

Reinforcement learning and metaheuristics for Feynman-integral reduction

Mao Zeng^{*}

Higgs Centre for Theoretical Physics, School of Physics and Astronomy, University of Edinburgh, Edinburgh EH9 3FD, United Kingdom



(Received 23 June 2025; accepted 5 December 2025; published 31 December 2025)

We propose new methods for optimizing the integration-by-parts (IBP) reduction of Feynman integrals, an important computational bottleneck in modern perturbative calculations in quantum field theory. Using the simple example of one-loop massive bubble integrals, we pose the problem of minimizing the number of arithmetic operations in reducing a target integral to master integrals via the Laporta algorithm. This is a nontrivial combinatorial optimization problem over the ordering of IBP equation generation (from pairs of seed integrals and IBP operators) and the ordering of integral elimination. Our first proposed method is reinforcement learning, which involves an agent interacting with an environment in a step-by-step manner and learning the best actions to take given an observation of the environment (in this case, the current state of the IBP reduction process). The second method is using metaheuristics, e.g., simulated annealing, to minimize the computational cost as a black-box function of numerical priority values that control the orderings. For large-scale problems, the number of free parameters can be compressed by using a small neural network to assign priority values. Remarkably, with almost no human guidance, both methods lead to IBP reduction schemes that are competitive with the most efficient human-designed algorithms. We also found interpretable features in the AI results that may be applicable to more complicated problems.

DOI: [10.1103/dmlf-jkfc](https://doi.org/10.1103/dmlf-jkfc)

I. INTRODUCTION

Perturbative quantum field theory as a precision science requires the evaluation of Feynman integrals from multi-loop amplitudes and correlation functions, with many applications in, e.g., collider physics, statistical mechanics, gravitational wave physics, and cosmology. Integration-by-parts (IBP) reduction [1] is used in almost all nontrivial modern calculations except for those in highly symmetric special theories, and is frequently the most computationally demanding part of the calculations. IBP reduction is commonly carried out using the Laporta algorithm [2] as implemented in publicly available computer programs [3–14]. While there are many proposed alternative methods for IBP reduction, e.g., Refs. [15–38], the Laporta algorithm remains arguably the most widely applied in practical calculations. While the Laporta algorithm traditionally required complicated manipulations of polynomials and rational functions in intermediate steps, the use of finite-field numerical techniques [39–52] allows reconstructing complicated analytic results using a large number of

numerical evaluations. With such a workflow, it is important to make each numerical IBP reduction run as fast as possible.

Applications of the Laporta algorithm involve many heuristic choices that are based on experience rather than derived from first principles, and adjusting such choices can have a tremendous impact on the performance of IBP reduction. For example, there is a choice about whether Lorentz-invariance relations [53] are included, and the selection of seed integrals and IBP operators can be optionally trimmed using Lie algebra relations [21]. Recently, several papers observed that widely used algorithms for selecting seed integrals can be improved to vastly reduce the number of IBP equations [54–56].

The goal of this paper is to present automated strategies for *blind* searches, with minimal human input, of IBP reduction schemes that outperform, and/or provide valuable feedback on, human-designed algorithms. Among recent interests in machine learning in perturbative scattering amplitudes [57–61], two papers [62,63] applied genetic algorithms and LLM-powered program searches to find optimal symbolic expressions controlling the various ordering choices in the Laporta algorithm. Our work takes several novel directions not explored in the previous literature. First, we initiate the study of reinforcement learning in the context of Feynman integrals, which breaks down the IBP reduction problem to a step-by-step decision process. Second, in parallel, we introduce metaheuristics

^{*}Contact author: mao.zeng@ed.ac.uk

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP³.

for minimizing the cost function over numerical, rather than symbolic, priority values that control the orderings. This provides a complementary perspective and allows leveraging the vast body of available numerical optimization algorithms. Third, we go beyond the binary inclusion [62] and ordering [63] of seed integrals, and take a fine-grained approach of ordering *seed-operator pairs*, which allows more aggressive optimizations, e.g., applying only a subset of IBP operators to a given seed integral [7,21]. Finally, we use the *number of arithmetic operations*, tracked by our custom IBP reduction program, as a more accurate proxy of computational complexity than the number of seed integrals considered in previous papers.

II. IBP REDUCTION AND SIMPLIFIED COST MODEL

Consider the family of one-loop massive *bubble integrals* used as an introductory example in the book [64]

$$I_{\nu_1, \nu_2} = \int \frac{d^D k}{i\pi^{D/2}} \frac{1}{(-k^2 + m^2)^{\nu_1} [(p-k)^2 + m^2]^{\nu_2}}, \quad (1)$$

where the two propagators in Fig. 1 are raised to non-negative integer powers ν_1 and ν_2 . The integral vanishes in dimensional regularization if both indices are nonpositive. When one index is nonpositive, the integral is referred to as a *tadpole integral*, otherwise a genuine bubble integral. IBP equations arise as total derivatives integrate to zero in dimensional regularization

$$0 = \int \frac{d^D k}{i\pi^{D/2}} \frac{\partial}{\partial k^\mu} \frac{q^\mu}{(-k^2 + m^2)^{\nu_1} [(p-k)^2 + m^2]^{\nu_2}}, \quad (2)$$

where q^μ is any Lorentz vector. With a slight abuse of notation, we refer to the above IBP equation as the outcome of applying the *IBP operator*, $(\partial/\partial k^\mu)q^\mu$, to the *seed integral*, I_{ν_1, ν_2} . Specifically, we have

(i) IBP operator 1: $(\partial/\partial k^\mu)k^\mu$.

(ii) IBP operator 2: $(\partial/\partial k^\mu)p^\mu$.

