Ensemble Filters for Parameter Estimation:

Jeffrey Anderson
NCAR Data Assimilation Research Section (DAReS)

IMAGe TOY Workshop: 27 February, 2006

Prelude: The Data Assimilation Problem:

Given: 1. A physical system (atmosphere, ocean...)

---

2. Observations of the physical system

Usually sparse and irregular in time and space.
Instruments have error of which we have a (poor) estimate.
Observations may be of ‘non-state’ quantities.
Many observations may have very low information content.

---

3. A model of the physical system

Usually thought of as approximating time evolution.
Could also be just a model of balance (attractor) relations.
Truncated representation of ‘continuous’ physical system.
Often quasi-regular discretization in space and/or time.
Generally characterized by ‘large’ systematic errors.
May be ergodic with some sort of ‘attractor’.

We want to increase our information about all three pieces:

1. Get an improved estimate of state of physical system

Includes time evolution and ‘balances’.
Initial conditions for forecasts.
High quality analyses (re-analyses).

2. Get better estimates of observing system error characteristics

Estimate value of existing observations.
Design observing systems that provide increased information.

3. Improve model of physical system

Evaluate model systematic errors.
Select appropriate values for model parameters.
Evaluate relative characteristics of different models.

Seminar Road Map:

1: Single variable and observation of that variable.

2: Single observed variable, single unobserved variable.

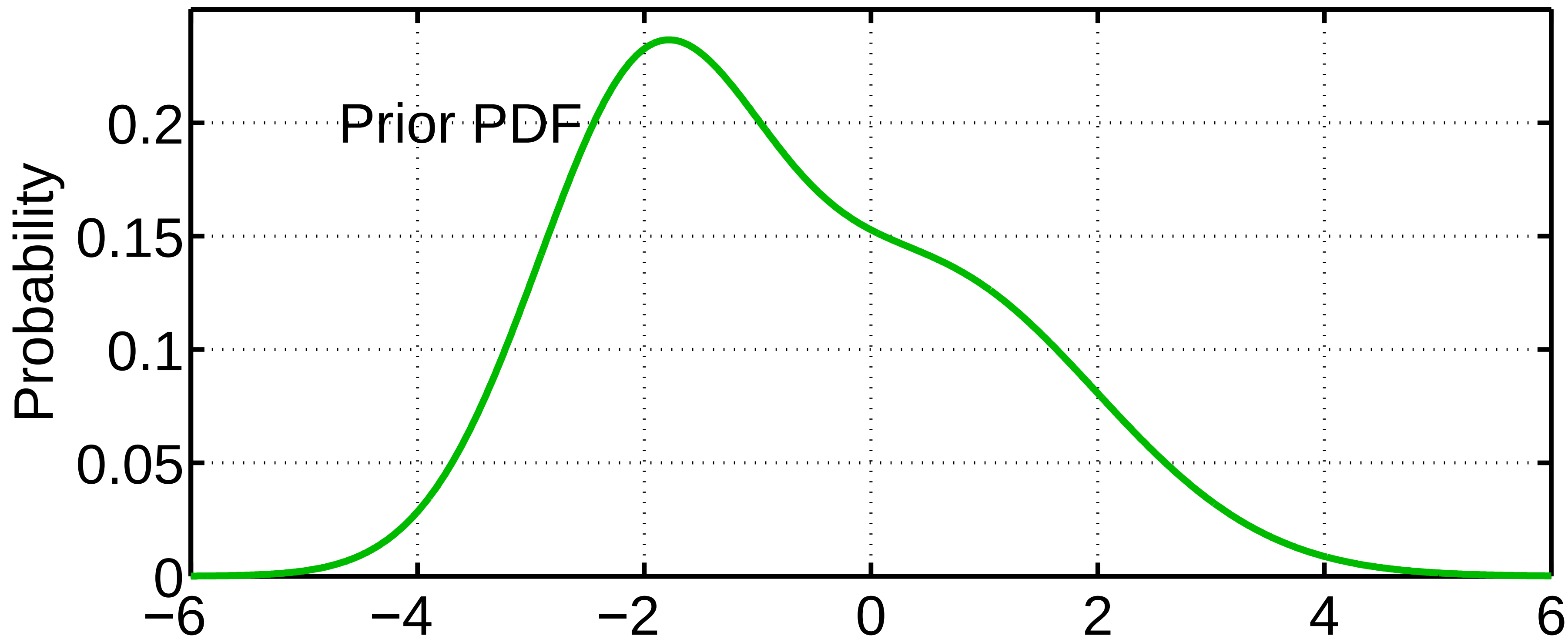
3: Generalize to geophysical models and observations.

4: Dealing with sampling and other errors.

5. Parameter estimation.

Tomorrow: Hierarchical Bayesian methods.

Bayes rule: $p(A|BC) = \frac{p(B|AC)p(A|C)}{p(B|C)} = \frac{p(B|AC)p(A|C)}{\int p(B|x)p(x|C)dx}$

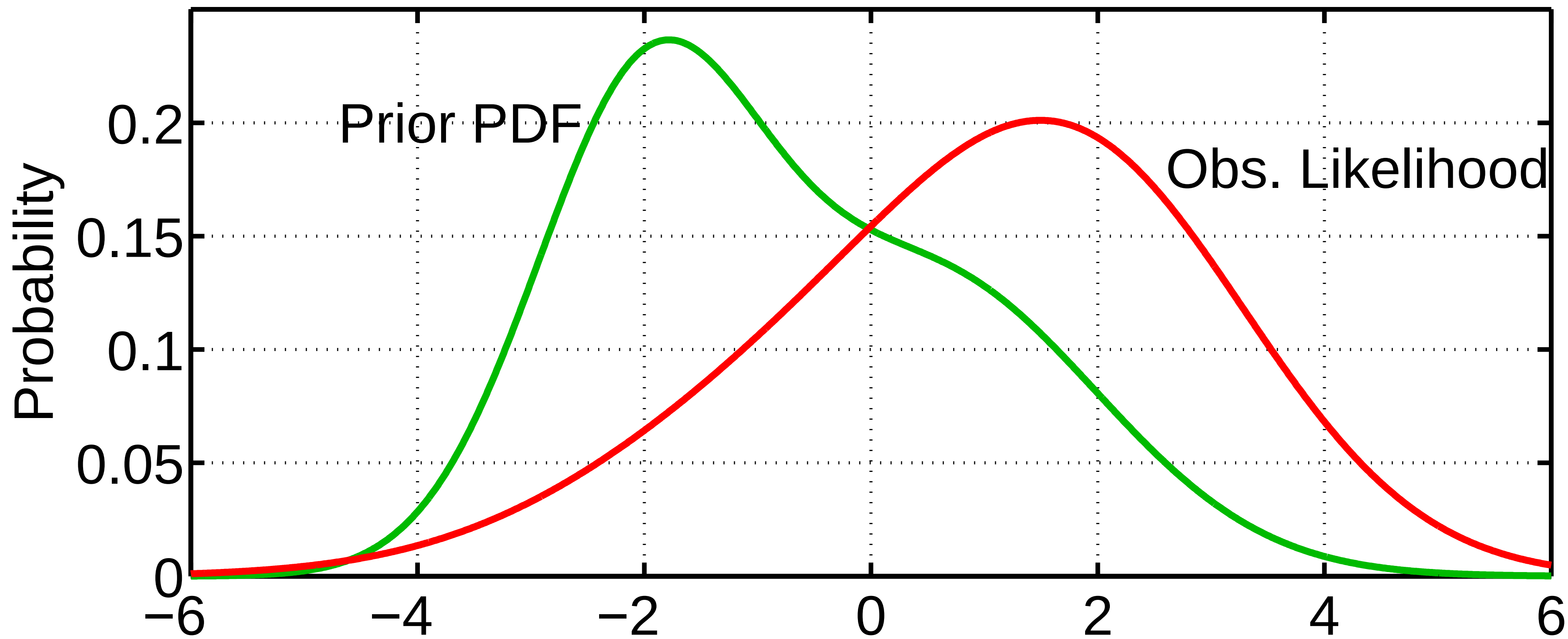


A: Prior estimate based on all previous information, C.

B: An additional observation.

$p(A|BC)$: Posterior (updated estimate) based on C and B.

Bayes rule: $p(A|BC) = \dfrac{p(B|AC)p(A|C)}{p(B|C)} = \dfrac{p(B|AC)p(A|C)}{\int p(B|x)p(x|C)dx}$

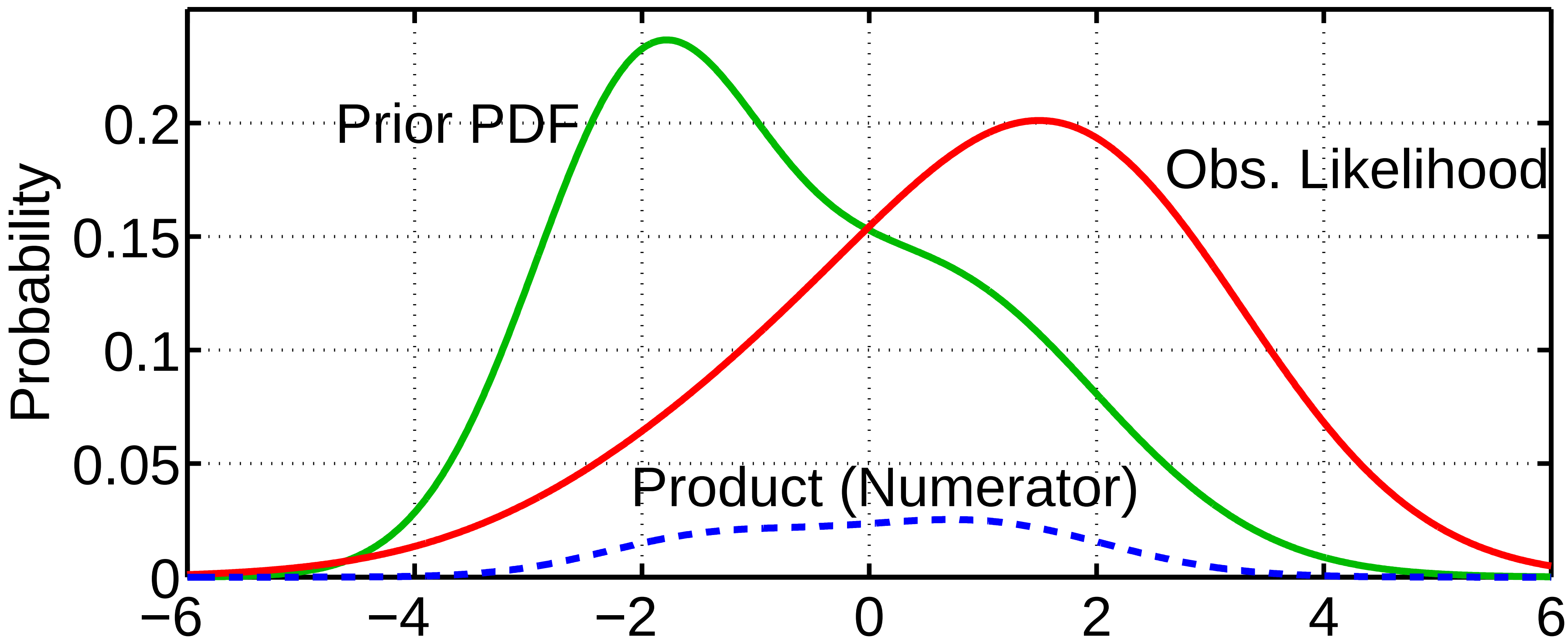


A: Prior estimate based on all previous information, C.

B: An additional observation.

$p(A|BC)$: Posterior (updated estimate) based on C and B.

Bayes rule: $p(A|BC) = \dfrac{p(B|AC)p(A|C)}{p(B|C)} = \dfrac{p(B|AC)p(A|C)}{\int p(B|x)p(x|C)dx}$

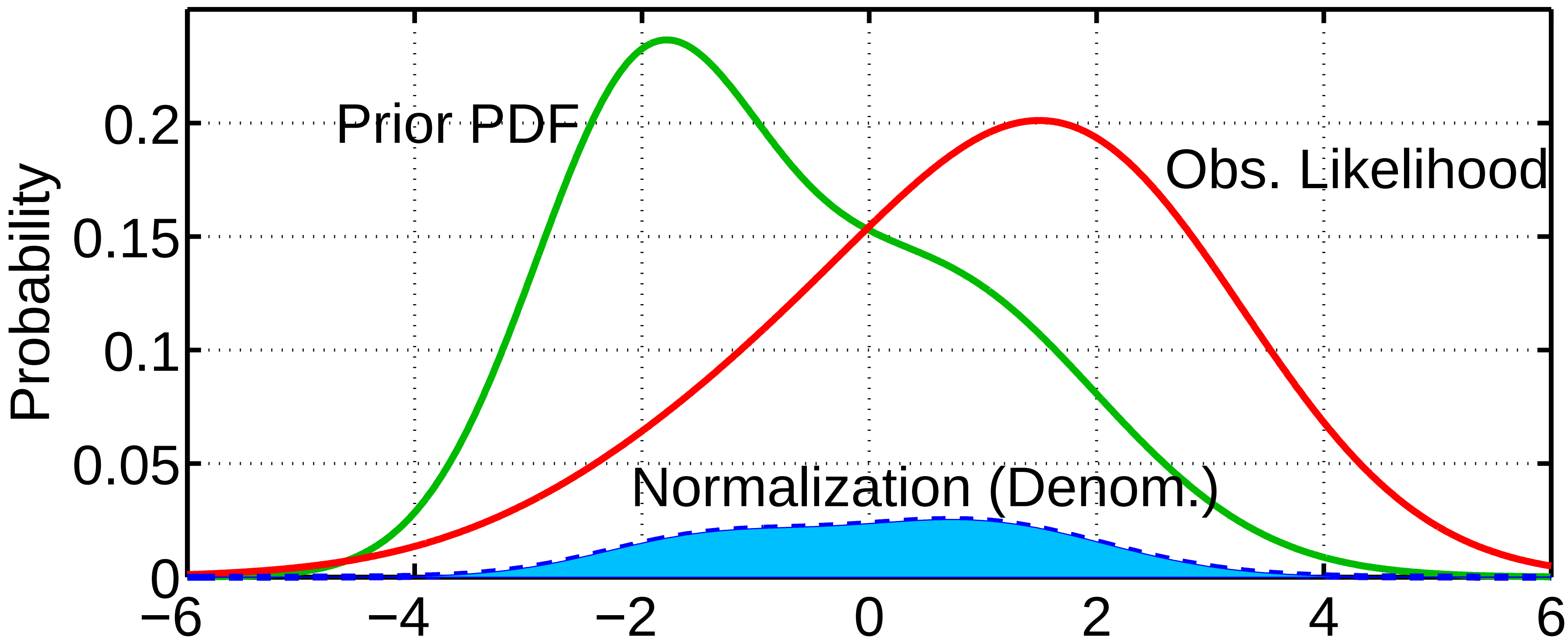


A: Prior estimate based on all previous information, C.

B: An additional observation.

$p(A|BC)$: Posterior (updated estimate) based on C and B.

Bayes rule: $p(A|BC) = \frac{p(B|AC)p(A|C)}{p(B|C)} = \frac{p(B|AC)p(A|C)}{\int p(B|x)p(x|C)dx}$

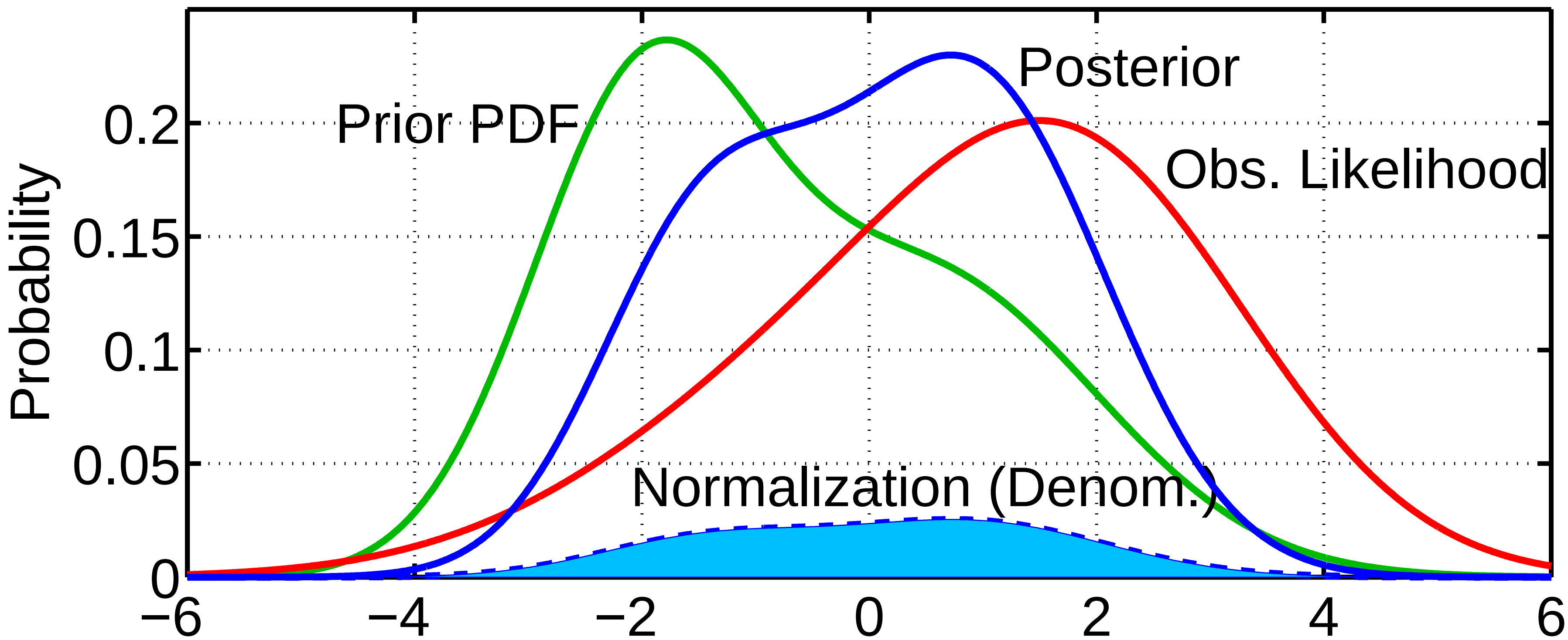


A: Prior estimate based on all previous information, C.

B: An additional observation.

$p(A|BC)$: Posterior (updated estimate) based on C and B.

Bayes rule: $p(A|BC) = \dfrac{p(B|AC)\,p(A|C)}{p(B|C)} = \dfrac{p(B|AC)\,p(A|C)}{\int p(B|x)\,p(x|C)\,dx}$



A: Prior estimate based on all previous information, C.

B: An additional observation.

$p(A|BC)$: Posterior (updated estimate) based on C and B.

Consistent Color Scheme Throughout Tutorial

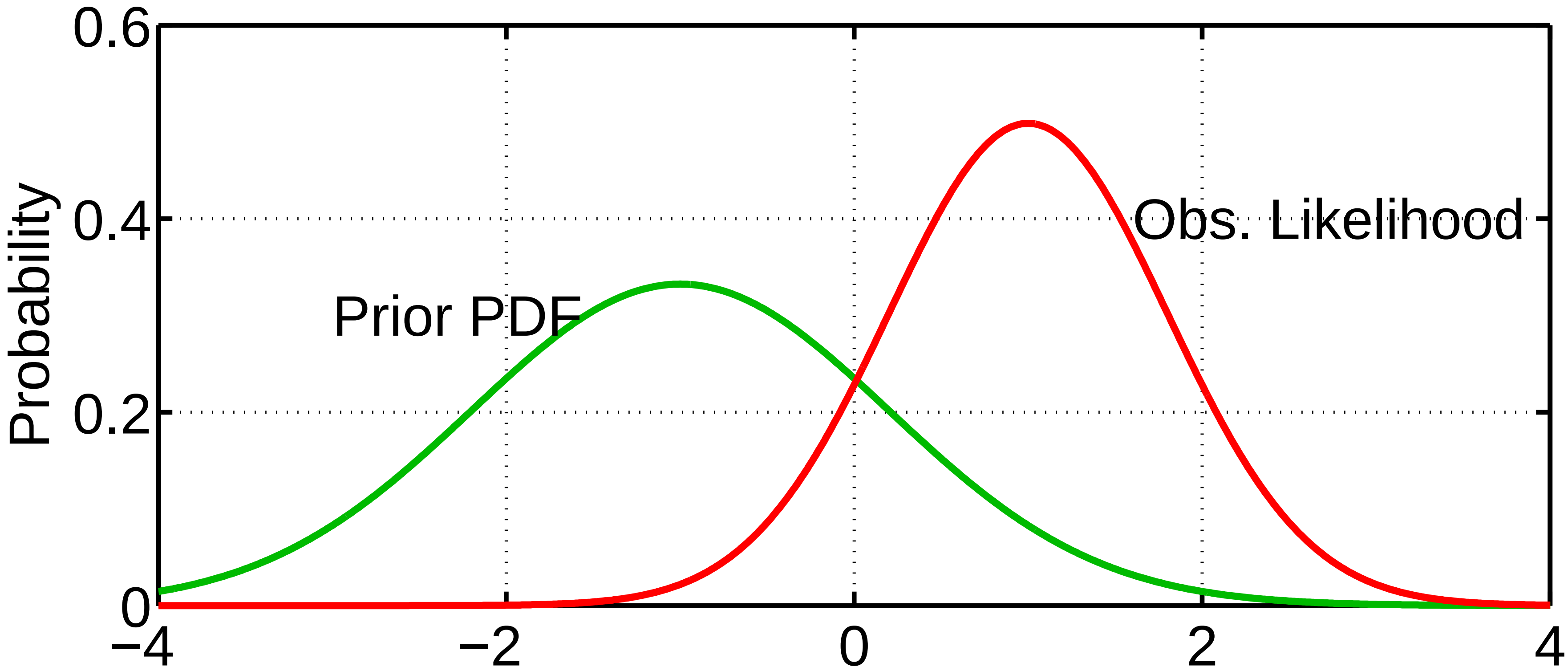
Green = Prior

Red = Observation

Blue = Posterior

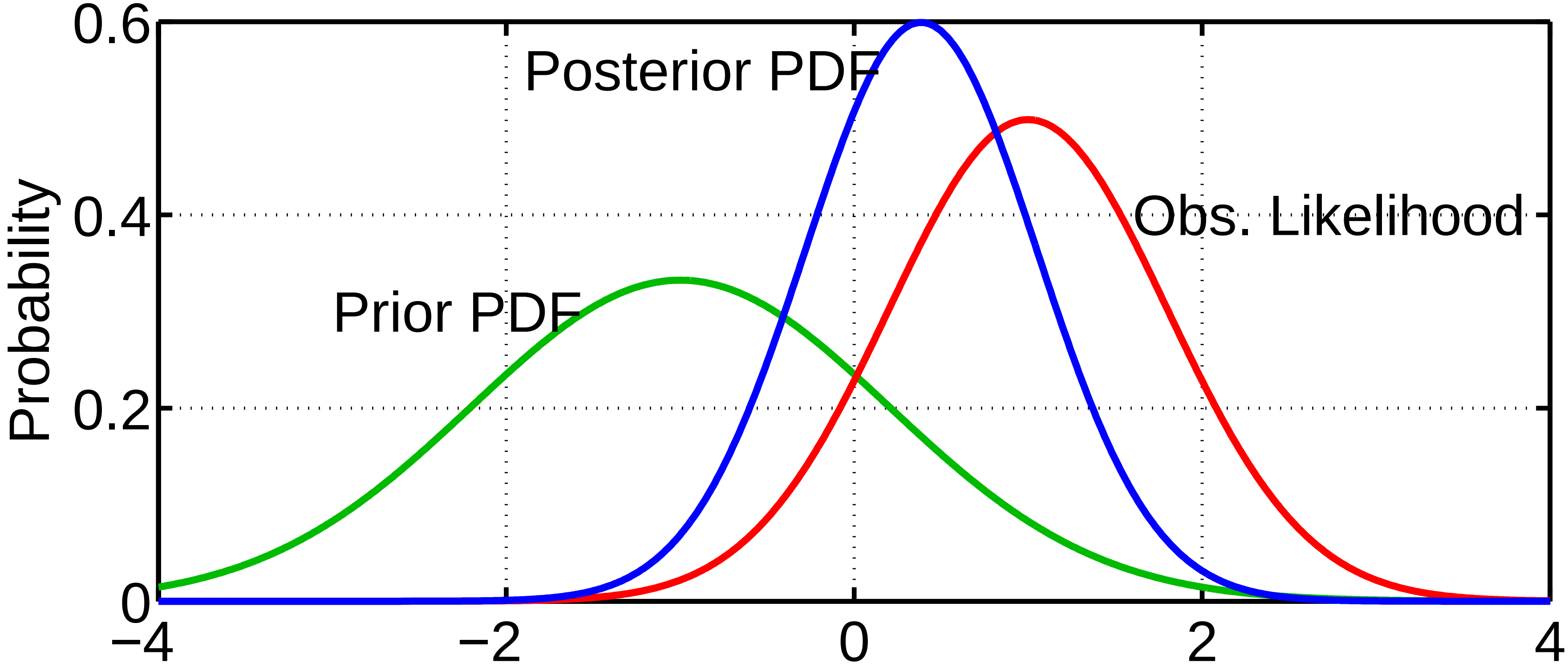
Bayes rule: $p(A|BC) = \dfrac{p(B|AC)p(A|C)}{p(B|C)} = \dfrac{p(B|AC)p(A|C)}{\int p(B|x)p(x|C)dx}$

This product is closed for Gaussian distributions.



Bayes rule: $p(A|BC) = \frac{p(B|AC)p(A|C)}{p(B|C)} = \frac{p(B|AC)p(A|C)}{\int p(B|x)p(x|C)dx}$

This product is closed for Gaussian distributions.



Product of two Gaussians:

Product of d-dimensional normals with means μ_1 and μ_2 and covariance matrices Σ_1 and Σ_2 is normal.

$$\mathbf{N}(\mu_1, \Sigma_1)\boldsymbol{N}(\mu_2, \Sigma_2) = c\mathbf{N}(\mu, \Sigma)$$

Product of two Gaussians:

Product of d-dimensional normals with means μ_1 and μ_2 and covariance matrices Σ_1 and Σ_2 is normal.

$$\mathbf{N}(\mu_1, \Sigma_1)N(\mu_2, \Sigma_2) = c\mathbf{N}(\mu, \Sigma)$$

Covariance: $$\Sigma = (\Sigma_1^{-1} + \Sigma_2^{-1})^{-1}$$

Mean: $$\mu = (\Sigma_1^{-1} + \Sigma_2^{-1})^{-1}(\Sigma_1^{-1}\mu_1 + \Sigma_2^{-1}\mu_2)$$

Product of two Gaussians:

Product of d-dimensional normals with means μ_1 and μ_2 and covariance matrices Σ_1 and Σ_2 is normal.

$$\mathbf{N}(\mu_1, \Sigma_1)\boldsymbol{N}(\mu_2, \Sigma_2) = c\mathbf{N}(\mu, \Sigma)$$

Covariance: $\Sigma = (\Sigma_1^{-1} + \Sigma_2^{-1})^{-1}$

Mean: $\mu = (\Sigma_1^{-1} + \Sigma_2^{-1})^{-1}(\Sigma_1^{-1}\mu_1 + \Sigma_2^{-1}\mu_2)$

Weight: $c = \frac{1}{(2\Pi)^{d/2}\left|\Sigma_1 + \Sigma_2\right|^{1/2}}\exp\left\{-\frac{1}{2}[(\mu_2 - \mu_1)^T(\Sigma_1 + \Sigma_2)^{-1}(\mu_2 - \mu_1)]\right\}$

We'll ignore the weight unless noted since we immediately normalize products to be PDFs.

Product of two Gaussians:

Product of d-dimensional normals with means μ_1 and μ_2 and covariance matrices Σ_1 and Σ_2 is normal.

$$\mathbf{N}(\mu_1, \Sigma_1)\mathbf{N}(\mu_2, \Sigma_2) = c\mathbf{N}(\mu, \Sigma)$$

Covariance: $\Sigma = (\Sigma_1^{-1} + \Sigma_2^{-1})^{-1}$

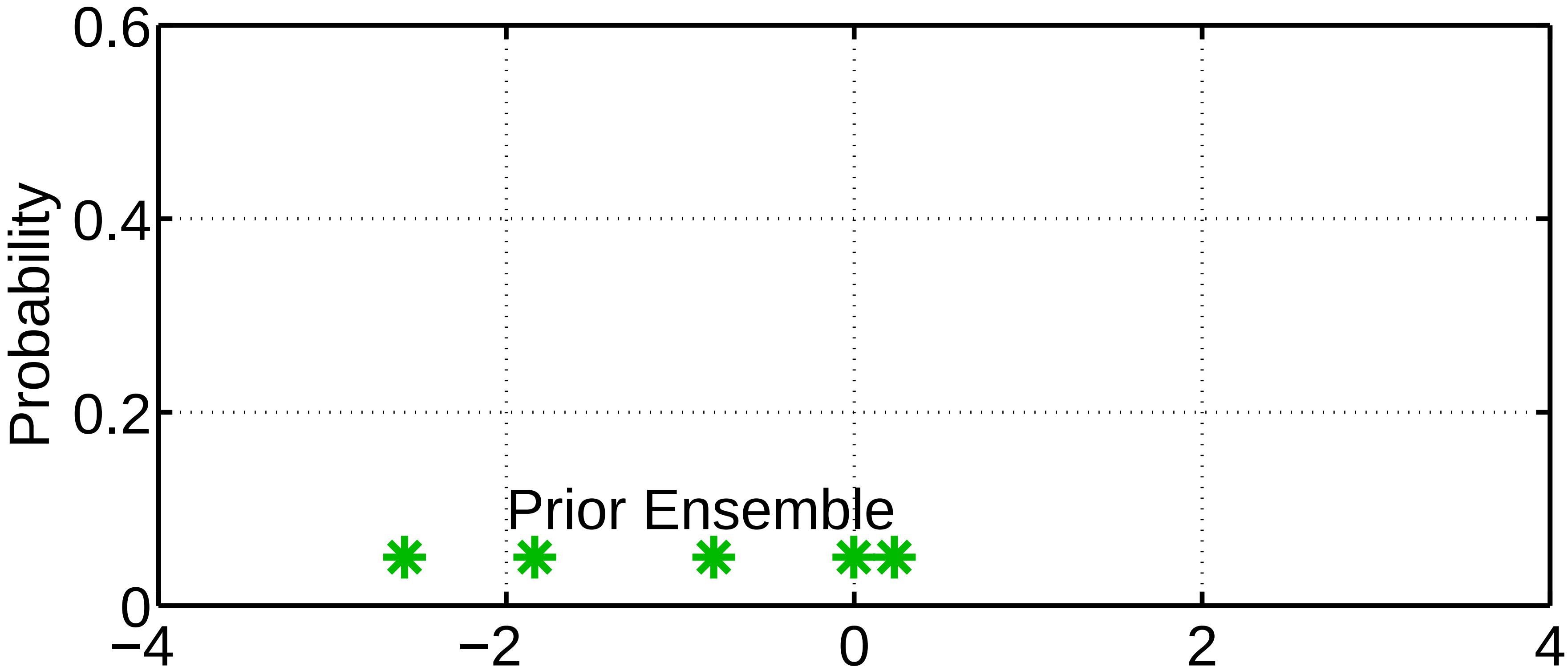
Mean: $\mu = (\Sigma_1^{-1} + \Sigma_2^{-1})^{-1}(\Sigma_1^{-1}\mu_1 + \Sigma_2^{-1}\mu_2)$

Weight: $c = \dfrac{1}{(2\Pi)^{d/2}|\Sigma_1 + \Sigma_2|^{1/2}} \exp\left\{-\frac{1}{2}[(\mu_2 - \mu_1)^T(\Sigma_1 + \Sigma_2)^{-1}(\mu_2 - \mu_1)]\right\}$

Easy to derive for 1-D Gaussians; just do products of exponentials.

Bayes rule: $p(A|BC) = \frac{p(B|AC)p(A|C)}{p(B|C)} = \frac{p(B|AC)p(A|C)}{\int p(B|x)p(x|C)dx}$

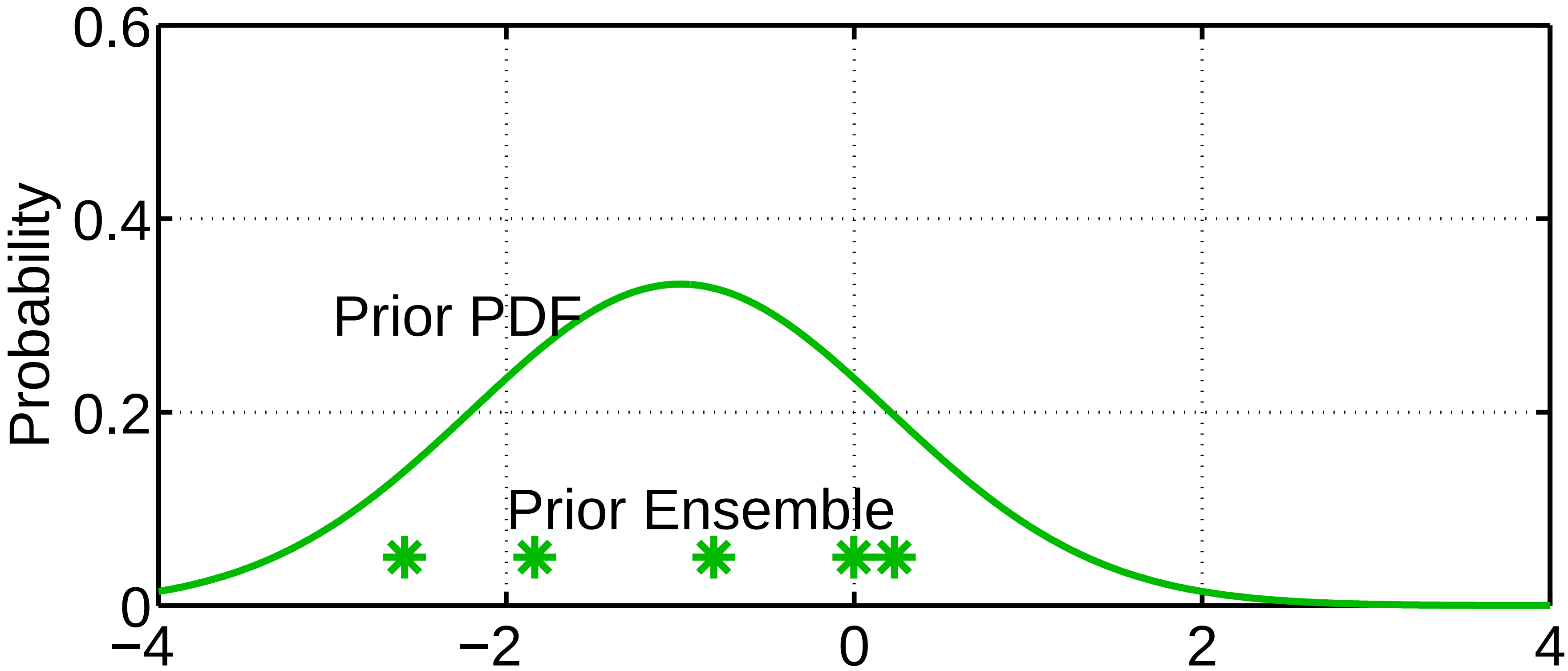
Ensemble filters: Prior is available as finite sample.



Don’t know much about properties of this sample.
May naively assume it is random draw from ‘truth’.

Bayes rule: $p(A|BC) = \frac{p(B|AC)p(A|C)}{p(B|C)} = \frac{p(B|AC)p(A|C)}{\int p(B|x)p(x|C)dx}$

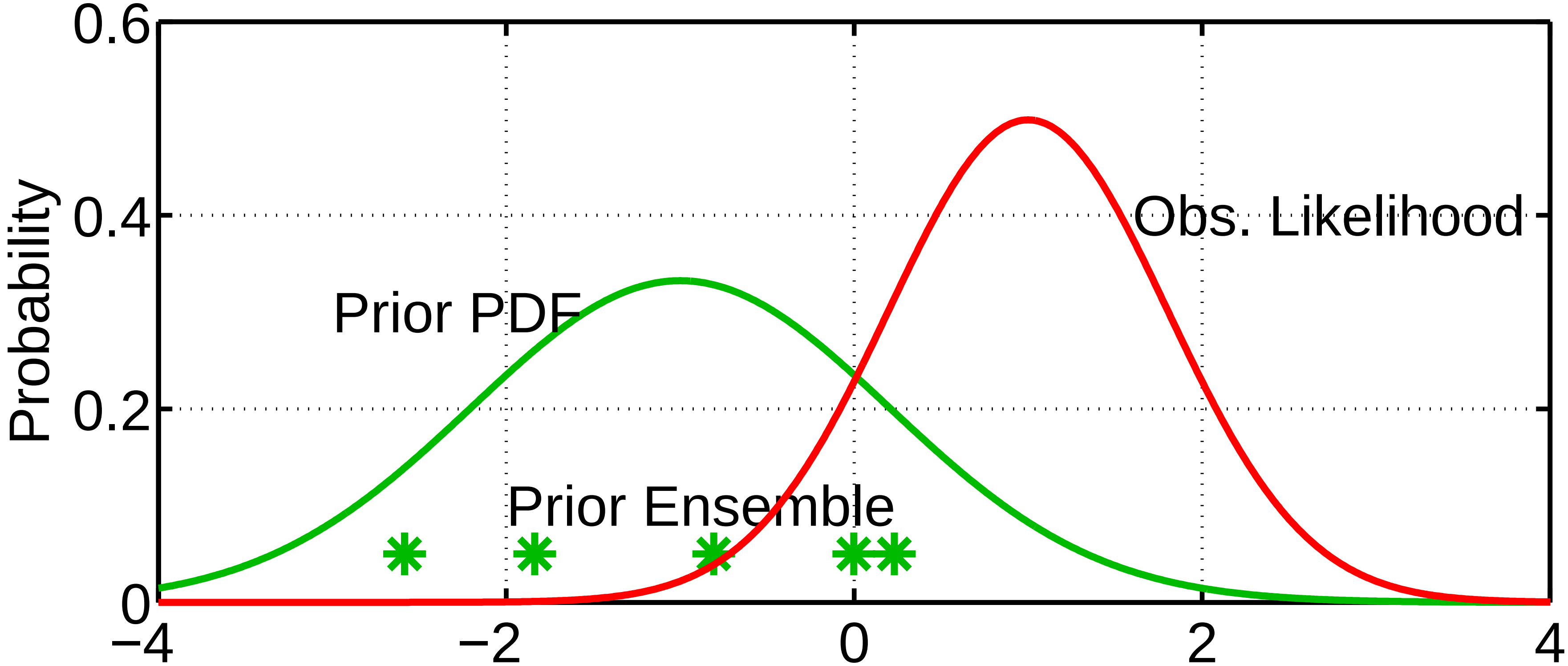
How can we take product of sample with continuous likelihood?



