Stable and fast inversion with large data sets and non-diagonal *R*

Geir Evensen

IRIS

Outline

IRIS

- Analysis equation
- Large data sets $m >> N$
- Measurements with correlated errors
- Stable inversion of $C=HPH' + R$
- Super-efficient subspace inversion

Analysis equation

IRIS

- Standard Kalman update equation

$$\boldsymbol{\Psi}^{\mathrm{a}} = \boldsymbol{\Psi}^{\mathrm{f}} + \boldsymbol{C}_{\Psi\Psi}\boldsymbol{M}^{\mathrm{T}}\left(\boldsymbol{M}\boldsymbol{C}_{\Psi\Psi}\boldsymbol{M}^{\mathrm{T}} + \boldsymbol{C}_{\epsilon\epsilon}\right)^{-1}\left(\boldsymbol{d} - \boldsymbol{M}\boldsymbol{\Psi}^{\mathrm{f}}\right)$$

Ensemble representation

IRIS

- Ensemble matrix: $$\boldsymbol{A} = (\boldsymbol{\Psi}_1, \boldsymbol{\Psi}_2, \ldots, \boldsymbol{\Psi}_N)$$
- Ensemble perturbations: $$\boldsymbol{A}' = \boldsymbol{A} - \overline{\boldsymbol{A}}$$
- Ensemble covariance: $$\boldsymbol{C}^{\mathrm{e}}_{\Psi\Psi} = \frac{\boldsymbol{A}'\boldsymbol{A}'^{\mathrm{T}}}{N-1}$$

Ensemble representation for measurements

IRIS

- Perturbed measurements: $\boldsymbol{d}_j = \boldsymbol{d} + \boldsymbol{\epsilon}_j$
- Measurement matrix: $\boldsymbol{D} = (\boldsymbol{d}_1, \boldsymbol{d}_2, \ldots, \boldsymbol{d}_N)$
- Measurement perturbations: $\boldsymbol{E} = (\boldsymbol{\epsilon}_1, \boldsymbol{\epsilon}_2, \ldots, \boldsymbol{\epsilon}_N)$
- Measurement error cov: $$\boldsymbol{C}^{\mathrm{e}}_{\epsilon\epsilon} = \frac{\boldsymbol{E}\boldsymbol{E}^{\mathrm{T}}}{N-1}$$

Analysis equation

IRIS

- Standard Kalman update equation

$$\boldsymbol{\Psi}^{\mathrm{a}} = \boldsymbol{\Psi}^{\mathrm{f}} + \boldsymbol{C}_{\psi\psi}\boldsymbol{M}^{\mathrm{T}}\left(\boldsymbol{M}\boldsymbol{C}_{\psi\psi}\boldsymbol{M}^{\mathrm{T}} + \boldsymbol{C}_{\epsilon\epsilon}\right)^{-1}\left(\boldsymbol{d} - \boldsymbol{M}\boldsymbol{\Psi}^{\mathrm{f}}\right)$$

- Ensemble formulation

$$\boldsymbol{A}^{\mathrm{a}} = \boldsymbol{A}^{\mathrm{f}} + \boldsymbol{A}'^{\mathrm{f}}(\boldsymbol{M}\boldsymbol{A}'^{\mathrm{f}})^{\mathrm{T}}\left(\boldsymbol{M}\boldsymbol{A}'^{\mathrm{f}}(\boldsymbol{M}\boldsymbol{A}'^{\mathrm{f}})^{\mathrm{T}} + \boldsymbol{E}\boldsymbol{E}^{\mathrm{T}}\right)^{-1}\left(\boldsymbol{D} - \boldsymbol{M}\boldsymbol{A}^{\mathrm{f}}\right)$$

$$\boldsymbol{A}' = \boldsymbol{A}(\boldsymbol{I} - \boldsymbol{1}_N) \qquad \boldsymbol{C} = \boldsymbol{S}\boldsymbol{S}^{\mathrm{T}} + (N-1)\boldsymbol{C}_{\epsilon\epsilon}$$

$$\boldsymbol{S} = \boldsymbol{M}\boldsymbol{A}', \qquad \boldsymbol{D}' = \boldsymbol{D} - \boldsymbol{M}\boldsymbol{A}$$

Update equation

IRIS

$$
\begin{aligned}
A^{\mathrm{a}} &= A + A' S^{\mathrm{T}} C^{-1} D' \\
&= A + A(I - \mathbf{1}_N) S^{\mathrm{T}} C^{-1} D' \\
&= A(I + (I - \mathbf{1}_N) S^{\mathrm{T}} C^{-1} D') \\
&= A(I + S^{\mathrm{T}} C^{-1} D') \\
&= AX,
\end{aligned}
\qquad \mathbf{1}_N S^{\mathrm{T}} \equiv \mathbf{0}
$$

- Solution searched for in the space spanned by prior realizations.
- Inversion of *C* with large *m>>N*?

$$C = SS^{\mathrm{T}} + (N-1) C_{\epsilon\epsilon}$$

SQRT scheme

IRIS

Update the mean and perturbations seperately

$$\overline{\boldsymbol{\psi}}^{\mathrm{a}} = \overline{\boldsymbol{\psi}}^{\mathrm{f}} + \boldsymbol{A}'\boldsymbol{S}^{\mathrm{T}}\boldsymbol{C}^{-1}\left(\boldsymbol{d} - \boldsymbol{M}\overline{\boldsymbol{\psi}}^{\mathrm{f}}\right)$$

$$\boldsymbol{A}^{\mathrm{a}\prime}\boldsymbol{A}^{\mathrm{a}\prime\mathrm{T}} = \boldsymbol{A}'\left(\boldsymbol{I} - \boldsymbol{S}^{\mathrm{T}}\boldsymbol{C}^{-1}\boldsymbol{S}\right)\boldsymbol{A}'^{\mathrm{T}}$$

Still need to invert *C*

Implementation issues

IRIS

- Need to invert

$$\boldsymbol{C} = \boldsymbol{S}\boldsymbol{S}^{\mathrm{T}} + (N-1)\boldsymbol{C}_{\epsilon\epsilon}.$$

