RESEARCH ARTICLE

Multi-level adaptive greedy algorithms for the reduced basis method

Jiahua Jiang$^1$ | Yanlai Chen$^*$$^2$

$^1$Department of Mathematics, Virginia Tech University, VA, U.S.A.
Email: jiahua@vt.edu

$^2$Department of Mathematics, University of Massachusetts Dartmouth, MA, U.S.A.
Email: Yanlai.Chen@umassd.edu

Correspondence
$^*$Corresponding author.

Summary
The reduced basis method (RBM) empowers repeated and rapid evaluation of parametrized partial differential equations through an offline-online decomposition, a.k.a. a learning-execution process. A key feature of the method is a greedy algorithm repeatedly scanning the training set, a fine discretization of the parameter domain, to identify the next dimension of the parameter-induced solution manifold along which we expand the surrogate solution space. Although successfully applied to problems with fairly high parametric dimensions, the challenge is that this scanning cost dominates the offline cost due to it being proportional to the cardinality of the training set which is exponential with respect to the parameter dimension. In this work, we review three recent attempts in effectively delaying this curse of dimensionality, and propose two new hybrid strategies through successive refinement and multilevel maximization of the error estimate over the training set. All five offline-enhanced methods and the original greedy algorithm are tested and compared on a typical elliptic equation as well as an indefinite Helmholtz problem.

KEYWORDS:
Reduced basis method, Greedy algorithm

1 INTRODUCTION

The reduced basis method (RBM) has proved to be a viable option for the purpose of designing fast numerical algorithms for parametrized systems. The parameters may include boundary conditions, material properties, amount of uncertainty, geometric configurations, source properties etc. To describe a realistic system, practitioners often resort to a large number of parameters making the generation of a reduced model strenuous.

The critical tool of RBM, toward attaining orders of magnitude gain in marginal (i.e. per parameter instance) computation time, is an offline-online decomposition process where the basis selection (i.e. surrogate solution space building) is performed offline by a greedy algorithm, see e.g. for details. The ultimate goal is that the complexity of the reduced solver, presumably to be called for an overwhelming number of times or even in a realtime fashion, is independent of the degrees of freedom of the high-fidelity approximation of the basis functions for this surrogate space.

The dimension-by-dimension construction of the surrogate solution space relies on a greedy scheme that keeps track of an efficiently-computable error estimator/indicator, denoted by $\Delta_N(\mu)$, indicating the discrepancy between the dimension-$n$ surrogate solution (RB solution) and the high-fidelity solution. This greedy procedure starts by selecting the first parameter $\mu^1$ randomly from the training set and obtaining its corresponding high-fidelity truth approximation $u^N(\mu^1)$. We then have a one-dimensional RB space $W_1 = \{u^N(\mu^1)\}$. Next, the scheme obtains an RB approximation $u^1_N(\mu)$ for each parameter in the training
set together with \( \Delta_1(\mu) \). The greedy choice for the \((n+1)\)th parameter \((n = 1, 2, \cdots)\) is made and the RB space augmented by

\[
\mu^{n+1} = \arg\max_{\mu \in \Xi_{\text{train}}} \Delta_n(\mu), \quad W_{n+1} = W_n \oplus \{u^N(\mu^{n+1})\}.
\]

The other parts of the offline process are devoted to the necessary preparations for the online reduced solver, a variational (i.e., Galerkin or Petrov-Galerkin) projection \(\Xi\) into the surrogate space, or a sparse representation of the solution at strategically chosen points in the physical domain. The greedy procedure implies that the offline cost is proportional to the cardinality of the training set. It is essential for this training set to be fine enough so that we are not missing critical phenomena not represented by \(\Xi_{\text{train}}\). In particular, its cardinality is exponential with respect to the parameter dimension. As a consequence, the “break-even” number of simulations for the parametric system (i.e., minimum number of simulations that make the offline preparation stage worthwhile) is exponential with respect to the parameter dimension as well, severely diminishing the attractiveness of RBM.

The need of delaying this curse of dimensionality motivates numerous recent attempts in designing offline-enhancement strategies for systems with high-dimensional parameter domain. Indeed, the authors of propose to decompose the given parameter training set \(\Xi_{\text{train}}\) \(a\ priori\) into a sequence of subsets that increase geometrically in size, and then perform the classical greedy on each of them. Authors of take a different route. They fix the cardinality of an active training set (much smaller than the full training set) with part of it replaced after each greedy step. Finally, promotes to recycle the information afforded by the error estimates \(\{\Delta_n(\mu)\}\) and to construct a surrogate training set that is adaptively reconstructed. There have been other techniques in the literature to alleviate the RB offline cost such as the parameter domain adaptivity, greedy sampling acceleration through nonlinear optimization, and local reduced basis method, etc.

In this paper, we focus on the three approaches of that are similar in nature. We review the essence of each algorithm, and more importantly, propose two new hybrid strategies through successive refinement and multilevel maximization of \(\{\Delta_n(\mu)\}\) over the training set. By testing all these five strategies and the classical greedy, we demonstrate that the two new hybrids outperform others some of which may even fail for the more challenging Helmholtz problems.

The paper is organized as follows. In Section 2 we review the three existing offline enhancement approaches mentioned above. Section 3 is devoted to the construction of the hybrid algorithms. Numerical results for two test problems to demonstrate the accuracy and efficiency of these offline improvement methods are shown in Section 4. Finally, concluding remarks are drawn in Section 5.

## 2 | BACKGROUND

In this section, we briefly review the classical reduced basis method and the three offline enhancement strategies. This also serves as a background and motivation for the subsequent discussion of our newly proposed multi-level greedy techniques leading to more efficient reduced basis methods. Table 1 outlines our notations.

### 2.1 | Reduced Basis Methods

Given \(\mu \in D\), the goal is to evaluate a certain output of interest

\[
s(\mu) = \ell^\prime(u(\mu); \mu),
\]

where the function \(u(\mu) \in X\) satisfies

\[
a(u(\mu), v; \mu) = f(v; \mu), \quad v \in X,
\]

which is a parametrized partial differential equation (pPDE) written in a weak form with \(\mu \in D\) being the (possibly multi-dimensional) parameter. Here \(X = X(\Omega)\) is a Hilbert space satisfying, e.g., \(H^1_0(\Omega) \subset X(\Omega) \subset H^1(\Omega)\). We denote by \(\langle \cdot, \cdot \rangle_X\) the inner product associated with the space \(X\), whose induced norm \(\|\cdot\|_X\) is equivalent to the usual \(H^1(\Omega)\) norm. We typically assume that \(a(\cdot, \cdot; \mu) : X \times X \to \mathbb{R}\) is continuous and coercive over \(X\) uniformly in \(\mu \in D\), that is,

\[
\gamma(\mu) := \sup_{u \in X} \sup_{v \in X} a(u, v; \mu) \frac{|u|_X |v|_X}{\|u\|_X \|v\|_X} < \infty, \quad \forall \mu \in D,
\]

\[
a(\mu) := \inf_{u \in X} a(u, v; \mu) \|u\|_X^2 \geq a_0 > 0, \forall \mu \in D.
\]
**Jiang, Chen: Multi-level Greedy Algorithms for RBM**

Parameter in $D \subseteq \mathbb{R}^p$

Function-valued solution of a parameterized PDE

Degrees of freedom (DoF) in PDE "truth" solver

Truth solution (finite-dimensional)

Number of reduced basis snapshots, $N \ll \mathcal{N}$

Terminal number of reduced bases

"Snapshot" parameter values, $j = 1, \ldots, N$

Span of $u^N(\mu)$ for $k = 1, \ldots, N$

Reduced basis solution, $u^N \in X_N^N$

Reduced basis solution error, equals $u^N(\mu) - u^N_N(\mu)$

Parameter training set, a finite subset of $\mathcal{X}$

Size of $\mathcal{X}_{train}$

Error estimate (upper bound) for $\|e_N(\mu)\|$

Error estimate stopping tolerance in greedy sweep

**A priori training set decomposition**

Number of sample sets

The $j$-th sample set

A posteriori error estimate with $N$ reduced bases at $j$-th sample set

**Adaptive enriching**

Sample parameters of fixed cardinality ($< n_{train}$)

