Localization and hybridization of sample covariances

15 December 2015

Benjamin Ménétrier
Introduction

Context:

• Data assimilation often relies on forecast error covariances.
• This matrix can be sampled from an ensemble of forecasts.
• Sampling noise arises because of the limited ensemble size.
• Question: how to tackle this sampling error?

Usual methods:
• Covariance localization → tapering with a localization matrix
• Covariance hybridization → linear combination with a static covariance matrix
Introduction

Context:

- Data assimilation often relies on forecast error covariances.
Introduction

Context:

- Data assimilation often relies on forecast error covariances.
- This matrix can be sampled from an ensemble of forecasts.
Introduction

Context:

• Data assimilation often relies on forecast error covariances.
• This matrix can be sampled from an ensemble of forecasts.
• Sampling noise arises because of the limited ensemble size.
Introduction

Context:

- Data assimilation often relies on forecast error covariances.
- This matrix can be sampled from an ensemble of forecasts.
- Sampling noise arises because of the limited ensemble size.
- Question: how to tackle this sampling error?
Introduction

Context:

• Data assimilation often relies on forecast error covariances.
• This matrix can be sampled from an ensemble of forecasts.
• Sampling noise arises because of the limited ensemble size.
• Question: how to tackle this sampling error?

Usual methods:
Introduction

Context:

- Data assimilation often relies on forecast error covariances.
- This matrix can be sampled from an ensemble of forecasts.
- Sampling noise arises because of the limited ensemble size.
- Question: how to tackle this sampling error?

Usual methods:

- Covariance localization
  → tapering with a localization matrix
Introduction

Context:

• Data assimilation often relies on forecast error covariances.
• This matrix can be sampled from an ensemble of forecasts.
• Sampling noise arises because of the limited ensemble size.
• Question: how to tackle this sampling error?

Usual methods:

• Covariance localization
  → tapering with a localization matrix
• Covariance hybridization
  → linear combination with a static covariance matrix
Introduction

Questions:
Introduction

Questions:

1. Can we compute an optimized localization with a method:
Introduction

Questions:

1. Can we compute an optimized localization with a method:
   - using data from the \textit{ensemble only},
Questions:

1. Can we compute an optimized localization with a method:
   - using data from the ensemble only,
   - affordable for high-dimensional systems.
Introduction

Questions:

1. Can we compute an optimized localization with a method:
   - using data from the ensemble only,
   - affordable for high-dimensional systems.

2. Can localization and hybridization be considered together, and optimized simultaneously?
Questions:

1. Can we compute an optimized localization with a method:
   - using data from the ensemble only,
   - affordable for high-dimensional systems.

2. Can localization and hybridization be considered together, and optimized simultaneously?

3. Is hybridization always improving the accuracy of forecast error covariances?
Covariance sampling

An ensemble of $N$ forecasts $\{\tilde{x}_p^b\}$ is used to sample $\tilde{B}$:

$$\tilde{B} = \frac{1}{N-1} \sum_{p=1}^{N} \delta \tilde{x}^b (\delta \tilde{x}^b)^T$$

where:

$$\delta \tilde{x}_p^b = \tilde{x}_p^b - \langle \tilde{x}^b \rangle \quad \text{and} \quad \langle \tilde{x}^b \rangle = \frac{1}{N} \sum_{p=1}^{N} \tilde{x}_p^b$$
Covariance sampling

An ensemble of $N$ forecasts $\{\tilde{x}_p^b\}$ is used to sample $\tilde{B}$:

$$\tilde{B} = \frac{1}{N-1} \sum_{p=1}^{N} \delta\tilde{x}_p^b (\delta\tilde{x}_p^b)^T$$

where:

$$\delta\tilde{x}_p^b = \tilde{x}_p^b - \langle \tilde{x}_p^b \rangle \quad \text{and} \quad \langle \tilde{x}_p^b \rangle = \frac{1}{N} \sum_{p=1}^{N} \tilde{x}_p^b$$

Asymptotic behavior: if $N \to \infty$, then $\tilde{B} \to \tilde{B}^*$
Covariance sampling

An ensemble of \( N \) forecasts \( \{ \tilde{x}_p^b \} \) is used to sample \( \tilde{B} \):

\[
\tilde{B} = \frac{1}{N-1} \sum_{p=1}^{N} \delta \tilde{x}_p^b (\delta \tilde{x}_p^b)^T
\]

where:

\[
\delta \tilde{x}_p^b = \tilde{x}_p^b - \langle \tilde{x}^b \rangle \quad \text{and} \quad \langle \tilde{x}^b \rangle = \frac{1}{N} \sum_{p=1}^{N} \tilde{x}_p^b
\]

Asymptotic behavior: if \( N \to \infty \), then \( \tilde{B} \to \tilde{B}^* \)

