Optimal Experimental Design for Sensor Placement and Acquisition of Highly-Correlated Data

Ahmed Attia and **Emil Constantinescu**

Mathematics and Computer Science Division (MCS), Argonne National Laboratory (ANL)

SIAM CSE-21
Texas, USA (*Virtual*)
March 5, 2021

Ahmed Attia, and Emil Constantinescu.
"Optimal Experimental Design for Inverse Problems in the Presence of Observation Correlations."
arXiv preprint arXiv:2007.14476 (2020).

Motivation: parameter identification

Consider the contaminant concentration u in domain $\Omega \in \mathbb{R}^2$

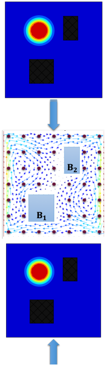
Advection-Diffusion

$$
\begin{aligned}
u_t - \kappa \Delta u + \mathbf{v} \cdot \nabla u &= 0 \quad \text{in } \mathcal{D} \times [0, T]\,, \\
u(x,\, 0) &= \theta \quad \text{in } \mathcal{D}\,, \\
\kappa \nabla u \cdot \mathbf{n} &= 0 \quad \text{on } \partial\mathcal{D} \times [0, T]
\end{aligned}
$$

Relevant SIAM CSE-21 Minisymposia:

MS154, MS214
MS195, MS230
MS320, MS351

- **Forward Problem:** given model state/parameter predict the expected model observations
- **Inverse Problem:** given spatiotemporal measurements, and a prior infer the **true QoI, e.g., function of the model state/parameter**
- **OED (sensor placement):** given a set of n_s candidate locations, determine the *optimal* positions, possibly under budget or sparsity constraints



Forward Problem and Bayesian Inversion

- **Forward problem:**

$$\mathbf{y} = \mathcal{F}(\theta) + \delta\,; \qquad \gamma = \mathbf{P}\theta$$

  - $\theta \in \mathbb{R}^{\mathrm{N_{state}}}$: discretized model parameter, e.g., IC
  - $\mathbf{y} \in \mathbb{R}^{\mathrm{N_{obs}}}$: spatiotemporal sensor observations
  - $\gamma \in \mathbb{R}^{\mathrm{N_{goal}}}$: goal/prediction/QoI

- **Bayesian inverse problem (goal-oriented):**
  - **The prior:** knowledge about the QoI γ prior to obtaining new observations

$$\gamma := \mathbf{P}\theta \sim \mathcal{N}(\mathbf{P}\theta_{\mathrm{pr}}, \mathbf{P}\mathbf{\Gamma}_{\mathrm{pr}}\mathbf{P}^*)$$

    where $\mathbf{P}$ here is a linear goal/prediction operator
  - **The likelihood:** Gaussian observation noise;

$$\mathcal{L}(\mathbf{y}|\theta) \propto \exp\left(-\frac{1}{2}\left\|\mathcal{F}(\theta) - \mathbf{y}\right\|^2_{\mathbf{\Gamma}^{-1}_{\mathrm{noise}}}\right); \qquad \|\mathbf{x}\|^2_{\mathbf{A}} = \mathbf{x}^{\mathsf{T}}\mathbf{A}\mathbf{x}$$

  - **The posterior**: distribution of the QoI γ conditioned on observations

Bayes' theorem: *Posterior* $\propto$ *Likelihood* $\times$ *Prior*

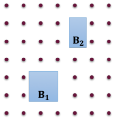
For a linear operator $\mathbf{F}$, the posterior is Gaussian $\mathcal{N}(\gamma_{\mathrm{post}}, \mathbf{\Sigma}_{\mathrm{post}})$ with

$$\gamma_{\mathrm{post}} = \mathbf{P}\mathbf{\Gamma}_{\mathrm{post}}\left(\mathbf{\Gamma}^{-1}_{\mathrm{pr}}\theta_{\mathrm{pr}} + \mathbf{F}^*\mathbf{\Gamma}^{-1}_{\mathrm{noise}}\,\mathbf{y}\right), \quad \mathbf{\Sigma}_{\mathrm{post}} = \underbrace{\mathbf{P}\left(\mathbf{F}^*\mathbf{\Gamma}^{-1}_{\mathrm{noise}}\mathbf{F} + \mathbf{\Gamma}^{-1}_{\mathrm{pr}}\right)^{-1}\mathbf{P}^*}_{\text{data-independent; only uncertainties}}$$

where $\mathbf{F}^*$, $\mathbf{P}^*$ is the adjoints of $\mathbf{F}$, $\mathbf{P}$

