A Fast Gibbs-like Symplectic Integrator for Hamiltonian Monte Carlo

Melissa Malcom

Abstract

This poster compares the convergence of Gibbs and HMC for Bayesian hierarchical models. The Hamiltonian dynamics in HMC is approximated by a Gibbs-like symplectic integrator adapted to the structure of hierarchical models. This integrator allows larger time step sizes than Verlet, which in turn, accelerates convergence of HMC.

Definitions

Gibbs was used with a HMC integrator based on the combination of Gibbs and HMC by Neal (2011). HMC was used to integrate the second component within Gibbs.

HMC was used with a Gibbs-like symplectic integrator based on the J-splitting method proposed by Bou-Rabee and Vanden-Eijnden (2012).

Funnel Distributions

\[
U(q_1, q_2) = \frac{q_1^2}{2\alpha_1^2} + \frac{q_2^2}{2\alpha_2^2} + \frac{q_1^2}{2\alpha_2} \tag{1}
\]

\[
U(q_1, q_2, q_3) = \frac{q_1^2}{2\alpha_1^2} + \frac{q_2^2}{2\alpha_2^2} + \frac{q_1^2}{2\alpha_2} - \rho q_2 q_3 \tag{2}
\]

Fig. 1: Hamiltonian dynamics for the bivariate funnel distribution This figure shows the solution of \(\dot{q}_1(t) = -\frac{\partial U(q_1(t), q_2(t))}{\partial q_1} \), \(\dot{q}_2(t) = -\frac{\partial U(q_1(t), q_2(t))}{\partial q_2} \) over the interval [0, 8] with initial conditions \(q_1(0) = 0.1 \), \(q_2(0) = q_1(0) = q_2(0) = 0 \). Several contour lines of \(U\) are plotted in solid black in the right panel. The color along the trajectory is related to \(t\): from blue (small \(t\)) through green and yellow to red (large \(t\)). Note that in the neck of the funnel the first component of the solution \(q_1(t)\) changes rapidly.

\[
\dot{q}_1(t) = -\frac{\partial U(q_1(t), q_2(t), q_3(t))}{\partial q_1} = -\frac{\partial U(q_1(t), q_2(t))}{\partial q_1} \tag{3}
\]

\[
\dot{q}_2(t) = -\frac{\partial U(q_1(t), q_2(t), q_3(t))}{\partial q_2} = -\frac{\partial U(q_1(t), q_2(t))}{\partial q_2} \tag{4}
\]

\[
\dot{q}_3(t) = -\frac{\partial U(q_1(t), q_2(t), q_3(t))}{\partial q_3} \tag{5}
\]

\[
\dot{q}_1(t) = -\frac{\partial U(q_1(t), q_2(t), q_3(t))}{\partial q_1} \tag{6}
\]

\[
\dot{q}_2(t) = -\frac{\partial U(q_1(t), q_2(t), q_3(t))}{\partial q_2} \tag{7}
\]

\[
\dot{q}_3(t) = -\frac{\partial U(q_1(t), q_2(t), q_3(t))}{\partial q_3} \tag{8}
\]

Fig. 2: Hamiltonian dynamics for the multivariable funnel distribution This figure shows the solution of \(\dot{q}_1(t) = -\frac{\partial U(q_1(t), q_2(t), q_3(t))}{\partial q_1} \), \(\dot{q}_2(t) = -\frac{\partial U(q_1(t), q_2(t), q_3(t))}{\partial q_2} \), \(\dot{q}_3(t) = -\frac{\partial U(q_1(t), q_2(t), q_3(t))}{\partial q_3} \) over the interval [0, 4] with initial conditions \(q_1(0) = 0.1 \), \(q_2(0) = q_3(0) = 1.0 \), \(q_1(0) = q_2(0) = q_3(0) = 0 \). Several contour lines of \(U\) in two dimensions are plotted in solid black in the middle panel. The sample path in three dimensions is plotted in the right panel. The color along the trajectory is related to \(t\): from blue (small \(t\)) through green and yellow to red (large \(t\)).

Fig. 3: Time step selection This figure displays the acceptance probabilities for Gibbs and HMC. The duration was fixed at \(\lambda = 8.0\) for the bivariate funnel and \(\lambda = 4.0\) for the multivariable funnel based on the sample paths. For the bivariate funnel, a time step of \(h = 0.5\) was chosen for Gibbs and a time step of \(h = 2.0\) was chosen for HMC. For the multivariable funnel, a time step of \(h = 0.08\) was chosen for Gibbs and a time step of \(h = 0.27\) was chosen for HMC.

Fig. 4: Autocorrelation time for the bivariate funnel This figure shows the autocorrelation time for a time step chosen with 99% acceptance probability, \(h = 0.1\), and a duration within the interval [2, 50]. The number of force evaluations for Gibbs was 100 samples and the number of force evaluations for HMC was 13 samples. Fixing the number of force evaluations considered the computational costs.

Fig. 5: Autocorrelation time for the multivariable funnel This figure shows the autocorrelation time for a time step chosen with 99% acceptance probability, \(h = 0.1\), and a duration within the interval [2, 50]. The number of force evaluations for Gibbs was 100 samples and the number of force evaluations for HMC was 16 samples. Fixing the number of force evaluations considered the computational costs.

Conclusions

• The proposed Gibbs-like symplectic integrator uses the structure of the hierarchical models to sample the funnel distributions.

• HMC with a Gibbs-like symplectic integrator can take larger time step sizes in comparison to Gibbs.

• HMC with a Gibbs-like symplectic integrator has smaller autocorrelation time for the correlated funnel distribution given equal computational costs to Gibbs.

Results

Fig. 6: Gibbs and HMC for the bivariate funnel The bivariate funnel distribution is simulated using Gibbs and HMC. The duration and time step for Gibbs, \(h = 0.50\) and \(\lambda = 8.00\), has 99.98% acceptance probability. The duration and time step for HMC, \(h = 2.00\) and \(\lambda = 8.00\), has 99.97% acceptance probability. A sample size of 1,000,000 was used to generate the marginal distributions for both figures and a sample size of 10,000 was used to generate the scatter plots for both figures.

Fig. 7: Gibbs and HMC for the multivariable funnel The multivariable funnel distribution is simulated using Gibbs and HMC. The duration and time step for Gibbs, \(h = 0.50\) and \(\lambda = 8.00\), has 99.13% acceptance probability. The duration and time step for HMC, \(h = 0.27\) and \(\lambda = 4.00\), has 99.08% acceptance probability. A sample size of 1,000,000 was used to generate the marginal distributions for both figures and a sample size of 10,000 was used to generate the scatter plots for both figures.

Acknowledgements

A special thank you to the Rutgers University – Camden Department of Mathematical Sciences, The National Science Foundation Grant (DMS-1816378), and Dr. Nawaf Bou-Rabee (Rutgers University – Camden Department of Mathematical Sciences).

References


