



Inside This Issue

- 1 Message from the Chair
- 1 Editor's Message
- 2 ICS Officers and Board
- 3 ICS-2013 Summary

Features

- 3 Message from JoC Editor
- 4 Sharda Received Harvey J. Greenberg Award
- 4 2012 ICS Prize
- 5 2012 ICS Student Prize
- 5 Members in the News

Articles

- 6 Highlights: Computational and Algebraic Aspects of Convexity
- 9 Highlights: Simulation Selection with Unknown Correlation Structures
- 13 Acknowledgments
- 13 Copyright notice

"It is cruel and unforgettable for not getting a chance to say goodbyes to closed ones ..."



Message from the Chair

William Cook
Combinatorics and Optimization
University of Waterloo
bico@uwaterloo.ca

Let me begin with a big thanks to Bill Hart and Jean-Paul Watson for putting together a beautiful ICS 2013 conference in Santa Fe. The venue was great, including the fabulous breakfasts sponsored by GAMS, Gurobi, AIMMS and AMPL. The program was highlighted by plenary talks delivered by Chris Beck, Jonathan Eckstein, and Mike Trick – if you are interested in learning the role the pope has played in scheduling professional baseball, Mike is the person to ask. A new item at ICS 2013 was the inclusion of advanced tutorials, given by Jeff Linderoth, Warren Powell, and Ted Ralphs. The three sessions were great and stimulated lots of discussion. Tutorials may well be on the menu in future years. The full program for ICS 2013 can be found on the conference Web page <https://www.informs.org/Community/Conferences/ICS2013>.

Speaking of future years, Ted Ralphs has been busy lining up potential sites and organizers for ICS 2015. It is not too late to lend a hand (or venue). If you have ideas for our follow up to Santa Fe, Ted would be delighted to hear from you.

I hope many of you will be able to attend our business meeting at the INFORMS Annual Meeting in Minneapolis. We will try to schedule the meeting in the standard Monday evening slot. The business meeting will feature the presentation of the ICS Prize and the ICS Student Paper Award. Beer, wine, soft drinks, and snacks will be served in our usual generous portions, with sufficient quantities to satisfy even the thirst of Jeff Linderoth.

A main item for discussion at the business meeting is the selection of subjects for our initial array of special-interest groups. The idea is to have an elected Vice Chair for each group who will take the primary responsibility for organizing clusters of sections at the INFORMS Annual Meetings. We have had several preliminary discussions at previous meetings, led by Warren Powell. This year we would like to iron out the details to put together a formal proposal to the society.

See you in Minneapolis.



Message from the Editor

Yongpei Guan
Industrial and Systems Engineering
University of Florida
guan@ise.ufl.edu

It is the time to share the news for the society again. In this letter, please be aware of the updates of the board of directors, ICS 2013 summary, the new data policy for IJOC, and the highlights and insights for the ICS awarding papers (special thanks to Amir Ali Ahmadi and Ilya Ryzhov for the contribution).

Officers

Chair:

William Cook
University of Waterloo
bico@uwaterloo.ca

Vice-Chair/Chair Elect:

Ted Ralphs
Lehigh University
ted@lehigh.edu

Secretary/Treasurer:

Simge Küçükyavuz
Ohio State University
kucukyavuz.2@osu.edu

Board of Directors

Richard Barr (-2013)
Southern Methodist University
barr@smu.edu

Cindy Phillips (-2013)
Sandia National Labs
caphill@sandia.gov

Dorit Hochbaum (-2014)
University of California Berkeley
hochbaum@ieor.berkeley.edu

Jill Hardin Wilson (-2014)
Northwestern University
jill.wilson@northwestern.edu

Karen Aardal (-2015)
Delft Institute
K.I.Aardal@tudelft.nl

Jeff Linderorth (-2015)
University of Wisconsin-Madison
linderorth@wisc.edu

Editors

Journal on Computing:

David Woodruff
University of California, Davis
editor_joc@mail.informs.org

ICS News:

Yongpei Guan
University of Florida
guan@ise.ufl.edu



SIAM Conference on Optimization

Miguel F. Anjos¹ and Michael Jeremy Todd²

¹ Polytechnique Montreal, Canada

² Cornell University, USA

The SIAM Conference on Optimization (OP14) will feature the latest research in theory, algorithms, software and applications in optimization problems. A particular emphasis will be put on applications of optimization in health care, biology, finance, aeronautics, control, operations research, and other areas of science and engineering. The conference brings together mathematicians, operations researchers, computer scientists, engineers, software developers and practitioners, thus providing an ideal environment to share new ideas and important problems among specialists and users of optimization in academia, government, and industry.

Location: Town and Country Resort & Convention Center, San Diego, CA

Dates: May 19-22, 2014

Organizing Committee Co-chairs:

Miguel F. Anjos, Polytechnique Montreal, Canada

Michael Jeremy Todd, Cornell University, USA

Organizing Committee:

Aharon Ben-Tal, Technion - Israel Institute of Technology, Israel

Andrew Conn, IBM Research, USA

Mirjam Dür, University of Trier, Germany

Michael Hintermüller, Humboldt-Universität zu Berlin, Germany

Etienne de Klerk, Nanyang University of Technology, Singapore

Jon Lee, University of Michigan, USA

Todd Munson, Argonne National Laboratory, USA

Warren Powell, Princeton University, USA

Daniel Ralph, University of Cambridge, United Kingdom

Ariela Sofer, George Mason University, USA

Akiko Yoshise, University of Tsukuba, Japan

The Call for Presentations is available at: <http://www.siam.org/meetings/op14/>

SUBMISSION DEADLINES

October 21, 2013: Minisymposium proposals

November 18, 2013: Abstracts for contributed and minisymposium speakers

TRAVEL FUND APPLICATION DEADLINE

November 4, 2013: SIAM Student Travel Award and Post-doc/Early Career Travel Award Applications

PLENARY SPEAKERS AND TWO MINITUTORIALS

<http://www.siam.org/meetings/op14/invited.php>

<http://www.siam.org/meetings/op14/mini.php>

For additional information, contact the SIAM Conference Department (meetings@siam.org).



INFORMS Journal on Computing News

David Woodruff
University of California, Davis
joc@mail.informs.org

New Data Policy

Attaching research data, such as instance data for experiments, to the electronically published versions of papers can be an important part of IJOC's role in the research community.

Associate Editor

When an Associate Editor recommends something other than rejection for a first submission of a paper they should also consider the data, if any, used in the paper. Here are the possibilities:

1. There is little or no meaningful data (e.g., the paper concerns theory of computing)
2. The data are readily available (e.g., instance data from MIPLIP, TSPLIB, etc.)
3. The computational experiments and research data are important, but not central to the contribution
4. Data are central to the contribution

In the case of 3, the AE should recommend to the authors that they consider publishing their data as an online supplement to the paper. Here is suggested wording:

"Your paper reports on computational experiments and the data used in these experiments would be helpful for subsequent researchers who wish to extend and cite your work. If the paper is ultimately accepted for publication at IJOC, we hope that you will provide the research data as an online supplement to the paper."