OED: sensor placement

Find the best subset (i.e., λ) of sensor location such as to optimize some utility function (e.g. identification accuracy, total uncertainties, information gain, etc.)



$$\xi := \left\{ \begin{matrix} \mathbf{x}_1, \ldots, \mathbf{x}_{n_s} \\ \zeta_1, \ldots, \zeta_{n_s} \end{matrix} \right\} \rightarrow \mathbf{W}(\zeta) \rightarrow$$

$\mathbf{\Gamma}_{\text{noise}}^{-1} \longrightarrow \mathbf{W}_{\Gamma}(\zeta)$ E.g.,

$$\mathbf{W}_{\Gamma}(\zeta) := \left\{ \begin{matrix} \mathbf{W}^{\frac{1}{2}}(\zeta)\mathbf{\Gamma}_{\text{noise}}^{-1}\mathbf{W}^{\frac{1}{2}}(\zeta), \text{ or,} \\ \mathbf{\Gamma}_{\text{noise}}^{-\frac{1}{2}}\mathbf{W}\mathbf{\Gamma}_{\text{noise}}^{-\frac{1}{2}} \end{matrix} \right.$$

- **Weighted-likelihood:**

$$\mathcal{L}(\mathbf{y}|\theta;\zeta) \propto \exp\left(-\frac{1}{2}\left\|\mathbf{F}(\theta) - \mathbf{y}\right\|^2_{\mathbf{W}_{\Gamma}(\zeta)}\right)$$

- **Weighted posterior covariance:**

$$\mathbf{\Sigma}_{\text{post}}(\zeta) = \mathbf{P}\left(\mathbf{F}^*\mathbf{W}_{\Gamma}(\zeta)\mathbf{F} + \mathbf{\Gamma}_{\text{pr}}^{-1}\right)^{-1}\mathbf{P}^*$$

- **OED optimization problem:**

Binary OED

$$\zeta^{\text{opt}} = \operatorname*{arg\,min}_{\zeta \in \{0,1\}^{n_s}} \mathcal{T}(\zeta) := \Psi(\zeta) + \alpha\,\Phi(\zeta)$$

Relaxed OED + Rounding

$$\zeta^{\text{opt}} = \operatorname*{arg\,min}_{\zeta \in [0,1]^{n_s}} \mathcal{T}(\zeta) := \Psi(\zeta) + \alpha\,\Phi(\zeta)$$

- Ψ: utility function; $\text{Tr}(\mathbf{\Sigma}_{\text{post}}) \rightarrow$ A-optimality, $\det(\mathbf{\Sigma}_{\text{post}}) \rightarrow$ D-optimality, etc.
- Φ: penalty function; ℓ_0, ℓ_1 etc.

OED: sensor placement

Solving OED optimization problem: gradient-based approach

- **A-optimality:** $\Psi := \Psi^{\mathrm{GA}}$

$$\frac{\partial \Psi^{\mathrm{GA}}}{\partial \zeta_i} = -\operatorname{Tr}\left(\mathbf{P} \left(\mathbf{H}(\zeta)\right)^{-1} \mathbf{F}^* \frac{\partial \mathbf{W}_{\Gamma}(\zeta)}{\partial \zeta_i} \mathbf{F} \left(\mathbf{H}(\zeta)\right)^{-1} \mathbf{P}^* \right)$$

