Journal of Computational Physics 281 (2015) 365–374

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

Extra Chance Generalized Hybrid Monte Carlo ☆

Cédric M. Campos [a,*], J.M. Sanz-Serna [b]

[a] Dept. Matemática Aplicada e IMUVA, Fac. Ciencias, UVA, Paseo de Belén 7, 47011 Valladolid, Spain
[b] Dept. Matemáticas, Universidad Carlos III de Madrid, Edificio Sabatini, 28911 Leganés (Madrid), Spain

ARTICLE INFO

Article history:
Received 28 July 2014
Received in revised form 25 September 2014
Accepted 30 September 2014
Available online 7 October 2014

Keywords:
Sampling methods
Hybrid Monte Carlo
Detailed balance
Delayed rejection
Rejection avoidance
Molecular dynamics

ABSTRACT

We study a method, Extra Chance Generalized Hybrid Monte Carlo, to avoid rejections in the Hybrid Monte Carlo method and related algorithms. In the spirit of delayed rejection, whenever a rejection would occur, extra work is done to find a fresh proposal that, hopefully, may be accepted. We present experiments that clearly indicate that the additional work per sample carried out in the extra chance approach clearly pays in terms of the quality of the samples generated.

© 2014 Elsevier Inc. All rights reserved.

1. Introduction

In this paper we study a technique, Extra Chance Generalized Hybrid Monte Carlo (XCGHMC), to avoid rejections in the Hybrid Monte Carlo (HMC) method [8] and its variants (e.g. [3,9,15–17]). In the spirit of what in the statistics literature is called *delayed rejection* [12,20,32], whenever a rejection would occur, additional work is done to find a fresh proposal that, hopefully, may be accepted. Rejections, while essential to ensure that the algorithms sample from the right target probability distribution, contribute to an increase of the correlation of the samples [22,31]. Furthermore, for algorithms with partial momentum refreshments [15,17], each rejection necessarily requires a flip of the momentum and interferes with the underlying Hamiltonian dynamics. We present experiments that clearly indicate that the additional work per sample carried out in the extra chance approach in order to avoid rejections clearly pays in terms of the quality of the samples generated.

The algorithm studied here is mathematically equivalent to that suggested by Sohl-Dickstein, Mudigonda and DeWeese in [30]. However the actual formulas for the acceptance probability in the present work are different from those in [30]. We believe that the formulas used here provide insight into the probabilities involved. Furthermore we prove that the extra chance algorithm actually satisfies detailed balance/stochastic reversibility; this is surprising, as [30] suggests that the avoidance of rejections comes at the price of the violation of detailed balance. Detailed balance is of course a valuable property in the performance and analysis of Markov Chain Monte Carlo algorithms [24]. For instance the estimation of the effective sample size or autocorrelation time used in our experiments (taken from [10]) relies on the chain being reversible with respect to the target distribution.

☆ A MATLAB implementation of the algorithm and other utilitarian scripts can be found at http://github.com/vitaminace33/xhmc.

* Corresponding author.
E-mail addresses: cedricmc@uva.es, cedricmc@icmat.es (C.M. Campos), jmsanzserna@gmail.com (J.M. Sanz-Serna).

http://dx.doi.org/10.1016/j.jcp.2014.09.037
0021-9991/© 2014 Elsevier Inc. All rights reserved.

```
00  given z = (x, y) ∈ ℝ^d × ℝ^d
01  y ← y − (Δt/2)∇V(x)
02  for ℓ = 1 to L − 1
03      x ← x + ΔtM^{-1} y
04      y ← y − Δt∇V(x)
05  end for
06  x ← x + ΔtM^{-1} y
07  y ← y − (Δt/2)∇V(x)
08  (x′, y′) = (x, y)
```

Fig. 1. Velocity Verlet integrator for advancing $L \geq 1$ time-steps of length $\Delta t > 0$ starting from the initial point z and producing the final point $z' = \mathbf{I}(z)$.