Size of the active training set $\mathcal{X}$, much smaller than $n_{train}$

**Adaptive construction of surrogate training sets**

Number of “outer” loops in the offline enhancement procedure, Algorithm 4

The largest error estimator at the beginning of outer loop $\ell$

Surrogate Training Set (STS), a subset of $\mathcal{X}_{train}$

A constant integer controlling the damping ratio of the error estimate when examining the STS

A constant integer adjusting the size of surrogate training set

**TABLE 1** Notation used throughout this article.

$f(\cdot)$ and $\ell(\cdot)$ are linear continuous functionals over $X$. We assume that there is a finite-dimensional discretization for the model problem [3]: The solution space $X$ is discretized by an $\mathcal{N}$-dimensional space $X^\mathcal{N}$ (i.e., $\dim(X^\mathcal{N}) = \mathcal{N}$) and (2) and (3) are discretized as

\[
\begin{align*}
\text{For } \mu \in D, \text{ solve} \quad & s^\mathcal{N} = \ell(u^N(\mu); \mu) \text{ where } u^N(\mu) \in X^\mathcal{N} \text{ satisfies} \\
& a(u^N, v; \mu) = f(v; \mu) \quad \forall v \in X^\mathcal{N}. 
\end{align*}
\]

The relevant quantities such as the coercivity constant [4b] are defined according to the discretization,

\[
\alpha^\mathcal{N}(\mu) = \inf_{u \in X^\mathcal{N}} \frac{a(u, u; \mu)}{||u||^2_X} \geq \alpha_0, \forall \mu \in D.
\]

In the RBM literature, any quantity associated to $\mathcal{N}$ is called a “truth”. E.g., $u^\mathcal{N}$ is called the “truth solution”, [5] “truth solver” i.e. Full Order Model (FOM). $\mathcal{N}$ is typically very large so that resolving the FOM gives highly accurate approximations for all $\mu \in D$.

For a training set $\mathcal{X}_{train} \subset D$ which consists of a fine discretization of $D$ of finite cardinality and a collection of $N$ parameters $S_N = \{\mu^1, \ldots, \mu^N\} \subset \mathcal{X}_{train}$, we define the reduced basis space as

\[
X^N_N := \text{span}\{u^N(\mu^n), 1 \leq n \leq N\}.
\]
The reduced basis approximation is now defined as: Given \( \mu \in D \), seek a surrogate RB solution \( u_N^\mathcal{N}(\mu) \) by solving the following reduced system

\[
\begin{align*}
\begin{cases}
\text{For } \mu \in D, \text{ evaluate} \\
\mathcal{N}^X \neq \mathcal{C}^\varepsilon(u_N^\mathcal{N}(\mu); \mu) \text{ s.t. } u_N^\mathcal{N}(\mu) \in X_N^\mathcal{N} \subset X^\mathcal{N} \text{ satisfies} \\
a(u_N^\mathcal{N}, v; \mu) = f(v) \quad \forall v \in X_N^\mathcal{N}.
\end{cases}
\end{align*}
\] (7)

\( \mathcal{N}^X \) is called the reduced solver, i.e. Reduced Order Model (ROM). The typical multiple orders of magnitude speedup of RBM manifests from that the assembly of ROM is independent of \( \mathcal{N} \), which crucially relies on the affine assumption of the parameter dependent problem \( \mathcal{N} \), such as,

\[
a(w, v; \mu) = \sum_{q=1}^{Q_e} \Theta^e_q(\mu)a_q(w, v), \quad \text{and} \quad f(v; \mu) = \sum_{q=1}^{Q_f} \Theta^f_q(\mu)f_q(v).
\] (8)

Here \( \Theta^e_q, \Theta^f_q \) are \( \mu \)-dependent functions, and \( a_q, f_q \) are \( \mu \)-independent forms. With this hypothesis, we can apply an offline-online decomposition to enable fast resolution of the ROM \( \mathcal{N} \). The offline stage, \( a_q(u_N^\mathcal{N}(\mu), u_N^\mathcal{N}(\mu')) \) and \( f_q(u_N^\mathcal{N}(\mu')) \) can be precomputed, which is relatively expensive but only done once. In the online stage, we construct the matrices and vectors in the reduced system \( \mathcal{N} \) and solve the resulting reduced basis problem. We remark that when this assumption is violated, we turn to Empirical Interpolation Method (EIM) to approximate the non-affine operators by affine ones. This is the case for our second test problem.

Leading to the key error estimator \( \Delta_N(\mu) \), we define the error \( e_N(\mu) := u_N^\mathcal{N}(\mu) - u_N^\mathcal{N}(\mu) \in X^\mathcal{N} \), linearity of \( a(\cdot, \cdot; \mu) \) yields the following error equation:

\[
a(e_N(\mu), v; \mu) = r_N(v; \mu) \quad \forall v \in X^\mathcal{N},
\] (9)

where the residual \( r_N(\cdot; \mu) \in (X^\mathcal{N})' \) (the dual of \( X^\mathcal{N} \) ) operated on \( v \in X^\mathcal{N} \) is defined as \( f(v; \mu) - a(u_N^\mathcal{N}(\mu), v; \mu) \). We define the \textit{a posteriori} error estimator for the solution as

\[
\Delta_N(\mu) = \frac{\|r_N(\cdot; \mu)\|_{(X^\mathcal{N})'}}{a_{LB}^N(\mu)}
\] (10)

where \( a_{LB}^N(\mu) \) is a lower bound of the coercivity constant. To build the parameter set \( S_N \) and the reduced basis space \( X_N^\mathcal{N} \), the classical greedy algorithm, outlined in Algorithm 1, is invoked.

**Algorithm 1** Classical Greedy, \((N, X_N^\mathcal{N}) = \mathcal{CG}(\Xi_{\text{train}}, \varepsilon_{\text{tol}}, N, X_N^\mathcal{N})\)

1. if \( N = 0 \) then
2. \quad Initialization: Choose an initial parameter value \( \mu^1 \in \Xi_{\text{train}}, \) set \( S_1 = \{ \mu^1 \} \), compute \( u_N^\mathcal{N}(\mu^1) \), and \( N = 1 \);
3. end if
4. while max \( \Delta_N(\mu) > \varepsilon_{\text{tol}} \) do
5. \quad Choose \( \mu^{N+1} = \arg \max_{\mu \in \Xi_{\text{train}}} \Delta_N(\mu) \);
6. \quad \( S_{N+1} = S_N \cup \{ \mu^{N+1} \} \);
7. \quad Compute \( u_N^\mathcal{N}(\mu^{N+1}) \) and augment RB space \( X_N^\mathcal{N+1} = X_N^\mathcal{N} \oplus \{ u_N^\mathcal{N}(\mu^{N+1}) \} \);
8. \quad \( N \leftarrow N + 1 \);
9. end while

### 2.2 Existing offline enhancement strategies

This greedy algorithm \( \mathcal{CG}(\Xi_{\text{train}}, \varepsilon_{\text{tol}}, N, X_N^\mathcal{N}) \) requires maximization of the \textit{a posteriori} error estimate over \( \Xi_{\text{train}} \). This becomes a bottleneck in the construction of the RB space \( X_N^\mathcal{N} \) , especially when the parameter domain \( D \) is of high dimension. Much recent research has concentrated on the schemes capable of accelerating this exact greedy algorithm. In this section, we review three such offline enhancement strategies as proposed in Section 2.2.1 to 2.2.3 respectively. The improved greedy algorithms are described in Algorithms 2 to 4 accordingly.
2.2.1 A priori training set decomposition (TSD)

In
A modified greedy algorithm is provided to address the many-parameter heat conduction problems. It attempts to run the classical greedy algorithm on a sequence of mutually exclusive subsets of \( \mathcal{X}_{train} = \{ \mathcal{X}_{train,1}, \ldots, \mathcal{X}_{train,J} \} \). They increase in size and form a partition of \( \mathcal{X}_{train} \). The idea is that the small sets are quick to run the algorithm on, and the large ones serve as a vanishing check. In the large sets, the algorithm is rerun only over the samples whose current error estimator is above the tolerance.