Fit a continuous (Gaussian for now) distribution to sample.

Bayes rule: $p(A|BC) = \dfrac{p(B|AC)p(A|C)}{p(B|C)} = \dfrac{p(B|AC)p(A|C)}{\int p(B|x)p(x|C)dx}$

Observation likelihood usually continuous (nearly always Gaussian).

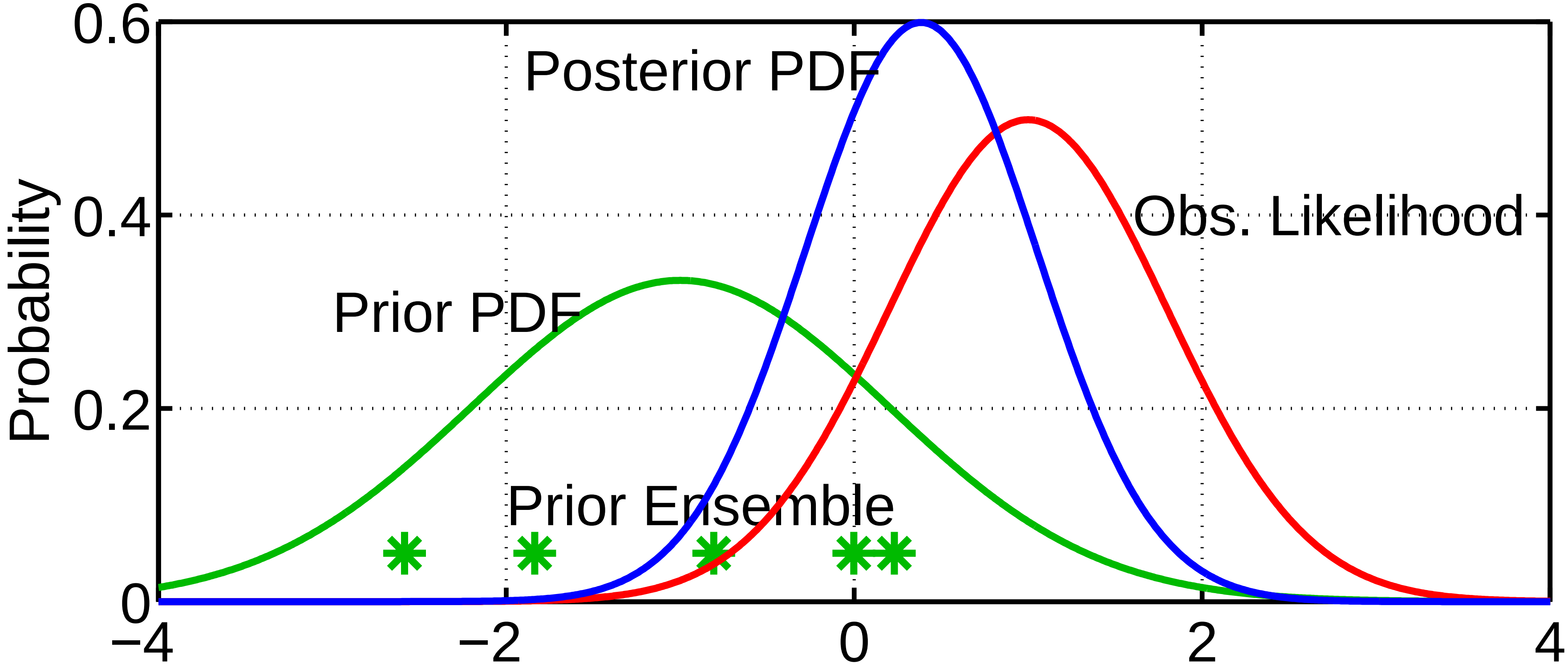


If Obs. Likelihood isn’t Gaussian, can generalize methods below.

For instance, can fit set of Gaussian kernels to obs. likelihood.

Bayes rule: $p(A|BC) = \dfrac{p(B|AC)p(A|C)}{p(B|C)} = \dfrac{p(B|AC)p(A|C)}{\int p(B|x)p(x|C)dx}$

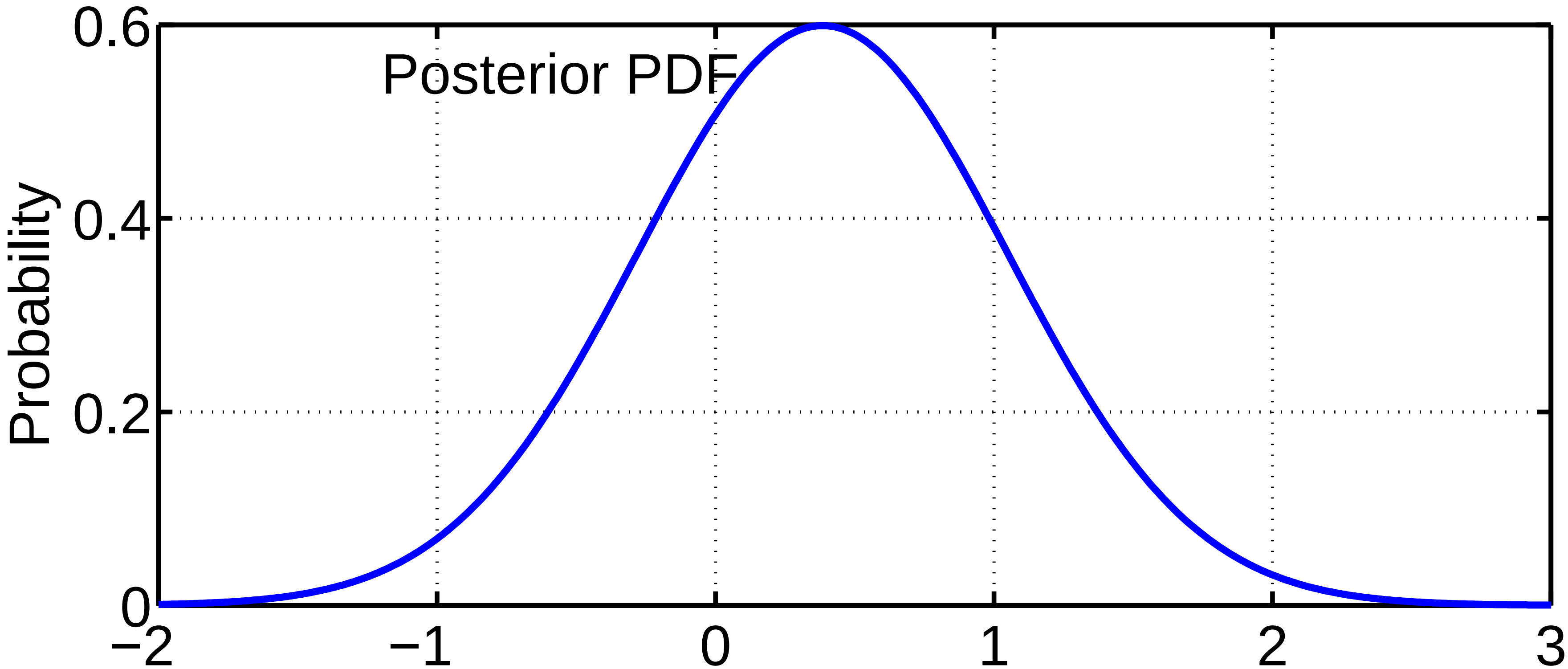
Product of prior Gaussian fit and Obs. likelihood is Gaussian.



Computing continuous posterior is simple.
BUT, need to have a SAMPLE of this PDF.

Sampling Posterior PDF:

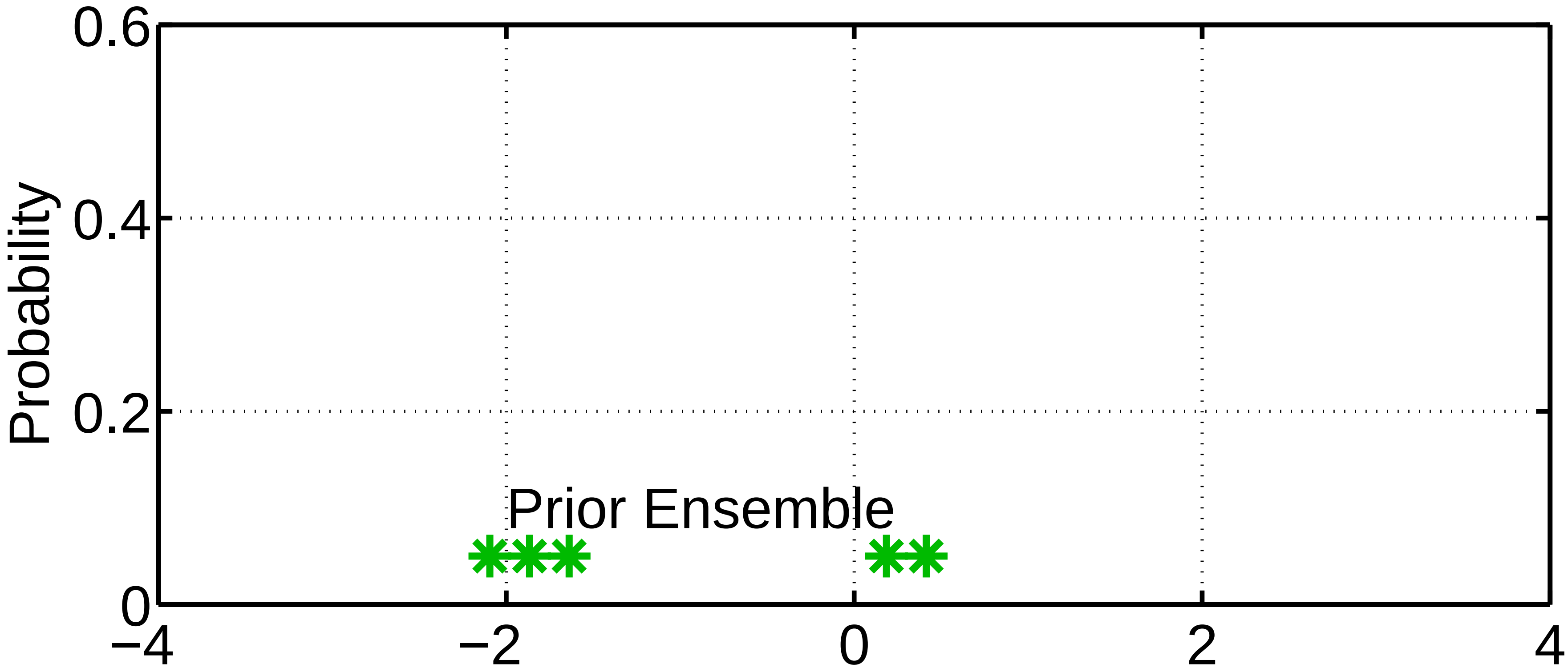
There are many ways to do this.



Exact properties of different methods may be unclear.
Trial and error still best way to see how they perform.
Will interact with properties of prediction models, etc.

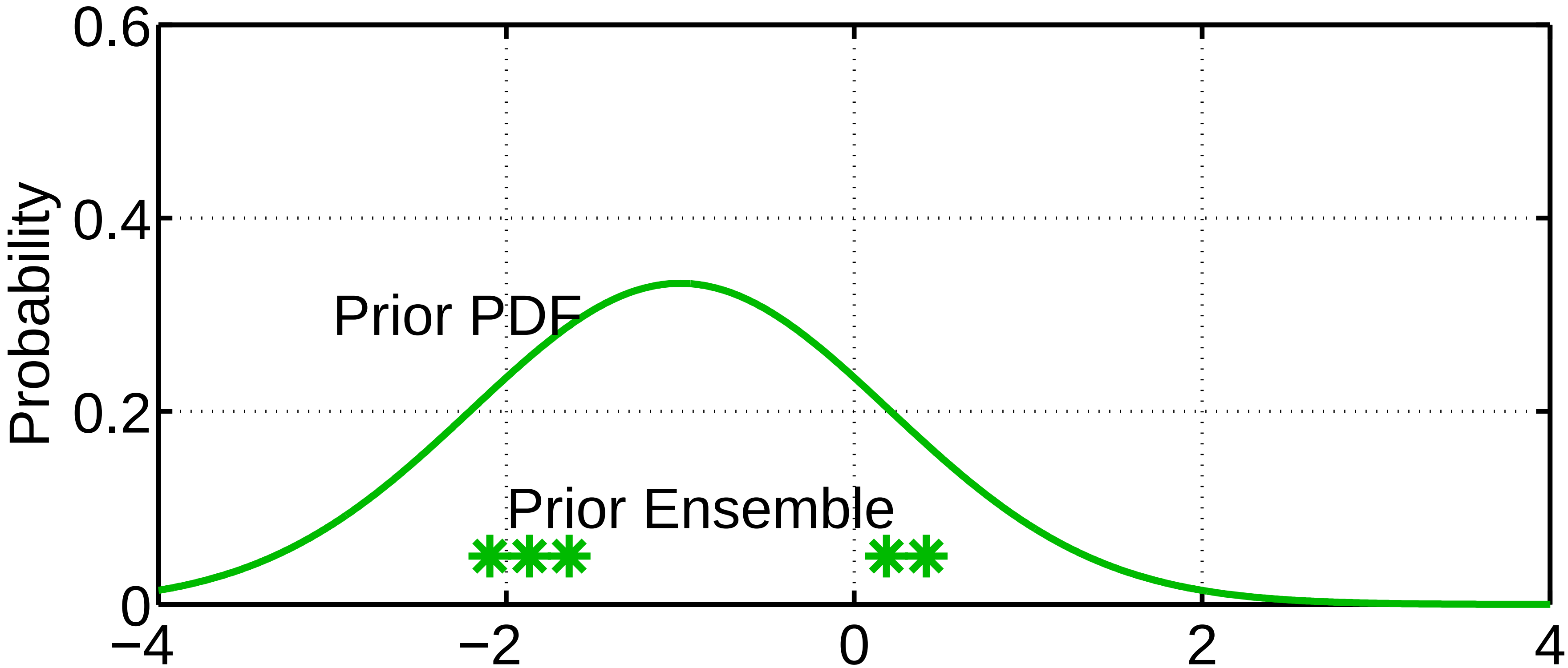
Ensemble Filter Algorithms:

3. Ensemble Adjustment (Kalman) Filter.



Ensemble Filter Algorithms:

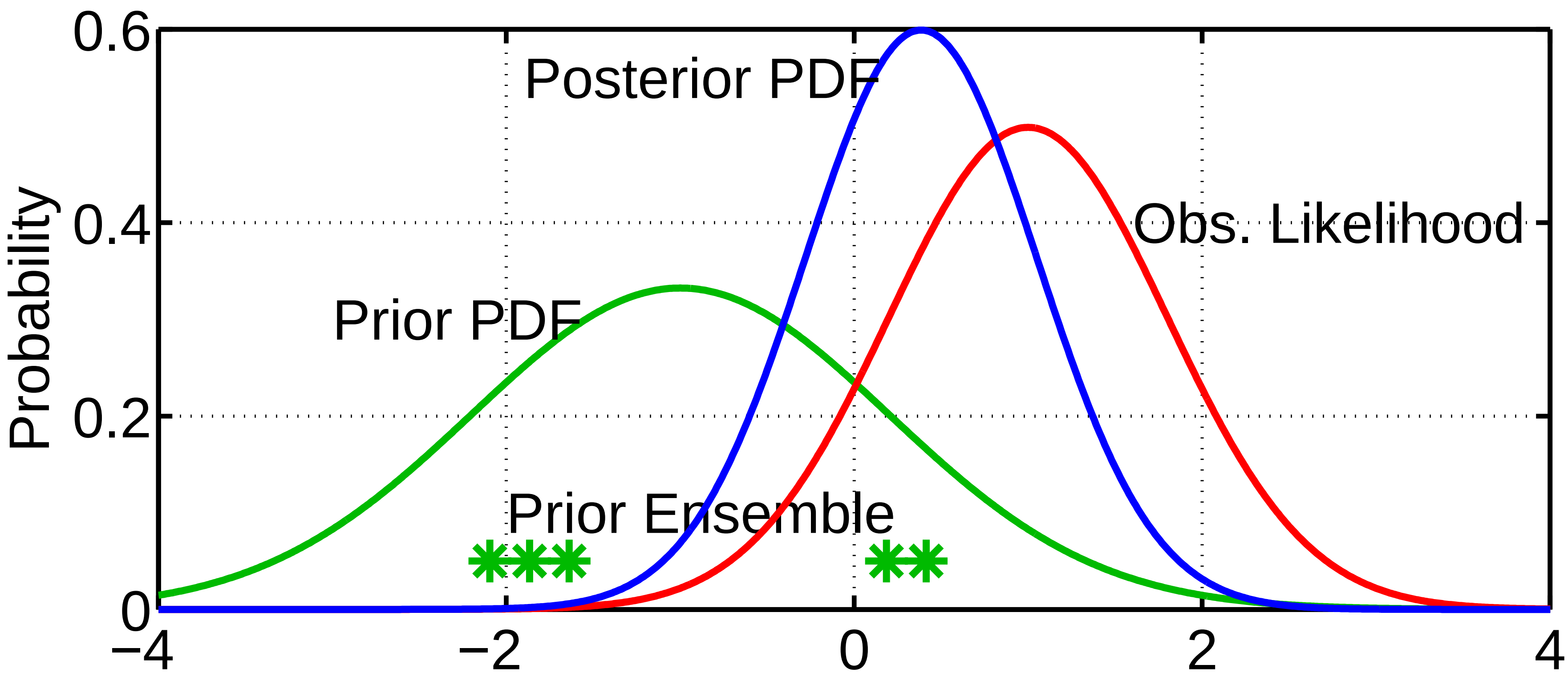
3. Ensemble Adjustment (Kalman) Filter.



Again, fit a Gaussian to sample.

Ensemble Filter Algorithms:

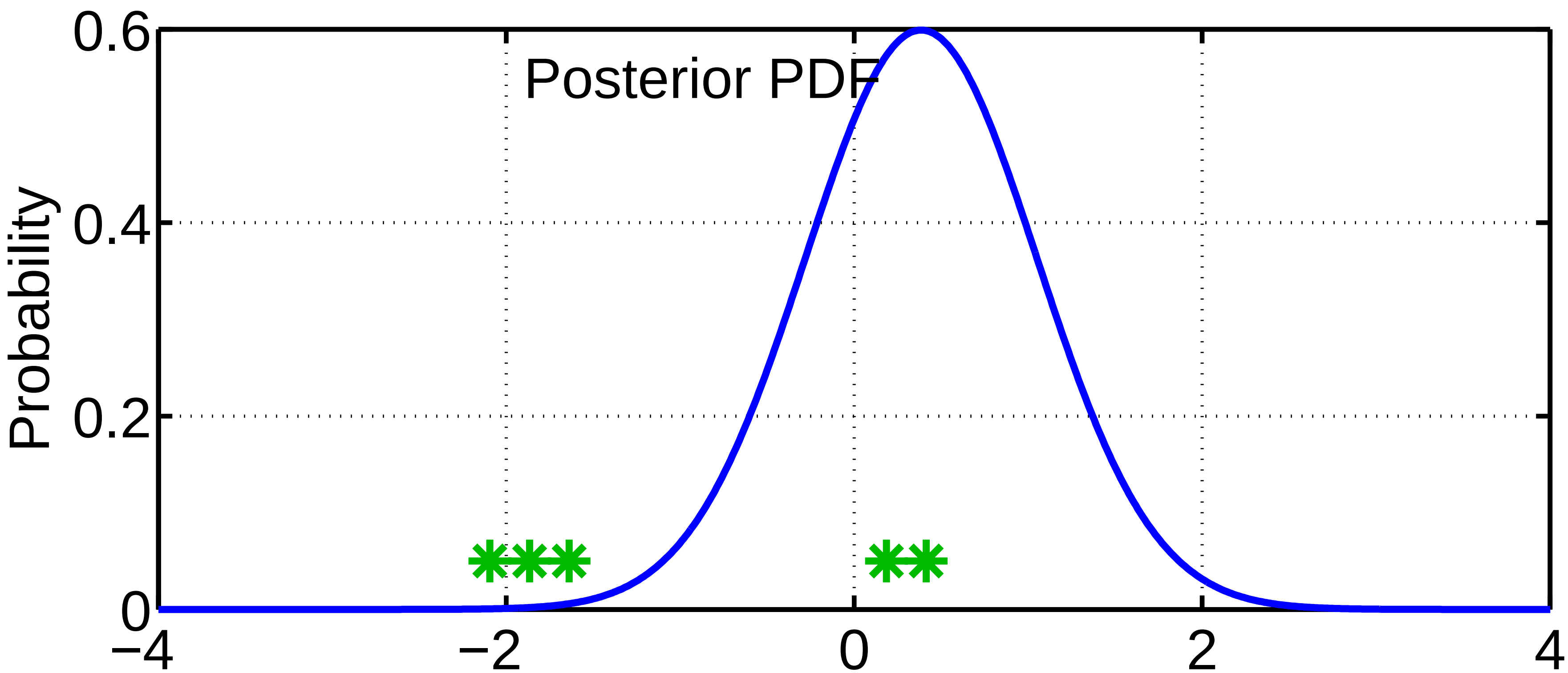
3. Ensemble Adjustment (Kalman) Filter.



Compute posterior PDF (same as previous algorithms).

Ensemble Filter Algorithms:

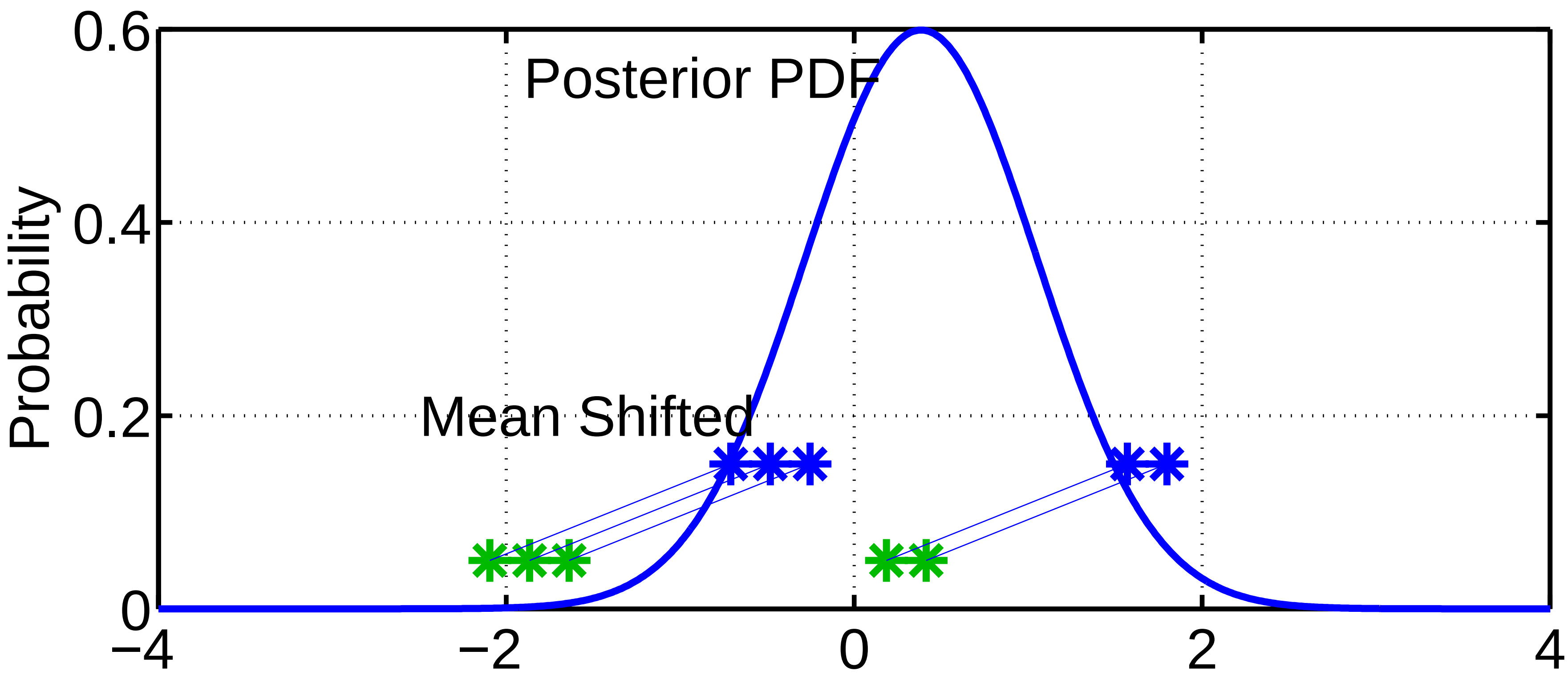
3. Ensemble Adjustment (Kalman) Filter.



Use deterministic algorithm to 'adjust' ensemble.

Ensemble Filter Algorithms:

3. Ensemble Adjustment (Kalman) Filter.

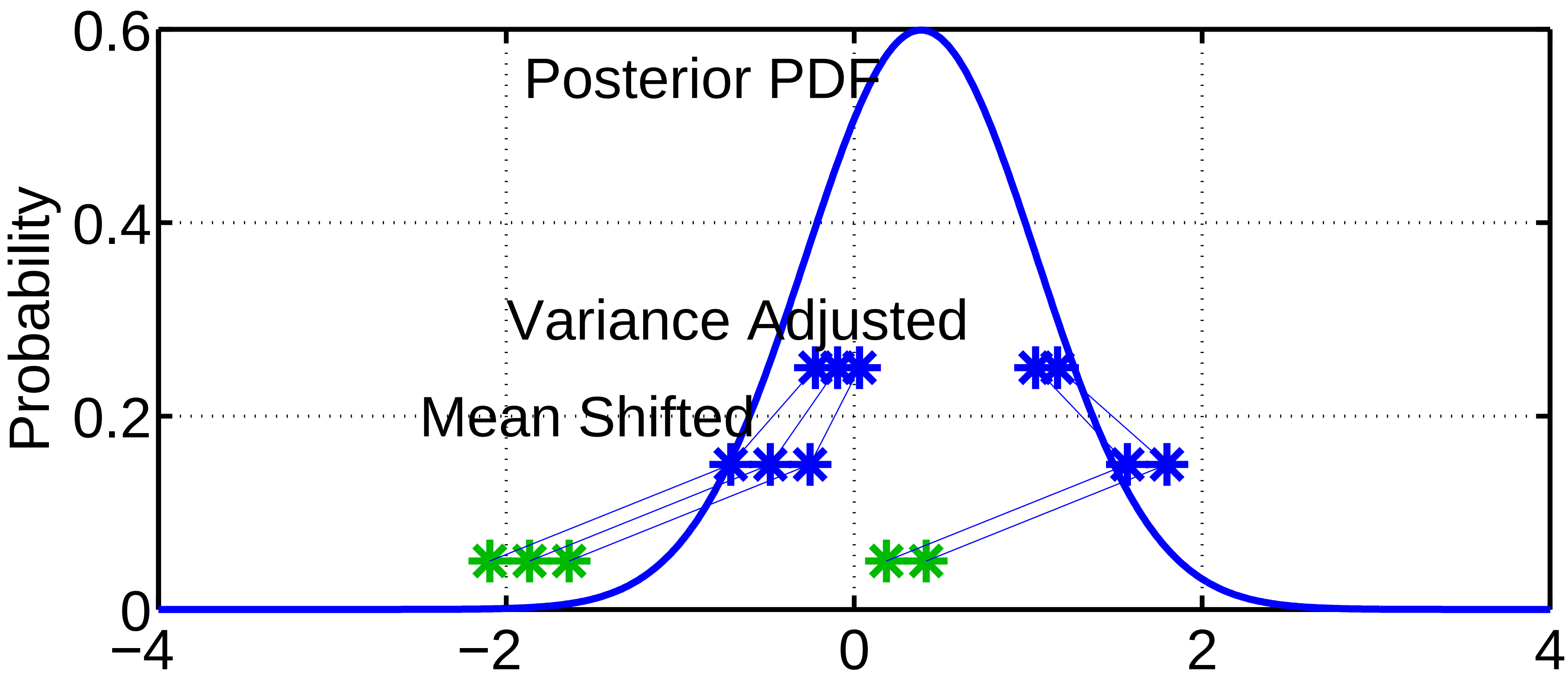


Use deterministic algorithm to ‘adjust’ ensemble.

First, ‘shift’ ensemble to have exact mean of posterior.

Ensemble Filter Algorithms:

3. Ensemble Adjustment (Kalman) Filter.



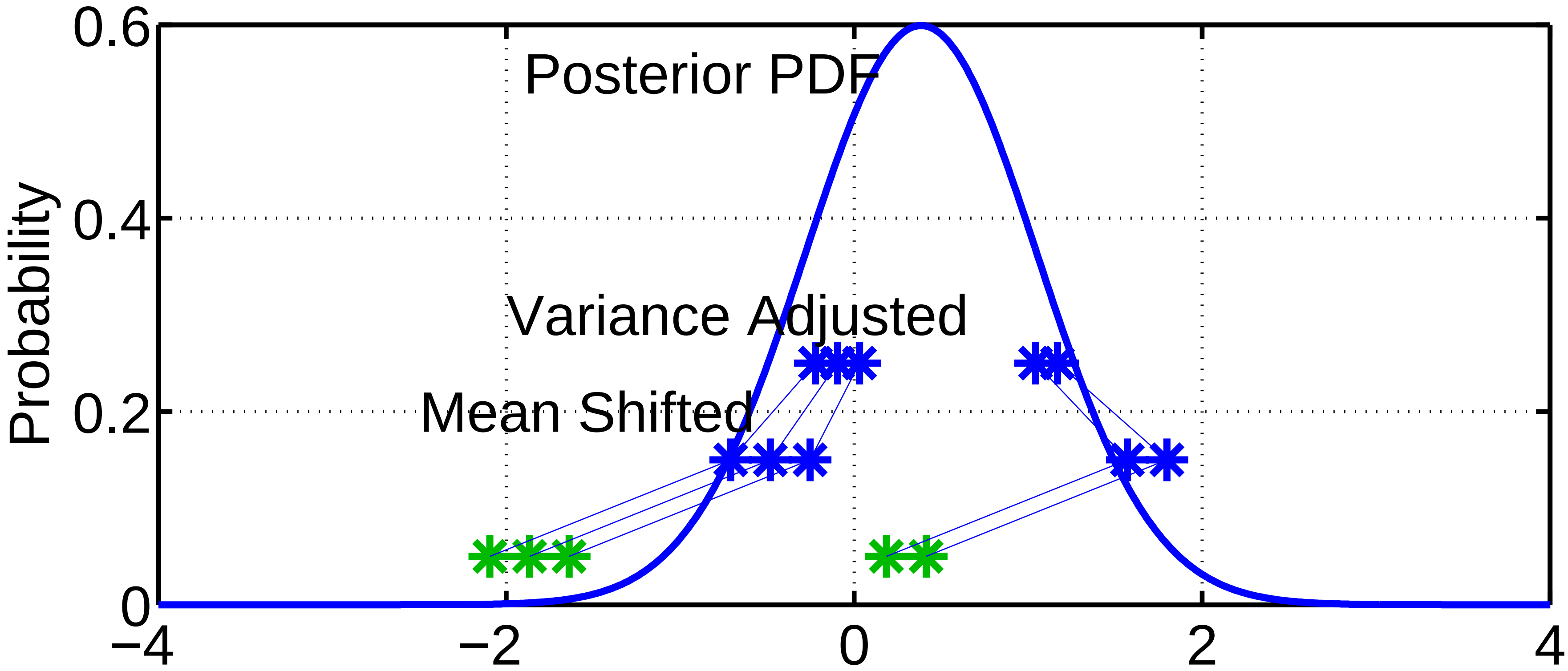
Use deterministic algorithm to ‘adjust’ ensemble.

First, ‘shift’ ensemble to have exact mean of posterior.

Second, use linear contraction to have exact variance of posterior.

Ensemble Filter Algorithms:

3. Ensemble Adjustment (Kalman) Filter.

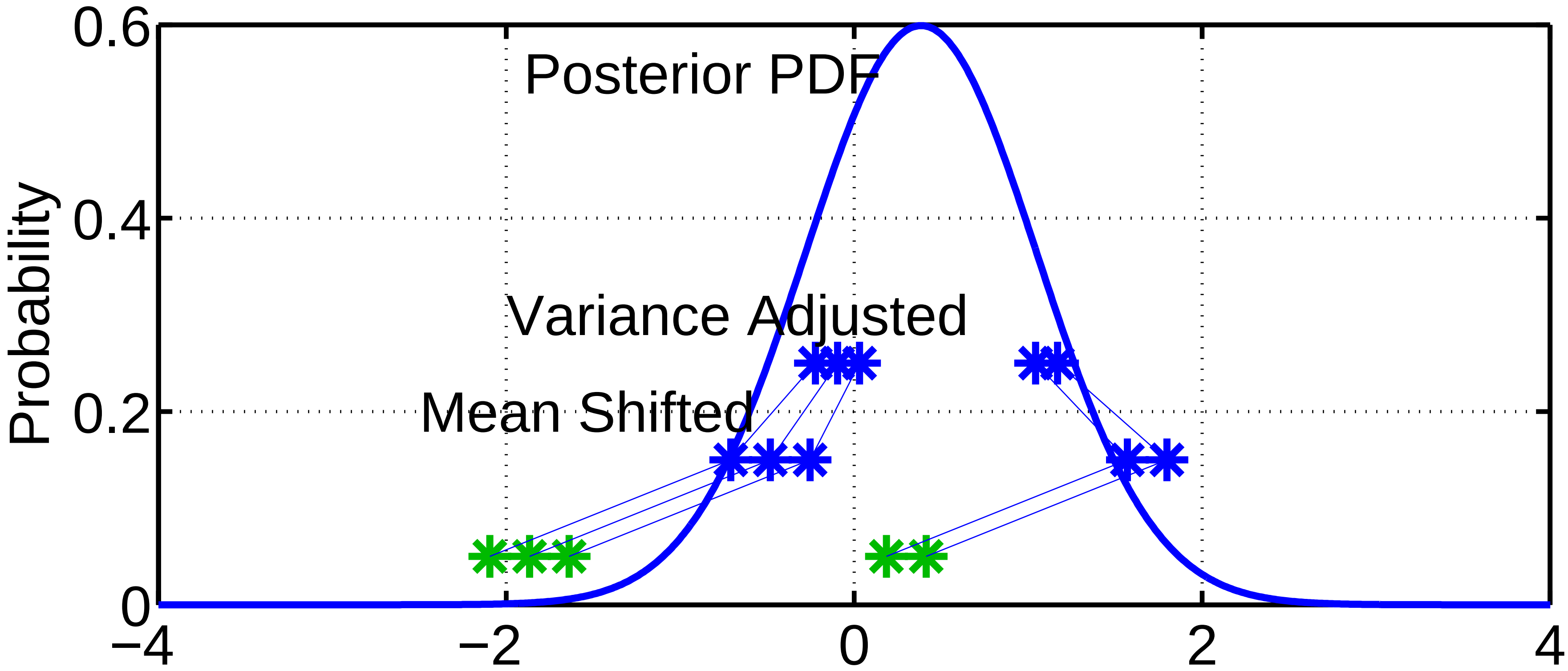


$$x_i^u = (x_i^p - \overline{x^p}) \cdot (\sigma^u / \sigma^p) + \overline{x^u} \quad \text{i = 1,..., ensemble size.}$$

p is prior, u is update (posterior), overbar is ensemble mean,
σ is standard deviation.

Ensemble Filter Algorithms:

3. Ensemble Adjustment (Kalman) Filter.



Bimodality maintained, but not appropriately positioned or weighted.
No problem with random outliers.

Phase 2: Single observed variable, single unobserved variable

So far, have known observation likelihood for single variable.

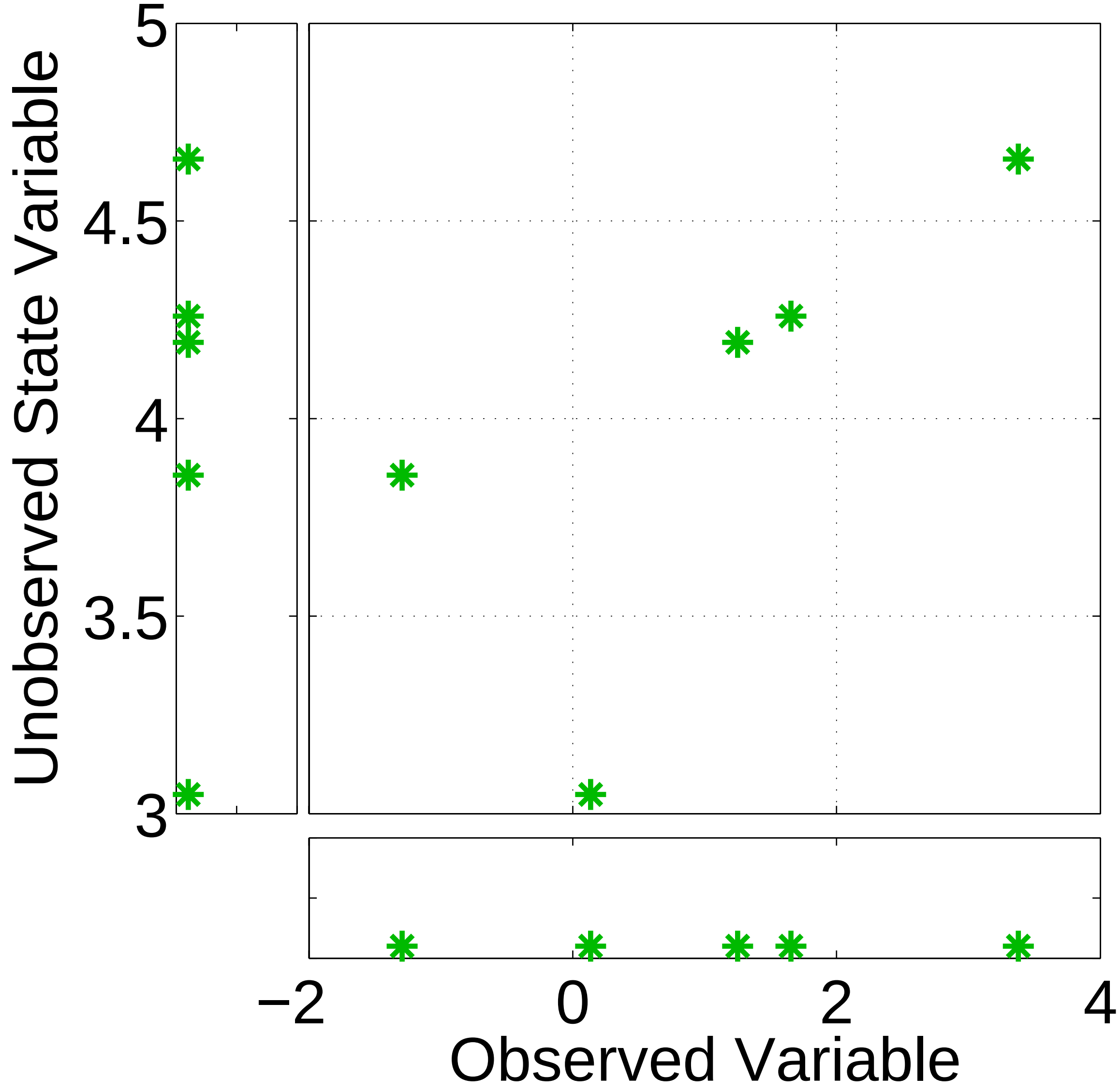
Now, suppose prior has an additional variable.

