A Goal-Oriented Adaptive Discrete Empirical Interpolation Method

Răzvan Ştefănescu[1] and Adrian Sandu[2]

[1] Spire Global, GVM Department, Boulder, Colorado,
[2] Computational Science Laboratory, Department of Computer Science, Virginia Polytechnic Institute and State University, Blacksburg, Virginia

Outline

Motivation and Introduction

- Enhance the accuracy of quantities of interests depending on reduced order model solutions.
- A posteriori error estimators employ the discrete solution itself to derive estimates of the actual solution errors.
- A posteriori error estimation results for the reduced order solution error - Dihlmann and Haasdonk (2014). Evaluate the error in some QoI computed via reduced order models - Carlberg (2014).
- The mechanism makes use of adjoint models and allows us to disentangle the QoI error contribution of each discrete space point at every time step.

Motivation and Introduction

- Knowing the largest error contributions we can than in turn tune ROMs by controling some of their features: DEIM points (nonlinear terms) - Chaturantabut and Sorensen (2010); DEIM indexes (Jacobians) - Wirtz and Sorensen (2014), Tonn (2011), Ştefănescu and Sandu (2014) and POD basis (modes, dimension) - Carlberg (2014).

- When using the adjoint approach in combination with ROMs, the reduced space has to be designed so that the adjoint solutions can be approximated well in this space (online estimation).

- Dual-weighted residuals to guide the selection of DEIM points for approximation of ROM nonlinear terms - Peherstorfer and Willcox (2015) - online optimal rank-one DEIM basis update with respect to the Frobenius norm; Feng et al. (2017) - update the dimensions of the ROM and DEIM bases using an a-posteriori error estimation result.

Reduced Order Modeling - POD

- The desired simulation is well approximated in the input collection Lumley(1967).
- Data analysis is conducted to extract basis functions, from experimental data or detailed simulations of high-dimensional systems.
- Galerkin and Petrov-Galerkin projections yield low dimensional dynamical models.
- Galerkin POD models - DEIM or QDEIM to address the efficiency of the nonlinear reduced order terms.

POD/DEIM justification and methodology

- Model order reduction : Reduce the computational complexity/time of large scale dynamical systems.
- Construct reduced-order model for different types of discretization method (finite difference (FD), finite element (FEM), finite volume (FV)) of unsteady and/or parametrized nonlinear PDEs. E.g., PDE:

$$\frac{\partial x}{\partial t}(z,\mu,t) = \mathrm{L}(x(z,\mu,t),\mu) + \mathrm{F}(x(z,\mu,t),\mu),\ \ t \in [0, T]$$

where L is a linear function and F a nonlinear one.

POD/DEIM justification and methodology

- The corresponding FD scheme is a n dimensional ordinary differential system
$$\frac{d}{dt}\mathbf{x}(t) = A\mathbf{x}(t) + \mathbf{F}(\mathbf{x}(t)),\ A \in \mathbb{R}^{n\times n},$$
where $\mathbf{x}(t) = [x_1(t,\mu), x_2(t,\mu), .., x_n(t,\mu)] \in \mathbb{R}^n$. $\mathbf{F}$ is a nonlinear function evaluated at $\mathbf{x}(t)$, i.e. $\mathbf{F} = [\mathrm{F}(x_1(t,\mu)), .., \mathrm{F}(x_n(t,\mu))]^T$, $\mathrm{F} : I \subset \mathbb{R} \to \mathbb{R}$.

- A common model order reduction method involves the Galerkin projection with basis $U_\mu \in \mathbb{R}^{n\times k}$ obtained from Proper Orthogonal Decomposition (POD), for $k \ll n$, i.e. $\mathbf{x} \approx \widehat{\mathbf{x}} = U_\mu\tilde{\mathbf{x}}(\mathbf{t})$, $\tilde{\mathbf{x}}(\mathbf{t}) \in \mathbb{R}^k$. Applying an inner product to the ODE discrete system we get

$$\frac{d}{dt}\tilde{\mathbf{x}}(\mathbf{t}) = \underbrace{U_\mu^T A U_\mu}_{k\times k}\tilde{\mathbf{x}}(\mathbf{t}) + \underbrace{U_\mu^T \mathbf{F}(U_\mu\tilde{\mathbf{x}}(\mathbf{t}))}_{\tilde{N}(\tilde{\mathbf{x}})} \qquad (1)$$