The size of the first sample set \( n_{tr,small} = |\mathcal{X}_{train,1}| \) serves as the only tuning parameter for this approach. As an example to build up this partition, let \( J = \lceil \log_2(\frac{N_{\text{max}}}{n_{tr,small}}) \rceil \). For \( j = 2 : J - 1 \), \( |\mathcal{X}_{train,j}| = 2^{j-1} n_{tr,small} \). The size of \( \mathcal{X}_{train,j} \) equals to \( n_{train} - \sum_{j=1}^{J-1} |\mathcal{X}_{train,j}| \). To make each of the subsets span the whole training set, the algorithm randomly sample \( n_{train} \) points from \( \mathcal{X}_{train} \). For \( j \)-th sample set, the \textit{a posteriori} error estimate computed by \( N \) reduced bases is denoted as \( \Delta_{N,j}(\mu) \). The pseudo-code of this methodology is provided in Algorithm 2.

\textbf{Remark 1.} This approach has \( n_{tr,small} \) as the only tuning parameter. Small \( n_{tr,small} \) likely leads to higher dimension of the surrogate solution space negatively impacting the online efficiency. Therefore, we usually choose relatively large \( n_{tr,small} \) in comparison to \( n_{train} \). Unfortunately, large \( n_{tr,small} \) will lead to more costly offline stage. This is obvious since when \( n_{tr,small} = n_{train} \), Algorithm 2 is exactly the same as classical greedy algorithm.

\begin{algorithm}[h]
\caption{Training set decomposition based classical greedy \( (N, X_N^N) = \text{TSD\_CG}(\{\mathcal{X}_{train,j}\}_{j=1}^J, \varepsilon_{tol}) \)}
\begin{algorithmic}[1]
\State \textbf{for} \( j = 1 : J \) \textbf{do}
\State \textbf{Call} \( (N, X_N^N) = \text{CG}(\mathcal{X}_{train,j}, \varepsilon_{tol}, N, X_N^N) \)
\State \textbf{end for}
\end{algorithmic}
\end{algorithm}

2.2.2 Adaptive enriching

Instead of decomposing the training set a priori,\(^{13}\) starts the classical greedy algorithm on a (relatively small) set \( \mathcal{X} \) of fixed cardinality (denoted by \( M_{sample} \)) that randomly samples the parameter domain. This set is dynamically updated, after each greedy step, by removing parameters that have error estimate below the tolerance and adding new parameter values from the training set. The idea is that it is not worthwhile to keep those parameters whose corresponding surrogate solutions are already accurate enough. More details of this algorithm are provided in Algorithm 3. Since the maximum \textit{a posteriori} error estimate in the stopping criteria is only the maximum in sample set \( \mathcal{X} \) (rather than full training set \( \mathcal{X}_{train} \)), the \( N_{safe} \) variable is introduced to make sure each parameter value of the full training set is checked once, albeit at different time of the greedy algorithm. For problems that we don’t have monotonic decay of error or error estimate for, a final “check” over the entire training set is necessary.\(^{13}\)

\textbf{Remark 2.} The adaptive enriching scheme has one tuning parameter, the size of the “fixed window” \( M_{sample} \). Two extremal case are worth mentioning:

- When \( M_{sample} = 1 \), the method becomes the approach taken by\(^{15}\) and\(^{16}\) where each parameter is examined once and decision on keep or toss made immediately. This approach may result in a larger-than-necessary surrogate space.

- When \( M_{sample} = n_{train} \), this algorithm is identical to the classical greedy algorithm, thus no saving realized.

2.2.3 Adaptive construction of surrogate training sets

The main idea of\(^{13}\) is to identify a subset of the training set, a “Surrogate Training Set” (STS) \textit{a posteriori}, and perform greedy algorithm until this STS is satisfactorily resolved. At that point, the algorithm returns to the full training set to generate a new STS. We let \( \mathcal{X}_{sur} \) denote a generic STS. Two techniques for STS construction are provided in\(^{13}\). The one called successive maximization method (SMM) will be described here.

SMM is motivated by the notion that the difference between the norm of the errors \( ||e(\mu_j)||_{\mathcal{X}} - ||e(\mu_k)||_{\mathcal{X}} || \) is partially indicative of the difference between the solutions. When selecting the \( N + 1 \)-th parameter value, we must compute the values \( \Delta_N(\mathcal{X}_{train}) = \{ \Delta_N(\mu) \mid \mu \in \mathcal{X}_{train} \} \). Instead of just exploiting its maximizer, we use the collection to identify the surrogate
training set. Specifically, we equidistantly sample values from $\Delta_N(\Xi_{train})$ to construct the surrogate training set. Indeed, with $\epsilon_{tol}$ the stopping tolerance for the RB sweep, let $\Delta^\text{max}_N = \max_{\mu \in \Xi_{train}} \Delta_N(\mu)$. We define $I^M_N$ as an equi-spaced set between $\epsilon_{tol}$ and $\Delta^\text{max}_N$:

$$I^M_N = \left\{ \nu_{n,m} : \epsilon_{tol} + (\Delta^\text{max}_N - \epsilon_{tol}) \frac{m}{M} : m = 0, \ldots, M - 1 \right\}.$$

(Roughly speaking, we attempt to construct $\Xi_{sur}$ as $\Xi_{sur} = \Delta^{-1}_N \left( I^M_N \right) \cap \Xi_{train}$. On outer loop round $\ell'$, we have $|\Xi_{sur}| \leq M_{\ell'}$ by this construction, where $M_{\ell}$ can be chosen as any monotonically increasing function with respect to $\ell'$. But in order to avoid excessively large $\Xi_{sur}$, we set $M_{\ell} = C_M(\ell' + 1)$, where $C_M$ is a constant. After this construction, we repeatedly sweep the current $\Xi_{sur}$ until

$$\max_{\mu \in \Xi_{sur}} \Delta_N(\mu) \leq E_{\ell'} \frac{1}{(\ell' + 1) \times K_{damp}}$$

where $E_{\ell'}$ is the starting (global) maximum error estimate for this outer loop iteration. The damping ratio for outer loop $\ell'$, $\frac{1}{(\ell' + 1) \times K_{damp}}$, enforces that the maximum error estimate over the $\Xi_{sur}$ decreases by a damping factor controlled by $K_{damp}$ which should be determined by the practitioner and the problem at hand. Following the choice in [13], we take $K_{damp}$ to be constant in this paper. Algorithm 4 details the adaptively constructed STS approach.

Remark 3. This approach has two tuning parameters $C_M$ and $K_{damp}$. The choice of $C_M$ indirectly controls the size of $\Xi_{sur}$. These surrogate training sets should be small enough compared to the full one $\Xi_{train}$ to offer considerable acceleration of the greedy sweep, yet large enough to capture the general landscape of the solution manifold that is iteratively learned. On the other hand, $K_{damp}$ controls how accurate we intend to resolve $u^N_N(\mu)$, where $\mu \in \Xi_{sur}$. The larger $C_M$ and $K_{damp}$ are, the faster algorithm 4 will be.