Will examine how ensemble methods update additional variable.

Basic method generalizes to any number of additional variables.

Methods related to Kalman filter in some sense, but not done here.

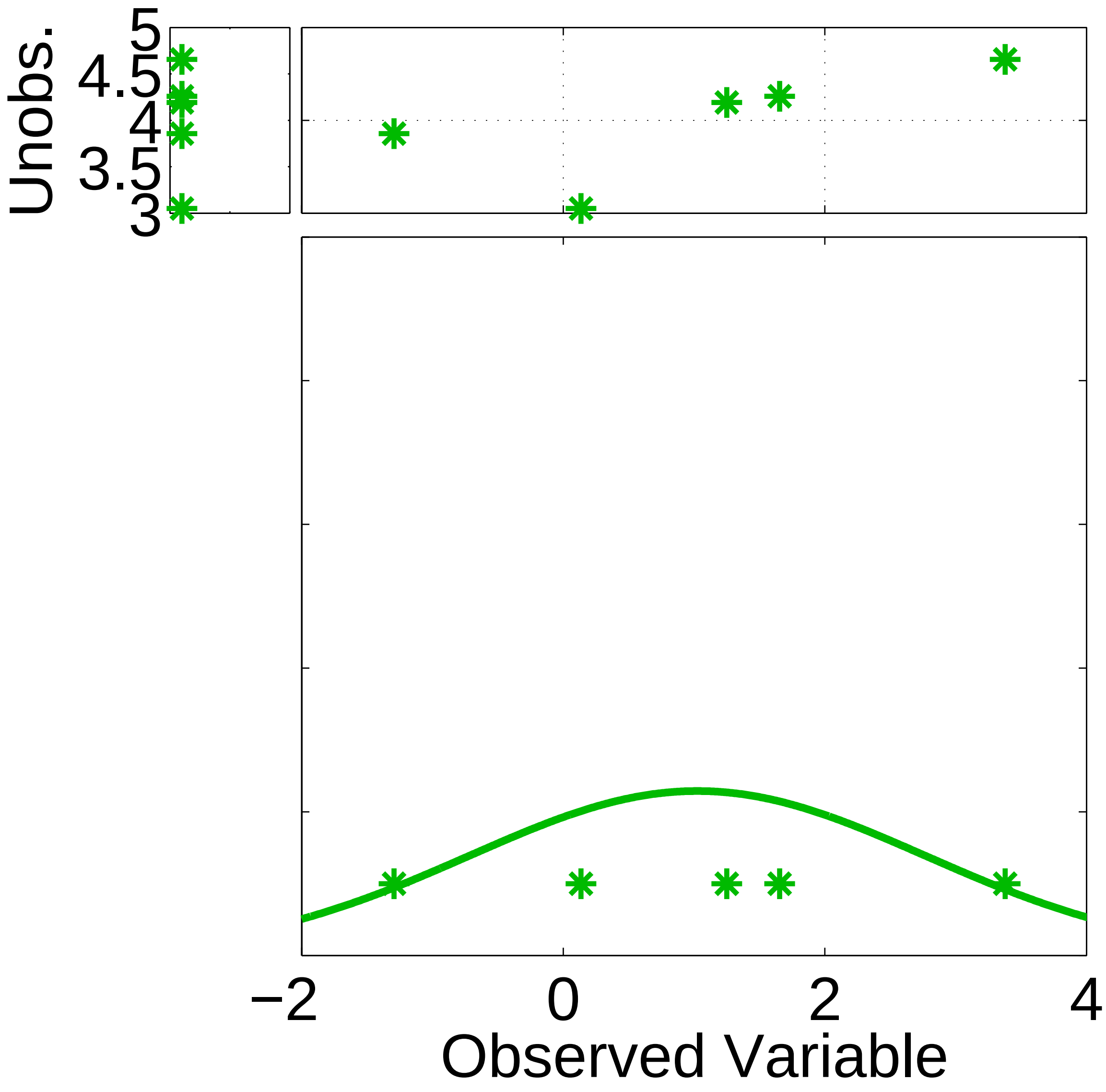
Ensemble filters: Updating additional prior state variables



Assume that all we know is prior joint distribution.

One variable is observed.
(Boulder temperature)
What should happen to unobserved variable?
(NCAR temp.)

Ensemble filters: Updating additional prior state variables

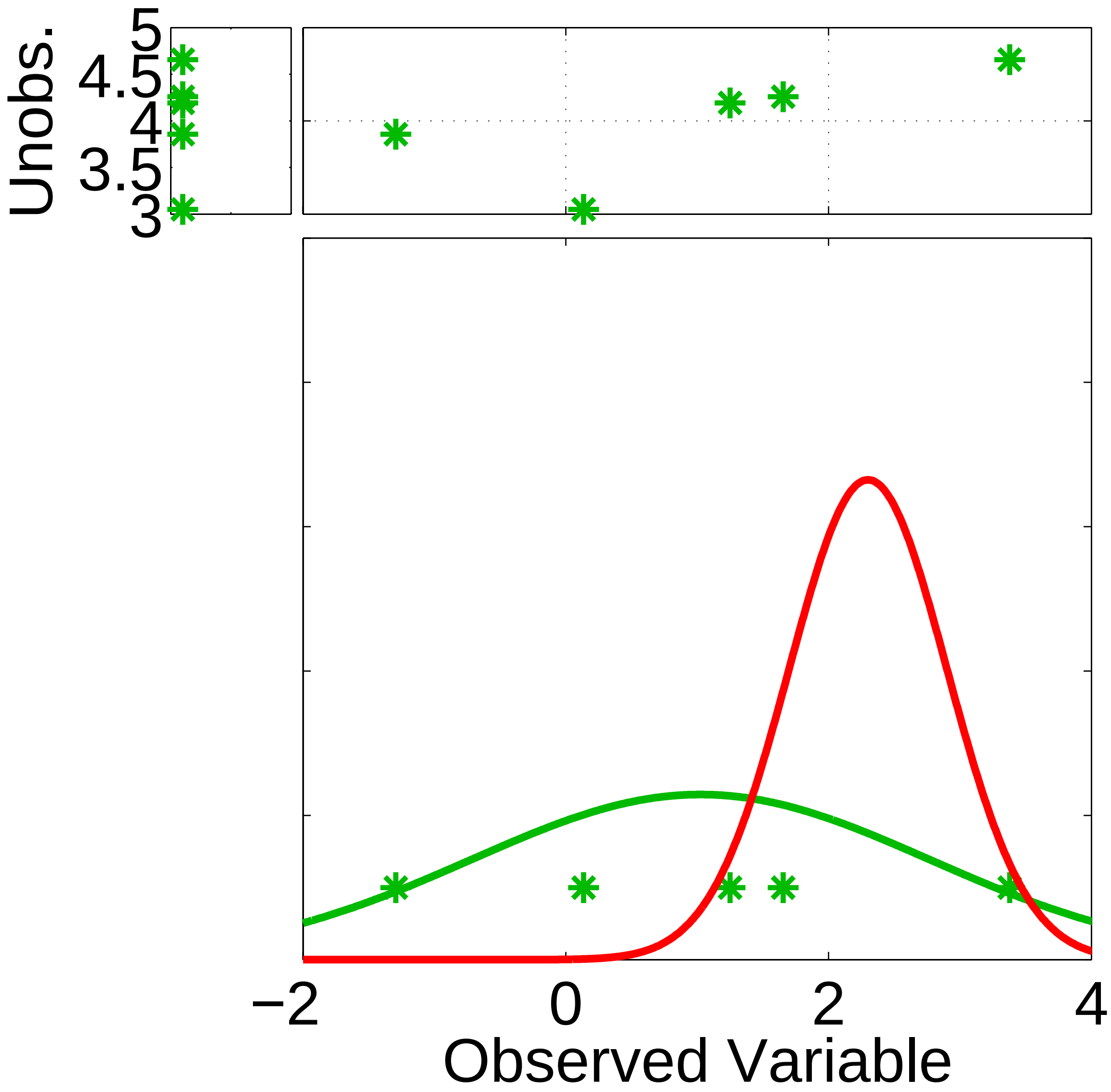


Assume that all we know is prior joint distribution.

One variable is observed.

Update observed variable with one of previous methods.

Ensemble filters: Updating additional prior state variables

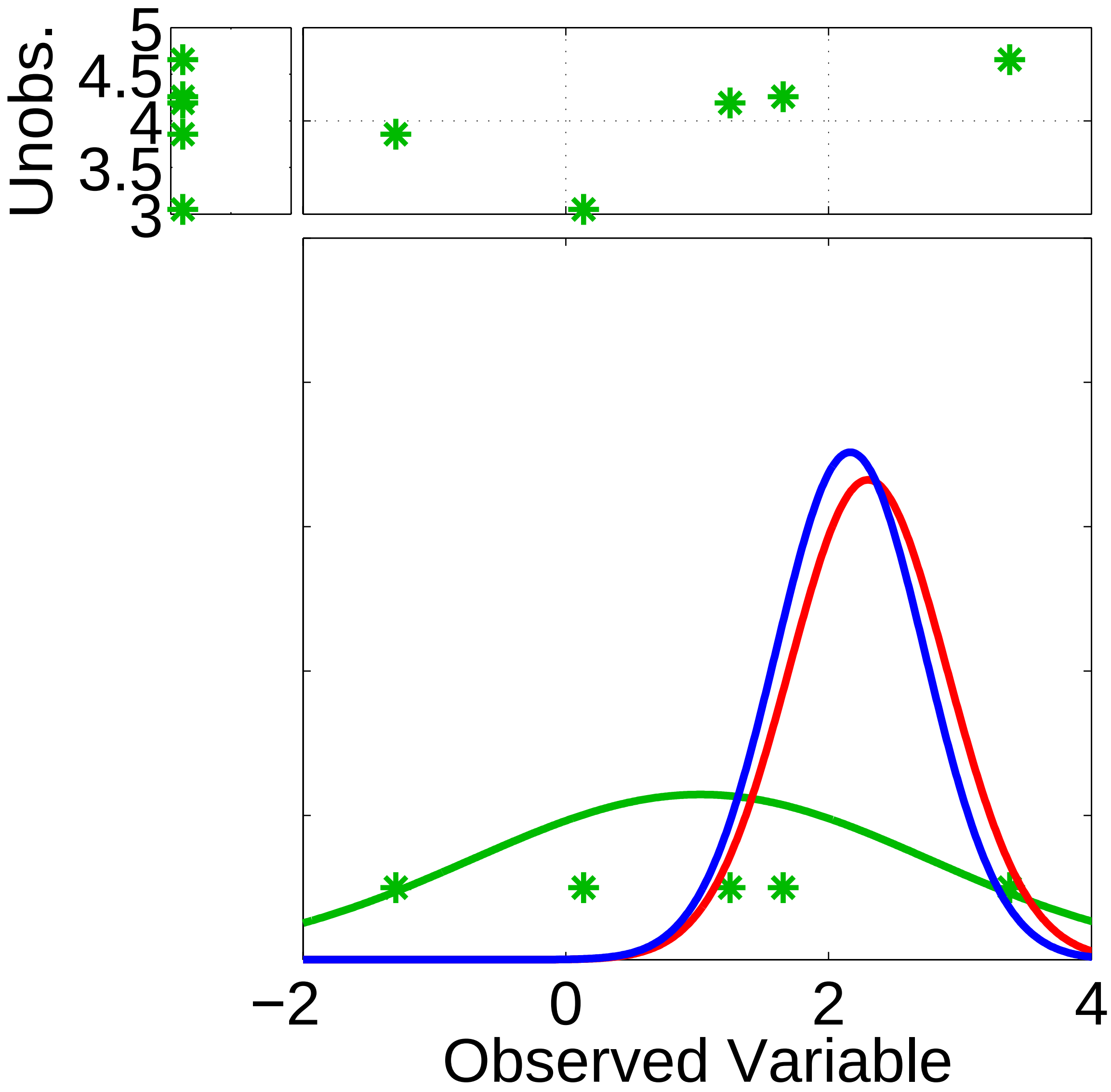


Assume that all we know is prior joint distribution.

One variable is observed.

Update observed variable with one of previous methods.

Ensemble filters: Updating additional prior state variables

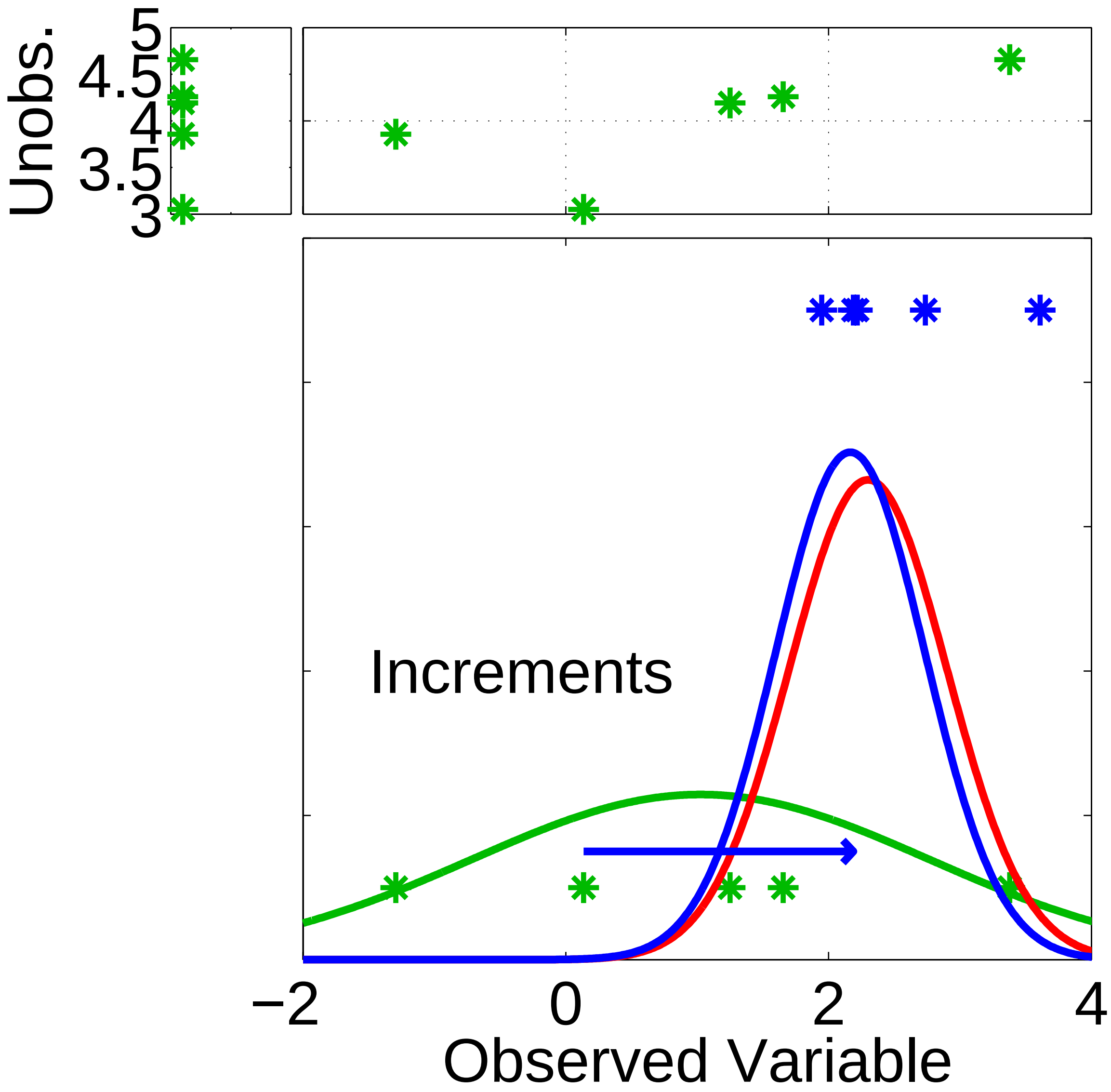


Assume that all we know is prior joint distribution.

One variable is observed.

Update observed variable with one of previous methods.

Ensemble filters: Updating additional prior state variables

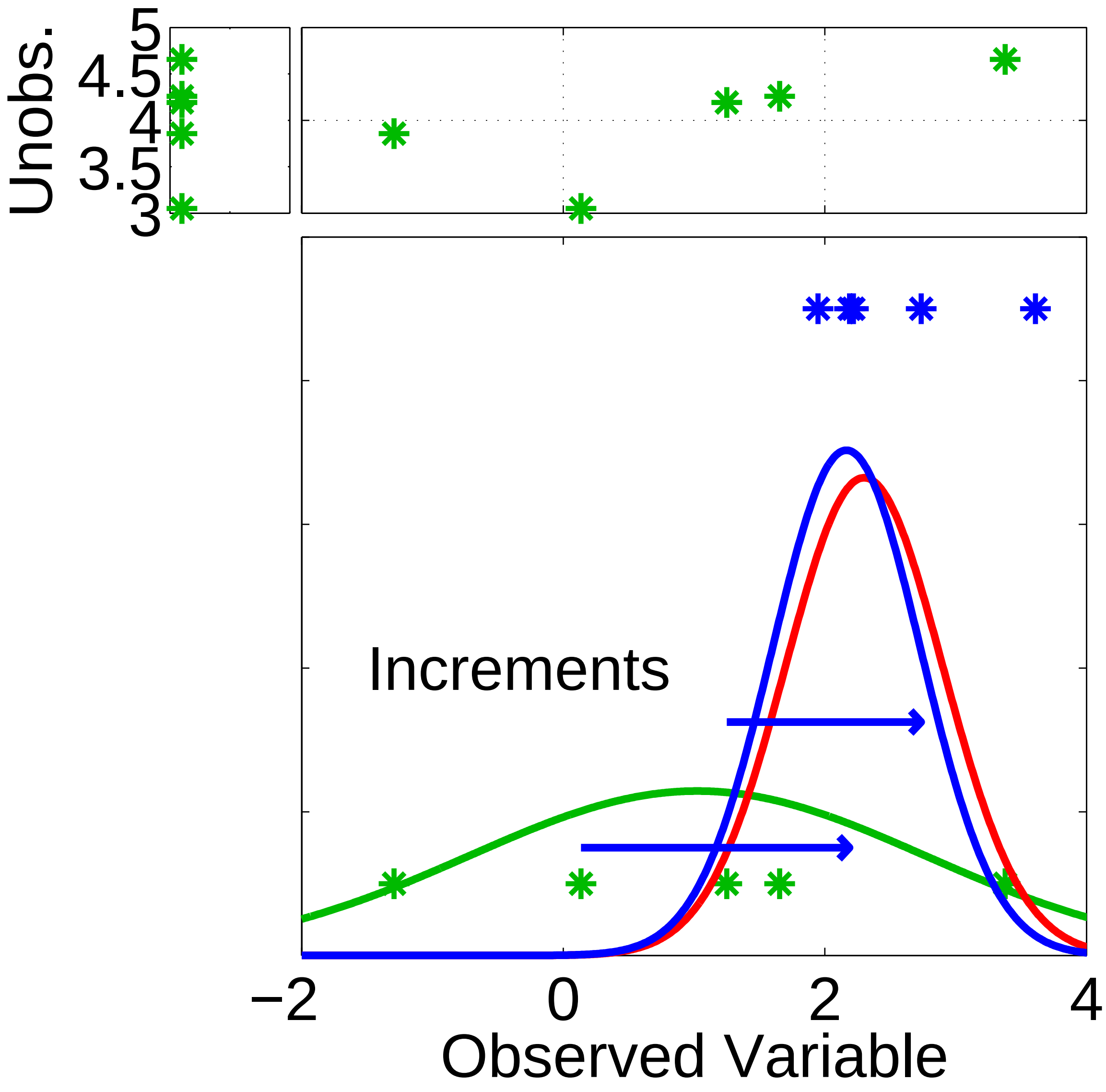


Assume that all we know is prior joint distribution.

One variable is observed.

Compute increments for prior ensemble members of observed variable.

Ensemble filters: Updating additional prior state variables

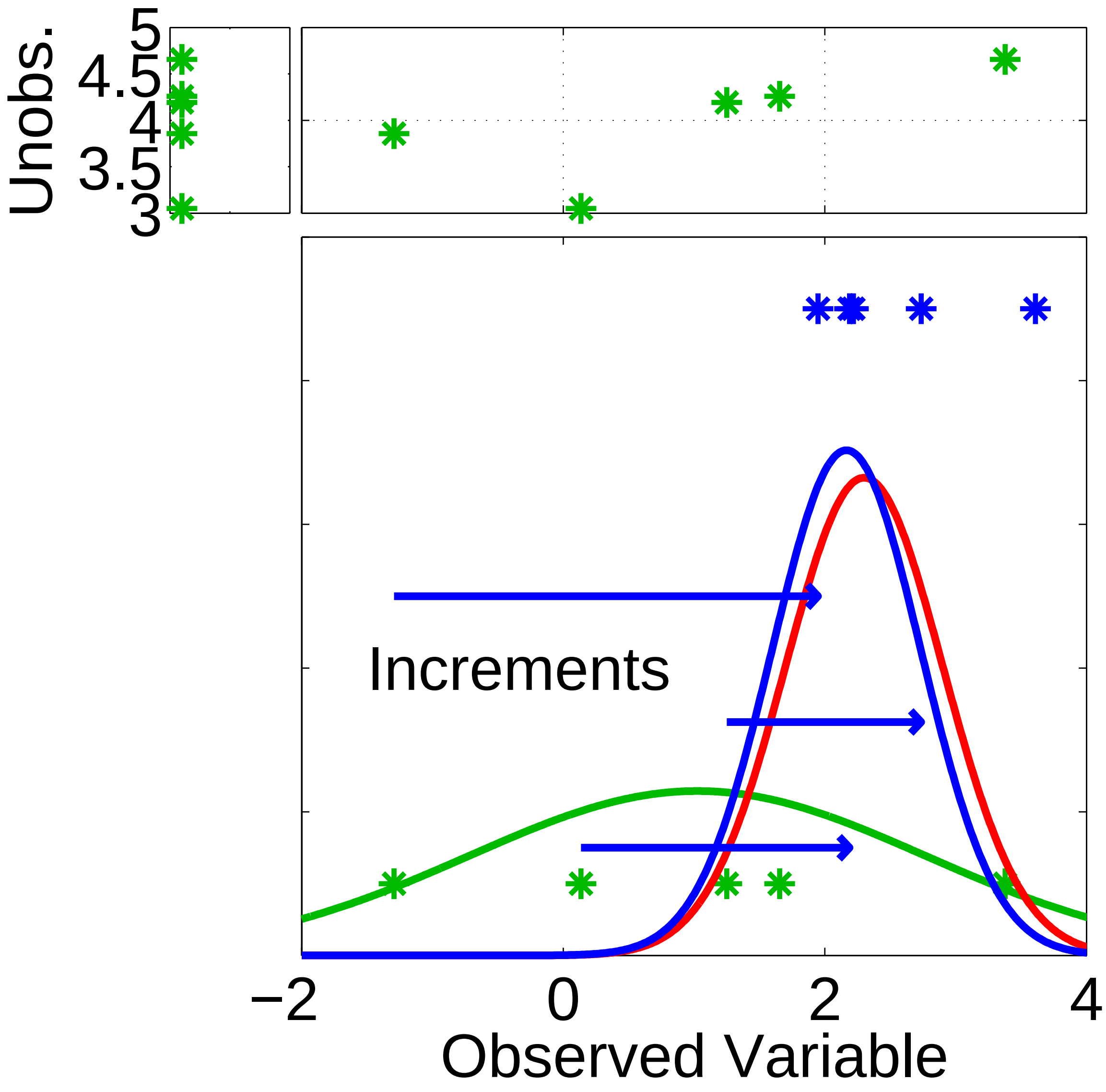


Assume that all we know is prior joint distribution.

One variable is observed.

Compute increments for prior ensemble members of observed variable.

Ensemble filters: Updating additional prior state variables

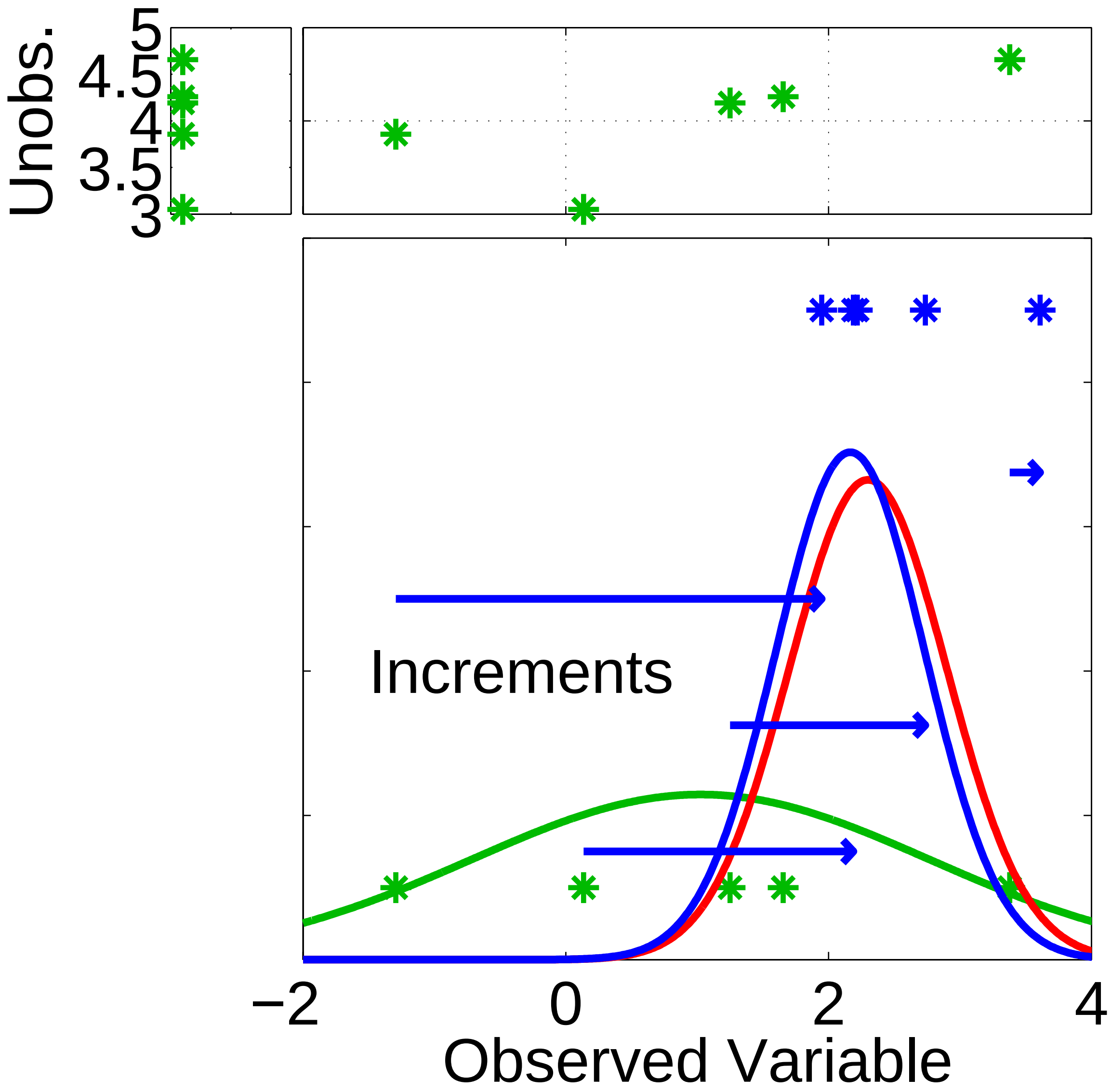


Assume that all we know is prior joint distribution.

One variable is observed.

Compute increments for prior ensemble members of observed variable.

Ensemble filters: Updating additional prior state variables

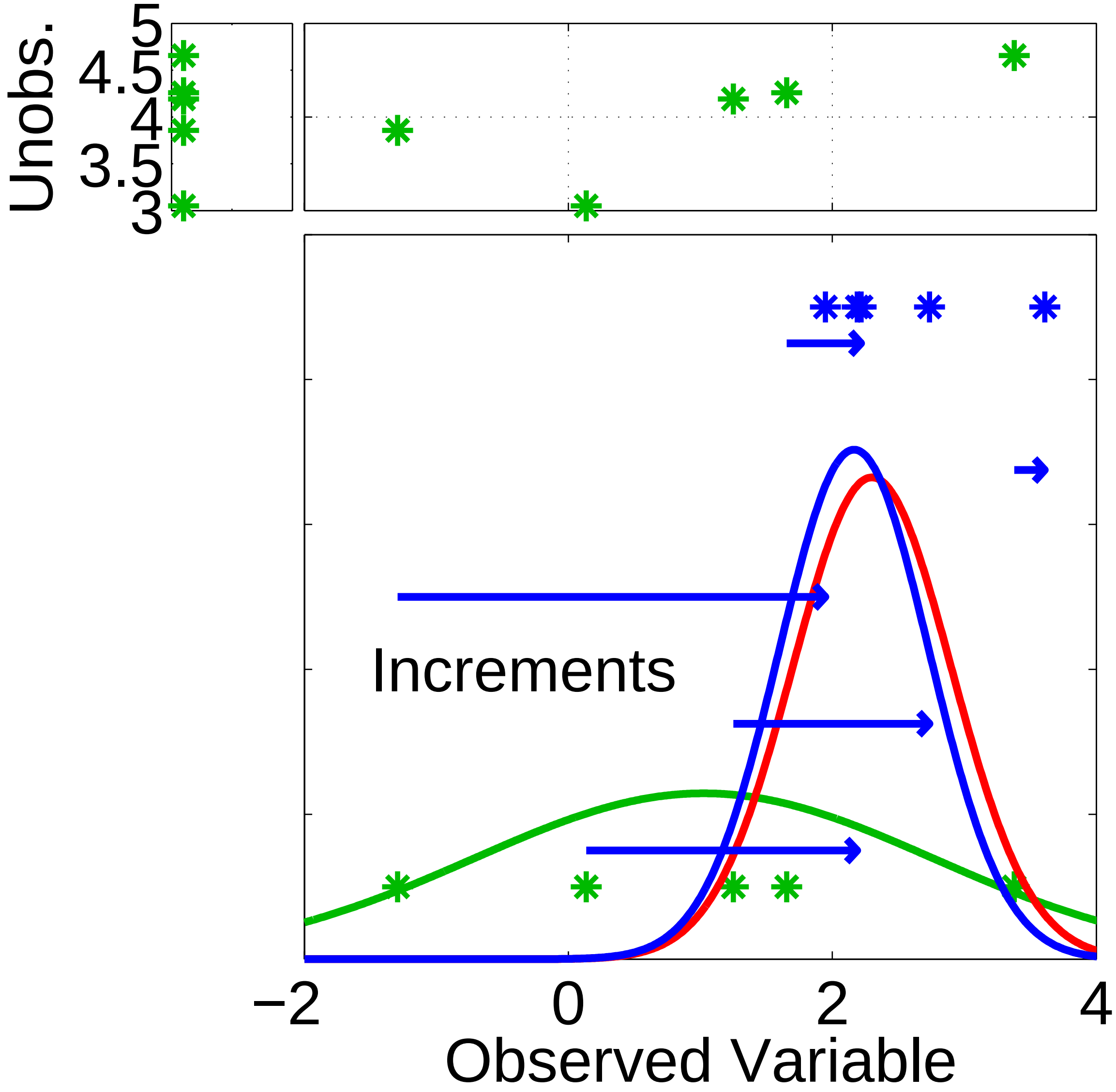


Assume that all we know is prior joint distribution.

One variable is observed.

Compute increments for prior ensemble members of observed variable.

Ensemble filters: Updating additional prior state variables

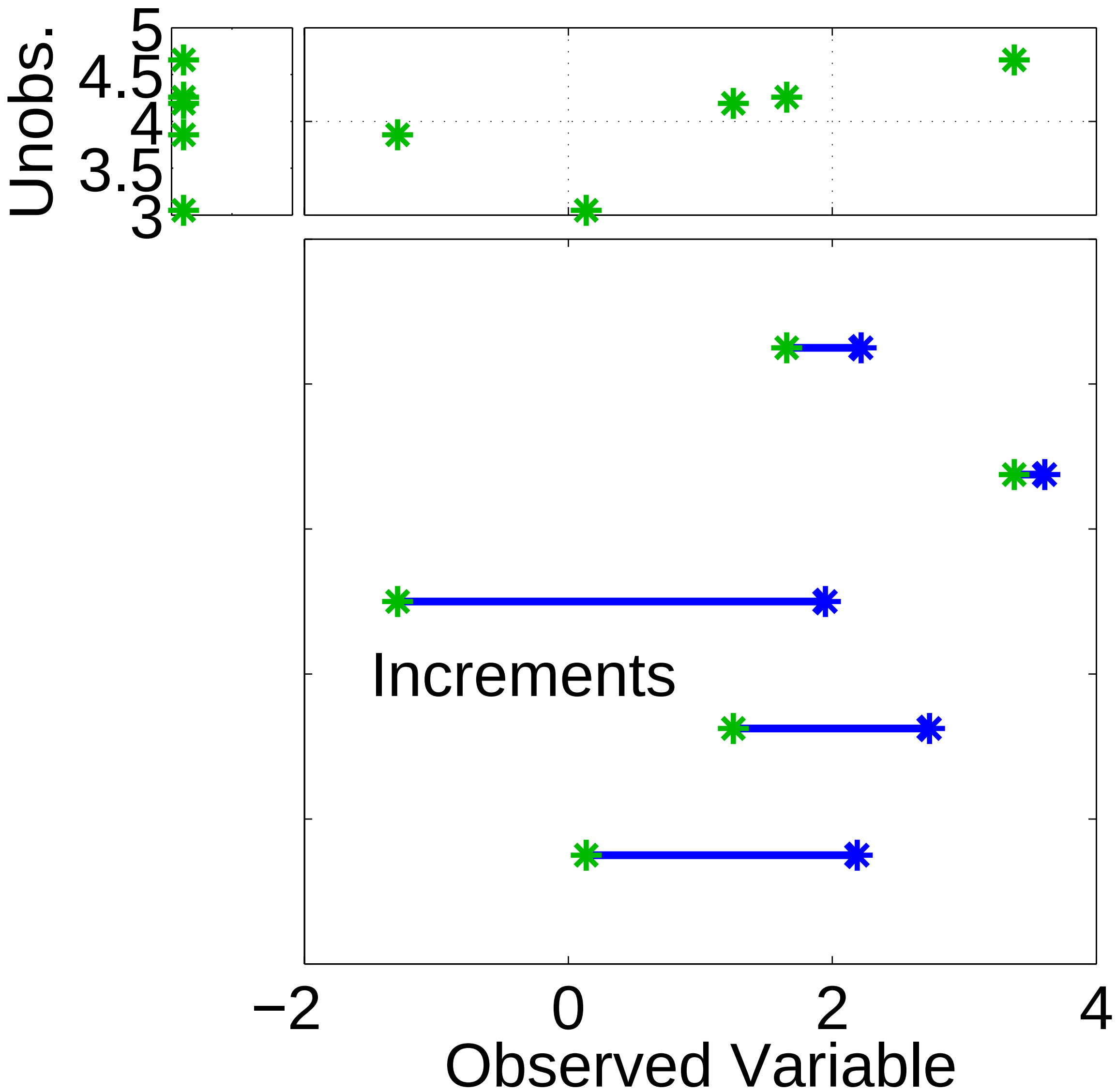


Assume that all we know is prior joint distribution.

One variable is observed.

Compute increments for prior ensemble members of observed variable.

