KU LEUVEN

STADIUS
Center for Dynamical Systems, Signal Processing and Data Analytics

Improved Tracking for Distributed Signal Fusion Optimization in a Fully-Connected Wireless Sensor Network

30th EUROPEAN SIGNAL PROCESSING CONFERENCE
EUSIPCO
BELGRADE/SERBIA
29 AUG – 2 SEP 2022

Cem Ates Musluoglu, Marc Moonen and Alexander Bertrand

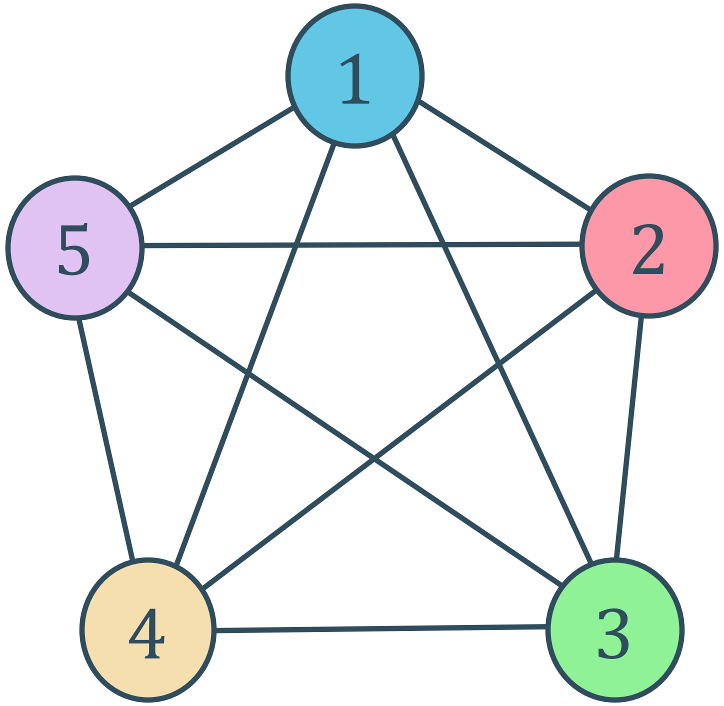
Presentation by Cem Ates Musluoglu

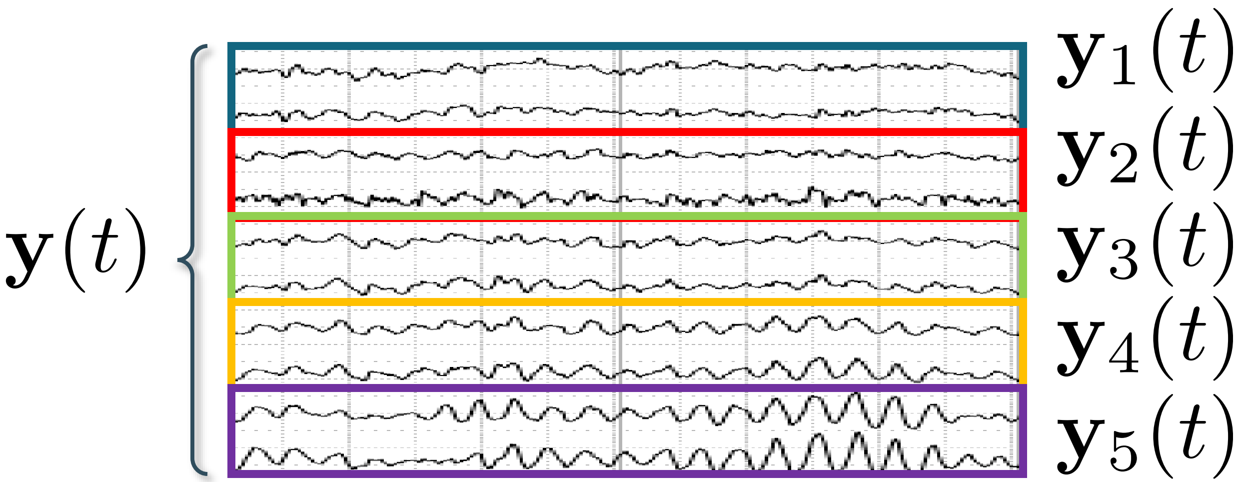
cemates.musluoglu@esat.kuleuven.be

European Research Council
Established by the European Commission

Wireless Sensor Network Setting

Each node k is a local sensor array measuring <u>multi-channel</u> signal $\boldsymbol{y}_k(t)$





Exploiting the spatial coherence

Goal: Exploit spatial coherence between channels of $\boldsymbol{y}$ by optimally combining them

using a linear filter $\boldsymbol{x}$ applied to $\boldsymbol{y}$: $\boldsymbol{x}^T\boldsymbol{y}$

Without data centralization

$\Downarrow$

Optimization problem to find $\boldsymbol{x}$

The scope of the framework

$$\underset{\mathbf{x}}{\text{minimize}}\ f(\mathbf{x}^T\mathbf{y}(t))$$

Example

LS / MMSE: Signal estimation

$$\underset{\mathbf{x}}{\text{minimize}}\ \mathbb{E}[||\mathbf{d}(t)-\mathbf{x}^T\mathbf{y}(t)||^2]$$

