Learning-at-Criticality in Large Language Models for Quantum Field Theory and Beyond

Xiansheng Cai(蔡贤盛)[1†], Sihan Hu(胡思寒)[2†], Tao Wang(王韬)[3,4*], Yuan Huang(黄远)[5*], Pan Zhang(张潘)[1,6,7*], Youjin Deng(邓友金)[2,8*], and Kun Chen(陈锟)[1*]

[1]*Institute of Theoretical Physics, Chinese Academy of Sciences, Beijing 100190, China*
[2]*Hefei National Laboratory, University of Science and Technology of China, Hefei 230088, China*
[3]*Department of Physics, University of Massachusetts, Amherst, MA 01003, USA*
[4]*Institute of Physics, Chinese Academy of Sciences, Beijing 100190, China*
[5]*DP Technology, Beijing 100080, China*
[6]*School of Fundamental Physics and Mathematical Sciences, Hangzhou Institute for Advanced Study, University of Chinese Academy of Sciences, Hangzhou 310024, China*
[7]*Hefei National Laboratory, Hefei 230088, China*
[8]*Hefei National Laboratory for Physical Sciences at the Microscale and Department of Modern Physics, University of Science and Technology of China, Hefei 230026, China*

(Received 16 August 2025; accepted manuscript online 30 October 2025)

Fundamental physics often confronts complex symbolic problems with few guiding exemplars or established principles. While artificial intelligence (AI) offers promise, its typical need for vast datasets to learn from hinders its use in these information-scarce frontiers. We introduce learning at criticality (LaC), a reinforcement learning scheme that tunes large language models (LLMs) to a sharp learning transition, addressing this information scarcity. At this transition, LLMs achieve peak generalization from minimal data, exemplified by 7-digit base-7 addition—a test of nontrivial arithmetic reasoning. To elucidate this peak, we analyze a minimal concept-network model designed to capture the essence of how LLMs might link tokens. Trained on a single exemplar, this model also undergoes a sharp learning transition. This transition exhibits hallmarks of a second-order phase transition, notably power-law distributed solution path lengths. At this critical point, the system maximizes a "critical thinking pattern" crucial for generalization, enabled by the underlying scale-free exploration. This suggests LLMs reach peak performance by operating at criticality, where such explorative dynamics enable the extraction of underlying operational rules. We demonstrate LaC in quantum field theory: an 8B-parameter LLM, tuned to its critical point by LaC using a few exemplars of symbolic Matsubara sums, solves unseen, higher-order problems, significantly outperforming far larger models. LaC thus leverages critical phenomena, a physical principle, to empower AI for complex, data-sparse challenges in fundamental physics.

DOI: 10.1088/0256-307X/42/12/120002 CSTR: 32039.14.0256-307X.42.12.120002

1. Introduction. Artificial intelligence (AI) has accelerated scientific discovery, yet its primary successes are in data-rich domains leveraging pattern recognition, a capability closely aligned with intuitive "System 1" thinking.[1–4] A distinct class of frontier scientific problems, particularly in theoretical physics, presents a different challenge: they often necessitate deriving complex analytical solutions through extended abstract, "System 2" reasoning, yet the very nature of these frontiers means training data for AI models is inherently limited,[5–8] seemingly placing them beyond the reach of conventional AI reliant on vast statistical correlations.

This chasm between current AI strengths and the needs of theoretical physics is starkly evident. In quantum electrodynamics, the electron's anomalous magnetic moment ($a_{\rm e}$) acts as a critical test of the standard model. While numerical evaluations of $a_{\rm e}$ coefficients are known to high precision (e.g., up to the fifth loop), their complete analytical derivation—essential for deep theoretical insight—is achieved only to the third loop after decades of effort.[9–13] Similarly, in many-electron systems in condensed matter, understanding phenomena like high-temperature superconductivity relies on analytically mastering the Fermi-surface complexities in Feynman diagrams. Here too, numerical methods like diagrammatic Monte Carlo (DiagMC) provide crucial estimates for higher-order terms,[14–17] but the analytical solution of even low-order diagrams remains a challenging task. For fields demanding generalizable symbolic reasoning from a few solved instances, AI models traditionally associated with "System 1" appear ill-equipped.

