Zaiwen Wen, Research Contribution

Zaiwen Wen's contributions have significantly impacted mathematical optimization and machine learning, enhancing both theoretical analysis and the design of algorithms for complex problems in science and engineering. His development of the semi-smooth Newton methods for nonsmooth optimization, adaptive regularized Newton methods for manifold optimization and eigenvalue optimization, and second-order optimization techniques for deep learning, have established new efficiency benchmarks and led to widely-used academic softwares including LMaFit, LMSVD, OptM, ARNT, Arrabit and MCPG, etc. His algorithms have solved industrial challenges from scheduling, wireless communication and railway timetabling. He has published over 80 papers, with 25 of them in the SIAM series. These advancements displayed a perfect fusion of mathematical theory, computational efficiency with practical applications across diverse disciplines.

Nonsmooth and nonlinear optimization

Zaiwen's first-order and second-order methods for nonsmooth composite optimization has notably advanced the field, particularly in addressing emergent problems arising from sparse optimization, low-rank matrix recovery, image and signal processing, and machine learning. He was one of the pioneers in applying the alternating direction method of multipliers to solve semidefinite programming problems, then enhanced its convergence with the concept of semi-smoothness. He observed that the solutions to the composite convex optimization problems and the solutions to the system of nonlinear equations corresponding to many first-order methods are equivalent. Although these systems are non-differentiable, they are semi-smooth. From the perspective of solving a single system of nonlinear equations corresponding to first-order methods, he and his collaborators systematically investigated the semi-smooth Newton methods. This framework exemplifies significant reductions in the computational cost, promising impressive efficacy across numerous applications. For large-scale stochastic problems, he designed a stochastic semi-smooth Newton method and proved global convergence and R-superlinear convergence locally with high probability. He developed a decomposition method within the augmented Lagrangian framework to address possibly nonlinear objective functions, nonsmooth regularization, and general linear equality/inequality constraints. These algorithms and software, accumulated over more than 10 years, have been integrated into OptSuite, a leading open-source software platform based on C++, capable of solving linear programming, second-order cone programming, semidefinite programming, nonlinear programming, integer programming, and etc.

1. Wang Yifei, Deng Kangkang, Liu Haoyang, **Wen Zaiwen,** A Decomposition Augmented Lagrangian Method for Low-rank Semidefinite Programming, **SIAM Journal on Optimization**, Vol. 33, No. 3, 1361-1390, 2023, **Google citation: 9**
2. Li Yongfeng, Liu Haoyang, **Wen Zaiwen**, and Yuan Yaxiang, Low-rank Matrix Optimization Using Polynomial-filtered Subspace Extraction, **SIAM Journal on Scientific Computing**, Vol 42, No. 3, A1686–A1713, 2020, **Google citation: 3**

3. Andre Milzarek, Xiao Xiantao, Cen Sicong, **Wen Zaiwen**, Michael Ulbrich; A stochastic semi-smooth Newton method for nonsmooth nonconvex optimization, **SIAM Journal on Optimization**, Vol 29, No. 4, pp. 2916–2948, 2019, **Google citation: 38**
4. Ma Chao, Liu Xin, **Wen Zaiwen**; Globally convergent Levenberg-Marquardt method for phase retrieval; **IEEE Transactions on Information Theory**; Vol 65, No 4, 2019, pp 2343 – 2359, **Google citation: 27**
5. Li Yongfeng, **Wen Zaiwen**, Yang Chao, Yuan Yaxiang; A Semi-smooth Newton Method For semidefinite programs and its applications in electronic structure calculations; **SIAM Journal on Scientific Computing**; Vol 40, No. 6, 2018, A4131–A4157, **Google citation: 15**
6. Zhang Junyu, Liu Haoyang, **Wen Zaiwen**, Zhang Shuzhong; A Sparse Completely Positive Relaxation of the Modularity Maximization for Community Detection; **SIAM Journal on Scientific Computing**; Vol 40, No. 5, 2018, A3091–A3120, **Google citation: 14**
7. Xiao Xiantao, Li Yongfeng, **Wen Zaiwen**, Zhang Liwei; A Regularized Semi-Smooth Newton Method with Projection Steps for Composite Convex Programs; **Journal of Scientific Computing**; 2018, Vol 76, No. 1, pp 364-389, **Google citation: 76**
8. Shen Yuan, **Wen Zaiwen**, Zhang Yin; Augmented Lagrangian alternating direction method for matrix separation based on low-rank factorization; **Optimization methods and Software**; Vol. 29, No. 2 (2014), pp 239–263, **Google citation: 344**
9. **Wen Zaiwen**, Yin Wotao, Zhang Yin; Solving A Low-Rank Factorization Model for Matrix Completion by A Nonlinear Successive Over-Relaxation Algorithm; **Mathematical Programming Computation**; 4 (2012), 333-361, **Google citation: 851**
10. Xu Yangyang, Yin Wotao, **Wen Zaiwen**, Zhang Yin; An Alternating Direction Algorithm for Matrix Completion with Nonnegative Factors; **Journal of Frontiers of Mathematics in China**; 7 (2012), 365-384, **Google citation: 387**
11. **Wen Zaiwen**, Yin Wotao, Donald Goldfarb, Zhang Yin; A fast algorithm for sparse reconstruction based on shrinkage, subspace optimization and continuation; **SIAM Journal on Scientific Computing**; 32 (2010) 1832-1857, **Google citation: 288**
12. **Wen Zaiwen**, Donald Goldfarb, Yin Wotao; Alternating Direction Augmented Lagrangian Methods for semidefinite programming; **Mathematical Programming Computation**; 2 (2010) 203-230, **Google citation: 454**

