SM1. Metropolis-adjusted sampling algorithms.

SM1.1. Metropolis-adjusted BAOAB. The BAOAB update is given as follows:

\[(SM1)\quad x^* = x_0 - (1 + c)\frac{\epsilon^2}{4} \nabla U(x_0) + (1 + c)\frac{\epsilon}{2} u_0 + \frac{\epsilon}{2} \sqrt{1 - c^2} Z_0,\]

\[(SM2)\quad u^* = cu_0 - \frac{\epsilon}{2} c \nabla U(x_0) - \frac{\epsilon}{2} \nabla U(x^*) + \sqrt{1 - c^2} Z_0,\]

where \(Z_0 \sim \mathcal{N}(0, 1)\), \(\eta \geq 0\) is a friction coefficient, and \(c = e^{-\eta \epsilon}\).

We first derive the acceptance probability for using BAOAB as a proposal scheme in the framework of generalized Metropolis–Hastings sampling (Song and Tan, 2021). Using (SM2), the noise \(Z_0\) can be expressed as

\[Z_0 = (1 - c^2)^{-1/2} \left[ u^* - cu_0 + \frac{\epsilon}{2} c \nabla U(x_0) + \frac{\epsilon}{2} \nabla U(x^*) \right].\]

Suppose that the same mapping is applied backward (with reversed momentum) from \((x^*, -u^*)\) to \((x_0, -u_0)\) using another noise \(-Z^*\). Then (SM2) with \((x_0, u_0, Z_0)\) exchanged with \((x^*, -u^*, -Z^*)\) shows that

\[(SM3)\quad Z^* = -(1 - c^2)^{-1/2} \left[ cu^* - u_0 + \frac{\epsilon}{2} c \nabla U(x_0) + \frac{\epsilon}{2} c \nabla U(x^*) \right].\]

Remarkably, it can be verified by direct calculation that (SM1) is also satisfied with \((x_0, u_0, Z_0)\) exchanged with \((x^*, -u^*, -Z^*)\):

\[x_0 = x^* - (1 + c)\frac{\epsilon^2}{4} \nabla U(x^*) - (1 + c)\frac{\epsilon}{2} u^* - \frac{\epsilon}{2} \sqrt{1 - c^2} Z^*.\]

Then the mapping \(\Phi\) from \((x_0, u_0, Z_0)\) to \((x^*, u^*, Z^*)\) given by (SM1)-(SM3) satisfies the following generalized reversibility:

forward: \[\begin{pmatrix} x_0 \\ u_0 \\ Z_0 \end{pmatrix} \xrightarrow{\Psi} \begin{pmatrix} x^* \\ u^* \\ Z^* \end{pmatrix},\]

backward: \[\begin{pmatrix} x^* \\ -u^* \\ -Z^* \end{pmatrix} \xrightarrow{\Psi} \begin{pmatrix} x_0 \\ -u_0 \\ -Z_0 \end{pmatrix}.\]

The forward and backward proposal densities are respectively,

\[Q(x^*, u^*|x_0, u_0) \propto \exp \left(-\frac{1}{2} Z_0^T Z_0\right),\]

\[Q(x_0, -u_0|x^*, -u^*) \propto \exp \left(-\frac{1}{2} Z^T Z^*\right).\]
By generalized Metropolis–Hastings sampling, we set \((x_1, u_1) = (x^*, u^*)\) with probability \(\alpha = \min(1, r)\) or \((x_1, u_1) = (x_0, -u_0)\) with the remaining probability, where
\[
r = \frac{\pi(x^*, -u^*)Q(x_0, -u_0|x^*, -u^*)}{\pi(x_0, u_0)Q(x^*, u^*|x_0, u_0)}
\]
\[
= \exp \left\{ H(x_0, u_0) - H(x^*, u^*) + \frac{1}{2}Z_0^TZ_0 - \frac{1}{2}Z^T\gamma Z^* \right\},
\]
(SM4) \[
= \exp \left\{ -[G(x^*, u^*, Z^*) - G(x_0, u_0, Z_0)] \right\} = \exp\{-\Delta G\}.
\]
To make \(\exp(-G)\) a proper density function, we define
\[
G(x, u, Z) = H(x, u) + \frac{1}{2}Z^TZ + \log(2\pi)^k.
\]
The expression for \(\Delta G\) can be calculated as
\[
\Delta G = U(x^*) - U(x_0) - \left[ \frac{\epsilon}{2}u^* + \frac{\epsilon^2}{8}\nabla U(x^*) \right]^T\nabla U(x_0)
\]
(SM5)
\[
= \left[ \frac{\epsilon}{2}u_0 - \frac{\epsilon^2}{8}\nabla U(x_0) \right]^T\nabla U(x_0).
\]
Next, we extend Proposition 4 to Metropolized BAOAB. The transition defined in (SM1)–(SM3), along with the acceptance probability (SM4), satisfy all the conditions used in the proof of Lemma 1. Hence when the chain is stationary, the expected acceptance rate is
(SM6)
\[E[\alpha] = 2P[\Delta G < 0],\]
where the \(P[\Delta = 0]\) term is always 0. Consider univariate Gaussian target density
\[\pi(x) = \mathcal{N}(0, \gamma^{-1}).\]
Then \(\Delta G\) in (SM5) evaluated at any \((x, u, Z)\) can be simplified as
(SM7)
\[\Delta G(x, u, Z) = \frac{\gamma \epsilon^2}{128}B_1 \cdot B_2,\]
where
\[
B_1 = (2 + 2c)u - \gamma(1 + c)x + 2\sqrt{1-c^2}Z,
\]
\[
B_2 = [8(1 - c) + 2\gamma c^2(1 + c)]u - \gamma^2 c^3 - 4\gamma c)(1 + c)x + 2\sqrt{1-c^2}(\gamma c^2 - 4)Z.
\]
In stationarity, \(x \sim \mathcal{N}(0, \gamma^{-1}), u \sim \mathcal{N}(0, 1), Z \sim \mathcal{N}(0, 1)\) independently. Therefore \((B_1, B_2)\) are bivariate normal with mean 0 and variance matrix
\[
\left( \begin{array}{c}
(1+c)(8 + (1+c)\gamma c^2) \\
(1+c)[4 - 4c + (1+c)\gamma c^2]
\end{array} \right)
\]
\[
\left( \begin{array}{c}
[4 - 4c + (1+c)\gamma c^2][32 + ((1+c)\gamma c^2 - 4 - 4c)\gamma c^2]\end{array} \right).
\]
The correlation coefficient between \(B_1, B_2\) is
\[
\rho = \frac{(1+c)[4 - 4c + (1+c)\gamma c^2]\gamma c^2}{\sqrt{(1+c)[8 + (1+c)\gamma c^2][4 - 4c + (1+c)\gamma c^2][32 + (1+c)(\gamma c^2 - 4)\gamma c^2]}},
\]
Because \(\rho > 0\), using Lemma SM1 (Section SM3.7), we have
\[
P[\Delta G < 0] = P[B_1B_2 < 0] = \frac{1}{2} - \frac{1}{\pi} \arcsin(\rho).
\]
Combining this with (SM6) shows that the expected acceptance rate is

\[ E[\alpha] = 2P[\Delta G < 0] = 1 - \frac{2}{\pi} \arcsin(\rho). \]  

(SM8)

Notice that \( E[\Delta G] = \gamma \epsilon^2 128 E[ B_1 B_2]. \) By direct calculation using the expressions from the variance matrix of \((B_1, B_2),\) it can be shown that

\[ \rho^2 = \frac{E^2[B_1 B_2]}{\text{Var}(B_1) \text{Var}(B_2)} = \frac{E[\Delta G]}{2 + E[\Delta G]} \]  

(SM9)

Because \( \rho \) is always positive, this implies

\[ \rho = \sqrt{\frac{E[\Delta G]}{2 + E[\Delta G]}}. \]

By this relation, \( E[\alpha] \) in (SM8) can be expressed in terms of \( E[\Delta G] \) as

\[ E[\alpha] = 1 - \frac{2}{\pi} \arctan \left( \sqrt{\frac{E[\Delta G]}{2 + E[\Delta G]}} \right). \]  

(SM10)

Letting \( c = e^{-\eta \epsilon} \) and taking a series expansion of (SM10) in \( \epsilon, \) we find that

\[ E[\alpha] = 1 - \frac{\sqrt{2}}{4\pi} \gamma \sqrt{\eta \epsilon^{5/2}} + O(\epsilon^{7/2}), \]

as stated in Corollary 4.

Finally, we prove the acceptance rate result for Metropolized BAOAB in Corollary 6. In the product Gaussian case, we have based on (SM7),

\[ E[\Delta G] = \sum_{i=1}^{d} E[\Delta G_i] = \frac{c^4(1 - c^2)}{32} \sum_{i=1}^{d} \gamma_i^2 + \frac{\epsilon^6(1 + c^2)}{128} \sum_{i=1}^{d} \gamma_i^3. \]

Recall that \( c = e^{-\eta \epsilon}, \) hence \( c^4(1 - c^2) = 2\eta \epsilon^5 + O(\epsilon^6) \) and \( \epsilon^6(1 + c^2) = O(\epsilon^6). \)

This observation gives rise to the assumption in Corollary 6. Notice that we require \( \epsilon = O(d^{-1/5}) \) because \( \eta > 0 \) (equivalently \( c < 1 \)). If \( \eta = 0 \) (or \( c = 1 \)) then this requirement can be relaxed, but this situation corresponds to a purely deterministic update and is ignored in our discussion. We now apply Lemma 2 to prove the result in Corollary 6. Conditions (40) and (41) are satisfied by the assumption in Corollary 6, where

\[ \mu = \lim_{d \to \infty} E[\Delta G] = \lim_{d \to \infty} \left( \frac{\eta \epsilon^5}{16} d \right) \lim_{d \to \infty} \left\{ \frac{1}{d} \sum_{i=1}^{d} \gamma_i^2 \right\} = m\tau. \]

According to (SM9), Metropolized BAOAB satisfies condition (SM47), so that Lemma SM2 applies. Consequently (39) also holds and the proof is complete.

