Strategies for Handling Applications with Nonconvexity (SHAWN)

Heinz Bauschke (UBC Okanagan), Radu Ioan Boţ (University of Vienna),
Regina Burachik (University of South Australia), and Walaa Moursi (University of Waterloo)

May 11–May 16, 2025

1 Overview and Goal

We live in exciting times, facing unparalleled opportunities and challenges. From the triumphs of machine learning and artificial intelligence to the challenges of sustainability and climate change, the underlying mathematical problems are optimization problems. Solving these optimization problems will allow us to make progress and mitigate serious problems. Mathematics plays a key role in the solution of these problems through the analysis and design of optimization algorithms, which are computer-implementable methods that allow us to find optimal solutions. Understanding how these algorithms perform in complicated real-world scenarios is critical to ensure that the answers the computer returns are accurate and obtained quickly.

The goal of the "Strategies for Handling Applications With Nonconvexity (SHAWN)" workshop was to understand these algorithms in all possible situations and to make them more efficient. This was achieved by surveying the latest research results, exchanging novel ideas and promising techniques, and discussing and identifying remaining challenges. Leading researchers, along with the next generation of researchers, explored new areas of applications, ranging from optics, crystallography, and satisfiability problems to linear programming problems. The expected outcome of this workshop is a new generation of research that will improve processes in industry and help train the next generation of applied scientists, equipping them with tools to address future challenges humankind will face.

2 Workshop Highlights

2.1 EDI

An Equity, Diversity, and Inclusion (EDI) networking event was held on Wednesday afternoon and was well attended. The discussion was lively, with many participants sharing their experiences and ideas on how to improve EDI in the mathematical sciences. Attendees shared insights from their own academic and professional journeys, highlighted barriers that underrepresented groups often face, and proposed practical strategies to foster a more inclusive and supportive environment.

2.2 Open Problem Session

The open problem session on Thursday provided a forum for attendees to present open problems in their research areas. Problems discussed ranged from subtle issues in Linear Algebra to finding the nearest zero of a sum of maximally monotone operators using splitting algorithms.

2.3 BIRS Lounge

The BIRS lounge was a great place for informal discussions in the evenings. Participants continued discussions from the day and connected with fellow researchers. In addition, participants explored topics such as the future of mathematics in view of the rise of artificial intelligence.

2.4 Talks

All talks were engaging and informative, making it difficult to highlight any single one. Below, we summarize them in their scheduled order.

Veit Elser: Life after the Book

The talk by Dr. Veit Elser provided a fascinating overview of his forthcoming book, *Solving Problems with Projections: From Phase Retrieval to Packing* [15], which explores the amazing power of projection algorithms to solve a wide range of problems in mathematics and physics.

Courtney Paquette: High-dimensional Optimization in Machine Learning with Applications to Scaling Limits and Compute-Optimal Neural Scaling Laws

Dr. Courtney Paquette introduced a framework using a power-law random features model, grounded in high-dimensional probability and random matrix theory, to analyze scaling laws in stochastic learning algorithms. This approach enables insights from smaller models to guide larger-scale ML training, addressing the compute-optimal selection of model size and hyper-parameters for efficient performance. Dr. Paquette also highlighted a scaling limit inspired by statistical physics and identified underexplored research directions in scaling laws with significant potential.

Ahmet Alacaoglu: Handling Generalized Monotone Operators with First-Order Algorithms

Dr. Ahmet Alacaoglu discussed first-order algorithms for finding zeros of monotone or generalized monotone operators, referencing the framework by Bauschke, Moursi, and Wang [5]. He highlighted complexity analysis of inexact fixed-point algorithms, offering insights into enhancing tolerance for nonmonotonicity and improving oracle complexity for stochastic problems. He also addressed implications for min-max problems. This work, conducted with Donghwan Kim and Stephen Wright (see [1]), advances the efficiency of first-order methods in optimization.

