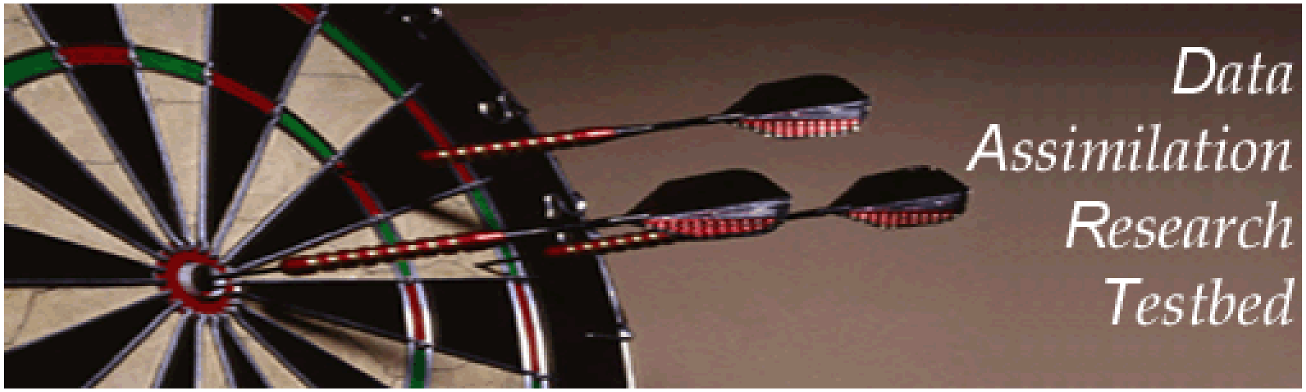
A Boxcar Kernel Filter for Assimilation of Discrete Structures (and Other Stuff)

Jeffrey Anderson
NCAR Data Assimilation Research Section (DAReS)



Background:

1. Most ensemble filters assume prior and likelihood are ~gaussian.

2. Particle filters do full non-gaussian, but don’t scale.

3. Assuming non-gaussian only in observation space is possible.

4. Gaussian kernel filters have been proposed but work poorly.

Requirements for an observation space kernel filter:

1. Low information content obs. can’t lead to large increments.

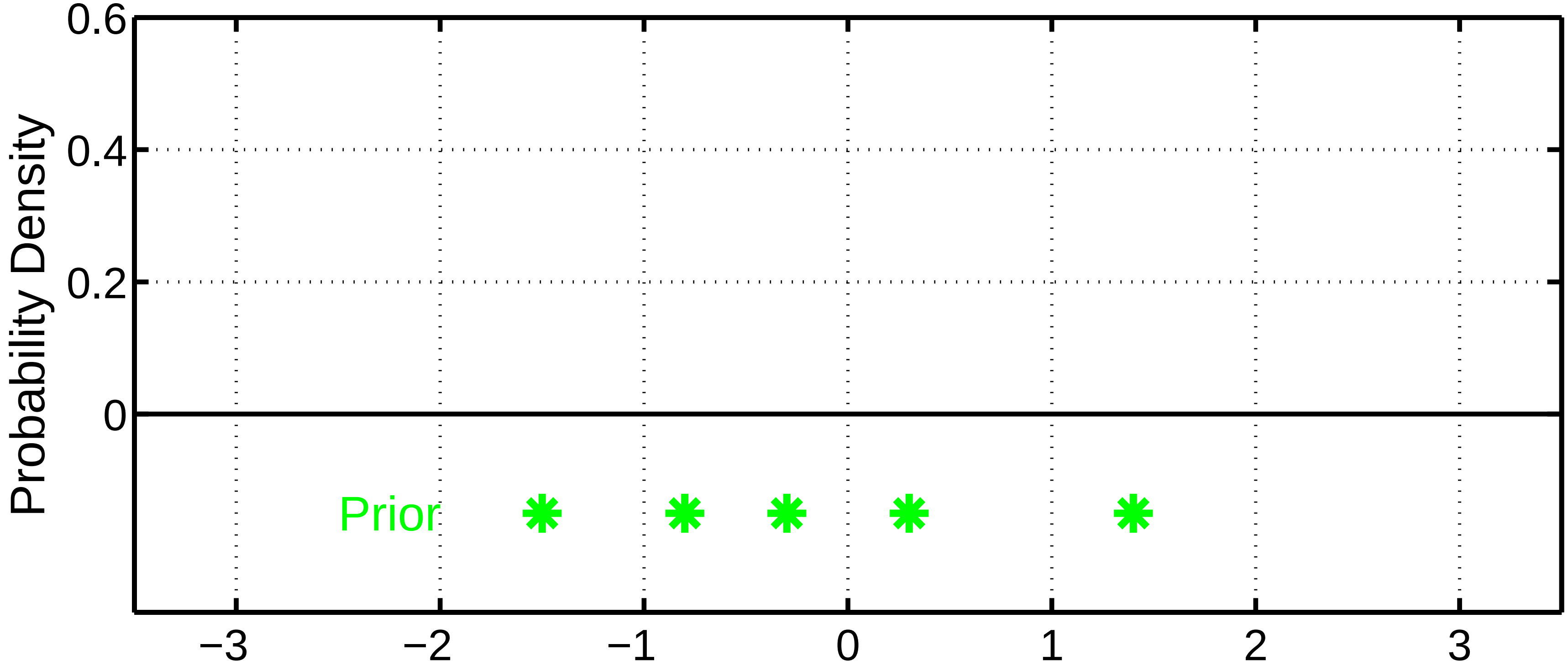
2. Want small increments for all cases.

3. Comparable to gaussian filters for ~gaussian cases.

4. Better than gaussian in non-gaussian cases.

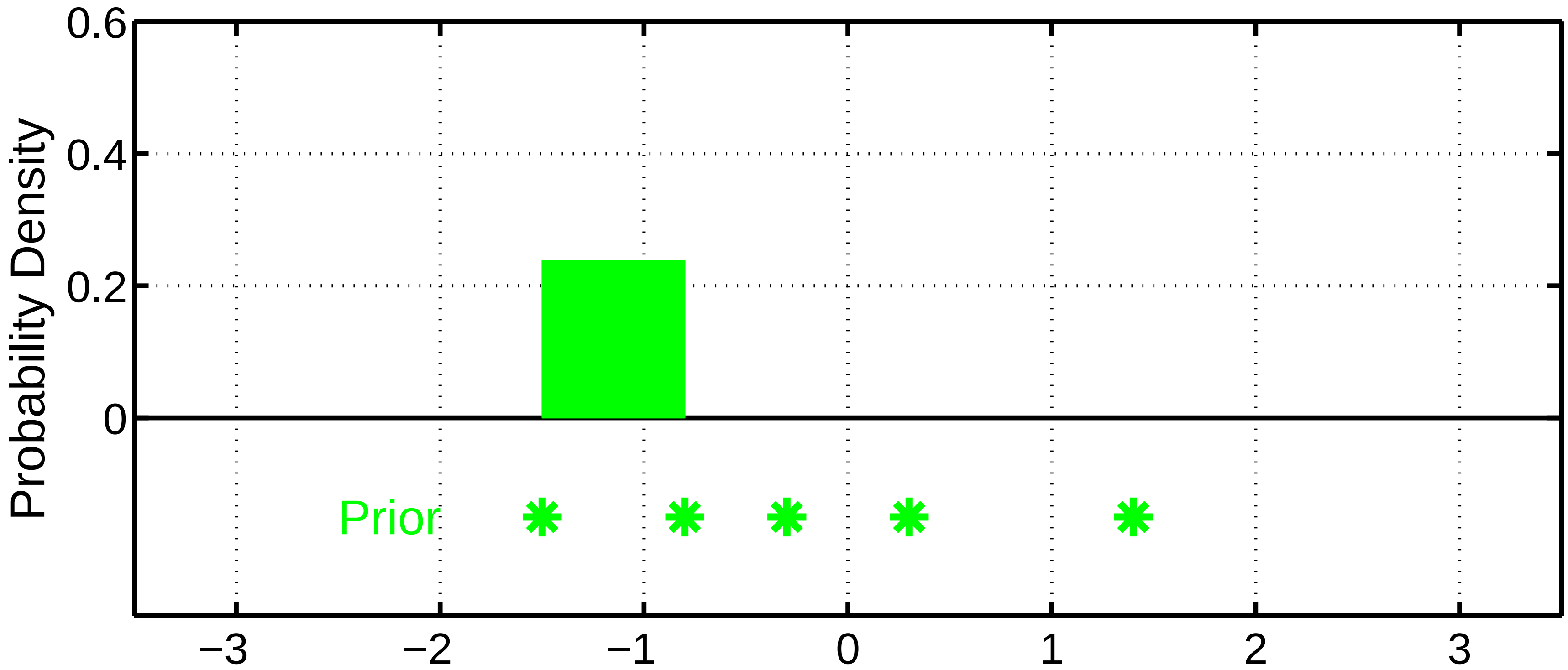
4. Must be computationally cheap.

