Optimization & Big Data 2015

Workshop, Trek and Colloquium

The age of "big data" is here: data of unprecedented sizes is becoming ubiquitous, which brings new challenges and new opportunities. With this comes the need to solve optimization problems of unprecedented sizes. Machine learning, compressed sensing, social network science and computational biology are some of many prominent application domains where it is increasingly common to formulate and solve optimization problems with billions of variables. Classical algorithms are not designed to scale to instances of this size and hence new approaches are needed. These approaches utilize novel algorithmic design involving tools such as distributed and parallel computing, randomization, asynchronicity, decomposition, sketching and streaming. This workshop aims to bring together researchers working on novel optimization algorithms and codes capable of working in the Big Data setting.

Organized by Zheng Qu and Peter Richtárik

EPSRC
Pioneering research and skills

LONDON MATHEMATICAL SOCIETY
150 YEARS

THE UNIVERSITY OF EDINBURGH

amazon.com

Keynote Speaker
Prof. Arkadi Nemirovski

Arkadi Nemirovski earned the Ph.D. in Mathematics (1974) from Moscow State University and the Doctor of Sciences in Mathematics (1990) from the Institute of Cybernetics of the Ukrainian Academy of Sciences, Kiev. He currently holds a Chair at the School of Industrial and Systems Engineering at Georgia Institute of Technology. Prof Nemirovski has made fundamental contributions in continuous optimization in the last thirty years that have significantly shaped the field.

In recognition of his contributions to convex optimization, Nemirovski was awarded the 1982 Fulkerson Prize from the Mathematical Programming Society and the American Mathematical Society (joint with Leonid Khachiyan and David Yudin), the 1991 Dantzig Prize from the Mathematical Programming Society and the Society for Industrial and Applied Mathematics (joint with Martin Grotschel). In recognition of his seminal and profound contributions to continuous optimization, Nemirovski was awarded the 2003 John von Neumann Theory Prize by the Institute for Operations Research and the Management Sciences (along with Michael J Todd). He is the only individual to have won all three of these prestigious prizes (Fulkerson, Dantzig, and von Neumann).

He continues to make significant contributions in almost all aspects of continuous optimization: complexity, numerical methods, stochastic optimization, and non-parametric statistics.

Honors & Awards

2009, Honorary Doctorate, Department of Mathematics, University of Waterloo
2006, Plenary talk @ International Congress of Mathematicians
2006, Semi-plenary talk @ International Symposium on Mathematical Programming
2003, John Von Neumann Theory Prize, INFORMS
1991, Dantzig Prize, Mathematical Programming Society and SIAM
1991, Plenary talk @ International Symposium on Mathematical Programming
1982, Fulkerson Prize, Mathematical Programming Society and the AMS

EPSRC
Pioneering research and skills

LONDON MATHEMATICAL SOCIETY
150 YEARS

THE UNIVERSITY OF EDINBURGH

amazon.com

Wednesday: May 6, 2015

Venue

Informatics Forum, room G.07
10 Crichton Street, George Square campus, University of Edinburgh (see attached map)

Schedule

09:15-09:35   *Registration*
Tea & Coffee with Sweet Danish

09:35-09:40   *Welcome*

09:40-10:15   **Arkadi Nemirovski** (Georgia Institute of Technology, USA)
Fenchel-type representations and large-scale problems with convex structure on difficult geometry domains

10:15-10:50   **Rodolphe Jenatton** (Amazon Berlin, Germany)
Sparse and spurious: dictionary learning with noise and outliers

10:50-11:20   *Break (Tea & Coffee with Biscuits)*

11:20-11:55   **Peter Richtárik** (Edinburgh of Edinburgh, UK)
SDNA: Stochastic Dual Newton Ascent for empirical risk minimization

11:55-12:45   *Spotlight Session 1*

    1. **Jose Vidal Alcala-Burgos** (CIMAT, Mexico)
        Affine invariant stochastic optimization
    2. **Dominik Csiba** (University of Edinburgh, UK)
        Stochastic dual coordinate ascent with adaptive probabilities
    3. **Kimon Fountoulakis** (University of Edinburgh, UK)
        A problem generator for big data optimization
    4. **Gordon Inverarity** (MetOffice, UK)
        Data assimilation for weather forecasting
    5. **Dimitris Kouzoupis** (University of Freiburg, Germany)
        First-order methods in nonlinear model predictive control
    6. **Wenting Long** (Sun Yat-sen University, China)
        Fixed point algorithm based on proximity and precondition operator for high resolution image reconstruction with displacement errors
    7. **Aurelien Lucchi** (ETH Zurich, Switzerland)
        Neighbourhood watch: variance reduction using nearest-neighbours

12:45-14:15   *Lunch + Poster Session 1 (Informatics Forum, Ground Floor)*

EPSRC
Pioneering research and skills

LONDON
MATHEMATICAL
SOCIETY
150 YEARS

THE UNIVERSITY OF EDINBURGH

amazon.com

14:15-14:50 **Martin Jaggi** (ETH Zurich, Switzerland)
Frank-Wolfe optimization algorithms: a brief tutorial

14:50-15:25 **Katya Scheinberg** (Lehigh University, USA)
Unconstrained trust region based stochastic optimization with biased and unbiased noise

15:25-15:55 *Break (Tea & Coffee with Biscuits)*

15:55-16:30 **Robert Gower** (University of Edinburgh, UK)
Randomized iterative point methods for linear systems and inverting matrices

16:30-17:05 **Mark Schmidt** (University of British Columbia, Canada)
Is greedy coordinate descent a terrible algorithm?

Arkadi Nemirovski (Georgia Institute of Technology, USA)

Fenchel-type representations and large-scale problems with convex structure on difficult geometry domains

In the talk, we present algorithms for solving problems with convex structure (*nonsmooth* convex minimization, convex-concave saddle point, and *variational* inequalities with monotone operators) on large-scale domains given by ``computationally cheap'' Linear Minimization Oracles (*LMO's*) capable to minimize linear forms over the domain. Our primary interest is in problems on large-scale nuclear/total variation norm balls, where the standard *proximal* type first order algorithms requiring at every step to minimize over problem's domain a specific, perhaps simple, but nonlinear convex function, become prohibitively time consuming, while minimizing a linear function over the domain still is relatively easy.

