Agentic-AI Detector Co-design and Optimization in Vertically-Integrated Differentiable Full Simulations

Wonyong Chung
Princeton University
Princeton, NJ 08544, USA
wonyongc@princeton.edu

Qibin Liu
SLAC National Accelerator Laboratory
Menlo Park, CA 94025, USA
qibin@slac.stanford.edu

Liangyu Wu
Stanford University
Stanford, CA 94305, USA
liangyu.wu@stanford.edu

Julia Gonski
SLAC National Accelerator Laboratory
Menlo Park, CA 94025, USA
jgonski@slac.stanford.edu

Abstract

We present the first implementation of AI agents into the design and optimization of detectors in high-energy physics experiments via a bi-level optimization framework that vertically integrates detector geometry, front-end digitization, and high-level reconstruction algorithm parameters in differentiable full simulations. Using the example of a dual-readout, segmented crystal EM calorimeter with a baseline resolution of $3\%/\sqrt{E}$, we investigate the capabilities and value propositions of AI agents in the identification and reduction of key detector parameters and in the nonlinear traversal of design space. We find that frontier LLM reasoning-models today, without being given additional experiment-specific context, are able to effectively execute complex workflows and proactively suggest generic but relevant avenues for further study or improvement. Here, we demonstrate an AI agent's ability to find an optimal design point amidst three competing performance criteria, showing that effective integration of agents into the complex workflows of frontier research areas can yield higher performance for key physics goals while reducing labor and compute. This study establishes the foundation for a future demonstration of the first fully AI-designed detector for future scientific facilities.

1 Introduction

High-energy physics (HEP) research today is organized around a set of key physics drivers, identified through global community strategic planning exercises such as Snowmass [1], P5 [2], and the European Strategy for Particle Physics [3]. These topics include precision measurements of the Higgs boson, broad discovery reach for beyond-the-Standard Model physics that can explain the nature of dark matter or the expansion of the universe, and the origin of neutrino mass. Successfully addressing each of these drivers requires a robust and broad landscape of future experimental infrastructure ranging from high-energy colliders to astrophysical observatories. Each of these types of experiments requires dedicated R&D in advanced detector technologies and designs.

Artificial intelligence and machine learning (AI/ML) have become primary workhorses for nearly every aspect of these research efforts [4, 5]. The next generation of HEP experiments will rely on AI/ML in fundamentally novel ways, integrated into the full experimental design cycle, ranging from the use of real-time AI-based readout for on-detector signal processing in radiation-hard front-end electronics [6, 7], to pattern-scoring and decision-making processes in rare-event triggers, to overall detector geometry design parameters. AI/ML also enables new methods of achieving optimal

Conference on Physics and AI at Stanford University (PAI 2026).

detector design, leveraging varying approaches such as differentiable detector elements [8, 9, 10], reinforcement learning [11], and synthetic data representations [12].

The introduction of AI agents that can autonomously reason about and execute workflows has dramatically accelerated progress in software development and closely adjacent fields but their introduction into HEP workflows is still in its infancy. Recent works have established the ability of AI agents across two wide swathes of the traditional training ground for PhD physicists in HEP: 1) established but unfinished theory calculations in a specific regime [13], and 2) standard analyses in collider physics [14, 15]. This work adds the example of phenomenological detector design. Using the example of a dual-readout, segmented crystal EM calorimeter for future colliders, we demonstrate the first application of an agentic-AI driven workflow based on a bilevel optimization framework designed to perform simultaneous optimization of detector geometry and reconstruction parameters for a given set of physics targets. We find that agents excel at complex execution workflows and significantly increase overall productivity in terms of labor and compute. Still, physics-informed human input is required to ensure relevance of the agent's proposed tasks to the experimental context.

2 Methods

This work debuts Vertically Integrated Bilevel geometry, Reconstruction, And TDAQ Optimizer (VIBRATO), a modular and extensible detector optimization framework [16] that iterates `DD4HEP` [17] full simulation runs with varying detector geometry parameters. It then evaluates simulation outputs against arbitrary downstream algorithm parameters such as front-end digitization or high-level reconstruction routines. Algorithm results are scored against a given physics- or statistics-objective and the final output is the loss landscape across the full geometry and algorithm parameter-space. Algorithms are designed to be plug-and-play with the full simulation outputs, with the only requirement that I/O be in ROOT [18] format. A menu of `scipy` [19] optimization methods are implemented for this study, with support for hyperparameter optimization for ML-methods in upcoming iterations. The plug-in interface is deliberately kept minimal to allow maximum flexibility. Figure 1 provides a diagram of the framework.