Observation Space Boxcar Kernel Filter



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

Observation Space Boxcar Kernel Filter

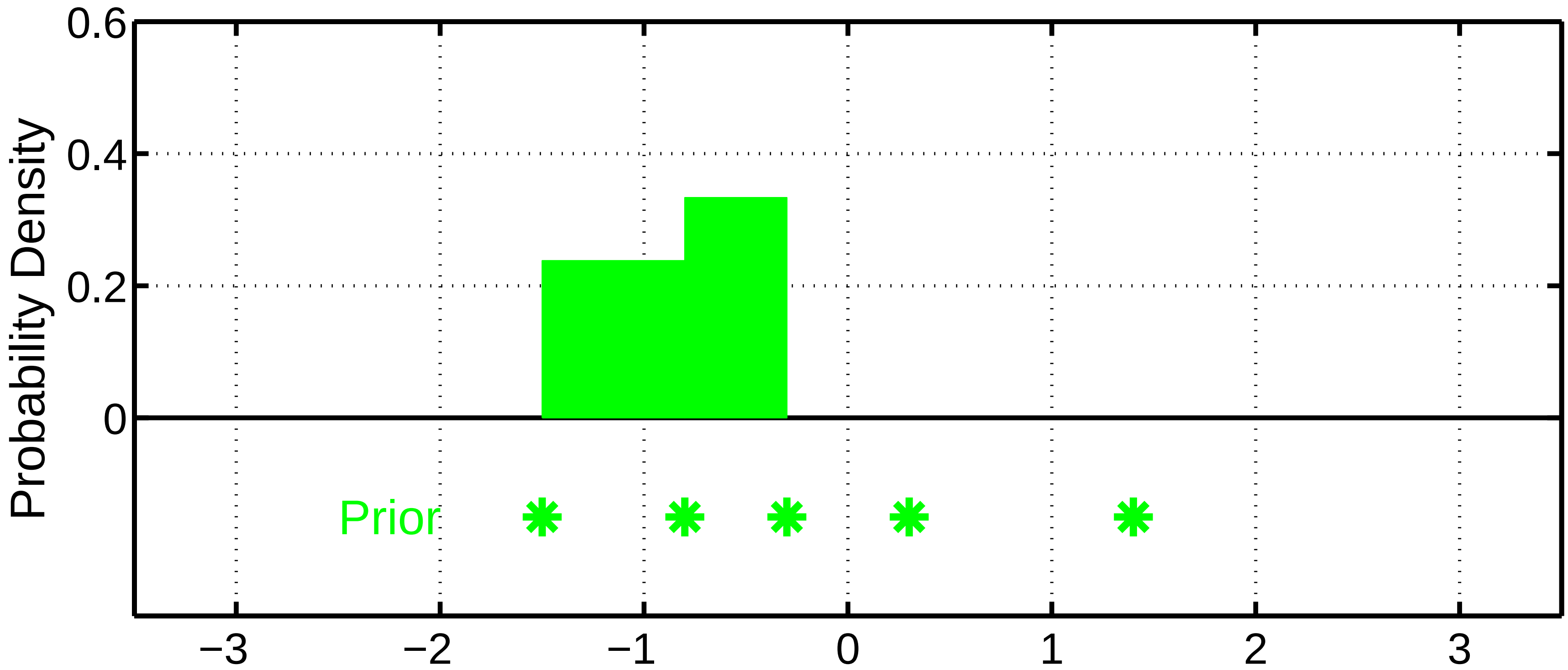


Step 1: Get continuous prior distribution density.

Place (ens_size + 1)$^{-1}$ mass between adjacent ensemble members.

Reminiscent of rank histogram evaluation method.

Observation Space Boxcar Kernel Filter

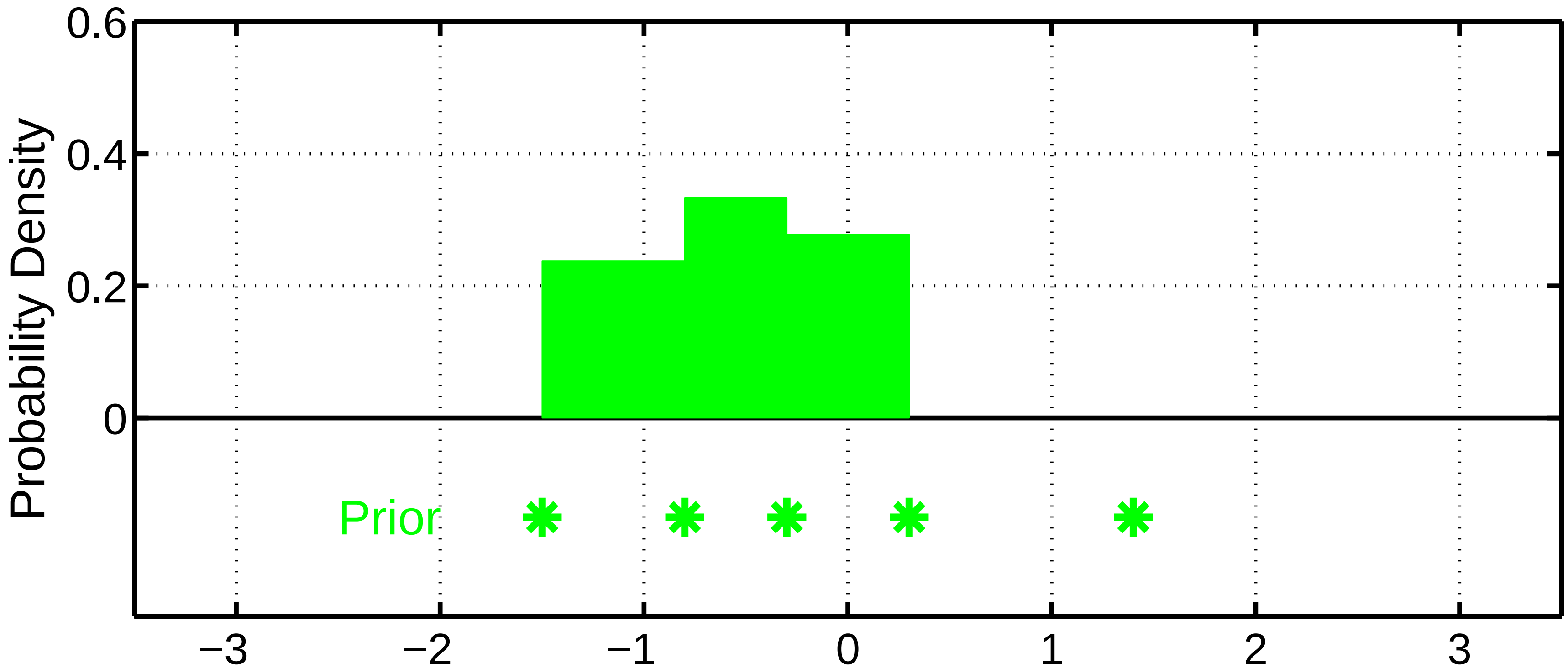


Step 1: Get continuous prior distribution density.

Place $(\text{ens\_size} + 1)^{-1}$ mass between adjacent ensemble members.

Reminiscent of rank histogram evaluation method.