The first ingredient of our approach is utilizing Fenchel-type representations of convex functions and monotone operators defined on *LMO*-represented domains. These representations allow to associate with the problem of interest its dual, which, in general, is a *variational* inequality with monotone operator on some other domain. Due to flexibility of Fenchel-type representations, the domain of the dual problem in many important cases can be made ``*proximal*-friendly,'' so that the dual problem can be solved by a *proximal*-type first-order algorithm. The second ingredient of the approach is formed by ``accuracy certificate'' techniques allowing to recover a good solution to the problem of interest from the information collected when building a good solution to the dual problem.

The talk is based on joint research with Anatoli Juditski (Université Joseph Fourier)

EPSRC
Pioneering research and skills

LONDON
MATHEMATICAL
SOCIETY
150 YEARS

THE UNIVERSITY OF EDINBURGH

amazon.com

Rodolphe Jenatton (Amazon Berlin, Germany)

Sparse and spurious: dictionary learning with noise and outliers

A popular approach within the signal processing and machine learning communities consists in modelling signals as sparse linear combinations of atoms selected from a learned dictionary. While this paradigm has led to numerous empirical successes in various fields ranging from image to audio processing, there have only been a few theoretical arguments supporting these evidences. In particular, sparse coding, or sparse dictionary learning, relies on a non-convex procedure whose local minima have not been fully analyzed yet.

In this paper, we consider a probabilistic model of sparse signals, and show that, with high probability, sparse coding admits a local minimum around the reference dictionary generating the signals. Our study takes into account the case of over-complete dictionaries, noisy signals, and possible outliers, thus extending previous work limited to noiseless settings and/or under-complete dictionaries. The analysis we conduct is non-asymptotic and makes it possible to understand how the key quantities of the problem, such as the coherence or the level of noise, can scale with respect to the dimension of the signals, the number of atoms, the sparsity and the number of observations.

Joint with Rémi Gribonval (INRIA) and Francis Bach (INRIA)

Peter Richtárik (University of Edinburgh, UK)

SDNA: Stochastic Dual Newton Ascent for empirical risk minimization

We propose a new algorithm for minimizing regularized empirical loss: Stochastic Dual Newton Ascent (SDNA). Our method is dual in nature: in each iteration we update a random subset of the dual variables. However, unlike existing methods such as stochastic dual coordinate ascent, SDNA is capable of utilizing all curvature information contained in the examples, which leads to striking improvements in both theory and practice - sometimes by orders of magnitude.

In the special case when an L2-regularizer is used in the primal, the dual problem is a concave quadratic maximization problem plus a separable term. In this regime, SDNA in each step solves a proximal subproblem involving a random principal submatrix of the Hessian of the quadratic function; whence the name of the method. If, in addition, the loss functions are quadratic, our method can be interpreted as a novel variant of the recently introduced Iterative Hessian Sketch.

Joint work with Zheng Qu (Edinburgh), Martin Takáč (Lehigh) and Olivier Fercoq (Telecom ParisTech)

EPSRC
Pioneering research and skills

LONDON MATHEMATICAL SOCIETY
150 YEARS

THE UNIVERSITY OF EDINBURGH

amazon.com

Martin Jaggi (ETH Zurich, Switzerland)

Frank-Wolfe optimization algorithms: a brief tutorial

Frank-Wolfe methods (a.k.a. Conditional Gradient) have been successfully applied to a wide range of large-scale learning and data processing applications, such as matrix factorizations, structured variable selection, sparse regression, multi-task learning, image denoising, structured prediction, just to name a few. The applications have in common that they utilize the atomic decomposition of the variable of interest, that is expanding it as a linear combination of the elements of a dictionary. In this brief tutorial, we provide an overview of the algorithm variants, their convergence rates and underlying assumptions.

Katya Scheinberg (Lehigh University, USA)

Unconstrained trust region based stochastic optimization with biased and unbiased noise

We will present a very general framework for unconstrained stochastic optimization, which is based on standard trust region framework using random models. In particular this framework retains the desirable features such step acceptance criterion, trust region adjustment and ability to utilize of second order models. We make assumptions on the stochasticity that are different from the typical assumptions of stochastic and simulation-based optimization. In particular we assume that our models and function values satisfy some good quality conditions with some probability fixed, but can be arbitrarily bad otherwise. We will analyze the convergence of this general framework and discuss the requirement on the models and function values. We will contrast our results with existing results from stochastic approximation literature.

We will then present computational results for several classes of noisy functions, including cases when noise is not i.i.d. and dominates the function values, when it occurs. We will show that our simple framework performs very well in that setting, while standard stochastic methods fail.

Robert Gower (University of Edinburgh, UK)

Randomized iterative methods for linear systems and inverting matrices

We present a novel and expressive framework for developing and analyzing randomized iterative methods for solving linear systems. Our generic method depends on two parameters: a fixed positive definite matrix and a random matrix (which can follow an almost arbitrary probability law). By varying these two parameters, we obtain several well-known and many new algorithms. We present two alternative and surprisingly simple complexity analyses of the generic method, proving linear convergence (i.e., exponential error decay) under very weak assumptions.

Each step of our method is written as an optimization problem admitting a closed-form solution. It turns out that this optimization problem can be written in two different but equivalent and mutually dual ways. This duality gives rise to two natural interpretations of the algorithm. In addition, our method can also be interpreted as a randomized fixed-point iteration.

EPSRC
Pioneering research and skills

LONDON MATHEMATICAL SOCIETY
150 YEARS

THE UNIVERSITY OF EDINBURGH

amazon.com

For particular choices of the two parameters, we recover, for instance, the celebrated randomized Kaczmarz and randomized coordinate descent methods. However, we naturally also obtain their block variants. Our complexity bounds match the best-known bounds. It was long observed that these two methods are related, but up to now a fully satisfying understanding of the relationship between them was not obtained. Our work sheds new light on this relationship.

