(X_t|Z_t)$ using weighted dist. of particles $\{X_t^i, W_t^i\}_{i=1}^N$

$$P(X_t|Z_t) \approx \sum_{i=1}^N W_t^i \delta_{X_t^i}$$

Update the weights based on the likelihood model (importance sampling)

---

**Particle filter**
Sequential Monte-Carlo method

**Idea:** Approximate the posterior $P(X_t|Z_t)$ using weighted dist. of particles $\{X^i_t, W^i_t\}_{i=1}^N$

$$P(X_t|Z_t) \approx \sum_{i=1}^N W^i_t \delta_{X^i_t}$$

Update the weights based on the likelihood model (importance sampling)

**Problems:**

1. Particle impoverishment for high-dimensional problems – $N \propto \exp(d)$
2. No control structure
3. No relation to Ensemble Kalman filter

---

**Feedback particle filter**  
Controlled interacting particle system

**Idea:** Monte-Carlo approximation + Control

\[
P(X_t|Z_t) \approx \frac{1}{N} \sum_{i=1}^{N} \delta_{X_t^i}
\]

**Update formula:**

\[
dX_t^i = (\text{dynamic model}) + K_t(X_t^i) \circ (dZ_t - \frac{h(X_t^i) + N^{-1} \sum_{j=1}^{N} h(X_t^j)}{2} dt), \quad X_0^i \sim p_0
\]

Gain function:  \( K_t(x) = \nabla \phi_t(x) \)  where \( \phi \) is the solution to a PDE

**Properties**

- Exact in the limit \( N = \infty \)
- Feedback control structure
- Simplify to EnKF in linear Gaussian setting

---

Feedback particle filter
Controlled interacting particle system

**Idea:** Monte-Carlo approximation + Control

\[ P(X_t | Z_t) \approx \frac{1}{N} \sum_{i=1}^{N} \delta_{X_t^i} \]

**Update formula:**

\[
dX_t^i = \text{(dynamic model)} + K_t(X_t^i) \circ \left( dZ_t - \frac{h(X_t^i) + N^{-1} \sum_{j=1}^{N} h(X_t^j)}{2} \right) dt, \quad X_0^i \overset{i.i.d.}{\sim} p_0
\]

Gain function: \( K_t(x) = \nabla \phi_t(x) \) where \( \phi \) is the solution to a PDE

**Properties**

- Exact in the limit \( N = \infty \)
- Feedback control structure
- Simplify to EnKF in linear Gaussian setting

---

Analysis of FPF

Recap:

**KF:** \[ dm_t = (\text{dynamical model}) + K_t (dZ_t - Hm_t \, dt) \]

**EnKF:** \[ dX^i_t = (\text{dynamical model}) + K^{(N)}_t (dZ_t - \frac{HX^i_t + N^{-1} \sum_{j=1}^{N}HX^j_t}{2} \, dt) \]

**FPF:** \[ dX^i_t = (\text{dynamical model}) + \mathcal{K}^{(N)}_t(X^i_t) \circ (dZ_t - \frac{h(X^i_t) + N^{-1} \sum_{j=1}^{N}h(X^j_t)}{2} \, dt) \]

Analysis of FPF:

<table>
<thead>
<tr>
<th>Design</th>
<th>Approximation</th>
<th>Error analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean-field limit ( \bar{X}_t, \bar{U}_t )</td>
<td>Finite-( N ) system ( {X^i_t, U^i_t}_{i=1}^{N} )</td>
<td></td>
</tr>
</tbody>
</table>
**Recap:**

**KF:** \[ dm_t = (\text{dynamical model}) + K_t (dZ_t - Hm_t \, dt) \]

**EnKF:** \[ dX^i_t = (\text{dynamical model}) + K^{(N)}_t (dZ_t - \frac{HX^i_t + N^{-1} \sum_{j=1}^N HX^j_t}{2} \, dt) \]

**FPF:** \[ dX^i_t = (\text{dynamical model}) + K^{(N)}_t (X^i_t) \circ (dZ_t - \frac{h(X^i_t) + N^{-1} \sum_{j=1}^N h(X^j_t)}{2} \, dt) \]