Ziyuan Wang: Bregman level proximal subdifferential and new insights into Bregman proximal operators

Dr. Ziyuan Wang introduced the left and right Bregman level proximal subdifferentials to address limitations of classical subdifferentials for nonconvex Bregman proximal operators. Dr. Wang demonstrated that these operators act as resolvents of the new subdifferentials under standard assumptions, establishing novel correspondences between the operators, underlying functions, and subdifferentials. He also explored asymmetry, duality gaps, existence, single-valuedness, and connections to relative smoothness, extending classical Euclidean results. See also the related preprint [26].

Stephen Vavasis: Complexity of optimal parameters for fitting samples: Maximum likelihood versus Wasserstein

Dr. Stephen Vavasis presented on the complexity of parameter estimation for high-dimensional distributions, focusing on a distribution supported on a finite union of hyperrectangles with mean and shrinkage parameters. He highlighted that the maximum likelihood estimator leads to a nonconvex, NP-hard optimization problem. In contrast, modeling the problem with Wasserstein distance minimization yields a polynomial-time approximation algorithm. This work, conducted with Valentio Iverson, demonstrates the computational advantages of Wasserstein-based approaches over traditional maximum likelihood methods. See also the related preprint [17].

Shambhavi Singh: Eckstein-Ferris-Pennanen-Robinson duality revisited: paramonotonicity, total Fenchel-Rockafellar duality, and the Chambolle-Pock operator

Dr. Shambhavi Singh presented a revisit of the Eckstein-Ferris-Pennanen-Robinson duality framework for finding zeros of the sum of two maximally monotone operators with a continuous linear operator. She emphasized paramonotonicity as a key condition ensuring saddle points align with the closed convex rectangle of primal and dual solutions. Dr. Singh also characterized total duality in the subdifferential context and derived projection formulas for sets relevant to the Chambolle-Pock algorithm, leveraging the recent framework by Bredies, Chenchene, Lorenz, and Naldi. See also the related preprint [4].

Sedi Bartz: Convex analysis in multi-marginal settings

Dr. Sedi Bartz explored the convex analytic foundations of multi-marginal optimal transport, presenting extensions of classical convex analysis and monotone operator theory. He discussed recent advancements, highlighted open questions, and shared insights from ongoing research, emphasizing the structural role of convexity in this field.

Jelena Diakonikolas: Gradient alignment, learning, and optimization

Dr. Jelena Diakonikolas focused on the challenges of learning Generalized Linear Models (GLMs) to recover an unknown vector from noisy labeled examples. She highlighted the nonconvex, often nonsmooth optimization problems arising in GLMs, particularly under realistic settings with arbitrary label deviations. Dr. Diakonikolas surveyed recent progress, introducing a unifying optimization-theoretic framework based on local error bounds. This framework leverages gradient field alignment with target solutions to enable efficient learning using first-order methods, despite nonconvexity and noise, emphasizing the role of structural assumptions on activations and data distributions.

Hristo Sendov: On Polar Convexity in Finite-Dimensional Euclidean Spaces

Dr. Hristo Sendov introduced a novel concept of polar convexity that extends classical convexity by incorporating a pole. This framework generalizes convexity from the complex plane to Euclidean spaces, where a set is polar convex if it satisfies a circular arc condition relative to a pole. As the pole approaches infinity, polar convexity reduces to traditional convexity. Sendov highlighted the richer structure of polar convexity, including a duality property absent in classical convexity, and presented polar analogues of classical results like Gordan's and Farkas' lemmas. He also characterized the convex hull for finitely many points with respect to multiple poles, showcasing the potential of this framework to advance convex analysis. See also the related paper [6].