†These authors contributed equally to this work.
*Corresponding authors. Email: taowang@iphy.ac.cn; huangyuan@dp.tech; panzhang@itp.ac.cn; yjdeng@ustc.edu.cn; chenkun@itp.ac.cn

However, recent advancements offer a new perspective. Large language models (LLMs), particularly when augmented with reinforcement learning (RL), are beginning to exhibit capabilities that transcend simple pattern matching.[18–25] RL enables LLMs to actively explore problem spaces, learn from feedback on their generated reasoning pathways, and refine strategies—processes that foster more coherent, goal-oriented, and multi-step thought, intriguingly reminiscent of "System 2" cognition. This opens a promising, albeit challenging, avenue for AI to assist in domains demanding true analytical depth.

While LLMs augmented by RL show promise for "System 2"-like reasoning, deploying them effectively in frontier science requires navigating inherent challenges. Firstly, the highly specialized nature of these problems means they are often statistical outliers to an LLM's general pre-training.[5–8] This misalignment can lead to unreliable outputs or "hallucinations"[26–28] when precise symbolic manipulation is critical, diminishing the utility of off-the-shelf models. A natural corrective is targeted RL fine-tuning, tailoring the LLM to the specific nuances of the cutting-edge problem. Yet, this essential fine-tuning step itself encounters a profound obstacle: the very frontier nature that necessitates such specialization also implies an extreme scarcity of existing data suitable for this RL process. This creates a central dilemma: if the indispensable RL fine-tuning must operate with exceptionally few examples, can these advanced AI models genuinely acquire the algorithmic understanding and robust generalization capabilities characteristic of "System 2" reasoning?[29,30] Addressing this question is crucial for determining AI's true potential in advancing fundamental science.

In this Letter, we resolve this central dilemma by introducing "learning at criticality" (LaC), a learning scheme inspired by critical phenomena in physics. LaC precisely guides an LLM[23] via RL to a narrow 'critical' training phase. At this threshold, analogous to a phase transition, the LLM attains strong problem-solving capabilities and optimal generalization. This targeted training to criticality imparts to the LLM genuine, generalizable algorithmic understanding from minimal data—even a single exemplar—circumventing the data-dependency that has historically limited conventional AI in numerous scientific domains. Consequently, LaC enables AI to address complex, abstract theoretical problems in fundamental science previously considered intractable due to data scarcity.

We first establish LaC by training an LLM on a single instance of 7-digit, base-7 addition, observing a peak in its generalization capability at a precise training stage (Fig. 1, left panel). To understand the physics underlying this peak, we propose the concept-network model (CoNet). This minimal model abstracts LLM reasoning, where cohesive sequences of autoregressively generated tokens form "concepts". The LLM's problem-solving is viewed as a stochastic traversal within an implicit network of these concepts, with RL training [e.g., a variation of group relative policy optimization (GRPO)[19–22]] adjusting inter-concept transition probabilities. Our CoNet embodies this as a Markovian walk on a random graph of such "concepts", where transitions are learned via reinforcement to find paths from question to answer. Remarkably, this simplified CoNet reproduces a sharp learning transition exhibiting hallmarks of a second-order phase transition: problem-solving accuracy (Fig. 2, upper panel) increases sigmoidally, while reasoning path length variance diverges, signaling critical fluctuations. Crucially, near this critical point, path lengths become power-law distributed ($L^{-\gamma}$, Fig. 3), indicating an emergent "critical thinking pattern" of diverse strategic exploration.

Finally, we apply LaC to the symbolic Matsubara frequency summation in Feynman diagrams—a challenging problem in finite-temperature many-body quantum field theory (QFT).[31–34] Remarkably, fine-tuning an 8-billion parameter LLM[24] to its critical point using only low-order diagrams enables it to learn the symbolic procedure and solve unseen, more complex diagrams (Fig. 1, right panel), outperforming models with nearly two orders of magnitude more parameters (Table 1). Our work establishes LaC as a data-efficient strategy for AI-driven discovery in theoretical physics and suggests that emergent reasoning in AI can be understood as a critical phenomenon.

2. Learning at Criticality: A Phenomenon. Before presenting a theoretical model, we empirically demonstrate that effective learning from sparse data, while achievable, critically depends on the training regime. We tasked a Qwen2.5-7B model,[23] initially unable to perform 7-digit base-7 addition, with learning this procedure from a *single* problem instance. This specific task was chosen to probe algorithmic reasoning. Its multi-step, rule-based nature (requiring sequential digit-by-digit processing and explicit base-7 carrying rules) serves as a strong proxy for deliberative "System 2" cognition. Furthermore, its presumed rarity in pre-training corpora, particularly when contrasted with arithmetic in common bases like 2 or 10, significantly minimizes the likelihood of the model recalling a memorized solution rather than learning the underlying algorithm.

