Data Assimilation
- P01. Local Particle Filter-

Shunji Kotsuki

Center for Environmental Remote Sensing / Institute of Advanced Academic Research

(shunji.kotsuki@chiba-u.jp)

CHIBA UNIVERSITY

Environmental Prediction Science Laboratory

Background Knowledge

Environmental
Prediction
Science
Laboratory

Local Ensemble Data Assimilations

	EnKF	PF	Hybrids of EnKF & LPF
Serial DA	Stochastic EnKF (PO) EAKF Serial EnSRF	Poterjoy et al.	Poterjoy et al.
Simultaneous DA w/ transform matrix	LETKF	Reich (2013) Penny and Miyoshi (2016)	LPFGM

This study aims at improving the PF with the transform matrix.

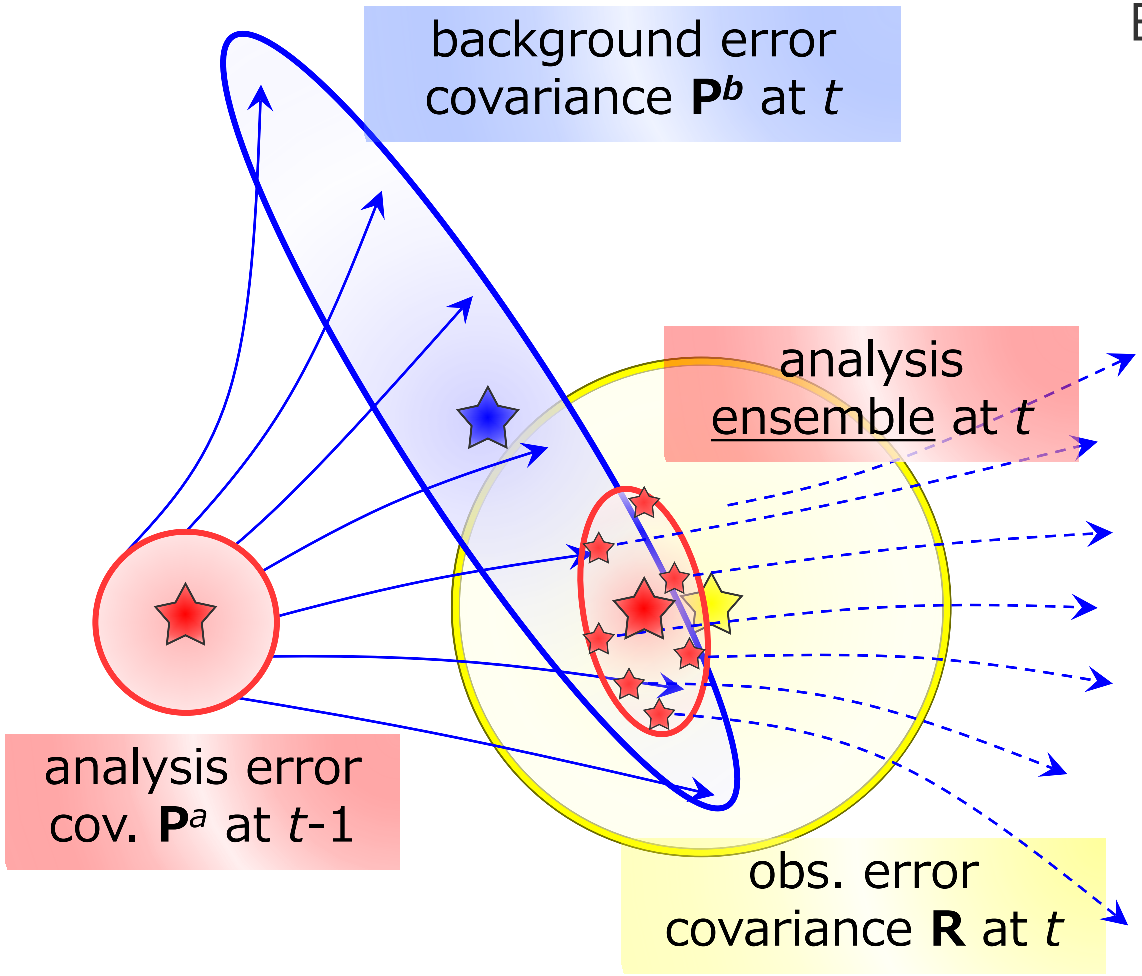
Ensemble Transform Matrix

Ensemble: $\mathbf{X}_t = \left[\mathbf{x}_t^{(1)}, \mathbf{x}_t^{(2)}, \ldots, \mathbf{x}_t^{(m)}\right]$

ens. members

Ens. perturbation $\delta\mathbf{X}_t = \left[\mathbf{x}_t^{(1)} - \bar{\mathbf{x}}_t, \ldots, \mathbf{x}_t^{(m)} - \bar{\mathbf{x}}_t\right]$

ens. members



Ensemble approximation of error cov.

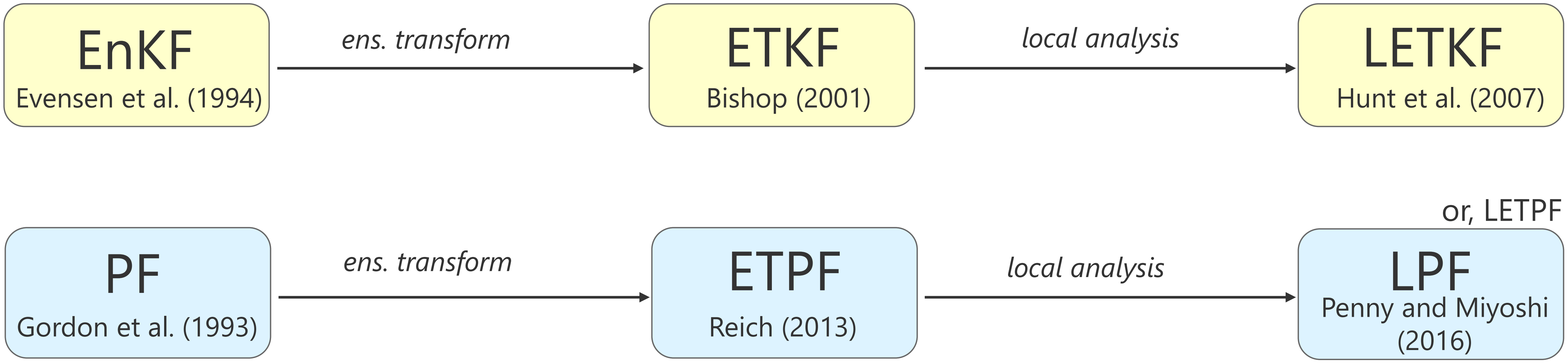
$$\mathbf{P}_t^b \approx \frac{\delta\mathbf{X}_t^b \left(\delta\mathbf{X}_t^b\right)^T}{m-1}$$

analysis is given by a linear combination of forecast ensemble

$$\mathbf{X}_t^a = \bar{\mathbf{x}}_t^b \cdot \mathbf{1}^T + \delta\mathbf{X}_t^b \hat{\mathbf{T}}$$

ensemble transform matrix

Local Ensemble Transform Kalman Filter (LETKF)



The analysis update equation of the LPF is represented by the ensemble transform matrix as the LETKF.

$$\mathbf{X}_t^a = \bar{\mathbf{x}}_t^b \cdot \mathbf{1}^T + \delta\mathbf{X}_t^b \widehat{\mathbf{T}}$$

Background Knowledge

Environmental Prediction Science Laboratory

Bayesian Estimation (Review)

