AFRL-AFOSR-VA-TR-2019-0120

Topology Optimization, Fabrication Adaptivity, and Model-Data Assimilation of Novel Photonic Materials

Robert Freund
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

04/24/2019
Final Report

Air Force Research Laboratory
AF Office Of Scientific Research (AFOSR)/ RTA2
Arlington, Virginia 22203
Air Force Materiel Command

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
09-04-2019	Final	01/08/2015-31/12/2018

4. TITLE AND SUBTITLE
Topology Optimization, Fabrication Adaptivity, and Model-Data Assimilation of Novel Photonic Materials

5a. CONTRACT NUMBER
FA9550-15-1-0276

5b. GRANT NUMBER
FA9550-15-1-0276

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Freund, Robert
Peraire, Jaime
Nguyen, Cuong

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Massachusetts Institute of Technology
77 Massachusetts Avenue
Cambridge, MA 02139

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Office of Science and Research
875 Randolph Street
Suite 325 Room 3112
Arlington, VA 22203

10. SPONSOR/MONITOR'S ACRONYM(S)
AFOSR

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Novel materials such as photonic crystals, nanoplasmonics, and metamaterials are engineered to interact with and control electromagnetic waves in ways that cannot be achieved with conventional materials. Such materials are very important in many optical devices including waveguides, fibers, lasers, optical lenses, cancer theranostics, biological and chemical sensing, energy harvesting, etc. Fundamental challenges abound about the design and fabrication of these materials. The issue of fabrication adaptivity (adapting a computed design to actual fabrication) is particularly important in practical applications. We have developed effective systematic numerical methods and tools for the optimal design of photonic materials with particular emphasis on novel applications which are of direct interest to the Air Force. We have worked on three research fronts: (1) robust, efficient and accurate numerical methods for solving acoustic and electromagnetic wave equations in realistic geometries, (2) Gaussian functional regression methods for quantifying uncertainties in the mathematical models, and (3) efficient and robust ``accelerated" first-order methods for solving large-scale PDE-constrained optimization problems arising from mathematical formulations of the design problem.

15. SUBJECT TERMS
wave propagation, metamaterials, periodic media, band-gap optimization, semidefinite programming, photonic crystal design, phononic crystal design, conic optimization, first-order methods, greedy methods, robust regularization

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(Include area code)*

Reset

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. DATES 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. F33615-86-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.

5d. PROJECT NUMBER. Enter all project numbers as they appear in the report, e.g. 1F665702D1257; ILIR.

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.

TOPOLOGY OPTIMIZATION, FABRICATION ADAPTIVITY, AND MODEL-DATA ASSIMILATION OF NOVEL PHOTONIC MATERIALS

Robert Freund and J. Peraire and C. N. Nguyen
Massachusetts Institute of Technology

Final Report
AFOSR Grant: FA9550-15-1-0276
Period covered: 8/1/15-12/31/18
Program Manager: Fariba Fahroo (formerly Jean-Luc Cambier)

Title: Topology Optimization, Fabrication Adaptivity, and Model-Data Assimilation of Novel Photonic Materials

Novel materials such as photonic crystals, nanoplasmonics, and metamaterials are scientifically engineered to interact with and control electromagnetic waves in ways that cannot be achieved with conventional materials. Such materials have proven to be very important as an integrated component in many optical devices including waveguides, fibers, lasers, optical lenses, cancer theranostics, biological and chemical sensing, energy harvesting, etc. These and other novel applications have attracted considerable research interest in photonic crystals, nanoplasmonics and metamaterials. However, fundamental challenges abound about the design and fabrication of these materials in order to yield a given set of prescribed properties. For instance, it is currently beyond the state-of-the-art to compute robust designs that exhibit prescribed properties subject to fabricability constraints. The issue of fabrication adaptivity (adapting a given computed design so that it is fabricable, without significantly deteriorating the design quality) is particularly important in practical applications since the desired length scales and material distributions are often at the limit of our fabrication capability and hence geometric design tolerances (in relative scale) need to be larger. Another important issue to address in the design optimization is the fidelity uncertainty arising in the mathematical model since physical phenomena can rarely be modeled with complete fidelity even under the best of circumstances. Increasing the fidelity of the mathematical model can lead to challenges in numerical simulations. The ability to accurately model and simulate electromagnetic wave propagation through these advanced materials requires capabilities beyond traditional simulation techniques because the problems of interest involve the interaction of long-wavelength light with nanometric structures, which are characterized by the extreme confinement and tight localization of the fields.