**SM1.2. Metropolis-adjusted ABOBA.** The ABOBA update can be stated as

\[ \hat{x} = x_0 + \frac{\epsilon}{2} u_0, \quad x^* = x_0 - (1 + c) \frac{\epsilon^2}{4} \nabla U(\hat{x}) + (1 + c) \frac{\epsilon}{2} u_0 + \frac{\epsilon}{2} \sqrt{1 - c^2} Z_0, \]  

(SM11)

\[ u^* = cu_0 - (1 + c) \frac{\epsilon}{2} \nabla U(\hat{x}) + \sqrt{1 - c^2} Z_0, \]  

(SM12)

where \( Z_0 \sim \mathcal{N}(0, 1), \) \( \eta \geq 0 \) is a friction coefficient, and \( c = e^{-\eta \epsilon}. \)
We first derive the acceptance probability for using ABOBA as a proposal scheme in generalized Metropolis-Hastings sampling. The process is similar to that in Section SM1.1 for BAOAB. Using (SM12), the noise \( Z_0 \) can be expressed as

\[
Z_0 = (1 - c^2)^{-1/2} \left[ u^* - cu_0 + (1 + c)\frac{\epsilon}{2} \nabla U(\hat{x}) \right].
\]

Then the noise \( Z^* \) required for the backward version of (SM12) after momentum reversal is

\[
Z^* = -(1 - c^2)^{-1/2} \left[ cu^* - u_0 + (1 + c)\frac{\epsilon}{2} \nabla U(\hat{x}) \right].
\]

The gradients in \( Z_0 \) and \( Z^* \) are evaluated at the same value because \( x_0 + \frac{\epsilon}{2} u_0 = x^* - \frac{\epsilon}{2} u^* = \hat{x} \). While the \( Z^* \) expression is derived by the \( u \) update, it can be verified by direct calculation that \( Z^* \) also gives the backward version of (SM11) from \( x^* \) to \( x_0 \):

\[
x_0 = x^* - (1 + c)\frac{\epsilon^2}{4} \nabla U(\hat{x}) - (1 + c)\frac{\epsilon}{2} u^* - \frac{\epsilon}{2} \sqrt{1 - c^2} Z^*.
\]

The acceptance probability is then \( \alpha = \min(1, r) \) where \( r \) has the same form as (SM4). For ABOBA, \( \Delta G \) can be calculated as

\[
\Delta G = U(x^*) - U(x_0) - \frac{\epsilon}{2} [u^* + u_0]^T \nabla U(\hat{x}).
\]

Next, we extend Proposition 4 to Metropolized ABOBA. The transition defined in (SM11)–(SM13), along with the acceptance probability (SM14), satisfy all the conditions used in the proof of Lemma 1. Hence when the chain is stationary, the expected acceptance rate is

\[
E[\alpha] = 2P[\Delta G < 0].
\]

Consider univariate Gaussian target density \( \pi(x) = \mathcal{N}(0, \gamma^{-1}) \). Then \( \Delta G \) in (SM14) evaluated at any \((x, u, Z)\) can be simplified as

\[
\Delta G(x, u, Z) = \frac{\gamma \epsilon^2}{128} B_1 : B_2,
\]

where

\[
B_1 = (1 + c)(4 - \gamma \epsilon^2)u - 2\gamma \epsilon(1 + c)x + 4\sqrt{1 - \epsilon^2}Z,
\]

\[
B_2 = (4c - 4 - \gamma \epsilon^2(1 + c))u - 2\gamma \epsilon(1 + c)x + 4\sqrt{1 - \epsilon^2}Z.
\]

In stationarity, \( x \sim \mathcal{N}(0, \gamma^{-1}), u \sim \mathcal{N}(0, 1), Z \sim \mathcal{N}(0, 1) \) independently. Therefore, \((B_1, B_2)\) are bivariate normal with mean 0 and variance matrix

\[
\begin{pmatrix}
(1 + c)[32 + (1 + c)\gamma^2(\gamma^2 - 4)] & (1 + c)[4 - 4c + (1 + c)\gamma^2]\gamma^2 \\
(1 + c)[4 - 4c + (1 + c)\gamma^2]\gamma^2 & [8 + (1 + c)\gamma^2][4 - 4c + (1 + c)\gamma^2]
\end{pmatrix}.
\]

The correlation coefficient between \( B_1, B_2 \) is given by

\[
\rho = \frac{(1 + c)[4 - 4c + (1 + c)\gamma^2]\gamma^2}{\sqrt{(1 + c)[8 + (1 + c)\gamma^2][4 - 4c + (1 + c)\gamma^2][32 + (1 + c)(\gamma^2 - 4)\gamma^2]}}.
\]
Notice that $\Delta G$ is of the same product form as $\Delta G$ in (SM7) for BAOAB, with the same coefficient $\frac{\gamma^2}{128}$. Moreover, the expression of the correlation coefficient $\rho$ and $E[B_1B_2]$ are also the same as those for BAOAB. Hence all the calculation in BAOAB directly applies to ABOBA, and the expected acceptance rate for ABOBA is

$$E[\alpha] = 1 - \frac{2}{\pi} \arctan \left( \sqrt{\frac{E[\Delta G]}{2}} \right) = 1 - \frac{\sqrt{2}}{4\pi} \gamma \eta^{c/2} + O(\epsilon^{7/2}),$$

as stated in Corollary 4.

Finally, because ABOBA agrees with BAOAB in $\rho$ and $E[B_1B_2]$ in the univariate Gaussian case, the acceptance rate result holds for Metropolized ABOBA in the same way as Metropolized BAOAB under a product Gaussian target as stated in Corollary 6.

**SM1.3. Metropolis-adjusted OBABO.** The OBABO proposal can be written as

$$x^* = x_0 - \frac{\epsilon^2}{2} \nabla U(x_0) + \epsilon \sqrt{c} u_0 + \epsilon \sqrt{1 - c} Z_0^{(1)},$$

$$u^* = cu_0 - \frac{\epsilon \sqrt{c}}{2} (\nabla U(x_0) + \nabla U(x^*)) + \sqrt{c(1 - c)} Z_0^{(1)} + \sqrt{1 - c} Z_0^{(2)}.$$

where $c = e^{-\eta \epsilon}$ and $Z_0^{(1)}, Z_0^{(2)} \sim \mathcal{N}(0, I)$ independently. Define

$$Z^{(1)*} = - (1 - c)^{-\frac{1}{2}} \left( \frac{x_0 - x^*}{\epsilon} + \frac{\epsilon \sqrt{c}}{2} \nabla U(x^*) + \sqrt{c} u^* \right),$$

$$Z^{(2)*} = - (1 - c)^{-\frac{1}{2}} \left( \sqrt{c} (x^* - x_0) \frac{\epsilon}{\sqrt{c}} \nabla U(x_0) - \sqrt{c} u_0 \right),$$

$$Z^* = ( (Z^{(1)*})^T, (Z^{(2)*})^T )^T, \quad Z_0 = ( (Z_0^{(1)})^T, (Z_0^{(2)})^T )^T.$$

Then the mapping from $(x_0, u_0, Z_0)$ to $(x^*, u^*, Z^*)$ satisfies the generalized reversibility. The acceptance probability is $\alpha = \min(1, r)$, where $r$ has the same form as (SM4) with $G$ given by

$$G(x, u, Z) = H(x, u) + \frac{1}{2} Z^T Z + \log(2\pi)^{\frac{1}{2}k}.$$

In this case, $\Delta G$ can be calculated as

$$\Delta G = U(x^*) - U(x_0) - \frac{(x^* - x_0)^T}{2} [\nabla U(x^*) + \nabla U(x_0)]$$

$$+ \frac{\epsilon^2}{8} [\nabla U(x^*)^T \nabla U(x^*) - \nabla U(x_0)^T \nabla U(x_0)].$$

Detailed calculation and discussions on the validity of the OBABO acceptance probability are included in Song and Tan (2021), where OBABO is referred to as UDL.

Next, we extend Proposition 4 to Metropolized OBABO. The transition defined in (SM15) through (SM17), along with the acceptance probability satisfy all the conditions used in the proof of Lemma 1. Hence when the chain is stationary, the expected acceptance rate is

$$E[\alpha] = 2P[\Delta G < 0],$$
as in Lemma 1. Consider univariate Gaussian target density \( \pi(x) = \mathcal{N}(0, \gamma^{-1}) \). Then \( \Delta G \) in (SM18) evaluated at any \((x, u, Z)\) can be simplified as

\[
\Delta G(x, u, Z) = \frac{\gamma^2 \epsilon^3}{32} B_1 \cdot B_2,
\]

where

\[
B_1 = 2\sqrt{c u} - \gamma \epsilon x + 2\sqrt{1-cZ^{(1)}},
B_2 = 2\sqrt{c \epsilon} - (\gamma \epsilon^2 - 4)x + 2\sqrt{1-c\epsilon Z^{(1)}}.
\]

Notice that \( \Delta G(x, u, Z) \) does not contain \( Z^{(2)} \). In stationarity, \( x \sim \mathcal{N}(0, \gamma^{-1}) \), \( u \sim \mathcal{N}(0, 1) \), \( Z^{(1)} \sim \mathcal{N}(0, 1) \) independently. Therefore, \((B_1, B_2)\) are bivariate normal with

\[
\begin{pmatrix} B_1 \\ B_2 \end{pmatrix} \sim \mathcal{N} \left( 0, \begin{pmatrix} 4 + \gamma \epsilon^2 & 16 \gamma \epsilon^3 \\ \gamma \epsilon^3 & 16 \gamma \epsilon^3 \end{pmatrix} \right).
\]

The correlation coefficient between \( B_1, B_2 \) is

\[
\rho = \sqrt{\frac{\gamma^3 \epsilon^6}{64 + \gamma^3 \epsilon^6}}.
\]

Because \( \rho > 0 \), using Lemma SM1 (Section SM3.7), we have

\[
P[\Delta G < 0] = P[B_1 B_2 < 0] = \frac{1}{2} - \frac{1}{\pi} \arcsin(\rho).
\]

Combining this with Lemma 1 shows that the expected acceptance is then

\[
E[\alpha] = 2P[\Delta G < 0] = 1 - \frac{2}{\pi} \arcsin(\rho).
\]

Notice that

\[
E[\Delta G] = \frac{\gamma^2 \epsilon^3}{32} E[B_1 B_2] = \frac{\gamma^3 \epsilon^6}{32}.
\]

Then apparently,

\[
\rho^2 = \frac{E[\Delta G]}{2 + E[\Delta G]} = \frac{\gamma^3 \epsilon^6}{64 + \gamma^3 \epsilon^6}.
\]

With the relation above, the expected acceptance rate can be expressed as

\[
E[\alpha] = 1 - \frac{2}{\pi} \arctan \left( \sqrt{\frac{E[\Delta G]}{2}} \right),
\]

which does not depend on \( c \) or \( \eta \) because \( E[\Delta G] \) does not. Moreover, \( E[\alpha] \) can be expanded as

\[
E[\alpha] = 1 - \frac{\gamma^{3/2}}{4\pi} \epsilon^3 + O(\epsilon^9),
\]

as stated in Corollary 4.

