Large-Scale PDE-Constrained Optimization: An Introduction

Omar Ghattas[1], Lorenz T. Biegler[1], Matthias Heinkenschloss[2], and Bart van Bloemen Waanders[3]

[1] Carnegie Mellon University, Pittsburgh, PA
[2] Rice University, Houston, TX
[3] Sandia National Laboratories[†], Albuquerque, NM 87185 USA

Abstract. Optimal design, optimal control, and parameter estimation of systems governed by partial differential equations (PDE) give rise to a class of problems known as PDE-constrained optimization, and the size and complexity of the discretized PDE model often poses significant challenges for contemporary optimization methods. Recent advances in algorithms, software, and high performance computing systems have resulted in PDE simulations that can often scale to millions of variables, thousands of processors, and multiple physics interactions. As PDE solvers mature, there is increasing interest in industry and the national labs in solving optimization problems governed by such large-scale simulations. This paper provides a brief introduction and overview to the *LNCSE volume on PDE Constrained Optimization*.

The volume contains nineteen papers that were initially presented at the *First Sandia Workshop on Large-Scale PDE Constrained Optimization*. The papers in this volume assess the state-of-the-art in PDE-constrained optimization, identify challenges to optimization presented by modern highly parallel PDE simulation codes and discuss promising algorithmic and software approaches to address them. . These papers represent current research of two strong scientific computing communities, in optimization and PDE modeling and simulation. This volume merges perspectives in two different areas and identifies interesting open questions for further research in both areas. We hope that these papers lead to greater synergy and collaboration between these communities.

1 Algorithmic Challenges for PDE-Constrained Optimization

PDE simulation models are widespread in science and engineering applications. Moreover, with increasing development and application of supercomputing hardware and advances in numerical methods, very large and detailed models can now be considered. An essential sequel to simulation models is their application in design, control, data assimilation and analysis. Most of these tasks are naturally stated as continuous variable optimization problems

[†] Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed-Martin Company, for the United States Department of Energy under Contract DE-AC04-94AL85000.

(i.e., nonlinear programs) and it is clear that superior designs and estimates result with efficient, high performance optimization strategies.

Research in optimization algorithms and numerical methods for PDEs has evolved along two parallel paths that are not always compatible with each other. To exploit recent advances in both fields a number of challenges need to be addressed. These can be summarized by the following topics.

1. Problem size of PDE constrained optimization problems.
2. Efficiency vs. intrusiveness of nonlinear programming (NLP) algorithms integrated with PDE solvers
3. "Physics-based" globalizations and inexact PDE solvers that employ various levels of model fidelity
4. Approximate Jacobians used in nonlinear PDE solvers and the application of sensitivity and adjoint methods
5. Implicitly-defined PDE residuals and non-smooth solutions
6. Treatment of inequality constraints in the NLP
7. Time-dependent PDE optimization

In the remainder of this section we briefly consider each of these topics. The next section provides a brief overview of the papers in this volume with the perspective of addressing these topics.

1.1 Problem size in PDE constrained optimization

The main characteristic of PDE constrained optimization problems is their large size. Nonlinear programs derived from PDE formulations are often orders of magnitude larger than other NLP problems. After choosing an appropriate meshing strategy and assembling the relevant finite element (or finite volume or finite difference) equations, the discretized PDE consists of $n = O(10^6)$ state variables and equations for typical production simulations. 'Heroic', leading edge simulations will have $n = O(10^9)$ state variables and equations.

Optimal design problems that incorporate these models will have decision variables that can consist of $O(1)$ physically-based design variables or $O(n^{2/3})$ variables for finely-parameterized shape optimization. Decision variables for optimal control problems scale in a similar manner with $O(n^{2/3})$ decisions if boundary control is applied and $O(n)$ decisions if volume control is used. Finally, inverse problems that incorporate PDE models have decisions that include $O(1)$ homogeneous material parameters, $O(n^{2/3})$ parameters for boundary conditions, and $O(n)$ parameters for heterogeneous materials.