In the case of 4, the AE should recommend to the Area Editor that publication of the paper should be conditional on publication of the data.

Area Editor

If they concur that the data should be required, the Area Editor should include the requirement in their letter to the authors. From that point on, the EiC will track the requirement and negotiate with the authors if that is needed. Here is suggested language for the Area Editor in this case:

"Your paper relies on computational experiments and the data used in these experiments would be critical for subsequent

researchers who wish to use and cite your work. They are also a critical part of the research that is described. If the paper is ultimately accepted for publication at IJOC, you will need to provide the research data as an online supplement to the paper. If you have questions or concerns about publishing the data, please let me or the Editor in Chief know."

Discussion

Data availability impacts the role of IJOC in the research community. At IJOC it will be up to the Associate and Area Editors, in consultation with the authors and referees, to decide if the data need to be included as an online supplement. The EiC and editorial staff will handle logistics and "enforcement." When published with a paper, the data will appear in the same way as online supplements, which appear as a link next to the link to the paper. In some cases, the data may already be hosted on a well-maintained site so the online supplement will simply contain a link to the site.

Of course, everything is negotiable. Some authors have good reasons to not want to publish their data and editors may agree that their paper is strong enough without the data. However, many papers published in IJOC should make their data easily accessible for the benefit of the research community. Benefit also accrues to the authors because their research will be more highly cited and impactful.



ICS-2013 Summary

William E. Hart and Jean-Paul Watson
Sandia National Laboratories
wehart,jwatson@sandia.gov

The 13th INFORMS Computing Society Conference (ICS) was held in January 2013, at the Eldorado Hotel in Santa Fe, New Mexico. Co-chaired by William E. Hart and Jean-Paul Watson (both at the nearby Sandia National Laboratories), the event attracted 205 registered attendees, who in aggregate presented over 200 technical talks, organized in 7 parallel tracks spanning 3 full days. This year's conference was supported in part by generous contributions from AIMMS, AMPL, GAMS, and Gurobi Optimization. The 13th ICS conference included plenary presentations by Michael Trick (Carnegie Mellon University), Jonathan Eckstein (Rutgers University), and J. Christopher Beck (University of Toronto). In addition to the plenary sessions, the co-chairs introduced the concept of "advanced tutorials" this year. The goal of these new sessions was to provide speakers the opportunity to provide extended (1.5 hour) "deep dive" technical introductions to topics that are of broad interest to the ICS community. Inaugural advanced tutorial sessions, which were heavily attended, included presentations by Warren Powell (Princeton University), Ted Ralphs

(Lehigh University), and Jeff Linderorth (University of Wisconsin-Madison). Finally, despite their best efforts to the contrary (the co-chairs did their best to maximize benefit to the attendees in the form of significant quantities of excellent food), the 13th ICS conference ended with a significant budget surplus. The co-chairs had a wonderful, if not stressful, time organizing the conference, and would like to thank the INFORMS staff and Sandia National Laboratories' Rachel Leyba, without which the event simply would not have transpired. The co-chairs are also happy to assist potential organizers for the 14th ICS conference in any way possible. Please let them know if you have any questions or would like advice on organizing the 14th ICS conference!



2013 Harvey Greenberg Service Award Goes to Sharda

Ramesh Sharda is currently Director of the Institute for Research in Information Systems (IRIS), ConocoPhillips Chair of Management of Technology, and a Regents Professor of Management Science and Information Systems in the Spears School of Business at Oklahoma State University. He received his B. Eng. degree from University of Udaipur, M.S. from The Ohio State University and an MBA and Ph. D. from the University of Wisconsin-Madison.

Ramesh's research on the comparison of linear and mixed integer programming software on personal computers in the mid 1980s led him join the Computer Science Technical Section (CSTS), the predecessor of the ICS. By 1988, he had become the General Chairperson for the first CSTS meeting held in Williamsburg, Virginia (this was the second CSTS meeting; the first was in Denver).

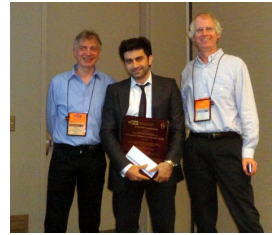
This 1989 meeting attracted nearly 160 OR/CS attendees. George Dantzig gave the plenary lecture. Ramesh co-edited the conference volume. Everything about the meeting was first rate, from the strong scientific program to the high quality of the conference venue. The success of this meeting created a buzz among the CSTS membership. Ramesh set the bar high for (and ensured that there would be) future meetings.

In 1992, Ramesh served as Program Co-chair of the third CSTS meeting and he again co-edited the conference volume. In addition, Ramesh was a member of the Organizing Committee for the 1994 CSTS meeting. In the early years (late 1980s to early 1990s), Ramesh was the key player in putting the CSTS meetings on the conference map. He, essentially, created the blueprint for successful ICS meetings.

Ramesh has contributed in numerous other ways, as well.

He served as Chair of CSTS and a Member of the ICS Board of Directors and the ICS Prize Committee. He has been a long-time Associate Editor of the INFORMS Journal on Computing. In the early 1990s, he launched the OR/CS Interfaces Series with Kluwer Academic Publishers. This book series has published more than 50 titles.

For nearly 25 years, Ramesh Sharda has helped govern, shape, and promote CSTS and ICS. We thank Ramesh for his extensive service to the Society by awarding him the 2013 Harvey J. Greenberg Service Award.



2012 ICS Prize Goes to Ahmadi, Olshevsky, Parrilo, and Tsitsiklis from MIT

The winners of the 2012 ICS Prize for the best English language paper dealing with the Operations Research/Computer Science interface are A.A. Ahmadi, A. Olshevsky, P. A. Parrilo, and J. N. Tsitsiklis for their papers

1. A. A. Ahmadi, A. Olshevsky, P. A. Parrilo, and J. N. Tsitsiklis. NP-hardness of deciding convexity of quartic polynomials and related problems. *Mathematical Programming*, 2011. Online version available at <http://arxiv.org/abs/1012.1908>,
2. A. A. Ahmadi and P. A. Parrilo. A convex polynomial that is not sos-convex. *Mathematical Programming*, 2011. Online version available at <http://arxiv.org/abs/0903.1287>, and
3. A. A. Ahmadi and P. A. Parrilo. A complete characterization of the gap between convexity and sos-convexity. *SIAM Journal on Optimization*, 2012. To appear. Online version available at <http://arxiv.org/abs/1111.4587>.

A common approach to solving optimization problems is to leverage convexity; linear and convex quadratic programming provide classical examples of polynomially solvable problems. The awarded papers explore the complexity frontier of optimization problems and its relationship to convexity.