In practice, \( N < \infty \) \( \Rightarrow \) sampling noise \( \tilde{B}^e = \tilde{B} - \tilde{B}^* \)
Covariance sampling

Randomization of $\mathbf{B}$:

$$\delta \tilde{\mathbf{x}}_p^b = \mathbf{U} \eta_p \quad \text{with} \quad \mathbf{U} \mathbf{U}^T = \mathbf{B} \quad \text{and} \quad \eta_p \sim \mathcal{N}(0, I).$$
Covariance sampling

Specified $B$

Randomized perturbations

Sampled $B$ ($N = 30$)

Randomization of $B$:

$$\tilde{x}_p^b = U \eta_p \quad \text{with} \quad UU^T = B \quad \text{and} \quad \eta_p \sim \mathcal{N}(0, I).$$
Covariance sampling

Specified $B$

Randomized perturbations

Sampled $B$ ($N = 50$)

Randomization of $B$:

$$\delta \tilde{x}_p^b = U \eta_p \quad \text{with} \quad U U^T = B \quad \text{and} \quad \eta_p \sim \mathcal{N}(0, I).$$
Covariance sampling

Randomization of $\mathbf{B}$:

$$\tilde{\delta x}_p = \mathbf{U} \eta_p$$

with

$$\mathbf{U} \mathbf{U}^T = \mathbf{B}$$

and

$$\eta_p \sim \mathcal{N}(0, \mathbf{I}).$$
Covariance sampling

Specified $B$

Randomized perturbations

Sampled $B$ ($N = 200$)

Randomization of $B$

\[
\tilde{\delta x}_p^b = U \eta_p \quad \text{with} \quad UU^T = B \quad \text{and} \quad \eta_p \sim \mathcal{N}(0, I).
\]
Covariance sampling

Specified $B$

Randomized perturbations

Sampled $B$ ($N = 500$)

Randomization of $B$:

$$\delta x_p^b = U \eta_p \quad \text{with} \quad UU^T = B \quad \text{and} \quad \eta_p \sim \mathcal{N}(0, I).$$
Covariance sampling

Randomization of $B$:

$$\tilde{\delta x}^b_p = U \eta_p$$
with

$$UU^T = B$$
and

$$\eta_p \sim \mathcal{N}(0, I).$$
Outline

Introduction

Objectively optimized localization

Practical computation

Jointly optimized localization and hybridization

Conclusions
<table>
<thead>
<tr>
<th>Outline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
</tr>
<tr>
<td>Objectively optimized localization</td>
</tr>
<tr>
<td>Practical computation</td>
</tr>
<tr>
<td>Jointly optimized localization and hybridization</td>
</tr>
<tr>
<td>Conclusions</td>
</tr>
</tbody>
</table>
Covariance localization

Localization = Schur product with a localization matrix $L$:

$$\hat{B} = L \odot \tilde{B} \iff \hat{B}_{ij} = L_{ij} \tilde{B}_{ij}$$
Covariance localization

Localization = Schur product with a localization matrix $L$:

$$\hat{B} = L \circ \tilde{B} \iff \hat{B}_{ij} = L_{ij} \tilde{B}_{ij}$$

Sampled $B$ (N = 10)  No localization  Localized $B$
Covariance localization

Localization = Schur product with a localization matrix $L$

$$\hat{B} = L \odot \tilde{B} \iff \hat{B}_{ij} = L_{ij} \tilde{B}_{ij}$$

Sampled $B$ ($N = 10$)  Localization ($L = 0.4$)  Localized $B$

Start reducing the sampling noise...
Covariance localization

Localization = Schur product with a localization matrix $L$:

$$
\hat{B} = L \circ \tilde{B} \quad \Leftrightarrow \quad \hat{B}_{ij} = L_{ij} \tilde{B}_{ij}
$$

Sampled $B$ ($N = 10$)  Localization ($L = 0.2$)  Localized $B$

Less and less sampling noise...
Covariance localization

Localization = Schur product with a localization matrix $L$:

$$\hat{B} = L \circ \tilde{B} \Leftrightarrow \hat{B}_{ij} = L_{ij} \tilde{B}_{ij}$$

Sampled $B$ ($N = 10$)  Localization ($L = 0.1$)  Localized $B$

Good! Almost no sampling noise anymore...
Covariance localization

Localization = Schur product with a localization matrix \( L \):

\[
\hat{B} = L \circ \tilde{B} \iff \hat{B}_{ij} = L_{ij} \tilde{B}_{ij}
\]

Sampled \( B \) (N = 10)  Localization (L = 0.05)  Localized \( B \)

Well, we are some loosing signal now...
Covariance localization

Localization = Schur product with a localization matrix $L$:

$$\hat{B} = L \circ \tilde{B} \iff \hat{B}_{ij} = L_{ij} \tilde{B}_{ij}$$

Sampled $B$ ($N = 10$)  Localization ($L = 0.025$)  Localized $B$

Hey, stop loosing signal!
Covariance localization

Localization = Schur product with a localization matrix $L$:

$$\hat{B} = L \circ \tilde{B} \iff \hat{B}_{ij} = L_{ij} \tilde{B}_{ij}$$

Sampled $B$ (N = 10)  Localization (L = 0.01)  Localized $B$

No more signal!
Covariance localization

Question: what is the optimal localization?
Covariance localization

Question: what is the optimal localization?

