Predictor-Corrector and Morphing Ensemble Filters for the Assimilation of Sparse Data into High-Dimensional Nonlinear Systems

Jan Mandel and Jonathan D. Beezley
University of Colorado
and National Center for Atmospheric Research

American Meteorological Society Annual Meeting
San Antonio, TX, January 16, 2007

Supported by National Science Foundation Grant CNS-0325314 and NCAR faculty fellowship.

Data Assimilation for Highly Nonlinear Problems

- Coherent features: fireline in wildfire, hurricane vortex
- Error in position of feature typically have close to Gaussian distribution
- Error in physical fields at a fixed point typically do not - multimodal
  - wildfire: concentrated about burning/not burning
  - hurricane: which side of the vortex are we on?
- The closer to Gaussian the better $\Rightarrow$
  transformation of variables to get closer to Gaussian: Morphing Filters
- EnKF based on the Gaussian assumption $\Rightarrow$
  combine with particle filters: Predictor-Corrector Filters

1. Morphing Ensemble Filters

Data Assimilation and Additive vs Positional Correction

- alternative error models including the position of features (Hoffman et al., 1995)

- additive correction to spatial transformation instead of original variables
  - global low order polynomial mapping for alignment (Alexander et al., 1998)
  - alignment as preprocessing to an additive correction (Lawson and Hansen, 2005; Ravela et al., 2006)

- **New Morphing Filter: a one-step method**
  - additive correction to spatial transformation **and** variable values
  - by automatic image registration, borrowed from image processing

Intermediate States by Morphing

Registration Given two functions u_0 and u_1, find transformation U such that

$$\|u_1 - u_0 \circ (I + T)\| + C_1\|T\| + C_2\| \bigtriangledown T\| \to \text{min}\,.$$

for suitable norms and constants (Gao, 1998), with modifications to speed up and decrease the chances of getting stuck in a local minimum.

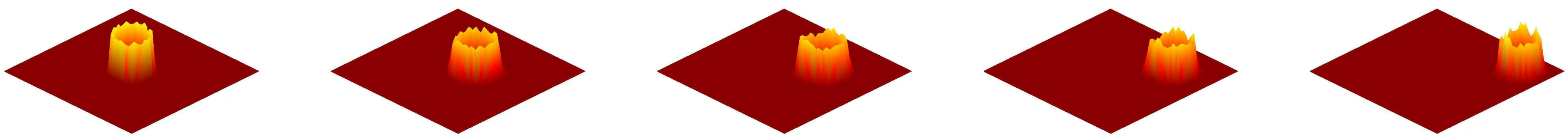
Create intermediate functions u_λ between u_0 and u_1, by

$$u_\lambda = (u_0 + \lambda r) \circ (I + \lambda T)\,, \quad 0 \leq \lambda \leq 1, \quad r = u_1 \circ (I + T)^{-1} - u_0$$

Morphing Ensemble Kalman Filter The EnKF is least squares on linear combinations of ensemble members.
Fix u_0 and **replace linear combinations by morphing:**