In paper [1], the authors show that the problem of deciding whether a 4-degree polynomial is convex to be NP-hard in the strong sense. The implication of this result is that unless $P=NP$, there cannot be a polynomial time or even pseudo-polynomial time algorithm for checking polynomial convexity. Although in many applications of convex optimization we design problems that are by construction convex, the result suggests that in general we cannot characterize convexity in optimization problems. These hardness results are ex-

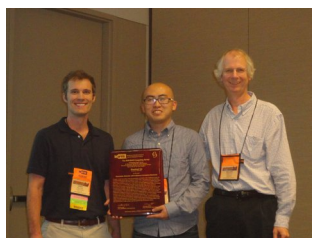
tended to the respective problems of deciding strict convexity, strong convexity, pseudoconvexity, and quasiconvexity of polynomials. Each of these well-known variants of convexity have their own special role in optimization theory. For example, strict convexity is useful for guaranteeing uniqueness of optimal solutions, strong convexity is a common assumption in convergence analysis of many iterative Newton-type algorithms, and quasiconvexity appears in the problem of deciding convexity of sets, and in many applications in economics and statistics. An interesting dichotomy here is that quasiconvexity and pseudoconvexity of odd degree polynomials can be decided in polynomial time, whereas the same questions for polynomials of even degree larger than two are strongly NP-hard.

Paper [2] shows the first known example of a convex polynomial that is not sos-convex. Such polynomials are generally not easy to construct. The authors resort to computational methods, involving semi-definite programming, to find their polynomial. Their computer assisted proof demonstrates the power of sum of squares certificates and SDP in automated theorem proving.

In [3], the authors give a complete characterization of all the degrees and dimensions for which convexity and sos-convexity are equivalent.

Readers are referred to the article “Highlights: Computational and Algebraic Aspects of Convexity” in this newsletter for further highlights and description.

2012 ICS Prize Committee: Daniel Bienstock, Pascal Van Hentenryck (chair), and Dorit Hochbaum



2012 ICS Best Student Paper Award Goes to Qu at U. of Maryland

The 2012 Student Paper Award Winner is Huashuai Qu (University of Maryland) for the paper, “Simulation Selection with Unknown Correlation Structures.” His advisors are Professors Michael Fu and Ilya Ryzhov.

This paper considers the problem of Bayesian optimization via simulation, with correlated prior beliefs and correlated sampling with an unknown sampling covariance matrix. This problem arises when performing optimization via simulation with common random numbers, and is important because sampling with common random numbers has the potential to allow better efficiency than does independent sampling. Analysis of this problem, however, is substantially more difficult than with independent sampling as there is no conjugate prior distribution permitting sequential sampling, making compu-

tation of the posterior distribution computationally challenging. This paper deftly steps around this difficulty by using an approximation based on minimizing the Kullback-Leibler divergence, which provides a computationally tractable approximate posterior distribution. Then, using this statistical technique as a foundation, this paper develops a new value of information sampling procedure that allows unknown correlation structures. This procedure has better performance than existing procedures on several problems, and it shows that modeling the unknown sampling covariance matrix can have a significant effect on the value of information. This work has broader implications for other problems in simulation optimization, and more broadly in sequential experimental design: it provides an appealing methodology for approximating posterior distributions in other sequential sampling problems; and it paves the way for unknown covariance matrices to be modeled explicitly, rather than assumed known, in other problems requiring sequential value-of-information analysis.

Readers are referred to the article “Highlights: Simulation Selection with Unknown Correlation Structures” in this newsletter for further highlights and description.



2012 ICS Best Student Paper Runner-up Goes to Takáč at U. of Edinburgh

The 2012 ICS Best Student Paper Runner-up is Martin Takáč, University of Edinburgh, for the paper “Iteration Complexity of Randomized Block-Coordinate Descent Methods for Minimizing a Composite Function.” His advisor is Peter Richtarik.

2012 ICS Student Paper Award Committee: Andreas S. Schulz (chair), Peter Frazier, and Cynthia A. Phillips

ICS Members in the News

James B Orlin(jorlin@mit.edu), Ph.D., Edward Pennell Brooks Professor of Operations Research at the Sloan School of Management, Massachusetts Institute of Technology, was awarded 2013 Harold Larnder Prize (<http://www.cors.ca/en/prizes/>). James B. Orlin is best known for his research on obtaining faster algorithms for problems in network and combinatorial optimization and for his text with Ravi Ahuja and Tom Magnanti entitled *Network Flows: Theory, Algorithms, and Applications*. The authors won the 1993 Lanchester Prize (given by INFORMS for the best publication in O.R. for the year) for this book. He has also won recognition for several co-authored publications that address (in one form or another) optimization under uncertainty. In particular, he has won the following

awards: the 2004 EXPLOR Award (for leadership in online marketing research), the 2007 INFORMS Computing Society Prize (for research in the interface of O.R. and computer science), the 2008 IEEE Leonard G. Abraham Prize (for research in communication theory), the 2008 INFORMS Koopman Prize (for research in military operations research), and the 2011 IEEE Bennett Prize (for research in communication theory).

The Harold Larnder Prize is awarded annually to an individual who has achieved international distinction in operational research. The prize winner delivers the Harold Larnder Memorial Lecture, on a topic of general interest to operational researchers, at the National Conference of the Canadian Operational Research Society.

Harold Larnder was a well-known Canadian in wartime OR. He played a major part in the development of an effective, radar-based, air defence system during the Battle of Britain. He returned to Canada in 1951 to join the Canadian Defence Research Board and was President of CORS in 1966-67.

Panos Pardalos (pardalos@ufl.edu), Ph.D., Distinguished Professor at the University of Florida, was awarded the 2013 EURO Gold Medal prize (<http://www.euro-online.org/web/pages/212/gold-medal-egm>), bestowed by the Association for European Operational Research Societies. This medal is the preeminent European award given to Operations Research (OR) professionals. The award was presented at the opening session of the EURO XXVI Conference in Rome on July 1st. The EURO association website (<http://www.euro-online.org/web/pages/605/announcements>) describes the criteria used to select Gold Medal prize winners.

“The EURO Gold Medal is awarded for a body of work in operational research, preferably published over a period of several years. Although recent work is not excluded, care should be taken to allow the contribution to stand the test of time. The potential prize recipient should have a recognized stature in the European OR community. Significance, innovation, depth, and scientific excellence should be stressed.”

Anna Nagurney (nagurney@isenberg.umass.edu), Ph.D., the John F. Smith Memorial Professor at the Isenberg School of Management at UMass Amherst, has been on sabbatical the 2012-2013 year. During this period she has been a Visiting Professor of Operations Management at the School of Business, Economics and Law at the University of Gothenburg, Sweden and also a Guest Professor at the Vienna University of Economics and Business in Austria.

