

The dependence of observables on action parameters

Guilherme Catumba^{a,*}, Alberto Ramos^a and Bryan Zaldivar^a

^a*Instituto de Física Corpuscular (IFIC) CSIC - Universitat de Valencia.
46071, Valencia, Spain*

E-mail: gtelo@ific.uv.es, alberto.ramos@ific.uv.es, b.zaldivar.m@csic.es

Many applications in Lattice field theory require to determine the Taylor series of observables with respect to action parameters. A primary example is the determination of electromagnetic corrections to hadronic processes. We show two possible solutions to this general problem, one based on reweighting, that can be considered a generalization of the RM123 method. The other based on the ideas of Numerical Stochastic Perturbation Theory (NSPT) in the Hamiltonian formulation. We show that 1) the NSPT-based approach shows a much reduced variance in the determination of the Taylor coefficients, and 2) That both approaches are related by a change of variables. Numerical results are shown for the case of $\lambda - \phi^4$ in 4 dimensions, but we expect these observations to be general. We conclude by commenting on the possible use of Machine Learning techniques to find similar change of variables that can potentially reduce the variance in Taylor coefficients.

*The 40th International Symposium on Lattice Field Theory,
31st July-4th August, 2023,
Fermilab, Batavia, Illinois, USA*

*Speaker

1. Introduction

The use of Monte Carlo (MC) techniques is a useful and common practice in physics, particularly in lattice field theory. Other than physics, MC methods are broadly used in the machine learning community, in statistical inference, general stochastic optimization problems, among others. In general, these systems aim at studying an underlying statistical distribution, $p_\theta(x) \propto e^{-S(x;\theta)}$, from which we extract observables. On the lattice these involve functional integrals of the form $\langle O \rangle \propto \int dx O(x) e^{-S(x;\theta)}$ that are stochastically estimated through MC sampling

$$\langle O(x; \theta) \rangle_{p_\theta} = \frac{1}{N} \sum_{\alpha=1}^N O(x^\alpha; \theta) + O\left(1/\sqrt{N}\right), \quad (1)$$

where the $\{x^\alpha\}_{\alpha=1}^N$ are samples from the distribution $p_\theta(x)$.

As implied by the notation, the distribution functions generally depend on parameters $\theta = \{\theta_i\}$ (e.g., bare couplings in case of a lattice field theory). Here we are interested in studying how the observables change under variations of the parameters¹. The aim of this work is to compute derivatives of the observables with respect to the parameters of the distribution function

$$\frac{\partial}{\partial \theta_i} \langle O(x; \theta) \rangle_{p_\theta}, \quad \frac{\partial^2}{\partial \theta_i \partial \theta_j} \langle O(x; \theta) \rangle_{p_\theta}, \quad \dots \quad (2)$$

using automatic differentiation techniques.

The available solutions to this problem (see [1] for examples) usually require the normalization of the distribution function, preventing its application in conjunction with MC. Additionally, the use of ordinary Automatic Differentiation [2] (suited for deterministic functions) is not possible due to the stochastic character of the problem. In order to extend automatic differentiation to stochastic problems we explore two possible approaches: using reweighting; and by modifying Hamiltonian type sampling algorithms.

1.1 Automatic Differentiation and Truncated polynomials

Conventional automatic differentiation (AD) is a method used to compute derivatives of deterministic functions exactly (up to machine precision). Forward AD is one of the possible implementations, it exploits the fact that any deterministic function can be built by a finite composition of elementary operations (addition, multiplication, etc.) and elementary functions (exp, sin, etc.). Defining the derivatives for each fundamental operation and employing the chain rule enables the straightforward computation of any derivative.

