Continuous Optimization:
Challenges and Applications

An international workshop in honor of
Ronny Ben-Tal's 70 birthday

September 5-8, 2016,
Bloomfield Building, Room 527 | Industrial Engineering & Management
The Technion, Israel

Participants

Hedy Attouch (University of Montpellier, France)
Francis Bach (INRIA, Paris, France)
Jonathan Barzilai (Dalhousie University, Canada)
Heinz Bauschke (UBC Okanagan, Canada)
Adi Ben-Israel (Rutgers, USA)
Dimitris Bertsimas (Massachusetts Institute of Technology, USA)
Chiranjib Bhattacharyya (Indian Institute of Science, India)
Jerome Bolte (Toulouse 1 Capitole University, France)
Yair Censor (University of Haifa, Israel)
Roberto Cominetti (University of Chile, Chile)
Yoel Drori (Google R & D, Tel Aviv, Israel)
Yonina Eldar (Technion, Israel)
Robert M. Freund (Massachusetts Institute of Technology, USA)
Dick den Hertog (Tilburg University, Netherlands)
Alexander D. Ioffe (Technion, Israel)
Jean Bernard Lasserre (LAAS-CNRS, Toulouse, France)
Xin Liu (Chinese Academy of Sciences, China)
Russell Luke (Goettingen, Germany)
Aaron Melman (Santa Clara University, USA)
Arkady Nemirovsky (Georgia Tech, USA)
Yurii Nesterov (Université catholique de Louvain, Belgium)
Jong-Shi Pang (University of Southern California, USA)
Edouard Pauwels (University of Toulouse 3, France)
Georgia Perakis (Massachusetts Institute of Technology, USA)
Thomas Pock (Graz University of Technology, Austria)
Roman Polyak (Technion, Israel)
Simeon Reich (Technion, Israel)
Katya Scheinberg (Lehigh University, USA)
Shimrit Shthern (Massachusetts Institute of Technology, USA)
Jean Philippe Vial (University of Geneva, Switzerland)
Henry Wolkowicz (University of Waterloo, Canada)
Yinyu Ye (Stanford, USA)
Xiaoming Yuan (Hong Kong Baptist University, Hong Kong)
Ya-Xiang Yuan (Chinese Academy of Sciences, China)

Organizing Committee

Amir Beck – The Technion
Shoham Sabach – The Technion
Marc Teboulle – Tel Aviv University

http://iew.technion.ac.il/coca2016/

The workshop is supported by the Israel Science Foundation
The international workshop

Continuous Optimization: Challenges and Applications

celebrating Ronny Ben-Tal’s 70th Birthday, will take place at the Technion (Israel) on September 5-8, 2016.

The aim of this four days workshop is to gather some of the top researchers working in the area of continuous optimization, highlighting recent advances in theory, algorithms and applications. It will provide a forum for exchanging ideas across fields, discussing new trends and challenges in the optimization discipline and its broad applications and for initializing new research cooperation. The conference is in honor of Ronny Ben-Tal on the occasion of his 70th birthday, who has made numerous pioneering and outstanding contributions to Optimization and its connection to various other mathematical areas and scientific applications.

Supported by the Israel Science Foundation

Location
Faculty of Industrial Engineering and Management
Bloomfield Building, 5th Floor – Room 527
Technion, Haifa

Website
http://iew.technion.ac.il/coca2016/

Organizing Committee
Amir Beck (The Technion)
Marc Teboulle (Tel Aviv University)
Shoham Sabach (The Technion)
Getting to your hotel/guesthouse from Ben-Gurion airport

To get to Haifa from Ben Gurion (Tel Aviv) airport, there are essentially three options (for additional information, please see the link of the conference website above.)

1. Shared Taxi

Amal Taxi operates shared taxi service from the Ben Gurion Airport to Haifa (costs 120NIS≈32 USD). Taxis are available (continuously regardless of Shabbat) coming out of gate 02 directly on the right side (Amal shuttle service Tel. 972-4-8662324). This taxi serves up to 10 passengers per vehicle and will leave the airport upon filling the cabin, so you may experience some waiting time until departure.

This service takes each passenger to her/his own destination, thus possibly prolonging the trip in Haifa itself.

2. Train

Please Note: Due to public transportation adherence to Shabbat observance, there is no train from Friday 14:30 to Saturday 19:30.

There is a direct train service between the Ben Gurion Airport and Haifa, with trains almost every hour all night long and, mostly, twice in an hour during the day.

The train station at Ben Gurion Airport is located on level S of the building, adjacent to the Greeters’ Hall. Tickets are purchased at the vendor or the automatic machines at the entrance to the train station (costs 45NIS≈$11).

Your destination is: “Haifa Hof HaKarmel” station. This is the first stop in the Haifa area.

From there take a taxi to the Dan Panorama Hotel or the Technion Guest House (The taxi should cost 60NIS≈15 USD).