David Torregrosa Belén: Enhanced randomized block proximal gradient methods beyond global Lipschitz continuity

Dr. David Torregrosa-Belén presented an advanced randomized block proximal gradient algorithm that extends beyond the standard global Lipschitz continuity assumption, focusing on locally Lipschitz continuous partial gradients. This innovation is crucial for tackling large-scale optimization challenges where full derivative computations are resource-intensive. The algorithm dynamically adjusts proximal stepsizes and incorporates an optional linesearch to guarantee convergence to stationary points in nonconvex scenarios, as evidenced by its successful application in image compression via nonnegative matrix factorization. See also the related preprint [24].

Yuan Gao: c-potential for infinite costs

Mr. Yuan Gao explored the existence of c-potential under c-path boundedness. He covered a general existence theorem without topological assumptions, results for ”locally finite” costs, and computation methods for a specific cost class, advancing optimal transport research.

Viktor Pavlovic: Accelerated gradient descent: A guaranteed bound for a heuristic restart strategy

Mr. Viktor Pavlovic explained that while accelerated gradient descent has an optimal $O(1/k^2)$ convergence rate, practical tweaks like restarts can make it faster. Focusing on the adaptive gradient restart method by O'Donoghue and Candès, Pavlovic showed it not only maintains but improves the $O(1/k^2)$ bound for one-dimensional functions, ensuring convergence. He also applied these findings to separable and nearly separable functions, highlighting broader uses. See also the related preprint [22].

Yura Malitsky: Entropic Mirror Descent for Linear Systems: Polyak's Stepsize and Implicit Bias

Dr. Yura Malitsky explored the application of entropic mirror descent to solve linear systems, discussing its implicit bias and analyzing its convergence properties. Despite its simple structure, the method's convergence behavior is subtle, but Malitsky demonstrated how a Polyak-type stepsize can overcome these challenges, leading to an explicit convergence rate. He also briefly covered generalizations of the entropic mirror descent method. See also the related preprint [20].

Renata Sotirov: Lagrangian duality for Mixed-Integer Semidefinite Programming

Dr. Renata Sotirov highlighted how mixed-integer semidefinite programming (MISDP) extends mixed-integer programming by using positive semidefinite matrix variables to model nonlinear optimization problems. She introduced Lagrangian duality-based bounds for MISDPs and discussed new projected bundle and subgradient algorithms, demonstrating that these yield stronger bounds than standard semidefinite programming relaxations for various optimization problems. See also the related preprint [13].

Aris Daniilidis: Pathological differentiable locally Lipschitz functions

Dr. Aris Daniilidis examined the distinction between differentiable and strictly differentiable locally Lipschitz functions in nonsmooth analysis. He demonstrated that, unlike strictly differentiable functions with singleton limiting Jacobians, certain differentiable locally Lipschitz functions have limiting Jacobians encompassing all nonempty compact connected subsets of matrices. For real-valued functions, he showed surjective limiting and Clarke subdifferentials, providing a concrete example to illustrate that such pathological functions are dense and spaceable. This work was conducted with R. Deville and S. Tapia-Garcia [11].

Robert Csetnek: Tikhonov regularization for monotone operators

Dr. Robert Csetnek explored the asymptotic behavior of a second-order dynamical system for solving monotone equations in Hilbert spaces. He introduced vanishing damping linked to the Tikhonov parameter and a correction term using the operator's time derivative, drawing connections to Newton and Levenberg–Marquardt methods. He demonstrated strong convergence to the minimal norm solution with fast rates for velocity and operator-related quantities, and applied these findings to a primal-dual system for linearly constrained convex optimization, achieving robust trajectory convergence and rapid feasibility and function-value rates. See also the related preprint [10].

Adriana Nicolae: Approaches to subgradient algorithms in nonpositively-curved spaces

Dr. Adriana Nicolae explored convex optimization in geodesic metric spaces of nonpositive curvature, extending beyond linear spaces and Riemannian manifolds. She discussed challenges in defining subgradients in such spaces, proposing subgradient-style methods, including a splitting variant, using horoballs and Busemann functions. Dr. Nicolae derived complexity bounds comparable to standard results, demonstrating the efficacy of these methods. This work, conducted with A. Goodwin, A.S. Lewis, and G. Lopez-Acedo, advances optimization in nontraditional geometric settings.