POD/DEIM justification and methodology

- The efficiency of POD - Galerkin technique is limited to the linear or bilinear terms. The projected nonlinear term still depends on the dimension of the original system

$$\tilde{N}(\tilde{\mathbf{x}}) = \underbrace{U_\mu^T}_{k\times n} \underbrace{\mathbf{F}(U_\mu \tilde{\mathbf{x}}(\mathbf{t}))}_{n\times 1}.$$

- To mitigate this inefficiency Chaturantabut and Sorensen (2010) introduces "Discrete Empirical Interpolation Method (DEIM)" for nonlinear approximation. For $m \ll n$

$$\tilde{N}(\tilde{\mathbf{x}}) \approx \underbrace{U_\mu^T V (P^T V)^{-1}}_{precomputed\ \ k\times m} \underbrace{\mathbf{F}(P^T U_\mu \tilde{\mathbf{x}}(\mathbf{t}))}_{m\times 1}.$$

Problem formulation

- We are interested in a particular aspect of the solution of the high-fidelity model defined by the smooth scalar function

$$\mathcal{Q}\left(\mathbf{x},\mu\right)=\sum_{i=0}^{N_t} r_i\left(\mathbf{x}_i,\mu\right). \tag{2}$$

- The reduced order approximation leads to an error in the computed QoI denoted by

$$\varepsilon(\mu)=\mathcal{Q}\left(\mathbf{x},\mu\right)-\mathcal{Q}\left(\widehat{\mathbf{x}},\mu\right)=\sum_{i=0}^{N_t} r_i\left(\mathbf{x}_i,\mu\right)-\sum_{i=0}^{N_t} r_i\left(\widehat{\mathbf{x}}_i,\mu\right), \tag{3}$$

where $\widehat{\mathbf{x}}_i = U_\mu\,\tilde{\mathbf{x}}_\mathbf{i},\ i=0,..,N_t$.

A posteriori error estimates

- Compact form of the high-fidelity model

$$\mathbf{x}_{i+1} = M_{i,i+1}(\mathbf{x}_i), \quad i = 0, \ldots, N_t - 1. \tag{4}$$

- Compact form of the reduced order model

$$\widehat{\mathbf{x}}_{i+1} = \widehat{M}_{i,i+1}(\widehat{\mathbf{x}}_i), \quad i = 0, 1, \ldots, N_t - 1. \tag{5}$$

A posteriori error estimates

Theorem (1)

Let $\mathbf{x} = \{\mathbf{x}_0,\ \mathbf{x}_1, \dots,\ \mathbf{x}_{N_t}\}$ *be the solution of the high-fidelity model, and* $\widehat{\mathbf{x}} = \{\widehat{\mathbf{x}}_0, \widehat{\mathbf{x}}_1, \dots, \widehat{\mathbf{x}}_{N_t}\}$ *the projection of reduced order model solution onto the high-fidelity space. Moreover, let* $\mathbf{x}^i = \{\widehat{\mathbf{x}}_i,\ \mathbf{x}^i_{i+1}, \dots, \mathbf{x}^i_{N_t}\}$ *be the partial trajectories obtained via the high-fidelity model using as initial conditions the solution of reduced order model at time* t_i *projected onto the high fidelity space, i.e.* $\widehat{\mathbf{x}}_i = U\,\tilde{\mathbf{x}}_{\mathbf{i}}$ *and*

$$\mathbf{x}^i_\ell = \mathcal{M}_{\ell-1,\ell}\left(\mathbf{x}^i_{\ell-1}\right), \quad \ell = i+1, \dots, N_t,\ \mathbf{x}^i_i = \widehat{\mathbf{x}}_i, \quad i = 0, \dots, N_t-1. \quad (6)$$

The partial trajectory $\mathbf{x}^i$ *contains only* $N_t - i + 1$ *time steps.*

A posteriori error estimates

Theorem (1(continuation))