She received the 2012 Walter Isard Award in Ottawa, Canada in November at the Annual North American Meetings of the Regional Science Association International. The Walter Isard Award is the top research award given by NARSC (North American Regional Science Council). The award was established in 1994 to pay tribute to regional scientists who have made significant theoretical and methodological contributions

to the field of Regional Science throughout their careers. The award is named after the founder of Regional Science, Professor Walter Isard, who passed away at the age of 91 in 2010. Anna was recognized for her sustained contributions to research on network systems.

While on sabbatical, she completed the book, “Networks Against Time: Supply Chain Analytics for Perishable Products,” co-authored with Min Yu, Amir H. Masoumi, and Ladimer S. Nagurney, which was published in 2013 by Springer Science and Business Media, NYC. On April 25, she was an invited panelist on Transport and Traffic at The New York Times Energy for Tomorrow Conference in NYC on Building Sustainable Cities. All panels are now available on video.

She continues to work on the NSF project: Network Innovation Through Choice, with collaborators from UMass Amherst, NCState, the University of Kentucky, and the Renaissance Institute (RENCI) at UNC Chapel Hill. This project was chosen by NSF as one of the five Future Internet Architecture (FIA) projects. She will be giving a plenary talk on this research at the Network Models in Economics and Finance Conference in Athens, Greece, June 13-15, 2013.

Erick Moreno-Centeno (e.moreno@tamu.edu), Ph.D., was named Montague-Center for Teaching Excellence Scholar for 2012-2013 (<http://cte.tamu.edu/content/montague-cte-scholars>). This award recognizes early-career excellence in undergraduate teaching, and is given to only one junior faculty per college; Dr. Moreno-Centeno was the faculty chosen from Texas A&M University’s College of Engineering. Dr. Moreno-Centeno is currently an Assistant Professor in the Industrial and Systems Engineering Department at Texas A&M University.

Arne Løkketangen, from the Department of Informatics, Molde College, died apparently of a heart attack while attending the Tristan Conference in Chile. He was an active member of ICS and an Associate Editor of the INFORMS Journal on Computing. He will be remembered for his numerous contributions to metaheuristic optimization and applications in transportation. He will also be remembered as an active participant at ICS conferences, most recently the ICS meeting in Santa Fe.



Highlights: Computational and Algebraic Aspects of Convexity

Ahmadi, Olshevsky, Parrilo,
and Tsitsiklis – MIT

We would like to take this opportunity to express our gratitude to the 2012 Informatics Computing Society Prize Committee, which was chaired by Pascal Van Hentenryck and included

Daniel Bienstock and Dorit Hochbaum. We are grateful to the committee for this very kind distinction and honored to receive the prize.

The ICS Prize was awarded for our work in [4], [5], [6] on the study of some very basic questions about *convexity*. The first paper [4] studies the computational complexity of recognizing convexity of functions and sets in polynomial optimization. The second and third papers [5], [6] are on an algebraic relaxation for convexity, known as *sum-of-squares-convexity* (sos-convexity), which has links to semidefinite programming and plays a central role in the emerging field of *convex algebraic geometry* [9]. Some of the results, we believe, turned out to be interesting—our complexity paper answered an open question of Naum Shor from 1992; our study of sos-convexity revealed an unforeseen connection with a classical result of Hilbert in real algebraic geometry. We welcome your feedback and hope that you also find some aspects of the work interesting. A brief description of the main contributions follows with more details to be found in references mentioned above or in [1].

1 Complexity of Deciding Convexity

Over the last century, the notion of convexity has established itself as a central concept in the theory of optimization and operations research. Extensive and greatly successful research in the applications of convex optimization has shown that surprisingly many problems of practical importance can be cast as convex optimization problems. Moreover, we have a fair number of rules based on the calculus of convex functions that allow us to design—whenever we have the freedom to do so—problems that are by construction convex. Nevertheless, in order to be able to exploit the potential of convexity in optimization in full, a very basic question is to understand whether we are even able to *recognize* the presence of convexity in optimization problems. In other words, can we have an efficient algorithm that tests whether a given optimization problem is convex?

A class of optimization problems that allows for a rigorous study of this question from a computational complexity viewpoint is the class of polynomial optimization problems. These are optimization problems where the objective is given by a polynomial function and the feasible set is described by finitely many polynomial inequalities. Our research in this direction was motivated by a concrete question of N. Z. Shor that

appeared as one of seven open problems in complexity theory for numerical optimization put together by Pardalos and Vavasis in 1992 [22]:

“Given a degree-4 polynomial in n variables, what is the complexity of determining whether this polynomial describes a convex function?”

The reason why Shor’s question is specifically about degree 4 polynomials is that deciding convexity of odd degree polynomials is trivial¹ and deciding convexity of degree 2 (quadratic) polynomials can be reduced to the simple task of checking whether a constant matrix is positive semidefinite. So, the first interesting case really occurs for degree 4 (quartic) polynomials. The main contribution of our first paper [4] is to show that deciding convexity of polynomials is strongly NP-hard already for polynomials of degree 4.

The implication of strong NP-hardness of this problem is that unless $P=NP$, there exists no algorithm that can take as input the (rational) coefficients of a quartic polynomial, have running time bounded by a polynomial in the numeric value of these coefficients (let alone in the number of bits needed to represent the coefficients), and output correctly on every instance whether or not the polynomial is convex. The reduction that establishes our result is purely algebraic. It shows that the NP-hard problem of testing global nonnegativity of so-called *biquadratic forms* (a special subclass of quartic polynomials) can be turned into the question of checking convexity by doubling the number of variables and without changing the degree.

Although the implications of convexity are very significant in optimization theory, our results suggest that unless additional structure is present, ensuring the mere presence of convexity is likely an intractable task. It is therefore natural to wonder whether there are other properties of optimization problems that share some of the attractive consequences of convexity, but are easier to check.

1.1 Complexity of deciding variants of convexity

In the same paper, we also study the complexity of recognizing some well-known variants of convexity, namely, the problems of deciding *strict convexity*, *strong convexity*, *pseudoconvexity*, and *quasiconvexity* of polynomials. The relationship between these notions is as follows (with none of the converse implications being true in general):

strong convexity \implies strict convexity \implies convexity \implies
pseudoconvexity \implies quasiconvexity.

Strict convexity is a property that is often useful to check because it guarantees uniqueness of the optimal solution in optimization problems. The notion of strong convexity is a

¹Independent of their coefficients, polynomials of odd degree are never (globally) convex, unless they are linear and hence always convex.