A possible definition: the optimal localization should minimize

\[ e = \mathbb{E} \left[ \| L \circ \tilde{B} - \tilde{B}^\star \|_F^2 \right] \]

where:

- \( \hat{B} = L \circ \tilde{B} \) is the localized covariance matrix
- \( \tilde{B}^\star = \lim_{N \to \infty} \tilde{B} \) is the target
- \( \| \cdot \|_F \) is the Frobenius norm
Covariance localization

Question: what is the optimal localization?

A possible definition: the optimal localization should minimize

\[ e = \mathbb{E} \left[ \| L \circ \tilde{B} - \tilde{B}^\star \|_F^2 \right] \]

where:

- \( \hat{B} = L \circ \tilde{B} \) is the localized covariance matrix
- \( \tilde{B}^\star = \lim_{N \to \infty} \tilde{B} \) is the target
- \( \| \cdot \|_F \) is the Frobenius norm

Set the partial derivatives of \( e \) to zero:

\[ \frac{\partial e}{\partial L_{ij}} = 0 \iff L_{ij} = \frac{\mathbb{E} \left[ \tilde{B}_{ij} \tilde{B}_{ij}^\star \right]}{\mathbb{E} \left[ \tilde{B}_{ij}^2 \right]} \]
Covariance localization

Question: what is the optimal localization?

A possible definition: the optimal localization should minimize

$$ e = \mathbb{E} \left[ \left\| L \circ \tilde{B} - \tilde{B}^* \right\|^2_F \right] $$

where:

- $\hat{B} = L \circ \tilde{B}$ is the localized covariance matrix
- $\tilde{B}^* = \lim_{N \to \infty} \tilde{B}$ is the target
- $\| \cdot \|_F$ is the Frobenius norm

Set the partial derivatives of $e$ to zero:

$$ \frac{\partial e}{\partial L_{ij}} = 0 \iff L_{ij} = \frac{\mathbb{E} \left[ \tilde{B}_{ij} \tilde{B}_{ij}^* \right]}{\mathbb{E} \left[ \tilde{B}_{ij}^2 \right]} \iff \text{OK} $$
Sampling noise properties

Homogeneous variance / length-scale

Asymptotic B ($N \to \infty$)  Sampled B ($N = 10$)  Sampling noise
Sampling noise properties

Heterogeneous variance / homogeneous length-scale

Asymptotic $B (N \to \infty)$  
Sampled $B (N = 10)$  
Sampling noise

Sampling noise amplitude related to the asymptotic variance
Sampling noise properties

Homogeneous variance / length-scale

Asymptotic B (N → ∞)  Sampled B (N = 10)  Sampling noise
Sampling noise properties

Homogeneous variance / heterogeneous length-scale

Asymptotic B (N → ∞)  Sampled B (N = 10)  Sampling noise

Sampling noise length-scale related to the asymptotic length-scale
Theory of sampling noise

Intuitive conclusion: statistical properties of the sampling noise are functions of the asymptotic covariance matrix.
Theory of sampling noise

Intuitive conclusion: statistical properties of the sampling noise are functions of the asymptotic covariance matrix.

Expectation of sample covariance: \( \mathbb{E}\left[ \tilde{B}_{ij} \right] = \mathbb{E}\left[ \tilde{B}_{ij}^* \right] \)
Theory of sampling noise

Intuitive conclusion: statistical properties of the sampling noise are functions of the asymptotic covariance matrix.

 Expectation of sample covariance: \( \mathbb{E}[\tilde{B}_{ij}] = \mathbb{E}[\tilde{B}^*_{ij}] \)

 Expectation of a product of sample covariances:

\[
\mathbb{E}[\tilde{B}_{ij} \tilde{B}_{kl}] = \frac{1}{N} \mathbb{E}[\tilde{\Xi}^*_{ijkl}] + \frac{N - 1}{N} \mathbb{E}[\tilde{B}^*_{ij} \tilde{B}^*_{kl}] \\
+ \frac{1}{N(N-1)} \left( \mathbb{E}[\tilde{B}^*_{ik} \tilde{B}^*_{jl}] + \mathbb{E}[\tilde{B}^*_{il} \tilde{B}^*_{jk}] \right)
\]

involving:

- the ensemble size \( N \),
- the asymptotic covariance \( \tilde{B}^* \),
- the asymptotic fourth-order centered moment \( \tilde{\Xi}^* \).
Theory of sampling noise