Assume that the reduced order model solution $\widehat{\mathbf{x}}_i$ *is in a neighborhood of the high-resolution model solution* $\mathbf{x}_i,\ i = 0, \ldots, N_t$ *and if the high-fidelity model is smooth, then*

$$\varepsilon \approx -\sum_{i=0}^{Nt} \widehat{\lambda}_i^T \cdot \Delta \mathbf{x}_i, \tag{7}$$

where the model residuals are

$$\Delta \mathbf{x}_0 = \mathbf{x}_0 - \widehat{\mathbf{x}}_0, \quad \Delta \mathbf{x}_i = \mathbf{x}_i^{i-1} - \widehat{\mathbf{x}}_i, \quad i = 1, \ldots, N_t, \tag{8}$$

and $\widehat{\lambda}_i,\ i = 0, \ldots, N_t$, *are the solutions of the high-fidelity adjoint models (partial) linearized about the trajectories* $\mathbf{x}^i,\ i = 0, \ldots, N_t - 1$.

A posteriori error estimates

- First order necessary optimality conditions of the problem

$$\min_{\mathbf{x}_0} \quad \mathcal{Q}\left(\mathbf{x}_0\right) = \sum_{i=0}^{N_t} \; r_i\left(\mathbf{x}_i\right), \tag{9a}$$

subject to the constraints posed by the high-fidelity model dynamics

$$\mathbf{x}_{i+1} = \mathcal{M}_{i,i+1}\left(\mathbf{x}_i\right), \quad i = 0, .., N_t - 1. \tag{9b}$$

- Adjoint model

$$\begin{aligned}
\lambda_N &= -\left(\frac{\partial r_{N_t}}{\partial \mathbf{x}_{N_t}}\right)^T (\mathbf{x}_{N_t}), \\
\lambda_i &= \mathbf{M}^*_{i+1,i}\lambda_{i+1} - \left(\frac{\partial r_i}{\partial \mathbf{x}_i}\right)^T (\mathbf{x}_i), \quad i = N_t - 1, .., 0.
\end{aligned} \tag{10}$$

- $\sum_{i=0}^{N_t} \; r_i\left(\mathbf{x}_i\right) - \sum_{i=0}^{N_t} \; r_i\left(\mathbf{x}_i^0\right) \approx -\lambda_0^T \Delta \mathbf{x}_0, \; \mathbf{x}_0^0 = \widehat{\mathbf{x}}_0.$

A posteriori error estimates

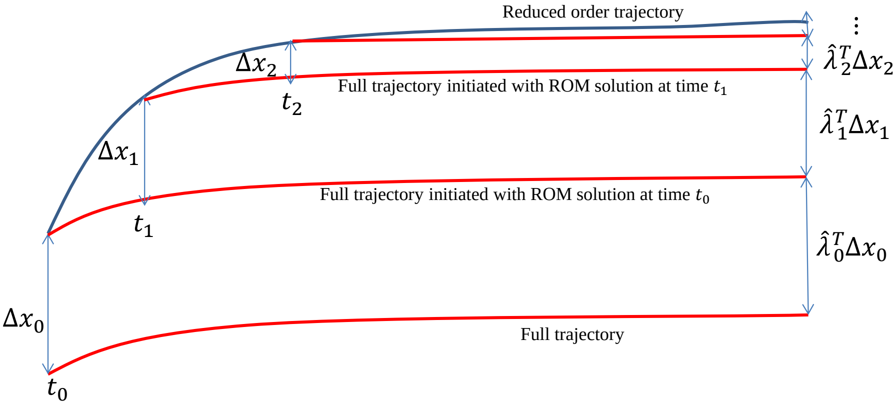


Figure: Geometrical Interpretation of aposteriori error estimates

A posteriori error estimates - efficient versions

- Discrete high fidelity explicit Euler scheme

$$\mathbf{x}_{i+1} = \mathbf{x}_i + h\,\mathbf{F}(\mathbf{x}_i), \quad i = 0, .., N_t - 1. \tag{11}$$

- One time step integration

$$\begin{aligned} \mathbf{x}^i_{i+1} &= U\tilde{\mathbf{x}}_i + h\,\mathbf{F}(U\tilde{\mathbf{x}}_i,), & (12)\\ \tilde{\mathbf{x}}_{i+1} &= \tilde{\mathbf{x}}_i + h\,U^T\,V\,(P^T V)^{-1}\,P^T\,\mathbf{F}(U\tilde{\mathbf{x}}_i). & (13) \end{aligned}$$