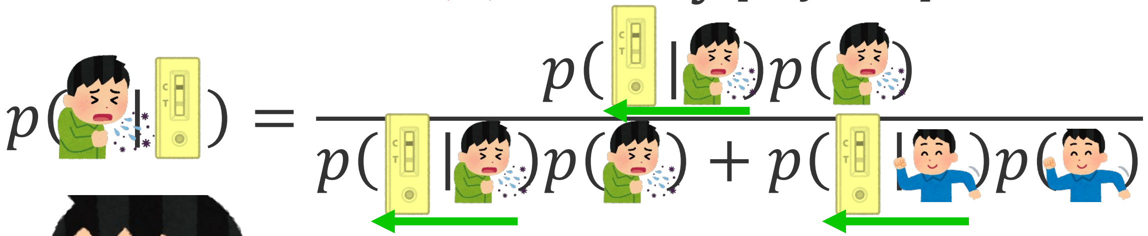
Bayesian Theorem

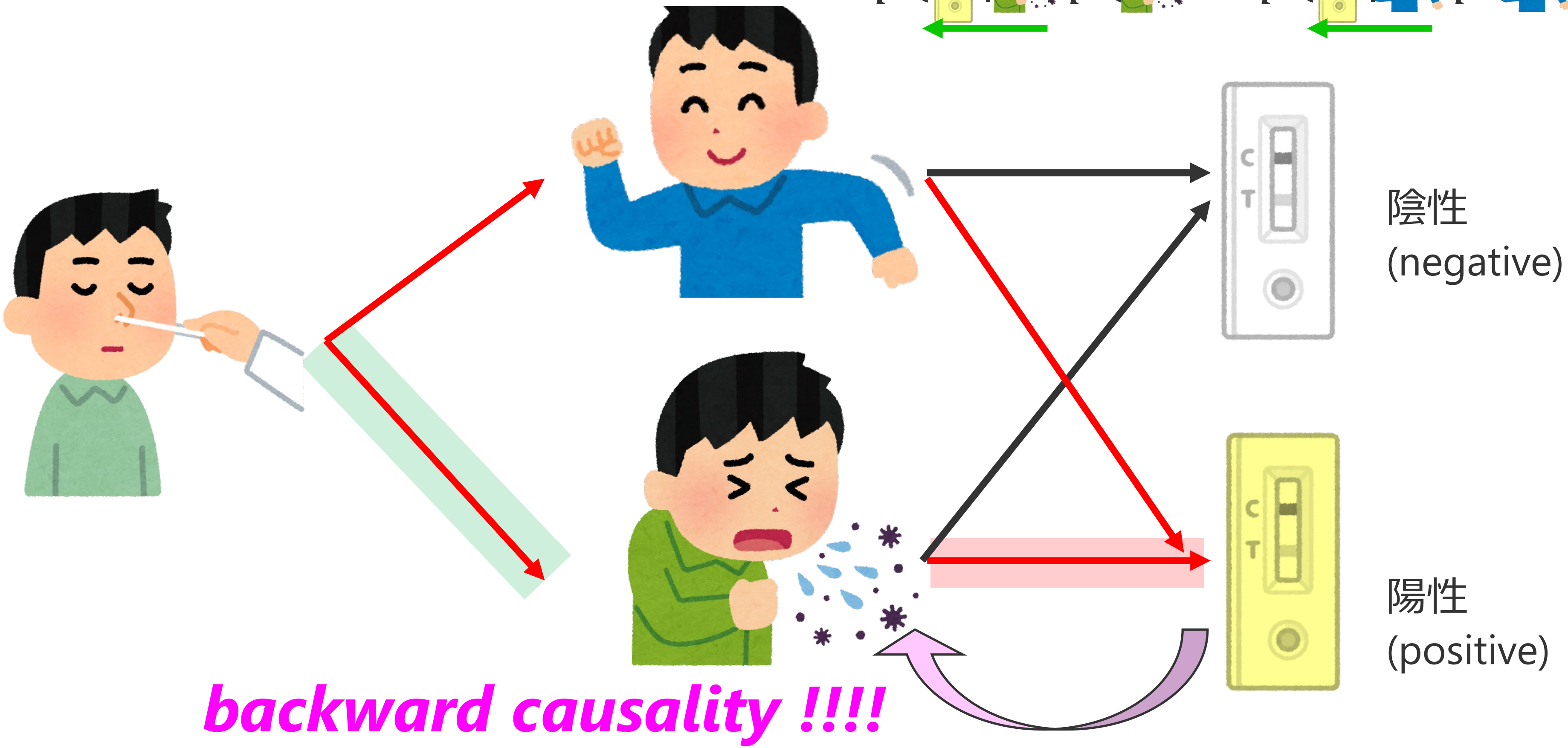
$$p(x|y) = \frac{p(y|x)p(x)}{p(y)}$$

: forward

: backward (結果➔原因)

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)} = \frac{p(y|x)p(x)}{\int p(y|x)p(x)dx}$$





Bayesian Estimation (Review)

Bayesian Theorem (discrete)

$$p(x_i|y) = \frac{p(y|x_i)p(x_i)}{p(y)} = \frac{p(y|x_i)p(x_i)}{\sum_{k=1}^{n} p(y|x_k)p(x_k)}$$

Bayesian Theorem (general)

Likelihood

Prior (uniform)

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)} = \frac{p(y|x)p(x)}{\int p(y|x)p(x)dx}$$

Posterior

constant (i.e., not a func. of x)

$$\left(= \frac{p(y|x)p(x)}{\int p(y|\theta)p(\theta)d\theta} \right)$$

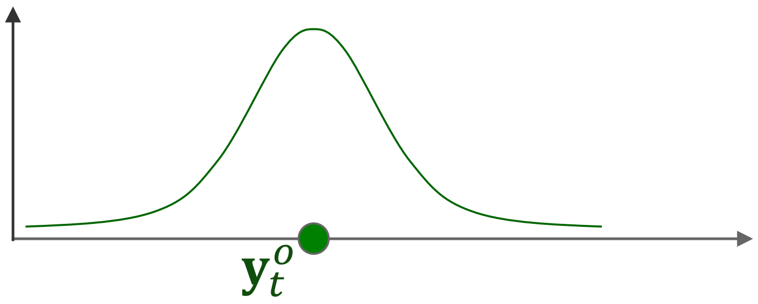
We would like to find x that maximizes $p(x|y)$

$$\text{maximize } p(x|y)$$

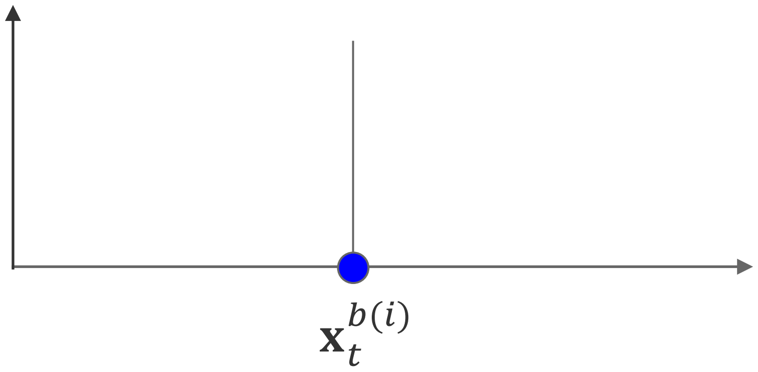
$$\Leftrightarrow \quad \text{maximize } p(y|x)$$

The Dirac's Delta function

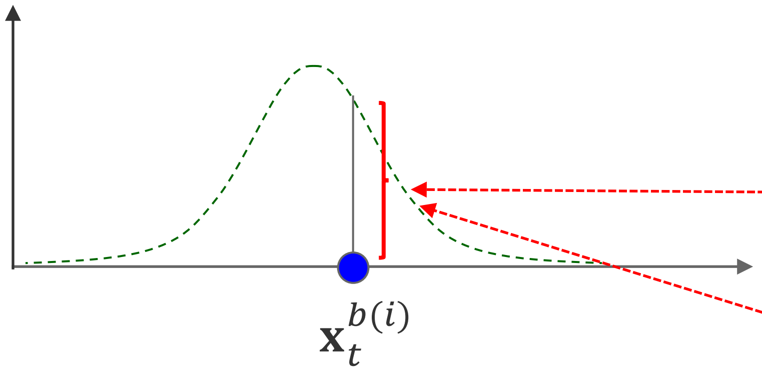
Obs. PDF



Prior PDF



Prior PDF



Observation Error PDF (a case of Gaussian)