Table 1: Summary of our complexity results. A yes (no) entry means that the question is trivial for that particular entry because the answer is always yes (no) independent of the input. By P, we mean that the problem can be solved in polynomial time.

property vs. degree	1	2	odd ≥ 3	even ≥ 4
strong convexity	no	P	no	strongly NP-hard
strict convexity	no	P	no	strongly NP-hard
convexity	yes	P	no	strongly NP-hard
pseudoconvexity	yes	P	P	strongly NP-hard
quasiconvexity	yes	P	P	strongly NP-hard

common assumption in convergence analysis of many iterative Newton-type algorithms in optimization theory; see, e.g., [10, Chaps. 9–11]. So, in order to ensure the theoretical convergence rates promised by many of these algorithms, one needs to first make sure that the objective function is strongly convex. The problem of checking quasiconvexity (convexity of sublevel sets) of polynomials also arises frequently in practice. For instance, if the feasible set of an optimization problem is defined by polynomial inequalities, by certifying quasiconvexity of the defining polynomials we can ensure that the feasible set is convex. In several statistics and clustering problems, we are interested in finding minimum volume convex sets that contain a set of data points in space. This problem can be tackled by searching over the set of quasiconvex polynomials [18]. In economics also, quasiconcave functions are prevalent as desirable utility functions [7]. Finally, the notion of pseudoconvexity is a natural generalization of convexity that inherits many of the attractive properties of convex functions. For example, every stationary point or every local minimum of a pseudoconvex function must be a global minimum. Because of these nice features, pseudoconvex programs have been studied extensively in nonlinear programming [19], [11].

Our complexity results as a function of the degree of the polynomial are listed in Table 1.1. As you can see, all of these properties are easy to decide for quadratics but hard for polynomials of even degree 4 or higher. Somewhat surprisingly, we were able to show that testing quasiconvexity and pseudoconvexity can be done in polynomial time if the degree is odd.

2 Convexity and SOS-Convexity

Of course, NP-hardness of a problem does not stop us from studying it, but on the contrary, stresses the need for finding good approximation algorithms that can deal with a large number of instances efficiently. Towards this goal, we study in [5], [6] a *semidefinite relaxation* for convexity of polynomials known as *sos-convexity*.

Definition 2.1. A polynomial $p(x) = p(x_1, \dots, x_n)$ is *sos-convex* if its Hessian $H(x)$ can be factored as $H(x) = M^T(x)M(x)$ with a possibly nonsquare polynomial matrix² $M(x)$.

It is easy to see that sos-convexity is a sufficient condition for convexity of polynomials. Indeed, if we have the factorization $H(x) = M^T(x)M(x)$, then $H(x)$ must be positive semidefinite for all x . Moreover, one can show that sos-convexity of a polynomial p can be decided by solving a single semidefinite program whose size is polynomial in the size of the coefficients of p ; see e.g. [5].

The idea behind sos-convexity is related to the concept of representing *nonnegative polynomials* as *sums of squares*³—a deep-rooted subject in real algebraic geometry that has found widespread recent applications in optimization theory. Just like a sum of squares (sos) decomposition produces an algebraic certificate for nonnegativity, sos-convexity can be thought of as an algebraic certificate for convexity. In [6, Thm. 3.1], we showed that if one applies the sos relaxation to the standard definition of convexity or its first order characterization (as opposed to Definition 2.1, which appeals to the second order characterization of convexity), then one ends up with conditions that are equivalent to sos-convexity. This is reassuring in that it demonstrates the legitimacy of sos-convexity as *the* right semidefinite relaxation for convexity.

2.1 The first example of a convex polynomial that is not sos-convex

The connection of sos-convexity to semidefinite programming has motivated its use in many application domains, such as statistics (convex regression, minimum volume shape fitting, etc., [18]) and control theory (stability of hybrid systems [3]). Aside from its computational implications, sos-convexity is a concept of interest in the field of *convex algebraic geometry* [9], which is devoted to the study of convex sets with algebraic structure. In particular, one of the early results in this area, due to Helton and Nie [14], states that any subset of \mathbb{R}^n defined as $\{x \mid g_i(x) \leq 0\}$, with each g_i an sos-convex polynomial, can be represented as the projection of the feasible set of a semidefinite program.

Motivated by results of this type, it had been speculated whether sos-convexity of a polynomial was in fact equivalent to its convexity. We showed in [5], via a concrete counterexample (a polynomial in 3 variables and degree 8), that the answer was negative. This example came before our NP-hardness result in [4]. Indeed, complexity considerations alone suggest (assuming $P \neq NP$) that convex but not sos-convex polynomials should exist, at least when the number of variables goes to infinity.

²A polynomial matrix is simply a matrix whose entries are (multivariate) polynomials.

³A polynomial p is nonnegative if $p(x) \geq 0$ for all $x \in \mathbb{R}^n$, and it is a sum of squares if it can be written as $p = \sum q_i^2$, for some polynomials q_i .

For the curious reader, we include here the first example of a convex but not sos-convex polynomial, published in [5]:

$$p(x) = 32x_1^8 + 118x_1^6x_2^2 + 40x_1^6x_3^2 + 25x_1^4x_2^4 - 43x_1^4x_2^2x_3^2 - 35x_1^4x_3^4 + 3x_1^2x_2^4x_3^2 - 16x_1^2x_2^2x_3^4 + 24x_1^2x_3^6 + 16x_2^8 + 44x_2^6x_3^2 + 70x_2^4x_3^4 + 60x_2^2x_3^6 + 30x_3^8.$$

We found this polynomial with the help of a computer and as a solution of a carefully-designed semidefinite program; see [5, Sect. 4] for details. In general, finding polynomials with such properties is a nontrivial task. For example, the following closely related problem is still open.

Open problem.

Find an explicit example of a convex, nonnegative polynomial that is not a sum of squares.

Blekherman [8] has shown via volume arguments that for degree $d \geq 4$ and asymptotically for large n such polynomials must exist, although no examples are known. It is known, however, that such a convex polynomial cannot be sos-convex [14], [6]. The question is particularly interesting from an optimization viewpoint since it implies that the well-known sum of squares relaxation for minimizing polynomials [26], [23] is not always exact, even in the easy case of minimizing convex polynomials.

2.2 A full characterization of cases where convexity equals sos-convexity

One of the cornerstones of real algebraic geometry is Hilbert's seminal paper in 1888 [15], where he gives a complete characterization of the degrees and dimensions for which nonnegative polynomials can be written as sums of squares of polynomials. In particular, Hilbert proves in [15] that there exist nonnegative polynomials that are not sums of squares, although explicit examples of such polynomials appeared only about 80 years later [21] and the study of the gap between nonnegative and sums of squares polynomials continues to be an active area of research to this day.

Once we produced the first example of a convex but not sos-convex polynomial in [5], it was natural to aim for a characterization of the dimensions and degrees for which such polynomials can exist, similar to the characterization that Hilbert provided for nonnegativity and sum of squares.

