SAND2024-01952C

Sandia National Laboratories

Exceptional service in the national interest

Schwarz-based Domain Decomposition Solutions for Data-Driven and Physics-Informed Models

SIAM UQ 2024

Chris Wentland, Francesco Rizzi,
Joshua Barnett, Irina Tezaur

February 28, 2024

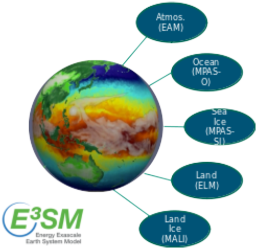
Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.
SAND NO. 2024-01952C

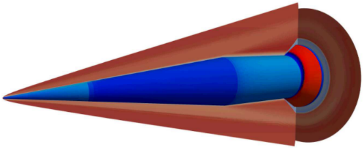
U.S. DEPARTMENT OF ENERGY

NNSA National Nuclear Security Administration

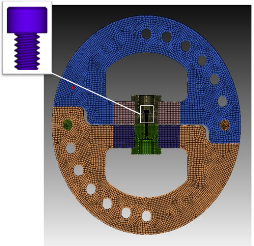
Multiscale, multiphysics systems

- Practical physical system solvers composed from subsystem modules
  - Governing equations (fluid, solid, multi-phase, radiation)
  - Mesh type, refinement (quad/hex, tri/tet, geodesic)
  - Time integrators, time step (implicit, explicit, IMEX)
- Coupling is labor-intensive, often reliant on empirical models





Blonigan *et al.*, 2022



Murugesen *et al.*, 2020

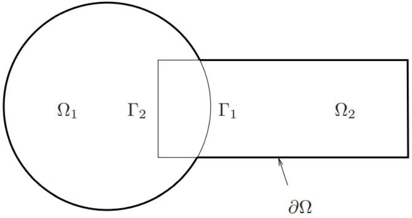
Overview

- The Schwarz alternating method
- Projection-based reduced-order models
- 2D shallow water equations – Gaussian pulse
  - Additive Schwarz parallelism
  - "Non-overlapping" finite-volume Schwarz
  - Hyper-reduction with domain decomposition
- Conclusions and outlook

The Schwarz alternating method[1,2]

- Monolithic (non-linear) ODE

$$\frac{d\mathbf{q}}{dt} = \mathbf{f}\left(\mathbf{q},\ t,\ \boldsymbol{\mu}\right) \quad \text{in } \Omega, \quad \mathbf{q} \in \mathbb{R}^N$$



- Decompose solution on **overlapping** $\Omega_1,\ \Omega_2$

$$\begin{cases} \dfrac{d\mathbf{q}_1^{k+1}}{dt} = \mathbf{f}\left(\mathbf{q}_1^{k+1},\ t,\ \boldsymbol{\mu}_1\right) & \text{in } \Omega_1 \\ \mathbf{q}_1^{k+1} = \mathbf{f}_b\left(\mathbf{q}_1^{k+1}\right) & \text{on } \partial\Omega_1 \setminus \Gamma_1 \\ \mathbf{q}_1^{k+1} = \mathbf{q}_2^{k} & \text{on } \Gamma_1 \end{cases} \qquad \begin{cases} \dfrac{d\mathbf{q}_2^{k+1}}{dt} = \mathbf{f}\left(\mathbf{q}_2^{k+1},\ t,\ \boldsymbol{\mu}_2\right) & \text{in } \Omega_2 \\ \mathbf{q}_2^{k+1} = \mathbf{f}_b\left(\mathbf{q}_2^{k+1}\right) & \text{on } \partial\Omega_2 \setminus \Gamma_2 \\ \mathbf{q}_2^{k+1} = \mathbf{q}_1^{k+1} & \text{on } \Gamma_2 \end{cases}$$

---

[1] Schwarz, H.A., *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich*, Vol. 15, 1870.

[2] Gander, *Electron. Trans. Numer. Anal*, Vol. 31, 2008.

Additive Schwarz

- Multiplicative Schwarz is strictly **serial**
- Additive: advance independently, communicate afterward

$$\begin{aligned} \frac{d\mathbf{q}_i^{k+1}}{dt} &= \mathbf{f}\left(\mathbf{q}_i^{k+1},\ t,\ \boldsymbol{\mu}_i\right) && \text{in } \Omega_i \\ \mathbf{q}_i^{k+1} &= \mathbf{f}_b\left(\mathbf{q}_i^{k+1}\right) && \text{on } \partial\Omega_i \setminus \Gamma_i \\ \mathbf{q}_i^{k+1} &= \mathbf{q}_j^{k} && \text{on } \Gamma_i \end{aligned}$$

- Empirically observed to slow convergence, no loss of accuracy

All results in this presentation use the **additive** Schwarz algorithm.
Parallelism helps balance additional cost due to Schwarz iterations.