apply EnKF to the morphing representation $[r, T]$

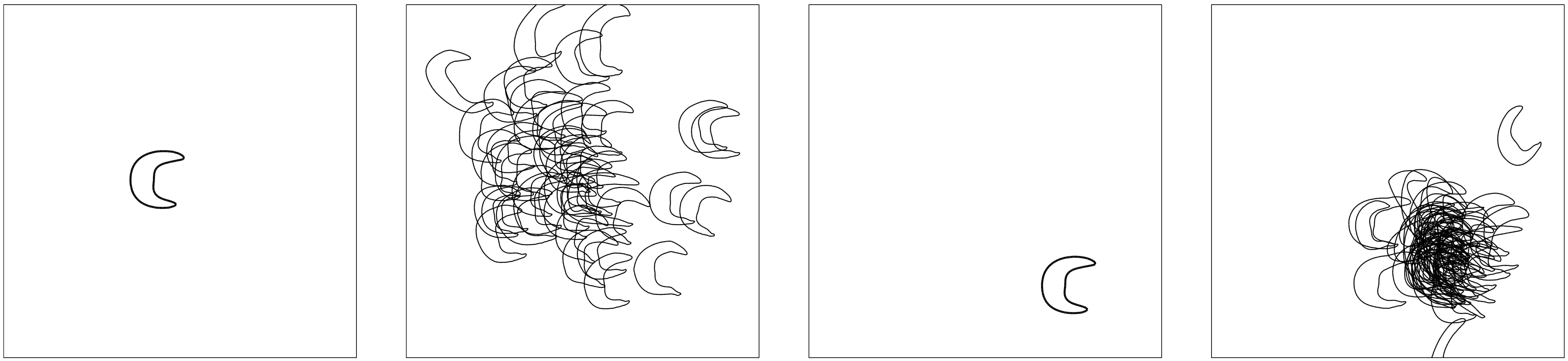
Intermediate States by Morphing



$\lambda = 0$ $\lambda = 0.25$ $\lambda = 0.5$ $\lambda = 0.75$ $\lambda = 1$

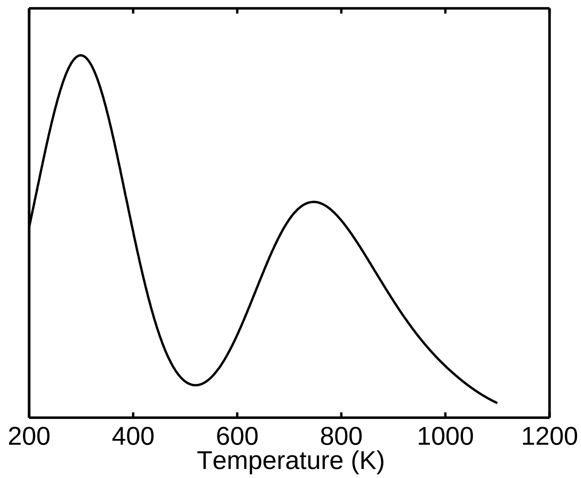
Morphing of two solutions of a reaction-diffusion equation system used in a wildfire simulation. The states with $\lambda = 0$ and $\lambda = 1$ are given. The intermediate states are created automatically. The horizontal plane is the earth surface. The vertical axis and the color map are the temperature. The morphing algorithm combines the values as well as the positions.

Data assimilation by the Morphing Ensemble Kalman Filter

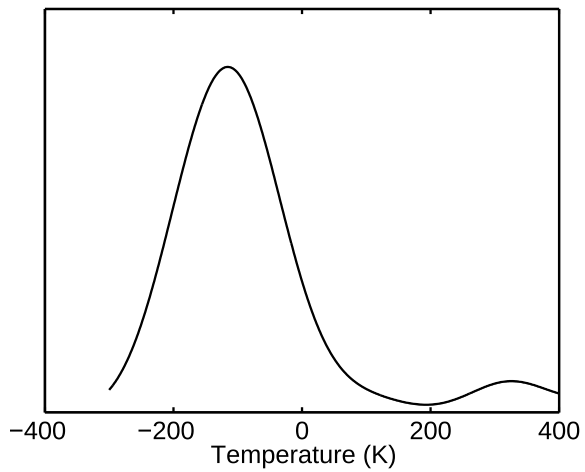

(a) (b) (c) (d)

The forecast ensemble (b) was created by smooth random morphing of the initial temperature profile (a). The analysis ensemble (d) was obtained by the EnKF applied to the transformed state. The data for the EnKF was the morphing transformation of the simulated data (c), and the observation function was the identity mapping. Contours are at $800K$, indicating the location of the fireline. The reaction zone is approximately between the two curves.

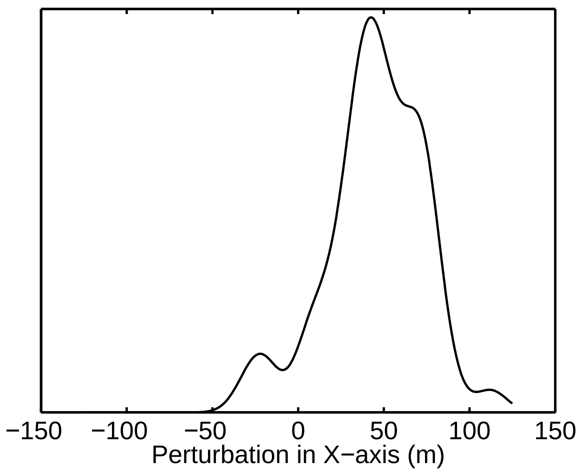
Morphing Transform Makes Distribution Closer to Gaussian



(a)