Manifold Optimization

Zaiwen's pioneering contributions have significantly advanced the field of manifold-constrained optimization, with notable achievements in both theory and application. His innovative usage of the Cayley transform to address orthogonality constraints has opened new research avenues, providing a computationally efficient method to preserve these non-convex constraints across various scientific and engineering problems. This approach, characterized by lower computational overhead compared to traditional methods based on projections and geodesics, has proven effective in identifying globally optimal solutions under certain conditions, demonstrating its utility in fields ranging from polynomial optimization to synchronization and community detection. Building on this foundation, Zaiwen has developed adaptive regularized Newton methods and structured quasi-Newton methods, which mark a substantial leap in solving manifold optimization problems. By approximating the objective function with a second-order Taylor expansion in Euclidean space while preserving Riemannian manifold constraints, these methods not only simplify the computational process, especially for costly Hessian calculations but also ensure global convergence and superlinear local convergence under mild conditions. The ARNT software package, emerging from this work, exemplifies the superior performance of the methods over existing first-order and second-order algorithms in numerous applications, including but not limited to correlation matrix estimation, electronic structure computation, and Bose-Einstein condensation. Furthermore, Zaiwen's approach to manifold optimization with additional nonnegativity constraints addresses the combinatorial challenges introduced by such constraints, offering practical and effective solutions.

1. Hu Jiang, Ao Ruicheng, Anthony Man-Cho So, Yang Minghang, **Wen Zaiwen,** Riemannian Natural Gradient Methods, **SIAM Journal on Scientific Computing**, Vol. 46, No. 1, A204-A231, 2024, **Google citation: 5**

2. Jiang Bo, Meng Xiang, **Wen Zaiwen**, Chen Xiaojun, An exact penalty approach for optimization with nonnegative orthogonality constraints, **Mathematical Programming**, Vol. 198, 855–897, 2023, **Google citation: 20**
3. Zhang Haixiang, Andre Milzarek, **Wen Zaiwen,** Yin Wotao; On the geometric analysis of a quartic-quadratic optimization problem under a spherical constraint, **Mathematical Programming**, Vol. 195, 421-473, 2022, **Google citation: 8**
4. Hu Jiang, Jiang Bo, Lin Lin, **Wen Zaiwen**, Yuan Yaxiang, Structured Quasi-Newton Methods for Optimization with Orthogonality Constraints, **SIAM Journal on Scientific Computing**, Vol. 41, No. 4, pp. A2239-A2269, 2019, **Google citation: 36**
5. Hu Jiang, **Wen Zaiwen**, Andre Milzarek, Yuan Yaxiang; Adaptive Quadratically Regularized Newton Method for Riemannian Optimization; **SIAM Journal on Matrix Analysis and Applications**; Vol 39, No. 3, 2018, pp 1181-1207, **Google citation: 64**
6. Lai Rongjie, **Wen Zaiwen**, Yin Wotao, Gu Xianfeng, Lui Lok Ming; Folding-Free Global Conformal Mapping for Genus-0 Surfaces by Harmonic Energy Minimization**; Journal of Scientific Computing**; 58(2014), 705-725, **Google citation: 54**
7. **Wen Zaiwen**, Yin Wotao; A Feasible method for Optimization with Orthogonality Constraints; **Mathematical Programming**; 142(2013), 397-434, **Google citation: 1106**
8. Donald Goldfarb, **Wen Zaiwen**, Yin Wotao; A Curvilinear Search Method for the p-Harmonic Flow on Spheres; **SIAM Journal on Imaging Sciences**; 2 (2009) 84-109, **Google citation: 71**

