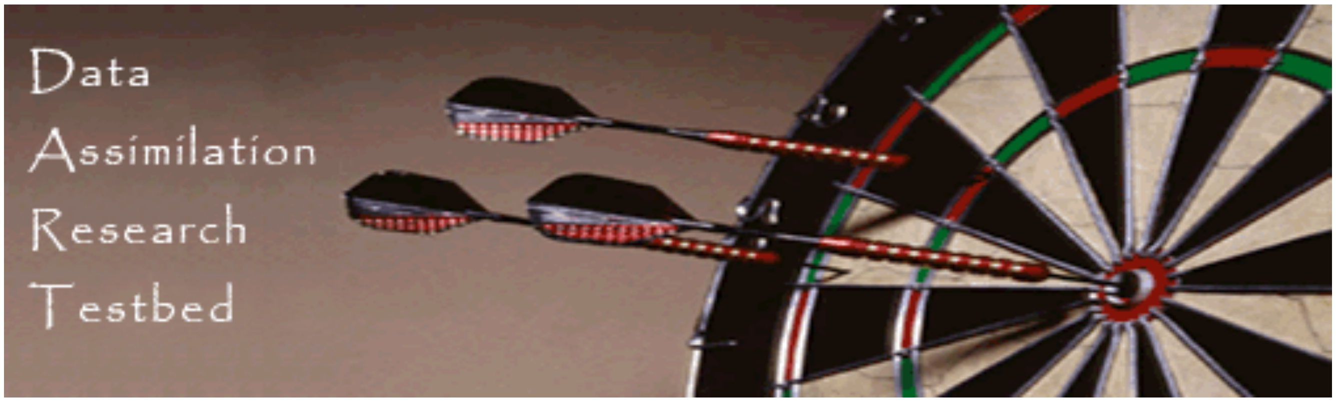
A Non-Gaussian Ensemble Filter for Data Assimilation



Jeffrey Anderson

NCAR Institute for Math Applied to Geophysics

The Ensemble Filter

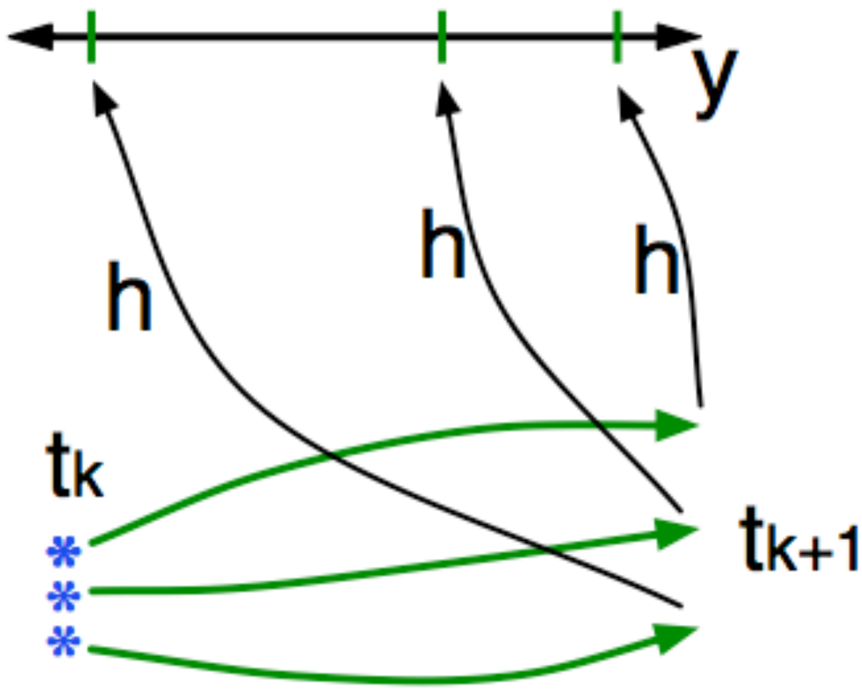
1. Model advances ensemble (3 members here) to time of next observation.

Ensemble state estimate after using previous observation (analysis)

Ensemble state at time of next observation (prior)

The Ensemble Filter

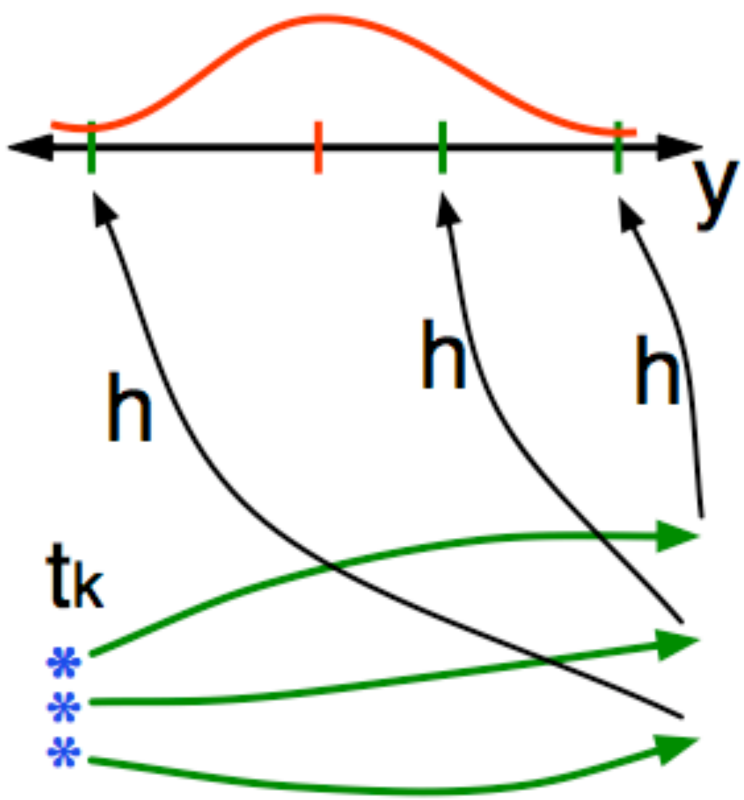
2. Get prior ensemble sample of observation, y = h(x), by applying forward operator **h** to each ensemble member.



Theory: observations from instruments with uncorrelated errors can be done sequentially.