In order to implement this idea we start by defining the order K truncated polynomials, \tilde{x} , as expansions in the variable ε of the form

$$\tilde{x} \equiv x_0 + x_1 \varepsilon + x_2 \varepsilon^2 + \dots + x_K \varepsilon^K. \quad (3)$$

¹In optimization problems one usually wants to find the optimal value for the parameters under some constraint; in Bayesian inference the sensitivity of the results w.r.t. the *hyper-parameters* is important; both cases require the computation of derivatives with respect to the parameters θ – see [1]

Secondly, all elementary operations and functions are defined such that all operations are exact at each order:

$$\begin{aligned}\tilde{x}\tilde{y} &= x_0y_0 + (x_0y_1 + x_1y_0)\varepsilon + (x_0y_2 + 2x_1y_1 + x_2y_0)\varepsilon^2 + \dots, \\ \exp(\tilde{x}) &= e^{x_0} + e^{x_0}x_1\varepsilon + e^{x_0}(x_1^2/2 + x_2)\varepsilon^2 + \dots, \\ &\dots\end{aligned}$$

Following this, and by virtue of the Taylor theorem, if we promote a variable x to the truncated polynomial $\tilde{x}(\varepsilon) = x + \varepsilon$ and evaluate a function $f(x)$ with this variable we get the truncated Taylor series of f at x

$$\tilde{f}(\varepsilon) = f(\tilde{x}(\varepsilon)) = \sum_n f_n \varepsilon^n, \quad f_n = \frac{1}{n!} \left. \frac{\partial^n f}{\partial x^n} \right|_x. \quad (4)$$

It is important to note that this expansion can be performed in multiple independent variables truncated at different orders – see [1] for a thorough introduction. The numerical implementation of these rules allows us to extend the ideas of AD to Monte Carlo – a freely available implementation of the algebra of truncated polynomials used in this work can be found in [3].

1.2 Reweighting & AD

The first extension of AD to Monte Carlo combines reweighting techniques with the truncated polynomials. Starting with samples distributed according to a distribution, $\{x^\alpha\} \sim p_\theta(x)$, reweighting is a way to compute expectation values w.r.t. a different distribution $p_{\theta'}(x)$. This is done by performing a weighted average

$$\langle f(x; \theta) \rangle_{p_{\theta'}} = \frac{\left\langle \frac{p_{\theta'}}{p_\theta} f(x; \theta) \right\rangle_{p_\theta}}{\left\langle \frac{p_{\theta'}}{p_\theta} \right\rangle} \quad (5)$$

where the weights $w^\alpha \equiv p_{\theta'}(x^\alpha)/p_\theta(x^\alpha)$ are computed for each sample α . This is an exact relation whose efficiency depends on some measure of similarity between the two distributions.

Since we are interested in computing derivatives w.r.t. the parameters θ , $\frac{\partial^n}{\partial \theta^n}$, our method amounts to modifying the target distribution by writing θ' as a truncated polynomial, $\tilde{\theta} = \theta + \varepsilon$. The new distribution $p_{\theta'}(x) = p_{\tilde{\theta}}(x)$ becomes a truncated polynomial, carrying the θ -dependence, and consequently the weights $\tilde{w}^\alpha(\varepsilon) \equiv p_{\tilde{\theta}}(x^\alpha)/p_\theta(x^\alpha)$ and the function $f(x; \tilde{\theta})$ both become a power series in ε^2 . When the reweighting average is performed, we obtain a truncated polynomial whose orders give stochastic estimates for the θ -derivatives.

1.3 Hamiltonian perturbative expansion

The second method to extend automatic differentiation for MC consists on a modification of Hamiltonian sampling algorithms. We consider the case of Hybrid Monte Carlo (HMC) [4], but the implementation is easily mimicked for other algorithms. The approach is inspired in Numerical

²Notice that the coefficients of this power series represent the derivatives w.r.t. θ instead of x as in eq. (4).

Stochastic Perturbation Theory [5–7], a method used to compute perturbative corrections to a free field theory by solving the molecular dynamics equations at each order in the perturbative coupling.