Large-scale PDE constrained nonlinear programs are typically solved using variants of Newton's method applied to the optimality (also known as the Karush-Kuhn-Tucker or KKT) conditions of the NLP. This Newton-based approach, extended to deal with inequality constraints, is known as Sequential Quadratic Programming (SQP) and variants that incorporate the elimination of the state variables and linearized equations are known as reduced space SQP, or rSQP methods. For optimization of 3D PDE systems an iterative linear solver is needed as the KKT matrix is too large to be factored. Here preconditioning of this system is essential and, as described in the next topics, a number of issues need to be addressed related to the calculation or approximation of first and second derivatives and the choice of linear and nonlinear solution algorithms.

1.2 Efficiency vs. Intrusiveness: PDE-solver viewpoint

PDE simulation environments have been developed to support the large-scale solution of PDE models, and a major challenge lies in the efficient implementation of nonlinear programming algorithms. To lend further insight into this challenge, we consider four levels of implementation of NLP algorithms with PDE solvers. At the most basic level, we consider a black-box or Nested Analysis and Design (NAND) implementation. This interface requires very little interaction of the NLP solver with the PDE model. Gradients for the NLP solver (e.g., SQP or rSQP) are obtained by finite difference, a nonlinear elimination of the state variables is assumed, and usually only few decision variables can be considered. On the other hand, the NAND implementation can suffer from repeated and time-consuming solution of the PDE model and intermediate convergence failures of the model.

At the other extreme we can consider a full space SQP method. This implementation allows access to the complete discretization of the PDE model and accurate first and second derivatives that make up the linear KKT system in the space of all of the state and decision variables. To solve this full system, an approximate elimination of states and adjoints is applied as a preconditioner. Therefore, because the iterative solution of the linearized PDE equations is now part of the NLP algorithm, this approach leads to a very fast *SAND (Simultaneous Analysis and Design)* optimization strategy. However, this fully open implementation requires the development of new PDE modeling systems along with rewriting components for meshing, Jacobian approximation and specialized residual calculations of the discretized PDEs.

One can also consider intermediate implementations that are tailored to the PDE modeling system using rSQP. These two approaches are distinguished by whether accurate Jacobian elements of the PDE model can be made available to the NLP algorithm. With the *direct tailored approach* reduced gradients and search directions for rSQP are calculated directly using Newton steps from the PDE modeling system and the "sensitivity" of Newton steps to the decision variables. Nevertheless, these steps require the iterative solution of preconditioned linear systems and are more more expensive than the SAND approach. On the other hand, the *adjoint tailored approach* calculates reduced gradients and search directions for rSQP by using both the model Jacobian and its transpose. This approach is still more expensive than

SAND but it requires far fewer iterative linear solutions than the direct approach. It also provides multiplier estimates for the NLP solver.

1.3 Physics-based globalizations

In addition to their large size, PDE models have a number of problem specific features. They can exhibit severe nonlinearities along with steep profiles and poorly conditioned solutions. As a result, a number of strategies are needed to promote convergence of PDE models. In particular, a wide variety of PDE physics-based globalizations are used. These include:

- mesh sequencing to reduce the problem nonlinearity, improve the conditioning on a coarse grid, and bootstrap a solution on a fine grid.
- continuation methods on nonlinear parameters that can involve pseudo-time stepping to the PDE solution
- lower-fidelity precursor models that provide good initializations for the discretized PDE model.

These globalization strategies lead to considerable deviations from Newton-based equation solving. As a result, they do not lend themselves to straightforward incorporation within simultaneous methods. The challenge here is to find the best ways to integrate these globalizations into SAND methods.