We only study IBP relations and disregard symmetry relations $I_{\nu_1, \nu_2} = I_{\nu_2, \nu_1}$ for simplicity. It turns out that all I_{ν_1, ν_2} can be reduced to three *master integrals*, $I_{1,1}, I_{1,0}, I_{0,1}$. Now we walk through an application of the Laporta algorithm to reduce the *target integral* $I_{1,2}$ to the three master integrals at numerical values of dimension and kinematic variables,

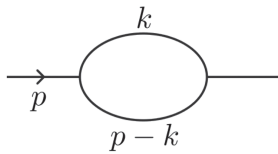


FIG. 1. One-loop massive bubble integrals.

$$p^2 = 1, \quad m^2 = 2, \quad D = 5/7, \quad (3)$$

which allows us to illustrate a simple cost model based on the number of arithmetic operations. The numerical values are unimportant, since the number of arithmetic operations is usually the same at different numerical values as long as the IBP system remain nonsingular, with occasional small differences due to accidental cancellations. An example sequence of IBP reduction steps, alternating between (a) generating an equation and reducing against all previous rules, and (b) eliminating an integral to form a reduction rule, is

- (i) Step 1a: generate an IBP equation from seed integral $I_{1,1}$ and IBP operator 2,

$$-I_{0,2} + I_{2,0} - I_{1,2} + I_{2,1} = 0. \quad (4)$$

We consider equation generation to be cheap, so do not track its cost. Now we reduce the above equation against previous reduction rules; currently there are none since we are at step 1, so there is no work to do and no row reduction cost (considering linear equations as rows in a matrix) is incurred.

- (ii) Step 1b: Choose an integral to eliminate. We choose $I_{1,2}$ to produce the reduction rule,

$$I_{1,2} \rightarrow -I_{0,2} + I_{2,0} + I_{2,1}. \quad (5)$$

In this step, we had to normalize the coefficient of the eliminated integral on the left-hand side (lhs) of the reduction rule to 1, and we count the normalization cost to be 4, i.e. the length of the equation. [In this case, the normalization is trivial, i.e. a sign flip, but this is a coincidence from our simple choice of kinematic variables Eq. (3).]

- (iii) Step 2a: generate an IBP equation from seed integral $I_{1,1}$ and IBP operator 1,

$$-I_{0,2} - (16/7)I_{1,1} + 3I_{1,2} + 4I_{2,1} = 0. \quad (6)$$

Reducing against all previous reduction rules, in this case the single rule Eq. (5), yields the reduced IBP equation,

$$-4I_{0,2} + 3I_{2,0} - (16/7)I_{1,1} + 7I_{2,1} = 0. \quad (7)$$

Since we reduced against a previous rule of length 4 (counting both lhs and rhs), we record a row reduction cost of 4. Generally, we need to reduce against more than one previous rule, and the cost is the sum of the lengths of all rules applied.

- (iv) Step 2b: Choose an integral to eliminate. We choose $I_{2,1}$ to produce the reduction rule,

$$I_{2,1} \rightarrow (4/7)I_{0,2} - (3/7)I_{2,0} + (16/49)I_{1,1}, \quad (8)$$

again with normalization cost 4. The chain of rules, Eqs. (5) and (8), reduces the target integral $I_{1,2}$ to the master integral $I_{1,1}$ and tadpole integrals. The cost incurred so far is 12, and continuing the process with tadpole seed integrals will finish the job of fully reducing $I_{1,2}$ the three master integrals.

III. REINFORCEMENT LEARNING

We start with some brief definitions. Reinforcement learning (RL) [65] is a branch of machine learning formulated in terms of Markov decision processes (MDPs), where an *agent* interacts with an environment and takes *actions* based on *observations* of the environment state. The actions cause transitions to other states and possibly positive feedback signals called *rewards*. When *terminal states* are reached, e.g., upon completion of some goal, a complete *episode* of interactions comes to an end. In *deep reinforcement learning*, specifically variants based on “policy gradients,” a deep neural network is trained to learn an optimized *policy*, i.e., a map between observations to the probability of taking each action, to maximize the expected future rewards. See Ref. [66] for a previous application of RL to theoretical high-energy physics.

To apply RL, we first specify the rewards. The reward received after each step is the negative of the cost (of row reduction or normalization) as illustrated in the previous section, until the episode terminates when we have succeeded in reducing $I_{3,3}$ to the three master integrals. Therefore, all the rewards in our environment are nonpositive, and maximizing the reward is equivalent to minimizing the cost. In principle, we could assign a positive reward for achieving the goal of reducing the target integral to master integrals, but since we have a finite system which always terminates with achieving this goal, such a positive reward would become an uninteresting constant for every episode.

Then we specify the observations. let us first revisit the IBP reduction of $I_{1,2}$ in the previous section before moving on to more nontrivial examples. We create graphical representations of the steps 1a and 2b, as two examples, in Fig. 2, which extends Fig. 3 with information about the IBP reduction process. An upper triangle Δ (or a lower triangle ∇) in a grid denotes that the seed integral has been used, in the most recent step or any previous step, with the IBP operator 1 (or IBP operator 2). The number in the triangle denotes the step at which a seed-operator is used to generate an IBP equation. Unfilled circles denote the integrals in the last reduced IBP equation (excluding the three master integrals) that are candidates to be eliminated to produce a new reduction rule. A filled circle denotes that an integral has been eliminated, and the number in it denotes the step at which this happened.

To target a more nontrivial integral $I_{3,3}$, we enlarge the board to have 37 squares in Fig. 3, where white square are candidate seed integrals, and gray squares are extra integrals that may appear in IBP equations. The

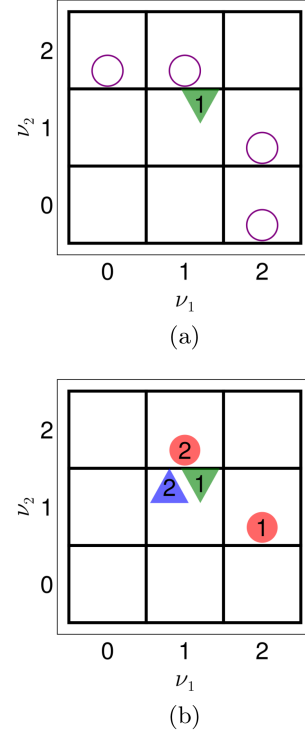


FIG. 2. Graphical representations of the state of the IBP reduction process for the target integral $I_{1,2}$ after step (1a), shown in subfigure (a), and after step (2b), shown in subfigure (b).

presence/absence and the step number in the three graphical features (upper/lower triangle, circle) are encoded in a $37 \times 3 = 111$ dimensional numerical vector as observation data. We use the simplest neural network architecture, a multilevel perceptron (MLP), which takes the 111 numbers as input and process the data through 3 hidden layers, each with 256 neurons, before generating an output whose

