Gain Function Approximation in the Feedback Particle Filter

5th Midwest Workshop on Control and Game Theory
Purdue University, May, 2016

Amirhossein Taghvaei
Joint work with P. G. Mehta

Coordinated Science Laboratory
University of Illinois at Urbana-Champaign

May 1, 2016

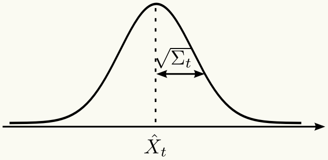
ILLINOIS

Kalman Filter:

$$\mathrm{d}X_t = AX_t\,\mathrm{d}t + \mathrm{d}B_t$$
$$\mathrm{d}Z_t = HX_t\,\mathrm{d}t + \mathrm{d}W_t$$

$\mathsf{P}(X_t|\mathcal{Z}_t) =$ Gaussian $N(\hat{X}_t, \Sigma_t)$,

$$\mathrm{d}\hat{X}_t = A\hat{X}_t\,\mathrm{d}t + \mathsf{K}_t(\mathrm{d}Z_t - H\hat{X}_t\,\mathrm{d}t)$$
$$\frac{\mathrm{d}\Sigma_t}{\mathrm{d}t} = \ldots \text{(Riccati equation)}$$



Feedback Particle Filter:

$$\mathrm{d}X_t = a(X_t)\,\mathrm{d}t + \mathrm{d}B_t$$
$$\mathrm{d}Z_t = h(X_t)\,\mathrm{d}t + \mathrm{d}W_t$$

$\mathsf{P}(X_t|\mathcal{Z}_t) \approx$ empirical dist. of $\{X^1, \ldots, X^N\}$,

$$\mathrm{d}X_t^i = a(X_t^i)\,\mathrm{d}t + \mathrm{d}B_t^i + \mathsf{K}_t(X_t^i) \circ (\mathrm{d}Z_t - \frac{h(X_t^i) + \hat{h}_t}{2}\,\mathrm{d}t)$$

Challenge: Compute the gain function $\mathsf{K}_t(.)$

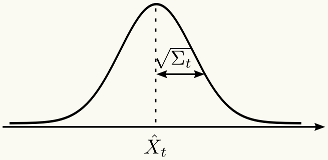
T. Yang, R. S. Laugesen, P. G. Mehta, and S. P. Meyn. Multivariable feedback particle filter, *Automatica*, 2015

Kalman Filter:

$$\mathrm{d}X_t = AX_t\,\mathrm{d}t + \mathrm{d}B_t$$
$$\mathrm{d}Z_t = HX_t\,\mathrm{d}t + \mathrm{d}W_t$$

$\mathsf{P}(X_t|\mathcal{Z}_t) =$ Gaussian $N(\hat{X}_t, \Sigma_t)$,

$$\mathrm{d}\hat{X}_t = A\hat{X}_t\,\mathrm{d}t + \mathsf{K}_t(\,\mathrm{d}Z_t - H\hat{X}_t\,\mathrm{d}t)$$
$$\frac{\mathrm{d}\Sigma_t}{\mathrm{d}t} = \ldots \text{(Riccati equation)}$$

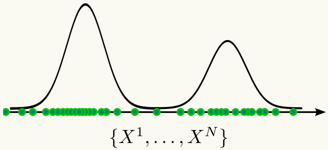


Feedback Particle Filter:

$$\mathrm{d}X_t = a(X_t)\,\mathrm{d}t + \mathrm{d}B_t$$
$$\mathrm{d}Z_t = h(X_t)\,\mathrm{d}t + \mathrm{d}W_t$$

$\mathsf{P}(X_t|\mathcal{Z}_t) \approx$ empirical dist. of $\{X^1, \ldots, X^N\}$,

$$\mathrm{d}X_t^i = a(X_t^i)\,\mathrm{d}t + \mathrm{d}B_t^i + \mathsf{K}_t(X_t^i) \circ (\,\mathrm{d}Z_t - \frac{h(X_t^i) + \hat{h}_t}{2}\,\mathrm{d}t)$$