δ: *the delta function*

$$P(\mathbf{y}_t^o | \mathbf{y}_{1:t-1}^o) = N(\mathbf{y}_t^o, \mathbf{R}_t)$$

A delta function $\delta(\mathbf{x}_t - \mathbf{x}_t^{b(i)})$

from the definition of the delta function $\int_{-\infty}^{\infty} \delta\left(\mathbf{x}_t - \mathbf{x}_t^{b(i)}\right) = 1$

Multiplication

$$N(\mathbf{y}_t^o, \mathbf{R}_t) \cdot \delta\left(\mathbf{x}_t - \mathbf{x}_t^{b(i)}\right)$$

$$= N\left(\mathbf{x}_t^{b(i)} | \mathbf{y}_t^o, \mathbf{R}_t\right) \cdot \delta\left(\mathbf{x}_t - \mathbf{x}_t^{b(i)}\right)$$

$N(\mathbf{y}_t^o, \mathbf{R}_t)$ evaluated at $\mathbf{x}_t^{b(i)}$ (a scalar)

$$= q_t^{(i)} \cdot \delta\left(\mathbf{x}_t - \mathbf{x}_t^{b(i)}\right)$$

likelihood

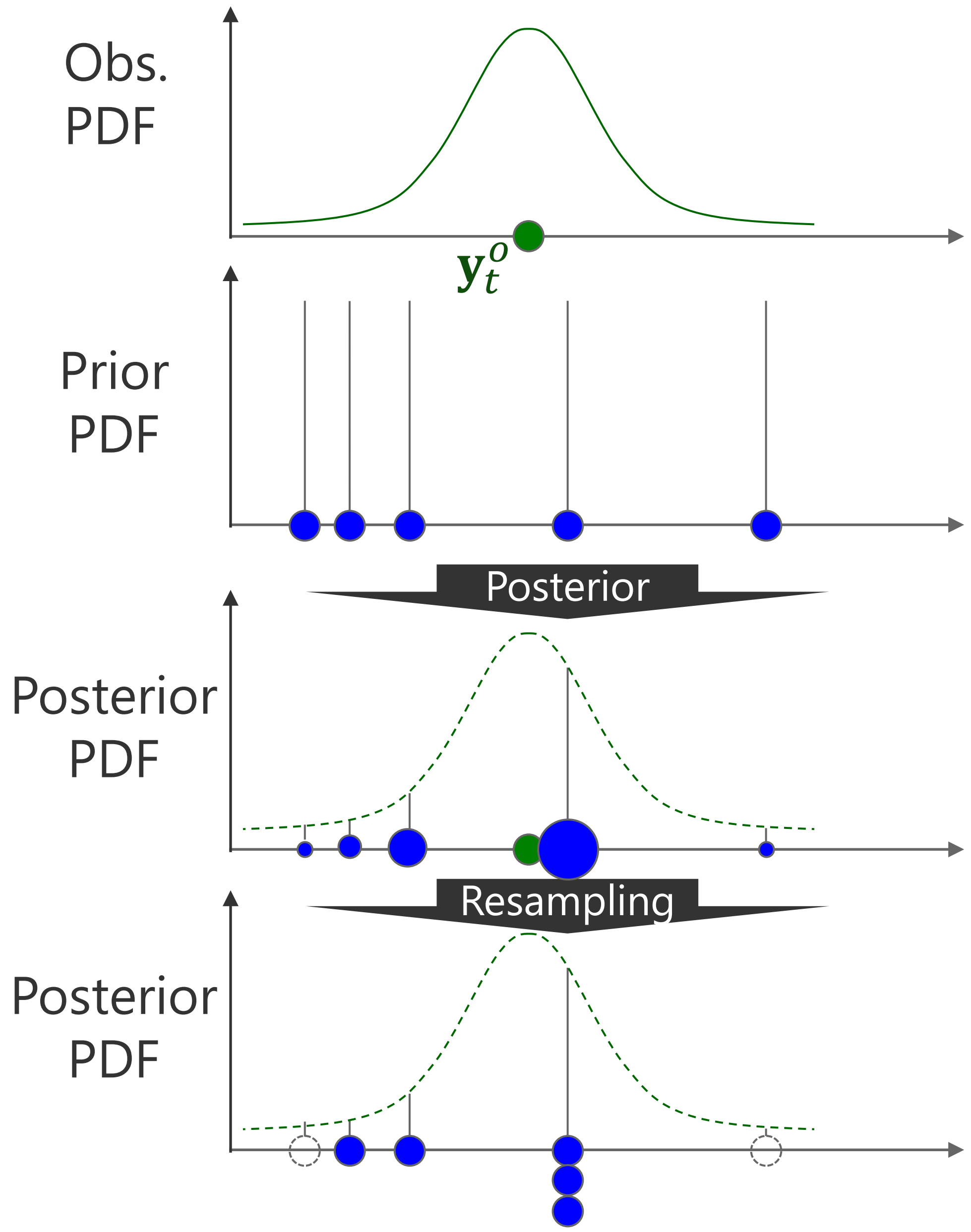
where $q_t^{(i)} = \exp\left[-\frac{1}{2}\left(\mathbf{d}_t^{(i)}\right)^T \mathbf{R}_t^{-1} \mathbf{d}_t^{(i)}\right]$

For simplicity, considering the case then $H() = \mathbf{I}$

Local Particle Filter

Environmental
Prediction
Science
Laboratory

Data Assimilation Steps of Particle Filter



Observation Error (a case of Gaussian)

$$P(\mathbf{y}_t^o|\mathbf{y}_{1:t-1}^o) = N(\mathbf{y}_t^o, \mathbf{R}_t)$$

これってこれでいいんだっけ？

δ: the delta function

Prior

$$P(\mathbf{x}_t|\mathbf{y}_{1:t-1}^o) \approx \frac{1}{m}\sum_{i=1}^{m}\delta(\mathbf{x}_t - \mathbf{x}_t^{b(i)})$$

Posterior from Bayes' theorem

$$P(\mathbf{x}_t|\mathbf{y}_{1:t}^o) = \frac{P(\mathbf{y}_t^o|\mathbf{x}_t)P(\mathbf{x}_t|\mathbf{y}_{1:t-1}^o)}{P(\mathbf{y}_t^o|\mathbf{y}_{1:t-1}^o)}$$

$$\approx \sum_{i=1}^{m}\mathbf{w}_t^{a(i)}\delta(\mathbf{x}_t - \mathbf{x}_t^{b(i)})$$

Posterior after resampling

$$P(\mathbf{x}_t|\mathbf{y}_{1:t}^o) \approx \frac{1}{m}\sum_{i=1}^{m}\delta(\mathbf{x}_t - \mathbf{x}_t^{a(i)})$$

Derivation of PF (1): Prior PDF

δ: *the delta function*

Prior

$$p(\mathbf{x}_t|\mathbf{y}^o_{1:t-1}) = \int p(\mathbf{x}_t|\mathbf{x}_{t-1})p(\mathbf{x}_{t-1}|\mathbf{y}^o_{1:t-1})d\mathbf{x}_{t-1}$$

prior PDF at time *t*

$$\approx \frac{1}{m}\sum_{i=1}^{m} p\left(\mathbf{x}_t|\mathbf{x}^{(i)}_{t-1|t-1}\right)$$