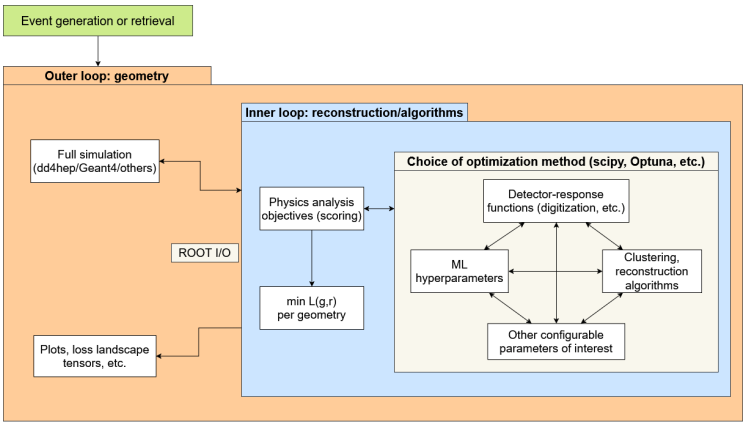

Figure 1: Bi-level detector optimization framework (VIBRATO), indicating geometry-based outer loop and reconstruction-based inner loop and their connection to user-defined optimization criteria.

Segmented dual-readout crystal electromagnetic (EM) calorimeters [20, 21, 22] are a promising technology for delivering excellent jet energy resolution at a future collider. Dual-readout calorimeters count the number of photons produced in the wake of a high-energy particle in both scintillation (S) and Cherenkov (C) channels, from which the EM fraction of hadronic showers, which suffers from a large intrinsic fluctuation due to asymptotic-freedom in hadronization, can be determined on an event-by-event basis. Calorimeters in colliders have historically only counted in the scintillation channel as a proxy for energy, and the addition of the Cherenkov channel, which counts the photons emitted in a distinctive wavefront as a particle travels faster than the speed of light in the medium, represents a genuinely new physical observable in detector hardware. This study uses the segmented crystal ECAL of the IDEA detector concept [23] and casts its design as a bilevel problem. The

geometry parameters chosen to be varied are the crystal angular pitch w_θ, setting the transverse granularity, a projective radial offset Δr, and the front-layer crystal lengths.

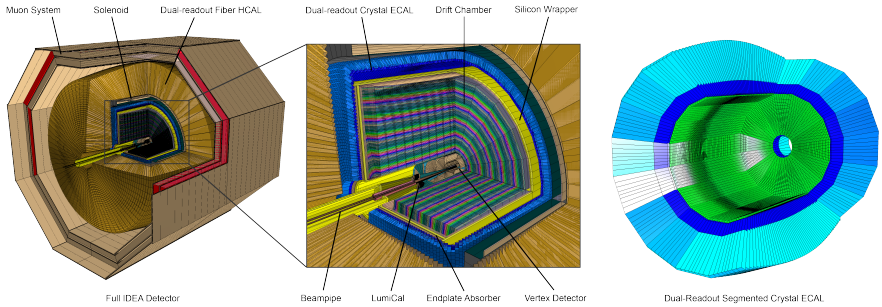

Figure 2: The IDEA detector concept with labeled subsystems (left, center). The Segmented Crystal EM Precision Calorimeter (SCEPCal) with exaggerated crystal sizes for visibility (right).

A `DD4HEP` full detector simulation running `Geant4` [24] under the hood is performed for each geometry iteration. This domain-specific tool models particle-material interactions and records realistic energy-deposition information. A digitization algorithm then converts simulated photon counts from ROOT files into realistic SiPM waveforms. For each readout channel, S/C photon counts are first extracted and aggregated with configurable granularity (per-crystal, per-layer, per-cluster, or per-event). Per-photon analog templates are then constructed analytically: the scintillation emission profile follows a two-component exponential decay model with a finite rise time, while the Cherenkov emission is modeled as a single fast exponential. Each emission profile is convolved via FFT with the SiPM single-photon response (SPR), which is parameterized as a double-exponential pulse shaped by a first-order RC high-pass filter representing AC coupling. The resulting per-photon templates are computed once and cached, enabling efficient reuse across all waveforms in a geometry sweep.

