CDC 2026 Workshop Proposal

Generative AI Meets Control and Optimization: Theory, Algorithms, and Systems

Runyu Zhang, Jiawei Zhang, Asuman Ozdaglar

Abstract

Recent advances in generative AI have led to remarkable empirical success, but their theoretical understanding remains limited, particularly with respect to efficiency, reliability, and controllability. Addressing these challenges calls for principled frameworks that can provide guarantees and systematic design methodologies. Control theory and optimization offer a natural foundation in this regard, with tools to analyze dynamical behavior, enforce constraints, and design stable and efficient algorithms.

This workshop brings together researchers from control, optimization, and machine learning to explore how system-theoretic and optimization-based perspectives can advance the understanding and design of generative models, including diffusion models and large language models. We are particularly interested in approaches that enable efficient and reliable training and inference, as well as provide insights into stability and convergence properties of generative systems. We also welcome contributions that leverage generative models within control and optimization pipelines, highlighting the growing interplay between these fields.

1 General Information

Type and duration: **Workshop, full day, December 14th, 2026**

Title: **Generative AI Meets Control and Optimization: Theory, Algorithms, and Systems**

A website for the workshop will be prepared and released.

2 Organizers

Dr. Runyu Zhang (primary point of contact): Postdoctoral scholar at MIT.
Address: 32 Vassar Street, D670, Cambridge, MA;
Phone: +1 6173352245
E-mail: runyuzha@mit.edu. Website: dianyu420376.github.io/runyu-cathy-zhang.github.io

Bio: Runyu (Cathy) Zhang is a postdoctoral researcher at MIT Laboratory for Information & Decision Systems. Before joining MIT, she obtained her Ph.D. degree at Harvard University, School of Engineering and Applied Sciences. Prior to pursuing her Ph.D., she earned a B.S. degree in mathematics from Peking University in 2019. Her research interests lie in learning-based control, reinforcement learning, game theory, and optimization, with particular focus on multi-agent systems. She has received the MIT Postdoctoral Fellowship for Engineering Excellence, was selected as an EECS Rising Star in 2024, and was a finalist for the Two Sigma Diversity PhD Fellowship in 2022.

Prof. Jiawei Zhang: Assistant Professor at University of Wisconsin–Madison
Address: 1205 University Ave, 701, Madison, WI 53706;
E-mail: jzhang2924@wisc.edu. Website: https://jwzhang.netlify.app/

Bio: Jiawei Zhang is an assistant professor in the Department of Computer Sciences at the University of Wisconsin–Madison. Prior to joining UW–Madison, he was a postdoctoral researcher in the Laboratory for Information & Decision Systems (LIDS) at the Massachusetts Institute of Technology, working with Asuman Ozdaglar and Saurabh Amin. He received his PhD in Computer and Information Engineering from the Chinese University of Hong Kong, Shenzhen under the supervision of Zhi-Quan Luo. He earned his BSc in Mathematics (Hua Loo-Keng Talent Program) from the University of Science and Technology of China.

His research interests include: • Optimization theory and algorithms • Theoretical foundations and algorithm design for generative models, including large language models and diffusion models • Optimization algorithms for engineering applications, including energy systems, communications, and manufacturing

Prof. Asuman Ozdaglar: MathWorks Professor of Electrical Engineering and Computer Science Department Head, Electrical Engineering and Computer Science & Deputy Dean of Academics, Schwarzman College of Computing at MIT
Address: Building 38-403, Cambridge, MA;
E-mail: asuman@mit.edu. Website: https://asu.mit.edu/

Bio: Professor Ozdaglar's research focuses on technical and societal aspects of large-scale data-driven systems. Her expertise includes optimization, machine learning, economics, and networks. In recent years, she has investigated issues of data ownership and markets, spread of misinformation on social media, economic and financial contagion, and social learning. In addition, she has an active research program on large-scale optimization, especially in the context of machine learning. Her recent work develops robust, efficient and decentralized machine learning models and algorithms. In optimization, her work has contributed to optimization duality, first-order scalable methods and new distributed algorithms for network resource allocation. Her work has also contributed to the study of adaptive game-theoretic dynamics and introduced alternative approaches to network games.