Finally, under a product Gaussian target, the acceptance rate result holds for Metropolized OBABO in Corollary 6. The proof uses Lemma 2. It is easy to verify that conditions (40) and (41) are satisfied by the assumption in Corollary 6. Moreover, because of the relation (SM19), Lemma SM2 also applies, which leads to (39) and hence completes the proof.
**SM2. Preconditioned sampling algorithms.** We present preconditioned versions of Metropolis-adjusted BAOAB, ABOBA, and OBABO algorithms, in addition to the preconditioned HAMS-A/B taken from Song and Tan (2021). These algorithms are used in our numerical experiments.

**Algorithm 1.** Preconditioned HAMS-A/HAMS-B

Initialize $x_0, u_0, \hat{x}_0 = L^* x_0$ and $\nabla U(\hat{x}_0) = L^{-1} \nabla U(x_0)$.

for $t = 0, 1, 2, \ldots, N_{\text{iter}}$ do

Sample $w \sim \text{Uniform}[0,1]$ and $\zeta \sim \mathcal{N}(0, I)$

$\xi = \sqrt{ab} u_t + \sqrt{a(2 - a - \beta)} \zeta$, $\hat{x}^* = \hat{x}_t - a \nabla U(\hat{x}_t) + \xi$

Propose $x^* = (L^T)^{-1} \hat{x}^*$

$\nabla U(\hat{x}^*) = L^{-1} \nabla U(x^*)$, $\tilde{\zeta} = \nabla U(\hat{x}^*) + \nabla U(\hat{x}_t)$

$\rho = \exp \left\{ U(x_t) - U(x^*) + \frac{1}{2a} (\tilde{\zeta})^T (\xi - \frac{\tilde{\zeta}}{\sqrt{a}}) \right\}$

if $w < \min(1, \rho)$ then

$x_{t+1} = x^*$, $\hat{x}_{t+1} = \hat{x}^*$, $\nabla U(\hat{x}_{t+1}) = \nabla U(\hat{x}^*)$ # Accept

if HAMS-A then

$u_{t+1} = \left( \frac{2b}{2-a} - 1 \right) u_t + \frac{2 \sqrt{b(2-a-b)}}{2-a} \zeta - \frac{\sqrt{ab}}{2-a} \tilde{\zeta}$

else

$u_{t+1} = u_t - \frac{\sqrt{ab}}{2-a} \tilde{\zeta}$

endif

endif

else

$x_{t+1} = x_t, u_{t+1} = -u_t, \hat{x}_{t+1} = \hat{x}_t, \nabla U(\hat{x}_{t+1}) = \nabla U(\hat{x}_t)$ # Reject

endif

**Algorithm 2.** Preconditioned BAOAB

Initialize $x_0, u_0, \hat{x}_0 = L^* x_0$ and $\nabla U(\hat{x}_0) = L^{-1} \nabla U(x_0)$.

for $t = 0, 1, 2, \ldots, N_{\text{iter}}$ do

Sample $w \sim \text{Uniform}[0,1]$ and $\zeta \sim \mathcal{N}(0, I)$

$\tilde{x}^* = \tilde{x}_t - (1 + c) \frac{\xi}{2} \nabla U(\tilde{x}_t) + (1 + c) \frac{\zeta}{2} u_t + \frac{\xi}{2} \sqrt{1 - c^2} \zeta$

Propose $x^* = (L^T)^{-1} \tilde{x}^*$

$\nabla U(\tilde{x}^*) = L^{-1} \nabla U(x^*)$

Propose $u^* = c u_t - \frac{\xi}{2} \nabla U(\tilde{x}_t) - \frac{\zeta}{2} \nabla U(\tilde{x}^*) + \sqrt{1 - c^2} \zeta$

$\rho = \exp \left\{ U(x_t) - U(x^*) + \left[ \frac{\xi}{2} u^* + \frac{\xi^2}{4} \nabla U(\tilde{x}^*) \right]^T \nabla U(\tilde{x}^*) + \left[ \frac{\zeta}{2} u_t - \frac{\zeta^2}{4} \nabla U(\tilde{x}_t) \right]^T \nabla U(\tilde{x}_t) \right\}$

if $w < \min(1, \rho)$ then

$x_{t+1} = x^*$, $u_{t+1} = u^*$, $\hat{x}_{t+1} = \hat{x}^*$, $\nabla U(\hat{x}_{t+1}) = \nabla U(\hat{x}^*)$ # Accept

else

$x_{t+1} = x_t, u_{t+1} = -u_t$, $\hat{x}_{t+1} = \hat{x}_t, \nabla U(\hat{x}_{t+1}) = \nabla U(\hat{x}_t)$ # Reject

endif

**SM3. Technical details.**

**SM3.1. Proof of Proposition 1.** First, consider the case $\phi = 0$. For notational simplicity, assume that the target density $\pi(x)$ is univariate. The proof can be easily extended to multivariate density $\pi(x)$. Then HAMS proposal given by (6)–(7) becomes...
Algorithm 3. Preconditioned ABOBA

Initialize $x_0, u_0, \tilde{x}_0 = L^T x_0$ and $\nabla U(\tilde{x}_0) = L^{-1} \nabla U \left( x_0 + \frac{c}{2} (L^T)^{-1} u_0 \right)$.

for $t = 0, 1, 2, \ldots, N_{\text{iter}}$ do

Sample $w \sim \text{Uniform}[0,1]$ and $\zeta \sim \mathcal{N}(0, I)$

$\hat{x}_t = \hat{x}_t - (1 + c) \frac{c}{2} \nabla U(\hat{x}_t) + (1 + c) \frac{c}{2} u_t + \frac{c}{2} \sqrt{1 - c^2} \zeta$

Propose $x^* = (L^T)^{-1} \hat{x}_t$

Propose $u^* = cu_t - (1 + c) \frac{c}{2} \nabla U(\hat{x}_t) + \sqrt{1 - c^2} \zeta$

$\rho = \exp \left\{ U(x_t) - U(x^*) + \frac{c}{2} (u^* + u_0)^T \nabla U(\hat{x}_t) \right\}$

if $w < \min(1, \rho)$ then

$x_{t+1} = x^*, u_{t+1} = u^*, \hat{x}_{t+1} = \hat{x}^*, \nabla U(\hat{x}_{t+1}) = L^{-1} \nabla U \left( x^* + \frac{c}{2} (L^T)^{-1} u^* \right)$

# Accept

else

$x_{t+1} = x_t, u_{t+1} = -u_t, \hat{x}_{t+1} = \hat{x}_t, \nabla U(\hat{x}_{t+1}) = L^{-1} \nabla U \left( x_t - \frac{c}{2} (L^T)^{-1} u_t \right)$

# Reject

Algorithm 4. Preconditioned OBABO

Initialize $x_0, u_0, \tilde{x}_0 = L^T x_0$ and $\nabla U(\tilde{x}_0) = L^{-1} \nabla U(x_0)$.

for $t = 0, 1, 2, \ldots, N_{\text{iter}}$ do

Sample $w \sim \text{Uniform}[0,1]$ and $\zeta_1, \zeta_2 \sim \mathcal{N}(0, I)$

$u^* = \sqrt{c} u_t + \sqrt{1 - c^2} \zeta_1, \quad \hat{x}^* = \hat{x}_t + c u^* - \frac{c^2}{2} \nabla U(\hat{x}_t)$

Propose $x^* = (L^T)^{-1} \hat{x}^*$

$\nabla U(\hat{x}^*) = L^{-1} \nabla U(x^*)$, $\quad \hat{\xi} = \nabla U(\hat{x}^*) + \nabla U(\hat{x}_t)$

$\rho = \exp \left\{ U(x_t) - U(x^*) + \frac{c}{2} \hat{\xi}^T (2u^* - \hat{\xi}) \right\}$

if $w < \min(1, \rho)$ then

$x_{t+1} = x^*, \quad \hat{x}_{t+1} = \hat{x}^*, \quad \nabla U(\hat{x}_{t+1}) = \nabla U(\hat{x}^*)$ \quad # Accept

$u_{t+1} = \sqrt{c} (u^* - \hat{\xi}) + \sqrt{1 - c^2} \zeta_2$

else

$x_{t+1} = x_t, u_{t+1} = -u_t, \quad \hat{x}_{t+1} = \hat{x}_t, \quad \nabla U(\hat{x}_{t+1}) = \nabla U(\hat{x}_t)$ \quad # Reject

(SM20) \[
\begin{pmatrix}
  x^* \\
  u^*
\end{pmatrix} = \begin{pmatrix}
  x_0 \\
  u_0
\end{pmatrix} - \tilde{A} \begin{pmatrix}
  \nabla U(x_0) \\
  u_0
\end{pmatrix} + \begin{pmatrix}
  Z_{0(1)}^T \\
  Z_{0(2)}^T
\end{pmatrix},
\]

where

\[
\tilde{A} = \begin{pmatrix}
  a_1 & -a_2 \\
  a_2 & 2 - a_1
\end{pmatrix}, \quad \begin{pmatrix}
  Z_{0(1)}^T \\
  Z_{0(2)}^T
\end{pmatrix} \sim \mathcal{N}(0, 2A - A^2).
\]