The scope of the framework

Example

PCA / GEVD: Dimensionality reduction

$$\begin{aligned} \underset{X}{\text{minimize}} \; & f(X^T\mathbf{y}(t), X^T\mathbf{v}(t)) \\ \text{subject to} \; & g_j(X^T\mathbf{y}(t), X^T\mathbf{v}(t)) = 0, \\ & h_j(X^T\mathbf{y}(t), X^T\mathbf{v}(t)) \leq 0 \end{aligned}$$

$$\begin{aligned} \underset{X}{\text{maximize}} \; & \mathbb{E}[||X^T\mathbf{y}(t)||]^2 \\ \text{subject to} \; & \mathbb{E}[X^T\mathbf{v}(t)\mathbf{v}(t)^T X] = I \end{aligned}$$

The scope of the framework

$$
\begin{aligned}
\underset{(X,W)}{\text{minimize}}\ & f(X^T\mathbf{y}(t), W^T\mathbf{v}(t)) \\
\text{subject to}\ & g_j(X^T\mathbf{y}(t), W^T\mathbf{v}(t)) = 0, \\
& h_j(X^T\mathbf{y}(t), W^T\mathbf{v}(t)) \le 0
\end{aligned}
$$

Example

CCA: Correlation between data sets

$$
\begin{aligned}
\underset{(X,W)}{\text{maximize}}\ & \mathbb{E}[\mathrm{tr}(X^T\mathbf{y}(t)\mathbf{v}^T(t)W)] \\
\text{subject to}\ & \mathbb{E}[X^T\mathbf{y}(t)\mathbf{y}(t)^T X] = I \\
& \mathbb{E}[W^T\mathbf{v}(t)\mathbf{v}(t)^T W] = I
\end{aligned}
$$

The scope of the framework

$$\begin{aligned}\underset{X}{\text{minimize}}\ & f(X^T\mathbf{y}(t), X^TB)\\ \text{subject to}\ & g_j(X^T\mathbf{y}(t), X^TB) = 0,\\ & h_j(X^T\mathbf{y}(t), X^TB) \leq 0\end{aligned}$$

Example

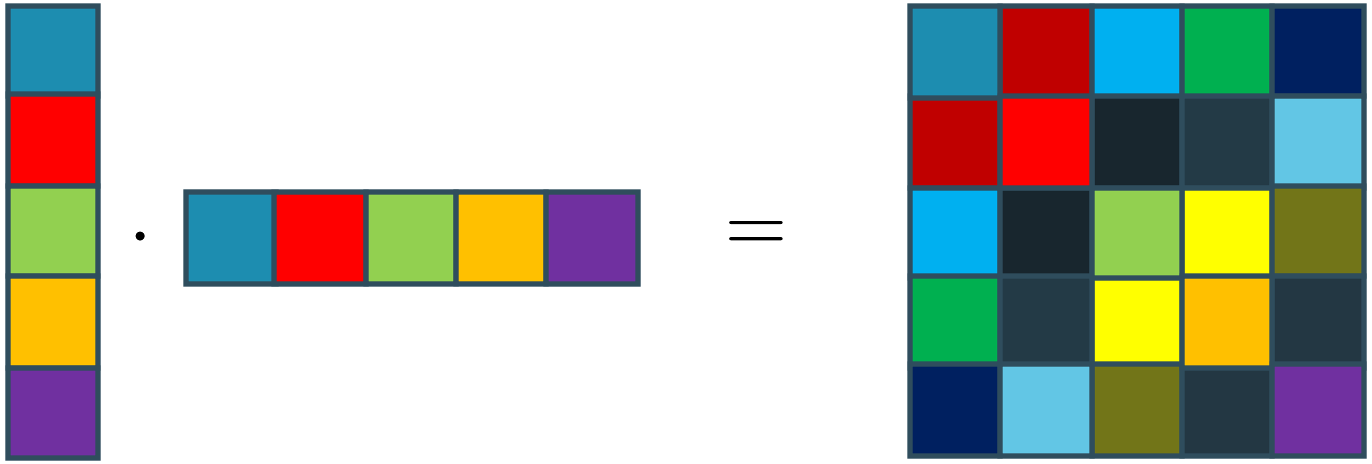
(Robust) **minimum variance beamforming**

$$\begin{aligned}\underset{X}{\text{minimize}}\ & \mathbb{E}[||X^T\mathbf{y}(t)||]^2\\ \text{subject to}\ & X^TB = H,\\ & \mathrm{tr}(X^TX) \leq \alpha^2\end{aligned}$$

$$X^TX = (X^T \cdot I)(X^T \cdot I)^T = (X^T \cdot B)(X^T \cdot B)^T$$

Exploiting the spatial coherence

- We want to exploit spatial coherence to optimally combine the data
- → Spatial covariance matrix