- **D-optimality:** $\Psi := \Psi^{\mathrm{GD}}$

$$\frac{\partial \Psi^{\mathrm{GD}}}{\partial \zeta_i} = \operatorname{Tr}\left(\mathbf{\Sigma}_{\mathrm{post}}^{-1}(\zeta) \mathbf{P} \left(\mathbf{H}(\zeta)\right)^{-1} \mathbf{F}^* \frac{\partial \mathbf{W}_{\Gamma}(\zeta)}{\partial \zeta_i} \mathbf{F} \left(\mathbf{H}(\zeta)\right)^{-1} \mathbf{P}^* \right)$$

$$\mathbf{H}(\zeta) = \mathbf{\Gamma}_{\mathrm{pr}}^{-1} + \mathbf{F}^* \mathbf{W}_{\Gamma}(\zeta) \mathbf{F}$$

- *No observation correlations:*
  - $\mathbf{\Gamma}_{\mathrm{noise}}$ and $\mathbf{W}$ are diagonal
  - $\mathbf{W}_{\Gamma}(\zeta) = \mathrm{diag}\left(\zeta\right) \rightarrow \mathbf{W}^{\frac{1}{2}}(\zeta)\mathbf{\Gamma}_{\mathrm{noise}}^{-1}\mathbf{W}^{\frac{1}{2}}(\zeta) \equiv \mathbf{\Gamma}_{\mathrm{noise}}^{-\frac{1}{2}}\mathbf{W}\mathbf{\Gamma}_{\mathrm{noise}}^{-\frac{1}{2}}$
- **Spatiotemporal observation correlations**
  - A general/flexible approach to weight observation variances/covariances; discussed next:

Schur-product OED formulation

We formulate the weighted likelihood as:

$$\mathcal{L}(\mathbf{y}|\theta;\zeta) \propto \exp\left(-\frac{1}{2}\left\|\mathbf{F}(\theta)-\mathbf{y}\right\|^2_{\mathbf{W}_\Gamma(\zeta)}\right),$$

$$\mathbf{W}_\Gamma(\zeta) := \mathbf{L}^\mathsf{T}\widetilde{\mathbf{\Gamma}}^{-1}_{\rm noise}(\zeta)\mathbf{L}\,; \qquad \widetilde{\mathbf{\Gamma}}_{\rm noise}(\zeta) := \mathbf{L}\Big(\mathbf{\Gamma}_{\rm noise}\odot\mathbf{W}(\zeta)\Big)\mathbf{L}^\mathsf{T}\,.$$

- $\odot$ is the Hadamard (Schur) product
- Entries of the localization/weighting matrix $\mathbf{W}(\zeta)$ influence the observation correlations
- $\mathbf{W}(\zeta)$ is a symmetric and doubly nonnegative weighting kernel, with entries:

$$\varpi\left(t_k, t_l;\, \zeta_i, \zeta_j\right); \quad k,l=1,2,\ldots,n_t\,;\; i,j=1,2,\ldots,n_{\rm s}$$

  * ϖ is a symmeteric weighting (localization) function, such that:

$$\varpi\left(t_k, t_l;\, \zeta_i, \zeta_j\right) = \varpi\left(t_l, t_k;\, \zeta_i, \zeta_j\right) = \varpi\left(t_k, t_l;\, \zeta_j, \zeta_i\right) = \varpi\left(t_l, t_k;\, \zeta_j, \zeta_i\right)$$

Let $n_t=2,\, n_{\rm s}=2$

$$\mathbf{\Gamma}_{\rm noise} = \left[\begin{array}{c|c} \mathbf{R}_{1,1} & \mathbf{R}_{1,2} \\ \hline \mathbf{R}_{2,1} & \mathbf{R}_{2,2} \end{array}\right] \rightarrow \mathbf{W}(\zeta) := \left[\begin{array}{cc|cc} \varpi(t_1,t_1;\,\zeta_1,\zeta_1) & \varpi(t_1,t_1;\,\zeta_1,\zeta_2) & \varpi(t_1,t_2;\,\zeta_1,\zeta_1) & \varpi(t_1,t_2;\,\zeta_1,\zeta_2) \\ \varpi(t_1,t_1;\,\zeta_2,\zeta_1) & \varpi(t_1,t_1;\,\zeta_2,\zeta_2) & \varpi(t_1,t_2;\,\zeta_2,\zeta_1) & \varpi(t_1,t_2;\,\zeta_2,\zeta_2) \\ \hline \varpi(t_2,t_1;\,\zeta_1,\zeta_1) & \varpi(t_2,t_1;\,\zeta_1,\zeta_2) & \varpi(t_2,t_2;\,\zeta_1,\zeta_1) & \varpi(t_2,t_2;\,\zeta_1,\zeta_2) \\ \varpi(t_2,t_1;\,\zeta_2,\zeta_1) & \varpi(t_2,t_1;\,\zeta_2,\zeta_2) & \varpi(t_2,t_2;\,\zeta_2,\zeta_1) & \varpi(t_2,t_2;\,\zeta_2,\zeta_2) \end{array}\right]$$