3. Taxi

Taxis to Haifa are available via an official dispatcher just outside the arrival hall, to your left (do not use an unauthorized taxi). The ride will cost about 600 NIS ≈ 150USD) (the night fare is higher).

4. Car

Note: To enter the Technion campus by car, you must show the first page of this program to the guard in the gate.
Transportation from Dan Panorama Hotel

A bus will take you from Dan Panorama Hotel to the conference venue everyday. The bus will also take you back to the hotel after the daily program ends.

The bus will depart the Dan Panorama Hotel every day according to the following schedule:

- Sunday (04.09) – 18:30 PM. (Opening reception at 19:00 PM).
- Monday (05.09) – 08:30 AM. (Special dinner at 18:30 PM).
- Tuesday (06.09) – 09:00 AM. The bus will take you back to the hotel after lunch (free afternoon and evening for this day).
- Wednesday (07.09) – 09:00 AM. (Farewell dinner at 18:00 PM).
- Thursday (08.09) – 08:00 AM. We have an excursion to Rosh Hanikra and Acco.

Please be on time for the bus rides!

Contact Information

If you need any help, please feel free to contact

- Shoham Sabach + 972-528289009
- Mili Harari + 972-545959519
For Participants staying at the Forchheimer guesthouse

Walking to the Conference Venue – Bloomfield Bldg ~ 400 m
Invited Participants

Hedy Attouch (University of Montpellier, France)
Francis Bach (INRIA, Paris, France)
Jonathan Barzilai (Dalhousie University, Canada)
Heinz Bauschke (UBC Okanagan, Canada)
Adi Ben-Israel (Rutgers, USA)
Dimitris Bertsimas (Massachusetts Institute of Technology, USA)
Chiranjib Bhattacharyya (Indian Institute of Science, India)
Jerome Bolte (Toulouse 1 Capitole University, France)
Yair Censor (University of Haifa, Israel)
Roberto Cominetti (University of Chile, Chile)
Yoel Drori (Google R & D, Tel Aviv, Israel)
Yonina Eldar (Technion, Israel)
Robert M. Freund (Massachusetts Institute of Technology, USA)
Dick den Hertog (Tilburg University, Netherlands)
Alexander D. Ioffe (Technion, Israel)
Jean Bernard Lasserre (LAAS-CNRS, Toulouse, France)
Xin Liu (Chinese Academy of Sciences, China)
Russell Luke (Goettingen, Germany)
Aaron Melman (Santa Clara University, USA)
Yuji Nesterov (Université catholique de Louvain, Belgium)
Jong-Shi Pang (University of Southern California, USA)
Edouard Pauwels (University of Toulouse 3, France)
Georgia Perakis (Massachusetts Institute of Technology, USA)
Thomas Pock (Graz University of Technology, Austria)
Roman Polyak (Technion, Israel)
Simeon Reich (Technion, Israel)
Kees Roos (Delft University of Technology, Netherlands)
Katya Scheinberg (Lehigh University, USA)
Shimrit Shtern (Massachusetts Institute of Technology, USA)
Jean Philippe Vial (University of Geneva, Switzerland)
Henry Wolkowicz (University of Waterloo, Canada)
Yinyu Ye (Stanford, USA)
Xiaoming Yuan (Hong Kong Baptist University, Hong Kong)
Ya-Xiang Yuan (Chinese Academy of Sciences, China)
Michael Zibulevsky (Technion, Israel)
Conference Schedule

• Sunday, September 4
  
  19:00 - 21:00 Opening reception in Bloomfield building (2nd Floor)

• Monday, September 5

  09:00 - 09:20 Registration
  09:20 - 09:30 Welcome
  09:30 - 10:00 Adi Ben-Israel
        Matrix volume and its applications
  10:00 - 10:30 Yinyu Ye
        Recent progresses on linear programming
  10:30 - 11:00 Coffee Break
  11:00 - 11:30 Yurii Nesterov
        Universal Newton method
  11:30 - 12:00 Ya-Xiang Yuan
        Theory and application of p-regularized subproblems for $p > 2$
  12:00 - 12:30 Jean Bernard Lasserre
        Convex optimization and parsimony of $L_p$-ball representation
  12:30 - 14:00 Lunch
  14:00 - 14:30 Yair Censor
        Weak and strong superiorization: between feasibility-seeking and minimization
  14:30 - 15:00 Francis Bach
        Linearly-convergent stochastic gradient algorithms
  15:00 - 15:30 Dick den Hertog
        Robust optimization with ambiguous stochastic constraints under mean and dispersion information
  15:30 - 16:00 Coffee Break
  16:00 - 16:30 Thomas Pock
        A highly parallel dual minorize-maximize algorithm for discrete optimization
16:30 - 17:00 Xin Liu
Column-wise block coordinate descent approach for quadratic
minimization with orthogonal constraints

