A-posteriori error estimation for DEIM reduced nonlinear dynamical systems

at
MoRePaS II Conference 2012
Schloss Reisensburg, Ulm

Dipl. Math. Dipl. Inf. Daniel Wirtz
October 16, 2012

Dipl. Math. Dipl. Inf. Daniel Wirtz, IANS, University of Stuttgart
Prof. Dr. D. Sorensen, CAAM, Rice University
Jun.-Prof. Dr. Bernard Haasdonk, IANS, University of Stuttgart
daniel.wirtz@mathematik.uni-stuttgart.de
Introduction

The Menu today

- Introduction
  - DEIM
  - Dynamical Systems
  - Model Order Reduction (MOR)
  - The Experiment!

- DEIM approximation error estimation
- MOR error estimation
- Matrix DEIM and Partial Similarity Transformations
- Discussion
Introduction
Discrete Empirical Interpolation Method (DEIM)

Notation

 Scalars $a, x, n$, vectors $y, z$ and matrices $A, V$

DEIM: Quick ’n Dirty

 Have nonlinear function $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$
 m-th order DEIM approximation of $f$:

$$\tilde{f}_m(y) := U_m(P_m^T U_m)^{-1} P_m^T f(y)$$

 $U_m = [u_1, \ldots, u_m]$ $m$-basis
 $(P_m^T U_m)^{-1}$ combination weights
 $P_m = [e_{g1}, \ldots, e_{gm}]$ component selection
 $e_i \in \mathbb{R}^d$ denotes the $i$-th unit vector in $\mathbb{R}^d$

Details: [Chaturantabut & Sorensen(2010)]
Introduction
Dynamical Systems

Base considered dynamical system structure

\[ y'(t) = f(y(t)), \quad y(0) = y_0 \]  

- System state \( y(t) \in \mathbb{R}^d \) for \( t \in [0, T] \)
- Initial state \( y_0 \in \mathbb{R}^d \)

More complex settings
Concepts readily extendable by time-/parameter affine components

- Initial values \( y_0(\mu) \)
- Inputs \( B(t, \mu), u(t) \)
- Output \( C(t, \mu) \)
- See Experiment!
Petrov-Galerkin projection of the full system (1)

\[ z'(t) = W^T \tilde{f}_m(Vz(t)) = \tilde{U}_m P_m^T f(Vz(t)) \]  \hspace{1cm} (2)

\[ z(0) = z_0 := W^T y_0, \]  \hspace{1cm} (3)

- Reduced variable \( z(t) \in \mathbb{R}^r \) at times \( t \in [0, T] \)
- Biorthogonal matrices \( V, W \in \mathbb{R}^{r \times d} \), \( V^T W = I_r \) with \( r \leq d \) (ideally \( r \ll d \))
- \( \tilde{U}_m := W^T U_m (P_m^T U_m)^{-1} \)
- Note: We have Galerkin case \( V = W \) in experiments.
Introduction

Experiment: 1D unsteady viscous Burgers’ equation

- Domain $\Omega := [0, 1]$ and time $t \in [0, T]$ with $T = 1$
- Governing system with parameter $\mu \in \mathcal{P} := [0.01, 0.06]$

$$\frac{\partial y}{\partial t}(x, t) = \mu \frac{\partial^2 y}{\partial x^2}(x, t) - \frac{\partial}{\partial x} \left( \frac{y(x, t)^2}{2} \right) + \langle \beta(x), u(t) \rangle,$$

- Boundary conditions

$$y(0, t) = y(1, t) = 0, \quad t \in [0, T], \quad y(x, 0) = 0, \quad x \in \Omega.$$

- External forces $u(t)$ at locations $\beta(x)$:

$$u_1(t) = \sin(2\pi t), \quad b_1(x) = \begin{cases} 4e^{-\left(\frac{x-0.2}{0.03}\right)^2} & x \in [0.1, 0.3], \\ 0 & \text{else}, \end{cases},$$

$$u_2(t) = \begin{cases} 1 & t \in [0.2, 0.4], \\ 0 & \text{else}, \end{cases}, \quad b_2(x) = \begin{cases} 4 & x \in [0.6, 0.7], \\ 0 & \text{else}. \end{cases}$$
Finite differences over a \( n = 500 \) grid gives dynamical system

\[
\frac{d}{dt} y(t) = \mu A y(t) + f(y(t)) + B u(t)
\]