Francisco J. Aragón Artacho: Forward-backward algorithms devised by graphs

Dr. Francisco J. Aragón Artacho presented on "Forward-Backward Algorithms Devised by Graphs," introducing a methodology for designing forward-backward methods to find zeros of sums of maximally monotone operators. Extending prior work of Bredies, Chenchene, and Naldi [8], he utilized three graphs to manage interactions and resolvent computations for a finite number of cocoercive operators, ensuring minimal lifting and frugal evaluations (one per operator per iteration). Dr. Aragón Artacho demonstrated how this framework recovers existing methods and generates novel ones, including a complete graph-induced algorithm, with numerical experiments highlighting the graphs' impact. This work was conducted with R. Campoy and C. López-Pastor [3].

Madeleine Udell: Online Scaled Gradient Methods

Dr. Madeleine Udell introduced a framework that uses online learning to dynamically scale gradients, accelerating gradient-based methods. Unlike traditional worst-case analyses, this approach offers strong convergence guarantees relative to the optimal stepsize trajectory. Dr. Udell provided the first convergence guarantee for Hypergradient Descent (HDM) and introduced variants with heavy-ball and Nesterov momentum (HDM-HB), which often rival L-BFGS performance with lower memory and computational costs. See also the related preprint [16].

Xianfu Wang: Level proximal subdifferential, variational convexity, and beyond

Dr. Xianfu Wang focused on the level proximal subdifferential introduced by Rockafellar for analyzing proximal mappings of potentially nonconvex functions. He characterized variational convexity through locally firm nonexpansiveness of proximal mappings and locally relative monotonicity of the level proximal subdifferential, applying these to study the local convergence of proximal gradient and related methods. Dr. Wang highlighted that variational sufficiency ensures convergence to local minimizers, not just to critical points, advancing optimization techniques. This work was conducted with H. Luo, Z. Wang, and X. Yang. See also the related preprint [19].

César David López-Pastor: A unifying graph-based analysis of projection algorithms for linear subspaces

Mr. César David López-Pastor presented on "A Unifying Graph-Based Analysis of Projection Algorithms for Linear Subspaces," focusing on a general analysis of fixed points for operators in graph-splitting methods developed by Bredies, Chenchene, and Naldi [8]. By specializing to projection algorithms for closed linear subspaces, he derived explicit formulas for the limit points of these schemes, unifying existing results and introducing new ones, enhancing the understanding of projection-based optimization techniques. See also the related preprint [2].

Enis Chenchene: A Fast Extra-Gradient Method with Flexible Anchoring

Dr. Enis Chenchene introduced a novel Extra-Gradient method with adaptable anchoring parameters derived from discretizing a dynamical system with Tikhonov regularization. He established strong convergence to specific solution points and derived convergence rates based on regularization parameters, achieving the standard $O(k^{-1})$ residual decay rate for typical settings. Numerical experiments highlighted the method's flexibility and competitive performance in optimization tasks. See also the related preprint [7].

Manish Krishan Lal: Learning, Sampling, and Inference Through the Lens of Projections

Dr. Manish Krishan Lal offered a unified perspective on learning, inference, and sampling algorithms using projection operators. He highlighted their theoretical elegance for structured problems, while noting they are lagging behind gradient methods in data science due to challenges with nonconvexity and noise. Dr. Lal proposed a refined framework linking projections to implicit regularization, dynamical systems, and sampling strategies, supported by empirical benchmark results and interpretability insights.