The contribution of our third paper [6] is to provide such a characterization for the inclusion relationship between convexity and sos-convexity. The results are summarized in Figure 1 and cover both the case of polynomials and forms (homogeneous polynomials). The entry $(n, d) = (3, 4)$ in the table on

the right is particularly challenging and is joint work with G. Blekherman [2].

convex=sos-convex?												
Polynomials						Forms						
n,d	2	4	≥6	n,d	2	4	≥6	n,d	2	4	≥6	n,d
1	yes	yes	yes	1	yes	yes	yes	1	yes	yes	yes	1
2	yes	yes	no	2	yes	yes	yes	2	yes	yes	yes	2
3	yes	no	no	3	yes	yes	no	3	yes	yes	no	3
≥4	yes	no	no	≥4	yes	no	no	≥4	yes	no	no	≥4

Figure 1: The tables characterize whether every convex polynomial (form) in n variables and of degree d is sos-convex [6]. A similar characterization for nonnegativity and sum of squares was done by Hilbert in 1888 [15].

An intriguing overall outcome of this research is that convex polynomials (resp. forms) were shown to be sos-convex *precisely* in the cases where nonnegative polynomials (resp. forms) are sums of squares, as shown by Hilbert. The proofs given in our work certainly do not use or depend on the results of Hilbert and it is not clear to us at this point whether this similarity is just a coincidence. It remains to be seen whether there is a more satisfying explanation of the resemblance of the two results, perhaps revealing a more fundamental connection between two basic concepts, nonnegativity and convexity.



Highlights: Simulation Selection with Unknown Correlation Structures

Qu, Ryzhov, and Fu – U. of Maryland

Simulation selection arises in problems where a decision-maker must choose one of a small number of high-impact alternatives whose values are unknown. For example, we may have a list of candidate locations for a new wind farm, and the problem is to choose the location where the average energy production will be highest. Or we may wish to choose one of several energy storage policies governing when to buy and store energy, and when to sell it back. We may be making staffing decisions for a call center, or choosing an investment strategy or a facility location decision. In all of these cases, we may have some prior information about the quality of each alternative, but there is still considerable uncertainty as to which is the best.

In many such applications, the decision-maker may use discrete-event or stochastic simulation to evaluate the performance of one or more alternatives before committing to a final selection or implementation. For example, Monte Carlo sim-

ulation may be used to estimate the energy production of a particular configuration of wind turbines [20]. The output of a stochastic simulation is uncertain, but repeated simulations will produce better estimates. The problem is that the total simulation budget is limited, by either money or time constraints. Using a simulation run to examine one wind farm location means having one less run to use on other locations.

As a result, it becomes important to allocate simulations efficiently. We would like to simulate alternatives with high potential to provide useful information about the problem and improve the quality of our final selection. Good simulation choices typically strike a balance between *exploration* (learning about something new) and *exploitation* (learning about something that seems to be good). On one hand, there is no sense simulating an obviously poor choice. On the other hand, if the same choice consistently appears to be the best after several simulations, but we have rarely tried the second-best choice, we may wish to revisit the latter on the chance that it is better than we think. The problem of balancing exploration and exploitation is called *optimal learning* [24]: we learn more about different alternatives as we simulate them, but information itself is a critical resource and must be collected in an optimal manner.

In simulation selection, a common approach [16] has been to model the alternatives independently. That is, we keep a set of sufficient statistics representing the estimated performance of an alternative (e.g. the mean and standard deviation of all simulations conducted thus far), and update them when a new simulation of that alternative is conducted. If we do not simulate an alternative, we do not update our estimate of it. While this learning model may work for a small number of alternatives (or a large number of simulations), it is less effective when the simulation budget is small, since we may end up with a large number of alternatives for which no information is available. Furthermore, in practice, we are unlikely to think of the alternatives as being independent. It is more likely that there are certain similarities between some alternatives. All else being equal, a call center with 10 representatives may be expected to perform similarly to one with 11, or at least, the similarity should be greater than for 10 versus 1. If two alternatives are similar, simulating one of them should provide some information about the other, thus increasing the power of a single simulation and reducing the need to simulate every alternative a large number of times.

Unfortunately, these similarities are themselves quite difficult to quantify. We might expect two wind farm locations to have similar energy production if they are closer together, but they may also have different elevation or other physical features. In ranking and selection, some recent algorithms are able to exploit similarities between alternatives, they assume that these similarities are correctly specified. This can become a serious issue: if we believe that alternatives x and y are very similar, and x seems to perform well, we will also expect y to

perform well. However, if they are in fact completely different (think negative vs. positive correlation in a probability distribution), a good result for x should lead us to expect a bad result for y . In this case, our incorrect beliefs will have led to misleading information about y . Our work addresses this problem by developing the first Bayesian statistical model for ranking and selection to treat the similarities (or “correlations”) between alternatives as unknown, and to learn these correlations from individual simulations.

The model works as follows. Suppose that there are K alternatives (e.g. candidate locations for our wind farm). Let Y be a random K -vector following a multivariate normal distribution with mean vector μ and covariance matrix Σ , with $R = \Sigma^{-1}$ being the relevant *precision matrix*. Think of Y as being the energy production that we would observe at all the locations on a given day, if we had the ability to simultaneously place a wind farm in every location. The mean vector μ represents the (unknown) average energy production at every location; our goal is to select the location x with the highest μ_x . The precision matrix R captures the similarities and differences between different locations. Crucially, our model treats R as unknown.

We adopt the Bayesian view, in which any unknown quantity is modeled as a random variable whose probability distribution represents our own uncertainty about the possible values of the quantity. Thus, we treat R as a random matrix following the Wishart distribution, parameterized by a scalar b and a matrix B that we have to specify. This distribution is commonly used in Bayesian statistics to model unknown correlations. The mean vector μ is also unknown and random; given R , the *conditional* distribution of μ is multivariate normal with mean vector θ and precision matrix qR , where θ and q are input parameters. Together, the pair (μ, R) is said to follow a normal-Wishart distribution with user-specified parameters (q, b, θ, B) .

The intuition for these parameters will become clear after the following example. Suppose that (μ, R) is normal-Wishart as above, and we then observe a sample $\hat{Y} \sim \mathcal{N}(\mu, R)$. The conditional distribution of (μ, R) , given \hat{Y} , is still normal-Wishart with parameters

$$q' = q + 1, \quad (1)$$

$$b' = b + 1, \quad (2)$$

$$\theta' = \frac{q\theta + \hat{Y}}{q + 1}, \quad (3)$$

$$B' = B + \frac{q}{q + 1} (\theta - \hat{Y})(\theta - \hat{Y})^T. \quad (4)$$

Notice that, since the posterior distribution stays within the normal-Wishart family, we can repeat this process and use the above equations to recursively update the distribution parameters after each new sample [12]. The parameters q and b behave analogously to sample sizes, while θ is analogous a sample mean of the observations, and B functions as a generalized

“sum of squares” matrix. Indeed, $\mathbb{E}(\Sigma) = \frac{1}{b-(K-1)}B$ is analogous to the empirical covariance matrix.