- May be singular or poorly conditioned.
- Pseudo inversion:

$$\boldsymbol{C} = \boldsymbol{Z}\boldsymbol{\Lambda}\boldsymbol{Z}^{\mathrm{T}} \Rightarrow \boldsymbol{C}^{+} = \boldsymbol{Z}\boldsymbol{\Lambda}^{+}\boldsymbol{Z}^{\mathrm{T}}$$

$$\text{diag}(\boldsymbol{\Lambda}^{+}) = (\lambda_1^{-1}, \ldots, \lambda_p^{-1}, 0, \ldots, 0).$$

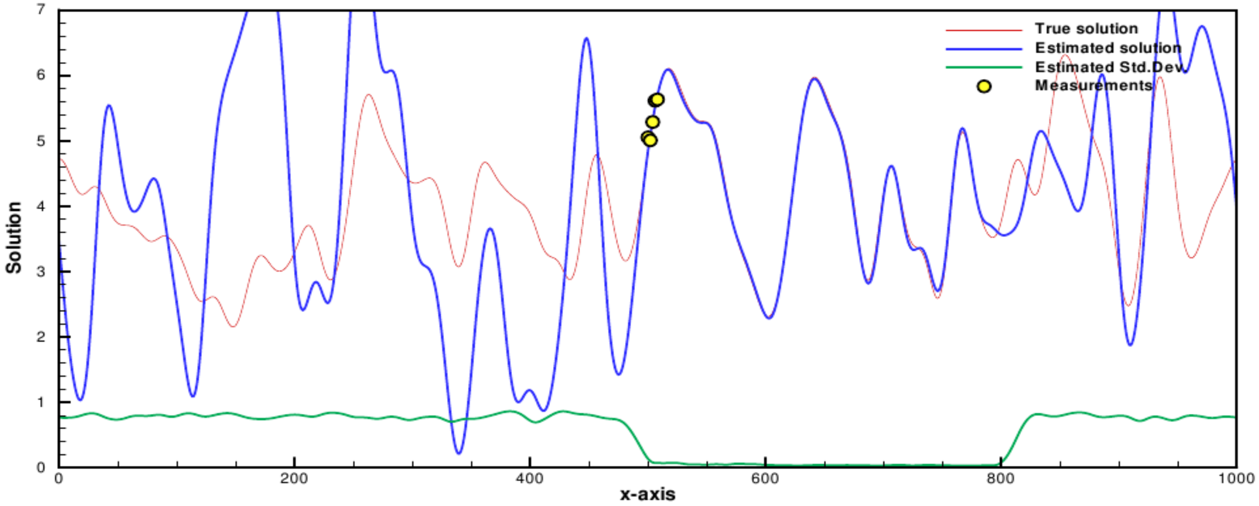
$$\mathcal{O}(m^3) \text{ operations.}$$

Example: EnKF with pseudo inversion of *C*

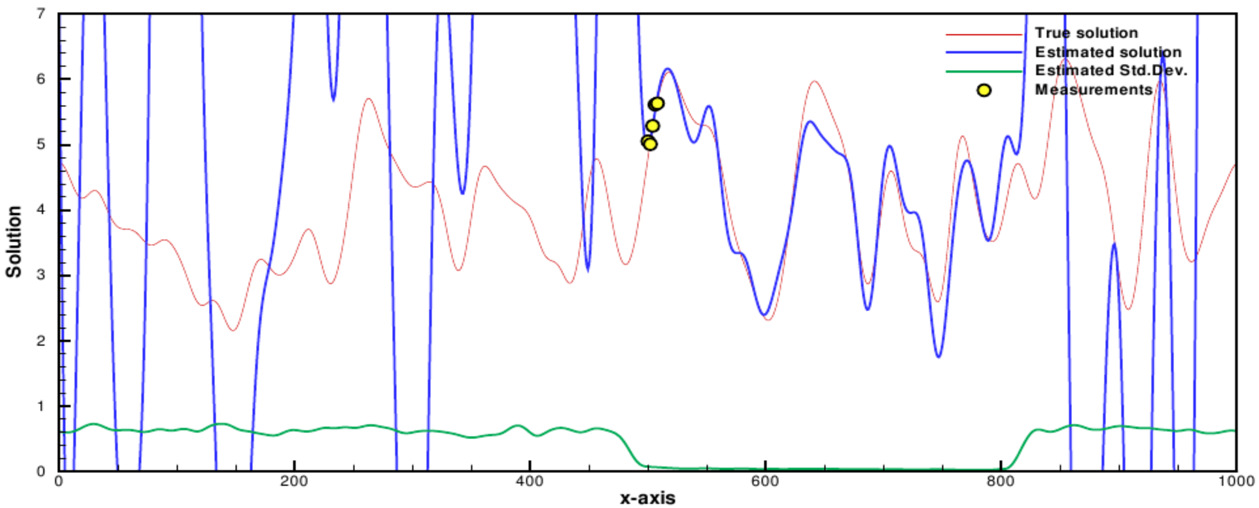
IRIS

Eig(1)/eig(5)= O(10^5)

Truncation at 90% retains one eigenvalue out of five.



Truncation at 99.9% retains four eigenvalues out of five.