**Analysis of FPF:**

<table>
<thead>
<tr>
<th>Design</th>
<th>Approximation</th>
<th>Error analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean-field limit ( \bar{X}_t, \bar{U}_t )</td>
<td>( \rightarrow )</td>
<td>Finite-( N ) system ( {X^i_t, U^i_t}_{i=1}^N )</td>
</tr>
</tbody>
</table>
Outline

- Background
  - Filtering problem
  - Kalman filter (1960's)
  - Monte-Carlo approximation: Ensemble Kalman filter & particle filters (1990s)

- Feedback Particle Filter
  - Design
  - Approximation
  - Error analysis
Design problem overview

The trajectory of the posterior distribution $p_t := P(X_t | Z_{[0,t]})$ in the probability space

**Design problem:** Construct a random process $\tilde{X}_t$ that follows the posterior, i.e.

$$\tilde{X}_t \sim p_t, \quad \forall t \geq 0$$

**Non-uniqueness:** There are infinitely many solutions
The trajectory of the posterior distribution $p_t := P(X_t | Z_{[0,t]})$ in the probability space.

**Design problem:** Construct a random process $\tilde{X}_t$ that follows the posterior, i.e.

$$\tilde{X}_t \sim p_t, \quad \forall t \geq 0$$

**Non-uniqueness:** There are infinitely many solutions.
The trajectory of the posterior distribution $p_t := P(X_t | Z_{[0,t]})$ in the probability space

**Design problem:** Construct a random process $\bar{X}_t$ that follows the posterior, i.e

$$\bar{X}_t \sim p_t, \quad \forall t \geq 0$$

**Non-uniqueness:** There are infinitely many solutions
Example:

**State process:** \( dX_t = dB_t, \quad X_0 \sim N(0, 1) \)

**Objective:** compute \( P(X_t) \)

Two solutions:

(I) \( dX_t^i = dB_t^i \)

(II) \( \frac{d}{dt} X_t^i = \frac{X_t^i}{\frac{2}{N} \sum_{j=1}^{N} (X_t^j)^2} \)

They both produce the same distribution \( N(0, 1 + t) \).
Example:

State process: \( \text{d}X_t = \text{d}B_t, \quad X_0 \sim N(0, 1) \)

Objective: compute \( P(X_t) \)

Two solutions:

(I) \( \text{d}X_t^i = \text{d}B_t^i \)

(II) \( \frac{\text{d}}{\text{d}t} X_t^i = \frac{X_t^i}{\frac{2}{N} \sum_{j=1}^{N} (X_t^j)^2} \)

They both produce the same distribution \( N(0, 1 + t) \).
Example:

**State process:** \( dX_t = dB_t, \quad X_0 \sim N(0, 1) \)

**Objective:** compute \( P(X_t) \)

**Two solutions:**

(I) \( dX_t^i = dB_t^i \)

(II) \( \frac{d}{dt} X_t^i = \frac{X_t^i}{\frac{2}{N} \sum_{j=1}^{N} (X_t^j)^2} \)

They both produce the same distribution \( N(0, 1 + t) \).
Optimal transportation approach

- Reason for non-uniqueness: Only the marginal distributions, at each time instant, are specified.
- Optimal transport maps provide a way to uniquely couple two distributions.

**Proposed solution:** Infinitesimal optimal transport maps

\[ \tilde{X}_{t+\Delta t} = T_t(\tilde{X}_t), \]

- \( T_t \) is the optimal transport map between \( p_t \) and \( p_{t+\Delta t} \).
- Take the limit as \( \Delta t \to 0 \).
Optimal transportation approach

Reason for non-uniqueness: Only the marginal distributions, at each time instant, are specified.

Optimal transport maps provide a way to uniquely couple two distributions.

**Proposed solution:** Infinitesimal optimal transport maps

\[ \bar{X}_{t+\Delta t} = T_t(\bar{X}_t), \]

- \(T_t\) is the optimal transport map between \(p_t\) and \(p_t+\Delta t\)
- Take the limit as \(\Delta t \to 0\)
The procedure is carried out in linear Gaussian setting.

Recall in this setting, only the mean and variance are important.