However, in simulation selection, we are not able to observe the entire random vector \hat{Y} . Rather, we choose a single alternative x and observe the marginal distribution $\hat{y}_x \sim \mathcal{N}(\mu_x, \Sigma_{xx})$. Unfortunately, the posterior distribution of (μ, R) given \hat{y}_x will no longer be normal-Wishart. This leads to a dilemma: we would like to use the Wishart distribution due to its ease of use for modeling unknown covariances, but we need a way to maintain the normal-Wishart distribution after each simulation observation, so that our beliefs can always be represented by a small number of parameters.

We solve this problem by using the technique of density projection. Let $\xi(\mu, R)$ be a normal-Wishart density with parameters (q', b', θ', B') , and let $f(\mu, R | \hat{y}_x)$ be the actual, non-normal-Wishart posterior density of (μ, R) given \hat{y}_x . Then, we choose (q', b', θ', B') to minimize the Kullback-Leibler divergence

$$D_{KL}(\xi \| f) = \mathbb{E}_\xi \left(\log \frac{\xi(\mu, R)}{f(\mu, R | \hat{y}_x)} \right),$$

and replace the posterior density by ξ with the optimal parameters. Essentially, D_{KL} measures the “distance” between two probability distributions. We create an artificial normal-Wishart density with the smallest possible distance from the actual, non-normal-Wishart posterior. This artificial density ξ can be viewed as the best possible approximation of f taken from the normal-Wishart family. We then simply replace f by this approximation, and proceed as if ξ were the actual posterior. It turns out that the parameters (q', b', θ', B') of the optimal approximation can be expressed in closed form, as

$$q' = q + \frac{1}{K}, \quad (5)$$

$$b' = b + \Delta b, \quad (6)$$

$$\theta' = \theta + \frac{\hat{y}_x - \theta_x}{\frac{qb'}{b'-(K-1)}B_{xx} + B_{xx}}, \quad (7)$$

$$B' = \frac{b'}{b}B + \frac{b'}{b+1} \left(\frac{q(\hat{y}_x - \theta_x)^2}{\frac{qb'}{b'-(K-1)} + 1} - \frac{B_{xx}}{b} \right) \frac{B_{xx}B_{xx}^T}{B_{xx}^2}. \quad (8)$$

This leads to a very fast approximation procedure. Now, our beliefs about (μ, R) are always characterized by four parameters (q, b, θ, B) , and every time we simulate an alternative x , equations (5-8) provide a quick and easy update. The updating scheme has two properties that are crucial for our purposes. First, in (7), a scalar observation \hat{y}_x is used to update the entire vector of beliefs through the matrix B . Second, in (8), the scalar squared deviation $(\hat{y}_x - \theta_x)^2$ is used as a stand-in for the matrix sum of squares, to learn about the entire correlation structure. Note also that, in (5), we increment q by $\frac{1}{K}$, suggesting that the scalar observation contributes a fraction $\frac{1}{K}$ of the information contributed by a complete observation in (1). The increment Δb has to be computed numerically via a bisection procedure, but typically also takes values close to $\frac{1}{K}$.

In this way, our statistical model is able to simultaneously learn unknown means and unknown correlations from scalar observations. It remains to address the question of how unknown correlations can be leveraged to make simulation allocation decisions, or in other words, how to choose which x to observe. Our approach uses an expected improvement criterion

$$x^* = \arg \max_x \mathbb{E}_x \left(\max_y \theta'_y - \max_y \theta_y \right), \quad (9)$$

where \mathbb{E}_x denotes an expectation given the decision to simulate x . Since we are trying to identify the alternative with the highest mean performance, the right-hand side of (9) can be viewed as the expected improvement obtained in our estimate of the best value as a result of simulating x . Expected improvement, also known as the “value of information” or “knowledge gradient,” has emerged as a powerful methodology for simulation selection [see 24, for a comprehensive survey], and is also widely used by the global optimization community [17]. In this context, under the assumption that θ'_y is calculated using (7), we can rewrite (9) in the equivalent form

$$x^* = \arg \max_x \mathbb{E}_x \left[\max_y \left(\theta_y + \tilde{s}_y(x) T \right) - \max_y \theta_y \right], \quad (10)$$

where $\tilde{s}(x)$ is a particular vector derived from the covariance parameter B , and T follows a scalar Student’s t -distribution with $b - (K - 1)$ degrees of freedom. Thus, although the entire vector θ changes after a single simulation, the randomness in that change still comes from a scalar quantity, represented by T . Just as in classical statistics, the normal distribution of \hat{y} is replaced by a t -distribution when the variance is unknown. Note that (10) takes the expectation of a piecewise linear function of T ; we can compute this expectation exactly using standard techniques [see 24].

Our goal throughout this work is to create a model that will learn unknown performance values and correlation structures quickly, as well as a practical algorithm that will calculate the information potential of the unknown correlation structure. A major advantage of our model and algorithm is their computational efficiency. In fact, the simulation policy in (10) requires the same computational cost as a value of information procedure that assumes *known* correlation structures. Likewise, the computational complexity of the updating equations in (5)-(8) is $O(K^2)$, the same as for a model with a known correlation structure. But speed, by itself, does not guarantee good performance, and so we briefly consider a numerical example.

Figure 2 considers a stylized version of the wind farm placement problem, with 64 candidate locations placed on an 8×8 grid. For the purposes of this example, we use wind speed as a stand-in for energy production, allowing us to create a realistic example using historical wind speed data. The upper-left section of the figure is a colour map showing the true average wind speeds in this region (historical averages over a large amount of data). Red indicates higher speeds, so the upper-left

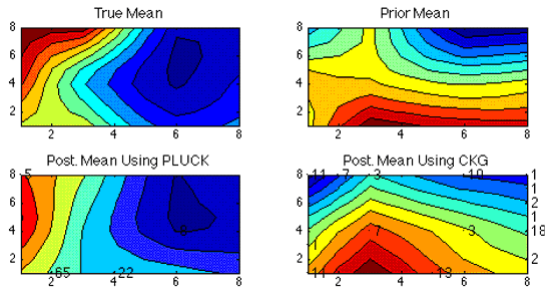


Figure 2: Illustration of the value of learning unknown correlation structures [25].

location (1,8) is the “best.” However, we do not know this; the upper-right section shows “prior averages” taken over a small sample of historical data, meant to represent a decision-maker with limited knowledge of the problem. According to this prior sample, locations in the bottom-center of the region are the best.