We trained the model on this one example using direct-advantage policy optimization (DAPO),[22] which is a variant of the GRPO algorithm.[19,20] It operates by having the LLM generate multiple potential reasoning paths for a given problem. These paths are evaluated, and their scores, relative to the group average, guide policy updates to the LLM's parameters. This process strengthens transitions on paths yielding above-average rewards and weakens those on paths performing below average, thereby preferentially guiding the policy towards effective reasoning strategies.

As shown in Fig. 1 (left panel), the training accuracy on the single problem instance (blue circles) displays a sharp, sigmoidal transition, indicating a sudden acquisition of the solution. This transition is not perfectly singular, likely due to multiple valid reasoning paths (e.g., different ways to handle carrying operations), but the qualitative feature is clear. More importantly, when we tested the model's performance on 128 new, unseen 7-digit base-7 addition problems, its generalization accuracy (orange triangles) peaked precisely in the vicinity of this learning transition.

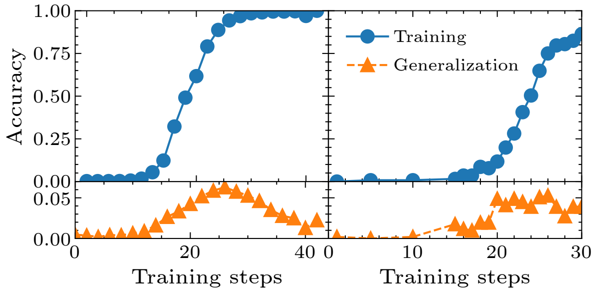


Fig. 1. Critical learning from a single training example. (left) Training a Qwen2.5-7B model on one 7-digit base-7 addition. Training accuracy (blue circles) shows a sharp transition. Generalization to unseen additions (orange triangles) peaks precisely at this critical point before overfitting. (right) Similar phenomenon for a Qwen3-8B model trained on Matsubara frequency summation (2-loop sunrise self-energy diagram). Generalization to other unseen 2-loop diagrams is maximized at the critical learning point.

Although this initial peak generalization rate was approximately 7%—a modest absolute value, yet significant given it stems from a single training example—we further validated that by carefully continuing the training within this critical regime, the model progressively enhanced its capability to solve unseen instances [see the Supplementary Materials (SM) for details] correctly. As training continued past this point (overfitting), the model's general ability declined, even as its performance on the single training example remained perfect. This finding is the cornerstone of our LaC proposal: it demonstrates the remarkable capacity of an LLM to acquire genuinely generalizable, algorithmic understanding from even a single exemplar, provided the training navigates and sustains it within a transient, optimal learning state.

3. CoNet Model for Learning Transitions. LLMs generate text autoregressively, predicting subsequent tokens based on prior context. When an LLM predicts a sequence of tokens where each next token has nearly 100% certainty, we consider these tokens to form a cohesive unit, which we abstract as a "concept". LLM reasoning can then be viewed as a stochastic traversal—akin to a random walk—within an underlying network of these concepts. GRPO training step effectively acts as an external "tuning parameter" that modifies the transition probabilities within the LLM's implicit concept network, optimizing pathways from the "question concepts" to the "answer concepts".

To model this process and understand the physics of LaC, we propose the CoNet. In this minimal model, the LLM's abstract concept space is represented as a K-regular random graph with N nodes (concepts). A reasoning task is modeled as finding a path from a source node Q (question) to a target node A (answer), constrained to a maximum path length of $L_{\max} = 200$. The LLM's token generation is simplified to a Markovian walk on this graph, where the transition probability from concept i to a neighboring concept j is

$$\pi_\theta(j|i) = \frac{\theta_{ij}}{\sum_{k\in\text{neighbors}(i)} \theta_{ik}}, \tag{1}$$

where $\theta_{ij} \in [0, 1)$ are learnable parameters representing transition strengths. For a given Q–A pair, $M = 10^4$ reasoning paths (indexed by m) are sampled. Each path receives a reward, and its advantage A_m (relative to the average reward) guides the update of θ_{ij} via a GRPO-variant rule,[19,21,22] $\Delta\theta_{ij} \propto \sum_m A_m \nabla_{\theta_{ij}} \log \pi_\theta(j|i)$. This reinforces transitions on above-average paths. It is important to clarify that CoNet is a minimalist abstraction of the emergent reasoning process, not the LLM's architecture. The model's concepts (nodes) correspond to stable, low-entropy token sequences, while its learnable transitions (links) represent the probabilistic choices made at high-entropy decision points during generation.