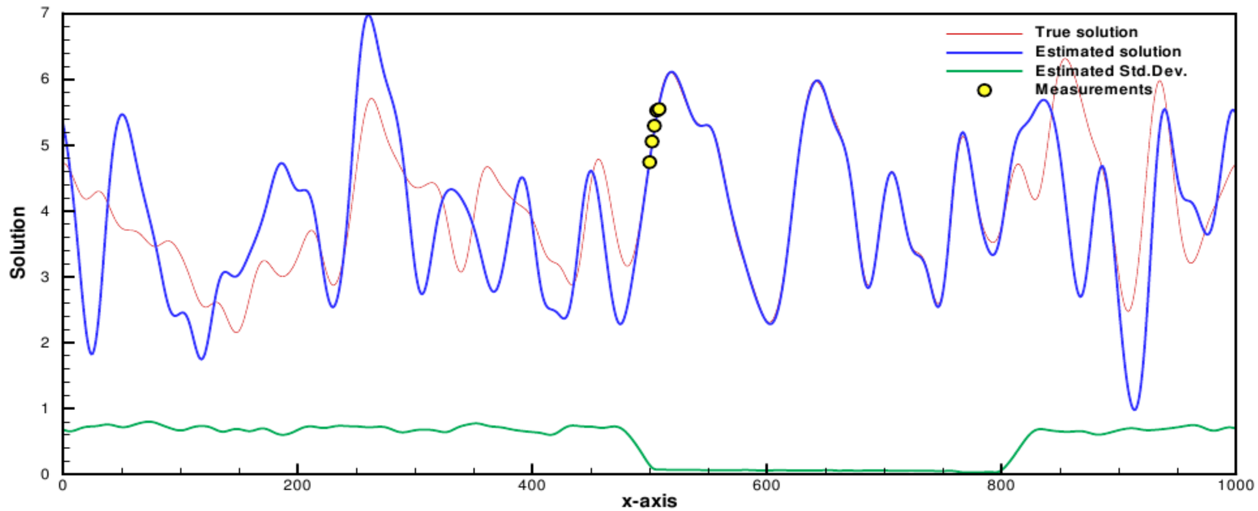


Example: EnSQRT with pseudo inversion of *C*

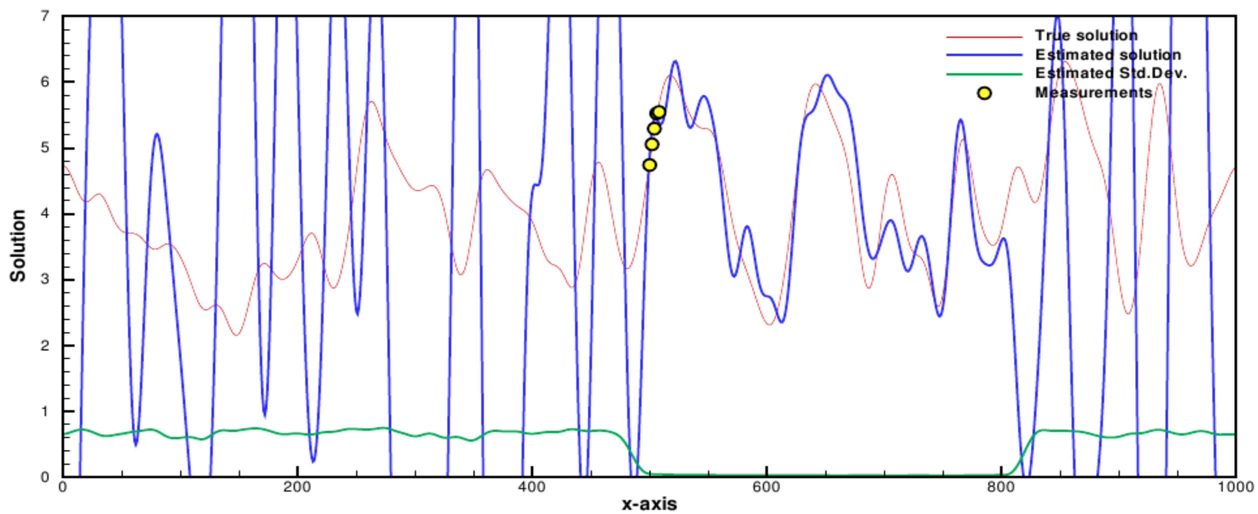
IRIS

Eig(1)/eig(5)= O(10^5)

Truncation at 90% retains one eigenvalue out of five.



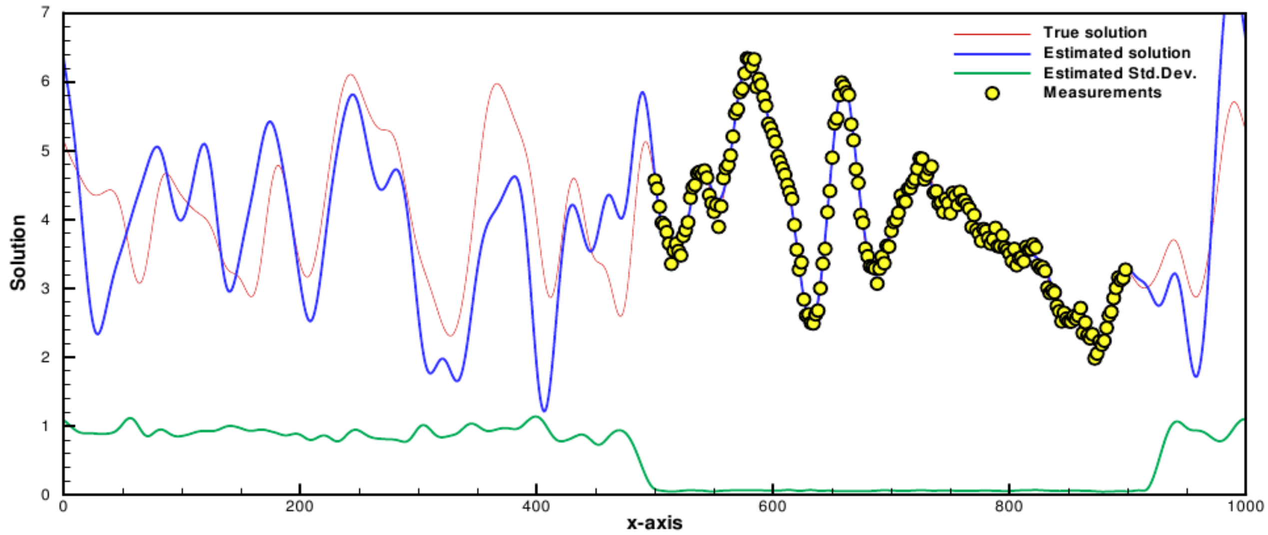
Truncation at 99.9% retains four eigenvalues out of five.



