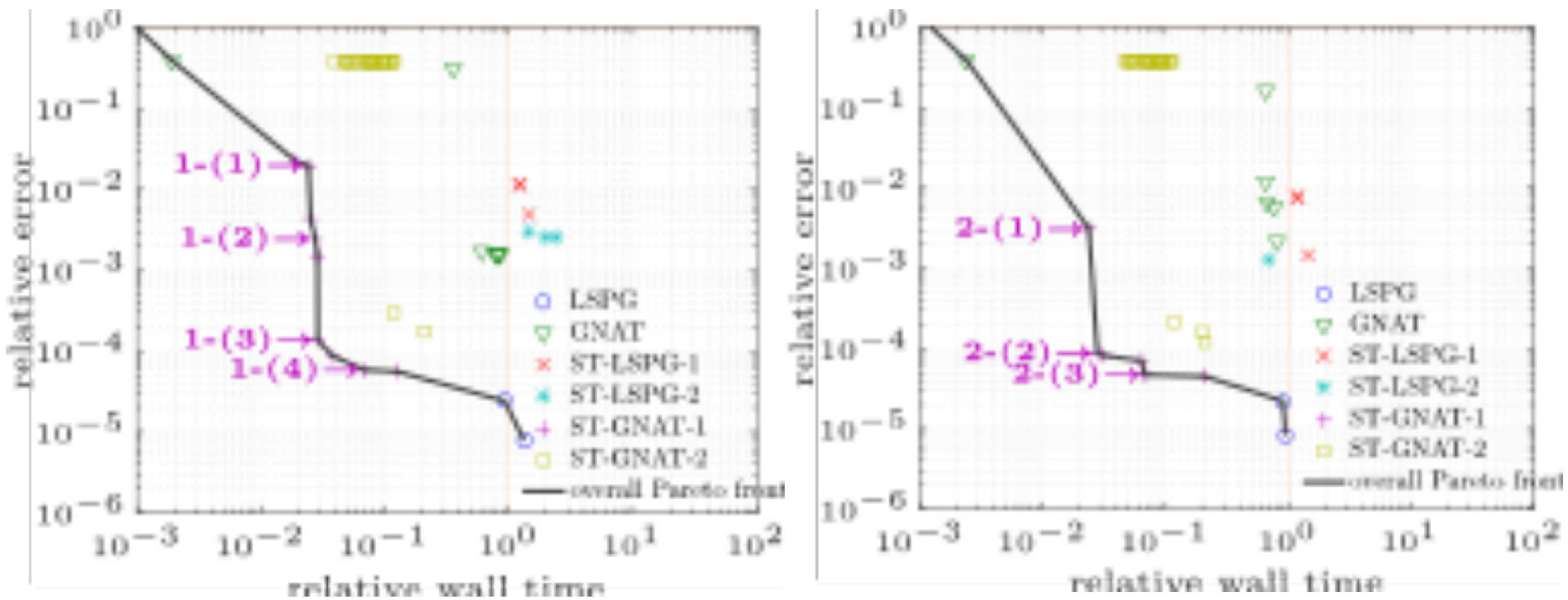
SAND2018-3549C

Space–time least-squares Petrov–Galerkin projection

for nonlinear model reduction



Youngsoo Choi, <u>Kevin Carlberg</u>

Sandia National Laboratories

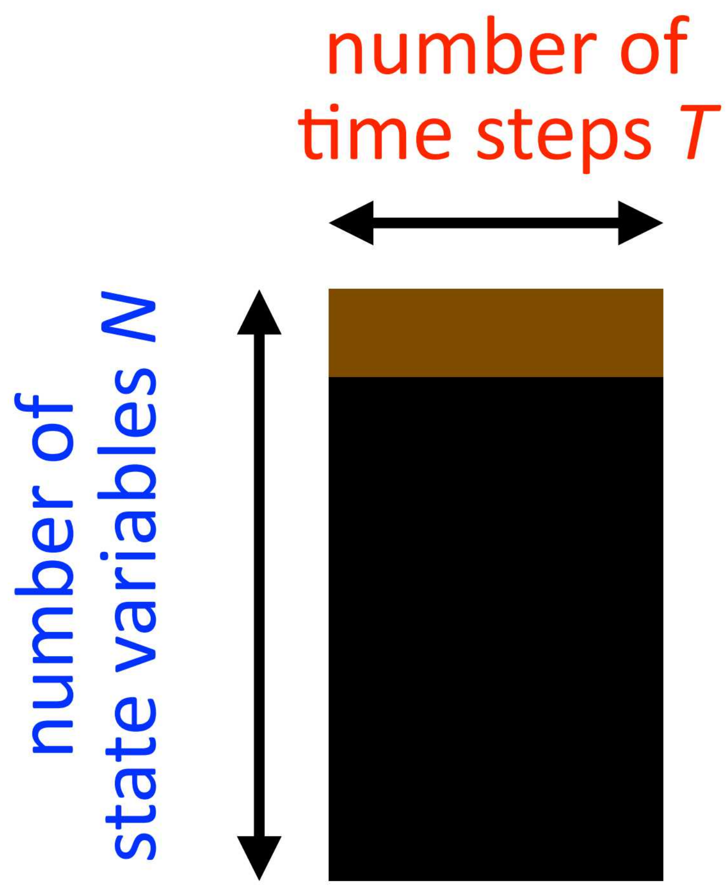
Livermore, California

MoRePaS IV, Nantes, France

April 12, 2018

Motivation

$$\mathbf{r}^n(\mathbf{x}^n) = 0, \;\; n = 1, \ldots, T$$



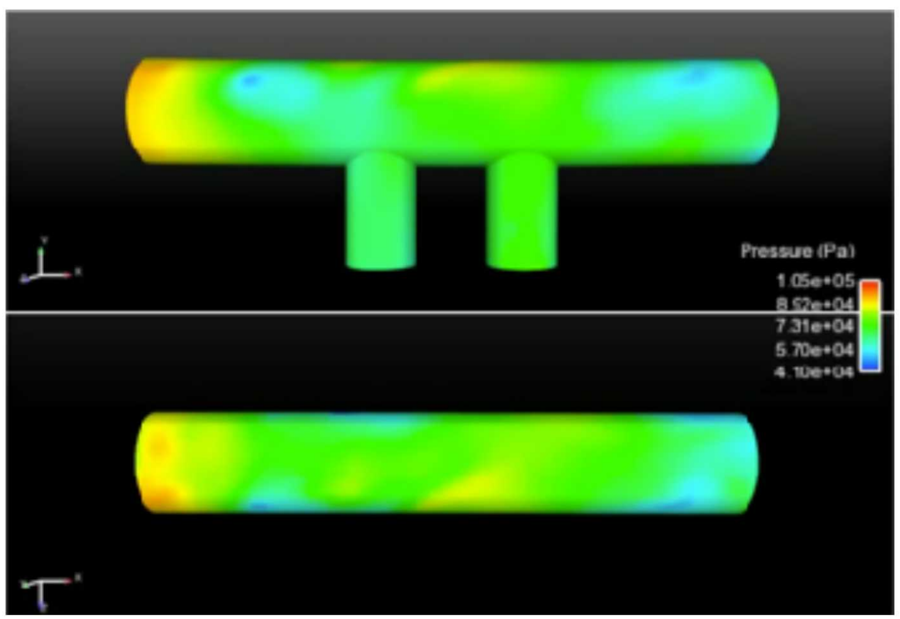
***Goal**: also reduce the tempo*

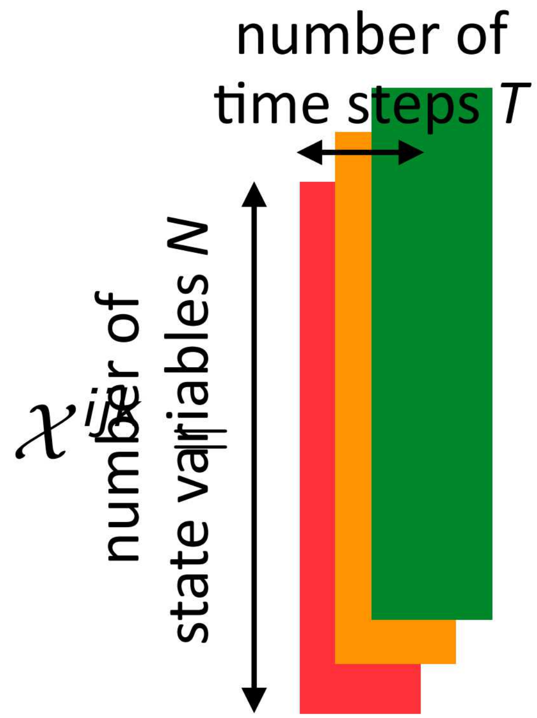
- *Most reduced-order models for nonlinear dynamical systems*
- *leverage spatial simulation data to reduce the spatial dimension*
- ***Goal**: reduce spatial dimension*

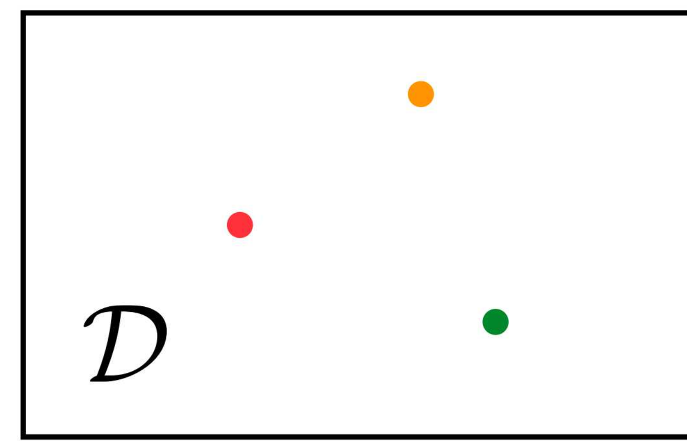
Training: state tensor

$$\text{ODE:} \quad \frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}; t, \boldsymbol{\mu})$$