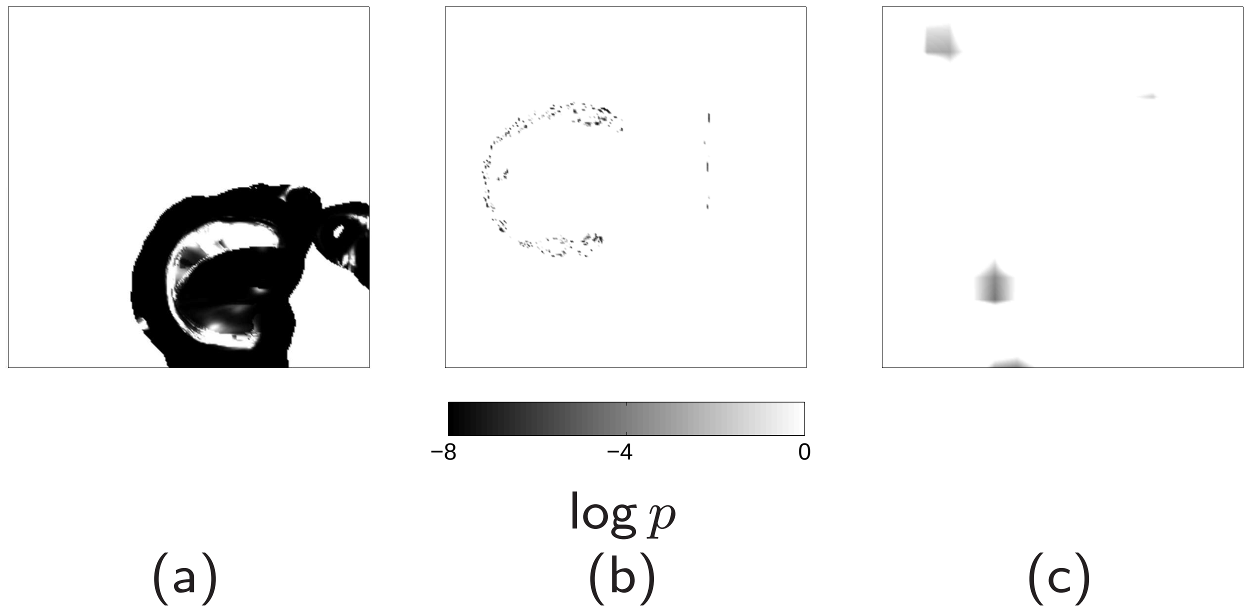
(b)



(c)

Typical pointwise densities near the reaction area of the original temperature (a), the residual component after the morphing transform, and (c) the spatial transformation component in the X-axis. The transformation has made bimodal distribution into unimodal.

Morphing Transform Makes Distribution Closer to Gaussian



The $p-$value of the data from the ensemble after five EnKF analysis cycles from the Anderson-Darling test to estimate the "Gaussian nature" of the point-wise densities throughout the domain. The shading indicates that the sample is highly Gaussian (white) or highly non-Gaussian (black) for the original temperature (a), the transformed residual temperature (b), and the morphing function (c).

2. Predictor-Corrector Ensemble Filters

Particle Filters

Model state (= probability density $p(u)$ of the system state u) is represented by a **weighted ensemble** (u_k, w_k), $k = 1, \ldots, N$

$\{u_k\}$ is a sample from some probability density p_π

w_k are positive weights, $\sum_{k=1}^{N} w_k = 1$, $\boxed{w_k \propto \frac{p(u_k)}{p_\pi(u_k)}}$

Given forecast ensemble (u_k^f, w_k^f) generate analysis ensemble by sampling from some p_π and get the weights from data likelihood

$$\boxed{u_k^a \sim p_\pi, \quad w_k^a \propto p\left(d|u_k^a\right) \frac{p_f\left(u_k^a\right)}{p_\pi\left(u_k^a\right)}}$$

SIS chooses $u_k^a = u_k^f$ (does not change the ensemble) and **only updates the weights** $\Longrightarrow$ already $\left(u_k^f\right) \sim p_\pi$, $w_k^f \propto \frac{p_f(u_k)}{p_\pi(u_k)} \Longrightarrow$

$$\boxed{w_k^a \propto p\left(d|u_k^f\right) w_k^f}$$

Sequential Importance Sampling with Resampling (SIR)

Trouble with SIS:

1. data likelihoods can be very small, then

2. the analysis ensemble represents the analysis density poorly

3. one or a few of the analysis weights will dominate

Standard solution: SIR (Gordon 1993,...)

1. Resample by choosing u_k^a with probability $\propto w_k^a$, set all weights equal

2. rely on stochastic behavior of the model to recover spread.

But there is still trouble:

1. huge ensembles (thousands) are needed because the analysis distribution is effectively approximated only by those ensemble members that have large weights, and a vast majority of weights is infinitesimal

2. if the model is not stochastic, need artificial perturbation to recover ensemble spread

Solution: Predictor-corrector filters (new)

Place the analysis ensemble so that the weights are all reasonably large.