Ensemble filters: Updating additional prior state variables

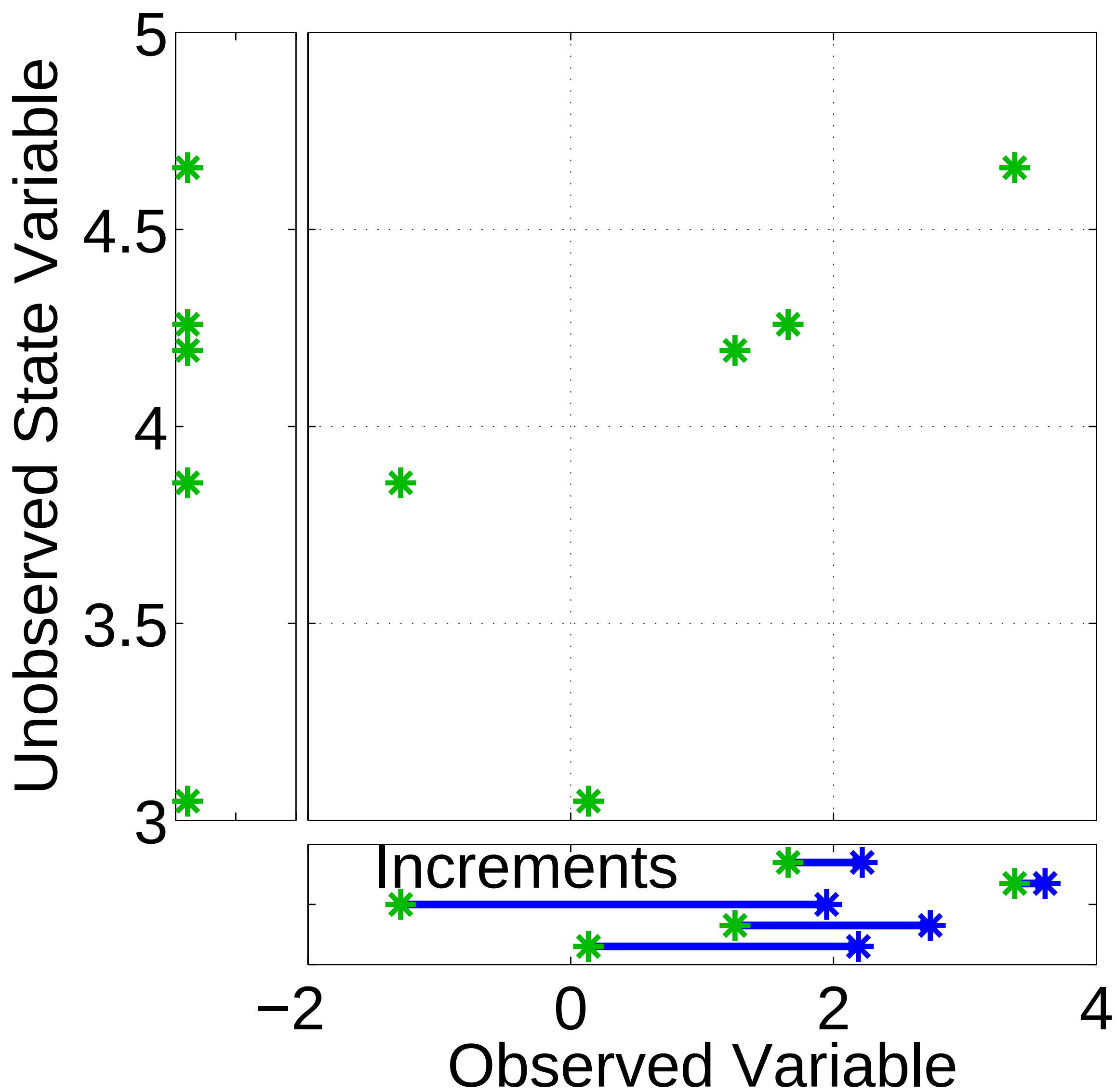


Assume that all we know is prior joint distribution.

One variable is observed.

Using only increments guarantees that if observation had no impact on observed variable, unobserved variable is unchanged (highly desirable).

Ensemble filters: Updating additional prior state variables



Assume that all we know is prior joint distribution.

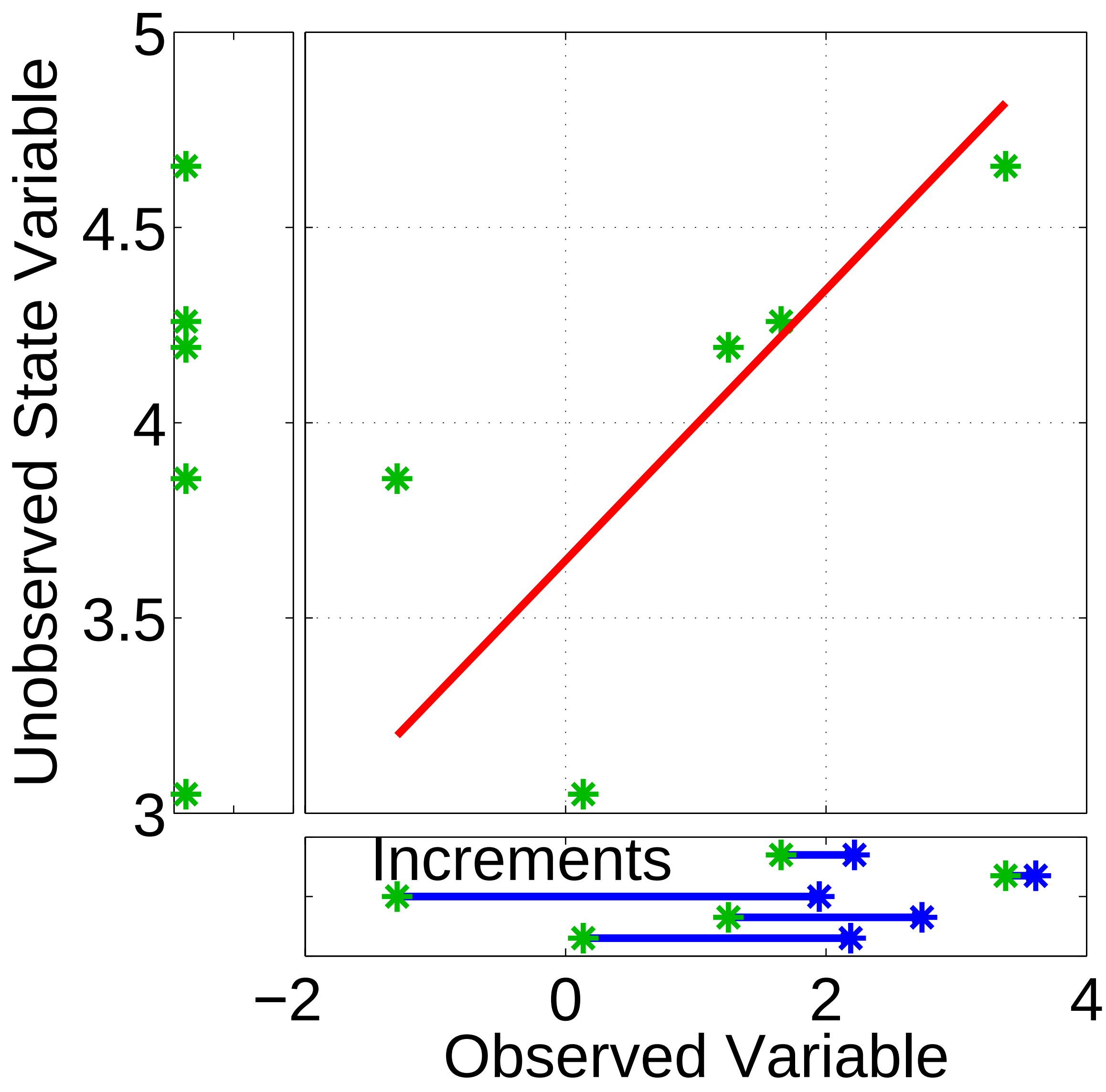
How should the unobserved variable be impacted?

First choice: least squares

Equivalent to linear regression.

Same as assuming binormal prior.

Ensemble filters: Updating additional prior state variables



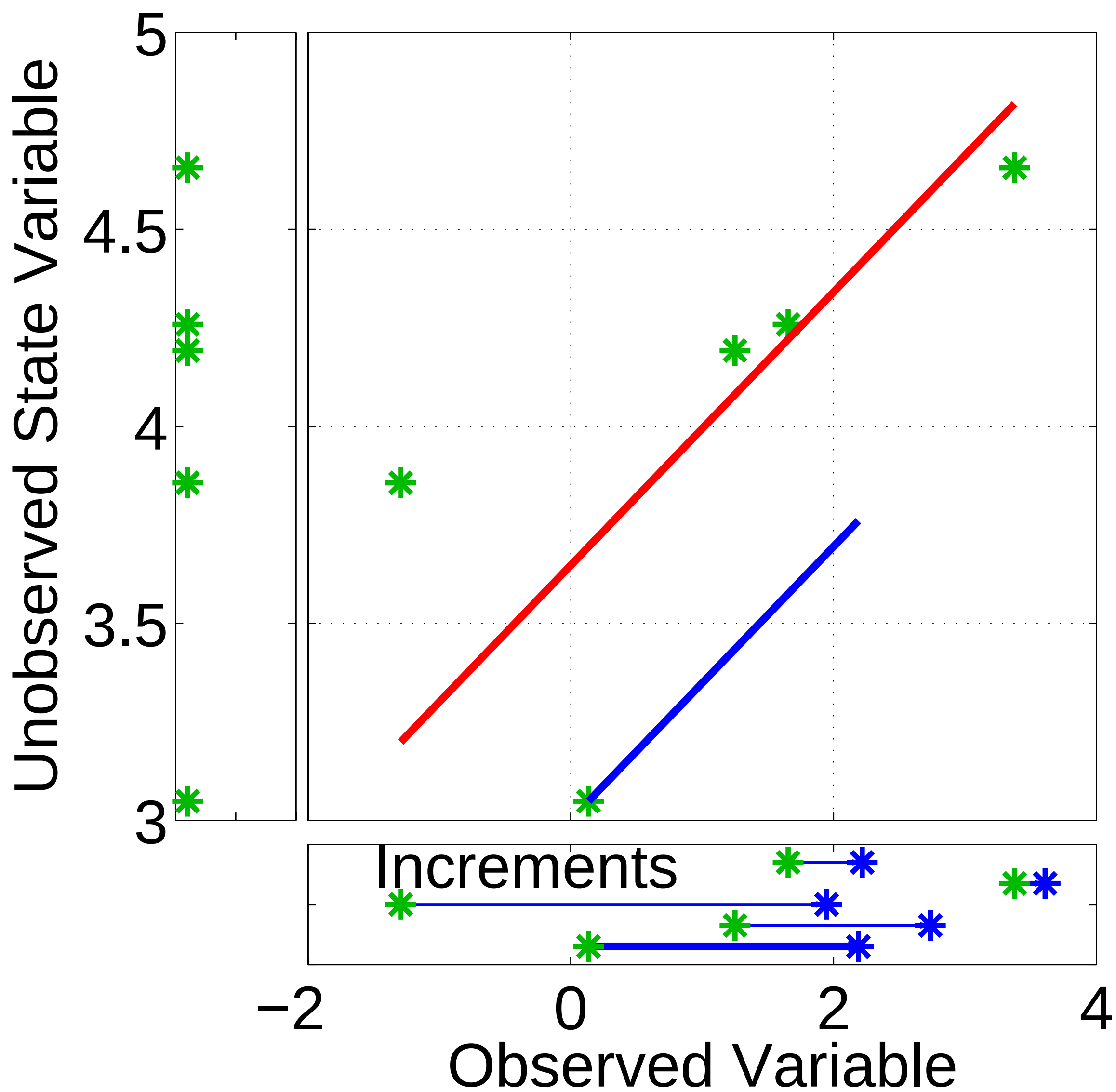
Have joint prior distribution of two variables.

How should the unobserved variable be impacted?

First choice: least squares

Begin by finding least squares fit.

Ensemble filters: Updating additional prior state variables

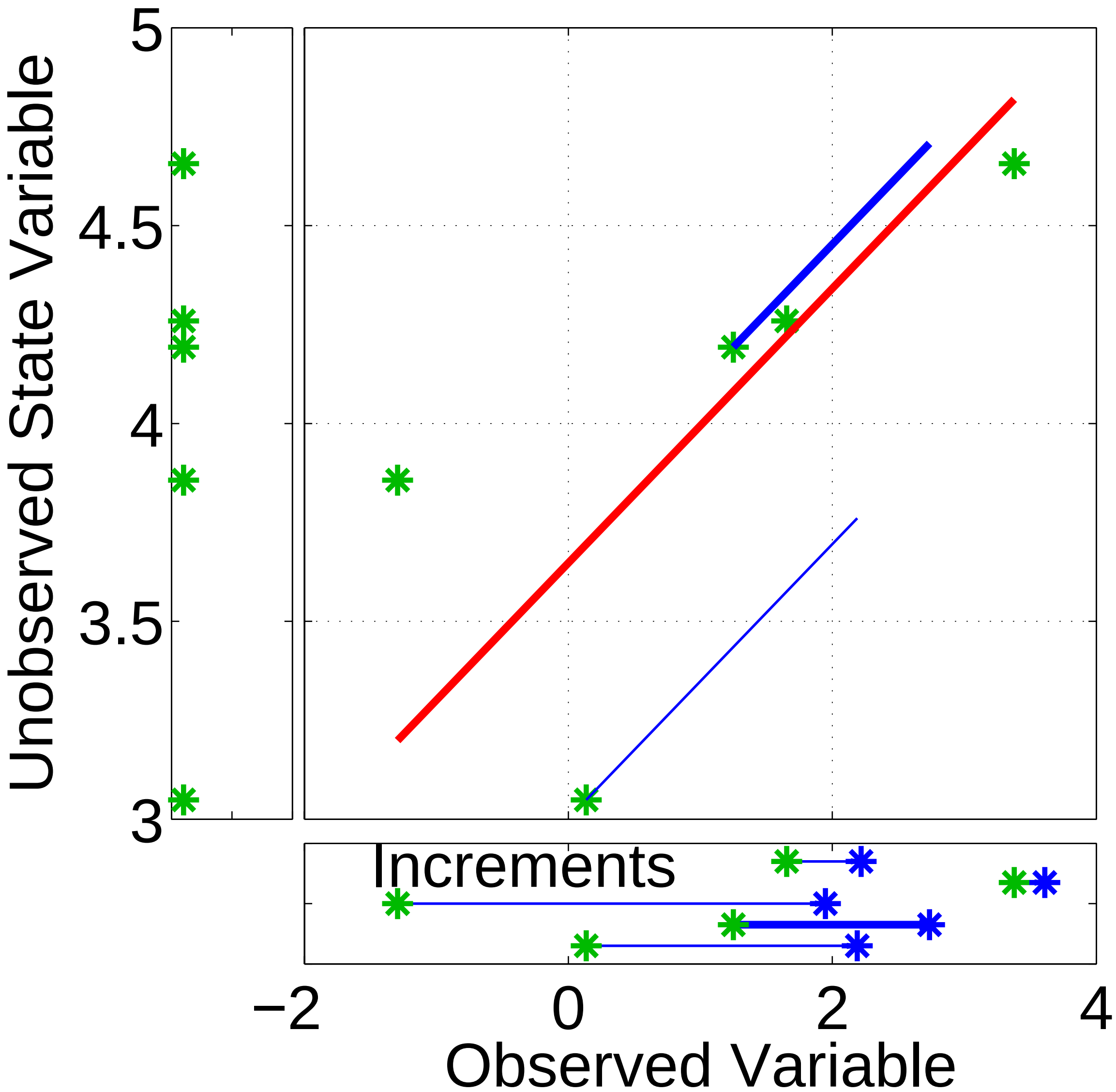


Have joint prior distribution of two variables.

Next, regress the observed variable increments onto increments for the unobserved variable.

Equivalent to first finding image of increment in joint space.

Ensemble filters: Updating additional prior state variables

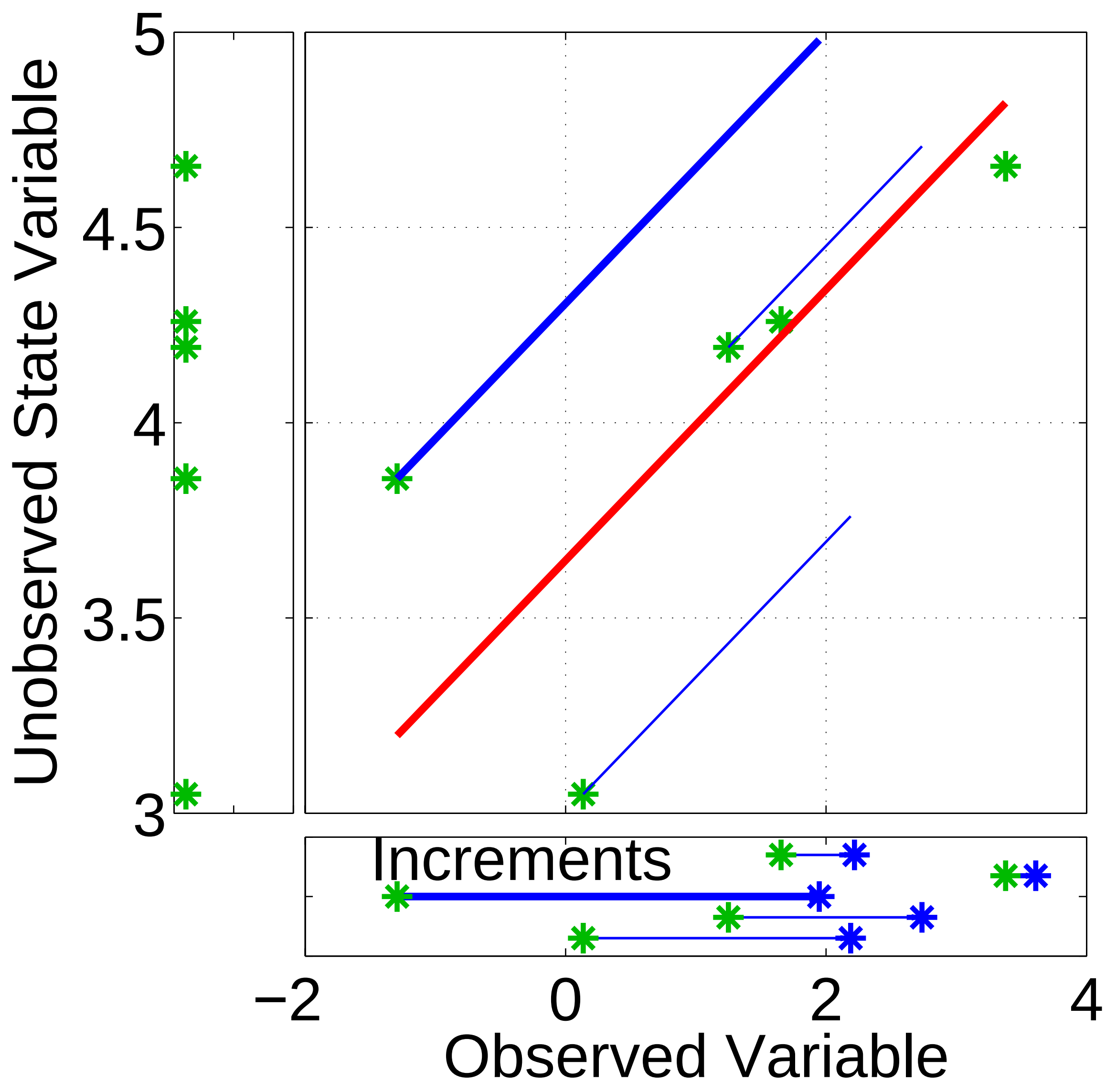


Have joint prior distribution of two variables.

Next, regress the observed variable increments onto increments for the unobserved variable.

Equivalent to first finding image of increment in joint space.

Ensemble filters: Updating additional prior state variables

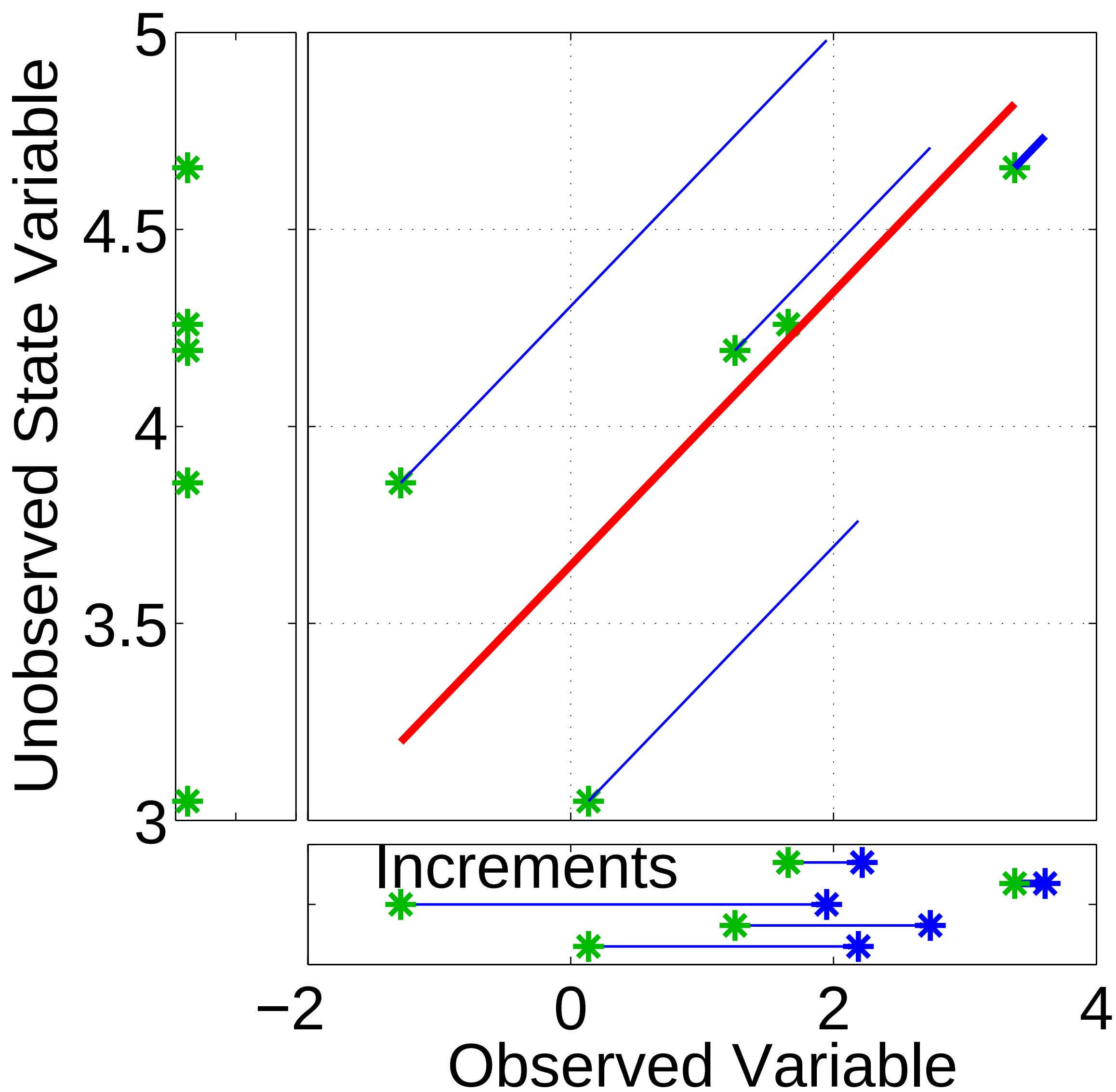


Have joint prior distribution of two variables.

Next, regress the observed variable increments onto increments for the unobserved variable.

Equivalent to first finding image of increment in joint space.

Ensemble filters: Updating additional prior state variables

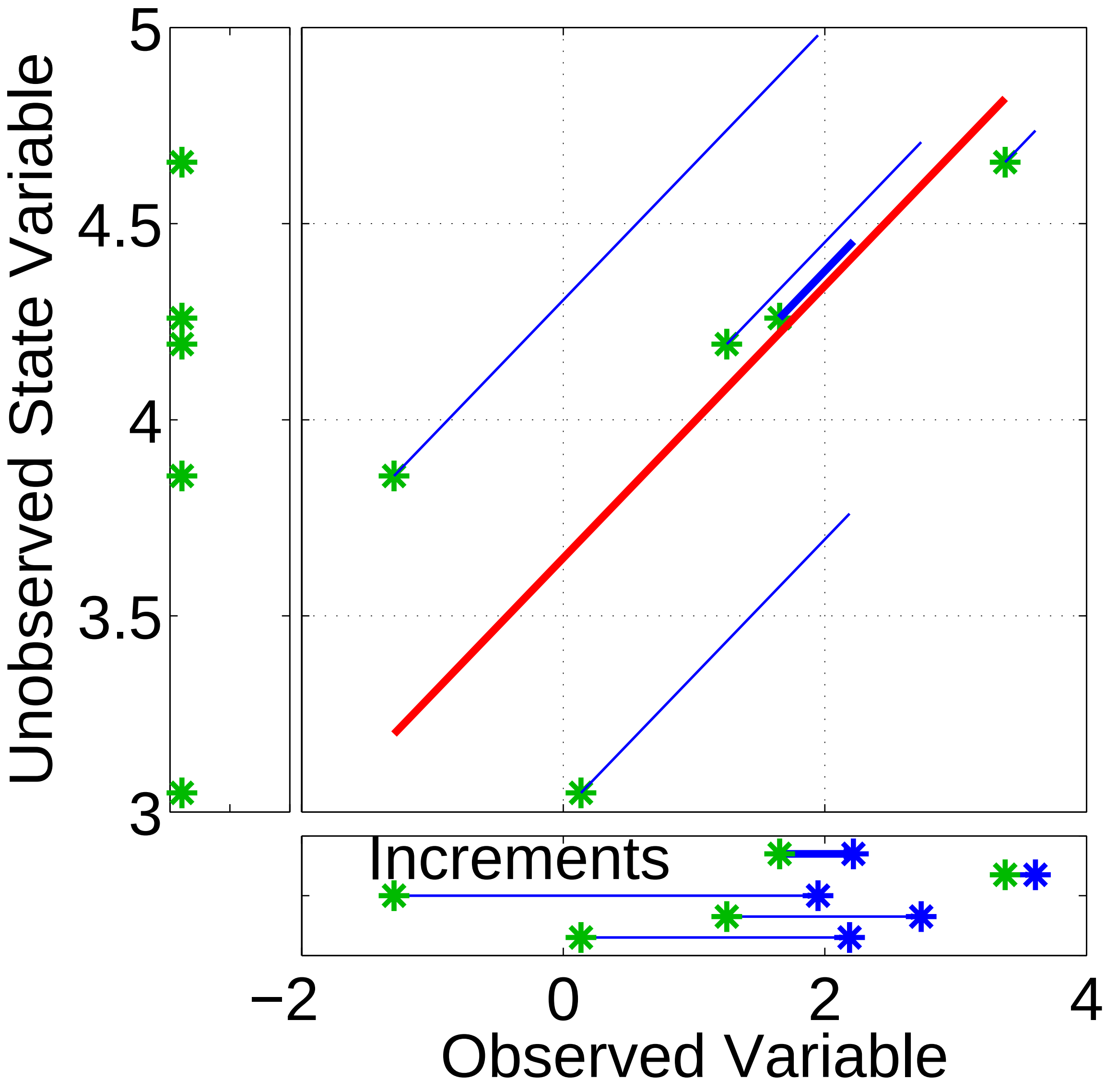


Have joint prior distribution of two variables.

Next, regress the observed variable increments onto increments for the unobserved variable.

Equivalent to first finding image of increment in joint space.

Ensemble filters: Updating additional prior state variables

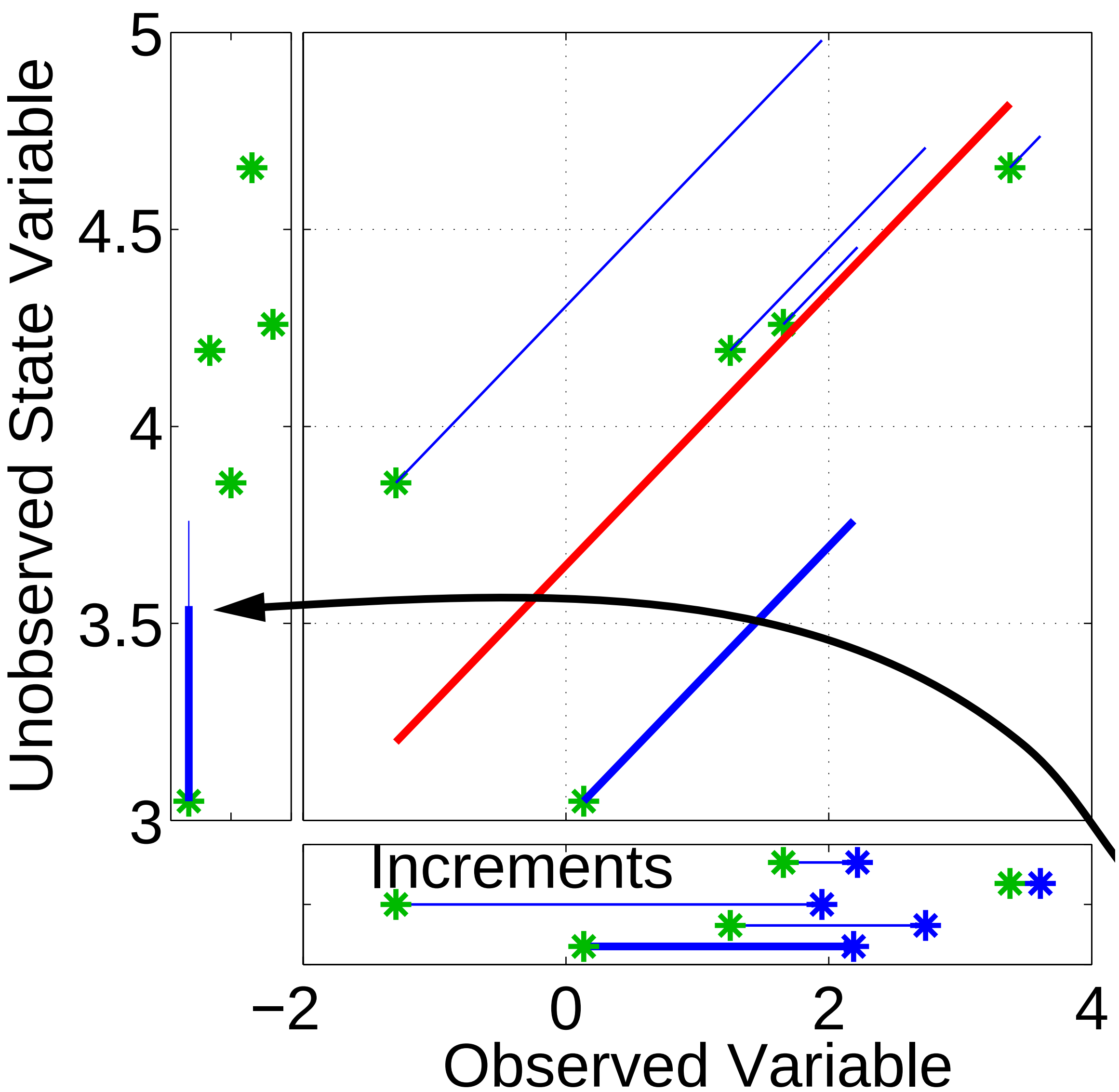


Have joint prior distribution of two variables.

Next, regress the observed variable increments onto increments for the unobserved variable.

Equivalent to first finding image of increment in joint space.

Ensemble filters: Updating additional prior state variables



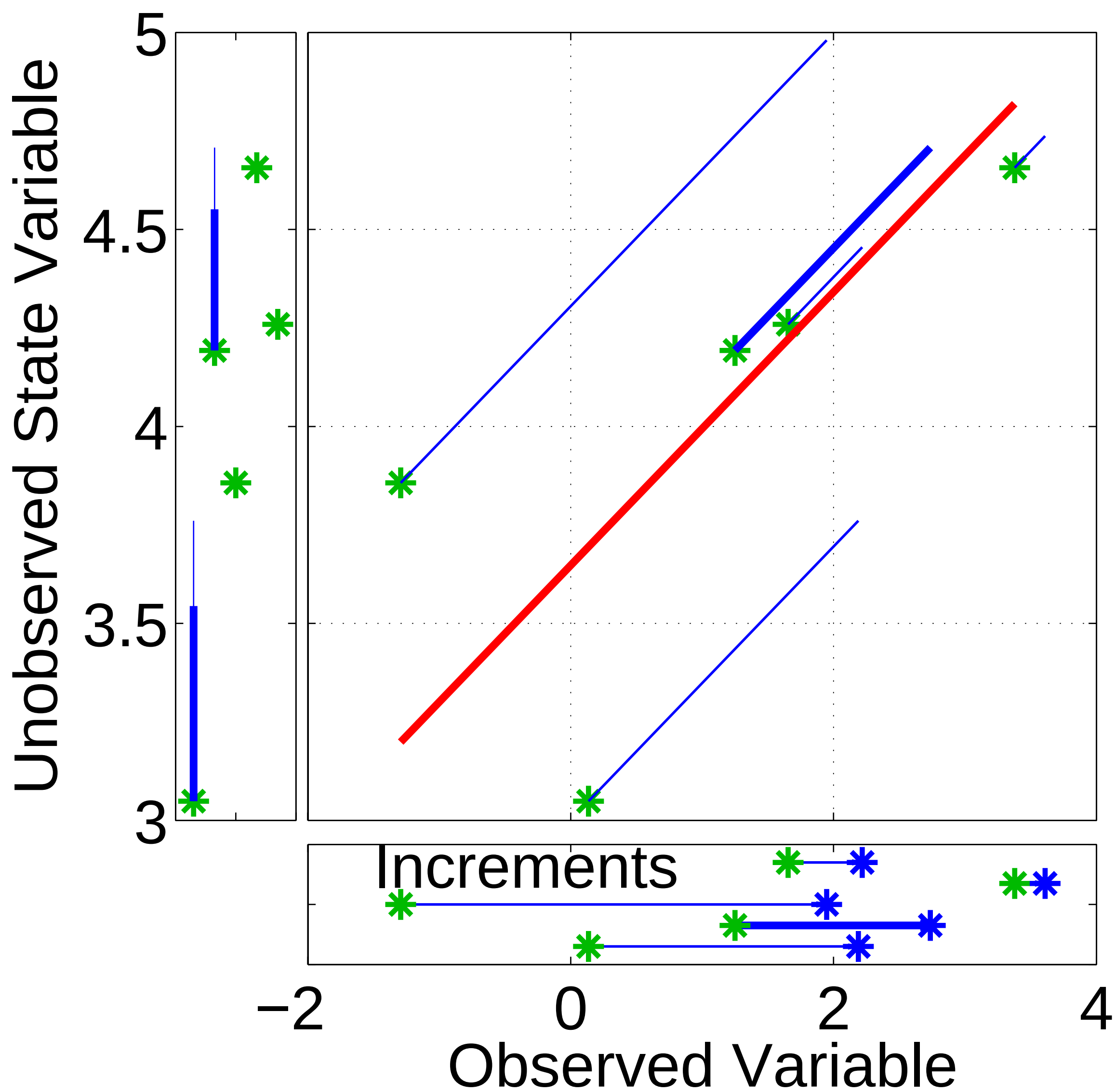
Have joint prior distribution of two variables.

Regression: Equivalent to first finding image of increment in joint space.

Then projecting from joint space onto unobserved priors.

Finally, multiply by prior sample correlation.

Ensemble filters: Updating additional prior state variables



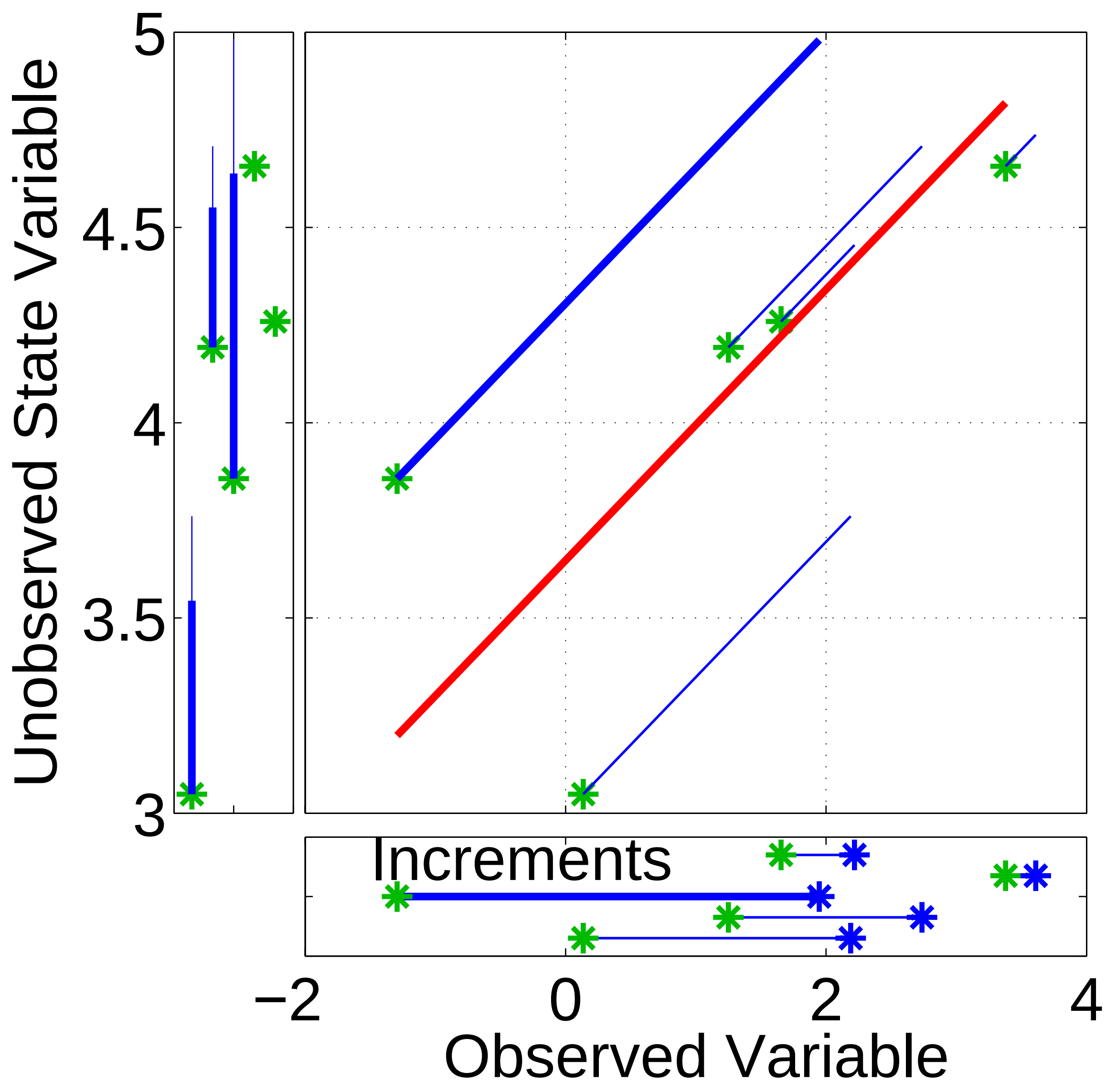
Have joint prior distribution of two variables.

Regression: Equivalent to first finding image of increment in joint space.

Then projecting from joint space onto unobserved priors.

Finally, multiply by prior sample correlation.

