Journal of Computational and Applied Mathematics 278 (2015) 101–109

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Summation through stochastic drawing of addends under steered morphing

Mirco Zerbetto, Diego Frezzato *

Dipartimento di Scienze Chimiche, Università degli Studi di Padova, via Marzolo 1, I-35131, Padova, Italy

HIGHLIGHTS

- A tool for stochastic evaluation of nested sums with many indexes is proposed.
- The Jarzynski equality is borrowed from the context of molecular thermodynamics.
- Addends build-up is seen as homologous of driven transformations in real systems.
- Numerical tests demonstrate the high performance of the method.

ARTICLE INFO

Article history:
Received 11 December 2012
Received in revised form 29 July 2014

Keywords:
Multiple summations
Stochastic sampling
Jarzynski's equality

ABSTRACT

A stochastic methodology for the numerical estimation of sums over a very large number of addends with huge spread of magnitudes is presented as continuation of our recent work in the field of multidimensional integration. The approach is based on the employment of Jarzynski's equality, borrowed from the physical context of thermodynamics of small systems (mainly macromolecular) subjected to driven transformations while all uncontrolled degrees of freedom freely fluctuate. An abstract interpretation of such an equality enables us to convert the sum into an exponential average over the "computational work" required to morph the addends from an initial set of values (taken all equal for simplicity) up to their actual values while the summation indexes are stochastically sampled by means of Importance Sampling Monte Carlo moves. A series of numerical tests reveals the high efficiency of the method in performing summations otherwise unfeasible.

1. Introduction

In a related work [1] we have presented a novel tool to perform multidimensional integration of bounded functions $f(\mathbf{x})$ in hyper-rectangles I, i.e., evaluation of $\int_I d\mathbf{x} f(\mathbf{x})$. The approach belongs to the category of "stochastic integration", but exploits an abstract interpretation-usage of Jarzynski's equality (JE in the following) which was derived in the thermodynamic context of free-energy-difference evaluations for (macro)molecular systems subjected to steered transformations along some coordinate(s) of interest [2,3]. Our starting point was the analogy between real deterministic transformations driven by an external apparatus, which controls some selected structural parameters of the system while all the other degrees of freedom are free to fluctuate (physical context), and guided morphing of the integrand function "while" the integration domain is stochastically explored (computational context). Such a parallelism has led us to "borrow" the JE from its original physical context and exploit it to evaluate the multidimensional integral from the amounts of "computational work" (i.e., the homologous of the proper work done by the external apparatus in real transformations) to perform the morphing of the

* Corresponding author. Tel.: +39 0498275658.
E-mail address: diego.frezzato@unipd.it (D. Frezzato).

http://dx.doi.org/10.1016/j.cam.2014.09.029
0377-0427/

integrand function in a series of repeated runs. The methodology was found to be very efficient in a series of tests, in the sense that high-dimensional integrations otherwise unfeasible could be performed with good accuracy in acceptable computational times.

In this communication we go forward in our project to treat a computational problem even farther from the original pertinence of the JE. Namely, we develop a strategy aimed to the efficient evaluation of sums $S = \sum_{i=1}^{n_{\text{adds}}} f_i$ over finite but possibly very large number of addends, n_{adds}. In all generality, each single term f_i, known or calculable on request, can be positive, negative, or null; the values can also show a very wide distribution of magnitudes. Although the approach here developed is rather general, we shall focus on a specific but widespread category of summations whose number of addends rapidly increases, namely nested sums of the kind

$$S = \sum_{i_1} \sum_{i_2} \sum_{i_3} \cdots \sum_{i_N} f_{\mathbf{i}} \tag{1}$$

where $\mathbf{i} = (i_1, i_2, \ldots, i_N)$ is the cumulative array of indexes which may run over different sets of entries. Also, N can be very large (several tens). We shall show that the cost of the computation can be highly lowered from the *a priori* power-law on N (which makes the calculation rapidly *unfeasible*) down to a roughly linear dependence on the number of indexes.