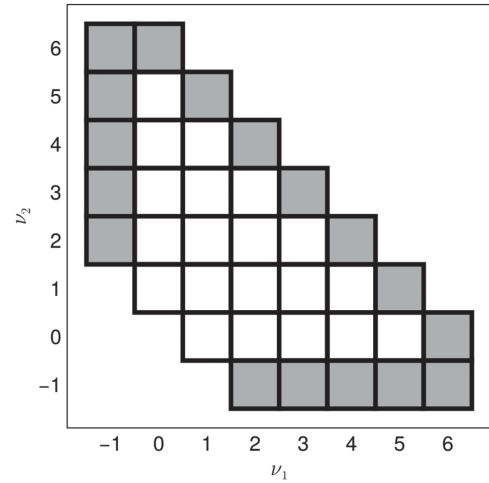


FIG. 3. The 20 candidate seed integrals (white squares) and 17 additional auxiliary integrals (light gray squares on the periphery).

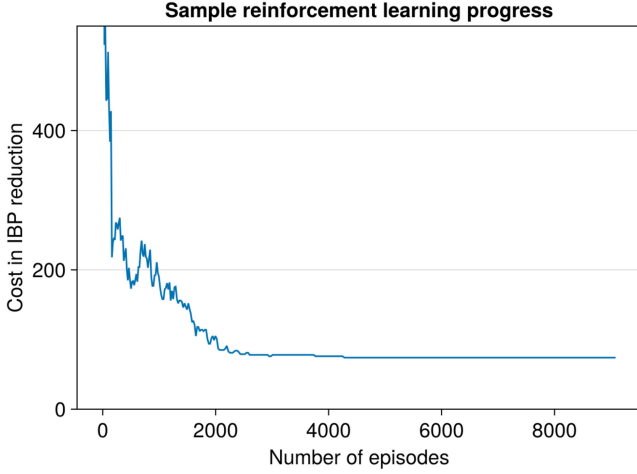


FIG. 4. Cost in reducing $I_{3,3}$ to master integrals versus the number of episodes used for training, in an example run that reached the best cost 74 after 4280 episodes. The cost at the beginning of training is 1234, outside the plot range, but drops below 550 after 30 episodes.

dimension is also 111, for probabilities for the next action. (The actual input dimension of nonzero data is 77, since the gray squares in Fig. 3 are not used as seed integrals. Similarly, the relevant output dimension for actions is at most 77, since the action must alternate between equation generation and elimination, and any seed-operator pair can only be used once.) We use the RL algorithm of proximal policy optimization [67]. Implementation details are covered by the Supplemental Material [68], which includes Refs. [69,70].

We run the algorithm 32 times, each with a maximum of 322560 steps. Typically a few dozen steps make up one episode. The costs achieved at the end of the 32 training runs have a median of 93.5 a minimum of 74. The minimum cost is achieved by 2 of the runs, and the learning curve showing the cost versus the number of episodes is in Fig. 4. The IBP reduction steps leading to the best cost of 74 is shown in our graphical notation in Fig. 5. Strikingly, starting from random exploration, RL has produced an IBP reduction scheme with distinct and interpretable features. First, the choice of seed integrals (denoted by triangles) adheres to a rectangular region with $\nu_1 \leq 3$, $\nu_2 \leq 2$. Second, the seed integrals are used in a strictly descending order in $\nu_1 + \nu_2$, ignoring the tadpole seed integrals. Third, elimination (denoted by numbers in circles) also follows a strictly descending order in $\nu_1 + \nu_2$ when ignoring tadpole integrals. The machine-generated log documenting the detailed steps is included in the Supplemental Material [68].

IV. METAHEURISTICS

We propose a second method for optimizing IBP reduction using the same cost model and action space. We assign numerical priority values to the 40 seed-operator pairs

corresponding to the white squares in Fig. 3 with either of the two IBP operators, which induce an ordering (in descending priority values) for generating IBP equations. Similarly, we assign numerical priority values to all 34 nonmaster integrals to order elimination. Combined, the 74 priority values completely specify the IBP reduction process, again for the example target integral $I_{3,3}$. IBP reduction proceeds with the prescribed order until a chain of reduction rules has been found which reduces the target integral to master integrals, and at this point the process terminates. The integer-valued IBP reduction cost is therefore a function of the 74 priority values, and we can feed this *black-box function* to numerical minimization algorithms.

Given the nondifferentiable (piecewise flat) nature of the cost function, we use metaheuristics, i.e., high-level search strategies for approximate optimal solutions. Specifically, we use *simulated annealing* (SA) [71], and treat the IBP reduction costs as energy levels in a thermodynamic system. We start with a population of 200 randomly initialized points in the 74-dimensional space of priority values bounded in the interval $[-2, 2]$. In each iteration, for each point, a random neighbor point at a small distance is probed. A transition to the neighbor point takes place if the latter has a lower cost, and if the cost is higher, the transition can still take place but with a probability that is suppressed by a Boltzmann factor dependent on the increase in cost. The temperature is initially set high to encourage exploration, and then gradually lowered to allow the points to settle down at values that minimize the black-box function. More implementation details are in the Supplemental Material [68], which includes Refs. [72,73]. We ran 32 times with different random numbers, with about 50 iterations and 9602 function evaluations in each run (comparable with the RL run in Fig. 4). The lowest cost is again 74, same as obtained with RL, found by 3 of the 32 SA runs. The median cost achieved is 84.5. The IBP reduction steps found by a run achieving the lowest cost is illustrated in Fig. 6. We can see that the seed integrals used and the IBP operators used with each seed integral are in complete agreement with the best RL solution in Fig. 5, despite some differences in the order of the steps.¹

V. PARAMETER COMPRESSION WITH A NEURAL NETWORK

Now we propose a way to scale up the simulated annealing method to handle large-scale problems without a prohibitively large increase in the number of free parameters. We use a tiny neural network illustrated in Fig. 7, with 15 free parameters, to assign priority values for

¹Note that when two computation steps are independent of each other, swapping them trivially has no effect. Therefore, a careful dependency analysis, not carried out in this work, would be needed to actually establish whether the solutions of Figs. 5 and 6 are equivalent or not.

