UTOPIÆ Uncertainty Treatment and Optimisation in Aerospace Engineering

THE FRAMEWORK PROGRAMME FOR RESEARCH AND INNOVATION HORIZON 2020

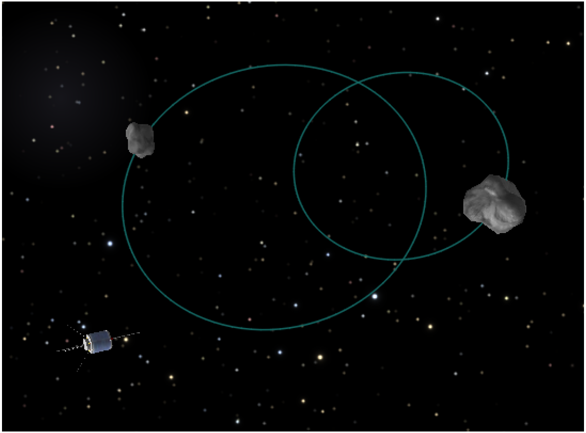
MARIE CURIE ACTIONS

Imprecise Filtering for Spacecraft Navigation

Tathagata Basu Cristian Greco Thomas Krak

Durham University Strathclyde University Ghent University

Filtering for Spacecraft Navigation



The General Problem

Dynamical system with state $x_t \in \mathcal{X}$, $t \geq 0$

- For simple orbital mechanics: $\mathcal{X} \simeq [0, 1]^6$

General statement given by (stochastic) differential equation

$$\frac{\mathrm{d}\, x_t}{\mathrm{d}\, t} = \underbrace{f(x_t, t \mid \theta)}_{\text{system dynamics}} + \underbrace{\frac{\mathrm{d}\, W_t}{\mathrm{d}\, t}}_{\text{Brownian motion noise}} \times \underbrace{\sigma(x_t, t \mid \theta)}_{\text{diffusion coefficient}}$$

with stochastic initial condition $x_0 \sim P(X_0 \mid \theta)$.

Collect all parameters in θ

For given θ induces a stochastic process $\{X_t|\theta\}_{t\geq 0}$

- This process is a *Markov chain*

The General Problem

Realisations of $\{X_t|\theta\}_{t\geq 0}$ cannot be observed directly

- I.e. it is a *hidden* Markov chain

But, *noisy measurement model*:

$$Y_t \sim \psi(\cdot \,|\, x_t, \theta)$$

(Just writing θ for the parameters again)

Filtering Problem

Given collection $y_{t_{1:n}} \equiv y_{t_1}, \ldots, y_{t_n}$, compute conditional probability

$$P_\theta(X_t \,|\, y_{t_{1:n}})$$

More generally: conditional expectation of function $h : \mathcal{X} \to \mathbb{R}$:

$$\mathbb{E}_\theta\left[h(X_t) \,\middle|\, y_{t_{1:n}}\right]$$

Imprecise Filtering

What if we don't know θ exactly?

If we can assume it belongs to a *set* ϑ:

- Induces an *imprecise* stochastic process
    - I.e. a *set* of stochastic processes

Imprecise Filtering Problem

Compute *lower*- and *upper* expectation

$$\underline{\mathbb{E}}\left[h(X_t) \,\middle|\, y_{t_{1:n}}\right] = \inf_{\theta \in \vartheta} \mathbb{E}_\theta\left[h(X_t) \,\middle|\, y_{t_{1:n}}\right]$$

$$\overline{\mathbb{E}}\left[h(X_t) \,\middle|\, y_{t_{1:n}}\right] = \sup_{\theta \in \vartheta} \mathbb{E}_\theta\left[h(X_t) \,\middle|\, y_{t_{1:n}}\right]$$

$\Rightarrow$ provide robust bounds on quantity of interest

Some Immediate Difficulties

- Even for fixed θ, no analytical solution for $\{X_t|\theta\}_{t\geq 0}$
  - Let alone for $\mathbb{E}_\theta\left[h(X_t)\,\middle|\,y_{t_{1:n}}\right]$
- Convergence of numerical approaches?
  - Uniformly w.r.t ϑ?
- The optimisation to compute $\underline{\mathbb{E}}$ or $\overline{\mathbb{E}}$ is difficult
  - Multi-modal objective surface, ...
- Semantic issues in problem statement
  - Physical interpretation of certain relaxations,...

Some Existing Approaches

Precise case:

- Kalman Filter (/variants)
  - Gaussian distr., linear dyn. $\Rightarrow P(X_t | y_{t_{1:n}})$ is Gaussian
  - *Extended* Kalman Filter: Just re-linearise over time
- Particle Filters
  - Essentially a Monte Carlo method

Imprecise case:

- Special cases: propagation of p-boxes, set-valued observations, ...
- "Imprecise Kalman Filter" [1]
  - Provides general solution...
  - ...but so far only solved in specific cases
    - E.g. Linear Gaussian-Vacuous Mixture

[1] A. Benavoli, M. Zaffalon, E. Miranda: *Robust filtering through coherent lower previsions*, IEEE Transactions on Automatic Control, 2011