**Proposition**

In linear Gaussian setting, the optimal transportation procedure result in the following process:

\[ d\bar{X}_t = \text{(terms effecting the mean)} + G_t(\bar{X}_t - \bar{m}_t) \, dt \]

where \( G_t \) is the unique symmetric solution to the Lyapunov equation

\[ G_t \bar{\Sigma}_t + \bar{\Sigma}_t G_t = \text{Ricc}(\bar{\Sigma}_t) \]

**Comparison:** A non-optimal (and stochastic) solution is

\[ d\bar{X}_t = \text{(terms effecting the mean)} + \left( A - \frac{1}{2} \bar{\Sigma}_t H^\top H \right)(\bar{X}_t - \bar{m}_t) \, dt + \sigma_B \, d\bar{B}_t \]

**Question:** What is the difference between the two forms of the solution? Does the optimal transport way result in a more stable procedure?

---

A. Taghvaei, P. G. Mehta, An optimal transport formulation for the linear feedback particle filter, (ACC) 2016
The procedure is carried out in linear Gaussian setting.

Recall in this setting, only the mean and variance are important.

Proposition

In linear Gaussian setting, the optimal transportation procedure result in the following process:

\[ d\bar{X}_t = (\text{terms effecting the mean}) + G_t(\bar{X}_t - \bar{m}_t) \, dt \]

where \( G_t \) is the unique symmetric solution to the Lyapunov equation

\[ G_t \bar{\Sigma}_t + \bar{\Sigma}_t G_t = \text{Ricc}(\bar{\Sigma}_t) \]

Comparison: A non-optimal (and stochastic) solution is

\[ d\tilde{X}_t = (\text{terms effecting the mean}) + (A - \frac{1}{2} \bar{\Sigma}_t H^\top H)(\tilde{X}_t - \tilde{m}_t) \, dt + \sigma_B \, d\tilde{B}_t \]

Question: What is the difference between the two forms of the solution? Does the optimal transport way result in a more stable procedure?

A. Taghvaei, P. G. Mehta, An optimal transport formulation for the linear feedback particle filter, (ACC) 2016
The procedure is carried out in linear Gaussian setting.

Recall in this setting, only the mean and variance are important.

**Proposition**

In linear Gaussian setting, the optimal transportation procedure result in the following process:

\[
d\tilde{X}_t = \text{(terms effecting the mean)} + G_t(\tilde{X}_t - \tilde{m}_t)\, dt
\]

where \( G_t \) is the unique symmetric solution to the Lyapunov equation

\[
G_t\tilde{\Sigma}_t + \tilde{\Sigma}_t G_t = \text{Ricc}(\tilde{\Sigma}_t)
\]

**Comparison:** A non-optimal (and stochastic) solution is

\[
d\tilde{X}_t = \text{(terms effecting the mean)} + \left(A - \frac{1}{2}\tilde{\Sigma}_t H^\top H\right)(\tilde{X}_t - \tilde{m}_t)\, dt + \sigma_B\, d\tilde{B}_t
\]

**Question:** What is the difference between the two forms of the solution? Does the optimal transport way result in a more stable procedure?

A. Taghvaei, P. G. Mehta, An optimal transport formulation for the linear feedback particle filter, (ACC) 2016
Outline

- Background
  - Filtering problem
  - Kalman filter (1960's)
  - Monte-Carlo approximation: Ensemble Kalman filter & particle filters (1990s)

- Feedback Particle Filter
  - Design
  - Approximation
  - Error analysis
Approximation
Problem formulation

**FPF update formula:**

\[ dX_t^i = (\text{dynamic model}) + K_t(X_t^i) \circ (dZ_t - \frac{1}{2}(h(X_t^i) + \hat{h}_t) \, dt) \]

Gain function  \( K_t(x) = \nabla \phi_t(x) \)  where  \( \phi \) solves the Poisson eq.

**Poisson equation:**

\[ -\frac{1}{p_t(x)} \nabla \cdot (p_t(x) \nabla \phi_t(x)) = h(x) - \hat{h}_t \]

**Computational problem:**

Given:  \( \{X_t^1, \ldots, X_t^N\} \ i.i.d \sim p_t \)

Approximate:  \( \{K_t(X_t^1), \ldots, K_t(X_t^N)\} \)
**Linear Gaussian setting**
Relation to ensemble the Kalman filter

General setting

\[ K(x) = ? \]

\[ \nabla \phi(x) \]

\[ \nabla \phi(X^i) \]

\[ X^i \]

\[ x \]

**Linear Gaussian setting**

\[ K(x) = K \] (Kalman gain)

\[ K_t = \text{Kalman gain} \]

\[ X_t^i \]

\[ x \]