The HMC algorithm introduces fictitious momenta π conjugated to the variables x . Instead of sampling $p_\theta(x) \propto \exp\{-S(x; \theta)\}$, it samples a modified distribution

$$q_\theta \propto e^{-H(\pi, x; \theta)}, \quad H(\pi, x; \theta) = \frac{1}{2}\pi^2 + S(x; \theta). \quad (6)$$

Due to the π -dependence of the Hamiltonian, $H(\pi, x; \theta)$, the momenta are Gaussian distributed. It is straightforward to see that expectation values that depend only on x remain the same when computed with respect to q_θ . The sampling then follows by repeated iteration of the following sequence: 1) sampling π from a Gaussian distribution, $\pi(0) \sim e^{-\pi^2/2}$; 2) solving Hamilton's equations of motion (EOM)

$$\dot{x} = \frac{\partial H}{\partial \pi} = \pi, \quad \dot{\pi} = -\frac{\partial H}{\partial x}, \quad (7)$$

for a time interval τ (randomly sampled for ergodicity); and finally 3) accepting the new configuration $(\pi(\tau), x(\tau))$ with probability $e^{-\Delta H}$. The energy violation $\Delta H = H(\pi(\tau), x(\tau)) - H(\pi(0), x(0))$ is due to numerical imprecision when solving the EOM (here we use higher order solvers [8]). The chain of values $x(\tau)$ obtained in this way are distributed according to $p_\theta(x)$.

The extension for automatic differentiation starts by 1) promoting the parameters to truncated polynomials, $\tilde{\theta}(\varepsilon) = \theta + \varepsilon$. This implies that the remaining variables also become expansions, $x, \pi \rightarrow \tilde{x}(\varepsilon), \tilde{\pi}(\varepsilon)$. Momentum refresh is done at zeroth order only, $\tilde{\pi}(0) \sim \exp(\pi^2(0)/2)$, while the remaining orders are initialized to zero. 2) The equations of motion are solved at each order (this is straightforward when using a numerical implementation of the algebra of truncated polynomials); 3) The energy violation ΔH is a truncated polynomial, not allowing to perform the accept/reject step. The results require an extrapolation to zero integration step size, however, in practice it is enough to work at a sufficiently precise step size such that the possible extrapolation is below the statistical error.

With this modification, we obtain samples $\{\tilde{x}^\alpha\}_{\alpha=1}^N$ which are truncated polynomials, carrying the θ dependence. Ordinary expectation values correspond to a Taylor expansion around θ

$$\frac{1}{N} \sum_{\alpha=1}^N f(\tilde{x}^\alpha; \tilde{\theta}) = \langle f(x; \theta) \rangle + \frac{\partial}{\partial \theta} \langle f(x; \theta) \rangle \varepsilon + \dots \quad (8)$$

such that we recover the corresponding derivatives.

2. Example in lattice 4D scalar theory

We consider a $\lambda - \phi^4$ lattice theory in 4 space-time dimensions to test both methods. The lattice Euclidean action is

$$S_{\text{lat}}(\phi; m, \lambda) = \sum_x \left\{ \frac{1}{2} \sum_\mu [\phi(x + \mu) - \phi(x)]^2 + \frac{m^2}{2} \phi^2(x) + \lambda \phi^4(x) \right\} \quad (9)$$

with the field ϕ and mass m being dimensionless quantities.

		λ					
		0.0	0.1	0.2	0.3	0.4	
$\langle \phi^2 \rangle$	∂_{m^2}	RW	-0.0428(20)	-0.0328(14)	-0.0270(13)	-0.0241(12)	-0.0220(11)
		HAD	-0.042526(41)	-0.030880(14)	-0.026273(10)	-0.0233672(82)	-0.0212721(72)
	∂_λ	RW	-0.0779(22)	-0.05227(94)	-0.04370(89)	-0.03534(61)	-0.03169(50)
		HAD	-0.077816(79)	-0.052499(24)	-0.042218(19)	-0.035830(14)	-0.031323(11)
	$\partial_{m^2, \lambda}^2$	RW	0.43(43)	0.03(16)	0.16(14)	-0.10(11)	0.116(77)
		HAD	0.2733(22)	0.061593(99)	0.035082(69)	0.024240(42)	0.018263(31)
$\langle s \rangle$	∂_{m^2}	RW	-0.0025(42)	-0.0006(34)	0.0027(35)	0.0028(36)	0.0057(32)
		HAD	-0.000003(22)	0.002623(16)	0.004218(20)	0.005397(14)	0.006265(16)
	∂_λ	RW	-0.0765(46)	-0.0567(26)	-0.0538(25)	-0.0447(23)	-0.0400(17)
		HAD	-0.069774(49)	-0.057738(34)	-0.050128(40)	-0.044530(27)	-0.040250(27)
	$\partial_{m^2, \lambda}^2$	RW	-1.9(2.8)	-0.16(39)	0.36(43)	-0.43(32)	0.16(26)
		HAD	0.038864(96)	0.019197(66)	0.013405(69)	0.010126(50)	0.007860(47)