Challenge: Compute the gain function $\mathsf{K}_t(.)$

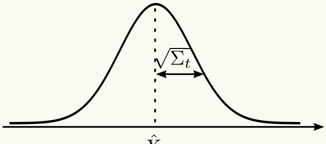
T. Yang, R. S. Laugesen, P. G. Mehta, and S. P. Meyn. Multivariable feedback particle filter, *Automatica*, 2015

Kalman Filter:

$$\mathrm{d}X_t = AX_t\,\mathrm{d}t + \mathrm{d}B_t$$
$$\mathrm{d}Z_t = HX_t\,\mathrm{d}t + \mathrm{d}W_t$$

$\mathsf{P}(X_t|\mathcal{Z}_t)$ = Gaussian $N(\hat{X}_t, \Sigma_t)$,

$$\mathrm{d}\hat{X}_t = A\hat{X}_t\,\mathrm{d}t + \mathsf{K}_t(\,\mathrm{d}Z_t - H\hat{X}_t\,\mathrm{d}t)$$
$$\frac{\mathrm{d}\Sigma_t}{\mathrm{d}t} = \ldots \text{(Riccati equation)}$$

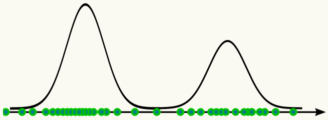


Feedback Particle Filter:

$$\mathrm{d}X_t = a(X_t)\,\mathrm{d}t + \mathrm{d}B_t$$
$$\mathrm{d}Z_t = h(X_t)\,\mathrm{d}t + \mathrm{d}W_t$$

$\mathsf{P}(X_t|\mathcal{Z}_t) \approx$ empirical dist. of $\{X^1, \ldots, X^N\}$,

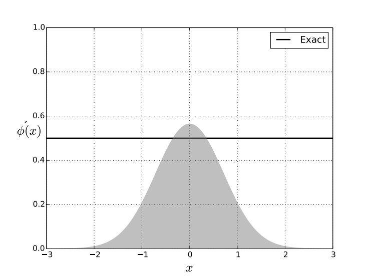
$$\mathrm{d}X_t^i = a(X_t^i)\,\mathrm{d}t + \mathrm{d}B_t^i + \mathsf{K}_t(X_t^i) \circ (\,\mathrm{d}Z_t - \frac{h(X_t^i) + \hat{h}_t}{2}\,\mathrm{d}t)$$



Challenge: Compute the gain function $\mathsf{K}_t(.)$

T. Yang, R. S. Laugesen, P. G. Mehta, and S. P. Meyn. Multivariable feedback particle filter, *Automatica*, 2015

Gaussian distribution
Linear observation

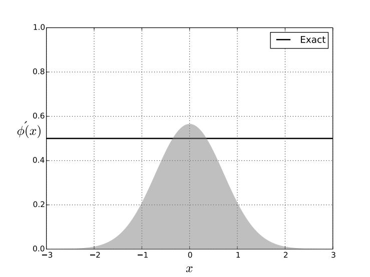


$\mathsf{K}_t(x) = \text{constant}$ (Kalman gain)

Non-Gaussian distribution
Nonlinear observation

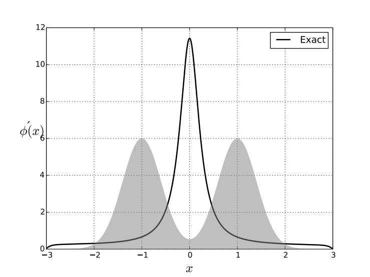
$\mathsf{K}_t(x) = \ldots$ (Nonlinear gain)

Gaussian distribution
Linear observation



$\mathsf{K}_t(x) = \text{constant}$ (Kalman gain)

Non-Gaussian distribution
Nonlinear observation



$\mathsf{K}_t(x) = \ldots$ (Nonlinear gain)