Woosuk Jung: Single Element Error Correction in a Euclidean Distance Matrix

Mr. Woosuk Jung addressed the exact correction of a noisy Euclidean distance matrix (EDM) with a single corrupted entry, given a known embedding dimension d. He introduced three divide-and-conquer strategies combined with facial-reduction techniques (exposing vectors, facial vectors, and Gale transforms), enabling solutions for up to 10^5 nodes in approximately one minute to machine precision. He also characterized conditions under which the perturbed matrix remains an EDM and when the nearest EDM problem approach succeeds.

Vera Roshchina: How to calculate generalised subdifferentials

Dr. Vera Roshchina focused on the role of generalised subdifferentials in providing local optimality conditions for nonsmooth nonconvex functions, akin to gradients and convex subdifferentials. She addressed their limitation in lacking convenient calculus rules and introduced a collection of practical techniques to simplify the computation of common generalised subdifferentials for structured functions, enhancing their applicability in nonsmooth optimization. For some background material, see [14].

Scott Lindstrom: Gabay Duality for Nonconvex Optimisation

Dr. Scott Lindstrom extended Daniel Gabay's finding that ADMM on convex primal problems applies Douglas–Rachford splitting to their duals. He generalized this duality to nonconvex problems, illustrating cases where ADMM iterates exhibit cycling or chaotic behavior. Dr. Lindstrom proposed leveraging Gabay's framework to design more robust nonconvex ADMM algorithms, offering insights to enhance performance in complex optimization scenarios.

Bethany Caldwell: Optimal control duality and the Douglas–Rachford algorithm

Dr. Bethany Caldwell explored the application of the Douglas–Rachford (DR) algorithm in infinite-dimensional settings for solving the dual of a weighted minimum-energy linear-quadratic control problem with constraints. She derived the DR fixed point, introduced an optimality check for numerical solutions, and proposed an algorithm generating both primal and dual sequences. Dr. Caldwell demonstrated the approach's effectiveness through two example control problems, advancing the understanding of DR in optimal control. See also the related paper [9].

Haihao Lu: GPU-Accelerated Linear Programming

Dr. Haihao Lu discussed first-order methods (FOMs) to enhance the scalability and speed of linear programming (LP) on GPUs. He highlighted the limitations of traditional LP solvers, which rely on memory-intensive matrix factorizations ill-suited for GPU parallelism, and contrasted these with FOMs that use efficient matrix-vector products. Dr. Lu outlined efforts to adapt FOMs for LP, including handling infeasible or unbounded cases, and presented benchmarks showcasing their convergence and GPU performance, advancing large-scale optimization. See also the related preprint [18].

Jane Ye: New second-order optimality conditions for set-constrained optimization problems

Dr. Jane Ye presented on "New Second-Order Optimality Conditions for Set-Constrained Optimization Problems," introducing novel second-order necessary and sufficient optimality conditions for set-constrained optimization problems with potentially nonconvex constraint sets. Utilizing classical and lower generalized support functions, she derived conditions through asymptotic second-order tangent cones and outer second-order tangent sets. Dr. Ye emphasized that these results do not require convexity, nonemptiness of tangent sets, or second-order regularity, significantly improving upon classical optimization theory. See also the related preprint [23].

Xiaoming Yuan: An Operator Learning Approach to Nonsmooth Optimal Control of Nonlinear PDEs

Dr. Xiaoming Yuan addressed the computational challenges of optimal control problems with nonsmooth objectives and nonlinear PDE constraints. He introduced a primal-dual operator-learning framework that constructs neural surrogates for PDEs, enabling efficient mesh-free forward passes and eliminating the need for repeated high-dimensional solves. Dr. Yuan highlighted the framework's ease of implementation, reusability across parameters, and significant reduction in computational cost, advancing nonsmooth optimal control. See also the related paper [25].

Matthew Tam: A first-order algorithm for decentralised min-max problems

Dr. Matthew Tam proposed a novel first-order method for solving convex-concave min-max problems in a decentralised network of agents. Combining elements of PG-EXTRA for decentralised minimization and the forward-reflected backward method for non-distributed min-max problems, the algorithm enables each agent to compute gradients and proximal steps locally while communicating with neighbors, offering an efficient approach to distributed optimization. See also the related preprint [21].

