ISBN: 978-3-95450-261-5 ISSN: 2673-7000 doi: 10.18429/JACoW-NAPAC2025-MOP076

PRELIMINARY STUDY OF AUTO-DIFFERENTIATION ALGORITHM IN BEAM DYNAMICS WITH STOCHASTIC PROCESS*

C. T. Ratcliff†, Y. Hao, Facility for Rare Isotope Beams, East Lansing, MI, USA

Abstract

Modern particle accelerator optimization requires sophisticated computational methods to address the inherently stochastic nature of beam dynamics. This research develops a framework that applies AD to stochastic differential equations that specifically addresses beam dynamics challenges in particle accelerators, focusing on accurately modeling and optimizing beam behavior in regimes dominated by stochastic processes. By incorporating key physical phenomena such as synchrotron radiation, wakefield effects, and quantum excitation, the framework aims to provide auto-differentiation on the figure of merit of the phase space evolution and beam dynamics. The methodology will enable an effective optimization method in a dynamic system with stochastic process.

INTRODUCTION

Optimizing particle accelerator performance in the presence of stochastic processes—such as synchrotron radiation, intra-beam scattering, or other random effects—can be particularly challenging [1]. The system's equilibrium behavior is typically determined through complex simulations, which often yield only the expected value of the stochastic process. When gradients are estimated using finite differences across multiple simulations, the resulting derivatives tend to exhibit high variance, making optimization less efficient and less reliable.

Automatic differentiation (AD) has emerged as a powerful tool for optimization in machine learning and scientific computing [2]. However, extending AD to stochastic differential equations (SDEs) presents challenges. Advances in stochastic AD, particularly the development of stochastic triple methods, offer new possibilities for efficient gradient computation in stochastic systems [3].

This work presents a study of applying stochastic AD algorithms to beam dynamics simulations in electron storage rings. The approach uses longitudinal dynamics with synchrotron radiation damping and excitation as a benchmark to validate the accuracy of gradient computations in a controlled setting where analytical solutions provide ground truth comparisons [4].

STOCHASTIC TRIPLE METHODOLOGY

For a stochastic program $X(p)$ with parameter p, a stochastic triple takes the form:

$$\text{stochastic triple} = v + d \cdot \varepsilon + (j \text{ with probability } w \cdot \varepsilon) \quad (1)$$

where v represents the primal value, d is the dual component analogous to traditional automatic differentiation, j captures the jump value for discrete transitions, and w denotes the jump probability weight. The parameter ε represents an infinitesimal perturbation.

The derivative of the expected value is computed through the relationship:

$$\frac{dE[X(p)]}{dp} = E[\delta + w(Y - X(p))], \quad (2)$$

where δ represents the deterministic derivative component, w is the weight factor, and Y captures finite perturbations with infinitesimal probabilities.

This methodology handles both discrete random variables such as Bernoulli and binomial distributions and continuous distributions within a unified framework, accounting for jump processes common in accelerator physics simulations including quantum excitation events and discrete radiation emission processes. Unlike conventional automatic differentiation, stochastic triples can propagate derivative information through non-differentiable stochastic operations while preserving statistical properties of the underlying random processes. The primary advantage over finite difference methods lies in providing derivative information through single program execution rather than requiring multiple function evaluations, significantly reducing computational overhead while maintaining statistical rigor in derivative estimates.

A comparison between various stochastic triple methods and other derivative methods for a simple random walk with 100 steps can be seen in Fig. 1. It can be seen that, excluding the score function, all of them grow at similar rates, but the two stochastic triples have the least growth in the standard deviation of the derivative. The smoothed stochastic triple is a method that would not very applicable accelerator methodologies, but it is a potential draw of the package.

ELECTRON STORAGE RING SIMULATION

The validation framework employs an electron storage ring (ESR) configuration designed to achieve equilibrium distributions that approximate the Haissinski distribution [4]. This semi-analytical benchmark provides a test case for later validation of derivative accuracy because the equilibrium properties can be computed in a semi-theoretical method and compared against simulation results.