Non-overlapping Schwarz

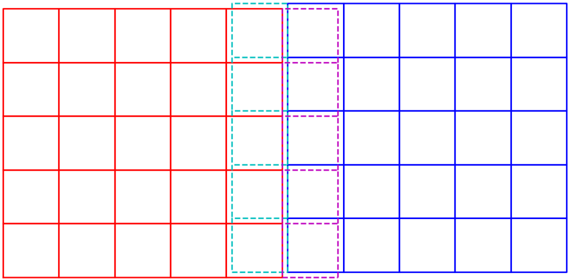
- Overlapping interfaces are uncommon in the physical world
  - Mechanical assemblies, multi-physics systems
- Non-overlapping Dirichlet–Dirichlet condition has no convergence guarantees
- **Dirichlet–Neumann** conditions resolve this issue

$$\begin{cases} \dfrac{d\mathbf{q}_1^{k+1}}{dt} = \mathbf{f}\left(\mathbf{q}_1^{k+1},\ t,\ \boldsymbol{\mu}_1\right) & \text{in } \Omega_1 \\ \mathbf{q}_1^{k+1} = \mathbf{f}_b\left(\mathbf{q}_1^{k+1}\right) & \text{on } \partial\Omega_1 \setminus \Gamma_1 \\ \mathbf{q}_1^{k+1} = \mathbf{q}_2^{k} & \text{on } \Gamma_1 \end{cases} \qquad \begin{cases} \dfrac{d\mathbf{q}_2^{k+1}}{dt} = \mathbf{f}\left(\mathbf{q}_2^{k+1},\ t,\ \boldsymbol{\mu}_2\right) & \text{in } \Omega_2 \\ \mathbf{q}_2^{k+1} = \mathbf{f}_b\left(\mathbf{q}_2^{k+1}\right) & \text{on } \partial\Omega_2 \setminus \Gamma_2 \\ \dfrac{\partial\mathbf{q}_2^{k+1}}{\partial\mathbf{n}} = \dfrac{\partial\mathbf{q}_1^{k+1}}{\partial\mathbf{n}} & \text{on } \Gamma_2 \end{cases}$$

- Robin–Robin also converges, but rare in production codes

A note on finite volume methods

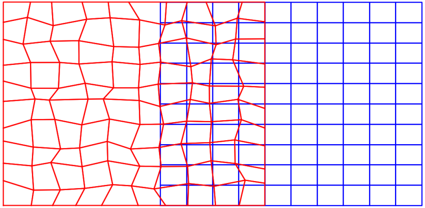
- Today's results use **cell-centered** FVM with **ghost cell** boundary conditions
  - BCs are enforced **approximately** by fictitious ghost cell state
- Can't apply Dirichlet–Neumann or Robin–Robin without significant effort
- Some literature suggests Dirichlet–Dirichlet works for hyperbolic problems[3]
- **Not truly non-overlapping**, but non-physical cells suggest interesting possibilities

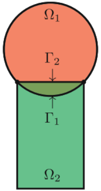


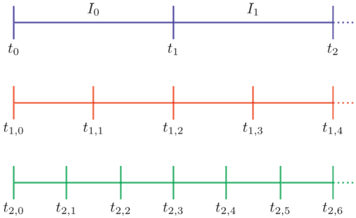
[3] Dolean et al., *SIAM J Sci Comput*, 2009

Not covered in this talk

- Non-conformal meshes
  - Interpolation/projection of transmission condition[4]
- Non-uniform time steps
  - "Controller" time-step concept[5]
- Physics-informed neural network (PINN) coupling[6]
  - Didn't work as well as we hoped







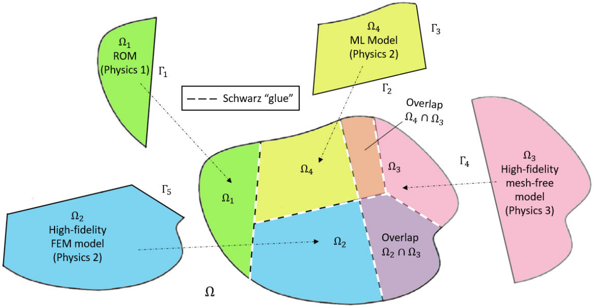
[4] Mota et al., *arXiv:2311.05643*, 2023