Expectation of the sample fourth-order centered moment $\tilde{\Xi}_{ijkl}$:

$$
\mathbb{E}\left[\tilde{\Xi}_{ijkl}\right] = \frac{(N - 1)(N^2 - 3N + 3)}{N^3} \mathbb{E}\left[\tilde{\Xi}_{ijkl}^*\right] + \frac{(N - 1)(2N - 3)}{N^3} \left( \mathbb{E}\left[\tilde{B}_{ij}^* \tilde{B}_{kl}^*\right] + \mathbb{E}\left[\tilde{B}_{ik}^* \tilde{B}_{jl}^*\right] + \mathbb{E}\left[\tilde{B}_{il}^* \tilde{B}_{jk}^*\right] \right)
$$
Theory of sampling noise

Expectation of the sample fourth-order centered moment $\tilde{\Xi}_{ijkl}$:

$$\mathbb{E}\left[\tilde{\Xi}_{ijkl}\right] = \frac{(N - 1)(N^2 - 3N + 3)}{N^3} \mathbb{E}\left[\tilde{\Xi}_{ijkl}^*\right] + \frac{(N - 1)(2N - 3)}{N^3} \left( \mathbb{E}\left[\tilde{B}^*_ij \tilde{B}^*_kl\right] + \mathbb{E}\left[\tilde{B}^*_ik \tilde{B}^*_jl\right] + \mathbb{E}\left[\tilde{B}^*_il \tilde{B}^*_jk\right] \right)$$

Mix everything and shake well to get:

$$\mathbb{E}\left[\tilde{B}^*_ij \tilde{B}^*_kl\right] = \frac{(N - 1)(N^2 - 3N + 1)}{N(N - 2)(N - 3)} \mathbb{E}\left[\tilde{B}^*_ij \tilde{B}^*_kl\right]$$

$$+ \frac{N - 1}{N(N - 2)(N - 3)} \left( \mathbb{E}\left[\tilde{B}^*_ik \tilde{B}^*_jl\right] + \mathbb{E}\left[\tilde{B}^*_il \tilde{B}^*_jk\right] \right)$$

$$- \frac{N}{(N - 2)(N - 3)} \mathbb{E}\left[\tilde{\Xi}_{ijkl}\right]$$
Localization: general case

Only assumption, independence of the sampling error and the asymptotic covariances:

\[ \mathbb{E} \left[ \left( \tilde{B}_{ij} - \tilde{B}_{ij}^* \right) \tilde{B}_{kl}^* \right] = 0 \iff \mathbb{E} \left[ \tilde{B}_{ij} \tilde{B}_{kl}^* \right] = \mathbb{E} \left[ \tilde{B}_{ij}^* \tilde{B}_{kl}^* \right] \]
Localization: general case

**Only assumption**, independence of the sampling error and the asymptotic covariances:

\[ \mathbb{E} \left[ \left( \tilde{B}_{ij} - \tilde{B}_{ij}^* \right) \tilde{B}_{kl}^* \right] = 0 \Leftrightarrow \mathbb{E} \left[ \tilde{B}_{ij} \tilde{B}_{kl}^* \right] = \mathbb{E} \left[ \tilde{B}_{ij}^* \tilde{B}_{kl}^* \right] \]

Final result for the localization:

\[
L_{ij} = \frac{\mathbb{E} \left[ \tilde{B}_{ij} \tilde{B}_{ij}^* \right]}{\mathbb{E} \left[ \tilde{B}_{ij}^2 \right]} = \frac{(N - 1)^2}{N(N - 3)} + \frac{N - 1}{N(N - 2)(N - 3)} \frac{\mathbb{E} \left[ \tilde{B}_{ii} \tilde{B}_{jj} \right]}{\mathbb{E} \left[ \tilde{B}_{ij}^2 \right]} - \frac{N}{(N - 2)(N - 3)} \frac{\mathbb{E} \left[ \tilde{\Xi}_{ijij} \right]}{\mathbb{E} \left[ \tilde{B}_{ij}^2 \right]} \]

Optimal localization is expressed without asymptotic quantities!
Localization: general case

**Only assumption**, independence of the sampling error and the asymptotic covariances:

$$
\mathbb{E}\left[ \left( \tilde{B}_{ij} - \tilde{B}_{ij}^* \right) \tilde{B}_{kl}^* \right] = 0 \iff \mathbb{E}\left[ \tilde{B}_{ij} \tilde{B}_{kl}^* \right] = \mathbb{E}\left[ \tilde{B}_{ij}^* \tilde{B}_{kl}^* \right]
$$