Schur product OED formulation: space correlations

$$\mathbf{\Gamma}_{\text{noise}} = \bigoplus_{m=1}^{n_t} (\mathbf{R}_m) := \text{diag}\left(\mathbf{R}_1, \ldots, \mathbf{R}_{n_t}\right) , \quad \mathbf{L} = \bigoplus_{m=1}^{n_t} (\mathbf{L}_m) , \quad \mathbf{W}_\Gamma(\zeta) = \bigoplus_{m=1}^{n_t} \left(\mathbf{V}_m^{\dagger}(\zeta)\right) ,$$

$$\mathbf{V}_m^{\dagger}(\zeta) = \mathbf{L}_m^{\mathsf{T}} \mathbf{V}_m^{-1}(\zeta) \mathbf{L}_m , \qquad \mathbf{V}_m(\zeta) = \mathbf{L}_m \left(\mathbf{R}_m \odot \left(\sum_{i,j=1}^{n_s} \omega(\zeta_i,\, \zeta_j) \mathbf{e}_i \mathbf{e}_j^{\mathsf{T}} \right) \right) \mathbf{L}_m^{\mathsf{T}} ,$$

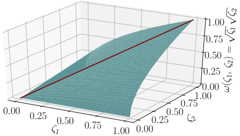
$$\Phi(\zeta) = \left\| \left(\omega(\zeta_1, \zeta_1),\, \ldots,\, \omega(\zeta_{n_s}, \zeta_{n_s}) \right)^{\mathsf{T}} \right\|_p$$

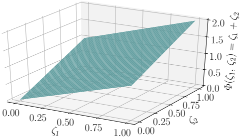
SQRT kernel:

$$\omega(\zeta_i, \zeta_j) = \sqrt{\zeta_i}\sqrt{\zeta_j};$$

$$\frac{\partial \omega(\zeta_i, \zeta_j)}{\partial \zeta_k} = \frac{1}{2} \left(\frac{\sqrt{\zeta_j}}{\sqrt{\zeta_i}} \delta_{i,k} + \frac{\sqrt{\zeta_i}}{\sqrt{\zeta_j}} \delta_{j,k} \right)$$

- $\zeta_i, \zeta_j \in [0, 1]$; $i, j = 1, 2, \ldots, n_s$,
- $\delta_{i,k}$: standard delta function
- $\zeta \in [0, 1] \rightarrow$ constrained-optimization (box-constraints)
- *Relates to traditional OED formulation*





Schur product OED formulation: space correlations

Solving Shur OED optimization problem: gradient-based approach

$$\frac{\partial \mathbf{W}_\Gamma(\zeta)}{\partial \zeta_i} = \frac{\partial \left(\mathbf{L}^\mathsf{T} \widetilde{\mathbf{\Gamma}}_{\text{noise}}^{-1}(\zeta) \mathbf{L}\right)}{\partial \zeta_i} = -\mathbf{L}^\mathsf{T} \widetilde{\mathbf{\Gamma}}_{\text{noise}}^{-1}(\zeta) \mathbf{L} \left(\mathbf{\Gamma}_{\text{noise}} \odot \frac{\partial \mathbf{W}(\zeta)}{\partial \zeta_i} \right) \mathbf{L}^\mathsf{T} \widetilde{\mathbf{\Gamma}}_{\text{noise}}^{-1}(\zeta) \mathbf{L}$$