Linear and nonlinear eigenvalue optimization

Zaiwen has enhanced the computational efficiency and scalability of eigenpair computations significantly. For linear eigenvalue problems, Zaiwen developed innovative block methods that surpass traditional Krylov subspace methods by enabling a higher degree of concurrency, crucial for modern multi-core computing environments. He introduced a limited memory block Krylov subspace optimization approach that accelerates the classic simultaneous iteration method and proposed an augmented Rayleigh-Ritz procedure combined with polynomial accelerators to optimize the computation of eigenpairs. These advancements in the software Arrabit have shown competitive, and often superior, performance against established eigensolvers such as ARPACK, FEAST, and SLEPc, as evidenced by extensive computational experiments. On the nonlinear front, Zaiwen and his coauthors addressed the challenges in quantum physics, quantum chemistry, and materials science, particularly with the Kohn-Sham (KS) and Hartree-Fock (HF) energy minimization problems. They discovered the violation of the Aufbau principle in certain models and established new analysis framework for the convergence of the widely used self-consistent field (SCF) method. He derived explicit expressions for the Hessian matrix and developed structured Newton-type methods that efficiently handle the complexities of these nonlinear eigenvalue problems, even outperforming SCF in many cases. His work also includes the first applicable algorithm for energy minimization in Bose-Einstein condensates (BECs) with arbitrary integer spin.

1. Tian Tonghua, Cai Yongyong, Wu Xinming, **Wen Zaiwen,** Ground States of Spin-F Bose–Einstein Condensates, **SIAM Journal on Scientific Computing**, Vol 42, No. 4, B983–B1013, 2020, **Google citation: 50**
2. **Wen Zaiwen,** Zhang Yin; Accelerating Convergence by Augmented Rayleigh-Ritz Projections For Large-Scale Eigenpair Computation; **SIAM Journal on Matrix Analysis and Applications**; Vol 38, No. 2 (2017), pp. 273-296, **Google citation: 9**
3. Jiang Bo, Liu Yafeng, **Wen Zaiwen**; Lp-norm regularization algorithms for optimization over permutation matrices; **SIAM Journal on Optimization**; Vol 26, No. 4(2016), pp. 2284-2313, **Google citation: 38**
4. Michael Ulbrich, **Wen Zaiwen,** Yang Chao, Dennis Klockner, Lu Zhaosong; A Proximal Gradient Method for Ensemble Density Functional Theory; **SIAM Journal on Scientific Computing**; Vol 37, No. 4 (2015), A1975–A2002, **Google citation: 24**
5. Liu Xin, **Wen Zaiwen**, Wang Xiao, Michael Ulbrich, Yuan Yaxiang; On the Analysis of the Discretized Kohn-Sham Density Functional Theory; **SIAM Journal on Numerical Analysis**; Vol 53, No. 4 (2015), 1758–1785, **Google citation: 42**
6. Liu Xin, **Wen Zaiwen**, Zhang Yin; An Efficient Gauss-Newton Algorithm for Symmetric Low-Rank Product Matrix Approximations; **SIAM Journal on Optimization**; Vol. 25, No. 3 (2015), 1571–1608, **Google citation: 45**