The Formal Setup

Basically we just need to apply Bayes' rule:

$$\mathbb{E}\big[h(X_t)\,|\,y_{t_{1:n}}\big] = \frac{\mathbb{E}\big[h(X_t)\prod_{i=1}^{n}\psi(y_{t_i}\,|\,X_{t_i})\big]}{\mathbb{E}\big[\prod_{i=1}^{n}\psi(y_{t_i}\,|\,X_{t_i})\big]}$$

In the imprecise case (let's say $\underline{\mathbb{E}}\big[\prod_{i=1}^{n}\mathbb{I}_{\{y_{t_i}\}}(Y_{t_i})\big] > 0$):

$$\underline{\mathbb{E}}\big[h(X_t)\,|\,y_{t_{1:n}}\big] = \mu \Leftrightarrow \underline{\mathbb{E}}\left[\big(h(X_t)-\mu\big)\prod_{i=1}^{n}\mathbb{I}_{\{y_{t_i}\}}(Y_{t_i})\right] = 0$$

So, we need to compute a joint expectation.

Iterated (Lower) Expectation

Write $g(X_{t_{1:n}}, X_t, Y_{t_{1:n}}) = \big(h(X_t) - \mu\big) \prod_{i=1}^{n} \mathbb{I}_{\{y_{t_i}\}}(Y_{t_i})$

Use epistemic irrelevance ⇒ "imprecise Markov property"

Recursively decompose the joint:

$$\begin{aligned}
\underline{\mathbb{E}}[g(\cdot)] &= \underline{\mathbb{E}}[\underline{\mathbb{E}}[g(\cdot) \mid X_0]] \\
&= \underline{\mathbb{E}}[\underline{\mathbb{E}}[\underbrace{\underline{\mathbb{E}}[g(\cdot) \mid X_0, X_{t_1}]}_{=\underline{\mathbb{E}}[g(\cdot) \mid X_{t_1}]} \mid X_0]] \\
&= \underline{\mathbb{E}}[\underline{\mathbb{E}}[\underline{\mathbb{E}}[g(\cdot) \mid X_{t_1}] \mid X_0]] \\
&= \underline{\mathbb{E}}[\underline{\mathbb{E}}[\underline{\mathbb{E}}[\underline{\mathbb{E}}[g(\cdot) \mid Y_{t_1}, X_{t_1}] \mid X_{t_1}] \mid X_0]] \\
&\vdots \\
&= \underline{\mathbb{E}}[\underline{\mathbb{E}}[\underline{\mathbb{E}}[\underline{\mathbb{E}}[\cdots \underline{\mathbb{E}}[\underline{\mathbb{E}}[g(\cdot) \mid X_{t_n}, Y_{t_n}] \mid X_{t_n}] \cdots \mid Y_{t_1}, X_{t_1}] \mid X_{t_1}] \mid X_0]]
\end{aligned}$$

Iterated (Lower) Expectation

Now resolve the remaining conditionals "inward-outward":

$$\underline{\mathbb{E}}\big[\underline{\mathbb{E}}\big[\underline{\mathbb{E}}\big[\underline{\mathbb{E}}\big[\cdots\underline{\mathbb{E}}\big[\underbrace{\underline{\mathbb{E}}\big[g(\cdot)\,|\,X_{t_n}, Y_{t_n}\big]}_{\text{start here}}\,|\,X_{t_n}\big]\cdots\,|\,Y_{t_1}, X_{t_1}\big]\,|\,X_{t_1}\big]\,|\,X_0\big]\big]$$

Fix all variables except X_t:

$$\begin{aligned}
&\underline{\mathbb{E}}\left[\left(h(X_t)-\mu\right)\prod_{i=1}^{n}\mathbb{I}_{y_{t_i}}(\gamma_{t_i})\,\middle|\,X_{t_n}=x_{t_n}, Y_{t_n}=\gamma_{t_n}\right]\\
&=\underline{\mathbb{E}}\left[\left(h(X_t)-\mu\right)|\,X_{t_n}=x_{t_n}, Y_{t_n}=\gamma_{t_n}\right]\prod_{i=1}^{n}\mathbb{I}_{y_{t_i}}(\gamma_{t_i})\\
&=\underline{\mathbb{E}}\left[\left(h(X_t)-\mu\right)|\,X_{t_n}=x_{t_n}\right]\prod_{i=1}^{n}\mathbb{I}_{y_{t_i}}(\gamma_{t_i})\\
&=: g'(\cdot)
\end{aligned}$$

Iterated (Lower) Expectation

We have a new function g' that no longer depends on X_t.
After substitution

$$\underline{\mathbb{E}}\Big[\underline{\mathbb{E}}\Big[\underline{\mathbb{E}}\Big[\underline{\mathbb{E}}\Big[\cdots \underbrace{\underline{\mathbb{E}}\big[g'(\cdot)\,\big|\,X_{t_n}\big]}_{\text{this part next}} \cdots \,\Big|\, Y_{t_1}, X_{t_1}\Big] \,\Big|\, X_{t_1}\Big] \,\Big|\, X_0\Big]\Big]$$

