REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
09/17/2019	Final Technical Report	06-01-2016 to 05-30-2019

4. TITLE AND SUBTITLE Data Assimilation and Parameter Estimation for Parametric Partial Differential Equations	**5a. CONTRACT NUMBER**
	5b. GRANT NUMBER N00014-16-1-2706
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Ronald DeVore	**5d. PROJECT NUMBER** 1000004430
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Texas A&M University 400 Harvey Mitchell Pkwy S. Suite 300 College Station, Texas 77845	**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Sponsored Research Services 400 Harvey Mitchell Pkwy S. Suite 300 College Station, Texas 77845	**10. SPONSOR/MONITOR'S ACRONYM(S)**
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Parametric partial differential equations (pdes) are used to model complex physical and biological systems and arise in optimal control and design. Their solution varies with the parameters in a complex way, especially when there are a large number of parameters. This project studied novel ways to understand the effect of changing parameters by using a technique called model reduction which isolates the important parameters. Various methods for model reduction were studied and evaluated for performance This included representation by high dimensional polynomials and interpolation of certain judiciously chosen parameter snapshots. A new class of algorithms for model reduction based on nonlinear approximation were introduced and evaluated for performance.
The project also studied the best way to incorporate data observations of the solution to improve efficiency. Certain algorithms for data assimilation were proven to be optimal. Several new methods were introduced to speed up computation. These included extracting random snapshots and employing various approaches to optimization. The project also studied the problem of how well the parameters can be determined when as observation of the state is given. Sufficient conditions on the coefficients of the pde were proved to ensure that the parameters are uniquely determined by the state. This was then employed to build algorithms for parameter estimation with certified error bounds.

15. SUBJECT TERMS
Parametric partial differentail equations, complex physical/biological systems, optimal computational algorithms

16. SECURITY CLASSIFICATION OF:
Unclassified

a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
			None		(979) 845-3261
					19b. TELEPHONE NUMBER *(Include area code)*

INSTRUCTIONS FOR COMPLETING SF 298

1. REPORT DATE. Full publication date, including day, month, if available. Must cite at least the year and be Year 2000 compliant, e.g. 30-06-1998; xx-06-1998; xx-xx-1998.

2. REPORT TYPE. State the type of report, such as final, technical, interim, memorandum, master's thesis, progress, quarterly, research, special, group study, etc.

3. DATE COVERED. Indicate the time during which the work was performed and the report was written, e.g., Jun 1997 - Jun 1998; 1-10 Jun 1996; May - Nov 1998; Nov 1998.

4. TITLE. Enter title and subtitle with volume number and part number, if applicable. On classified documents, enter the title classification in parentheses.

5a. CONTRACT NUMBER. Enter all contract numbers as they appear in the report, e.g. F33315-86-C-5169.

5b. GRANT NUMBER. Enter all grant numbers as they appear in the report. e.g. AFOSR-82-1234.

5c. PROGRAM ELEMENT NUMBER. Enter all program element numbers as they appear in the report, e.g. 61101A.

5e. TASK NUMBER. Enter all task numbers as they appear in the report, e.g. 05; RF0330201; T4112.

5f. WORK UNIT NUMBER. Enter all work unit numbers as they appear in the report, e.g. 001; AFAPL30480105.

6. AUTHOR(S). Enter name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. The form of entry is the last name, first name, middle initial, and additional qualifiers separated by commas, e.g. Smith, Richard, J, Jr.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES). Self-explanatory.

8. PERFORMING ORGANIZATION REPORT NUMBER. Enter all unique alphanumeric report numbers assigned by the performing organization, e.g. BRL-1234; AFWL-TR-85-4017-Vol-21-PT-2.

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES). Enter the name and address of the organization(s) financially responsible for and monitoring the work.

10. SPONSOR/MONITOR'S ACRONYM(S). Enter, if available, e.g. BRL, ARDEC, NADC.

11. SPONSOR/MONITOR'S REPORT NUMBER(S). Enter report number as assigned by the sponsoring/ monitoring agency, if available, e.g. BRL-TR-829; -215.

12. DISTRIBUTION/AVAILABILITY STATEMENT. Use agency-mandated availability statements to indicate the public availability or distribution limitations of the report. If additional limitations/ restrictions or special markings are indicated, follow agency authorization procedures, e.g. RD/FRD, PROPIN, ITAR, etc. Include copyright information.