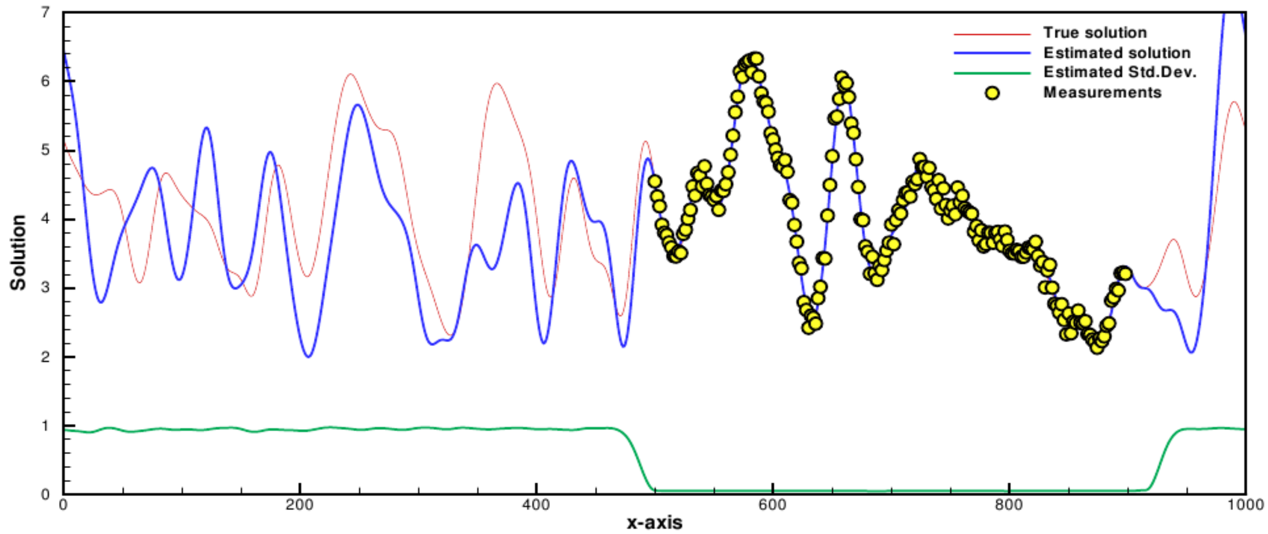
Example with 200 measurements

IRIS

EnKF: Truncation at 99% retains about 40 eigenvalues.



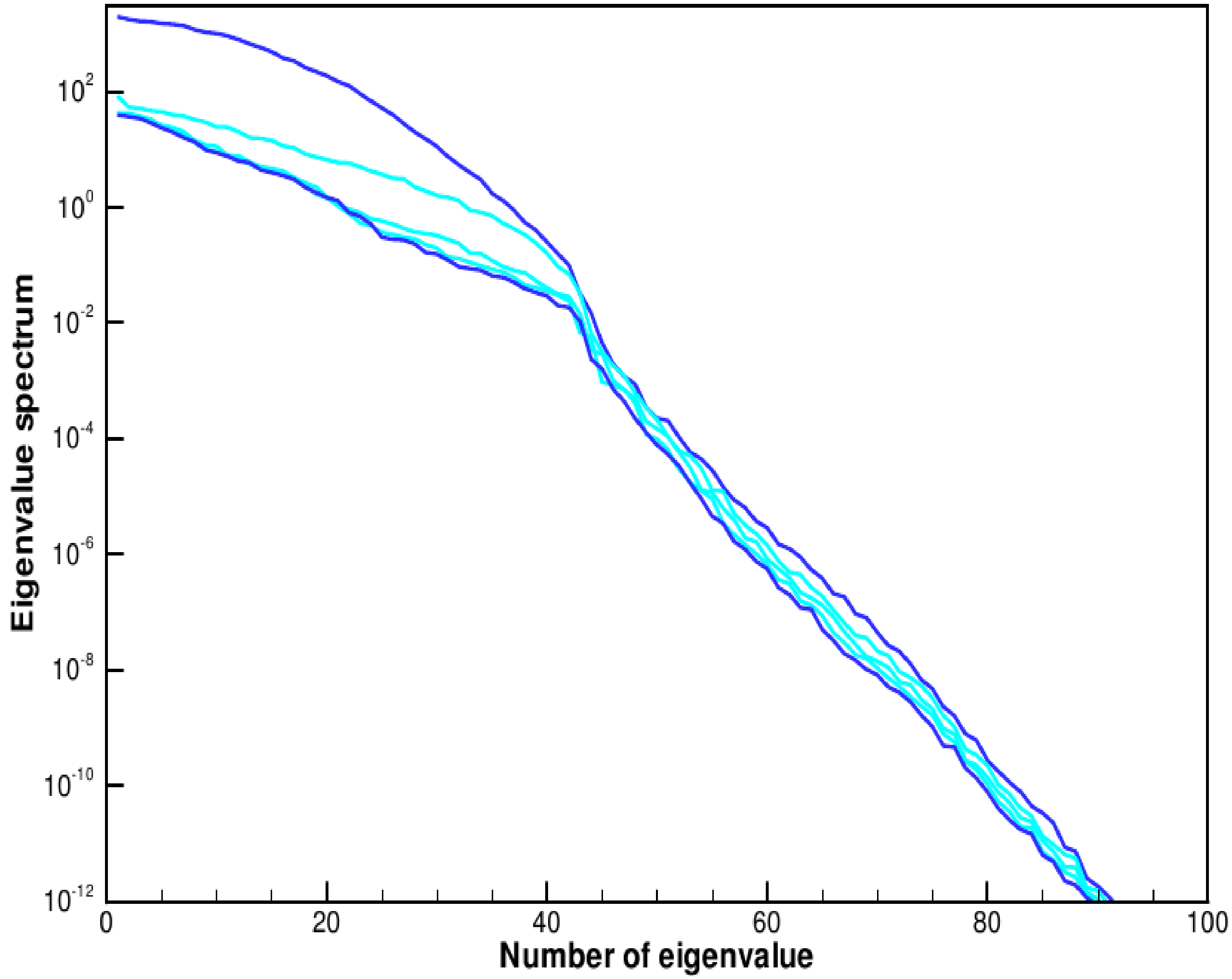
EnSQRT: Truncation at 99% retains about 40 eigenvalues.