Fix all variables except Y_{t_n},

$$\begin{aligned}
&\underline{\mathbb{E}}\big[g'(\cdot)\,\big|\,X_{t_n} = x_{t_n}\big] \\
&= \underline{\mathbb{E}}\Big[\underline{\mathbb{E}}\big[(h(X_t) - \mu)|X_{t_n} = x_{t_n}\big]\, \mathbb{I}_{y_{t_n}}(Y_{t_n}) \textstyle\prod_{i=1}^{n-1} \mathbb{I}_{y_{t_i}}(\gamma_{t_i}) \Big| X_{t_n} = x_{t_n}\Big] \\
&= \textstyle\prod_{i=1}^{n-1} \mathbb{I}_{y_{t_i}}(\gamma_{t_i}) \underline{\mathbb{E}}\big[(h(X_t) - \mu)|X_{t_n} = x_{t_n}\big] \\
&\quad \times \begin{cases} \underline{\mathbb{E}}\big[\mathbb{I}_{y_{t_n}}(Y_{t_n})|X_{t_n} = x_{t_n}\big] & \text{if } \underline{\mathbb{E}}\big[(h(X_t) - \mu)|X_{t_n} = x_{t_n}\big] \geq 0 \\ \overline{\mathbb{E}}\big[\mathbb{I}_{y_{t_n}}(Y_{t_n})|X_{t_n} = x_{t_n}\big] & \text{otherwise} \end{cases} \\
&=: g''(\cdot)
\end{aligned}$$

Iterated (Lower) Expectation

We have

$$\underline{\mathbb{E}}\big[\underline{\mathbb{E}}\big[\underline{\mathbb{E}}\big[\underline{\mathbb{E}}\big[\cdots g''(\cdot)\cdots \mid Y_{t_1}, X_{t_1}\big] \mid X_{t_1}\big] \mid X_0\big]\big]$$

$g''(\cdot)$ depends on $X_{t_1}, \ldots, X_{t_{n-1}}$, X_{t_n} and $Y_{t_1}, \ldots, X_{t_{n-1}}$

$\Rightarrow$ This is what we started with, but reduced

Now just recurse

$$\underline{\mathbb{E}}\big[\underbrace{\underline{\mathbb{E}}\big[\underline{\mathbb{E}}\big[\underline{\mathbb{E}}\big[\cdots g''(\cdot)\cdots \mid Y_{t_1}, X_{t_1}\big] \mid X_{t_1}\big] \mid X_0\big]}_{=:g^*(\cdot)}\big]$$

$$= \underline{\mathbb{E}}\big[g^*(X_0)\big]$$

Remaining Problem, in Summary

Lower- and upper likelihoods of states given measurements:

- $\underline{\mathbb{E}}\big[\mathbb{I}_{y_t}(Y_t)|X_t = x_t\big]$ and $\overline{\mathbb{E}}\big[\mathbb{I}_{y_t}(Y_t)|X_t = x_t\big]$

For arbitrary $h : \mathcal{X} \to \mathbb{R}$, need to evaluate

- $\underline{\mathbb{E}}\big[h(X_0)\big]$ (okay)
- $\underline{\mathbb{E}}\big[h(X_{t_n})\,|\,X_{t_{n-1}}\big]$ (not okay)

Some Simplifying Assumptions

Assume dynamics are *precise* and *time-homogeneous*:

$$\frac{\mathrm{d}\, x_t}{\mathrm{d}\, t} = f(x_t) + \frac{\mathrm{d}\, W_t}{\mathrm{d}\, t} \times \sigma(x_t)$$

Let $\mathcal{X} \simeq [0,1]^d$ and $h_t : x \mapsto \mathbb{E}[h(X_t) \,|\, X_0 = x]$. Then

$$\frac{\mathrm{d}\, h_t(x)}{\mathrm{d}\, t} = \sum_{i=1}^{d} f_i(x) \frac{\partial\, h_t(x)}{\partial\, x_i} + \frac{1}{2} \sum_{i=1}^{d} \sum_{j=1}^{d} (\sigma(x)\sigma(x)^\top)_{ij} \frac{\partial^2\, h_t(x)}{\partial x_i \partial x_j}$$

with $h_0 = h$.

To get $h_t(x) = \mathbb{E}[h(X_t) \,|\, X_0 = x]$ we need to solve this PDE.

In general no analytical solution.

Some Further Simplifying Assumptions

Assume the dynamics are also *deterministic*

- I.e. the process is degenerate and described by $f(x_t)$
- We keep stochastic (and imprecise) initial distribution and measurements

We get

$$\frac{\mathrm{d}\, h_t(x)}{\mathrm{d}\, t} = \sum_{i=1}^{d} f_i(x) \frac{\partial\, h_t(x)}{\partial\, x_i}$$

with $h_0 = h$.