Ensemble filters: Updating additional prior state variables



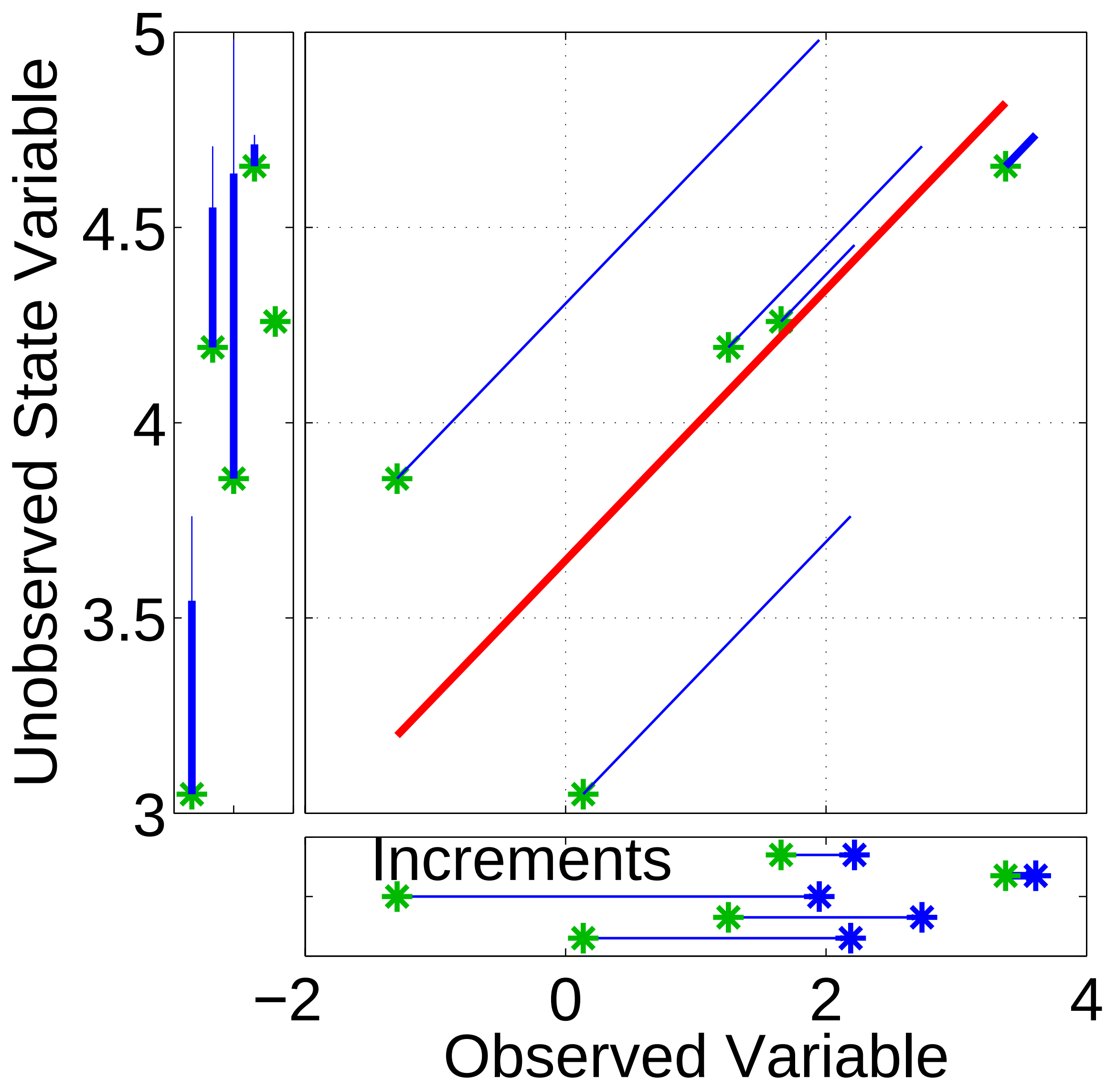
Have joint prior distribution of two variables.

Regression: Equivalent to first finding image of increment in joint space.

Then projecting from joint space onto unobserved priors.

Finally, multiply by prior sample correlation.

Ensemble filters: Updating additional prior state variables



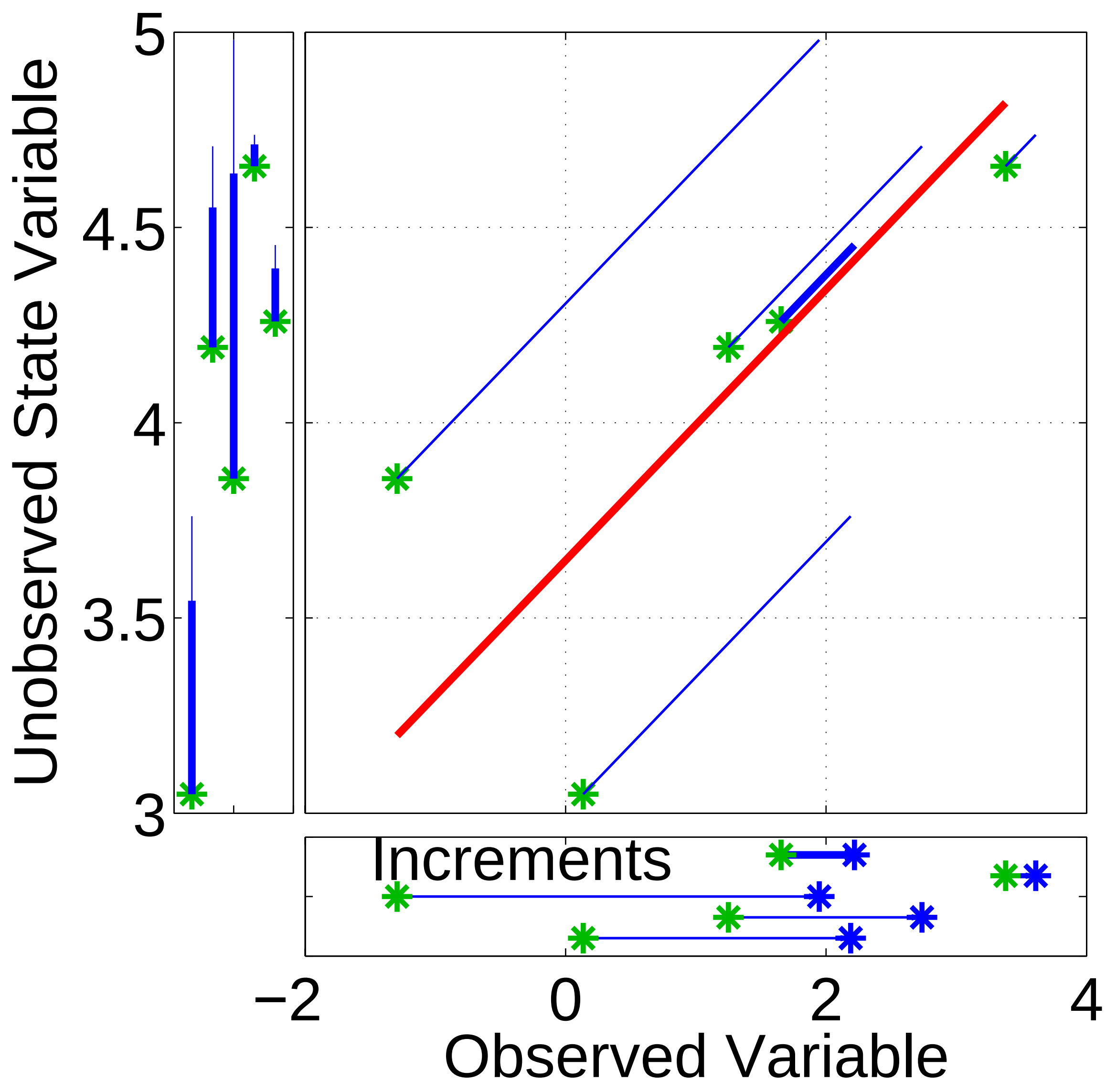
Have joint prior distribution of two variables.

Regression: Equivalent to first finding image of increment in joint space.

Then projecting from joint space onto unobserved priors.

Finally, multiply by prior sample correlation.

Ensemble filters: Updating additional prior state variables



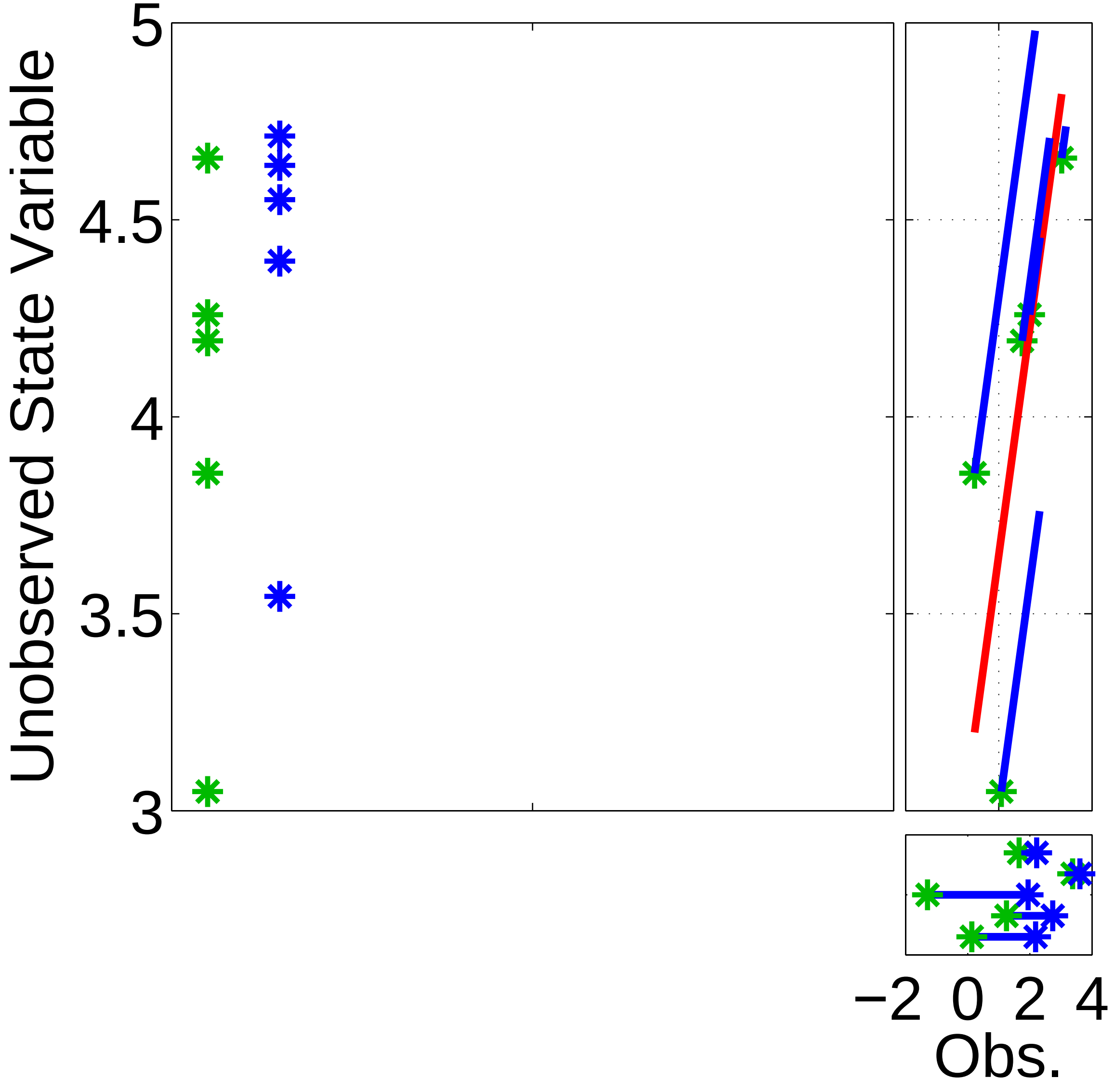
Have joint prior distribution of two variables.

Regression: Equivalent to first finding image of increment in joint space.

Then projecting from joint space onto unobserved priors.

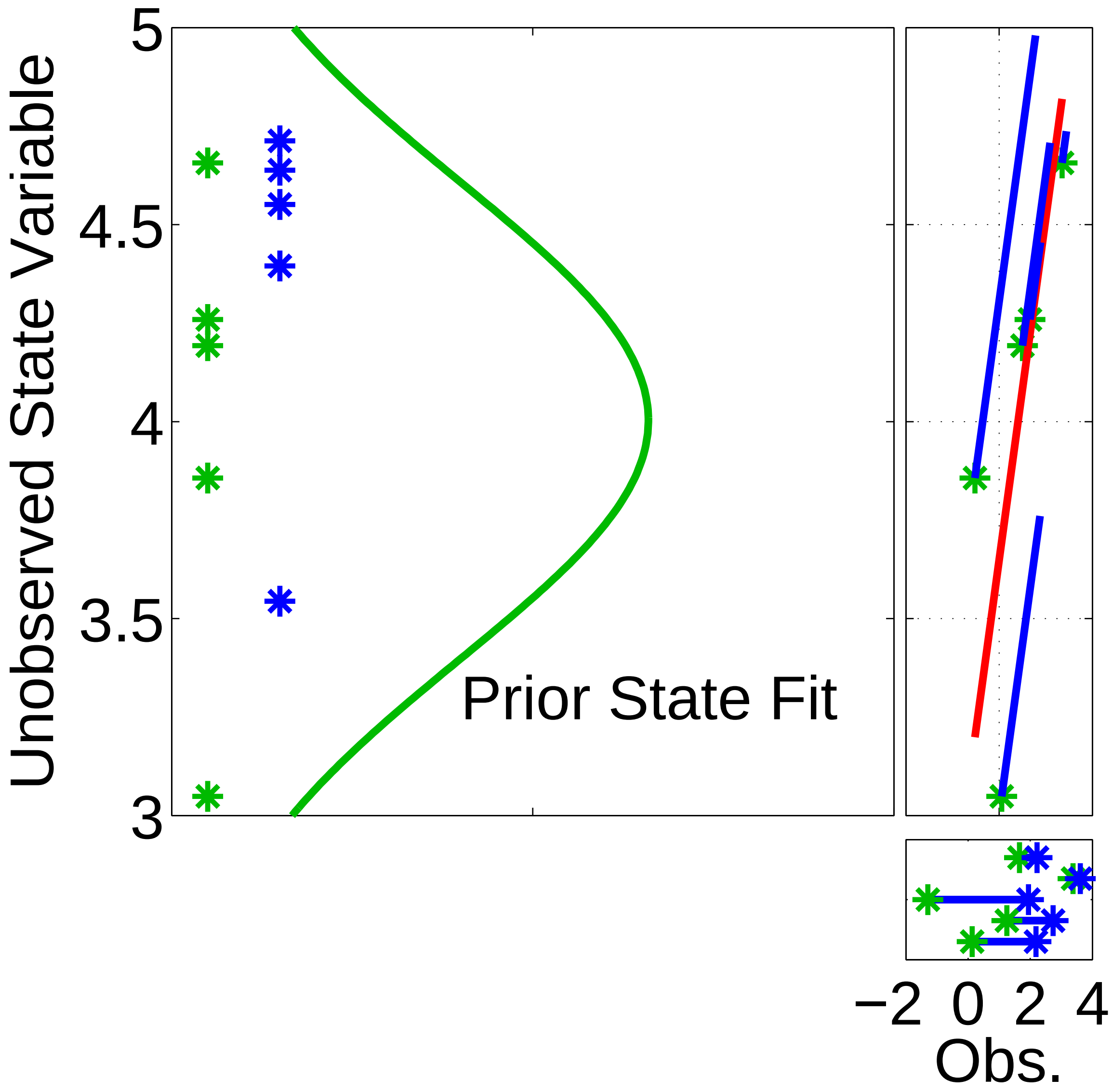
Finally, multiply by prior sample correlation.

Ensemble filters: Updating additional prior state variables



Now have an updated (posterior) ensemble for the unobserved variable.

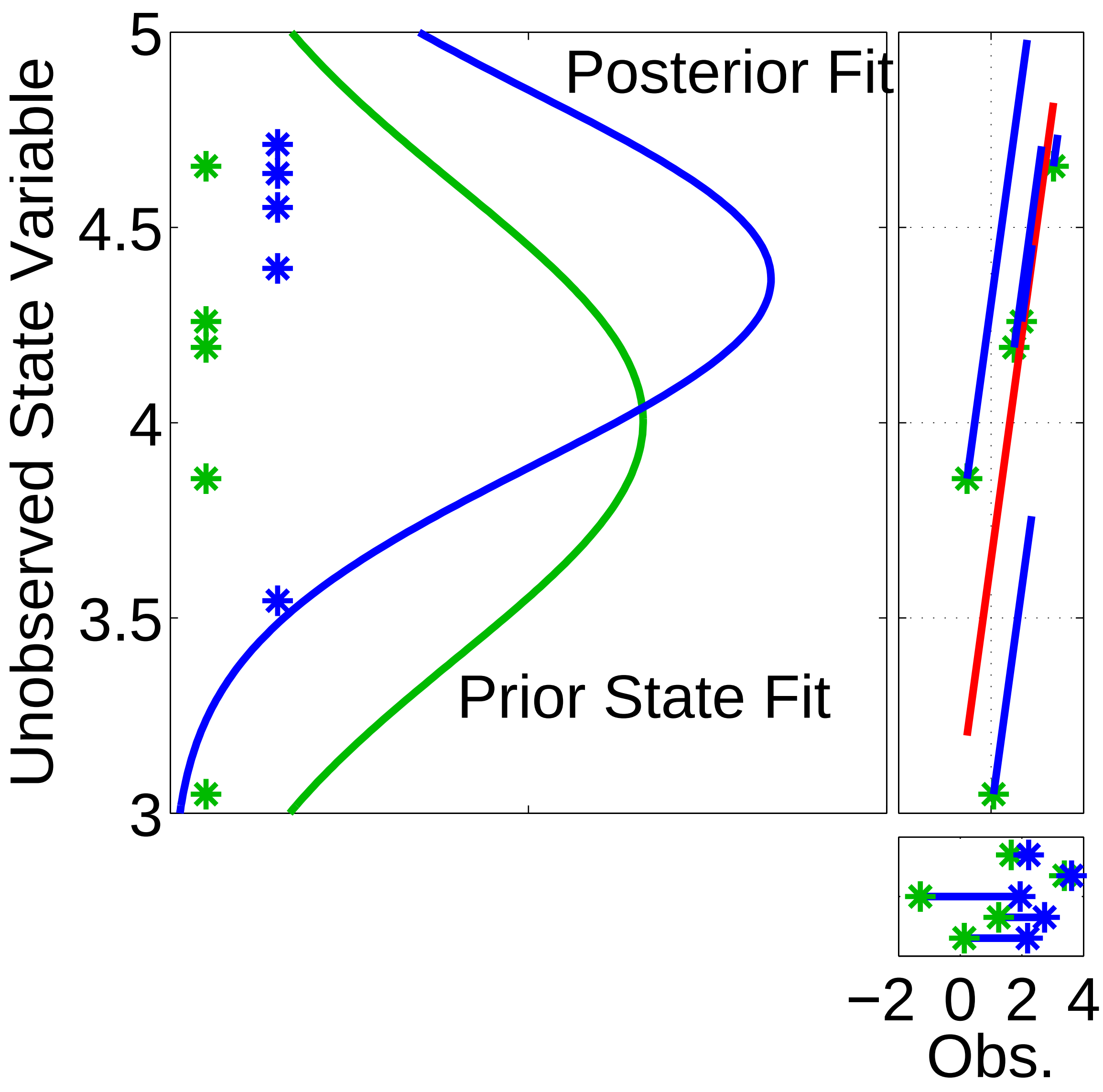
Ensemble filters: Updating additional prior state variables



Now have an updated (posterior) ensemble for the unobserved variable.

Fitting Gaussians shows that mean and variance have changed.

Ensemble filters: Updating additional prior state variables

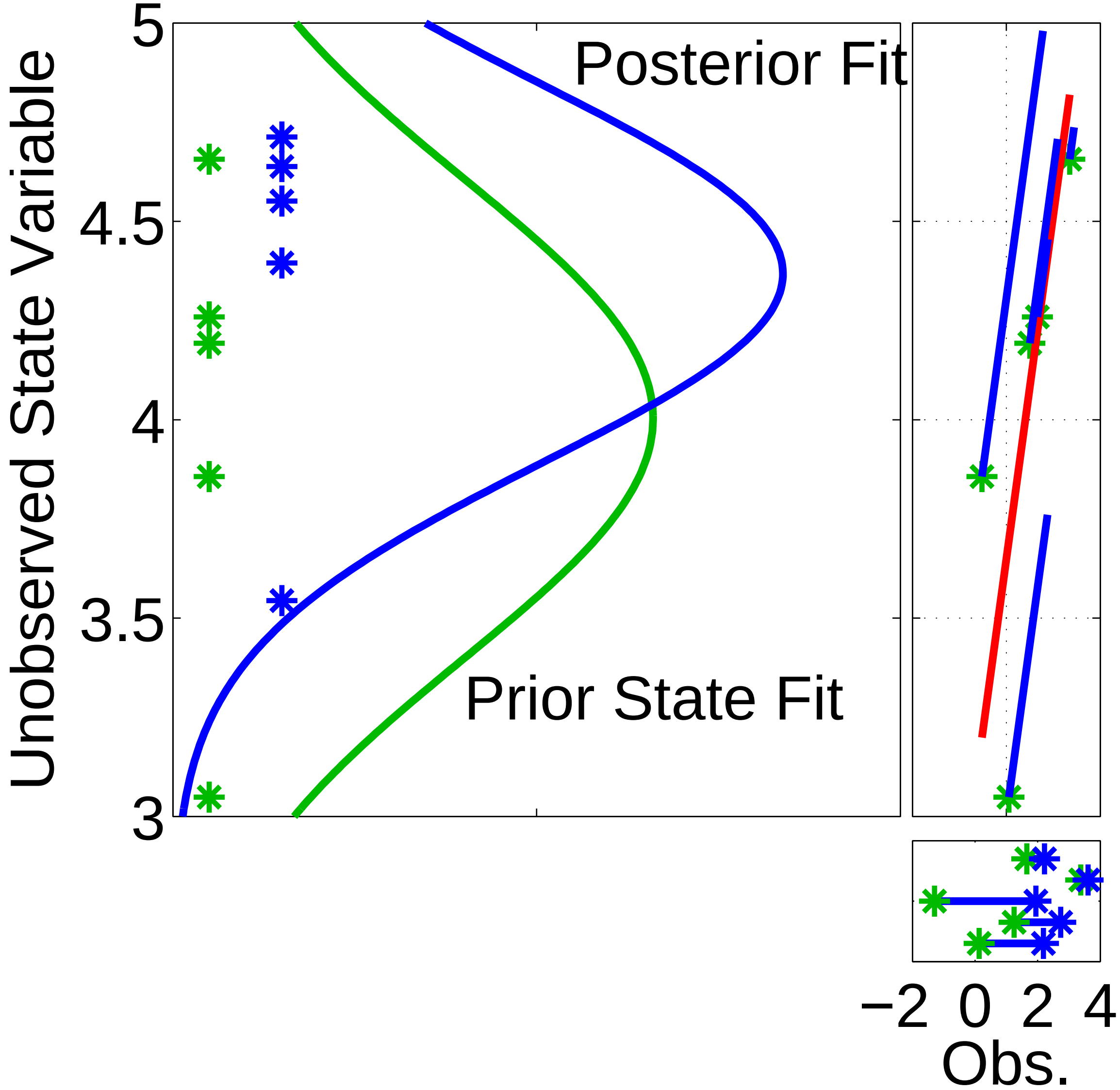


Now have an updated (posterior) ensemble for the unobserved variable.

Fitting Gaussians shows that mean and variance have changed.

Other features of the prior distribution may also have changed.

Ensemble filters: Updating additional prior state variables



CRITICAL POINT:

Since impact on unobserved variable is simply a linear regression, can do this INDEPENDENTLY for any number of unobserved variables!

Could also do many at once using matrix algebra as in traditional Kalman Filter.

Ensemble filters: Updating additional prior state variables

Two primary error sources:

<u>These are major issues for parameter estimation.</u>

1. Linear approximation is invalid.
   Substantial nonlinearity in 'true' relation over range of prior.

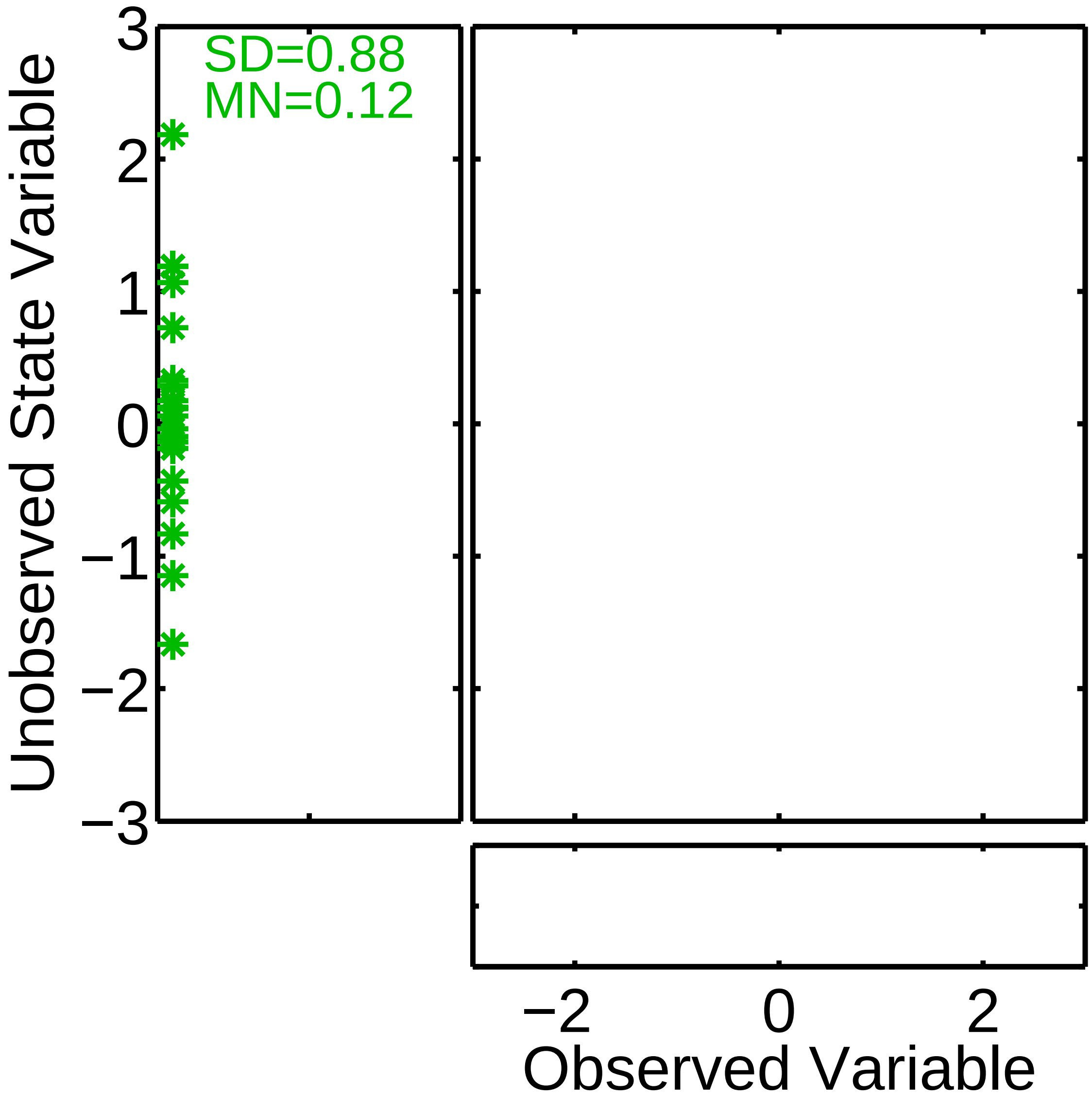
2. Sampling error due to noise.
   Even if linear relation, sample regression coefficient imprecise.

May need to address both issues for good performance.

<u>Parameter Estimation Questions:</u>

1. Are parameters 'linearly' related to state and observations?
2. Are parameters strongly or weakly related to observations?

Regression sampling error and filter divergence



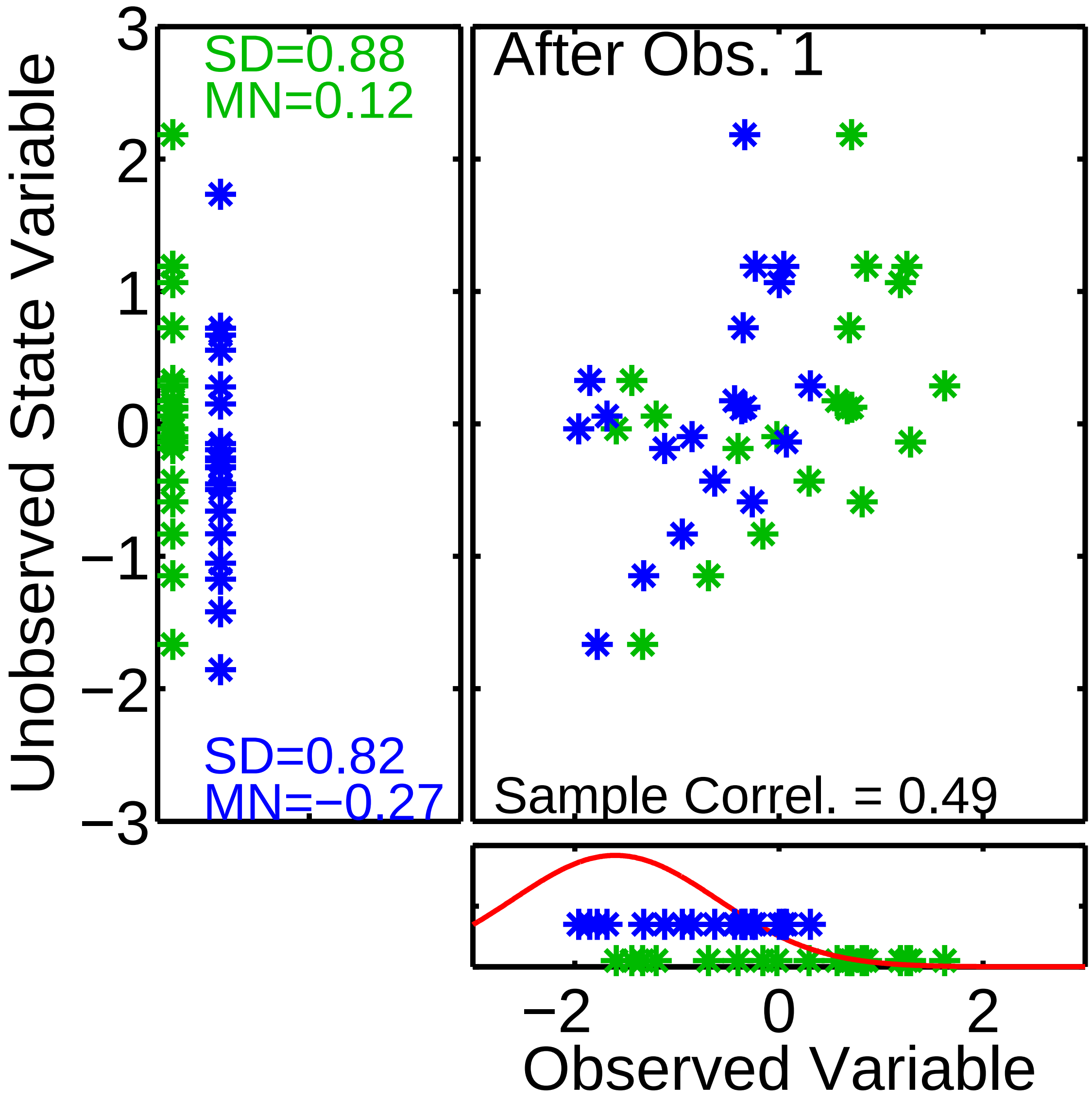
Suppose unobserved state variable is known to be unrelated to set of observed variables.

Unobserved variable should remain unchanged.

Let observations be of Antarctic wind velocity.

State variable is NCAR temp.

Regression sampling error and filter divergence

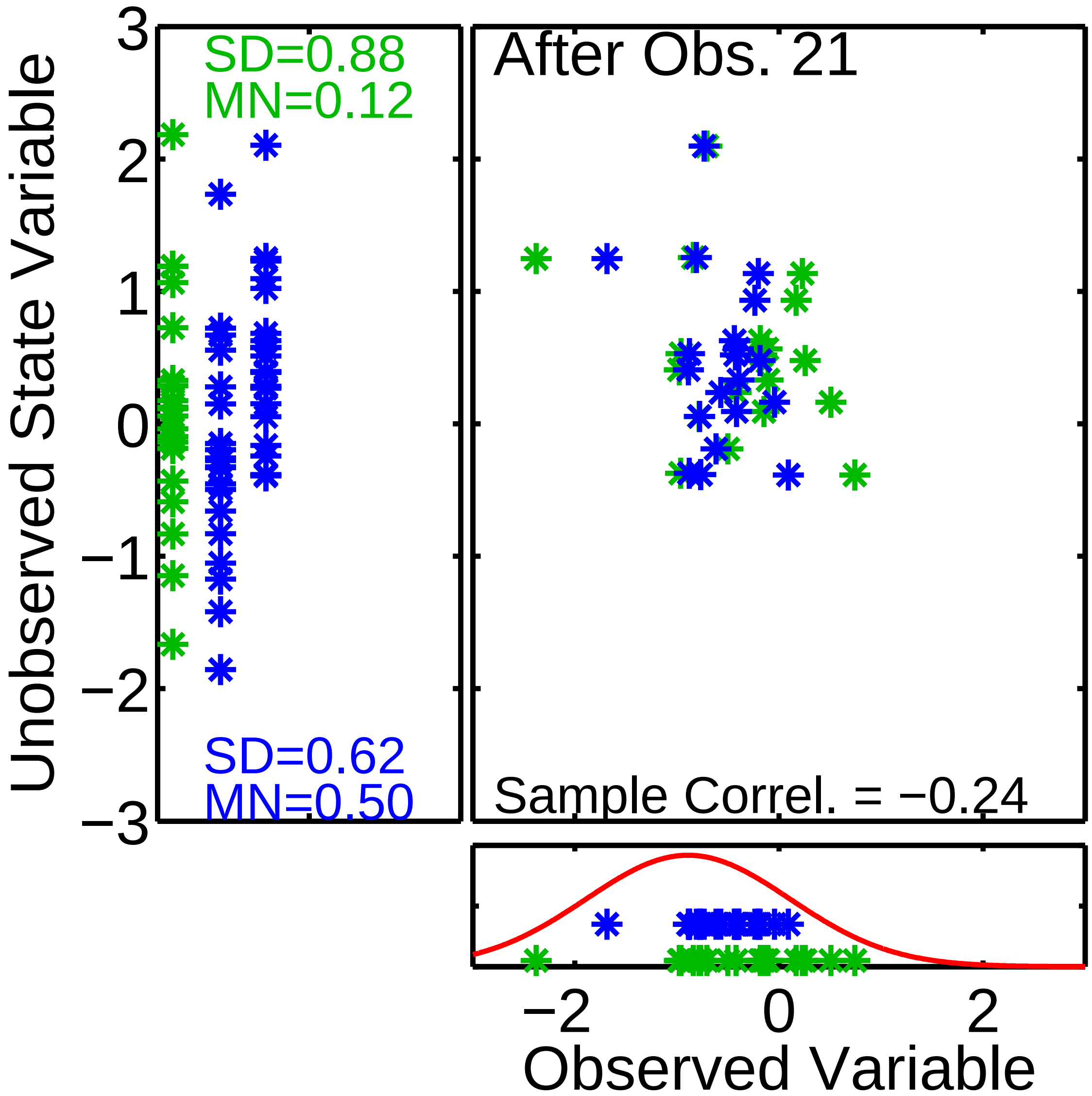


Suppose unobserved state variable is known to be unrelated to set of observed variables.

Finite samples from joint distribution will have non-zero correlation (expected |corr| = 0.19 for 20 samples).

After one observation, unobs. variable mean and S.D. change.

Regression sampling error and filter divergence

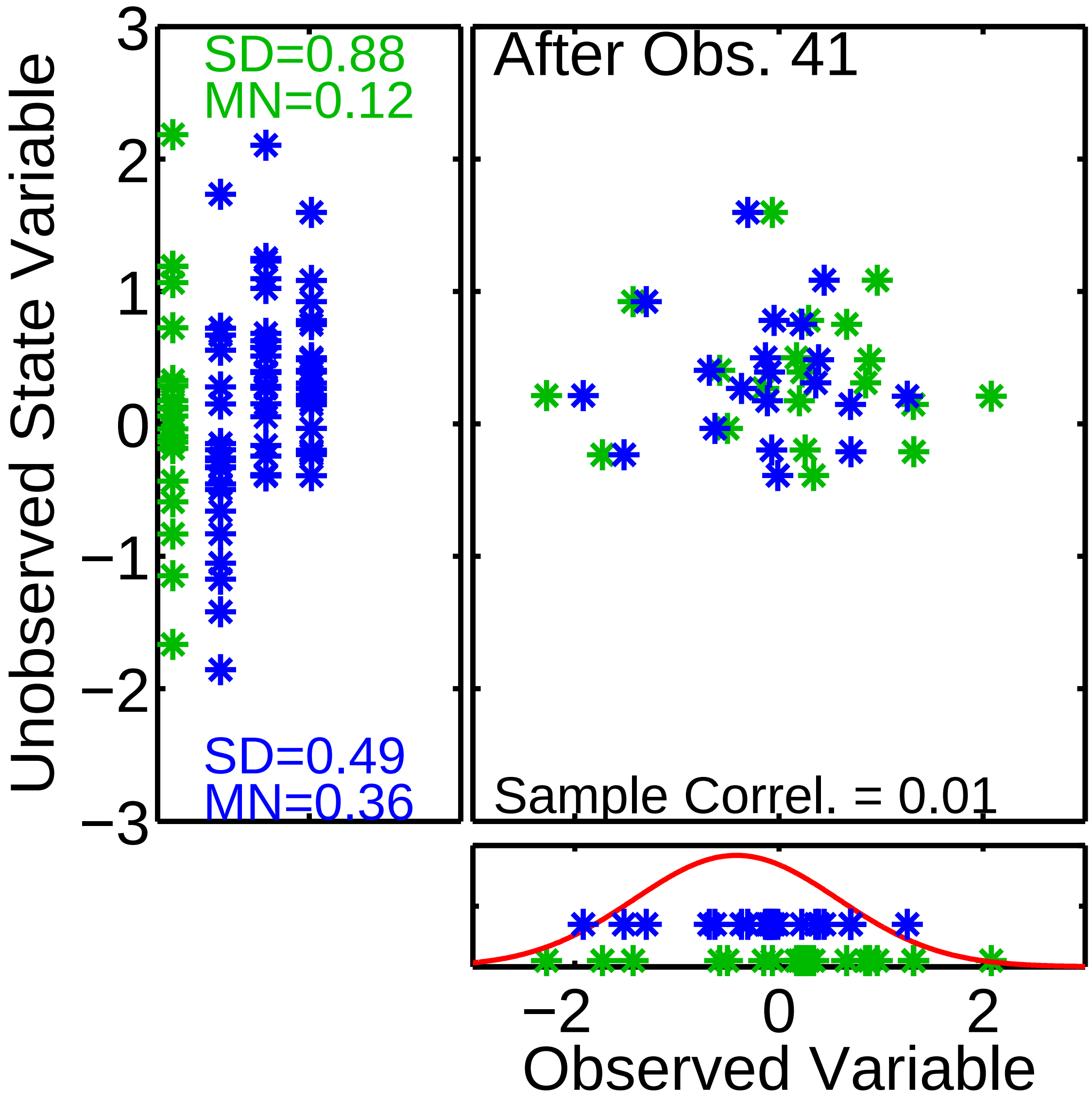


Suppose unobserved state variable is known to be unrelated to set of observed variables.