1. *Training:* Solve ODE for $\boldsymbol{\mu} \in \mathcal{D}_{\text{training}}$ and collect simulation data
2. *Machine learning:* Identify structure in data
3. *Reduction:* Reduce the cost of solving ODE for $\boldsymbol{\mu} \in \mathcal{D}_{\text{query}} \setminus \mathcal{D}_{\text{training}}$





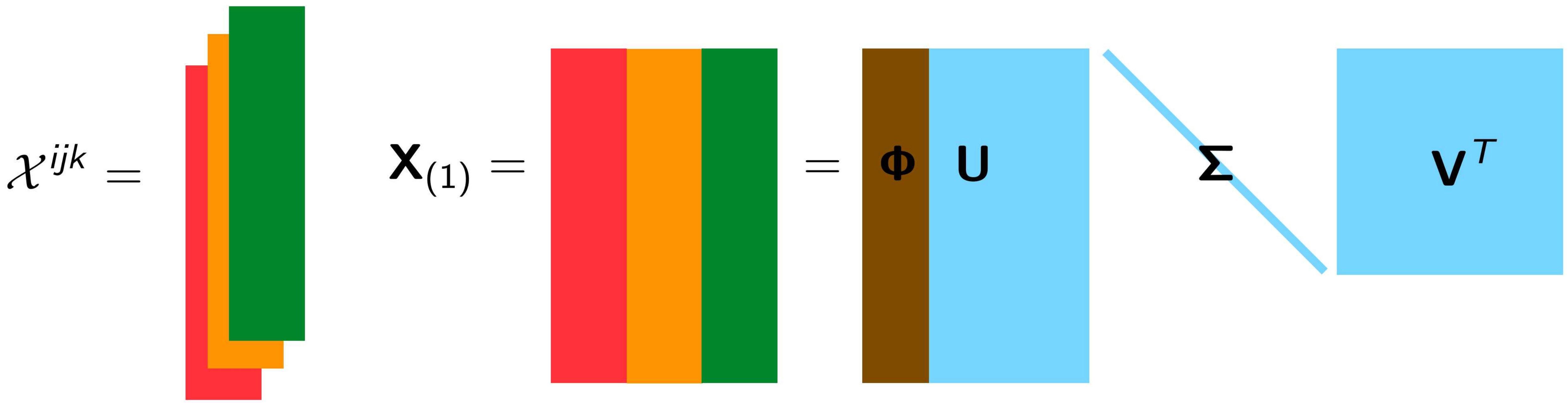


Machine learning: PCA of spatial state data

$$\text{ODE:} \quad \frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}; t, \boldsymbol{\mu})$$

1. *Training:* Solve ODE for $\boldsymbol{\mu} \in \mathcal{D}_{\text{training}}$ and collect simulation data
2. *Machine learning:* Identify structure in data
3. *Reduction:* Reduce the cost of solving ODE for $\boldsymbol{\mu} \in \mathcal{D}_{\text{query}} \setminus \mathcal{D}_{\text{training}}$

Compute dominant left singular values of mode-1 unfolding

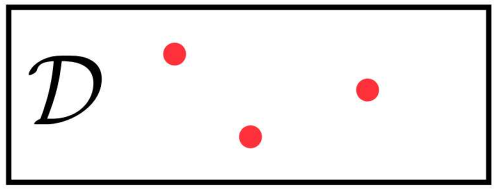


$\mathbf{\Phi}$ *columns are principal components of the spatial simulation data*

How to integrate these data with the computational model?

Reduction: Petrov–Galerkin projection

ODE: $$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}; t, \boldsymbol{\mu})$$



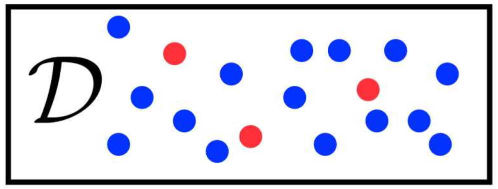
1. *Training:* Solve ODE for $\mu \in \mathcal{D}_{\text{training}}$ and collect simulation data
2. *Machine learning:* Identify structure in data
3. *Reduction:* Reduce the cost of solving ODE for $\boldsymbol{\mu} \in \mathcal{D}_{\text{query}} \setminus \mathcal{D}_{\text{training}}$

1. Reduce the number of unknowns
2. Reduce the number of equations

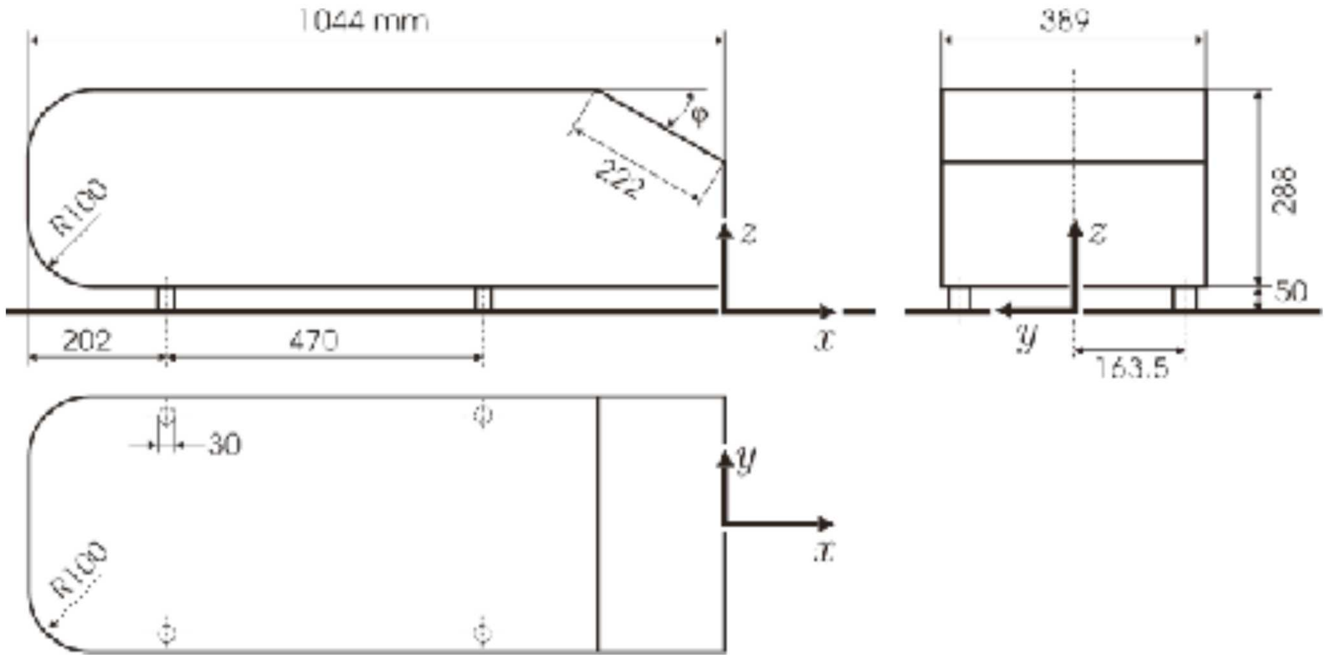
$$\mathbf{x}(t) \approx \tilde{\mathbf{x}}(t) = \boldsymbol{\Phi}\,\hat{\mathbf{x}}(t)$$

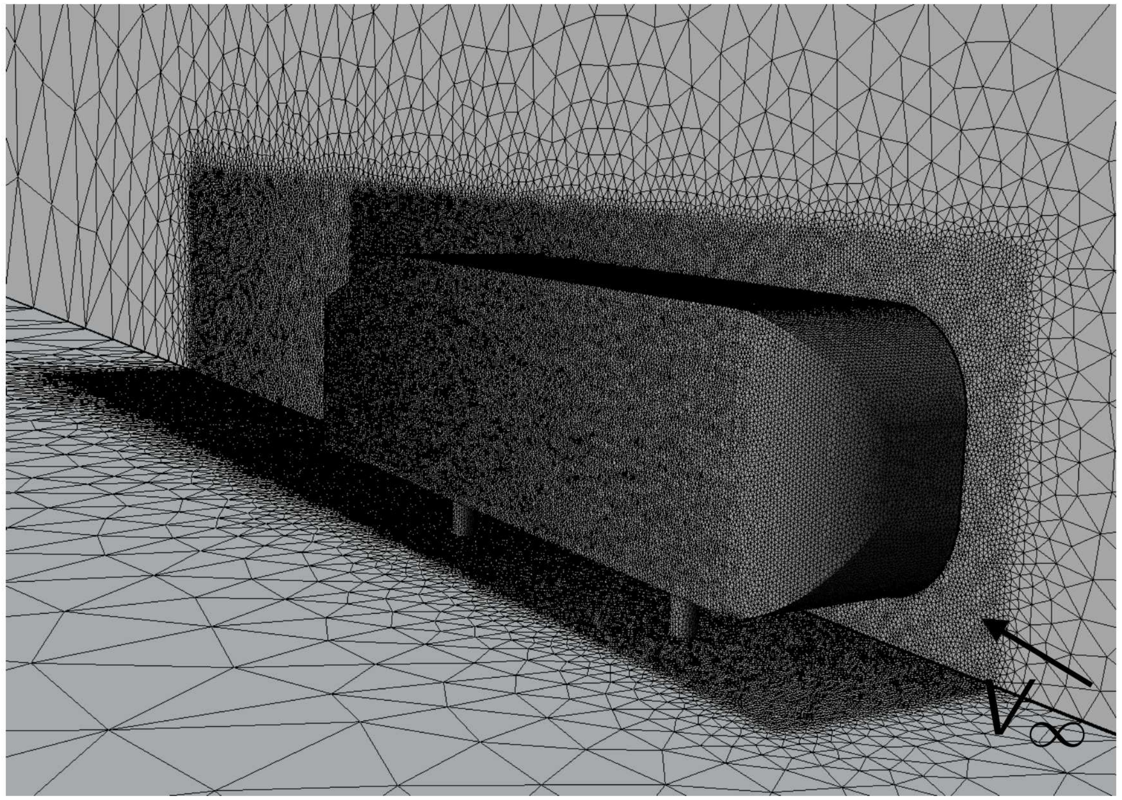
$$\boldsymbol{\Psi}(\hat{\mathbf{x}}, t)^T\left(\mathbf{f}(\boldsymbol{\Phi}\hat{\mathbf{x}}; t, \boldsymbol{\mu}) - \boldsymbol{\Phi}\,\frac{d\hat{\mathbf{x}}}{dt}\right) = 0$$