By the parametrization (18) for $A$, we have

\[
\tilde{A} = \begin{pmatrix}
  2 - c_1 (1 + \sqrt{1 - \epsilon^2}) & -c_1 \epsilon c_2 \\
  c_1 \epsilon c_2 & 2 - c_2 (1 + \sqrt{1 - \epsilon^2})
\end{pmatrix},
\]

\[
\text{Var}(Z_{0(1)}) = c_1[2 - c_2 \epsilon^2 + 2 \sqrt{1 - \epsilon^2} + c_1 \{ \epsilon^2 - 2(1 + \sqrt{1 - \epsilon^2}) \}],
\]

\[
\text{Var}(Z_{0(2)}) = c_2[2(1 + \sqrt{1 - \epsilon^2}) - c_1 \epsilon^2 - c_2 (1 + \sqrt{1 - \epsilon^2})^2],
\]

(SM21) \[
\text{Cov}(Z_{0(1)}^T, Z_{0(2)}^T) = (c_1 - c_2) \sqrt{c_2 c_2} (1 + \sqrt{1 - \epsilon^2}) \epsilon.
\]
Moreover, using (19) and taking Taylor expansions with respect to \( \epsilon \) around 0 lead to

\[
\hat{A} = \begin{pmatrix}
\hat{\eta}_1 + \left( \frac{1}{2} - \frac{\nu_1}{4} \right) \epsilon^2 + \mathcal{O}(\epsilon^3)
& -\epsilon + \frac{1}{4} (\hat{\eta}_1 + \eta_2) \epsilon^2 + \mathcal{O}(\epsilon^3) \\
\epsilon - \frac{1}{4} (\hat{\eta}_1 + \eta_2) \epsilon^2 + \mathcal{O}(\epsilon^3)
& \eta_2 \epsilon + \left( \frac{1}{2} - \frac{\nu_2}{4} \right) \epsilon^2 + \mathcal{O}(\epsilon^3)
\end{pmatrix},
\]

where

\[
\text{Var}(Z_0^{(1)}) = 2\hat{\eta}_1 \epsilon - \frac{3}{2} \hat{\eta}_1^2 \epsilon^2 + \mathcal{O}(\epsilon^3),
\]

\[
\text{Var}(Z_0^{(2)}) = 2\eta_2 \epsilon - \frac{3}{2} \eta_2^2 \epsilon^2 + \mathcal{O}(\epsilon^3),
\]

and

\[
\text{Cov}(Z_0^{(1)}, Z_0^{(2)}) = (\eta_2 - \hat{\eta}_1) \epsilon^2 + \mathcal{O}(\epsilon^3).
\]

From (SM22) with all \( \mathcal{O}(\epsilon^2) \) terms as remainders, we obtain

\[
\hat{A} = \begin{pmatrix}
\hat{\eta}_1 \epsilon & -\epsilon \\
\epsilon & \eta_2 \epsilon
\end{pmatrix} + \mathcal{O}(\epsilon^2),
\]

\[
\text{Var} \begin{pmatrix} Z_0^{(1)} \\ Z_0^{(2)} \end{pmatrix} = \begin{pmatrix} 2\hat{\eta}_1 \epsilon & 0 \\ 0 & 2\eta_2 \epsilon \end{pmatrix} + \mathcal{O}(\epsilon^2).
\]

Using this approximation, the update (SM20) becomes

\[
(x^* \quad u^*) = \left( x_0 \quad u_0 \right) - \begin{pmatrix} \hat{\eta}_1 & -1 \\ 1 & \eta_2 \end{pmatrix} \left( \nabla U(x_0) \quad u_0 \right) \epsilon + \begin{pmatrix} \sqrt{2\hat{\eta}_1} \zeta_1 \\ \sqrt{2\eta_2} \zeta_2 \end{pmatrix}, \quad \zeta_1, \zeta_2 \sim i.i.d. \mathcal{N}(0, \epsilon),
\]

which is Euler’s discretization, hence solving SDE (20) as \( \epsilon \to 0 \).

Next, we handle the case of nonzero \( \phi \), which appears in the HAMS update only through \( \phi(\tilde{Z}^{(1)} + \nabla U(x_0) - \nabla U(x^*)) \) in the update (11) for \( u^* \). The term \( \tilde{Z}^{(1)} \) is of order \( \mathcal{O}_p(\sqrt{\epsilon}) \):

\[
\tilde{Z}^{(1)} = Z_0^{(1)} - a_1 \nabla U(x_0) + a_2 u_0 = \mathcal{O}_p(\sqrt{\epsilon}),
\]

because by (SM22), \( a_1 = \mathcal{O}(\epsilon) \), \( a_2 = \mathcal{O}(\epsilon) \), \( \text{Var}(Z_0^{(1)}) = \mathcal{O}(\epsilon) \), and hence \( Z_0^{(1)} = \mathcal{O}_p(\sqrt{\epsilon}) \). Moreover, by the assumption that \( \|\nabla^2 U(x)\| \leq M \) for a constant \( M \) and the mean value theorem,

\[
|\nabla U(x_0) - \nabla U(x^*)| \leq M|x_0 - x^*| = M|\tilde{Z}^{(1)}| = \mathcal{O}_p(\sqrt{\epsilon}).
\]

For \( \phi = \mathcal{O}(\epsilon) \), combining the preceding results yields \( \phi(\tilde{Z}^{(1)} + \nabla U(x_0) - \nabla U(x^*)) = \mathcal{O}_p(\epsilon^{3/2}) \). Hence the additional term depending on \( \phi \) does not affect the approximation (SM23) for (SM20). This completes the proof of Proposition 1.

**SM3.2. Parametrization for HAMS-B.** We discuss the relationship between the parametrization (25) and that used in Song and Tan (2021), Section 3.4. The latter, with \( (\epsilon, c) \) renamed \( (\delta, d) \), can be stated as

\[
(a_1, a_2, a_3) = \begin{pmatrix} 1 - \sqrt{1 - \delta^2} & \delta \sqrt{d} & 2 - d(1 - \sqrt{1 - \delta^2}) \end{pmatrix}, \quad \phi = \frac{\delta \sqrt{d}}{1 + \sqrt{1 - \delta^2}},
\]

where \( a_1, a_2, \) and \( \phi \) would be the same as in (24) with \( (\epsilon, c_2) \) replaced by \( (\delta, d) \). By matching the expressions in (25) and (SM24), \( (\epsilon, c_1) \) and \( (\delta, d) \) are related as follows:

\[
1 - \sqrt{1 - \delta^2} = 2 - c_1(1 + \sqrt{1 - \epsilon^2}),
\]

\[
2 - d(1 - \sqrt{1 - \delta^2}) = 1 + \sqrt{1 - \epsilon^2}.
\]
Solving for \((\delta, d)\) from the above equations gives

\[
\delta^2 = 1 - \left[ c_1 (1 + \sqrt{1 - \epsilon^2}) - 1 \right]^2,
\]

\[
d = \frac{1 - \sqrt{1 - \epsilon^2}}{2 - c_1 (1 + \sqrt{1 - \epsilon^2})}.
\]

For \(c_1 = e^{-n\epsilon/2}\) in (19), taking Taylor expansions yields

\[
\delta^2 = 2\eta_1 \epsilon + (1 - \frac{3}{2} \eta_1) \epsilon^2 + \mathcal{O}(\epsilon^2),
\]

\[
d = \frac{\epsilon + \mathcal{O}(\epsilon^3)}{2\eta_1 + (1 - \frac{3}{2} \eta_1) \epsilon + \mathcal{O}(\epsilon^2)}.
\]

For any fixed \(\eta_1 > 0\), it follows that \(\delta^2 = \mathcal{O}(\epsilon)\) and \(d = \mathcal{O}(\epsilon)\). Hence a \(\delta\) value translates into a much smaller value for \(\epsilon\) in the new parametrization, and the \(d\) value also tends to be much smaller than 1. Nevertheless, if \(\eta_1 = 0\), then, by the leading terms, \(\delta \approx \epsilon\) and \(d \approx 1\), which are expected for the corresponding Hamiltonian dynamics.

**SM3.3. Proof of Proposition 2.** When \(\pi(x) = \mathcal{N}(0, \gamma^{-1})\), we have \(\nabla U(x) = \gamma x\) and the HAMS proposal becomes

\[(SM25)\]
\[x^* = (1 - a_1 \gamma)x_0 + a_2 u_0 + Z_0^{(1)},\]

\[(SM26)\]
\[u^* = \{a_1 \phi \gamma (\gamma - 1) - a_2 \gamma\} x_0 + \{a_3 - a + \phi a_2 (1 - \gamma)\} u_0 + \phi (1 - \gamma) Z_0^{(1)} + Z_0^{(2)},\]

\[(SM27)\]
\[Z^{(1)*} = Z_0^{(1)} - a_1 \gamma (x_0 + x^*) + a_2 (u_0 - u^*),\]

\[(SM28)\]
\[Z^{(2)*} = Z_0^{(2)} - a_2 \gamma (x_0 + x^*) + a_3 (u_0 - u^*).\]

We can also simplify \(\Delta G\) as

\[(SM29)\]
\[\Delta G = \frac{\gamma}{2} ((x^*)^2 - x_0^2) + \frac{1}{2}((u^*)^2 - u_0^2) + \frac{1}{2} Z^{*\gamma}(2A - A^2)^{-1} Z^* - \frac{1}{2} Z_0^{(1)} Z_0^{(2)}(2A - A^2)^{-1} Z_0.
\]

Combining (SM25) – (SM29) shows that \(\Delta G\) can be expressed as a quadratic form of \((x_0, u_0, Z_0^{(1)}, Z_0^{(2)})^\gamma\) stated in Proposition 2:

\[
\Delta G = (x_0, u_0, Z_0^{(1)}, Z_0^{(2)}) D(\gamma) (x_0, u_0, Z_0^{(1)}, Z_0^{(2)})^\gamma.
\]

The entries of \(D(\gamma)\) are then computed using *Mathematica*. We find that \(\Delta G\) does not contain any \([Z_0^{(2)}]^2\) terms, therefore \(d_{44}(\gamma) = 0\). For the other diagonal entries, the order of \(d_{11}(\gamma)\) is \(\gamma^4\) and the orders of \(d_{22}(\gamma), d_{33}(\gamma)\) are both \(\gamma^2\). For \(d_{11}(\gamma)\), the coefficient of \(\gamma^4\) is

\[
\frac{a_1^2 h(\phi)}{2a_1(a_3 - 2) - 2(a_2^2 + 2a_4 - 4)}.
\]
For $d_{22}(\gamma)$, the coefficient of $\gamma^2$ is

\[(\text{SM31}) \quad \frac{a_{2}^2h(\phi)}{2a_1(a_3 - 2) - 2(a_{2}^2 + 2a_3 - 4)}.\]

For $d_{33}(\gamma)$, the coefficient of $\gamma^2$ is

\[(\text{SM32}) \quad \frac{h(\phi)}{2a_1(a_3 - 2) - 2(a_{2}^2 + 2a_3 - 4)}.\]

The same function of $\phi$ that appears in (SM30), (SM31) and (SM32) is quadratic in $\phi$,

\[h(\phi) = \phi^2(4 - 2a_1) - \phi a_2 + 2a_1 + a_{_2}^2 - a_1a_3.\]

This quadratic function $h(\phi)$ is always $\geq 0$, with the discriminant

\[
16a_1^2 - 4(4 - 2a_1)(2a_1 + a_{_2}^2 - a_1a_3) \\
= -8a_1(4 - 2a_1 - 2a_3 + a_1a_3 - a_{_2}^2) \leq 0,
\]

because $a_1 \geq 0$, and $(2 - a_1)(2 - a_3) \geq a_{_2}^2$ due to the constraint $0 \leq A \leq 2I$. Therefore $|h(\phi)|$ is minimized when $\phi = \frac{a_2}{2 - a_1}$.

