ESI Erwin Schrödinger International Institute for Mathematics and Physics

universität wien

DVR 0065528

Programme on

"Modern Maximal Monotone Operator Theory: From Nonsmooth Optimization to Differential Inclusions"

January 28 – March 8, 2019

organized by

Heinz H. Bauschke (U of British Columbia), Radu I. Boţ (U Vienna), Hélène Frankowska (U Paris VI), Michael Hintermüller (Weierstrass Inst. Berlin), D. Russell Luke (U Göttingen)

Workshop 2 on

"Numerical Algorithms in Nonsmooth Optimization"

February 25 to March 1, 2019

Abstracts

Hedy Attouch (University of Montpellier)

Relaxed inertial proximal algorithms for monotone inclusions

In a Hilbert space H, given A a maximally monotone operator acting on H, we study the convergence properties of a general class of relaxed inertial proximal algorithms. This study aims to extend to the case of monotone inclusions the acceleration techniques initially introduced by Nesterov in the case of convex minimization. As a guideline, we use the interpretation of these algorithms as temporal discretized versions of inertial dynamical systems with vanishing damping. In the general monotone case, the relaxed form of the proximal algorithms plays a central role. It comes naturally with the regularization of the operator A by its Yosida approximation with a variable parameter, a technique recently introduced by Attouch-Peypouquet for a particular class of inertial proximal algorithms. Our study provides an algorithmic version of the convergence results obtained by Attouch-Cabot in the case of continuous dynamical systems. Then, we specialize our study on nonsmooth convex optimization, and show fast convergence properties of these algorithms. We also present new accelerated proximal algorithms based on inertial gradient dynamics which have been rescaled in time. In doing so, we improve and obtain a dynamic interpretation of the seminal papers of Güler on the convergence rate of the proximal methods for convex optimization.

Sebastian Banert (KTH Royal Institute of Technology)

How to accelerate convex optimisation with machine learning

In this talk, we are going to present some ideas how to design algorithms for convex optimisation with possibly nonsmooth functions and choose optimal parameters for them with the help of recent deep learning techniques.

The main point of this talk will be that one still can get convergence guarantees for the neural networks resulting from this procedure. We will demonstate their performance in variational regularisation of inverse problems in imaging.

This talk will present joint work with Axel Ringh, Jonas Adler, Jevgenija Rudzusika, Johan Karlsson, and Ozan Öktem.

Amir Beck (Tel-Aviv University)

On the convergence to stationary points of deterministic and randomized feasible descent directions methods

We study the class of nonsmooth nonconvex problems in which the objective is to minimize the difference between a continuously differentiable function (possibly nonconvex) and a convex (possibly nonsmooth) function over a convex polytope. This general class contains many types of problems, including difference of convex functions (DC) problems, and as such, can be used to model a variety of applications. Our goal is to attain a point satisfying the stationarity necessary condition, defined as the lack of feasible descent directions. In this work we prove that stationarity in our model can be characterized by a finite number of directions that positively span the set of feasible directions. We then use the latter to develop two methods, one deterministic and one random, whose accumulation points are proved to be stationary points. Supplementing discussion on positively spanning sets, and corresponding methods to obtain members of such sets, are also presented. Numerical experiments illustrate the appeal of obtaining a stationary point and the advantage of using the random method to do so.

Jérôme Bolte (University Toulouse I Capitole)

A multi-proximal method for convex composite optimization

Composite minimization involves a collection of smooth functions which are aggregated in a nonsmooth manner. We design an algorithm by linearizing each smooth component in accordance with its main curvature. The resulting method, called the multiprox method, consists in solving successively simple problems (e.g. constrained quadratic problems) which can also feature some proximal operators. To study the complexity and the convergence of this method we introduce a new type of convex qualification condition. We also obtain explicit complexity results involving new types of constant terms. A distinctive feature of our approach is to be able to cope with oracles featuring moving constraints. Our method generalizes the proximal Gauss-Newton's method, the moving balls method or the forward-backward splitting, for which we recover known complexity results or establish new ones.