*i*th posterior particle at time *t*-1

prior prob. of *i*th particle at time *t* (i.e., stochastic fcst)

$$\approx \frac{1}{m}\sum_{i=1}^{m} p\left(\mathbf{x}^{b(i)}_t\right)$$

can be any PDF

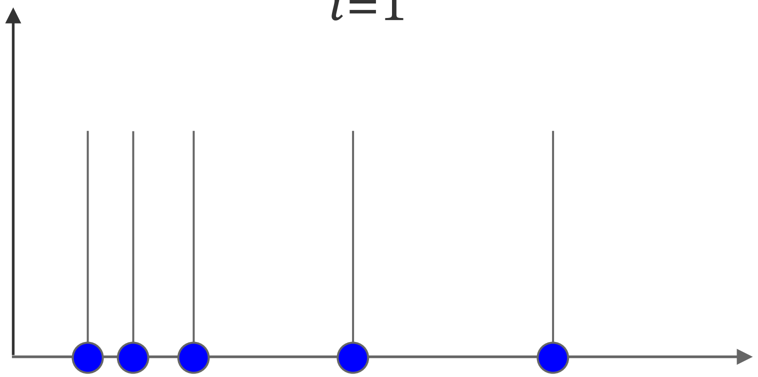
$$p(\mathbf{x}_t|\mathbf{y}^o_{1:t-1}) = p\left(\mathbf{x}^b_t\right)$$

Standard Particle Filters

δ: *the delta function*

$$p(\mathbf{x}_t|\mathbf{y}^o_{1:t-1}) \approx \frac{1}{m}\sum_{i=1}^{m} \delta(\mathbf{x}_t - \mathbf{x}^{b(i)}_t)$$

Prior PDF

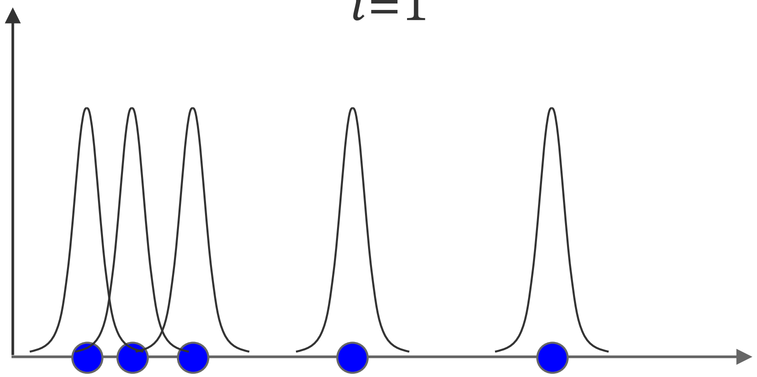


LPF with Gaussian Mixture

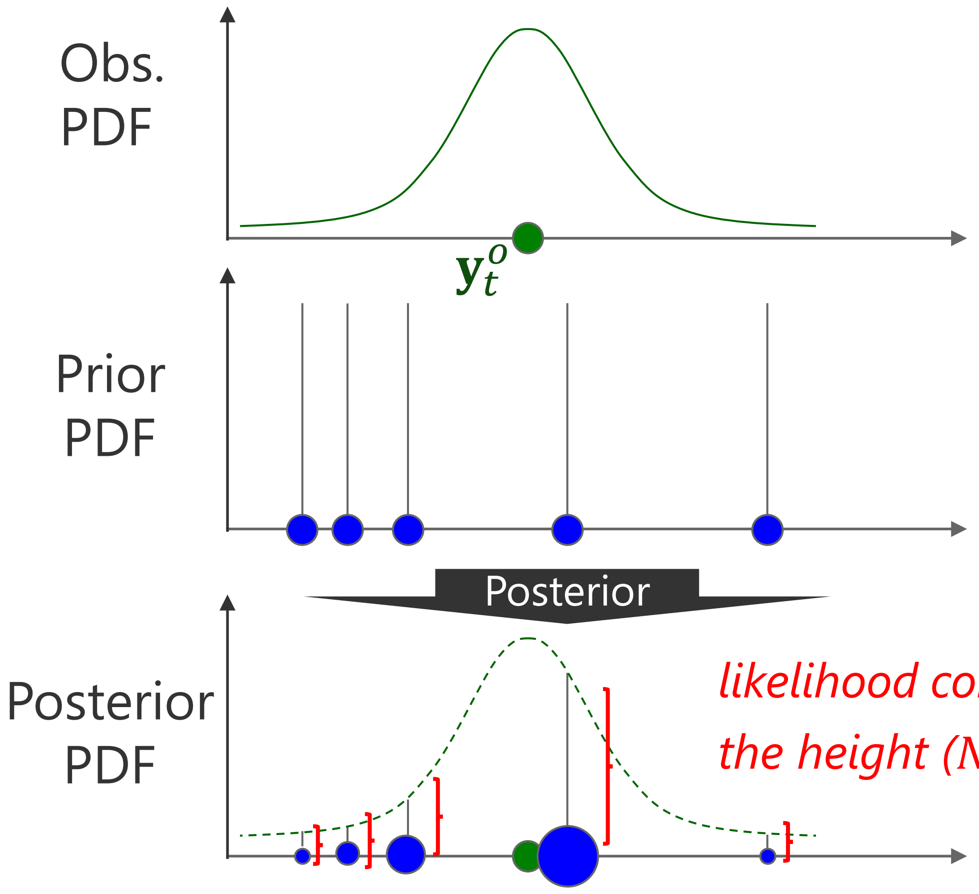
Hoteit et al. (2008)
Stordal et al. (2011)

$$p(\mathbf{x}_t|\mathbf{y}^o_{1:t-1}) \approx \frac{1}{m}\sum_{i=1}^{m} N(\mathbf{x}^{b(i)}_t, \mathbf{B})$$

Prior PDF



Derivation of PF (2): Posterior PDF



likelihood corresponds to the height ($N\left(\mathbf{x}_t^{b(i)}|\mathbf{y}_t^o,\mathbf{R}_t\right)$)

$$q_t^{(i)} = \exp\left[-\frac{1}{2}\left(\mathbf{d}_t^{(i)}\right)^T \mathbf{R}_t^{-1}\mathbf{d}_t^{(i)}\right]$$

$$\mathbf{d}_t^{(i)} \coloneqq \mathbf{y}_t^o - H_t\left(\mathbf{x}_t^{b(i)}\right)$$

Posterior from Bayes' theorem

δ: the delta function

$$P(\mathbf{x}_t|\mathbf{y}_{1:t}^o) = \frac{P(\mathbf{y}_t^o|\mathbf{x}_t)P(\mathbf{x}_t|\mathbf{y}_{1:t-1}^o)}{P(\mathbf{y}_t^o|\mathbf{y}_{1:t-1}^o)}$$

$$\approx \frac{p(\mathbf{y}_t^o|\mathbf{x}_t)\frac{1}{m}\sum_{i=1}^m \delta\left(\mathbf{x}_t - \mathbf{x}_t^{b(i)}\right)}{\int p(\mathbf{y}_t^o|\mathbf{x}_t)\frac{1}{m}\sum_{i=1}^m \delta\left(\mathbf{x}_t - \mathbf{x}_t^{b(i)}\right) d\mathbf{x}_t}$$

numerator $p(\mathbf{y}_t^o|\mathbf{x}_t)p(\mathbf{x}_t|\mathbf{y}_{1:t-1}^o)$

$$= \frac{1}{m}\sum_{i=1}^m p\left(\mathbf{y}_t^o|\mathbf{x}_t^{b(i)}\right)p\left(\mathbf{x}_t^{b(i)}\right)$$

$$= \frac{1}{m}\sum_{i=1}^m p\left(\mathbf{y}_t^o|\mathbf{x}_t^{b(i)}\right)\delta\left(\mathbf{x}_t - \mathbf{x}_t^{b(i)}\right)$$