Still need to solve a PDE, and still no analytical solution

We're going to approximate this

Approximate Solution

Easy enough to solve pointwise: for any $x_0 \in \mathcal{X}$,

$$\mathbb{E}[h(X_t) \,|\, X_0 = x_0] = h(x_t)\,,$$

with

$$x_t = x_0 + \int_0^t f(x_\tau)\,\mathrm{d}\tau$$

$\Rightarrow$ Just do this for a lot of different x_0

Then extend pointwise estimates to entire domain $\mathcal{X}$

Dynamics Propagation

Collection of starting points $\mathbf{Z} = z_1, \ldots, z_m$

Compute $\mathbf{Z}^* = z_1^{(t)}, \ldots, z_m^{(t)}$, with

$$z_i^{(t)} = z_i + \int_0^t f\left(z_i^{(\tau)}\right) \mathrm{d}\tau, \quad i = 1, \ldots, m$$

This is just an ODE $\Rightarrow$ use your favourite numerical integrator

- I.e. Explicit/implicit Euler, Runge-Kutta schemes,...

Note: this could be "embarrassingly parallelized"

- A modern GPU can solve hundreds to thousands in parallel

Propagated Dynamics

We now know

$$h_t(z_i) = h\left(x_i^{(t)}\right) = \mathbb{E}[h(X_t) \mid X_0 = z_i] \quad \text{for } i = 1, \ldots, m$$

Now extend this to $\hat{h}_t$ on entire $\mathcal{X}$

Radial Basis Function Interpolation

Training points $\mathbf{Z} = z_1, \ldots, z_m \in \mathcal{X}$.

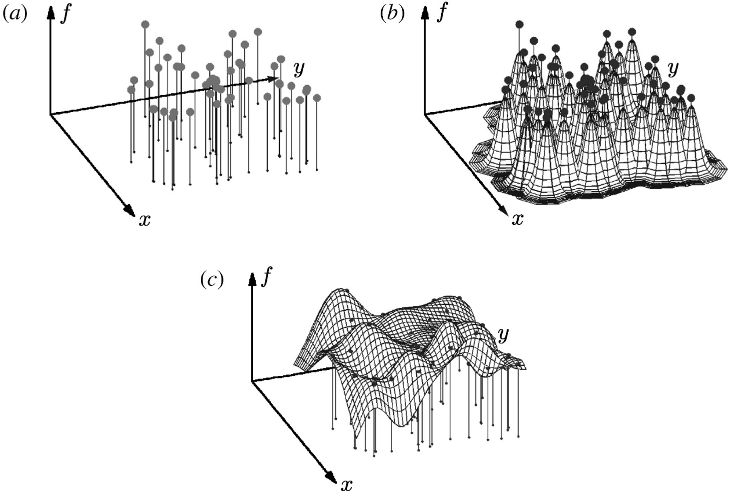
Radial basis functions $\phi : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$, $\phi(z, c) = \phi(\|z - c\|)$

Basis expansion $\Phi_z \in \mathbb{R}^m$ of any $z \in \mathcal{X}$:

$$\Phi_z = [\phi(z, z_1) \cdots \phi(z, z_m)]^\top$$

Can find $\mathbf{w}_h \in \mathbb{R}^m$ to approximate any ("nice") function $h : \mathcal{X} \to \mathbb{R}$:

$$h(z) \approx \hat{h}(z) = \sum_{i=1}^{m} w_i \phi(z, z_i) = \Phi_z^\top \mathbf{w}_h$$



N. Flyer, G.B. Wright: *A radial basis function method for the shallow water equations on a sphere*, Proceedings Royal Soc A, 2009

Radial Basis Function Interpolation

Construct design(/kernel/feature/...) matrix

$$\mathbf{\Phi}_\mathbf{Z} = \begin{bmatrix} \Phi_{z_1} \\ \vdots \\ \Phi_{z_m} \end{bmatrix} = \begin{bmatrix} \phi(z_1, z_1) & \cdots & \phi(z_1, z_m) \\ \vdots & \ddots & \vdots \\ \phi(z_m, z_1) & \cdots & \phi(z_m, z_m) \end{bmatrix}$$

Then $\mathbf{\Phi}_\mathbf{Z}$ is usually invertible if all $z_1, \ldots, z_m$ are unique.

Let $h(\mathbf{Z}) = [h(z_1) \cdots h(z_m)]^\top$

Find $\mathbf{w}_h$ that interpolates h on $\mathbf{Z}$:

$$\mathbf{\Phi}_\mathbf{Z}\mathbf{w}_h = h(\mathbf{Z})$$

In other words, we get $\mathbf{w}_h = \mathbf{\Phi}_\mathbf{Z}^{-1} h(\mathbf{Z})$

Radial Basis Function Interpolation

Pros:

- Straightforward (conceptually and to implement)
  - Meshless (i.e. arbitrary $\mathbf{Z}$), so no need to partition $\mathcal{X}$
  - Just solve a linear system
- Well-developed theory
- Good performance for approximating PDE solutions
  - Near spectral (exponential) convergence
- Many parameters to fine tune performance
  - Basis functions, hyperparameters, regularisation, ...