Immanuel Bomze (University of Vienna)

Non-convex min-max fractional quadratic problems under quadratic constraints: copositive relaxations

We address a min-max problem of fractional quadratic (not necessarily convex) over linear functions on a feasible set described by linear and (not necessarily convex) quadratic functions. We propose a conic reformulation on the cone of completely positive matrices. By relaxation, a doubly non negative conic formulation is used to provide lower bounds with evidence of very small gaps. It is known that in many solvers using Branch and Bound the optimal solution is obtained in early stages and a heavy computational price is paid in the next iterations to obtain the optimality certificate. To reduce this effort, tight lower bounds are crucial. We will show empirical evidence that lower bounds provided by the copositive relaxation are able to substantially speed up a well known solver in obtaining the optimality certificate.

This is joint work with Paula Amaral (Univ. Nova de Lisboa, Portugal).

Coralia Carţiş (University of Oxford)

Dimensionality reduction techniques for global optimization

We show that the scalability challenges of Global Optimisation (GO) algorithms can be overcome for functions with low effective dimensionality, which are constant along certain linear subspaces. Such functions can often be found in applications, for example, in hyper-parameter optimization for neural networks, heuristic algorithms for combinatorial optimization problems and complex engineering simulations. We propose the use of random subspace embeddings within a(ny) global minimisation algorithm, extending the approach in Wang et al (2013). Using tools from random matrix theory and conic integral geometry, we investigate the success rates of our low-dimensional embeddings of the original problem, in both a static and adaptive formulation, and show their independence on the (large) ambient dimension of the problem. We illustrate our algorithmic proposals and theoretical findings numerically, using state of the art global solvers.

This work is joint with Adilet Otemissov (Turing Institute, London and Oxford University).

Volkan Cevher (École Polytechnique Fédérale de Lausanne)

Storage optimal semidefinite programming

Semidefinite convex optimization problems often have low-rank solutions that can be represented with O(p)-storage. However, semidefinite programming methods require us to store the matrix decision variable with size O(p^2), which prevents the application of virtually all convex methods at large scale.

Indeed, storage, not arithmetic computation, is now the obstacle that prevents us from solving large-scale optimization problems. A grand challenge in contemporary optimization is therefore to design storage-optimal algorithms that provably and reliably solve large-scale optimization problems in key scientific and engineering applications. An algorithm is called storage optimal if its working storage is within a constant factor of the memory required to specify a generic problem instance and its solution.

So far, convex methods have completely failed to satisfy storage optimality. As a result, the literature has largely focused on storage optimal non-convex methods to obtain numerical solutions.

To this end, my talk introduces a new convex optimization algebra to obtain numerical solutions to semidefinite programs with a low-rank matrix streaming model. This streaming model provides us an opportunity to integrate sketching as a new tool for developing storage optimal convex optimization methods that go beyond semidefinite programming to more general convex templates.

We then propose a practical inexact augmented Lagrangian method for non-convex problems with nonlinear constraints and contrast this approach to the convex one. We characterize the total computational complexity of the non-convex method subject to a verifiable geometric condition.

Patrick Combettes (North Carolina State University)

Between subdifferentials and monotone operators

We discuss various aspects of the gap between subdifferentials and monotone operators as well as the necessity of the general theory of monotone operators in convex optimization.

Aris Daniilidis (University of Chile)

Self-contracted curves and extensions

The class of self-contracted curves encompasses all (smooth) curves that are orbits of the gradient flow of a smooth quasiconvex function, all (absolutely continuous) curves that that orbits of the subgradient flow of a continuous convex function and all polygonal curves that are obtained by the proximal sequence on a convex continuous function. Self-contracted curves enjoy a simple metric definition and have been the object of the study of several mathematicians working on different domains. In this talk we shall present main ideas, extensions and consequences in numerical optimization.

Minh Đào (University of Newcastle)

Adaptive Douglas-Rachford splitting algorithm and applications

The Douglas–Rachford algorithm is a classical and powerful splitting method for minimizing the sum of two convex functions and, more generally, finding a zero of the sum of two maximally monotone operators. Although this algorithm has been well understood when the involved operators are monotone or strongly monotone, the convergence theory for weakly monotone settings is far from being complete. In this work, we propose an adaptive Douglas–Rachford splitting algorithm for the sum of two operators, one of which is strongly monotone while the other one is weakly monotone. With appropriately chosen parameters, the algorithm converges globally to a fixed point from which we derive a solution of the problem. When one operator is Lipschitz continuous, we prove global linear convergence which sharpens recent known results.