3 Content

3.1 Motivation and relevance to the CDC community

Recent years have witnessed rapid progress in generative artificial intelligence, including diffusion models and large language models, which have demonstrated remarkable capabilities across a wide range of domains. Despite these empirical successes, a principled understanding of these systems—particularly in terms of stability, efficiency, reliability, and controllability—remains limited. This gap presents a significant opportunity for the control and optimization community.

From a control and optimization perspective, modern generative models can be naturally viewed as high-dimensional dynamical systems, where learning and inference correspond to iterative stochastic processes. For example, diffusion models evolve through a sequence of denoising steps that can be interpreted as discretizations of stochastic differential equations [1], while large language models generate outputs via autoregressive dynamics [2]. These perspectives connect generative modeling to classical topics in control and dynamical systems, including stability, convergence, adaptivity and robustness of iterative processes. Further, the pre-training and fine-tuning of modern generative AI models can be naturally formulated as large-scale stochastic optimization problems [3–5], which raises new challenges in understanding convergence, generalization, and efficiency in highly nonconvex and overparameterized regimes. Moreover, even inference-time deployment can often be interpreted through an optimization lens. For example, diffusion-based generative models perform iterative updates that resemble stochastic gradient-based sampling procedures, progressively guiding samples toward regions of higher data likelihood [1]. More broadly, recent approaches to controlled generation and reinforcement learning leverage optimization-based formulations to guide outputs toward user-specified objectives while remaining consistent with the underlying data distribution [6, 7]. These perspectives highlight both the opportunities and challenges in designing principled, efficient, and reliable optimization-based methods for generative systems. As such, the above discussions raise fundamental questions that are central to the CDC community: how to analyze and guarantee stability and convergence of these high-dimensional stochastic dynamics, how to enforce constraints and ensure safety during generation, and how to design more efficient and robust algorithms leveraging insights from optimization and control.

At the same time, generative models are increasingly being integrated into control and decision-making pipelines. Recent works have demonstrated their effectiveness in trajectory generation and planning [6], as well as in representing policies and value functions in reinforcement learning [8]. Generative models have also been explored for system identification and modeling of complex dynamical systems [9]. While these approaches offer new capabilities for handling high-dimensional and uncertain environments, they also introduce challenges related to reliability, interpretability, and deployment in safety-critical settings. Addressing these challenges requires methodologies that go beyond purely data-driven approaches and instead leverage structure, guarantees, and principled design tools from control and optimization.

Control theory and optimization provide a natural foundation to tackle these issues. Concepts such as stability/convergence analysis, feedback design, constrained optimization, adaptive control/optimization and robustness offer powerful lenses through which generative systems can be better understood and improved. Conversely, the complexity and scale of modern generative models motivate new theoretical developments in control and optimization, including scalable algorithms, learning-based system identification, and analysis of nonconvex and stochastic dynamical systems.

This workshop aims to promote a dialogue between these communities by bringing together researchers working at the intersection of control, optimization, and generative AI. It is highly relevant to the CDC audience, as it highlights emerging research directions where control-theoretic thinking can play a central role in shaping the next generation of intelligent systems. By connecting foundational theory with modern machine learning practice, the workshop seeks to broaden the scope of control research and identify impactful directions for future work.

3.2 Objectives

The primary objective of this workshop is to build a bridge between the control and optimization community and the rapidly evolving field of generative AI. In particular, the workshop seeks to promote a deeper, system-theoretic understanding of generative models while exploring how these models can be effectively incorporated into control and decision-making frameworks.