Waveform synthesis proceeds by scaling the cached S/C templates by the respective photon counts, shifting them by a sampled signal arrival time t_0, then superimposes Poisson-distributed SiPM dark counts (each convolved with the SPR), additive Gaussian electronic noise, and a configurable DC baseline offset. The continuous analog waveform is sampled at a configurable rate determined by the sampling interval and readout time window. An electronics chain then applies ADC quantization by mapping the sampled signal to discrete codes within a specified dynamic range and bit depth.

We apply a novel, science-focused agentic-AI framework, SciFi [25], to the task of calorimeter design optimization. This fully open-source workflow is designed to offload and streamline common scientific workloads, including data processing, visualization, analysis through agentic loops, and interface them with scientific computing environments. It ensures safe execution and allows choice of the LLM backbone. Combined with the bilevel optimization framework above, which provides the domain-specific toolchain and incorporates detector-domain knowledge, this work demonstrates the first agent-driven optimization workflow for a realistic experimental detector design. *Claude Code Opus 4.6* is used as the reasoning LLM model together with human input to draft a high-level initial plan, and the SciFi agentic framework is the execution system.

The bilevel optimization is performed in a high-dimensional parameter space, where grid search is highly computationally intensive and complex correlations are present. The detector geometry includes three tunable parameters: offset, crystal length, and crystal width. The pre-digitization analysis, digitization, and post-digitization analysis then introduce a total of nine algorithmic parameters, including the reconstruction window size (R-size), the energy threshold for pre-digitization reconstruction, the digitization sampling rate, the ADC bit depth, and waveform-fitting parameters that control the "C", "S" and "t0" fit bound and χ^2 threshold. Since these are evaluated only after each point in the simulation loop is completed, they are defined as outer-loop parameters.

In total, the workflow involves 11 parameters, with 5 detector-design parameters that must be fixed before construction, while the reconstruction and fitting parameters remain adjustable during runtime. The optimization target is multi-objective and includes the energy signal-noise ratio (SNR) from the pre-digitization analysis, the detector-level C/S fitting error from waveform fitting, and the ADC cost, defined as the total power consumption summed over all channels and determined jointly by the detector geometry and the digitization configuration.

3 Results

The final optimized result is shown in Figure 3 and is presented together with an analysis of the agent-driven optimization trajectory. Early probing by the agent revealed an important piece of domain knowledge: the offset parameter has little effect on the final metrics and can be treated as a nuisance parameter. Subsequent iterations therefore focused on tuning crystal width and length, substantially reducing optimization effort. Over the next several rounds, labeled A through E, the autonomous agentic loop progressively narrowed down the optimal region in this two-dimensional parameter plane, while the remaining inner-loop parameters were scanned on the fly. The scan ranges, step sizes, and event statistics were all proposed adaptively by the AI agent.

Furthermore, the agent learned that the metrics associated with the pre-digitization analysis and the digitization stage are effectively decoupled. In particular, SNR optimization is found to be independent of the other six outer-loop parameters. Digitization and waveform-fitting can be studied afterward, decomposing the full 11D parameter space into smaller, more tractable subspaces. Additionally, the four fit parameters act as nuisance parameters: only unreasonable choices degrade performance, while any reasonable configuration yields a valid fit and a good S_{err}.

Through this iterative reasoning and trial-and-feedback process, the original 11-dimensional optimization problem was effectively reduced by the agent into two simpler stages: first, a two-dimensional scan in crystal length and crystal width with a single objective, together with three nuisance parameters (E threshold, R-size, and offset); and second, a two-dimensional scan in ADC bit depth and sampling rate with two objectives (S_{photon} error and ADC cost), together with four nuisance fitting parameters. Once the optimal SNR point is fixed, the optimal ADC bit depth and sampling rate can be determined to achieve the best S_{photon} error, while the correlation with the third metric, ADC cost, is also clearly exposed. The final SciFi-determined optimal design assigns values of front crystal length $= 4\mathrm{cm}$ and crystal width $= 1.9\mathrm{cm}$, which gives a best SNR of 0.616.