Petrov–Galerkin ODE: $$\frac{d\hat{\mathbf{x}}}{dt} = \left(\boldsymbol{\Psi}(\hat{\mathbf{x}}, t)^T\boldsymbol{\Phi}\right)^{-1}\boldsymbol{\Psi}(\hat{\mathbf{x}}, t)^T\mathbf{f}(\boldsymbol{\Phi}\hat{\mathbf{x}}; t, \boldsymbol{\mu})$$



Ahmed body [Ahmed, Ramm, Faitin, 1984]





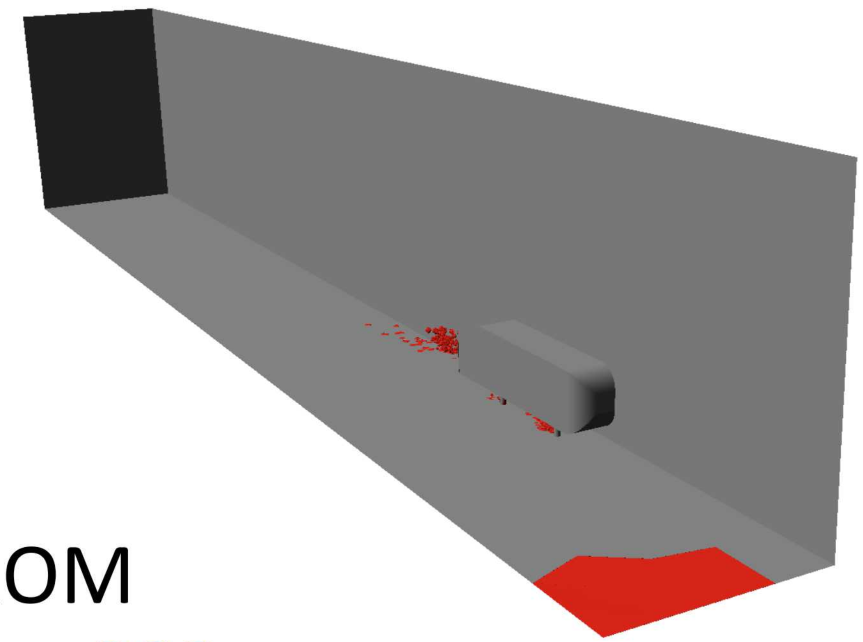
- Unsteady Navier–Stokes
- Re = 4.3 x 10^6
- M_∞ = 0.175

Spatial discretization
- 2nd-order finite volume
- DES turbulence model
- 1.7×10^7 degrees of freedom

Temporal discretization
- 2nd-order BDF
- Time step $\Delta t = 8 \times 10^{-5}\mathrm{s}$
- 1.3×10^3 time instances

Ahmed body results [Carlberg, Farhat, Cortial, Amsallem, 2013]

sample mesh

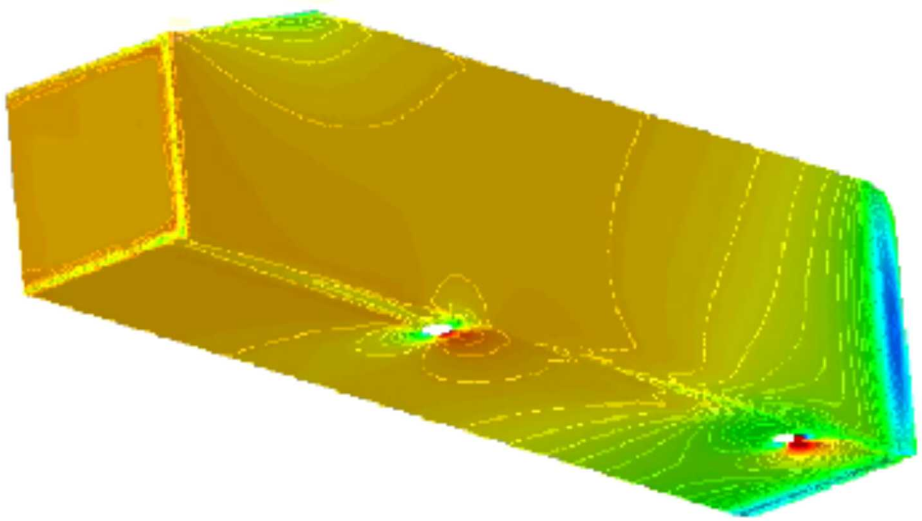


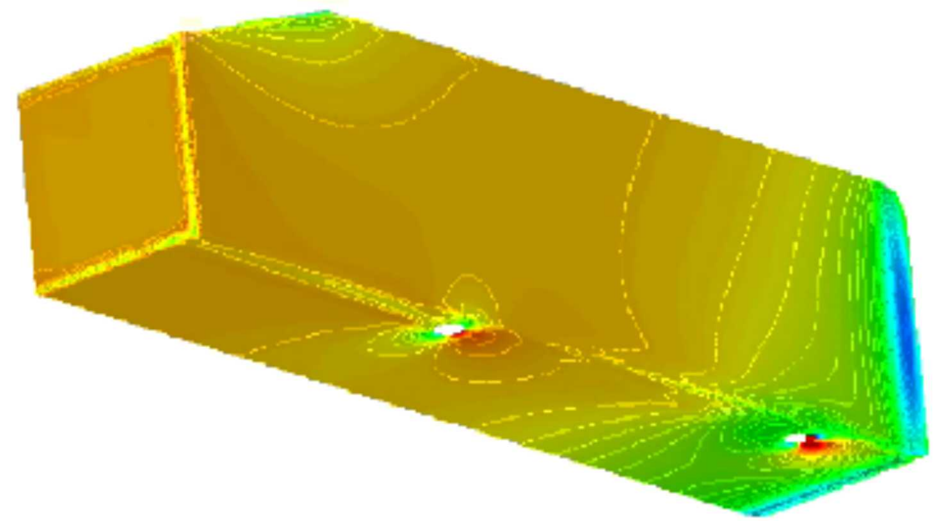
+ *HPC on a laptop*

GNAT ROM
spatial dim: 283
time dim: 1.3×10^3
4 hours, 4 cores

high-fidelity model
spatial dim: 1.7×10^7
time dim: 1.3×10^3
13 hours, 512 cores

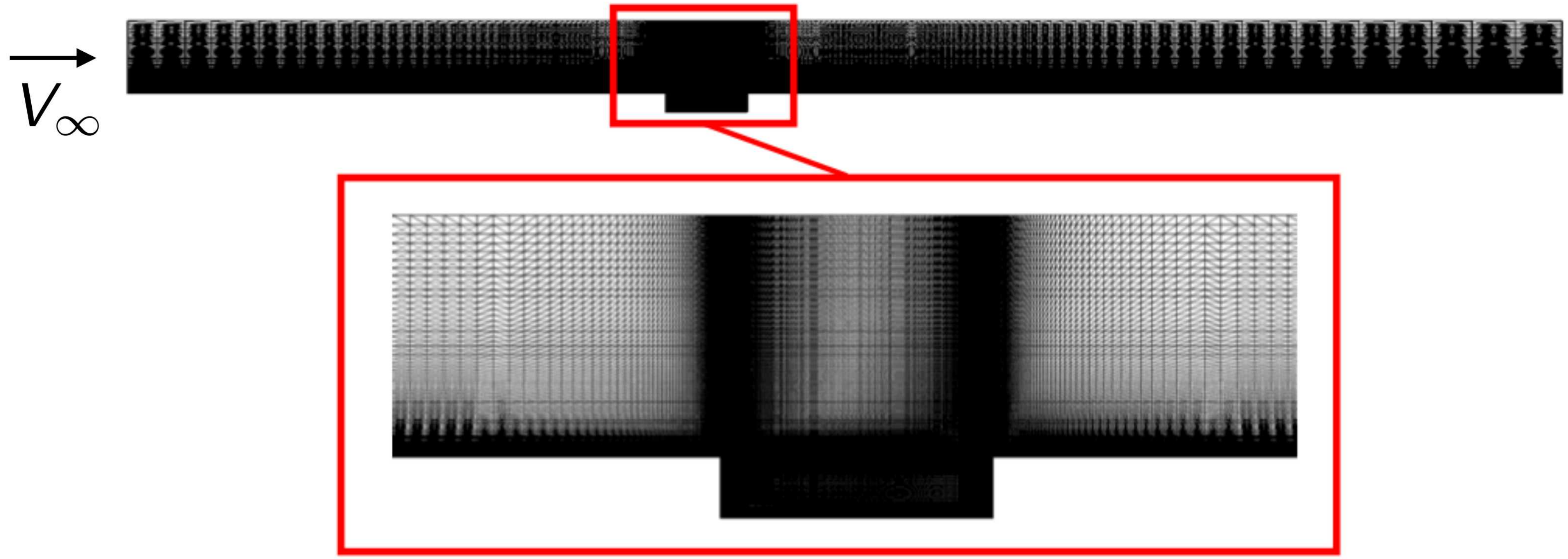
pressure field





+ *438X computational-cost reduction*
+ *60,500X spatial-dimension reduction*
- *Zero temporal-dimension reduction*