Example with 200 measurements

EnKF: Truncation at 99% retains about 40 eigenvalues.

EnSQRT: Truncation at 99% retains about 40 eigenvalues.



Ensemble subspace inversion (1)

IRIS

- Compute inversion in *N*-dim ensemble space rather than the *m*-dim meas. space.

- SVD of *S*

$$\boldsymbol{U}_0\boldsymbol{\Sigma}_0\boldsymbol{V}_0^{\mathrm{T}} = \boldsymbol{S}$$

- Pseudo inverse of *S*

$$\boldsymbol{S}^{+} = \boldsymbol{V}_0\boldsymbol{\Sigma}_0^{+}\boldsymbol{U}_0^{\mathrm{T}}$$

$$\begin{aligned}
\boldsymbol{C} &= \left(\boldsymbol{U}_0\boldsymbol{\Sigma}_0\boldsymbol{\Sigma}_0^{\mathrm{T}}\boldsymbol{U}_0^{\mathrm{T}} + (N-1)\boldsymbol{C}_{\epsilon\epsilon}\right) \\
&= \boldsymbol{U}_0\left(\boldsymbol{\Sigma}_0\boldsymbol{\Sigma}_0^{\mathrm{T}} + (N-1)\boldsymbol{U}_0^{\mathrm{T}}\boldsymbol{C}_{\epsilon\epsilon}\boldsymbol{U}_0\right)\boldsymbol{U}_0^{\mathrm{T}} \\
&\approx \boldsymbol{U}_0\boldsymbol{\Sigma}_0\left(\boldsymbol{I} + (N-1)\boldsymbol{\Sigma}_0^{+}\boldsymbol{U}_0^{\mathrm{T}}\boldsymbol{C}_{\epsilon\epsilon}\boldsymbol{U}_0\boldsymbol{\Sigma}_0^{+\mathrm{T}}\right)\boldsymbol{\Sigma}_0^{\mathrm{T}}\boldsymbol{U}_0^{\mathrm{T}} \\
&= \boldsymbol{S}\boldsymbol{S}^{\mathrm{T}} + (N-1)(\boldsymbol{S}\boldsymbol{S}^{+})\boldsymbol{C}_{\epsilon\epsilon}(\boldsymbol{S}\boldsymbol{S}^{+})^{\mathrm{T}}.
\end{aligned}$$

Ensemble subspace inversion (2)

IRIS

- New expression for *C*

$$\boldsymbol{C} \approx \boldsymbol{U}_0\boldsymbol{\Sigma}_0(\boldsymbol{I} + \boldsymbol{X}_0)\boldsymbol{\Sigma}_0^{\mathrm{T}}\boldsymbol{U}_0^{\mathrm{T}}$$

$$\boldsymbol{X}_0 = (N-1)\boldsymbol{\Sigma}_0^{+}\boldsymbol{U}_0^{\mathrm{T}}\boldsymbol{C}_{\epsilon\epsilon}\boldsymbol{U}_0\boldsymbol{\Sigma}_0^{+\mathrm{T}}$$

- Eigenvalue decomposition

$$\boldsymbol{Z}_1\boldsymbol{\Lambda}_1\boldsymbol{Z}_1^{\mathrm{T}} = \boldsymbol{X}_0$$

- *C* becomes

$$\begin{aligned}\boldsymbol{C} &\approx \boldsymbol{U}_0\boldsymbol{\Sigma}_0(\boldsymbol{I} + \boldsymbol{Z}_1\boldsymbol{\Lambda}_1\boldsymbol{Z}_1^{\mathrm{T}})\boldsymbol{\Sigma}_0^{\mathrm{T}}\boldsymbol{U}_0^{\mathrm{T}}\\ &= \boldsymbol{U}_0\boldsymbol{\Sigma}_0\boldsymbol{Z}_1(\boldsymbol{I} + \boldsymbol{\Lambda}_1)\boldsymbol{Z}_1^{\mathrm{T}}\boldsymbol{\Sigma}_0^{\mathrm{T}}\boldsymbol{U}_0^{\mathrm{T}}\end{aligned}$$

Ensemble subspace inversion (3)

IRIS

$$\begin{aligned} \boldsymbol{C} &\approx \boldsymbol{U}_0\boldsymbol{\Sigma}_0(\boldsymbol{I} + \boldsymbol{Z}_1\boldsymbol{\Lambda}_1\boldsymbol{Z}_1^{\mathrm{T}})\boldsymbol{\Sigma}_0^{\mathrm{T}}\boldsymbol{U}_0^{\mathrm{T}} \\ &= \boldsymbol{U}_0\boldsymbol{\Sigma}_0\boldsymbol{Z}_1(\boldsymbol{I} + \boldsymbol{\Lambda}_1)\boldsymbol{Z}_1^{\mathrm{T}}\boldsymbol{\Sigma}_0^{\mathrm{T}}\boldsymbol{U}_0^{\mathrm{T}} \end{aligned}$$