As the ability to learn key parameter correlations expedites the SciFi optimization process, a better final result is achievable with lower computation cost. Compared to a conventional grid search, achieving the same optimal point with a 0.1 cm resolution in the length–width parameter plane and including offset requires 2255 configurations (41 × 11 × 5) for a full 3D scan. In contrast, the progressive optimization driven by the agentic loop converges within 5 rounds and 46 total runs. Given that at each configuration a relatively large amount of computation is required, the dimensionality of the problem presented here already represents a complexity that is nearly intractable for grid search, solidifying this use case as representative of the new frontier of design afforded by agentic workflow adoption. Nonetheless, expert physicist input is required to ensure the agent remains in technically and practically realizable design space.

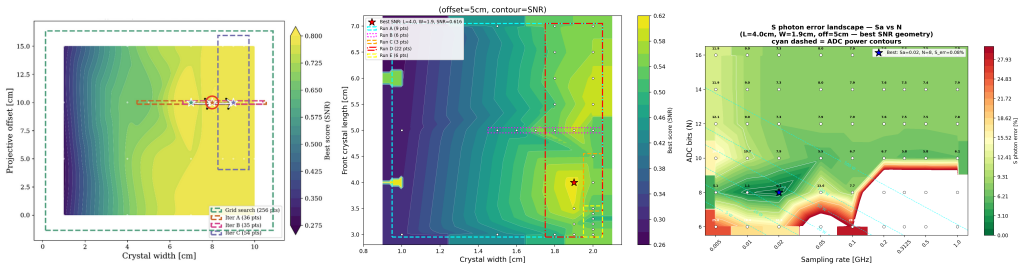

Figure 3: Optimization for the detector geometry as driven by SciFi agentic application: crystal offset and crystal width, using SNR as the primary metric. (left) offset is identified as not relevant to SNR metric and a value of $5\,\mathrm{cm}$ is taken. (center) Scan points and step sizes proposed adaptively by the agentic tool. The two artifacts on the left arise from fluctuations in first round probe points and are not relevant to the parameter region of interest. (right) Optimization results for the digitization parameters using S_{photon} error as the primary metric. ADC cost is indicated by the dashed line.

Acknowledgments and Disclosure of Funding

This work is supported by the U.S. Department of Energy under contract numbers DE-AC02-76SF00515, DE-SC0007968, DE-SC0025956, and used resources of the National Energy Research Scientific Computing Center (NERSC), a Department of Energy User Facility using NERSC award HEP-ERCAP 0037461.

References

[1] Joel N. Butler, R. Sekhar Chivukula, André de Gouvêa, Tao Han, Young-Kee Kim, Priscilla Cushman, Glennys R. Farrar, Yury G. Kolomensky, Sergei Nagaitsev, Nicolás Yunes, Stephen Gourlay, Tor Raubenheimer, Vladimir Shiltsev, Kétévi A. Assamagan, Breese Quinn, V. Daniel Elvira, Steven Gottlieb, Benjamin Nachman, Aaron S. Chou, Marcelle Soares-Santos, Tim M. P. Tait, Meenakshi Narain, Laura Reina, Alessandro Tricoli, Phillip S. Barbeau, Petra Merkel, Jinlong Zhang, Patrick Huber, Kate Scholberg, Elizabeth Worcester, Marina Artuso, Robert H. Bernstein, Alexey A. Petrov, Nathaniel Craig, Csaba Csáki, Aida X. El-Khadra, Laura Baudis, Jeter Hall, Kevin T. Lesko, John L. Orrell, Julia Gonski, Fernanda Psihas, and Sara M. Simon. Report of the 2021 U.S. Community Study on the Future of Particle Physics (Snowmass 2021) Summary Chapter, 2023.

[2] Hitoshi Murayama, Shoji Asai, Karsten Heeger, Amalia Ballarino, Tulika Bose, Kyle Cranmer, Francis-Yan Cyr-Racine, Sarah Demers, Cameron Geddes, Yuri Gershtein, Beate Heinemann, JoAnne Hewett, Patrick Huber, Kendall Mahn, Rachel Mandelbaum, Jelena Maricic, Petra Merkel, Christopher Monahan, Peter Onyisi, Mark Palmer, Tor Raubenheimer, Mayly Sanchez, Richard Schnee, Sally Seidel, Seon-Hee Seo, Jesse Thaler, Christos Touramanis, Abigail Vieregg, Amanda Weinstein, Lindley Winslow, Tien-Tien Yu, and Robert Zwaska. *Exploring the Quantum Universe: Pathways to Innovation and Discovery in Particle Physics*. 2023.