Final result for the localization:

$$
L_{ij} = \frac{\mathbb{E}\left[ \tilde{B}_{ij} \tilde{B}_{ij}^* \right]}{\mathbb{E}\left[ \tilde{B}_{ij}^2 \right]} = \frac{(N - 1)^2}{N(N - 3)} + \frac{N - 1}{N(N - 2)(N - 3)} \frac{\mathbb{E}\left[ \tilde{B}_{ii} \tilde{B}_{jj} \right]}{\mathbb{E}\left[ \tilde{B}_{ij}^2 \right]}
$$

$$
- \frac{N}{(N - 2)(N - 3)} \frac{\mathbb{E}\left[ \tilde{\Xi}_{ijij} \right]}{\mathbb{E}\left[ \tilde{B}_{ij}^2 \right]}
$$

Optimal localization is expressed without asymptotic quantities!
Localization: Gaussian case

In the case of a Gaussian distributed ensemble, possible simplification

\[ L_{Gij} = \frac{(N - 1)^2}{2} \left( \frac{(N - 2)(N + 1)}{N - 1} \right) - \frac{N - 1}{(N - 2)(N + 1)} \left( \mathbb{E}[\tilde{B}_{ii}\tilde{B}_{jj}] - \mathbb{E}[\tilde{B}_{ij}^2] \right) \]

No fourth-order moment anymore!
Localization: Gaussian case

In the case of a Gaussian distributed ensemble, possible simplification

The Wick-Isserlis theorem gives:

\[ \tilde{\Xi}^{*}_{ijkl} = \tilde{B}^{*}_{ij}\tilde{B}^{*}_{kl} + \tilde{B}^{*}_{ik}\tilde{B}^{*}_{jl} + \tilde{B}^{*}_{il}\tilde{B}^{*}_{jk} \]
Localization: Gaussian case

In the case of a Gaussian distributed ensemble, possible simplification

The Wick-Isserlis theorem gives:

\[ \tilde{\Xi}_{ijkl}^* = \tilde{B}_{ij}^* \tilde{B}_{kl}^* + \tilde{B}_{ik}^* \tilde{B}_{jl}^* + \tilde{B}_{il}^* \tilde{B}_{jk}^* \]

The optimal localization can be simplified into:

\[ L_{ij}^G = \frac{(N - 1)^2}{(N - 2)(N + 1)} - \frac{N - 1}{(N - 2)(N + 1)} \frac{\mathbb{E}[\tilde{B}_{ii} \tilde{B}_{jj}]}{\mathbb{E}[\tilde{B}_{ij}^2]} \]

No fourth-order moment anymore!
### Outline

<table>
<thead>
<tr>
<th>Introduction</th>
<th>Localization</th>
<th>Computation</th>
<th>Hybridization</th>
<th>Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>Objectively optimized localization</td>
<td>Practical computation</td>
<td>Jointly optimized localization and hybridization</td>
<td>Conclusions</td>
</tr>
</tbody>
</table>

---

**Introduction**

Objectively optimized localization

**Practical computation**

Jointly optimized localization and hybridization

**Conclusions**
An ergodicity assumption is required to estimate the statistical expectations $\mathbb{E}[\cdot]$ in practice:

Estimation of correlation and localization (30 members, 5 m temperature)
Practical computation

An ergodicity assumption is required to estimate the statistical expectations $\mathbb{E}[\cdot]$ in practice:

Estimation of correlation and localization (30 members, 5 m temperature)
Practical computation

An ergodicity assumption is required to estimate the statistical expectations $\mathbb{E}[\cdot]$ in practice:

Estimation of correlation and localization (30 members, 5 m temperature)
An ergodicity assumption is required to estimate the statistical expectations $\mathbb{E}[\cdot]$ in practice:

Estimation of correlation and localization (30 members, 5 m temperature)
Practical computation

An ergodicity assumption is required to estimate the statistical expectations $\mathbb{E}[\cdot]$ in practice:

Estimation of correlation and localization (30 members, 5 m temperature)
An ergodicity assumption is required to estimate the statistical expectations $\mathbb{E}[\cdot]$ in practice:

Estimation of correlation and localization (30 members, 5 m temperature)
Practical computation

An ergodicity assumption is required to estimate the statistical expectations $\mathbb{E}[\cdot]$ in practice:

Estimation of correlation and localization (30 members, 5 m temperature)
Practical computation

An ergodicity assumption is required to estimate the statistical expectations $\mathbb{E}[\cdot]$ in practice:

Estimation of correlation and localization (30 members, 5 m temperature)
Homogeneity issue

Random points are drawn from the coordinates vector. This sampling should take the grid structure into account.
Localization length-scale regional diagnostic

Sampling

Localization (km)

Global

Latitude bands
Isotropy issue

Global random sampling for the class around 1350 km
Isotropy issue

Regional random sampling ($-40^\circ$ $-$ $+40^\circ$) for 1350 km
Isotropy issue

Regional random sampling \((-20^\circ - +20^\circ\) for 1350 km)
Regional random sampling ($-10^\circ - +10^\circ$) for 1350 km
Isotropy issue