Observation Space Boxcar Kernel Filter

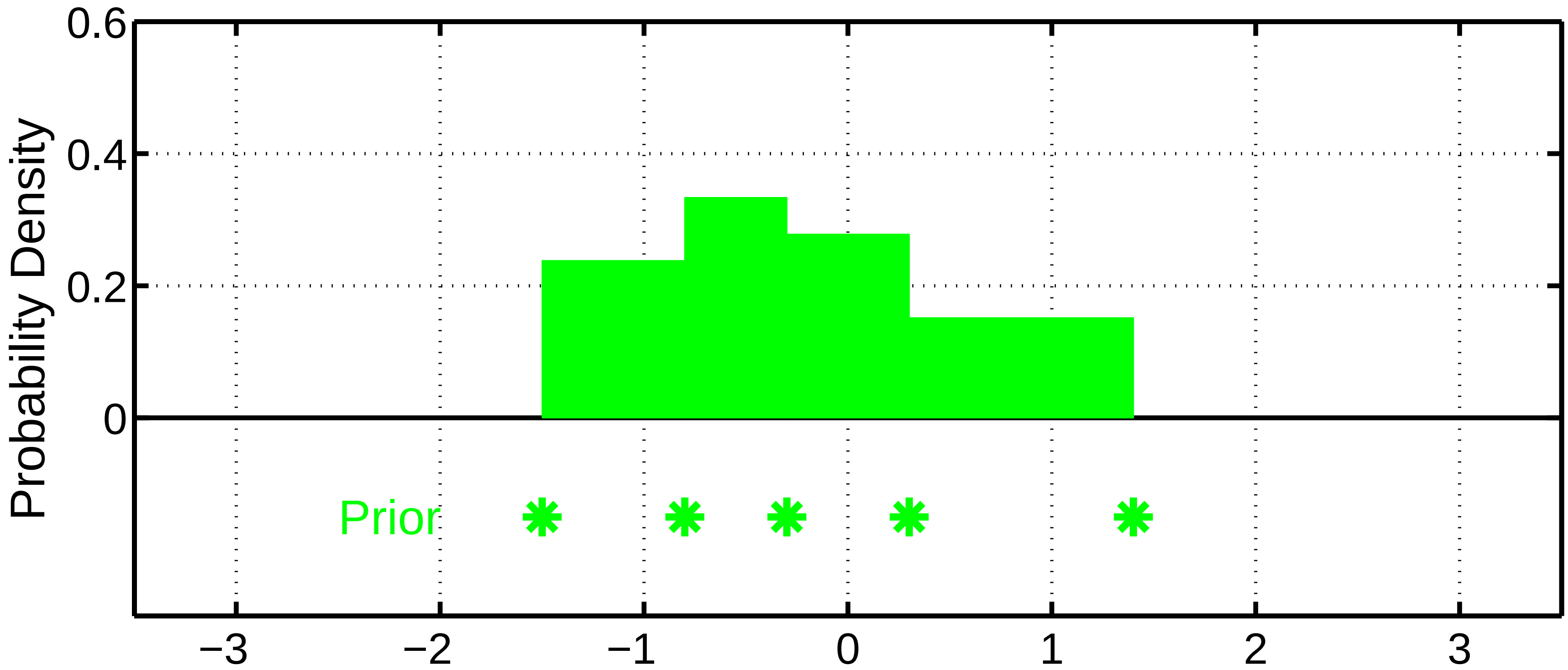


Step 1: Get continuous prior distribution density.

Place $(\text{ens\_size} + 1)^{-1}$ mass between adjacent ensemble members.

Reminiscent of rank histogram evaluation method.

Observation Space Boxcar Kernel Filter

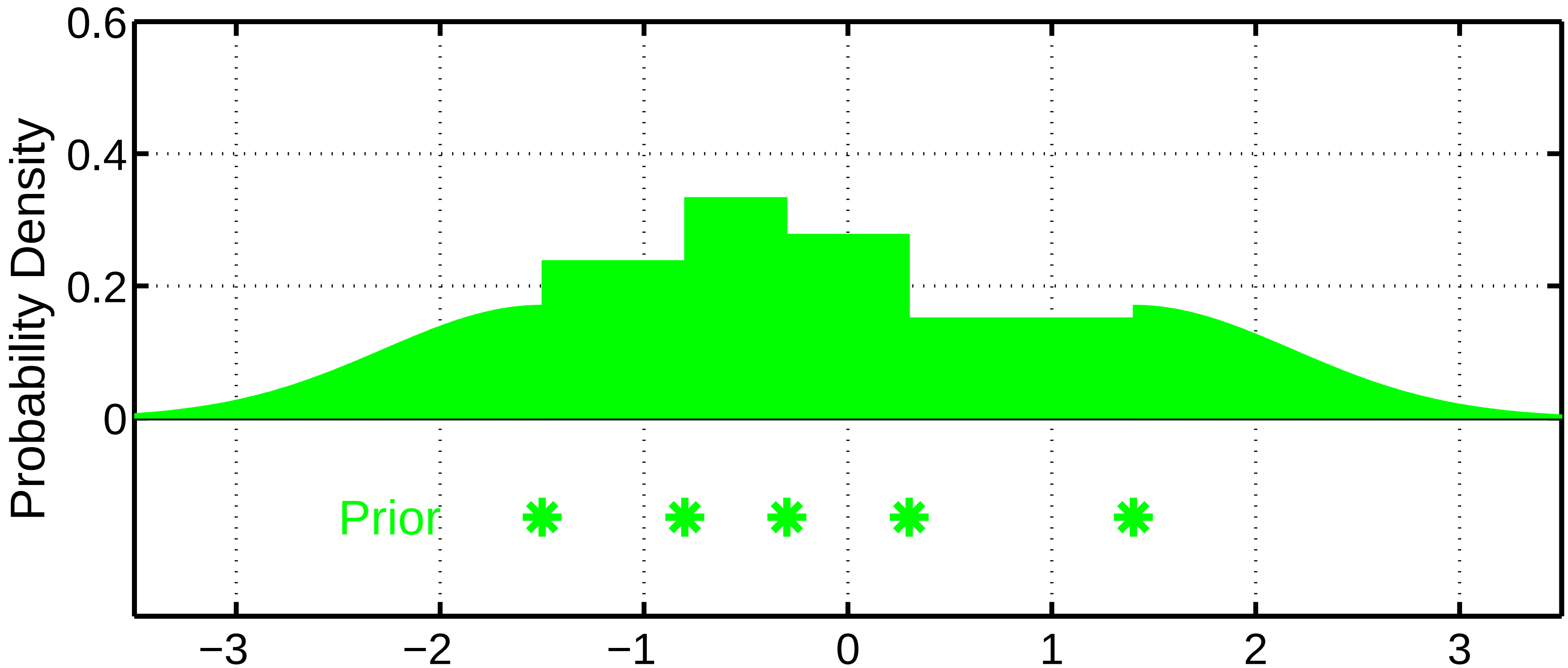


Step 1: Get continuous prior distribution density.

Place (ens_size + 1)$^{-1}$ mass between adjacent ensemble members.

Reminiscent of rank histogram evaluation method.

Observation Space Boxcar Kernel Filter



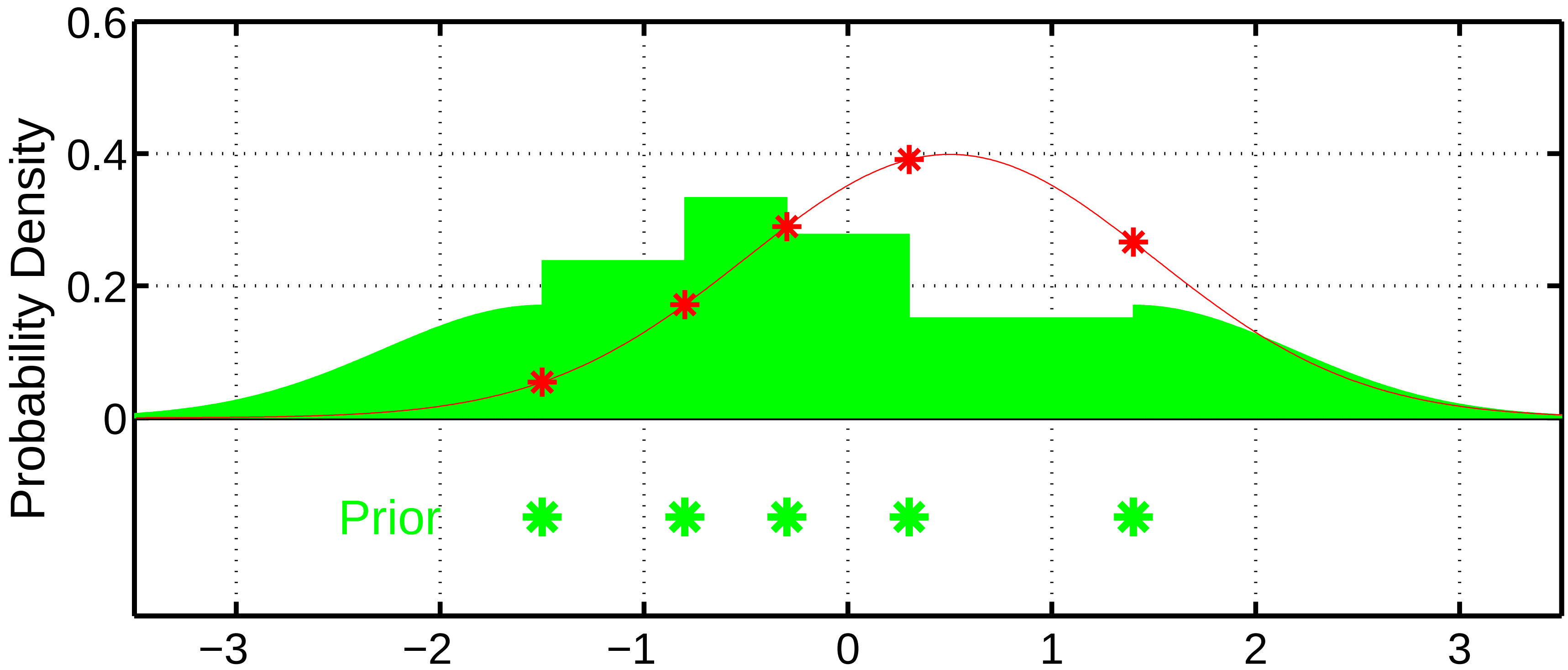
Step 1: Get continuous prior distribution density.