$$= -\mathbf{W}_\Gamma(\zeta) \left(\mathbf{\Gamma}_{\text{noise}} \odot \frac{\partial \mathbf{W}(\zeta)}{\partial \zeta_i} \right) \mathbf{W}_\Gamma(\zeta)$$

$$\mathbf{W}' = [\eta_1, \, \eta_2, \ldots, \, \eta_{n_s}] \,, \qquad \eta_j = \left(\frac{1}{1+\delta_{1,j}} \frac{\partial \omega(\zeta_1, \zeta_j)}{\partial \zeta_j}, \, \ldots, \, \frac{1}{1+\delta_{n_s,j}} \frac{\partial \omega(\zeta_{n_s}, \zeta_j)}{\partial \zeta_j} \right)^\mathsf{T}$$

- **A-optimality:** $\zeta^{\text{opt}} = \underset{\zeta \in [0,\,1]^{n_s}}{\arg\min}\, \mathcal{T}(\zeta) := \operatorname{Tr}(\mathbf{\Sigma}_{\text{post}}(\zeta)) + \alpha\, \Phi(\zeta)$

$$\nabla_\zeta \Psi^{\text{GA}}(\zeta) = 2 \sum_{m=1}^{n_t} \text{diag}\left(\mathbf{V}_m^\dagger(\zeta) \mathbf{F}_{0,m} \mathbf{H}^{-1}(\zeta) \mathbf{P}^* \mathbf{P} \mathbf{H}^{-1}(\zeta) \mathbf{F}_{m,0}^* \mathbf{V}_m^\dagger(\zeta) \left(\mathbf{R}_m \odot \mathbf{W}' \right) \right)$$

- **D-optimality:** $\zeta^{\text{opt}} = \underset{\zeta \in [0,\,1]^{n_s}}{\arg\min}\, \mathcal{T}(\zeta) := \log\det(\mathbf{\Sigma}_{\text{post}}(\zeta)) + \alpha\, \Phi(\zeta)$

$$\nabla_\zeta \Psi^{\text{GD}}(\zeta) = 2 \sum_{m=1}^{n_t} \text{diag}\left(\mathbf{V}_m^\dagger(\zeta) \mathbf{F}_{0,m} \mathbf{H}^{-1}(\zeta) \mathbf{P}^* \mathbf{\Sigma}_{\text{post}}^{-1}(\zeta) \mathbf{P} \mathbf{H}^{-1}(\zeta) \mathbf{F}_{m,0}^* \mathbf{V}_m^\dagger(\zeta) \left(\mathbf{R}_m \odot \mathbf{W}' \right) \right)$$

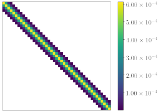
Experimental Settings

- Numerical model (AD): u solves:

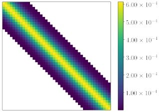
$$
\begin{aligned}
u_t - \kappa \Delta u + \mathbf{v} \cdot \nabla u &= 0 \quad \text{in } \mathcal{D} \times [0, T], \\
u(x, 0) &= \theta \quad \text{in } \mathcal{D}, \\
\kappa \nabla u \cdot \mathbf{n} &= 0 \quad \text{on } \partial\mathcal{D} \times [0, T]
\end{aligned}
$$

  * $\Omega \in \mathbb{R}^2$ is an open and bounded domain
  * u the concentration of a contaminant in the domain Ω
  * κ is the diffusivity,
  * $\mathbf{v}$ is the velocity field

- Observational setup: $n_s = 43$ candidate sensor locations, with
  * observation times $t_k := t_1 + s\Delta t$
  * $\Delta t = 0.2$ is the model simulation time step;
  * $t_1 = 1,\ s = 0, 1, \ldots, 5$
  * Observation correlations; synthetic, created with Gaspari-Cohn, and 5% noise level