The flexibility of our approach leads to a number of other (existing and new) methods. For instance, our generic method specializes to a method, which in every iteration projects the current iterate onto the hyperplane formed by a linear combination, with Gaussian weights, of the equations in the system (a "Gaussian", as opposed to row-wise, Kaczmarz method). Another interesting scheme is based on exact line search for the least squares function, in a Gaussian direction (a "Gaussian", as opposed to coordinate, descent).

Lastly, using an analogous framework but in a matrix space, we develop new randomized iterative methods for inverting matrices, which, combined with suitable randomized linear solvers, pave the way for randomized preconditioned methods.

Joint work with Peter Richtárik (Edinburgh)

Mark Schmidt (University of British Columbia, Canada)

Is greedy coordinate descent a terrible algorithm?

There has been significant recent work on the theory and application of randomized coordinate-descent algorithms, beginning with the work of Nesterov [2010], who showed that a random-coordinate selection rule achieves the same convergence rate as the Gauss-Southwell selection rule. This result suggests that we should never use the Gauss-Southwell rule, as it is typically much more expensive than random selection. However, the empirical behaviours of these algorithms contradict this theoretical result: in applications where the computational costs of the selection rules are comparable, the Gauss-Southwell selection rule tends to perform substantially better than random coordinate selection.

We give a simple analysis of the Gauss-Southwell rule showing that---except in extreme cases---it is always faster than choosing random coordinates. Further, in this work we (i) show that exact coordinate optimization improves the convergence rate for certain sparse problems, (ii) propose a Gauss-Southwell-Lipschitz rule that gives an even faster convergence rate given knowledge of the Lipschitz constants of the partial derivatives, (iii) analyze the effect of approximate Gauss-Southwell rules, and (iv) analyze a proximal-gradient variant of the Gauss-Southwell rule.

Joint work with Michael Friedlander (University of California, Davis)

EPSRC
Pioneering research
and skills

LONDON
MATHEMATICAL
SOCIETY
150 YEARS

THE UNIVERSITY OF EDINBURGH

amazon.com

Thursday: May 7, 2015

Venue

Informatics Forum, room G.07
10 Crichton Street, George Square campus, University of Edinburgh (see attached map)

Schedule

09:15-09:35 *Registration*
Tea & Coffee with Sweet Danish

09:35-09:40 *Welcome*

09:40-10:15 **Donald Goldfarb** (Columbia University, USA)
Low-rank matrix and tensor recovery: theory and algorithms

10:15-10:50 **Zheng Qu** (University of Edinburgh, UK)
Randomized dual coordinate ascent with arbitrary sampling

10:50-11:20 *Break (Tea & Coffee with Biscuits)*

11:20-11:55 **Jonathan Eckstein** (Rutgers University, USA)
Object-parallel solution of large-scale lasso problems

11:55-12:45 *Spotlight Session 2*

8. **Rodrigo Mendoza Smith** (University of Oxford, UK)
    Expander L0 decoding
9. **Christian Mueller** (Simons Center for Data Analysis, USA)
    Auto-tuned high-dimensional regression with the TREX: theoretical guarantees and non-convex global optimization
10. **Thomas Prescott** (University of Oxford, UK)
    Layered synthetic biomolecular systems
11. **Daniel Robinson** (Johns Hopkins University, USA)
    A hybrid ADMM algorithm
12. **Chee-Wei Tan** (City University of Hong Kong, Hong Kong)
    A probabilistic approach to rumor source detection and graph-based message passing algorithms
13. **Simon Tett** (University of Edinburgh, UK)
    Optimising parameter values in climate models: observational/model synthesis
14. **Weiqi Zhou** (Jacobs University, Germany)
    Kaczmarz iteration with random row permutation

12:45-14:15 *Lunch + Poster Session 2 (Informatics Forum, Ground Floor)*

14:15-14:50   **Patrick Louis Combettes** (Paris 6, France)
Splitting techniques in the face of huge problem sizes: block-coordinate and block-iterative approaches

14:50-15:25   **Francois Glineur** (Louvain, Belgium)
Smooth strongly convex interpolation and exact worst-case performance of first-order methods

15:25-15:55   *Break (Tea & Coffee with Biscuits)*

15:55-16:30   **Jakub Konečný** (University of Edinburgh, UK)
Distributed optimization with arbitrary local solvers

16:30-17:05   **Garud Iyengar** (Columbia University, USA)
A distributed proximal method for composite convex optimization

Donald Goldfarb (Columbia University, USA)

Low-rank matrix and tensor recovery: theory and algorithms

Recovering a low-rank matrix from incomplete or corrupted observations is a recurring problem in signal processing and machine learning. For problems in which the intrinsic structure of incomplete or corrupted data is more than 3-dimensional, low-rank completion and RPCA convex models for matrices have been extended to tensors. Here we establish recovery guarantees for both tensor completion and tensor RPCA, show that using the most popular convex relaxation for the tensor Tucker rank can be substantially sub-optimal in terms of the number of observations needed for exact recovery and introduce a very simple new convex relaxation that is theoretically and empirically much better. We also propose algorithms to solve these models that are based on Alternating Direction Augmented Lagrangian (ADAL), Frank-Wolfe and prox-gradient methods, and empirically study their performance on both simulated and real data.

Joint work with: Cun Mu, Bo Huang and Tony Qin (current and former IEOR PhD students at Columbia University) and John Wright (E.E. faculty member at Columbia University)

Zheng Qu (University of Edinburgh, UK)

Randomized dual coordinate ascent with arbitrary sampling

We study the problem of minimizing the average of a large number of smooth convex functions penalized with a strongly convex regularizer. We propose and analyze a novel primal-dual method (Quartz) which at every iteration samples and updates a random subset of the dual variables, chosen according to an *arbitrary distribution*. In contrast to typical analysis, we directly bound the decrease of the *primal-dual error* (in expectation), without the need to first analyze the dual error. Depending

on the choice of the sampling, we obtain efficient serial, parallel and distributed variants of the method. In the serial case, our bounds match the best-known bounds for SDCA (both with uniform and importance sampling).