Table 1: First derivatives with respect to m^2 (∂_{m^2}), λ (∂_λ) and the cross derivative ($\partial_{m^2, \lambda}^2$) for $\langle \phi^2 \rangle$ and $\langle s \rangle$ from a L^4 lattice with $L/a = 32$ and $m^2 = 0.05$. The reweighting method is denoted RW while HAD corresponds to the Hamiltonian expansion.

The simulations³ are performed on a L^4 lattice with $L/a = 32$. This is done for several values of λ at fixed $m^2 = 0.05$. We are interested in estimating the dependence of expectation values of this theory on the couplings m^2 and λ . For this we promote $m^2 \rightarrow \tilde{m}^2$ and $\lambda \rightarrow \tilde{\lambda}$ (a double expansion around m^2 and λ) and perform reweighting from samples computed at a fixed value of the couplings. On the other hand, we also generate samples, $\{\tilde{\phi}^\alpha(x)\}_{\alpha=1}^N$, from the Hamiltonian expansion, where $\tilde{\phi}$ is a truncated polynomial in m^2 and λ (a single HMC simulation allows to compute both derivatives).

As observables, we consider expectation values of the field and the action density

$$\langle \phi^2(x) \rangle \quad \langle s(x; m, \lambda) \rangle = \frac{1}{2} \langle [\phi(x + \mu) - \phi(x)]^2 \rangle + \frac{m^2}{2} \langle \phi^2(x) \rangle + \lambda \langle \phi^4(x) \rangle. \quad (10)$$

Notice that in this way we test the method for observables with and without ‘connected’ contributions, *i.e.*, contributions for the derivative coming from the explicit dependence of the observables w.r.t. the parameters m^2 or λ . In table 1 we show the results for the first derivatives, ∂_{m^2} , ∂_λ , and cross derivative $\partial_{m^2, \lambda}^2$ computed with the reweighting and Hamiltonian approaches at various values of λ and for fixed $m^2 = 0.05$.

The first remarkable thing to notice is that the Hamiltonian approach delivers more precise results across all available parameter space. In fact, although both methods were obtained with the same statistics, compared to the reweighting method the Hamiltonian expansion shows a factor of ~ 100 in precision for the derivatives. The m^2 -derivative of the action density, which should be zero for $\lambda = 0$, is a good example of the level of precision of both methods. The situation is even more severe for the cross derivatives where the reweighting approach is not able to deliver a consistent signal while the Hamiltonian method maintains some precision.

It is important to notice that both methods are valid for derivatives of arbitrary orders. In the original work [1] the truncated Taylor expansion was shown to properly reconstruct the parameter

³The code used to generate the data can be found at <https://igit.ific.uv.es/alramos/lambdaphi4.jl/>

dependence for a significant range, which is dictated by the underlying function, the truncation order, and the available statistics.

2.1 Comparison of methods

What is the reason for the difference in the precision of both approaches? To understand this it is worth summarizing their properties and attempt to build a connection between the methods. We start with the reweighting for an observable O with weights $w = e^{-\delta S}$, whose formula at leading order is

$$\partial_\theta \langle O \rangle = \langle \partial_\theta O \rangle + [\langle O \rangle \langle \partial_\theta S \rangle - \langle O \partial_\theta S \rangle]. \quad (11)$$

The first term is the connected part, while the one in parenthesis corresponds to the disconnected contribution. The latter comes from the implicit dependence of the distribution function on the parameters θ and usually suffers from large variance leading to less precise results. While the reweighting method re-utilizes conventional samples from $p_\theta(x)$, the Hamiltonian samples, $\{\tilde{\phi}\}$, carry their dependence on the parameters $\theta = \{m^2, \lambda\}$ in the form of a Taylor series, thereby avoiding the computation of disconnected terms.