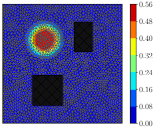


$\mathbf{R}_k$ with $\ell = 1$

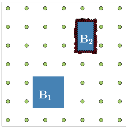


$\mathbf{R}_k$ with $\ell = 3$

- QoI: γ is the contaminant concentration predicted around the second building at time $t_p = 2.2$

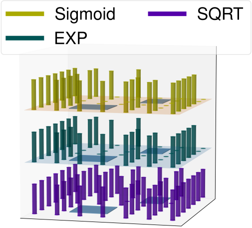


Ground truth (initial parameter) θ

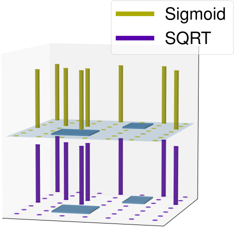


QoI (prediction) gridpoints.

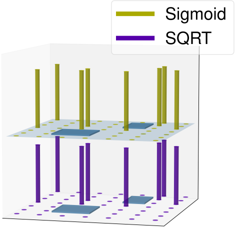
Numerical Results: space-correlation I



$\ell = 0$ (No correlation)



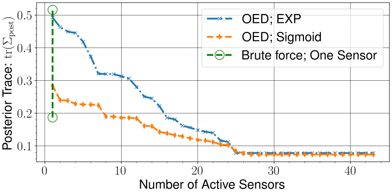
$\ell = 1$



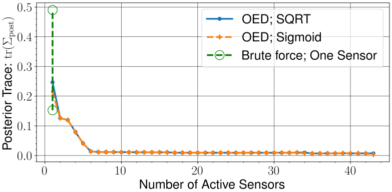
$\ell = 3$

Optimal design weights resulting from solving the OED problem, with $\alpha = 0$.
Sensors with weights above 0.5 are activated.

Numerical Results: space-correlation II



$\ell = 0.$



$\ell = 3.$

Objective value (post-trace) evaluated for increasing number of sensors, corresponding to highest optimal weights. Here, $\alpha = 0$.

Schur product OED formulation: space-time correlations

Entries of $\boldsymbol{\Gamma}_{\rm noise}^{-1}$ are weighted by

$$\varpi(\zeta_i, \zeta_j;\, t_m, t_n) := \rho(t_m,\, t_n)\, \omega(\zeta_i, \zeta_j)\,; \qquad \begin{matrix} i, j = 1, 2, \ldots, n_{\rm s}\,, \\ m, n = 1, 2, \ldots, n_t\,. \end{matrix}$$

$\rho(t_m,\, t_n)$ is a symmetric temporal decorrelation function:

- $\rho(t_m,\, t_m) = 1$
- $\rho(t_m,\, t_m)$ is conversely related with $d(t_m,\, t_n)$, the distance between t_m and t_n

---

- Gauss:

$$\rho(t_m,\, t_n) := \exp\left(\frac{-d(t_m,\, t_n)}{2\ell^2}\right)$$

- Gaspari-Cohn

$$\rho(t_m, t_n) = \begin{cases} -\frac{\upsilon^5}{4} + \frac{\upsilon^4}{2} + \frac{5\upsilon^3}{8} - \frac{5\upsilon^2}{3} + 1\,, & 0 \leq \upsilon \leq 1 \\ \frac{\upsilon^5}{12} - \frac{\upsilon^4}{2} + \frac{5\upsilon^3}{8} + \frac{5\upsilon^2}{3} - 5\upsilon + 4 - \frac{2}{3\upsilon}\,, & 1 \leq \upsilon \leq 2 \\ 0\,, & \upsilon \geq 2\,, \end{cases}$$

$\upsilon := \frac{d(t_m, t_n)}{\ell}$ and, ℓ is a predefined correlation length scale