To this end, the workshop will pursue the following goals:

- **Develop a control- and optimization-based perspective on generative models.** Highlight how tools such as dynamical systems analysis, stability theory, and constrained optimization can be used to understand and design generative processes, including both training and inference dynamics. Identify and discuss open problems related to stability, robustness, safety, and computational efficiency in large-scale generative systems, particularly in high-dimensional and stochastic settings.
- **Explore the integration of generative models into control pipelines.** Examine how generative AI can be used for system modeling, trajectory generation, policy design, and decision-making under uncertainty, while emphasizing the need for principled and reliable approaches. Present recent progress on topics such as diffusion-based control, generative modeling for planning and reinforcement learning, optimization-based architectures, and provable guarantees for learning in dynamical systems.
- **Encourage interdisciplinary collaboration and community building.** Bring together researchers from control, optimization, and machine learning to exchange ideas, identify common challenges, and stimulate new collaborations, particularly among early-career researchers.

Overall, the workshop aims to articulate a coherent research agenda at the intersection of generative AI, control, and optimization—one that combines empirical advances with rigorous theoretical foundations and aligns closely with the core interests of the CDC community.

3.3 Detailed Schedule

The workshop is organized into three sessions that reflect the central themes outlined in Section 3.1. The morning session focuses on the learning dynamics and foundational aspects of generative AI, followed by a session on the interplay of genAI with control- and dynamical-systems perspectives. The afternoon session addresses optimization, learning, and decision-making with generative models. Each talk is allocated 30 minutes (25-minute presentation and 5-minute Q&A). The more detailed schedule is reported in Table 1.[1]

In the following, we present the list of invited speakers together with their biographies. To address the motivation outlined in Section 3.1, we have invited a group of researchers who are leading contributors to the emerging interface of generative AI, control, and optimization. The speakers bring complementary expertise spanning dynamical systems, learning theory, optimization, and decision-making, ensuring both depth and breadth in the perspectives represented.

We are particularly pleased with the diversity of the speaker lineup, not only in terms of research areas and methodologies, but also across career stages and institutional backgrounds. This diversity is essential for fostering a rich and balanced discussion on the challenges and opportunities at this rapidly evolving intersection. [2]

Elad Hazan Elad Hazan is a professor of computer science at Princeton University. His research focuses on the design and analysis of algorithms for basic problems in machine learning and optimization. Among his contributions are the co-invention of the AdaGrad algorithm for deep learning, the first sublinear-time algorithms for convex optimization, and online nonstochastic control theory. He is the recipient of the Bell Labs Prize, the IBM Goldberg best paper award twice, a European Research Council grant, a Marie Curie fellowship, Google Research Award and is an ACM fellow. He served on the steering committee of the Association for Computational Learning and was program chair for the Conference on Learning Theory 2015. He is the co-founder and director of Google AI Princeton.

Kaiqing Zhang Kaiqing Zhang is currently an Assistant Professor at the Department of Electrical and Computer Engineering (ECE) and the Institute for Systems Research (ISR), with also joint appointment at the Department of Computer Science (CS), at the University of Maryland, College Park. He is also a member of the Center for Machine Learning and Maryland Robotics Center. Prior to joining Maryland, he was a postdoctoral scholar affiliated with LIDS and CSAIL at MIT, and a Research Fellow at the Simons Institute for the Theory of Computing at Berkeley. He obtained his Ph.D. from the Department of ECE at the University of Illinois at Urbana-Champaign (UIUC). He also received M.S. in both ECE and Applied Mathematics from UIUC, and B.E. in Automation with a second major in Economics from Tsinghua University. His research interests lie in Control and Decision Theory, Game Theory, Reinforcement/Machine Learning, Robotics, Computation, and their intersections. His work has been recognized by several awards, including Simons-Berkeley Research Fellowship, Coordinated Science Lab Thesis Award, ICML Outstanding Paper Award, AAAI New Faculty Highlights, NSF CAREER Award, AFOSR YIP Award, faculty research awards from Cisco, Open Philanthropy, and JP Morgan, and the George Corcoran Memorial Award for Educational Leadership.