---

Linear Gaussian setting
Relation to ensemble the Kalman filter

General setting

$K(x) = ?$

Linear Gaussian setting

$K(x) = K$ (Kalman gain)

**Idea:** Projection into a finite-dim subspace

\[ \phi \in H^1_0(\rho, \mathbb{R}^d) \]

\[ K = \nabla \phi \]

\[ S = \text{span}\{\ldots\} \]

Choice of basis function is difficult
Galerkin approximation

**Idea:** Projection into a finite-dim subspace

\[ \phi \in H^1_0(\rho, \mathbb{R}^d) \]

\[ K = \nabla \phi \]

\[ S = \text{span}\{\ldots\} \]

\[ S = \text{span}\{1, x, \ldots, x^M\} \]

Choice of basis function is difficult
**Galerkin approximation**

**Idea:** Projection into a finite-dim subspace

\[ \phi \in H^1_0(\rho, \mathbb{R}^d) \]

\[ K = \nabla \phi \]

\[ S = \text{span}\{\ldots\} \]

\[ S = \text{span}\{1, x, \ldots, x^M\} \]

Choice of basis function is difficult
**Galerkin approximation**

**Idea:** Projection into a finite-dim subspace

\[ \phi \in H^1_0(\rho, \mathbb{R}^d) \]

\[ S = \text{span}\{\ldots\} \]

\[ K = \nabla \phi \]

\[ S = \text{span}\{1, x, \ldots, x^M\} \]

Choice of basis function is difficult
**Galerkin approximation**

**Idea:** Projection into a finite-dim subspace

\[ \phi \in H_0^1(\rho, \mathbb{R}^d) \]

\[ K = \nabla \phi \]

\[ S = \text{span}\{\ldots\} \]

\[ S = \text{span}\{1, x, \ldots, x^M\} \]

Choice of basis function is difficult
**Idea:** Projection into a finite-dim subspace

\[ \phi \in H_0^1(\rho, \mathbb{R}^d) \]

\[ K = \nabla \phi \]

\[ S = \text{span}\{\ldots\} \]

\[ S = \text{span}\{1, x, \ldots, x^M\} \]

Choice of basis function is difficult
Idea: Projection into a finite-dim subspace

$\phi \in H_0^1(\rho, \mathbb{R}^d)$

$K = \nabla \phi$

$S = \text{span}\{\ldots\}$

$S = \text{span}\{1, x, \ldots, x^M\}$

Choice of basis function is difficult
**Stochastic formulation:**

\[ \phi = P_{\epsilon} \phi + \int_{0}^{\epsilon} P_{s}(h - \hat{h}) \, ds \]

where \( \{P_{t}\} \) is the semigroup for \( \Delta_{\rho} := \frac{1}{\rho} \nabla \cdot (\rho \nabla) \)

Approximate \( P \) with a Markov matrix using particles (Coifman & Lafon, 2006)

The resulting approximation takes the form \( K(X^i) \approx \sum_{j=1}^{N} s_{ij} X^j \)
Diffusion map approximation
Numerical analysis

Error estimates: $\text{r.m.s.e} = O(\epsilon) + O(\frac{1}{\epsilon^{1+d/2}N^{1/2}})$

Question: How does the error effect the distribution?

A. Taghvaei, P. G. Mehta, and S. P. Meyn. Gain function approximation in the feedback particle filter, SIAM (under review)
Diffusion map approximation
Numerical analysis

Error estimates: \( \text{r.m.s.e} = O(\epsilon) + O\left(\frac{1}{\epsilon^{1+d/2}N^{1/2}}\right) \)

Question: How does the error affect the distribution?

A. Taghvaei, P. G. Mehta, and S. P. Meyn. Gain function approximation in the feedback particle filter, SIAM (under review)
Error estimates: \[ \text{r.m.s.e} = O(\epsilon) + O\left(\frac{1}{\epsilon^{1+d/2} N^{1/2}}\right) \]

Question: How does the error effect the distribution?

A. Taghvaei, P. G. Mehta, and S. P. Meyn. Gain function approximation in the feedback particle filter, SIAM (under review)
Diffusion map approximation
Numerical analysis

Error estimates: \( \text{r.m.s.e} = O(\epsilon) + O\left(\frac{1}{\epsilon^{1+d/2} N^{1/2}}\right) \)

Question: How does the error effect the distribution?