Jonathan Eckstein (Rutgers University)

Projective splitting with cocoercive operators

This talk describes a new variant of projective splitting in which cocoercive operators can be processed with a single forward step per iteration. This result establishes a symmetry between projective splitting algorithms, the classical forward-backward method, and Tseng's forward-backward-forward method: in a situation in which Lipschitz monotone operators require two forward steps, cocoercive operators may be processed with a single forward step. The single forward step may be interpreted as a single step of the classical gradient method for the standard "prox" problem for the cocoercive operator, starting at the previous known point in the operator graph. Proving convergence of the algorithm requires some departures from the usual proof framework for projective splitting.

Joint work with Patrick Johnstone

Pontus Giselsson (Lund University)

Performance estimation in operator splitting methods

We propose a methodology for studying the performance of common operator splitting methods through semidefinite programming. We prove tightness of the methodology, meaning that the semidefinite program is guaranteed to exactly capture the worst case behavior of the studied splitting method. We demonstrate the value of the approach by using it as a tool for computer-assisted proofs to prove tight analytical contraction factors for Douglas–Rachford splitting that are likely too complicated to find bare-handed.

Michael Hintermüller (Humboldt-University of Berlin and WIAS)

(Pre)Dualization, dense embeddings of convex sets, and applications in image processing

For a class of non smooth minimization problems in Banach spaces, predualization results and their connection to dense embeddings of convex sets are discussed. Motivating applications are related to nonsmooth filters in mathematical image processing. For this problem class also some numerical aspects are highlighted including primal/dual splitting or ADMM-type methods as well as proper (numerical) dissipation reducing discretization.

Christian Kanzow (Universität Würzburg)

Safeguarded augmented Lagrangian methods in finite and infinite dimensions

The classical augmented Lagrangian method belongs to the standard approaches for the solution of constrained optimization problems. Recent modifications by Andreani, Birgin, Martinez and Co-workers for finite-dimensional optimization problems turn out to have stronger properties than the classical method. In this talk, we begin with a review of these classical and modified augmented Lagrangian methods. We then present an extension of the modified method to optimization problems in Banach spaces. The global and local convergence properties are discussed including a result on strong local convergence under a second-order sufficiency condition (without assuming any constraint qualification). Some numerical results illustrate the reliability of the proposed technique.

This talk is based on joint work with Daniel Steck.

The research was supported by the German Research Foundation (DFG) within the priority program “Non-smooth and Complementarity-based Distributed Parameter Systems: Simulation and Hierarchical Optimization” (SPP 1962) under grant number KA 1296/24-1.

Karl Kunisch (University of Graz)

Monotone and primal-dual algorithms for optimization problems involving ℓ^p and ℓ^p-like functionals with $p \in [0, 1)$

Nonsmooth nonconvex optimization problems involving the ℓ^p quasi-norm, $p \in [0, 1)$, are the focus of this talk. Two schemes are presented and analyzed, and their performance in practice is discussed: a monotonically convergent scheme and a primal dual active set scheme. The latter heavily relies on a non-standard formulation of the first order optimality conditions. Numerical tests include an optimal control problem, models from fracture mechanics and microscopy image reconstruction. We also remark infinite horizon closed loop optimal control problems with ℓ^p cost.

Szilárd László (Technical University of Cluj-Napoca)

A gradient type algorithm with backward inertial steps for a nonconvex minimization

We investigate an algorithm of gradient type with a backward inertial step in connection with the minimization of a nonconvex differentiable function. We show that the generated sequences converge to a critical point of the objective function, if a regularization of the objective function satisfies the Kurdyka-Łojasiewicz property. Further, we provide convergence rates for the generated sequences and the objective function values formulated in terms of the Łojasiewicz exponent. Finally, some numerical experiments are presented in order to compare our numerical scheme with some algorithms well known in the literature.

Joint work with A. Viorel and C. Alecsa.

D. Russell Luke (University of Göttingen)

Convergence analysis of algorithms for inconsistent nonconvex feasibility

A recent paper by Luke, Thao and Tam (Math. Oper Res. 2018) establishes a framework for local analysis of fixed point algorithms in nonconvex settings. Motivated by problems in nonconvex imaging, we apply this analysis to the cyclic projections and relaxed Douglas-Rachford algorithms for inconsistent nonconvex feasibility problems. To address the relaxed Douglas-Rachford algorithm, we introduce a new type of regularity of sets, called super-regular at a distance, to establish sufficient conditions for local linear convergence of the corresponding sequences. Our results subsume and extend existing results for both of these algorithms.