- Discrete Laplacian \( A \in \mathbb{R}^{n \times n} \) and \( B \in \mathbb{R}^{n \times 2} \)
- \( f(y) = -y \cdot A_x y \) with 1st-order discrete diff. operator \( A_x \in \mathbb{R}^{n \times n} \)

Time integration with semi-implicit Euler scheme

\[
(I_n - \Delta t \mu A) y(t_{i+1}) = y(t_i) + \Delta t (f(y(t_i)) + B u(t_i))
\]

- \( n_t = 100 \) equidistant time-steps \( t_i := (i - 1) \Delta t \), \( \Delta t = \frac{T}{n_t - 1}, i = 1 \ldots n_t \)
- Training for 100 log-spaced parameters
- \( V = W \) by POD on trajectories, incl. \( f(x_i) \) values and \( \text{span} \{ B \} \)
Introduction

Experiment plots

Figure: Simulation results for both ends of the parameter range $\mathcal{P}$ and their difference

Figure: Full (left), reduced simulation (middle) and the absolute error (right) for $m = 12, \mu = 0.04$
DEIM approximation error estimation

MoRePaS

DEIM approximation error estimation
First central question: How to measure $f - \tilde{f}$?

**Matrix decomposition lemma**

- $d, m, m' \in \mathbb{N}, m + m' \leq d$
- $U_m, P_m \in \mathbb{R}^{d \times m}, U_m', P_m' \in \mathbb{R}^{d \times m'}$
- $U := [U_m, U_m'], P := [P_m, P_m']$ lin. indep. col.
- $P_m^T U_m$ is nonsingular

Then

$$U(P^T U)^{-1} P^T = U_m(P_m^T U_m)^{-1} P_m^T$$

$$+ (U_m A^{-1} B - U_m') F^{-1} (EP_m^T - P_m'^T),$$

with suitable $A, B, C, D, E, F$. 

[34x231]DEIM approximation error estimation
Preparations
**DEIM approximation error estimation**

**Approximation error**

**Theorem (DEIM error representation)**

- **DEIM basis** $\mathcal{U} := \{u_1, \ldots, u_M\}$, points $\mathcal{E} := \{\varrho_1, \ldots, \varrho_M\}$
- **Assume** $\tilde{f}_M \equiv f$ on $\Omega$. (exactness for $m = M$!)
- $P_m := [e_{\varrho_1}, \ldots, e_{\varrho_m}]$, $P_{m'} := [e_{\varrho_{m+1}}, \ldots, e_{\varrho_{m+m'}}]$
- $U_m := [u_1, \ldots, u_m]$, $U_{m'} := [u_{m+1}, \ldots, u_{m+m'}]$

Then with $m \leq M - 1$, $m' := M - m$ the approximation error is

$$f(y) - \tilde{f}_m(y) = (U_m A^{-1} B - U_{m'}) F^{-1} (E P_m^T - P_{m'}^T) f(y),$$

$$= (M_1 P_m^T - M_2 P_{m'}^T) f(y),$$

with $M_1, M_2$ suitable.

- Already mentioned in [Barrault et al. (2004)Barrault, Maday, Nguyen & Patera], used with $m' > 1$ in [Drohmann et al. (2012)Drohmann, Haasdonk & Ohlberger]
DEIM approximation error estimation

Experiments

- Nonlinearity $f$ of Burgers equation
- Approximation errors over DEIM training data $Y$, $|Y| = 10\,100$
- Full dimension $d = 500$, max. DEIM order $M = 200$
- Using $L^2$ in state space and $L^\infty$ norm over $Y$

Figure: True (left) and estimated (right) absolute DEIM approximation errors for different orders/error orders
DEIM approximation error estimation