Moreover, large-scale PDE solvers are often inexact. Both the linear and nonlinear solvers applied to the discretized PDE equations are iterative and one is often content only to reduce the (linear, nonlinear) residual by several orders of magnitude. While inexact solvers can be used to advantage far from optimum, care must be taken in constructing a convergent and efficient optimization strategy that incorporates these solvers.

1.4 Approximate Jacobians

The sources of inexactness in PDE models and solution strategies are most evident in the construction of (approximate) Jacobians from the discretized PDE equations. For convergence and even for successful performance, PDE solvers do not require exact Jacobians of the model equations, and most PDE solvers do not form the exact Jacobian. For PDE optimization, on the other hand, the exact PDE Jacobian is required to compute the reduced gradients of the objective (and constraint functions) and the KKT conditions for termination.

Approximate Jacobians in PDE solvers result from a number of sources. Often, derivative terms are dropped due to complexities in differentiating constitutive models and dependence of PDE coefficients on state variables in the PDE model. Neglected Jacobian elements may also be due to moving mesh schemes and stabilization schemes in the PDE solver. Other sources of inaccuracy may be due to incomplete linear factorizations that sacrifice coupling in the PDE system and lower-order approximations. Moreover, it should be noted that modern Newton PDE solvers may not form the Jacobian either. While previously the approximate Jacobian was used directly within a Newton-like solver, current methods use the approximate Jacobian to precondition matrix-free Newton-Krylov methods, wherein the Jacobian-vector product can be generated by directional differencing of the discretized equations.

The generation of approximate Jacobians has important implications on PDE-based optimization strategies and algorithms. Exact Jacobians and Hessians are essential for the SAND approach but not necessary for NAND. Moreover, approximate Jacobians influence the two intermediate tailored methods in different ways. For the *direct rSQP approach*, reduced gradients are generated by the iterative solution of linear sensitivity equations which have a structure analogous to the state equations. Thus the same matrix-free Krylov solver and preconditioner (represented by the approximate Jacobian) that is applied to the state equations can be used for the sensitivity equations. For the *adjoint rSQP approach*, which is essential for large problems, a matrix-free approach cannot be applied without the ability to evaluate matrix-vector products with the *transpose* of the Jacobian. Hence, exact derivatives are a necessity here.

1.5 Implicitly-defined and non-smooth PDE residuals

Obtaining accurate Jacobians and higher derivatives for optimization is further complicated by PDE equations that are implicitly defined and potentially nonsmooth. Evaluating the PDE residual may require complex internal calculations that are due to solver specific features such as solution-adaptivity, discontinuity-capturing and front-tracking. An example of this can be seen in the ALE/moving-mesh treatment of dynamic interfaces. Model features of PDEs also contribute to difficulties with implicit residuals. These can be found in subgrid-scale models, material property calculations, evaluation of nondifferential contact problems, and transitions for elasto(visco)plasticity. In many cases, the global residual may not even be formed, as in code-coupled multiphysics problems.

A key conceptual problem with implicitly defined functions is that the PDE residual may not depend smoothly (or even continuously!) on state variables. Moreover, some non-smoothness effects are an artifact of the numerical method, including adaptivity, unstructured meshers or nonsmoothness due to the solid modeler or mesh generator. Other nonsmooth effects can possibly be reformulated away by promoting them to level of optimizer, as with variational inequalities that arise in contact and plasticity models.

1.6 Treatment of Inequalities

Many optimal design problems have bounds on decision variables and also inequalities on state variables at points in the PDE domain; this can lead to $O(n)$ inequalities, which are often nonlinear. Examples of these include avoiding material failure, reducing vorticity, and limits on temperature, flow and concentration.

In classical SQP and rSQP methods, inequalities are handled through the active set solution of the quadratic programming subproblem. However, determination of the active set of inequalities is combinatorial and may be extremely expensive for large systems. Moreover, in the rSQP method, reduced gradients are required for all of the inequalities as well as for the objective function. While this information is provided for free with the direct rSQP method, the calculation of these reduced gradients with the adjoint rSQP method adds a large computational cost, and removes any advantage of the adjoint approach over the direct approach.