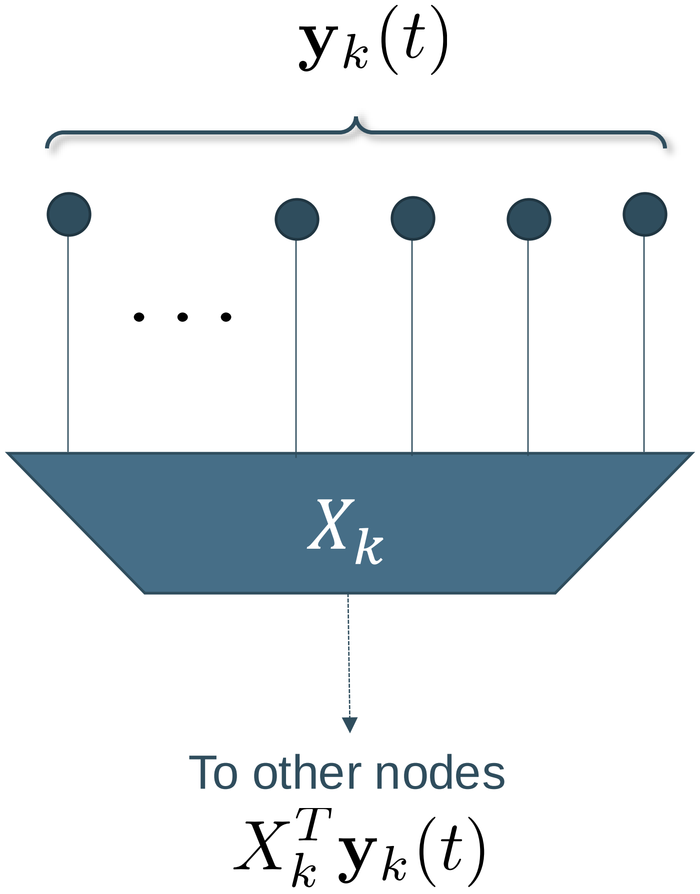
$$\cdot \quad = \quad = K_{\mathbf{yy}}$$

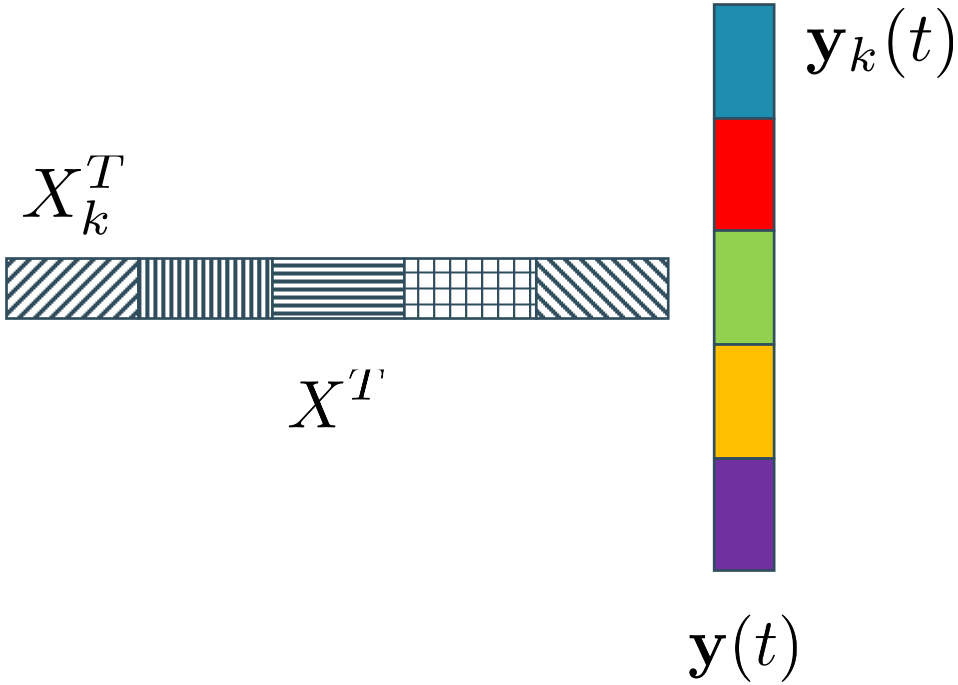
$$\mathbf{y} \cdot \mathbf{y}^T$$

Unavailable if no exchange of data!