- By multiplying (13) with U and subtracting the result from (12)

$$\begin{aligned} \Delta\mathbf{x}_{i+1} &= \mathbf{x}^i_{i+1} - U\,\tilde{\mathbf{x}}_{i+1} & (14)\\ &= h\left(\mathbf{I} - UU^T V\,(P^T V)^{-1}\,P^T\right)\mathbf{F}(\widehat{\mathbf{x}}_i) \\ &= -\phi_{i+1}, \quad i = 0, .., N_t - 1, \end{aligned}$$

$$\phi_{i+1} = \widehat{\mathbf{x}}_{i+1} - \widehat{\mathbf{x}}_i - h\mathbf{F}(\widehat{\mathbf{x}}_i),\ \ i = 0, .., N_t - 1. \tag{15}$$

A posteriori error estimates - efficient versions

- If accurate reduced order model is available; i.e, $\mathbf{x} \approx U\tilde{\mathbf{x}}$, then the partial trajectories can be approximated by truncated trajectories obtained using one single high-fidelity model run. Then for estimating $\widehat{\lambda}_i^T, \; i = 0, \ldots, N_t$ only a single high-fidelity adjoint model run is required.
- Unlike the Galerkin POD residual, the DEIM based residual (15) is not orthogonal to the reduced manifold U. As such we can make use of a reduced order adjoint model solution to estimate

$$\varepsilon \approx -\left[U\tilde{\lambda}_0\right]^T (\mathbf{x}_0 - \widehat{\mathbf{x}}_0) + \sum_{i=1}^{N_t} \left[U\tilde{\lambda}_i\right]^T \phi_i. \tag{16}$$

- The new error estimate requires only one reduced forward and one adjoint model runs as well as evaluating the residuals.

A posteriori error estimates - efficient versions

- Discrete high fidelity implicit Euler scheme

$$\mathbf{x}_{i+1} = \mathbf{x}_i + h\,\mathbf{F}(\mathbf{x}_{i+1}), \quad i = 0,..,N_t - 1, \tag{17}$$

- The error in the quantity of interest

$$\varepsilon \approx -\left[U\tilde{\lambda}_0\right]^T(\mathbf{x}_0 - \widehat{\mathbf{x}}_0) + \sum_{i=0}^{N_t-1} \phi_{i+1}^T\left[U\tilde{\lambda}_i + \frac{\partial r_i}{\partial \mathbf{x}_i}(U\tilde{\mathbf{x}}_\mathbf{i})\right], \tag{18}$$

where ϕ_{i+1} is now the residual associated with the implicit full model

$$\phi_{i+1} = \widehat{\mathbf{x}}_{i+1} - \widehat{\mathbf{x}}_i - h\mathbf{F}(\widehat{\mathbf{x}}_{i+1}),\ \ i = 0,..,N_t - 1. \tag{19}$$

DEIM: Algorithm for Interpolation Indices

INPUT: $\{v_l\}_{l=1}^{m} \subset \mathbb{R}^n$ (linearly independent):

OUTPUT: $\vec{\rho} = [\rho_1, .., \rho_m] \in \mathbb{N}^m$

1. $[|\psi| \quad \rho_1] = max|v_1|, \psi \in \mathbb{R}$ and ρ_1 is the component position of the largest absolute value of v_1, with the smallest index taken in case of a tie.
2. $V = [v_1],\ P = [e_{\rho_1}],\ \vec{\rho} = [\rho_1]$.
3. For $l = 2, .., m$ do
   - a. Solve $(P^T V)c = P^T v_l$ for c
   - b. $r = u_l - Vc$
   - c. $[|\psi| \quad \rho_l] = max\{|r|\}$
   - d. $U \leftarrow [V \quad v_l],\ P \leftarrow [P \quad e_{\rho_l}],\ \vec{\rho} \leftarrow \begin{bmatrix} \vec{\rho} \\ \rho_l \end{bmatrix}$
4. end for.