Recently, large-scale interior point (or barrier) NLP algorithms have been considered to address these challenges. Here inequality constraints are lumped into a barrier term and the combinatorial problem is transformed to solving a set of nonlinear equations. In addition, a reduced gradient is required only for the barrier function and this is obtained cheaply with an adjoint method. Nevertheless, there are still some open questions with this approach. Barrier problems may be ill-conditioned and more efficient ways to incorporate second order information must be found.

1.7 Time-dependent problems

Because dynamic PDE optimization extends directly from the optimization of static PDE systems, all of the above issues are captured in time dependent problems. These systems are represented by a large problem size and the forward dynamic structure of these problems needs to be exploited.

Here, SAND methods require the storage of entire time history of the state variables; this is clearly not practical for large problems. On the other hand, NAND methods can be applied using final-value adjoint equations. Here reduced gradient are calculated with state variables (forward-in-time) and adjoints (backward-in-time). To avoid storage of the complete state-time history, checkpointing and compression schemes can be used. Nevertheless, there are a number of open questions that remain in design optimization algorithms that best combine the advantages of SAND and NAND.

1.8 Summary comments

Optimization of PDE-constrained problems will lead to tremendous improvements in the design, control, estimation and analysis of devices and systems in science and engineering. Moreover, tackling the above challenges leads to a rich source of exciting research problems.

The emergence of PDE-constrained optimization comes at a time when many legacy PDE codes are being rewritten to take advantage of distributed memory parallel machines as well as advances in object oriented software development. This therefore provides a window of opportunity to influence design of PDE solvers to make them optimization-friendly. As a result, we hope that these circumstances provide significant opportunities for cross-fertilization between optimization and simulation communities. It is in this spirit that we introduce the following papers in this volume.

2 Overview of Workshop Papers

Collectively, the papers in this volume assess the current state of the art in PDE-constrained optimization, and address some of the difficulties expressed above that are posed by PDE constraints and modern PDE solvers. These include discretization issues, unavailability of exact Jacobians, the expense of computing sensitivities, the need to accommodate approximate solvers and inexact problem information, large sets of inequality constraints that may arise from discretized pointwise constraints, opportunities to exploit models of varying fidelity, large storage problems presented by time-dependent PDEs, and considerable software engineering challenges.

2.1 State of the Art

The first three papers deal with the leading edge applications of PDE-constrained optimization. The first paper by Young et al. [19] discusses a number of issues related to implementing an SQP optimization strategy to the TRANAIR CFD code at Boeing. Here a nonlinear elimination (NAND) strategy is presented along with a detailed algorithm and examples drawn from airfoil design. The second paper by Salinger et al. [17] deals with a rSQP tailored optimization implementation on MPSalsa, a reacting flow code at Sandia National Laboratories. This approach is illustrated with the design optimization of a chemical vapor deposition reactor. The third paper by Biros and Ghattas [6] describes the Lagrange-Newton-Krylov-Schur (LNKS) method for SAND optimization. The paper provides a conceptual derivation of the method, a number of illustrative examples and a discussion of how LNKS can incorporate inexact derivative information within the KKT matrix.

2.2 Multifidelity Models

To deal with inexactness and physics-based globalizations in PDE models, the next two papers develop optimization strategies that use multifidelity models, i.e. less accurate low fidelity models to deal with most of the search space and high fidelity models to refine the search and obtain convergence. The first paper by Alexandrov and Lewis [1] deals with a multifidelity model management scheme that enforces first order consistency among the models and guarantees convergence for the high fidelity model. The approach is demonstrated on the optimization of a multi-element airfoil operating in transonic conditions. The second paper by Cox and Haftka [8] deals with a multifidelity extension of the DIRECT optimization algorithm, a variation of a Lipschitzian method. The paper details this approach and describes its extension to deal with search among models with different levels of accuracy, including polynomial interpolants. The approach is illustrated on a small test problem.