Yura Malitsky (University of Göttingen)

On a new method for monotone inclusions

We present a novel method for monotone inclusions and discuss its connection with other popular methods. The method needs only one forward and one backward step in every iteration. The inspiration for our method comes from dynamical system approach.

Shin-ya Matsushita (Akita Prefectural University)

Rates of asymptotic regularity for the forward-backward splitting algorithm

We consider the forward-backward splitting algorithm for minimizing the sum of a proper, lower semi-continuous and convex function and a differentiable and convex function whose gradient is Lipschitz continuous. It is known that the forward-backward splitting algorithm can be interpreted as the Krasnosel'skiĭ-Mann (KM) iteration and the rates of asymptotic regularity for the KM iteration have been studied. The aim of this talk is to give rates of asymptotic regularity for the forward-backward splitting algorithm in Hilbert spaces.

Walaa Moursi (Stanford University)

Reflected resolvents in the Douglas-Rachford algorithm: order of the operators and linear convergence

The Douglas-Rachford algorithm is a popular method for finding zeros of sums of monotone operators. By its definition, the Douglas-Rachford operator is not symmetric with respect to the order of the two operators. In this talk we report on a systematic study of the two possible Douglas-Rachford operators. We show that the reflectors of the underlying operators act as bijections between the fixed points sets of the two Douglas-Rachford operators. Some elegant formulae arise under additional assumptions. Results on linear rates of convergence are also presented. Various examples illustrate our results.

Ion Necoară (Politehnica University of Bucharest)

Minibatch stochastic first order methods for composite convex optimization

Many problems from engineering, statistics and machine learning can be formulated as stochastic composite optimization problems. We propose a general framework for the analysis of stochastic composite convex optimization that includes the most well-known classes of objective functions analyzed in the literature: non-smooth functions, and composition of a (potentially) non-smooth function and a smooth function, with or without a quadratic functional growth property. Based on this framework we derive convergence analysis for the most known classes of minibatch stochastic first order methods, that is the minibatch stochastic proximal gradient and minibatch proximal point algorithms. These schemes are typically the method of choice in practice due to their cheap iteration and superior empirical performance. Usually, the convergence theory of these methods have been derived for simple stochastic optimization models, the rates are in general sublinear and hold only for decreasing stepsizes.

In this talk we analyze the convergence rates of minibatch stochastic first order methods with constant or variable stepsize for composite convex optimization problems, expressed in terms of expectation operator. We show that these methods can achieve linear convergence rate up to a constant proportional the stepsize and under some strong stochastic bounded gradient condition even pure linear convergence. When the strong stochastic bounded gradient condition does not hold we show that restarted variants of these methods can achieve linear convergence. Moreover, when variable stepsize is chosen we derive sublinear convergence rates that show an explicit dependence on the minibatch size. Applications of these results to convex feasibility problems will be also discussed.

Dominikus Noll (Paul Sabatier University of Toulouse)

Optimization strategies to control infinite-dimensional systems

Performance and robustness specifications in controlled systems are generally assured by the use nonlinear non-convex and often non-smooth optimization methods. Strategies suited even for challenging high technology control design problems have been identified during the past decade and work well for finite-dimensional (real-rational) systems. In this presentation we investigate whether, or to what extent, similar approaches can be put to work for the control of infinite-dimensional systems. Applications include boundary and distributed control of parabolic and hyperbolic PDEs, and control of delay or fractional order systems.

Panos Patrinos (Catholic University of Leuven)

A universal majorization-minimization framework for the convergence analysis of nonconvex proximal algorithms

Although originally designed and analyzed for convex problems, many splitting algorithms have been observed to perform well when applied to certain classes of structured nonconvex optimization problems. Without convexity, however, the elegant link with monotone operator theory onto which the convergence of many splitting algorithms is based no longer holds, and attempts to extend the analysis beyond convexity have only led to case-specific and unrelated results. The universal interpretation of convex splitting algorithms as relaxed fixed-point iterations of averaged operators serves a role much more important than mere aesthetics, as it furnishes a key insight for the understanding of how this type of methods functions and can possibly be improved.