Place $(\text{ens\_size} + 1)^{-1}$ mass between adjacent ensemble members.

Half-gaussian kernels on tails, N(Outer ensemble, σ_{ens} /2).

These prevent filter divergence in presence of model error.

Observation Space Boxcar Kernel Filter

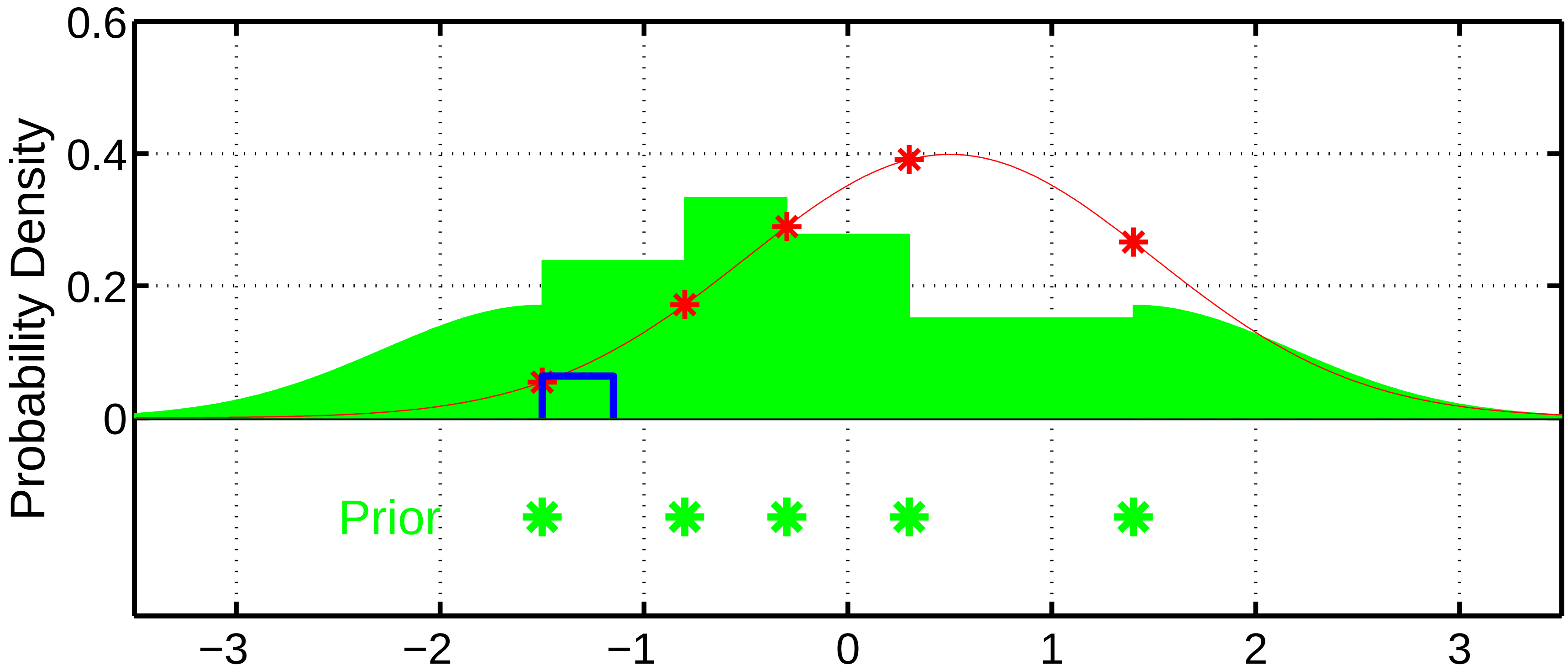


Step 2: Use likelihood to compute weight for each ensemble member.

Analogous to classical particle filter.

Can be easily extended to non-gaussian obs. likelihoods.

Observation Space Boxcar Kernel Filter

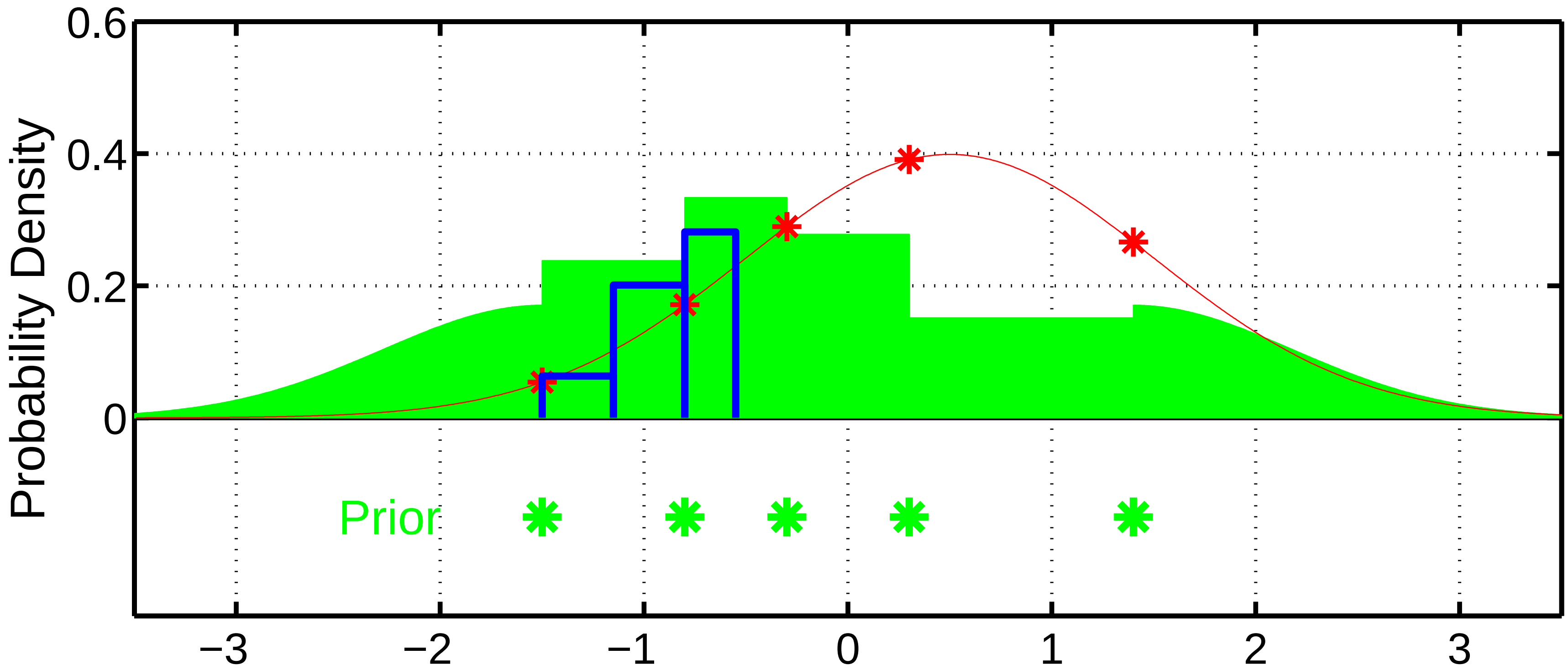


Step 3: Compute continuous posterior distribution.