**SM3.4. Proof of Corollary 2.** When $\phi = a_2/(2 - a_1)$, we have

\[(\text{SM33}) \quad u^* = \frac{a_1 + a_{_2}^2 + 2a_3 - a_1a_3 - 2}{2 - a_1} u_0 - \frac{a_2}{2 - a_1} \nabla U(x_0) - \frac{a_2}{2 - a_1} \nabla U(x^*) + \frac{a_2}{2 - a_1} Z_0^{(1)} + Z_0^{(2)}

\[Z^{(1)*} = a_2 \left(2 - a_3 - \frac{a_{_2}^2}{2 - a_1}\right) u_0 + \left(\frac{a_{_2}^2}{2 - a_1} - a_1\right) \nabla U(x_0) + \left(\frac{a_{_2}^2}{2 - a_1} - a_1\right) \nabla U(x^*)

(\text{SM34})

\[+ (1 - \frac{a_{_2}^2}{2 - a_1}) Z_0^{(1)} - a_2 Z_0^{(2)}

\[Z^{(2)*} = a_3 \left(2 - a_3 - \frac{a_{_2}^2}{2 - a_1}\right) u_0 + \frac{a_2(a_1 + a_3 - 2)}{2 - a_1} \nabla U(x_0) + \frac{a_2(a_1 + a_3 - 2)}{2 - a_1} \nabla U(x^*)

(\text{SM35})

\[- \frac{a_2a_3}{2 - a_1} Z_0^{(1)} + (1 - a_3) Z_0^{(2)}.

Furthermore, because $A$ matrix is block diagonal, $(2A - A^2)^{-1}$ is also block diagonal given by

\[(\text{SM36}) \quad (2A - A^2)^{-1} = \left[ (a_{1}a_3 - a_{_2}^2)(4 + a_1a_3 - a_{_2}^2 - 2a_1 - 2a_3) \right]^{-1} \left[\begin{array}{cc} (2a_3 - a_{_2}^2 - a_{_2}^2) I & a_2(a_1 + a_3 - 2) I \\ a_2(a_1 + a_3 - 2) I & (2a_3 - a_{_2}^2 - a_{_2}^2) I \end{array} \right].

Recall that $\Delta G$ is defined as

\[\Delta G = U(x^*) - U(x_0) + \frac{1}{2} \{u^* \nabla U(x_0) + Z^{*T} (2A - A^2)^{-1} Z^* - Z_0^{*T} (2A - A^2)^{-1} Z_0\}.

Substituting (SM33) – (SM36) into the above, we obtain results in Corollary 2.
SM3.5. Proof of Proposition 3. When the target density is $\mathcal{N}(0, \gamma^{-1})$, the HAMS proposal can be equivalently written in the form of (28) with

$$
\Phi = \begin{pmatrix} 1 - a_1 \gamma \\ a_1 \phi \gamma (\gamma - 1) - a_3 - 1 + \phi a_2 (1 - \gamma) \end{pmatrix}, \quad \zeta \sim \mathcal{N}(0, W),
$$

$$
W = \begin{pmatrix} 2a_1 - a_1^2 - a_2^2 & \phi (1 - \gamma)(2a_1 - a_1^2 - a_2^2) + 2a_2 - a_1a_2 - a_2a_2 \\ \phi (1 - \gamma)(2a_1 - a_1^2 - a_2^2) & \phi^2 (1 - \gamma)^2 (2a_1 - a_1^2 - a_2^2) + 2a_2 - a_1a_2 - a_2a_2 \\ +2a_1 - a_1a_2 - a_2a_2 & +2\phi (1 - \gamma)(2a_2 - a_1a_2 - a_2a_2) \\ \end{pmatrix}.
$$

Similarly as in Burrage et al. (2007), Section 2, taking the variance of both sides in (28) shows that the stationary variance $V$ satisfies the following equation,

$$
V = \Phi V \Phi^\top + W.
$$

Substituting the default choice $\phi = a_2/(2 - a_1)$ in $\Phi$ and $W$ and solving for $V$ in (SM38), we obtain (29) stated in Proposition 3. Note that without using the default choice $\phi$, a general expression of $V$ in terms of $a_1, a_2, a_3$ and $\phi$ can also be obtained from (SM38). But the expression is complicated and not informative, hence not presented here.

SM3.6. Proof of Lemma 1. Denote as $\Psi$ the mapping from $(x_0, u_0, Z_0)$ to $(x^*, u^*, Z^*)$ given by (9)–(12). Then $\Psi$ satisfies the following generalized reversibility:

$$
\text{(SM39)} \quad \text{forward: } \begin{pmatrix} x_0 \\ u_0 \\ Z_0 \end{pmatrix} \xrightarrow{\Phi} \begin{pmatrix} x^* \\ u^* \\ Z^* \end{pmatrix}, \quad \text{backward: } \begin{pmatrix} x^* \\ -u^* \\ -Z^* \end{pmatrix} \xrightarrow{\Phi} \begin{pmatrix} x_0 \\ -u_0 \\ -Z_0 \end{pmatrix}.
$$

Let $S$ be the mapping that changes the signs of $u$ and $Z$, that is, $S(x, u, Z) = (x, -u, -Z)$. Define the composite $\Psi = S \circ \Psi$. Then (SM39) can be equivalently stated as

$$
\text{(SM40)} \quad \Psi(S(x, u, Z)) = S(x, u, Z).
$$

Moreover, because the function $G$ is even in $u$ and $Z$, we have

$$
\text{(SM41)} \quad G \circ S = G.
$$

According to (14), the acceptance rate at any current value $(x, u, Z)$ is

$$
\alpha(x, u, Z) = \min[1, \exp(-\Delta G(x, u, Z))],
$$

where $\Delta G(x, u, Z) = G(\Psi(x, u, Z)) - G(x, u, Z)$, and $G$ is redefined as

$$
G(x, u, Z) = H(x, u) + \frac{1}{2} Z^\top (2A - A^2)^{-1} Z + \frac{1}{2} \log(2\pi)^k + \frac{1}{2} \log [(2\pi)^{2k}/\text{Det}(2A - A^2)]
$$

The determinant terms are included to make $\exp(-G)$ a valid density function. If the target density is $\mathcal{N}(0, I)$, then HAMS is rejection free with $\Delta G \equiv 0$ and hence Lemma 1 trivially holds: $E[\alpha] = P[\Delta G = 0] = 1$. If the target density is not $\mathcal{N}(0, I)$, then $P[\Delta G = 0] = 0$. Hence it suffices to show that $E[\alpha] = 2P[\Delta G < 0]$. 

At stationarity, the density of \((x, u, Z)\) is \(\exp(-G(x, u, Z))\). Then

\[
E[\alpha] = \int \min[1, \exp(-\Delta G(x, u, Z))] \cdot \exp(-G(x, u, Z)) \, dx \, du \, dZ
\]

\[
= \int_{\Delta G<0} \exp(-G(x, u, Z)) \, dx \, du \, dZ \tag{I}
\]

\[
+ \int_{\Delta G>0} \exp(-\Delta G(x, u, Z)) \exp(-G(x, u, Z)) \, dx \, du \, dZ. \tag{II}
\]

Apparently \((I) = P[\Delta G < 0]\). In the following, we show that \((I) = (II)\). On one hand, \((II)\) can be directly calculated as

\[
(II) = \int_{\Delta G>0} \exp[-(\Delta G(x, u, Z) + G(x, u, Z))] \, dx \, du \, dZ
\]

\[
= \int_{\Delta G>0} \exp[-G(\Psi(x, u, Z))] \, dx \, du \, dZ \tag{by (SM41)}
\]

\[
= \int_{\Delta G>0} \exp[-G(\hat{\Psi}(x, u, Z))] \, dx \, du \, dZ. \tag{SM42}
\]

On the other hand, \((I)\) can be shown to be

\[
(I) = \int_{\Delta G<0} \exp(-G(\hat{x}, \hat{u}, \hat{Z})) \, d\hat{x} \, d\hat{u} \, d\hat{Z}
\]

\[
= \int_{\Delta G>0} \exp(-G(\hat{\Psi}(x, u, Z))) \, dx \, du \, dZ. \tag{SM43}
\]

The first step follows by replacing \((x, u, Z)\) with \((\hat{x}, \hat{u}, \hat{Z})\) in the notation. The second step involves a change of variables in the integration: \((\hat{x}, \hat{u}, \hat{Z}) = \Psi(x, u, Z)\). The mapping \(\hat{\Psi}\) can be expressed as a series of shear mappings and sign changes, and hence has a unit Jacobian. Moreover, there is a one-to-one correspondence between \((x, u, Z)\) with \(\Delta G(x, u, Z) > 0\) and \((\hat{x}, \hat{u}, \hat{Z})\) with \(\Delta G(\hat{x}, \hat{u}, \hat{Z}) < 0\) under the change of variables:

\[
\Delta G(\hat{x}, \hat{u}, \hat{Z}) = G(\hat{\Psi}(\hat{x}, \hat{u}, \hat{Z})) - G(\hat{x}, \hat{u}, \hat{Z})
\]

(by definition) \(= G(\Psi(\hat{x}, \hat{u}, \hat{Z})) - G(\hat{x}, \hat{u}, \hat{Z})\)

(by (SM40)) \(= G(S(x, u, Z)) - G(\hat{\Psi}(x, u, Z))\)

(by (SM41)) \(= G(x, u, Z) - G(\Psi(x, u, Z)) = -\Delta G(x, u, Z)\).