In this talk we provide a novel unified interpretation of (possibly nonconvex) splitting algorithms as compositions of Lipschitzian and outer semicontinuous set-valued mappings. Specifically, the (possibly set-valued) fixed-point iteration black box A yielding $s \mapsto s^+ \in A(s)$ for the minimization of a lower semicontinuous (lsc) externded-real-valued function φ is identified with a pair $A \sim (G, M)$, in such a way that s^+ can be retrieved as the composition of the Lipschitz and strongly monotone (single-valued) mapping G and the (set-valued) parametric minimization of the *majorizing model* M, after a suitable relaxation step. Possibly under additional assumptions to compensate the lack of convexity, this interpretation is general enough to cover popular splitting algorithms such as the ADMM, the forward-backward splitting, the Douglas-Rachford splitting, and the Davis-Yin three-term splitting, together with their (over-)relaxed variants.

The convergence analysis is based on a suitable merit function, named proximal envelope, that generalizes the Moreau envelope and its connections with the proximal point algorithm. By combining its lower boundedness, being its infimum the same as that of the original cost function, with a sufficient decrease property, subsequential convergence to stationary points of φ as well as the vanishing of the fixed-point residual can be easily inferred.

Furthermore, this framework naturally leads to the integration with fast local methods applied to the nonlinear inclusion $s \in A(s)$ encoding optimality conditions of the minimization problem. Global convergence can be guaranteed thanks to a novel line-search strategy which is solely based on continuity properties of the proximal envelope and on the sufficient decrease condition. The particular choice of quasi-Newton schemes can yield up to superlinear rates of convergence while maintaining the oracle complexity of the original splitting algorithm.

Jean-Christophe Pesquet (University Paris-Saclay)

Deep unfolded proximal interior point algorithm

Variational methods are widely applied to ill-posed inverse problems as they have the ability to embed prior knowledge about the solution. However, these methods depend on a set of parameters, which need to be estimated through computationally expensive methods. In contrast, deep learning offers very generic architectures, at the expense of explainability and without any fine control over its output.

Deep unfolding provides a convenient approach to combine variational-based and deep learning approaches.

Starting from a variational formulation for image restoration, we develop iRestNet, a neural network architecture obtained by unfolding a proximal interior point algorithm. Hard constraints, encoding desirable properties for the restored image, are incorporated into the network thanks to a logarithmic barrier, while the barrier parameter, the stepsize, and the penalization weight are learned by the network. We derive explicit expressions for the gradient of the proximity operator for various choices of constraints, which allows training iRestNet with gradient backpropagation. In addition, we provide theoretical results regarding the stability of the network for a common inverse problem example. Numerical experiments on image deblurring problems show that the good performance of the proposed approach.

This is a joint work with M.-C. Corbineau and E. Chouzenoux (University Paris-Saclay), and with C. Bertocchi and M. Prato (University of Modena).

Juan Peypouquet (University of Chile)

Lagrangian penalization scheme with parallel forward-backward splitting

We propose a new iterative algorithm for the numerical approximation of the solutions to convex optimization problems and constrained variational inequalitites, especially when the functions and operators involved have a separable structure on a product space, and exhibit some dissymmetry in terms of their component-wise regularity. Our method combines Lagrangian techniques and a penalization scheme with bounded parameters, with parallel forward-backward iterations. Conveniently combined, these techniques allow us to take advantage of the particular structure of the problem. We prove the weak convergence of the sequence generated by this scheme, along with worst-case convergence rates in the convex optimization setting, and for the strongly nondegenerate monotone operator case. Implementation issues, and areas of application are discussed.

Elena Resmeriţă (University of Klagenfurt)

Sparsity regularization: a general non-convex approach

The classical framework for sparsity promoting regularization of ill-posed problems employs ℓ^p penalties with $p \in (0,2)$. In this talk, we recall a more flexible way of (performing/achieving) sparse regularization by varying exponents, which seems effective in several applications. Furthermore, we propose a more general penalty setting for which we provide a few theoretical results. Some algorithmic approaches for the involved optimization problems are also outlined.