Adaptive DEIM

- In Peherstorfer and Willcox (2015), the adaptivity mechanism changes the non-linear term reduced basis via rank-one updates and points.
- The individual contribution at each spatial location and time step to the error in the quantity of interest can be calculated by using the Hadamard product $\odot$ instead of the scalar products in a-posteriori error results. The Hadamard products are the dual weighted residuals.
- For the explicit case the dual weighted residuals are defined as

$$z_0 = [U\tilde{\lambda}_0] \odot (\mathbf{x}_0 - \widehat{\mathbf{x}}_0); \quad z_i = [U\tilde{\lambda}_i] \odot \phi_i, \;\; i = 1, .., N_t,$$

- For the implicit case these are defined as

$$z_0 = [U\tilde{\lambda}_0] \odot (\mathbf{x}_0 - \widehat{\mathbf{x}}_0); \quad z_i = \phi_i \odot \left[U\tilde{\lambda}_{i-1} + \frac{\partial r_{i-1}}{\partial \mathbf{x}_{i-1}}(U\tilde{\mathbf{x}}_{\mathbf{i-1}})\right], \;\; i = 1, .., N_t.$$

- Singular vector decomposition is applied to extract the left singular vectors of the dual weighted residuals denoted by

$$W = \{w_0, \; w_1, \ldots, w_m\}.$$

DEIM adaptive: Algorithm for Interpolation Indices

INPUT: $\{v_\ell\}_{\ell=1}^{m} \subset \mathbb{R}^{N}_{\text{state}}$ (linearly independent), $\{w_\ell\}_{\ell=1}^{m} \subset \mathbb{R}^{N}_{\text{state}}$ (linearly independent), $\alpha \in [0,1]$:

OUTPUT: $\boldsymbol{\rho} = [\rho_1, .., \rho_m] \in \mathbb{N}^m$

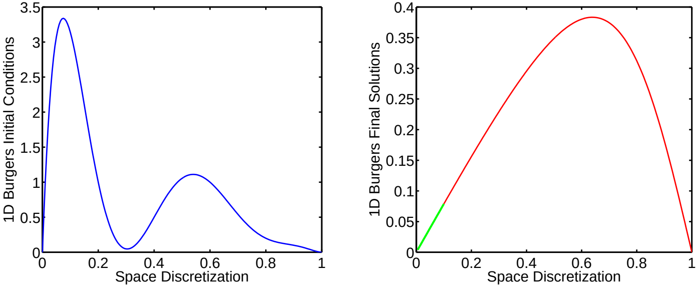
1. $\{\psi^v, \rho_1^v\} = \max|v_1|$; $\psi^v \in \mathbb{R}$ is the largest absolute value among entries of v_1, and ρ_1^v is its position.
2. $\{\psi^w, \rho_1^w\} = \max|w_1|$, $\psi^w \in \mathbb{R}$.
3. Set $\rho_1 = \rho_1^v$ if $\psi^v \geq \psi^w$, or $\rho_1 = \rho_1^w$ otherwise.
4. $V := [v_1] \in \mathbb{R}^n$, $P := [e_{\rho_1}] \in \mathbb{R}^n$, $\boldsymbol{\rho} := [\rho_1] \in \mathbb{N}$.
5. For $\ell = 2, .., m$
   a. Solve $(P^T V)\, c^v = P^T\, v_\ell$ for $c^v \in \mathbb{R}^{\ell-1}$; $V, P \in \mathbb{R}^{N_{\text{state}} \times (\ell-1)}$.
   b. $r^v := v_\ell - V c^v$, $r^v \in \mathbb{R}^{N_{\text{state}}}$.
   c. Solve $(P^T V)\, c^w = P^T\, w_\ell$ for $c^w \in \mathbb{R}^{\ell-1}$; $V, P \in \mathbb{R}^{N_{\text{state}} \times (\ell-1)}$.
   d. $r^w := w_\ell - V\, c^w$, $r^w \in \mathbb{R}^{N_{\text{state}}}$.
   e. $\{\psi, \rho_\ell\} = \max\{\alpha|r^v| + (1-\alpha)|r^w|\}$.
   f. $V := [V \;\; v_\ell]$, $P := [P \;\; e_{\rho_\ell}]$, $\boldsymbol{\rho} := \left[\boldsymbol{\rho}^T \;\; \rho_\ell\right]^T$.
6. end for.

Numerical Results

- Burgers' equation

$$\frac{\partial u}{\partial t} + u\,\frac{\partial u}{\partial x} = \mu\,\frac{\partial^2 u}{\partial x^2}, \quad x \in [0,1], \quad t \in (0,1]. \tag{20}$$

- $u(0,t) = u(1,t) = 0, \ t \in (0,1]$; Implicit Euler method.