Comparing (SM42) and (SM43) completes the proof.

**SM3.7. Proof of Proposition 4.** First, we prove the following result about bivariate normal random variables.

**Lemma SM1.** For \(\tau \in [-1, 1]\), let

\[
\begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathcal{N}\left(0, \begin{pmatrix} 1 & \tau \\ \tau & 1 \end{pmatrix}\right).
\]

Then

\[
P[X > 0 \text{ and } Y > 0] = \frac{1}{4} + \frac{1}{2\pi} \arcsin(\tau).
\]
Proof of Lemma SM1. Define $\zeta = (Y - \tau X)/\sqrt{1 - \tau^2}$. Then $\zeta$ and $X$ are jointly normal with $(X, \zeta)^T \sim \mathcal{N}(0, I)$. Using the fact that $Y > 0$ is equivalent to $\zeta > -\frac{\tau}{\sqrt{1 - \tau^2}}X$ and making a change to polar coordinates, we have

$$P[X > 0 \text{ and } Y > 0] = P \left[ X > 0 \text{ and } \zeta > -\frac{\tau}{\sqrt{1 - \tau^2}}X \right]$$

$$= \int_{\pi/2}^{\pi/2} \frac{1}{2\pi} \exp \left( -\frac{x^2}{2} - \frac{\zeta^2}{2} \right) d\zeta \, dx = \int_{\theta = \arctan \left( -\frac{\sqrt{1 - \tau^2}}{\tau} \right)}^{\pi/2} \int_{r = 0}^{\infty} \frac{1}{2\pi} e^{-r^2/2} \, dr \, d\theta$$

$$= \int_{\theta = \arctan \left( -\frac{\sqrt{1 - \tau^2}}{\tau} \right)}^{\pi/2} \frac{d\theta}{2\pi} = \frac{1}{4} + \frac{1}{2\pi} \arcsin(\tau).$$

Next, to apply Lemma 1 for $E[\alpha]$, we calculate $P[\Delta G < 0]$ for a univariate normal target. According to Corollary 2,

$$\Delta G(x, u, Z) = \frac{a_1 \gamma (\gamma - 1)}{2(2 - a_1)} B_1 \cdot B_2,$$

where $B_1 = (a_2 u + Z^{(1)} - a_1 \gamma x)$ and $B_2 = (a_2 u + Z^{(1)} + (2 - a_1 \gamma)x)$. At stationarity, we have $x \sim \mathcal{N}(0, \gamma^{-1}), u \sim \mathcal{N}(0, 1)$ and $Z^{(1)} \sim \mathcal{N}(0, 2a_1 - a_1^2 - a_2^2)$. Then $(B_1, B_2)$ are jointly normal with

$$(SM44) \quad \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} \sim \mathcal{N} \left( 0, \begin{pmatrix} a_1^2(\gamma - 1) & + 2a_1 \\ a_1^2(\gamma - 1) & a_1^2(\gamma - 1) - 2a_1 + \frac{4}{\gamma} \end{pmatrix} \right).$$

The correlation coefficient between $B_1$ and $B_2$ is

$$\rho = \frac{a_1^2(\gamma - 1)\sqrt{\gamma}}{a_1(a_1^2(\gamma - 1)^2 - 4a_1 + 8)}.$$

Using Lemma SM1, we have

$$P(B_1 > 0 \text{ and } B_2 > 0) = \frac{1}{4} + \frac{1}{2\pi} \arcsin(\rho),$$

which leads to

$$P(B_1 B_2 > 0) = P(B_1 > 0 \text{ and } B_2 > 0) + P(B_1 > 0 \text{ and } B_2 < 0)$$

$$= 2P(B_1 > 0 \text{ and } B_2 > 0) = \frac{1}{2} + \frac{1}{2\pi} \arcsin(\rho),$$

and $P(B_1 B_2 < 0) = \frac{1}{2} - \frac{1}{2\pi} \arcsin(\rho)$. Clearly $\text{Sign}(\rho) = \text{Sign}(\gamma - 1)$. Then depending on whether $\gamma > 1$, we have

$$P[\Delta G < 0] = \begin{cases} P[B_1 B_2 < 0] = \frac{1}{2} - \frac{1}{2\pi} \arcsin(\rho) & \text{if } \gamma > 1, \\ P[B_1 B_2 > 0] = \frac{1}{2} + \frac{1}{2\pi} \arcsin(\rho) & \text{if } 0 < \gamma < 1, \\ 0 & \text{if } \gamma = 1. \end{cases}$$

Combining this with Lemma 1 shows that the expected acceptance rate is

$$(SM45) \quad E[\alpha] = \begin{cases} 1 - \frac{2}{\pi} \arcsin(\rho) & \text{if } \gamma > 1 \\ 1 + \frac{2}{\pi} \arcsin(\rho) & \text{if } 0 < \gamma < 1 \\ 1 & \text{if } \gamma = 1. \end{cases}$$
Finally we relate the above expression to the expected change $E[\Delta G]$. Because $E[B_1B_2] = a_1^2(\gamma - 1)$ by (SM44), we have

$$E[\Delta G] = \frac{a_1\gamma(\gamma - 1)}{2(2 - a_1)} E[B_1B_2] = \frac{a_1^2\gamma(\gamma - 1)^2}{2(2 - a_1)},$$

which, by direct calculation, is related to $\rho$ as follows:

$$\rho^2 = \frac{E[\Delta G]}{2 + E[\Delta G]} \iff \rho = \text{Sign}(\gamma - 1) \sqrt{\frac{E[\Delta G]}{2 + E[\Delta G]}}.$$

Substituting (SM46) into (SM45), we obtain the unified expression in Proposition 4:

$$E[\alpha] = 1 - \frac{2}{\pi} \arcsin \left( \frac{E[\Delta G]}{2 + E[\Delta G]} \right) = 1 - \frac{2}{\pi} \arctan \left( \frac{\sqrt{E[\Delta G]}}{2} \right).$$

**SM3.8. Proof of Corollary 5.** The assumption stated in Corollary 5 implies (40) and (41) in Lemma 2. It remains to verify the moments conditions stated as (39). To this end, we present the following general result which is also used in the proof of Corollary 6.

**Lemma SM2.** For $\tau \in [-1, 1]$, let

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathcal{N} \left( 0, \begin{pmatrix} 1 & \tau \\ \tau & 1 \end{pmatrix} \right).$$

Let $Z = \theta XY$, where $\theta$ and $\tau$ satisfy

$$\tau^2 = \frac{\theta \tau}{2 + \theta \tau}.$$

Denote $\mu = E[Z] = \theta \tau$. Then we have

$$E[Z^2] = 2\mu + 3\mu^2 = 2\mu + O(\mu^2),$$
$$E[Z^3] = 18\mu^2 + 15\mu^3 = O(\mu^2),$$
$$E[Z^4] = 36\mu^2 + 180\mu^3 + 105\mu^4 = O(\mu^2).$$

**Proof of Lemma SM2.** Using the formulas in Kan (2008), we directly calculate

$$E[XY] = \tau, \ E[(XY)^2] = 1 + 2\tau^2, \ E[(XY)^3] = 9\tau + 6\tau^3, \ E[(XY)^4] = 9 + 72\tau^2 + 24\tau^4.$$

From the relation (SM47), we get

$$\theta^2 = 2\tau \theta + \theta^2 \tau^2 = 2\mu + \mu^2.$$

Use the above together with $\mu = \theta \tau$, we then have

$$E[Z^2] = \theta^2 E[(XY)^2] = \theta^2(1 + 2\tau^2) = 2\mu + 3\mu^2,$$
$$E[Z^3] = \theta^3 E[(XY)^3] = \theta^3(9\tau + 6\tau^3) = 18\mu^2 + 15\mu^3,$$
$$E[Z^4] = \theta^4 E[(XY)^4] = \theta^4(9 + 72\tau^2 + 24\tau^4) = 36\mu^2 + 180\mu^3 + 105\mu^4.$$

This completes the proof of Lemma SM2.
Now to apply Lemma SM2 to HAMS, we write
\[
\Delta G_i = \frac{a_1 \gamma_i (\gamma_i - 1)}{2(2 - a_1)} B_{1i} B_{2i},
\]
similarly as in the proof of Proposition 4. Random variables $B_{1i}, B_{2i}$ can be standardized so that $\Delta G_i$ is a product of a constant and two unit variance Gaussian variables. The relation (SM47) holds for $\Delta G_i$ because of (SM46). By Lemma SM2, condition (39) is then satisfied for HAMS. This completes the proof of Corollary 5.