Schur product OED formulation: space-time correlations

Solving Shur-OED optimization problem: gradient-based approach

$$\mathbf{W} := \left[\varpi \left(\zeta_{(k-1)\%n_{\mathrm{s}}+1}, \zeta_{(h-1)\%n_{\mathrm{s}}+1};\, t_{\left\lfloor \frac{k-1}{n_t} \right\rfloor +1}, t_{\left\lfloor \frac{h-1}{n_t} \right\rfloor +1} \right) \right]_{k,h=1,2,\ldots,\mathrm{N_{obs}}}$$

$$\vartheta_{i,m}[k] := \frac{1}{1+\delta_{i,\,(k-1)\%n_{\mathrm{s}}+1}} \frac{\partial}{\partial \zeta_i} \varpi \left(\zeta_i, \zeta_{(k-1)\%n_{\mathrm{s}}+1};\, t_m, t_{\left\lfloor \frac{k-1}{n_t} \right\rfloor +1} \right) ; \qquad \begin{array}{c} i = 1, \ldots, n_{\mathrm{s}}\,, \\ m = 1, \ldots, n_t\,, \\ k = 1, \ldots, \mathrm{N_{obs}} \end{array}$$

$$\frac{\partial \mathbf{W}_\Gamma}{\partial \zeta_i} = -\mathbf{W}_\Gamma(\zeta) \left[\sum_{m=1}^{n_t} \mathbf{e}_q \left(\left(\mathbf{\Gamma}_{\mathrm{noise}}^{-1} \mathbf{e}_q \right) \odot \vartheta_{i,m} \right)^{\mathsf{T}} + \sum_{m=1}^{n_t} \left(\left(\mathbf{\Gamma}_{\mathrm{noise}}^{-1} \mathbf{e}_q \right) \odot \vartheta_{i,m} \right) \mathbf{e}_q^{\mathsf{T}} \right] \mathbf{W}_\Gamma(\zeta)\,, \quad \begin{array}{c} q = i + (m-1) n_{\mathrm{s}}\,, \\ i = 1, \ldots, n_{\mathrm{s}} \end{array}$$

- A-optimality: $\Psi := \Psi^{\mathrm{GA}}$

$$\nabla_\zeta \Psi^{\mathrm{GA}}(\zeta) = 2 \sum_{i=1}^{n_{\mathrm{s}}} \sum_{m=1}^{n_t} \mathbf{e}_i \mathbf{e}_q^{\mathsf{T}} \mathbf{W}_\Gamma(\zeta) \mathbf{F} \mathbf{H}^{-1}(\zeta) \mathbf{P}^* \mathbf{P} \mathbf{H}^{-1}(\zeta) \mathbf{F}^* \mathbf{W}_\Gamma(\zeta) \left(\left(\mathbf{\Gamma}_{\mathrm{noise}}^{-1} \mathbf{e}_q \right) \odot \vartheta_{i,m} \right)$$

- D-optimality: $\Psi := \Psi^{\mathrm{GD}}$

$$\nabla_\zeta \Psi^{\mathrm{GD}}(\zeta) = 2 \sum_{i=1}^{n_{\mathrm{s}}} \sum_{m=1}^{n_t} \mathbf{e}_i \left(\mathbf{\Gamma}_{\mathrm{noise}}^{-1} \mathbf{e}_q \odot \vartheta_{i,m} \right)^{\mathsf{T}} \mathbf{W}_\Gamma(\zeta) \mathbf{F} \mathbf{H}^{-1}(\zeta) \mathbf{P}^* \mathbf{\Sigma}_{\mathrm{post}}^{-1}(\zeta) \mathbf{P} \mathbf{H}^{-1}(\zeta) \mathbf{F}^* \mathbf{W}_\Gamma(\zeta) \mathbf{e}_q$$

Schur product OED formulation

Efficient evaluation of the OED objective Ψ:

- reduced-order approximation of $\mathbf{F}$
- randomized approximation of Ψ

A-optimality: randomized-trace approximation: recall: $\mathbf{W}_{\Gamma}(\zeta) = \mathbf{L}^{\mathsf{T}} \left(\mathbf{L} \left(\mathbf{\Gamma}_{\text{noise}} \odot \mathbf{W}(\zeta) \right) \mathbf{L}^{\mathsf{T}} \right)^{-1} \mathbf{L}$