Batches of signals required → Large communication costs

Fuse and Forward





Only compressed data is communicated

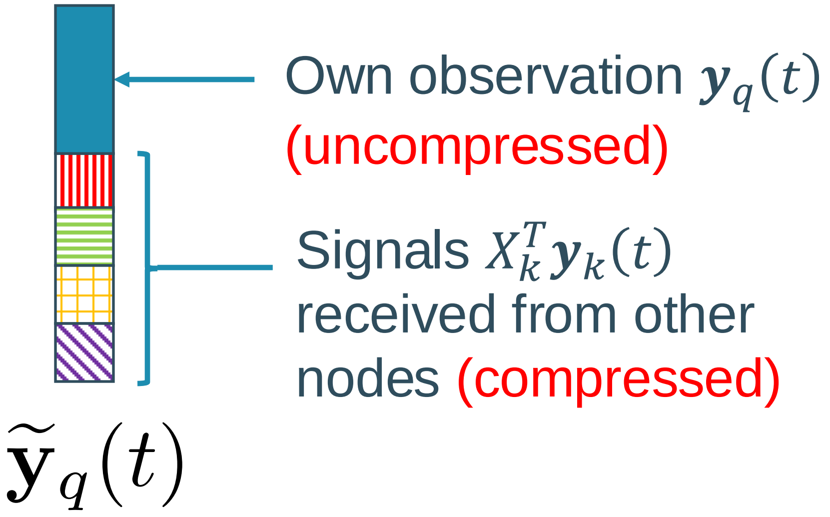
Compression ratio of M_k/Q

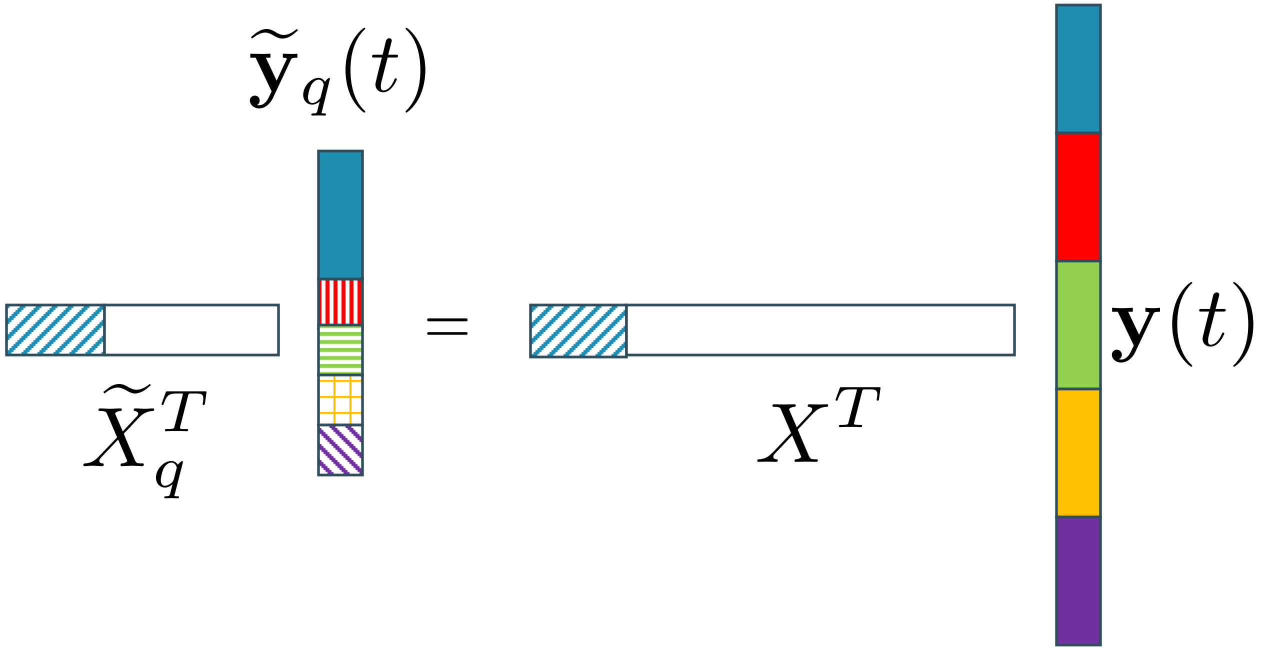
M_k: Number of channels for node k

Q: Number of filters (columns of X)

Local data – Local variable – Local problem

Available local data at node q:





$$f(X^T \mathbf{y}(t)) = f(\tilde{X}_q^T \tilde{\mathbf{y}}_q(t)) \quad \rightarrow$$