13. SUPPLEMENTARY NOTES. Enter information not included elsewhere such as: prepared in cooperation with; translation of; report supersedes; old edition number, etc.

14. ABSTRACT. A brief (approximately 200 words) factual summary of the most significant information.

15. SUBJECT TERMS. Key words or phrases identifying major concepts in the report.

16. SECURITY CLASSIFICATION. Enter security classification in accordance with security classification regulations, e.g. U, C, S, etc. If this form contains classified information, stamp classification level on the top and bottom of this page.

17. LIMITATION OF ABSTRACT. This block must be completed to assign a distribution limitation to the abstract. Enter UU (Unclassified Unlimited) or SAR (Same as Report). An entry in this block is necessary if the abstract is to be limited.

Final Report: N00014-16-1-2706
Data Assimilation and Parameter Estimation for Parametric Partial Differential Equations
PI: Ronald DeVore

September 17, 2019

This project studied numerical algorithms for solving partial differential equations (PDEs) which depend on many variables or parameters. Such PDEs arise in several important areas of applied mathematics including the modeling of complex physical and biological systems as well as in optimal design in engineering. In describing the accomplishments of this proposal, we let $u(y) = u(y, x)$ be the solution to the PDE where y is the vector of parameters taken from a parameter domain Y and x is the physical variable taken from a domain D. We denote by V the energy space associated with the PDE. The solution $u(y)$ is in V for all y and is to be computed with respect to the norm of V. In most situations V is a Hilbert space such as $H_0^1(D)$ in the case of elliptic problems.

The project had three main goals as described in the next three sections of this report.

1 Build fast forward solvers through methods of model reduction.

Given a parameter query y, the solver should efficiently compute an accurate approximation to $u(y)$ on line. Such forward solvers are used to simulate the process and are also used in optimal control and design. They are built off-line utilizing model reduction based on the smoothness of the mapping $y \to u(y)$. The paper **1** gave a state of the art survey of model reduction and numerical methods for parametric PDEs. The main theme of that paper was to show that the solution map for elliptic (and certain more general) parametric PDEs has analytic and anisotropic properties sufficient to guarantee efficient model reduction.

The model reduction is typically built by finding a low dimensional linear space X_n in V such that $\operatorname{dist}(u(y), X_n) \le \varepsilon$ for all parameters $y \in Y$ where ε is a user prescribed accuracy. The space X_n is typically built in one of two ways. The first is through a sparse polynomial $P(y)$ taking values in V which approximates $u(y)$ to accuracy ε in the norm of V. The span of the coefficients of P then provide the linear space X_n. We have given in our earlier work [2, 3] theorems that prove the existence of such polynomials with quantitative estimates on how large n has to be to satisfy the prescribed error accuracy ε. In **6**, we extend this theory to include the log normal assumption on diffusion coefficients that are often used in stochastic settings. In **13**, we improve on the guarantees for rates of convergence (approximation error versus the size of n) by exploiting connections with results in number theory, in particular additive and multiplicative partitioning of integers.

A second construction of reduced models is through greedy algorithms which adaptively select snapshots $u(y^{(j)})$, $j = 1, \dots, n$, whose span provides the space X_n. We have given in our earlier work ([1, 4]) the most rigorous and far reaching results for model reduction via greedy algorithms and in particular we show that these algorithms are near optimal for selecting a space X_n. The intimate connections between greedy algorithms and foundational quantities such as widths and entropy is fully described in **5**. In **2**, we prove a general principal that widths are almost preserved under analytic maps. This fact has been exploited to prove optimality of greedy algorithms in many other settings.

The implementation of the greedy algorithm is done off-line and is expensive since it requires the computation of a fine net in the parameter domain Y. The size of the net typically grows exponentially with the dimension n of the reduced space. This makes its off-line implementation impossible when n is large. To circumvent this difficulty, we have studied two new approaches.

The first approach was to replace the discretization of Y by a random selection of points rather than a deterministic one. We have shown in **11** that such random selections can reduce the off-line discretization from exponential in n to polynomial in n. However, the performance guarantees are now with high probability rather than certainty. Nevertheless, the numerical evidence is overwhelming that this leads to tremendous speed up of off-line constructions. In a related work **12**, we give an algorithm for finding best reduced model affine spaces through convex optimization.