Predictor-Corrector Filters

Given forecast ensemble $p_f \sim (u_\ell^f, w_\ell^f)$ and a **proposal ensemble** (by a **predictor, EnKF)** $(u_k^a) \sim p_\pi$

Apply Bayes theorem, get weights by estimating $\boxed{\frac{p_f(u_k^a)}{p_\pi(u_k^a)}}$

Trouble:

1. density estimates in high dimension are intractable

2. need to estimate far away from and outside of the span of the sample

Solution:

1. the probability densities are not arbitrary: they come from probability measures on spaces of smooth functions, low effective dimension

2. nonparametric estimation that depends only the concept of distance

Nonparametric Density Ratio Estimation

$$\frac{p_f\left(u_k^a\right)}{p_\pi\left(u_k^a\right)} \approx \frac{\sum_{\ell:\left\|u_\ell^f - u_k^a\right\| < h} w_\ell^f}{\sum_{\ell:\left\|u_\ell^f - u_k^a\right\| < h} \frac{1}{N}}$$

The norm is linked to the underlying measure via the measure of small balls.

The probability for a function in the state space to fall into a ball should be positive and have limit zero for small balls.

The probability for a function in the state space to fall in a set of measure ν zero should be zero.

In infinite dimension, this is far from automatic, and restricts the choice of the norm in the density estimate! The initial ensemble is constructed by a perturbation of initial condition by smooth random fields; this gives the underlying measure ν and associated norm.

Numerical results for predictor-corrector ensemble filters

Choose:

predictor by EnKF (a new version of EnKF for weighted ensemble): algorighm called EnKF+SIS

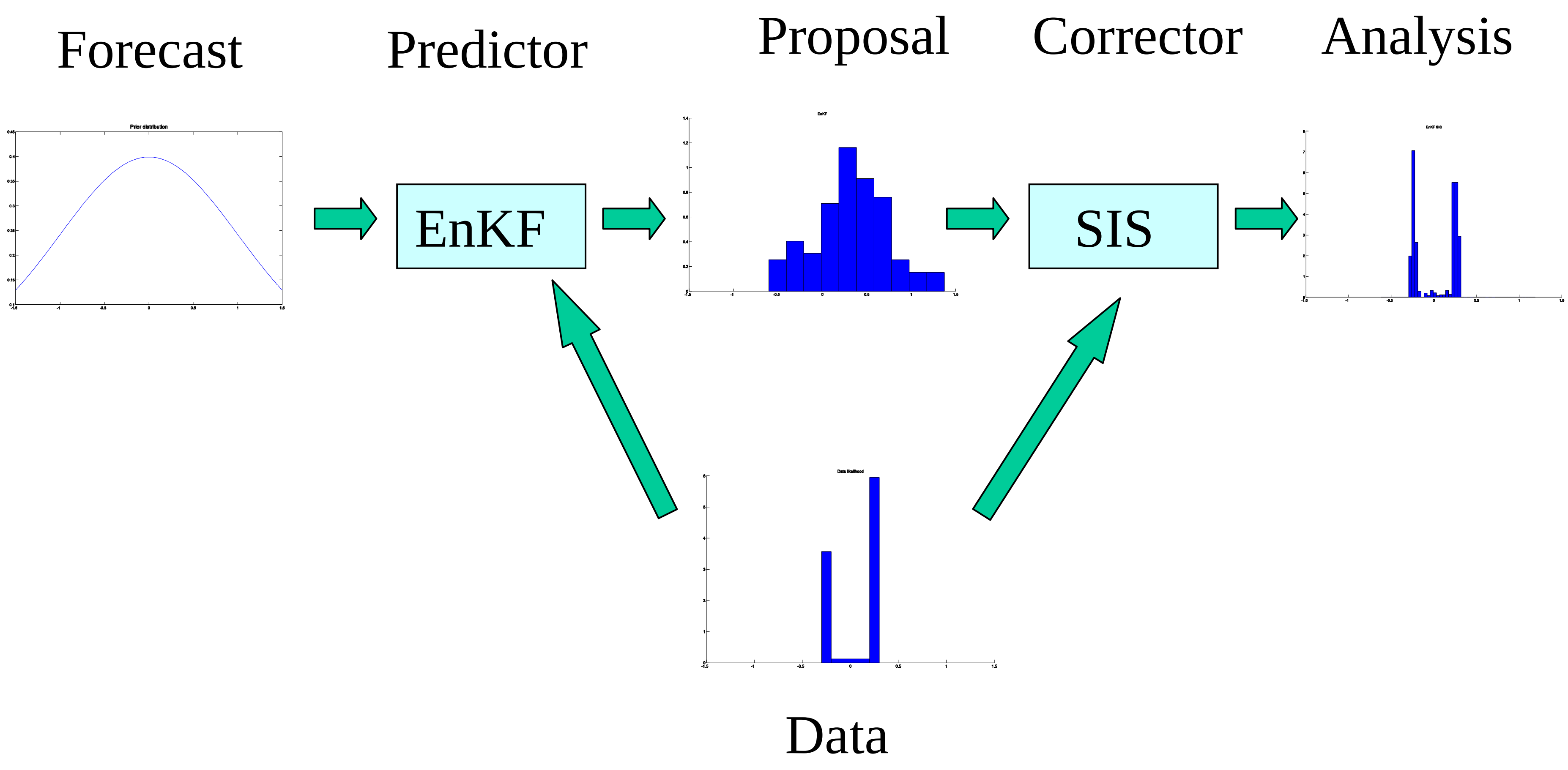
corrector with density estimation with bandwidth by k-th nearest in the proposal ensemble, $k = N^{1/2}$