Another way to see this and connect both methods uses the reparametrization trick. Consider the following toy model, a Gaussian distribution function, $p_\sigma(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}$ where $\theta = \sigma$. Having samples $\{x^\alpha\} \sim p_{\sigma^*=1}$ we can obtain samples from another value of σ by a change of variables, $y^\alpha = \sigma x^\alpha$, $\{y^\alpha\} \sim p_\sigma$. If we apply reweighting with AD on the $\{x^\alpha\}$ samples, using an expansion around $\sigma^* = 1$ as $\tilde{\sigma} = \sigma^* + \varepsilon$, the weights become

$$\tilde{w}(\varepsilon) \propto \exp\left(-\frac{x^2}{2} \left(\frac{1}{(\sigma^* + \varepsilon)^2} - \frac{1}{(\sigma^*)^2}\right)\right). \quad (12)$$

However, if we apply reweighting, $\sigma^* = 1 \rightarrow \sigma$, after changing to the y -variables, the reweighting factors become trivially y -independent

$$\tilde{w}(\varepsilon) = \frac{p_{\tilde{\sigma}}(y)}{p_{\sigma^*}(y)} = \frac{1}{\sigma^* + \varepsilon} \quad (13)$$

thus being irrelevant in the computation of the expectation values. The fact that there is no effective reweighting is clear since we use the correctly sampled variables. To understand the connection with the Hamiltonian expansion, we start by writing the EOM for the first two orders

$$\ddot{x}_0 = -\frac{x_0}{\sigma^2}, \quad (14)$$

$$\ddot{x}_1 = -\frac{x_1}{\sigma^2} + 2\frac{x_0}{\sigma^3} \quad (15)$$

(here we suppressed the momentum variables in favour of a second order differential equation). While the zeroth order is an harmonic oscillator, the leading order equation corresponds to an undamped driven oscillator whose driving force has the same frequency as the oscillator⁴ – this leads to both orders having the same average value (over various trajectories), $\langle x_1 \rangle = \langle x_0 \rangle$. This can be seen in fig. 1 where the solutions of the EOM for the first two orders are shown to be similar.

⁴This also leads to a resonant phenomena, where the oscillation amplitude grows with the integration time. Although for any finite integration time it does not affect the central value of the expectation value, it may lead to large variance. The average trajectory length was optimized to reduce the variance and autocorrelations.

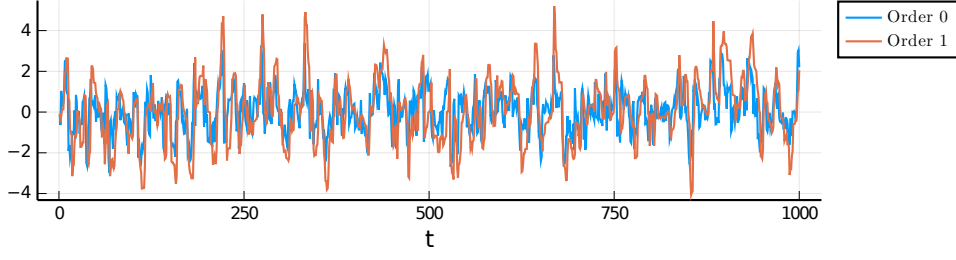


Figure 1: Solution of the equations eqs. (14) and (15) with a trajectory length tuned to reduce the variance and autocorrelations.

The key to understand the relation between reweighting, Hamiltonian expansion, and the change of variables lies in the truncated polynomial corresponding to y^α from the reparametrization, $y^\alpha = \sigma x^\alpha \rightarrow \tilde{y}^\alpha = x^\alpha + x^\alpha \varepsilon$. Notice how the first two orders are the same, exactly as in the solutions of the EOM above. We conclude that the EOM from the Hamiltonian approach find the transformation/change of variables that leads to constant reweighting factors, and thus the variance of the method is not affected by the disconnected contributions arising from the weighting factors.