Ernest K. Ryu (University of Edinburgh and KAUST)

Uniqueness of DRS as the 2 Operator Resolvent-Splitting and Impossibility of 3 Operator Resolvent-Splitting

TBA

Peter Richtarik (University of Edinburgh and KAUST)

SEGA: Variance reduction via gradient sketching

We propose a randomized first order optimization method–SEGA (SkEtched GrAdient method)– which progressively throughout its iterations builds a variance-reduced estimate of the gradient from random linear measurements (sketches) of the gradient obtained from an oracle. In each iteration, SEGA updates the current estimate of the gradient through a sketch-and-project operation using the information provided by the latest sketch, and this is subsequently used to compute an unbiased estimate of the true gradient through a random relaxation procedure. This unbiased estimate is then used to perform a gradient step. Unlike standard subspace descent methods, such as coordinate descent, SEGA can be used for optimization problems with a non-separable proximal term. We provide a general convergence analysis and prove linear convergence for strongly convex objectives. In the special case of coordinate sketches, SEGA can be enhanced with various techniques such as importance sampling, minibatching and acceleration, and its rate is up to a small constant factor identical to the best-known rate of coordinate descent.

Shoham Sabach (Technion - Israel Institute of Technology)

Lagrangian-based methods for nonconvex optimization problems

Lagrangian-based methods have a long history, and recently, there has been an intensive renewed interests in these methods and, in particular, within the Alternating Direction of Multipliers (ADM) scheme. The recent literature on Lagrangian-based methods in general and ADM, in particular, is voluminous in the setting of convex problems. However, the situation in the nonconvex setting is far from being well-understood, and analysis of Lagrangian-based methods in the nonconvex setting remain scarce and challenging. In this talk, some recent work on the analysis and applications of Lagrangian-based methods in the nonconvex setting will be presented.

Otmar Scherzer (University of Vienna)

Convergence rates of first and higher order dynamics for solving linear ill-posed problems

Recently, there has been a great interest in analysing dynamical flows, where the stationary limit is the minimiser of a convex energy. Particular flows of great interest have been continuous limits of Nesterov's algorithm and the Fast Iterative Shrinkage-Thresholding Algorithm (FISTA), respectively. Here we approach the solutions of linear ill-posed problems by dynamical flows. Because the squared norm of the residuum of a linear operator equation is a convex functional, the theoretical results from convex analysis for energy minimising flows are applicable. The proposed flows for minimising the residuum of a linear operator equation are optimal regularisation methods and that they provide optimal convergence rates for the regularised solutions.

This is joint work with R.I. Boţ, G. Dong and P. Elbau.

Mathias Staudigl (Maastricht University)

On the convergence of stochastic forward-backward-forward algorithms with variance reduction in pseudo-monotone variational inequality problems

We develop a new stochastic algorithm with variance reduction for solving pseudo-monotone stochastic variational inequalities. Our method builds on Tseng's forward-backward-forward algorithm, which is known in the deterministic literature to be a valuable alternative to Korpelevich's extragradient method when solving variational inequalities over a convex and closed set governed with pseudo-monotone and Lipschitz continuous operators. The main computational advantage of Tseng's algorithm is that it relies only on a single projection step, and two independent queries of a stochastic oracle. Our algorithm incorporates a variance reduction mechanism, and leads to a.s. convergence to solutions of a merely pseudo-monotone stochastic variational inequality problem. To the best of our knowledge, this is the first stochastic algorithm achieving this by using only a single projection at each iteration. Specifically, we focus on the following algorithm for solving stochastic variational inequality (SVI) problems: Find $x^* \in X$ such that $(T(x^*), x - x^*) \geq 0$ for all $x \in X$, in which $T(x) = Ex_\xi[F(x, \xi)]$ is a pseudo-monotone operator given as expected value of a map $F(x, z)$, and $X \subset R^n$ is a closed convex set. In all but the simplest situations, the evaluation of T is impossible since either the computation of the expected value is intractable due to high-dimensionality of the problem, or because the law of the random element ξ is not known (as is the rule in statistics and machine learning). We therefore have to rely on stochastic approximation (SA) schemes for solving the SVI. In this paper we consider a new SA scheme inspired by Tseng's modified forward-backward-forward scheme, given by

$$Y_n = \Pi_X \left[X_n - \frac{\alpha_n}{m_{n+1}} \sum_{k=1}^{m_{n+1}} F(X_n, \xi_{n+1}^{(k))}) \right],$$

$$X_{n+1} = Y_n + \frac{\alpha_n}{m_{n+1}} \sum_{k=1}^{m_{n+1}} (F(X_n, \xi_{n+1}^{(k)}) - F(Y_n, \eta_{n+1}^{(k)})),$$

