University of Potsdam

SFB 1114

Scaling Cascades in Complex Systems

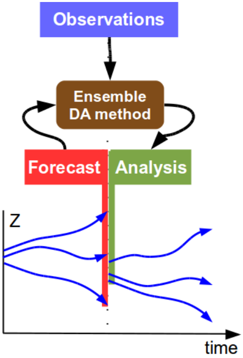
Nonlinear data assimilation via hybrid particle-Kalman filters and optimal coupling

Walter Acevedo, Sebastian Reich and Maria Reinhardt

Universität Potsdam, Germany

RIKEN International Symposium on Data Assimilation 2017
27 February – 2 March, 2017, AICS, Kobe, Japan

University of Potsdam



Traditional approaches

- Ensemble Kalman Filters
  - + Robust and moderately affordable
  - − Biased for non-Gaussian PDFs
- Traditional Particle Filters
  - + Non-Gaussianity properly handled
  - − Liable to the "Curse of Dimensionality"

Some recent developments

- Localized particle filters [Lei and Bickel, 2011][Poterjoy, 2015]
- Hybrid Kalman-particle filters [Frei and Künsch, 2013][Chustagulprom et al., 2016]
- Optimal transportation based particle filters [Reich, 2013]
- Intrinsic Dimension [Agapiou et al, arXiv:1511.06196]

Outline

Motivation

Linear Ensemble Transform Filters

Hybrid Ensemble Transform Filter

Single one-dimensional DA step

Non-spatially extended dynamical systems

Spatially extended systems

Conclusions and Prospect

- Forecast and Analysis ensembles:

$$z_i^{\mathrm{f}}, z_i^{\mathrm{a}}, \qquad i = 1, \ldots, M$$

- Update step via a General linear transformation [Reich and Cotter, 2015]

$$z_j^{\mathrm{a}} = \sum_{i=1}^{M} z_i^{\mathrm{f}} d_{ij}$$

- Transform coefficients satisfy

$$\sum_{i=1}^{M} d_{ij} = 1$$

$$\sum_{j=1}^{M} d_{ij} = M w_i$$

ESRF as a LETF:

$$z_j^{\mathrm{a}} = \sum_{i=1}^{M} z_i^{\mathrm{f}} \hat{w}_i + \sum_{i=1}^{M} (z_i^{\mathrm{f}} - \bar{z}^{\mathrm{f}}) s_{ij} = \sum_{i=1}^{M} z_i^{\mathrm{f}} d_{ij}^{\mathrm{KF}}$$

- Transform coefficients:

$$d_{ij}^{\mathrm{KF}} = d_{ij}^{\mathrm{KF}}(\{z_l^{\mathrm{f}}\}, y_{\mathrm{obs}}) := s_{ij} + \hat{w}_i - \frac{1}{M},$$

- Square root matrix: $S = \{s_{ij}\}$

$$S = \left\{ I + \frac{1}{M-1} (HA^{\mathrm{f}})^{\mathrm{T}} R^{-1} HA^{\mathrm{f}} \right\}^{-1/2},$$

- ESRF “Weights”:

$$\hat{w}_i = \frac{1}{M} - \frac{1}{M-1} e_i^T S^2 (HA^{\mathrm{f}})^{\mathrm{T}} R^{-1} (H\bar{z}^{\mathrm{f}} - y_{\mathrm{obs}}), \qquad \sum_{i=1}^{M} \hat{w}_i = 1$$

Analysis Mean

$$\bar{z}^{\mathrm{a}} = \sum_{i=1}^{M} w_i z_i^{\mathrm{f}}, \quad w_i \sim \exp\left(-\frac{1}{2}(Hz_i^{\mathrm{f}} - y_{\mathrm{obs}})^{\mathrm{T}} R^{-1} (Hz_i^{\mathrm{f}} - y_{\mathrm{obs}})\right)$$

ETPF transform update $DF : \{d_{ij}^{\mathrm{PF}}\}$

$$z_j^{\mathrm{a}} = \sum_{i=1}^{M} z_i^{\mathrm{f}} d_{ij}^{\mathrm{PF}}$$

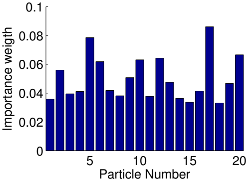
Transform constraints:

$$\sum_{i=1}^{M} d_{ij} = 1, \quad \sum_{j=1}^{M} d_{ij} = w_i M, \quad d_{ij} \geq 0,$$

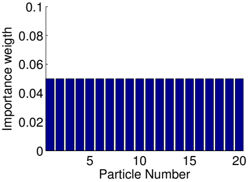
Minimal Transportation Cost:

$$D^{PF} = \arg\min\, J(D), \quad J(D) = \sum_{i,j=1}^{M} d_{ij} \|z_i^{\mathrm{f}} - z_j^{\mathrm{f}}\|^2$$