Our goal is the development of effective systematic numerical methods and tools for the optimal design of photonic materials with particular emphasis on novel applications which are of direct interest to the Air Force. Towards this goal, we have been working on three research fronts: (1) robust, efficient and accurate numerical methods for solving acoustic and electromagnetic wave equations in realistic geometries, (2) Gaussian functional regression methods for quantifying uncertainties in the mathematical models, and (3) efficient and robust "accelerated" first-order methods for solving large-scale PDE-constrained optimization problems arising from mathematical formulations of the design problem. Below, we describe the main research accomplishments during this past year.

I. Research Results

A. Nanophotonic modeling and simulation

The interaction of light with metallic nanostructures produces a collective excitation of electrons at the metal surface, also known as surface plasmons. These collective excitations lead to resonances that enable the confinement of light in deep-subwavelength regions, thereby leading to large near-field enhancements. Because of the enormous mismatch in length scales between long waves and nanostructures, conventional modeling techniques cannot readily be used for analysis. We developed Hybridizable Discontinuous Galerkin (HDG) method for full three-dimensional modeling of the resonant transmission of THz waves through an annular gap that is 2 nm in width and 32 μm in diameter [1]. Our simulation and atomic layer lithography enable a series of new investigations in THz nanophotonics that has not been possible before. Our recent work [2] demonstrates our method on practical applications simulating ectromagnetic waves in metamaterials with extreme subwavelength features such as single-digit-nanometer gaps.

In a recent paper [5], we develop the hybridizable discontinuous Galerkin (HDG) method to solve Maxwells equations augmented with the hydrodynamic model for the conduction-band electrons in noble metals. This method enables the efficient simulation of plasmonic nanostructures while accounting for the nonlocal interactions between electrons and the incident light. We introduce a novel postprocessing scheme to recover superconvergent solutions and demonstrate the convergence of the proposed HDG method for the simulation of a 2D gold nanowire and a 3D periodic annular nanogap structure. The results of the hydrodynamic model are compared to those of a simplified local response model, showing that differences between them can be significant at the nanoscale.

The interaction of electromagnetic waves with metallic nanostructures generates resonant oscillations of the conduction-band electrons at the metal surface. These resonances can lead to large enhancements of the incident field and to the confinement of light to small regions, typically several orders of magnitude smaller than the incident wavelength. The accurate prediction of these resonances entails several challenges. Small geometric variations in the plasmonic structure may lead to large variations in the electromagnetic field responses. Furthermore, the material parameters that characterize the optical behavior of metals at the nanoscale need to be determined experimentally and are consequently subject to measurement errors. It then becomes essential that any predictive tool for the simulation and design of plasmonic structures accounts for fabrication tolerances and measurement uncertainties. In [3], we develop a reduced order modeling framework that is capable of real-time accurate electromagnetic responses of plasmonic nanogap structures for a wide range of geometry and material parameters. The main ingredients of the proposed method are: (i) the hybridizable discontinuous Galerkin method to numerically solve the equations governing electromagnetic wave propagation in dielectric and metallic media, (ii) a reference domain formulation of the time-harmonic Maxwells equations to account for arbitrary geometry variations; and (iii) proper orthogonal decomposition and empirical interpolation techniques to construct an efficient reduced model. To demonstrate effectiveness of the models developed, we analyze geometry sensitivities and explore optimal designs of a 3D periodic coaxial nanogap structure as shown in Figure 1.

Electromagnetic resonances in nanoslits of various types have been studied by many researchers. A rectangular slit in a metal film exhibits a resonance spectrally close to the cutoff frequency determined by the aperture geometries, such as the slit length and width. Coaxial apertures also possess such cutoff resonances along with higher-order Fabry-Perot (FP) modes. At the cutoff resonance, the real component of the wavevector along the length of the waveguide, is greatly diminished. Numerical simulations of our devices were performed using the finite element method [6]. The simulation results for 10 nm wide coaxial nanogap arrays with four different diameters (Figure 2d) agree well with the measured FTIR spectra (Figure 2e). While we did not explore multiband SEIRA sensing in this work, these results show that it will be possible to mix coaxial apertures with different diameters side-by-side for such applications. Unlike nonzeroth-order FP modes, which show phase variation along the gap, this zero-mode has uniform