The lower half of the figure shows the posterior estimates of the different wind speeds obtained by following two procedures for 100 simulations. The PLUCK procedure (Projected Learning of Unknown Correlations with Knowledge gradients) uses (10) to make simulation decisions and (5)-(8) to update the beliefs. The CKG procedure of [13] uses a similar expected improvement criterion to make simulation decisions, but assumes a known correlation structure. Thus, the difference between the two procedures can be viewed as the value of incorporating unknown correlations into the expected improvement logic. The other numbers on the two plots represent the number of simulations allocated to different locations on the grid. We see that CKG improves over the initial beliefs: we no longer think that the bottom-right area is promising, and the red part of the graph has generally shifted left compared to the prior. However, the shift is relatively slow, and the policy needs to sample numerous locations on the grid, including relatively irrelevant locations in the upper-right. By contrast, the PLUCK policy has already settled on the upper-left region, and has focused much more on a few key locations during simulation.

We offer a few closing thoughts. By using correlations to extract more information from each individual simulation, we can greatly extend the power of a small simulation budget: if one simulation provides information about the entire set of alternatives, we will need fewer simulations in total to identify the best. In some applications, standard assumptions on the correlation structure can lead to good results; for instance, algorithms in global optimization can work very well with simple distance-based covariances. However, our work provides the insight that, when our beliefs about the correlations are wrong, this can lead to incorrect or misleading results. In these situations, we can obtain much better performance by

including the correlations among the unknown parameters to be learned. Because the fairly small number of standard learning models limits the kinds of problems that we can handle, we propose to move beyond these models and create computationally efficient *approximate* learning schemes that incorporate new information with an optimal degree of accuracy (using a metric such as the KL divergence). We hope that the ideas in our paper may be useful for addressing other complex learning problems.

References

- [1] A. A. Ahmadi. *Algebraic relaxations and hardness results in polynomial optimization and Lyapunov analysis*. PhD thesis, Massachusetts Institute of Technology, September 2011. Available at <http://aaa.lids.mit.edu/publications>.
- [2] A. A. Ahmadi, G. Blekherman, and P. A. Parrilo. Convex ternary quartics are sos-convex. In preparation, 2013.
- [3] A. A. Ahmadi and R. Jungers. SOS-convex Lyapunov functions with applications to nonlinear switched systems. Submitted to the *2013 IEEE Conference on Decision and Control*, 2013.
- [4] A. A. Ahmadi, A. Olshevsky, P. A. Parrilo, and J. N. Tsitsiklis. NP-hardness of deciding convexity of quartic polynomials and related problems. *Mathematical Programming*, 137(1-2):453–476, 2013.
- [5] A. A. Ahmadi and P. A. Parrilo. A convex polynomial that is not sos-convex. *Mathematical Programming*, 135(1-2):275–292, 2012.
- [6] A. A. Ahmadi and P. A. Parrilo. A complete characterization of the gap between convexity and sos-convexity. *SIAM Journal on Optimization*, 23(2):811–833, 2013.
- [7] K. J. Arrow and A. C. Enthoven. Quasi-concave programming. *Econometrica*, 29(4):779–800, 1961.
- [8] G. Blekherman. Convex forms that are not sums of squares. arXiv:0910.0656, 2009.
- [9] G. Blekherman, P. A. Parrilo, and R. Thomas. *Semidefinite optimization and convex algebraic geometry*. SIAM Series on Optimization, 2013.
- [10] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.
- [11] R. W. Cottle and J. A. Ferland. On pseudo-convex functions of nonnegative variables. *Math. Programming*, 1(1):95–101, 1971.
- [12] M. H. DeGroot. *Optimal Statistical Decisions*. John Wiley and Sons, 1970.
- [13] P. I. Frazier, W. B. Powell, and S. Dayanik. The knowledge-gradient policy for correlated normal rewards. *INFORMS Journal on Computing*, 21(4):599–613, 2009.
- [14] J. W. Helton and J. Nie. Semidefinite representation of

convex sets. *Mathematical Programming*, 122(1, Ser. A):21–64, 2010.

- [15] D. Hilbert. Über die Darstellung Definiten Formen als Summe von Formenquadraten. *Math. Ann.*, 32, 1888.
- [16] L. J. Hong and B. L. Nelson. A Brief Introduction To Optimization Via Simulation. In M. Rosetti, R. Hill, B. Johansson, A. Dunkin, and R. Ingalls, editors, *Proceedings of the 2009 Winter Simulation Conference*, pages 75–85, 2009.
- [17] D. R. Jones, M. Schonlau, and W. J. Welch. Efficient global optimization of expensive black-box functions. *Journal of Global Optimization*, 13(4):455–492, 1998.
- [18] A. Magnani, S. Lall, and S. Boyd. Tractable fitting with convex polynomials via sum of squares. In *Proceedings of the 44th IEEE Conference on Decision and Control*, 2005.
- [19] O. L. Mangasarian. Pseudo-convex functions. *J. Soc. Indust. Appl. Math. Ser. A Control*, 3:281–290, 1965.
- [20] G. Marmidis, S. Lazarou, and E. Pyrgioti. Optimal placement of wind turbines in a wind park using Monte Carlo simulation. *Renewable Energy*, 33(7):1455–1460, 2008.
- [21] T. S. Motzkin. The arithmetic-geometric inequality. In *Inequalities (Proc. Sympos. Wright-Patterson Air Force Base, Ohio, 1965)*, pages 205–224. Academic Press, New York, 1967.
- [22] P. M. Pardalos and S. A. Vavasis. Open questions in complexity theory for numerical optimization. *Mathematical Programming*, 57(2):337–339, 1992.
- [23] P. A. Parrilo and B. Sturmfels. Minimizing polynomial functions. *Algorithmic and Quantitative Real Algebraic Geometry, DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, 60:83–99, 2003.
- [24] W. B. Powell and I. O. Ryzhov. *Optimal Learning*. John Wiley and Sons, 2012.
- [25] H. Qu, I. O. Ryzhov, and M. C. Fu. Ranking and selection with unknown correlation structures. In C. Laroque, J. Himmelspach, R. Pasupathy, O. Rose, and A. M. Uhrmacher, editors, *Proceedings of the 2012 Winter Simulation Conference*, pages 144–155, 2012.
- [26] N. Z. Shor. Class of global minimum bounds of polynomial functions. *Cybernetics*, 23(6):731–734, 1987. (Russian orig.: Kibernetika, No. 6, (1987), 9–11).

Copyright© 2013 by the Institute for Operations Research and the Management Sciences (INFORMS). Permission to make digital/hard copy of part or all of this work for personal or classroom use is granted provided that copies are not made or distributed for profit or commercial advantage, the copyright notice, the title of the publication, and its date appear, and notice is given that copying is by permission of INFORMS. Distribution through course packs is permitted provided that permission is obtained from INFORMS. To republish, post on servers, or redistribute to lists requires specific permission. Address requests regarding reprint permission to permissions@informs.org, or to INFORMS, 7240 Parkway Drive, Suite 310, Hanover, MD 21076.

Acknowledgments

The Editor would like to thank all contributors who helped to make this newsletter available.