Optimal Transportation (OT) Problem



⇒ **Optimal Coupling** ⇒



Optimal transportation cost: min $J(D)$

- also known as Earth Mover distance
- takes into account shape of the PDFs

Optimal coupling: D^{PF}

- maximizes correlation between prior and posterior ensembles
- is deterministic, preserving the regularity of the state fields
- consistent with Bayes' formula [Reich 2013]
- convergence (regarding ensemble size) is faster than SIR-PFs at the cost of solving an optimal transportation problem

Exact solution

- Computational complexity $\mathcal{O}(M^3 \ln(M))$
- Efficient Earth Mover Distance algorithms available, e.g. FastEMD [Pele and Werman, 2009].

Entropic Regularized Approximation [Cuturi, 2013]

$$J(D) = \sum_{i,j=1}^{M} \left\{ d_{ij} \| z_i^{\mathrm{f}} - z_j^{\mathrm{f}} \|^2 + \frac{1}{\lambda} d_{ij} \ln d_{ij} \right\}$$

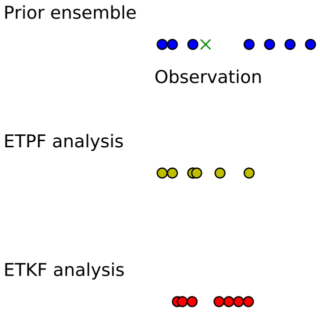
- Sinkhorn's fixed point iteration can be used.
- Computational complexity $\mathcal{O}(M^2 \cdot C(\lambda))$

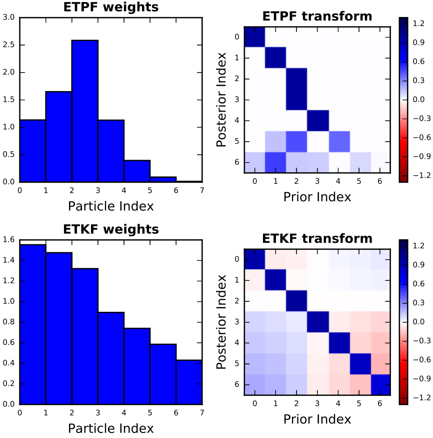
1D Approximation(Each variable independently updated)

- OT problem reduces to reordering.
- Computational complexity $\mathcal{O}(M \ln(M))$
- No particle distance needed

University of Potsdam

Observation within the convex hull





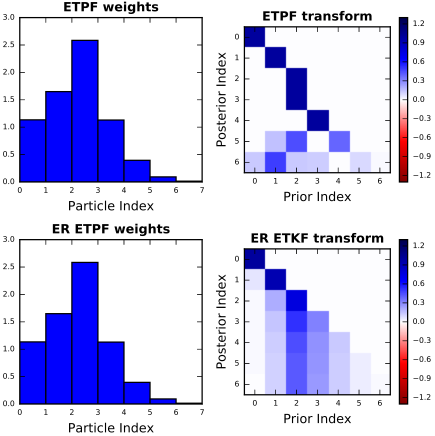
Observation within the convex hull

Prior ensemble

Observation

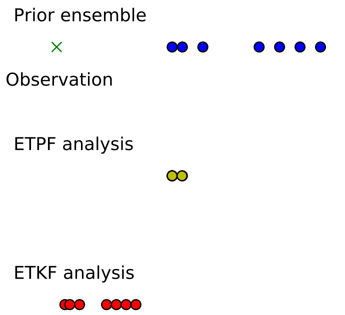
ETPF analysis

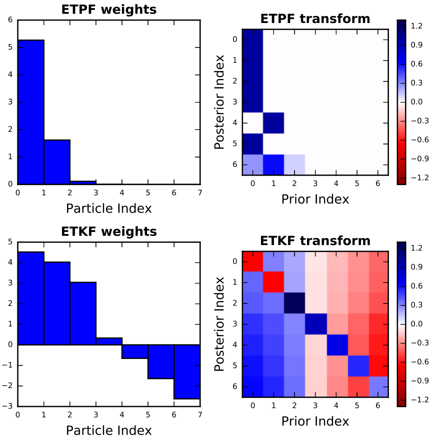
Entropic Reg. ETPF analysis



University of Potsdam

Observation outside the convex hull





University of Potsdam

Likelihood splitting [Frei and Künsch, 2013]

$$\pi_Y(y_{\text{obs}}|z) \propto \exp\left(-\frac{1}{2}(\tilde{y}_z)^{\mathrm{T}}R^{-1}\tilde{y}_z\right), \quad \tilde{y}_z = Hz - y_{\text{obs}}$$

$$\propto \quad \exp\left(\alpha\left(-\frac{1}{2}(\tilde{y}_z)^{\mathrm{T}}R^{-1}\tilde{y}_z\right)\right) \quad \times \exp\left((1-\alpha)\left(-\frac{1}{2}(\tilde{y}_z)^{\mathrm{T}}R^{-1}\tilde{y}_z\right)\right)$$

$$\propto \quad \exp\left(-\frac{1}{2}(\tilde{y}_z)^{\mathrm{T}}\left(\frac{R}{\alpha}\right)^{-1}\tilde{y}_z\right) \quad \times \exp\left(-\frac{1}{2}(\tilde{y}_z)^{\mathrm{T}}\left(\frac{R}{1-\alpha}\right)^{-1}\tilde{y}_z\right)$$