1. Split each uniform box in half; multiply mass by adjacent likelihood.
   (Outermost ensemble members have only one associated half box).

This is just quadrature computation of posterior.

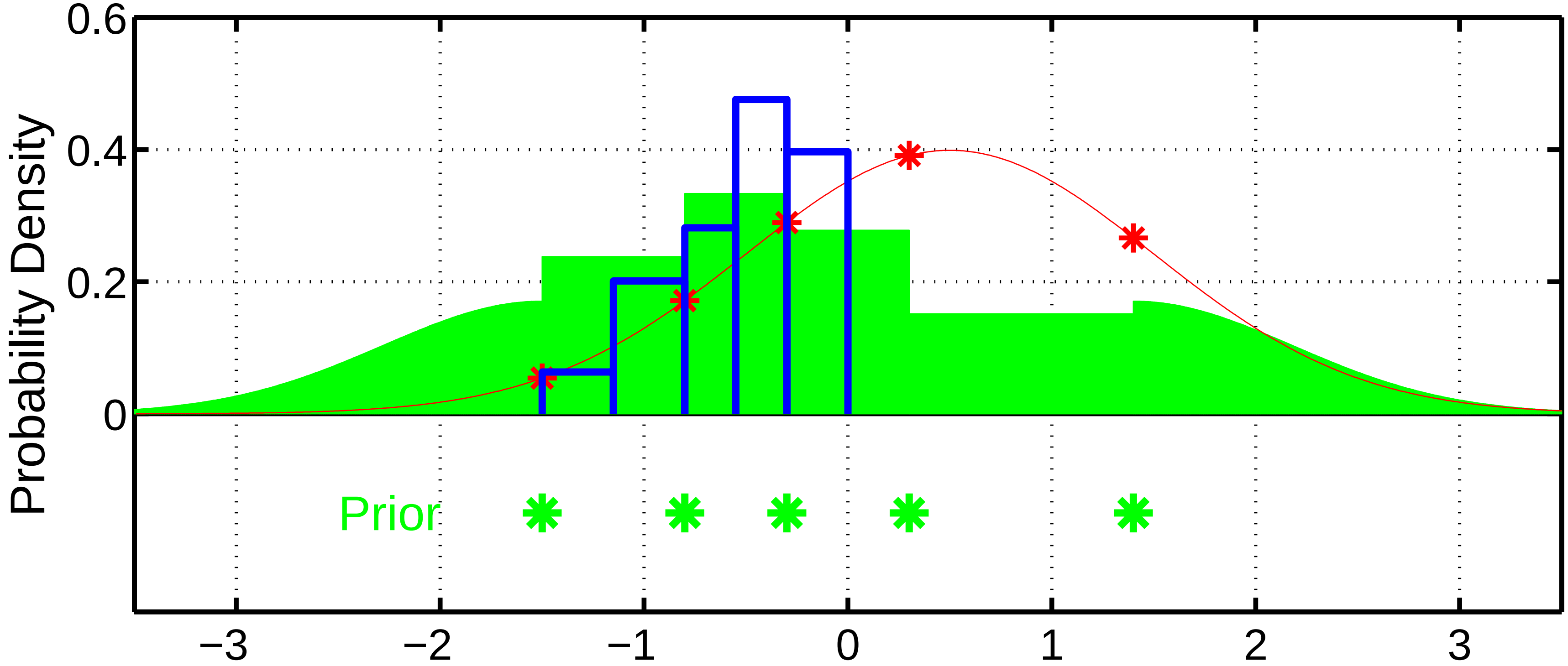
Observation Space Boxcar Kernel Filter



Step 3: Compute continuous posterior distribution.

1. Split each uniform box in half; multiply mass by adjacent likelihood.

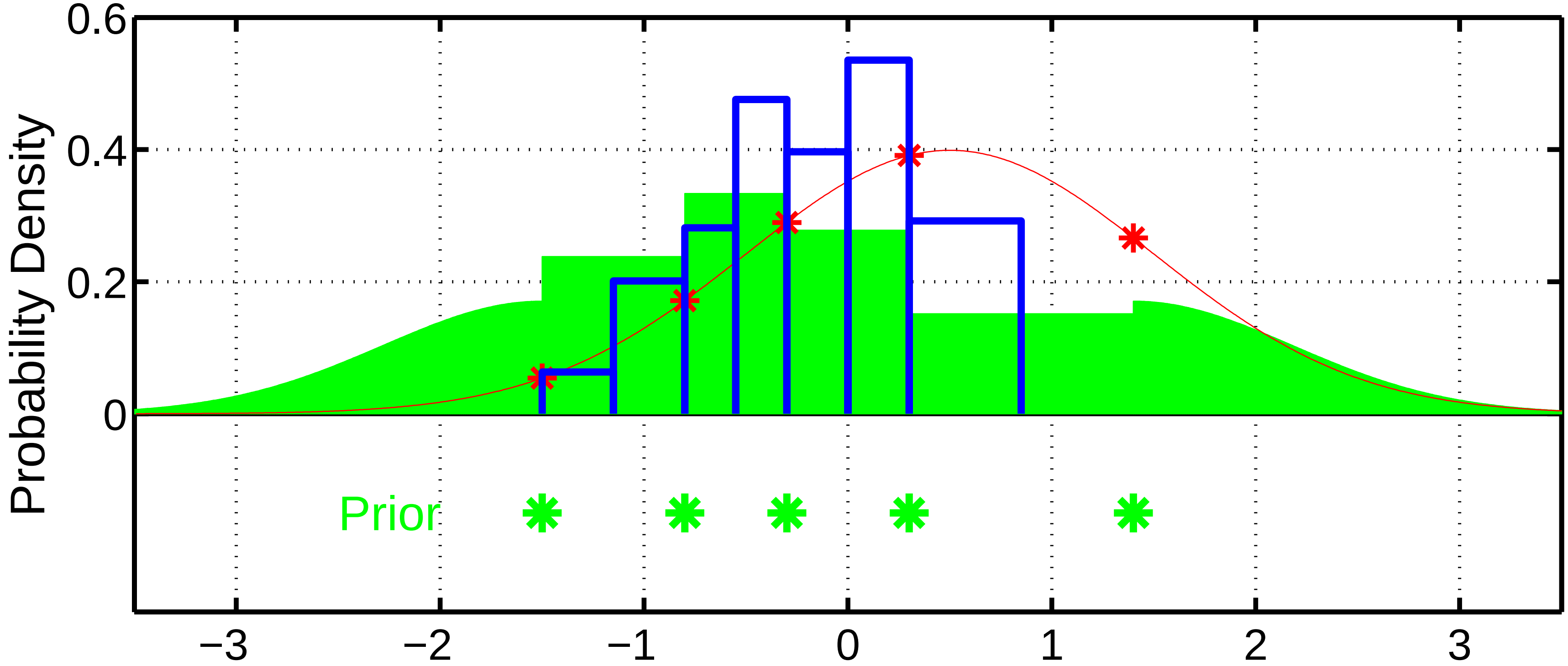
Observation Space Boxcar Kernel Filter



Step 3: Compute continuous posterior distribution.