Cons:

- Many parameters to fine tune
- Numerical instability (system is typically ill-conditioned)

Approximation in Summary

We want to compute $\hat{h}_t(X_0) \approx h_t(X_0) := \mathbb{E}[h(X_t) \mid X_0]$

We know that $h_t(x_0) = h(x_t)$, where

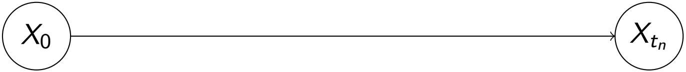
$$x_t = x_0 + \int_0^t f(x_\tau)\,\mathrm{d}\tau$$

We proceed as follows

1. Choose some $\mathbf{Z}$
2. Compute $\mathbf{Z}^*$ for time t
3. Compute design matrix $\mathbf{\Phi}_{\mathbf{Z}}$
4. Compute $\mathbf{w}_{h_t} = \mathbf{\Phi}_{\mathbf{Z}}^{-1} h(\mathbf{Z}^*)$
5. Let $\hat{h}_t(x) := \Phi_x^\top \mathbf{w}_{h_t}$ for any $x \in \mathcal{X}$

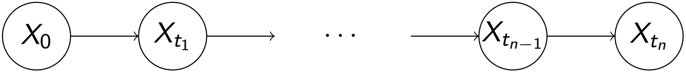
Iterated Expectation, Again

$$\mathbb{E}[h(X_{t_n})] = \mathbb{E}\big[\mathbb{E}\big[h(X_{t_n})|X_0\big]\big]$$



Iterated Expectation, Again

$$\mathbb{E}[h(X_{t_n})] = \mathbb{E}\Big[\mathbb{E}\Big[\mathbb{E}\Big[\cdots \mathbb{E}[h(X_{t_n})|X_{t_{n-1}}] \cdots |X_{t_1}\Big]|X_0\Big]\Big]$$

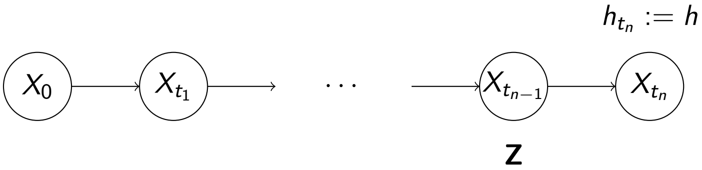


Iterated Expectation, Again

$$\mathbb{E}[h(X_{t_n})] = \mathbb{E}\Big[\mathbb{E}\Big[\mathbb{E}\Big[\cdots \mathbb{E}[h(X_{t_n})|X_{t_{n-1}}]\cdots |X_{t_1}\Big]|X_0\Big]\Big]$$

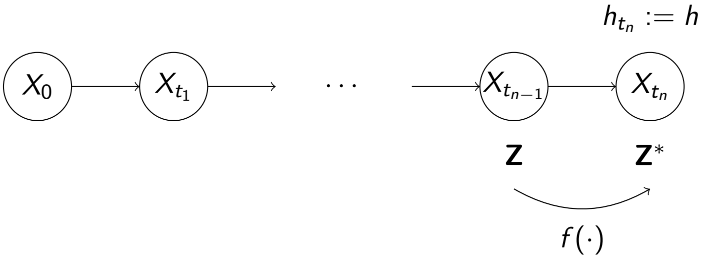
Iterated Expectation, Again

$$\mathbb{E}[h(X_{t_n})] = \mathbb{E}\Big[\mathbb{E}\Big[\mathbb{E}\Big[\cdots \mathbb{E}[h(X_{t_n})|X_{t_{n-1}}]\cdots |X_{t_1}\Big]|X_0\Big]\Big]$$



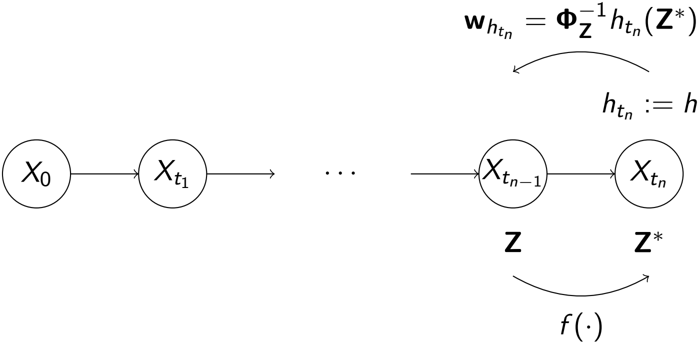
Iterated Expectation, Again

$$\mathbb{E}[h(X_{t_n})] = \mathbb{E}\Big[\mathbb{E}\Big[\mathbb{E}\Big[\cdots \mathbb{E}[h(X_{t_n})|X_{t_{n-1}}]\cdots|X_{t_1}\Big]|X_0\Big]\Big]$$



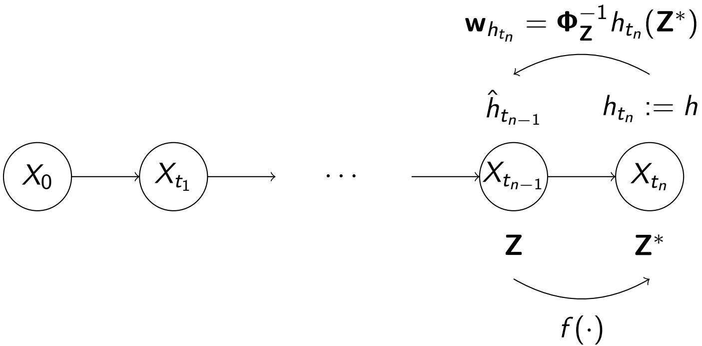
Iterated Expectation, Again

$$\mathbb{E}[h(X_{t_n})] = \mathbb{E}\Big[\mathbb{E}\Big[\mathbb{E}\Big[\cdots \mathbb{E}[h(X_{t_n})|X_{t_{n-1}}]\cdots |X_{t_1}\Big]|X_0\Big]\Big]$$