[3] The European Strategy for Particle Physics: 2026 Update - Recommendations by the European Strategy Group. Technical report, Monte Verità/Ascona, Switzerland, 2025.

[4] Kim Albertsson et al. Machine learning in high energy physics community white paper, 2019.

[5] Thea Klaeboe Aarrestad et al. Building an ai-native research ecosystem for experimental particle physics: A community vision, 2026.

[6] Jieun Yoo, Jennet Dickinson, Morris Swartz, Giuseppe Di Guglielmo, Alice Bean, Douglas Berry, Manuel Blanco Valentin, Karri DiPetrillo, Farah Fahim, Lindsey Gray, James Hirschauer, Shruti R Kulkarni, Ron Lipton, Petar Maksimovic, Corrinne Mills, Mark S Neubauer, Benjamin Parpillon, Gauri Pradhan, Chinar Syal, Nhan Tran, Dahai Wen, and Aaron Young. Smart pixel sensors: towards on-sensor filtering of pixel clusters with deep learning. *Machine Learning: Science and Technology*, 5(3):035047, aug 2024.

[7] D. Yilmaz, L. Wu, J. Gonski, D. Rankin, and C. Herwig. Edge machine learning for cluster counting in next-generation drift chambers. In *Proceedings of the Machine Learning for the Physical Sciences Workshop at NeurIPS 2025*, 2025. arXiv:2511.10540.

[8] Atılım Güneş Baydin, Kyle Cranmer, Pablo de Castro Manzano, Christophe Delaere, Denis Derkach, Julien Donini, Tommaso Dorigo, Andrea Giammanco, Jan Kieseler, Lukas Layer, Gilles Louppe, Fedor Ratnikov, Giles C. Strong, Mia Tosi, Andrey Ustyuzhanin, Pietro Vischia, and Hevjin Yarar. Toward machine learning optimization of experimental design. *Nuclear Physics News*, 31(1):25–28, 2021.

[9] MODE Collaboration. Mode: Machine-learning optimized design of experiments. `https://mode-collaboration.github.io/`, 2026.

[10] Max Aehle, Lorenzo Arsini, R. Belén Barreiro, Anastasios Belias, Alexey Boldyrev, Florian Bury, Susana Cebrian, Alexander Demin, Jennet Dickinson, Julien Donini, Tommaso Dorigo, Michele Doro, Nicolas R. Gauger, Andrea Giammanco, Lindsey Gray, Borja S. González, Verena Kain, Jan Kieseler, Lisa Kusch, Marcus Liwicki, Gernot Maier, Federico Nardi, Fedor Ratnikov, Ryan Roussel, Roberto Ruiz de Austri, Fredrik Sandin, Michael Schenk, Bruno

Scarpa, Pedro Silva, Giles C. Strong, and Pietro Vischia. Progress in end-to-end optimization of fundamental physics experimental apparata with differentiable programming. *Reviews in Physics*, 13:100120, December 2025.

[11] Shah Rukh Qasim, Patrick Owen, and Nicola Serra. Physics instrument design with reinforcement learning, 2024.

[12] Wonyong Chung. Synthetic training and representation bridging in reconstruction domains, 2025.

[13] Matthew D. Schwartz. Resummation of the c-parameter sudakov shoulder using effective field theory, 2026.

[14] Eric A. Moreno, Samuel Bright-Thonney, Andrzej Novak, Dolores Garcia, and Philip Harris. AI Agents Can Already Autonomously Perform Experimental High Energy Physics. 3 2026.

[15] Joschka Birk, Gregor Kasieczka, Siddharth Mishra-Sharma, Benjamin Nachman, Dennis Noll, and Tanvi Wamorkar. A scientific human-agent reproduction pipeline, 2026.

[16] Wonyong Chung. bilevel_det_opt: Bilevel optimization of detector geometry and reconstruction algorithm parameters. `https://github.com/wonyongc/bilevel_det_opt`, 2026. Accessed: 2026-04-16.