Gain function $\mathsf{K}_t(x) := \nabla\phi(x)$ where ϕ satisfies

Poisson equation: $$-\frac{1}{\rho(x)}\nabla\cdot(\rho(x)\nabla\phi(x)) = h(x) - \hat{h}$$

- ρ is a probability density function
- h is a real-valued function, $\hat{h} = \int h\rho\,\mathrm{d}x$

Poisson equation also appears in:

- Simulation and optimization theory for Markov models [Meyn, Tweedie, 2012]
- Other filtering algorithms [Daum, et. al. 2010]

Problem:

Given: $\{X^1,\ldots,X^N\} \overset{\text{i.i.d}}{\sim} \rho$

Find: $\{\nabla\phi(X^1),\ldots,\nabla\phi(X^N)\}$ (approximately)

Related work: [Berntorp, et. al. 2016],[Radhakrishnan, et. al. 2016]

R. S. Laugesen, P. G. Mehta, S. P. Meyn, and M. Raginsky. Poisson Equation in Nonlinear Filtering. SICON, 2015

Gain Function Approximation in FPF
Problem formulation

Gain function $\mathsf{K}_t(x) := \nabla\phi(x)$ where ϕ satisfies

$$\textbf{Poisson equation:} \quad -\frac{1}{\rho}\nabla\cdot(\rho\nabla\phi) = h$$

- ρ is a probability density function
- h is a real-valued function, $\hat{h} = \int h\rho\,\mathrm{d}x$

Poisson equation also appears in:

- Simulation and optimization theory for Markov models [Meyn, Tweedie, 2012]
- Other filtering algorithms [Daum, et. al. 2010]

Problem:

Given: $\{X^1,\ldots,X^N\} \overset{\text{i.i.d}}{\sim} \rho$

Find: $\{\nabla\phi(X^1),\ldots,\nabla\phi(X^N)\}$ (approximately)

Related work: [Berntorp, et. al. 2016],[Radhakrishnan, et. al. 2016]

R. S. Laugesen, P. G. Mehta, S. P. Meyn, and M. Raginsky. Poisson Equation in Nonlinear Filtering. SICON, 2015

Gain function $\mathsf{K}_t(x) := \nabla\phi(x)$ where ϕ satisfies

$$\textbf{Poisson equation:} \quad -\frac{1}{\rho}\nabla\cdot(\rho\nabla\phi) = h$$

- ρ is a probability density function
- h is a real-valued function, $\hat{h} = \int h\rho\,\mathrm{d}x$

Poisson equation also appears in:

- Simulation and optimization theory for Markov models [Meyn, Tweedie, 2012]
- Other filtering algorithms [Daum, et. al. 2010]

Problem:

Given: $\{X^1,\ldots,X^N\} \overset{\text{i.i.d}}{\sim} \rho$

Find: $\{\nabla\phi(X^1),\ldots,\nabla\phi(X^N)\}$ (approximately)

Related work: [Berntorp, et. al. 2016],[Radhakrishnan, et. al. 2016]

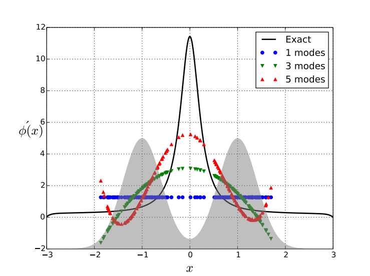
R. S. Laugesen, P. G. Mehta, S. P. Meyn, and M. Raginsky. Poisson Equation in Nonlinear Filtering. SICON, 2015

Method I: Galerkin Approximation

1. Write ϕ as linear combination of basis functions

$$\phi = c_1\psi_1 + \ldots + c_M\psi_M$$

2. Construct an M-dimensional approximation of the Poisson equation
3. Solve the system of M linear equations for $c = [c_1, \ldots, c_M]$



Issues:

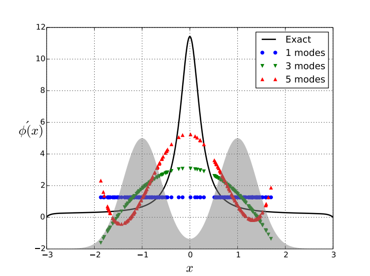
- Choice of basis functions
- Gibbs phenomenon

Method I: Galerkin Approximation

1. Write ϕ as linear combination of basis functions

$$\phi = c_1\psi_1 + \ldots + c_M\psi_M$$

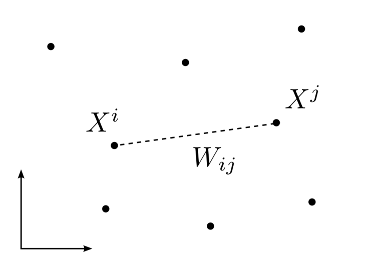
2. Construct an M-dimensional approximation of the Poisson equation
3. Solve the system of M linear equations for $c = [c_1, \ldots, c_M]$



Issues:

- Choice of basis functions
- Gibbs phenomenon

Method II: Kernel Approximation
Basic idea



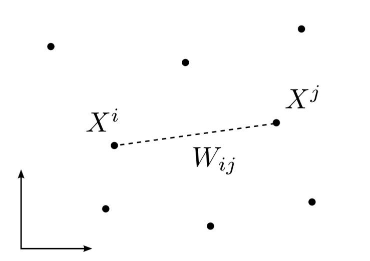
Data: $\{X^1, \ldots, X^N\} \overset{\text{i.i.d}}{\sim} \rho$

Graph Laplacian: $L := \frac{1}{\epsilon}(I - D^{-1}W), \quad W_{ij} = k_\epsilon(X^i, X^j)$

Asymptotics of the graph Laplacian: [Belkin, 2003], [Coifman, Lafon, 2006], [Hein, 2007]

$$L\phi \;\to\; -\frac{1}{\rho}\nabla\cdot(\rho\nabla\phi) \quad \text{as} \quad N\to\infty, \epsilon\to 0$$

Leads to discretization of the Poisson equation



Data: $\{X^1, \ldots, X^N\} \overset{\text{i.i.d}}{\sim} \rho$

Graph Laplacian: $L := \frac{1}{\epsilon}(I - D^{-1}W), \quad W_{ij} = k_\epsilon(X^i, X^j)$

Asymptotics of the graph Laplacian: [Belkin, 2003], [Coifman, Lafon, 2006], [Hein, 2007]

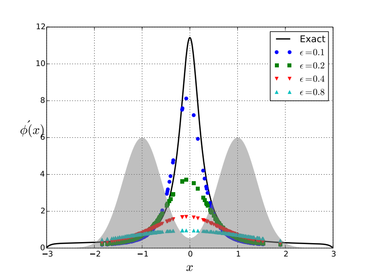
$$L\phi \;\rightarrow\; -\frac{1}{\rho}\nabla \cdot (\rho \nabla \phi) \quad \text{as} \quad N \rightarrow \infty, \epsilon \rightarrow 0$$

Leads to discretization of the Poisson equation

- Form an N-dimensional approximation of the Poisson equation,

$$-\frac{1}{\rho}\nabla\cdot(\rho\nabla\phi) = h \quad \approx \quad \begin{bmatrix} & & \\ & L_{(X^i,X^j)} & \\ & & \end{bmatrix}\begin{bmatrix} \phi_{(X^1)} \\ \vdots \\ \phi_{(X^N)} \end{bmatrix} = \begin{bmatrix} h_{(X^1)} \\ \vdots \\ h_{(X^N)} \end{bmatrix}$$