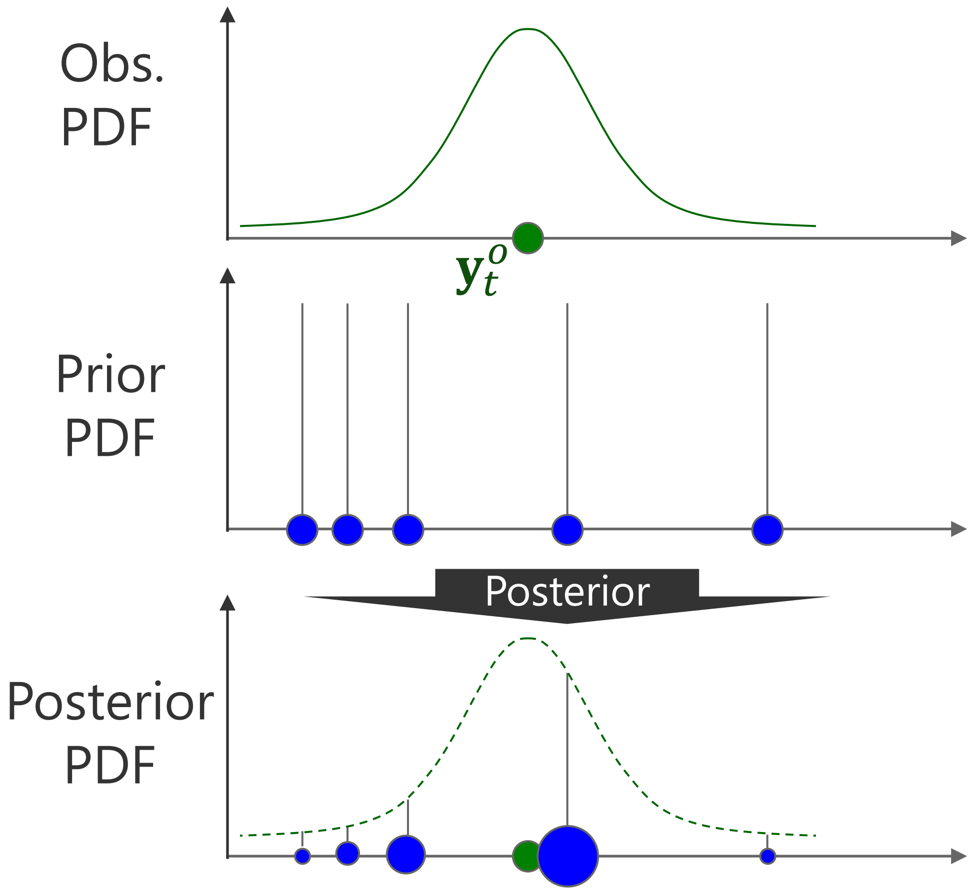
$$= \frac{1}{m}\sum_{i=1}^m \underline{N\left(\mathbf{x}_t^{b(i)}|\mathbf{y}_t^o,\mathbf{R}_t\right)} \cdot \delta\left(\mathbf{x}_t - \mathbf{x}_t^{b(i)}\right)$$

$N(\mathbf{y}_t^o,\mathbf{R}_t)$ evaluated at $\mathbf{x}_t^{b(i)}$ (a scalar)

$$= \frac{1}{m}\sum_{i=1}^m \underline{q_t^{(i)}} \cdot \delta\left(\mathbf{x}_t - \mathbf{x}_t^{b(i)}\right)$$

Likelihood (a scalar)

Derivation of PF (2): Posterior PDF (cont'd)



$$q_t^{(i)} = \exp\left[-\frac{1}{2}\left(\mathbf{d}_t^{(i)}\right)^T \mathbf{R}_t^{-1}\mathbf{d}_t^{(i)}\right]$$

$$\mathbf{d}_t^{(i)} \coloneqq \mathbf{y}_t^o - H_t\left(\mathbf{x}_t^{b(i)}\right)$$

Posterior from Bayes' theorem

δ: the delta function

$$P(\mathbf{x}_t|\mathbf{y}_{1:t}^o) = \frac{P(\mathbf{y}_t^o|\mathbf{x}_t)P(\mathbf{x}_t|\mathbf{y}_{1:t-1}^o)}{P(\mathbf{y}_t^o|\mathbf{y}_{1:t-1}^o)}$$

$$\approx \frac{p(\mathbf{y}_t^o|\mathbf{x}_t)\frac{1}{m}\sum_{i=1}^m \delta\left(\mathbf{x}_t - \mathbf{x}_t^{b(i)}\right)}{\int p(\mathbf{y}_t^o|\mathbf{x}_t)\frac{1}{m}\sum_{i=1}^m \delta\left(\mathbf{x}_t - \mathbf{x}_t^{b(i)}\right) d\mathbf{x}_t}$$

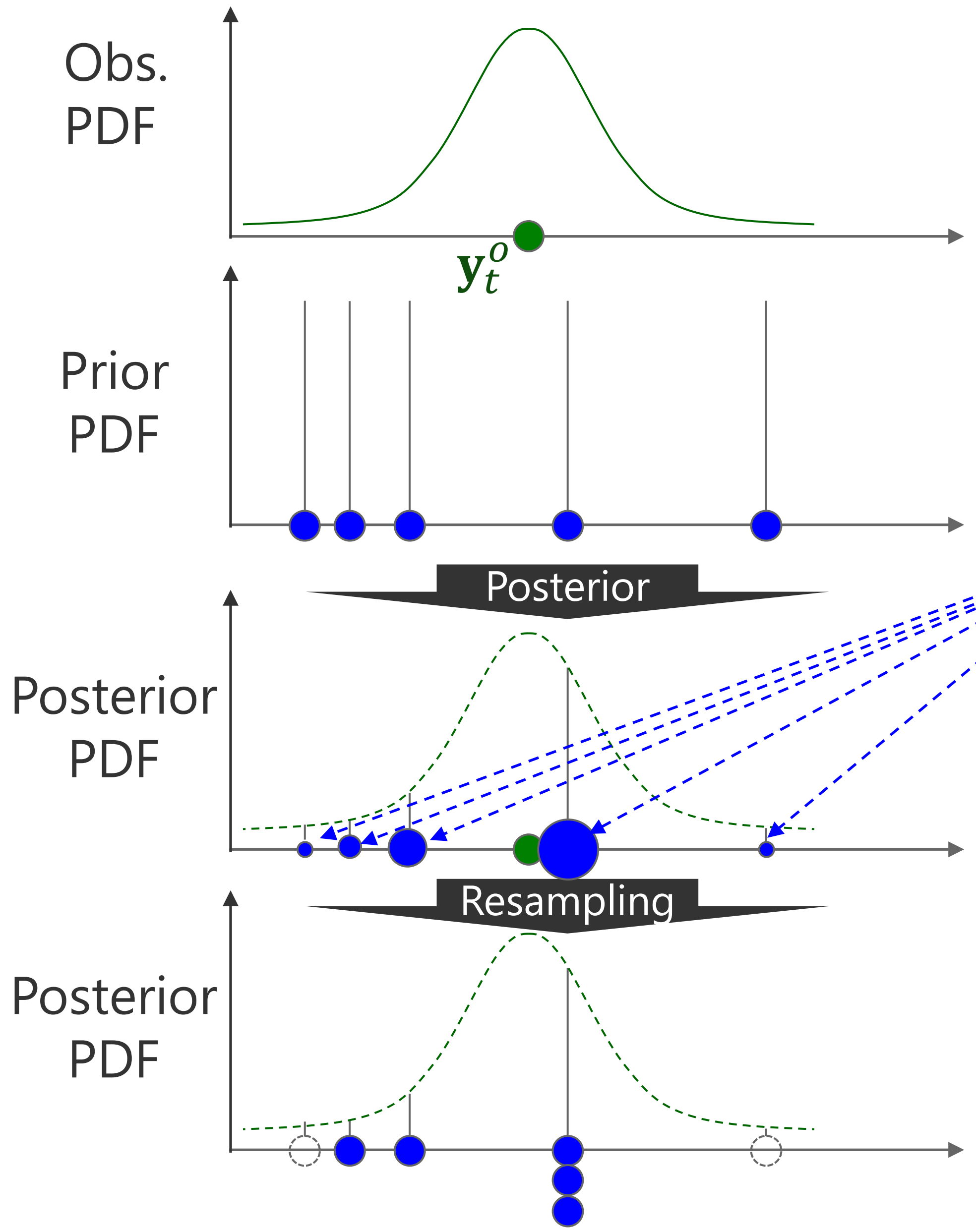
numerator $$= \frac{1}{m}\sum_{i=1}^m q_t^{(i)} \cdot \delta\left(\mathbf{x}_t - \mathbf{x}_t^{b(i)}\right)$$

denominator $$= \int \frac{1}{m}\sum_{i=1}^m q_t^{(i)} \cdot \delta\left(\mathbf{x}_t - \mathbf{x}_t^{b(i)}\right)\ d\mathbf{x}_t$$

$$= \frac{1}{m}\sum_{i=1}^m q_t^{(i)} \int \delta\left(\mathbf{x}_t - \mathbf{x}_t^{b(i)}\right) d\mathbf{x}_t$$

$$= \frac{1}{m}\sum_{i=1}^m q_t^{(i)}$$

Derivation of PF (2): Posterior PDF (cont'd)