With standard mini-batching, our bounds predict initial data-independent speedup as well as additional data-driven speedup, which depends on spectral and sparsity properties of the data. We calculate theoretical speedup factors and find that they are excellent predictors of actual speedup in practice. Moreover, we illustrate that it is possible to design an efficient mini-batch importance sampling. The distributed variant of Quartz is the first distributed SDCA-like method with an analysis for non-separable data.

Joint work with Peter Richtárik (Edinburgh) and Tong Zhang (Rutgers and Baidu)

Jonathan Eckstein (Rutgers University, USA)

Object-parallel solution of large-scale lasso problems

This talk describes an "object-parallel" philosophy to implementing first-order optimization algorithms, and some C++ software tools we have begun developing to support it. Our tools allow algorithms to be expressed with MATLAB-like simplicity, but still attain high performance on a variety of serial and parallel platforms. In particular, we use a novel "symbolic temporaries" delayed-evaluation technique to improve the efficiency of C++ operator overloading.

As an example of our approach, we solve some large-scale Lasso problems on a distributed-memory supercomputer, employing the spectral projected gradient (SPG) method and considering both dense and sparse problem instances. We describe both the dense implementation, which is relatively straightforward, and the techniques we use to efficiently handle unbalanced sparsity patterns in the sparse case.

Joint work with György Mátyásfalvi (doctoral candidate, Rutgers University)

Patrick Louis Combettes (Paris 6, France)

Splitting techniques in the face of huge problem sizes: block-coordinate and block-iterative approaches

Monotone operator splitting technology constitutes the theoretical and algorithmic foundation of a wide array of numerical methods in data-driven problems. Fueled by new developments in abstract duality for monotone inclusions and product space techniques, significant advances have been made in splitting methods in recent years. In particular, it is now possible to solve highly structured monotone inclusions with algorithms which guarantee the convergence of the iterates. In problems of huge sizes, the implementation of these algorithms faces significant challenges which often render them inapplicable.

We present two approaches to face this issue, which both preserve the splitting and convergence properties of the algorithms. First (joint work with Jean-Christophe Pesquet), we propose a general

stochastic block-coordinate fixed-point framework, from which we derive flexible block-coordinate versions of common splitting algorithms. Second, (joint work with Jonathan Eckstein), we propose a block-coordinate primal-dual splitting framework for composite monotone inclusions in which only subgroups of operators need to be activated at each iteration.

Joint work with Jean-Christophe Pesquet (Université Paris-Est) and Jonathan Eckstein (Rutgers)

Francois Glineur (Universite Catholique de Louvain, Belgium)

Smooth strongly convex interpolation and exact worst-case performance of first-order methods

We show that the exact worst-case performance of fixed-step first-order methods for smooth (possibly strongly) convex functions can be obtained by solving convex programs. Finding the worst-case performance of a black-box first-order method is formulated as an optimization problem over a set of smooth (strongly) convex functions and initial conditions. We develop closed-form necessary and sufficient conditions for smooth (strongly) convex interpolation, which provide a finite representation for those functions. This allows us to reformulate the worst-case performance estimation problem as an equivalent finite dimension-independent semidefinite optimization problem, whose exact solution can be recovered up to numerical precision. Optimal solutions to this performance estimation problem provide both worst-case performance bounds and explicit functions matching them, as our smooth (strongly) convex interpolation procedure is constructive.

Our works build on those of Drori and Teboulle who introduced and solved relaxations of the performance estimation problem for smooth convex functions. We apply our approach to different fixed-step first-order methods with several performance criteria, including objective function accuracy and residual gradient norm. We conjecture several numerically supported worst-case bounds on the performance of the gradient, fast gradient and optimized fixed-step methods, both in the smooth convex and the smooth strongly convex cases, and deduce tight estimates of the optimal step size for the gradient method.

Joint work with Adrien B. Taylor and Julien M. Hendrickx (UCL, Belgium)

Jakub Konečný (University of Edinburgh, UK)

Distributed optimization with arbitrary local solvers

In situations when the data describing an optimization problem is so large that it cannot be stored on a single computer---such situations are becoming increasingly common---it is necessary to design method capable of running on a cluster (i.e., in a memory-distributed fashion). However, "naively" distributed implementations of traditional optimization algorithms often suffer from significant decrease in computational efficiency caused by the excessive cost of communication between individual computers/nodes.

We present a novel optimization framework, complete with theoretical complexity bounds, for

efficient distributed optimization with arbitrary local solvers (local to individual compute nodes of a cluster). Our framework allows the user (e.g., a data science company) to solve big data problems by making it possible for them to utilize existing, trusted and often fine-tuned solvers they have access to/own, but which are limited to single-node implementation only. Many companies are now in precisely this situation. We provide a distribution and communication protocol on top of their existing local solver, and prove complexity bounds for the overall algorithm. The efficiency of the method can be optimized by fine-tuning a certain parameter of the method to the computing architecture being employed.

In particular, we first split the data describing the optimization problem into as many parts as there are computers in the cluster. Our framework then prescribes what subproblems should each of the nodes solve in each iteration. These subproblems are fully described by the data stored and owned locally by each node, and hence are solvable by that node. A particular feature of this framework is that we allow the subproblems to be solved by an *arbitrary solver* (possibly a randomized one), and inexactly. Once all nodes solve their local problems, our framework prescribes how the local solutions should be combined and what data should be communicated before the next iteration can start. We prove complexity bounds revealing a trade-off between the level of inexactness and the unit communication cost. The inexactness parameter can then be fine-tuned for optimal performance. This effect is strong and can be clearly observed in practice.

Joint work with Peter Richtárik (Edinburgh University), Martin Jaggi (ETH Zurich), Chenxin Ma (Lehigh University) and Martin Takáč (Lehigh University)

Garud Iyengar (Columbia University, USA)

A distributed proximal method for composite convex optimization

We propose a distributed first-order augmented Lagrangian (DFAL) algorithm for minimizing the sum of composite convex functions, where each term in the sum is only known at one of the nodes, and only nodes connected by an edge can directly communicate with each other. We also allow for global constraints on the solution. This optimization model abstracts a number of applications in distributed sensing and machine learning.