Simulations of CoNet ($N = 8000$ nodes, $K = 5$) reveal a sharp learning transition (Fig. 2), a behavior that proves robust across various model parameters as shown in the SM. The accuracy (fraction of successful paths, serving as an order parameter) displays a steep, sigmoidal increase. Concurrently, the variance of the reasoning path lengths exhibits a pronounced peak at the transition, analogous to diverging susceptibility at a critical point in physical systems. These features are hallmarks of a continuous phase transition and mirror the empirical learning peak.

The microscopic origin of this transition is revealed in the hybrid nature of the reasoning path distribution $P(L)$ at criticality. As shown in Fig. 3 (around step 30), the distribution is effectively a superposition of two distinct search strategies. A peak at short lengths signifies local exploration around an emergent optimal path, a behavior that consolidates into a purely exponential decay [$P(L) \sim \mathrm{e}^{-\alpha L}$] post-transition (step 36). Coexisting with this is a pronounced power-law tail, $P(L) \sim L^{-\gamma}$ ($\gamma \approx 0.16$), a canonical signature of scale-free, critical phenomena (see the SM for similar results with a larger maximum response length). The prominence of this exploratory, power-law mode is maximized at the transition,

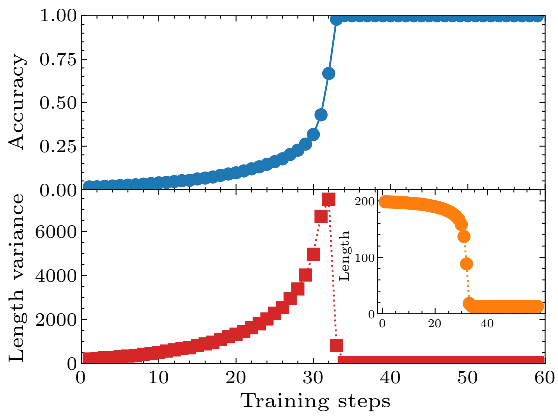


Fig. 2. Training dynamics of the minimal CoNet. The figure shows the accuracy, average response length, and the response length's variance of the minimal model on the training problem, plotted against training steps. The accuracy (blue) increases and the average response length (orange) decreases in a sigmoidal manner. Concurrently, the response length's variance (red) exhibits a lambda-shape discontinuity at the learning transition, mirroring the behavior of specific heat at the lambda point marking the normal-to-superfluid helium phase transition.

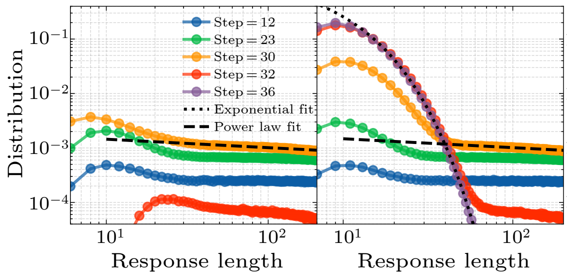
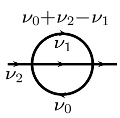


Fig. 3. Distinct reasoning dynamics: critical power-law search transitions to post-convergence exponential exploration. Toy model reasoning response length distributions $P(L)$ across training epochs. During the critical learning transition (e.g., step 30, orange), long exploratory responses exhibit characteristic power-law decay $P(L) \sim L^{-\gamma}$ with $\gamma \approx 0.16$ (dashed fit, left panel; also evident for early-stage odd paths, right panel), a signature of scale-invariant critical search. Post-transition, as the policy converges (e.g., to a 7-step optimal odd path, step 34, dark grey), local perturbations around this path display exponential decay $P(L) \sim \mathrm{e}^{-\alpha L}$ (dotted fit, right panel). These distinct scaling regimes characterize the evolution from broad, critical exploration to refined exploitation, crucial for learning. Distributions are truncated at the maximum allowed response length.

directly causing the large path-length variance (susceptibility). This "critical thinking pattern" represents a diverse, long-range exploratory search, and its coexistence with efficient exploitation is the hallmark of the learning transition, enabling the discovery of generalizable strategies before the system collapses into a single, non-exploratory state.