The simulation incorporates the physical processes governing electron beam evolution in storage rings [1]. RF cavity interactions provide longitudinal acceleration, while synchrotron radiation introduces both damping and quantum

* Work supported by the U.S. Department of Energy Office of Science under award number DE-SC0024170 and Cooperative Agreement DE-SC0023633, the State of Michigan, and Michigan State University.
† ratcliff@frib.msu.edu

ISBN: 978-3-95450-261-5 ISSN: 2673-7000 doi: 10.18429/JACoW-NAPAC2025-MOP076

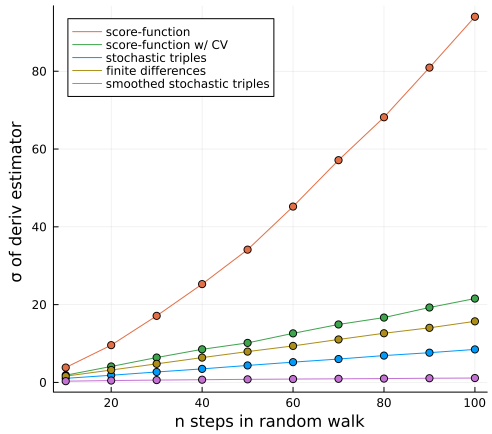

Figure 1: Conceptual diagram showing stochastic triple propagation through beam dynamics operations.

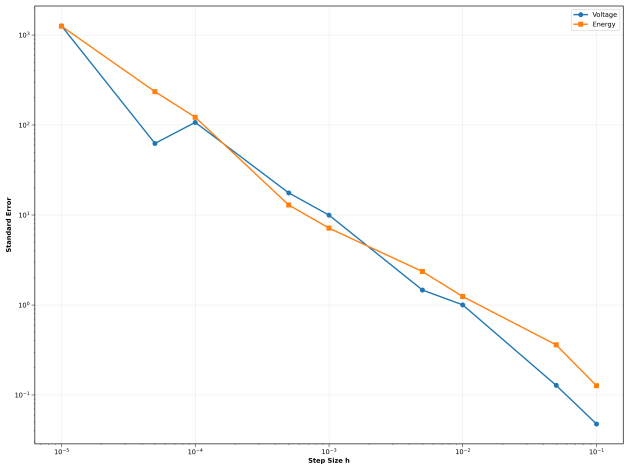

Figure 2: Comparison between various step sizes and the standard errors in the derivative estimates that they lead to.

excitation. The balance between these effects determines the equilibrium bunch properties.

The simulation parameters correspond to an electron storage ring with beam energy of 4 GeV, cavity voltage of 5 MV, and harmonic number of 360. The synchrotron phase is set to 150 degrees to ensure stable operation, with momentum compaction factor $\alpha_c = 3.68 \times 10^{-4}$.

Particles are initialized with Gaussian distributions in both longitudinal position and energy, with parameters chosen to match the expected equilibrium values. The evolution proceeds through multi-turn tracking, with each turn applying RF kicks, synchrotron radiation, and quantum excitation in sequence.

COMPARATIVE ANALYSIS WITH FINITE DIFFERENCE

The accuracy and efficiency of stochastic AD were evaluated through comparison with finite difference methods across parameter sensitivities. Two parameters were investigated: RF voltage scaling and beam energy scaling, both critical for accelerator optimization.

For finite difference calculations, central difference formulas were employed with step sizes ranging from 10^{-5} to 10^{-1}. The step size of $h = 5 \times 10^{-3}$ was determined through convergence analysis, balancing truncation error against numerical noise. As demonstrated in Fig. 2, finite difference methods exhibit strong dependence on step size selection, with derivative estimates varying significantly across the tested range, with a clear steady increase in error when the step size is decreased. This is a clear advantage that AD presents, as it is not affected by the stochasticity to near the same degree as the finite difference implementation. This sensitivity necessitates extensive parameter studies to identify optimal step sizes, consuming substantial computational resources before meaningful optimization can begin.

The uncertainty of the derivative when the number of particles in the simulation is scaled can be seen in Figs. 3 and 4. They decrease at a similar rate, but the StochasticAD method has a lower uncertainty the entire time.