We show that any limit point of DFAL iterates is optimal; and the iterates are epsilon-optimal in $O(\log(1/epsilon))$ iterations, which require $O(psi\_max^{3/2}/d\_min*1/epsilon)$ communications steps, where psi_max denotes the largest eigenvalue of the graph Laplacian, and d_min is the degree of smallest degree node.

We also propose an asynchronous version of DFAL by incorporating randomized block coordinate descent methods; and demonstrate the efficiency of DFAL on large scale sparse-group LASSO problems.

Joint work with Necdet Serhat Aybat (Penn State University)

Friday: May 8, 2015

Trek to Arthur's Seat

On Friday morning there is an optional walk / light trek to the top of Arthur's Seat—a 350 million year old volcano in the city centre rising 251 meters above the sea level—offering magnificent views of Edinburgh and the Firth of Forth. We are departing at 9:30 sharp from the entrance of Informatics Forum, please plan to arrive 10 minutes earlier.

"The views from the summit [of Arthur's seat] are awesome. The city itself is a real gem, Edinburgh is in my own opinion the most beautiful and fascinating city in Britain by a mile."
[www.TrekkingBritain.com]

While the walk is reasonably light, good footwear is recommended as there will be some ascending to do. It is recommended that you bring along a light weather-proof jacket; it can get windy up on the hill. Also, please bring enough water and light refreshments with you.

Schedule

09:20-09:30   *Meeting in front of the Informatics Forum*

09:30-12:30   Light trek to the top of Arthur's Seat (and back to Royal Mile)

12:30-14:00   *Lunch (individual)*

EPSRC
Pioneering research and skills

LONDON MATHEMATICAL SOCIETY
150 YEARS

THE UNIVERSITY OF EDINBURGH

amazon.com

The Route

We are not going to walk straight to the Arthur's Seat. Instead, we will take a detour and first walk through the George Square campus towards the National Museum of Scotland (entrance free, great exhibits) and the Elephant house, the birthplace cafe of Harry Potter. Continuing along the 300m long George IV Bridge built in 1832, we enter the Royal Mile (a 1 mile long backbone of Edinburgh's historic centre joining the Edinburgh Castle and the Holyrood Abbey) near the High Court of Justiciary, Scotland's supreme criminal court. From that point we will walk along the Royal Mile, away from the Castle and towards the Holyrood Abbey, i.e., down the hill, passing many points of interest which you might want to visit if your are staying in Edinburgh a bit longer: St. Gile's Cathedral (Mother Church of Presbyterianism), Mary King's Close, North Bridge, Canongate Kirk, Scottish Parliament Building, Dynamic Earth Science Centre, Palace of Holyroodhouse and Queen's Gallery. Once at the end of the Royal Mile we enter the Holyrood Park, walking along a foot path (called Radical Road) below Salisbury Crags—a stretch of over 46 meters high cliffs of dolerite and columnal basalt offering spectacular views of the city centre. At the end of the path is the foot of Arthur's Seat; from there it is just a 20 min walk up to the summit point. We will return to the Royal Mile via a different route through the Holyrood Park, passing by the St. Margaret's Loch and the ruins of St Anthony's Chapel.

Friday: May 8, 2015

Colloquium

Venue

The Colloquium is held in the James Clerk Maxwell Building (JCMB) of the King's Buildings campus of the University of Edinburgh, in the southern part of the city. The lecture starts in Lecture Theatre C at 14:00 (on the third floor of JCMB). Following the lecture there will be a Networking Reception in the open space area of the third floor of JCMB at 15:10.

Schedule

14:00-15:00	**Prof. Arkadi Nemirovski** (Georgia Institute of Technology)
	On Statistical Inference via Convex Optimization
	(University of Edinburgh, James Clerk Maxwell Building, Lecture Theatre C)
15:00-15:10	Best Contribution Award (sponsored by Amazon Berlin)
15:10-17:00	*Networking Reception (JCMB 3rd floor, Magnet Café Area)*

Arkadi Nemirovski (Georgia Institute of Technology, USA)

On statistical inference via convex optimization

We discuss a general computation-oriented approach to near-optimal testing of composite hypotheses and some other statistical inference problems. Our main building block is the notion of a ``good'' observation scheme (o.s.); examples include the cases where an observation is

(a) affine image of a vector (''signal'') corrupted by white Gaussian noise, which is the standard model in Signal Processing,

(b) a vector with independent Poisson entries with parameters affinely depending on the signal (situation arising in Poisson Imaging covering Positron Emission Tomography and Nanoscale Fluorescent Microscopy),

(c) random variable taking finitely many values with probabilities affinely parameterized by the signal, and

(d) naturally defined direct products of o.s.'s (a) - (c) reflecting what happens when passing from a single observation to a collection of independent observations.

We show that given a good o.s., a near-optimal test deciding on a (finite) collection of convex hypotheses (i.e., hypotheses stating that the signal underlying observations belongs to a convex compact set associated with the hypothesis) and the risk of the test can be computed efficiently via Convex Optimization. We discuss sequential and dynamical versions of the tests and outline some applications, including near-optimal in the minimax sense recovery of a linear form (or a slightly more general function) of a signal known to belong to the union of finitely many convex sets and observed via a good o.s.

