I. Summary of Completed Project

This proposal addresses the development of the theory and implementation of algorithms for semidefinite and cone programming. More specifically, its objectives consist in:

1) developing new and/or improving existing algorithms and implementations for first-order smooth and non-smooth methods for SDP and more general CP problems, as well as VI and saddle point problems;

2) developing inexact first-order augmented Lagrangian methods applicable to cone programming;

3) studying the geometry of the central path and its implication on the theoretical complexities of path-following IP algorithms for LP and other cone programs.

Fifteen works [1, 2, 3, 4, 5, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16] acknowledge the ONR grant N00014-11-1-0062. The works [2, 3, 7, 8, 9, 10, 13, 14, 15, 16] were written during the duration of this project. Moreover, the works [1, 4, 5, 11, 12] were written during the duration of the previous ONR grant but are still being revised and updated during the duration of current grant. Eight among the above fifteen works have already appeared or been accepted for publication.

II. General Activities

The P.I. has been collaborating with the following investigators during the course of this grant: Guanghui Lan (Univ. of Florida), Zhaosong Lu (Simon Fraser, Canada), Mauricio Sicre (Universidade Federal da Bahia, Brazil), Benar Svaiter (IMPA, Brazil), Haesun Park (Georgia Tech), and Takashi Tsuchiya (Institute of Statistical Mathematics, Japan).
The P.I. is currently serving as Associate Editor for the journal *Mathematics of Operations Research* (2003-present).

This grant has partially supported one Ph.D. student, namely Camilo Ortiz. He has already co-authored three papers with the PI and is currently working on a fourth paper about block-decomposition methods for cone programming in which the conjugate gradient method is used to solve reoccurring linear system of equations of the form $(\mathcal{A}\mathcal{A}^*)u = d$, where $\mathcal{A}$ is the coefficient matrix of the cone programming problem.

Finally, a total of thirteen presentations on the research work supported by this grant were given during the period under consideration.

### III. New Research

This section briefly discusses the results of the papers [7, 8, 9, 10, 12, 13, 14, 15, 16].

Paper [12] analyzes the iteration-complexity of the hybrid proximal extragradient (HPE) method for finding a zero of a maximal monotone operator recently proposed by Solodov and Svaiter. One of the key points of our analysis is the use of a new termination criteria based on the $\varepsilon$-enlargement of a maximal monotone operator. The advantages of using this termination criterion is that its definition does not depend on the boundedness of the domain of the operator. We then show that Korpelevich's extragradient method for solving monotone variational inequalities falls in the framework of the HPE method. As a consequence, using the complexity analysis of the HPE method, we obtain new complexity bounds for Korpelevich's extragradient method which do not require the feasible set to be bounded, as assumed in the recent paper by Nemirovski [17]. Another feature of our analysis is that the derived iteration complexity bounds are proportional to the distance of the initial point to the solution set. Using also the framework of the HPE method, we study the complexity of a variant of a Newton-type extragradient algorithm proposed by Solodov and Svaiter for finding a zero of a smooth monotone function with Lipschitz continuous Jacobian.

Paper [14] considers both a variant of Tseng’s modified forward-backward splitting method and an extension of Korpelevich’s method for solving hemi-variational inequalities with Lipschitz continuous operators. By showing that these methods are special cases of the HPE method, it derives iteration-complexity bounds for them to obtain different types of approximate solutions. In the context of saddle-point problems, it also derives complexity bounds for these methods to obtain another type of an approximate solution, namely that of an approximate saddle point. Finally, it illustrates the usefulness of the above results by applying them to a large class of linearly
constrained convex programming problems, including for example cone programming and problems whose objective functions converge to infinity as the boundary of its domain is approached.