Experiments

<table>
<thead>
<tr>
<th>m</th>
<th>0.1</th>
<th>0.01</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>14/4</td>
<td>25/14</td>
</tr>
<tr>
<td>8</td>
<td>14/4</td>
<td>25/13</td>
</tr>
<tr>
<td>16</td>
<td>15/6</td>
<td>21/15</td>
</tr>
<tr>
<td>24</td>
<td>13/5</td>
<td>23/13</td>
</tr>
<tr>
<td>32</td>
<td>12/5</td>
<td>19/12</td>
</tr>
<tr>
<td>40</td>
<td>13/6</td>
<td>20/11</td>
</tr>
<tr>
<td>48</td>
<td>12/6</td>
<td>20/12</td>
</tr>
<tr>
<td>56</td>
<td>16/4</td>
<td>21/12</td>
</tr>
<tr>
<td>64</td>
<td>13/3</td>
<td>28/13</td>
</tr>
<tr>
<td>72</td>
<td>18/5</td>
<td>28/16</td>
</tr>
<tr>
<td>80</td>
<td>18/8</td>
<td>32/17</td>
</tr>
<tr>
<td>88</td>
<td>21/4</td>
<td>33/20</td>
</tr>
<tr>
<td>96</td>
<td>23/12</td>
<td>33/22</td>
</tr>
<tr>
<td>104</td>
<td>20/5</td>
<td>32/18</td>
</tr>
<tr>
<td>112</td>
<td>17/7</td>
<td>36/17</td>
</tr>
<tr>
<td>120</td>
<td>23/6</td>
<td>32/21</td>
</tr>
</tbody>
</table>

Figure: Maximum relative error $||f - \tilde{f} - (M_1 P_{m}^T - M_2 P_{m'\epsilon}^T) f|| ||f - \tilde{f}||^{-1}$ between true and estimated DEIM approximation error over $Y$. Transparent plane located at $10^{-2}$. Table: Req. min $m'$-values for max/avg rel. err smaller than column header value.
DEIM approximation error estimation

Experiments

Figure: Minimum required $m'$ values for relative error < 0.1 (left) and < 0.01 (right) for different discretizations.
MOR error estimation

MoRePaS

MOR error estimation
Let $G \in \mathbb{R}^{d \times d}$ denote a symmetric positive definite weighting matrix.

**Definition (Logarithmic Lipschitz constants)**

For a function $f : \mathbb{R}^d \to \mathbb{R}^d$ we define the logarithmic Lipschitz constant with respect to $G$ by

$$L_G[f] := \lim_{h \to 0^+} \frac{1}{h} \left( \sup_{x \neq y \in \mathbb{R}^d} \frac{\|x - y + h(f(x) - f(y))\|_G}{\|x - y\|_G} - 1 \right).$$

For Lipschitz-continuous functions we have the equivalent

$$L_G[f] = \sup_{x \neq y \in \mathbb{R}^d} \frac{\langle x - y, f(x) - f(y) \rangle_G}{\|x - y\|_G^2}. \quad \text{Background: [Dahlquist(1959), Söderlind(2006)]}$$
MOR error estimation

Main estimation result

- \( y^r(t) := V z(t), \quad e(t) := y(t) - y^r(t), \quad t \in [0, T] \)
- Obtained by estimation of error system [incl. Comp. Lemma]

\[
e'(t) = f(y(t)) - VW^T \tilde{f}_m(y^r(t)), \quad e(0) = y_0 - VW^T y_0.
\]
MOR error estimation

Main estimation result

- \( y^r(t) := Vz(t), \ e(t) := y(t) - y^r(t), \ t \in [0, T] \)
- Obtained by estimation of error system [incl. Comp. Lemma]

\[
e'(t) = f(y(t)) - VW^T \tilde{f}_m(y^r(t)), \quad e(0) = y_0 - VW^T y_0.
\]

Theorem (A-post. error estimation for DEIM reduced systems)