1. Split each uniform box in half; multiply mass by adjacent likelihood.

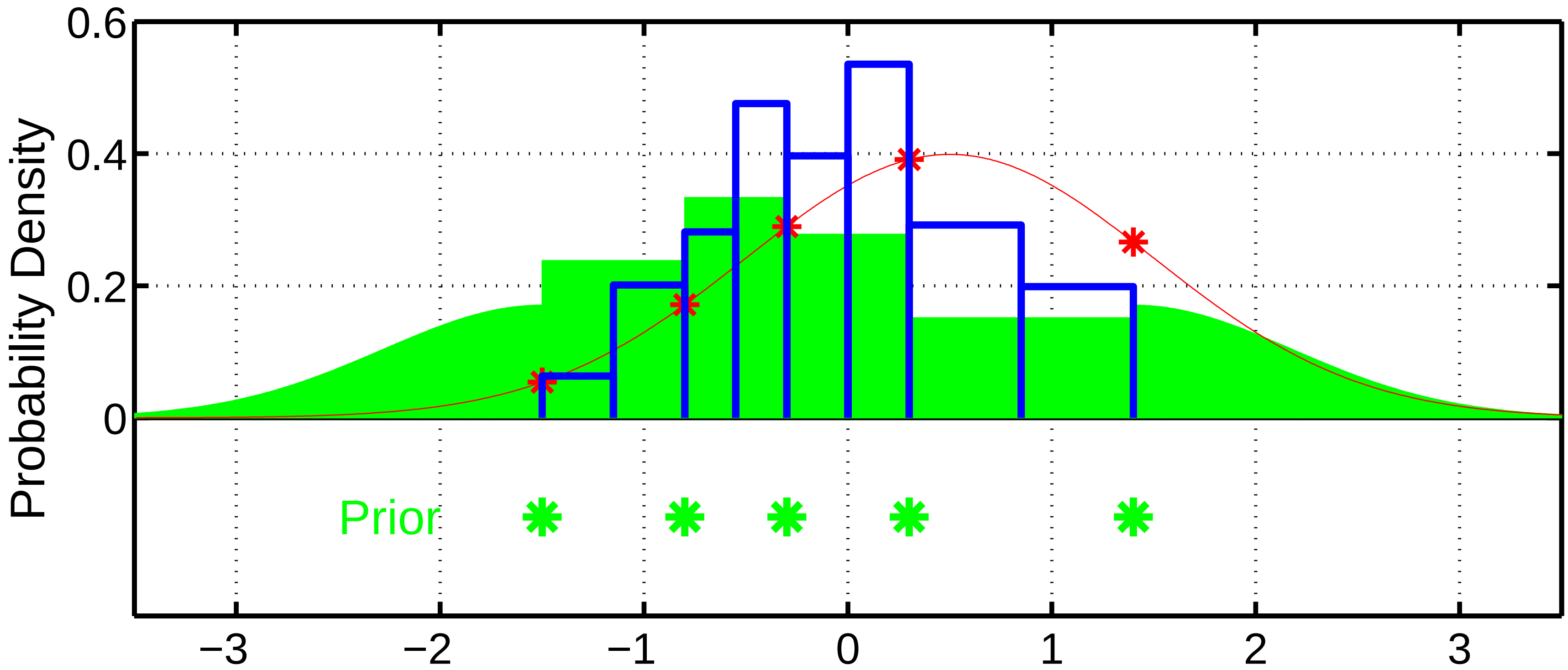
Observation Space Boxcar Kernel Filter



Step 3: Compute continuous posterior distribution.

1. Split each uniform box in half; multiply mass by adjacent likelihood.

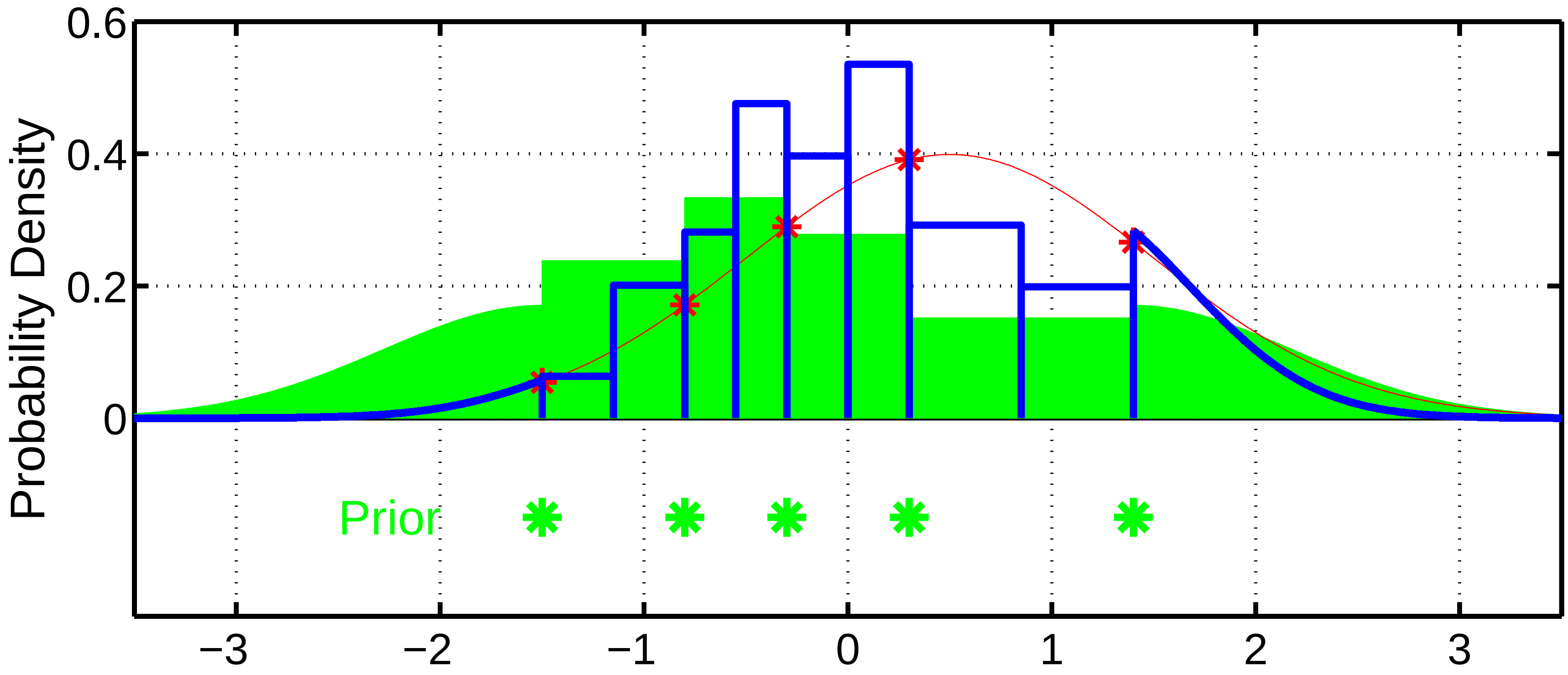
Observation Space Boxcar Kernel Filter



Step 3: Compute continuous posterior distribution.

1. Split each uniform box in half; multiply mass by adjacent likelihood.
   (Outermost ensemble members have only one associated half box).