$$\Psi^{\text{GA}}(\zeta) \approx \widetilde{\Psi}^{\text{GA}}(\zeta) = \frac{1}{n_r} \sum_{r=1}^{n_r} \mathbf{z}_r^{\mathsf{T}} \mathbf{P} \left(\mathbf{F}\, \mathbf{W}_{\Gamma}(\zeta)\, \mathbf{F}^* + \mathbf{\Gamma}_{\text{pr}}^{-1} \right)^{-1} \mathbf{P}^* \mathbf{z}_r \,, \qquad \mathbf{z}_r \in \mathbb{R}^{\mathrm{N}_{\text{goal}}}$$

$$\frac{\partial \widetilde{\Psi}^{\text{GA}}(\zeta)}{\partial \zeta_i} = \frac{1}{n_r} \sum_{r=1}^{n_r} \mathbf{z}_r^{\mathsf{T}} \mathbf{P} \mathbf{H}^{-1}(\zeta) \mathbf{F}^* \mathbf{W}_{\Gamma}(\zeta) \left(\mathbf{\Gamma}_{\text{noise}} \odot \frac{\partial \mathbf{W}(\zeta)}{\partial \zeta_i} \right) \mathbf{W}_{\Gamma}(\zeta) \mathbf{F} \mathbf{H}^{-1}(\zeta) \mathbf{P}^* \mathbf{z}_r$$

- space correlations:

$$\nabla_{\zeta} \widetilde{\Psi}^{\text{GA}}(\zeta) = \frac{2}{n_r} \sum_{r=1}^{n_r} \sum_{m=1}^{n_t} \psi_{r,m}^* \odot \left(\left(\mathbf{R}_m \odot \mathbf{W}' \right) \psi_{r,m} \right) \; ; \qquad \begin{aligned} \psi_{r,m} &:= \mathbf{V}_m^{\dagger}(\zeta) \mathbf{F}_{0,m} \mathbf{H}^{-1}(\zeta) \mathbf{P}^* \mathbf{z}_r \,, \\ \psi_{r,m}^* &:= \left(\mathbf{z}_r^{\mathsf{T}} \mathbf{P} \mathbf{H}^{-1}(\zeta) \mathbf{F}_{m,0}^* \mathbf{V}_m^{\dagger}(\zeta) \right)^{\mathsf{T}} \end{aligned}$$

- space-time correlations:

$$\nabla_{\zeta} \widetilde{\Psi}^{\text{GA}}(\zeta) = 2 \sum_{r=1}^{n_r} \sum_{i=1}^{n_s} \sum_{m=1}^{n_t} \mathbf{e}_i \psi_r^* \mathbf{e}_q \left(\left(\mathbf{\Gamma}_{\text{noise}} \mathbf{e}_q \right) \odot \vartheta_{i,m} \right)^{\mathsf{T}} \psi_r \; ; \qquad \begin{aligned} \psi_r &:= \mathbf{W}_{\Gamma}(\zeta) \mathbf{F} \mathbf{H}^{-1}(\zeta) \mathbf{P}^* \mathbf{z}_r \,, \\ \psi_r^* &:= \mathbf{z}_r^{\mathsf{T}} \mathbf{P} \mathbf{H}^{-1}(\zeta) \mathbf{F}^* \mathbf{W}_{\Gamma}(\zeta) \end{aligned}$$

References

1. *Further details:*
   Ahmed Attia, and Emil Constantinescu. "Optimal Experimental Design for Inverse Problems in the Presence of Observation Correlations." arXiv preprint arXiv:2007.14476 (2020).
   - → Additional weighting kernels
   - → Randomized trace approximation
   - → More empirical results
   - → D-optimality extension is straight forward
2. *Recent work; stochastic learning for binary OED* **(Differentiability w.r.t. ζ is NOT required)**:
   Ahmed Attia, Sven Leyffer, and Todd Munson. "Stochastic Learning Approach to Binary Optimization for Optimal Design of Experiments." arXiv preprint arXiv:2101.05958 (2021).

U.S. DEPARTMENT OF ENERGY | Office of Science

- These slides are available
  https://www.mcs.anl.gov/~attia/conferences.html
- Questions are welcome
  attia@mcs.anl.gov