Unobserved variable should remain unchanged

Unobserved mean follows a random walk as more obs. are used.

Regression sampling error and filter divergence



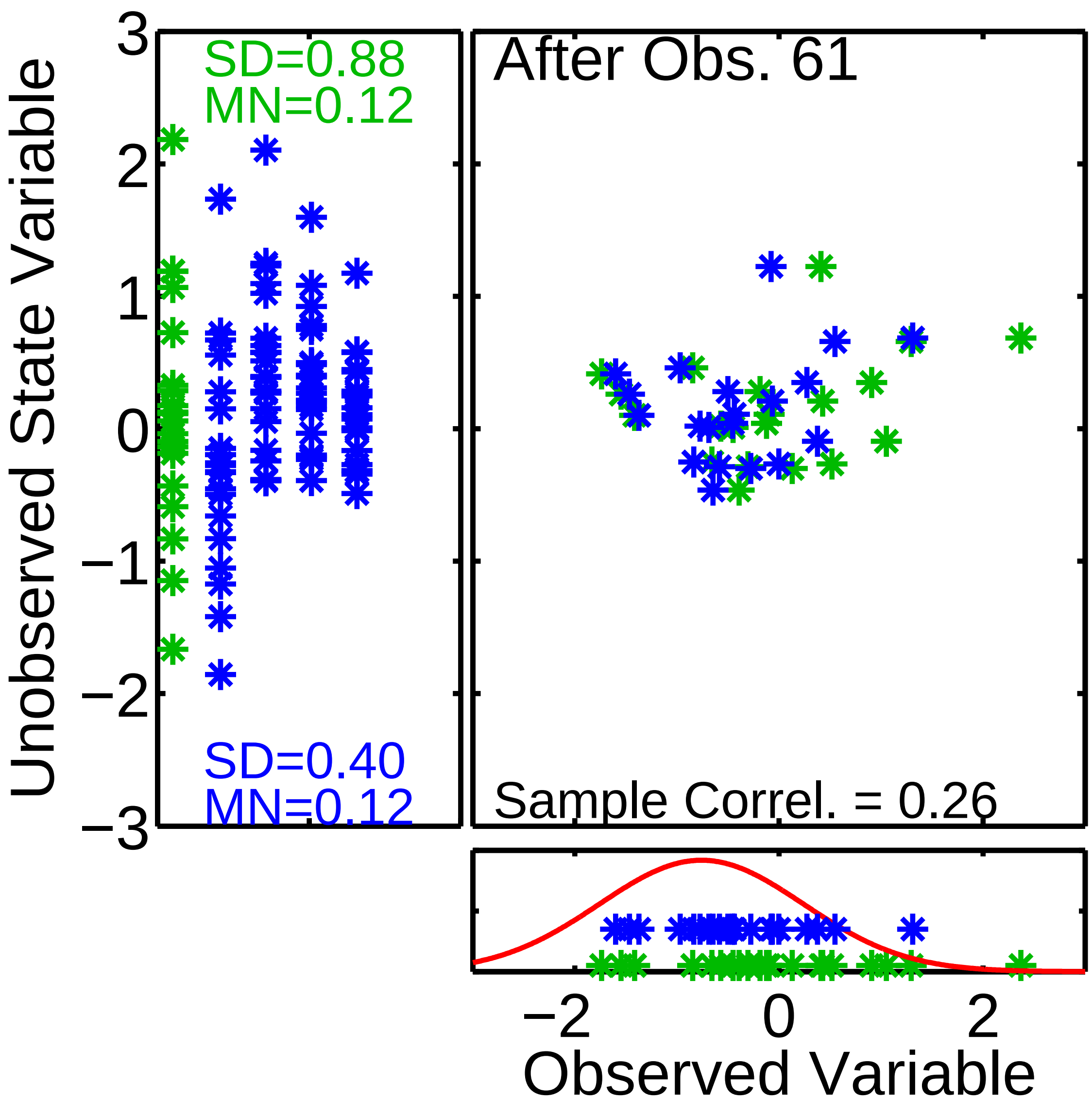
Suppose unobserved state variable is known to be unrelated to set of observed variables.

Unobserved variable should remain unchanged

Unobserved standard deviation is persistently decreased.

Expected change in |SD| is negative for any non-zero sample correlation!

Regression sampling error and filter divergence



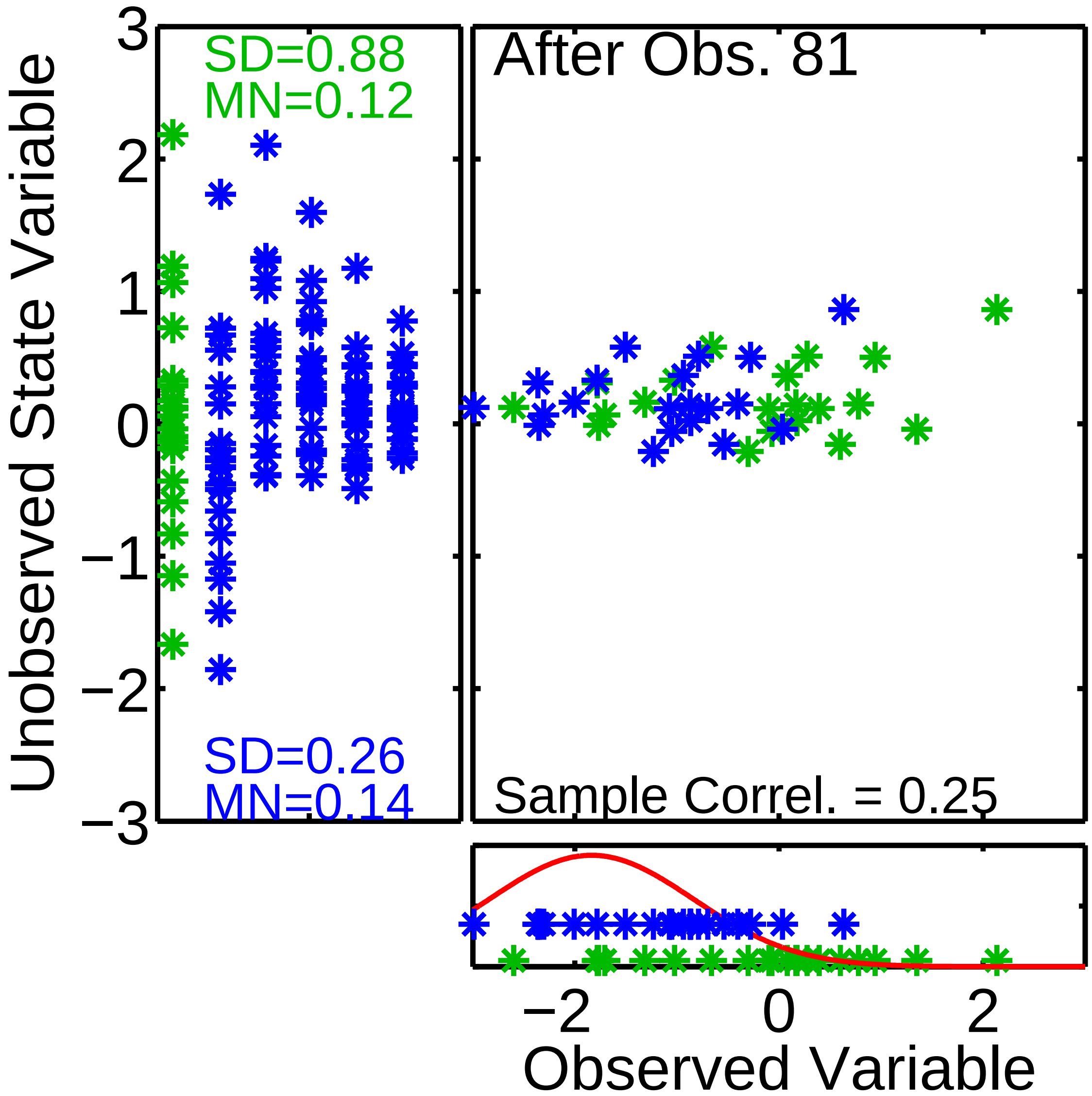
Suppose unobserved state variable is known to be unrelated to set of observed variables.

Unobserved variable should remain unchanged

Unobserved standard deviation is persistently decreased.

Expected change in |SD| is negative for any non-zero sample correlation!

Regression sampling error and filter divergence



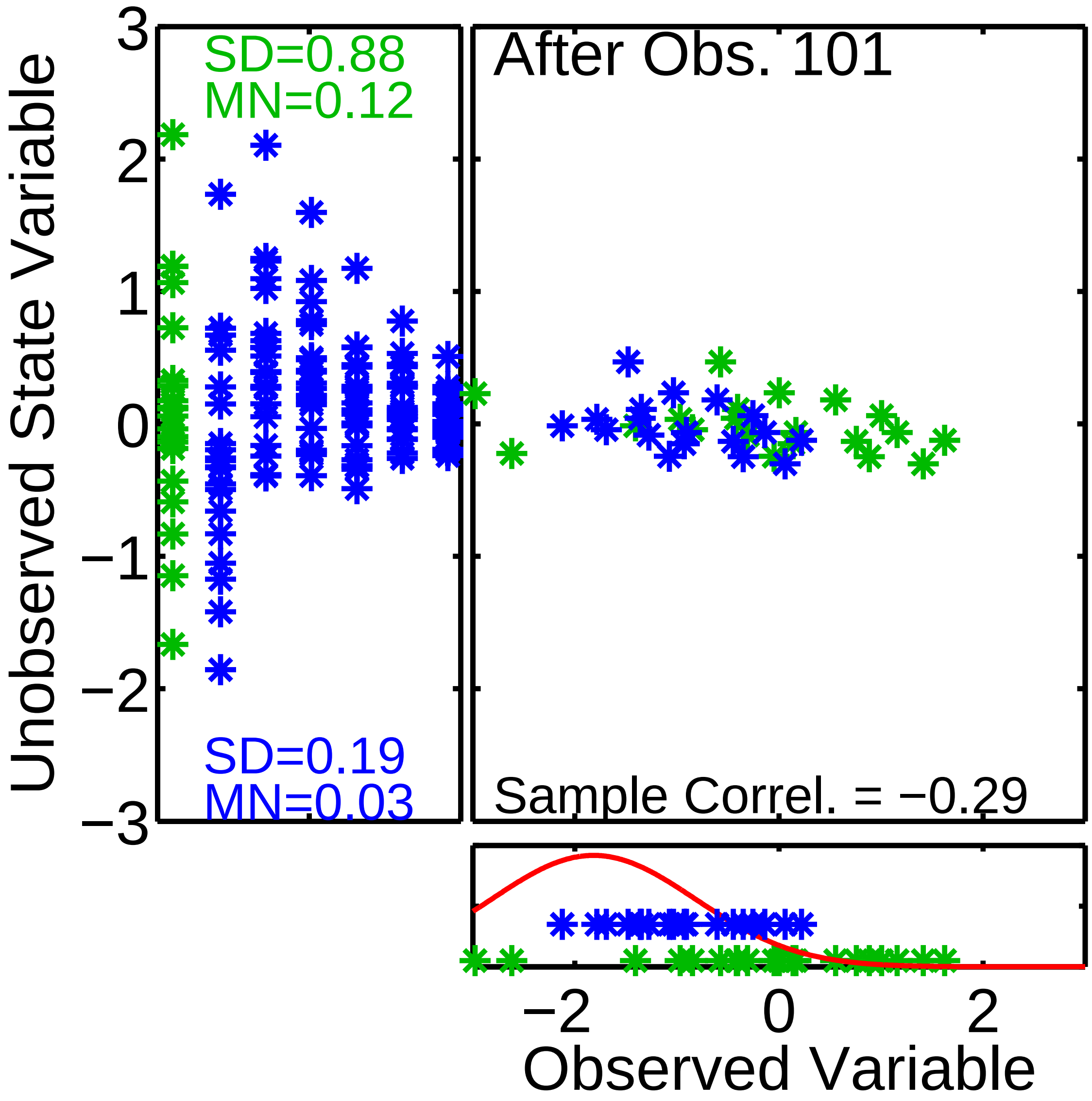
Suppose unobserved state variable is known to be unrelated to set of observed variables.

Unobserved variable should remain unchanged

Unobserved standard deviation is persistently decreased.

Expected change in |SD| is negative for any non-zero sample correlation!

Regression sampling error and filter divergence



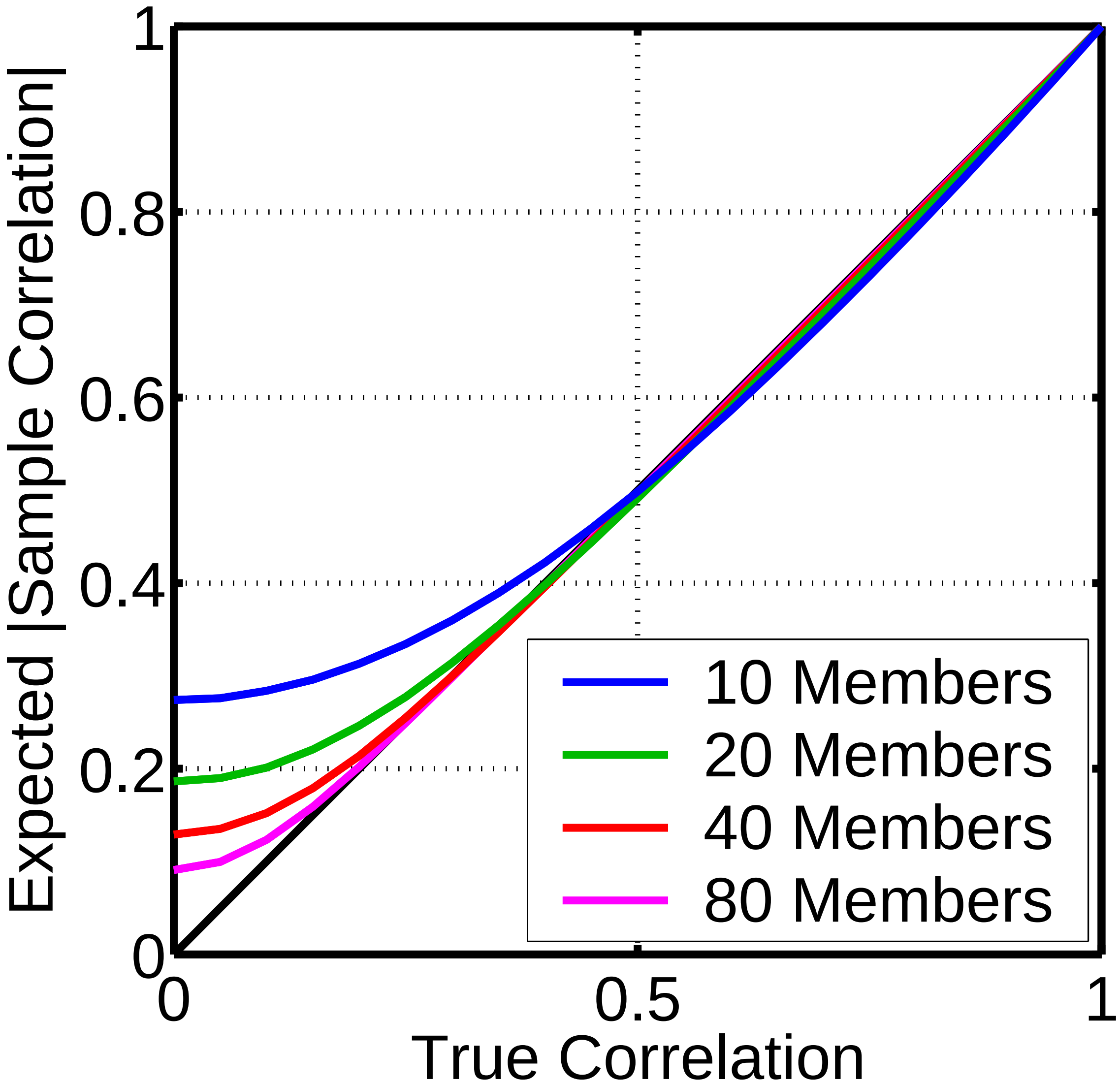
Suppose unobserved state variable is known to be unrelated to set of observed variables.

Estimates of unobs. become too confident

Give progressively less weight to any meaningful observations.

End result can be that meaningful obs. are essentially ignored.

Regression sampling error and filter divergence



Plot shows expected absolute value of sample correlation vs. true correlation.

Errors decrease with sample size and for large |real correlations|.

Do parameters have large or small expected correlation with observations?

Ways to deal with regression sampling error:

1. Ignore it: if number of unrelated observations is small
   and there is some way of maintaining variance in priors.

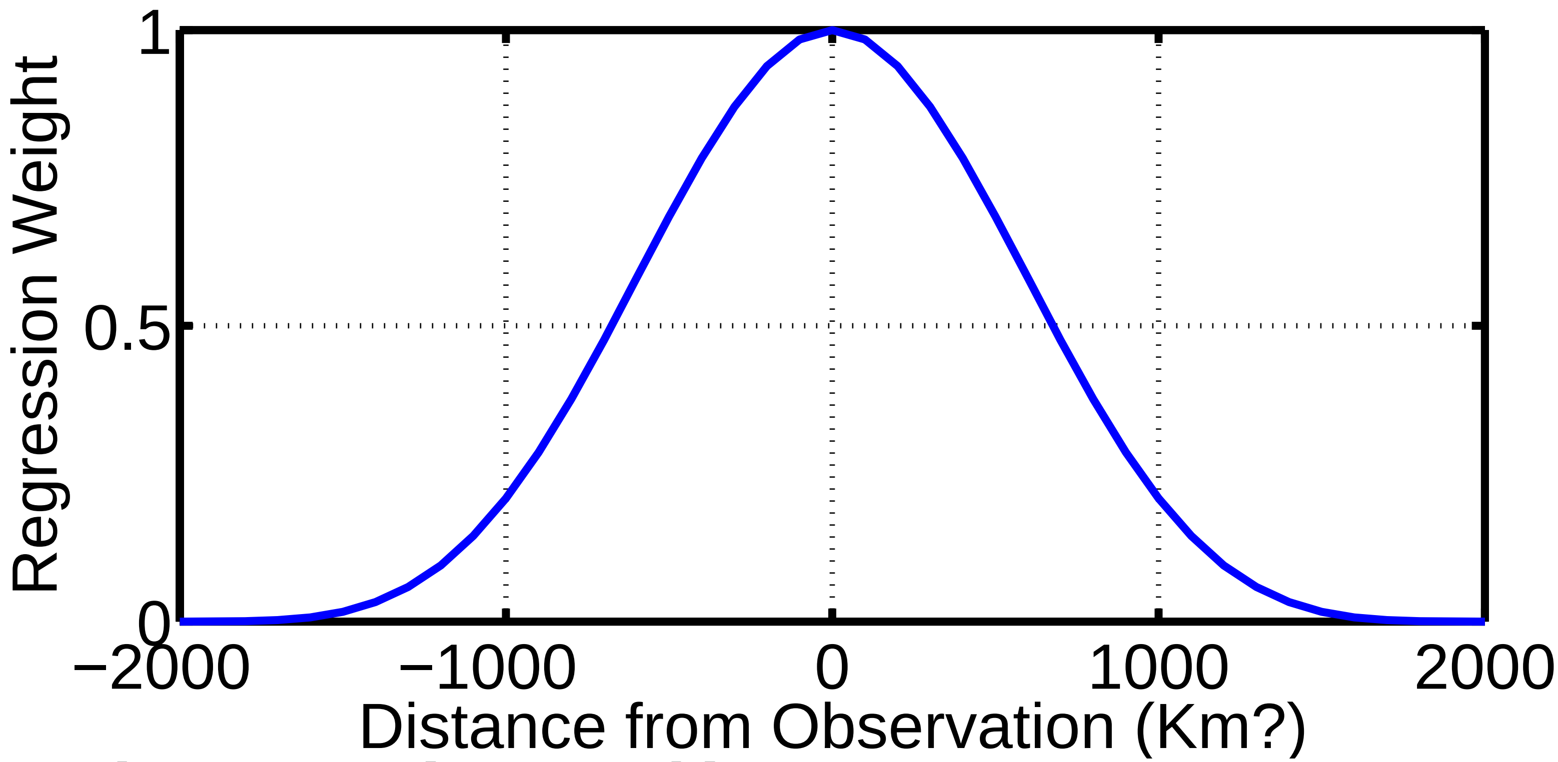
2. Use larger ensembles to limit sampling error.

3. Use additional a priori information about relation between
   observations and state variables.

4. Try to determine the amount of sampling error and correct for it.

Ways to deal with regression sampling error:

3. Use additional a priori information about relation between observations and state variables.



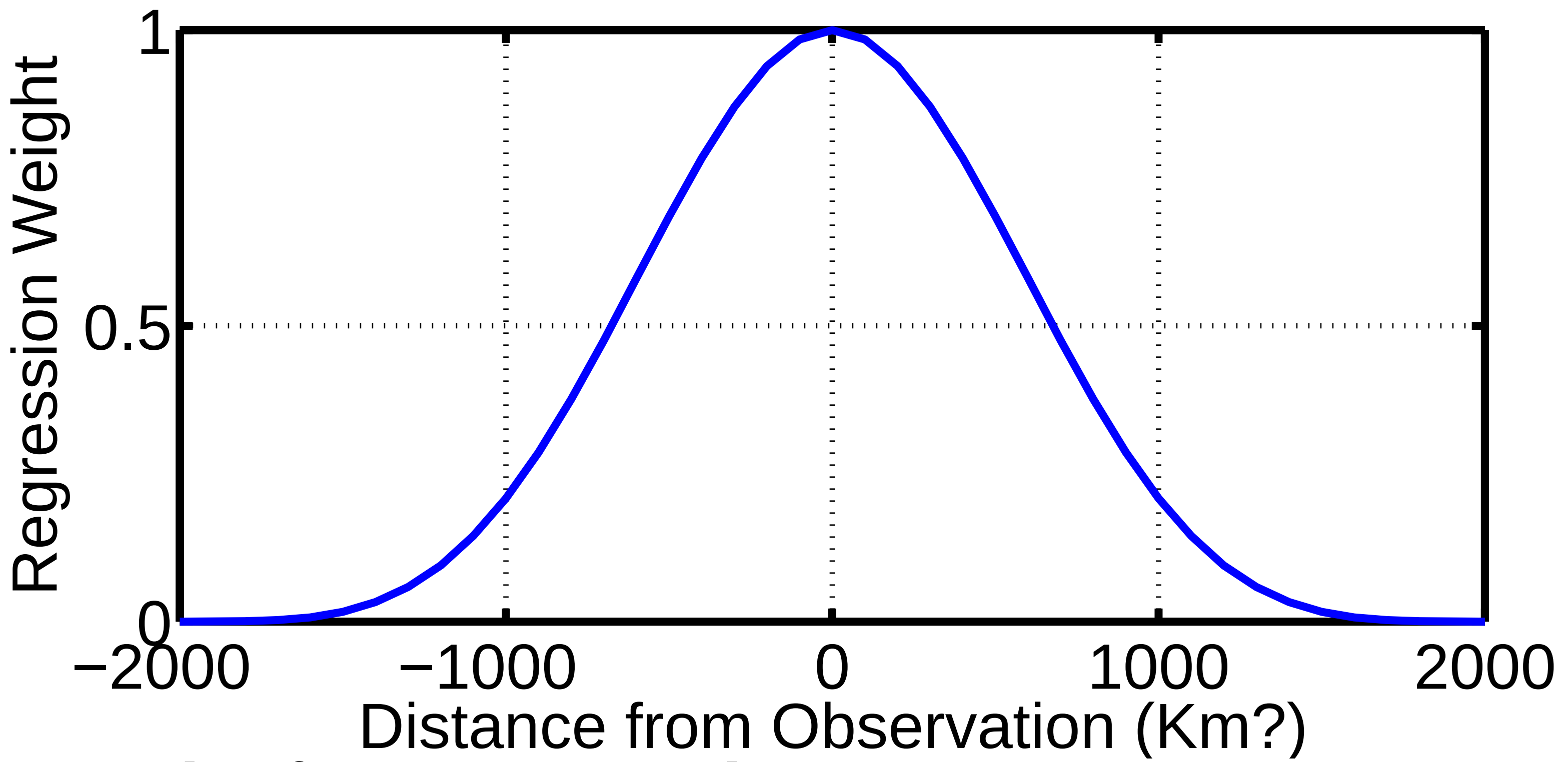
Atmospheric assimilation problems.

Weight regression as function of horizontal *distance* from observation.

Gaspari-Cohn: 5th order compactly supported polynomial.

Ways to deal with regression sampling error:

3. Use additional a priori information about relation between observations and state variables.



Can use other functions to weight regression.

Unclear what *distance* means for some obs./state variable pairs.

Referred to as LOCALIZATION.

Phase 3: Generalize to geophysical models and observations

Dynamical system governed by (stochastic) Difference Equation:

$$dx_t = f(x_t, t) + G(x_t, t)d\beta_t, \quad t \geq 0 \qquad (1)$$

Observations at discrete times:

$$y_k = h(x_k, t_k) + v_k; \; k = 1, 2, \ldots; \qquad t_{k+1} > t_k \geq t_0 \qquad (2)$$

Observational error white in time and Gaussian (nice, not essential).

$$v_k \rightarrow N(0, R_k) \qquad (3)$$

Complete history of observations is:

$$Y_\tau = \{y_l; \; t_l \leq \tau\} \qquad (4)$$

Goal: Find probability distribution for state at time t:

$$p(x, t | Y_t) \qquad (5)$$

Phase 3: Generalize to geophysical models and observations

State between observation times obtained from Difference Equation.
Need to update state given new observation:

$$p(x, t_k | Y_{t_k}) = p(x, t_k | y_k, Y_{t_{k-1}}) \qquad (6)$$

Apply Bayes rule:

$$p(x, t_k | Y_{t_k}) = \frac{p(y_k | x_k, Y_{t_{k-1}}) p(x, t_k | Y_{t_{k-1}})}{p(y_k | Y_{t_{k-1}})} \qquad (7)$$

Noise is white in time (3) so:

$$p(y_k | x_k, Y_{t_{k-1}}) = p(y_k | x_k) \qquad (8)$$

Integrate numerator to get normalizing denominator:

$$p\left(y_k | Y_{t_{k-1}}\right) = \int p(y_k | x) p(x, t_k | Y_{t_{k-1}}) dx \qquad (9)$$

Phase 3: Generalize to geophysical models and observations

Probability after new observation:

$$p\left(x, t_k \middle| Y_{t_k}\right) = \frac{p(y_k|x)\,p(x, t_k|Y_{t_{k-1}})}{\int p(y_k|\xi)\,p(\xi, t_k|Y_{t_{k-1}})\,d\xi} \quad (10)$$

Exactly analogous to earlier derivation except that x and y are vectors.

EXCEPT, no guarantee we have prior sample for each observation.

SO, let’s make sure we have priors by ‘extending’ state vector.

Phase 3: Generalize to geophysical models and observations

Extending the state vector to joint state-observation vector.

Recall: $y_k = h(x_k, t_k) + v_k$; $k = 1, 2, \ldots$; $\quad t_{k+1} > t_k \geq t_0 \qquad (2)$

Applying h to x at a given time gives expected values of observations.

Get prior sample of obs. by applying h to each sample of state vector x.

Let z = [x, y] be the combined vector of state and observations.

Phase 3: Generalize to geophysical models and observations

NOW, we have a prior for each observation:

$$p\left(z, t_k \middle| Y_{t_k}\right) = \frac{p(y_k|z)\,p(z, t_k|Y_{t_{k-1}})}{\int p(y_k|\xi)\,p(\xi, t_k|Y_{t_{k-1}})\,d\xi} \qquad (10.\text{ext})$$

Phase 3: Generalize to geophysical models and observations

One more issue: how to deal with many observations in set y_k?

Let y_k be composed of s subsets of observations: $y_k = \{y_k^1, y_k^2, \ldots, y_k^s\}$

Observational errors for obs. in set i independent of those in set j.

Then: $p(y_k|z) = \prod_{i=1}^{s} p(y_k^i|z)$

Can rewrite (10.ext) as series of products and normalizations.

Phase 3: Generalize to geophysical models and observations

One more issue: how to deal with many observations in set y_k?

Implication: can assimilate observation subsets sequentially.

If subsets are scalar (individual obs. have mutually independent error distributions), can assimilate each observation sequentially.

If not, have two options:

1. Repeat everything above with matrix algebra.

2. Do singular value decomposition; diagonalize obs. error covariance.
   Assimilate observations sequentially in rotated space.
   Rotate result back to original space.

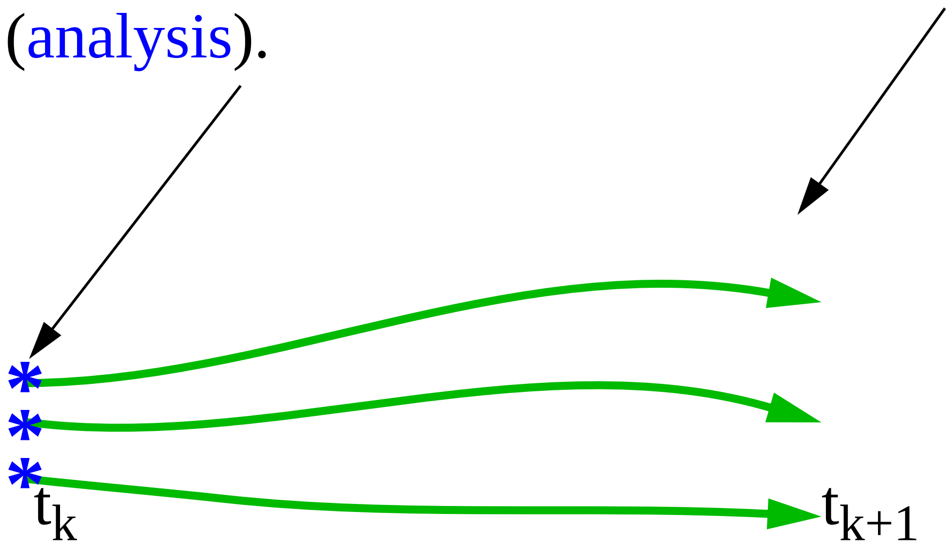
Good news: Most geophysical obs. have independent errors!

How an Ensemble Filter Works for Geophysical Data Assimilation

1. Use model to advance ensemble (3 members here) to time at which next observation becomes available.

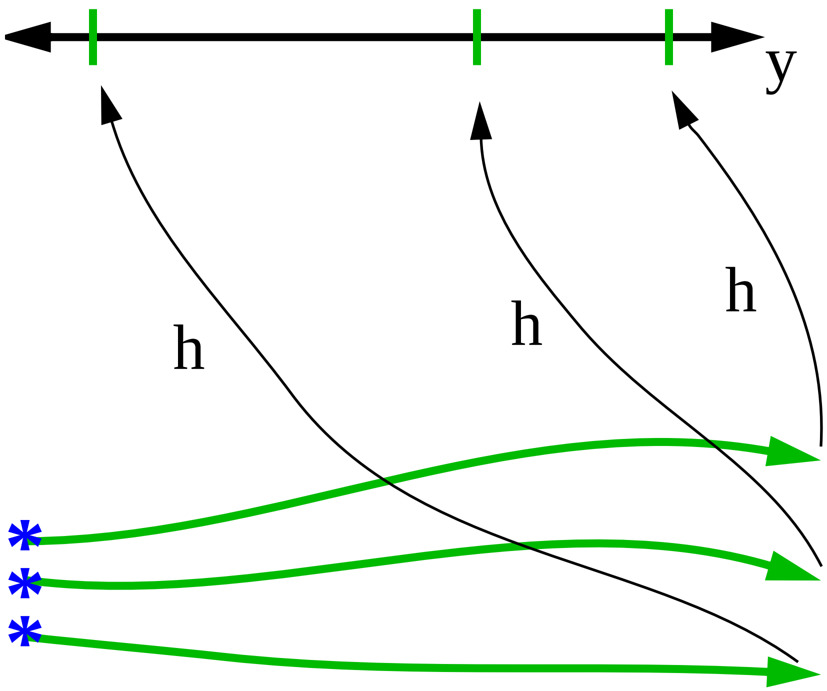
Ensemble state estimate after using previous observation (analysis).

Ensemble state at time of next observation (prior).



How an Ensemble Filter Works for Geophysical Data Assimilation

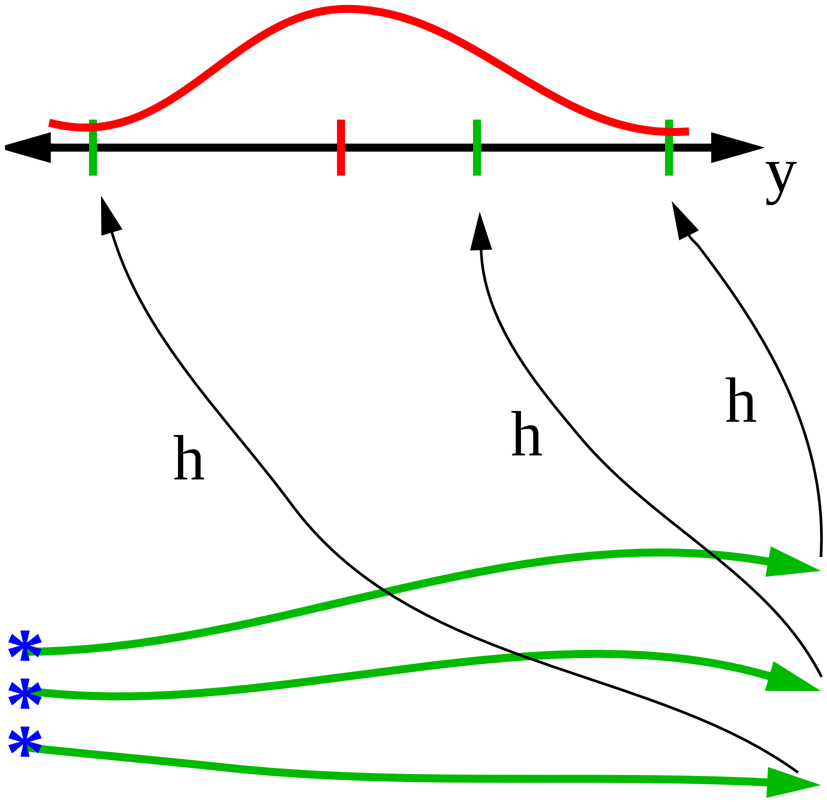
2. Get prior ensemble sample of observation, y=h(x), by applying forward operator h to each ensemble member.