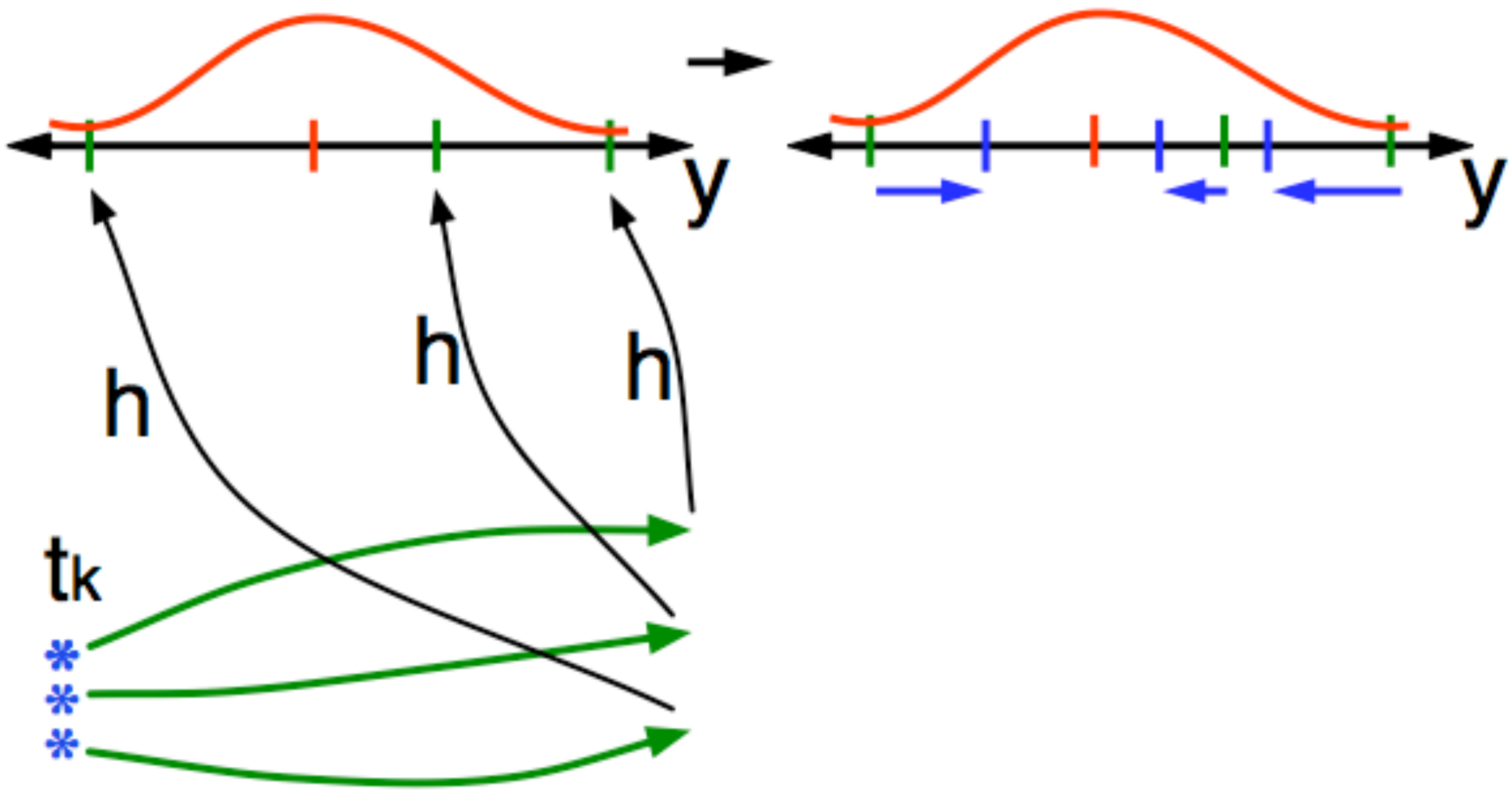
The Ensemble Filter

3. Get observed value and observational error distribution from observing system.



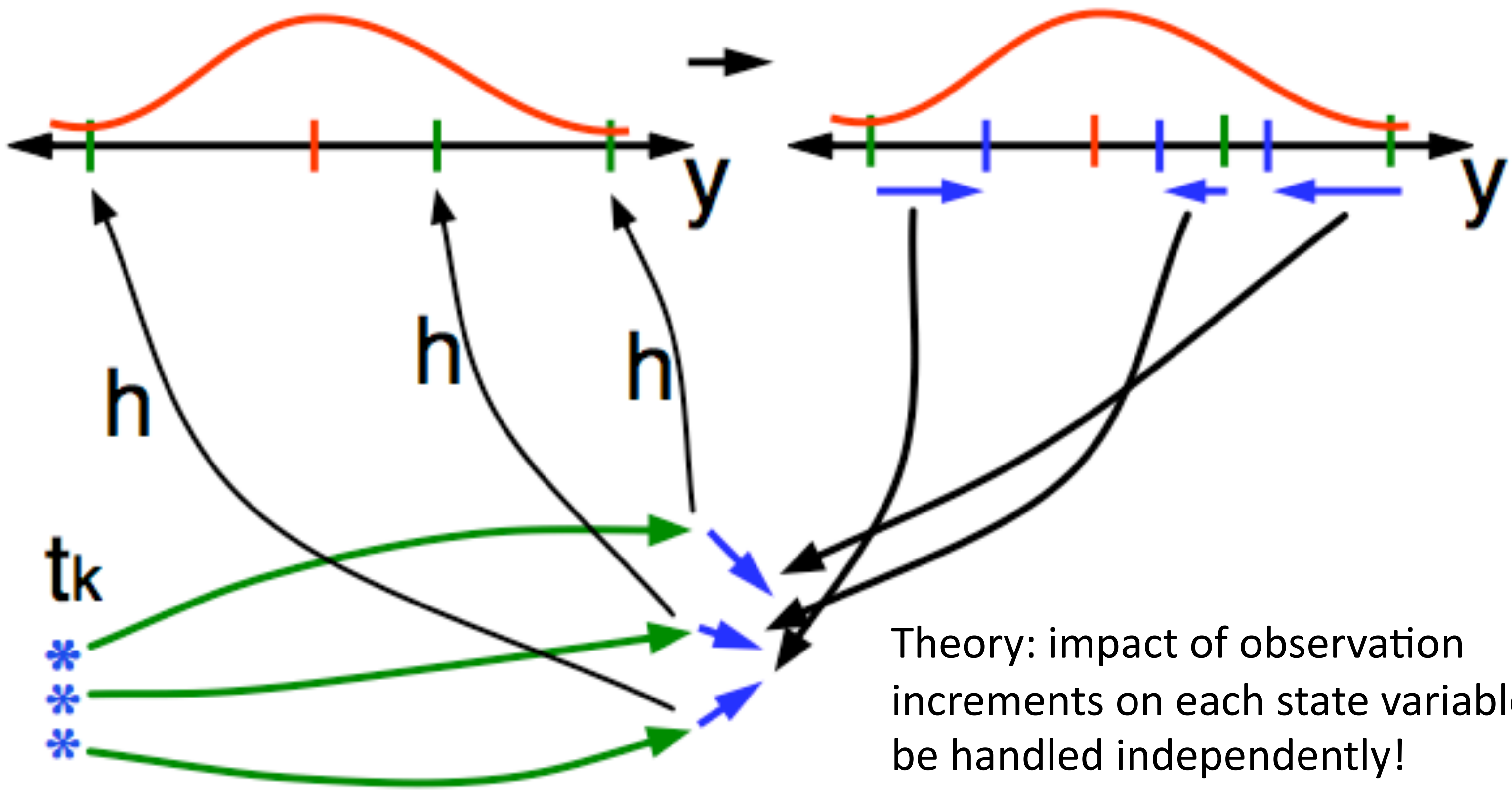
The Ensemble Filter

4. Compute increments for prior observation ensemble (a scalar problem for uncorrelated observation errors). This talk discusses this step.



The Ensemble Filter

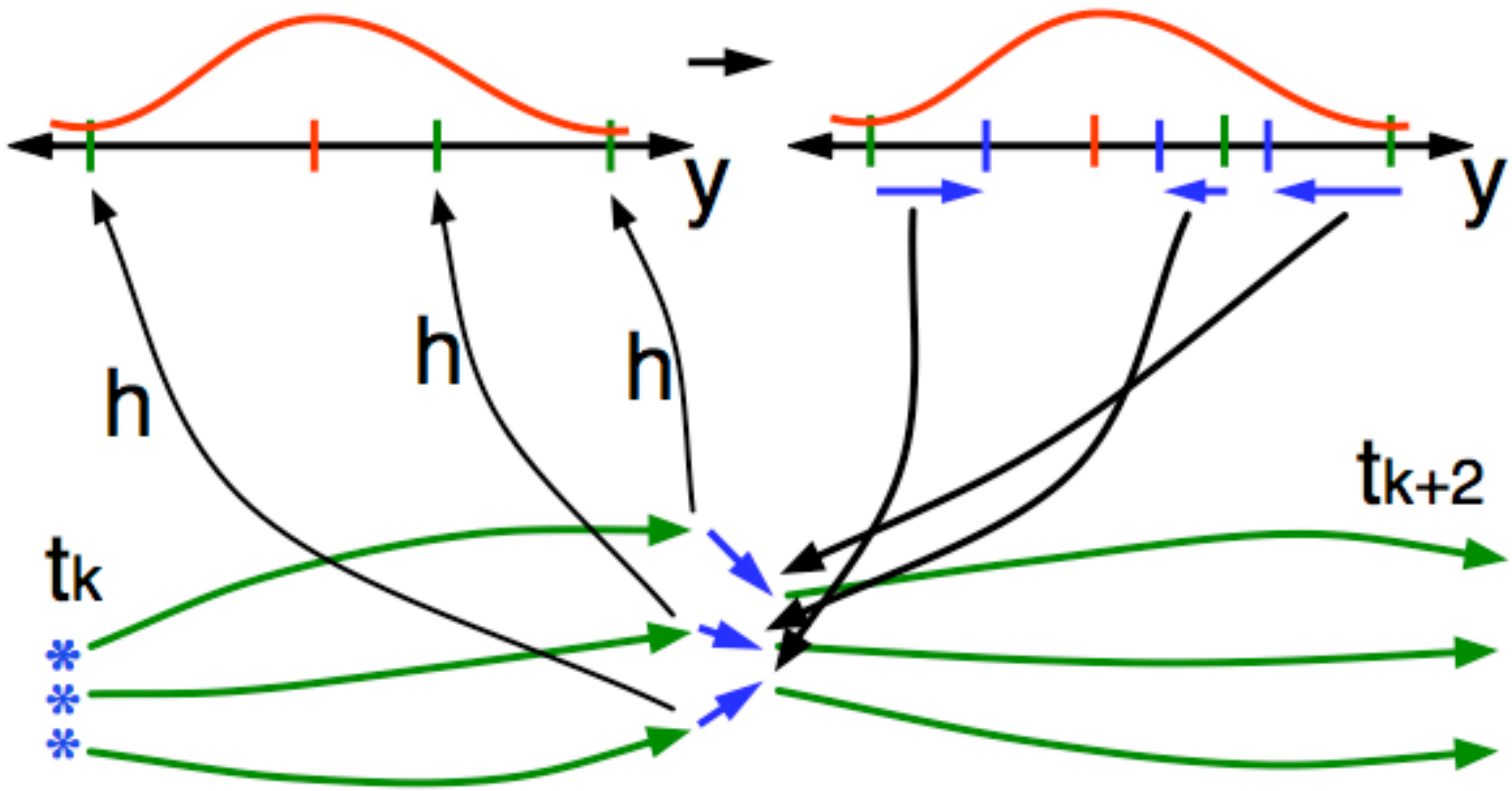
5. Use ensemble samples of **y** and each state variable to linearly regress observation increments onto state variable increments.



Theory: impact of observation increments on each state variable can be handled independently!

The Ensemble Filter

6. Repeat steps 2 through 5 for all observations at this time. Then advance to time of next observation.



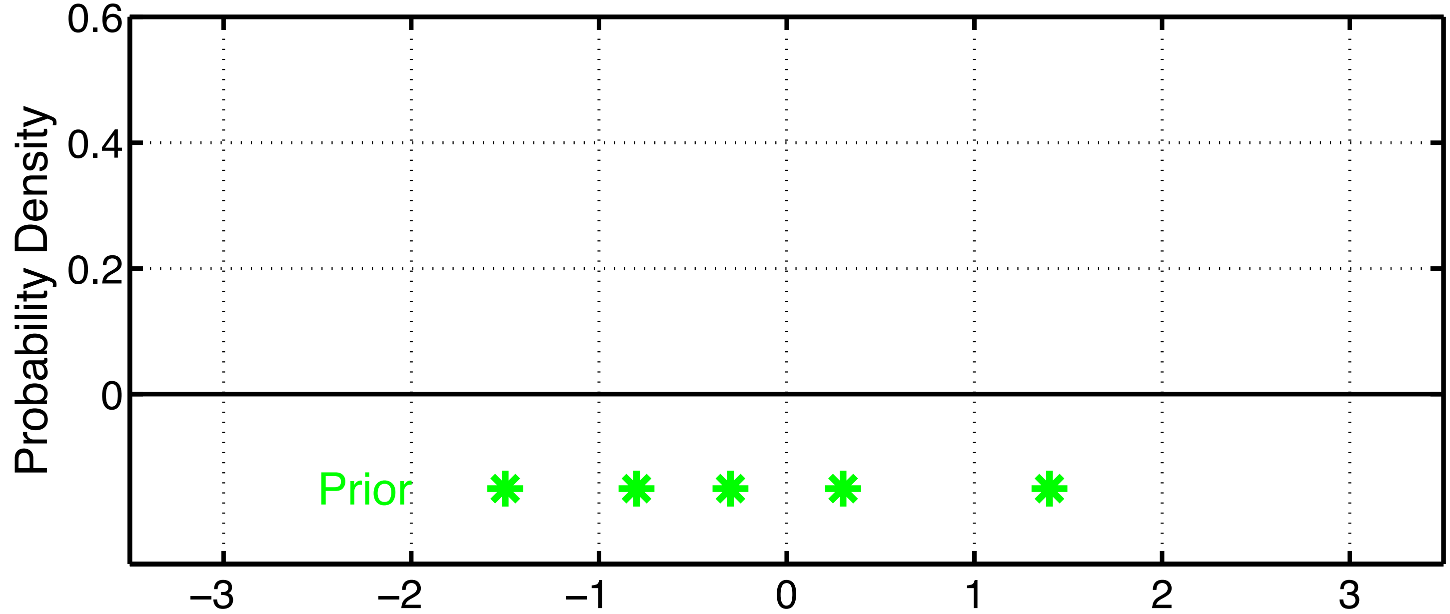
A Deterministic Non-Gaussian Observation Space Update

- Most ensemble filters assume prior and likelihood are approximately gaussian.
- Particle filters do full non-gaussian, but don't scale.
- Assuming non-gaussian in observation space is possible.
- Gaussian kernel filters have been proposed but work poorly.