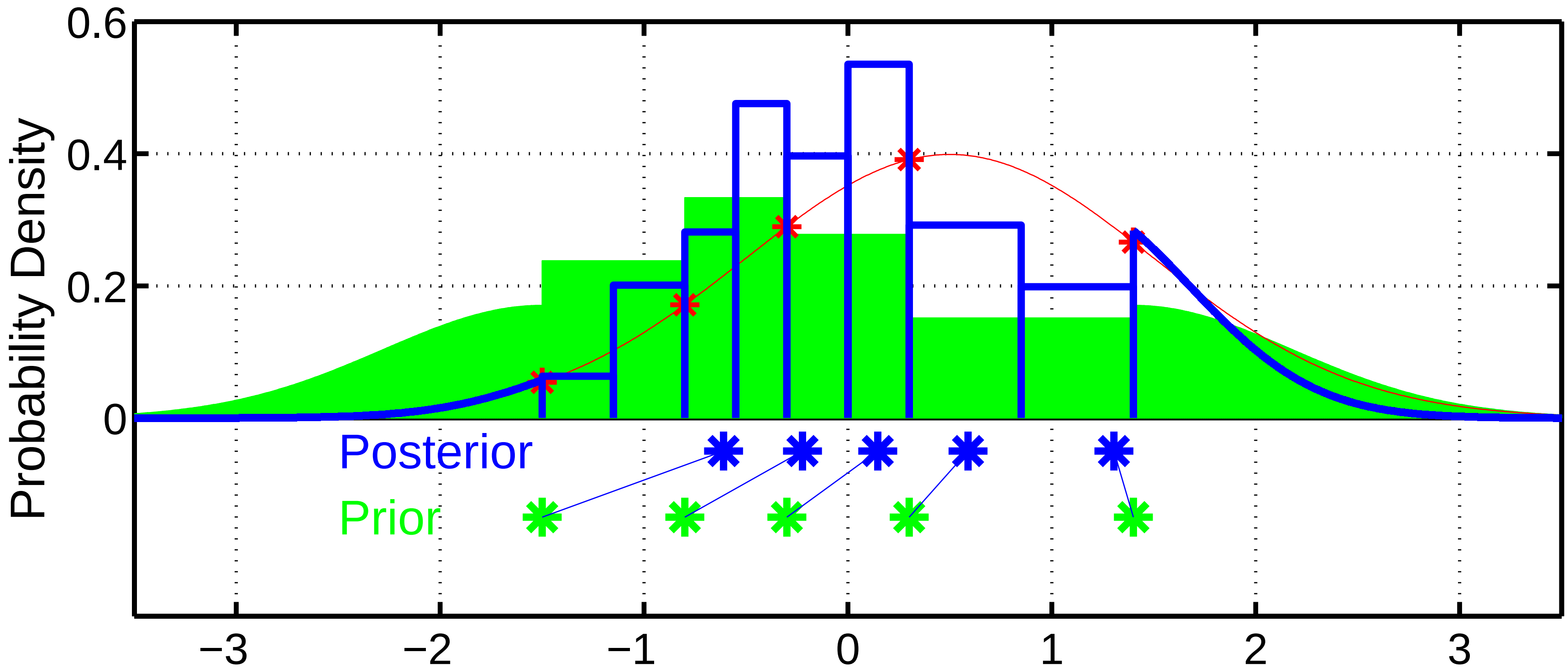
Observation Space Boxcar Kernel Filter



Step 3: Compute continuous posterior distribution.

1. Split each uniform box in half; multiply mass by adjacent likelihood.
2. Product of prior gaussian kernel with likelihood for tails.
    Easy for gaussian likelihood.
    More quadrature if non-Gaussian likelihood.

Observation Space Boxcar Kernel Filter



Step 4: Compute updated ensemble members:

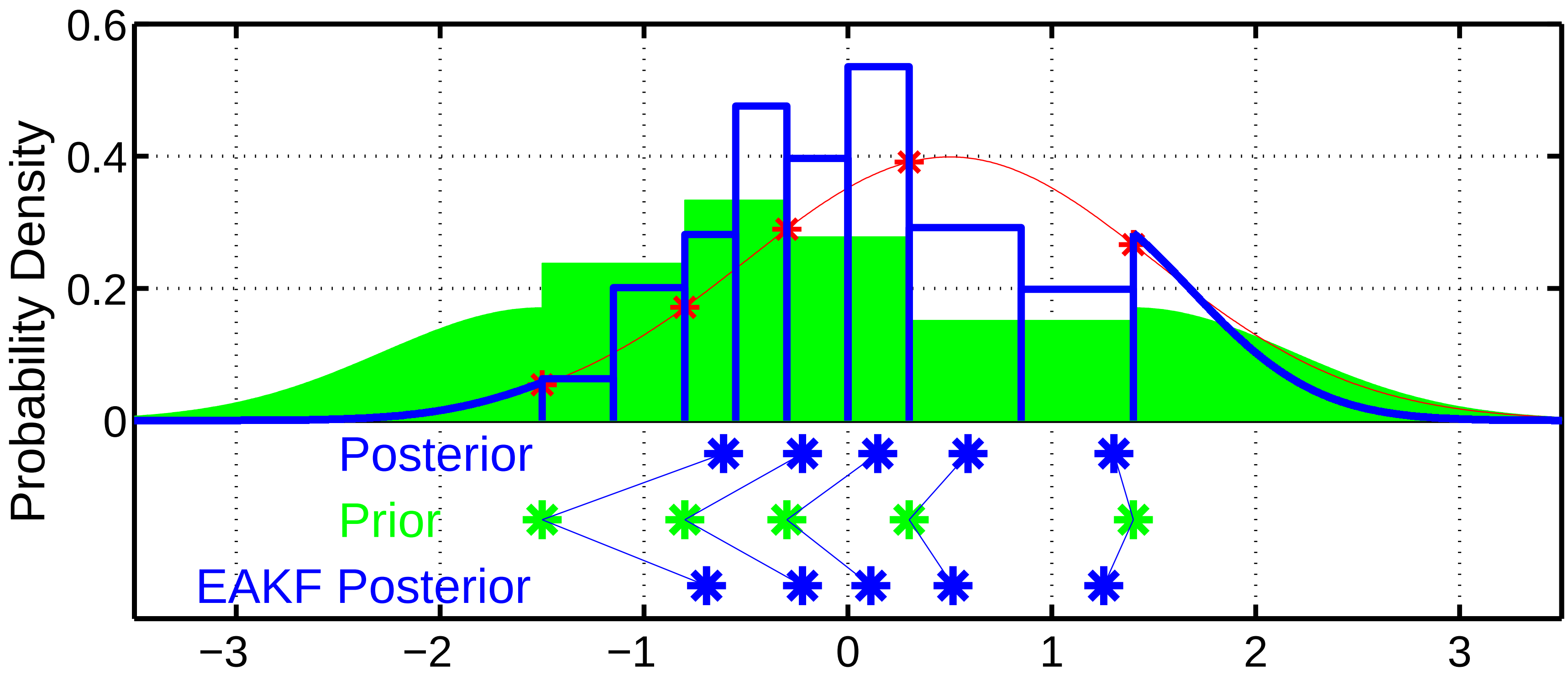
$(\text{ens\_size} + 1)^{-1}$ of posterior mass between each ensemble pair.

$(\text{ens\_size} + 1)^{-1}$ in each wing.

Trivial to compute cumulative density if likelihood is gaussian.

Uninformative observation has no impact.

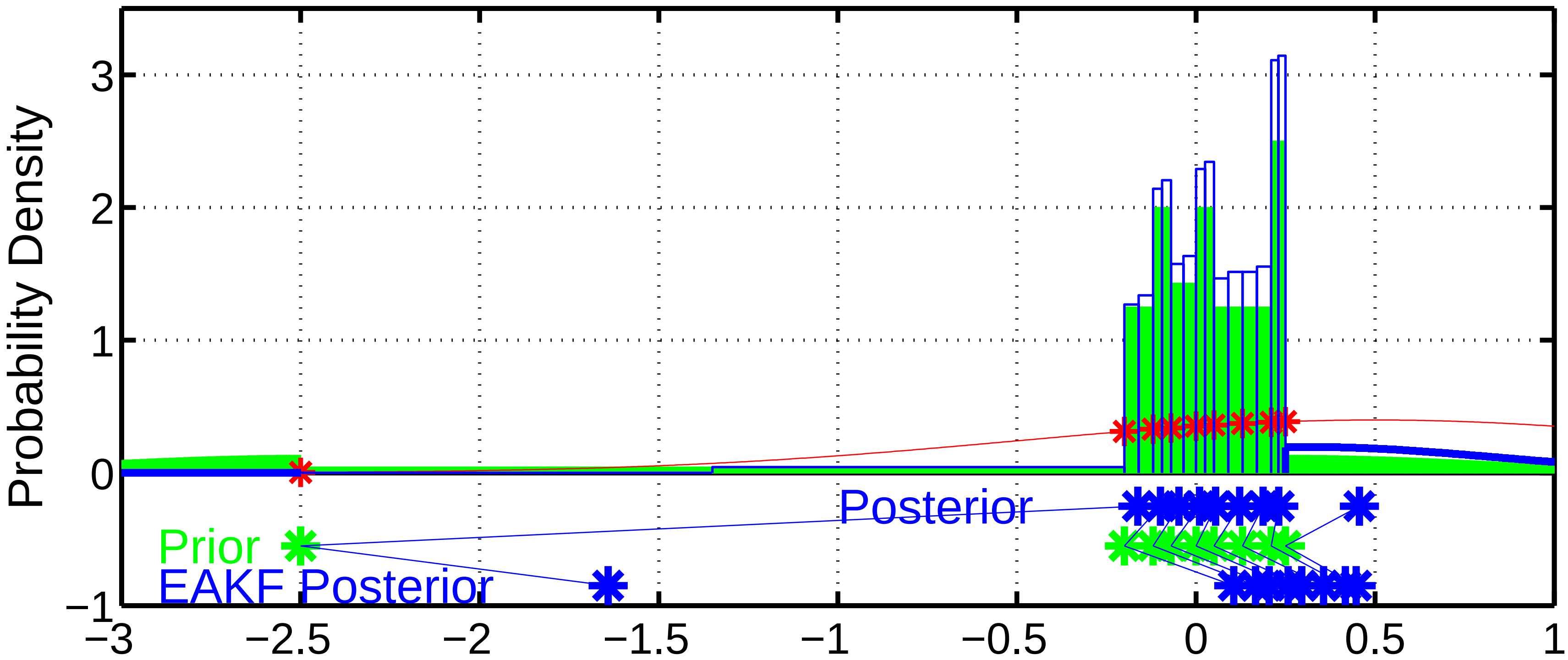
Observation Space Boxcar Kernel Filter



Compare to standard Ensemble Adjustment Filter (EAKF).