B61 captive carry



- Unsteady Navier–Stokes
- Re = 6.3 x 10^6
- $M_\infty = 0.6$

Spatial discretization

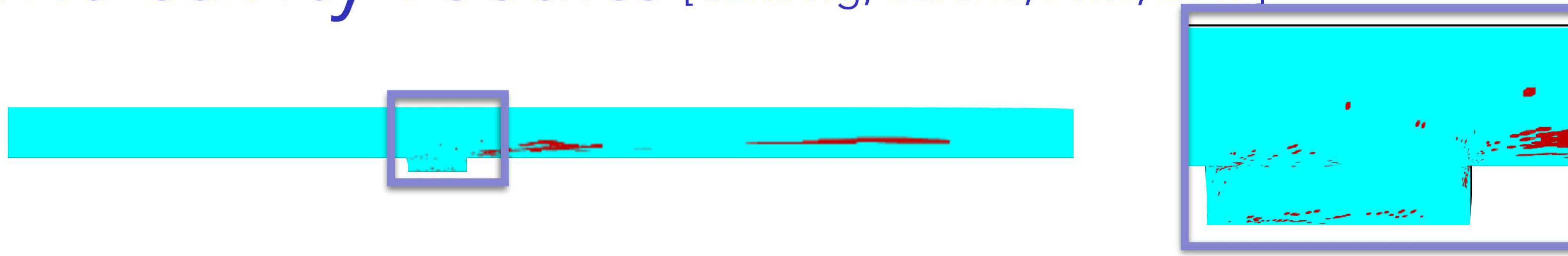
- 2nd-order finite volume
- DES turbulence model
- 1.2×10^6 degrees of freedom

Temporal discretization

- 2nd-order BDF
- Verified time step $\Delta t = 1.5 \times 10^{-3}$
- 8.3×10^3 time instances

Turbulent-cavity results [Carlberg, Barone, Antil, 2017]

sample mesh



+ *HPC on a laptop*

vorticity field *pressure field*

GNAT ROM

spatial dim: 179

temporal dim: 458

32 min, 2 cores

high-fidelity

spatial dim: 1.2M

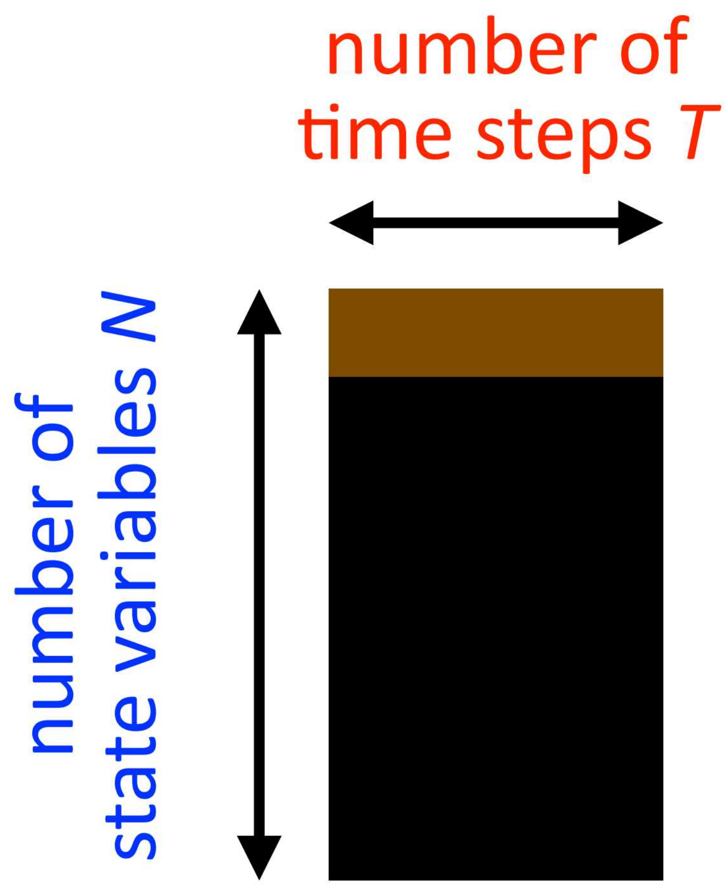
time dim: 3,700

5 hours, 48 cores

+ *229X computational-cost reduction*
+ *6,500X spatial-dimension reduction*
- *8X temporal-dimension reduction*

Temporal complexity

$$\mathbf{r}^n(\mathbf{x}^n) = 0, \;\; n = 1, \ldots, T$$



So far, we have focused on reducing the spatial complexity

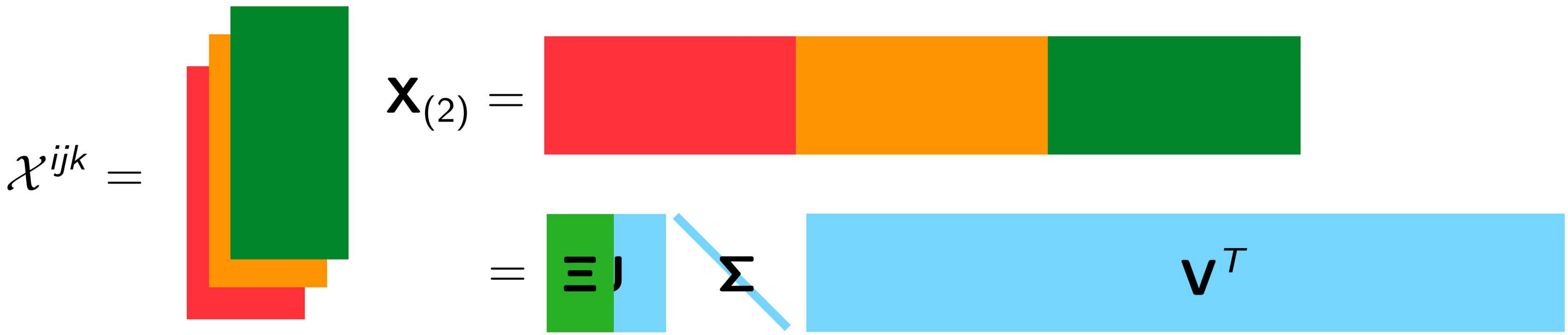
What about the temporal complexity?

Tensor decomposition

$$\text{ODE:} \quad \frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}; t, \boldsymbol{\mu})$$

1. *Training:* Solve ODE for $\boldsymbol{\mu} \in \mathcal{D}_{\text{training}}$ and collect simulation data
2. *Machine learning:* Identify structure in data
3. *Reduction:* Reduce the cost of solving ODE for $\boldsymbol{\mu} \in \mathcal{D}_{\text{query}} \setminus \mathcal{D}_{\text{training}}$

Compute dominant left singular values of ***mode-2*** *unfolding*



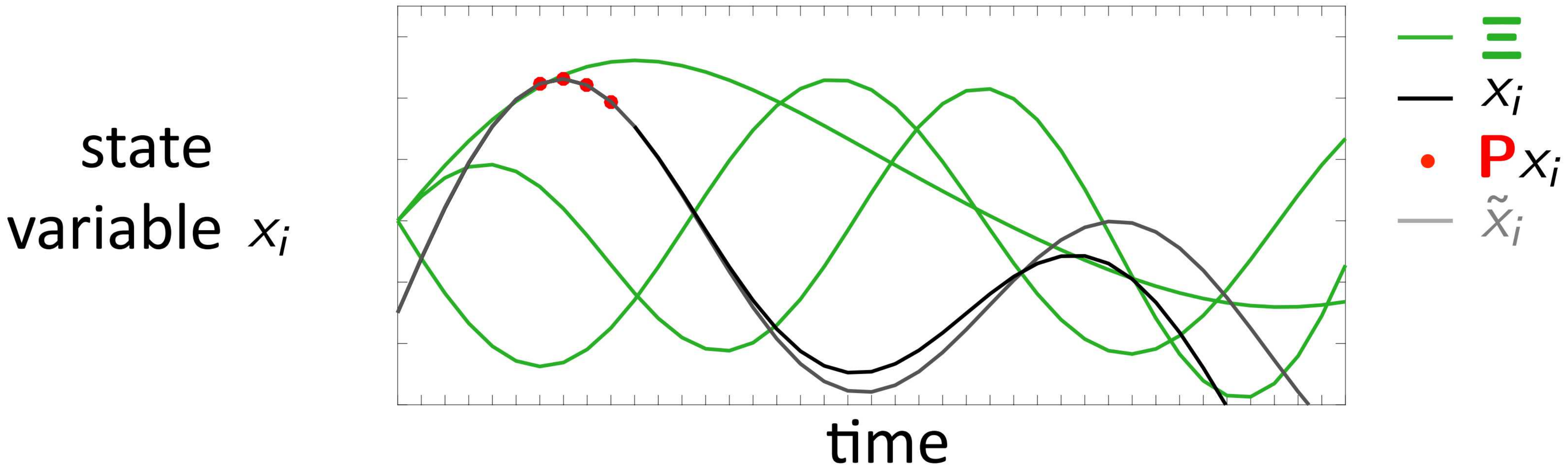
$\boldsymbol{\Xi}$ *columns are principal components of the* ***temporal*** *simulation data*

How to integrate these data with the computational model?