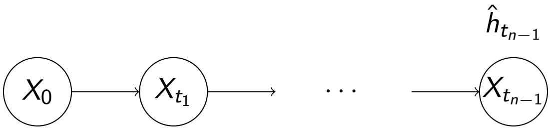
Iterated Expectation, Again

$$\mathbb{E}[h(X_{t_n})] = \mathbb{E}\Big[\mathbb{E}\Big[\mathbb{E}\Big[\cdots \mathbb{E}[h(X_{t_n})|X_{t_{n-1}}]\cdots|X_{t_1}\Big]|X_0\Big]\Big]$$



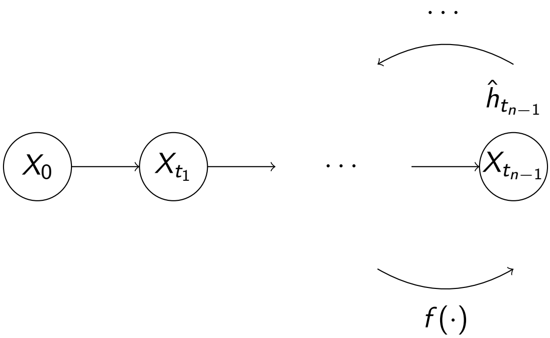
Iterated Expectation, Again

$$\mathbb{E}[h(X_{t_n})] \approx \mathbb{E}\left[\mathbb{E}\left[\mathbb{E}\left[\cdots \hat{h}_{t_{n-1}}(X_{t_{n-1}}) \cdots | X_{t_1}\right] | X_0\right]\right]$$



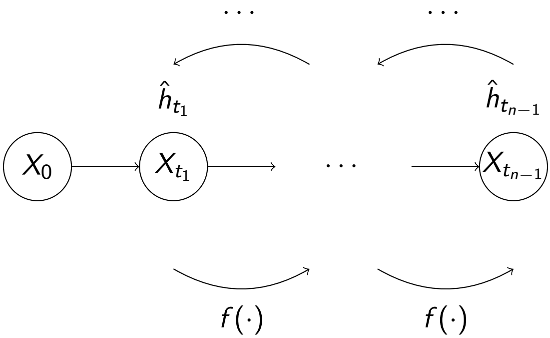
Iterated Expectation, Again

$$\mathbb{E}[h(X_{t_n})] \approx \mathbb{E}\left[\mathbb{E}\left[\mathbb{E}\left[\cdots \hat{h}_{t_{n-1}}(X_{t_{n-1}}) \cdots | X_{t_1}\right] | X_0\right]\right]$$



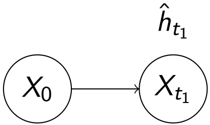
Iterated Expectation, Again

$$\mathbb{E}[h(X_{t_n})] \approx \mathbb{E}\left[\mathbb{E}\left[\mathbb{E}\left[\cdots \hat{h}_{t_{n-1}}(X_{t_{n-1}}) \cdots |X_{t_1}\right]|X_0\right]\right]$$



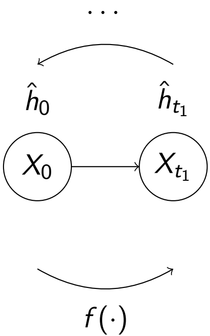
Iterated Expectation, Again

$$\mathbb{E}[h(X_{t_n})] \approx \mathbb{E}\left[\mathbb{E}\left[\hat{h}_{t_1}(X_{t_1})|X_0\right]\right]$$



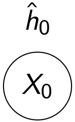
Iterated Expectation, Again

$$\mathbb{E}[h(X_{t_n})] \approx \mathbb{E}\left[\mathbb{E}\left[\hat{h}_{t_1}(X_{t_1})|X_0\right]\right]$$



Iterated Expectation, Again

$$\mathbb{E}[h(X_{t_n})] \approx \mathbb{E}\left[\hat{h}_0(X_0)\right]$$



Iterated Expectation, Again

$$\mathbb{E}[h(X_{t_n})] \approx \mathbb{E}\left[\hat{h}_0(X_0)\right] = \int_{\mathcal{X}} \hat{h}_0(x_0)\,\mathrm{d}P(X_0 = x_0)$$

The Crucial Part

How to make this tractable:

1. Make $0 =: t_0, t_1, \ldots, t_n$ a *uniform* partition of $[0, t]$
   - So $t_i - t_{i-1} = \Delta$ for all $i = 1, \ldots, n$
2. Choose the same $\mathbf{Z}$ at every step