Paper [16] considers the monotone inclusion problem consisting of the sum of a continuous monotone map and a point-to-set maximal monotone operator with a separable two-block structure and introduces a framework of block-decomposition prox-type algorithms for solving it which allows for each one of the single-block proximal subproblems to be solved in an approximate sense. Moreover, by showing that any method in this framework is also a special instance of the HPE method introduced by Solodov and Svaiter, it derives corresponding convergence rate results. It also describes some instances of the framework based on specific and inexpensive schemes for solving the single-block proximal subproblems. Finally, it considers some applications of our methodology to: i) propose new algorithms for the monotone inclusion problem consisting of the sum of two maximal monotone operators, and; ii) study the complexity of the classical alternating minimization augmented Lagrangian method for a class of linearly constrained convex programming problems with proper closed convex objective functions.

The previous paper also uses the HPE framework to study the iteration-complexity of a first-order (or, in the context of optimization, second-order) method for solving monotone nonlinear equations. The purpose of the paper [15] is to extend this analysis to study a prox-type first-order method for monotone smooth variational inequalities and inclusion problems consisting of the sum of a smooth monotone map and a maximal monotone point-to-set operator. Each iteration of the method computes an approximate solution of a proximal subproblem, obtained by linearizing the smooth part of the operator in the corresponding proximal equation for the original problem, which is then used to perform an extragradient step as prescribed by the HPE framework. Both pointwise and ergodic iteration-complexity results are derived for the aforementioned first-order method using corresponding results obtained here for a subfamily of the HPE framework.

Paper [10] presents a primal interior-point hybrid proximal extragradient (HPE) method for solving a monotone variational inequality over a closed convex set endowed with a self-concordant barrier and whose underlying map has Lipschitz continuous derivative. In contrast to the method of [15] in which each iteration required an approximate solution of a linearized variational inequality over the original feasible set, the present one only requires solving a Newton linear system of equations. The method performs two types of iterations, namely: those which follow an ever changing
path within a certain “proximal interior central surface” and those which correspond to a large-step HPE iteration of the type described in [15]. Due to its first-order nature, the iteration-complexity of the method is shown to be faster than its 0-th order counterparts such as Korpelevich’s algorithm and Tseng’s modified forward-backward splitting method.

Paper [13] presents an accelerated variant of the HPE method for convex optimization, referred to as the accelerated HPE (A-HPE) method. Iteration-complexity results are established for the A-HPE method, as well as a special version of it, where a large stepsize condition is imposed. Two specific implementations of the A-HPE method are described in the context of a structured convex optimization problem whose objective function consists of the sum of a smooth convex function and an extended real-valued non-smooth convex function. In the first implementation, a generalization of a variant of Nesterov’s method is obtained for the case where the smooth component of the objective function has Lipschitz continuous gradient. In the second implementation, an accelerated Newton proximal extragradient (A-NPE) method is obtained for the case where the smooth component of the objective function has Lipschitz continuous Hessian. It is shown that the A-NPE method has a $O(1/k^{7/2})$ convergence rate, which improves upon the $O(1/k^3)$ convergence rate bound for another accelerated Newton-type method presented by Nesterov. Finally, while Nesterov’s method is based on exact solutions of subproblems with cubic regularization terms, the A-NPE method is based on inexact solutions of subproblems with quadratic regularization terms, and hence is potentially more tractable from a computational point of view.

Paper [7] considers block-decomposition first-order methods for solving large-scale conic semidefinite programming problems. Several ingredients are introduced to speed-up the method in its pure form such as: an aggressive choice of stepsize for performing the extragradient step; use of scaled inner products in the primal and dual spaces; dynamic update of the scaled inner product in the primal space for properly balancing the primal and dual relative residuals; and proper choices of the initial primal and dual iterates, as well as the initial parameter for the primal scaled inner product. Finally, it presents computational results showing that the proposed method substantially outperforms the two most competitive codes for large-scale conic semidefinite programs, namely: the boundary point method introduced by Povh et al. and the Newton-CG augmented Lagrangian method by Zhao et al.