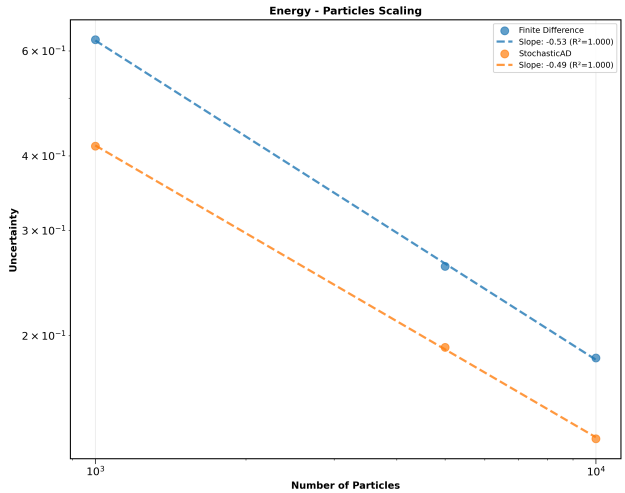

Figure 3: Comparison of derivative estimates between stochastic AD and finite difference methods for energy sensitivities by scaling the number of particles.

Figures 5 and 6 show the uncertainty of the derivative that occurs when the number of samples are decreased. Similar to when the number of samples was scaled, they decrease at a similar rate but the uncertainty in the StochasticAD method remained lower everywhere.

FUTURE WORK

The next phase of development focuses on integrating stochastic AD capabilities into established accelerator simulation codes. JuTrack.jl, a Julia-based particle tracking library, represents the primary target for refactoring to support stochastic triple propagation [5, 6]. This integration requires modifying the core tracking routines to handle stochastic triple arithmetic while maintaining compatibility with existing lattice definitions and beam dynamics calculations.

ISBN: 978-3-95450-261-5 ISSN: 2673-7000 doi: 10.18429/JACoW-NAPAC2025-MOP076

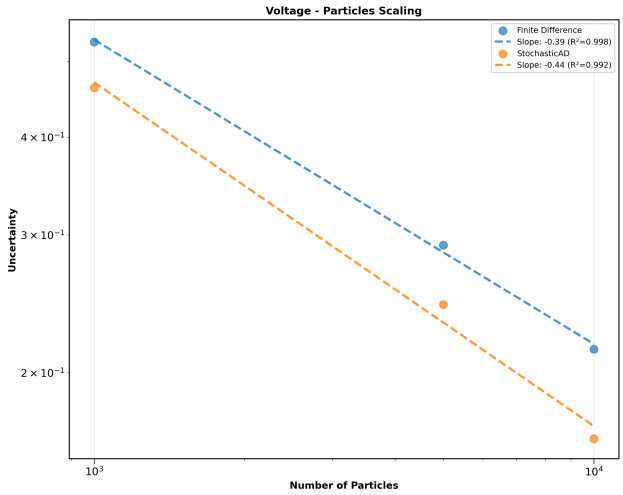

Figure 4: Comparison of derivative estimates between stochastic AD and finite difference methods for voltage sensitivities by scaling the number of particles.

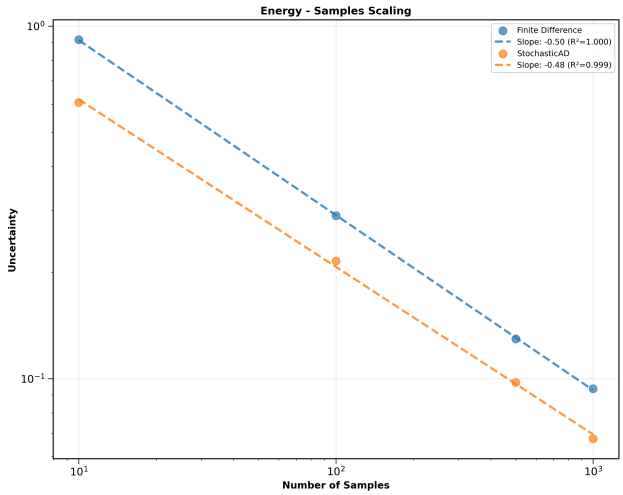

Figure 5: Comparison of derivative estimates between stochastic AD and finite difference methods for energy sensitivities by scaling the number of samples that were averaged.

Performance optimization remains a consideration for practical deployment. Current implementations show promise but require further development to handle large-scale simulations with realistic lattices. Memory management and computational efficiency must be optimized to compete with highly-tuned conventional tracking codes.