where Π_X is the Euclidean projector onto the feasible set X, and ξ_n, η_n are iid copies of ξ. (m_n) is the size of the sample of the data drawn at each iteration anew, and the estimator of the mean operator $T(X_n)$ is thus a Monte-Carlo average of the sample observations $F(X_n, \xi_{n+1}^{(k)})$. This structure of the *stochastic oracle* yields online variance reduction. We establish the following facts: We prove that the process (X_n) is bounded in $L^2(P^r)$, and the distance to the set of solutions of the SVI diminishes almost surely. Furthermore, we establish optimal convergence rates for methods where the step-size α_n is bounded away from zero. Indeed, we prove an $O(1/n)$ convergence rate in term of the mean squared residual function, which shows that the method is competitive with the extragradient method. However, since our algorithm only requires a single projection step at each iteration, improved per-iteration complexity is to be expected. Numerical experiments on stochastic fractional programming and random bi-matrix games reveal this computational advantage.

Thomas Surowiec (Philipps University of Marburg)

A primal-dual algorithm for PDE-constrained optimization under uncertainty

A wide array of applications in the natural sciences and engineering require the solution of PDE-constrained optimization problems with uncertain inputs. It is therefore crucial to determine optimal solutions that are robust to uncertainty. One possibility to obtain such robust or risk-averse solutions uses risk measures. However, many commonly used risk measures, e.g., the coherent risk measures, are nonsmooth. As expected, the nonsmoothness complicates both the numerical approximation and solution of these optimization problems. In order to address these challenges, we propose a primal-dual algorithm for solving nonsmooth risk-averse optimization problems in Banach spaces. The algorithm is motivated by the well-known method of multipliers and exploits a number of results on the epigraphical regularization of risk measures. We prove convergence of the algorithm and conclude with numerical examples demonstrating the efficiency of our method.

Matthew Tam (University of Göttingen)

Forward-backward splitting without cocoercivity

In this talk, I will discuss a simple modification of the forward-backward splitting method for finding a zero in the sum of two monotone operators. The modified method converges under the same assumptions as Tseng's forward-backward-forward method, namely, it does not require cocoercivity of the single-valued operator. Moreover, each of its iterations only require one forward evaluation rather than two as is the case in Tseng's method. Variants of the method incorporating a linesearch, an inertial term, or a structured three operator inclusion will also be discussed.

Based on joint work with Yura Malitsky (University of Göttingen).

Marc Teboulle (Tel-Aviv University)

Analysis of proximal methods for composite minimization

Proximal based methods are nowadays starring in optimization algorithms, and are effectively used in a wide spectrum of applications. This talk will present some recent work on the impact of the proximal framework for composite minimization, with a particular focus on convergence analysis and applications.

Michael Ulbrich (Technical University of Munich)

An inexact bundle method for nonconvex nondifferentiable minimization in Hilbert space

Motivated by optimal control problems for elliptic variational inequalities we develop an inexact bundle method for nonsmooth nonconvex optimization subject to general convex constraints. The proposed method requires only approximate (i.e., inexact) evaluations of the cost function and of an element of Clarke's subdifferential. The algorithm allows for incorporating curvature information while aggregation techniques ensure that an approximate solution of the piecewise quadratic subproblem can be obtained efficiently. A global convergence theory in a suitable infinite-dimensional Hilbert space setting is presented. For adaptive inexactness control, error estimates for the cost function and the solution of the subproblem are developed. We discuss the application of our framework to optimal control problems for the obstacle problem and present numerical results.

This is joint work with Lukas Hertlein.

The project is funded by the DFG within the SPP 1962.

Stefan Ulbrich (Technical University of Darmstadt)

Computing a subgradient for the solution operator of the obstacle problem and numerical realization

The obstacle problem is an important prototype of an elliptic variational inequality and it appears in the mathematical formulation of applications from physics, finance and other fields. When dealing with constraints of obstacle type in optimization problems, the main difficulty is the nondifferentiability of the corresponding solution operator. In this talk we determine and characterize a specific element of the Bouligand subdifferential of the solution operator of the obstacle problem. We construct an abstract sequence of differentiability points whose derivatives converge to a subgradient. In order to show this convergence, a precise analysis of the relevant set-valued mappings connected to the Gâteaux derivatives is necessary. The limit and thus the subgradient itself is determined by the solution of a variational equation, which is independent of the abstract approximating sequence. We suggest how the resulting PDE can be tackled from a numerical point of view and we investigate problems that have to be taken care of when applying discretization or approximation schemes in order to obtain inexact subgradients e.g. for bundle-type methods.