- Pseudo inverse of *C* becomes

$$\begin{aligned} \boldsymbol{C}^{+} &\approx (\boldsymbol{U}_0\boldsymbol{\Sigma}_0^{+\mathrm{T}}\boldsymbol{Z}_1)(\boldsymbol{I} + \boldsymbol{\Lambda}_1)^{-1}(\boldsymbol{U}_0\boldsymbol{\Sigma}_0^{+\mathrm{T}}\boldsymbol{Z}_1)^{\mathrm{T}} \\ &= \boldsymbol{X}_1(\boldsymbol{I} + \boldsymbol{\Lambda}_1)^{-1}\boldsymbol{X}_1^{\mathrm{T}}, \end{aligned}$$

$$\boldsymbol{X}_1 = \boldsymbol{U}_0\boldsymbol{\Sigma}_0^{+\mathrm{T}}\boldsymbol{Z}_1.$$

$$m^2N + mN^2$$

EnKF analysis by subspace pseudo inversion

$$\boldsymbol{A}^{\mathrm{a}} = \boldsymbol{A}^{\mathrm{f}}\left(\boldsymbol{I} + \boldsymbol{S}^{\mathrm{T}}\boldsymbol{X}_1(\boldsymbol{I} + \boldsymbol{\Lambda}_1)^{-1}\boldsymbol{X}_1^{\mathrm{T}}(\boldsymbol{D} - \mathcal{M}[\boldsymbol{A}^{\mathrm{f}}])\right).$$

Low rank ensemble subspace inversion (1)

IRIS

$$\mathbf{C}^{\mathrm{e}}_{\epsilon\epsilon} = \mathbf{E}\mathbf{E}^{\mathrm{T}}/(N-1) \qquad \mathbf{C} = \mathbf{S}\mathbf{S}^{\mathrm{T}} + \mathbf{E}\mathbf{E}^{\mathrm{T}}$$

$$\begin{aligned}\mathbf{C} &\approx \mathbf{U}_0\mathbf{\Sigma}_0\big(\mathbf{I} + \mathbf{\Sigma}_0^{+}\mathbf{U}_0^{\mathrm{T}}\mathbf{E}\mathbf{E}^{\mathrm{T}}\mathbf{U}_0\mathbf{\Sigma}_0^{+\mathrm{T}}\big)\mathbf{\Sigma}_0^{\mathrm{T}}\mathbf{U}_0^{\mathrm{T}} \\ &= \mathbf{U}_0\mathbf{\Sigma}_0\big(\mathbf{I} + \mathbf{X}_0\mathbf{X}_0^{\mathrm{T}}\big)\mathbf{\Sigma}_0^{\mathrm{T}}\mathbf{U}_0^{\mathrm{T}},\end{aligned}$$

$$\mathbf{X}_0 = \mathbf{\Sigma}_0^{+}\mathbf{U}_0^{\mathrm{T}}\mathbf{E}$$

$$\mathbf{U}_1\mathbf{\Sigma}_1\mathbf{V}_1^{\mathrm{T}} = \mathbf{X}_0$$

Low rank ensemble subspace inversion (2)

IRIS

$$
\begin{aligned}
\boldsymbol{C} &\approx \boldsymbol{U}_0\boldsymbol{\Sigma}_0(\boldsymbol{I} + \boldsymbol{U}_1\boldsymbol{\Sigma}_1^2\boldsymbol{U}_1^{\mathrm{T}})\boldsymbol{\Sigma}_0^{\mathrm{T}}\boldsymbol{U}_0^{\mathrm{T}} \\
&= \boldsymbol{U}_0\boldsymbol{\Sigma}_0\boldsymbol{U}_1(\boldsymbol{I} + \boldsymbol{\Sigma}_1^2)\boldsymbol{U}_1^{\mathrm{T}}\boldsymbol{\Sigma}_0^{\mathrm{T}}\boldsymbol{U}_0^{\mathrm{T}}
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{C}^{+} &\approx (\boldsymbol{U}_0\boldsymbol{\Sigma}_0^{+\mathrm{T}}\boldsymbol{U}_1)(\boldsymbol{I} + \boldsymbol{\Sigma}_1^2)^{-1}(\boldsymbol{U}_0\boldsymbol{\Sigma}_0^{+\mathrm{T}}\boldsymbol{U}_1)^{\mathrm{T}} \\
&= \boldsymbol{X}_1(\boldsymbol{I} + \boldsymbol{\Sigma}_1^2)^{-1}\boldsymbol{X}_1^{\mathrm{T}},
\end{aligned}
$$

$$
\boldsymbol{X}_1 = \boldsymbol{U}_0\boldsymbol{\Sigma}_0^{+\mathrm{T}}\boldsymbol{U}_1 \qquad mN^2
$$

EnKF subspace analysis with low-rank $\boldsymbol{C}_{\epsilon\epsilon}$

$$
\boldsymbol{A}^{\mathrm{a}} = \boldsymbol{A}^{\mathrm{f}}\left(\boldsymbol{I} + \boldsymbol{S}^{\mathrm{T}}\boldsymbol{X}_1(\boldsymbol{I} + \boldsymbol{\Sigma}_1^2)^{-1}\boldsymbol{X}_1^{\mathrm{T}}(\boldsymbol{D} - \mathcal{M}[\boldsymbol{A}^{\mathrm{f}}])\right).
$$

Summary of EnKF analyses implementations

IRIS

Standard EnKF analysis

$$\boldsymbol{A}^{\mathrm{a}} = \boldsymbol{A}^{\mathrm{f}}\left(\boldsymbol{I} + \boldsymbol{S}^{\mathrm{T}}\boldsymbol{C}^{-1}(\boldsymbol{D} - \mathcal{M}[\boldsymbol{A}^{\mathrm{f}}])\right),$$

$$\boldsymbol{C}^{+} = \boldsymbol{Z}\boldsymbol{\Lambda}^{+}\boldsymbol{Z}^{\mathrm{T}}$$

$\mathcal{O}(m^3)$ operations.