**SM3.9. Proof of Proposition 5.** The two eigenvalues of $\Phi$ are given by
\[
\frac{1}{2} \left( a_3 - a_1 + \sqrt{(a_1 + a_3 - 2)^2 - 4a_2^2} \right).
\]
The spectral radius (i.e., maximum modulus of the two eigenvalues) is
\[
\rho(\Phi) = \begin{cases} 
\sqrt{a_1 + a_3 - a_1 a_3 + a_2^2 - 1}, & \text{if } 4a_2^2 \geq (a_1 + a_3 - 2)^2, \\
\frac{1}{2} \left( |a_3 - a_1| + \sqrt{(a_1 + a_3 - 2)^2 - 4a_2^2} \right), & \text{if } 4a_2^2 < (a_1 + a_3 - 2)^2.
\end{cases}
\]
For fixed $a_1$ and $\nu = a_2^2/a_3$, we write the spectral radius of $\Phi$ as a function of $a_3$ (SM48)
\[
\rho(a_3) = \begin{cases} 
\sqrt{a_1 + a_3 - a_1 a_3 + \nu a_3 - 1}, & \text{if } R_1 \leq a_3 \leq R_2, \\
\frac{1}{2} \left( |a_3 - a_1| + \sqrt{(a_1 + a_3 - 2)^2 - 4\nu a_3} \right), & \text{if } a_3 < R_1 \text{ or } a_3 > R_2,
\end{cases}
\]
where $R_1$ and $R_2$ are the roots of $(a_1 + a_3 - 2)^2 - 4\nu a_3$ as a function of $a_3$, given by
\[
R_1 = 2\nu + 2 - a_1 - 2\sqrt{\nu(\nu + 2 - a_1)} = (\sqrt{\nu + 2 - a_1} - \sqrt{\nu})^2, \\
R_2 = 2\nu + 2 - a_1 + 2\sqrt{\nu(\nu + 2 - a_1)} = (\sqrt{\nu + 2 - a_1} + \sqrt{\nu})^2.
\]
Assume that $0 < a_1 < 2$ and $\nu \leq a_1 \leq 1 + \nu$. Then as we show later, the function $\rho(a_3)$ is nonincreasing when $a_3 < R_1$ and nondecreasing where $a_3 \geq R_2$. In the intermediate case $R_1 \leq a_3 \leq R_2$, the function $\rho(a_3)$ is nondecreasing because by the condition $a_1 \leq \nu + 1$, we have
\[
\rho(a_3) = \sqrt{a_1 + a_3 - a_1 a_3 + \nu a_3 - 1}.
\]
Consequently, the spectral radius of $\Phi$ is minimized at
\[
a_3^* = R_1 = (\sqrt{\nu + 2 - a_1} - \sqrt{\nu})^2,
\]
with the minimum spectral radius
\[
\frac{|a_3^* - a_1|}{2}
\]
and the implied choice of $a_2$
\[
a_2^* = \pm \sqrt{\nu a_3^*}.
\]
The condition $0 \leq A \leq 2I$ holds if and only if $0 \leq a_1, a_3 \leq 2$, $a_1 a_3 \geq a_3^2$, and $(2 - a_1)(2 - a_3) \geq a_3^2$. With $a_3^2 = \nu a_3$ and $\nu \leq a_1$, this condition dictates that

$$0 \leq a_3 \leq \frac{2(2 - a_1)}{\nu + 2 - a_1},$$

which is always satisfied by $a_3^2$, because

$$a_3 \leq (\sqrt{\nu + 2 - a_1} - \sqrt{\nu}) (\sqrt{\nu + 2 - a_1} + \sqrt{\nu}) = 2 - a_1 \leq \frac{2(2 - a_1)}{\nu + 2 - a_1}.$$

The last inequality follows with $\nu \leq a_1$ and $\nu + 2 - a_1 \geq 2$.

In the remainder of this section, we demonstrate the monotonicity of $\rho(a_3)$ for $a_3 < R_1$ or $a_3 > R_2$ as mentioned above. We distinguish four cases.

**Case 1:** $a_3 \geq a_1, a_3 > R_2$. Then $\rho(a_3)$ becomes

$$\rho(a_3) = \frac{1}{2} \left( a_3 - a_1 + \sqrt{a_1 + a_3 - 2} + 4\nu a_3 \right),$$

with the derivative

$$\frac{d\rho}{da_3} = \frac{1}{8} \left( (a_1 + a_3 - 2)^2 - 4\nu a_3 \right)^{-1/2} \left( a_1 + a_3 - 2 - 2\nu + \sqrt{a_1 + a_3 - 2} - 4\nu a_3 \right).$$

Then $\frac{d\rho}{da_3} > 0$ because

$$a_1 + a_3 - 2 - 2\nu + \sqrt{a_1 + a_3 - 2} - 4\nu a_3
\geq a_1 + R_2 - 2 - 2\nu + \sqrt{a_1 + a_3 - 2} - 4\nu a_3
= 2\sqrt{\nu^2 + 2\nu - a_1\nu} + \sqrt{a_1 + a_3 - 2} - 4\nu a_3 \geq 0.$$

**Case 2:** $a_1 \leq a_3 < R_1$. Then $\rho(a_3)$ and $\frac{d\rho}{da_3}$ are the same as in (SM49) and (SM50). For $0 < a_1 \leq 2$, it holds that $\nu \leq \sqrt{\nu(2 - a_1 + \nu)}$. Then $\frac{d\rho}{da_3} \leq 0$ because

$$a_1 + a_3 < a_1 + R_1 = 2 + 2\nu - \sqrt{(2 - a_1 + \nu)} \leq 2
\Rightarrow 2 - a_1 - a_3 + 2\nu \geq 2 - a_1 - a_3 \geq \sqrt{(a_1 + a_3 - 2) - 4\nu a_3}
\Rightarrow a_1 + a_3 - 2 - 2\nu + \sqrt{(a_1 + a_3 - 2) - 4\nu a_3} \leq 0.$$

**Case 3:** $a_3 < a_1, a_3 < R_1$. Then $\rho(a_3)$ becomes

$$\rho(a_3) = \frac{1}{2} \left( a_1 - a_3 + \sqrt{a_1 + a_3 - 2} + 4\nu a_3 \right),$$

with the derivative

$$\frac{d\rho}{da_3} = \frac{1}{8} \left( (a_1 + a_3 - 2)^2 - 4\nu a_3 \right)^{-1/2} \left( a_1 + a_3 - 2 - 2\nu - \sqrt{a_1 + a_3 - 2} - 4\nu a_3 \right).$$

Then $\frac{d\rho}{da_3} < 0$ because

$$a_1 + a_3 - 2 - 2\nu - \sqrt{a_1 + a_3 - 2} - 4\nu a_3
< a_1 + R_1 - 2 - 2\nu - \sqrt{(a_1 + a_3 - 2) - 4\nu a_3}
= -2\sqrt{\nu^2 + 2\nu - a_1\nu} - \sqrt{(a_1 + a_3 - 2)^2 - 4\nu a_3} \leq 0.$$
Case 4: $2 > a_1 \geq a_3 > R_2$. Then $\rho(a_3)$ and $d\rho/da_3$ are the same as in (SM51) and (SM52). Notice that

$$a_1 + a_3 > a_1 + R_2 = 2 + 2\nu + 2\sqrt{\nu(2 - a_1 + \nu)}$$

$$\implies a_1 + a_3 - 2 - 2\nu > 0.$$ 

By the condition $a_1 \leq 1 + \nu$, we have $2 + \nu - a_1 > 0$. Then $\frac{d\rho}{da_3} \geq 0$ because

$$4\nu(2 + \nu - a_1)^2 \geq 0$$

$$\implies (a_1 + a_3 - 2)^2 - 4\nu a_3 + 4\nu(2 + \nu - a_1) \geq (a_1 + a_3 - 2)^2 - 4\nu a_3$$

$$\implies (a_1 + a_3 - 2 - 2\nu)^2 \geq (a_1 + a_3 - 2)^2 - 4\nu a_3$$

$$\implies a_1 + a_3 - 2 - 2\nu \geq \sqrt{(a_1 + a_3 - 2)^2 - 4\nu a_3}.$$

Combining all four cases shows that

$$\frac{d\rho}{a_3} \leq 0 \text{ if } a_3 < R_1, \quad \text{ and } \frac{d\rho}{a_3} \geq 0 \text{ if } a_3 \geq R_2.$$ 

This completes the proof of Proposition 5.

SM3.10. Proof of Proposition 6. Consider a transformation $\tilde{a}_1 = 2 - a_1$, $\tilde{a}_2 = a_2$, and $\tilde{a}_3 = 2 - a_3$. Then the $\Phi$ matrix becomes

$$\Phi = \begin{pmatrix} 1 - a_1 & a_2 \\ -a_2 & a_3 - 1 \end{pmatrix} = \begin{pmatrix} \tilde{a}_1 - 1 & \tilde{a}_2 \\ -\tilde{a}_2 & 1 - \tilde{a}_3 \end{pmatrix}.$$ 

The eigenvalues of $\Phi$, hence also the spectral radius, depend on $(\tilde{a}_3, \tilde{a}_2, \tilde{a}_1)$ in the same way as $\Phi$ depends on $(a_1, a_2, a_3)$. Moreover, fixed $a_3$ and $\tilde{\nu} = a_2^2/(2 - a_1)$ translate into fixed $\tilde{a}_3$ and $\tilde{\nu} = \tilde{a}_2^2/\tilde{a}_1$. The condition $\tilde{\nu} \leq 2 - a_3 \leq 1 + \tilde{\nu}$ translates into $\tilde{\nu} \leq \tilde{a}_3 \leq 1 + \tilde{\nu}$. Hence Proposition 5 can be applied to obtain that for fixed $\tilde{a}_3$ and $\tilde{\nu}$, the spectral radius of $\Phi$ is minimized over $(\tilde{a}_1, \tilde{a}_2)$ by the choice

$$\tilde{a}_1^* = (\sqrt{\tilde{\nu} + 2} - \tilde{a}_3 - \sqrt{\tilde{\nu}})^2,$$

which leads to $2 - a_1^* = (\sqrt{\tilde{\nu} + a_3} - \sqrt{\tilde{\nu}})^2$ as stated in Proposition 6.