This is joint work with Anne-Therese Rauls.

Tuomo Valkonen (National Polytechnic University Quito)

First-order methods and model splitting techniques for non-convex non-smooth optimisation

Convex optimisation problems can frequently be solved more efficiently by converting their original, primal, form into a dual form, or, better yet, a saddle-point form; first-order algorithms for the latter including the Chambolle-Pock method or primal-dual proximal splitting (PDPS), as well as the classical ADMM and its preconditioned variants. Until recently, non-convex problems were most commonly solved by second-order methods in their primal form. In this talk, we present ways to reformulate non-convex non-smooth problems as general saddle point problems that split into a convex non-smooth part, and a non-convex smooth part. We then study extensions of the PDPS to such problems, illustrating the performance on practical inverse problems.

Silvia Villa (University of Genova)

Thresholding gradient algorithms in Hilbert spaces: support identification and linear convergence

We study ell_1 regularized least squares optimization problem in a separable Hilbert space. We show that the iterative soft-thresholding algorithm (ISTA) converges linearly, without making any assumption on the linear operator into play or on the problem. The result is obtained combining two key concepts: the notion of extended support, a finite set containing the support, and the notion of conditioning over finite dimensional sets. We prove that ISTA identifies the solution extended support after a finite number of iterations, and we derive linear convergence from the conditioning property, which is always satisfied for ell_1 regularized least squares problems.

Isao Yamada (Tokyo Institute of Technology)

An approximate simultaneous matrix-diagonalization via alternating projection

The approximate simultaneous diagonalization (ASD) of a given set of matrices has been a key for many computational strategies in modern signal processing and multi-way data analytics. In this talk, we present a novel formulation of the ASD as a nonconvex feasibility problem to find a simultaneously diagonalizable matrix set, in the neighborhood of the given matrix set which is not necessarily diagonalizable simultaneously, followed by an algebraic simultaneous diagonalization scheme.

Unlike the existing algorithms for the ASD, e.g., Jacobi-like methods, the proposed algorithm can enjoy effectively a central property, i.e., the pairwise commutativity, of simultaneously diagonalizable matrices.

This talk is based on a recent joint work with Riku Akema and Masao Yamagishi.

Wotao Yin (University of California, Los Angeles)

Scaled relative graph: a rigorous $2D$ geometric tool for contractive and nonexpansive operators

Many iterative algorithms can be thought of as fixed-point iterations of contractive or nonexpansive operators. Traditionally, such algorithms and operators are analyzed analytically, with inequalities. Since Eckstein and Bertsekas (Figure 1, Math Program 55:293-318, 1992), circles and half-spaces have been used to geometrically illustrate operator theoretic notions although the actual analyses, proofs, and the computation of optimal stepsizes were done analytically with inequalities.

In this talk, we formalize a correspondence between operators and geometric objects on the $2D$ plane and use elementary Euclidean geometry to rapidly prove many useful results of operator theory. The formalism maps various subclasses of operators to sets on the $2D$ plane and maps algebraic operations such as scaling, inversion, addition, and composition of operators to geometric operations on sets on the $2D$ plane. Equipped with these tools, we use geometric arguments to review classic results and obtain novel results on operator theory.

Joint work with Ernest Ryu and Robert Hannah.

Xiaoming Yuan (The University of Hong Kong)

An inexact Uzawa algorithmic framework for nonlinear saddle point problems

Saddle point problems are fundamental in many areas; their particular applications arise in various mathematical fields such as optimization, scientific computing, data science, control and game theory. While a rich set of literature exists for linear saddle point problems, the research for nonlinear saddle point problems is still in its infancy. In this talk, we will discuss an algorithmic framework derived from an inexact Uzawa method for a class of application-driven nonlinear saddle point problems. We uniformly establish its convergence and linear convergence rate, and show how it can be specified as some highly implementable splitting algorithms for special convex optimization problems. Some numerical results for solving elliptic optimal control problems will be shown.

This is a joint work with Yongcun Song and Hangrui Yue.