Norm (distance function) from absolute value for scalar problems, Sobolev norm for problems involving functions.

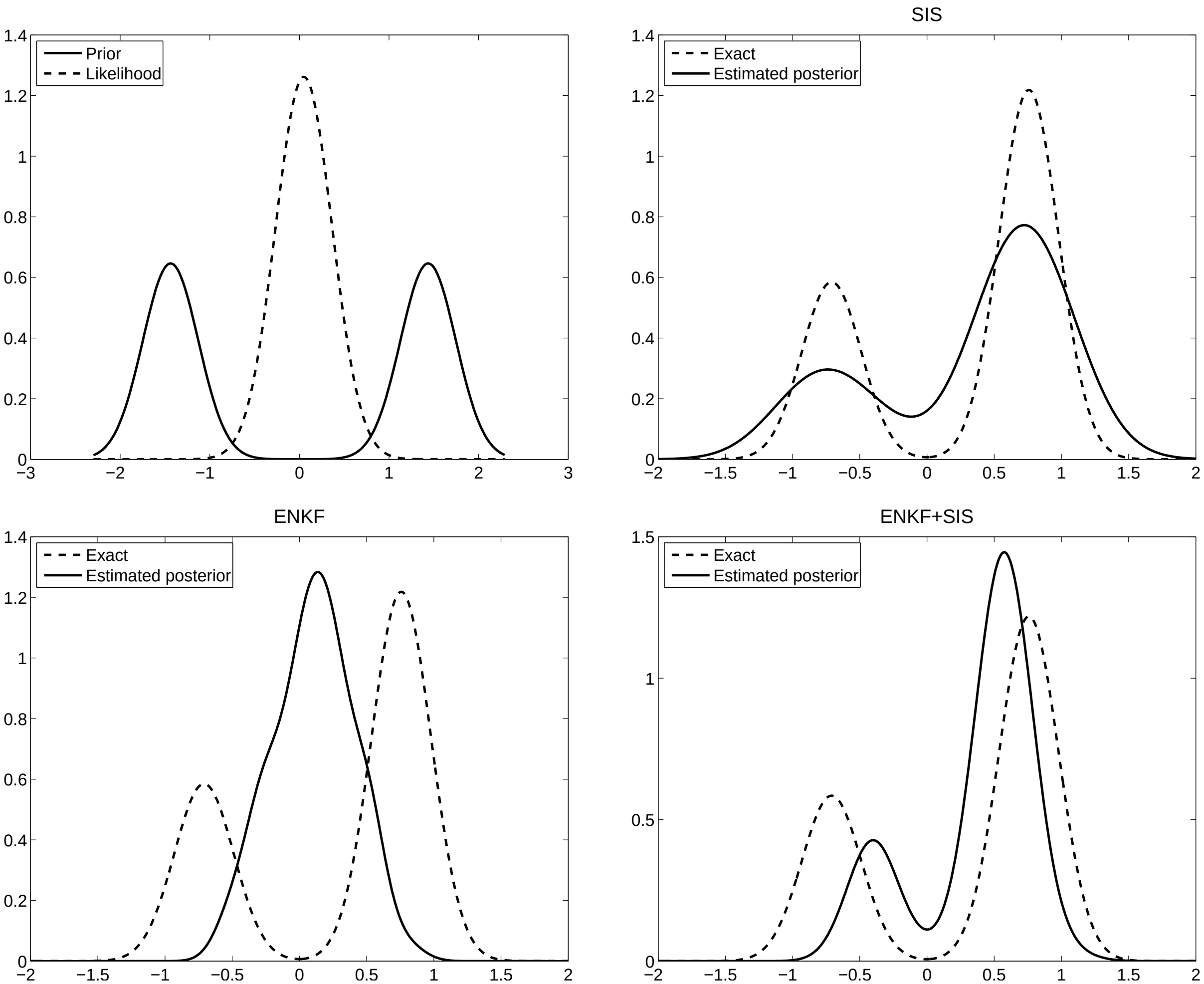
For comparison, using SIS with very large ensemble for “exact” solution.