The state space error is rigorously bounded via

\[
\|e(t)\|_G \leq \Delta(t) \quad \forall \ t \in [0, T],
\]

with

\[
\Delta(t) := \int_0^t \alpha(s)e^{L_G[f](t-s)}ds + e^{L_G[f]}t \left\|y_0 - VW^T y_0\right\|_G,
\]

\[
\alpha(t) := \left\|M_1P_m^T f(y^r(t), t) - M_2P_m' f(y^r(t), t)\right\|_G,
\]

- \( M_1, M_2 \) suitable, \( \alpha(t) \) offline/online decomposable
MOR error estimation

Estimation results

- All estimations use “true local log. Lip. const.”

\[ L_G[f] \Rightarrow \frac{\langle e(t), f(y(t)) - f(y^r(t)) \rangle}{\|e(t)\|_G^2} \]

- Reference estimate: Uses true DEIM approximation error

Figure: Absolute errors over time for different DEIM \( m' \) approximation error orders, \( \mu = 0.04, m = 12 \)
MOR error estimation

More workarounds!

- Always: We don't have $y(t)$ at reduced simulations
- We most likely also don't have $L_G[f]$

Key idea with approximation and localization

$$\frac{\langle e(t), f(y(t)) - f(y^r(t)) \rangle}{\|e(t)\|_G^2} = \frac{\langle e(t), J(y^r(t))e(t) + O\left(\|e(t)\|_G^2\right) \rangle}{\|e(t)\|_G^2}$$

$$= \frac{\langle e(t), J(y^r(t))e(t) \rangle}{\|e(t)\|_G^2} + O\left(\|e(t)\|_G\right)$$

$$\leq L_G[J(y^r(t))] + O\left(\|e(t)\|_G\right).$$

+ Only $y^r(t)$ required & localized
- Need Jacobian $J$ and its logarithmic norm $L_G[J]$
MOR error estimation

Estimation results

Now replace

$$\langle e(t), f(y(t)) - f(y^r(t)) \rangle_G \Rightarrow L_G[J(y^r(t))]$$

with

$$L_G[A] = \max \left\{ \sigma \left( \frac{1}{2} (\tilde{A} + \tilde{A}^T) \right) \right\}$$

with \( \tilde{A} := C^T AC^{-T} \) and \( G = CC^T \).

Figure: Absolute errors over time for different DEIM \( m' \) approximation error orders, \( \mu = 0.04, m = 12 \)
Partial Similarity Transformations and Matrix DEIM

MoRePaS

Partial Similarity Transformations and Matrix DEIM
Lemma (Approx. of eigenvalues for a family of sym. matrices)

- $H(t) \in \mathbb{R}^{d \times d}$ fam. of sym. matrices over $t \in [a, b]$
- $[\lambda(t), q(t)] := \lambda_{\text{max}}(H(t))$ largest e-val $\lambda(t)$ / normed e-vec $q(t)$
- $\sup_{t \in [a, b]} \|H(t)\| \leq C_H$ holds
- Define $R = \int_a^b q(t)q(t)^T dt$ and let $Q\Sigma^2Q^T = R$ with $Q^TQ = I$
- Let $\Sigma = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_d)$ with $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_d > 0$
- For $k \leq d$ let $Q_k := Q(:, 1:k)$ and $\lambda_k(t) := \lambda_{\text{max}}(Q_k^TH(t)Q_k)$

Then

$$\int_a^b |\lambda(t) - \lambda_k(t)| dt \leq 4C_H \sum_{j > k} \sigma_j^2. \quad (4)$$
Partial Similarity Transformations and Matrix DEIM

Application of transformation to Jacobian $\mathbf{J}$

- Consider

$$H(t) := \frac{1}{2} \left( \mathbf{C}^T \mathbf{J}(\mathbf{y}^r(t)) \mathbf{C}^{-T} + (\mathbf{C}^T \mathbf{J}(\mathbf{y}^r(t)) \mathbf{C}^{-T})^T \right)$$

- Have the corresponding values $C_H > 0$, $\{\sigma_i\}_{i=1}^d$, $\mathbf{Q} \in \mathbb{R}^{d \times d}$