Requirements for an observation space update

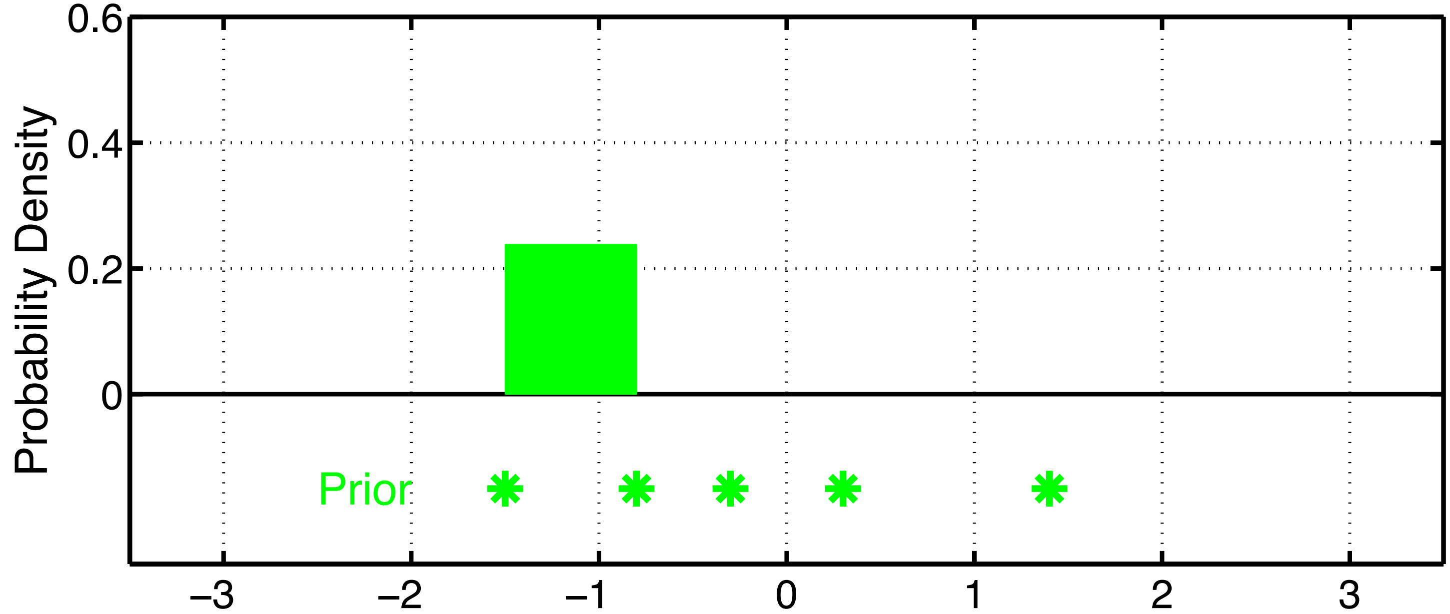
- Low information content obs. can't lead to large increments.
- Want smallest possible increments for all cases.
- Comparable to gaussian filters for ~gaussian cases.
- Better than gaussian in non-gaussian cases.
- Computationally cheap.

Observation Space Rank Histogram Filter



- Apply forward operator to each ensemble member.
- Get prior ensemble in observation space.

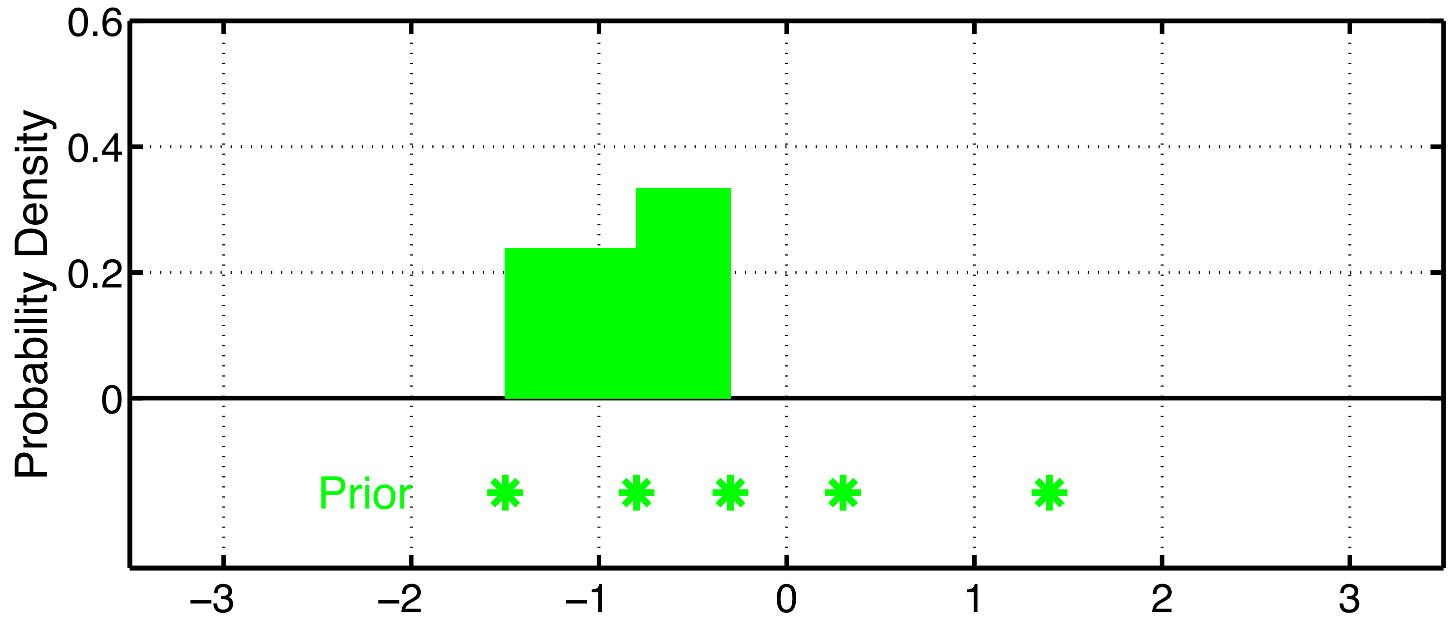
Observation Space Rank Histogram Filter



Step 1: Get continuous prior distribution density.

- Place (ens_size + 1)$^{-1}$ mass between adjacent ensemble members.
- Reminiscent of rank histogram evaluation method.

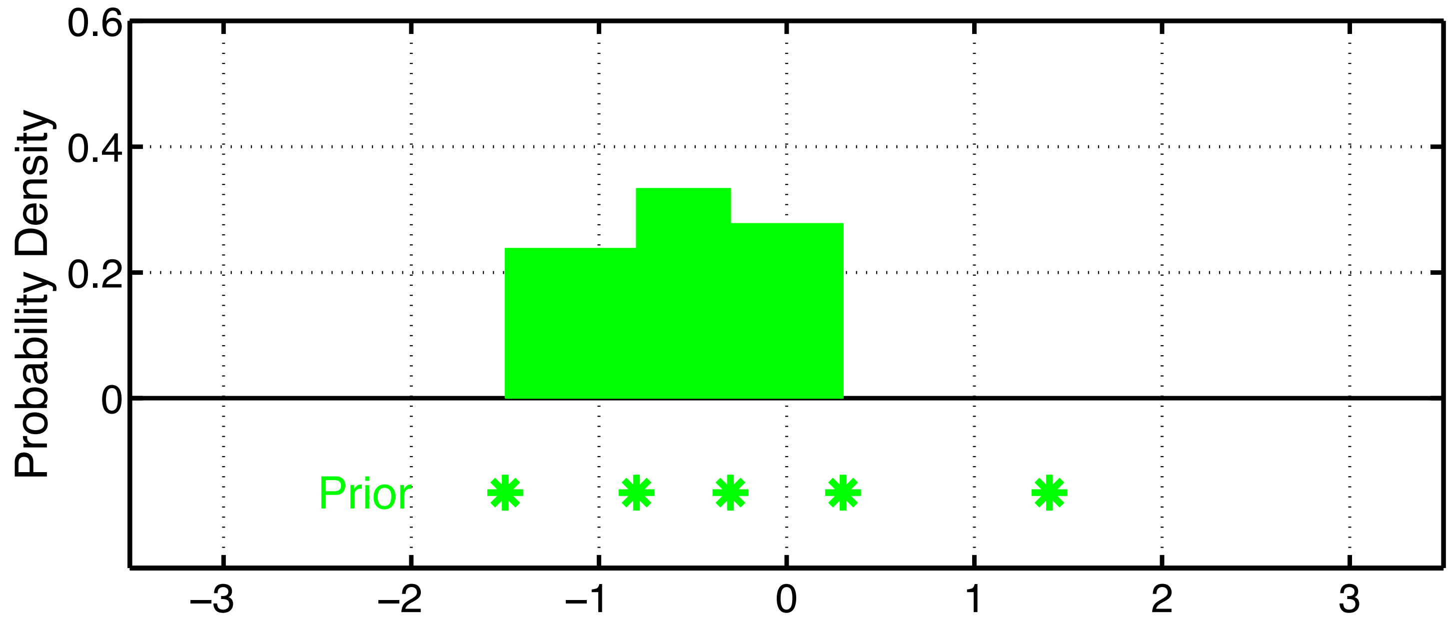
Observation Space Rank Histogram Filter



Step 1: Get continuous prior distribution density.