EnKF analysis by subspace pseudo inversion

$$\boldsymbol{A}^{\mathrm{a}} = \boldsymbol{A}^{\mathrm{f}}\left(\boldsymbol{I} + \boldsymbol{S}^{\mathrm{T}}\boldsymbol{X}_1(\boldsymbol{I} + \boldsymbol{\Lambda}_1)^{-1}\boldsymbol{X}_1^{\mathrm{T}}(\boldsymbol{D} - \mathcal{M}[\boldsymbol{A}^{\mathrm{f}}])\right).$$

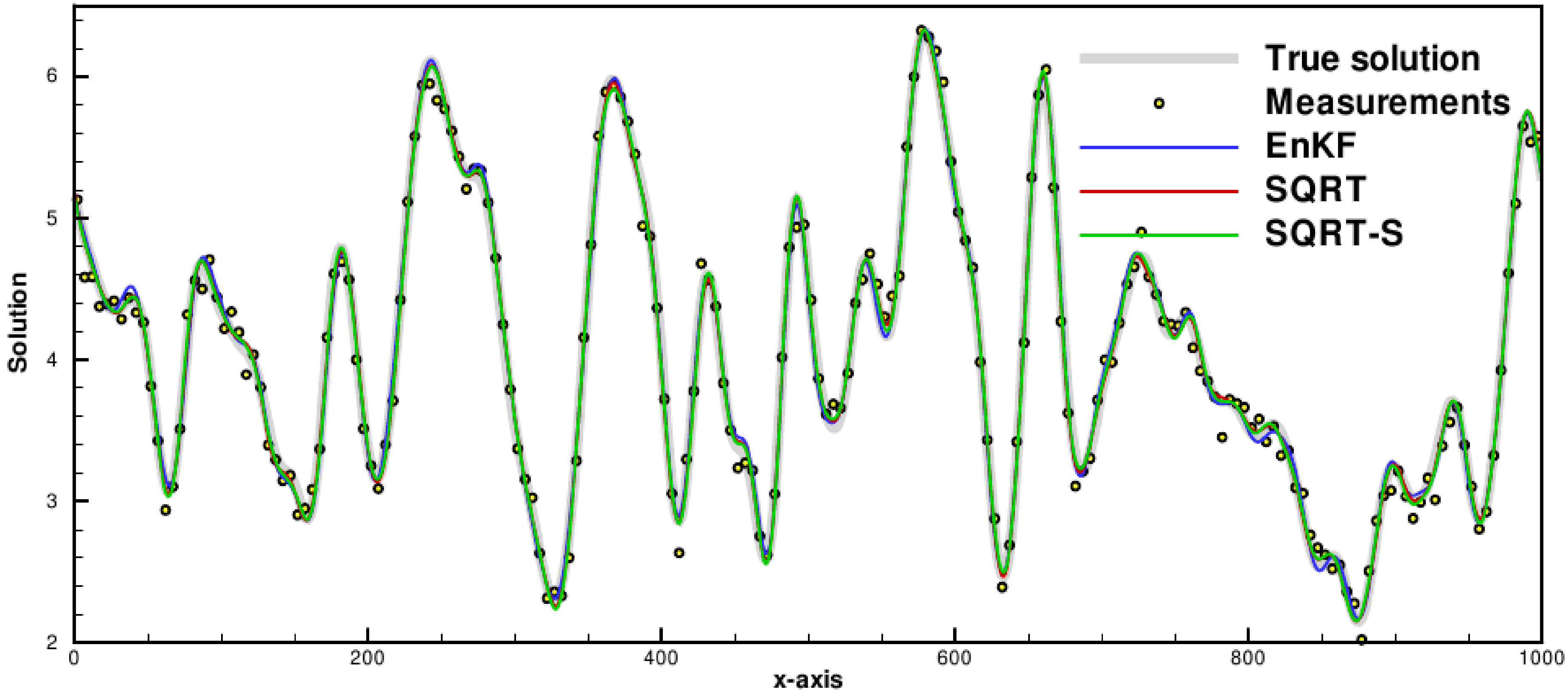
$m^2N + mN^2$

EnKF subspace analysis with low-rank $\boldsymbol{C}_{\epsilon\epsilon}$

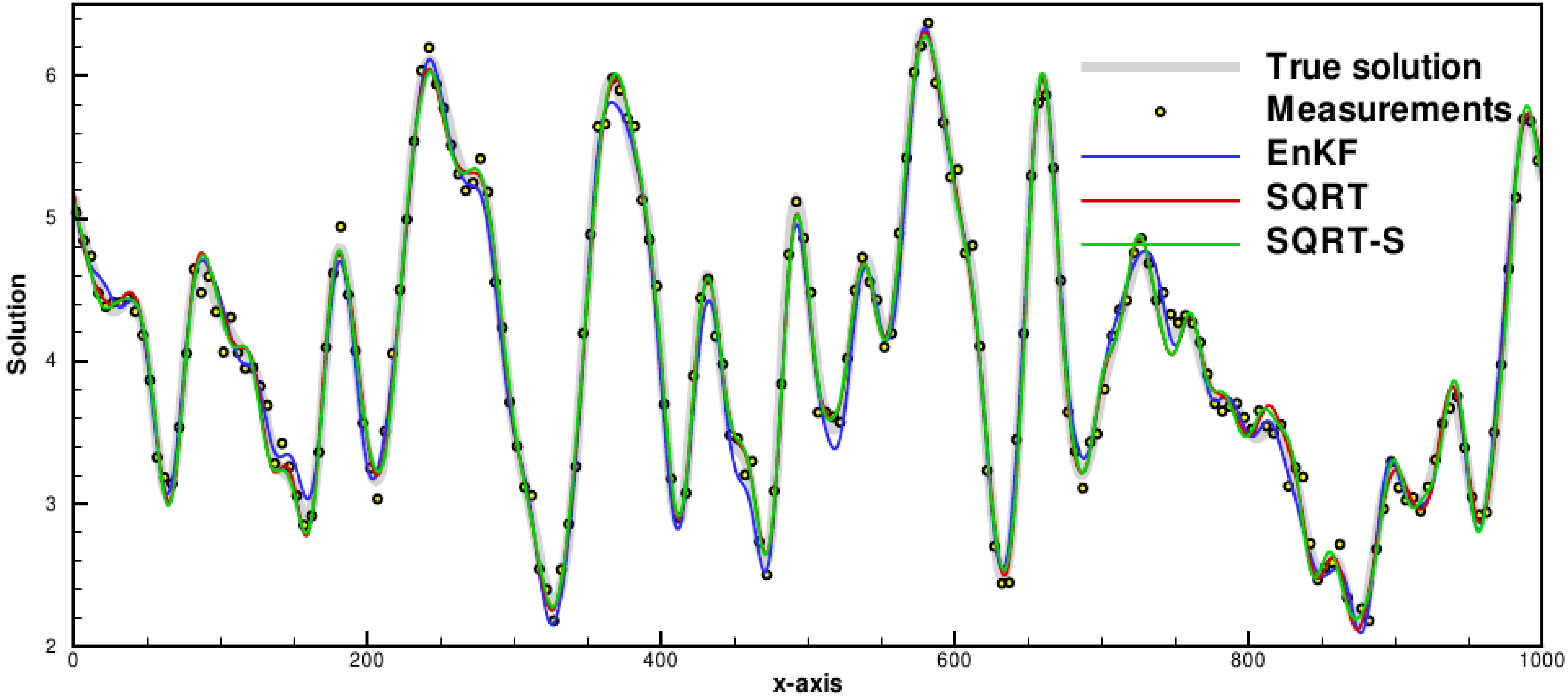
$$\boldsymbol{A}^{\mathrm{a}} = \boldsymbol{A}^{\mathrm{f}}\left(\boldsymbol{I} + \boldsymbol{S}^{\mathrm{T}}\boldsymbol{X}_1(\boldsymbol{I} + \boldsymbol{\Sigma}_1^2)^{-1}\boldsymbol{X}_1^{\mathrm{T}}(\boldsymbol{D} - \mathcal{M}[\boldsymbol{A}^{\mathrm{f}}])\right).$$