## 3 HYBRID OFFLINE ENHANCEMENTS FOR THE GREEDY ALGORITHMS

The unifying theme of the three approaches in the last section is the construction of a small-size subset of the full training set on which the classical greedy algorithm is performed. Algorithms 2 and 3 adopt random sampling that does not take into consideration the specific problem at hand and *a posteriori* error estimates already obtained. On the other hand, Algorithm 4 takes full advantage of these information when constructing the small subsets. Nonetheless, the construction of the surrogate training sets relying on the evaluation of $\Delta_N(\mu)$ for all $\mu$ in the training set $\Xi_{train}$ is still a bottleneck of Algorithm 4. To address this, one idea is building up surrogate training set $\Xi_{sur}$ by only computing $\Delta_N(\mu)$ for a limited number of $\mu$. This idea motivates us to develop the following two algorithms. Specifically,
Algorithm 4: Adapting STS-based classical greedy algorithm \((N, X_N^\mathcal{N}) = \text{STS}_\mathcal{CG}(\Xi_{\text{train}}, \epsilon_{\text{tol}}, C_M, K_{\text{damp}}, N, X_N^\mathcal{N})\)

1: if \(N = 0\) then
2:     Randomly select the first sample \(\mu^1 \in \Xi_{\text{train}}\), and set \(N = 1, \ell = 0\) and \(E_\ell = 2\epsilon_{\text{tol}}\);
3:     Obtain truth solution \(u_N^\mathcal{X}(\mu^1)\), and set \(S_1 = \{\mu^1\}, X_1^\mathcal{N} = \text{span} \{u_N^\mathcal{X}(\mu^1)\}\);
4: end if
5: while \((E_\ell > \epsilon_{\text{tol}})\) do
6:     Set \(\ell \leftarrow \ell + 1\);
7:     for each \(\mu \in \Xi_{\text{train}}\) do
8:         Obtain RBM solution \(u_N^\mathcal{X}(\mu) \in X_N^\mathcal{N}\) and error estimate \(\Delta_N(\mu)\);
9:     end for
10: \(\mu^{N+1} = \arg\max_{\mu \in \Xi_{\text{train}}} \Delta_N(\mu), \epsilon = \Delta_N(\mu^{N+1}), E_\ell = \epsilon;\)
11: Augment RB space \(X_{N+1}^\mathcal{N} = X_N^\mathcal{N} \oplus \{u_N^\mathcal{X}(\mu^{N+1})\}\);
12: \(S_{N+1} = S_N \cup \{\mu^{N+1}\}\);
13: Set \(N \leftarrow N + 1\);
14: Construct STS \(\Xi_{\text{Sir}}\) based on \(\{(u_N^\mathcal{X}(\mu), \Delta_N(\mu)) : \mu \in \Xi_{\text{train}}\}\) with SMM;
15: while \((\epsilon > \epsilon_{\text{tol}})\) and \((\frac{\epsilon}{E_\ell} > \frac{1}{K_{\text{damp}}(\ell+1)})\) do
16:     for each \(\mu \in \Xi_{\text{Sir}}\) do
17:         Obtain RBM solution \(u_N^\mathcal{X}(\mu) \in X_N^\mathcal{N}\) and error estimate \(\Delta_N(\mu)\);
18:     end for
19: \(\mu^{N+1} = \arg\max_{\mu \in \Xi_{\text{Sir}}} \Delta_N(\mu), \epsilon = \Delta_N(\mu^{N+1});\)
20: Augment RB space \(X_{N+1}^\mathcal{N} = X_N^\mathcal{N} \oplus \{u_N^\mathcal{X}(\mu^{N+1})\}\);
21: \(S_{N+1} = S_N \cup \{\mu^{N+1}\}\);
22: Set \(N \leftarrow N + 1\);
23: end while
26: end while

Hybrid Training Set Decomposition, Algorithm 5 integrating Algorithms 2 and 4

Given a full training set \(\Xi_{\text{train}}\), first we construct a training set decomposition according to Algorithm 2 to obtain \(\{\Xi_{\text{train},j}\}_{j=1}^J\). Then, we perform Algorithm 4 on each sample set \(\Xi_{\text{train},j}\) sequentially. As verified by our numerical results, this simple hybrid approach tends to reduce the greedy algorithm’s sensitivity to \(n_{\text{tr},\text{small}}\), while retaining the ease of picking \(C_M\) and \(K_{\text{damp}}\) as outlined in section 2.2.3.

Algorithm 5: Hybrid training set decomposition based greedy algorithm \((N, X_N^\mathcal{N}) = \text{HTSD}_\mathcal{CG}(\Xi_{\text{train}}, \epsilon_{\text{tol}}, C_M, K_{\text{damp}}, \{\Xi_{\text{train},j}\}_{j=1}^J)\)

1: Choose \(\mu^1 \in \Xi_{\text{train},1}\) at random. Denote \(S_1 = \{\mu^1\}, X_1^\mathcal{N} = \text{span}\{u_N^\mathcal{X}(\mu^1)\}\) and \(N = 1\);
2: for \(j = 1 : J\) do
3:     Call \((N, X_N^\mathcal{N}) = \text{STS}_\mathcal{CG}(\Xi_{\text{train},j}, \epsilon_{\text{tol}}, C_M, K_{\text{damp}}, N, X_N^\mathcal{N})\)
4: end for

Hybrid Adaptive Enriching, Algorithm 6 assimilating Algorithms 3 and 4

The randomly-generated sample set \(\Xi\) in Algorithm 3, albeit not covering all of \(\Xi_{\text{train}}\), keeps updating itself after each round of greedy searching by replacing parameter values whose \textit{a posteriori} error estimate falls below \(\epsilon_{\text{tol}}\) by those not seen yet. To avoid having too many rounds of replacements, \(\Xi\) must be rich enough to be representative of \(\Xi_{\text{train}}\). However, large \(\Xi\) adversely...
impact the algorithm’s efficiency. The proposed hybrid strikes a balance between efficiency and richness. Toward that end and as outlined in Algorithm 6, we adaptively construct $\Xi_{\text{sur}}$ for the active training set $\Xi$ which is only fully examined when (the much smaller) $\Xi_{\text{sur}}$ is sufficiently resolved. This three-level approach enables a faster examination of $\Xi$ than the Adaptive Enriching algorithm [3] which, in turn, allows $\Xi$ to be larger thus more representative of $\Xi_{\text{train}}$.

Algorithm 6 Hybrid adaptive enriching greedy algorithm $H_{\text{AE-CG}}(\Xi_{\text{train}}, \varepsilon_{\text{tol}}, C_M, K_{\text{damp}}, M_{\text{sample}})$