Max Simchowitz Max Simchowitz is an assistant professor at the Machine Learning Department at Carnegie Mellon University with a courtesy appointment in the Robotics Institute. His work studies theoretical foun-

[1] Specific times can be adapted based on the detailed conference requirements (e.g., to synchronize coffee breaks)

[2] As the workshop is currently in the proposal stage, some speakers have indicated tentative availability due to potential scheduling conflicts. To ensure a robust and well-balanced program, we have intentionally invited a slightly larger pool of speakers. The final schedule and length of the talks will be adjusted accordingly while maintaining coherence and sufficient discussion time.

Time	Speaker	Title
09:00-09:10	Runyu Zhang & Jiawei Zhang & Asuman Ozdaglar	*Welcome and introduction to the workshop*
		Session I: Learning Foundations
09:00-09:30	Prof. Yuejie Chi	On the Learning Dynamics of RLVR at the Edge of Competence
09:30-10:00	Prof. Meisam Razaviyayn	From Theory to Throughput: Unifying Architectures and Scaling Deep Memory via Online Optimization
10:00-10:30	Dr. Jingfeng Wu	Towards a Less Conservative Theory of Machine Learning: Unstable Optimization and Implicit Regularization
		Coffee break
		Session II: Interplay of GenAI with Control and Dynamical Systems
11:00-11:30	Dr. Carmen Amo Alonso	Controllable AI: Artificial Intelligence Meets Control Theory
11:30-12:00	Prof. Max Simchowitz	A Mathematical Basis for Moravec's Paradox, and Some Open Problems
		Lunch break
14:00-14:30	Prof. Elad Hazan	Provably Efficient Learning in Nonlinear Dynamical Systems via Spectral Transformers
		Session III: Interplay of GenAI with Optimization and Decision-Making
14:30-15:00	Prof. Jiawei Zhang	Optimization for Efficient Training and Inference
		Coffee break
15:30-16:00	Prof. Kaiqing Zhang	Towards Understanding and Post-training LLMs as Decision-Making Agents: A Regret-based Approach
16:00-16:30	Andrey Kharitenko	Constrained Optimization through Denoising and Applications to Data-Driven Control
16:30-17:00	Runyu Zhang & Jiawei Zhang & Asuman Ozdaglar	*Concluding and Discussions*

Table 1: Proposed schedule for the workshop, organized around themes spanning learning dynamics, control-theoretic perspectives, and optimization and decision-making in generative systems.

dations and new methodologies for machine learning problems with an interactive, sequential, or dynamical component, currently focusing on reinforcement learning and applications to robotics. His past work has ranged broadly across control, theoretical reinforcement learning, optimization and algorithmic fairness. He received his PhD from University of California, Berkeley in 2021 under Ben Recht and Michael I. Jordan, and completed his postdoctoral research under Russ Tedrake in the Robot Locomotion Group at MIT. His work has been recognized with an ICML 2018 Best Paper Award, ICML 2022 Outstanding Paper Award, and RSS 2023 and ICRA 2024 Best Paper Finalist designations.

Meisam Razaviyayn Meisam Razaviyayn is an Associate Professor (with the Andrew and Erna Viterbi Early Career Chair) of Industrial and Systems Engineering, Computer Science, Quantitative and Computational Biology, and Electrical Engineering at the University of Southern California. He also serves as Associate Director of the USC–Meta Center for Research and Education in AI and Learning (REAL@USC) and is a part-time Research Scientist at Google Research. His research focuses on the design and study of scalable, trustworthy optimization algorithms for modern data science and machine learning applications. Prior to joining USC, he was a postdoctoral research fellow in the Department of Electrical Engineering at Stanford University and a Visiting Scientist at the Simons Institute for the Theory of Computing at UC Berkeley. He received his Ph.D. in Electrical Engineering (minor in Computer Science) from the University of Minnesota, where he also earned his M.Sc. in Mathematics. He is the recipient of the 2022 NSF CAREER Award, the