2.3 Sensitivity for PDE-based Optimization

The next three papers deal with the computation of sensitivities for PDE-constrained optimization. The first paper by Li et al. [13] deals with a mesh refinement scheme in space that is combined with a BDF solver in time and also adapted to yield sensitivities. They describe mesh refinement and time integration as well as a number of implementation details. The approach is illustrated on a flame propagation problem. The next paper by Hovland et al. [12] discuss the state of the art for automatic differentiation (AD) using ADIFOR. It also discusses challenges and planned developments in advancing these methods within a number of object oriented toolkits for PDE modeling and optimization. Experimental performance results are presented using the PETSc and SensPVODE toolkits. The third paper by Griewank and Faure [11] presents new results on properties of automatic differentiation. In particular, they analyze the convergence of AD for fixed point iterations for both direct and adjoint schemes. Moreover, the rapid performance of their AD algorithm is demonstrated on a 2D Euler code. The results of these papers have important implications for obtaining first and second derivatives for PDE-constrained optimization problems.

2.4 NLP Algorithms and Inequality Constraints

The next five papers deal with advances for SQP-type NLP algorithms, particularly with respect to the treatment of inequality constraints. The first paper by Morales et al. [15] presents an experimental case study that compares two interior point SQP codes, KNITRO and LOQO, with two active set SQP codes, FilterSQP and SNOPT. Applied to unconstrained, equality constrained and general constrained test problems, the authors include a number of observations on these four algorithms and note the advantages of interior point methods. The second paper by Betts et al. [4] describes recent advances with the *SOCS* NLP solver at Boeing. Here an primal-dual interior point is implemented and a filter line search is incorporated for globalization. The approach is illustrated on a highly constrained, high index PDE where the interior point has significant advantages over the active set strategy. The next paper by Biegler and Wächter [5] considers a number of options for rSQP strategies including second order information, type of line search strategy and active set vs. interior point methods. A comparison of options is given on discretized optimal control problems and some conclusions are offered for their use with PDE-constrained optimization.

The last two papers in this set deal with particular NLP applications. The paper by Bank et al. [2] deals with the application of primal- dual interior methods to a class of elliptic variational inequalities. Previously, these problems were posed as unconstrained optimization problems and are normally solved with a preconditioned conjugate-based Newton method applied to the finite element equations. The interior point algorithm is shown to extend this method naturally to include obstacle inequality constraints, and the paper presents several examples with these constraints. The fifth paper by Boggs et al. [7] deals with the hierarchical control of a linear diffusion equation. This is posed as a multicriterion optimization problem and reformulated as a highly constrained nonlinear program. Several cases are considered and solved using an SQP approach.

2.5 Time dependent problems

This set of papers addresses the challenges due to the optimization of large time-dependent PDE problems. The first paper by Muske and Howse [16] deals with a Lagrangian method for simultaneous nonlinear model predictive control. Here time dependent advection-diffusion equations are considered and a SAND optimization is proposed using a Newton-Krylov solver and block preconditioner. The next paper by Fahl and Sachs [9] describes an approach to PDE-constrained optimization based on Proper Orthogonal Decomposition (POD). POD offers a systematic way to simplify the PDE problem but safeguards are needed to ensure convergence for the original model. Here a POD-based trust region framework is proposed. The subsequent algorithm is applied to control of a time dependent heat conduction. The third paper by Gockenbach and Symes [10] deals with optimization problems that depend on adaptive marching simulations in time. Here, the computed states are often nonsmooth even though the underlying problem is well defined. The authors describe an abstract framework is described that allows the computation of sensitivities as part of the Hilbert Class Library.