- Place (ens_size + 1)$^{-1}$ mass between adjacent ensemble members.
- Reminiscent of rank histogram evaluation method.

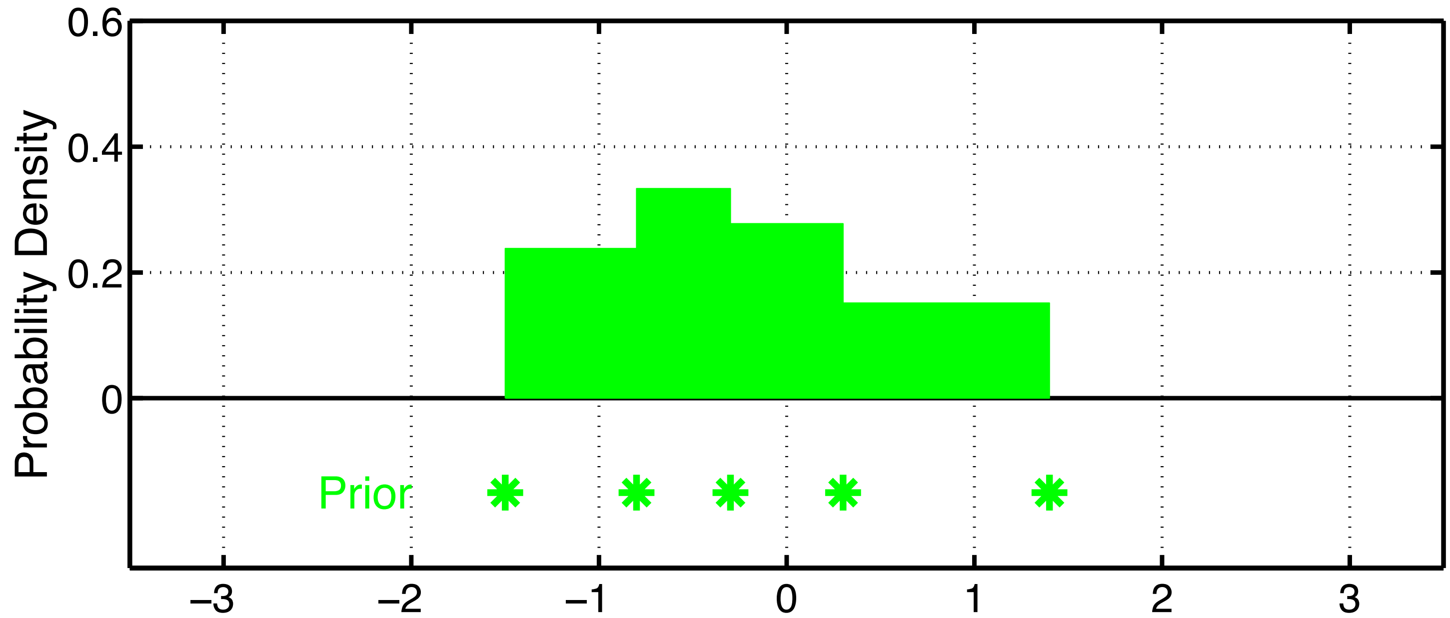
Observation Space Rank Histogram Filter



Step 1: Get continuous prior distribution density.

- Place (ens_size + 1)$^{-1}$ mass between adjacent ensemble members.
- Reminiscent of rank histogram evaluation method.

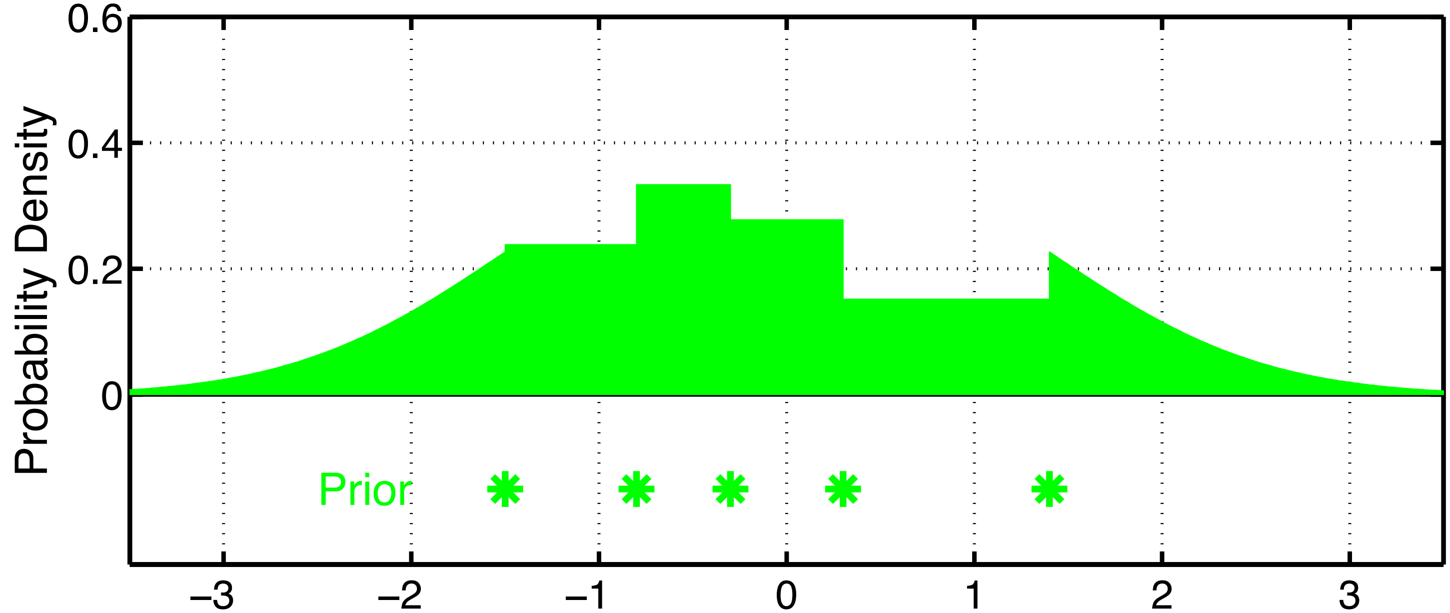
Observation Space Rank Histogram Filter



Step 1: Get continuous prior distribution density.

- Place (ens_size + 1)$^{-1}$ mass between adjacent ensemble members.
- Reminiscent of rank histogram evaluation method.

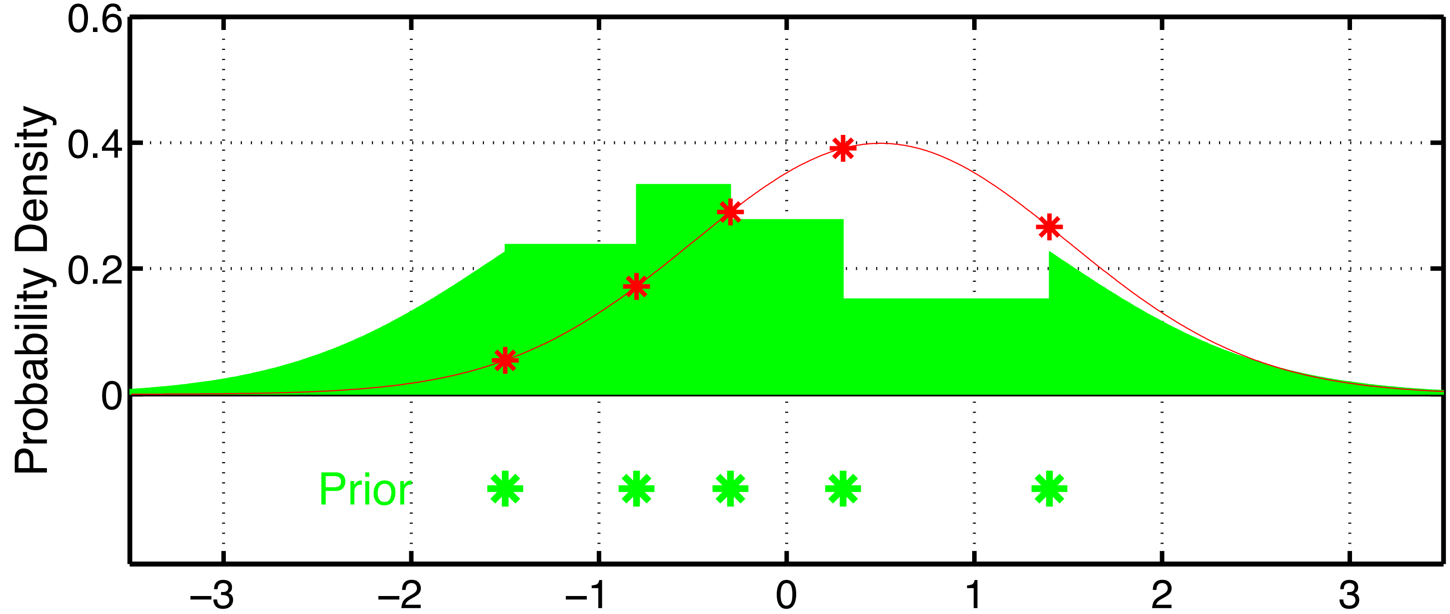
Observation Space Rank Histogram Filter



Step 1: Get continuous prior distribution density.