17:00 - 17:30 Katya Scheinberg
Convergence rate analysis of a stochastic trust region method for
nonconvex optimization

18:30 - 21:00 Conference Dinner Celebrating Ronny’s Birthday

• Tuesday, September 6

09:30 - 10:00 Jerome Bolte
From error-bounds to the complexity of first-order descent methods
for convex functions

10:00 - 10:30 Roberto Cominetti
Optimal convergence rates for Karsnoselskii-Mann fixed-point
iterations

10:30 - 11:00 Coffee Break

11:00 - 11:30 Robert M. Freund
New computational guarantees for first-order methods for convex
optimization, via a function growth constant

11:30 - 12:00 Yonina Eldar
Phase retrieval with application to optical imaging

12:00 - 13:30 Lunch

Free Afternoon/Evening

• Wednesday, September 7

09:30 - 10:00 Jong-Shi Pang
Computing B-stationary solutions of nonsmooth difference-of-convex
optimization problems with applications

10:00 - 10:30 Hedy Attouch
Fast splitting algorithms for convex optimization. Beyond Nesterov
complexity bound $O(1/k^2)$

10:30 - 11:00 Coffee Break
11:00 - 11:30 Alexander D. Ioffe
On the method of alternating projections for nonconvex sets

11:30 - 12:00 Heinz Bauschke
On the Douglas-Rachford algorithm

12:00 - 12:30 Dimitris Bertsimas
Machine learning and statistics via a modern optimization lens

12:30 - 14:00 Lunch

14:00 - 14:30 Simeon Reich
Outer approximation methods for solving variational inequalities in Hilbert space

14:30 - 15:00 Roman Polyak
Lagrangian transform and interior ellipsoid methods

15:00 - 15:30 Russell Luke
Quantifying convergence of Picard iterations

15:30 - 16:00 Coffee Break

16:00 - 16:30 Henry Wolkowicz
The many faces of degeneracy in conic optimization

16:30 - 17:00 Xiaoming Yuan
Some recent advances in primal-dual methods for saddle-point problems

17:00 - 17:30 Georgia Perakis
Optimization models for promotion planning; from vendor deals to pricing and promotion vehicle selection

18:00 - 20:30 Farewell dinner

• Thursday, September 8

08:30 - 16:00 Excursion
Abstracts

Hedy Attouch (University of Montpellier, France)

Fast splitting algorithms for convex optimization. Beyond Nesterov complexity bound $O(1/k^2)$

Abstract. Large scale optimization problems naturally appear in the modeling of many scientific and engineering situations. To meet the challenges posed by these issues, in recent years, considerable efforts have been devoted to the study of first-order splitting algorithms. The forward-backward algorithm, which is one of the most important, is a powerful tool for solving optimization problems with an additively separable and smooth plus nonsmooth structure. In the convex setting, a simple but ingenious acceleration scheme developed by Nesterov, and Beck-Teboulle improves the theoretical rate of convergence for the function values from the standard $O(k^{-1})$ down to $O(k^{-2})$. In this lecture, we show that the rate of convergence of a slight variant of this accelerated forward-backward method, which produces convergent sequences, is actually $o(k^{-2})$, rather than $O(k^{-2})$. Our arguments are based on the connection between this algorithm and a second-order differential inclusion with vanishing damping, recently introduced by Su, Boyd and Candès. The key point is the introduction of energy-like Lyapunov functions, with adapted scaling. Linking algorithms with dynamical systems provide connections between different areas, and a valuable guide for the proofs. Finally, we consider the hierarchical multi-objective problem which consists in finding by rapid methods the solution with minimum norm of a convex minimization problem. To this end, we introduce into the dynamics and algorithms a Tikhonov regularization term with vanishing coefficient. Applications are given in sparse optimization for signal/imaging processing, and inverse problems. We conclude by showing some recent directions of research, in particular the developments of these methods to nonconvex nonsmooth semi-algebraic problems, based on Kurdyka-Lojasiewicz inequality.

Francis Bach (INRIA - Ecole Normale Supérieure, Paris, France)

Linearly-convergent stochastic gradient algorithms

Abstract. Many machine learning and signal processing problems are traditionally cast as convex optimization problems where the objective function is a sum of many simple terms. In this situation, batch algorithms compute gradients of the objective function by summing all individual gradients at every iteration and exhibit a linear convergence rate for strongly-convex problems. Stochastic methods rather select a single function at random at every iteration, classically leading to cheaper iterations but with a convergence rate which decays only as the inverse of the number of iterations. In this talk, I will present the stochastic averaged gradient (SAG) algorithm which is dedicated to minimizing finite sums of smooth functions; it has a linear convergence rate for strongly-convex problems, but with an iteration cost similar to stochastic gradient descent, thus leading to faster convergence for machine learning problems. I will also mention several extensions, in particular to saddle-point problems, showing that this new class of incremental algorithms applies more generally.
Heinz Bauschke (University of British Columbia, Kelowna, Canada)

On the Douglas-Rachford algorithm

Abstract. The Douglas-Rachford algorithm is a popular splitting method for finding a minimizer of the sum of two convex (possibly nonsmooth) functions; or, more generally, a zero of the sum of two maximally monotone operators. My talk will focus on recent joint works on this algorithm.