Posterior from Bayes' theorem

$$P(\mathbf{x}_t|\mathbf{y}_{1:t}^o) = \frac{P(\mathbf{y}_t^o|\mathbf{x}_t)P(\mathbf{x}_t|\mathbf{y}_{1:t-1}^o)}{P(\mathbf{y}_t^o|\mathbf{y}_{1:t-1}^o)} \approx \sum_{i=1}^{m} \mathbf{w}_t^{a(i)} \delta(\mathbf{x}_t - \mathbf{x}_t^{b(i)})$$

weight

$$\mathbf{w}_t^{a(i)} = q_t^{(i)} \Big/ \left\{\sum_{k=1}^{m} q_t^{(k)}\right\}$$

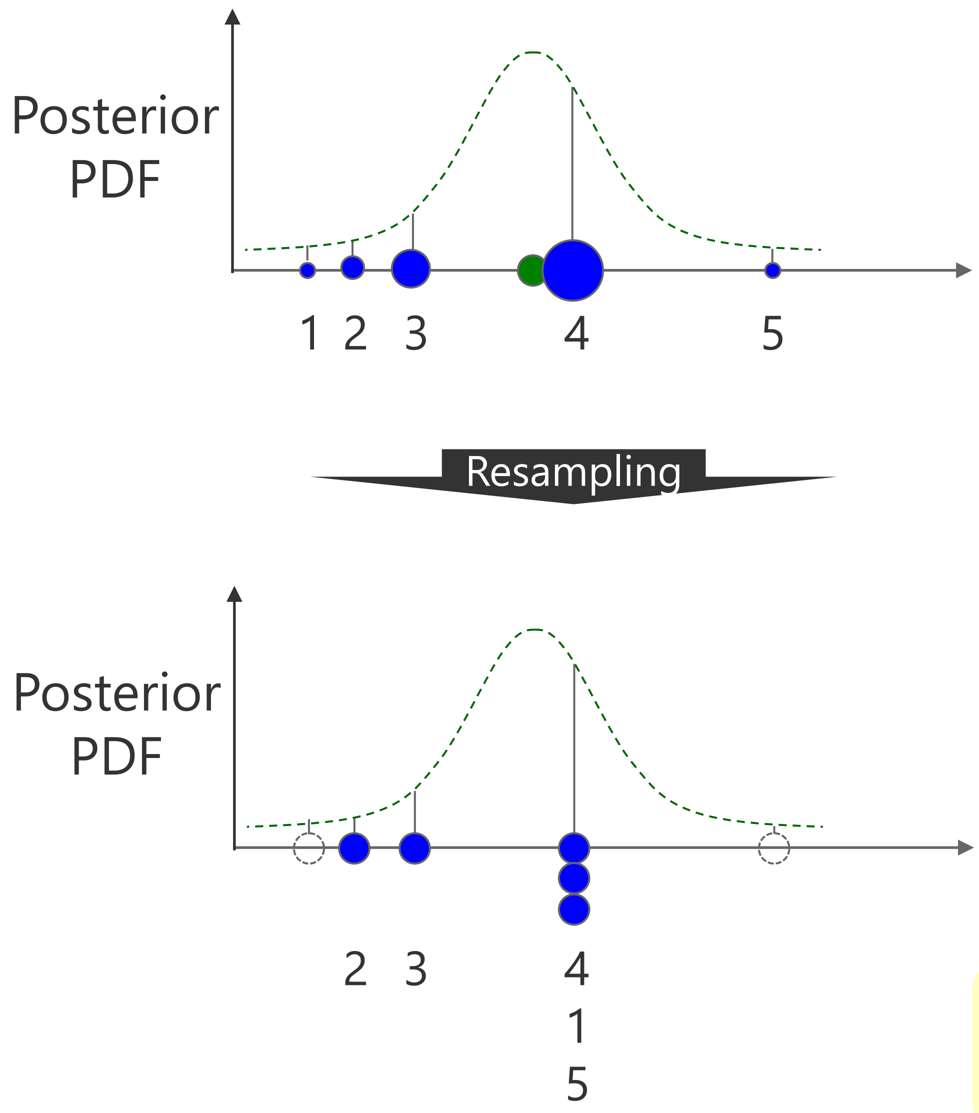
Likelihood

$$q_t^{(i)} = \exp\left[-\frac{1}{2}\left(\mathbf{d}_t^{(i)}\right)^T \mathbf{R}_t^{-1} \mathbf{d}_t^{(i)}\right]$$

can be localized (LPF)
by R Localization as in the LETKF

where $\mathbf{d}_t^{(i)} := \mathbf{y}_t^o - H_t\left(\mathbf{x}_t^{b(i)}\right)$

Derivation of PF (3): Resampling



$$\mathbf{X}_t^b \coloneqq \left[\mathbf{x}_t^{b(1)}, \ldots, \mathbf{x}_t^{b(m)}\right] \qquad \mathbf{X}_t^a \coloneqq \left[\mathbf{x}_t^{a(1)}, \ldots, \mathbf{x}_t^{a(m)}\right]$$

transform matrix

$$\mathbf{X}_t^a = \mathbf{X}_t^b \widehat{\mathbf{T}}_t$$

Index of posterior particle (columns); Index of prior particle (rows)

$$\widehat{\mathbf{T}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$= \left(\bar{\mathbf{x}}_t^b \cdot \mathbf{1}^T + \delta \mathbf{X}_t^b\right) \widehat{\mathbf{T}}_t$$

$$= \bar{\mathbf{x}}_t^b \cdot \mathbf{1}^T + \delta \mathbf{X}_t^b \widehat{\mathbf{T}}_t \quad \textit{because} \sum_{i=1}^{m} \left(\widehat{\mathbf{T}}_t\right)_{i,j} = \mathbf{1}^T$$

LPF can be developed in a form as the LETKF
➔ We can develop LPF based on LETKF code!

An Example of Implementation

Environmental Prediction Science Laboratory

C:¥Users¥kotsuki-vaio2¥AppData¥Local¥Temp¥scp10879¥fefs¥data¥o¥o140¥speedy¥letkf-master_jss2_LPF¥speedy¥LTERPLUT_letkf¥letkf_tools.f90 - Sublime Text 2 (UNREGISTERED)

File Edit Selection Find View Goto Tools Project Preferences Help

letkf_tools.f90

```
  !!WRITE(6,'(A,I3)') 'ilev = ',ilev
  DO ij=1,nij1
    DO n=1,nv3d
      IF(var_local_n2n(n) < n) THEN
        trans(:,:,n)                  = trans(:,:,var_local_n2n(n))
        work3d(ij,ilev,n)             = work3d(ij,ilev,var_local_n2n(n))
        wvec3d(ij,ilev,n,1:nbv)       = wvec3d(ij,ilev,var_local_n2n(n),1:nbv)           !<------ 6/6/2018 A.
        wmat3d(ij,ilev,n,1:nbv,1:nbv)= wmat3d(ij,ilev,var_local_n2n(n),1:nbv,1:nbv)      !<------ 6/6/2018 A.
        nobs3d(ij,ilev,n)             = nobs3d(ij,ilev,var_local_n2n(n))                 !

        pvec3d(ij,ilev,n,1:nbv)       = pvec3d(ij,ilev,var_local_n2n(n),1:nbv)
        pmat3d(ij,ilev,n,1:nbv,1:nbv)= pmat3d(ij,ilev,var_local_n2n(n),1:nbv,1:nbv)
        peff3d(ij,ilev,n)             = peff3d(ij,ilev,var_local_n2n(n))
        pnum3d(ij,ilev,n)             = pnum3d(ij,ilev,var_local_n2n(n))
      ELSE
        CALL obs_local(ij,ilev,n,hdxf,rdiag,rloc,dep,nobsl,logpfm)
        parm = work3d(ij,ilev,n)
        nobs3d(ij,ilev,n) = REAL( nobsl, r_size )
        CALL letkf_core(nobstotal,nobsl,hdxf,rdiag,rloc,dep,parm,trans(:,:,n),wvec3d(ij,ilev,n,1:nbv),wmat3
        work3d(ij,ilev,n) = parm

        !if( myrank==1 ) & !debug!
        CALL lpf_core  (nobstotal,nobsl,hdxf,rdiag,rloc,dep,parm,peff3d(ij,ilev,n),pnum3d(ij,ilev,n),pvec3d
      END IF
    END DO
    IF(ilev == 1) THEN !update 2d variable at ilev=1
```