- Partial gaussian kernels on tails, N(*tail_mean, ens_sd).*
- *tail_mean selected so that (ens_size + 1)$^{-1}$ mass is in tail.*

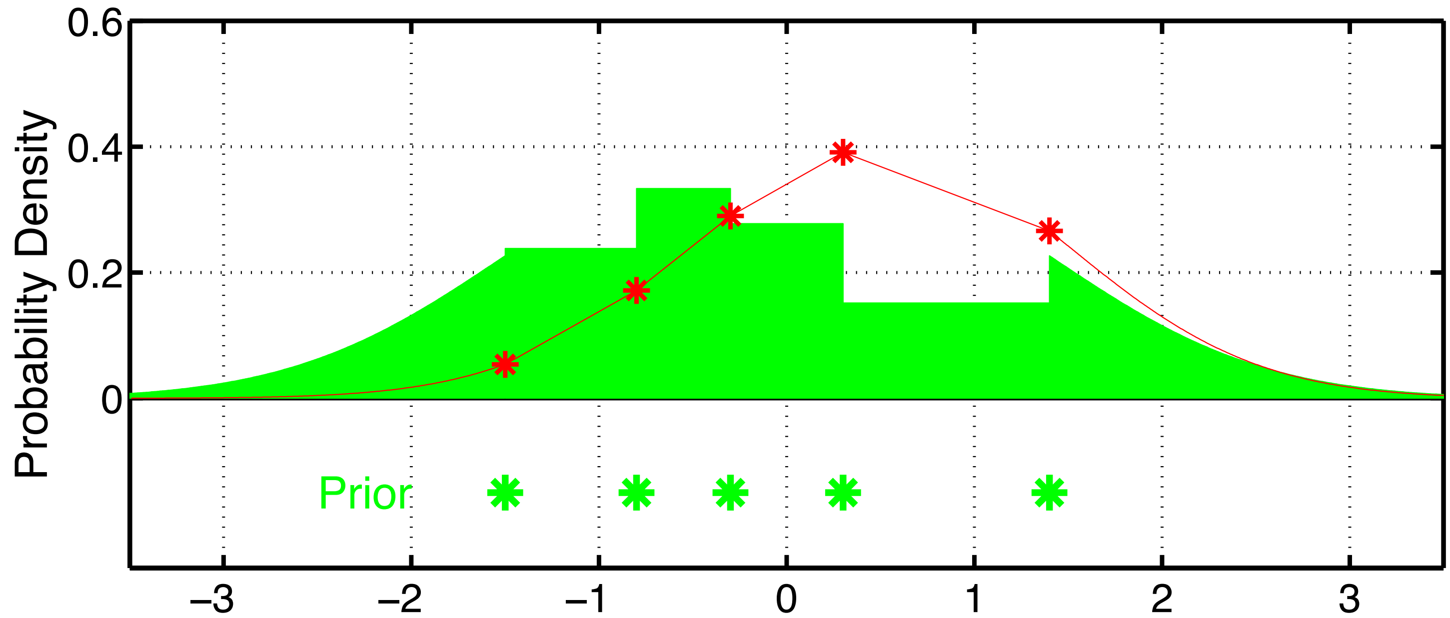
Observation Space Rank Histogram Filter



Step 2: Use likelihood to compute weight for each ensemble member.

- Analogous to classical particle filter.
- Can be extended to non-gaussian obs. likelihoods.

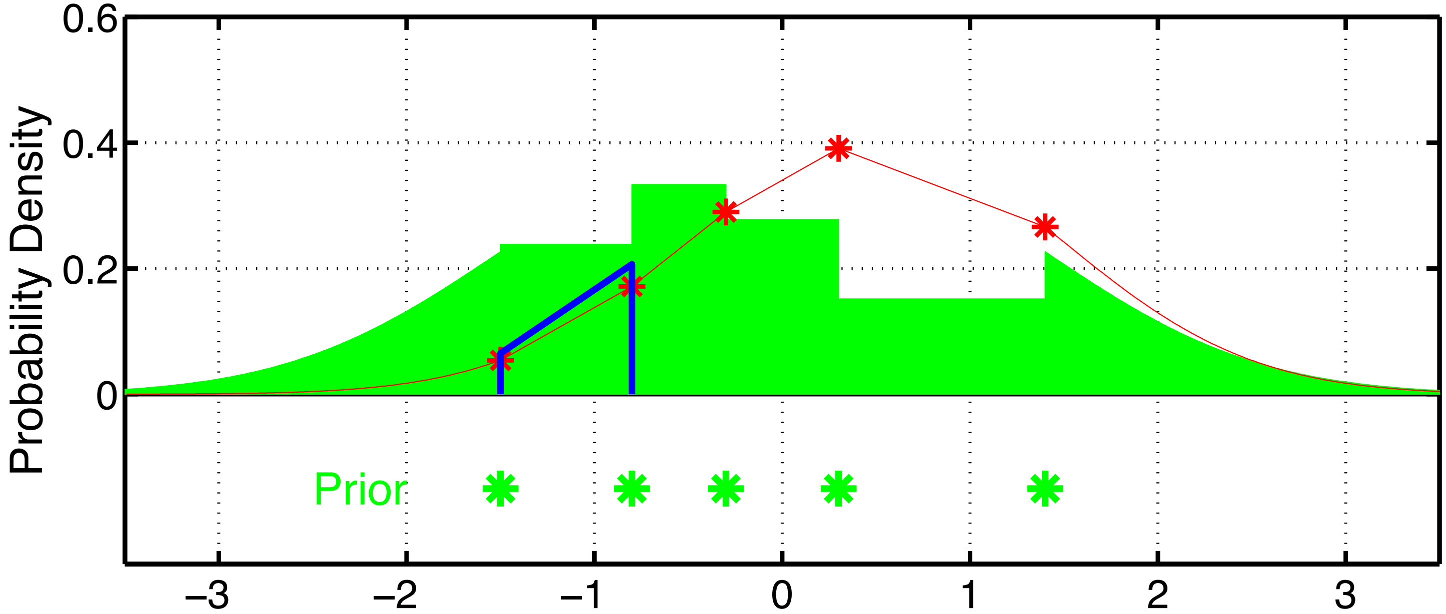
Observation Space Rank Histogram Filter



Step 2: Use likelihood to compute weight for each ensemble member.

- Can approximate interior likelihood with linear fit; for efficiency.

Observation Space Rank Histogram Filter

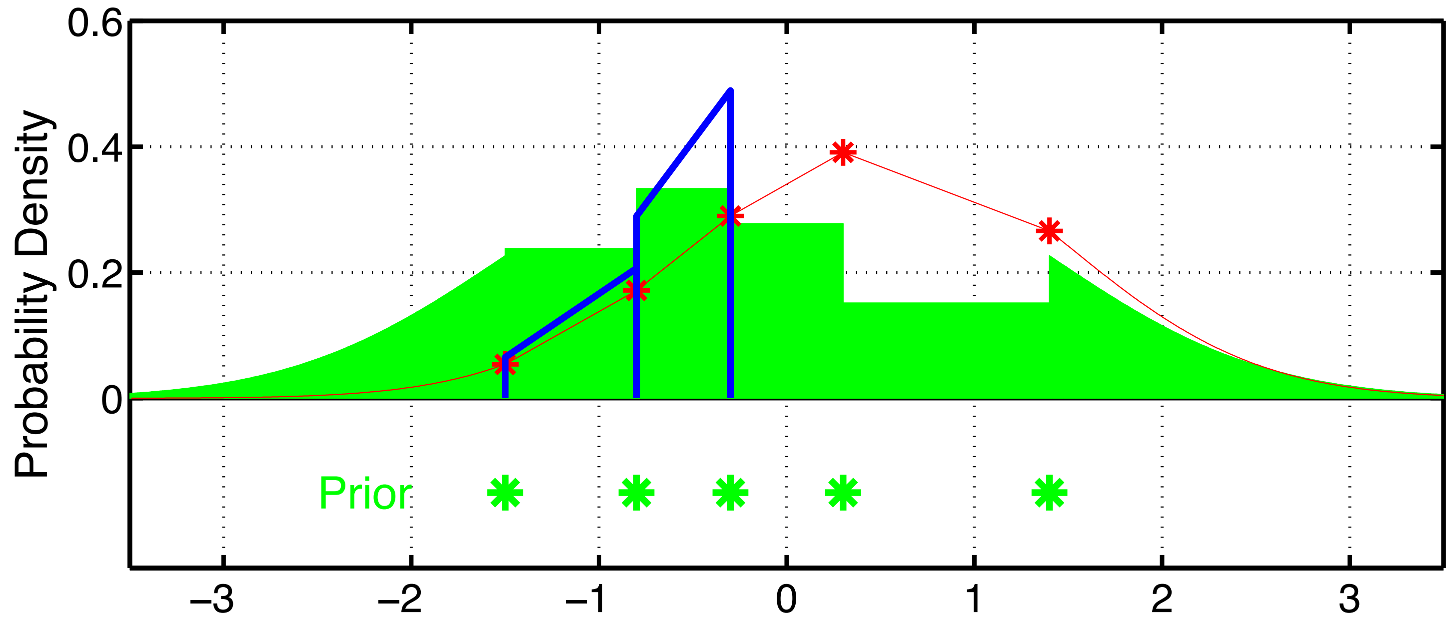


Step 3: Compute continuous posterior distribution.

- Approximate likelihood with trapezoidal quadrature, take product.

  (Displayed product normalized to make posterior a PDF).