- Identify

$$\lambda(t) = L_G[\mathbf{J}(\mathbf{y}^r(t))],$$

$$\lambda_k(t) = L_{I_k} \left[ \mathbf{Q}_k^T \mathbf{C}^T \mathbf{J}(\mathbf{y}^r(t)) \mathbf{C}^{-T} \mathbf{Q}_k \right],$$

**Jacobian partial similarity transform**

$$\int_0^T \left| L_G[\mathbf{J}(\mathbf{y}^r(t))] - L_{I_k} \left[ \mathbf{Q}_k^T \mathbf{C}^T \mathbf{J}(\mathbf{y}^r(t)) \mathbf{C}^{-T} \mathbf{Q}_k \right] \right| dt \leq C_H \sum_{j>k} \sigma_j^2.$$
Take a breath..

What do we have?

- **Already done**: Replaced $L_G[f]$ by $L_G[J(y^r(t))]$
  - Localized
  - Likely available

Reduced cost for eigenvalue problem: Size $d$ to size $k \ll d$

But possibly high offline cost

Still unanswered:

$Q^T_k C^T J(y^r(t)) C - T Q_k \in \mathbb{R}^{k \times k}$, but

$J(y^r(t)) \in \mathbb{R}^{d \times d}$!
Partial Similarity Transformations and Matrix DEIM

Take a breath..

What do we have?

- Already done: Replaced $L_G[f]$ by $L_G[J(y^r(t))]$
  - Localized
  - Likely available
- New: Replaced $L_G[J(y^r(t))]$ by $L_{I_k}[Q_k^T C^T J(y^r(t)) C^{-T} Q_k]$
  - Reduced cost for eigenvalue problem: Size $d$ to size $k \ll d$
  - But possibly high offline cost
Partial Similarity Transformations and Matrix DEIM

Take a breath..

What do we have?

- **Already done**: Replaced \( L_G[f] \) by \( L_G[J(y^r(t))] \)
  - Localized
  - Likely available
- **New**: Replaced \( L_G[J(y^r(t))] \) by \( L_{I_k}[Q_k^T C^T J(y^r(t)) C^{-T} Q_k] \)
  - Reduced cost for eigenvalue problem: Size \( d \) to size \( k \ll d \)
  - But possibly high offline cost
- **Still unanswered**: \( Q_k^T C^T J(y^r(t)) C^{-T} Q_k \in \mathbb{R}^{k \times k} \), but \( J(y^r(t)) \in \mathbb{R}^{d \times d} \)
- **We need some offline/online separability of \( J \)!**
Partial Similarity Transformations and Matrix DEIM

Matrix DEIM

For $A \in \mathbb{R}^{d \times d}$ define the transformation (vec-operation)

$$\mathcal{V} : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^{d^2}$$

$$A \mapsto \mathcal{V}[A] := (A^T_1, A^T_2, \ldots, A^T_d)^T,$$

Matrix DEIM

- Choose $M_J \leq d$
- Let $U_{M_J}, P_{M_J} \in \mathbb{R}^{d^2 \times M_J}$ be basis/point matrices for $\mathcal{V}[J(y)]$

Then, for $m_J \leq M_J$, the $m_J$-th order MDEIM approximation of $J$ is given via

$$\tilde{J}_{m_J}(y) := \mathcal{V}^{-1} \left[ (U_{m_J}(P_{m_J} U_{m_J})^{-1} P_{m_J}^T \mathcal{V}[J(y)]) \right],$$

where $U_{m_J} := U_{M_J}(\cdot, 1:m_J)$ and $P_{m_J} := P_{M_J}(\cdot, 1:m_J).$
off1 Compute $U_{M,J}, P_{M,J}, Q$ once, then for each new $m_j, k$ do
Partial Similarity Transformations and Matrix DEIM