Adi Ben-Israel (Rutgers, USA)

Matrix volume and its applications

Abstract. The volume $\text{vol}(A)$ of an $m \times n$ matrix $A$ of rank $r$ is defined as

(a) the product of the $r$ singular values of $A$, or

(b) the square root of the sum of squares of all $r \times r$ subdeterminants of $A$, or

(c) the volume of the image under $A$ of a unit cube in the range of $A$ transpose.

Definition (b) is applicable to non-numerical matrices, in particular to rectangular Jacobians. Some representative applications will be discussed.

References:


Matrix volume was suggested by work of Ronny and Marc some 25 years ago, where squares of subdeterminants appeared unexpectedly.

Dimitris Bertsimas (Massachusetts Institute of Technology, USA)

Machine learning and statistics via a modern optimization lens

Abstract. The field of Statistics has historically been linked with Probability Theory. However, some of the central problems of classification, regression and estimation can naturally be written as optimization problems. While continuous optimization approaches has had a significant impact in Statistics, mixed integer optimization (MIO) has played a very limited role, primarily based on the belief that MIO models are computationally intractable.

The period 1991-2015 has witnessed a) algorithmic advances in mixed integer optimization (MIO), which coupled with hardware improvements have resulted in an astonishing 800 billion
factor speedup in solving MIO problems, b) significant advances in our ability to model and solve very high dimensional robust and convex optimization models.

In this talk, we demonstrate that modern convex, robust and especially mixed integer optimization methods, when applied to a variety of classical Machine Learning (ML)/Statistics (S) problems can lead to certifiable optimal solutions for large scale instances that have often significantly improved out of sample accuracy compared to heuristic methods used in ML/S. Specifically, we report results on

1) The classical variable selection problem in regression currently solved by Lasso heuristically.

2) We show that robustness and not sparsity is the major reason of the success of Lasso in contrast to widely held beliefs in ML/S.

3) A systematic approach to design linear and logistic regression models based on MIO.

4) Optimal trees for classification solved by CART heuristically.

5) Robust classification including robust Logistic regression, robust optimal trees and robust support vector machines.


In all cases we demonstrate that optimal solutions to large scale instances (a) can be found in seconds, (b) can be certified to be optimal in minutes and (c) outperform classical approaches. Most importantly, this body of work suggests that linking ML/S to modern optimization will lead to significant advantages.

Jerome Bolte (Toulouse 1 Capitole University, France)

*From error-bounds to the complexity of first-order descent methods for convex functions*

**Abstract.** We will show that error bounds can be turned into effective tools for deriving complexity results of first order methods for convex minimization. This led us to study the interplay between error bounds and KL inequality. We showed the equivalence between the two concepts for functions having a profile moderately flat near the minimizers set (as those of functions with Hölderian growth). A counterexample show the relevance of our approach since the equivalence is no longer true for extremely flat functions.

In a second stage, we show how KL inequalities can in turn be used to compute new complexity bounds for a wealth of descent methods. Our method is completely original and makes use of a one dimensional worst case proximal sequence in the spirit of the famous majorant method of Kantorovich. Our result applies to a very simple abstract scheme that covers a very wide class of descent methods.
Our approach inaugurates a simple methodology: derive an error bound, compute the desingularizing function whenever possible, identify essential constants in the descent method and finally compute the complexity using the one dimensional worst case proximal sequence. Our method is illustrated through the famous iterative thresholding algorithm, also known as ISTA, for which we show that the complexity bound is of the form $O(q^{2k})$ where the constituents of the bound only depend on error bounds constants obtained for the usual objective. This talk is based on joint work with T.P Nguyen, J. Peypouquet and B. Suter.

Yair Censor (University of Haifa, Israel)

*Weak and strong superiorization: between feasibility-seeking and minimization*

**Abstract.** Superiorization can be thought of, in some cases, as lying between feasibility-seeking and constrained minimization. It is not trying to solve the full fledged constrained minimization problem; rather, the task is to find a feasible point which is superior, with respect to an objective function value, to one returned by a feasibility-seeking only algorithm. We distinguish between weak superiorization and strong superiorization and clarify their nature. Superiorization has been successfully used in some important practical applications such as image reconstruction from projections, intensity-modulated radiation therapy and nondestructive testing, and awaits to be implemented and tested in additional fields.