A. Taghvaei, P. G. Mehta, and S. P. Meyn. Gain function approximation in the feedback particle filter, SIAM (under review)
Error estimates: \( r.m.s.e = O(\epsilon) + O\left(\frac{1}{\epsilon^{1+d/2}N^{1/2}}\right) \)

Question: How does the error effect the distribution?

A. Taghvaei, P. G. Mehta, and S. P. Meyn. Gain function approximation in the feedback particle filter, SIAM (under review)
**Diffusion map approximation**

**Numerical analysis**

![Graph showing error estimates and question](image)

**Error estimates:**  
\[ \text{r.m.s.e} = O(\epsilon) + O\left(\frac{1}{\epsilon^{1+d/2}N^{1/2}}\right) \]

**Question:** How does the error effect the distribution?

---

**A. Taghvaei, P. G. Mehta, and S. P. Meyn.** Gain function approximation in the feedback particle filter, SIAM (under review)
Diffusion map approximation
Numerical analysis

Error estimates: \( \text{r.m.s.e} = O(\epsilon) + O\left(\frac{1}{\epsilon^{1+d/2} N^{1/2}}\right) \)

Question: How does the error effect the distribution?

A. Taghvaei, P. G. Mehta, and S. P. Meyn. Gain function approximation in the feedback particle filter, SIAM (under review)
Error estimates: \[ \text{r.m.s.e} = O(\epsilon) + O\left(\frac{1}{\epsilon^{1+d/2} N^{1/2}}\right) \]

**Question:** How does the error effect the distribution?
Outline

- Background
  - Filtering problem
  - Kalman filter (1960's)
  - Monte-Carlo approximation: Ensemble Kalman filter & particle filters (1990s)

- Feedback Particle Filter
  - Design
  - Approximation
  - Error analysis
**Error analysis of finite-$N$ system**
Linear Gaussian setting

**Motivation:**
- Simulating Kalman filter is computationally expensive for high-dimensional problems

  if state dimension is $d$  \[ \Rightarrow \]  covariance matrix is $d \times d$
  \[ \Rightarrow \]  computational complexity is $O(d^2)$
  \[ \Rightarrow \]  Not scalable for high-dim problems
  (e.g. weather prediction)

- However EnKF computationally scales better with dimension $O(Nd)$

**Question:** What is the error of the EnKF for finite number of particles?
Error analysis of finite-$N$ system
Linear Gaussian setting

Motivation:
- Simulating Kalman filter is computationally expensive for high-dimensional problems

  if state dimension is $d$  $\Rightarrow$  covariance matrix is $d \times d$
  $\Rightarrow$  computational complexity is $O(d^2)$
  $\Rightarrow$  Not scalable for high-dim problems
  (e.g. weather prediction)

- However EnKF computationally scales better with dimension $O(Nd)$

Question: What is the error of the EnKF for finite number of particles?
Motivation:

- Simulating Kalman filter is computationally expensive for high-dimensional problems

  if state dimension is $d \Rightarrow$ covariance matrix is $d \times d$

  $\Rightarrow$ computational complexity is $O(d^2)$

  $\Rightarrow$ Not scalable for high-dim problems

  (e.g weather prediction)

- However EnKF computationally scales better with dimension $O(Nd)$

Question: What is the error of the EnKF for finite number of particles?
Error analysis of finite-$N$ system

Problem formulation

**Finite-$N$ system:**

$$dX^i_t = (\text{linear dynamics}) + K^{(N)}_t(dZ_t - \frac{1}{2}H(X^i_t + m^{(N)}_t)\,dt), \quad X^i_0 \sim p_0$$

$$K^{(N)}_t = \Sigma^{(N)}_t H^\top$$

with empirical mean $m^{(N)}_t$ and covariance $\Sigma^{(N)}_t$

**Mean-field limit:**

$$d\bar{X}_t = (\text{linear dynamics}) + \bar{K}_t(d\bar{Z}_t - \frac{1}{2}H(\bar{X}_t + \bar{m}_t)\,dt), \quad \bar{X}_0 \sim p_0$$

$$\bar{K}_t = \bar{\Sigma}_t H^\top$$

with mean-field mean $\bar{m}_t = E[\bar{X}_t|Z_t]$ and covariance $\bar{\Sigma}_t = \text{Cov}(\bar{X}_t|Z_t)$