Idea: forecasting via gappy POD in time

[Carlberg, Ray, van Bloemen Waanders, 2015]

$$x_i(t) \approx \tilde{x}_i(t) = \boldsymbol{\Xi}(t)(\mathbf{P}\boldsymbol{\Xi}(t))^{+}\mathbf{P}x_i(t)$$



<u>MoRePaS II</u>: Data-driven initial guess

[Carlberg, Ray, van Bloemen Waanders, 2015]

- use forecast $\tilde{x}_i$ as accurate initial guess for the Newton solver
- \+ 50% speedup improvement observed; no accuracy loss

<u>MoRePaS III</u>: Data-driven time-parallel solver

[Carlberg, Brencher, Haasdonk, Barth, 2016]

- use forecast $\tilde{x}_i$ as accurate coarse propagator
- \+ provably stable; superlinear convergence; ideal speedups possible
- \+ 10x speedup improvements observed; no accuracy loss

This work

Introduce a space-time ROM that satisfies:

- + Dimension and complexity reduction in both space and time
- + Amenable to any time integrators
- + Slow time growth in error bound

Parameterized nonlinear dynamical system

$$\frac{d\boldsymbol{w}}{dt} = \boldsymbol{f}(\boldsymbol{w}, t; \boldsymbol{\mu})$$
$$\boldsymbol{w}(0; \boldsymbol{\mu}) = \boldsymbol{w}_0(\boldsymbol{\mu})$$

- Parameter vector, $\boldsymbol{\mu} \in \mathbb{R}^{N_p}$
- State, $\boldsymbol{w} : [0, T] \times \mathbb{R}^{N_p} \rightarrow \mathbb{R}^{N_s}$
- Initial condition, $\boldsymbol{w}_0 : \mathbb{R}^{N_p} \rightarrow \mathbb{R}^{N_s}$
- Nonlinear function, $\mathbf{f} : \mathbb{R}^{N_s} \times [0, T] \times \mathbb{R}^{N_p} \rightarrow \mathbb{R}^{N_s}$

Linear multistep schemes

$$\boldsymbol{r}^n(\boldsymbol{w}^n) := \sum_{j=0}^{k} \alpha_j \boldsymbol{w}^{n-j} - \Delta t \sum_{j=0}^{k} \beta_j \boldsymbol{f}(\boldsymbol{w}^{n-j}, t^{n-j}, \boldsymbol{\mu})$$

- Total time steps: $N_t := T/\Delta t \in \mathbb{N}$

Truncated higher-order SVD [Tucker et al.,2012]

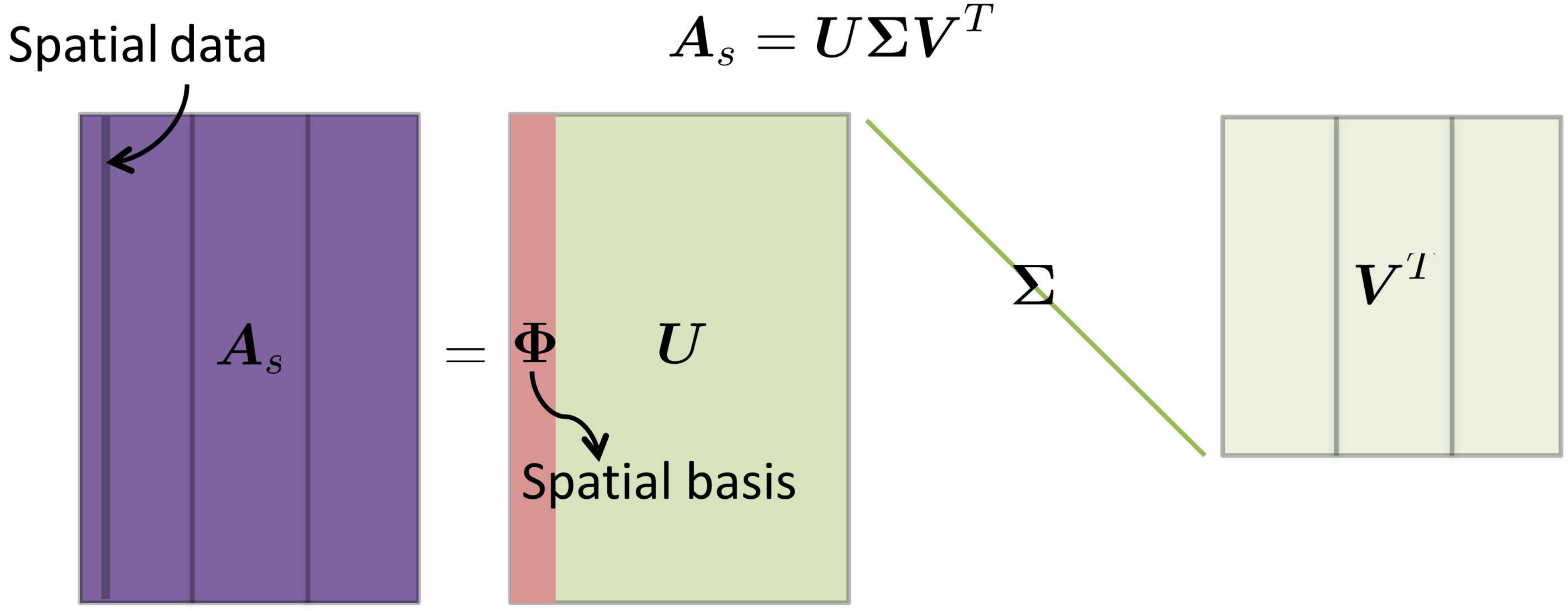
Snapshot matrix, $\boldsymbol{A}_s$

$$\boldsymbol{A}_s = \begin{bmatrix}\boldsymbol{W}_1 & \boldsymbol{W}_2 & \boldsymbol{W}_3\end{bmatrix} - \boldsymbol{W}_{ref}$$

- Solution, $\boldsymbol{W}_i := \begin{bmatrix}\boldsymbol{w}^1(\boldsymbol{\mu}_i) & \cdots & \boldsymbol{w}^{N_t}(\boldsymbol{\mu}_i)\end{bmatrix}, \quad i \in \{1, 2, 3\}$

Singular value decomposition

$$\boldsymbol{A}_s = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T$$



Global temporal basis, T-HOSVD

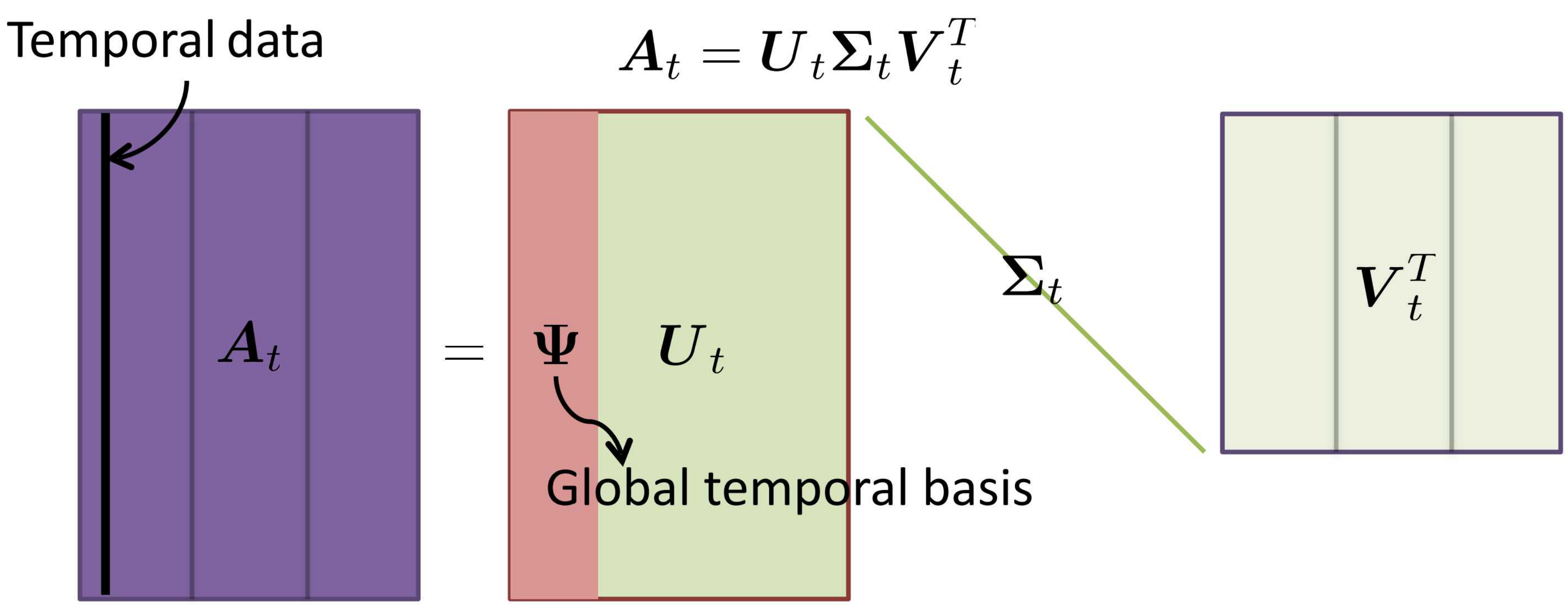
Snapshot matrix, $\boldsymbol{A}_t$

$$\boldsymbol{A}_t = \begin{bmatrix} \boldsymbol{W}_1^T & \boldsymbol{W}_2^T & \boldsymbol{W}_3^T \end{bmatrix} - \boldsymbol{W}_{ref}^T$$

- Solution, $\boldsymbol{W}_i := \begin{bmatrix} \boldsymbol{w}^1(\boldsymbol{\mu}_i) & \cdots & \boldsymbol{w}^{N_t}(\boldsymbol{\mu}_i) \end{bmatrix}, \quad i \in \{1, 2, 3\}$