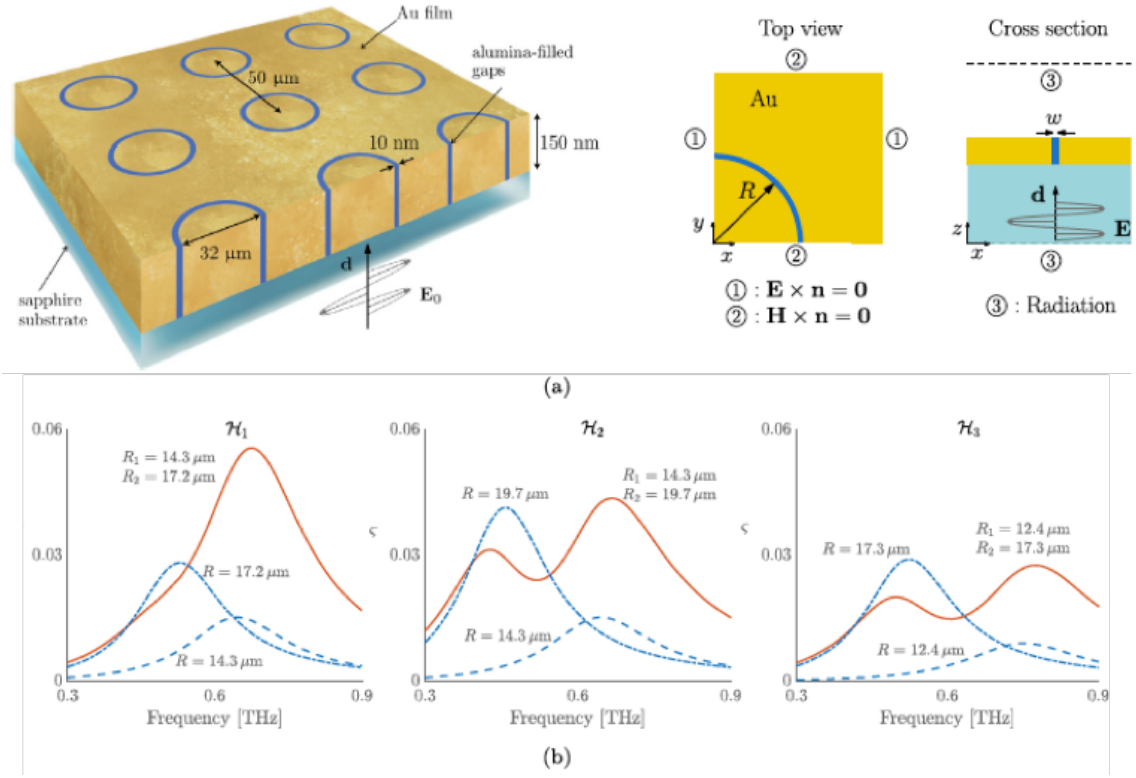

Fig. 1. (a) Schematic diagram of thin gold film on silica substrate patterned with periodic square array of alumina gaps under plane wave THz illumination. (b) Frequency-transmission profiles of optimal configurations for different objective functions.

and strong electric fields inside the entire length of the coaxial nanogap. As these simulations show, by funneling incident IR radiation through annular nanogaps at the resonance condition, it is possible to create extended ring-shaped hot spots along each aperture and obtain a strong electric field $|\mathsf{E}|$ enhancement of 100 fold (Figure 2c).

B. Multiscale continuous Galerkin (MSCG) method

In a recent work [4], we develop a multiscale continuous Galerkin (MSCG) method for the fast and accurate stochastic simulation and optimization of time-harmonic wave propagation through photonic crystals. The MSCG method exploits repeated patterns in the geometry to drastically decrease computational cost and incorporates the following ingredients: (1) a reference domain formulation that allows us to treat geometric variability resulting from manufacturing uncertainties; (2) a reduced basis approximation to solve the parametrized local subproblems; (3) a gradient computation of the objective function; and (4) a model and variance reduction technique that enables the accelerated computation of statistical outputs by exploiting the statistical correlation between the MSCG solution and the reduced basis approximation. The proposed method is thus well suited for both deterministic and stochastic simulations, as well as robust design of photonic crystals. In Fig 3 we show the electric field for a frequency in the first bandgap, a frequency between the two bandgaps and a frequency in the second bandgap. An attractive feature of waveguiding with photonic crystals are the low losses that occur even for sharp bends, thus enabling the efficient manipulation of electromagnetic waves.