Matrix DEIM: Offline/online decomposition

**off1** Compute $U_{M_J}, P_{M_J}, Q$ once, then for each new $m_J, k$ do

**off2** Compute offline vectors for Matrix DEIM of Jacobian via

\[
\hat{U}_{m_J} := U_{M_J}(:, 1:m_J), \quad \hat{P}_{m_J} := P_{M_J}(:, 1:m_J), \\
\hat{U} := \hat{U}_{m_J} \left( \hat{P}_{m_J}^T \hat{U}_{m_J} \right)^{-1}.
\]
Partial Similarity Transformations and Matrix DEIM

Matrix DEIM: Offline/online decomposition

1. **off1** Compute $U_{M,J}$, $P_{M,J}$, $Q$ once, then for each new $m_J, k$ do

2. **off2** Compute offline vectors for Matrix DEIM of Jacobian via

   \[
   \hat{U}_{m_J} := U_{M,J}(:,1:m_J),\quad \hat{P}_{m_J} := P_{M,J}(:,1:m_J),
   \]

   \[
   \hat{U} := \hat{U}_{m_J}(\hat{P}_m^T \hat{U}_{m_J})^{-1}.
   \]

3. **off2** Select partial similarity transform matrix of size $k$ as

   \[
   Q_k := Q(:,1:k)
   \]

   and compute

   \[
   \tilde{U}(:,j) := \mathcal{V}_k [Q_k^T C^T \mathcal{V}_k^{-1} [\hat{U}(:,j)] C^{-T} Q_k] \in \mathbb{R}^{k^2}, j = 1 \ldots m_J,
   \]

   where $\mathcal{V}_k$ denotes the same transformation as $\mathcal{V}$ but for $k \times k$ matrices.
Partial Similarity Transformations and Matrix DEIM

**Matrix DEIM: Offline/online decomposition**

**off1** Compute $U_{m_J}, P_{m_J}, Q$ once, then for each new $m_J, k$ do

**off2** Compute offline vectors for Matrix DEIM of Jacobian via

\[ \hat{U}_{m_J} := U_{m_J}(:, 1:m_J), \quad \hat{P}_{m_J} := P_{m_J}(:, 1:m_J), \]
\[ \hat{U} := \hat{U}_{m_J}(\hat{P}_m^T \hat{U}_{m_J})^{-1}. \]

**off2** Select partial similarity transform matrix of size $k$ as
\[ Q_k := Q(:, 1:k) \]
and compute
\[ \tilde{U}(::, j) := V_k [Q_k^T C^T V_k^{-1} \hat{U}(::, j)] C^{-T} Q_k \in \mathbb{R}^{k^2}, j = 1 \ldots m_J, \]
where $V_k$ denotes the same transformation as $V$ but for $k \times k$ matrices.

**online** For new $y^r(t)$ compute $L_{I_k} [V_k^{-1} [\hat{U} \hat{P}_m^T V[J(y^r(t))]]]$ in $O(k^3 + g_f(m_J))$ ($g_f$ “internal $f$ complexity”)
Partial Similarity Transformations and Matrix DEIM

Some estimation results

Figure: Absolute errors over time for different matrix DEIM orders and partial similarity transformation sizes
Partial Similarity Transformations and Matrix DEIM

Some estimation results

Figure: Absolute errors over time for different matrix DEIM orders and partial similarity transformation sizes
Summary

... any questions?

Short summary

- DEIM error estimation using next basis functions
- Local approximation of Lipschitz constants via Jacobian logarithmic norms
- Efficient computation of Jacobian logarithmic norm via
  - Partial similarity transformation
  - Matrix DEIM
- Bottom line: Non-rigorous but applicable a-posteriori error estimator!
Summary

...any questions?

Short summary

- DEIM error estimation using next basis functions
- Local approximation of Lipschitz constants via Jacobian logarithmic norms
- Efficient computation of Jacobian logarithmic norm via
  - Partial similarity transformation
  - Matrix DEIM
- Bottom line: Non-rigorous but applicable a-posteriori error estimator!

Thank you for your attention! Any questions/remarks?


[Drohmann et al. (2012) Drohmann, Haasdonk & Ohlberger]


Image sources

- http://www.bildungsstiftung.org/05_aktionen.htm