The second approach to off-line speed up has been to investigate whether nonlinear methods can be used in model reduction. The main theme of our work was to investigate the advantages of replacing the linear space X_n by a low dimensional manifold $\mathcal{M}_n$ for the approximation of $u(y)$. We have approached this problem along two tracks. The first was to examine how to approximate $u(y)$ locally by very low dimension polynomials. That is, we want to approximate u by a polynomial not on all of Y but only on a local hypercube - in this way we can keep the degree of the polynomial small. The analysis for this approximation is given in **14**, where it gives a bound on the number of hyperrectangles needed in a partition when the number of terms $m << n$ of the local reduced space is specified along with the desired error tolerance ε.

2 Data Assimilation

In some settings, one is not given a parameter and asked to compute the solution u but rather one observes data about u for an unknown parameter and one wishes to fully recover u. This type of state estimation occurs for example when taking core samples for fluid simulation through porous media. In such settings, one wants to merge the information that u is a solution to the parametric PDE (but with unknown parameters) together with the observed data in order to approximate u. Such problems are loosely referred to as data assimilation. In **4**, we show that a certain least squares procedure proposed in [5] is essentially optimal for recovery of u from the measurements and that the optimal performance can be exactly computed. We have extended these results for data assimilation from Hilbert spaces to general Banach spaces in **7** with the solution manifold of the parametric PDE replaced by a general compact set $\mathcal{K}$. We provide in that paper, optimal algorithms for data assimilation and a priori determination of the performance of these optimal algorithms.

In some cases, one is not interested in a full approximation to the target function u but rather only in evaluating a quantity of interest $Q(u)$. Typically, Q is a linear functional such as a weighted average of u. We have studied this problem in **9** where we establish optimal algorithms for evaluation $Q(u)$ and establish performance bounds for these algorithms. We note that these algorithms exhibit a reduction in computational cost by having only to compute $Q(u)$ rather than give a full approximation to u and then evaluate $Q(u)$.

3 Parameter estimation

Another important problem occurs when one observes the solution to the parametric PDE and wishes to determine the parameters associated to this solution This is a classical inverse problem that occurs in various contexts. The first challenge is to determine whether the state u uniquely determines the parameter. For elliptic parametric PDEs in one physical variable, we prove this is the case whenever the right side of the PDE does not change sign. In higher physical space dimension d, we have not settled this inverse problem but we have given in **8** state of the art results under additional assumptions on the smoothness of the state in the form of a very weak Besov regularity assumption.

We are currently investigating how one can combine our algorithms for data assimilation with the inverse parameter estimation results to obtain parameter estimates when only data about the state is known, i.e., when we do not have full knowledge of the state. We have pursued two directions which we have not yet reported on since we are still evaluating their numerical implementation. The first method is a divide and conquer method based on discretizing the parameter domain in pursuit of quantifying which parameters may give rise to the data. While these discretizations are in common use they are expensive when the number of parameters is large. We are able to restrict the search by using our inverse theorems on parameter estimation. A second approach is to view the parameter search as a constrained optimization problem applied to the residual. The celebrated Nesterov algorithm gives some certified convergence rates for such constrained optimization using gradient descent. The guaranteed convergence rates are slow but we are investigating whether they can be improved in the particular context of parameter estimation by exploiting properties of the quadratic minimization that are particular to the parametric PDE setting.

4 Related topics

Finally, we mention some topics that we have investigated that are related to the three major goals delineated above. In **10**, we provide greedy strategies for convex optimization in high dimensions and give certifiable bounds on their convergence rate. Such optimization occurs in many contexts, for example in parameter estimation, and more generally in learning problems such as deep learning. In **3**, we provide a simplified proof of the fact that the Restricted Isometry Property is sufficient to guarantee that Orthogonal Matching Pursuit algorithms for dictionary approximation perform as well as best n-term approximation.

Regarding deep learning, there is currently much interest in determining if these methods can have a major impact in the field of numerical PDEs and in particular in parametric PDEs. However, there is a lack of analytical results on the performance of deep neural networks. In **15**, we investigate whether deep networks provide a new method of approximation which out performs more traditional methods of approximation such as finite elements, wavelet and Fourier methods. We show large classes of functions that are not captured by traditional approximation but well captured by deep neural networks. We also show, however, that some of the success of deep networks is due to its instability in selecting the parameters of the network. As such, we are led to understand what limitations are imposed on approximation rates when one requires stability of the numerical algorithms. In particular, this is an issue in understanding the performance of Stochastic Gradient Descent (SGD) which is the common optimization scheme used in deep learning applications.