1. $N_{\text{safe}} = \ceil{(\Xi_{\text{train}})} {M_{\text{sample}}}$;
2. Randomly generate an initial training set $\Xi$ with $M_{\text{sample}}$ parameters from $\Xi_{\text{train}}$;
3. Choose an initial parameter $\mu^1 \in \Xi_{\text{train}}$. Set $S_1 = \{\mu^1\}, X_{1}^{N} = \text{span}(u_{n}^{N}(\mu^1))$ and $N = 1$;
4. Set $\text{safe} = 0, \ell' = 1, \varepsilon = 2\varepsilon_{\text{tol}}$ and $r = n_{\text{train}}$;
5. while $(E_{\ell'} > \varepsilon_{\text{tol}})$ or safe $\leq N_{\text{safe}}$ and $(r > 0)$ do
   6. One-step greedy scan on $\Xi_{\text{train}}$
      7. for each $\mu \in \Xi$ do
         8. Obtain RBM solution $u_{n}^{N}(\mu) \in X_{N}^{N}$ and error estimate $\Delta_{N}(\mu)$;
      end for
   9. $\mu^{N+1} = \arg\max_{\mu \in \Xi} \Delta_{N}(\mu), \varepsilon = \Delta_{N}(\mu^{N+1}), E_{\ell'} = \varepsilon$;
10. Augment RB space $X_{N+1}^{N} = X_{N}^{N} \oplus \{u_{n}^{N}(\mu^{N+1})\}$;
11. $S_{N+1} = S_{N} \cup \{\mu^{N+1}\}$;
12. Set $N \leftarrow N + 1$;
13. Multi-step greedy scan on $\Xi_{\text{sur}}$
14. Construct STS $\Xi_{\text{sur}}$ based on $\{(u_{n}^{N}(\mu), \Delta_{N}(\mu)) : \mu \in \Xi\}$ with SMM;
   16. while $(\varepsilon > \varepsilon_{\text{tol}})$ and $(\frac{1}{E_{\ell'}} > \frac{1}{K_{\text{damp}}(\ell'+1)})$ do
      17. for each $\mu \in \Xi_{\text{sur}}$ do
         18. Obtain RBM solution $u_{n}^{N}(\mu) \in X_{N}^{N}$ and error estimate $\Delta_{N}(\mu)$;
      end for
   19. $\mu^{N+1} = \arg\max_{\mu \in \Xi_{\text{sur}}} \Delta_{N}(\mu), \varepsilon = \Delta_{N}(\mu^{N+1})$;
20. Augment RB space $X_{N+1}^{N} = X_{N}^{N} \oplus \{u_{n}^{N}(\mu^{N+1})\}$;
21. $S_{N+1} = S_{N} \cup \{\mu^{N+1}\}$;
22. Set $N \leftarrow N + 1$;
23. end while
24. end while
25. Truncate $\Xi$ by $\Xi_{<r} = \{\mu \in \Xi : \Delta_{N}(\mu) < \varepsilon_{\text{tol}}\}$;
26. Truncate $\Xi_{\text{train}}$ by $\Xi_{<r}$, and set $r = r - |\Xi_{<r}|$;
27. if $|\Xi_{<r}| = M_{\text{sample}}$ then
   28. Set safe $= \text{safe} + 1$, and $\ell' = 1$;
29. else
   30. $\ell' = \ell' + 1$;
31. end if
32. Randomly choose $M_{\text{sample}} - |\Xi|$ parameters from $\Xi_{\text{train}}$ for addition to $\Xi$;
33. end while

4 | NUMERICAL TESTS

In this section, we test and compare these six greedy algorithms, namely classical greedy, three existing improvements reviewed in section 2 and our newly designed two hybrids in section 3. For the sake of clarity, we list these methods in one place, Table 2.

Two types of examples will be presented to demonstrate the efficiency enhancement of our proposed approaches without sacrificing the quality of the reduced bases. The corresponding results are presented in the subsections below. The CPU times reported in this paper refer to computations performed on a workstation with 3.1 GHz Intel Core i5 processor and 16GB memory,
in the MATLAB environment adopting redaKIT library. Except the multi-paramter case in section 4.2 if an algorithm includes random sampling, we run it 10 times for fixed tuning parameters and present the resulting mean and standard deviation.

4.1 Two-dimensional diffusion problem

We first test them on the following elliptic equation which becomes degenerate at the corners of a rectangular two-dimensional parameter domain.

\[(1 + \mu_1 x)u_{xx} + (1 + \mu_2 y)u_{yy} = e^{4xy} \quad \text{on } \Omega.\]

Here \(\Omega = [-1, 1] \times [-1, 1]\) and we impose homogeneous Dirichlet boundary conditions on \(\partial \Omega\). The truth solver is a spectral Chebyshev collocation method based on \(N_x^2 = 35\) degrees of freedom in each direction, with \(N^2_x = N_y\). The parameter domain \(D\) for \((\mu_1, \mu_2)\) is taken to be \([-0.99, 0.99]^2\). For the \(\Xi_{\text{train}}\) we discretize \(\mathbb{D}\) using a tensorial \(160 \times 160\) Cartesian grid with 160 equi-spaced points in each dimension.

First, we fix \(K_{\text{damp}} = 5\), \(C_M = 10\), \(N_{\text{sample}} = 1024\) and \(n_{\text{tr, small}} = 1024\) and test 6 different tolerances. Table 3 shows numbers of bases at convergence for each method, and its speedup factor in comparison to the classical greedy algorithm. For relative time, defined as the corresponding running time scaled by the running time of classical greedy algorithm, we observe that the worst case scenario of \(\mathbb{H}_{\text{TSD CG}}\) is comparable to, or better than, the best case scenario of any other approach. Moreover, the range of number of bases for \(\mathbb{H}_{\text{TSD CG}}\) is the smallest compared to \(\text{TSD CG}\) or \(\text{AE CG}\). This clearly shows that, thanks to the hybridization, \(\mathbb{H}_{\text{TSD CG}}\) is the least sensitive to the randomness in sampling.

Next, we consider the impact of tweaking the tuning parameters. Toward that end, for each of the three tolerances \(\varepsilon_{\text{tol}} = 10^{-3}, 10^{-4}, 10^{-5}\), we choose \((K_{\text{damp}}, C_M) = (5, 10), (10, 20), (20, 40)\). \(N_{\text{sample}} = 1024, 2048, 4096\) and \(n_{\text{tr, small}} = 1024, 2048, 4096\) and record the relative time and numbers of bases at convergence. The results for \(\text{STS CG, TSD CG, AE CG}\) and \(\mathbb{H}_{\text{TSD CG}}\) are shown in table 4 and 5. For fixed tolerance, it is evident that the performance of \(\mathbb{H}_{\text{TSD CG}}\) is very stable for various selection of tuning parameters both in time and the size of RB space. On the other hand, \(\text{TSD CG}\) and \(\text{AE CG}\) appear to show non-negligible variations between the runs for different tuning parameter settings. Note that \(\text{STS CG}\) is deterministic. In comparison to its two components \(\text{STS CG}\) and \(\text{TSD CG}\), our new hybrid approach \(\mathbb{H}_{\text{TSD CG}}\) clearly improves both efficiency and stability. This reduction in sensitivity makes the hybrid particularly attractive.

4.2 Helmholtz equation on a parametrized domain

We consider the propagation of pressure wave \(P(x, t)\) into the acoustic horn illustrated in Figure 1, the same example considered in [25]. Under the assumption of waves being time harmonic, the acoustic pressure \(P\) can be separated as \(P(x, t) = \Re(p(x) \exp^{i\omega t})\) where the complex amplitude \(p(x)\) satisfies the following Helmholtz equation:

\[
p_{xx} + \frac{\omega^2}{c^2} p = 0.
\]
(a) Efficiency in terms of relative computation time

| $\epsilon_{\text{tol}}$ | CG   | STS_CG | TSD_CG | AE_CG | $H_{\text{TSD}}$ | $H_{\text{AE}}$ |
|------------------------|------|--------|--------|-------|-----------------|----------------|
| $10^{-3}$              | 1    | 0.258  | [0.248,0.27] | [0.345,0.411] | [0.216,0.234] | [0.282,0.333] |
| $5 \times 10^{-4}$     | 1    | 0.228  | [0.218,0.232] | [0.34,0.368]  | [0.19,0.208]  | [0.239,0.354] |
| $10^{-4}$              | 1    | 0.222  | [0.25,0.262]  | [0.222,0.258] | [0.174,0.182] | [0.162,0.205] |
| $5 \times 10^{-5}$     | 1    | 0.173  | [0.193,0.207] | [0.28,0.348]  | [0.154,0.164] | [0.214,0.295] |
| $10^{-5}$              | 1    | 0.159  | [0.164,0.176] | [0.254,0.294] | [0.131,0.141] | [0.188,0.225] |
| $5 \times 10^{-6}$     | 1    | 0.154  | [0.136,0.244] | [0.195,0.343] | [0.117,0.151] | [0.143,0.250] |

(b) Number of bases at convergence

| $\epsilon_{\text{tol}}$ | CG   | STS_CG | TSD_CG | AE_CG | $H_{\text{TSD}}$ | $H_{\text{AE}}$ |
|------------------------|------|--------|--------|-------|-----------------|----------------|
| $10^{-3}$              | 94   | 96     | [98,103] | [93,102] | [98,101]        | [99,106]       |
| $5 \times 10^{-4}$     | 105  | 108    | [111,113] | [106,111] | [110,114]       | [110,119]      |
| $10^{-4}$              | 132  | 134    | [144,146] | [135,142] | [141,144]       | [143,152]      |
| $5 \times 10^{-5}$     | 146  | 150    | [158,162] | [147,156] | [156,158]       | [155,168]      |
| $10^{-5}$              | 178  | 183    | [195,200] | [182,192] | [191,194]       | [190,207]      |
| $5 \times 10^{-6}$     | 193  | 202    | [198,270] | [186,247] | [210,237]       | [192,256]      |

**TABLE 3** Comparison of efficiency and size of the resulting reduced basis space for the diffusion problem. Shown for non-deterministic methods are their 95% confidence intervals.