In the short term, the inclusion of wakefield effects represents a physics extension that will test the robustness of the stochastic AD approach. Collective effects introduce additional complexity through coupled particle motion..

Higher-order derivative calculations will enable advanced optimization algorithms and uncertainty quantification studies. The stochastic triple framework can be extended to compute second and higher-order derivatives, potentially offering computational advantages for optimization methods that require Hessian information.

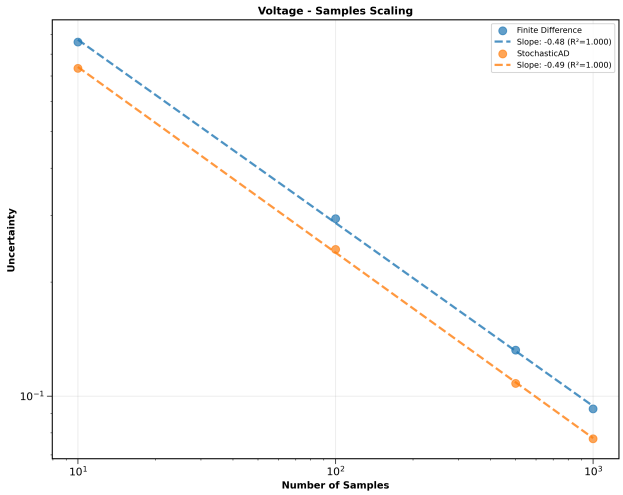

Figure 6: Comparison of derivative estimates between stochastic AD and finite difference methods for voltage sensitivities by scaling the number of samples.

Another short term decision to be made is whether to continue with the StochasticAD package, or shift to an custom implementation of the underlying math [3]. Switching to a custom method would represent a siginifant upfront time investment, but gains would be made by improving backend efficency and allow it to better worth with the simulation codes.

CONCLUSION

This study demonstrates the feasibility and advantages of applying stochastic automatic differentiation to beam dynamics simulations [3]. The stochastic triple methodology handles the quantum stochastic processes in electron storage rings while providing derivative information for optimization.

The work establishes a foundation for comprehensive studies incorporating collective effects and multi-objective optimization. The demonstrated accuracy and efficiency suggest that stochastic AD could become a tool for accelerator design and operation optimization, particularly in scenarios where gradient-based methods offer advantages over derivative-free approaches.

Future developments will focus on expanding the physics coverage and demonstrating optimization applications. The goal is to enable efficient optimization of accelerator systems by leveraging the advantages of stochastic automatic differentiation in handling the probabilistic nature of beam dynamics [3].

REFERENCES

[1] M. Sands, “The physics of electron storage rings: An introduction”, Stanford Linear Accelerator Center, Calif., Tech. Rep. SLAC-R-121, 1970. doi:10.2172/4064201

ISBN: 978-3-95450-261-5 ISSN: 2673-7000 doi: 10.18429/JACoW-NAPAC2025-MOP076

[2] A. G. Baydin, B. A. Pearlmutter, A. A. Radul, and J. M. Siskind, "Automatic differentiation in machine learning: A survey", *J. Mach. Learn. Res.*, vol. 18, no. 1, pp. 5595–5637, 2017.

[3] G. Arya, M. Schauer, F. Schäfer, and C. Rackauckas, "Automatic differentiation of programs with discrete randomness", in *Proc. NIPS'22*, New Orleans, USA, Nov.-Dec. 2022, pp. 10 435–10 447.

[4] J. Haissinski, "Exact longitudinal equilibrium distribution of stored electrons in the presence of self-fields", *Nuovo Cim. B*, vol. 18, pp. 72–82, 1973. doi:10.1007/BF02832640

[5] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah, "Julia: A fresh approach to numerical computing", *SIAM Review*, vol. 59, no. 1, pp. 65–98, 2017. doi:10.1137/141000671

[6] J. Wan, H. Alamprese, C. Ratcliff, J. Qiang, and Y. Hao, "Jutrack: A julia package for auto-differentiable accelerator modeling and particle tracking", *Comput. Phys. Commun*, vol. 309, p. 109 497, 2025. doi:10.1016/j.cpc.2024.109497