[5] Mota et al., *Int J Numer Meth Eng*, Vol. 123, 2022

[6] Snyder et al., *CSRI Summer Proceedings*, 2023

Data-driven methods

- Accelerate workflows for **many-query problems**
- Coupling traditional numerical methods with
  - Projection-based ROMs (PROMs)
  - Operator inference ROMs
  - Neural network surrogates
- Schwarz lends to "plug-and-play" architectures



Projection-based reduced-order models (PROMs)

- Recall governing (non-linear) ODE

$$\frac{d\mathbf{q}}{dt} = \mathbf{f}\left(\mathbf{q},\ t,\ \boldsymbol{\mu}\right), \quad \mathbf{q} \in \mathbb{R}^N$$

- Low-dimensional affine representation ($K \ll N$)

$$\mathbf{q}(t) \approx \widetilde{\mathbf{q}}(t) = \overline{\mathbf{q}} + \mathbf{\Phi}\widehat{\mathbf{q}}(t)$$

$$\widehat{\mathbf{q}} \in \mathbb{R}^K, \quad \mathbf{\Phi} \in \mathbb{R}^{N\times K}$$

- The proper orthogonal decomposition

$$\mathbf{\Phi} = \underset{\mathbf{A}\in\mathbb{R}^{N\times K}}{argmin} \left\|\mathbf{Q} - \mathbf{A}\mathbf{A}^\top\mathbf{Q}\right\|_2^2$$

$$\mathbf{Q} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top, \quad \mathbf{\Phi} \leftarrow \mathbf{U}[:,:K]$$

Least-squares Petrov–Galerkin[7]

- No dimension reduction yet, project by test basis $\mathbf{\Psi} \in \mathbb{R}^{N \times K}$ and rearrange

$$\frac{d\widehat{\mathbf{q}}}{dt} = \left[\mathbf{\Psi}^\top \mathbf{\Phi}\right]^{-1} \mathbf{\Psi}^\top \mathbf{f}\left(\overline{\mathbf{q}} + \mathbf{\Phi}\widehat{\mathbf{q}}\right)$$

- For fully-discrete residual (e.g. BDF1)

$$\mathbf{r}\left(\widehat{\mathbf{q}}^n\right) = \mathbf{\Phi}\left(\widehat{\mathbf{q}}^n - \widehat{\mathbf{q}}^{n-1}\right) - \Delta t \mathbf{f}\left(\overline{\mathbf{q}} + \mathbf{\Phi}\widehat{\mathbf{q}}^n\right) = \mathbf{0}$$

- Minimize with Gauss–Newton

$$\delta\widehat{\mathbf{q}}^n = \underset{\mathbf{y} \in \mathbb{R}^K}{argmin} \left\| \frac{\partial \mathbf{r}\left(\mathbf{y}\right)}{\partial \mathbf{y}} \mathbf{y} + \mathbf{r}\left(\mathbf{y}\right) \right\|_2^2$$

- Equivalent to Petrov–Galerkin projection with test basis

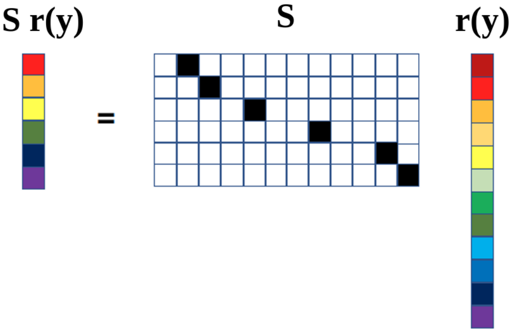
$$\mathbf{\Psi} = \frac{\partial \mathbf{r}\left(\widehat{\mathbf{q}}\right)}{\partial \mathbf{q}} \mathbf{\Phi}$$

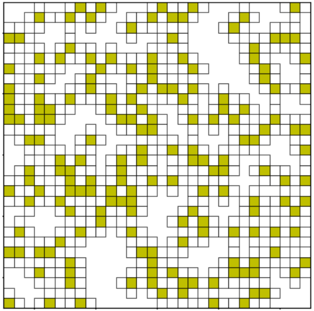
[7] Carlberg et al., *Int J Numer Meth Eng*, Vol. 86, 2011.