$\Downarrow$ ETPF $\qquad$ $\Downarrow$ ESRF

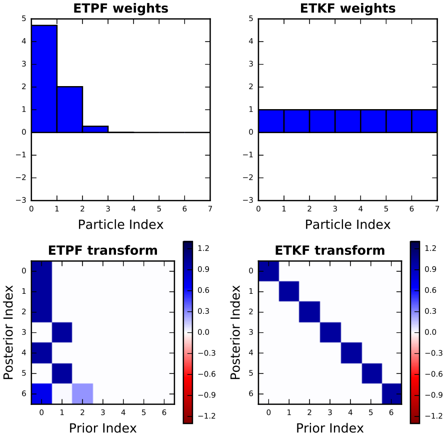
Bridging parameter extrems:

- $\alpha = 0 \Rightarrow$ Pure Kalman Filter
- $\alpha = 1 \Rightarrow$ Pure Particle Filter

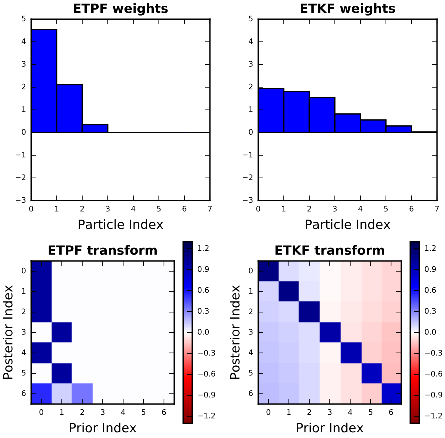
Models explored:

- Single DA step
- Lorenz 63
- Lorenz 96
- Lorenz 96 coupled to a wave equation
- Shallow water equations (2D)
- Modified shallow water equations (1D)