- Solve for $\{\phi(X^1),\ldots,\phi(X^N)\}$



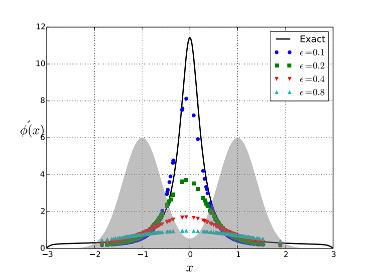
Issue: Computational cost (N is large) $\to$ Use sparsity of L (future work)

A. Taghvaei, and P. G. Mehta, Gain Function Approximation in the Feedback Particle Filter, *CDC*, 2016 (submitted)

- Form an N-dimensional approximation of the Poisson equation,

$$-\frac{1}{\rho}\nabla\cdot(\rho\nabla\phi)=h \quad \approx \quad \begin{bmatrix} L_{(X^i,X^j)} \end{bmatrix}\begin{bmatrix}\phi_{(X^1)}\\ \vdots \\ \phi_{(X^N)}\end{bmatrix}=\begin{bmatrix}h_{(X^1)}\\ \vdots \\ h_{(X^N)}\end{bmatrix}$$

- Solve for $\{\phi(X^1),\dots,\phi(X^N)\}$



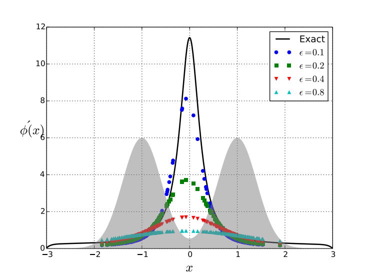
Issue: Computational cost (N is large) → Use sparsity of L (future work)

A. Taghvaei, and P. G. Mehta, Gain Function Approximation in the Feedback Particle Filter, *CDC*, 2016 (submitted)

- Form an N-dimensional approximation of the Poisson equation,

$$-\frac{1}{\rho}\nabla\cdot(\rho\nabla\phi) = h \quad \approx \quad \begin{bmatrix} & & \\ & L_{(X^i,X^j)} & \\ & & \end{bmatrix}\begin{bmatrix}\phi_{(X^1)}\\ \vdots \\ \phi_{(X^N)}\end{bmatrix} = \begin{bmatrix}h_{(X^1)}\\ \vdots \\ h_{(X^N)}\end{bmatrix}$$

- Solve for $\{\phi(X^1),\ldots,\phi(X^N)\}$



Issue: Computational cost (N is large) $\rightarrow$ Use sparsity of L (future work)

A. Taghvaei, and P. G. Mehta, Gain Function Approximation in the Feedback Particle Filter, *CDC*, 2016 (submitted)

Final slide

Thank you!

Method II: Kernel Approximation

Idea behind the proof

Weighted **Heat equation:** $\Delta f(t,x) = \frac{\partial f(t,x)}{\partial t}$ **(I)**, $f(0,x) = \phi(x)$ **(II)**

Weighted **Kernel solution:** $f(t,x) = \int g(t,x,y) f(0,y)\, \mathrm{d}y$ **(III)**

Kernel approximation of $\Delta \phi$

$$\text{(I)} \quad \Rightarrow \quad \Delta f(0,x) \approx \frac{1}{t}(f(t,x) - f(0,x))$$

$$\overset{\text{(III)}}{\Rightarrow} \quad \Delta f(0,x) \approx \frac{1}{t}\Big(\int g(t,x,y) f(0,y)\, \mathrm{d}y - f(0,x)\Big)$$

$$\overset{\text{(II)}}{\Rightarrow} \quad \Delta \phi(x) \approx \frac{1}{t}\Big(\int g(t,x,y) \phi(y)\, \mathrm{d}y - \phi(x)\Big)$$

Empirical approximation $\Delta \phi$

$$\Delta \phi(x) \approx \frac{1}{t}\Big(\frac{1}{N} \sum_{j=1}^{N} g(t,x,X^j) \phi(X^j) - \phi(x)\Big)$$

Weighted **Heat equation:** $\Delta f(t,x) = \frac{\partial f(t,x)}{\partial t}$ **(I)**, $f(0,x) = \phi(x)$ **(II)**