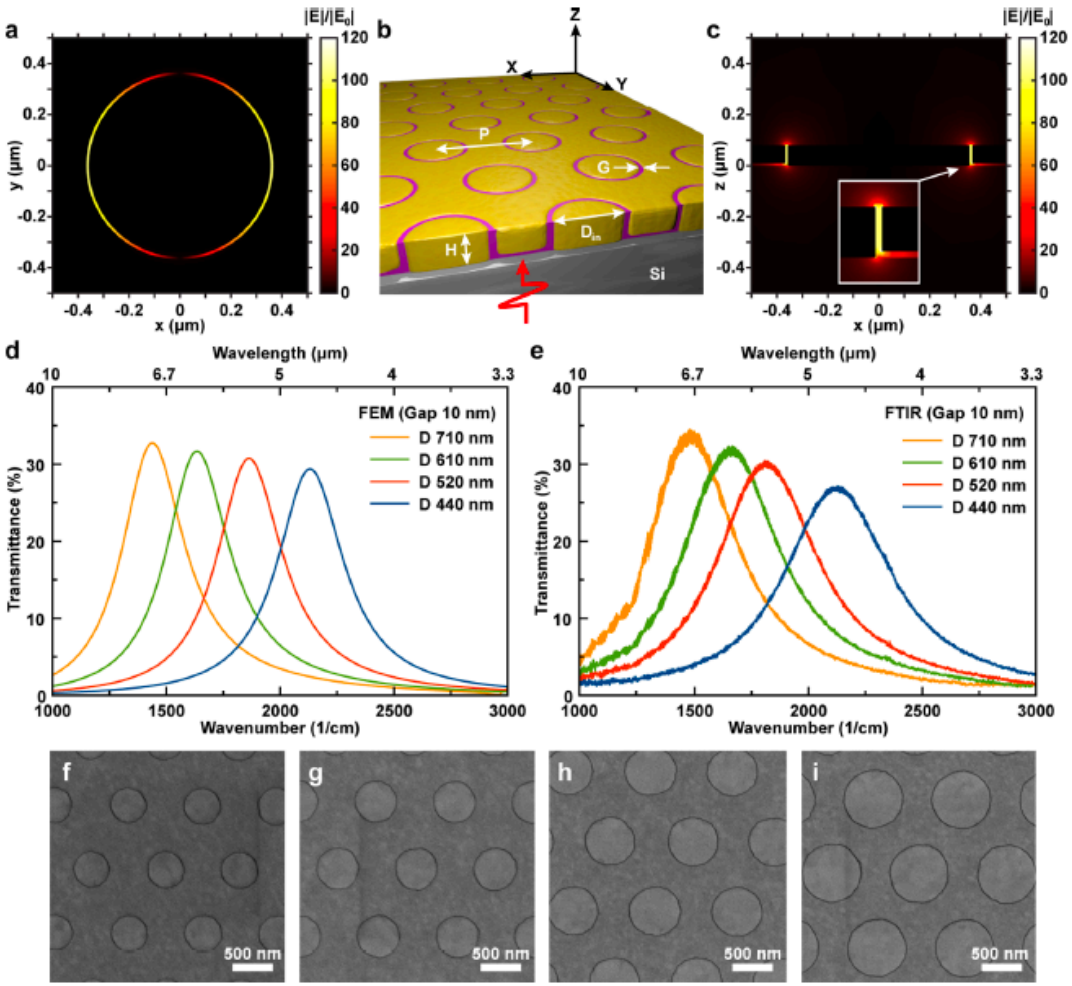

Figure 2. (a–c) Schematics of a coaxial nanoaperture array on a Si substrate. Horizontal and vertical distributions of electric field in the coaxial nanoaperture with a 10 nm gap (G), 710 nm diameter (D_{in}), and 80 nm thickness (H). Simulations were performed using linearly polarized light launched from a Si substrate. (d) Spectra with different diameters calculated using FEM modeling, (e) Spectra with different diameters measured experimentally using FTIR. (f–i) SEM images of 10 nm wide coaxial aperture arrays with 440, 520, 610, and 710 nm diameters, respectively. Scale bars: (f–i) 500 nm.