Bridging Parameter = 1

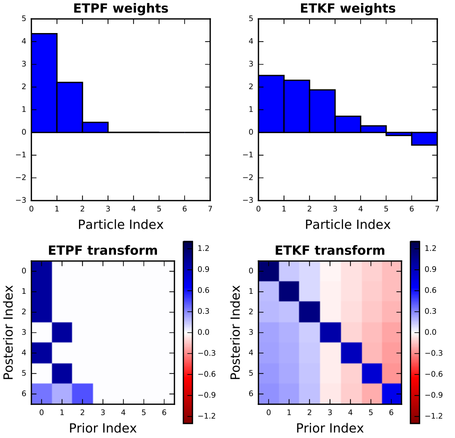


Bridging Parameter = 0.9



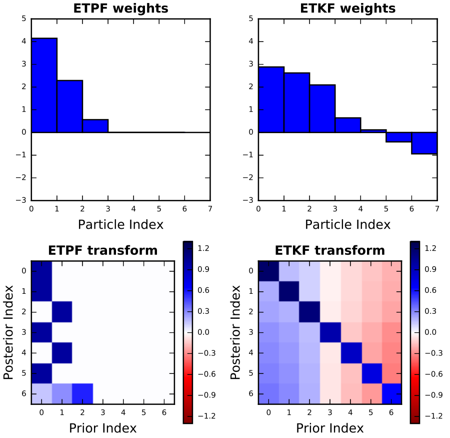
Single one-dimensional DA step

Bridging Parameter = 0.8

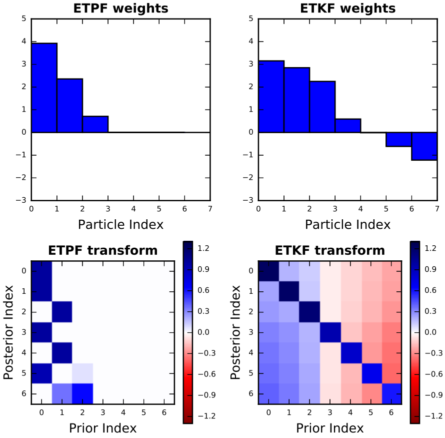


Single one-dimensional DA step

Bridging Parameter = 0.7

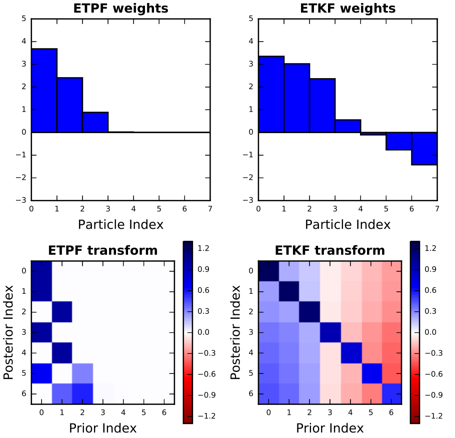


Bridging Parameter = 0.6

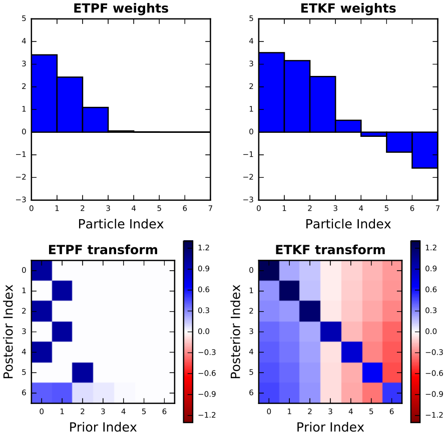


Single one-dimensional DA step

Bridging Parameter = 0.5

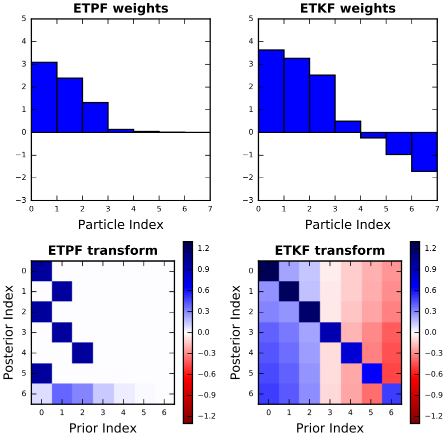


Bridging Parameter = 0.4

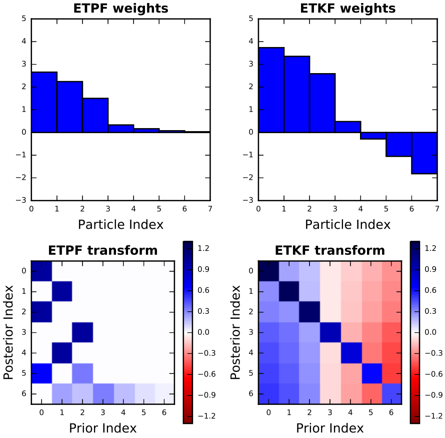


Single one-dimensional DA step

Bridging Parameter = 0.3

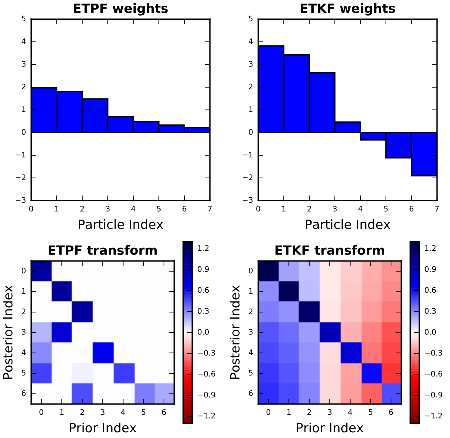


Bridging Parameter = 0.2

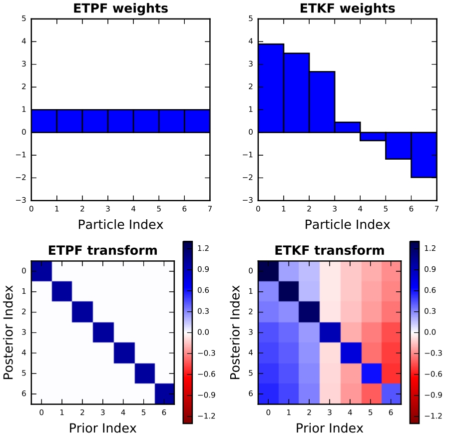


Single one-dimensional DA step

Bridging Parameter = 0.1

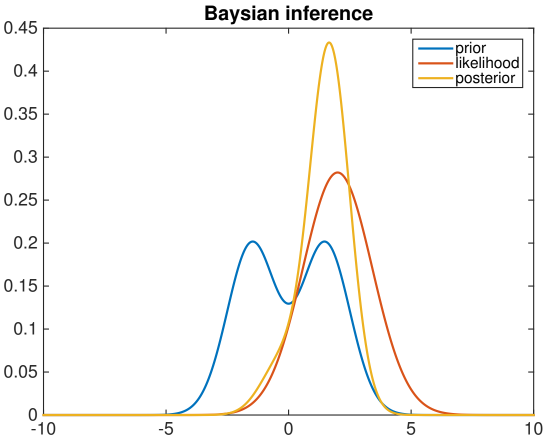


Bridging Parameter = 0



University of Potsdam

Bayesian Inference
for bimodal prior
and
Gaussian likelihood

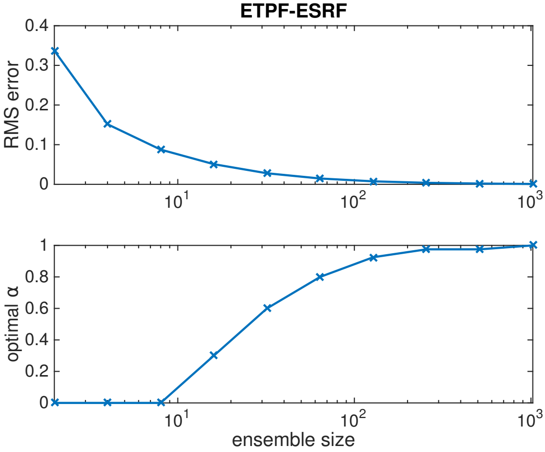


University of Potsdam

ETPF-ESRF

performance vs

ensemble size



Particle rejuvenation:

$$z_j^{\mathrm{a}} \to z_j^{\mathrm{a}} + \sum_{i=1}^{M} (z_i^{\mathrm{f}} - \bar{z}^{\mathrm{f}}) \frac{\beta \Xi_{ij}}{\sqrt{M-1}}$$