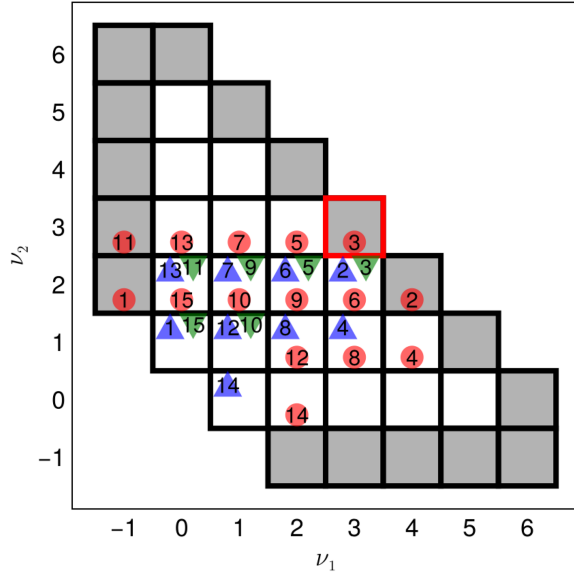


FIG. 5. The optimal IBP reduction steps found by reinforcement learning, for reducing $I_{3,3}$ marked as the red square.

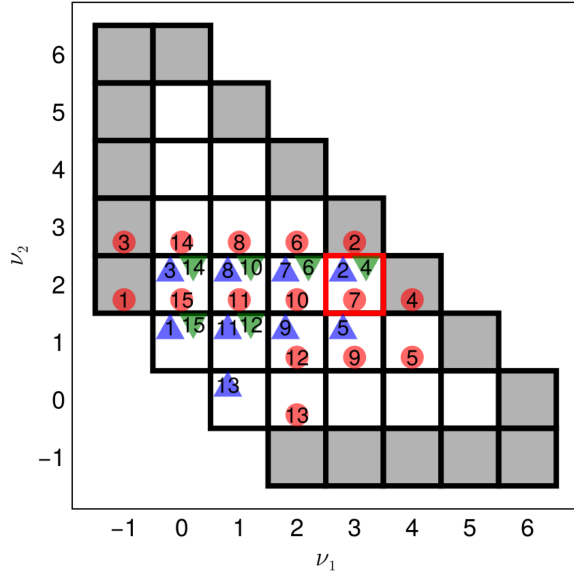


FIG. 6. The optimal IBP reduction steps found by simulated annealing, for reducing $I_{3,3}$ marked as the red square.

the integral I_{ν_1, ν_2} , paired with an IBP operator labeled $n > 0$ for seeding or with $n = 0$ for elimination. Constant normalization factors are applied to rescale the three input numbers to between 0 and 1. Then we use simulated annealing again to optimize the neural network parameters.

For $I_{3,3}$, we obtain a minimum cost of 121 and median cost of 162.5 in 32 runs. Though not as good as results from RL and SA with the full 74 priority parameters, the minimum cost here still beats common human-designed algorithms, as we will see later.

We stress-test the method with the much harder problem of reducing $I_{50,50}$. In this case, we restrict the seed integral

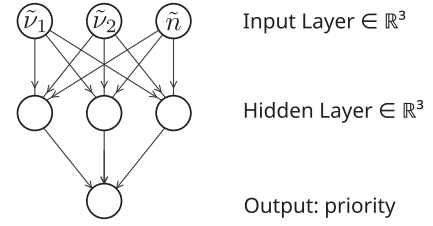


FIG. 7. A tiny neural network for assigning priority values, with normalized inputs $\tilde{\nu}_i = \nu_i / \max(\nu_i)$, $\tilde{n} = n / \max(n)$.

candidates to a rectangular range $0 \leq \nu_{1,2} \leq 50$ rather than a larger triangular range of the type in Fig. 3 which restricts $|\nu_1| + |\nu_2|$. Despite starting from a slightly more informed choice that reduces the number of seed integral candidates by one half, the automated algorithm still needs to discover the nontrivial optimized orderings of seed-operator pairs and elimination. The number of priority values is about 7500, which makes it very difficult to apply SA directly, so the parameter compression of the neural network becomes essential. The seed integral choice of the best solution found in 32 runs is shown in Fig. 8. A striking and interpretable feature of Fig. 8 is that only a small part of the square region is used as seed integrals, concentrated in two narrow strips, one at the top and one on the left, showing a dramatic reduction of the number of seed integrals compared with the human-designed rectangular seeding scheme $-1 \leq \nu_1, \nu_2 \leq 50$, i.e., filling up the full area in the figure. We refer to this as the “double strip” seeding scheme.

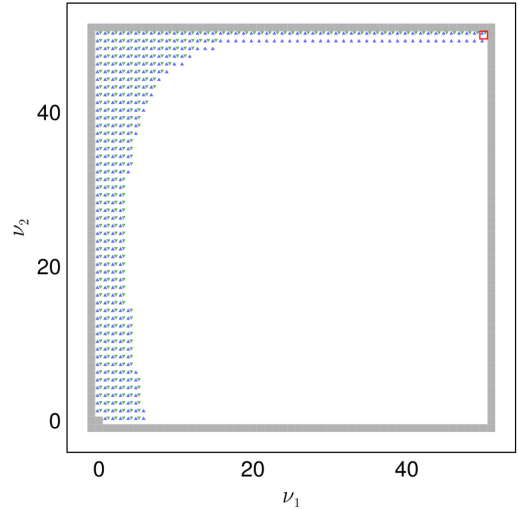


FIG. 8. The optimal IBP reduction steps found by simulated annealing with parameter compression by a neural network, for reducing the target integral $I_{50,50}$. The tiny red square in the top right corner marks the target integral. The bottom left gray square corresponds to $(-1, -1)$. Auxiliary integrals not allowed to be seed integrals are colored gray. Grid lines are omitted to avoid cluttering the figure.