Roberto Cominetti (Universidad Adolfo Ibáñez, Chile)

*Optimal convergence rates for Krasnoselskii-Mann fixed-point iterations*

**Abstract.** A classical method to approximate a fixed point for a non-expansive map $T : C \to C$ is the Krasnoselskii-Mann iteration

\[(KM) \quad x_{n+1} = (1 - \alpha_{n+1})x_n + \alpha_{n+1}Tx_n.\]

We use optimal transport to establish a recursive formula to estimate the distance between iterates $\|x_m - x_n\| \leq d_{mn}$. These bounds $d_{mn}$ induce a metric on the integers that allows to study the rate of convergence of $(KM)$. As a result, we settle Baillon-Brucks conjecture: for every non-expansive map in any normed space the following estimate holds with $\kappa = 1/\sqrt{\pi}$

\[\|x_n - Tx_n\| \leq \kappa \frac{\text{diam}(C)}{\sqrt{\sum_{i=1}^{n} \alpha_k(1 - \alpha_k)}}.\]

The analysis exploits an unexpected connection with discrete probability and combinatorics, related to the Gambler’s ruin for sums of non-homogeneous Bernoulli trials. We also show that the constant $\kappa = 1/\sqrt{\pi}$ is sharp for nonlinear maps, and we provide an improved sharp bound for linear affine maps.
Yonina Eldar (Technion, Israel)

Phase retrieval with application to optical imaging

Abstract. The problem of phase retrieval, namely the recovery of a function given the magnitude of its Fourier transform - arises in various fields of science and engineering, including electron microscopy, crystallography, astronomy, and optical imaging. Due to the loss of Fourier phase information, this problem is generally ill-posed. In this talk we review several modern methods for treating the phase retrieval problem including matrix lifting, structured illumination and short-time Fourier measurements. We also consider techniques that exploit sparsity on the input together with contemporary optimization tools to further facilitate recovery. We then illustrate the use of these methods in several different problems arising in optical imaging.

Robert M. Freund (Massachusetts Institute of Technology, USA)

New computational guarantees for first-order methods for convex optimization, via a function growth constant

Abstract. We present new algorithms and complexity bounds for solving convex optimization problems using first-order methods. We presume we are given a strict lower bound on the optimal value $f^*$. We introduce a new functional measure called the growth constant $G$ for $f(x)$ that measures how quickly the function level sets grow and that plays a fundamental role in the complexity analysis. We present new computational guarantees for non-smooth and smooth optimization that improves on existing complexity bounds in many ways. This talk is based on joint work with Haihao (Sean) Lu.

Dick den Hertog (Tilburg University, The Netherlands)

Robust optimization with ambiguous stochastic constraints under mean and dispersion information

Abstract. We consider ambiguous stochastic constraints under partial information consisting of means and dispersion measures of the underlying random parameters. Whereas the past literature used the variance as the dispersion measure, here we use the mean absolute deviation

References


[2] R. Cominetti, M. Bravo [2015], The optimal constant of asymptotic regularity is $\frac{1}{\sqrt{\pi}}$, in preparation.
from the mean (MAD). This makes it possible to use the 1972 result of Ben-Tal and Hochman (BH) in which tight upper and lower bounds on the expectation of a convex function of a random variable are given. First, we use these results to treat ambiguous expected feasibility constraints. This approach requires, however, the independence of the random variables and, moreover, may lead to an exponential number of terms in the resulting robust counterparts. We then show how upper bounds can be constructed that alleviate the independence restriction and require only a linear number of terms, by exploiting models in which random variables are linearly aggregated. Moreover, using the BH bounds we derive new safe tractable approximations of chance constraints. In a numerical study, we demonstrate the efficiency of our methods in solving stochastic optimization problems under mean-MAD ambiguity.

This is joint work with Krzysztof Postek (Tilburg University), Aharon Ben-Tal (Technion), and Bertrand Melenberg (Tilburg University).

We also study several classes of multi-stage stochastic programming problems under distributional uncertainty, under the same distributional information as above. We show how the above mentioned BH bounds can be applied to these problems. We also propose convex approximations of problems with integer later-stage variables for which explicit performance bounds are derived. The BH bounds can be applied to these convex approximations. Several techniques that can be used to reduce the computational effort related to the distributional uncertainty are discussed.

This is joint work with Krzysztof Postek (Tilburg University), Ward Romeijnders (University of Groningen), and Maarten van der Vlerk (University of Groningen).

Alexander D. Ioffe (Technion, Israel)

On the method of alternating projections for nonconvex sets

Abstract. I shall discuss recent results and open questions connected with linear convergence of the method of alternating projections for nonconvex sets. This talk is based on joint work with D. Drusvyatskiy and A. Lewis.