- β: Rejuvenation parameter
- $\{\Xi_{ij}\}$: i.i.d. Gaussian random variables with mean zero and variance one

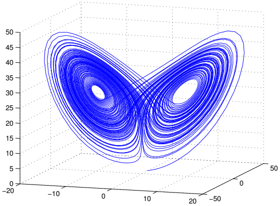
Properties:

- Increase ensemble spread while preserving ensemble mean
- Ameliorates particle degeneracy
- Does not destroy spatial regularity

University of Potsdam

Dynamical system:

$$\dot{x}_1 = 10(x_2 - x_1)$$
$$\dot{x}_2 = x_1(28 - x_3) - x_2$$
$$\dot{x}_3 = x_1 x_2 - 8/3 x_3$$

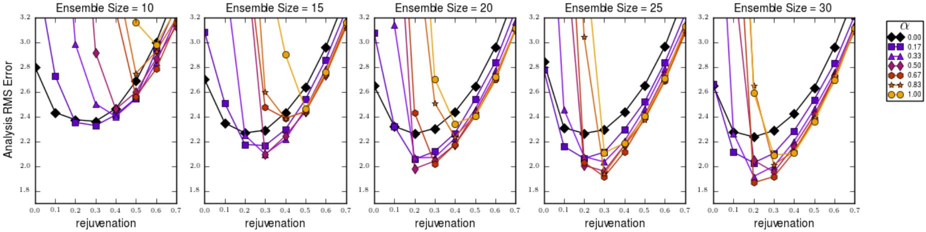


Perfect model DA experiments:

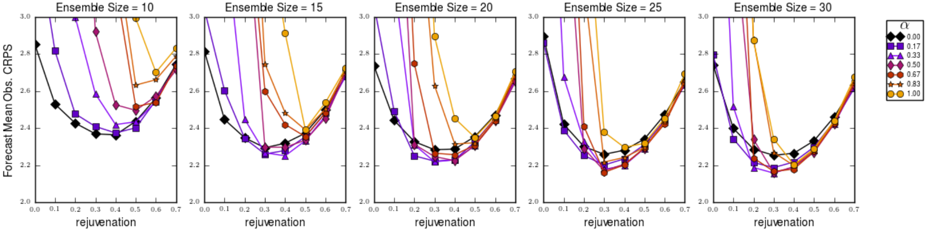
- x_1 observed every 12 time-steps
- Observation error variance $R = 8$

Example: Lorenz 63 model

PF-KF RMSE with exact OT solver



PF-KF CRPS with exact OT solver



University of Potsdam

Possible updating sequences:

- ETPF-ESRF

$$z_j^{\mathrm{h}} = \sum_{i=1}^{M} z_i^{\mathrm{f}} d_{ij}^{\mathrm{PF}}, \qquad d_{ij}^{\mathrm{PF}} := d_{ij}^{\mathrm{PF}}(\alpha, \{z_l^{\mathrm{f}}\}, y_{\mathrm{obs}}),$$

$$z_j^{\mathrm{a}} = \sum_{i=1}^{M} z_i^{\mathrm{h}} d_{ij}^{\mathrm{KF}}, \qquad d_{ij}^{\mathrm{KF}} := d_{ij}^{\mathrm{KF}}(\alpha, \{z_l^{\mathrm{h}}\}, y_{\mathrm{obs}}).$$