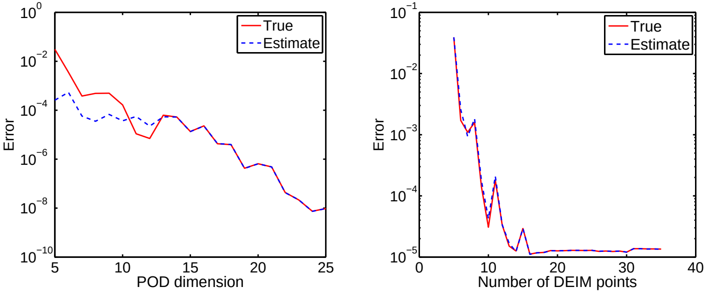
(a) Initial Conditions

(b) Final time solutions

Figure: $\mu = 0.1$

Numerical Results

$$\mathcal{Q}(u) = \sum_{i=2}^{21} u(x_i, t_{N_t})^2, \;\; [x_2, x_{21}] = [0.05, 0.1]. \tag{21}$$



(a) Number of DEIM points $= 40$ (b) Dimension of POD basis $= 15$

Figure: A-posteriori error estimates for the same parametric configuration - $\mu = 0.1$.

Numerical Results

- Computed the dual weighted residuals, performed a singular value decomposition and collected 15 singular vectors.
- The parameter α was set to 0.5.

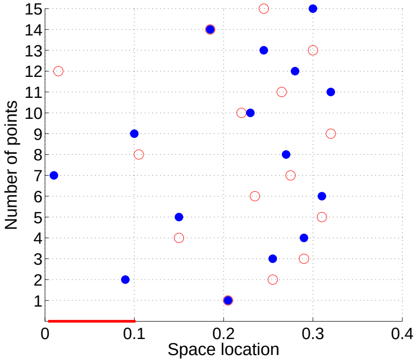
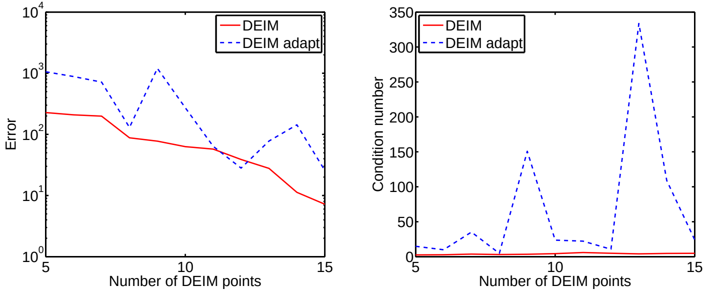


Figure: DEIM points locations.

Numerical Results



(a) Non-linear approximation (b) Condition numbers

Figure: Comparison between traditional and adaptive DEIM strategies - Global non-linear term error at time step t_2 in the Euclidian norm (left panel); Condition number of matrix $P^T V$ (right panel).

Numerical Results

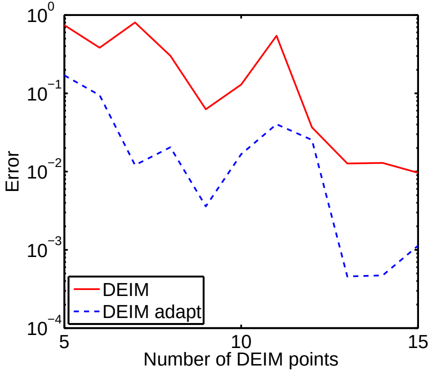


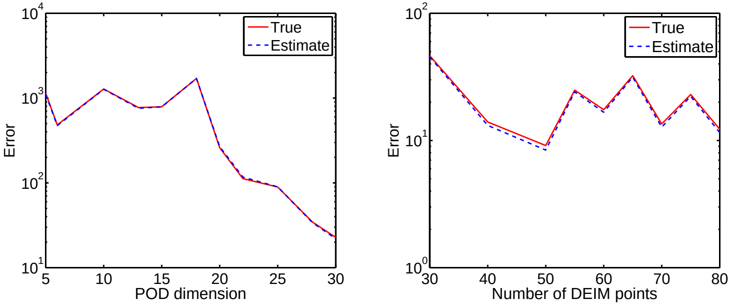
Figure: Adaptive vs traditional DEIM errors approximation errors of the quantity of interest.