2022 Northrop Grumman Excellence in Teaching Award, the 2021 AFOSR Young Investigator Award, the 2021 3M Nontenured Faculty Award, the 2020 ICCM Best Paper Award in Mathematics, the 2019 IEEE Data Science Workshop Best Paper Award, and the 2014 IEEE Signal Processing Society Young Author Best Paper Award. He was among the 40 selected attendees of the 2023 German–American Frontiers of Engineering Symposium organized by the National Academy of Engineering, and is a silver medalist of Iran's National Mathematics Olympiad.

Yuejie Chi Yuejie Chi is the Charles C. and Dorothea S. Dilley Professor of Statistics and Data Science at Yale University, with a secondary appointment in Computer Science, and a member of Yale Institute for Foundations of Data Science. She received her Ph.D. and M.A. from Princeton University, and B. Eng. (Hon.) from Tsinghua University, all in Electrical Engineering. Her research interests lie in the theoretical and algorithmic foundations of data science, generative AI, reinforcement learning, and signal processing, motivated by applications in scientific and engineering domains. Among others, Dr. Chi received the Presidential Early Career Award for Scientists and Engineers (PECASE), SIAM Activity Group on Imaging Science Best Paper Prize, IEEE Signal Processing Society Young Author Best Paper Award, and the inaugural IEEE Signal Processing Society Early Career Technical Achievement Award. She is an IEEE Fellow (Class of 2023) for contributions to statistical signal processing with low-dimensional structures.

Carmen Amo Alonso Carmen Amo Alonso is a Schmidt Science Fellow affiliated with the Autonomous Systems Lab (led by Prof. Marco Pavone) at Stanford University. Her research lies at the intersection of control theory, optimization, and artificial intelligence (AI) with a focus on the principled design of reliable and controllable AI technologies. Carmen's work seeks to adapt mathematical principles from control theory, traditionally used to ensure safety and predictability in engineered systems, to understand, control, and ultimately improve the behavior of AI systems, including generative models for language applications and embodied systems. The impact of her research has led to various collaborations with industry, including Google and Apple. At Stanford, Carmen was named an Emerson Consequential Scholar for the potential of her research to positively impact society. Prior to joining Stanford, she was a Fellow at the Artificial Intelligence Center at ETH Zurich. Carmen earned her Ph.D. in Control and Dynamical Systems from Caltech in 2023, where she was advised by Prof. John Doyle. Her thesis was awarded the Milton and Francis Clauser Doctoral Prize, which recognizes the best Ph.D. dissertation of the year across all disciplines at Caltech. During her Ph.D., her research received two IEEE best paper awards, was partially funded by Amazon and D. E. Shaw fellowships, and earned her three Rising Star titles (EECS, Cyber-Physical Systems, and Brain and Cognitive Sciences). Besides her research collaborations across academia and industry, Carmen is committed to education for all. As a member of Clubes de Ciencia, she travels to Mexico in the summer to teach underserved students.

Jingfeng Wu Jingfeng Wu is a postdoctoral fellow at the Simons Institute for the Theory of Computing at UC Berkeley. His research focuses on deep learning theory, optimization, and statistical learning. He earned his Ph.D. in Computer Science from Johns Hopkins University in 2023. Prior to that, he received a B.S. in Mathematics (2016) and an M.S. in Applied Mathematics (2019), both from Peking University. In 2023, he was recognized as a Rising Star in Data Science by the University of Chicago and UC San Diego.

Andrey Kharitenko Andrey Kharitenko is a PhD student at the Optimization & Decision Intelligence Group at ETH Zurich under the supervision of Niao He and Florian Dorfler. He received this B.Sc. and M.Sc. in mathematics from the University of Stuttgart in 2019 and 2022 respectively. In 2024 he additionally received his M.Sc. degree in control engineering from TU Eindhoven. His interests include the generative modeling for optimization and its applications to data-driven control.