Observation Space Rank Histogram Filter

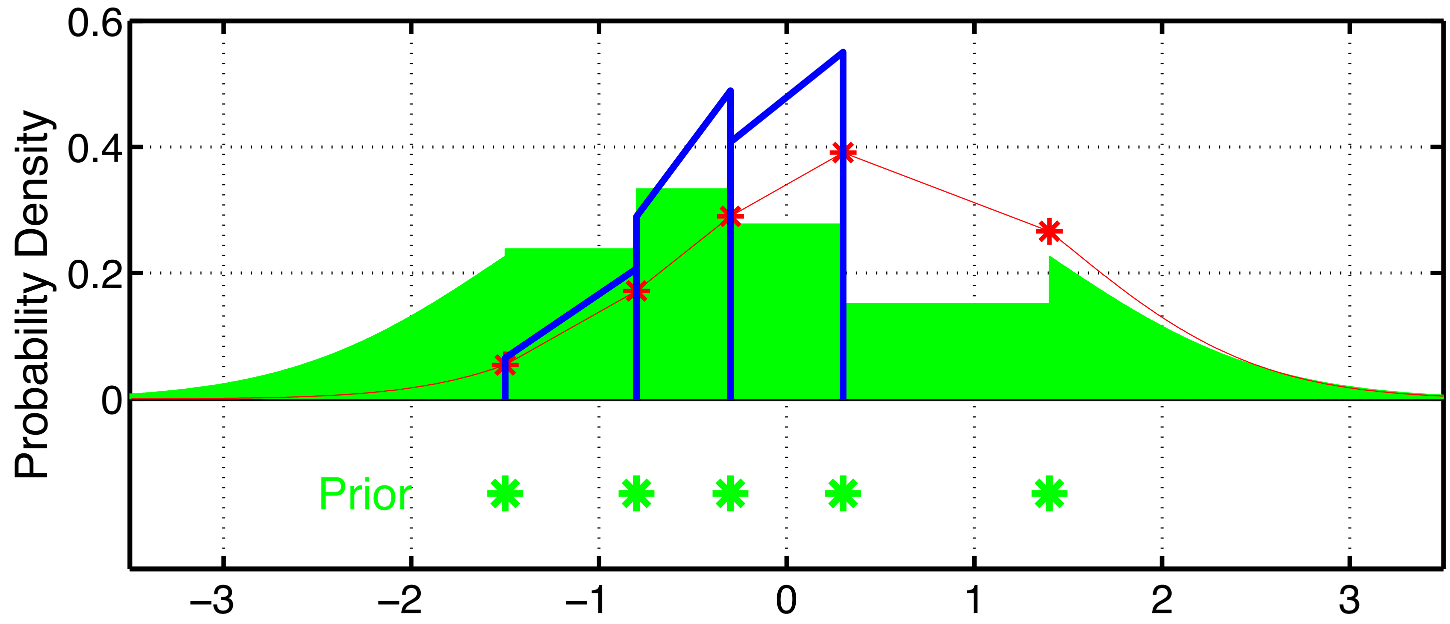


Step 3: Compute continuous posterior distribution.

- Approximate likelihood with trapezoidal quadrature, take product.

  (Displayed product normalized to make posterior a PDF).

Observation Space Rank Histogram Filter

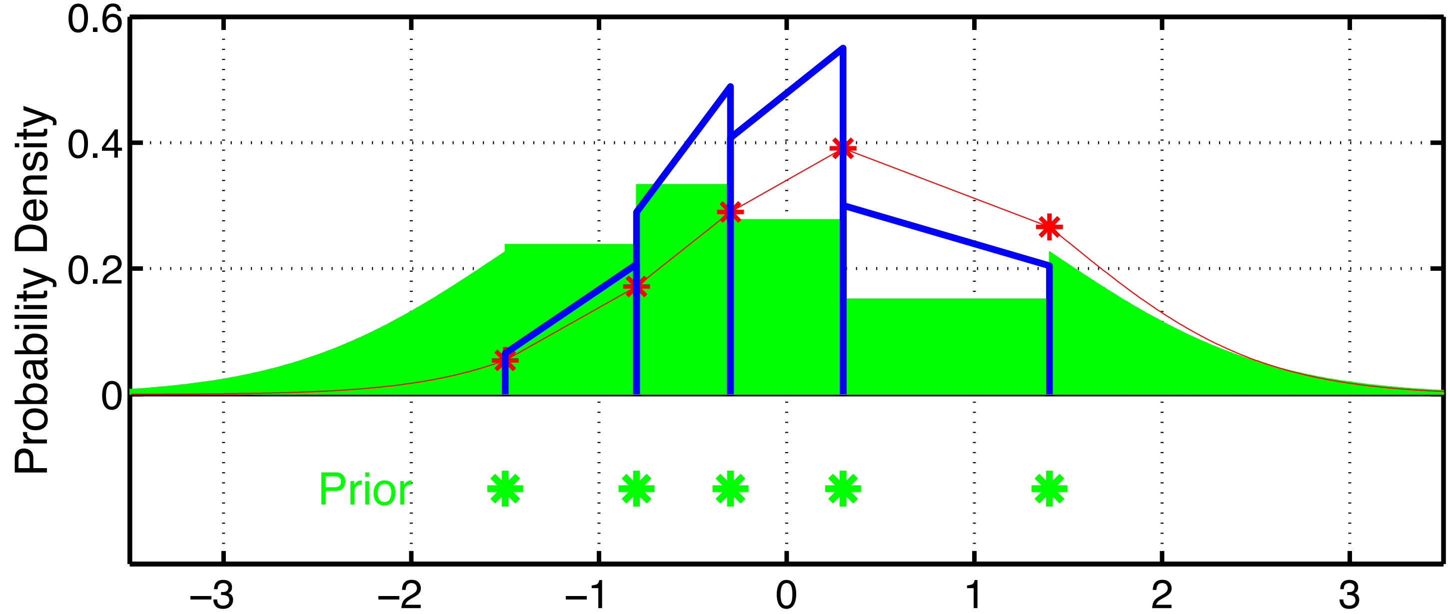


Step 3: Compute continuous posterior distribution.

- Approximate likelihood with trapezoidal quadrature, take product.

  (Displayed product normalized to make posterior a PDF).

Observation Space Rank Histogram Filter

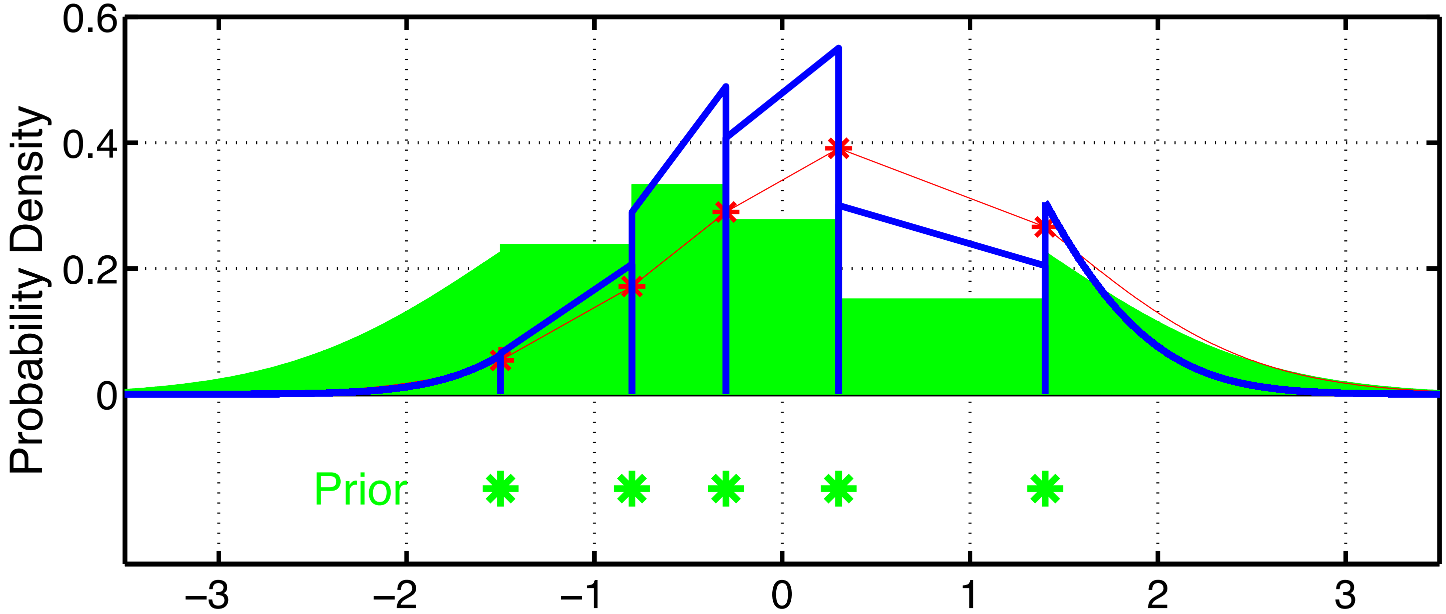


Step 3: Compute continuous posterior distribution.

- Approximate likelihood with trapezoidal quadrature, take product.

  (Displayed product normalized to make posterior a PDF).

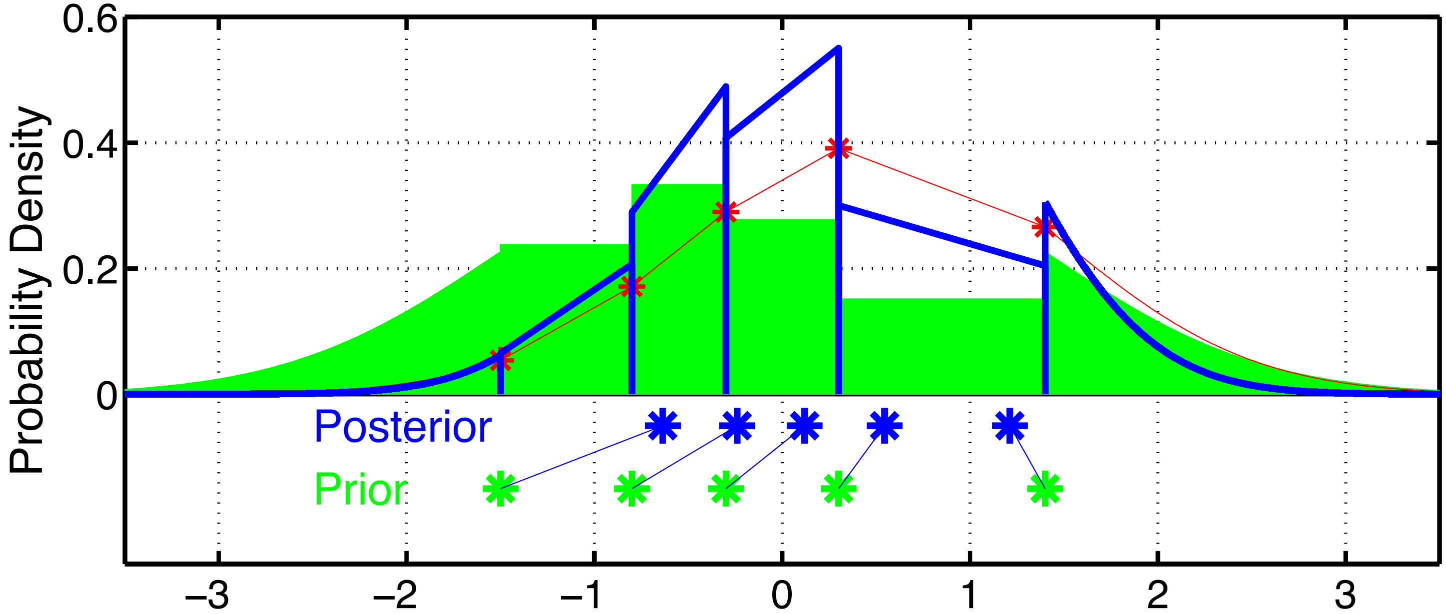
Observation Space Rank Histogram Filter



Step 3: Compute continuous posterior distribution.