loops for variables and model grids

search local observations

local analysis of the LETKF

local analysis of the LPF

LPF can be implemented easily if one had the code of LETKF

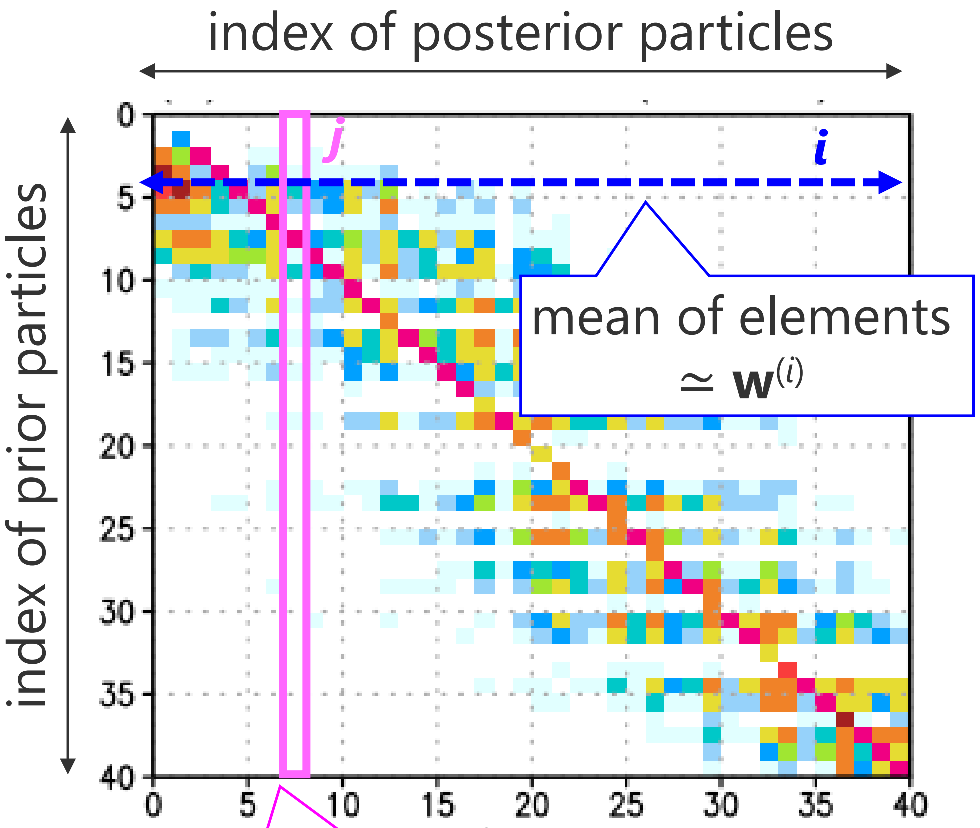
Line 297, Column 59 Spaces: 2 Fortran - Modern

On Resampling Matrix

Environmental
Prediction
Science
Laboratory

Transform (or Resampling) matrix in LPF

$$\mathbf{X}_t^a = \bar{\mathbf{x}}_t^b \cdot \mathbf{1} + \delta\mathbf{X}_t^b \widehat{\mathbf{T}}_t$$



an example of resampling matrix

*j*th posterior particle is given by linear combination of prior particles

Some requirements for T

(1) Posterior particle is given by linear combination of prior particles *(necessary)*

$$\sum_{i=1}^{m} \widehat{\mathbf{T}}_t^{(i,j)} = 1 \qquad j=1,\ldots,m$$

(2) Posterior mean is given by weighted average of prior particles *(preferable)*

$$\bar{\mathbf{x}}_t^a = \sum_{j=1}^{m} \mathbf{x}_t^{b(j)} \cdot \mathbf{w}_t^{(j)}$$

$$\Leftrightarrow \frac{1}{m}\sum_{j=1}^{m} \widehat{\mathbf{T}}_t^{(i,j)} = \mathbf{w}_t^{(i)} \qquad i=1,\ldots,m$$

(3) Spatially smooth **T** as in LETKF *(preferable)*

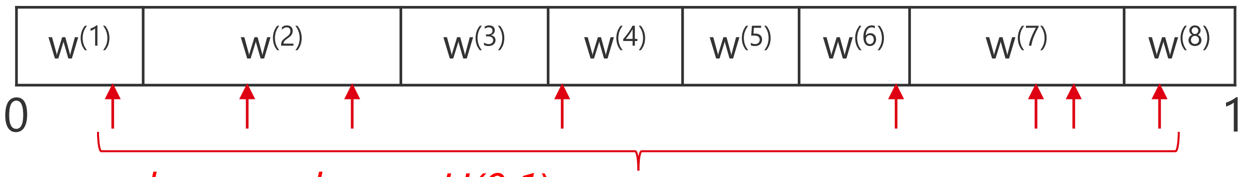
Infinite solutions

Q: how to determine T?

Stochastic Multinomial Resampling (SMR)

Used in Kotsuki et al. 2022 (GMD)

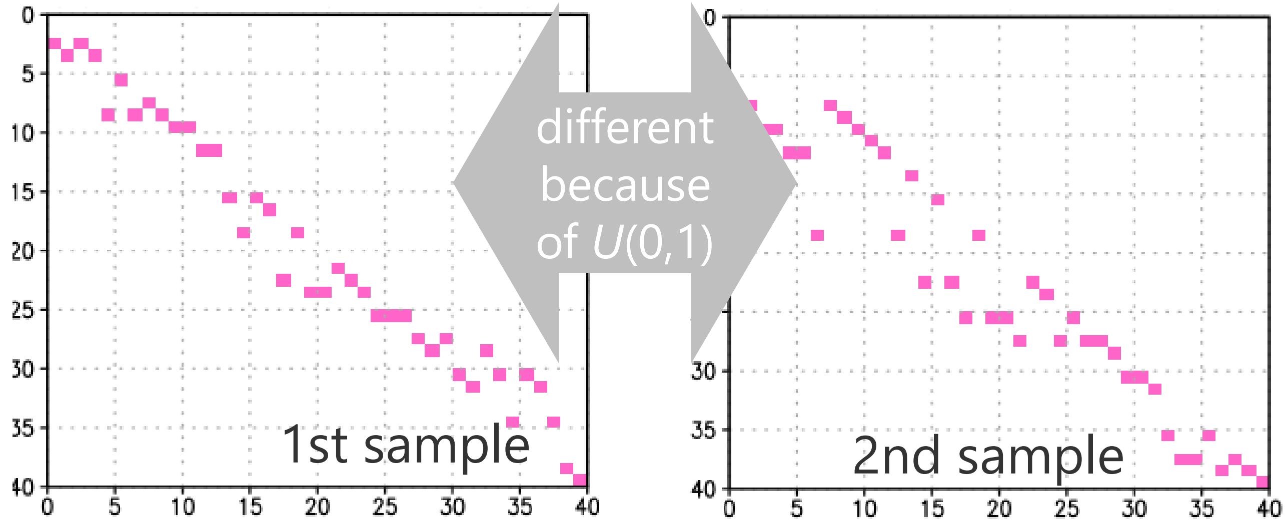
Multinomial Resampling (MR)



accumulated weight (a case of m=8)