[17] M Frank, F Gaede, C Grefe, and P Mato. DD4hep: A Detector Description Toolkit for High Energy Physics Experiments. *Journal of Physics: Conference Series*, 513(2):022010, jun 2014.

[18] Rene Brun and Fons Rademakers. ROOT - an object oriented data analysis framework. In *Proceedings of the AIHENP'96 Workshop*, Lausanne, September 1996. Published in Nucl. Instrum. Meth. A 389 (1997) 81–86.

[19] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020.

[20] Sehwook Lee, Michele Livan, and Richard Wigmans. Dual-readout calorimetry. *Rev. Mod. Phys.*, 90:025002, Apr 2018.

[21] R. Hirosky, T. Anderson, G. Cummings, M. Dubnowski, C. Guinto-Brody, Y. Guo, A. Ledovskoy, D. Levin, C. Madrid, C. Martin, and J. Zhu. Dual-readout calorimetry with homogeneous crystals. In *Proceedings of CALOR2024*, EPJ Web of Conferences, 2024. arXiv:2408.11973.

[22] S. Eno, L. Wu, M. Y. Aamir, S. V. Chekanov, S. Nabili, and C. Palmer. On the resolution of dual readout calorimeters. *Nucl. Instrum. Meth. A*, 1083:171080, 2026.

[23] Marco T. Lucchini, Wonyong Chung, Sarah C. Eno, Yihui Lai, Lorenzo Lucchini, Minh-Thi Nguyen, and Christopher G. Tully. New perspectives on segmented crystal calorimeters for future colliders. *JINST*, 15(11):P11005, 2020.

[24] S. Agostinelli, J. Allison, K. Amako, J. Apostolakis, H. Araujo, P. Arce, M. Asai, D. Axen, S. Banerjee, G. Barrand, F. Behner, L. Bellagamba, J. Boudreau, L. Broglia, A. Brunengo, H. Burkhardt, S. Chauvie, J. Chuma, R. Chytracek, G. Cooperman, G. Cosmo, P. Degtyarenko, A. Dell'Acqua, G. Depaola, D. Dietrich, R. Enami, A. Feliciello, C. Ferguson, H. Fesefeldt, G. Folger, F. Foppiano, A. Forti, S. Garelli, S. Giani, R. Giannitrapani, D. Gibin, J.J. Gómez Cadenas, I. González, G. Gracia Abril, G. Greeniaus, W. Greiner, V. Grichine, A. Grossheim, S. Guatelli, P. Gumplinger, R. Hamatsu, K. Hashimoto, H. Hasui, A. Heikkinen, A. Howard, V. Ivanchenko, A. Johnson, F.W. Jones, J. Kallenbach, N. Kanaya, M. Kawabata, Y. Kawabata, M. Kawaguti, S. Kelner, P. Kent, A. Kimura, T. Kodama, R. Kokoulin, M. Kossov, H. Kurashige,

E. Lamanna, T. Lampén, V. Lara, V. Lefebure, F. Lei, M. Liendl, W. Lockman, F. Longo, S. Magni, M. Maire, E. Medernach, K. Minamimoto, P. Mora de Freitas, Y. Morita, K. Murakami, M. Nagamatu, R. Nartallo, P. Nieminen, T. Nishimura, K. Ohtsubo, M. Okamura, S. O'Neale, Y. Oohata, K. Paech, J. Perl, A. Pfeiffer, M.G. Pia, F. Ranjard, A. Rybin, S. Sadilov, E. Di Salvo, G. Santin, T. Sasaki, N. Savvas, Y. Sawada, S. Scherer, S. Sei, V. Sirotenko, D. Smith, N. Starkov, H. Stoecker, J. Sulkimo, M. Takahata, S. Tanaka, E. Tcherniaev, E. Safai Tehrani, M. Tropeano, P. Truscott, H. Uno, L. Urban, P. Urban, M. Verderi, A. Walkden, W. Wander, H. Weber, J.P. Wellisch, T. Wenaus, D.C. Williams, D. Wright, T. Yamada, H. Yoshida, and D. Zschiesche. Geant4—a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250–303, 2003.

[25] Qibin Liu and Julia Gonski. Scifi: A safe, lightweight, user-friendly, and fully autonomous agentic ai workflow for scientific applications, 2026.