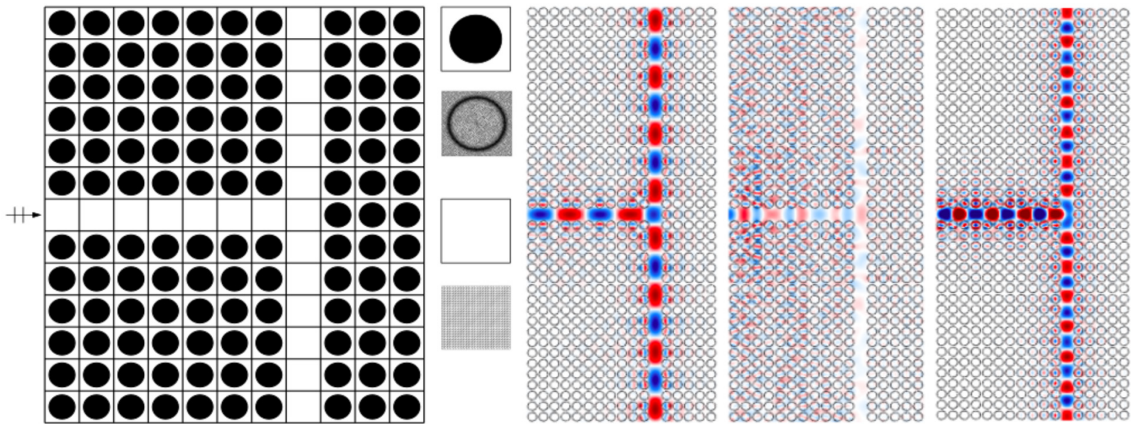

Fig. 2. MSCG simulation of a TM waveguide splitter of GaAs (ϵ=11.4) rods of radii $R_0 = 0.4a$ in air. Computational domain with subdomain decomposition (excluding PMLs) and meshes (left). Numerical simulation at frequencies $\omega = 0.39$, $\omega = 0.46$, and $\omega = 0.53$ (left to right).

C. Hybridizable and embedded discontinuous Galerkin methods

In a recent work [9], we develop hybridizable and embedded discontinuous Galerkin (DG) methods for wave propagation problems in fluids, solids, and electromagnetism. In each of these areas, we describe the methods, discuss their main features, display numerical results to illustrate their performance, and conclude with bibliography notes. The main ingredients in devising these DG methods are (1) a local Galerkin projection of the underlying partial differential equations at the element level onto spaces of polynomials of degree k to parametrize the numerical solution in terms of the numerical trace; (2) a judicious choice of the numerical flux to provide stability and consistency; and (3) a global jump condition that enforces the continuity of the numerical flux to obtain a global system in terms of the numerical trace. These DG methods are termed hybridized DG methods, because they are amenable to hybridization (static condensation) and hence to more efficient implementations. They share many common advantages of DG methods and possess some unique features that make them well-suited to wave propagation problems.

D. Accelerated first-order methods

In a recent work [7], we develop accelerated residual methods for solving linear and nonlinear systems of equations, which leverage both past and recent developments in accelerated gradient methods in convex optimization. First, we propose a modification of Nesterovs accelerated gradient method to obtain an accelerated residual scheme that can be applied to systems of equations. We show that the scheme can be viewed as a finite difference approximation (FDA) of a second-order ordinary differential equation (ODE) system, which turns out to coincide with the ODE system derived from Nesterovs method. In practice, our scheme converges faster than Nesterovs method even though it requires an additional residual evaluation per iteration. Next, we discuss stability properties of our scheme and Nesterovs method for solving linear systems of equations. We then propose an adaptive restarting and a judicious selection of the acceleration parameter to further improve the (empirical) convergence rate of our scheme. Last of all, we generalize the scheme to encompass a family of accelerated residual methods, thereby providing an opportunity to devise improved methods in future work. We demonstrate the usefulness of our scheme on systems of equations resulting from the finite element approximation of linear and nonlinear partial differential equations (PDEs). On a variety of test cases, our numerical results show that the proposed method outperforms pseudo-time marching method, Nesterovs method, and Newton-Krylov methods. We consider this research very relevant

for practical applications. The Hessian-free character of these methods is ideally suited for both parallel simulations and large-scale optimization problems considered in this project.