Hoa Bui: Finding Epsilon-Solutions of Constrained Convex Optimization with Single Projection

Dr. Hoa Bui demonstrated that an ε-solution for convex optimization problems with a linear objective in a Hilbert space can be achieved through a single projection of a carefully chosen infeasible point onto the feasible set. She quantified the necessary distance of the infeasible point to ensure its projection yields an exact solution, offering a streamlined approach to convex optimization.

Henry Wolkowicz: A Peaceman-Rachford Splitting Method for the Protein Side-Chain Positioning Problem and other Hard Problems

Dr. Henry Wolkowicz tackled the NP-hard protein side-chain positioning (SCP) problem through an integer quadratic program. He introduced a doubly nonnegative (DNN) relaxation refined by facial reduction to address strict infeasibility, followed by a Peaceman-Rachford splitting method (PRSM) that delivers strong approximate solutions, proven optimal for nearly all Protein Data Bank instances. Dr. Wolkowicz also explored the method's application to the Wasserstein barycenter problem, broadening its impact on complex optimization challenges.

Minh Dao: Projected proximal gradient trust-region algorithm for nonsmooth optimization

Dr. Minh Dao focused on trust-region methods for objectives combining a smooth nonconvex function with a nonsmooth convex regularizer. He extended global convergence theory to account for worst-case complexity under unbounded Hessian growth and introduced a novel subproblem solver that integrates proximal gradient iterations with a single projection step. Dr. Dao demonstrated that the method satisfies sufficient-descent conditions and exhibits promising numerical performance in nonsmooth optimization tasks. See also the related preprint [12].

3 Scientific Progress Made and Outcome of the Meeting

The SHAWN workshop was a great success, with participants attending from around the world (South America, North America, Europe, Asia, and Australia). One third of the participants were female scientists—a proportion that notably exceeds the typical representation of women in mathematics departments.

The participants were able to present their latest research results and discuss them with their peers. The workshop provided a platform for collaboration and networking, which is essential for the advancement of the field. The organizers are confident that the SHAWN workshop will lead to new collaborations and research projects that will advance the field of optimization algorithms.

Several researchers have indicated their interest in contributing to the proceedings volume, which will be published by Springer. Manuscripts will undergo a peer-review process, and the organizers will ensure that the volume is of high quality. The expected volume will contain a collection of papers providing a representative snapshot of the current state of the field, and illustrating the diversity of the contributors to the meeting.

References

[1] A. Alacaoglu, D. Kim, and S. Wright: Revisiting Inexact Fixed-Point Iterations for Min-Max Problems: Stochasticity and Structured Nonconvexity, `https://arxiv.org/abs/2402.05071`.

[2] F.J. Aragón Artacho, H.H. Bauschke, R. Campoy, and C. López-Pastor: Graph splitting methods: Fixed points and strong convergence for linear subspaces, `https://arxiv.org/abs/2505.16564`.

[3] F.J. Aragón Artacho, R. Campoy, C. López-Pastor: Forward-backward algorithms devised by graphs, `https://arxiv.org/abs/2406.03309`.

[4] H.H. Bauschke, W.M. Moursi, and S. Singh: Eckstein-Ferris-Pennanen-Robinson duality revisited: paramonotonicity, total Fenchel-Rockallar duality, and the Chambolle-Pock operator, `https://arxiv.org/abs/2412.11880`.

[5] H.H. Bauschke, W.M. Moursi, and X. Wang: Generalized monotone operators and their averaged resolvents, *Mathematical Programming (Series B)* **189** (2021), 55–74.

[6] S. Bhatt and H.S. Sendov: On polar convexity in finite-dimensional Euclidean spaces, *Canadian Journal of Mathematics*, `https://doi.org/10.4153/S0008414X24000671`.