mN^2

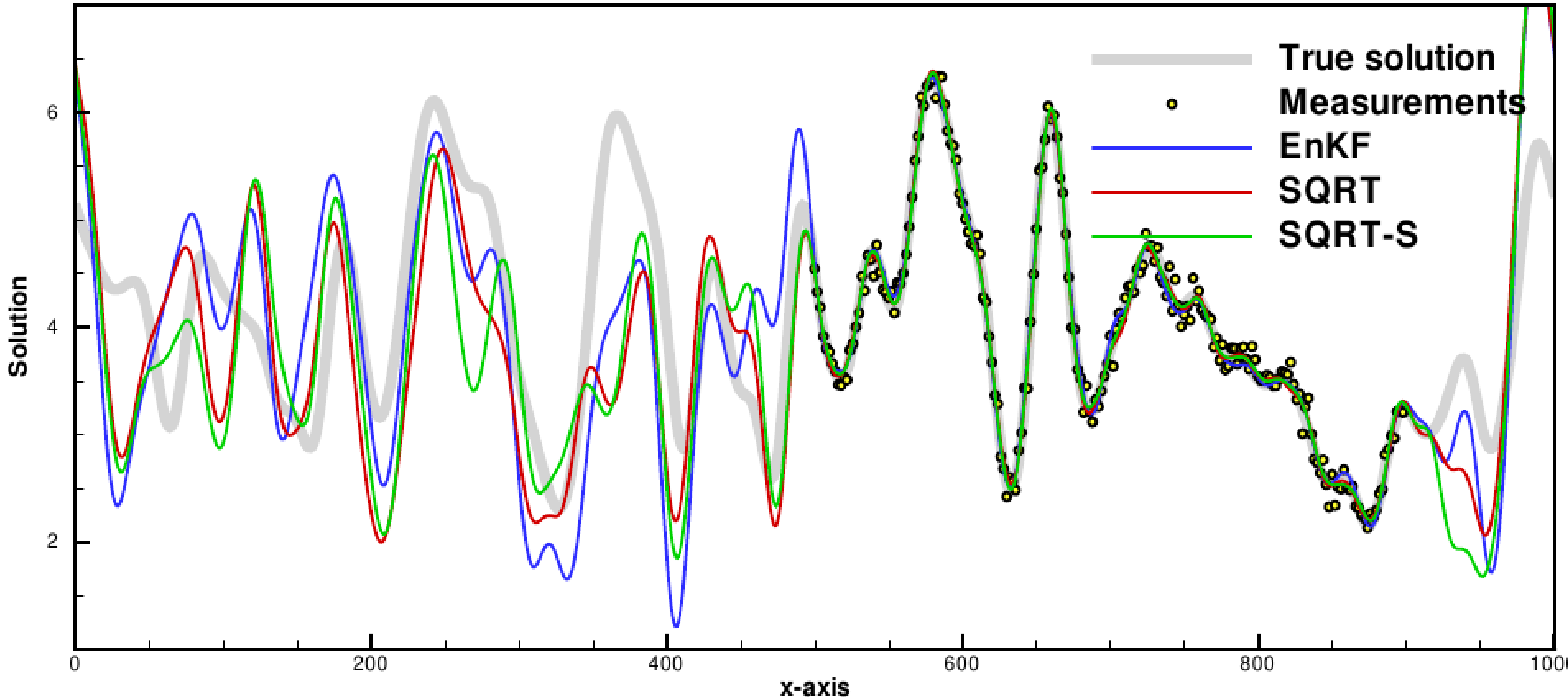
Examples: Diagonal *R*



Examples: Non-diagonal *R*



Examples: Non-diagonal *R* and clustered data



Summary

IRIS

- Necessary to deal with the rank of *C.*
  - Pseudo inversion is key.
  - Adding a diagonal *R* may improve conditioning but does not add more "information."

- No need to restrict analysis to a diagonal *R.*
  - Analysis with full R as efficient.
  - No need to actually construct measurement error-covariance matrix.
  - Easier and faster to sample measurement perturbations in *E.*

- Stable and efficient SQRT schemes that handle non-diagonal *R* are available.
  - Ensemble subspace pseudo inversion and representation of *R* using measurement perturbations.
  - Represent the measurement errors within the ensemble subspace!

- Details in Evensen (2009), Chapter 14.