Singular value decomposition

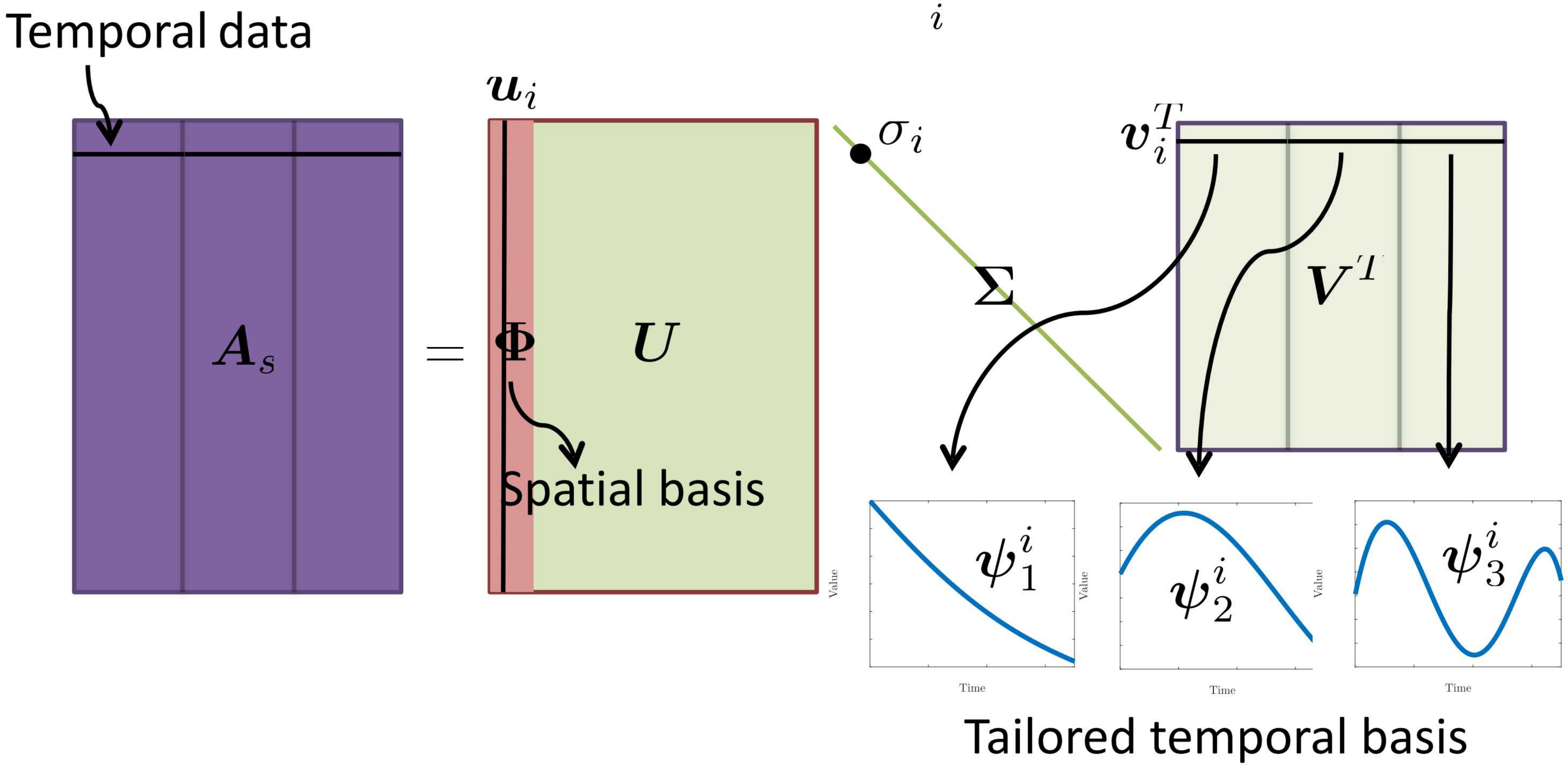
$$\boldsymbol{A}_t = \boldsymbol{U}_t \boldsymbol{\Sigma}_t \boldsymbol{V}_t^T$$



Tailored ST-HOSVD

Singular value decomposition

$$\boldsymbol{A}_s = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T = \sum_i \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^T$$



Trial subspace

Spatial ROM

- Solution approximation:

$$\tilde{\boldsymbol{w}} \in \underbrace{\boldsymbol{w}_{ref} \otimes \mathbf{1}_{N_t} + \mathcal{S} \otimes \mathbb{R}^{N_t}}_{\text{space–time subspace}} \subseteq \underbrace{\mathbb{R}^{N_s} \otimes \mathbb{R}^{N_t}}_{\text{FOM subspace}}$$

$$\Updownarrow$$

$$\tilde{\boldsymbol{w}}(t^n) = \boldsymbol{w}_{ref} + \boldsymbol{\Phi}\hat{\boldsymbol{w}}(t^n)$$

- Spatial subspace

$$\mathcal{S} = \mathrm{Ran}(\boldsymbol{\Phi}) \subseteq \mathbb{R}^{N_s}, \ \dim(\mathcal{S}) = n_s$$

- # degrees of freedom

$$\dim(\mathcal{S} \otimes \mathbb{R}^{N_t}) = n_s N_t$$

– Ignores temporal reduction

Spatiotemporal ROM

- Solution approximation:

$$\tilde{\boldsymbol{w}} \in \underbrace{\boldsymbol{w}_{ref} \otimes \mathbf{1}_{N_t} + \oplus_{i=1}^{n_s} \mathcal{S}_i \otimes \mathcal{T}_i}_{\text{space–time subspace}} \subseteq \underbrace{\mathbb{R}^{N_s} \otimes \mathbb{R}^{N_t}}_{\text{FOM subspace}}$$

$$\Updownarrow$$

$$\tilde{\boldsymbol{w}} = \boldsymbol{w}_{ref} \otimes \mathbf{1}_{N_t} + \sum_{i=1}^{n_s} \sum_{j=1}^{n_t^i} \left(\boldsymbol{\psi}_j^i \otimes \boldsymbol{\phi}_i\right) \hat{w}_{ij}$$

- Spatial subspace

$$\mathcal{S}_i := \mathrm{span}(\boldsymbol{\phi}_i) \subset \mathbb{R}^{N_s}, i = 1, \ldots, n_s, \dim \mathcal{S}_i = 1$$

- Temporal subspace

$$\mathcal{T}_i := \mathrm{Ran}\left(\begin{bmatrix} \boldsymbol{\psi}_1^i & \cdots & \boldsymbol{\psi}_{n_t^i}^i \end{bmatrix}\right) \subset \mathbb{R}^{N_t}, i = 1, \ldots, n_s$$

- # degrees of freedom

$$\dim(\oplus_{i=1}^{n_s} \mathcal{S}_i \otimes \mathcal{T}_i) = \sum_{i=1}^{n_s} n_t^i$$

+ Fewer degrees of freedom

LSPG projection

Spatial LSPG

- discrete-residual minimization

$$\hat{\boldsymbol{w}}^n = \arg\min_{\boldsymbol{y}} \|\boldsymbol{G}\boldsymbol{r}^n(\boldsymbol{w}_{ref} + \boldsymbol{\Phi}\boldsymbol{y})\|_2^2$$

- LSPG: $\boldsymbol{G} = \boldsymbol{I} \in \mathbb{R}^{N_s \times N_s}$
- Collocation: $\boldsymbol{G} = \boldsymbol{Z} \in \mathbb{R}^{n_z \times N_s}$

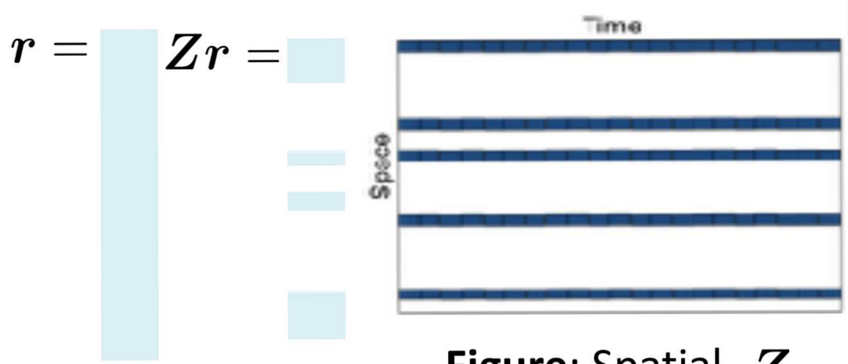


Figure: Spatial $\boldsymbol{Z}$

- GNAT: $\boldsymbol{G} = (\boldsymbol{Z}\boldsymbol{\Phi}_r)^{\dagger}\boldsymbol{Z} \in \mathbb{R}^{n_r \times N_s}$

− No temporal complexity reduction

Spatiotemporal LSPG

- discrete-residual minimization

$$\hat{\boldsymbol{w}}_{\mathrm{st}} = \arg\min_{\hat{\boldsymbol{y}}} \left\| \overline{\boldsymbol{G}}\bar{\boldsymbol{r}} \left(\boldsymbol{w}_{ref} \otimes \boldsymbol{1}_{N_t} + \sum_{i=1}^{n_s} \sum_{j=1}^{n_t^i} (\boldsymbol{\psi}_j^i \otimes \boldsymbol{\phi}_i)\hat{y}_{ij} \right) \right\|_2^2$$

- ST-LSPG: $\overline{\boldsymbol{G}} = \boldsymbol{I} \in \mathbb{R}^{N_s N_t \times N_s N_t}$
- ST-Collocation: $\overline{\boldsymbol{G}} = \overline{\boldsymbol{Z}} \in \mathbb{R}^{n_{\bar{z}} \times N_s N_t}$

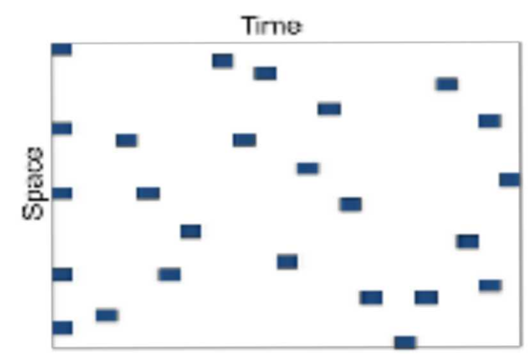


Figure: Spatiotemporal $\overline{\boldsymbol{Z}}$

- ST-GNAT: $\overline{\boldsymbol{G}} = (\overline{\boldsymbol{Z}\boldsymbol{\Phi}_r})^{\dagger}\overline{\boldsymbol{Z}} \in \mathbb{R}^{\bar{n}_r \times N_s N_t}$

\+ More complexity reduction

Error bound

Spatial LSPG [Carlberg et al., 2016] (linear multistep schemes)

$$\|\boldsymbol{w}_\star^n - \tilde{\boldsymbol{w}}_{PG}^n\|_2 \leq K(e^{ct^n} - 1) \max_{j \in \mathbb{N}(n), \ell \in \mathbb{N}_0(\ell_\epsilon^\star)} \|(\boldsymbol{I} - \mathbb{P}^j)\boldsymbol{f}(\tilde{\boldsymbol{w}}_{PG}^{j-\ell})\|_2$$