2.2 Improved reweighting

Given that the conventional reweighting suffers from large variance, and knowing that the variance for the derivatives is dependent on the change of variables, we may attempt to improve the reweighting method by finding a suitable transformation. A similar approach, using a gradient flow transformation for QCD can be found in [9].

Let us again consider the m^2 -derivatives for the scalar theory. The $m^2 \rightarrow \tilde{m}^2$ reweighting factors are

$$\tilde{w}^\alpha = \exp\{-S(x^\alpha; \tilde{m}^2, \lambda) + S(x^\alpha; m^2, \lambda)\} \quad (16)$$

(here we consider only an expansion in m^2). If we first change variables to $\tilde{y} = x + f(x)\varepsilon$ with a unknown function f , the weights become

$$\tilde{w}^\alpha = \exp\{S(x^\alpha; m^2, \lambda) - S(\tilde{y}^\alpha; \tilde{m}^2, \lambda) - \log |\tilde{J}|\}, \quad (17)$$

where the last term comes from the Jacobian of the transformation. The aim is to find the a suitable transformation f to reduce the effect of reweighting.

We can find this transformation exactly for the free theory, whose lattice action in momentum space reads

$$S_{\text{latt}}(\phi_p; m) = \sum_p \phi_p^* \left[\sum_\mu \hat{p}_\mu^2 + m^2 \right] \phi_p, \quad \hat{p} = 2 \sin(ap/2). \quad (18)$$

The change of variables in momentum space becomes $\tilde{\varphi}_p = \phi_p + f_p \varepsilon$. The term ΔS in the reweighting factor

$$S(\phi_p; m^2) - S(\tilde{\varphi}_p; \tilde{m}^2) = -\varepsilon \left[(\hat{p}^2 + \tilde{m}^2)(\phi_p^* f_p + \phi_p f_p^*) + \phi_p^* \phi_p \right] \quad (19)$$

		λ				
		0.0	0.1	0.2	0.3	0.4
$\langle \phi^2 \rangle$	RW	-0.0428(20)	-0.0328(14)	-0.0270(13)	-0.0241(12)	-0.0220(11)
	TRW	-0.042528(30)	-0.03069(32)	-0.02604(37)	-0.02328(54)	-0.02079(56)
	HAD	-0.042526(41)	-0.030880(14)	-0.026273(10)	-0.0233672(82)	-0.0212721(72)
$\langle s \rangle$	RW	-0.0025(42)	-0.0006(34)	0.0027(35)	0.0028(36)	0.0057(32)
	TRW	$3(28) \times 10^{-19}$	0.00344(85)	0.0052(11)	0.0052(17)	0.0072(15)
	HAD	-0.000003(22)	0.002623(16)	0.004218(20)	0.005397(14)	0.006265(16)

Table 2: Derivatives with respect to \hat{m}^2 for $\langle \phi^2 \rangle$ and $\langle s \rangle$ from a L^4 lattice with $L/a = 32$ and $\hat{m}^2 = 0.05$. The derivatives were computed with conventional reweighting, transformed reweighting (TRW), and with the Hamiltonian expansion (HAD).

can be eliminated with the $f_p = -\frac{\phi_p}{2(\hat{p}^2+m^2)}$. With this choice the weights have the Jacobian contribution only which turns out to be ϕ -independent $w(\phi_p; m) = \prod_p -\frac{1}{2(\hat{p}^2+m^2)}$, and thus not affecting the reweighting.

In table 2 the m^2 -derivative from reweighting after the transformation is compared with the normal reweighting and the Hamiltonian approach. For the free theory we obtain the best results since the transformation is exact. In fact, for $\langle \phi^2 \rangle$ the precision of the Hamiltonian expansion is similar to the improved reweighting, which supports the idea that the former method finds the correct change of variables. Nonetheless, although the transformation is not exact for non-zero λ values, it still constitutes an improvement in relation with conventional reweighting for all available values of the coupling, with a decrease in precision for the largest values.