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

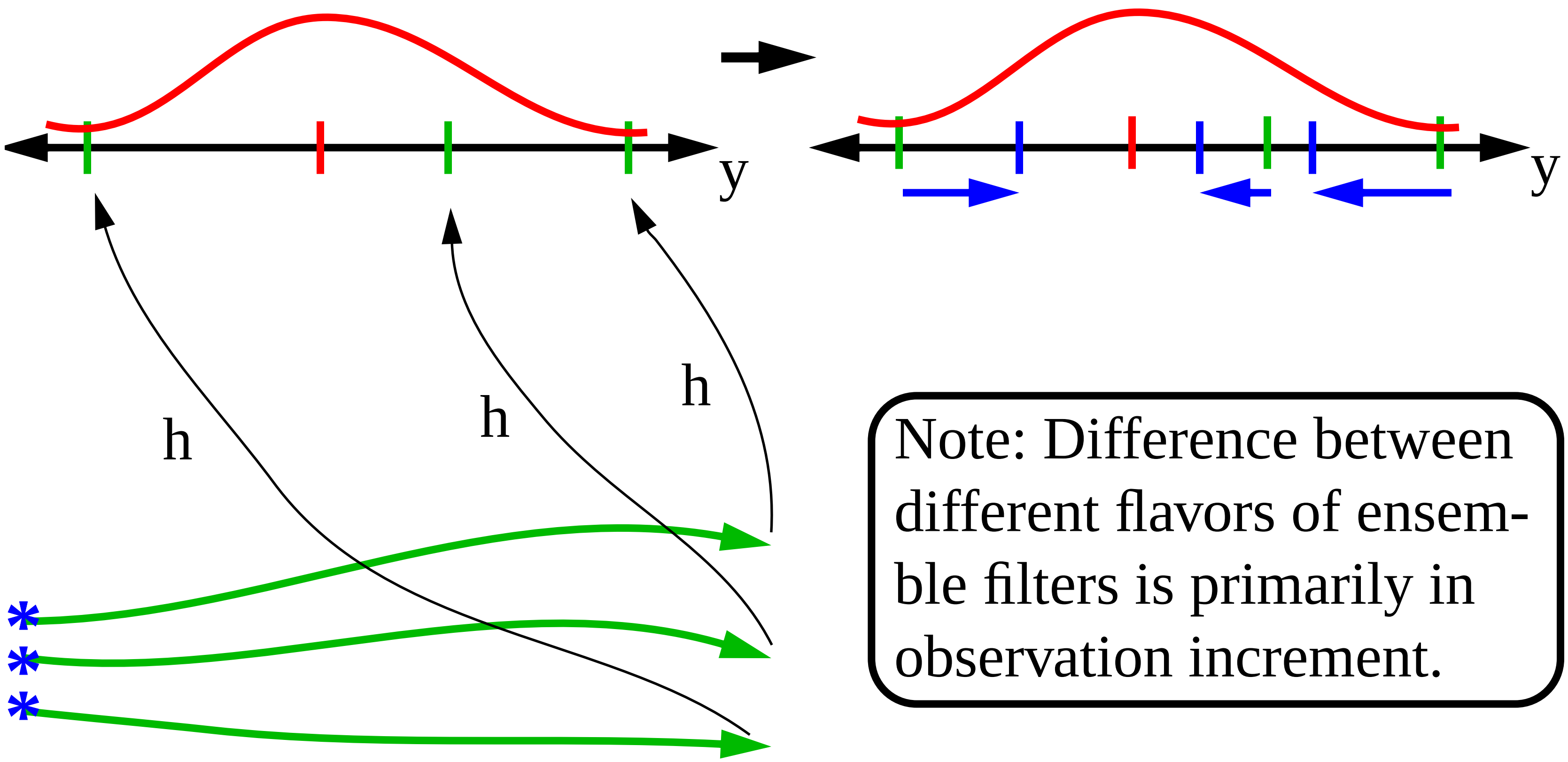
How an Ensemble Filter Works for Geophysical Data Assimilation

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



How an Ensemble Filter Works for Geophysical Data Assimilation

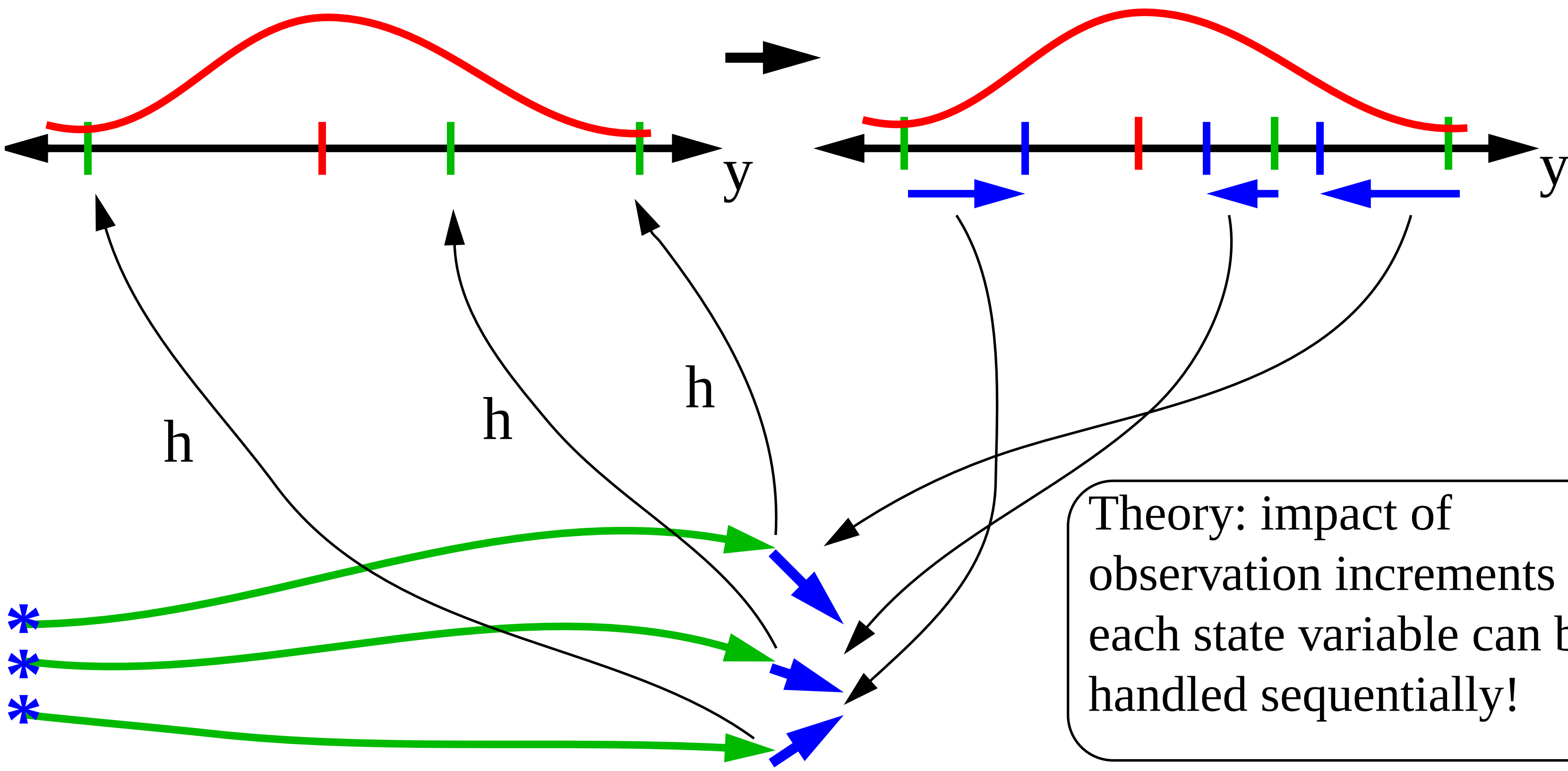
4. Find increment for each prior observation ensemble
(this is a scalar problem for uncorrelated observation errors).



Note: Difference between different flavors of ensemble filters is primarily in observation increment.

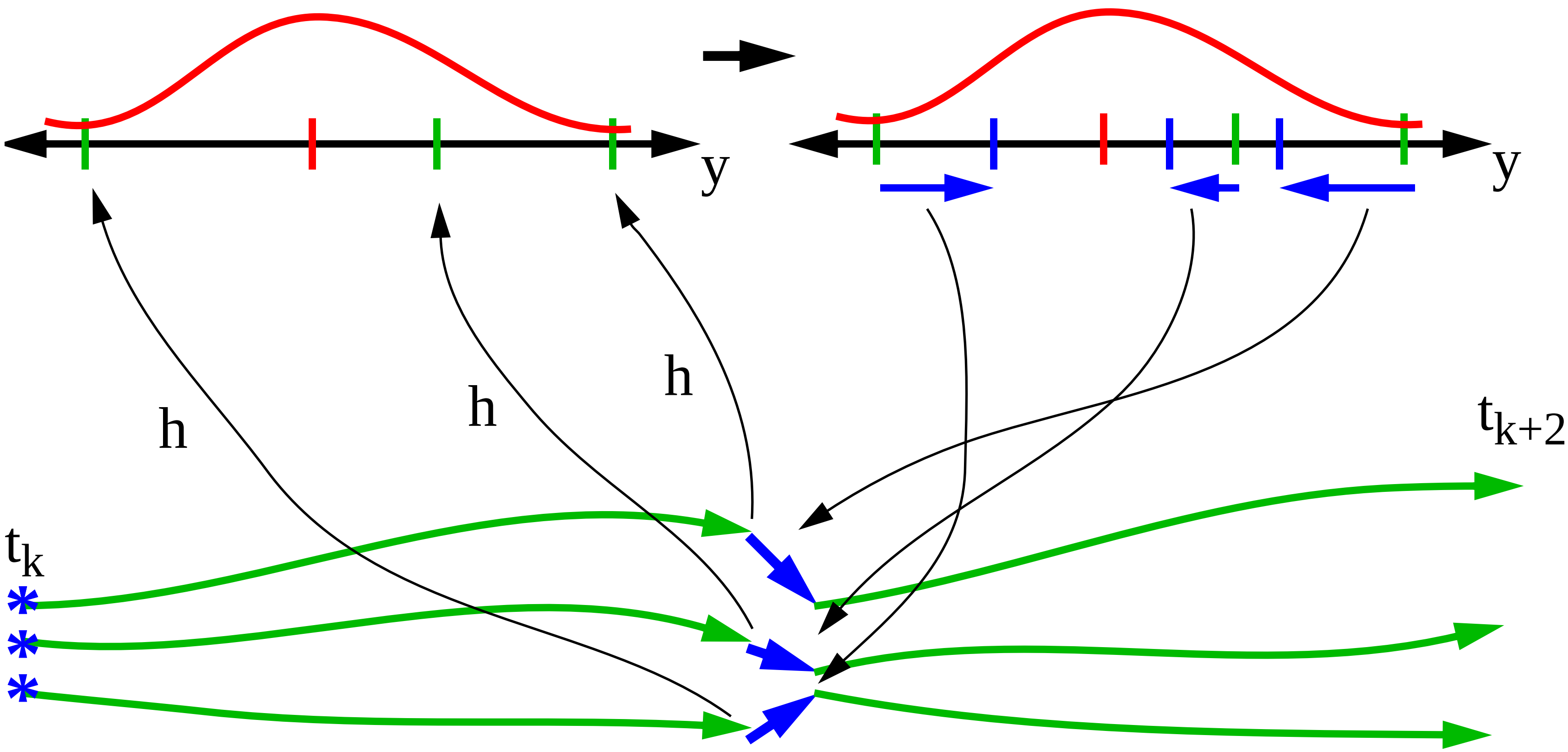
How an Ensemble Filter Works for Geophysical Data Assimilation

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



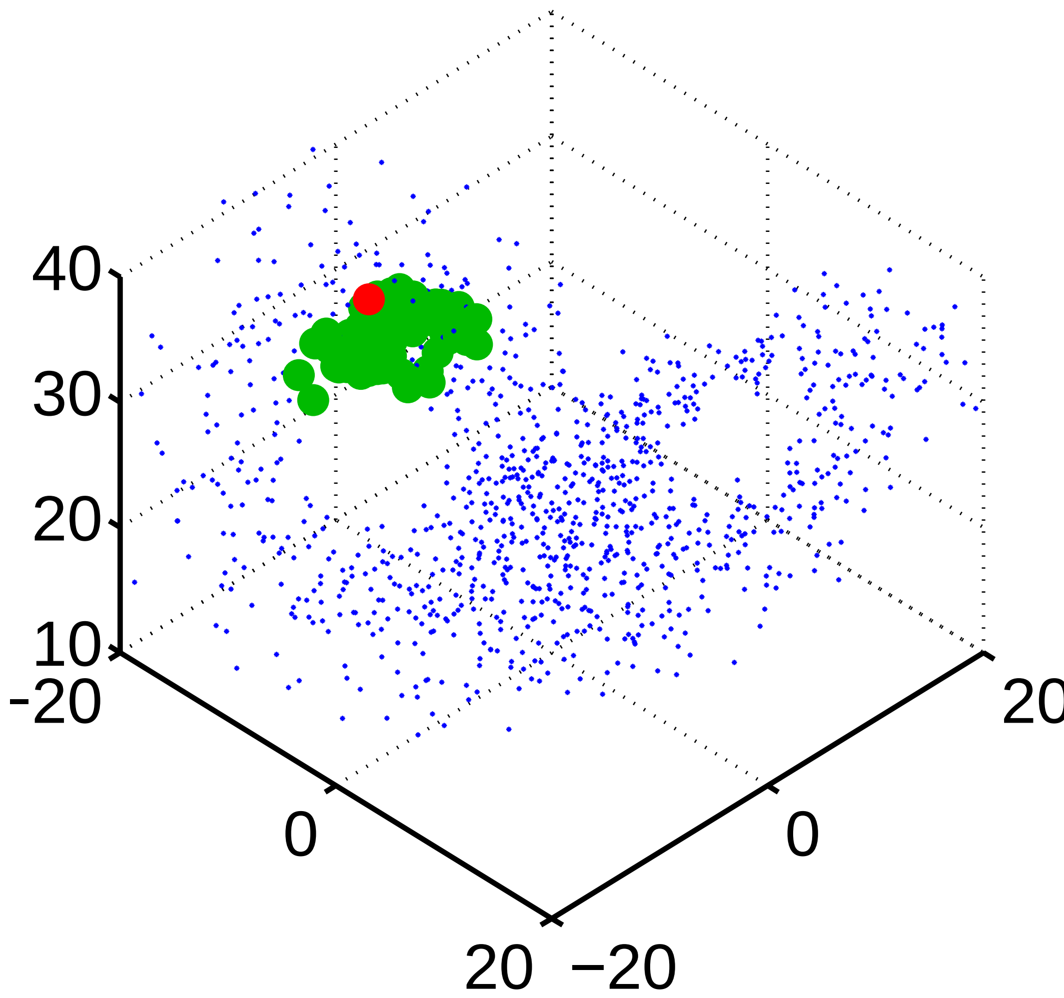
Theory: impact of observation increments on each state variable can be handled sequentially!

6. When all ensemble members for each state variable are updated, have a new analysis. Integrate to time of next observation...



Phase 3: Generalize to geophysical models and observations

Simple example: Lorenz-63 3-variable chaotic model.



Observation in red.

Prior ensemble in green.

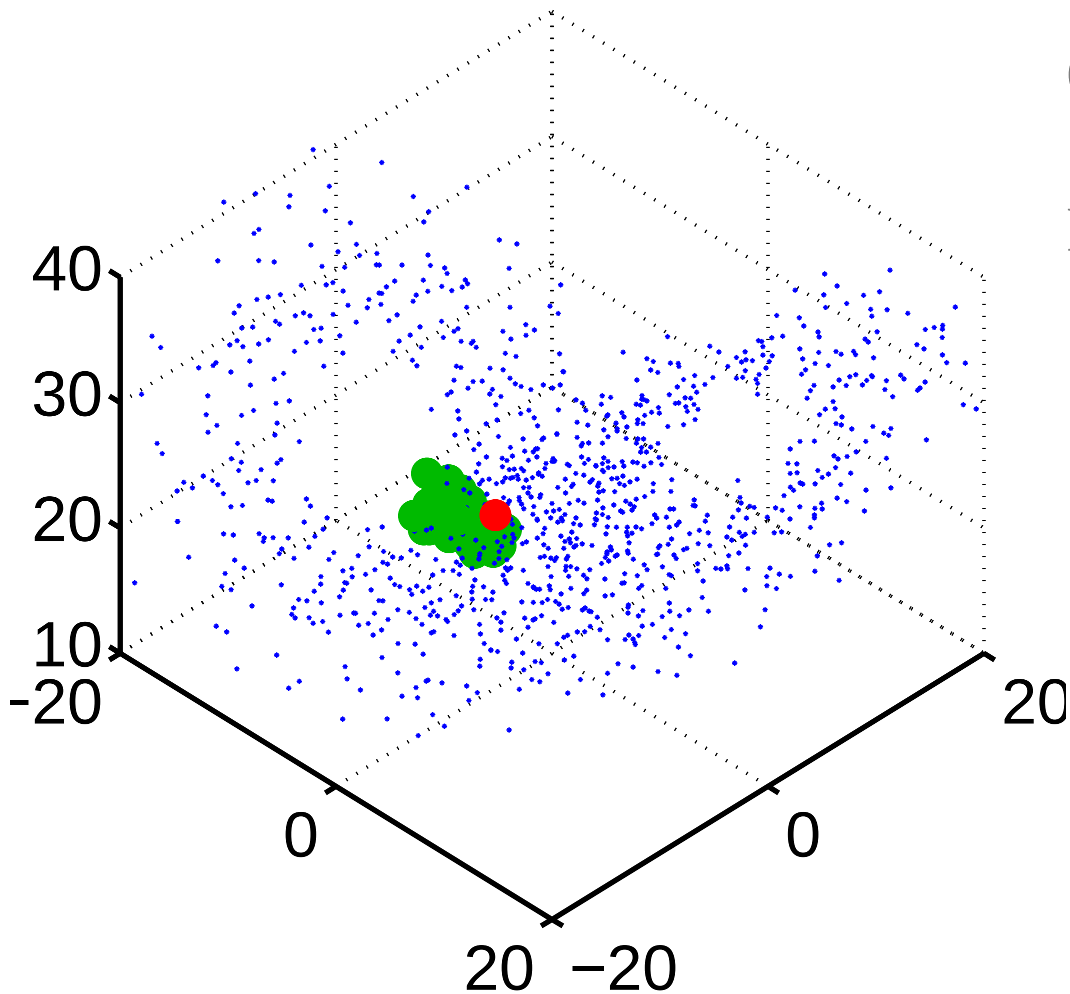
Observing all three state variables.

Obs. error variance = 4.0.

4 20-member ensembles.

Phase 3: Generalize to geophysical models and observations

Simple example: Lorenz-63 3-variable chaotic model.

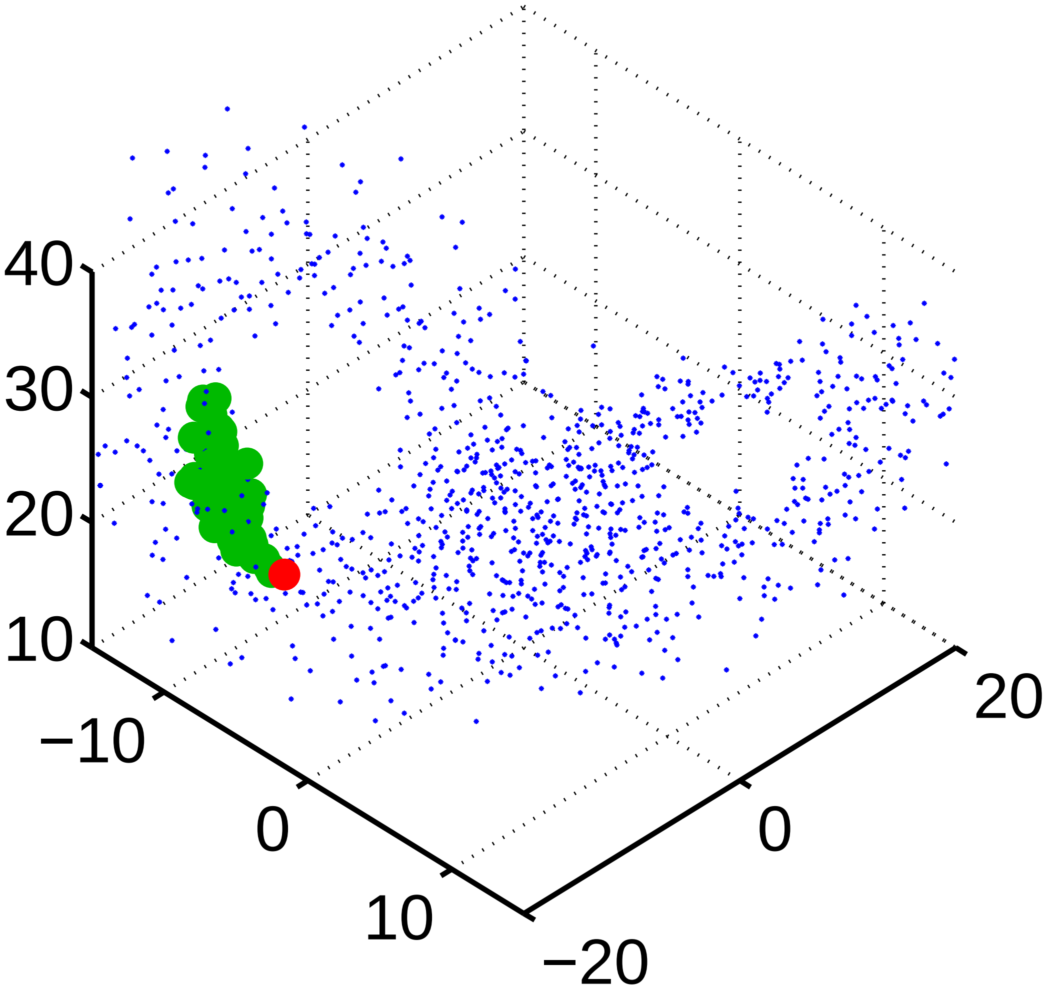


Observation in red.

Prior ensemble in green.

Phase 3: Generalize to geophysical models and observations

Simple example: Lorenz-63 3-variable chaotic model.

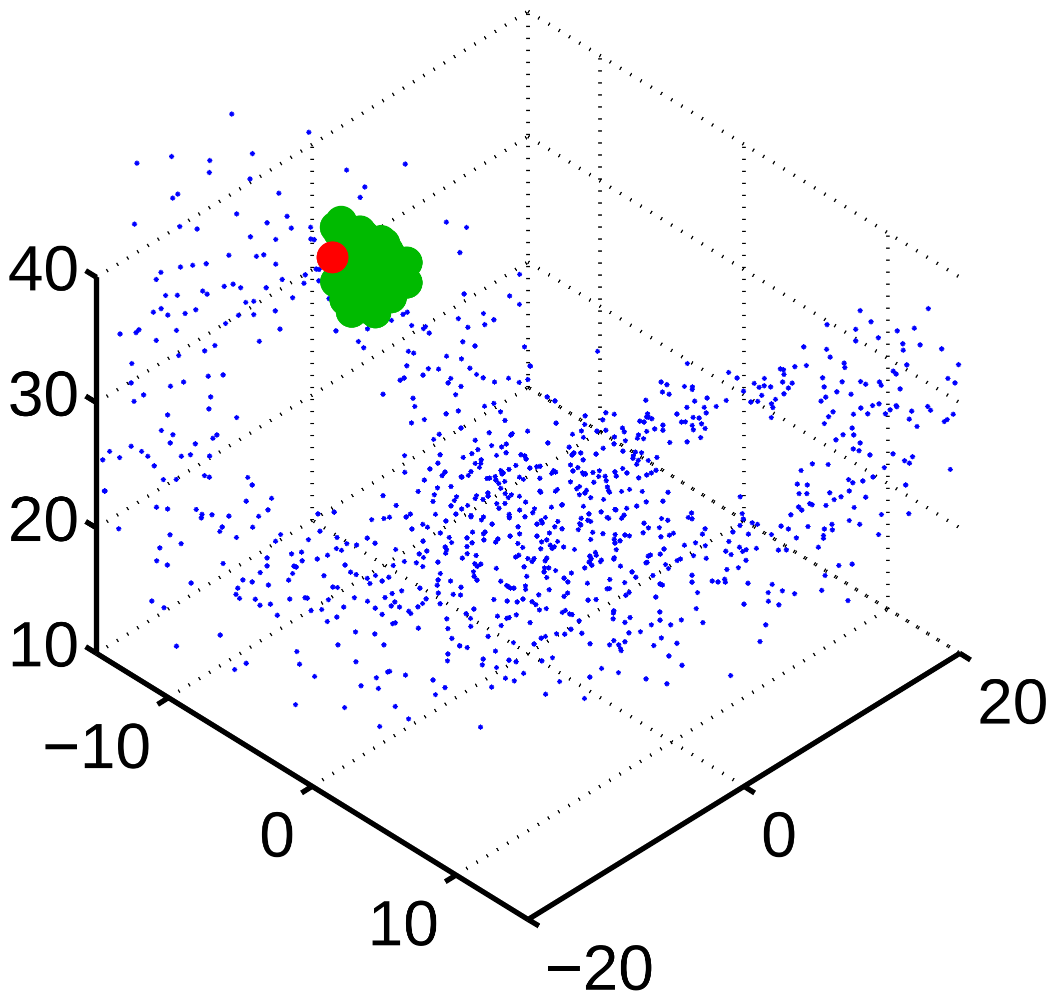


Observation in red.

Prior ensemble in green.

Phase 3: Generalize to geophysical models and observations

Simple example: Lorenz-63 3-variable chaotic model.

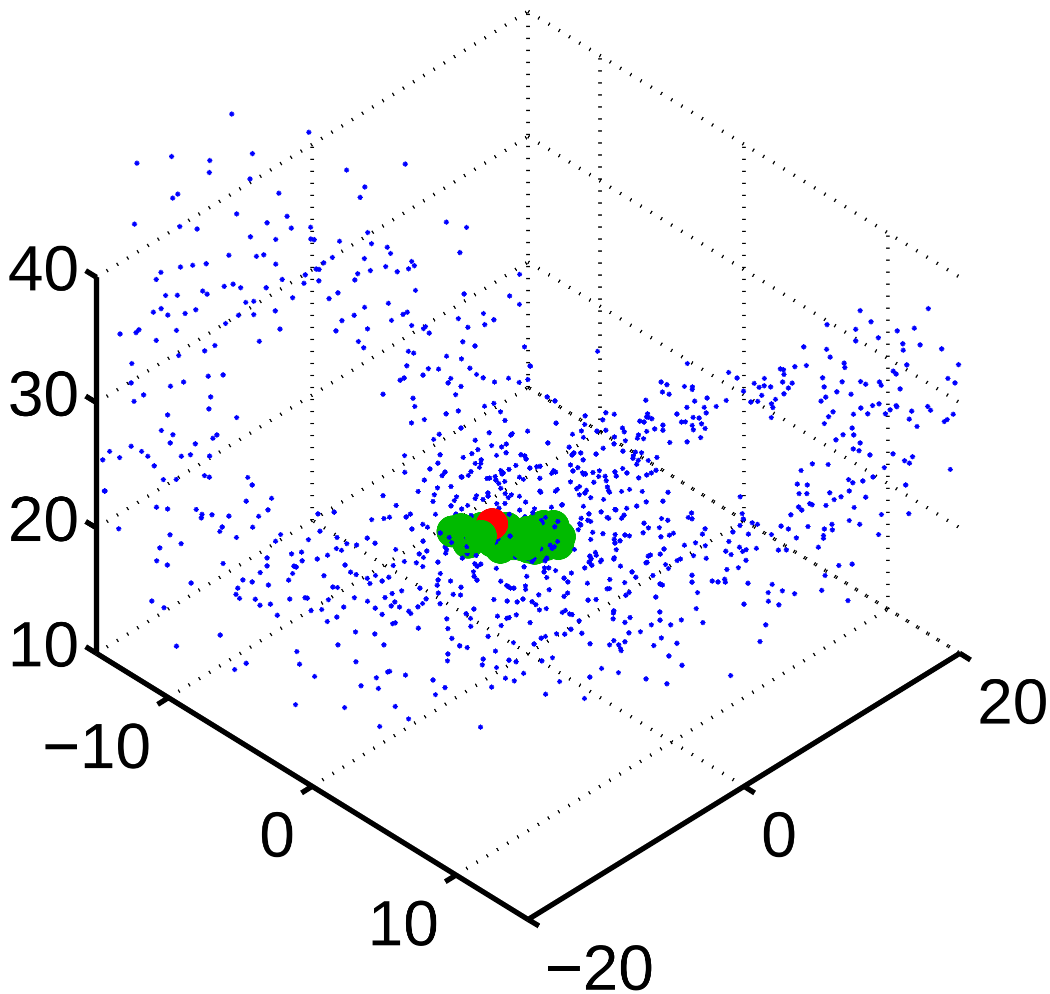


Observation in red.

Prior ensemble in green.

Phase 3: Generalize to geophysical models and observations

Simple example: Lorenz-63 3-variable chaotic model.



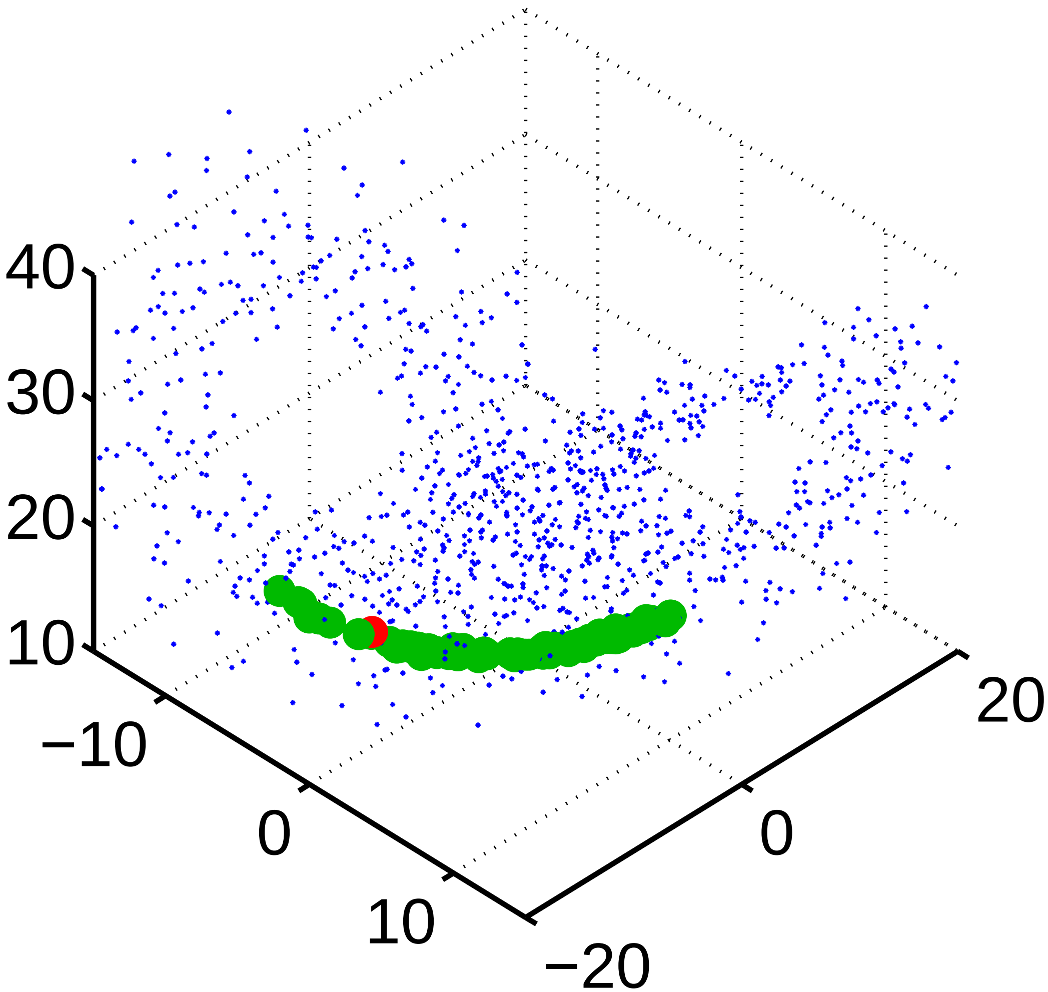
Observation in red.

Prior ensemble in green.

Ensemble is passing through unpredictable region.

Phase 3: Generalize to geophysical models and observations

Simple example: Lorenz-63 3-variable chaotic model.



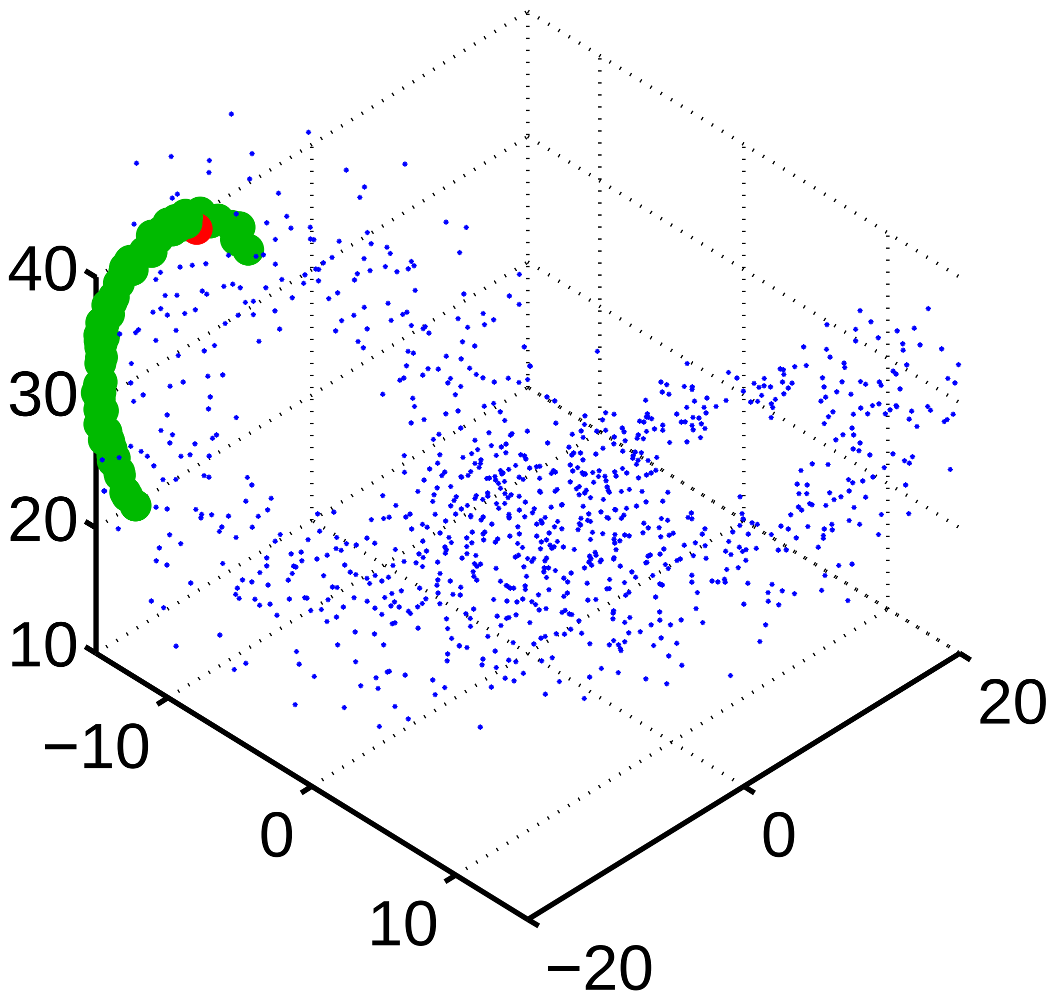
Observation in red.

Prior ensemble in green.

Part of ensemble heads for one lobe, the rest for the other.

Phase 3: Generalize to geophysical models and observations

Simple example: Lorenz-63 3-variable chaotic model.



Observation in red.

Prior ensemble in green.

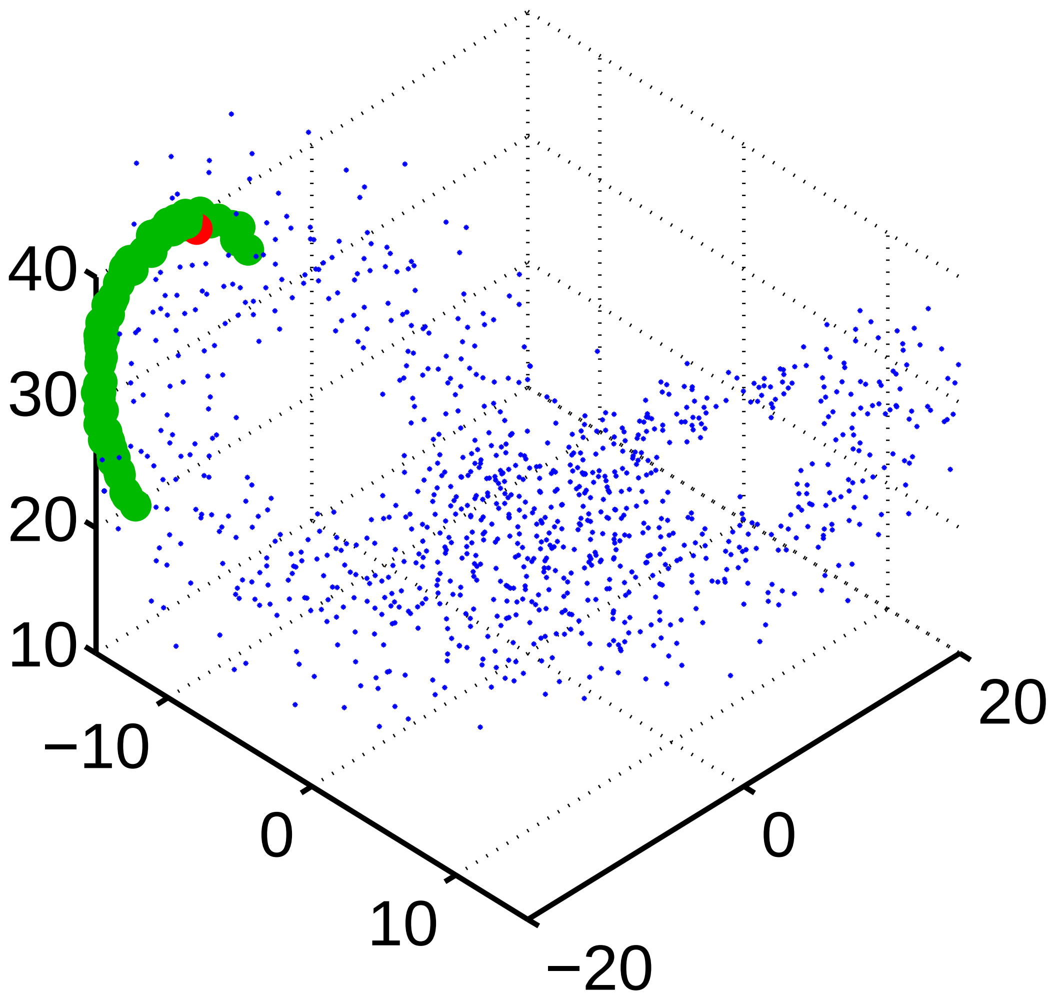
The prior is not linear here.

Standard regression might be pretty bad.

Covariance inflation might also be bad, pushing ensemble off the attractor.

Phase 3: Generalize to geophysical models and observations

Simple example: Lorenz-63 3-variable chaotic model.



Observation in red.

Prior ensemble in green.

The prior is not linear here.

On the other hand...

Hard to contrive examples this bad.

Behavior like this not apparent in real assimilations.

Phase 4: Quick look at real atmospheric applications...

Results from CAM Assimilation: January, 2003

Model:

CAM 3.1 T85L26

U,V, T, Q and PS state variables impacted by observations.

Land model (CLM 2.0) not impacted by observations.

Climatological SSTs.

Assimilation / Prediction Experiments:

80 member ensemble divided into 4 equal groups.

Adaptive error correction algorithm.

Initialized from a climatological distribution (huge spread).