4. AI for QFT. To demonstrate LaC's potential for addressing data-scarce,[5–8] symbolically complex problems in theoretical physics, we apply it to a challenging representative task: the symbolic evaluation of Matsubara frequency sums.[33,34] This multi-step, algorithmically rich procedure is foundational to finite-temperature QFT[31,32] and methods like DiagMC for many-electron problems.[14–17,35–41] An example is the 2-loop sunrise self-energy diagram Eq. (2), whose symbolic solution involves contour integration and complex analysis.



$$= \sum_{\nu_0,\nu_1} \frac{1}{\nu_0+\varepsilon_0}\frac{1}{\nu_1+\varepsilon_1}\frac{1}{\nu_0+\nu_2-\nu_1+\varepsilon_2}$$
$$= \oint_C \frac{\mathrm{d}z'}{2\pi \mathrm{i}}\frac{f(z')}{z'+\varepsilon_1}\oint_C \frac{\mathrm{d}z}{2\pi \mathrm{i}}\frac{f(z)}{z+\varepsilon_0}\frac{1}{z+\nu_2-z'+\varepsilon_2}, \tag{2}$$

where ν are fermionic Matsubara frequencies, ε are energy parameters, and $f(z)$ is the Fermi–Dirac distribution. This task is an ideal LaC testbed due to its structured complexity, availability of low-order diagrams for sparse training, and verifiable analytical solutions, allowing us to probe for emergent algorithmic understanding.

The well-defined, symbolic nature of this procedure makes it an ideal testbed for LaC. Specifically: (i) low-order diagrams offer sparse training instances; (ii) complexity scales with diagram order, naturally testing generalization; and (iii) exact analytical solutions permit unambiguous verification. Our objective is to investigate if an LLM, trained via staged LaC on minimal examples, can acquire the algorithmic reasoning inherent in this task, highlighting LaC's potential for data-efficient learning of complex physics procedures.

Our training employed a staged LaC strategy with a Qwen3-8B model[24] using GRPO. Phase 1: LaC training on only tree-level and 1-loop Matsubara sums. This not only enabled solutions for these simpler diagrams but also induced initial generalization to higher-order diagrams decomposable into these learned components. Notably, while the base model had zero success on the challenging sunrise self-energy diagram Eq. (2)—a 2-loop graph with nested frequencies—this first LaC phase elevated its performance to a non-zero baseline. Phase 2: This crucial improvement enabled a second LaC phase focused specifically on the sunrise diagram. As shown in Fig. 1 (right panel), the model again exhibited a sharp learning transition for this complex task. Critically, at this new critical point, the ability gained from mastering the sunrise diagram is generalized to other 2-loop topologies (e.g., polarization, vertex functions), where performance was previously negligible, confirming the LaC hypothesis. Phase 3: Finally, starting from this enhanced critical state, we incorporated all 2-loop diagrams into the LaC training until a critical point was reached for this comprehensive set.

The results are shown in Table 1. The base Qwen3-8B model's success is modest for 1-loops (45.0%) and 2-loops (18.4%), and minimal for 3-loops (0.9%) and 4-loops (0.2%). The latter non-zero values reflect occasional solutions to simple, decomposable higher-order diagrams, not general complex reasoning. In contrast, the fully fine-tuned model (Qwen3-8B-2-loop, post-Phase 3) greatly improved accuracy on its 2-loop training set. Crucially, it then achieved strong accuracy on complex 3-loop problems, despite no explicit training on them, and generalized to 4-loop cases. This LaC-trained model far surpasses its base and larger models (DeepSeek-R1, Qwen3-32B), underscoring LaC's efficiency. This success demonstrates that navigating critical points enabled the LLM to develop an emergent, algorithmic grasp of Matsubara summation.

Table 1. Critical learning transition enables generalization for Matsubara sums. Qwen3-8B, fine-tuned on problems up to 2-loop complexity, exhibits a sharp increase in success rate. This critical learning phase facilitates robust generalization to unseen 3-loop overlapped sums. Consequently, this fine-tuned 8B model markedly outperforms significantly larger base models (Qwen3-32B, DeepSeek-R1 671B) on this task, indicating data-efficient acquisition of algorithmic reasoning.