E. Reduced basis methods for solving linear systems

In [8], we present a class of reduced basis (RB) methods for the iterative solution of parametrized symmetric positive-definite (SPD) linear systems. The essential ingredients are a Galerkin projection of the underlying parametrized system onto a reduced basis space to obtain a reduced system; an adaptive greedy algorithm to efficiently determine sampling parameters and associated basis vectors; an offline-online computational procedure and a multi-fidelity approach to decouple the construction and application phases of the reduced basis method; and solution procedures to employ the reduced basis approximation as a stand-alone iterative solver or as a preconditioner in the conjugate gradient method. We present numerical examples to demonstrate the performance of the proposed methods in comparison with multigrid methods. Numerical results show that, when applied to solve linear systems resulting from discretizing the Poissons equations, the speed of convergence of our methods matches or surpasses that of the multigrid-preconditioned conjugate gradient method, while their computational cost per iteration is significantly smaller providing a feasible alternative when the multigrid approach is out of reach due to timing or memory constraints for large systems. Moreover, numerical results verify that this new class of reduced basis methods, when applied as a stand-alone solver or as a preconditioner, is capable of achieving the accuracy at the level of the truth approximation which is far beyond the RB level.

F. Condition Number Analysis of Logistic Regression, and its Implications for Standard First-Order Solution Methods

The elementary probabilistic model underlying logistic regression implies that it is most natural to consider logistic regression when the data is not (linearly) separable. Building on this basic intuition, in [10] we introduce a pair of condition numbers that measure the degree of non-separability or separability of a given dataset in the setting of binary classification. When the training data is not separable, we show that the degree of non-separability naturally enters the analysis and informs the properties and convergence guarantees of two standard first-order methods: steepest descent (for any given norm) and stochastic gradient descent. Expanding on the work of Bach, we also show how the degree of non-separability enters into the analysis of linear convergence of steepest descent (without needing strong convexity), as well as the adaptive convergence of stochastic gradient descent. When the training data is separable – in which case many properties of logistic regression essentially break down – we demonstrate how the degree of separability enters into the analysis of ℓ_2 steepest descent and stochastic gradient descent for delivering approximate-maximum-margin solutions with associated computational guarantees as well.

G. A New Perspective on Boosting in Linear Regression via Subgradient Optimization and Relatives

In [16] we analyze boosting algorithms in linear regression from a new perspective: that of modern first-order methods in convex optimization. We show that classic boosting algorithms in linear regression, namely the incremental forward stagewise algorithm (FS_ε) and least squares boosting (LS-BOOST(ε)), can be viewed as subgradient descent to minimize the loss function defined as the maximum absolute correlation between the features and residuals. We also propose a minor modification of FS_ε that yields an algorithm for the LASSO, and that may be easily extended to an algorithm that computes the LASSO path for different values of the regularization parameter. Furthermore, we show that these new algorithms for the LASSO may also be interpreted as the same master algorithm (subgradient descent), applied to a regularized version of the maximum absolute correlation loss function. We derive novel, comprehensive computational guarantees for several boosting algorithms in linear regression (including LS-BOOST(ε) and FS_ε) by using techniques of first-order methods in convex optimization. Our computational guarantees inform us about the statistical properties of boosting algorithms. In particular they provide, for the first time, a precise

theoretical description of the amount of data-fidelity and regularization imparted by running a boosting algorithm with a prespecified learning rate for a fixed but arbitrary number of iterations, for *any* dataset.

H. An Extended Frank-Wolfe Method with "In-Face" Directions, and its Application to Low-Rank Matrix Completion

Motivated principally by the low-rank matrix completion problem, we present in [15] an extension of the Frank-Wolfe Method that is designed to induce near-optimal solutions on low-dimensional faces of the feasible region. This is accomplished by a new approach to generating "in-face" directions at each iteration, as well as through new choice rules for selecting between in-face and "regular" Frank-Wolfe steps. Our framework for generating in-face directions generalizes the notion of away-steps introduced by Wolfe. In particular, the in-face directions always keep the next iterate within the minimal face containing the current iterate. We present computational guarantees for the new method that trade off efficiency in computing near-optimal solutions with upper bounds on the dimension of minimal faces of iterates. We apply the new method to the matrix completion problem, where low-dimensional faces correspond to low-rank matrices. We present computational results that demonstrate the effectiveness of our methodological approach at producing nearly-optimal solutions of very low rank. On both artificial and real datasets, we demonstrate significant speed-ups in computing very low-rank nearly-optimal solutions as compared to the Frank-Wolfe Method (as well as several of its significant variants).