Jean Bernard Lasserre (LAAS-CNRS, Toulouse, France)

Convex optimization and parsimony of $L_p$-ball representation

Abstract. With fixed integer $2p$, we consider the family $F$ of homogeneous polynomials of degree $2p$ such that the associated “unit-ball” $G = \{x : g(x) \leq 1\}$ has fixed volume (e.g., the Lebesgue volume of the $L_{2p}$-unit ball; in particular each “$g$” is nonnegative). We then consider the optimization problem $P$: Find the polynomial “$g$” in $F$ that minimizes the parsimony-inducing $\ell_1$-norm of coefficients, i.e., $\min\{\|g\|_1 : g \in F\}$. We then show that the polynomial $g^*$ associated with the $L_{2p}$-unit ball is the unique solution of this problem. And indeed $g^*$ is very “sparse”! We next discuss the solution of the same problem with now the $\ell_2$-norm and the trace norm in the sub-family of sums-of-squares homogeneous polynomials. Surprisingly the standard Euclidean unit-ball is the unique optimal solution.
Xin Liu (Chinese Academy of Sciences, China)

Column-wise block coordinate descent approach for quadratic minimization with orthogonal constraints

Abstract. We propose a column-wise block coordinate descent (BCD) approach for solving general quadratic minimization problems with orthogonal constraints. This approach consists of a Gauss-Seidel type BCD iteration with a multiplier symmetrisation step. We can prove the subsequence convergence of the proposed approach to the stationarity. Iterate convergence can be guaranteed in some special cases. Preliminary numerical experiments illustrate that our new algorithm performs much better than the existent solvers in solving a large class of testing problems.

Russell Luke (Goettingen, Germany)

Quantifying convergence of Picard iterations

Abstract. It can be observed in many applications that algorithms converge unexpectedly well, despite any theory to explain their success. An analysis of algorithm complexity in terms of function values is widespread and can be applied readily in many practical instances. However, this strategy of analysis cannot explain the observed convergence to stationary points of iterates in many cases, nor can it yield error estimates on the distance to locally optimal solutions as opposed to locally optimal values. We propose a framework for quantifiable local convergence analysis of iterations of expansive fixed point mappings. The key to this analysis is the well-established notion of metric subregularity of the mapping at fixed points, what we simply call metric regularity on a semi-pointwise set on the graph of the mapping. Within this framework we show that a relaxed version of metric regularity together with a type of calmness of the fixed point mapping is necessary for local linear convergence. To demonstrate the theory, we prove for the first time a number of results showing local linear convergence of cyclic projections for (possibly) inconsistent feasibility problems, local linear convergence of the forward-backward algorithm for structured optimization without convexity, strong or otherwise, and local linear convergence of a nonconvex application of the Douglas-Rachford algorithm for minimization. Known results, convex and nonconvex, are recovered in this framework.

Yurii Nesterov (CORE/INMA UCL, Belgium)

Universal Newton method

Abstract. In this talk we present a second-order method for unconstrained minimization of convex functions. It can be applied to functions with Holder continuous Hessians. Our main scheme is the Cubic Regularization of Newton Method, equipped with a special line-search procedure. We show that the global rate of convergence of this scheme depends continuously on the smoothness parameter. Thus, our method can be used even for minimizing functions with discontinuous Hessians. At the same time, the line-search procedure is very efficient: the average number of calls of oracle per iteration is equal to two. We show that for finding a
point with small norm of the gradient, the Universal Newton Method must be equipped with a special termination criterion for the line-search.

Jong-Shi Pang (University of Southern California, USA)

*Computing B-stationary solutions of nonsmooth difference-of-convex optimization problems with applications*

**Abstract.** We discuss a nonsmooth, difference-of-convex (dc) minimization problem and its applications to statistical learning and digital communication. Topics include: (i) clarification of several kinds of stationary solutions and their relations; (ii) development and demonstration of the convergence of a novel algorithm for computing a d-stationary solution of a problem with a convex feasible set that is arguably the sharpest kind among the various stationary solutions; (iii) extension of the algorithm in several directions including: a randomized choice of the subproblems that could help the practical convergence of the algorithm, a distributed penalty approach for problems whose objective functions are sums of dc functions, and problems with a specially structured (nonconvex) dc constraint. For the latter class of problems, a pointwise Slater constraint qualification is introduced that facilitates the verification and computation of a B(ouligand)-stationary point.

Georgia Perakis (Massachusetts Institute of Technology, USA)

*Optimization models for promotion planning; from vendor deals to pricing and promotion vehicle selection*