Regional random sampling \((-10^\circ \rightarrow +10^\circ\)\) for 1350 km

The shape of the domain can influence the azimuth distribution for large separations (no impact for small separations).
Diagnostic robustness

Localization accuracy depends on the sampling size \((n_s)\):

Localization statistics over different sets of sampling
Every quantity depends on the ensemble size:

\[
L = \frac{(N-1)^2}{N(N-3)} + \frac{N-1}{N(N-2)(N-3)} \frac{\mu[\tilde{v}^2]}{\mu[\tilde{B}^2]} - \frac{N}{(N-2)(N-3)} \frac{\mu[\tilde{\Xi}]}{\mu[\tilde{B}^2]}
\]

**Correlation (black) and localization (colors)**

**Ensemble size:**
- Red: 10
- Orange: 30
- Green: 50
Outline

Introduction

Objectively optimized localization

Practical computation

Jointly optimized localization and hybridization

Conclusions
From localization to hybridization

Localization by $\mathbf{L}$ (Schur product):

$$\hat{\mathbf{B}} = \mathbf{L} \circ \tilde{\mathbf{B}}$$
From localization to hybridization

Localization by $\mathbf{L}$ (Schur product):

$$\hat{\mathbf{B}} = \mathbf{L} \circ \tilde{\mathbf{B}}$$

Localization by $\mathbf{L} +$ hybridization with $\overline{\mathbf{B}}$:

$$\hat{\mathbf{B}}^h = \beta^{e2} \mathbf{L}' \circ \tilde{\mathbf{B}} + \beta^{c2} \overline{\mathbf{B}}$$
From localization to hybridization

Localization by $\mathbf{L}$ (Schur product):

$$\hat{\mathbf{B}} = \mathbf{L} \circ \tilde{\mathbf{B}}$$

Localization by $\mathbf{L} +$ hybridization with $\overline{\mathbf{B}}$:

$$\hat{\mathbf{B}}^h = \beta^{e2} \mathbf{L}' \circ \tilde{\mathbf{B}} + \beta^{c2} \overline{\mathbf{B}}$$

Gain $\mathbf{L}^h$ Offset $\beta^{c2}$
From localization to hybridization

Localization by $\mathbf{L}$ (Schur product):

$$\hat{\mathbf{B}} = \mathbf{L} \circ \tilde{\mathbf{B}}$$

Localization by $\mathbf{L}$ + hybridization with $\bar{\mathbf{B}}$:

$$\hat{\mathbf{B}}^h = \beta e^2 \mathbf{L}' \circ \tilde{\mathbf{B}} + \beta c^2 \bar{\mathbf{B}}$$

Gain $\mathbf{L}^h$ and Offset $\beta c^2$ have to be optimized together

Localization + hybridization = linear filtering of $\tilde{\mathbf{B}}$

$L^h$ and $\beta c^2$ have to be optimized together
From localization to hybridization

Localization by $L$ (Schur product):

$$\hat{B} = L \circ \tilde{B}$$

Localization by $L$ + hybridization with $\overline{B}$:

$$\hat{B}^h = \beta e^{2} L' \circ \tilde{B} + \beta c^2 \overline{B}$$

Gain $L^h$ and Offset $\beta c^2$ have to be optimized together.

Localization + hybridization = linear filtering of $\tilde{B}$

$L^h$ and $\beta c^2$ have to be optimized together.

The optimal localization/hybridization should minimize

$$e^h = \mathbb{E} \left[ \| L^h \circ \tilde{B} + \beta c^2 \overline{B} - \tilde{B}^* \|_F^2 \right]$$
Joint optimization of localization and hybridization

At optimality:

\[
\frac{\partial e^h}{\partial L^h_{ij}} = 0 \quad \text{and} \quad \frac{\partial e^h}{\partial \beta^{c2}} = 0
\]
Joint optimization of localization and hybridization

At optimality:

$$\frac{\partial e^h}{\partial L^h_{ij}} = 0 \quad \text{and} \quad \frac{\partial e^h}{\partial \beta^{c2}} = 0$$

With the same assumptions as before:

$$\beta^{c2} = \frac{\sum_{ij} \left( 1 - L_{ij} \right) \mathbb{E} \left[ \tilde{B}_{ij} \right] \bar{B}_{ij}}{\sum_{ij} \frac{\text{Var} \left[ \tilde{B}_{ij} \right]}{\mathbb{E} \left[ \tilde{B}^2_{ij} \right]} \bar{B}^2_{ij}}$$

and

$$L^h_{ij} = L_{ij} - \frac{\mathbb{E} \left[ \tilde{B}_{ij} \right]}{\mathbb{E} \left[ \tilde{B}^2_{ij} \right]} \beta^{c2} \bar{B}_{ij}$$

where $L_{ij}$ is the localization optimized alone.
Hybridization benefits

Comparison of:

- $\hat{B} = L \circ \tilde{B}$, with an optimal $L$ minimizing $e$
- $\hat{B}^h = L^h \circ \tilde{B} + \beta^{c2} \tilde{B}$, with optimal $L^h$ and $\beta^{c2}$ minimizing $e^h$
Hybridization benefits

Comparison of:

- $\hat{B} = L \circ \tilde{B}$, with an optimal $L$ minimizing $e$
- $\hat{B}^h = L^h \circ \tilde{B} + \beta c^2 \bar{B}$, with optimal $L^h$ and $\beta c^2$ minimizing $e^h$

We can show that:

$$e^h_{\text{opt}} - e_{\text{opt}} = - \sum_{ij} \frac{\text{Var}(\tilde{B}_{ij})}{\mathbb{E}[\tilde{B}_{ij}^2]} \left( \beta c^2 \right)^2 \bar{B}_{ij}^2 \leq 0$$
Hybridization benefits

Comparison of:

- $\hat{B} = L \circ \tilde{B}$, with an optimal $L$ minimizing $e$
- $\hat{B}^h = L^h \circ \tilde{B} + \beta c^2 \overline{B}$, with optimal $L^h$ and $\beta c^2$ minimizing $e^h$

We can show that:

$$e^h_{\text{opt}} - e_{\text{opt}} = - \sum_{ij} \frac{\text{Var}\left(\tilde{B}_{ij}\right)}{\mathbb{E}\left[\tilde{B}_{ij}^2\right]} \left(\beta c^2\right)^2 \overline{B}_{ij}^2 \leq 0$$

With optimal parameters, whatever the static $\overline{B}$: Localization + hybridization is better than localization alone.
As expected:

- $\beta_e^2$ increases with the ensemble size,
- $\beta_c^2$ decreases with the ensemble size.
Localization transformation

Localization optimized alone:

\[ L = \frac{(N - 1)^2}{N(N - 3)} + \frac{N - 1}{N(N - 2)(N - 3)} \frac{\mu[\tilde{\nu}^2]}{\mu[\tilde{B}^2]} - \frac{N}{(N - 2)(N - 3)} \frac{\mu[\tilde{\Xi}]}{\mu[\tilde{B}^2]} \]

Ensemble size:
- Red: 10
- Orange: 30
- Green: 50

Correlation (black) and localization (colors)
Localization adaptation to the hybridization:

\[ L^h = L - \frac{\mu [\tilde{B}]}{\mu [\tilde{B}^2]} \beta^2 \mu [\tilde{B}] \]
Autocorrelation function in North-Atlantic

- the "truth" from which the ensemble is generated,
- the static covariance matrix,
- the localized-only sample covariance matrix,
- the localized and hybridized sample covariance matrix.
Autocorrelation function in North-Atlantic

- the "truth" from which the ensemble is generated,
- the static covariance matrix,
- the localized-only sample covariance matrix,
- the localized and hybridized sample covariance matrix.
Autocorrelation function in North-Atlantic

- the "truth" from which the ensemble is generated,
- the static covariance matrix,
- the localized-only sample covariance matrix,
- the localized and hybridized sample covariance matrix.
Outline

Introduction

Objectively optimized localization

Practical computation

Jointly optimized localization and hybridization

Conclusions
Conclusions

1. Localization can be optimized with a method:
Conclusions

1. Localization can be optimized with a method:
   - based on properties of the ensemble only,
Conclusions

1. Localization can be optimized with a method:
   - based on properties of the ensemble only,
   - affordable for high-dimensional systems,
Conclusions

1. Localization can be optimized with a method:
   - based on properties of the ensemble only,
   - affordable for high-dimensional systems,
   - tackling the sampling noise issue only.
Conclusions

1. Localization can be optimized with a method:
   - based on properties of the ensemble only,
   - affordable for high-dimensional systems,
   - tackling the sampling noise issue only.

2. Localization and hybridization are two joint aspects of the linear filtering of sample covariances and can be optimized simultaneously.
1. Localization can be optimized with a method:
   - based on properties of the ensemble only,
   - affordable for high-dimensional systems,
   - tackling the sampling noise issue only.

2. Localization and hybridization are two joint aspects of the linear filtering of sample covariances and can be optimized simultaneously.

3. If done optimally, hybridization always improves the accuracy of forecast error covariances.

Conclusions
Conclusions

1. Localization can be optimized with a method:
   - based on properties of the ensemble only,
   - affordable for high-dimensional systems,
   - tackling the sampling noise issue only.

2. Localization and hybridization are two joint aspects of the linear filtering of sample covariances and can be optimized simultaneously.