– Exponential growth in time

Spatiotemporal LSPG

- A priori error bound with respect to ℓ_2-optimal solution

$$\|\bar{\boldsymbol{w}}_\star^n - \bar{\boldsymbol{w}}_{PG}^n\|_2 \leq \sqrt{N_t}(1 + \Lambda) \min_{\bar{\boldsymbol{y}} \in \mathcal{ST}} \max_{k \in \mathbb{N}(N_t)} \|\bar{\boldsymbol{w}}_\star^k - \bar{\boldsymbol{y}}^k\|_2$$

+ Lebesgue constant grows polynomially in time with degree of 3/2

Quasi 1D Euler equation

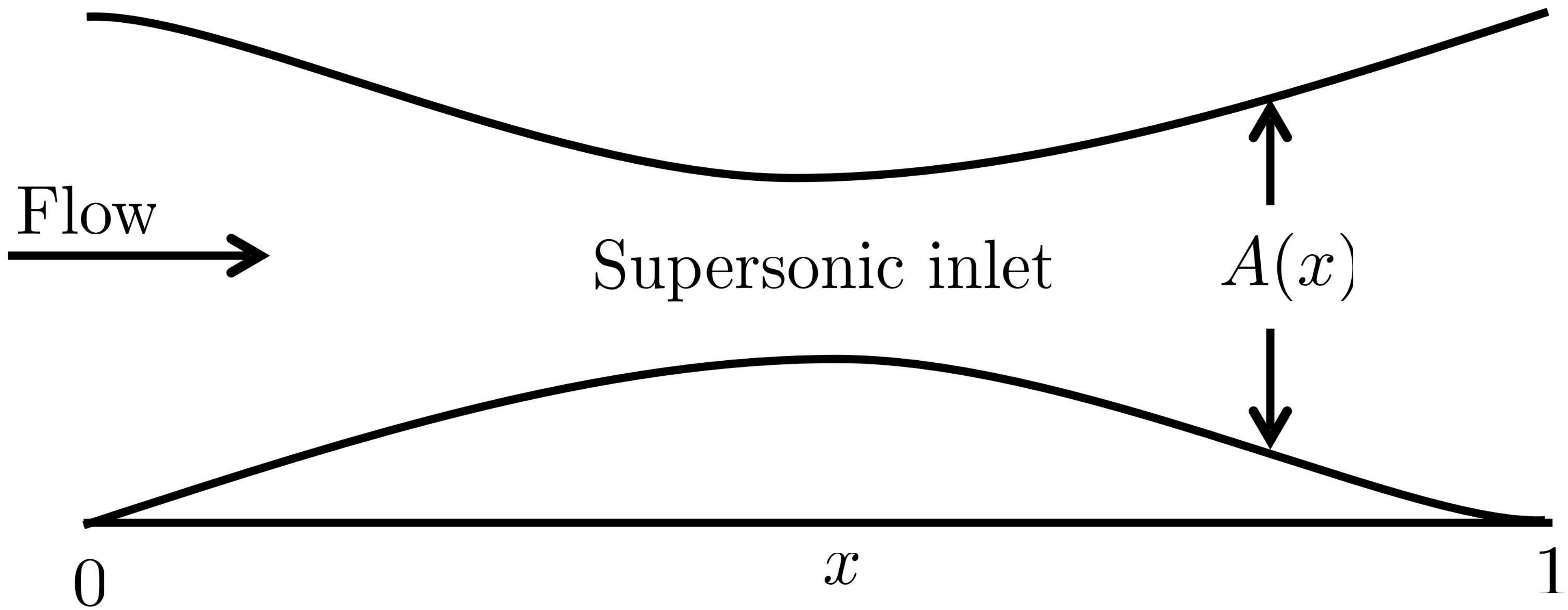
- Governing equation

$$\frac{\partial \boldsymbol{w}}{\partial t} + \frac{1}{A}\frac{\partial \boldsymbol{f} A}{\partial x} = \boldsymbol{q}$$

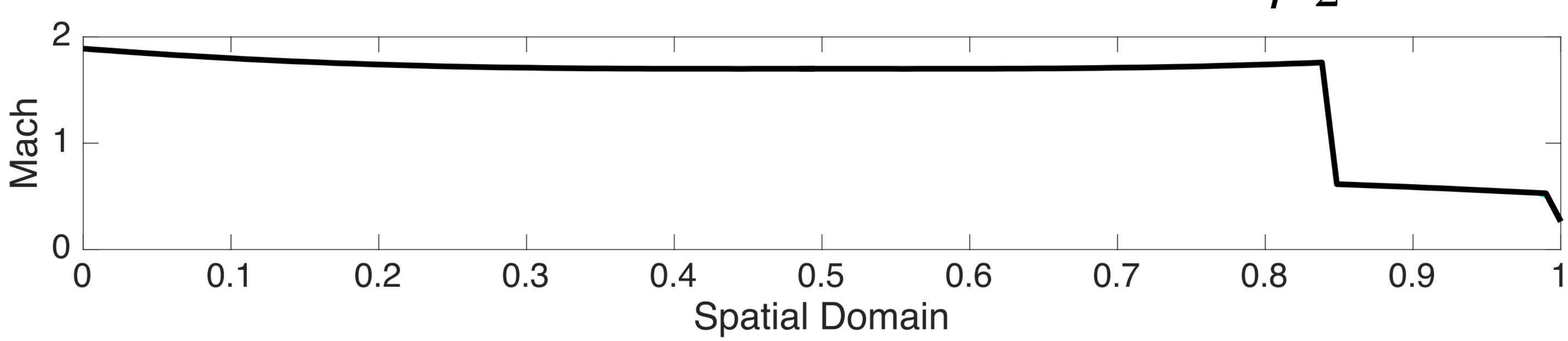
$$\boldsymbol{w} = \begin{bmatrix} \rho \\ \rho u \\ e \end{bmatrix} \qquad \boldsymbol{f} = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ (e+p)u \end{bmatrix} \qquad \boldsymbol{q} = \begin{bmatrix} 0 \\ \frac{p}{A}\frac{\partial A}{\partial x} \\ 0 \end{bmatrix}$$

- Spatial discretization: a finite volume method, $\Delta x = 10^{-2}$
- Time integrator: implicit Row method, $\Delta t = 10^{-3}$
- Parameters: exit pressure increase & Mach number at the middle

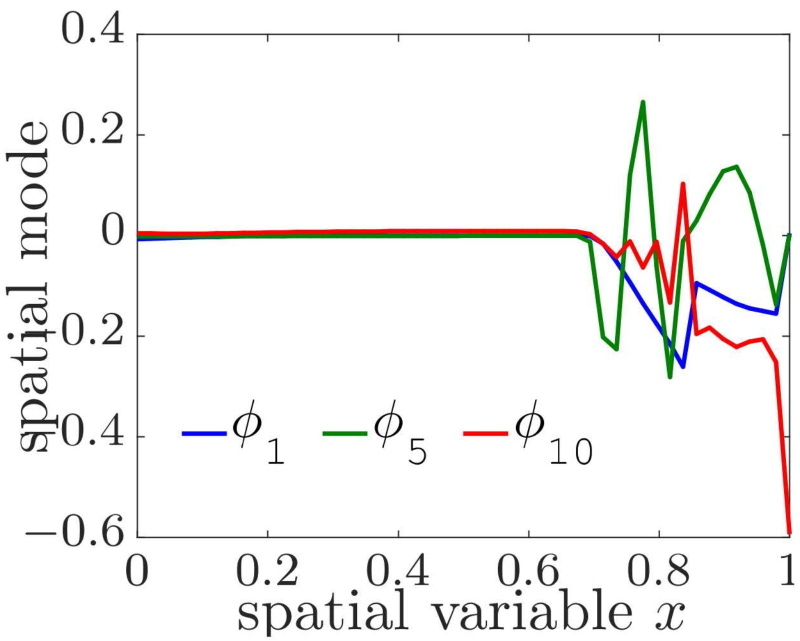
Moving Shock Problem



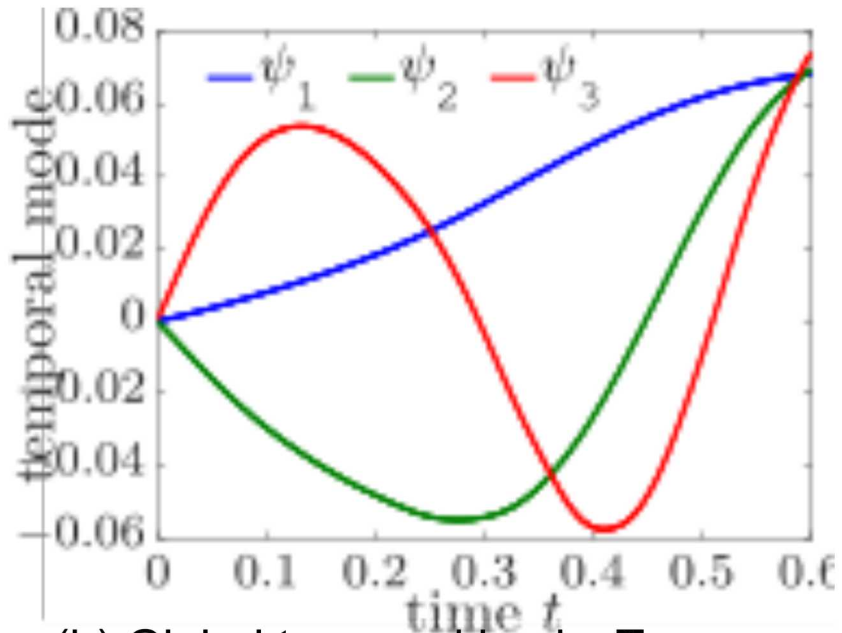
Initial flow field: Isentropic relation, shock at , $x = 0.85$
pressure increase at exit, μ_1
Mach number at the middle, μ_2