**Abstract.** Promotions are a key instrument for driving sales for retailers. The promotion planning process is complex. It starts with vendor fund deals between manufacturers and retailers in order to encourage the retailer to promote their products. The next step of the process involves pricing, that is, when to promote, which items and how deeply to reduce their price. The final step involves scheduling what vehicle is best to use each time period to promote which item. In this talk we will discuss how to use optimization for these three steps of the promotion planning process. We will focus mostly on the first and last step of the process; namely, vendor fund deals and scheduling promotion vehicles. We will discuss a model for the vendor fund selection problem of the retailer taking into account the impact on promotional pricing. We will propose and analyze an approximation algorithm with analytical guarantees for solving this problem. Furthermore, for the vehicle selection problem we will introduce a greedy algorithm as well as an integer optimization approach in order to solve the vehicle selection problem in a tractable way. We will illustrate that the methods we introduce are computationally efficient and “easy to use in practice. We will also illustrate the performance of our methods using data through our collaboration with the Oracle Retail Business Unit. This talk is based on joint work with Lennart Baardman (ORC PhD student), Maxime Cohen, (recently graduated ORC PhD), Jeremy Kalas (EECS Undergraduate), Zachary Leung (recently graduated ORC PhD), Danny Segev (U. Haifa) as well as Kiran Panchamgam (Oracle RGBU).
Thomas Pock (Graz University of Technology, Austria)

A highly parallel dual minorize-maximize algorithm for discrete optimization

Abstract. In this work we propose a new highly parallel algorithm to solve a class of discrete optimization problems frequently arising in computer vision and image processing. The idea is to solve a dual of the problem that is obtained from a Lagrangian decomposition into horizontal and vertical chains. We then propose a new parallel coordinate ascent on the dual that is based on maximizing a sequence of minorants of the dual problem. The algorithm allows for a highly parallel implementation on recent graphics processing units, which leads to real-time performance for stereo problems. This talk is based on joint work with A. Shekhovtsov, G. Graber, and C. Reinbacher.

Roman Polyak (Technion, Israel)

Lagrangian transform and interior ellipsoid methods

Abstract. We consider a class $\Psi$ of monotone increasing, strictly concave and smooth functions with particular properties. The Lagrangian Transform (LT) scheme employs $\Psi$ to transform terms of the Classical Lagrangian, which correspond to the constraints. The transformation is scaled by a positive scaling parameter. The LT method at each step finds the primal minimizer of LT following by the Lagrange multipliers update, while the scaling parameter can be fixed or updated from step to step. The LT is equivalent to the prox-method with Bregman’s or Bregman’s type distance based on the kernel for the dual problem. The equivalence of the prox to the interior ellipsoid methods is the key ingredient of the convergence analysis. In particular, the MBF transformation leads to the dual prox with Bregman distance, moreover the MBF kernel is a Self-Concordant function, therefore the correspondent interior ellipsoids are Dikins ellipsoids. Application of LT for LP calculations leads to the affine scaling type method for the dual LP.

Simeon Reich (Technion, Israel)

Outer approximation methods for solving variational inequalities in Hilbert space

Abstract. We study variational inequalities in a real Hilbert space, which are governed by a strongly monotone and Lipschitz continuous operator $F$ over a closed and convex set $C$. We assume that the set $C$ can be outerly approximated by the fixed point sets of a sequence of certain quasi-nonexpansive operators called cutters. We propose an iterative method the main idea of which is to project at each step onto a particular super-half-space constructed by using the input data. Our approach is based on a method presented by Fukushima in 1986, which has recently been extended by several authors. We establish strong convergence in Hilbert space. To the best of our knowledge, Fukushima’s method has so far been considered only in the Euclidean setting with different conditions on $F$. We also provide numerical illustration of our theoretical results. This talk is based on joint work with Aviv Gibali and Rafal Zalas.
Katya Scheinberg (Lehigh University, USA)

*Convergence rate analysis of a stochastic trust region method for nonconvex optimization*

**Abstract.** We introduce a variant of a traditional trust region method which is aimed at stochastic optimization. While traditional trust region method relies on exact computations of the gradient and values of the objective function, our method assumes that these values are available up to some dynamically adjusted accuracy. Moreover, this accuracy is assumed to hold only with some sufficiently large, but fixed, probability, without any additional restrictions on the variance of the errors. We show that our assumptions apply to the standard stochastic setting assumed in the machine learning problems, but also include more general settings. We provide a bound on the expected number of iterations the stochastic algorithm requires to reach accuracy $\|\nabla f(x)\| \leq \epsilon$, for any $\epsilon > 0$. We recover the standard first order complexity $O(\epsilon^{-2})$, under the assumption of sufficiently accurate stochastic gradient. This talk is based on joint work with Jose Blanchet, Coralia Cartis and Matt Menickelly.

Henry Wolkowicz (University of Waterloo, Canada)

*The many faces of degeneracy in conic optimization*

**Abstract.** New optimization modeling techniques and convex relaxations for hard nonconvex problems have shown that the loss of strict feasibility is a much more pronounced phenomenon than has previously been realized. It occurs often in semidefinite programming relaxations of hard combinatorial problems. We describe the various reasons for the loss of strict feasibility, whether due to poor modeling choices or (more interestingly) rich underlying structure, and describe ways to cope with it and, in many pronounced cases, how to use it as an advantage.