Model	Accuracy (%) Problem			
	1-loop	2-loop	3-loop	4-loop
DeepSeek-R1-0120(671B)	12.5	10.0	1.1	0.4
Qwen3-32B	47.8	25.5	1.6	0.6
Qwen3-8B	45.0	18.4	0.9	0.2
Qwen3-8B-1-loop	90.2	25.5	3.3	0.8
Qwen3-8B-2-loop	**97.5**	**56.9**	**9.5**	**1.7**

5. Conclusion. We addressed the challenge of "System 2" AI reasoning in data-scarce science by introducing "learning at criticality" (LaC), where LLM fine-tuning via RL induces a critical phase transition towards generalizable algorithmic understanding from minimal data.

This LaC principle was first shown with 7-digit base-7 addition: generalization from a single example peaked at the sharp learning transition. Our minimal physical model *CoNet* indicates that it is a continuous phase transition, revealing power-law scaling in reasoning paths—a signature of emergent "critical thinking". Critically, applying LaC to symbolic Matsubara frequency summation, an 8-billion parameter LLM, trained on a few low-order diagrams, learn the procedure and solves unseen, more complex diagrams, substantially outperforming models nearly two orders of magnitude larger. Important open questions remain regarding how the LaC phenomenon scales with LLM size. While our CoNet model provides a tractable framework for exploring network parameters, the equivalent scaling experiments on LLMs are computationally expensive. We therefore leave this investigation for future work.

Critically, this LaC-induced phase transition occurs in an implicit concept network linked to "System 2" reasoning, differentiating it from grokking phenomena which occur during pre-training or supervised fine-tuning and are associated with "System 1" pattern matching.[42,43] LaC thus provides a novel mechanism for cultivating advanced AI capabilities.

This work provides both a physical framework for understanding emergent AI reasoning and a practical, data-efficient strategy for applying it to data-scarce domains like theoretical physics. The implications of LaC extend deeply into domains like QFT, where navigating immense symbolic complexity with limited examples of optimal strategies is paramount.[44,45] Beyond theoretical physics, LaC provides a pathway towards AI systems as powerful scientific collaborators, equipped to tackle complex symbolic manipulations and accelerate discovery across diverse frontiers of knowledge where data is sparse but the demand for deep reasoning is high.

Acknowledgements. The authors thank Haijun Zhou, Lei Wang, Linfeng Zhang, and Pengcheng Hou for inspiring discussions. This work was supported by the National Key Research and Development Program of China (Grant No. 2024YFA1408604 for K. C. and X. C.), the National Natural Science Foundation of China (Grant Nos. 12047503, 12447103 for K. C. and X. C., 12325501 for P. Z., and 12275263 for Y. D. and S. H.), the Innovation Program for Quantum Science and Technology (Grant No. 2021ZD0301900 for Y. D. and S. H.), and the Natural Science Foundation of Fujian Province of China (Grant No. 2023J02032 for Y. D. and S. H.).

References

[1] Stanovich K E and West R F 2000 *Behav. Brain Sci.* **23** 645

[2] Kahneman D 2011 *Thinking, fast and slow Thinking, fast and slow* (New York, NY, US: Farrar, Straus and Giroux)

[3] Hua W and Zhang Y 2022 *System 1 + System 2 = Better World: Neural-Symbolic Chain of Logic Reasoning* (Abu Dhabi, United Arab Emirates: Association for Computational Linguistics)

[4] Wei J, Wang X, Schuurmans D, Bosma M, Ichter B, Xia F, Chi E, Le Q, and Zhou D 2023 arXiv:2201.11903 [cs.CL]

[5] Zhang H, Liu C, Wang Z, Wei X, Liu S, Zheng N, Shao B, and Liu T Y 2024 arXiv:2403.09560 [cs.LG]

[6] Hörmann L, Stark W G, and Maurer R J 2025 arXiv:2503.19814 [cond-mat.mtrl-sci]

[7] McGiff J and Nikolov N S 2025 arXiv:2505.04531 [cs.CL]

[8] Pang Y T, Kuo K M, Yang L, and Gumbart J C 2025 bioRxiv 2025.02.27.640693

[9] Schwinger J 1948 *Phys. Rev.* **73** 416

[10] Laporta S and Remiddi E 1996 *Phys. Lett. B* **379** 283

[11] Aoyama T, Hayakawa M, Kinoshita T, and Nio M 2015 *Phys. Rev. D* **91** 033006