I. New Computational Guarantees for Solving Convex Optimization Problems with First Order Methods, via a Function Growth Condition Measure

In [14] we present new computational methods and associated computational guarantees for solving convex optimization problems using first-order methods. Our problem of interest is the general convex optimization problem $f^* = \min_{x \in Q} f(x)$, where we presume knowledge of a strict lower bound $f_{slb} < f^*$. [Indeed, f_{slb} is naturally known when optimizing many loss functions in statistics and machine learning (least-squares, logistic loss, exponential loss, total variation loss, etc.) as well as in Renegar's transformed version of the standard conic optimization problem; in all these cases one has $f_{slb} = 0 < f^*$.] We introduce a new functional measure called the growth constant G for $f(\cdot)$, that measures how quickly the level sets of $f(\cdot)$ grow relative to the function value, and that plays a fundamental role in the complexity analysis. When $f(\cdot)$ is non-smooth, we present new computational guarantees for the Subgradient Descent Method and for smoothing methods, that can improve existing computational guarantees in several ways, most notably when the initial iterate x^0 is far from the optimal solution set. When $f(\cdot)$ is smooth, we present a scheme for periodically restarting the Accelerated Gradient Method that can also improve existing computational guarantees when x^0 is far from the optimal solution set, and in the presence of added structure we present a scheme using parametrically increased smoothing that further improves the associated computational guarantees.

J. Relatively Smooth Convex Optimization by First-Order Methods, and Applications

The usual approach to developing and analyzing first-order methods for smooth convex optimization assumes that the gradient of the objective function is uniformly smooth with some Lipschitz constant L. However, in many settings the differentiable convex function $f(\cdot)$ is not uniformly smooth – for example in D-optimal design where $f(x) := -\ln\det(HXH^T)$ and $X := \mathbf{D}\text{iag}(x)$, or even the univariate setting with $f(x) := -\ln(x) + x^2$. In our paper [13] we develop a notion of "relative smoothness" and relative strong convexity that is determined relative to a user-specified "reference function" $h(\cdot)$ (that should be computationally tractable for algorithms), and we show that many differentiable convex functions are relatively smooth with respect to a correspondingly fairly-simple reference function $h(\cdot)$. We extend two standard algorithms – the primal gradient scheme and the dual averaging scheme – to our new setting, with associated computational guarantees. We apply our new approach to develop a new first-order method for the D-optimal design problem, with associated computational complexity analysis.

K. Accelerating Greedy Coordinate Descent Methods

In [11] we study ways to accelerate greedy coordinate descent in theory and in practice, where "accelerate" refers either to $O(1/k^2)$ convergence in theory, in practice, or both. We introduce and study two algorithms: Accelerated Semi-Greedy Coordinate Descent (ASCD) and Accelerated Greedy Coordinate Descent (AGCD). While ASCD takes greedy steps in the x-updates and randomized steps in the z-updates, AGCD is a straightforward extension of standard greedy coordinate descent that only takes greedy steps. On the theory side, our main results are for ASCD: we show that ASCD achieves $O(1/k^2)$ convergence, and it also achieves accelerated linear convergence for strongly convex functions. On the empirical side, we observe that both AGCD and ASCD outperform Accelerated Randomized Coordinate Descent on a variety of instances. In particular, we note that AGCD significantly outperforms the other accelerated coordinate descent methods in numerical tests, in spite of a lack of theoretical guarantees for this method. To complement the empirical study of AGCD, we present a Lyapunov energy function argument that points to an explanation for why a direct extension of the acceleration proof for AGCD does not work; and we also introduce a technical condition under which AGCD is guaranteed to have accelerated convergence. Last of all, we confirm that this technical condition holds in our empirical study.

L. Generalized Stochastic Frank-Wolfe Algorithm with Stochastic "Substitute" Gradient for Structured Convex Optimization

The stochastic Frank-Wolfe method has recently attracted much general interest in the context of optimization for statistical and machine learning due to its ability to work with a more general feasible region. However, there has been a complexity gap in the guaranteed convergence rate for stochastic Frank-Wolfe compared to its deterministic counterpart. In [12], we present a new stochastic Frank-Wolfe method which closes this gap by introducing the notion of a "substitute" gradient that is a not-necessarily unbiased sample of the gradient. Moreover, we show that this new approach is equivalent to a randomized coordinate mirror descent algorithm applied to the dual problem, which in turn provides a new interpretation of dual coordinate descent method in the primal space. When the regularizer is furthermore strongly convex, we show that the generalized stochastic Frank-Wolfe method as well as the randomized dual coordinate descent present linear convergence. These new results are benefited from the understanding that first-order methods can inherently minimize the primal-dual gap.