[7] R.I. Boţ and E. Chenchene: Extra-gradient method with flexible anchoring: strong convergence and fast residual decay, `https://arxiv.org/abs/2410.14369`.

[8] K. Bredies, E. Chenchene, and E. Naldi: Graph and distributed extensions of the Douglas–Rachford method, *SIAM Journal on Optimization* **34** (2024), 1569–1594.

[9] R.S. Burachik, B.I. Caldwell, C.Y. Kaya, and W.M. Moursi: Optimal control duality and the Douglas–Rachford algorithm, *SIAM Journal on Control and Optimization* **62** (2024), 1–24.

[10] E.R. Csetnek and S.C. László: Strong convergence and fast rates for systems with Tikhonov regularization, `https://arxiv.org/abs/2411.17329`.

[11] A. Daniilidis, R. Deville, and S. Tapia-Garcia: All convex bodies are in the subdifferential of some everywhere differentiable locally Lipschitz function, `https://arxiv.org/abs/2405.09206`.

[12] M.N. Dao, H.M. Phan, and L. Roberts: Projected proximal gradient trust-region algorithm for nonsmooth optimization, `https://arxiv.org/abs/2501.04889`.

[13] F. de Meijer and R. Sotirov: Lagrangian Duality for Mixed-Integer Semidefinite Programming: Theory and Algorithms, `https://arxiv.org/abs/2501.11397`.

[14] V.F. Demyanov and V.A. Roshchina: Optimality conditions in terms of upper and lower exhausters, *Optimization* **55** (2006), 525–540.

[15] V. Elser: *Solving Problems with Projections: From phase retrieval to packing*, Cambridge University Press, to appear in 2025.

[16] W. Gao, Y.-C. Chu, Y. Ye, and M. Udell: Gradient Methods with Online Scaling Part I. Theoretical Foundations, `https://www.arxiv.org/abs/2505.23081`.

[17] V. Iverson and S. Vavasis: Mean and Variance Estimation Complexity in Arbitrary Distributions via Wasserstein Minimization, `https://arxiv.org/abs/2501.10172`.

[18] H. Lu and J. Yang: An Overview of GPU-based First-Order Methods for Linear Programming and Extensions, `https://arxiv.org/abs/2506.02174`.

[19] H. Luo, X. Wang, Z. Wang, and X. Yang: Level proximal subdifferential, variational convexity, and pointwise Lipschitz smoothness, `https://arxiv.org/abs/2406.00648`.

[20] Y. Malitsky and A. Posch: Entropic Mirror Descent for Linear Systems: Polyak's Stepsize and Implicit Bias, `https://arxiv.org/abs/2505.02614`.

[21] Y. Malitsky and M. Tam: Entropic Mirror Descent for LA First-order algorithm for decentralised min-max problems, `https://arxiv.org/abs/2308.11876`.

[22] W.M. Moursi, V. Pavlovic, and S. Vavasis: Accelerated gradient descent: A guaranteed bound for a heuristic restart strategy, `https://arxiv.org/abs/2310.07674`.

[23] W. Ouyang, J. Ye, and B. Zhang: New second-order optimality conditions for directional optimality of a general set-constrained optimization problem, `https://arxiv.org/abs/2404.17696`.

[24] P. Pérez-Aros and D. Torregrosa-Belén: Randomized block proximal method with locally Lipschitz continuous gradient, `https://arxiv.org/abs/2504.11410`.

[25] Y. Song, X. Yuan, and H. Yue: The ADMM-PINNs algorithmic framework for nonsmooth PDE-constrained optimization: a deep learning approach, *SIAM Journal on Scientific Computing* **46** (2024), C659—C687.

[26] Z. Wang and A. Themelis: Bregman level proximal subdifferentials and new characterizations of Bregman proximal operators, `https://arxiv.org/abs/2506.07333`.