- Product of prior gaussian kernel with likelihood for tails.
- Easy for gaussian likelihood.
- More quadrature if non-Gaussian likelihood.

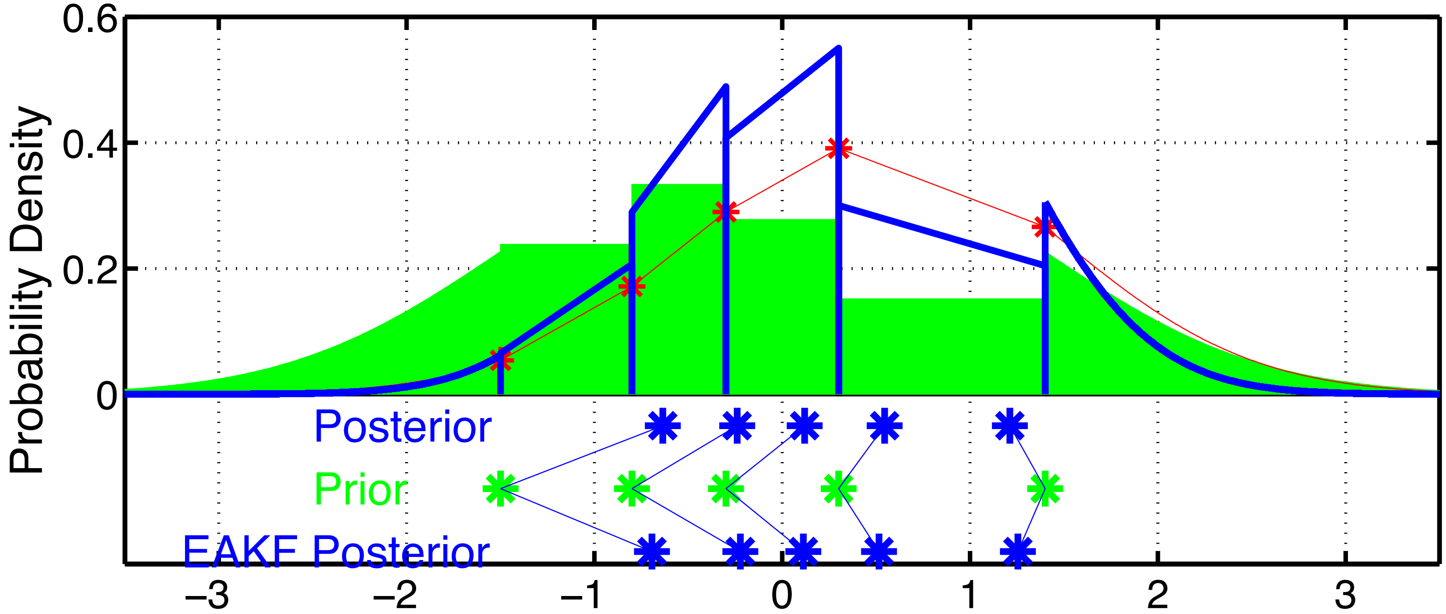
Observation Space Rank Histogram Filter



Step 4: Compute updated ensemble members:

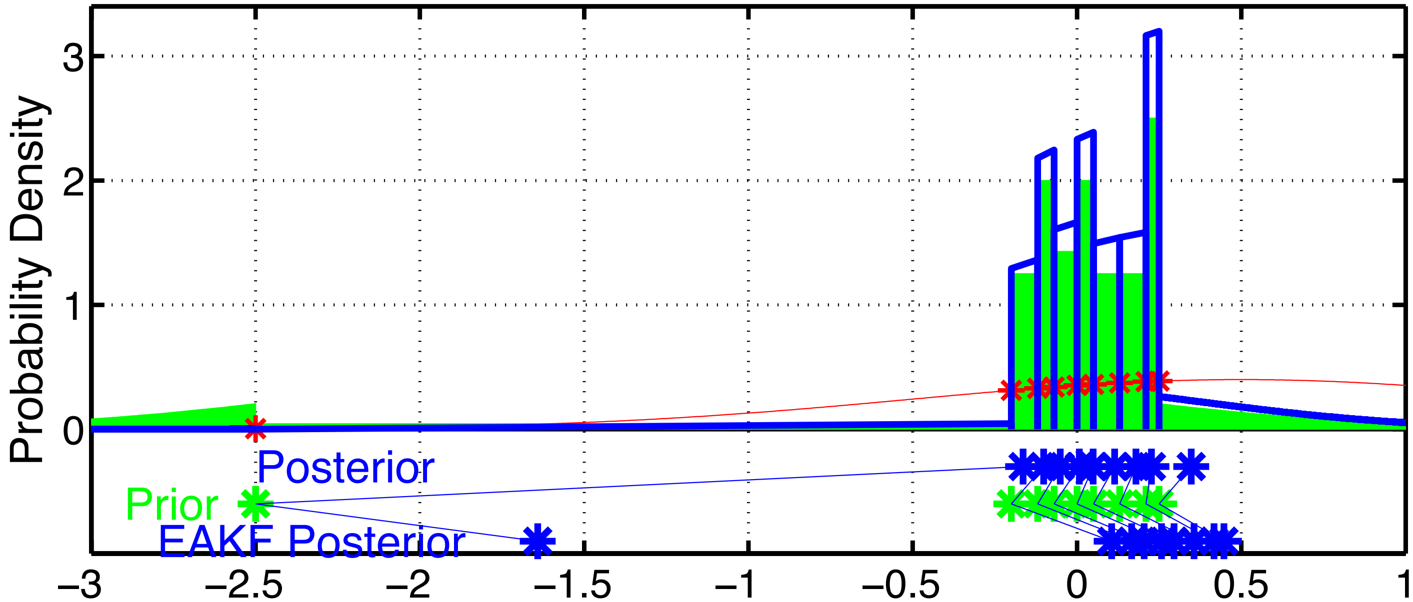
- (ens_size + 1)$^{-1}$ of posterior mass between each ensemble pair.
- (ens_size + 1)$^{-1}$ in each tail.
- Uninformative observation has no impact.

Observation Space Rank Histogram Filter



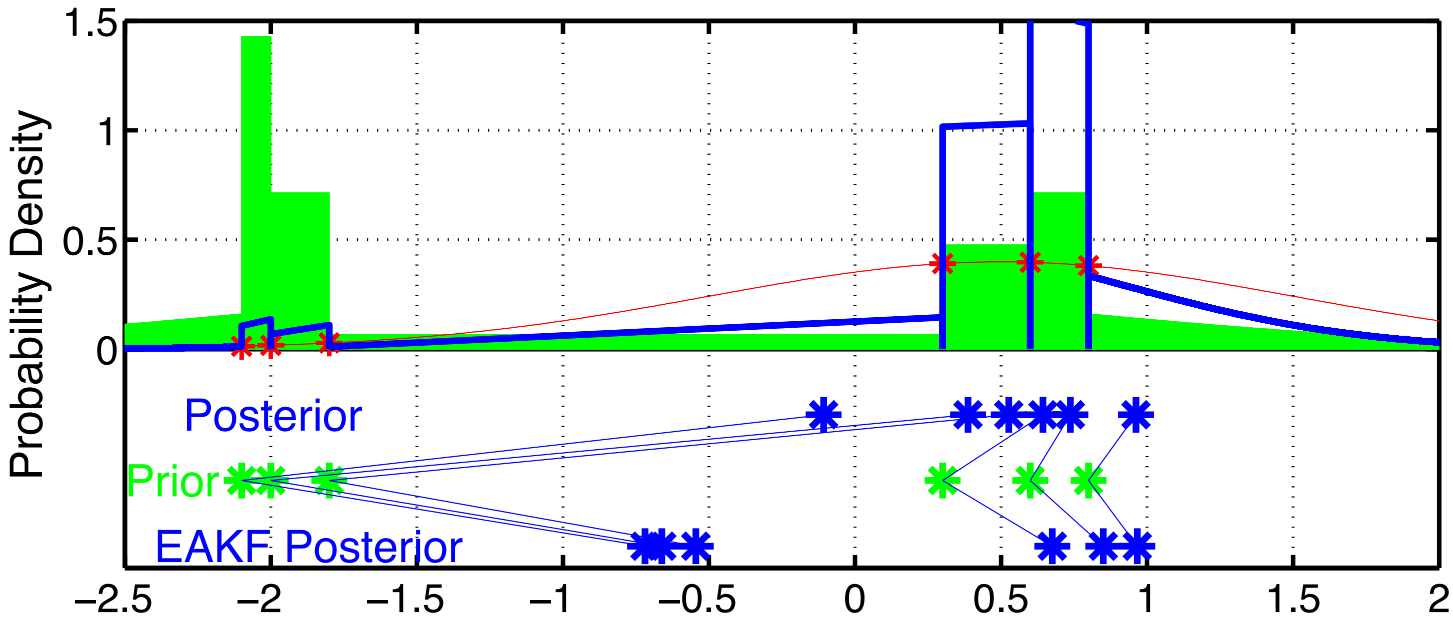
- Compare to standard Ensemble Adjustment Filter (EAKF).
- Nearly gaussian case, differences are small.