The Crucial Part

- $\mathbf{Z}^*$ is the same at each step (by homogeneity)
- You can precompute $\mathbf{Z}$, $\mathbf{Z}^*$, $\boldsymbol{\Phi}_{\mathbf{Z}}$, $\boldsymbol{\Phi}_{\mathbf{Z}^*}$, $\boldsymbol{\Phi}_{\mathbf{Z}}^{-1}$, $(\boldsymbol{\Phi}_{\mathbf{Z}^*}\boldsymbol{\Phi}_{\mathbf{Z}}^{-1}) =: \mathbf{K}$
    - Only *once*, *offline*, and $\mathcal{O}(\Delta m) + \mathcal{O}(m^3)$

The Crucial Part

- $\mathbf{Z}^*$ is the same at each step (by homogeneity)
- You can precompute $\mathbf{Z}$, $\mathbf{Z}^*$, $\mathbf{\Phi_Z}$, $\mathbf{\Phi_{Z^*}}$, $\mathbf{\Phi_Z^{-1}}$, $(\mathbf{\Phi_{Z^*}\Phi_Z^{-1}}) =: \mathbf{K}$
  - Only *once*, *offline*, and $\mathcal{O}(\Delta m) + \mathcal{O}(m^3)$
- Remains to solve the recursion:

$$
\begin{aligned}
\hat{h}_{t_n}(x) &= h(x) \\
\hat{h}_{t_{n-1}}(x) &= \Phi_x^\top \mathbf{w}_{\hat{h}_{t_n}} = \Phi_x^\top \mathbf{\Phi_Z^{-1}} h(\mathbf{Z}^*) = \Phi_x^\top \mathbf{\Phi_Z^{-1}} \mathbf{K}^0 h(\mathbf{Z}^*) \\
\hat{h}_{t_{n-2}}(x) &= \Phi_x^\top \mathbf{w}_{\hat{h}_{t_{n-1}}} = \Phi_x^\top \mathbf{\Phi_Z^{-1}} \mathbf{\Phi_{Z^*}} \mathbf{\Phi_Z^{-1}} h(\mathbf{Z}^*) = \Phi_x^\top \mathbf{\Phi_Z^{-1}} \mathbf{K}^1 h(\mathbf{Z}^*) \\
&\vdots \\
\hat{h}_{t_{n-\ell}}(x) &= \Phi_x^\top \mathbf{\Phi_Z^{-1}} \mathbf{K}^{\ell-1} h(\mathbf{Z}^*) \quad \text{for } \ell = 1, \dots, n
\end{aligned}
$$

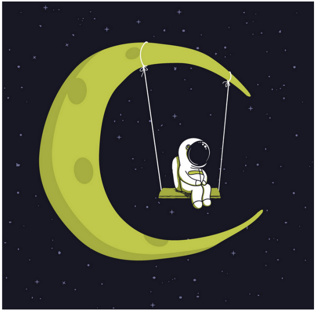
After precomputing, can find $\hat{h}_0$ for *any* h in $\mathcal{O}(nm^2)$

But does it work?

Dynamics of a Spacecraft

https://www.vectorstock.com/royalty-free-vectors/vectors-by_mirquurius

Dynamics of a ~~Spacecraft~~ Pendulum



https://www.vectorstock.com/royalty-free-vectors/vectors-by_mirquurius

Pendulum Setup

Pendulum dynamics model:

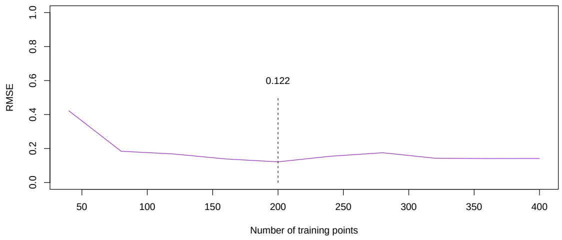
$$\frac{\mathrm{d}\,x_t}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}\alpha_t \\ \dot{\alpha}_t\end{bmatrix} = \begin{bmatrix}\dot{\alpha}_t \\ {}^{g}\!/\!{}_{\ell} \sin \alpha_t - \rho\dot{\alpha}_t\end{bmatrix}$$

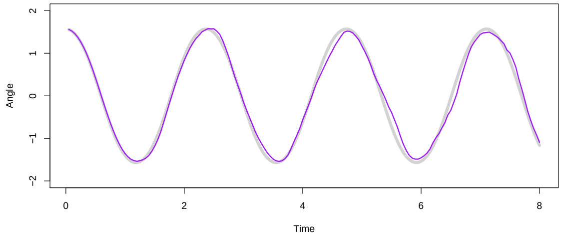
With $\ell = 1, g \approx -9.8$ and $\rho \in \{0, 0.5\}$.