3. If done optimally, hybridization always improves the accuracy of forecast error covariances.

Ménétrier, B. and T. Auligné, 2015: Optimized Localization and Hybridization to Filter Ensemble-Based Covariances
*Monthly Weather Review*, **143**(10), 3931-3947
Other applications

- Multivariate case, optimization of a common localization $\bar{L}$:

$$
\begin{pmatrix}
\hat{B}_{11} & \cdots & \hat{B}_{1P} \\
\vdots & \ddots & \vdots \\
\hat{B}_{P1} & \cdots & \hat{B}_{PP}
\end{pmatrix}
= 
\begin{pmatrix}
\bar{L} & \cdots & \bar{L} \\
\vdots & \ddots & \vdots \\
\bar{L} & \cdots & \bar{L}
\end{pmatrix}
\circ 
\begin{pmatrix}
\tilde{B}_{11} & \cdots & \tilde{B}_{1P} \\
\vdots & \ddots & \vdots \\
\tilde{B}_{P1} & \cdots & \tilde{B}_{PP}
\end{pmatrix}
$$

and of a common hybridization coefficient $\beta^{c2}$. 
Other applications

- Multivariate case, optimization of a common localization $\bar{L}$:

$$
\begin{pmatrix}
\hat{B}_{11} & \cdots & \hat{B}_{1P} \\
\vdots & \ddots & \vdots \\
\hat{B}_{P1} & \cdots & \hat{B}_{PP}
\end{pmatrix} =
\begin{pmatrix}
\bar{L} & \cdots & \bar{L} \\
\vdots & \ddots & \vdots \\
\bar{L} & \cdots & \bar{L}
\end{pmatrix}
\circ
\begin{pmatrix}
\tilde{B}_{11} & \cdots & \tilde{B}_{1P} \\
\vdots & \ddots & \vdots \\
\tilde{B}_{P1} & \cdots & \tilde{B}_{PP}
\end{pmatrix}
$$

and of a common hybridization coefficient $\beta^{c^2}$.

- Optimization of the localization $L^C$ for a correlation matrix:

$$
\tilde{C}_{ij} = \tilde{B}_{ij} / \sqrt{\tilde{B}_{ii} \tilde{B}_{jj}}
$$

and

$$
\hat{C} = L^C \circ \tilde{C}
$$
Other applications

- Multivariate case, optimization of a common localization $\overline{L}$:
  \[
  \begin{pmatrix}
    \hat{B}_{11} & \cdots & \hat{B}_{1P} \\
    \vdots & \ddots & \vdots \\
    \hat{B}_{P1} & \cdots & \hat{B}_{PP}
  \end{pmatrix}
  =
  \begin{pmatrix}
    \overline{L} & \cdots & \overline{L} \\
    \vdots & \ddots & \vdots \\
    \overline{L} & \cdots & \overline{L}
  \end{pmatrix}
  \circ
  \begin{pmatrix}
    \tilde{B}_{11} & \cdots & \tilde{B}_{1P} \\
    \vdots & \ddots & \vdots \\
    \tilde{B}_{P1} & \cdots & \tilde{B}_{PP}
  \end{pmatrix}
  \]

  and of a common hybridization coefficient $\beta^{c2}$.

- Optimization of the localization $L^C$ for a correlation matrix:
  \[
  \tilde{C}_{ij} = \frac{\tilde{B}_{ij}}{\sqrt{\tilde{B}_{ii}\tilde{B}_{jj}}}
  \quad \text{and} \quad
  \hat{C} = L^C \circ \tilde{C}
  \]

- Optimization of a multi-resolution ensemble:
  - small size - high resolution ensemble: $\tilde{B}_{HR}$
  - large size - low resolution ensemble: $\tilde{B}_{LR}$

  Multi-resolution hybridization:
  \[
  \hat{B}^m = L^m_{HR} \circ \tilde{B}_{HR} + L^m_{LR} \circ \tilde{B}_{LR}
  \]
Code available!

A code computing all the localization and hybridization diagnostics is already available:

• simple Fortran 90, NetCDF input/output, OpenMP,
• generic and model-agnostic computation core, for any grid,
• adding a new model takes a few minutes only!

Contact: benjamin.menetrier@cerfacs.fr

Thank you for your attention!

Any question?
A code computing all the localization and hybridization diagnostics is already available:

- simple Fortran 90, NetCDF input/output, OpenMP,
### Code available!

A code computing all the localization and hybridization diagnostics is already available:

- simple Fortran 90, NetCDF input/output, OpenMP,
- generic and model-agnostic computation core, for any grid,
Code available!

A code computing all the localization and hybridization diagnostics is already available:

- simple Fortran 90, NetCDF input/output, OpenMP,
- generic and model-agnostic computation core, for any grid,
- adding a new model takes a few minutes only!
Code available!

A code computing all the localization and hybridization diagnostics is already available:

- simple Fortran 90, NetCDF input/output, OpenMP,
- generic and model-agnostic computation core, for any grid,
- adding a new model takes a few minutes only!

Contact: benjamin.menetrier@cerfacs.fr

Thank you for your attention!
Any question?