**FIGURE 1** Left: the acoustic horn domain and boundaries (background coloring given by $\Re(p)$ for $f = 900$Hz). Right: RBF control points (red circles) whose vertical displacement is treated as a parameter. [Figure taken from 28].

\[
\Delta p + \kappa^2 p = 0 \quad \text{in} \quad \Omega \\
(i\kappa + \frac{1}{2R})p + \nabla p \cdot \mathbf{n} = 0 \quad \text{on} \quad \Gamma_o \\
i\kappa p + \nabla p \cdot \mathbf{n} = 2i\kappa A \quad \text{on} \quad \Gamma_i \\
\nabla p \cdot \mathbf{n} = 0 \quad \text{on} \quad \Gamma_h \cup \Gamma_s = \Gamma_n,
\]

where $\kappa = \frac{w}{c}$ is the wave number, $w = 2\pi f$ the angular frequency and $c = 340\text{ cm/s}$ the speed of sound. A radiation condition is prescribed on the boundary $\Gamma_i$ imposing an inner-going wave with amplitude $A = 1$ and absorbing the outer-going planar waves. A Neumann boundary condition is applied on the walls $\Gamma_h$ of the device as well as on the symmetry boundary $\Gamma_s$. Finally, an absorbing condition is placed on the far-field boundary $\Gamma_o$ with radius $R = 1$. 
We consider up to five parameters. The first is the frequency $f$. The other four describe the shape of the horn, representing the vertical displacement of the RBF control points in Figure 1. As a result, we have the five-dimensional parameter vector $\mu = [f, \mu_x]$. The output of interest is the index of reflection intensity (IRI) defined as

$$J(\mu) = \left| \frac{1}{\Gamma_1} \int_{\Gamma_1} p(\mu) d\Gamma - 1 \right|,$$

which measures the transmission efficiency of the device. More details about the construction of this problem can be found in [28].

For this problem, TSD\_CG and its corresponding hybrid H\_TSD\_CG fail. As a result, we report the comparison of CG, STS\_CG, AE\_CG and H\_AE\_CG.

### TABLE 4 Tuning parameter study for the diffusion problem. Comparison of efficiency in terms of relative computation time.

The number before the ± sign is the mean value while that after the ± sign is twice the standard deviation.

| $\varepsilon_{\text{tol}}/(K_{\text{damp}}, C_M)$ | STS\_CG | TSD\_CG | AE\_CG | H\_TSD\_CG |
|----------------------------------|---------|---------|---------|------------|
| $10^{-3}$                         | 0.258   | 0.243   | 0.232   | 0.225 ± 0.009 |
| $10^{-4}$                         | 0.222   | 0.193   | 0.191   | 0.222 ± 0.004 |
| $10^{-5}$                         | 0.159   | 0.148   | 0.147   | 0.17 ± 0.006 |

| $\varepsilon_{\text{tol}}/n_{\text{tr,small}}$ | STS\_CG | TSD\_CG | AE\_CG |
|-----------------------------------------------|---------|---------|---------|
| $10^{-3}$                                     | 0.259 ± 0.011 | 0.258 ± 0.007 | 0.312 ± 0.006 |
| $10^{-4}$                                     | 0.256 ± 0.006 | 0.235 ± 0.004 | 0.285 ± 0.007 |
| $10^{-5}$                                     | 0.17 ± 0.006 | 0.184 ± 0.006 | 0.24 ± 0.003 |

| $\varepsilon_{\text{tol}}/M_{\text{sample}}$ | STS\_CG | TSD\_CG | AE\_CG |
|-----------------------------------------------|---------|---------|---------|
| $10^{-3}$                                     | 0.378 ± 0.034 | 0.412 ± 0.04 | 0.429 ± 0.056 |
| $10^{-4}$                                     | 0.24 ± 0.018 | 0.262 ± 0.038 | 0.278 ± 0.046 |
| $10^{-5}$                                     | 0.274 ± 0.02 | 0.303 ± 0.018 | 0.343 ± 0.016 |

| $\varepsilon_{\text{tol}}/(K_{\text{damp}}, C_M)/n_{\text{tr,small}}$ | STS\_CG | TSD\_CG | AE\_CG |
|---------------------------------------------------------------------|---------|---------|---------|
| $(5, 10)$                                                           | 0.225 ± 0.009 | 0.221 ± 0.009 | 0.228 ± 0.006 |
| $(10, 20)$                                                          | 0.222 ± 0.004 | 0.221 ± 0.01 | 0.229 ± 0.009 |
| $(20, 40)$                                                          | 0.178 ± 0.004 | 0.185 ± 0.006 | 0.190 ± 0.009 |
| $(5, 10)$                                                           | 0.136 ± 0.005 | 0.139 ± 0.006 | 0.138 ± 0.004 |
| $(10, 20)$                                                          | 0.136 ± 0.004 | 0.141 ± 0.004 | 0.134 ± 0.006 |
| $(20, 40)$                                                          | 0.139 ± 0.005 | 0.142 ± 0.006 | 0.132 ± 0.007 |
(a) $\text{STS}_\text{CG}$

| $\varepsilon_{\text{tol}}/(K_{\text{damp}}, C_M)$ | (5, 10) | (10, 20) | (20, 40) |
|---|---|---|---|
| $10^{-3}$ | 96 | 95 | 96 |
| $10^{-4}$ | 134 | 136 | 136 |
| $10^{-5}$ | 183 | 181 | 183 |

(b) $\text{TSD}_\text{CG}$

| $\varepsilon_{\text{tol}}/n_{\text{tr,small}}$ | 1024 | 2048 | 4096 |
|---|---|---|---|
| $10^{-3}$ | 100.6 ± 2.2 | 97.8 ± 0.98 | 96.3 ± 0.9 |
| $10^{-4}$ | 145.2 ± 1.0 | 140.9 ± 1.58 | 96.3 ± 0.9 |
| $10^{-5}$ | 197.5 ± 2.7 | 189 ± 1.8 | 186.5 ± 1.6 |

(c) $\text{AE}_\text{CG}$

| $\varepsilon_{\text{tol}}/M_{\text{sample}}$ | 1024 | 2048 | 4096 |
|---|---|---|---|
| $10^{-3}$ | 97.4 ± 4.2 | 94.7 ± 2.2 | 93.4 ± 2.2 |
| $10^{-4}$ | 138.4 ± 3.2 | 135.3 ± 4.2 | 133.5 ± 2.4 |
| $10^{-5}$ | 187.1 ± 4.8 | 182.9 ± 2.8 | 181.8 ± 3 |

(d) $\text{H}_\text{TSD}_\text{CG}$

| $\varepsilon_{\text{tol}}$ | $(K_{\text{damp}}, C_M)/n_{\text{tr,small}}$ | 1024 | 2048 | 4096 |
|---|---|---|---|---|
| $10^{-3}$ | (5, 10) | 99.7 ± 1.5 | 98.5 ± 1.5 | 98.9 ± 1.3 |
| (10, 20) | 100 ± 1.3 | 98.7 ± 1.8 | 99.5 ± 1.2 |
| (20, 40) | 99.9 ± 1.8 | 98.6 ± 2 | 99.7 ± 1.2 |
| $10^{-4}$ | (5, 10) | 142.1 ± 1.6 | 140.8 ± 1.7 | 141.9 ± 2.1 |
| (10, 20) | 143.2 ± 2.2 | 141.6 ± 1.4 | 142 ± 2.1 |
| (20, 40) | 142.3 ± 1.5 | 140.3 ± 1.6 | 142.2 ± 2.6 |
| $10^{-5}$ | (5, 10) | 192.4 ± 1.8 | 192.7 ± 5.8 | 189.6 ± 2.7 |
| (10, 20) | 192.2 ± 1.9 | 191.3 ± 3.4 | 189 ± 2.3 |
| (20, 40) | 193.4 ± 2.5 | 193.2 ± 3.4 | 189.9 ± 4.1 |

**TABLE 5** Tuning parameter study for the diffusion problem. Comparison of number of bases at convergence. The number before the ± sign is the mean value, while that after the ± sign is twice the standard deviation.