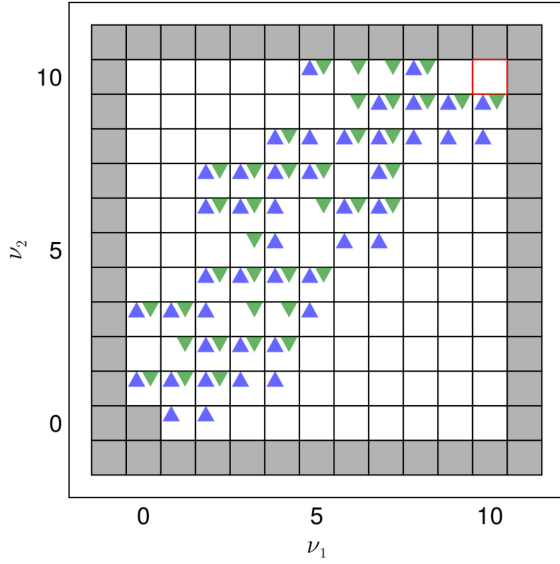


FIG. 9. The optimal IBP reduction steps found by simulated annealing for the full priority values, i.e., without parameter compression, for reducing the target integral $I_{10,10}$ (marked as a red square).

It is informative to look at another example, the reduction of $I_{10,10}$, which is a moderately sized problem which is still small enough to allow the application of simulated annealing both with and without the parameter compression by a neural network, while demonstrating more interesting features than the small-scale problem of reducing $I_{3,3}$. In Fig. 9, we show the seed-operator choices found by a single SA run optimizing the full set of priority values, and Fig. 10 shows the choices found by a single SA run with parameter compression. Each run starts with 300 randomly initialized

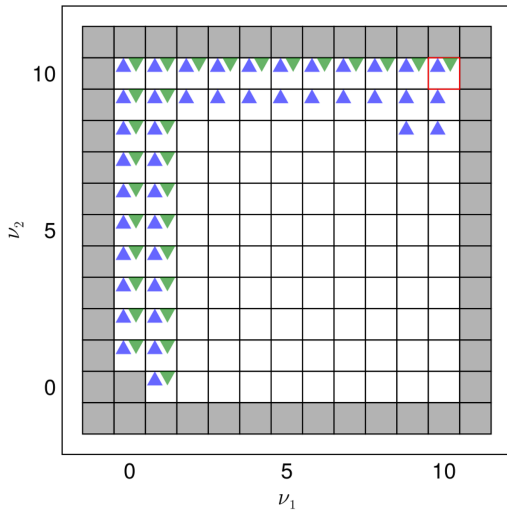


FIG. 10. The optimal IBP reduction steps found by simulated annealing with parameter compression by a tiny neural network, for reducing the target integral $I_{10,10}$ (marked as a red square).

points and uses about 100,000 IBP reduction runs. The two schemes found achieve costs of 961 (without parameter compression) and 919 (with parameter compression), respectively. Both figures omit the step numbers and the elimination choices in the interest of simplicity. Again, both schemes use significantly fewer seed integrals than the human-designed rectangular seeding scheme and leave large blank areas in the figures due to unused seed integrals. The scheme of Fig. 10 has a similar “double strip” structure as Fig. 8 for the larger-scale $I_{50,50}$ problem but has a more regular appearance, possibly because a larger number of steps would need to be run for the $I_{50,50}$ problem to converge to the same regular pattern.

VI. COMPARISON WITH HUMAN-DESIGNED ALGORITHMS

As a baseline algorithm resembling those in widely available public IBP codes, we follow Laporta [2] and generate IBP equations first from tadpole seed integrals and then genuine bubble seed integrals, chosen from the candidates indicated as white squares in Fig. 3, ordered with ascending complexity of the seed integrals. The complexity can be taken as $|\nu_1| + |\nu_2|$, using ν_1 to break ties. Each seed integral is used with the IBP operator 1 followed by operator 2. As a more optimized human-designed algorithm, we further restrict seed integrals to the rectangular range $\nu_1 \leq n_1, \nu_2 \leq n_2$ for reducing I_{n_1, n_2} . Optionally, we remove redundant IBP equations, i.e., those linearly dependent on previous equations, using the idea of Refs. [39,44] and discard the costs associated with the redundant equations. Finally, as the most dedicated human optimization effort, for $I_{3,3}$ only, we apply the advanced ordering algorithms of Kira [12–14], which requires generating all IBP equations at once for a global analysis at the cost of higher initial memory usage, using separate hand-tuned cutoffs for the two propagator powers following the recent work of Ref. [54].

Table I summarizes the IBP reduction costs achieved for reducing $I_{3,3}$, $I_{10,10}$, and $I_{50,50}$ found from reinforcement learning, two variants of simulated annealing, with or without parameter compression using a tiny neural network, and human-designed algorithms. Note that applying RL to problems larger than $I_{3,3}$ is left to future work due to technical limitations. For $I_{3,3}$, RL and SA both outperform most human-designed algorithms and achieve parity with the most optimized one with the aforementioned tradeoffs in memory usage and manual tuning effort.

For $I_{50,50}$ and $I_{10,10}$, we have already seen, in Figs. 8–10, that significantly fewer seed integrals are used than human-designed algorithms, and indeed incur lower reduction costs according to Table I. In the case of $I_{50,50}$ only, the steps found by SA contain redundant equations, and simply truncating redundant equations further improves the result.

TABLE I. Comparison of the best IBP reduction costs found by reinforcement learning (RL), simulated annealing (SA), possibly with a reduced parameter space using a neural network (NN), and human-designed algorithms. RL and SA runs depend on random number generators and are repeated for 32 times for $I_{3,3}$, with the lowest costs tabulated. Unlike the case of $I_{3,3}$, all $I_{10,10}$ and $I_{50,50}$ runs, including the baseline human algorithm, start from a rectangular seed range, possibly subject to further optimization by RL or SA algorithms.