Paper [8] considers a first-order block-decomposition-type method for solving a minimization problem whose objective function is the sum of a finite everywhere
convex function with Lipschitz continuous gradient and two proper closed convex 
(possibly, nonsmooth) functions with easily computable resolvents. The method 
presented contains two important ingredients from a computational point of view, 
namely: an adaptive choice of stepsize for performing the extragradient step; and, 
the use of a scaling factor to balance the blocks. Its specialization to the context 
of conic semidefinite programming (SDP) problems consisting of two easy blocks of 
constraints is also discussed. Without putting them in standard form, it is shown 
that four important classes of graph-related SDP problems automatically possess the 
above two-easy-block structure, namely: SDPs for $\theta$-functions and $\theta_+$-functions of 
graph stable set problems, and SDP relaxations of binary integer quadratic and 
frequency assignment problems. It is also shown that SDPs in standard form can be 
viewed as possessing the two-easy-block structure as long as it is easy to project onto 
the associated affine manifold constraint. Finally, computational results on the afore-
mentioned classes of SDPs are presented which show that the method outperforms 
the three most competitive codes for large-scale conic semidefinite programs, namely: 
the boundary point (BP) method introduced by Povh et al., a Newton-CG augmented 
Lagrangian method, called SDPNAL, by Zhao et al., and a variant of the BP method, 
called the SPDAD method, by Ma et al.

Paper [9] presents a new accelerated variant of Nesterov’s method for solving 
a class of convex optimization problems, in which certain acceleration parameters 
are adaptively (and aggressively) chosen so as to: preserve the theoretical iteration-
complexity of the original method, and; substantially improve its practical performance 
in comparison to the other existing variants. Computational results are pre-
sented to demonstrate that the proposed adaptive accelerated method performs quite 
well compared to other variants proposed earlier in the literature.

IV. Related or Continuing Research

This section briefly discusses the results of the papers [1, 2, 3, 4, 5, 11]

Paper [5] considers a special but broad class of convex programing (CP) problems 
whose feasible region is a simple compact convex set intersected with the inverse 
image of a closed convex cone under an affine transformation. Two first-order penalty 
methods for solving the above class of problems are studied, namely: the quadratic 
penalty method and the exact penalty method. In addition to one or two gradient 
evaluations, an iteration of these methods requires one or two projections onto the 
simple convex set. The paper establishes iteration-complexity bounds for the above 
methods to obtain two types of near optimal solutions, namely: near primal and near
primal-dual optimal solutions. Finally, the paper presents variants, with possibly better iteration-complexity bounds than the aforementioned methods, which consist of applying penalty-based methods to the perturbed problem obtained by adding a suitable perturbation term to the objective function of the original CP problem.

Paper [4] studies the iteration-complexity of a first-order augmented Lagrangian method for finding near primal-dual optimal solutions of a special but broad class of convex programming (CP) problems whose feasible regions consist of a simple compact convex set intersected with an affine manifold. In addition to one or two gradient evaluations, an iteration of the above method requires one or two projections onto the simple convex set. It also presents a variant, with a better iteration-complexity bound than the aforementioned method, which consists of applying a first-order augmented Lagrangian method to the perturbed problem obtained by adding a suitable perturbation term to the objective function of the original CP problem.

Paper [11] considers a framework of inexact proximal point methods for convex optimization that allows a relative error tolerance in the approximate solution of each proximal subproblem and establish its convergence rate. It is shown that the well-known forward-backward splitting algorithm for convex optimization belongs to this framework. Finally, an inexact forward-backward splitting algorithm is proposed for solving optimization problems whose objective functions are obtained by maximizing convex-concave saddle functions.

Sparse principal component analysis (PCA) imposes extra constraints or penalty terms to the original PCA to achieve sparsity. Paper [1] introduces an efficient algorithm to find a single sparse principal component with a specified cardinality. The algorithm consists of two stages. In the first stage, it identifies an active index set with desired cardinality corresponding to the nonzero entries of the principal component. In the second one, it finds the best direction with respect to the active index set, using the power iteration method. Experiments on both randomly generated data and real-world data sets show that our algorithm is very fast, especially on large and sparse data sets, while the numerical quality of the solution is comparable to other methods.