Hyper-reduction

- Non-linear residual still scales with full dimension N
- Collocation: minimize residual at **small subset** of DOFs, $N_s \ll N$

$$\widehat{\mathbf{q}}^n = \underset{\mathbf{y}\in\mathbb{R}^K}{argmin} \left\|\mathbf{S}\mathbf{r}\left(\mathbf{y}\right)\right\|_2^2$$





Yellow: residual cells, White: stencil cells

Open-source software

- Projection-based ROM interface provided by **pressio**

  https://github.com/Pressio/pressio

- Experiments run with **pressio-demoapps**

  https://github.com/Pressio/pressio-demoapps

- Schwarz coupling for **pressio-demoapps**

  https://github.com/cwentland0/pressio-demoapps-schwarz

Rizzi, F. *et al.*, "Pressio: Enabling Projection-based Model Reduction for Large-scale Nonlinear Dynamical Systems," arXiv:2003.07798v3, 2021.

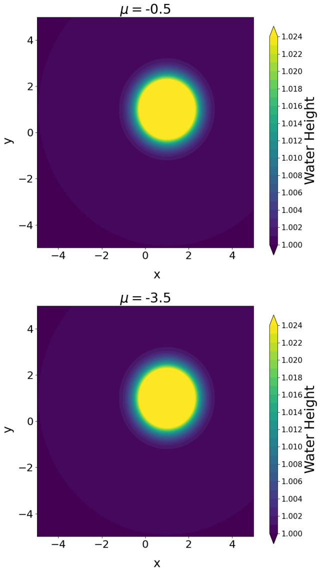
2D shallow water equations: Gaussian pulse

$$\frac{\partial h}{\partial t} + \frac{\partial (hu)}{\partial x} + \frac{\partial (hv)}{\partial y} = 0$$

$$\frac{\partial (hu)}{\partial t} + \frac{\partial}{\partial x}\left(hu^2 + \frac{1}{2}gh^2\right) + \frac{\partial}{\partial y}(huv) = -\mu v$$

$$\frac{\partial (hv)}{\partial t} + \frac{\partial}{\partial x}(huv) + \frac{\partial}{\partial y}\left(hv^2 + \frac{1}{2}gh^2\right) = \mu u$$

- Uniform 300×300 Cartesian grid
- First-order in space and time
- Vary "Coriolis" parameter μ
  - $\mu_{\text{train}} \in [-4.0,\ -3.0,\ -2.0,\ -1.0,\ 0.0]$
  - $\mu_{\text{test}} \in [-3.5,\ -2.5,\ -1.5,\ -0.5]$
- $T = 10.0$, $\Delta t = 0.01$

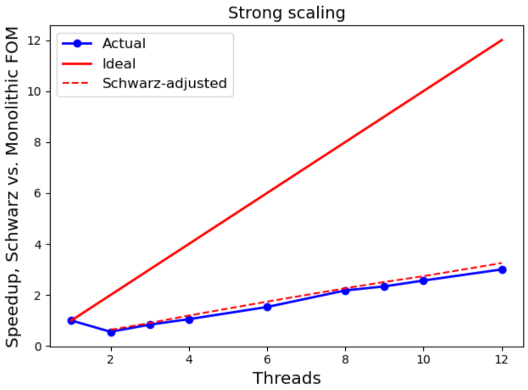


2D SWE: Parallelism

- Straightforward OpenMP parallelism, one thread per subdomain
- Schwarz iterations incur significant costs ($n_t \approx 3.5$ iterations on average here)
- Adjusting ideal scaling by Schwarz iterations shows parallelism works as intended