Integral	$I_{3,3}$	$I_{10,10}$	$I_{50,50}$
RL	74		
SA	74	919	
SA + NN	121	961	17843
SA + NN + truncation	121	961	12775
Human (baseline)	365	3951	92471
Human(rectangular seeding + truncation)	237	2239	43579
Human(advanced ordering heuristics + separate cutoffs for ν_1 and ν_2)	74		

VII. CONCLUSION

IBP reduction of Feynman integrals is an important computational bottleneck in difficult calculations in perturbative QFT. We have demonstrated two new methods, namely reinforcement learning and metaheuristics, for optimizing IBP reduction by adjusting various orderings in the Laporta algorithm. One-loop massive bubble integrals are used to illustrate the methods. Despite the apparent simplicity of such integrals, the problem is surprisingly nontrivial, and reinforcement learning and metaheuristics both succeed in finding computational steps that are competitive with the most efficient human-designed algorithms. In fact, for the simplest example of reducing the bubble integral with each propagator raised to a cubic power, the two methods agree with each other in the lowest cost reached, and an example scheme that achieves this lowest cost is documented in detail in the Supplemental Material [68] as a benchmark for future investigations. We also took initial steps in tackling large-scale problems, again outperforming common human-designed algorithms for reducing $I_{50,50}$, which required tens of thousands of arithmetic operations. This was done by using a tiny neural network to reduce the number of free parameters and then applying simulated annealing. The application of reinforcement learning to large-scale problems is left as an open question for future studies. Since a first preprint version of the paper appeared, a moderate-sized example of applying simulated annealing to the reduction of $I_{10,10}$ has been added. The results in Figs. 10 and 8 suggest a novel “double-strip” seeding scheme with apparent linear rather than quadratic scaling of the number of seed integrals. It would be interesting to see if such a scheme can be generalized to more complicated families of integrals. Meanwhile, the result in Fig. 9 has some resemblance to the ellipse scheme found in Ref. [63], where seed integrals are concentrated in a diagonal area. The interpretability of the AI results may also inform better designs of algorithms by humans.

The two methods have their advantages and disadvantages. The ease of implementation is a major advantage of

using metaheuristics to optimize the IBP reduction cost as a black-box function of the orderings (encoded in numerical priority values). Though we focused on simulated annealing inspired by thermodynamics, a vast body of other metaheuristic algorithms, including particle swarm optimization and evolutionary/genetic algorithms, can be applied in an essentially plug-and-play manner. Reinforcement learning introduces the perspective of dynamically adjusting the chosen actions by reacting to changes in the environment. This brings the possibility that a single agent can be trained to adjust and adapt to a large number of different IBP reduction problems, e.g., involving different families of Feynman integrals.

Since the methods presented, in the current state, require thousands of IBP runs with different ordering choices before converging to optimal ones, the most likely practical application would be to the IBP reduction of Feynman integrals with a large number of kinematic scales, which could require hundreds of thousands or even more numerical finite-field runs for analytic results to be reconstructed. The cost savings in such a large number of later runs would make the initial optimization cost worthwhile. It is of course highly desirable to improve our methods and achieve better *sample efficiency*, i.e., be able to find highly optimized IBP reduction schemes using a small number of reduction runs.

In the future, we plan to publish our custom IBP reduction program that provides step-by-step feedback on computational costs with a standard interface (e.g., Ref. [74]), to facilitate follow-up investigations by both physics and machine learning communities. The present study focuses on pedagogical examples that can be quickly tested on a laptop in the matter of minutes, leaving many possible future enhancements, e.g., more sophisticated neural network architectures like convolutional neural networks, graph neural networks and transformers, as well as refining the details of metaheuristic algorithms, e.g., the temperature reduction scheme in simulated annealing. It would also be interesting to apply reinforcement learning

and metaheuristics to other problems in Feynman integral evaluations besides IBP reduction.

We use the software Makie.jl [75] for general visualization and NN-SVG [76] for the neural network illustration.

ACKNOWLEDGMENTS

We thank Johann Usovitsch for assistance in reanalyzing internal computation steps of Kira against our cost model

to provide insightful comparisons with AI results. M. Z.'s work is supported in part by the U.K. Royal Society through Grant No. URF/R1\20109.

DATA AVAILABILITY

The data supporting this study's findings are available within the article.