Centralized global problem

~

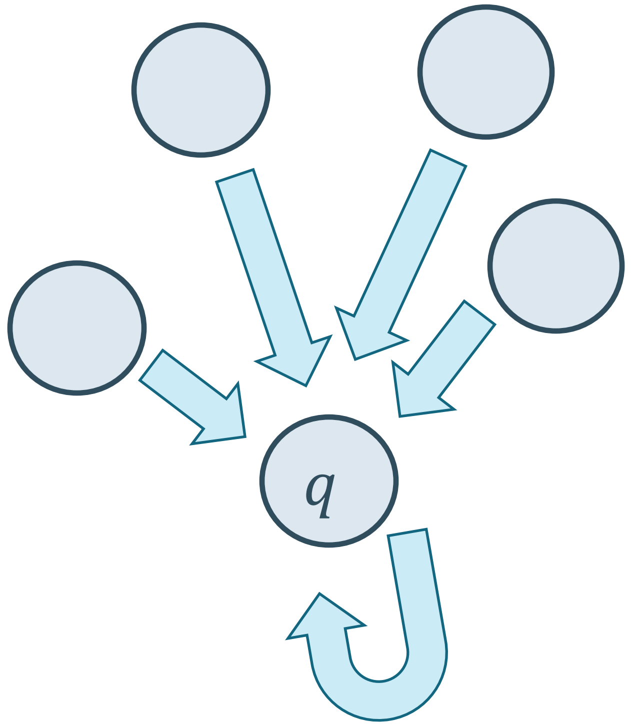
Local problems

Global and local problems

	Global	Local
LS:	$\underset{X}{\text{minimize}}\ \mathbb{E}[\|\|\mathbf{d} - X^T\mathbf{y}\|\|^2]$	$\underset{\widetilde{X}}{\text{minimize}}\ \mathbb{E}[\|\|\mathbf{d} - \widetilde{X}^T\widetilde{\mathbf{y}}\|\|^2]$
CCA:	$\underset{X,W}{\text{maximize}}\ \mathbb{E}[\text{tr}(X^T\mathbf{y}\mathbf{v}^T W)]$ $\text{subject to } \mathbb{E}[X^T\mathbf{y}\mathbf{y}^T X] = I_Q$ $\mathbb{E}[W^T\mathbf{v}\mathbf{v}^T W] = I_Q$	$\underset{\widetilde{X},\widetilde{W}}{\text{maximize}}\ \mathbb{E}[\text{tr}(\widetilde{X}^T\widetilde{\mathbf{y}}\widetilde{\mathbf{v}}^T\widetilde{W})]$ $\text{subject to } \mathbb{E}[\widetilde{X}^T\widetilde{\mathbf{y}}\widetilde{\mathbf{y}}^T\widetilde{X}] = I_Q$ $\mathbb{E}[\widetilde{W}^T\widetilde{\mathbf{v}}\widetilde{\mathbf{v}}^T\widetilde{W}] = I_Q$
PCA:	$\underset{X}{\text{maximize}}\ \mathbb{E}[\|\|X^T\mathbf{y}\|\|^2]$ $\text{subject to } X^T X = I_Q$	$\underset{\widetilde{X}}{\text{maximize}}\ \mathbb{E}[\|\|\widetilde{X}^T\widetilde{\mathbf{y}}\|\|^2]$ $\text{subject to } \widetilde{X}^T K\widetilde{X} = I_Q$

The DASF framework [1]

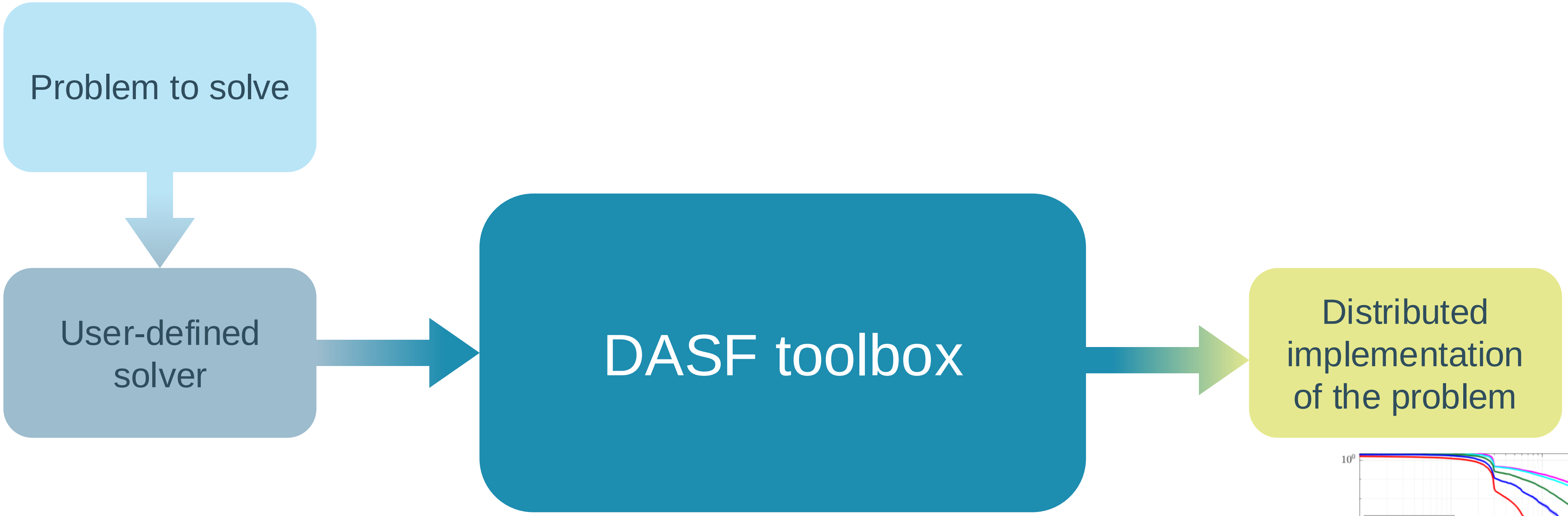
- Compress signals measured at nodes k using current estimate X^i of the filter: $\hat{\boldsymbol{y}}_k^i(t) = X_k^{iT}\boldsymbol{y}_k(t)$.
- Send them to node q.
- At node q, build a compressed version of the network-wide problem using the available local data and solve it to obtain new estimate X^{i+1}.
- Repeat for other nodes.
- Convergence to X^*, the global solution of the centralized problem (for any network topology) [1].