7. Liu Xin, Wang Xiao, **Wen Zaiwen**, Yuan Yaxiang; On the Convergence of the Self-Consistent Field Iteration in Kohn-Sham Density Functional Theory; **SIAM Journal on Matrix Analysis and Applications**; Vol. 35, No. 2(2014), pp. 546–558, **Google citation: 64**
8. Zhang Xin, Zhu Jinwei, **Wen Zaiwen**, Zhou Aihui; Gradient-type Optimization Methods for Electronic Structure Calculation; **SIAM Journal on Scientific Computing**; Vol. 36, No. 3 (2014), pp. C265–C289, **Google citation: 56**
9. **Wen Zaiwen**, Michael Ulbrich, Andre Milzarek, Zhang Hongchao; Adaptive Regularized Self-Consistent Field Iteration with Exact Hessian for Electronic Structure Calculation; **SIAM Journal on Scientific Computing**; Vol 35, No. 3 (2013) A1299–A1324, **Google citation: 21**
10. Liu Xin, **Wen Zaiwen**, Zhang Yin; Limited Memory Block Krylov Subspace Optimization for Computing Dominant Singular Value Decompositions; **SIAM Journal on Scientific Computing**; 35 (2013) A1641–A1668, **Google citation: 81**

Deep learning and Reinforcement learning

Zaiwen has provided approaches in the optimization landscape of deep learning and reinforcement learning (RL), focusing on leveraging second-order structures to enhance the efficiency and stability of training methodologies. His contribution is prominent in the development of the NG+ method, a generalized natural gradient technique that simplifies the approximation of the Fisher information matrix by directly working with matrix-type variables. This method substantially reduces the computational overhead associated with matrix inversion, making second-order optimization methods more feasible for large-scale deep learning applications. In addition, Zaiwen introduced the sketch-based empirical natural gradient method and structured stochastic quasi-Newton methods by incorporating local curvature information into optimization processes without the prohibitive costs of explicit Hessian storage. His development of the stochastic augmented Lagrangian function algorithm for RL, by ingeniously overcoming sampling difficulties related to conditional expectations and quadratic penalties, is new for solving the optimal Bellman equation. Further contributions include establishing a sample complexity lower bound for constrained Markov decision processes (CMDPs) and proposing an approximately optimal primal-dual learning algorithm, which guarantees zero constraint violation and exhibits near-optimal sample complexity. Moreover, his methods for state aggregation in finite-state Markov chains have been efficient for decision-making with compressed state spaces.

1. Wu Jiayuan, Hu Jiang, Hongchao Zhang, **Wen Zaiwen,** Convergence analysis of an adaptively regularized natural gradient method, **IEEE Transactions on Signal Processing**, 2024
2. Ke Zhifa, Zhang Junyu, **Wen Zaiwen**, An Improved Finite-time Analysis of Temporal Difference Learning with Deep Neural Networks, **ICML** 2024
3. Yang Minghan, Xu Dong, **Wen Zaiwen**, Chen Mengyun, Xu Pengxiang, Sketchy-based Empirical Natural Gradient Methods for Deep Learning, **Journal of Scientific Computing**, Vol. 9, No. 2, 94, 2023, **Google citation: 23**
4. Li Yongfeng, Zhao Mingming, Chen Weijie, **Wen Zaiwen**, A Stochastic Composite Augmented Lagrangian Method For Reinforcement Learning, **SIAM Journal on Optimization**, Vol. 33, No.2, 921-949, 2023, **Google citation: 2**
5. Yang Minghan, Andre Milzarek, **Wen Zaiwen**, Zhang Tong; A Stochastic Extra-Step Quasi-Newton Method for Nonsmooth Nonconvex Optimization, **Mathematical Programming**, Vol. 194, 257-303, 2022, **Google citation: 22**
6. Yang Minghan, Xu Dong, Cui Qiwen, **Wen Zaiwen,** Xu Pengxiang, An Efficient Fisher Matrix Approximation Method for Large-Scale Neural Network Optimization, **IEEE Transactions on Pattern Analysis and Machine Intelligence**, 2022, **Google citation: 7**
7. Chen Fan, Zhang Junyu, **Wen Zaiwen,** A Near-Optimal Primal-Dual Method for Off-Policy Learning in CMDP, **NeurIPS** 2022, **Google citation: 6**
8. Duan Yaqi, Wang Mengdi, **Wen Zaiwen**, Yuan Yaxiang; Adaptive Low Rank Approximation for State Aggregation of Markov Chain, **SIAM Journal on Matrix Analysis and Applications**, Vol 41, No. 1, pp. 244-278, 2020, **Google citation: 9**