-
- [1] K. G. Chetyrkin and F. V. Tkachov, Integration by parts: The algorithm to calculate β -functions in 4 loops, *Nucl. Phys. B* **192**, 159 (1981).
 - [2] S. Laporta, High-precision calculation of multiloop Feynman integrals by difference equations, *Int. J. Mod. Phys. A* **15**, 5087 (2000).
 - [3] Charalampos Anastasiou and Achilleas Lazopoulos, Automatic integral reduction for higher order perturbative calculations, *J. High Energy Phys.* **07** (2004) 046.
 - [4] C. Studerus, Reduze-Feynman integral reduction in C++, *Comput. Phys. Commun.* **181**, 1293 (2010).
 - [5] Roman N. Lee, LiteRed 1.4: A powerful tool for reduction of multiloop integrals, *J. Phys. Conf. Ser.* **523**, 012059 (2014).
 - [6] A. V. Smirnov, Algorithm FIRE—Feynman Integral REDuction, *J. High Energy Phys.* **10** (2008) 107.
 - [7] A. V. Smirnov and V. A. Smirnov, FIRE4, LiteRed and accompanying tools to solve integration by parts relations, *Comput. Phys. Commun.* **184**, 2820 (2013).
 - [8] Alexander V. Smirnov, FIRE5: A C++ implementation of Feynman Integral REDuction, *Comput. Phys. Commun.* **189**, 182 (2015).
 - [9] A. V. Smirnov and F. S. Chukharev, FIRE6: Feynman Integral REDuction with modular arithmetic, *Comput. Phys. Commun.* **247**, 106877 (2020).
 - [10] Alexander V. Smirnov and Mao Zeng, FIRE 6.5: Feynman integral reduction with new simplification library, *Comput. Phys. Commun.* **302**, 109261 (2024).
 - [11] Philipp Maierhöfer, Johann Usovitsch, and Peter Uwer, Kira—A Feynman integral reduction program, *Comput. Phys. Commun.* **230**, 99 (2018).
 - [12] Philipp Maierhöfer and Johann Usovitsch, Kira 1.2 release notes, [arXiv:1812.01491](https://arxiv.org/abs/1812.01491).
 - [13] Jonas Klappert, Fabian Lange, Philipp Maierhöfer, and Johann Usovitsch, Integral reduction with Kira 2.0 and finite field methods, *Comput. Phys. Commun.* **266**, 108024 (2021).
 - [14] Fabian Lange, Johann Usovitsch, and Zihao Wu, Kira 3: Integral reduction with efficient seeding and optimized equation selection, [arXiv:2505.20197](https://arxiv.org/abs/2505.20197).
 - [15] R. N. Lee, Presenting LiteRed: A tool for the Loop InTEgrals REDuction, [arXiv:1212.2685](https://arxiv.org/abs/1212.2685).
 - [16] O. V. Tarasov, Computation of Grobner bases for two loop propagator type integrals, *Nucl. Instrum. Methods Phys. Res., Sect. A* **534**, 293 (2004).
 - [17] Vladimir P. Gerdt and Daniel Robertz, A Maple package for computing Grobner bases for linear recurrence relations, *Nucl. Instrum. Methods Phys. Res., Sect. A* **559**, 215 (2006).
 - [18] A. V. Smirnov and Vladimir A. Smirnov, Applying Grobner bases to solve reduction problems for Feynman integrals, *J. High Energy Phys.* **01** (2006) 001.
 - [19] A. V. Smirnov, An Algorithm to construct Grobner bases for solving integration by parts relations, *J. High Energy Phys.* **04** (2006) 026.
 - [20] A. V. Smirnov and V. A. Smirnov, S-bases as a tool to solve reduction problems for Feynman integrals, *Nucl. Phys. B, Proc. Suppl.* **160**, 80 (2006).
 - [21] R. N. Lee, Group structure of the integration-by-part identities and its application to the reduction of multiloop integrals, *J. High Energy Phys.* **07** (2008) 031.
 - [22] Mohamed Barakat, Robin Brüser, Claus Fieker, Tobias Huber, and Jan Piclum, Feynman integral reduction using Gröbner bases, *J. High Energy Phys.* **05** (2023) 168.
 - [23] Janusz Gluza, Krzysztof Kajda, and David A. Kosower, Towards a basis for planar two-loop integrals, *Phys. Rev. D* **83**, 045012 (2011).
 - [24] Robert M. Schabinger, A new algorithm for the generation of unitarity-compatible integration by parts relations, *J. High Energy Phys.* **01** (2012) 077.
 - [25] Harald Ita, Two-loop integrand decomposition into master integrals and surface terms, *Phys. Rev. D* **94**, 116015 (2016).
 - [26] Kasper J. Larsen and Yang Zhang, Integration-by-parts reductions from unitarity cuts and algebraic geometry, *Phys. Rev. D* **93**, 041701(R) (2016).
 - [27] S. Abreu, F. F. Cordero, H. Ita, M. Jaquier, B. Page, and M. Zeng, Two-loop four-gluon amplitudes from numerical unitarity, *Phys. Rev. Lett.* **119**, 142001 (2017).
 - [28] Samuel Abreu, Fernando Febres Cordero, Harald Ita, Ben Page, and Mao Zeng, Planar two-loop five-gluon amplitudes from numerical unitarity, *Phys. Rev. D* **97**, 116014 (2018).
 - [29] Janko Böhm, Alessandro Georgoudis, Kasper J. Larsen, Hans Schönemann, and Yang Zhang, Complete integration-by-parts reductions of the non-planar hexagon-box via module intersections, *J. High Energy Phys.* **09** (2018) 024.
 - [30] Dominik Bendle, Janko Böhm, Wolfram Decker, Alessandro Georgoudis, Franz-Josef Pfreundt, Mirko Rahn, Pascal Wasser, and Yang Zhang, Integration-by-parts reductions of Feynman integrals using Singular and GPI-Space, *J. High Energy Phys.* **02** (2020) 079.