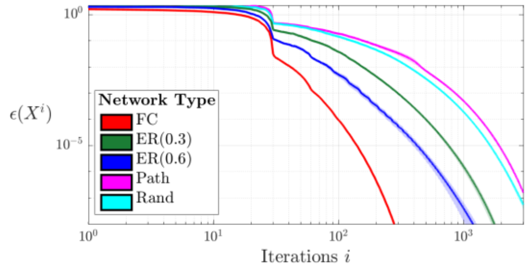


[1] C. A. Musluoglu and A. Bertrand, "A unified algorithmic framework for distributed adaptive signal and feature fusion problems — part I: Algorithm derivation," Submitted for publication - preprint available at https://arxiv.org/abs/2208.08867

Matlab and Python Toolbox

https://github.com/AlexanderBertrandLab/DASF_toolbox



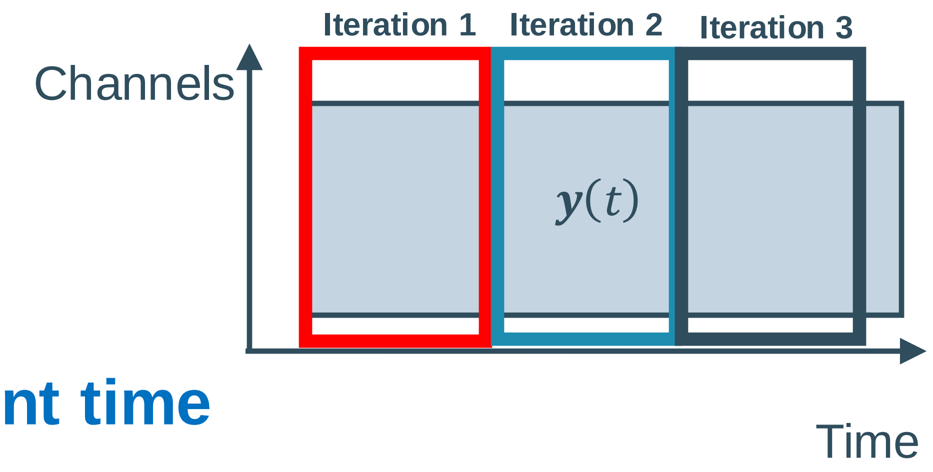


Practical setting

- Use samples of measured signals to estimate covariance matrices (spatial statistics)

- Divide signal into window of samples

- **Iterations of DASF spread out over different time windows (cfr. adaptive filters)**