Forecast also called prior, and analysis is posterior.

Predictor-corrector ensemble filter EnKF+SIS with bimodal data likelihood



Scalar bimodal prior - ensemble size 100



SIS was not close, EnKF did not see nongaussian density. EnKF-SIS was best.

High-dimensional example, bimodal prior

Space of functions on $[0, \pi]$ of the form

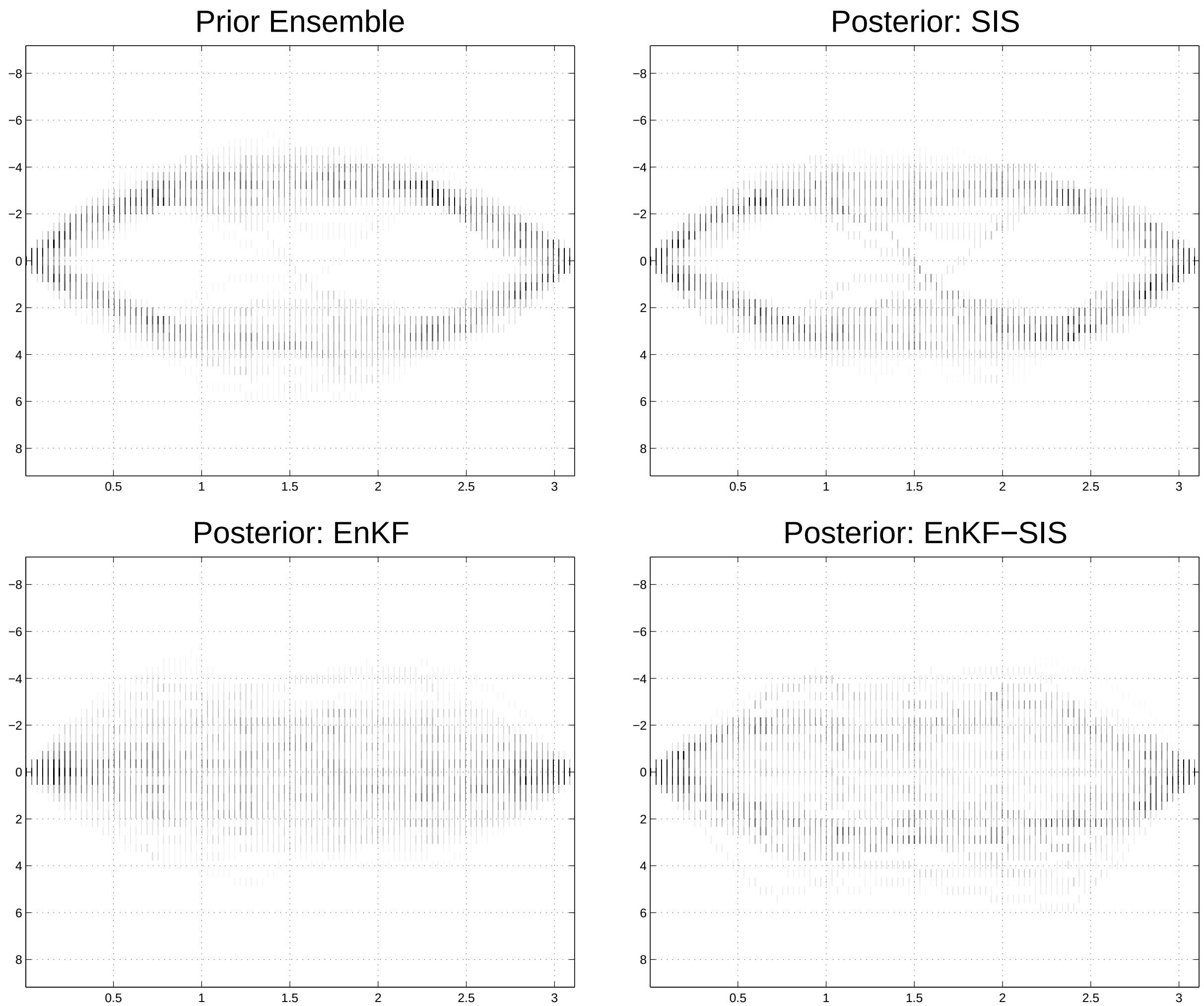
$$u = \sum_{n=1}^{d} c_n \sin(nx)$$

The ensemble size $N = 100$

The dimension of the state space $d = 500$

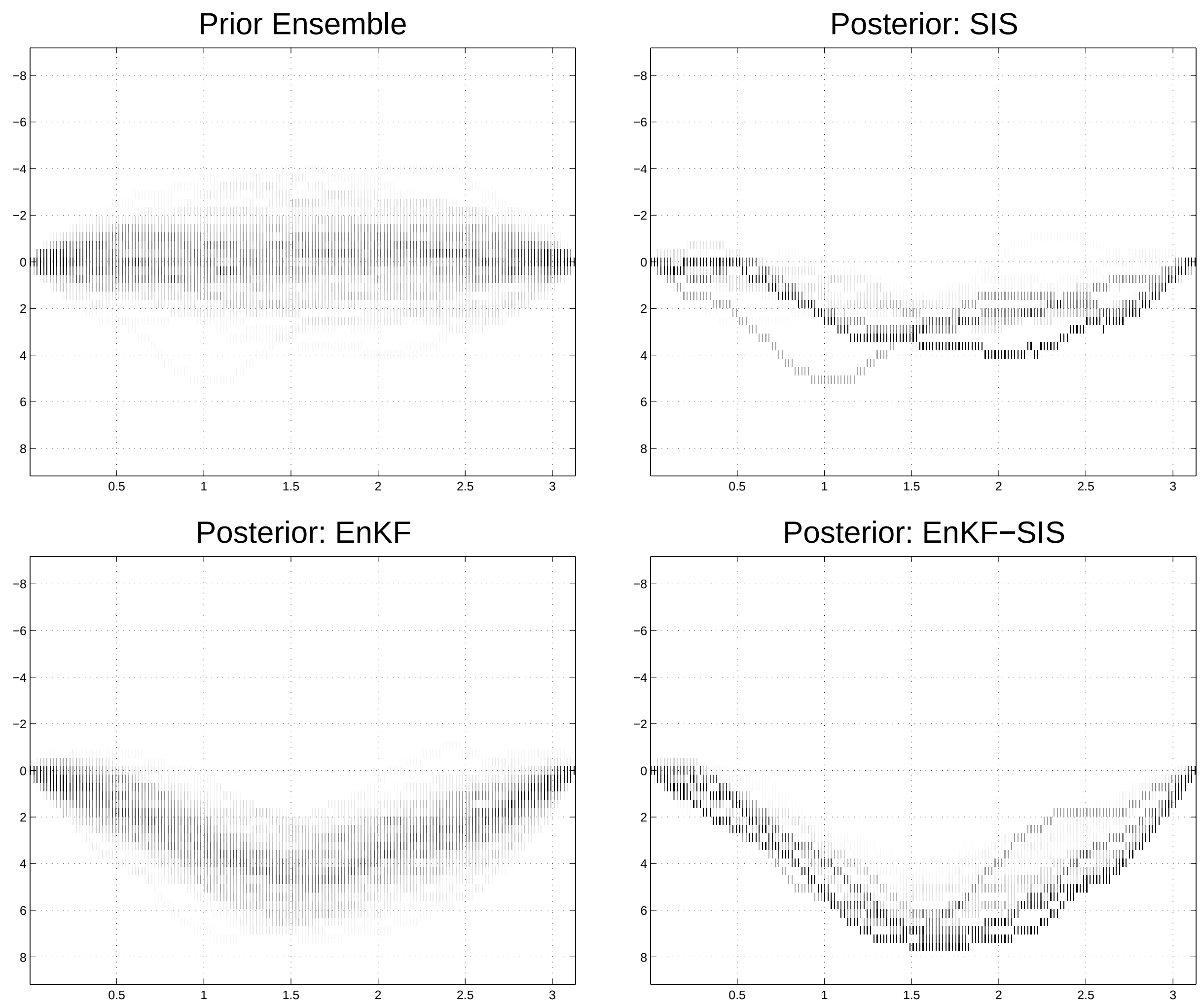
The eigenvalues of the covariance $\lambda_n = n^{-3}$ to generate the initial ensemble and $\lambda_n = n^{-2}$ for density estimation.