Yinyu Ye (Stanford, USA)

*Recent progresses on linear programming*

**Abstract.** We describe recent algorithmic progresses on linear programming as well as research explorations of alternative algorithms. The topics include polynomial-time simplex and policy iteration methods for the Markov Decision-Game Process, faster interior-point algorithms for network flows and linear programming, convergence of the multi-block alternating direction method of multipliers for constrained convex optimization, and, in particular, first-order potential reduction algorithms for linear programming and their preliminary computational results.

Xiaoming Yuan (Hong Kong Baptist University, Hong Kong)

*Some recent advances in primal-dual methods for saddle-point problems*

**Abstract.** The primal-dual hybrid gradient (PDHG) method is not new, but recently has found many sparsity-driven applications in various areas, particularly in image processing and statistical learning. Our earlier work in 2012 shows that the PDHG can be explained as an
application of the proximal point algorithm; this fact has enabled us to find more results of PDHG such as its convergence when one function of a saddle-point problem is strongly convex. In this talk, I will present some of our recent results on this topic, including designing an algorithmic framework for generalized PDHG methods and proposing inexact versions of PDHG with provable convergence.

Ya-Xiang Yuan (Chinese Academy of Sciences, China)

Theory and application of p-regularized subproblems for \( p > 2 \)

Abstract. The \( p \)-regularized subproblem \((p-RS)\) is the key content of a regularization technique in computing a Newton-like step for unconstrained optimization. The idea is to globally minimize a local quadratic approximation of the objective function at each iteration, while incorporating a weighted regularization term \( \frac{\sigma}{p} \| x \|_p^p \). In this paper, we establish a complete theory of the \( p \)-regularized subproblems for general \( p > 2 \) that cover previous known results on \( p = 3 \) or \( p = 4 \). The theory features necessary and sufficient optimality conditions for the global and also for the local non-global minimizers of \((p-RS)\). It gives a closed-form expression of the global minimum set, which facilitates the computation of \((p-RS)\) at each step. We prove that \((p-RS), p > 2\) can have at most one local non-global minimizer. In the end, we develop an algorithm for solving \((p-RS)\) with \( m \) additional linear inequality constraints, denoted by \((p-RS_m)\). Our theory indicates that \((p-RS)\) have all properties that the trust region subproblems (TRS) do. The free interchange between \((p-RS)\) and (TRS) thus provides flexibility in formulation and approximation for optimization models. In application, we point out that \((p-RS)\) can appear in natural formulation for optimization problems. There are two examples. One is to utilize the Tikhonov regularization to stabilize the least square solution for an over-determined linear system; and the other comes from numerical approximations to the generalized Ginzburg-Landau functionals. We also show that the partition problem and the \( k \)-dispersion-sum problem in combinatorial optimization are both equivalent to special types of \((p-RS_m)\) with \( p = 4 \).
Further Practical Informations

Currency: New Israeli Shekel (NIS). In short ”Shekel”. All major currencies are exchangeable. Current exchange varies: 1 USD = 3.8 to 4.0NIS, 1 Euro = 4.2 to 4.4NIS

All major credit cards are accepted in shops, restaurants ect. (not always by Taxi drivers) and are not accepted in public transportation.

Automatic bank machines for cash on Campus: Bank Leumi located on campus (see map). For any other transactions such as traveller checks. The Bank hours are: Sunday, Tuesday and Wednesday: 08:30-13:30; Monday and Thursday: 08:30-12:30 and 15:00-17:00.

Electricity: The electric current in Israel is 220 volts AC, single phase, 50 Hz.

Most Israeli sockets are of the three-pronged variety (Type H). In most cases, it accepts the European two-pronged plugs (Types E and F), but does not accept the US Type A. See http://www.worldstandards.eu/electricity/plugs-and-sockets/

Internet: You have free access to internet at the hotel/guest house, and public WiFi at the Technion Campus.

Post Office: Situated on campus below the main library Building. It is open Sunday to Thursday from 08:00 to 15:00.

Student Store Michlol: Is located on campus and is open Sunday-Thursday 8:00-18:00; Friday 8:00-12.00.

Medical First Aid on Campus: A campus medical clinic (including a doctor) is situated below the Central Library (next to the post office). For Medical Emergencies dial 101 from any phone. Haifa has three excellent hospitals and many medical clinics. Campus security can be reached at 04-8294242 or 04-8293590 or internal extension 4242 or 3590 from any campus phone.

Tourism: For special attractions in Haifa, please visit the following website information. More places that worth visit (among other) are Jerusalem, Massada and Dead Sea. You might wish to have a look at the following site https://www.beinharimtours.com/ which organizes very nice tours and where you will find much more options.

Public Buses

There are several bus lines to the Technion. In particular, there is an hourly direct bus #31 between the Central Carmel district (where Dan Panorama Hotel is located) and the Technion. A single bus ride costs 6.6NIS. For more information please visit Egged public transportation web site where you can find the appropriate bus line that runs from the station closest to you (see conference website link).