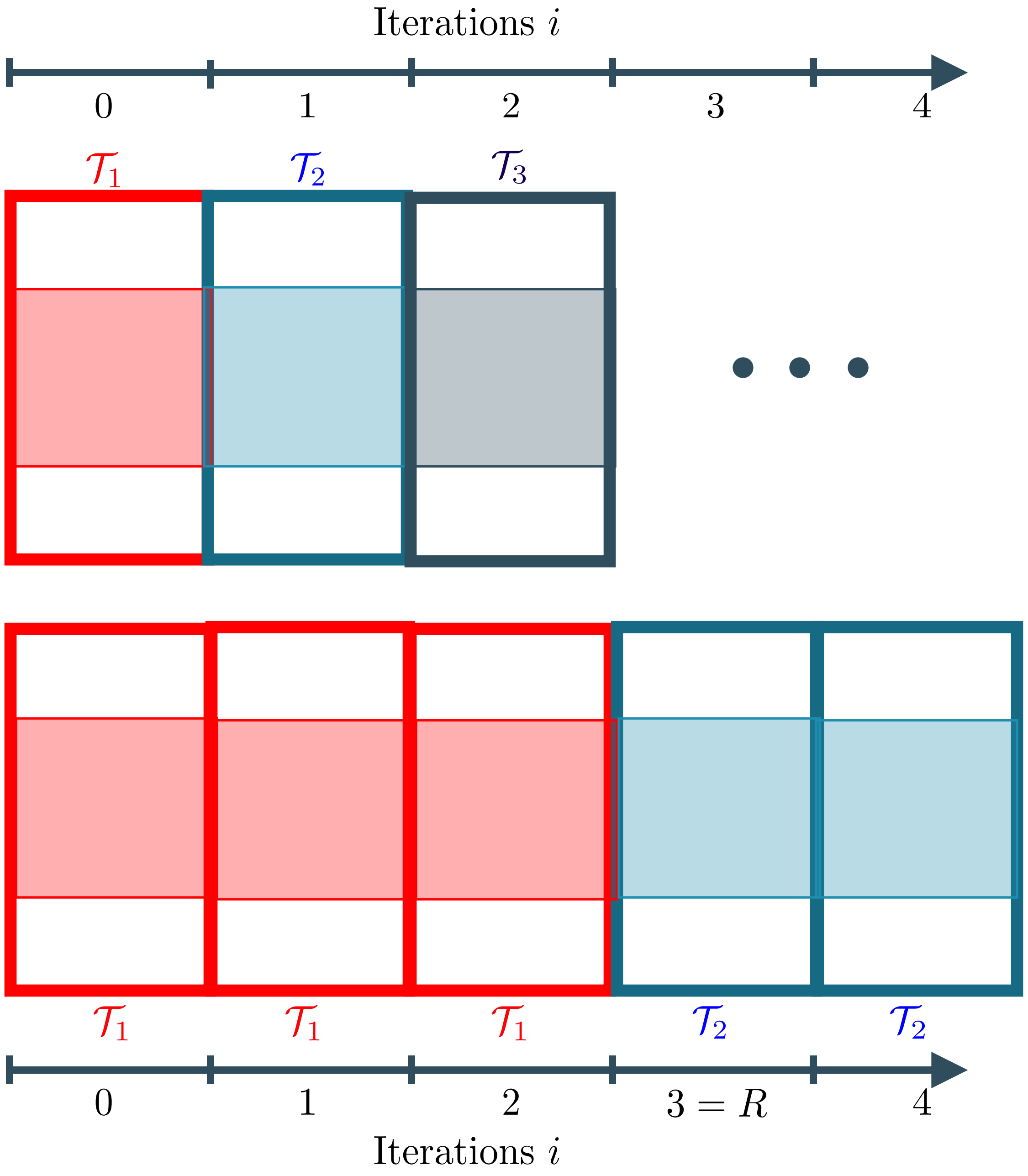


Accuracy and tracking problem if stationarity is broken across windows

Re-use the same data (window of samples) R times

+

Broadcast the signals across the network in an efficient way (see paper)



Communication efficient scheme

Original DASF	Straightforward approach	Efficient scheme
$\mathcal{O}(NKQ)$	$\mathcal{O}(NKQR)$	$\mathcal{O}(NQ(K+R-1))$

N: Number of time samples per window
K: Number of nodes in the network
Q: Number of columns of X
R: Number of times each window is used

Efficient bandwidth vs. accuracy trade-off:

R times faster convergence, only $\left(1+\frac{R-1}{K}\right)$ times more bandwidth

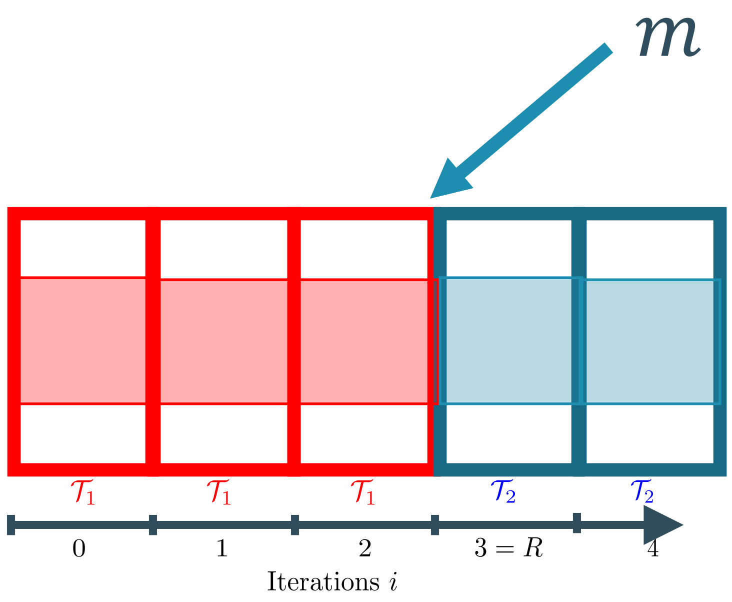
e.g. for $K = R = 10$: $10\ \times$ faster convergence for $< 2\ \times$ the bandwidth