### 4.2.1 One parameter case (frequency)

For the first case, we keep the geometrical parameter $\mu_s$ fixed to the reference configuration, and let the frequency $f = \mu_1$ vary in the range $D = [10, 1800]$. $\Xi_{\text{train}}$ is obtained from random sampling in $D$. Since the shape parametrization is not considered here, the problem exhibits a trivial affine decomposition. The computational details are shown in Table 6. The numbers of bases at convergence for each algorithm are shown in Table 8, which showed that the RB spaces constructed by these three enhanced greedy algorithms match that generated by the classical algorithm $\text{CG}$. Therefore, $\text{H}_\text{AE}_\text{CG}$ speeds up the greedy search process without degrading the quality of RB space.
### Multi-level Greedy Algorithms for RBM

#### Table 6: Experiment setup for the acoustic horn.

| Variable | Value (one parameter) | Value (two parameters) | Value (five parameters) |
|----------|-----------------------|------------------------|-------------------------|
| Size of $\xi_{\text{train}}$ | 50000 | 10000 | 10000 |
| Number of vertices | 4567 | 4567 | 4567 |
| Number of elements | 8740 | 8740 | 8740 |
| Number of nodes | 4567 | 4567 | 4567 |
| Number of finite element dofs | 4567 | 4567 | 4567 |
| Interpolation procedure in (8) | DEIM | MDEIM | MDEIM |
| Interpolation points used in (8) | 20 | 30 | 250 |
| $Q_a$ in (8) | 4 | 3 | 95 |
| $Q_f$ in (8) | 1 | 1 | 4 |
| $(K_{\text{damp}}, C_M, M_{\text{sample}})$ | Various | (5, 10, 2048) | (5, 10, 2048) |
| Tolerance $\epsilon_{\text{tol}}$ | $10^{-6}$ | various | various |

#### Table 7: Tuning parameter study for the acoustic horn: efficiency in terms of relative computation time. The number before the $\pm$ sign is the mean value. The number after the $\pm$ sign is twice the standard deviation.

(a) $\text{STS}_\text{CG}$

| $(K_{\text{damp}}, C_M)$ | (5, 10) | (10, 20) | (20, 40) |
|--------------------------|----------|----------|----------|
|                          | 0.225    | 0.206    | 0.214    |

(b) $\text{AE}_\text{CG}$

| $M_{\text{sample}}$ | 1024 | 2048 | 4096 |
|---------------------|------|------|------|
|                     | 0.162 ± 0.013 | 0.198 ± 0.011 | 0.262 ± 0.028 |

(c) $\text{H}_{\text{AE}_\text{CG}}$

| $(K_{\text{damp}}, C_M) / M_{\text{sample}}$ | 1024 | 2048 | 4096 |
|---------------------------------------------|------|------|------|
| (5, 10) | 0.145 ± 0.014 | 0.149 ± 0.013 | 0.169 ± 0.019 |
| (10, 20) | 0.140 ± 0.012 | 0.144 ± 0.019 | 0.162 ± 0.012 |
| (20, 40) | 0.149 ± 0.013 | 0.152 ± 0.018 | 0.158 ± 0.012 |

#### 4.2.2 Multi-parameter case (frequency plus RBF control points)

Here, we include the parameters delineating the geometrical configurations of the acoustic horn, namely the vertical displacements of the RBF control points in Figure [1]. The problem becomes nonaffine, so MDEIM is employed, and we report one single run for each method.

**Two parameters case.** We consider frequency and the vertical displacement of the right-most RBF control point in Figure [1]. The parameter domain is given by $D = [50, 1000] \times [-0.03, 0.03]$ and $\Xi_{\text{train}}$ is generated by operating Latin hypercube sampling in $D$. Other computational parameters are listed in Table [6] third column.

Figure [2] demonstrates the performance of $\text{STS}_\text{CG}$, $\text{AE}_\text{CG}$ and $\text{H}_{\text{AE}_\text{CG}}$ for five different tolerances $\epsilon_{\text{tol}} = 10^{-3}, 5 \times 10^{-4}, 10^{-4}, 5 \times 10^{-5}, 10^{-5}$. For $\text{AE}_\text{CG}$ and $\text{STS}_\text{CG}$, smaller tolerance leads to more runtime. This is much less severe for $\text{H}_{\text{AE}_\text{CG}}$. This observation suggests that $\text{H}_{\text{AE}_\text{CG}}$ has great potential to handle larger-scale problems. In addition, the sizes
TABLE 8 Tuning parameter study for the acoustic horn: number of bases. The number before the ± sign is the mean value. The number after the ± sign is twice the standard deviation.

| (K_damp, C_M)         | (5, 10)  | (10, 20) | (20, 40) |
|-----------------------|----------|----------|----------|
| Easing                | 58       | 60       | 58       |

(b) AE_CG

| M_sample | 1024 | 2048 | 4096 |
|----------|------|------|------|
| Time(s)  | 59.2 ±1.3 | 59.2 ± 2.3 | 59.5 ± 1.1 |

(d) H_AE_CG

| (K_damp, C_M) / M_sample | 1024 | 2048 | 4096 |
|---------------------------|------|------|------|
| (5, 10)                   | 60.3 ± 2.3 | 60.3 ± 2.9 | 59.7 ± 1.5 |
| (10, 20)                  | 60.5 ± 1.3 | 59.8 ± 1.8 | 59.8 ± 2.5 |
| (20, 40)                  | 59.9 ± 2.5 | 59.8 ± 1.8 | 60 ± 1.9 |

FIGURE 2 Two parameter case for the acoustic horn with different tolerance: computational runtime (left) and number of bases (right).

of these three RB spaces are very similar, indicating that H_AE_CG does not appear to suffer from online efficiency degradation for this example.

Five parameters case. We now let all five parameters \( \{ f, \mu, \gamma \} \) vary. The parameter domain is given by \( D = [50, 1000] \times D_\varepsilon \), where \( D_\varepsilon = [-0.03, 0.03]^4 \). The full training set \( \mathcal{X}_{\text{train}} \) is also from Latin hypercube sampling in \( D \). The corresponding computational settings are detailed in Table 6 column 4.

This five parameter case is highly non-affine and much more complicated than the previous tests. Due to the computation of CG being much more demanding for this case, we test 3 cases: \( \varepsilon_{\text{tol}} = 10^{-3}, 5 \times 10^{-4}, 10^{-4} \) to verify the performance of STS_CG, AE_CG and H_AE_CG. The results are shown in Table 9 which demonstrate that the new hybrid approach H_AE_CG is the most efficient.
(a) Runtime(s)

| Methods/ε_{tol} | 10^{-3} | 5 × 10^{-4} | 10^{-4} |
|-----------------|---------|-------------|---------|
| STS_CG          | 33849   | 42903       | 78998   |
| AE_CG           | 33779   | 47473       | 114030  |
| H AE_CG         | 27522   | 37466       | 51003   |

(a) Number of bases

| Methods/ε_{tol} | 10^{-3} | 5 × 10^{-4} | 10^{-4} |
|-----------------|---------|-------------|---------|
| STS_CG          | 109     | 130         | 193     |
| AE_CG           | 105     | 127         | 187     |
| H AE_CG         | 116     | 135         | 199     |

TABLE 9 Results for the five-parameter acoustic horn with different tolerances.