Weighted **Kernel solution:** $f(t,x) = \int g(t,x,y) f(0,y)\,\mathrm{d}y$ **(III)**

Kernel approximation of $\Delta\phi$

$$\text{(I)} \quad \Rightarrow \quad \Delta f(0,x) \approx \frac{1}{t}(f(t,x) - f(0,x))$$

$$\overset{\text{(III)}}{\Rightarrow} \quad \Delta f(0,x) \approx \frac{1}{t}\Big(\int g(t,x,y) f(0,y)\,\mathrm{d}y - f(0,x)\Big)$$

$$\overset{\text{(II)}}{\Rightarrow} \quad \Delta \phi(x) \approx \frac{1}{t}\Big(\int g(t,x,y) \phi(y)\,\mathrm{d}y - \phi(x)\Big)$$

Empirical approximation $\Delta\phi$

$$\Delta\phi(x) \approx \frac{1}{t}\Big(\frac{1}{N}\sum_{j=1}^{N} g(t,x,X^j)\phi(X^j) - \phi(x)\Big)$$

Method II: Kernel Approximation

Idea behind the proof

Weighted Heat equation: $\Delta_\rho f(t,x) = \dfrac{\partial f(t,x)}{\partial t}$ **(I)**, $\quad f(0,x) = \phi(x)$ **(II)**

Weighted Kernel solution: $f(t,x) \approx \int g_\rho(t,x,y) f(0,y) \rho(y)\, \mathrm{d}y$ **(III)**

Kernel approximation of $\Delta_\rho \phi$

$$\text{(I)} \quad \Rightarrow \quad \Delta_\rho f(0,x) \approx \frac{1}{t}\big(f(t,x) - f(0,x)\big)$$

$$\overset{\text{(III)}}{\Rightarrow} \quad \Delta_\rho f(0,x) \approx \frac{1}{t}\Big(\int g_\rho(t,x,y) f(0,y)\rho(y)\,\mathrm{d}y - f(0,x)\Big)$$

$$\overset{\text{(II)}}{\Rightarrow} \quad \Delta_\rho \phi(x) \approx \frac{1}{t}\Big(\int g_\rho(t,x,y)\phi(y)\rho(y)\,\mathrm{d}y - \phi(x)\Big)$$

Empirical approximation $\Delta_\rho \phi$

$$\Delta_\rho \phi(x) \approx \frac{1}{t}\Big(\frac{1}{N}\sum_{j=1}^{N} g_\rho(t,x,X^j)\phi(X^j) - \phi(x)\Big)$$

Weighted Heat equation: $\Delta_\rho f(t,x) = \dfrac{\partial f(t,x)}{\partial t}$ **(I)**, $f(0,x) = \phi(x)$ **(II)**

Weighted Kernel solution: $f(t,x) \approx \int g_\rho(t,x,y) f(0,y)\rho(y)\,\mathrm{d}y$ **(III)**

Kernel approximation of $\Delta_\rho \phi$

$$\text{(I)} \quad \Rightarrow \quad \Delta_\rho f(0,x) \approx \frac{1}{t}(f(t,x) - f(0,x))$$

$$\overset{\text{(III)}}{\Rightarrow} \quad \Delta_\rho f(0,x) \approx \frac{1}{t}(\int g_\rho(t,x,y) f(0,y)\rho(y)\,\mathrm{d}y - f(0,x))$$

$$\overset{\text{(II)}}{\Rightarrow} \quad \Delta_\rho \phi(x) \approx \frac{1}{t}(\int g_\rho(t,x,y)\phi(y)\rho(y)\,\mathrm{d}y - \phi(x))$$

Empirical approximation $\Delta_\rho \phi$

$$\Delta_\rho \phi(x) \approx \frac{1}{t}(\frac{1}{N}\sum_{j=1}^{N} g_\rho(t,x,X^j)\phi(X^j) - \phi(x))$$

Numerical Results