[12] Aoyama T, Hayakawa M, Kinoshita T, and Nio M 2012 *Phys. Rev. Lett.* **109** 111807

[13] Fan X, Myers T G, Sukra B A D, and Gabrielse G 2023 *Phys. Rev. Lett.* **130** 071801

[14] Prokof'ev N V and Svistunov B V 1998 *Phys. Rev. Lett.* **81** 2514

[15] Prokof'ev N and Svistunov B 2008 *Phys. Rev. B* **77** 020408

[16] Van Houcke K, Kozik E, Prokof'ev N, and Svistunov B 2010 *Phys. Procedia* **6** 95

[17] Kozik E 2024 *Nat. Commun.* **15** 7916

[18] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A N, Kaiser L, and Polosukhin I 2023 arXiv:1706.03762 [cs.CL]

[19] Shao Z, Wang P, Zhu Q, Xu R, Song J, Bi X, Zhang H, Zhang M, Li Y K, Wu Y, and Guo D 2024 arXiv:2402.03300 [cs.CL]

[20] Guo D, Yang D, Zhang H *et al.* 2025 *Nature* **645** 633

[21] Liu Z, Chen C, Li W, Qi P, Pang T, Du C, Lee W S, and Lin M 2025 arXiv:2503.20783 [cs.LG]

[22] Yu Q, Zhang Z, Zhu R *et al.* 2025 arXiv:2503.14476 [cs.LG]

[23] Yang A, Yang B, Zhang B *et al.* 2024 arXiv:2412.15115 [cs.CL]

[24] Yang A, Li A, Yang B *et al.* 2025 arXiv:2505.09388 [cs.CL]

[25] Kimi Team, Du A, Gao B *et al.* 2025 arXiv:2501.12599 [cs.AI]

[26] Huang L, Yu W, Ma W, Zhong W, Feng Z, Wang H, Chen Q, Peng W, Feng X, Qin B, and Liu T 2025 *ACM Trans. Inf. Syst.* **43** 1

[27] Liu M and Fang J 2025 arXiv:2504.09440 [cs.AI]

[28] Yu T, Jing Y, Zhang X, Jiang W, Wu W, Wang Y, Hu W, Du B, and Tao D 2025 arXiv:2503.04550 [cs.AI]

[29] Wang Y, Yang Q, Zeng Z, Ren L, Liu L, Peng B, Cheng H, He X, Wang K, Gao J, Chen W, Wang S, Du S S, and Shen Y 2025 arXiv:2504.20571 [cs.LG]

[30] Gao Z, Chen L, Zhou J, and Dai B 2025 arXiv:2505.20282 [cs.CL]

[31] Abrikosov A A, Gorkov L P and Dzyaloshinski I E 1963 *Methods of Quantum Field Theory in Statistical Physics* (Dover Publications Inc.)

[32] Mahan G D Mahan G D 2000 *Many-Particle Physics* 3rd ed. (Springer New York, NY)

[33] Nieto A 1995 *Comput. Phys. Commun.* **92** 54

[34] Espinosa O 2009 *Math. Comput.* **79** 1709

[35] Tupitsyn I S and Prokof'ev N V 2025 *Phys. Rev. B* **111** L041106

[36] Rossi R 2017 *Phys. Rev. Lett.* **119** 045701

[37] Chen K and Haule K 2019 *Nat. Commun.* **10** 3725

[38] Haule K and Chen K 2022 *Sci. Rep.* **12** 2294

[39] Burke M D and LeBlanc J P F 2025 *Comput. Phys. Commun.* **308** 109437

[40] LeBlanc J P F, Chen K, Haule K, Prokof'ev N V, and Tupitsyn I S 2022 *Phys. Rev. Lett.* **129** 246401
[41] Taheridehkordi A, Curnoe S H, and LeBlanc J P F 2019 *Phys. Rev. B* **99** 035120
[42] Power A, Burda Y, Edwards H, Babuschkin I, and Misra V 2022 arXiv:2201.02177 [cs.LG]
[43] Lin Y, Zhang Y, Feng S, and Zhao H 2025 arXiv:2501.02436 [cs.LG]
[44] Song Z Y, Yang T Z, Cao Q H, Luo M X, and Zhu H X 2025 arXiv:2502.09544 [hep-ph]
[45] von Hippel M and Wilhelm M 2025 arXiv:2502.05121 [hep-th]