High-dimensional example: bimodal prior



SIS does OK, EnKF cannot see bimodal distribution, EnKF+SIS is OK.

High-dimensional example: sparse data, Gaussian case



SIS cannot make a large update, EnKF and EnKF+SIS are fine.

Filtering for a Stochastic ODE

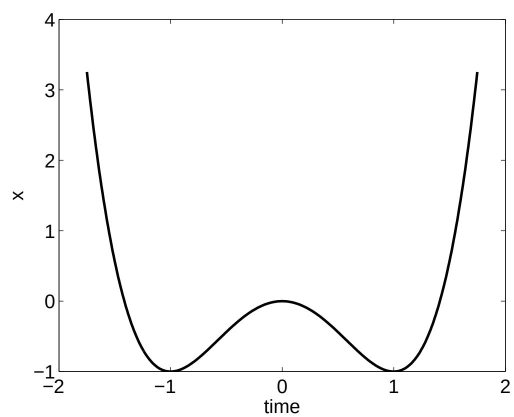
Simplest ODE model problem (PDE fire model adds spatial diffusion)

$$\dot{x} = -f'(x) + \kappa \text{ white noise},$$

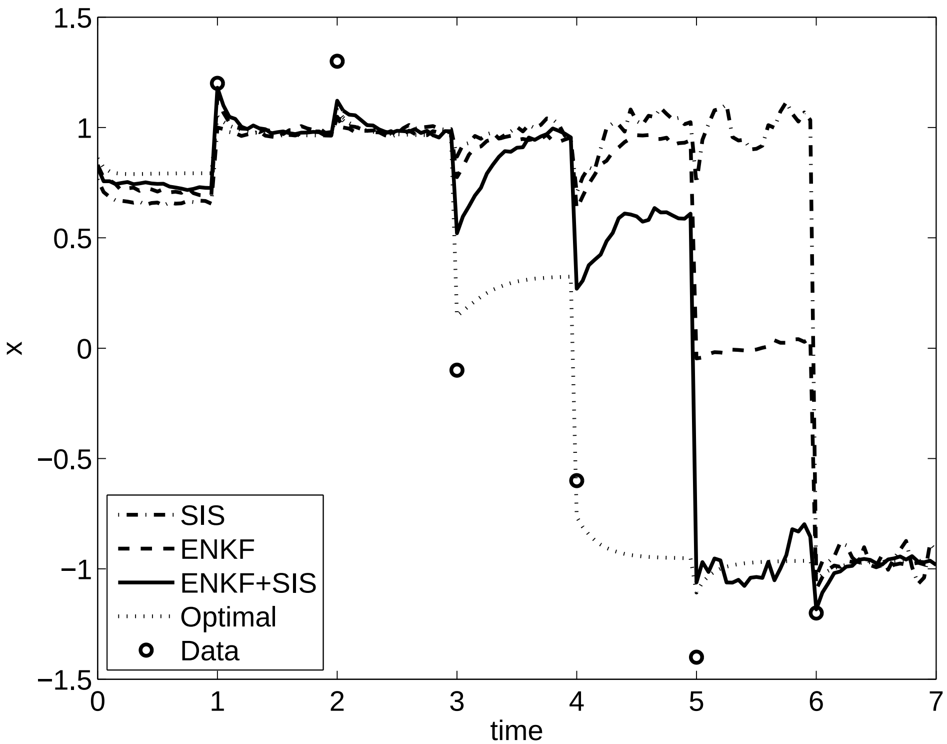
The potential

$$f(x) = -2x^2 + x^4$$

has 3 equilibria and the solution switches between the stable equilibria ± 1.



Filtering results for a stochastic ODE



• One transition from +1 to -1

• data are sampled from one "reference" solution

• ensemble size 100

• mean shown

• optimal solution advances probability distribution "exactly" by numerical solution of the Fokker-Planck equation and applies the Bayes theorem numerically to probability densities discretized by piecewise linear functions.

EnKF+SIS tracks the solution far better than either SIS or EnKF, which lag the transition by more analysis cycles.

Conclusion

- Tested on toy problems
  - better posterior in single analysis step
  - infrequent data, larger state changes
  - better convergence over multiple analysis steps
- Future
  - production quality parallel implementation
  - test on real problems: apply to WRF for hurricanes, WRF+wildfire
  - integrate into DART
- Theory: prove convergence incl. in infinite dimension