Paper [2] extends the approach proposed in [1] to multiple PCA. By combining the algorithm for computing a single sparse PC proposed in [1] with the Schur complement deflation scheme, [1] presents an algorithm which sequentially computes multiple PCs by greedily maximizing the adjusted variance explained by them. Moreover, to address the difficulty of choosing the proper sparsity and parameter in various sparse PCA algorithms, a new PCA formulation is also proposed whose aim is to minimize
the sparsity of the PCs while requiring that their relative adjusted variance is larger than a given prespecified fraction. It is also shown that a slight modification of the aforementioned multiple component PCA algorithm can also find sharp solutions of the latter formulation.

A recent challenge in data analysis for science and engineering is that data are often represented in a structured way. In particular, many data mining tasks have to deal with group-structured prior information, where features or data items are organized into groups. Paper [3] develops group sparsity regularization methods for nonnegative matrix factorization (NMF). NMF is an effective data mining tool that has been widely adopted in text mining, bioinformatics, and clustering, but a principled approach to incorporating group information into NMF has been lacking in the literature. Motivated by an observation that features or data items within a group are expected to share the same sparsity pattern in their latent factor representation, mixednorm regularization is proposed to promote group sparsity in the factor matrices of NMF. Group sparsity improves the interpretation of latent factors. Efficient convex optimization methods for dealing with the mixed-norm term are presented along with computational comparisons between them. Application examples of the proposed method in factor recovery, semi-supervised clustering, and multilingual text analysis are demonstrated.

V. Highlights

Recently, the P.I. has demonstrated in [7] and [8] that the block-decomposition methods studied in [16] are extremely efficient for solving large scale cone programming problems. Computational results are presented in [7] and [8] showing that a cleverly implemented block-decomposition method significantly outperforms the three most competitive codes for solving large-scale conic semidefinite programs, namely: the boundary point method introduced by Povh et al. [6, 19], the efficient alternating direction augmented Lagrangian method by Wen et al. [20], and the Newton-CG augmented Lagrangian method by Zhao et al. [21]. During the last year of this project, we have released two software packages, namely DSA-BD and 2EBD-HPE (see http://www2.isye.gatech.edu/~monteiro/software), which implement block-decompositions algorithms for conic SDPs and extensions. More specifically, DSA-BD expects as input any conic SDP in standard form and hence is a general purpose code for solving conic SDPs. On the other hand, 2EBD-HPE is a specialized code which takes advantage of the classes of conic SDPs for which it is intended to solve, namely: SDPs for $\theta$-functions and $\theta_+$-functions of graph stable set problems,
and SDP relaxations of binary integer quadratic and frequency assignment problems. It should be noted though that the idea behind the efficient implementation of 2EBD-HPE can be extended to many other classes of conic SDPs, and more generally, convex optimization problems (see [8] for more details).

Paper [4] establishes for the first time the iteration-complexity of a first-order version of the augmented Lagrangian (AL) method. It is well-known that the AL method can be very effective to solve large scale convex and/or nonconvex optimization problems. For example, the low-rank method developed by the P.I. and Sam Burer (see http://www2.isye.gatech.edu/~monteiro/software) is one of the most efficient codes for solving large scale SDPs in which the primal matrix variable $X$ is extremely large. The backbone of this code is a clever implementation of the AL method applied to a nonlinear programming reformulation of the original SDP. The latter code is being constantly revised by Sam Burer and the P.I. and is always being used by many practitioners in the areas of science and engineering.

Finally, the PI has been doing quite novel work in the design of first-order methods for VIs (which is the equivalent of second-order methods for optimization). Combining ideas from interior-point theory and proximal point methodology, paper [10] extends the theory of self-concordant barrier methods to a broader class of convex optimization than that considered by Nemirovski and Nesterov in their landmark work (see for example [18]). The PI expects to extend the ideas of [10] to develop second-order methods in the context of convex optimization where more structure is present, and hence better convergence rate results are expected.
References


Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0205, USA, July 2012. Submitted to *SIAM Journal on Optimization*.


