SGNHT Relativistic HMC version 2

Xiaoyu Lu

March 2016

Consider the Hamiltonian

$$H(q,p,\zeta) = U(q) + K(p) + \frac{d}{2}(\zeta - A)^2$$

where A is the injected noise. Suppose the gradient of the kinetic energy can be represented by $\nabla_p K(p) = M^{-1}(p)p$ where $M(p)$ is a matrix or scalar which can depend on the momentum p, we assume it to be symmetric for simplicity. In the Relativistic Hamiltonian, we have $K(p) = mc^2 \left(\frac{p^T p}{m^2c^2} + 1\right)^{\frac{1}{2}}$ and $\nabla_p K(p) = M^{-1}(p)p$ with $M(p) := m\left(\frac{p^T p}{m^2c^2} + 1\right)^{\frac{1}{2}}$ a scalar, and in the Newtonian Hamiltonian, we have $K(p) = \frac{p^T M^{-1} p}{2}$ with $\nabla_p K(p) = M^{-1}p$ where $M(p)$ is simply the mass matrix M.

We have

$$\nabla H = \begin{pmatrix} \nabla U(q) \\ M^{-1}(p)p \\ d(\zeta - A) \end{pmatrix}$$

Under the Ma/Chen/Fox framework, we define

$$D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & A & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad Q = \begin{pmatrix} 0 & -I & 0 \\ I & 0 & \frac{1}{d}M^{-1}(p)p \\ 0 & -\frac{1}{d}p^T M^{-1}(p) & 0 \end{pmatrix}, \quad \text{hence} \quad \Gamma = \begin{pmatrix} 0 \\ 0 \\ -tr\left(\frac{d}{dp}\left(\frac{1}{d}M^{-1}(p)p\right)\right) \end{pmatrix}$$

Hence

$$-(D+Q)\nabla H + \Gamma = \begin{pmatrix} 0 & I & 0 \\ -I & -A & -\frac{1}{d}M^{-1}(p)p \\ 0 & \frac{1}{d}p^T M^{-1}(p) & 0 \end{pmatrix} \begin{pmatrix} \nabla U(q) \\ M^{-1}(p)p \\ d(\zeta - A) \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ -tr\left(\frac{d}{dp}\left(\frac{1}{d}M^{-1}(p)p\right)\right) \end{pmatrix}$$

leading to the following SDE:

$$d\begin{pmatrix} q \\ p \\ \zeta \end{pmatrix} = \begin{pmatrix} M^{-1}(p)p \\ -\nabla\tilde{U} - \zeta M^{-1}(p)p \\ \frac{1}{d}p^T M^{-2}(p)p - tr\left(\frac{d}{dp}\left(\frac{1}{d}M^{-1}(p)p\right)\right) \end{pmatrix} dt + \begin{pmatrix} 0 \\ \sqrt{2A - Bdt} \\ 0 \end{pmatrix} dW_t$$

where B is the noise from the stochastic gradient.

In the case of relativistic HMC, we have

$$d\begin{pmatrix} q \\ p \\ \zeta \end{pmatrix} = \begin{pmatrix} M^{-1}(p)p \\ -\nabla\tilde{U} - \zeta M^{-1}(p)p \\ \frac{p^T p}{d}\left(M^{-2}(p) + c^{-2}M^{-3}(p)\right) - M^{-1}(p) \end{pmatrix} dt + \begin{pmatrix} 0 \\ \sqrt{2A - Bdt} \\ 0 \end{pmatrix} dW_t$$