Numerical Results

- SWE model using the β-plane approximation on a rectangular domain.
- The alternating direction fully implicit (ADI) scheme.
- The domain is discretized using a mesh of $31 \times 17 = 527$ points, with $\Delta x = 200\text{km}$ and $\Delta y = 275\text{km}$. We select the integration time window to be 24h and we use 181 time steps corresponding to $\Delta t = 480\text{s}$.
- The considered quantity of interest depends on some particular components of the geopotential ϕ at the final time step

$$\mathcal{Q}(\phi) = \sum_{i=1}^{6} \sum_{j=2}^{8} \phi(x_i, y_j, t_{N_t}),\ [x_1, x_6] \times [y_2, y_8] = [0, 1000]\mathrm{km}\ \times [275, 1925]\mathrm{km}. \tag{22}$$

Numerical Results



(a) Number of DEIM points = 35 (b) Dimension of POD basis = 30

Figure: A-posteriori error estimates for the same parametric configuration.

Numerical Results

- The continuity equation dual weighted residuals are employed together with the non-linear basis of the non-linear term F_{31}.

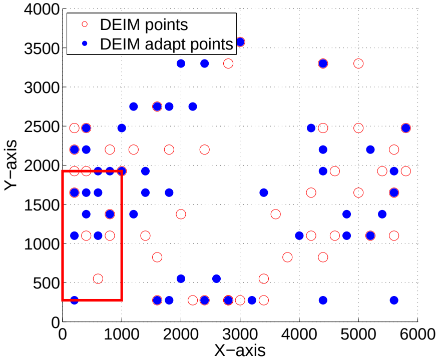
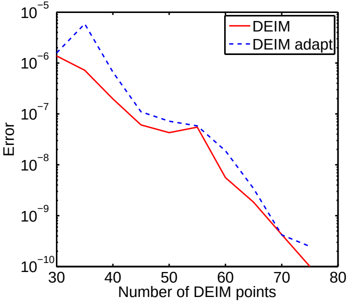
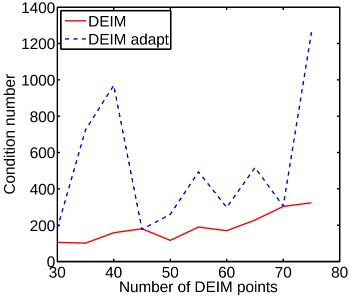


Figure: Adaptive vs traditional DEIM points - $F_{31} = \phi u_x + \phi_x u$.

Numerical Results



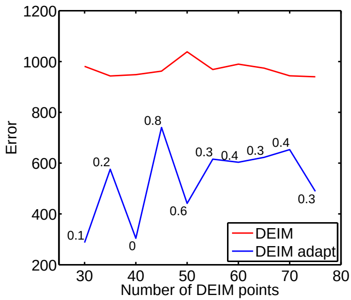
(a) Non-linear approximation



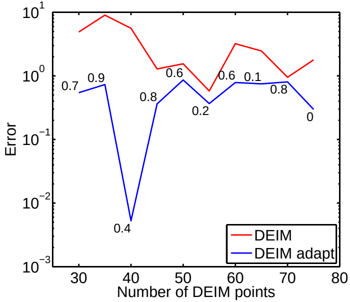
(b) Condition numbers

Figure: Comparison between traditional and adaptive DEIM strategies - Global non-linear term error at time step t_2 in the Euclidian norm (left panel); Condition number of matrix $P^T V$ for non-linear term F_{31} (right panel).

Numerical Results



(a) Dimension of POD basis = 10



(b) Dimension of POD basis = 20

Figure: Adaptive vs traditional DEIM errors of the quantity of interest.

Discussions and Conclusions

- Stabilization issues - condition number of the $(P^T V)^{-1}$; greedy algorithm that relaxes the condition of selecting the location of the largest absolute value of the residuals;
- The error bounds proposed by Chaturantabut and Sorensen (2010) are still valid;
- Comparison with the recent proposed updated optimized rank-one approximation -Peherstorfer and Willcox (2015) - using basis vectors of dual weighted residuals;
- Extension to ROM optimization and adapt on the fly the DEIM interpolation location using aposteriori error estimates for the sub-optimal solution.