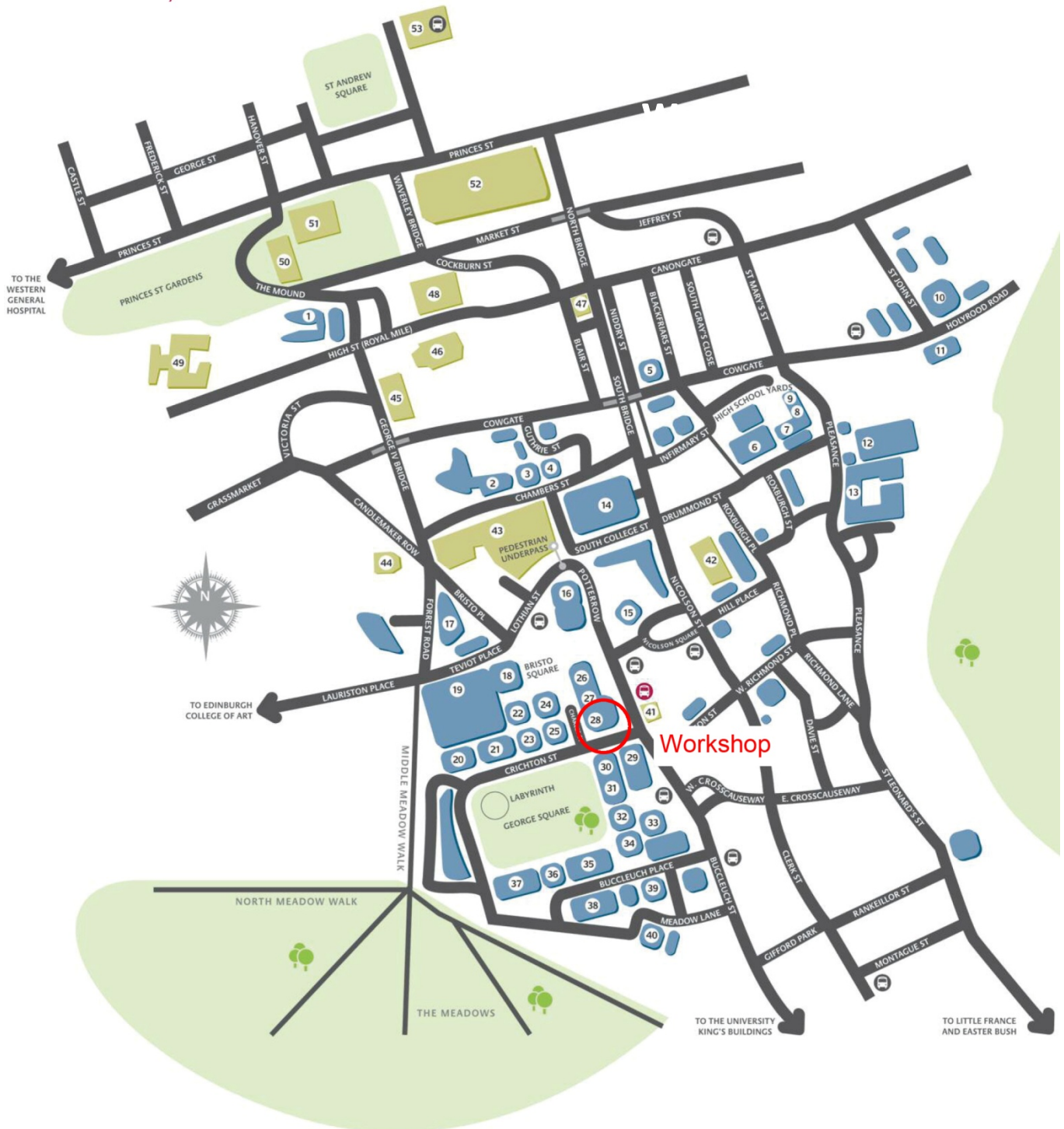*This talk is based on joint work with Alexander Goldenshluger (Haifa) and Anatoli Juditski (Université Joseph Fourier)*

EPSRC
Pioneering research and skills

LONDON MATHEMATICAL SOCIETY
150 YEARS

THE UNIVERSITY OF EDINBURGH

amazon.com

Maps

The University in the city

The University Central Area

Workshop

Map labels (street names and markers):
ST ANDREW SQUARE, PRINCES ST, GEORGE ST, CASTLE ST, FREDERICK ST, HANOVER ST, WAVERLEY BRIDGE, MARKET ST, NORTH BRIDGE, JEFFREY ST, CANONGATE, ST MARY'S ST, ST JOHN ST, HOLYROOD ROAD, THE MOUND, COCKBURN ST, PRINCES ST GARDENS, TO THE WESTERN GENERAL HOSPITAL, HIGH ST (ROYAL MILE), BLAIR ST, NIDDRY ST, BLACKFRIARS ST, SOUTH GRAY'S CLOSE, COWGATE, VICTORIA ST, GEORGE IV BRIDGE, GRASSMARKET, COWGATE, GUTHRIE ST, CHAMBERS ST, SOUTH BRIDGE, INFIRMARY ST, HIGH SCHOOL YARDS, PLEASANCE, DRUMMOND ST, ROXBURGH ST, ROXBURGH PL, RICHMOND PL, CANDLEMAKER ROW, FORREST ROAD, BRISTO PL, LOTHIAN ST, POTTERROW, SOUTH COLLEGE ST, NICOLSON ST, HILL PLACE, RICHMOND LANE, TO EDINBURGH COLLEGE OF ART, LAURISTON PLACE, TEVIOT PLACE, BRISTO SQUARE, NICOLSON SQUARE, W. RICHMOND ST, DAVIE ST, CRICHTON ST, MIDDLE MEADOW WALK, LABYRINTH GEORGE SQUARE, W. CROSSCAUSEWAY, E. CROSSCAUSEWAY, ST LEONARD'S ST, CLERK ST, BUCCLEUCH PLACE, BUCCLEUCH ST, NORTH MEADOW WALK, MEADOW LANE, GIFFORD PARK, RANKEILLOR ST, MONTAGUE ST, THE MEADOWS, TO THE UNIVERSITY KING'S BUILDINGS, TO LITTLE FRANCE AND EASTER BUSH, PEDESTRIAN UNDERPASS