**Error analysis:**

1. Analysis of the mean-field system
2. Analysis of the convergence of the finite-$N$ system to the mean-field limit

$$\begin{align*}
\text{Finite-}N \text{ system} &\overset{2}{\approx} \text{mean-field system} &\overset{1}{=} \text{Kalman filter}
\end{align*}$$
Error analysis of finite-$N$ system
Problem formulation

**Finite-$N$ system:**

\[
\begin{align*}
\text{d}X^i_t &= (\text{linear dynamics}) + K^{(N)}_t (\text{d}Z_t - \frac{1}{2} H (X^i_t + m^{(N)}_t) \text{d}t), \quad X^i_0 \sim p_0 \\
K^{(N)}_t &= \Sigma^{(N)}_t H^\top
\end{align*}
\]

with empirical mean $m^{(N)}_t$ and covariance $\Sigma^{(N)}_t$

**Mean-field limit:**

\[
\begin{align*}
\text{d}\bar{X}_t &= (\text{linear dynamics}) + \bar{K}_t (\text{d}\bar{Z}_t - \frac{1}{2} H (X^i_t + \bar{m}_t) \text{d}t), \quad \bar{X}_0 \sim p_0 \\
\bar{K}_t &= \bar{\Sigma}_t H^\top
\end{align*}
\]

with mean-field mean $\bar{m}_t = \mathbb{E}[\bar{X}_t|\bar{Z}_t]$ and covariance $\bar{\Sigma}_t = \text{Cov}(\bar{X}_t|\bar{Z}_t)$

**Error analysis:**

1. Analysis of the mean-field system
2. Analysis of the convergence of the finite-$N$ system to the mean-field limit

Finite-$N$ system $\underset{(2)}{\approx}$ mean-field system $\underset{(1)}{=} \text{Kalman filter}$
Error analysis of finite-$N$ system
Problem formulation

**Finite-$N$ system:**

$$dX^i_t = (\text{linear dynamics}) + K^{(N)}_t (dZ_t - \frac{1}{2} H(X^i_t m^{(N)}_t) dt), \quad X_0^i \overset{i.i.d.}{\sim} p_0$$

$$K^{(N)}_t = \Sigma^{(N)}_t H^\top$$

with empirical mean $m^{(N)}_t$ and covariance $\Sigma^{(N)}_t$

**Mean-field limit:**

$$d\bar{X}_t = (\text{linear dynamics}) + \bar{K}_t (d\bar{Z}_t - \frac{1}{2} H(\bar{X}_t + \bar{m}_t) dt), \quad \bar{X}_0 \sim p_0$$

$$\bar{K}_t = \bar{\Sigma}_t H^\top$$

with mean-field mean $\bar{m}_t = E[\bar{X}_t|Z_t]$ and covariance $\bar{\Sigma}_t = \text{Cov}(\bar{X}_t|Z_t)$

**Error analysis:**

1. Analysis of the mean-field system
2. Analysis of the convergence of the finite-$N$ system to the mean-field limit

$$\text{Finite-}N\text{ system} \overset{(2)}{\approx} \text{mean-field system} \overset{(1)}{=} \text{Kalman filter}$$
Assumption  The system is stable and the observation matrix is full rank.

Error analysis

Under the assumption, EnKF admits the following error estimates:

\[
\begin{align*}
E[|m_t - m_t^{(N)}|^2] & \leq \frac{(\text{const.})}{N} \\
E[|\Sigma_t - \Sigma_t^{(N)}|^2] & \leq \frac{(\text{const.})}{N}
\end{align*}
\]

where the constant does not depend on time.

Question:

- Kalman filter is stable when the system is stabilizable and detectable
- Can we prove uniform error estimates of EnKF under these conditions?

P Del Moral, J Tugaut. On the stability and the uniform propagation of chaos properties of ensemble Kalman–Bucy filters, 2018
A. Taghvaei, P. G. Mehta, Error analysis of the stochastic linear feedback particle filter, CDC, 2018
Error analysis of the EnKF

**Assumption**  The system is stable and the observation matrix is full rank.

**Error analysis**

Under the assumption, EnKF admits the following error estimates:

\[
E[|m_t - m_t^{(N)}|^2] \leq \frac{(\text{const.})}{N}
\]

\[
E[|\Sigma_t - \Sigma_t^{(N)}|^2] \leq \frac{(\text{const.})}{N}
\]

where the constant does not depend on time.