The posterior particles are determined by m random numbers ~ U(0,1)

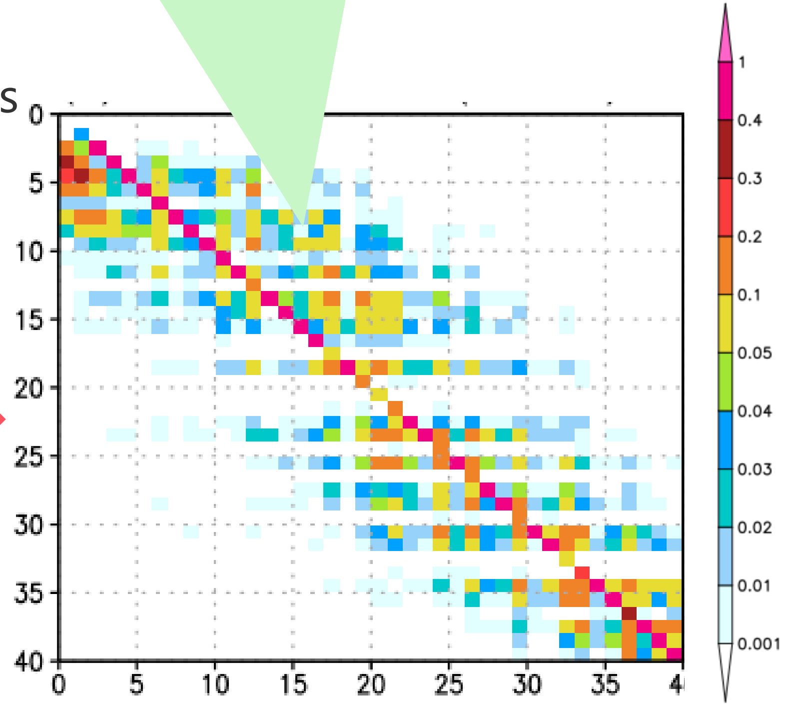
MR w/ <u>prioritized insertion to diagonal elements</u>



N times

• • •

average



Noise in the matrix would be also important to maintain the diversity of posterior particles.

this matrix approximates

$$(2) \quad \bar{\mathbf{x}}_t^a = \sum_{j=1}^{m} \mathbf{x}_t^{b(j)} \cdot \mathbf{w}_t^{(j)}$$

Optimal Transport-based Resampling

ETKF (symmetric square root matrix) (Bishop et al. 2001; Wang et al. 2004, Hunt et al. 2007)

to minimize mean square distance b/w **I** and **W** where $\delta\mathbf{X}_t^a = \delta\mathbf{X}_t^b\mathbf{W}$

(1) to reduce updates of ensemble perturbation
(2) to ensure spatially smooth transition of **W**

It would be important in LPF, too

LPF w/ optimal transport (OT) (Reich 2013; Farchi and Bocquet 2016)

$$\underset{\mathbf{T}}{\operatorname{argmin}} \sum_{j=1}^{m}\sum_{i=1}^{m} \mathbf{C}\hat{\mathbf{T}}_t^{(i,j)}$$

(1) minimize analysis increments of particles
(2) solved exactly (=uniquely) by simplex method

LPF w/ Sinkhorn algorithm (Oishi and Kotsuki 2023)

$$\underset{\mathbf{T}}{\operatorname{argmin}} \sum_{j=1}^{m}\sum_{i=1}^{m} \mathbf{C}\hat{\mathbf{T}}_t^{(i,j)} + \varepsilon \sum_{j=1}^{m}\sum_{i=1}^{m} \hat{\mathbf{T}}_t^{(i,j)}\log(\hat{\mathbf{T}}_t^{(i,j)})$$

ε: regularization term

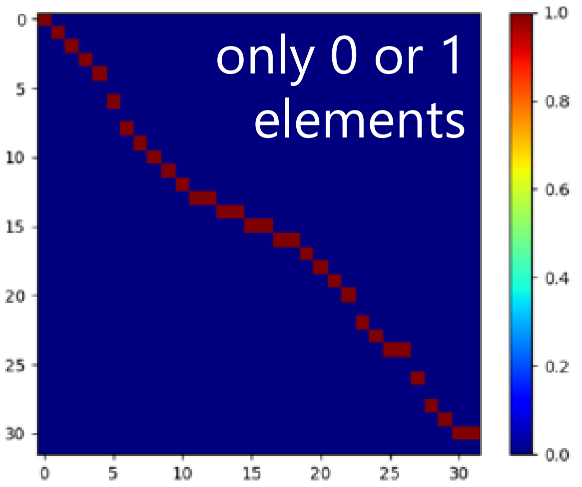
(1) approximately minimize analysis increments of particles
(2) solved iteratively (=non-uniquely) by an internal method

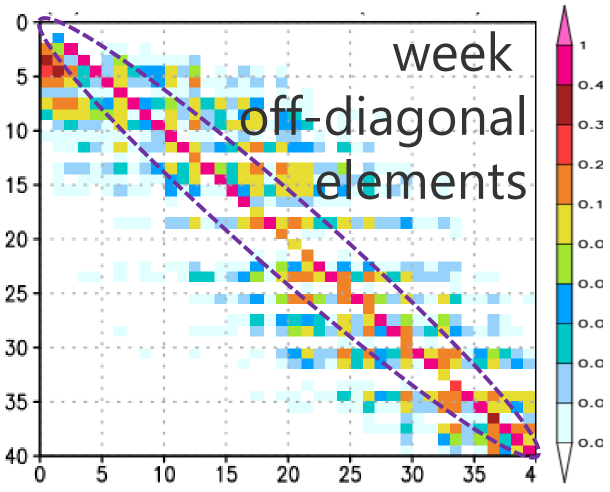
C is *mxm* cost matrix whose (*i,j*)th element is distance b/w *i*th and *j*th particles
(this study takes L2 norm within localization radius)

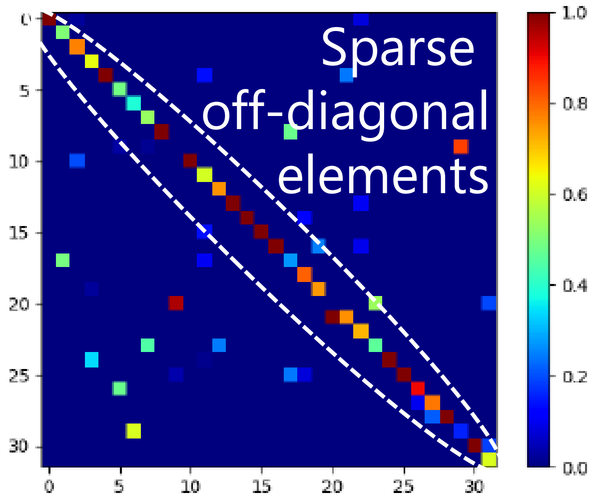
Comparison of Resampling methods

Resampling method	SU (e.g. Penny and Miyoshi 2016)	Stochastic MR (e.g. Kotsuki et al. 2022)	Sinkhorn (e.g. Oishi and Kotsuki 2023)
(1) $\sum_{i=1}^{m} \mathbf{T}_t^{(i,j)} = 1$	◎	◎	◎
(2) $\frac{1}{m}\sum_{j=1}^{m} \mathbf{T}_t^{(i,j)} = \boldsymbol{w}_t^{(i)}$	△	○	◎
(3) Spatially smooth T	△	○ (due to diagonal insertion)	◎ (due to optimal transport)
Comp. complexity	$O(m)$	$O(N_{SMR} \times m\log(m))$	$O(m^2)$

Examples of resampling matrices







Thank you for your attention!

Presented by Shunji Kotsuki
(shunji.kotsuki@chiba-u.jp)

Further information is available at
https://kotsuki-lab.com/

Environmental Prediction Science Laboratory