- [31] Zihao Wu, Janko Boehm, Rourou Ma, Hefeng Xu, and Yang Zhang, NeatIBP 1.0, a package generating small-size integration-by-parts relations for Feynman integrals, *Comput. Phys. Commun.* **295**, 108999 (2024).
- [32] Zihao Wu, Janko Böhm, Rourou Ma, Johann Usovitsch, Yingxuan Xu, and Yang Zhang, Performing integration-by-parts reductions using NeatIBP 1.1 + Kira, *Comput. Phys. Commun.* **316**, 109798 (2025).
- [33] Pierpaolo Mastrolia and Sebastian Mizera, Feynman integrals and intersection theory, *J. High Energy Phys.* **02** (2019) 139.
- [34] Hjalte Frellesvig, Federico Gasparotto, Stefano Laporta, Manoj K. Mandal, Pierpaolo Mastrolia, Luca Mattiazzi, and Sebastian Mizera, Decomposition of Feynman integrals by multivariate intersection numbers, *J. High Energy Phys.* **03** (2021) 027.
- [35] Xiao Liu and Yan-Qing Ma, Determining arbitrary Feynman integrals by vacuum integrals, *Phys. Rev. D* **99**, 071501(R) (2019).
- [36] Xin Guan, Xiao Liu, and Yan-Qing Ma, Complete reduction of integrals in two-loop five-light-parton scattering amplitudes, *Chin. Phys. C* **44**, 093106 (2020).
- [37] David A. Kosower, Direct solution of integration-by-parts systems, *Phys. Rev. D* **98**, 025008 (2018).
- [38] Bo Feng, Chang Hu, Jiyuan Shen, and Yaobo Zhang, General one-loop generating function by IBP relations, *Phys. Rev. D* **111**, 076026 (2025).
- [39] Philipp Kant, Finding linear dependencies in integration-by-parts equations: A Monte Carlo approach, *Comput. Phys. Commun.* **185**, 1473 (2014).
- [40] Andreas von Manteuffel and Robert M. Schabinger, A novel approach to integration by parts reduction, *Phys. Lett. B* **744**, 101 (2015).
- [41] Tiziano Peraro, Scattering amplitudes over finite fields and multivariate functional reconstruction, *J. High Energy Phys.* **12** (2016) 030.
- [42] S. Abreu, J. Dormans, F.F. Cordero, H. Ita, and B. Page, Analytic form of planar two-loop five-gluon scattering amplitudes in QCD, *Phys. Rev. Lett.* **122**, 082002 (2019).
- [43] Jonas Klappert and Fabian Lange, Reconstructing rational functions with FireFly, *Comput. Phys. Commun.* **247**, 106951 (2020).
- [44] Tiziano Peraro, FiniteFlow: Multivariate functional reconstruction using finite fields and dataflow graphs, *J. High Energy Phys.* **07** (2019) 031.
- [45] Giuseppe Laurentis and Daniel Maître, Extracting analytical one-loop amplitudes from numerical evaluations, *J. High Energy Phys.* **07** (2019) 123.
- [46] Jonas Klappert, Sven Yannick Klein, and Fabian Lange, Interpolation of dense and sparse rational functions and other improvements in FireFly, *Comput. Phys. Commun.* **264**, 107968 (2021).
- [47] Giuseppe De Laurentis and Ben Page, Ansätze for scattering amplitudes from p-adic numbers and algebraic geometry, *J. High Energy Phys.* **12** (2022) 140.
- [48] Vitaly Magerya, Rational tracer: A tool for faster rational function reconstruction, *arXiv:2211.03572*.
- [49] A. V. Belitsky, A. V. Smirnov, and R. V. Yakovlev, Balancing act: Multivariate rational reconstruction for IBP, *Nucl. Phys.* **B993**, 116253 (2023).
- [50] Herschel A. Chawdhry, p-adic reconstruction of rational functions in multiloop amplitudes, *Phys. Rev. D* **110**, 056028 (2024).
- [51] Xiao Liu, Reconstruction of rational functions made simple, *Phys. Lett. B* **850**, 138491 (2024).
- [52] Andreas Maier, Scaling up to multivariate rational function reconstruction, *Comput. Phys. Commun.* **317**, 109827 (2025).
- [53] T. Gehrmann and E. Remiddi, Differential equations for two-loop four-point functions, *Nucl. Phys.* **B580**, 485 (2000).
- [54] Mathias Driesse, Gustav Uhre Jakobsen, Gustav Mogull, Jan Plefka, Benjamin Sauer, and Johann Usovitsch, Conservative black hole scattering at fifth post-Minkowskian and first self-force order, *Phys. Rev. Lett.* **132**, 241402 (2024).
- [55] Xin Guan, Xiao Liu, Yan-Qing Ma, and Wen-Hao Wu, Blade: A package for block-triangular form improved Feynman integrals decomposition, *Comput. Phys. Commun.* **310**, 109538 (2025).
- [56] Zvi Bern, Enrico Herrmann, Radu Roiban, Michael S. Ruf, Alexander V. Smirnov, Vladimir A. Smirnov, and Mao Zeng, Amplitudes, supersymmetric black hole scattering at $\mathcal{O}(G^5)$, and loop integration, *J. High Energy Phys.* **10** (2024) 023.
- [57] Aurélien Dersy, Matthew D. Schwartz, and Xiaoyuan Zhang, Simplifying polylogarithms with machine learning, *Int. J. Data Sci. Math. Sci.* **1**, 135 (2024).
- [58] Ryusuke Jinno, Gregor Kälin, Zhengwen Liu, and Henrique Rubira, Machine learning post-Minkowskian integrals, *J. High Energy Phys.* **07** (2023) 181.
- [59] Francesco Calisto, Ryan Moodie, and Simone Zoia, Learning Feynman integrals from differential equations with neural networks, *J. High Energy Phys.* **07** (2024) 124.
- [60] Tianji Cai, Garrett W. Merz, François Charton, Niklas Nolte, Matthias Wilhelm, Kyle Cranmer, and Lance J. Dixon, Transforming the bootstrap: Using transformers to compute scattering amplitudes in planar $\mathcal{N} = 4$ super Yang–Mills theory, *Mach. Learn. Sci. Tech.* **5**, 035073 (2024).
- [61] Clifford Cheung, Aurélien Dersy, and Matthew D. Schwartz, Learning the simplicity of scattering amplitudes, *SciPost Phys.* **18**, 040 (2025).
- [62] Matt von Hippel and Matthias Wilhelm, Refining integration-by-parts reduction of Feynman integrals with machine learning, *J. High Energy Phys.* **05** (2025) 185.
- [63] Zhuo-Yang Song, Tong-Zhi Yang, Qing-Hong Cao, Mingxing Luo, and Hua Xing Zhu, Explainable AI-assisted Optimization for Feynman integral reduction, *arXiv:2502.09544*.
- [64] Stefan Weinzierl, *Feynman Integrals. A Comprehensive Treatment for Students and Researchers*, UNITEXT for Physics (Springer, Cham, 2022).
- [65] Andrew G. Barto and Richard S. Sutton, *Reinforcement Learning: An Introduction* (MIT Press, 1998).
- [66] Andrei Constantin, Thomas R. Harvey, and Andre Lukas, Heterotic string model building with monad bundles and reinforcement learning, *Fortschr. Phys.* **70**, 2100186 (2022).
- [67] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov, Proximal policy optimization algorithms, *arXiv:1707.06347*.

- [68] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/dmlf-jkfc> for implementation details and lengthy data.
- [69] Jeff Bezanson, Alan Edelman, Stefan Karpinski, and Viral B Shah, Julia: A fresh approach to numerical computing, *SIAM Rev.* **59**, 65 (2017).
- [70] Anthony Corso and Robert Moss, Crux.jl, GitHub (2025), <https://github.com/sisl/Crux.jl>.
- [71] Emile H. L. Aarts and Peter J. M. Van Laarhoven, Simulated annealing: A pedestrian review of the theory and some applications, in *Pattern Recognition Theory and Applications* (Springer, New York, 1987), pp. 179–192.
- [72] Jesús-Adolfo Mejía de Dios and Efrén Mezura-Montes, Metaheuristics: A Julia package for single- and multi-objective optimization, *J. Open Source Software* **7**, 4723 (2022).
- [73] Héctor Corte, Simulated annealing optimization, MATLAB Central File Exchange (2011) (retrieved: April 21, 2025).
- [74] Mark Towers, Ariel Kwiatkowski, Jordan Terry, John U. Balis, Gianluca De Cola, Tristan Deleu, Manuel Goulao, Andreas Kallinteris, Markus Krimmel, Arjun KG *et al.*, Gymnasium: A standard interface for reinforcement learning environments, [arXiv:2407.17032](https://arxiv.org/abs/2407.17032).
- [75] Simon Danisch and Julius Krumbiegel, Makie.jl: Flexible high-performance data visualization for Julia, *J. Open Source Software* **6**, 3349 (2021).
- [76] Alexander LeNail, NN-SVG: Publication-ready neural network architecture schematics, *J. Open Source Software* **4**, 747 (2019).