Outliers are a Challenge for Gaussian Filters



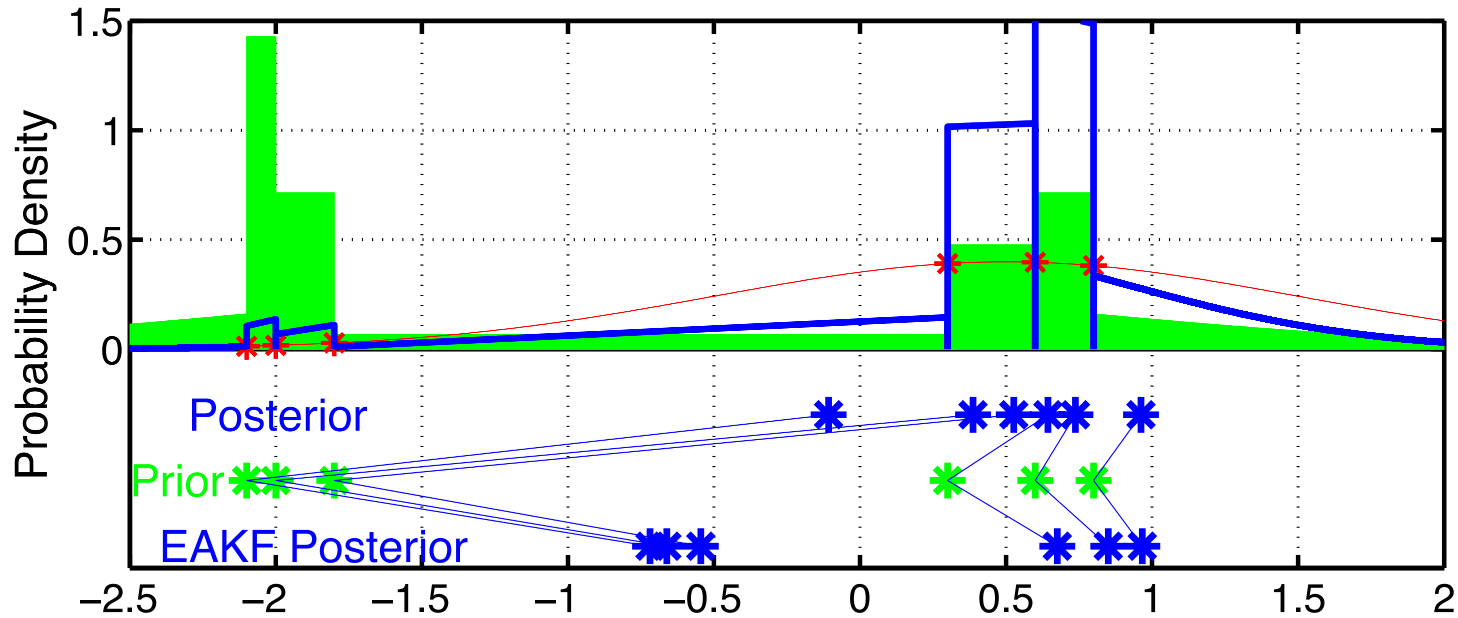
- Rank Histogram gets rid of outlier that is clearly inconsistent with obs.
- EAKF can't get rid of outlier.
- Large prior variance from outlier causes EAKF to shift all members too much towards observation.

Multimodal Prior Distributions



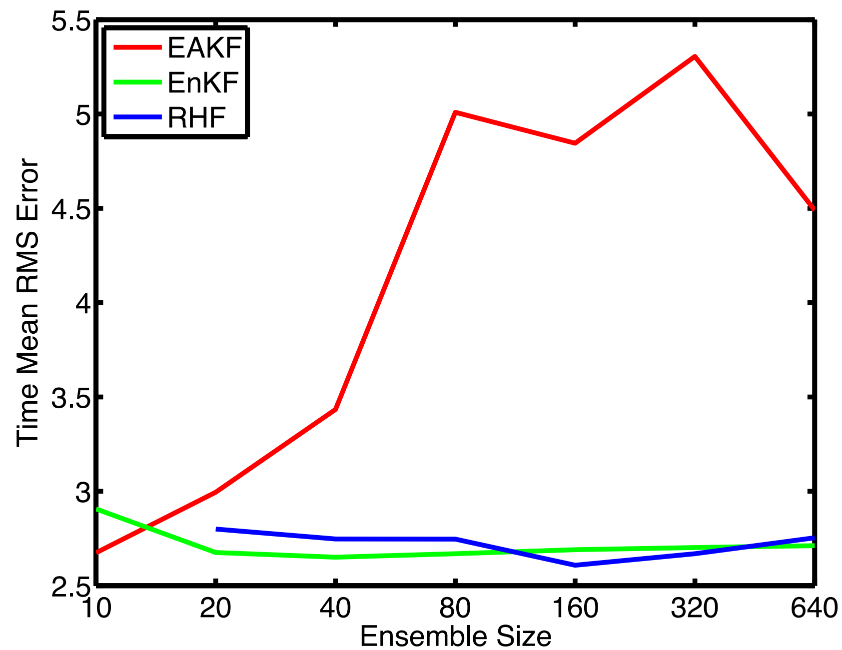
- Rank Histogram handles multimodal prior and compelling observation.
- EAKF still bimodal; left mode is inconsistent with everything.
- Lorenz_63 can have priors like this.

Multimodal Prior Distributions



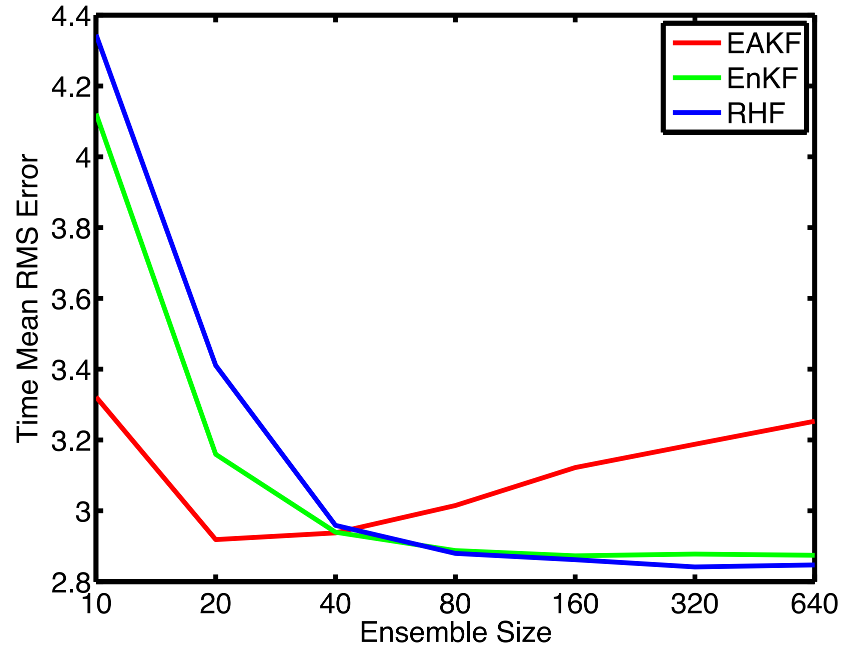
- Convective scale models (and land models) have analogous behavior.
- Convection may fire at ‘random’ locations.
- Subset of ensembles will be in right place, rest in wrong place.
- Want to aggressively eliminate convection in wrong place.

Results: Lorenz63 RMS



- All 3 state variables observed, error variance 1.0.
- RHF and EnKF comparable.
- EAKF gets progressively worse (but pretty good for 10 members).

Results: Lorenz96 RMS



- 40 Observations, average of adjacent state variables, Error var = 4.
- Localization halfwidth 0.3 of domain, adaptive inflation.
- RHF comparable to EnKF.

Results: Global NWP in Finite Volume CAM

- Prior fit to observations as metric:
- 80-member EAKF and RHF virtually indistinguishable.
  (Comparable to NCEP operational, better in tropics, near sfc.).
- 80-member EnKF significantly worse.

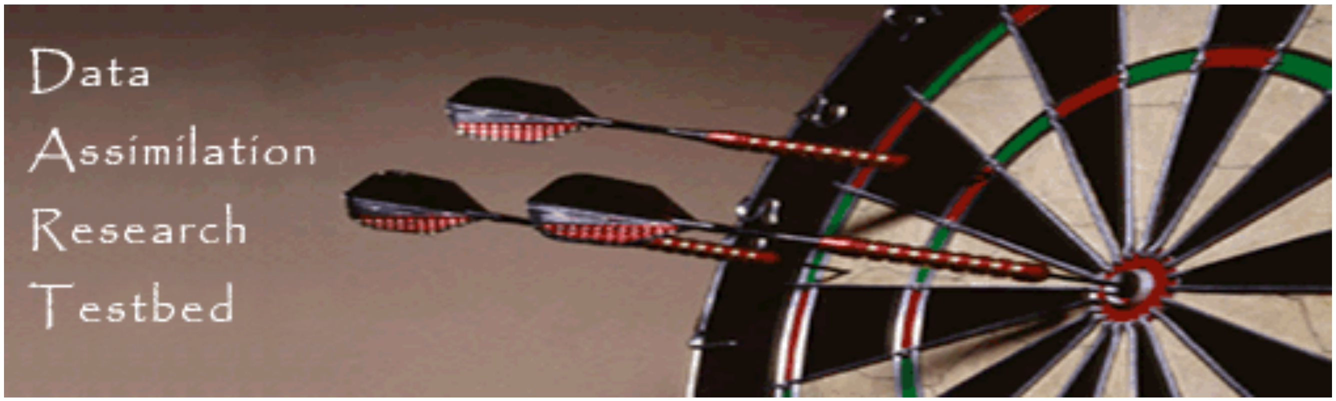
Additional Capabilities of RHF

- 1. Observations with highly non-gaussian observation likelihoods:
  Bounded quantities like RH, precip., or reflectivity.
  Just need to evaluate likelihood at prior locations (caveat tails).

- 2. Priors that are highly non-gaussian:
  Non-linear forward operators like radiances.

- 3. Ability to deal with discrete structure priors:
  Example: Convective scale.
  Subset of priors may have convection in a given location.
  Posterior should be either yes or no, not maybe.

Code to implement all of the algorithms discussed are freely available from:



http://www.image.ucar.edu/DAReS/DART/