In turn the basic idea in [30] is not essentially different from delayed rejection [12,20,32]. However some peculiarities of HMC (momentum flip, modified detailed balance vs. standard detailed balance, proposals being obtained via a deterministic flow) make it difficult, or even impossible, to apply the material of [12,20,32] to devise or analyze algorithms similar to those considered in [30] or here.

The literature is not lacking in suggestions to avoid rejection/momentum flips in HMC, see e.g. [1–3,29,33]. Comparisons between those techniques and the extra chance approach are not within our scope here. Neither shall we be concerned with comparing HMC with alternative sampling algorithms as done in [7].

In Section 2 we present the extra chance algorithm. Section 3 is devoted to an analysis of the acceptance probabilities. Proof of concept numerical experiments is reported in Section 4 and Section 5 concludes. Some more mathematical results are given in Appendices A–C.

2. Algorithm

The aim is to obtain samples x_n from a target probability distribution in the *state space* $\mathbb{R}^d$ with density

$$\frac{1}{Z}\exp\bigl(-\beta V(x)\bigr), \qquad Z = \int_{\mathbb{R}^d} \exp\bigl(-\beta V(x)\bigr)\,\mathrm{d}x. \tag{1}$$

The algorithms considered here do not require that Z be known; they merely need to evaluate V and its gradient ∇V. They generate a Markov chain $x_0 \to x_1 \to \cdots \to x_N$ [24] that has (1) as an invariant distribution, in such a way that under suitable ergodic behavior, it is possible to estimate averages $\langle A\rangle$ with respect to the target by taking means of the values of A along a realization of the chain:

$$\langle A\rangle = \frac{1}{Z}\int_{\mathbb{R}^d} A(x)\exp\bigl(-\beta V(x)\bigr)\,\mathrm{d}x \approx \frac{1}{N+1}\sum_{n=0}^{N} A(x_n). \tag{2}$$

The format in (1) implies that the density is everywhere positive. As shown in e.g. [9], it is not difficult to extend the algorithms and analysis to cases where the density vanishes in a subset of the state space.

Regardless of the details of the application in mind, HMC and its variants use the Hamiltonian formalism of classical mechanics. The components of x are interpreted as generalized co-ordinates describing the configuration of a mechanical system and an auxiliary d-dimensional vector y is introduced whose components represent the associated conjugated momenta. We shall use the notations $z = (x, y)$, $z' = (x', y')$, etc. to refer to points in the *phase space* $\mathbb{R}^d \times \mathbb{R}^d$. If M is a user-specified, symmetric positive-definite $d \times d$ mass matrix, the algorithms use the Hamiltonian function (total mechanical energy)

$$H(z) = \frac{1}{2}y^T M^{-1} y + V(x), \tag{3}$$

and the probability distribution in the phase space defined by the unnormalized density

$$\rho(z) = \exp\bigl(-\beta H(z)\bigr) = \exp\bigl(-(\beta/2)y^T M^{-1} y\bigr) \times \exp\bigl(-\beta V(x)\bigr). \tag{4}$$

The product structure of ρ implies that x and y are stochastically independent; x is distributed according to the target (1) and $y \sim \mathcal{N}(0, M)$, i.e. y is Gaussian with zero mean and covariance matrix M.

The dynamics associated with (3) is given by

$$\frac{\mathrm{d}}{\mathrm{d}t}x = M^{-1}y, \qquad \frac{\mathrm{d}}{\mathrm{d}t}y = -\nabla V(x), \tag{5}$$

a system of differential equations whose solution flow exactly preserves the distribution (4), see e.g. [26]. In practice this flow cannot be computed in closed form and one has to resort to numerical approximations; the Störmer–Verlet/leapfrog integrator [13,18,27,28] is the method of choice. Fig. 1 shows pseudocode for computing the result $z' = (x', y') = \mathbf{I}(z)$ of L integration time-steps of length $\Delta t > 0$ starting from the initial point z. The transformation $\mathbf{I}$, which maps the phase space into itself, is both *volume preserving* (i.e. has unit Jacobian determinant) and *reversible*. Reversibility means that, for each z',

Download English Version:

https://daneshyari.com/en/article/6932023

Download Persian Version:

https://daneshyari.com/article/6932023

Daneshyari.com