2.6 Frameworks for PDE-constrained optimization

This last section deals with future developments and opportunities within computational frameworks for PDE-constrained optimization. The first paper by Stewart and Edwards [18] details the *Sierra* framework developed at Sandia National Labs. The main goal is to integrate mesh management, field variable management and algorithm support services along with a suite of existing application codes within a distributed parallel computing environment. The second paper by Bartlett and Biegler [3] describes *rSQP++* an object-oriented framework for SQP. This framework supports different variations of SQP algorithms as well as specific linear algebra solvers. It leads to algorithm configurations that exploit the structure of PDE based models as well as the architecture of the computing environment. Finally, the paper by Long [14] describes the *Sundance* framework, a rapid prototyping tool for parallel PDE optimization. Here the parallel PDE solver can be accessed through a high level problem description. Using a small amount of user-level code the variational PDE formulation, the discretization method and even operators for Jacobians and Hessians can be specified for use in optimization algorithms.

3 Concluding Remarks and Acknowledgements

The papers in this volume represent a rich spectrum of research in PDE-constrained optimization. In particular we hope that they will spur future research in this exciting field, especially through collaborations between the PDE and optimization communities.

Finally, we are pleased to acknowledge the following groups and individuals, without whom this volume would not have been realized. Our thanks go to:

- the authors, reviewers and session chairs for the workshop papers and their efforts in producing the papers in this volume,
- Ms. Linda Benavides, Ms. Deanna Ceballos and Ms. Barbara DeLap at Sandia National Labs for administrative support,
- the Computer Science Research Institute at Sandia National Labs for sponsorship and support of the workshop,
- the National Science Foundation under Grant ACI 0116984 for providing travel support for graduate students and researchers,
- Dr. Martin Peters, editor of the LNCSE series at Springer Verlag for his support and assistance related to this volume.

References

1. Alexandrov, Natalia and R. Michael Lewis, "First-Order Approximation and Model Management in Optimization"

2. Bank, Randolph E., Philip E. Gill and Roummel F. Marcia, "Interior Methods For a Class of Elliptic Variational Inequalities"

3. Bartlett, Roscoe A. and Lorenz T. Biegler, "rSQP++ : An Object-Oriented Framework for Successive Quadratic Programming"

4. Betts, J. T., Samuel Eldersveld, Paul D. Frank, John G. Lewis, "An Interior-Point Algorithm for Large Scale Optimization"

5. Biegler, Lorenz T., and Andreas Wächter, "SQP SAND Strategies that Link to Existing Modeling Packages"

6. Biros, George and Omar Ghattas, "Inexactness Issues in Lagrange-Newton-Krylov- Schur Methods"

7. Boggs, P. T., Kearsley, A. J. and J. W. Tolle, "Hierarchical Control of a Linear Diffusion Equation"

8. Cox, Steven E., and Rafael Haftka, "Multifidelity Global Optimization using DIRECT"

9. Fahl, Marco and Ekkehard W. Sachs, "Reduced Order Modelling Approaches to PDE-Constrained Optimization Based on Proper Orthogonal Decomposition"

10. Gockenbach, Mark S. and William W. Symes, "Adaptive Simulation, the Adjoint State Method and Optimization"

11. Griewank, Andreas and Christele Faure, "Piggyback Differentiation and Optimization"

12. Hovland, Paul, Steven Lee, Lois McInnes, Boyana Norris and Barry Smith, "Challenges and Opportunities in Using Automatic Differentiation with Object-Oriented Toolkits for Scientific Computing "

13. Li, Shengtai, Linda R. Petzold and James M. Hyman, "Solution Adapted Mesh Refinement and Sensitivity Analysis for Parabolic Partial Differential Equation Systems"

14. Long, Kevin, R., "Sundance: A Rapid Prototyping Tool for Parallel PDE Optimization"

15. Morales, J. L., J. Nocedal, R. A. Walz, G. Liu and J-P. Goux, "Assessing the Potential of Interior Methods for Nonlinear Optimization"