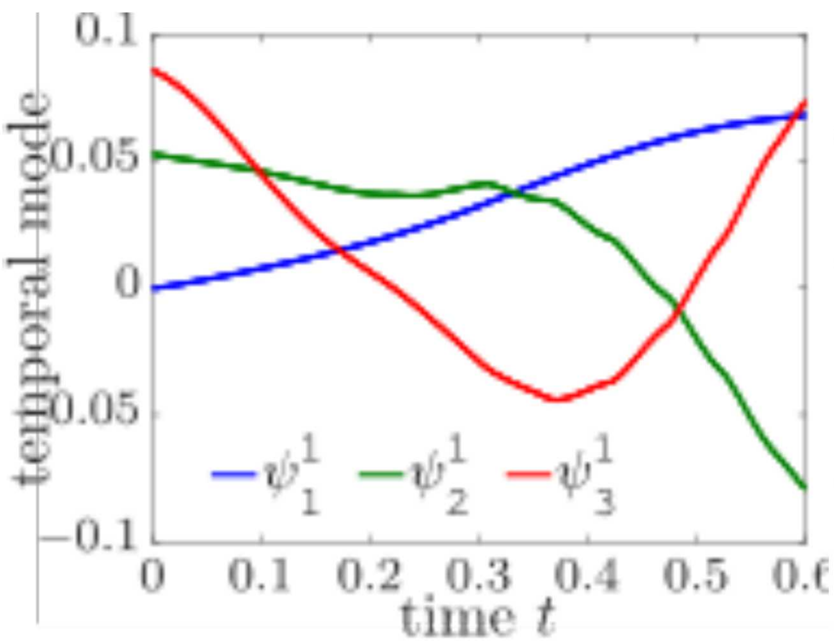
Spatiotemporal bases



(a) Spatial basis

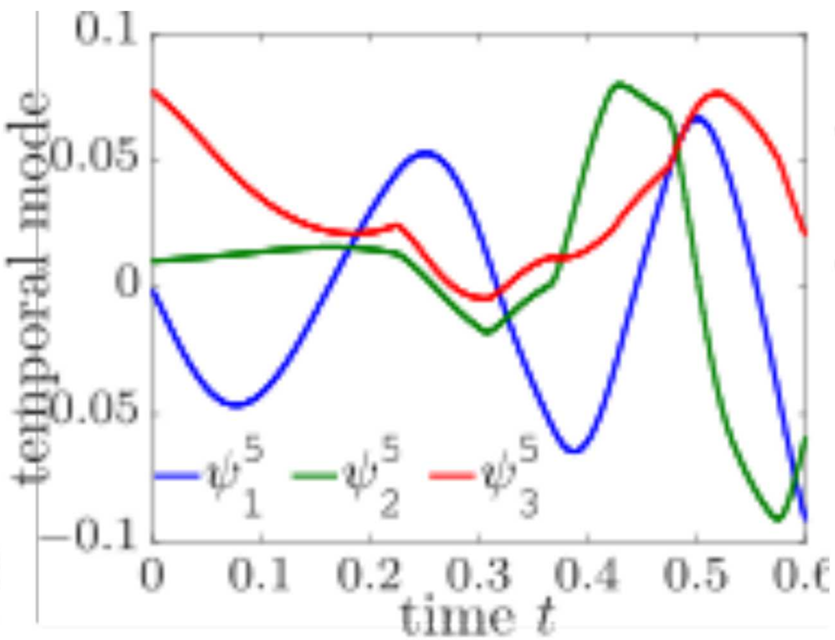


(b) Global temporal basis, T-HOSVD



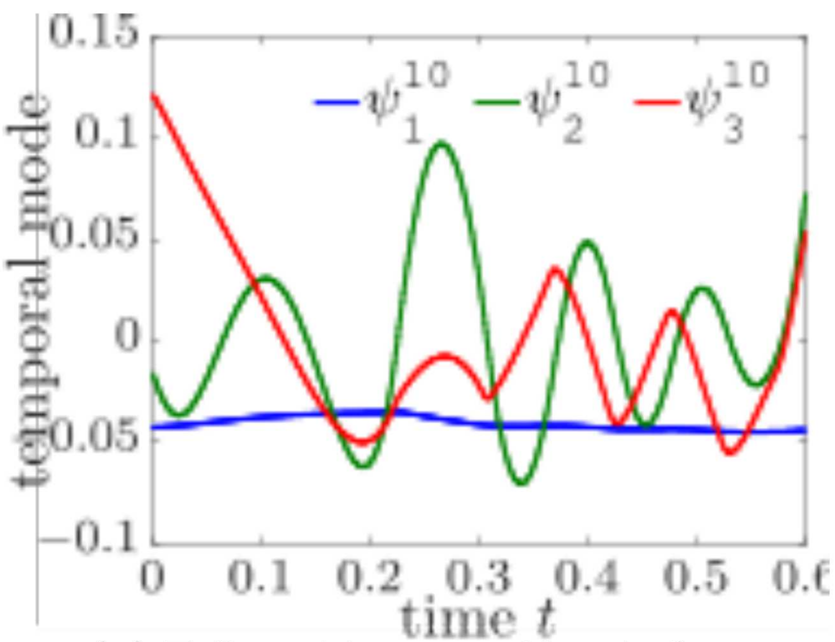
(c) Tailored temporal basis for

ϕ_1



(d) Tailored temporal basis for

ϕ_5



(e) Tailored temporal basis for

ϕ_{10}

Accuracy and speedup

$$\boldsymbol{\mu} = (1.7125, 1.71) \notin \mathcal{D}_{\text{train}}$$

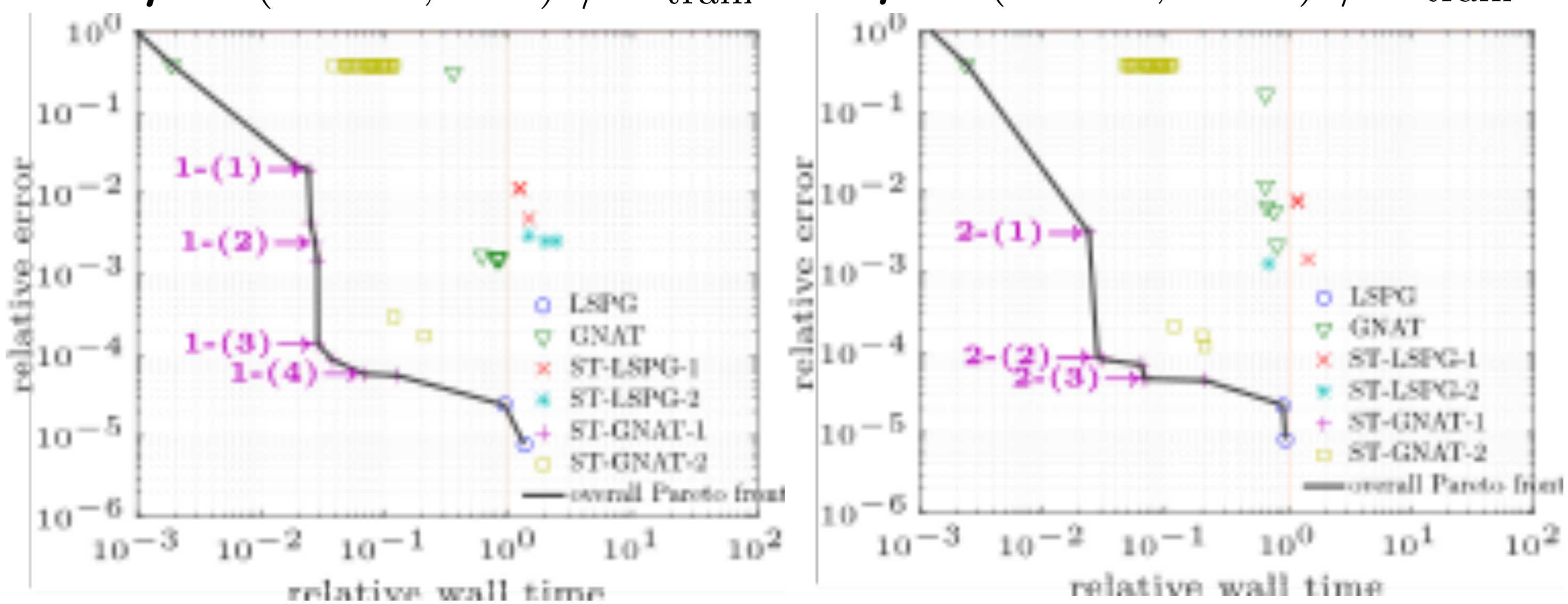
Method	LSPG	GNAT	ST-LSPG1	ST-LSPG2	ST-GNAT1	ST-GNAT2
Rel. Error	7.78×10^{-6}	0.55	0.012	6.3×10^{-4}	0.0023	0.0048
Speedup	0.77	0.94	0.84	0.58	**21.79**	0.49

$$\boldsymbol{\mu} = (1.7225, 1.705) \notin \mathcal{D}_{\text{train}}$$

Method	LSPG	GNAT	ST-LSPG1	ST-LSPG2	ST-GNAT1	ST-GNAT2
Rel. Error	8.31×10^{-6}	0.026	0.0076	0.0021	0.0025	0.0040
Speedup	0.81	1.00	0.85	0.41	**22.52**	**2.79**

Performance Pareto front

$\boldsymbol{\mu} = (1.7125, 1.71) \notin \mathcal{D}_{\text{train}}$ $\boldsymbol{\mu} = (1.7225, 1.705) \notin \mathcal{D}_{\text{train}}$



+ ST-GNAT-1 method is Pareto-optimal ROM
+ ST-GNAT-1 method produces same accuracy with 100x lower wall time than GNAT

Conclusions

Accomplishment

+ Dimension and complexity reduction in both space and time
+ Amenable to any time integrators
+ Slower time growth in error bound

Future work

- Implement in high performance computing codes
- Apply the method in PDE-constrained optimization and UQ

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-NA-0003525