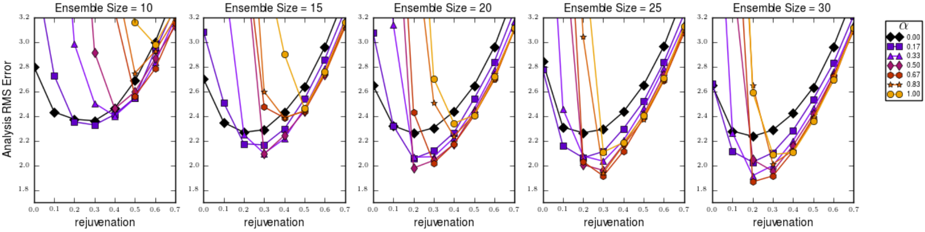
- ESRF-ETPF

$$z_j^{\mathrm{h}} = \sum_{i=1}^{M} z_i^{\mathrm{f}} d_{ij}^{\mathrm{KF}}, \qquad d_{ij}^{\mathrm{KF}} := d_{ij}^{\mathrm{KF}}(\alpha, \{z_l^{\mathrm{f}}\}, y_{\mathrm{obs}}),$$

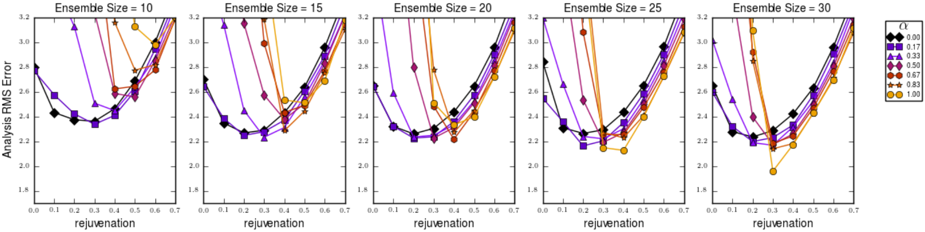
$$z_j^{\mathrm{a}} = \sum_{i=1}^{M} z_i^{\mathrm{h}} d_{ij}^{\mathrm{PF}}, \qquad d_{ij}^{\mathrm{PF}} := d_{ij}^{\mathrm{PF}}(\alpha, \{z_l^{\mathrm{h}}\}, y_{\mathrm{obs}}).$$