5 | CONCLUSIONS

In this paper, we review three recent offline enhancement approaches for the reduced basis method which share the overarching theme of constructing a small-size subset of the full training set and then perform the classical greedy algorithm on it. In addition, we propose two new hybrid approaches in the same vein. It provides a multi-level framework that can be further generalized or integrated with new approaches. Through extensive numerical tests, we show that the new hybrids are particularly well-suited for large-scale or high dimensional parameterized problems thanks to their robustness, ease of setting tuning parameters, and significant speedup over the classical or improved greedy algorithms.

ACKNOWLEDGMENTS

The authors want to express their sincere gratitude to Prof. Anthony Patera of MIT for suggesting to test the algorithms on the Helmholtz equation. The second author was partially supported by National Science Foundation grant DMS-1719698.

References

1. Quarteroni A, Manzoni A, Negri F. Reduced basis methods for partial differential equations: An introduction. 92. Springer . 2015.
2. Hesthaven JS, Rozza G, Stamm B. Certified reduced basis methods for parametrized partial differential equations. SpringerBriefs in MathematicsSpringer, Cham; BCAM Basque Center for Applied Mathematics, Bilbao . 2016. BCAM SpringerBriefs
3. Rozza G, Huynh DBP, Patera AT. Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations. Arch. Comput. Methods Eng. 2008; 15(3): 229-275.
4. Haasdonk B. Chapter 2: Reduced Basis Methods for Parametrized PDEs - A Tutorial Introduction for Stationary and Instationary Problems: 65-136;
5. Prud’homme C, Rovas D, Veroy K, Maday Y, Patera AT, Turinici G. Reliable real-time solution of parametrized partial differential equations: Reduced-basis output bound methods. Journal of Fluids Engineering 2002; 124(1): 70–80.
6. Nguyen NC, Veroy K, Patera AT. Certified Real-Time Solution of Parametrized Partial Differential Equations. In: Yip S., ed. Handbook of Materials ModelingSpringer Netherlands. 2005 (pp. 1529–1564).
7. Benner P, Gugercin S, Willcox K. A survey of projection-based model reduction methods for parametric dynamical systems. *SIAM review* 2015; 57(4): 483–531.

8. Carlberg K, Bou-Mosleh C, Farhat C. Efficient non-linear model reduction via a least-squares Petrov–Galerkin projection and compressive tensor approximations. *International Journal for Numerical Methods in Engineering* 2011; 86(2): 155-181. [doi: 10.1002/nme.3050](https://doi.org/10.1002/nme.3050)

9. Carlberg K, Barone M, Antil H. Galerkin v. least-squares Petrov–Galerkin projection in nonlinear model reduction. *Journal of Computational Physics* 2017; 330: 693 - 734. [doi: https://doi.org/10.1016/j.jcp.2016.10.033](https://doi.org/10.1016/j.jcp.2016.10.033)

10. Chen Y, Gottlieb S. Reduced Collocation Methods: Reduced Basis Methods in the Collocation Framework. *J. Sci. Comput.* 2013; 55(3): 718–737.

11. Chen Y, Gottlieb S, Ji L, Maday Y, Xu Z. L1-ROC and R2-ROC: L1- and R2-based Reduced Over-Collocation methods for parametrized nonlinear partial differential equations. *arXiv:1906.07349* 2019.

12. Sen S. Reduced-basis approximation and a posteriori error estimation for many-parameter heat conduction problems. *Numerical Heat Transfer, Part B: Fundamentals* 2008; 54(5): 369–389.

13. Hesthaven JS, Stamm B, Zhang S. Efficient greedy algorithms for high-dimensional parameter spaces with applications to empirical interpolation and reduced basis methods. *ESAIM: Mathematical Modelling and Numerical Analysis* 2014; 48(1): 259–283.

14. Jiang J, Chen Y, Narayan A. Offline-enhanced reduced basis method through adaptive construction of the surrogate training set. *Journal of Scientific Computing* 2017; 73(2-3): 853–875.

15. Haasdonk B, Dihlmann M, Ohlberger M. A training set and multiple bases generation approach for parameterized model reduction based on adaptive grids in parameter space. *Mathematical and Computer Modelling of Dynamical Systems* 2011; 17(4): 423–442.

16. Iapichino L, Quarteroni A, Rozza G. Reduced basis method and domain decomposition for elliptic problems in networks and complex parametrized geometries. *Computers & Mathematics with Applications* 2016; 71(1): 408–430.

17. Urban K, Volkwein S, Zeeb O. Greedy sampling using nonlinear optimization. In: Springer. 2014 (pp. 137–157).

18. Ohlberger M, Schindler F. Error control for the localized reduced basis multiscale method with adaptive on-line enrichment. *SIAM Journal on Scientific Computing* 2015; 37(6): A2865–A2895.

19. Kaulmann S, Flemisch B, Haasdonk B, Lie KA, Ohlberger M. The localized reduced basis multiscale method for two-phase flows in porous media. *International Journal for Numerical Methods in Engineering* 2015; 102(5): 1018–1040.

20. Ohlberger M, Rave S, Schindler F. True error control for the localized reduced basis method for parabolic problems. In: Springer. 2017 (pp. 169–182).

21. Zou Z, Kouri D, Aquino W. An adaptive local reduced basis method for solving PDEs with uncertain inputs and evaluating risk. *Computer Methods in Applied Mechanics and Engineering* 2019; 345: 302–322.

22. Barrault M, Nguyen NC, Maday Y, Patera AT. An “empirical interpolation” method: Application to efficient reduced-basis discretization of partial differential equations. *C. R. Acad. Sci. Paris, Série I* 2004; 339: 667–672.

23. Grepl MA, Maday Y, Nguyen NC, Patera AT. Efficient reduced-basis treatment of nonaffine and nonlinear partial differential equations. *Mathematical Modelling and Numerical Analysis* 2007; 41(3): 575–605.

24. Chaturantabut S, Sorensen DC. Nonlinear model reduction via discrete empirical interpolation. *SIAM J. Sci. Comput.* 2010; 32(5): 2737–2764. [doi: 10.1137/090766498](https://doi.org/10.1137/090766498)

25. Elman HC, Liao Q. Reduced basis collocation methods for partial differential equations with random coefficients. *SIAM/ASA Journal on Uncertainty Quantification* 2013; 1(1): 192–217.
26. Liu Y, Chen T, Chen Y, Shu CW. Certified Offline-Free Reduced Basis (COFRB) Methods for Stochastic Differential Equations Driven by Arbitrary Types of Noise. *Journal of Scientific Computing* 2019. doi: 10.1007/s10915-019-00976-5

27. Negri F. redbKIT. [http://redbkit.github.io/redbKIT/](http://redbkit.github.io/redbKIT/) 2016. Version 2.2.

28. Negri F, Manzoni A, Amsallem D. Efficient model reduction of parametrized systems by matrix discrete empirical interpolation. *Journal of Computational Physics* 2015; 303: 431–454.

29. Bángtsson E, Noreland D, Berggren M. Shape optimization of an acoustic horn. *Computer methods in applied mechanics and engineering* 2003; 192(11-12): 1533–1571.

30. Kasolis F, Wadbro E, Berggren M. Fixed-mesh curvature-parameterized shape optimization of an acoustic horn. *Structural and Multidisciplinary Optimization* 2012; 46(5): 727–738.

31. Buhmann MD. *Radial basis functions: theory and implementations*. 12. Cambridge university press. 2003.