Legend

1 New College: School of Divinity
2 Minto House
3 Charles Stewart House
4 Adam House
5 St Cecilia's Hall
6 Old Infirmary (Geography)
7 Old Surgeon's Hall
8 Drummond Library (GeoSciences)
9 Chisholm House
10 The Moray House School of Education
11 St Leonard's Land
12 Centre for Sport and Exercise
13 The Pleasance
14 Old College
15 Alison House
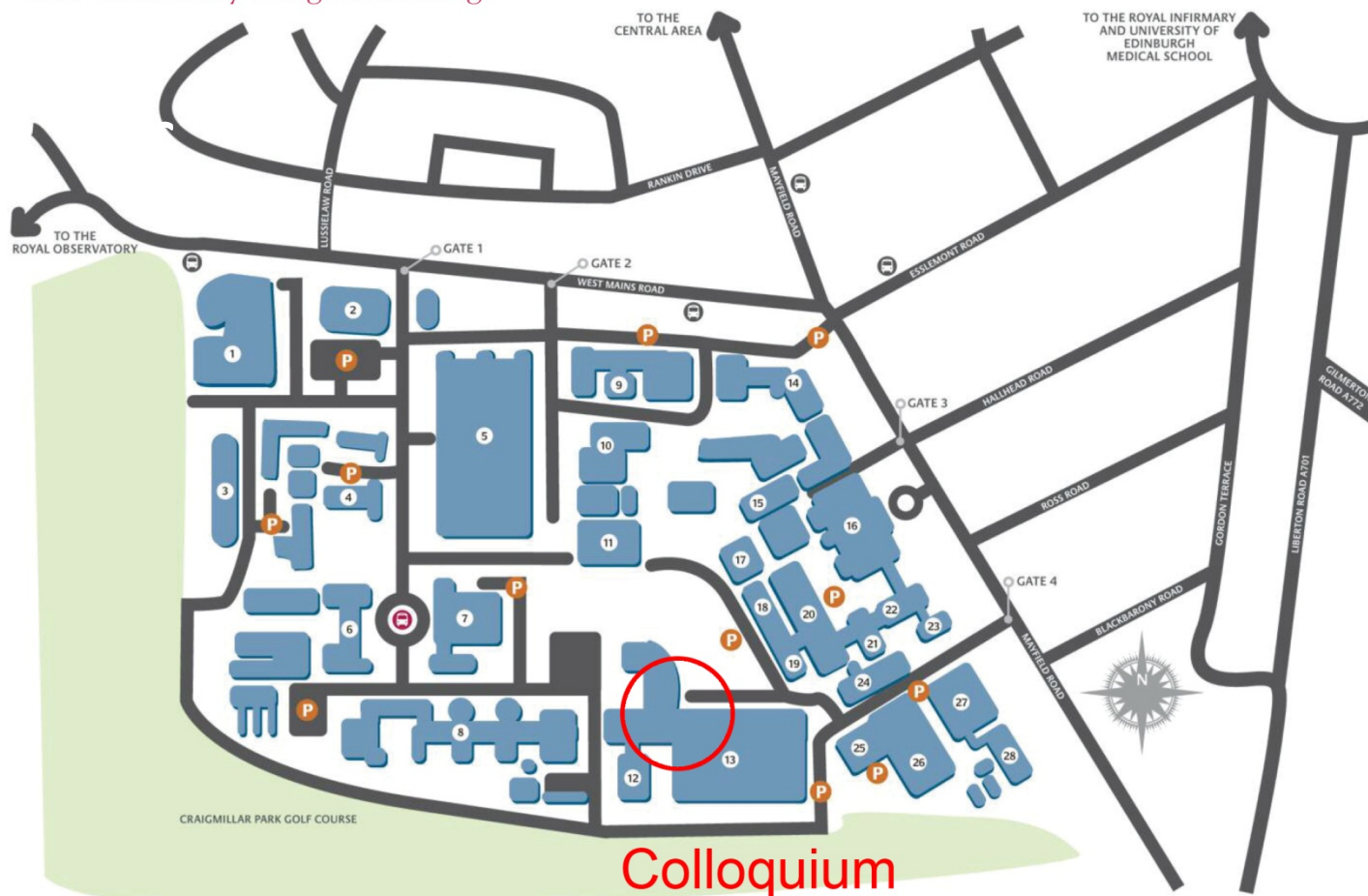
16 The Potterrow Student Centre: Students' Association (EUSA); Chaplaincy Centre; Health Centre
17 Bedlam Theatre
18 McEwan Hall
19 Medical School
20 Chrystal Macmillan Building
21 Hugh Robson Building
22 Reid Concert Hall
23 7 George Square
24 Teviot Row House: Student Union
25 1 George Square
26 Dugald Stewart Building
27 Visitor Centre: Information, Exhibition and Shop
28 Informatics Forum

29 Appleton Tower
30 International Office and Student Recruitment and Admissions
31 College of Humanities & Social Science Office
32 50 George Square
33 David Hume Tower Lecture Theatres
34 David Hume Tower
35 Business School
36 George Square Lecture Theatre
37 Main Library; Careers Service; Student Counselling Service; Student Disability Service
38 International Student Centre
39 Office of Lifelong Learning
40 Institute for Advanced Studies in the Humanities (IASH)

Shuttle bus to King's Buildings Campus
41 Edinburgh Central Mosque
42 Surgeon's Hall
43 Royal Scottish Museum
44 Greyfriars Kirk
45 The National Library of Scotland
46 St Giles' Cathedral
47 Tron Kirk
48 City Chambers
49 Edinburgh Castle
50 Royal Scottish Academy
51 National Gallery of Scotland
52 Edinburgh Waverley train station
53 St Andrew's bus station
Public bus stop

The University King's Buildings

TO THE CENTRAL AREA

TO THE ROYAL OBSERVATORY

TO THE ROYAL INFIRMARY AND UNIVERSITY OF EDINBURGH MEDICAL SCHOOL

RANKIN DRIVE

LUSSIELAW ROAD

MAYFIELD ROAD

ESSLEMONT ROAD

HALLHEAD ROAD

ROSS ROAD

GORDON TERRACE

CILMERTON ROAD A772

LIBERTON ROAD A701

BLACKBARONY ROAD

GATE 1

GATE 2

WEST MAINS ROAD

GATE 3

GATE 4

CRAIGMILLAR PARK GOLF COURSE

Colloquium

The regular bus service between the Central area and King's Buildings can be viewed at
www.ed.ac.uk/schools-departments/transport/public-transport/buses/shuttle-bus

1 British Geological Survey
2 Student residences
3 Weir Building: Science & Engineering College Office; Careers Service; Student Counselling Service
4 Crew Building
5 Joseph Black Building
6 Scottish Micro Electronics Centre
7 Roger Land Building
8 Peter Wilson Building
9 Grant Institute
10 KB House: Students' Association (EUSA)
11 KB Centre
12 Erskine Williamson Building
13 James Clerk Maxwell Building
14 Ashworth Laboratories