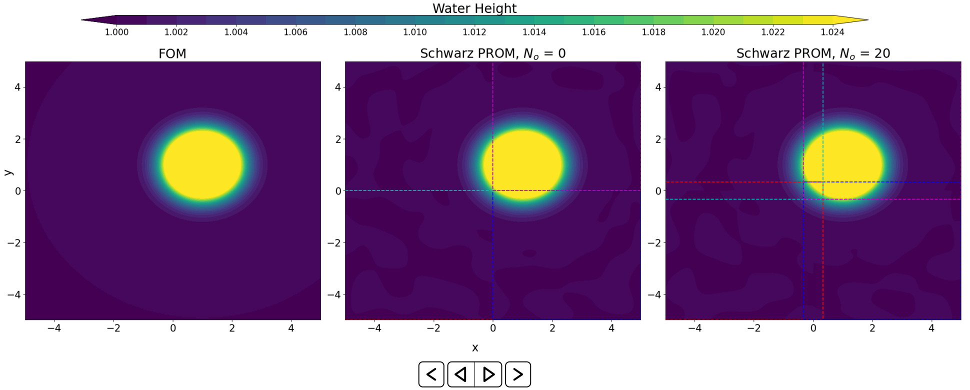
$$\text{Adjusted} = \frac{\#\ \text{threads}}{\frac{1}{N_t}\sum_{i=1}^{N_t} n_{t,i}}$$

$$n_t := \#\ \text{Schwarz iterations}$$



2D SWE: PROMs

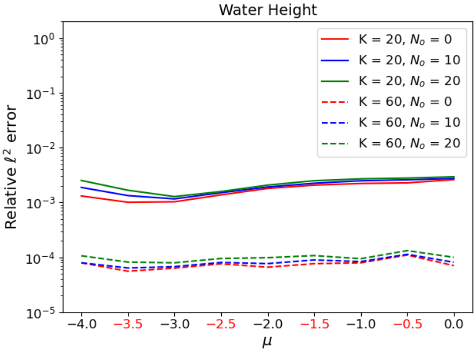
- $N_o :=$ cell width of overlap region (not including ghost cells)
- "Non-overlapping" Dirichlet–Dirichlet FVM performs well, no convergence issues

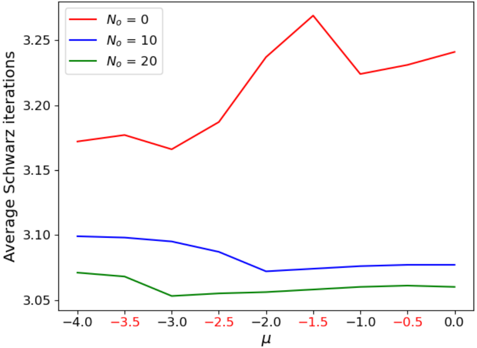


μ = -0.5, K = 80

2D SWE: PROMs, domain overlap

- Decreasing overlap has no effect or improves solution
  - Minimizing residual on trial space **does not guarantee matching** in overlap region
- Schwarz iterations decrease (very roughly) with $N_o^{0.25}$, but $\mathbf{r}(\mathbf{q})$ scales with N_o^2

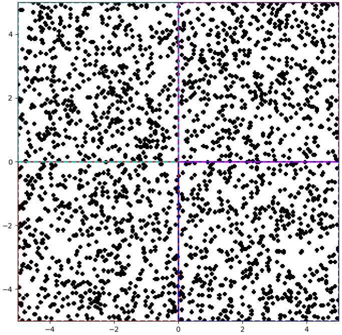




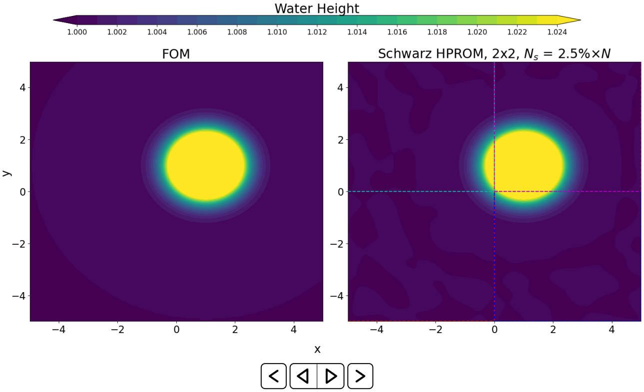
Black μ: training set, Red μ: testing set

2D SWE: Hyper-reduced PROMs (HPROMs)

- Naive random sampling invariably results in an unstable solution
- Sparsely-sampled boundary fails to transmit strong gradients