Question:

- Kalman filter is stable when the system is stabilizable and detectable
- Can we prove uniform error estimates of EnKF under these conditions?

---

P Del Moral, J Tugaut. On the stability and the uniform propagation of chaos properties of ensemble Kalman–Bucy filters, 2018

A. Taghvaei, P. G. Mehta, Error analysis of the stochastic linear feedback particle filter, CDC, 2018
**Assumption**  The system is stable and the observation matrix is full rank.

**Error analysis**

Under the assumption, EnKF admits the following error estimates:

\[
E[|m_t - m_t^{(N)}|^2] \leq \frac{(\text{const.})}{N}
\]

\[
E[|
\Sigma_t - \Sigma_t^{(N)}|^2] \leq \frac{(\text{const.})}{N}
\]

where the constant does not depend on time.

**Question:**

- Kalman filter is stable when the system is stabilizable and detectable
- Can we prove uniform error estimates of EnKF under these conditions?

---

P Del Moral, J Tugaut. *On the stability and the uniform propagation of chaos properties of ensemble Kalman–Bucy filters*, 2018

A. Taghvaei, P. G. Mehta, *Error analysis of the stochastic linear feedback particle filter*, CDC, 2018
Conclusion

Part I: FPF is generalization of Kalman filter

Kalman filter $\rightarrow$ Ensemble Kalman filter $\rightarrow$ Feedback Particle filter

Part II: Analysis of FPF

- Design: Infinitesimal optimal transport maps
- Approximation: Galerkin and Diffusion map approximation
- Error analysis: Convergence in linear Gaussian setting under strong conditions

Question: The three aspects are disjoint. Can they be carried out in a single framework?

Thanks for your attention!
Conclusion

Part I: FPF is generalization of Kalman filter

Kalman filter $\rightarrow$ Ensemble Kalman filter $\rightarrow$ Feedback Particle filter

Part II: Analysis of FPF

<table>
<thead>
<tr>
<th>Design</th>
<th>Approximation</th>
<th>Error analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean-field limit $\bar{X}_t, \bar{U}_t$</td>
<td>$\rightarrow$ Finite-$N$ system ${X_t^i, U_t^i}_{i=1}^N$</td>
<td></td>
</tr>
</tbody>
</table>

- Design: Infinitesimal optimal transport maps
- Approximation: Galerkin and Diffusion map approximation
- Error analysis: Convergence in linear Gaussian setting under strong conditions

Question: The three aspects are disjoint. Can they be carried out in a single framework?

Thanks for your attention!
Conclusion

Part I: FPF is generalization of Kalman filter

Kalman filter $\rightarrow$ Ensemble Kalman filter $\rightarrow$ Feedback Particle filter

Part II: Analysis of FPF

<table>
<thead>
<tr>
<th>Design</th>
<th>Approximation</th>
<th>Error analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean-field limit $\bar{X}_t, \bar{U}_t$</td>
<td>$\longrightarrow$</td>
<td>Finite-$N$ system ${X^i_t, U^i_t}_{i=1}^N$</td>
</tr>
</tbody>
</table>

- Design: Infinitesimal optimal transport maps
- Approximation: Galerkin and Diffusion map approximation
- Error analysis: Convergence in linear Gaussian setting under strong conditions

Question: The three aspects are disjoint. Can they be carried out in a single framework?

Thanks for your attention!
Part I: FPF is generalization of Kalman filter

Kalman filter $\rightarrow$ Ensemble Kalman filter $\rightarrow$ Feedback Particle filter

Part II: Analysis of FPF

<table>
<thead>
<tr>
<th>Design</th>
<th>Approximation</th>
<th>Error analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean-field limit $\bar{X}_t, \bar{U}_t$</td>
<td>$\rightarrow$ Finite-$N$ system ${X^i_t, U^i_t}_{i=1}^N$</td>
<td></td>
</tr>
</tbody>
</table>

- Design: Infinitesimal optimal transport maps
- Approximation: Galerkin and Diffusion map approximation
- Error analysis: Convergence in linear Gaussian setting under strong conditions

Question: The three aspects are disjoint. Can they be carried out in a single framework?

Thanks for your attention!