15 Kenneth Denbigh Building
16 Sanderson Building
17 John Muir Building
18 William Rankine Building
19 Alexander Graham Bell Building
20 Fleeming Jenkin Building
21 Faraday Building
22 Hudson Beare Building
23 Engineering Lecture Theatre
24 Alrick Building
25 Michael Swann Building
26 Darwin Building
27 Daniel Rutherford Building
28 CH Waddington Building

🚌 Shuttle bus to Central Area Campus
🚌 Public bus stop
🅿 Parking

Optimization & Big Data 2015
Workshop, Trek and Colloquium

EPSRC
Pioneering research and skills

LONDON MATHEMATICAL SOCIETY
150 YEARS

THE UNIVERSITY OF EDINBURGH

amazon.com

Internet Access

eduroam

Please use eduroam if you have signed up for it at your home institution.

The eduroam network is a secure world-wide roaming access service for the international research and education community. Eduroam is available across Europe, North America and Asia/Pacific. If you are registered for eduroam at your home institution, then you should be able to connect automatically to the wireless network here. If this does not work, it may be because your computer has already cached site-specific details for a different eduroam location, and these conflict with the settings here. Try deleting your eduroam connection entirely and letting it bootstrap to pick up the local connection details.

- Network name (SSID): **eduroam**
- Wireless Network Authentication type: **WPA2 Enterprise**
- Data Encryption method: **AES**

If your wireless device does not support WPA2/AES, select WPA authentication and TKIP encryption.

Individual Wireless Accounts

Individual wireless access accounts will be provided at registration to those who are not able to use eduroam. Please ask at the registration desk or talk to the organizers if you have not received one.

Participants

Kerem Akartunali	*University of Strathclyde, Management Science, UK*
Jose Vidal Alcala-Burgos	*Centro de Investigacion en Matematicas, Mexico*
Razvan Anisca	*Lakehead University, Mathematical Sciences, Canada*
Francesco Bagattini	*Università di Firenze, Information Engineering, Italy*
Richard Barraclough	*Cloudbuy, UK*
Luca Bravi	*Università di Firenze, Information Engineering, Italy*
Waqquas Bukhsh	*University of Strathclyde, Institute for Energy and Environment, UK*
Coralia Cartis	*University of Oxford, Mathematical Institute, UK*
Gillian Chin	*Twitter, Data Scientist, USA*
Patrick Louis Combettes	*Université Pierre et Marie Curie - Paris 6, Mathematics, France*
Dominik Csiba	*University of Edinburgh, Mathematics, UK*
Boris Defourny	*Lehigh University, Industrial and Systems Engineering, USA*
Jonathan Eckstein	*Rutgers University, Management Science and Information Systems, USA*
Zhou Fang	*Biomathematics and Statistics Scotland (BIOSS), UK*
Kimon Fountoulakis	*University of Edinburgh, Mathematics, UK*
Sergio Garcia Quiles	*University of Edinburgh, Mathematics, UK*
Richard Gibbens	*University of Cambridge, Computer Laboratory, UK*
Francois Glineur	*Université Catholique de Louvain, CORE, Belgium*
Donald Goldfarb	*Columbia University, Industrial Engineering and Operations Research, USA*
Robert Gower	*University of Edinburgh, Mathematics, UK*
Harry Hill	*Smartodds, UK*
Julian Hall	*University of Edinburgh, Mathematics, UK*
Gordon Inverarity	*Met Office, UK*
Garud Iyengar	*Columbia University, Industrial Engineering and Operations Research, USA*
Martin Jaggi	*ETH Zurich, Data Analytics Lab, Switzerland*
Rodolphe Jenatton	*Amazon, Germany*
Jakob Kisiala	*University of Edinburgh, Mathematics, UK*
Jakub Konečný	*University of Edinburgh, Mathematics, UK*
Maria Koroliuk	*University of Warwick, Complex Systems Science, UK*
Dimitris Kouzoupis	*University of Freiburg, Microsystems Engineering, Germany*
Wenting Long	*Sun Yat-sen University, China*
Aurelien Lucchi	*ETH Zurich, Switzerland*
Belen Martin-Barragan	*University of Edinburgh, Management Science, UK*
Iain MacLeod	*Tesco Bank, UK*
Rodrigo Mendoza Smith	*University of Oxford, Mathematical Institute, UK*
Somayeh Moazeni	*Stevens Institute of Technology, Systems and Enterprises, USA*
Christian Mueller	*Simons Center for Data Analysis, USA*
Mojmir Mutný	*University of Edinburgh, Mathematical Physics, UK*
Arkadi Nemirovski	*Georgia Institute of Technology, Industrial and Systems Engineering, USA*
Alexandru Onose	*Heriot-Watt University, UK*
Thomas Prescott	*University of Oxford, Engineering Science, UK*
Zheng Qu	*University of Edinburgh, Mathematics, UK*
Rajeev Ratan	*University of Edinburgh, UK*

EPSRC
Pioneering research
and skills

LONDON
MATHEMATICAL
SOCIETY
150 YEARS

THE UNIVERSITY OF EDINBURGH

amazon.com

Peter Richtárik	*University of Edinburgh, Mathematics, UK*
Daniel Robinson	*Johns Hopkins University, Applied Mathematics and Statistics, USA*
Arnold Pimentel Salas	*University of Oxford, Quantitative Finance / Machine Learning, UK*
Katya Scheinberg	*Lehigh University, Industrial and Systems Engineering, USA*
Mark Schmidt	*University of British Columbia, Computer Science, Canada*
Sebastian Stich	*Universite Catholique de Louvain, Belgium*
Chee-Wei TAN	*City University of Hong Kong, Computer Science, UK*
Simon Tett	*University of Edinburgh, Geosciences, UK*
Andrew Thompson	*Dell Secureworks, UK*
Neil Urquhart	*Edinburgh Napier University, UK*
Lian Zheng	*City University of Hong Kong, Hong Kong*
Weiqi Zhou	*Jacobs University, Mathematics, Germany*

Acknowledgements