3. Conclusion

While automatic differentiation is a very important tool across many areas, there are no simple extensions when Monte Carlo sampling is involved in the definition of the function. These problems involve an underlying distribution function that depends on some parameters (couplings in the context of lattice field theory). The dependence on these parameters is very useful and usually difficult to determine. In this work we suggest two possibilities that tackle the problem by using truncated polynomials.

One of the methods is a straightforward modification of reweighting averages with the implementation of truncated polynomials. It has the advantage of being agnostic of the sampling algorithm, but may in general lead to large variance problems which are in general hard to predict.

The second method consists on a modification of Hamiltonian sampling algorithms. In this case we produce samples that carry the dependence on the parameters in the form of Taylor expansions. We showed (for HMC) that this method is able to reach a much higher precision with the same level of statistics. It is important to refer that the convergence of the equations of motion is model dependent. For the example considered in this work the convergence is guaranteed.

Both methods were tested on a 4-dimensional lattice scalar theory, where the dependence on the mass and quartic coupling was studied. Additionally, we have proposed that the Hamiltonian approach can be seen as a transformation to a basis where the reweighting factors are constant.

This sustains the claim that all parameter dependence is in the samples and that the absence of disconnected contributions is the reason behind lower variance in the second method.

This connection allowed us to improve the reweighting method by first transforming the variables. Although finding these transformations is, in general, a highly non-trivial task, this is a suitable problem for the application of Machine-Learning techniques, which can significantly improve the computation of derivatives of observables.

Acknowledgements

AR and GT acknowledge financial support from the Generalitat Valenciana (CIDEAGENT/2019/040). Similarly, BZ acknowledges the support from CIDEAGENT/2020/055. The authors gratefully acknowledge as well the support from the Ministerio de Ciencia e Innovacion (PID2020-113644GB-I00) and computer resources at Artemisa, funded by the European Union ERDF and Comunitat Valenciana as well as the technical support provided by the Instituto de Física Corpuscular, IFIC (CSIC-UV). The authors acknowledge the financial support from the MCIN with funding from the European Union NextGenerationEU (PRTR-C17.I01) and Generalitat Valenciana. Project “ARTEMISA”, ref. ASFAE/2022/024.

References

- [1] Catumba, G., Ramos, A. & Zaldivar, B., *Stochastic automatic differentiation for Monte Carlo processes*, [arXiv:2307.15406 [hep-lat]]
- [2] Haro, A., *Automatic differentiation tools in computational dynamical systems*, (2011,10)
- [3] Ramos, A. FormalSeries.jl., <https://igit.ific.uv.es/alramos/formalseries.jl>.
- [4] Duane, S., Kennedy, A., Pendleton, B. & Roweth, D., *Hybrid Monte Carlo*, *Physics Letters B*. **195**, 216-222 (1987)
- [5] Di Renzo, F., Marchesini, G., Marenzoni, P. & Onofri, E., *Lattice perturbation theory on the computer*, *Nuclear Physics B - Proceedings Supplements*. **34** pp. 795-798 (1994).
- [6] Brida, M. & Lüscher, M. SMD-based numerical stochastic perturbation theory. *Eur. Phys. J. C*. **77**, 308 (2017,5), <http://arxiv.org/abs/1703.04396>, arXiv:1703.04396 [hep-lat]
- [7] Brida, M., Garofalo, M. & Kennedy, A. Investigation of New Methods for Numerical Stochastic Perturbation Theory in φ^4 Theory. *Phys. Rev. D*. **96**, 054502 (2017,9), <http://arxiv.org/abs/1703.04406>, arXiv:1703.04406 [hep-lat]
- [8] Omelyan, I., Mryglod, I. & Folk, R., *Symplectic analytically integrable decomposition algorithms: classification, derivation, and application to molecular dynamics, quantum and celestial mechanics simulations*, *Computer Physics Communications*. **151**, 272-314 (2003)
- [9] Bacchio, S., *A novel approach for computing gradients of physical observables*, (2023), [arXiv:2305.07932 [hep-lat]].