Publications resulting from this award:

1. A. Cohen, R. DeVore, *Approximation of high-dimensional parametric PDEs*, Acta Numerica, **24** (2015), 1–159.

2. A. Cohen, R. DeVore, *Kolmogorov widths under holomorphic mappings*, IMA Journal of Numerical Analysis, **36** (2016), 1–12.

3. A. Cohen, W. Dahmen, and R. DeVore, *Orthogonal Matching Pursuit under the Restricted Isometry Property*, Constructive Approximation, **45** (2017), 113–127.

4. P. Binev, A. Cohen, W. Dahmen, R. DeVore, G. Petrova, and P. Wojtaszczyk, *Data Assimilation in Reduced Modeling*, SIAM UQ **5** (2017), 1–29.

5. R. A. DeVore, *The Theoretical Foundation of Reduced Basis Methods*, in "Model Reduction and Approximation: Theory and Algorithms."(P. Benner, A. Cohen, M. Ohlberger, and K. Willcox, Eds.), pages 136–167, SIAM, Philadelphia, PA, 2016.

6. M. Bachmayr, A. Cohen, R. DeVore, and G. Migliorati, *Sparse polynomial approximation of parametric elliptic PDEs. Part II: lognormal coefficients*, M2AN, **51** (2017), 321–340.

7. R. DeVore, G. Petrova, and P. Wojtaszczyk, *Data assimilation and sampling in Banach spaces*, Calcolo, **54** (2017), 963–1007.

8. A. Bonito, A. Cohen, R. DeVore, G. Petrova, and G. Welper, *Diffusion Coefficients Estimation for Elliptic Partial Differential Equations*, SIAM J. Math. Anal., **49** (2017), 1570–1592.

9. R. DeVore, S. Foucart, G. Petrova, and P. Wojtaszczyk, *Computing a Quantity of Interest from Observational Data*, Constructive Approximation **49** (2019), 461-508

10. R. DeVore, V. Temlyakov, *Convex optimization on Banach Spaces*, FOCM, **16** (2016), 369–394.

11. A. Cohen, W. Dahmen, R. DeVore, *Reduced Basis Greedy Selection Using Random Training Sets* , arXiv preprint, arXiv:1810.09344.

12. A. Cohen, W. Dahmen, J. Fadili, O. Mula and J. Nichols), *Optimal reduced model algorithms for data-based state estimation*, arXiv preprint, arXiv:1903.07938.

13. A. Bonito, R. DeVore, D. Guignard, P. Jantsch, and G. Petrova, *Polynomial Approximation of Anisotropic Analytic Functions of Several Variables*, arXiv preprint, arXiv:1904.12105.

14. A. Bonito, A. Cohen, R. DeVore, D. Guignard, P. Jantsch, and G. Petrova, *Nonlinear methods for model reduction*, in preparation

15. I. Daubechies, S. Foucart, B. Hanin and G. Petrova), *Nonlinear Approximation and (Deep) ReLU Networks*, arXiv preprint arXiv:1905.02199

References

BCDDPW [1] P. Binev, A. Cohen, W. Dahmen, R. DeVore, G. Petrova, and P. Wojtaszczyk, *Convergence rates for greedy algorithms in reduced basis methods*, SIAM J. Math. Anal., **43** (2011), 1457–1472.

CDS [2] A Cohen, R DeVore, and C Schwab, *Analytic regularity and polynomial approximation of parametric and stochastic elliptic PDE's* Analysis and Applications **9** (2011), 11–47.

CCDS [3] A Chkifa, A Cohen, R DeVore, and C Schwab, *Adaptive algorithms for sparse polynomial approximation of parametric and stochastic elliptic pdes* M2AN **47**(2013), 253-280.

DPW [4] R. DeVore, G. Petrova, and P. Wojtaszczyk, *Greedy agorithms for reduced bases in Banach spaces*, Constructive Approximation, **37** (2013), 455–466.

MPPY [5] Y. Maday, A. Patera, J. Penn, and M. Yano, *A parametrized-background data-weak approach to variational data assimilation: Formulation, analysis, and application to acoustics*, Int. J. Numer. Meth. Eng., **102** (2015), 933–965.