REFERENCES

[1] H. R. PARK, X. CHEN, N. C. NGUYEN, J. PERAIRE, AND S. H. OH Nanogap-enhanced terahertz sensing of 1 nm thick ($\lambda/10^6$) dielectric films. *ACS Photonics*, 2 (3), 417–424, 2015.

[2] D. YOO, N. C. NGUYEN, L MARTIN-MORENO, D. A. MOHR, S. CARRETERO-PALACIOS, J. SHAVER, J. PERAIRE, T. W. EBBESEN, AND S. H. OH High-throughput fabrication of resonant metamaterials with ultrasmall coaxial apertures via atomic layer lithography. *Nano letters*, 16 (3), 2040–2046, 2016.

[3] F. VIDAL-CODINA, N. C. NGUYEN, AND J. PERAIRE, *Computing parametrized solutions for plasmonic nanogap structures*, J. Comp. Phys., 366 (2018) 89–106.

[4] F. VIDAL-CODINA, N. C. NGUYEN, AND J. PERAIRE, *A multiscale continuous Galerkin method for stochastic simulation and robust design of photonic crystals*, J. Comp. Phys., vol. 2 (2019) 100016.

[5] F. VIDAL-CODINA, N. C. NGUYEN, S. H. OH, AND J. PERAIRE, *A hybridizable discontinuous Galerkin method for computing nonlocal optical responses in metallic nanostructures*, J. Comp. Phys., 355 (2018): 548-565.

[6] YOO, D., MOHR, D.A., VIDAL-CODINA, F., JOHN-HERPIN, A., JO, M., KIM, S., MATSON, J., CALDWELL, J.D., JEON, H., NGUYEN, N.C. AND MARTIN-MORENO, L., *High-Contrast Infrared Absorption Spectroscopy via Mass-Produced Coaxial Zero-Mode Resonators with Sub-10 nm Gaps*, Nano letters 18, no. 3 (2018): 1930-1936.

[7] N. C. NGUYEN, P. FERNANDEZ, R.M. FREUND, AND J. PERAIRE, *Accelerated residual methods for the iterative solution of systems of equations*, SIAM J. Sci. Comput., 40(5), A3157A3179.

[8] N. C. NGUYEN AND Y. CHEN, *Reduced-basis method for the iterative solution of parametrized symmetric positive-definite linear systems*. arXiv preprint arXiv:1804.06363 (2018).

[9] P. FERNANDEZ, A. CHRISTOPHE, S. TERRANA, N. C. NGUYEN, AND J. PERAIRE, *Hybridized discontinuous Galerkin methods for wave propagation*, Journal of Scientific Computing 77 (3), 1566-1604, 2019.

[10] P. GRIGAS, R. MAZUMDER, AND R.M. FREUND, *Condition Number Analysis of Logistic Regression, and its Implications for Standard First-Order Solution Methods*, submitted to Mathematics of Operations Research, 2018.

[11] H. LU AND R.M. FREUND, *Accelerating Greedy Coordinate Descent Methods*, proceedings of International Conference on Machine Learning (ICML), Stockholm, Sweden, 2018.
[12] H. LU AND R.M. FREUND, *Generalized Stochastic Frank-Wolfe Algorithm with Stochastic "Substitute" Gradient for Structured Convex Optimization*, submitted to Mathematical Programming A.
[13] H. LU, Y. NESTEROV, AND R.M. FREUND, *Relatively Smooth Convex Optimization by First-Order Methods, and Applications*, SIAM Journal on Optimization (2018), vol. 28 (1), pp. 333-354.
[14] H. LU AND R.M. FREUND, *New Computational Guarantees for Solving Convex Optimization Problems with First Order Methods, via a Function Growth Condition Measure*, to appear in Mathematical Programming.
[15] P. GRIGAS, R. MAZUMDER, AND R.M. FREUND, *An Extended Frank-Wolfe Method with "In-Face" Directions, and its Application to Low-Rank Matrix Completion*, SIAM Journal on Optimization 27(1), pp. 319-346, 2017.
[16] P. GRIGAS, R. MAZUMDER, AND R.M. FREUND, *A New Perspective on Boosting in Linear Regression via Subgradient Optimization and Relatives*, Annals of Statistics 45(6), pp. 2328-2364, 2017.