4 Engaging the Research Community

4.1 Plan to solicit participation

The organizers have extensive experience in hosting workshops at major conferences and organizing international seminar series. Through these experiences, we have developed effective strategies to solicit broad participation that extend well beyond traditional mailing lists. Our outreach plan typically includes:

- **Direct invitations**: personally contacting research groups and inviting interested students and researchers to participate. The organizers and speakers maintain active connections with leading institutions such as MIT, Harvard University, CMU, Stanford, Berkeley, University of Wisconsin-Madison, Princeton, ETH, USC.
- **Social media outreach**: promoting the event through professional platforms including Facebook, LinkedIn, and X (formerly Twitter), which has proven effective in engaging a wide audience.

To ensure accessibility and continued engagement, we also plan to provide a comprehensive "learning packet" (see Section 4.3) containing lecture notes, slides, and recordings from the workshop. This will allow participants—especially those unable to attend in person—to benefit from the discussions and materials, thereby extending the workshop's impact beyond the event itself.

4.2 Plan to encourage interaction among participants

The core of the proposed workshop is active interaction among participants, which is particularly important given the diverse communities it brings together. The program is designed to encourage engagement through a combination of in-depth talks and structured discussion. Each session allocates time for questions and discussion, encouraging participants to engage with speakers and exchange ideas. Rather than purely lecture-style presentations, the sessions are intended to promote dialogue and interaction across different backgrounds and areas of expertise.

To further support interaction, we will:

- allocate dedicated discussion and question time following presentations;
- facilitate informal interactions during coffee and lunch breaks;
- promote follow-up discussions among participants and speakers after the sessions.

4.3 Dissemination

To extend the impact of the workshop beyond the event itself, we will prepare a concise set of materials that summarize key insights and resources discussed during the sessions. These materials are intended to support both attendees and those unable to participate in person.

Specifically, we plan to provide:

Curated references to relevant literature and resources highlighted by the speakers;
A brief summary of key themes, open challenges, and emerging directions discussed during the workshop.

These materials will be made available on the workshop website following the event.

5 Diversity Statement

The workshop features a carefully curated group of speakers whose expertise spans multiple disciplines at the intersection of control, optimization, and modern machine learning. In particular, the invited speakers bring complementary perspectives ranging from dynamical systems and control theory to learning theory, optimization, and generative AI, reflecting the interdisciplinary nature of the topic. The speaker lineup also represents diversity across career stages, including graduate student, postdoctoral researchers, early-career faculty, and established leaders in the field, which helps create a balanced exchange of ideas and perspectives. In addition, we have made an intentional effort to include speakers from a variety of institutional backgrounds and to promote broader participation in the community.

References

[1] Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole, "Score-based generative modeling through stochastic differential equations," in *ICLR*, 2021.

[2] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," *NeurIPS*, 2017.

[3] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *ICLR*, 2015.

[4] L. Bottou, F. E. Curtis, and J. Nocedal, "Optimization methods for large-scale machine learning," *SIAM Review*, 2018.

[5] K. Jordan. Muon: An optimizer for hidden layers in neural networks. Blog post. [Online]. Available: https://kellerjordan.github.io/posts/muon/

[6] M. Janner, Y. Du, J. Tenenbaum, and S. Levine, "Planning with diffusion for flexible behavior synthesis," *ICML*, 2022.

[7] H. Chung, B. Sim, D. Ryu, and J. C. Ye, "Improving diffusion models for inverse problems using manifold constraints," *Advances in Neural Information Processing Systems*, vol. 35, pp. 25 683–25 696, 2022.

[8] Z. Wang, J. J. Hunt, and M. Zhou, "Diffusion policies as an expressive policy class for offline reinforcement learning," in *ICLR*, 2023.

[9] R. T. Chen, Y. Rubanova, J. Bettencourt, and D. K. Duvenaud, "Neural ordinary differential equations," *NeurIPS*, 2018.