16. Muske, Kenneth R. and James W. Howse, "A Lagrangian Method for Simultaneous Nonlinear Model Predictive Control"

17. Salinger, A.G., R.P. Pawlowski, J. N. Shadid, B. van Bloemen Waanders, R. Bartlett G.C. Itle and L. Biegler, "Optimization of Large-Scale Reacting Flows using MPSalsa and Sequential Quadratic Programming"

18. Stewart, James, R. and H. Carter Edwards, " The Sierra Framework for Developing Advanced Parallel Mechanics Applications "

19. Young, D. P., W. P. Huffman, R. G. Melvin, C. L. Hilmes and F. T. Johnson, "Nonlinear Elimination in Aerodynamic Analysis and Design Optimization"

4 Suggested Table of Contents

- State of the Art
  - Young, D. P., W. P. Huffman, R. G. Melvin, C. L. Hilmes and F. T. Johnson, "Nonlinear Elimination in Aerodynamic Analysis and Design Optimization"
  - Salinger, A.G., R.P. Pawlowski, J. N. Shadid, B. van Bloemen Waanders, R. Bartlett G.C. Itle and L. Biegler, "Optimization of Large-Scale Reacting Flows using MPSalsa and Sequential Quadratic Programming"
  - Biros, George and Omar Ghattas, "Inexactness Issues in Lagrange-Newton-Krylov- Schur Methods"
- Multifidelity Models
  - Alexandrov, Natalia and R. Michael Lewis, "First-Order Approximation and Model Management in Optimization"
  - Cox, Steven E., and Rafael Haftka, "Multifidelity Global Optimization using DIRECT"
- Sensitivity for PDE-based Optimization
  - Li, Shengtai, Linda R. Petzold and James M. Hyman, "Solution Adapted Mesh Refinement and Sensitivity Analysis for Parabolic Partial Differential Equation Systems"
  - Hovland, Paul, Steven Lee, Lois McInnes, Boyana Norris and Barry Smith, "Challenges and Opportunities in Using Automatic Differentiation with Object-Oriented Toolkits for Scientific Computing "
  - Griewank, Andreas and Christele Faure, "Piggyback Differentiation and Optimization"
- NLP algorithms and inequality constraints
  - Morales, J. L., J. Nocedal, R. A. Walz, G. Liu and J-P. Goux, " Assessing the Potential of Interior Methods for Nonlinear Optimization "
  - Betts, J. T., Samuel Eldersveld, Paul D. Frank, John G. Lewis, "An Interior-Point Algorithm for Large Scale Optimization"
  - Biegler, Lorenz T., and Andreas Wächter, "SQP SAND Strategies that Link to Existing Modeling Packages"

    - Bank, Randolph E., Philip E. Gill and Roummel F. Marcia, "Interior Methods For a Class of Elliptic Variational Inequalities"

    - Boggs, P. T., Kearsley, A. J. and J. W. Tolle, "Hierarchical Control of a Linear Diffusion Equation"

- Time dependent problems
    - Muske, Kenneth R. and James W. Howse, "A Lagrangian Method for Simultaneous Nonlinear Model Predictive Control "

    - Fahl, Marco and Ekkehard W. Sachs, "Reduced Order Modelling Approaches to PDE-Constrained Optimization Based on Proper Orthogonal Decomposition"

    - Gockenbach, Mark S. and William W. Symes, Adaptive Simulation, "the Adjoint State Method and Optimization"

- Frameworks for PDE-constrained optimization
    - Stewart, James, R. and H. Carter Edwards, " The Sierra Framework for Developing Advanced Parallel Mechanics Applications "

    - Bartlett, Roscoe A. and Lorenz T. Biegler, "rSQP++ : An Object-Oriented Framework for Successive Quadratic Programming"

    - Long, Kevin, R., "Sundance: A Rapid Prototyping Tool for Parallel PDE Optimization"