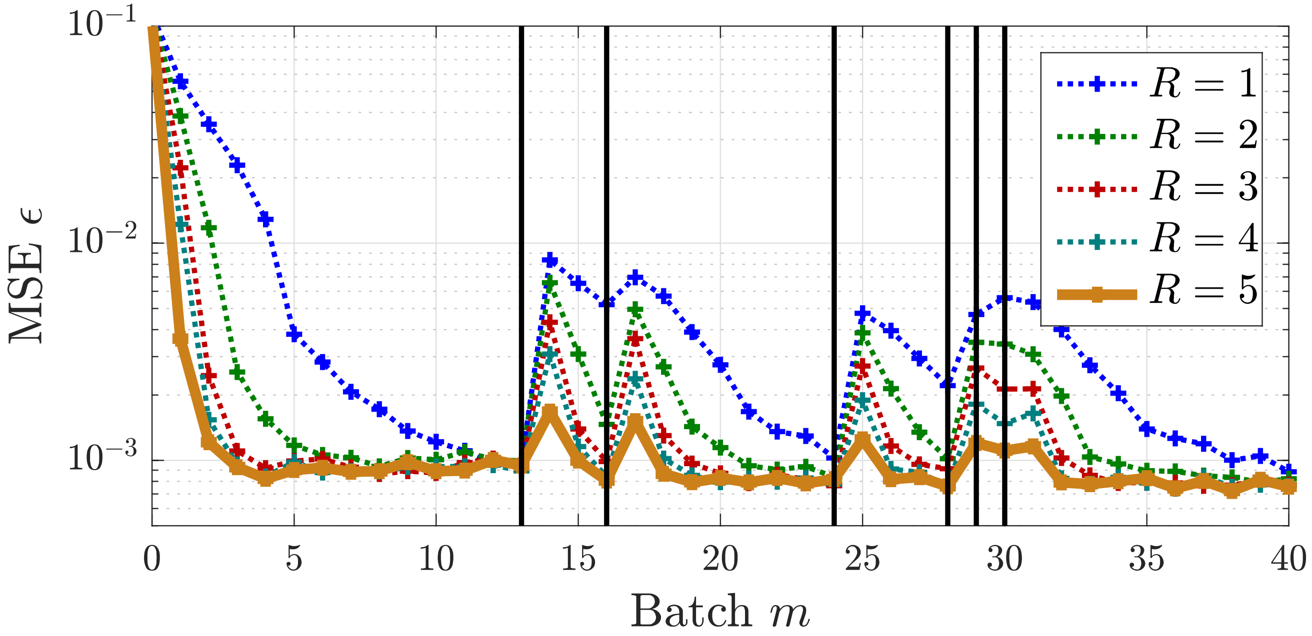
Results – LCMV beamforming

$$\underset{X\in\mathbb{R}^{M\times Q}}{\text{minimize}} \quad \mathbb{E}[||X^T\mathbf{y}(t)||^2] = \text{trace}(X^T K_{\mathbf{yy}} X)$$

$$\text{subject to} \quad X^T B = F^T.$$

$$\epsilon(m) = \frac{1}{MQ}||X^m - X^*||_F^2$$

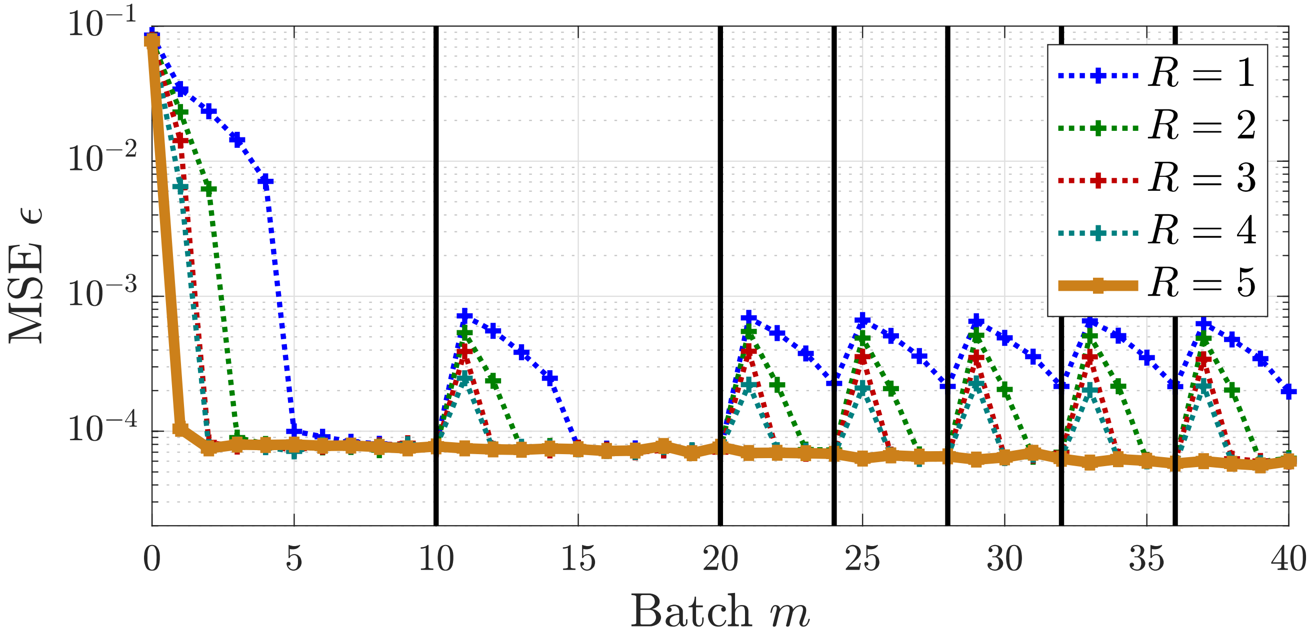




Results – PCA

$$\begin{aligned} \underset{\mathbf{x}\in\mathbb{R}^M}{\text{maximize}} \quad & \mathbb{E}\left[|\mathbf{x}^T\mathbf{y}(t)|^2\right] = \mathbf{x}^T K_{\mathbf{yy}}\mathbf{x} \\ \text{subject to} \quad & \mathbf{x}^T\mathbf{x} = 1. \end{aligned}$$

$$\epsilon(m) = \frac{1}{M}||\mathbf{x}^m - \mathbf{x}^*||^2$$



Conclusion

- In practical settings, the DASF framework can encounter accuracy and tracking problems

- Re-using each signal window multiple times

- Efficient bandwidth-vs-accuracy tradeoff method for fully-connected networks

Future work:

- Extension to any network topology

Thank you