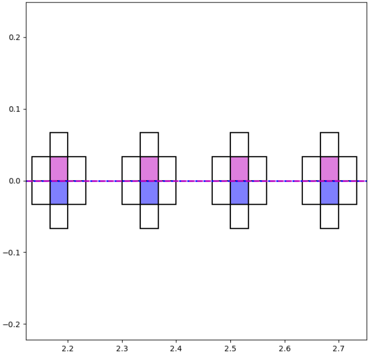
$N_s = 2.5\% \times N$



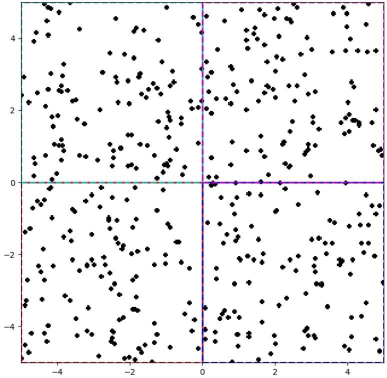
μ = -0.5, K = 100

2D SWE: Hyper-reduced PROM boundary sampling

- Sample boundaries at fixed interval N_b
- For fixed budget N_s, boundary samples draw points away from interior

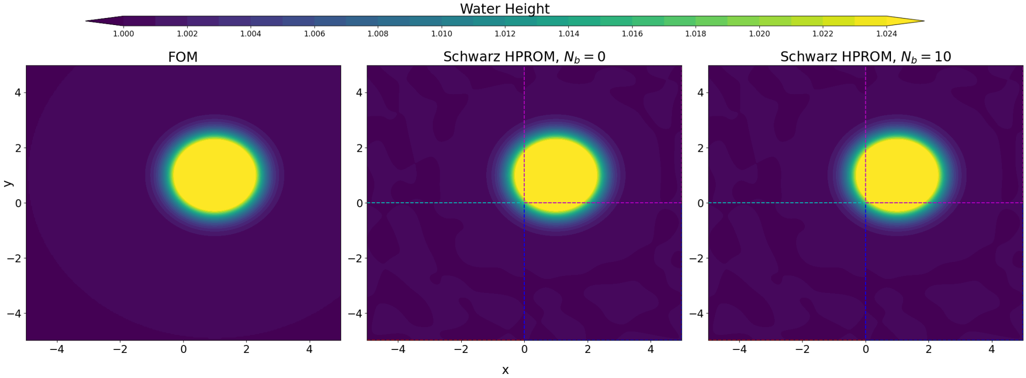


$N_b = 5$, boundary zoom view



2D SWE: Hyper-reduced PROM boundary sampling

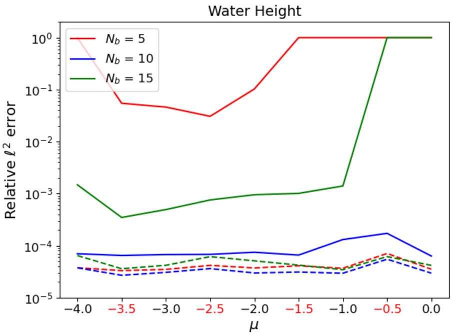
- Can often get away with $< 10\%$ boundary sampling

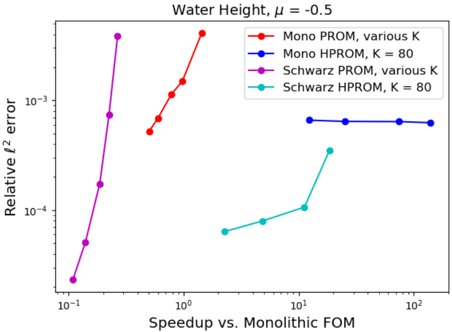


μ = -0.5, K = 100, $N_s = 0.5\% \times N$

2D SWE: Hyper-reduced PROM performance

- Balancing act between capturing dynamics and boundary transmission
- Noticeable cost savings **relative to monolithic FOM**
- For fixed PROM dimension K, Schwarz delivers lower error and comparable cost





Solid: $N_s = 0.5\% \times N$, Dash: $N_s = 1.0\% \times N$

Lower-right is better

Conclusions

- Schwarz provides a flexible framework for coupling disparate models
  - Difficult to recover traditional parallel performance due to Schwarz iterations
- "Non-overlapping" Dirichlet–Dirichlet FVM treatment performs well
  - Requires extension and verification for non-conformal meshes
- Schwarz HPROMs require special attention to boundary treatment
  - Fixed-interval boundary sampling method works, but is not cost-effective
- Ongoing work
  - "Optimal" spatial decomposition
  - Online multi-fidelity modeling adaptivity
  - Extensions to production Sandia applications

Acknowledgements

Francesco Rizzi

Joshua Barnett

Irina Tezaur

Alejandro Mota

LABORATORY DIRECTED RESEARCH & DEVELOPMENT