In this nearly gaussian case, differences in increments are small.

Outliers are a Challenge for Gaussian Filters

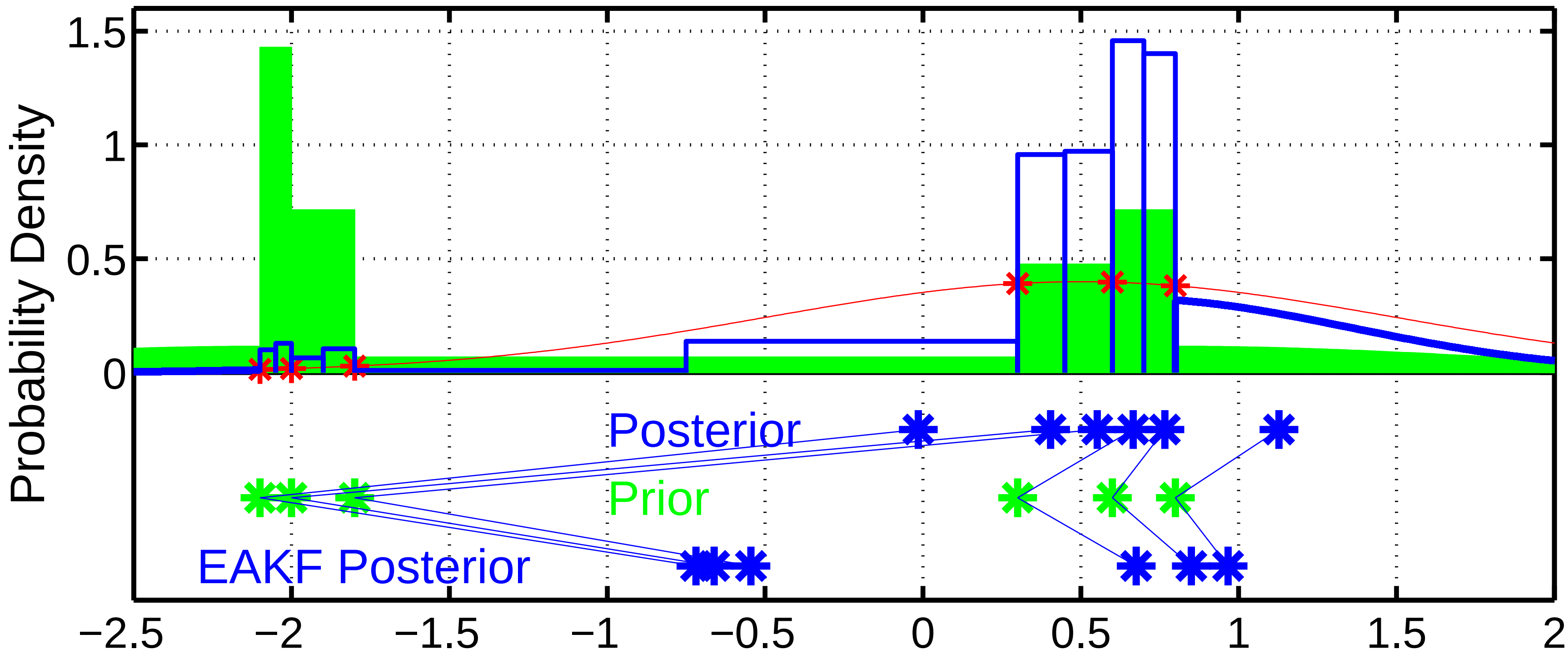


Boxcar gets rid of outlier that is clearly inconsistent with obs.

EAKF can't get rid of outlier.

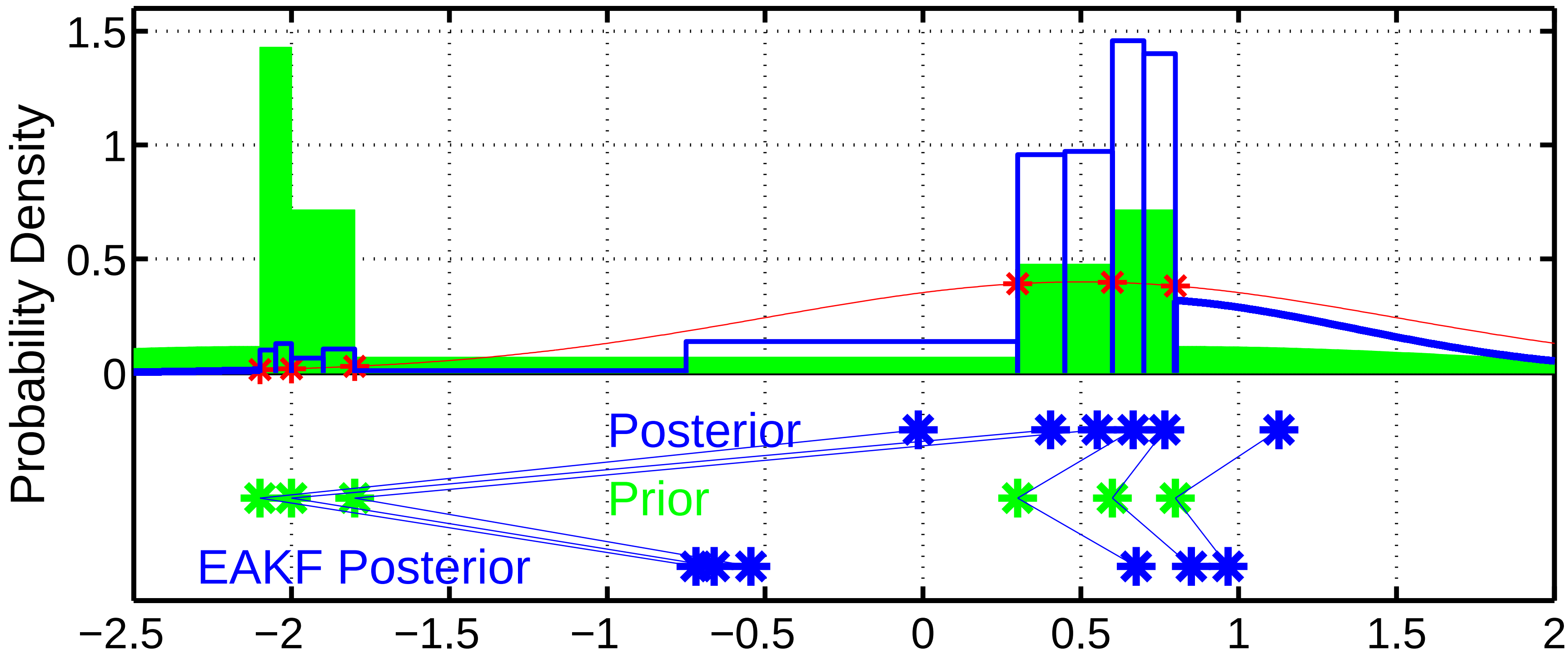
Large prior variance from outlier causes EAKF to shift all members too much towards observation.

Multimodal Prior Distributions



Boxcar can deal with multimodal prior with a compelling observation.
EAKF still bimodal; left mode is inconsistent with everything.

Multimodal Prior Distributions



Convective scale 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.

Boxcar results compared to conventional EAKF and EnKF:

1. Lorenz-63, infrequent observations (ideally suited for boxcar).
   Boxcar significantly better for all ensemble sizes.

2. Lorenz-96 (40 variable) with significant model error:
   Boxcar slightly worse for 20 members.
   Boxcar slightly better for 80 members.

3. T85L26 CAM with *in situ* BUFR observations:
   Boxcar somewhat worse for 20 members.
   Still competitive with NCEP 2003 operational.

Future work:

1. Need to test in models with discrete structures.

2. Understand occurrence of outliers in ensemble assimilation.

3. Study enhancements for tails.

4. Non-gaussian likelihoods?

5. Bounded observed quantities like tracer concentration.
   Appropriate priors/likelihoods may be log normal or beta.

Want to try it out?

The boxcar and 6 other ensemble update variants are in DART.

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