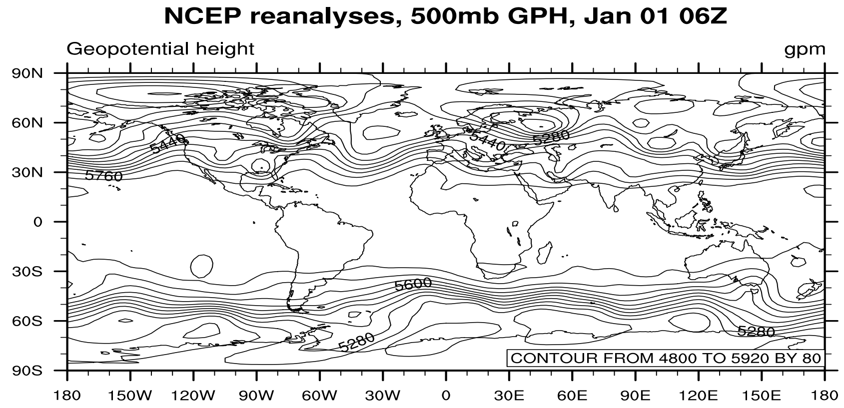
Uses most observations used in reanalysis

(Radiosondes, ACARS, Sat. Winds..., no surface obs. or retrievals).

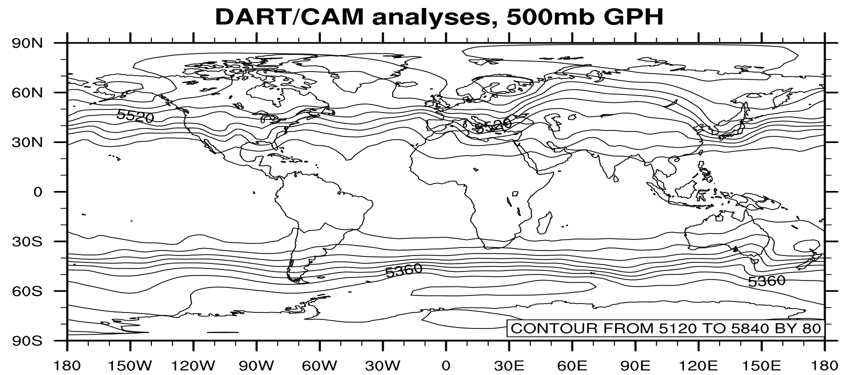
Assimilated every 6 hours; +/- 1.5 hour window for obs.

After 6 hours.

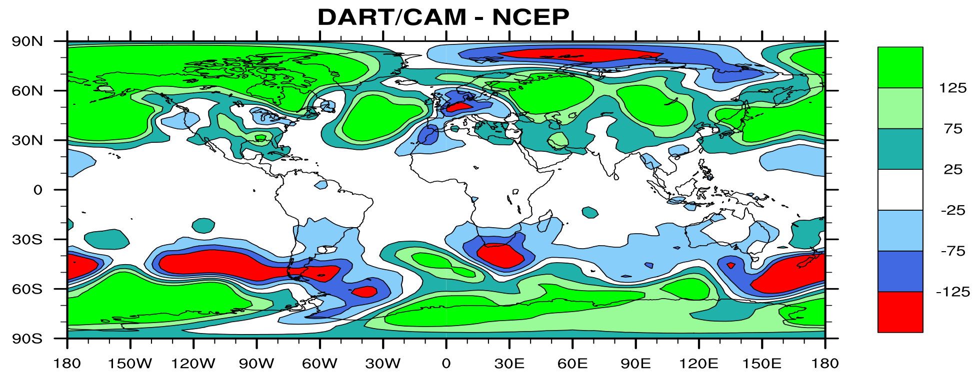
NCEP



DART/CAM



Difference.

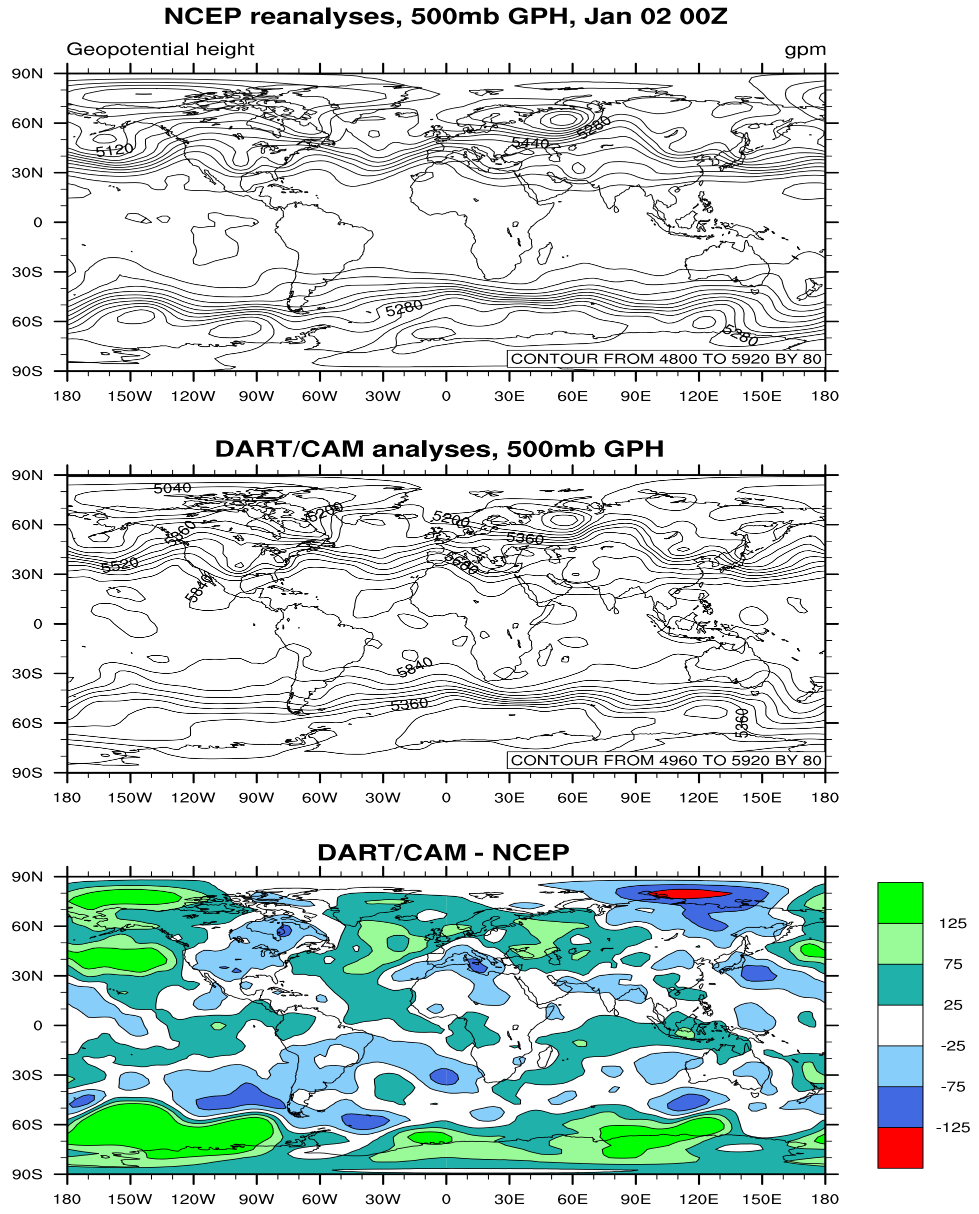


After 1 day.

NCEP

DART/CAM

Difference.

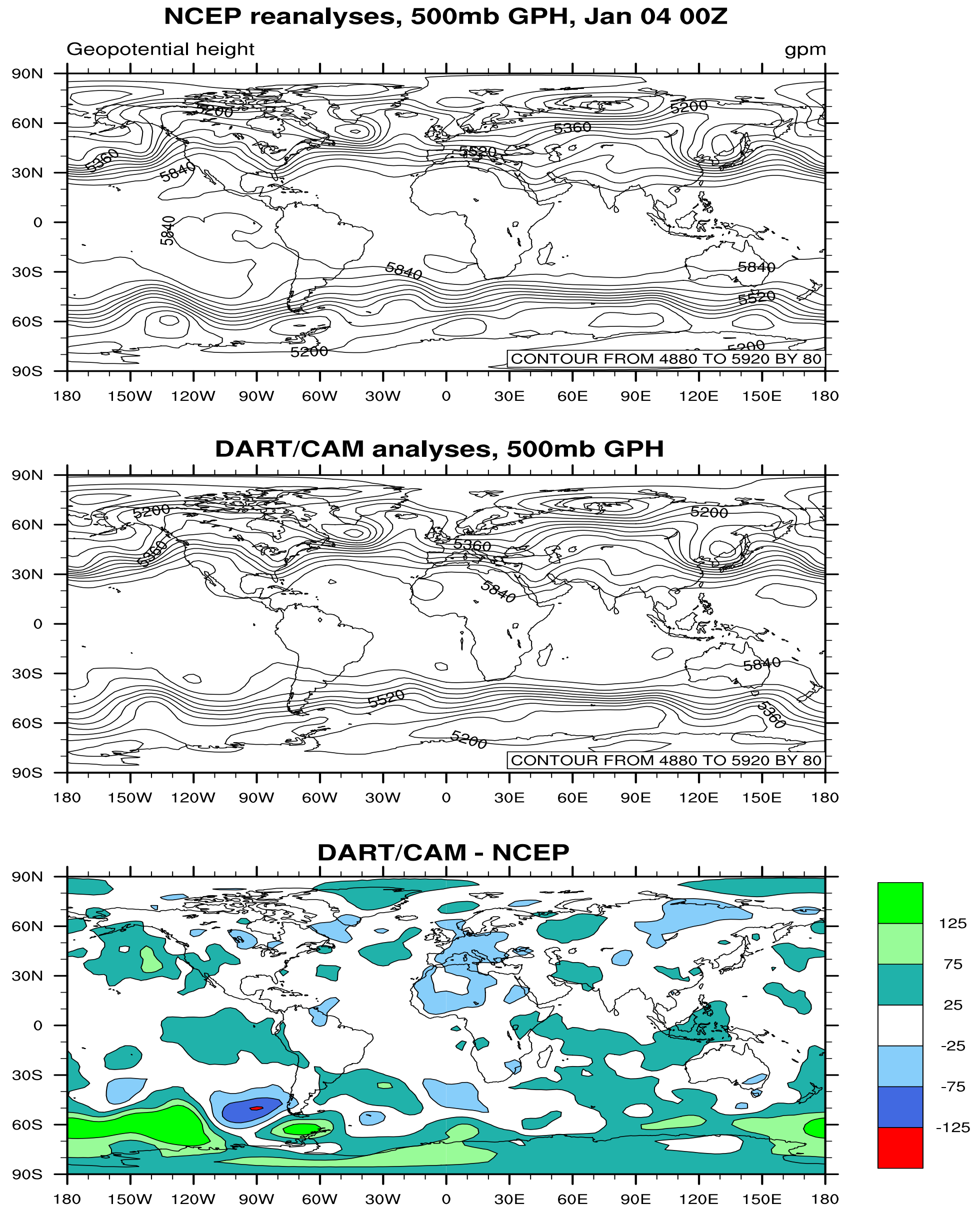


After 3 days.

NCEP

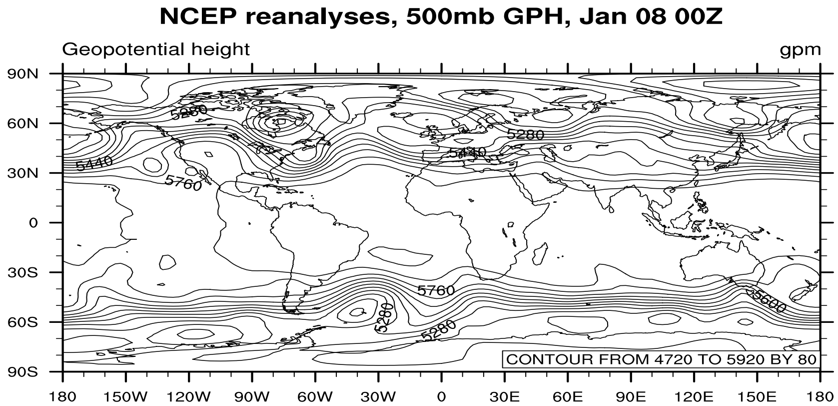
DART/CAM

Difference.

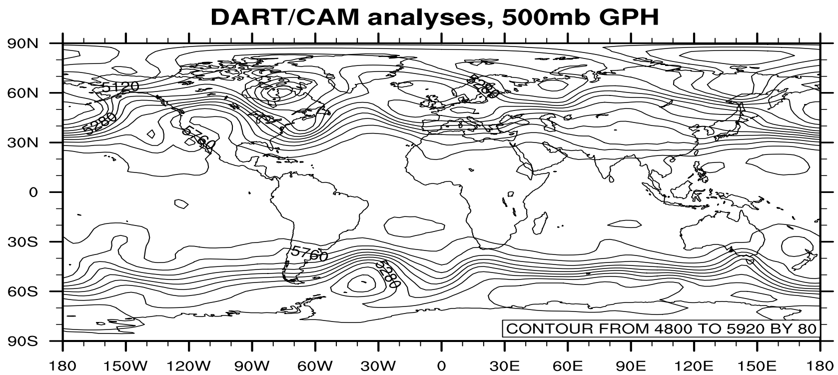


After 7 days.

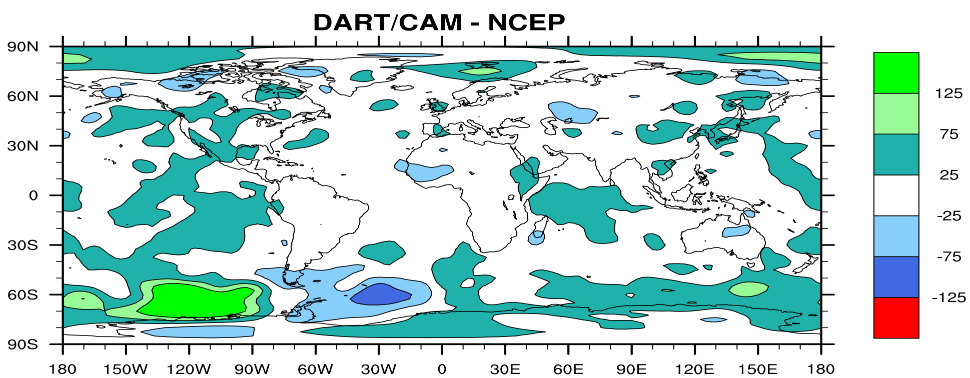
NCEP



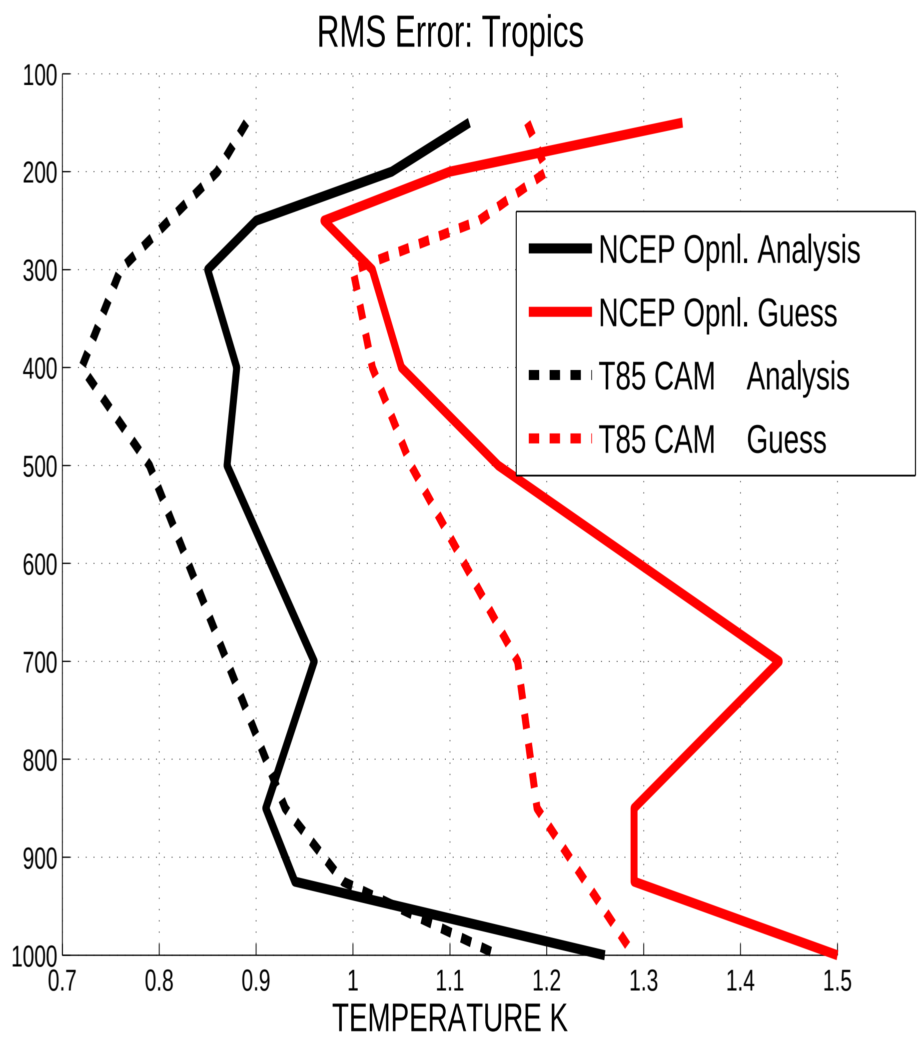
DART/CAM



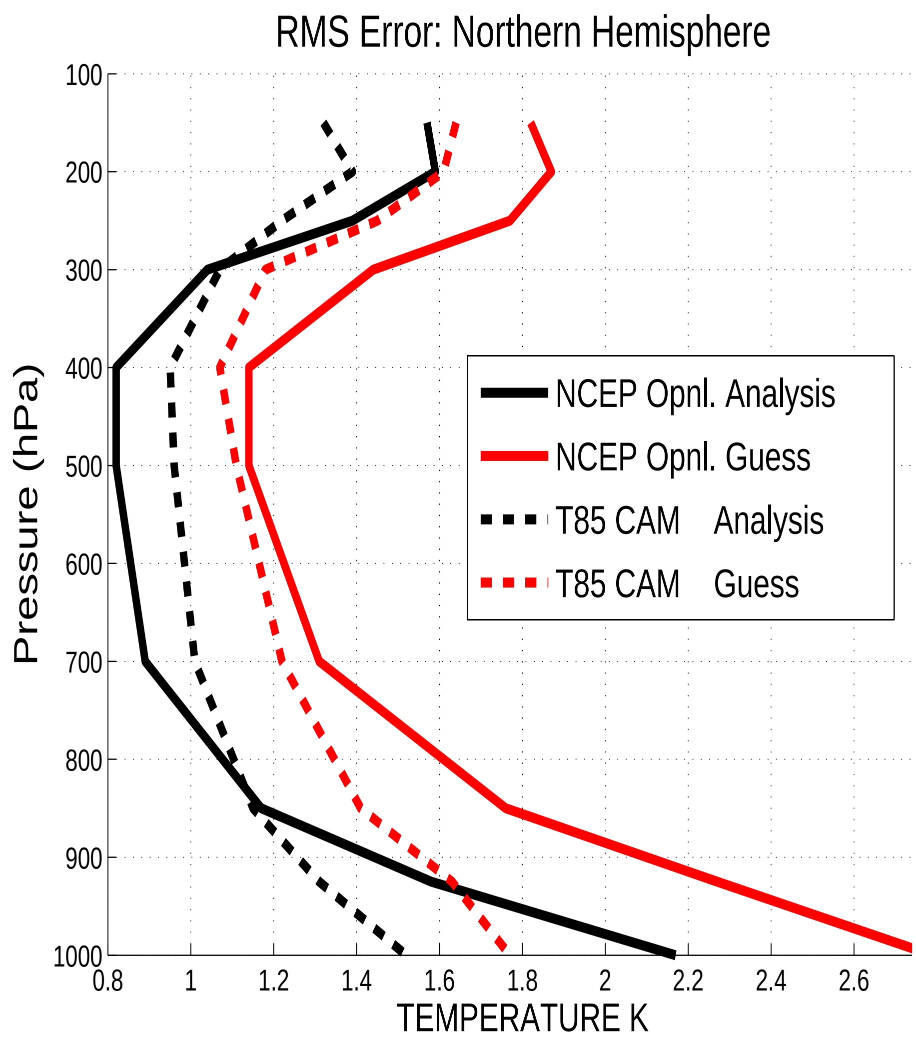
Difference.



6-Hour Forecast and Analysis Observation Space Temperature RMS

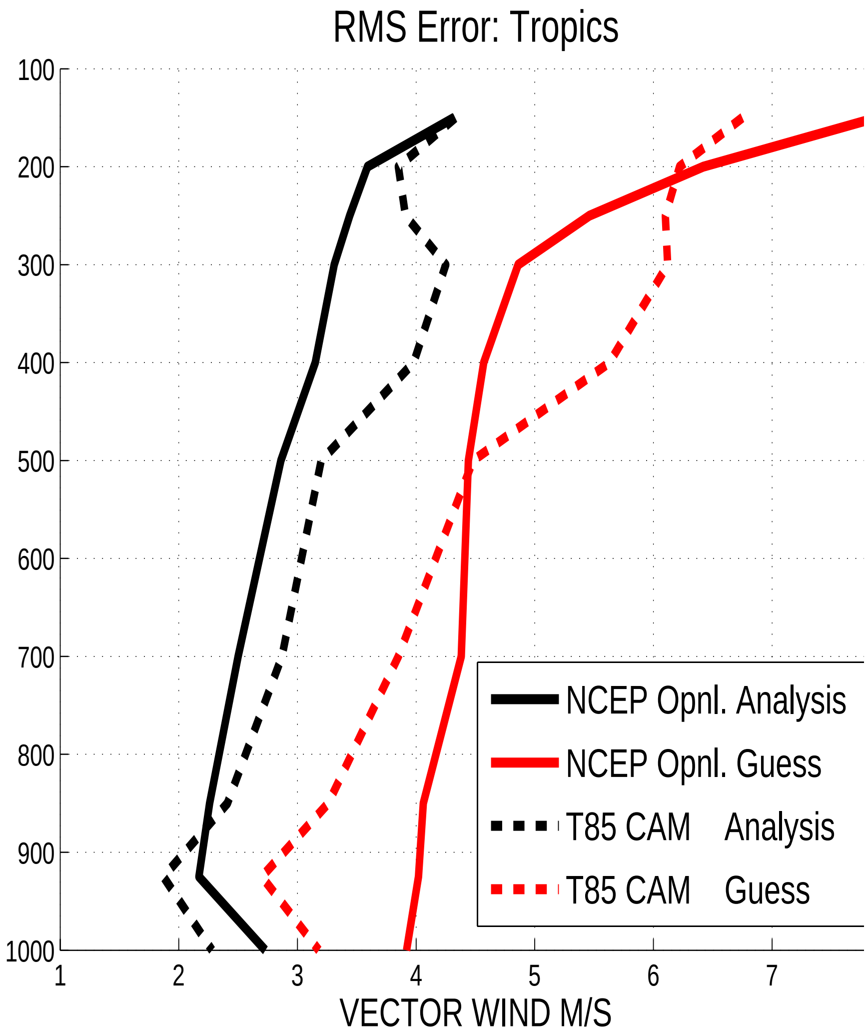


Tropics

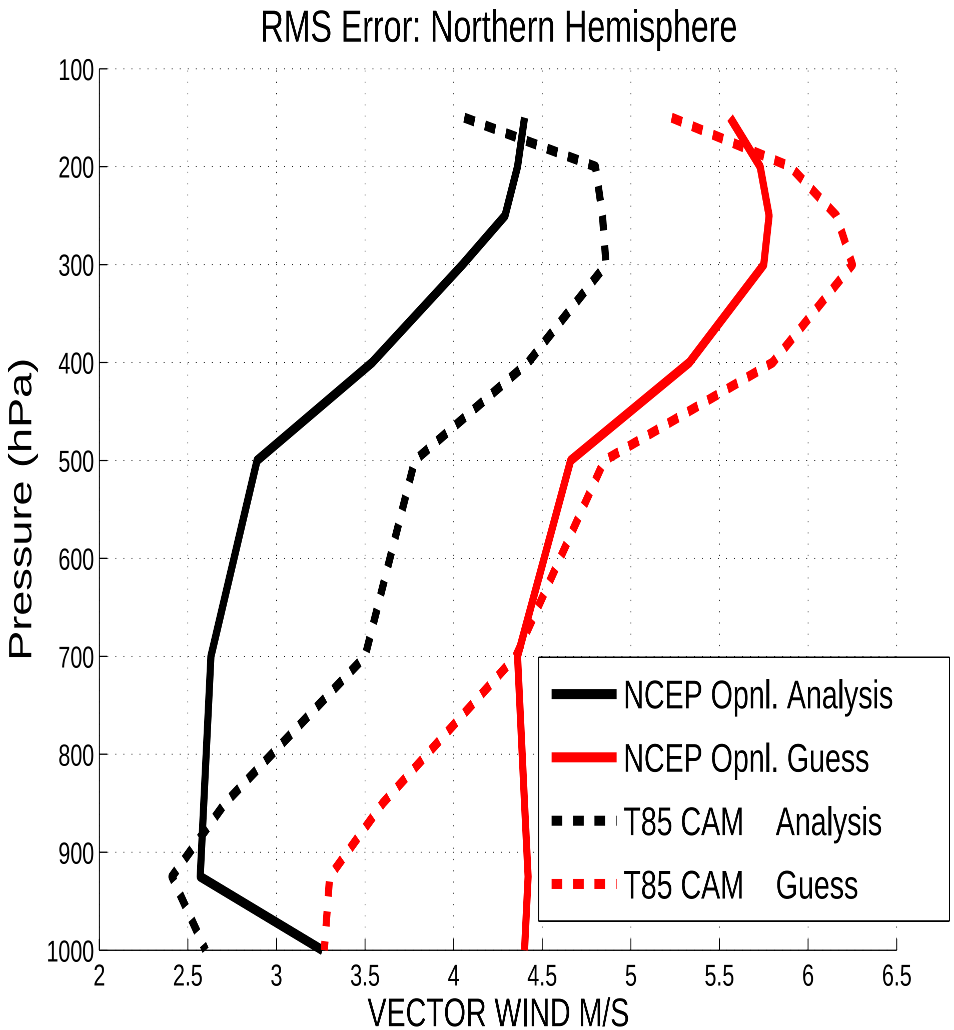


Northern Hemisphere

6-Hour Forecast and Analysis Observation Space Wind RMS



Tropics



Northern Hemisphere

Finally..., Parameter Estimation.

Suppose model is governed by (stochastic) Difference Equation:

$$dx_t = f(x_t, t; u) + G(x_t, t; w)d\beta_t, \quad t \geq 0 \qquad (1)$$

where u and w are vectors of parameters.

Also, suppose we really don't know values of parameters (very well).

Can use observations with assimilation to help constrain these values.

Rewrite (1) as:

$$dx_t^A = f^A(x_t^A, t) + G^A(x_t^A, t)d\beta_t, \quad t \geq 0 \qquad (2)$$

where the augmented state vector includes x_t, u, and w.

Model is modified so values of u and w can be changed by assimilation.

From ensemble filter perspective:

Just add any parameters of interest to the model state vector;
Proceed to assimilate as before.

Possible difficulties:

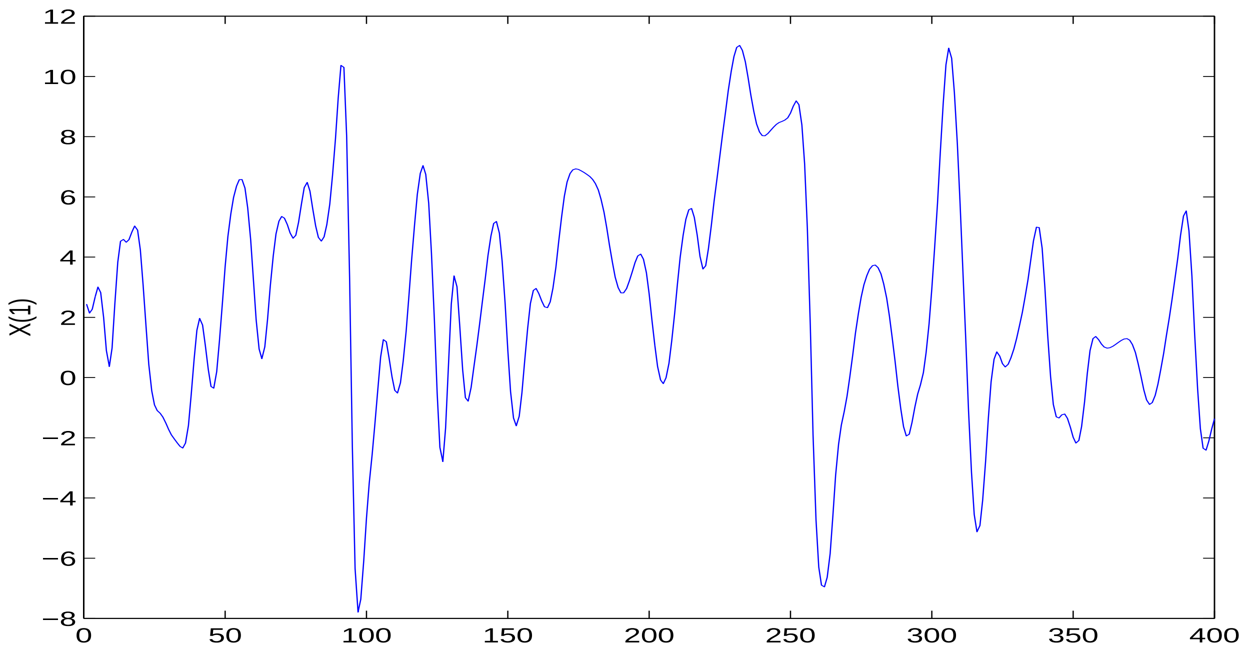
1. Where are parameters ‘located’ for localization?
2. Parameters won’t have any error growth in time
   (unless we add some): could lead to filter divergence.
3. Parameters may not be strongly correlated with any observations.

Testing Parameter Estimation in Lorenz-96 40-variable Model

Variable size low-order dynamical system

N variables, X_1, X_2,..., X_N

Use N = 40, F = 8.0, 4th-order Runge-Kutta with dt=0.0



Time series of x_1

Testing Parameter Estimation in Lorenz-96 40-variable Model

$dX_i / dt = (X_{i+1} - X_{i-2})X_{i-1} - X_i + F$ $\qquad$ $i=1, ...40$ $\qquad$ (1)

(cyclic indices).

Have observations of functions of state variables.
Generated by model with fixed but unknown value of F.

Recast F as a state variable.
Single additional variable, or 40 as follows:

$$dX_i / dt = (X_{i+1} - X_{i-2})X_{i-1} - X_i + F_i \qquad (2)$$

$$dF_i/dt = N(0, \sigma_{noise}) \qquad (3)$$

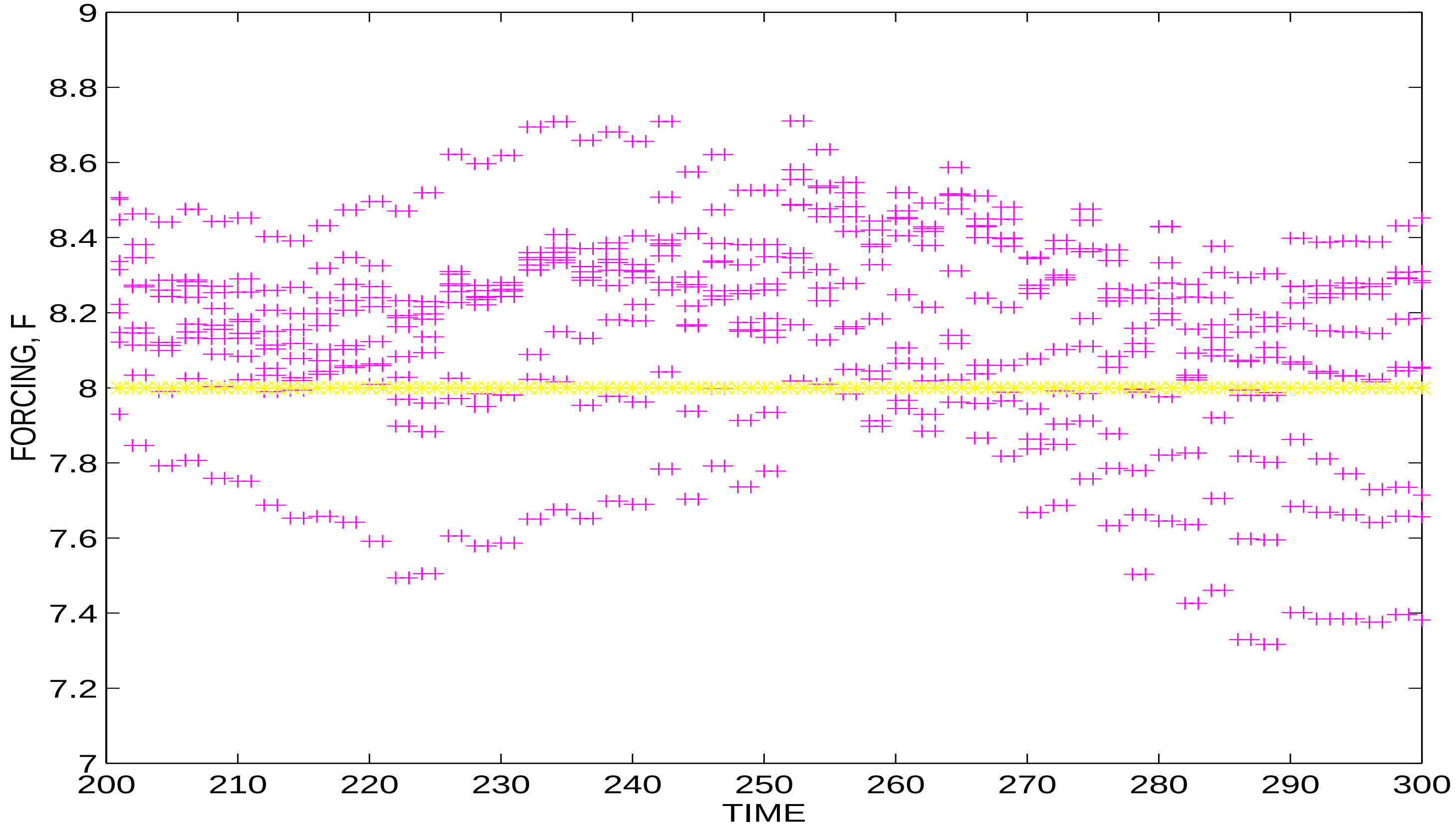
Can observations of some function of state variables constrain F?

Testing Parameter Estimation in Lorenz-96 40-variable Model

20 Member Ensemble (10 Plotted).

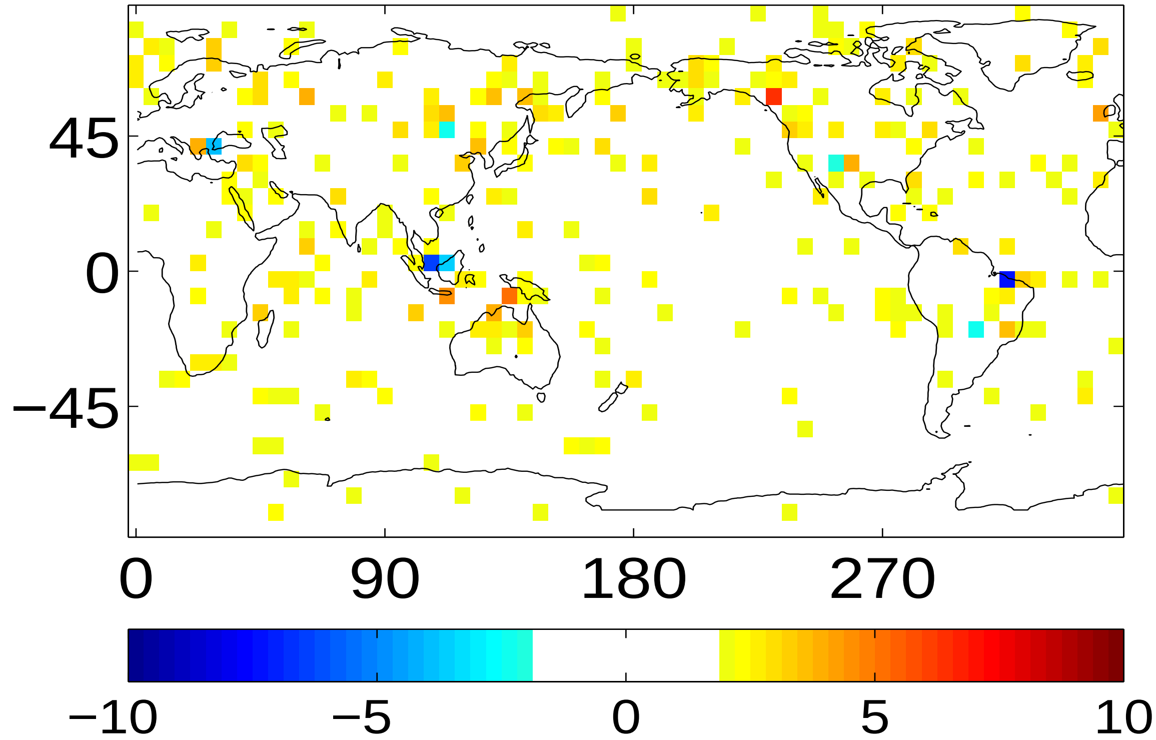
40 Randomly Located Observations (interpolated state) every step.

Truth in Yellow (8.0) Ensemble



Intriguing Fact: Best assimilations of state come when F_i varies, even better than when F_i is set to exact value, 8!

Climate Model Parameter Estimation via Ensemble Data Assimilation



T21 CAM assimilation of gravity wave drag efficiency parameter.

Oceanic values are noise (should be 0).

0< efficiency< ~4 suggested by modelers.

Positive values over NH land expected.
Problem: large negative values over SH land near convection.
Reduces model bias, but for ‘Wrong Reason’.

Assimilation tries to use free parameter to fix ALL model problems

Ensemble Assimilation / Parameter Estimation Issues

1. Dealing with sampling noise.
   A. How to localize.
   B. What are expected correlations of obs. with parameters?
   C. If these are small, can things be ‘rotated’ to get signal?

2. What question is being answered?
   A. Filters try to minimize RMS error.
   B. Hard to specify complex prior constraints on parameters.
   C. Model designers may be asking a different question.

3. Observability.
   A. How much can one get from available obs?
   B. Is it hard to estimate many parameters at once?