Known true initial position $x_0 = [{}^{\pi}\!/\!{}_{2}\ 0]^\top$

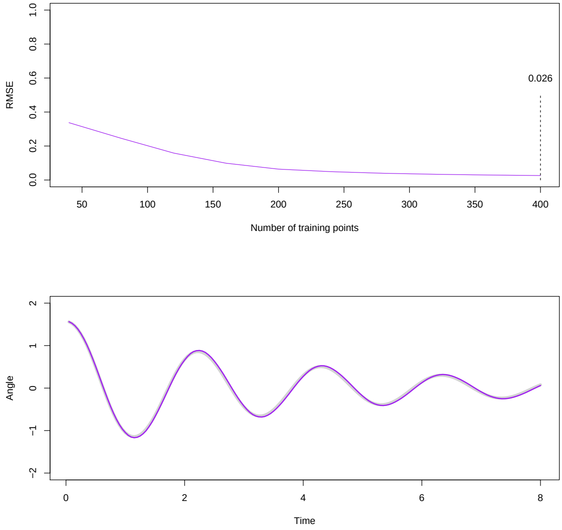
Sequential reinitialisation with best guess after 1 second

No Friction, Known Start, No Measurements





With Friction, Known Start, No Measurements



Pendulum Setup, With Measurements

Model thinks friction term $\rho = 0$

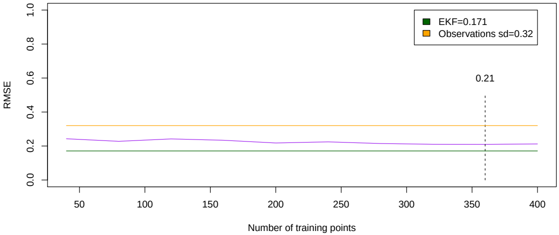
Measurement model:

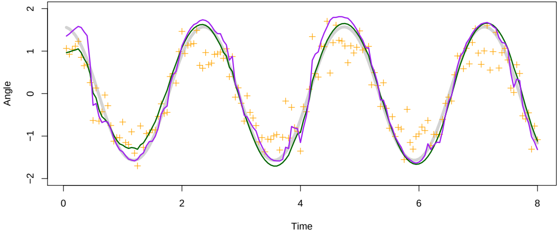
$$Y_t \sim \mathcal{N}\left(\cdot \mid \ell \sin \alpha_t, 0.32^2\right)$$

Unknown initial position with uniform $P(X_0) = U(-2, 2) \times U(-5, 5)$

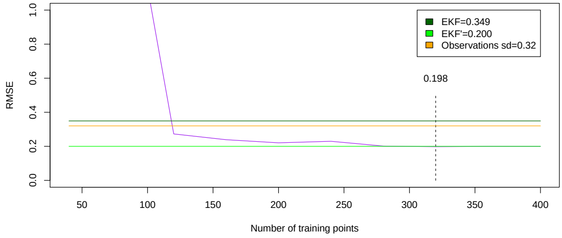
Sequential reinitialisation with same uniform after 1 second

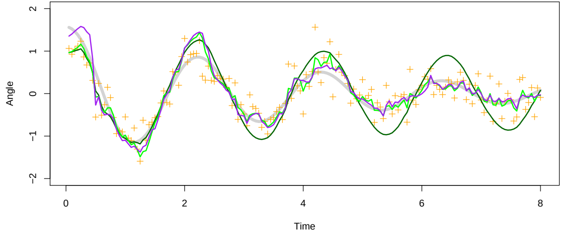
No Friction, Uniform Start, 20 Measurements/s





Friction, Uniform Start, 20 Measurements/s





Pendulum Setup, Imprecise

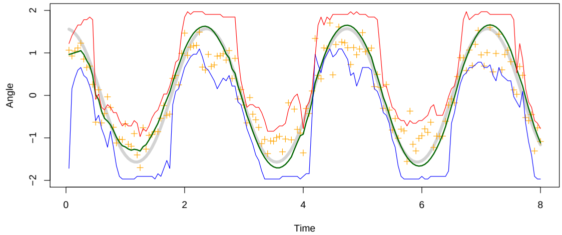
Imprecise measurement model with lower- and upper likelihood:

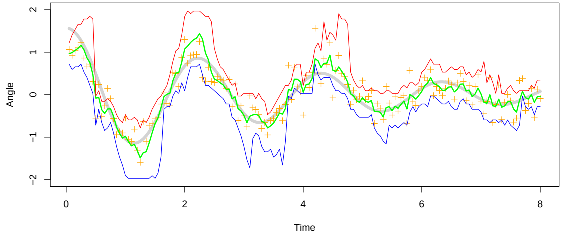
$$\underline{\psi}(y_t \mid x_t) = 0.5\mathcal{N}\left(y_t \mid \ell \sin \alpha_t, 0.32^2\right)$$
$$\overline{\psi}(y_t \mid x_t) = 1.5\mathcal{N}\left(y_t \mid \ell \sin \alpha_t, 0.32^2\right)$$

Unknown initial position with *vacuous* on $[-2, 2] \times [-5, 5]$

Sequential reinitialisation with same vacuous after 1 second

Vacuous Start, 20 Measurements/s, Imprecise Measurement Models





Open Questions

- Does this scale??
- Numerical stability issues?
    - Preconditioning
    - Smart selection/pruning of training points?
    - E.g. regularisation, LASSO
- Generalisation to stochastic (and/or imprecise) dynamics?
- Find informative "sequential prior"?
- Interpretation of $\mathbf{K} = \mathbf{\Phi}_{\mathbf{Z}^*}\mathbf{\Phi}_{\mathbf{Z}}^{-1}$?