Here we anticipate the main lines of the strategy. The basic idea is to apply a suitable "growth" of the addends starting from the situation where they are all equal to 1, up to the final situation (morphed addends) where they take the actual values. "During" such a controlled morphing, the addends are stochastically drawn by means of Metropolis–Hastings Importance Sampling Monte Carlo (IS-MC in the following) moves [4–6]. The well-known IS-MC sampling scheme allows the random exploration of the ensemble of sites by means of moves whose acceptance depends on the relative weight ("stability") of the starting–arrival configurations. The term "site" will be used through the text as synonymous of a configuration $\mathbf{i}$. At the beginning, the moves are unbiased and almost all accepted due to the flat landscape of values, but, as the morphing distinguishes the terms, a "trajectory" visits preferentially the sites corresponding to most relevant addends and possibly remains entrapped in one of them (or within a subset of quasi-degenerate addends mutually reachable via IS-MC moves). Intuitively, the gradual growing of the addends, that is the gradual development of their spread of magnitudes, allows the jumping among sites to last on average for a larger number of steps (with respect to standard IS-MC sampling) before the entrapping becomes permanent. Up to here, the morphing would only allow a more effective exploration of the addends pattern. The further step consists in freely using the JE to get the sum from the exponential average (taken over a sufficiently large ensemble of runs) of the "computational work" of morphing (to be defined and evaluated along each IS-MC chain).

To make this presentation self-contained, a brief outlook of the JE in its proper physical context is due.[1] In the essence, the JE enables one to evaluate the ratio Z_2/Z_1 between the so-called configurational partition functions [8] of a physical system in two equilibrium states "1" and "2" at the same temperature but structurally distinguished by the values of a given set of parameters $\mathbf{\Lambda}$. Namely, the JE gives $Z_2/Z_1 = \langle \exp(-w/k_B T) \rangle$ where w is the amount of work that has to be done by an external apparatus to drive the system from $\mathbf{\Lambda}_1$ to $\mathbf{\Lambda}_2$ following a certain (but freely chosen) protocol $\mathbf{\Lambda}(t)$ starting from initial configurations picked at the equilibrium "1". T is the absolute temperature, and k_B is the Boltzmann constant. The control is exerted only on some parameters, while all other variables are free to fluctuate during the transformation. For this reason a distribution of path-dependent works is found, over which the above mentioned average is taken. To ensure the validity of the JE, the following conditions must be met: (i) any starting configuration must be sampled within the equilibrium probability at "1", (ii) the fluctuations on the uncontrolled variables are a Markovian (memory-less) process [9], and (iii) if the protocol is stopped at some time t^*, the system relaxes to the underlying equilibrium distribution corresponding to the set of structural parameters $\mathbf{\Lambda}(t^*)$. In practice, the above average is performed over a finite number of runs, that is $Z_2/Z_1 \simeq N_{\text{runs}}^{-1} \sum_{j=1}^{N_{\text{runs}}} \exp(-w_j/k_B T)$. The rate of convergence to the exact limit depends on the intrinsic nature of the system and on the kind of morphing protocol; however, it is normally found that a good estimate of Z_2/Z_1 can be obtained with a limited number of runs, so that such a non-equilibrium tool outperforms the much more demanding direct calculation (via integrations or summations) of the single Z_1 and Z_2. The interested Reader can find the presentation of the JE in excellent reviews of Refs. [3,10,11]; a clear derivation of the JE can be found in Section II of Ref. [12].

In the present computational context, by analogy, Z_1 and Z_2 become respectively the values of the summation on the initial non-morphed state and on the final morphed state of addends; the driven transformation becomes the chosen morphing of the addends; the fluctuations over the uncontrolled variables become the IS-MC stochastic jumps amongst the sites of the summation indexes. The "computational work" needs to be properly expressed as it will be shown in the next section. The elaboration of such a naïve picture is given what follows.

In the next section we illustrate the method, first starting from the case of positive-valued addends and then extending the approach to the general case of sums with positive, negative and null terms. Since most of the key-steps have been presented in Ref. [1] for the related problem of multidimensional integration, and since the adaptation is straightforward, here we give only the main lines. In Section 3 we test the approach on some case-studies (conceived in a way that the reference exact result can be easily obtained), and show that summations otherwise unfeasible via addend-by-addend

[1] The Reader could be interested in comparing the JE with the statistical tool known as "Annealed Importance Sampling" [7] when it is applied to evaluate normalizing constants of distribution functions [1].