Updating order dependency

PF-KF RMSE with exact OT solver

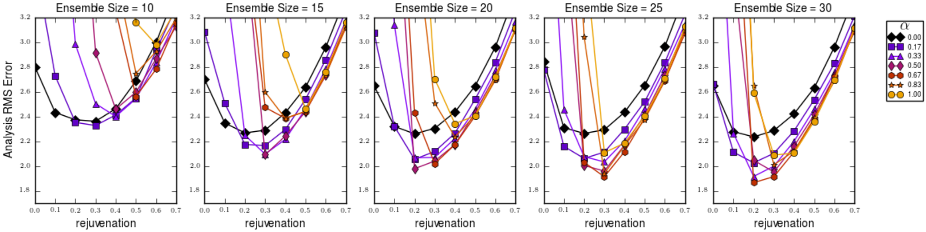


KF-PF RMSE with exact OT solver

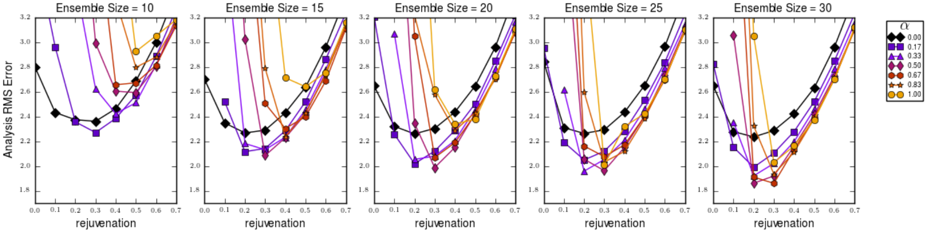


University of Potsdam

PF-KF RMSE with exact OT solver

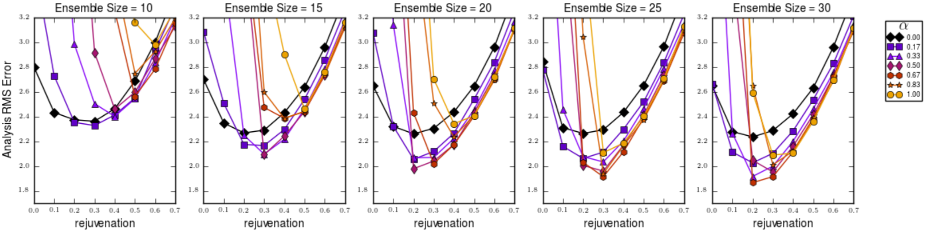


PF-KF RMSE solving entropic regularized OT problem

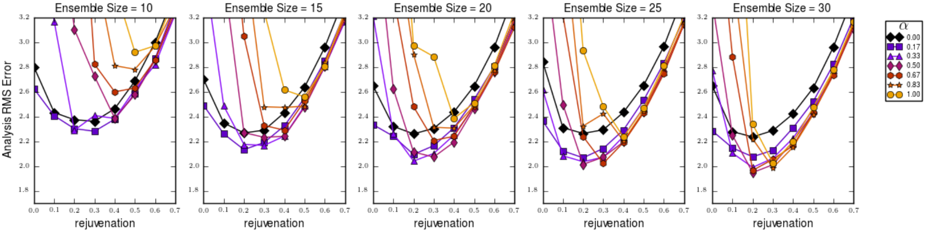


OT solver dependency

PF-KF RMSE with exact OT solver



PF-KF RMSE solving 1D OT problem

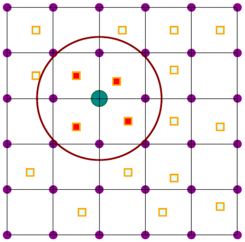


University of Potsdam

R-Localization

$$r_{qq}^{LOC}(x_k) = \frac{r_{qq}}{\rho\left(\frac{\|x_k - x_q\|}{R_{\text{loc}}}\right)},$$

with ρ a compactly supported tempering function.



- Directly applicable to Kalman Filters
- Directly applicable to importance weights

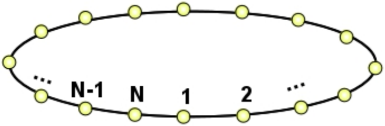
Particle distance Localization [Cheng and Reich, 2015]

$$c_{ij}(x) = \int_{\mathbb{R}} \rho\left(\frac{\|x_k - x_q\|}{C_{\text{loc}}}\right) \|z_i^{\text{f}}(x_q) - z_j^{\text{f}}(x_q)\|^2 dx_q$$

Dynamical system:

$$\dot{x}_j = (x_{j+1} - x_{j-2})x_{j-1} - x_j + F,$$
$$x_j = x_{j+N}$$

where $F = 8$ and $N = 40$.

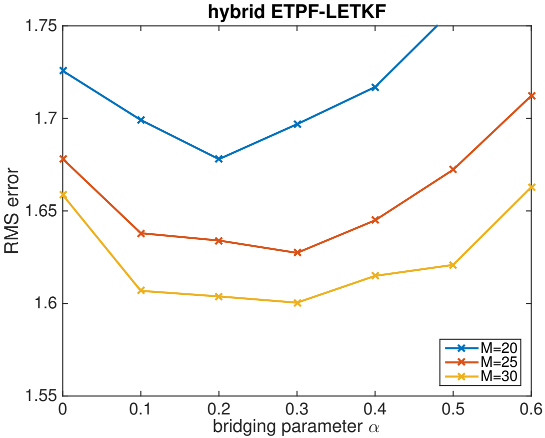


Perfect model DA experiments:

- Odd variables observed every 22 time-steps
- Observation error variance $R = 8$
- Particle rejuvenation $\beta = 0.2$
- Localization radius is $R_{\text{loc}} = 4$
- OTP solved using FastEMD algorithm

University of Potsdam

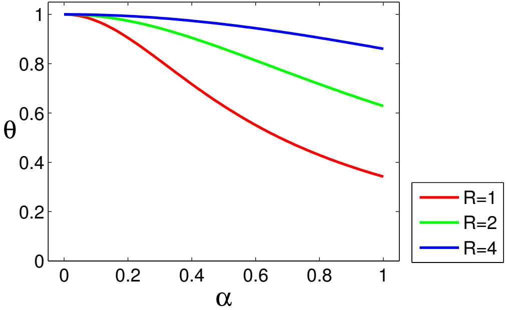
Skill dependence on bridging parameter for different ensemble sizes using fixed likelihood splitting



University of Potsdam

Normalized effective sample size: $\theta(\alpha) = \dfrac{1}{M\sum_{i=1}^{M} w_i(\alpha)^2}$

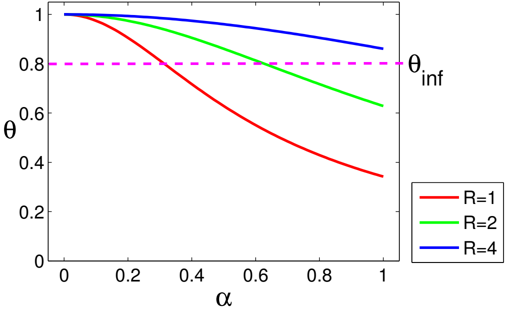
θ vs α for different obs. error levels



University of Potsdam

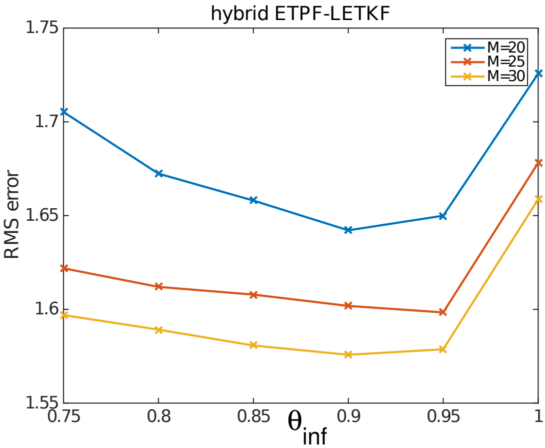
Normalized effective sample size: $\theta(\alpha) = \frac{1}{M\sum_{i=1}^{M} w_i(\alpha)^2}$

θ vs α for different obs. error levels



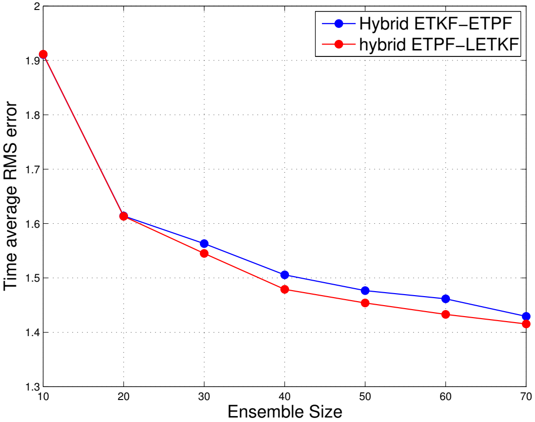
Adaptivity criterion: Set α to the maximal value for which $\theta \geq \theta_{inf}$.

Skill dependence on minimum effective sample size for different ensemble sizes using adaptive likelihood splitting



University of Potsdam

Skill dependence on ensemble size for both update orders using optimal fixed bridging parameter



- Proposed Hybrid scheme
  allows consistent localization for both filters,
  preserves model state regularity,
  outperforms both ETKF and ETPF for a suitable α
- Approximated OT solution approaches are promising cheaper options, in particular 1D approximation
- Update ordering sensitivity is model-dependent
- Adaptive likelihood splitting benefits spatially extended systems.

- ETPF and hybrid scheme being implemented for the German Weather Service (DWD)
- Second order accurate Ensemble Transform Particle Filter (under review)
- 1D OT approximation and then multivariate dependence recovery via Ensemble copula coupling [Schefzik et al., 2013]
- Generation of unbalanced fields through localization currently being addressed via mollification.

Cheng, Y. and Reich, S. (2015).
Nonlinear Data Assimilation, chapter Assimilating data into scientific models: An optimal coupling perspective, pages 75–118.

Chustagulprom, N., Reich, S., and Reinhardt, M. (2016).
A hybrid ensemble transform particle filter for nonlinear and spatially extended dynamical systems.
SIAM/ASA J. Uncertainty Quantification, 4(1):592–608.

Cuturi, M. (2013).
Sinkhorn distances: Lightspeed computation of optimal transport.
In Burges, C., Bottou, L., Welling, M., Ghahramani, Z., and Weinberger, K., editors, *Advances in Neural Information Processing Systems 26*, pages 2292–2300. Curran Associates, Inc.

Frei, M. and Künsch, H. R. (2013).
Bridging the ensemble kalman and particle filters.
Biometrika.

Lei, J. and Bickel, P. (2011).
A moment matching ensemble filter for nonlinear non-gaussian data assimilation.
Mon. Wea. Rev., 139(12):3964–3973.

Pele, O. and Werman, M. (2009).
Fast and robust earth mover's distances.
In *ICCV*.

Poterjoy, J. (2015).
A localized particle filter for high-dimensional nonlinear systems.
Mon. Wea. Rev., 144(1):59–76.

Reich, S. (2013).
A non-parametric ensemble transform method for bayesian inference.
SIAM J Sci Comput, 35:A2013–A2014.

Reich, S. and Cotter, C. (2015).
Probabilistic Forecasting and Bayesian Data Assimilation.
Cambridge University Press.

Schefzik, R., Thorarinsdottir, T. L., and Gneiting, T. (2013).
Uncertainty quantification in complex simulation models using ensemble copula coupling.
pages 616–640.

Tippett, M. K., Anderson, J. L., Bishop, C. H., Hamill, T. M., and Whitaker, J. S. (2003).
Ensemble square root filters.
Monthly Weather Review, 131(7):1485–1490.

Thanks!