SM3.11. Proof of Corollary 7. To obtain HAMS-A, we set $\nu = a_1 = 1 - \sqrt{1 - \epsilon^2}$ in (44). Then

$$a_3^* = (\sqrt{2} - \sqrt{1 - \sqrt{1 - \epsilon^2}})^2.$$ 

In the SDE parameterization, we also have

$$a_3^* = e^{-\frac{\nu_2}{2}}(1 + \sqrt{1 - \epsilon^2}).$$

Therefore

$$e^{-\frac{\nu_2}{2}}(1 + \sqrt{1 - \epsilon^2}) = (\sqrt{2} - \sqrt{1 - \sqrt{1 - \epsilon^2}})^2$$

$$\implies \eta_2 = 2\epsilon \log \left[ \frac{1 + \sqrt{1 - \epsilon^2}}{(\sqrt{2} - \sqrt{1 - \sqrt{1 - \epsilon^2}})^2} \right].$$
Similarly for HAMS-B, let \( \tilde{\nu} = 2 - a_3 = 1 - \sqrt{1 - \epsilon^2} \) in (47). Then
\[
2 - a_1^* = (\sqrt{2} - \sqrt{1 - \sqrt{1 - \epsilon^2}})^2.
\]
According to SDE parameterization, we have
\[
a_1^* = 2 - e^{-\frac{\eta_1}{2}}(1 + \sqrt{1 - \epsilon^2}) \Rightarrow 2 - a_1^* = e^{-\frac{\eta_1}{2}}(1 + \sqrt{1 - \epsilon^2}).
\]
Hence
\[
e^{-\frac{\eta_1}{2}}(1 + \sqrt{1 - \epsilon^2}) = (\sqrt{2} - \sqrt{1 - \sqrt{1 - \epsilon^2}})^2
\]
\[
\Rightarrow \eta_1 = \frac{2}{\epsilon} \log \left[ \frac{1 + \sqrt{1 - \epsilon^2}}{(\sqrt{2} - \sqrt{1 - \sqrt{1 - \epsilon^2}})^2} \right].
\]
Thus \( \eta_1 = \eta_2 \). Taking expansions we have
\[
\eta_1 = \eta_2 = 2 + \frac{5}{12} \epsilon^2 + O(\epsilon^4).
\]

**SM3.12. Proof of Corollary 8.** According to (18), (19) and (45), for fixed \( \eta_1 \geq 0 \), the optimal choice of \( \eta_2 \) satisfies
\[
e^{-\eta_2 \epsilon/2}(1 + \sqrt{1 - \epsilon^2}) = \left\{3 - \sqrt{1 - \epsilon^2} - 2 \sqrt{2\epsilon}(1 + \sqrt{1 - \epsilon^2})^{-1/2}\right\} e^{-\eta_1 \epsilon/2}
\]
\[
\Rightarrow \eta_2 = \frac{2}{\epsilon} \log \left[ \frac{(1 + \sqrt{1 - \epsilon^2}) e^{\eta_1 \epsilon/2}}{3 - \sqrt{1 - \epsilon^2} - 2 \sqrt{2\epsilon}(1 + \sqrt{1 - \epsilon^2})^{-1/2}} \right].
\]
Then we take expansions with \( \eta_1 \) fixed,
\[
\eta_2 = 2 + \eta_1 + \frac{5}{12} \epsilon^2 + O(\epsilon^4).
\]
This proves the first statement in Corollary 8. Furthermore, the SDE studied in Bou-Rabee and Eberle (2020), Section 4.2, matches SDE (20) by setting \( \epsilon \) to \( \eta_1 \), \( \gamma \) to \( \eta_2 \), \( \beta = -1 \), and \( u = v = 1 \). Then a direct application of Lemma 4.12 in Bou-Rabee and Eberle (2020) shows that the optimal convergence rate is achieved at \( \eta_2 = 2 + \eta_1 \) for fixed \( \eta_1 \).

**SM3.13. Proofs of Propositions 7–9.** For a univariate target density \( \pi(x) \), the HAMS updates (6)–(8) can be equivalently stated as follows:

\[
x^* = x_0 - a_1 \nabla U(x_0) + a_2 u_0 + Z_0^{(1)},
\]
\[
u^* = (a_3 + \phi a_2 - 1) u_0 + (\phi - \phi a_1 - a_2) \nabla U(x_0) - \phi \nabla U(x^*) + \phi Z_0^{(1)} + Z_0^{(2)},
\]
\[
Z_0^{(1)}, Z_0^{(2)} \text{ are zero mean Gaussian with } \text{Var}(Z_0^{(1)}) = 2a_1 - a_1^2 - a_2^2,
\]
\[
\text{Var}(Z_0^{(2)}) = 2a_3 - a_3^2 - a_2^2 \text{ and } \text{Cov}(Z_0^{(1)}, Z_0^{(2)}) = 2a_2 - a_1 a_2 - a_2 a_3.
\]
The variance and covariance of \((x^*, u^*)\) in (SM53) and (SM54) given \((x_0, u_0)\) are then

\[
\begin{align*}
\text{Var}(x^*) &= \text{Var}(Z_0^{(1)}) , \\
\text{Cov}(x^*, u^*) &= \phi \text{Var}(Z_0^{(1)}) + \text{Cov}(Z_0^{(1)}, Z_0^{(2)}) , \\
\text{Var}(u^*) &= \phi^2 \text{Var}(Z_0^{(1)}) + 2\phi \text{Cov}(Z_0^{(1)}, Z_0^{(2)}) + \text{Var}(Z_0^{(2)}) .
\end{align*}
\]

Throughout this section, \(\text{Var}(x^*), \text{Var}(u^*)\) and \(\text{Cov}(x^*, u^*)\) are understood to be conditional on \((x_0, u_0)\). For matching between HAMS and existing algorithms, we first identify \(a_1, a_2, a_3, \phi\) to match the coefficients for \(\nabla U(x_0)\) and \(u_0\) in (SM53) and those of \(\nabla U(x_0), \nabla U(x^*),\) and \(u_0\) in (SM54) and then we compare \(\text{Var}(x^*), \text{Var}(u^*)\) and \(\text{Cov}(x^*, u^*)\). Even with modification to existing algorithms, the matching is nontrivial because there are five coefficients of \(\nabla U(x_0), \nabla U(x^*)\), and \(u_0\), but only four tuning parameters \(a_1, a_2, a_3, \phi\).

**Rescaled GJF.** The rescaled GJF update is

\[
\begin{align*}
\text{SM57} & \quad x^* = x_0 - \frac{\epsilon^2}{2 + \eta \epsilon} \nabla U(x_0) + \frac{\epsilon \sqrt{4 - \epsilon^2}}{2 + \eta \epsilon} u_0 + \frac{\epsilon}{2 + \eta \epsilon} W , \\
\text{SM58} & \quad u^* = \frac{2 - \eta \epsilon}{2 + \eta \epsilon} u_0 + \frac{\eta \epsilon^2 - 2 \epsilon}{\sqrt{4 - \epsilon^2}(2 + \eta \epsilon)} \nabla U(x_0) - \frac{\epsilon}{\sqrt{4 - \epsilon^2}} \nabla U(x^*) + \frac{4}{\sqrt{4 - \epsilon^2}(2 + \eta \epsilon)} W ,
\end{align*}
\]

where \(W \sim \mathcal{N}(0, 2\eta \epsilon)\). The coefficients of \(\nabla U(x_0), \nabla U(x^*),\) and \(u_0\) between (SM53)–(SM54) and (SM57)–(SM58) are matched, remarkably, by setting

\[
\begin{align*}
\text{SM59} & \quad a_1 = \frac{\epsilon^2}{2 + \eta \epsilon} , \\
\text{SM59} & \quad a_2 = \frac{\epsilon \sqrt{4 - \epsilon^2}}{2 + \eta \epsilon} , \\
\text{SM59} & \quad a_3 = \frac{4 - \epsilon^2}{2 + \eta \epsilon} , \\
\text{SM59} & \quad \phi = \frac{\epsilon}{\sqrt{4 - \epsilon^2}} .
\end{align*}
\]

Then using the HAMS formulas in (SM55)–(SM56), we find

\[
\begin{align*}
\text{Var}(x^*) &= \frac{2\eta \epsilon^4}{(2 + \eta \epsilon)^2} , \\
\text{Var}(u^*) &= \frac{32\eta \epsilon}{(4 - \epsilon^2)(2 + \eta \epsilon)^2} , \\
\text{Cov}(x^*, u^*) &= \frac{8\eta \epsilon^2}{(2 + \eta \epsilon)(2 + \alpha \epsilon)^2 \sqrt{4 - \epsilon^2}} ,
\end{align*}
\]

which match exactly the variances and covariance of \((x^*, u^*)\) in (SM57)–(SM58). Moreover, the choices of \((a_1, a_2, a_3)\) in (SM59) satisfy \(a_1 a_3 = a_2^2\), corresponding to a singular matrix \(A\) in HAMS-A. Thus the rescaled GJF is identical to HAMS-A except for the choice of \(\phi\). As stated by Proposition 2, HAMS-A uses \(\phi = a_2/(2 - a_1)\), which by the values in (SM59) leads to

\[
\phi = \frac{\epsilon^2}{2 + \eta \epsilon} + \frac{2\eta \epsilon}{\sqrt{4 - \epsilon^2}} .
\]

The difference in \(\phi\) is \(O(\epsilon^2)\).

**Rescaled BAOAB.** The rescaled BAOAB update is

\[
\begin{align*}
\text{SM60} & \quad x^* = x_0 - \frac{\epsilon^2}{4} (1 + e^{-\eta \epsilon}) \nabla U(x_0) + \frac{\epsilon \sqrt{4 - \epsilon^2}}{4} (1 + e^{-\eta \epsilon}) u_0 + \frac{\epsilon \sqrt{1 - e^{-2\eta \epsilon}}}{2} W , \\
\text{SM61} & \quad u^* = e^{-\eta \epsilon} u_0 - \frac{\epsilon e^{-\eta \epsilon}}{\sqrt{4 - \epsilon^2}} \nabla U(x_0) - \frac{\epsilon}{\sqrt{4 - \epsilon^2}} \nabla U(x^*) + 2 \sqrt{\frac{1 - e^{-2\eta \epsilon}}{4 - \epsilon^2}} W ,
\end{align*}
\]
SUPPLEMENTARY MATERIALS: IRREVERSIBLE METROPOLIS SAMPLING  SM21

where $W \sim \mathcal{N}(0, 1)$. The coefficients of $\nabla U(x_0)$, $\nabla U(x^*)$, and $u_0$ between (SM53)–(SM54) and (SM60)–(SM61) are matched, remarkably, by setting (SM62)

$$a_1 = \frac{\epsilon^2}{4}(1 + e^{-\eta})$$

where $W$ agrees with (SM60) and (SM61). By (SM62), the rescaled BAOAB corresponds to a singular $A$ matrix and only differs from HAMS-A by $O(\epsilon^2)$ in $\phi$. The $\phi$ value implied by HAMS-A is

$$\phi = \frac{\sqrt{4 - \epsilon^2}}{\sqrt{4 - \epsilon^2}} \neq \frac{\epsilon}{\sqrt{4 - \epsilon^2}}.$$

The difference in $\phi$ is $O(\epsilon^2)$.

**IL with full-step momentum.** The IL update (53)–(54) can be rewritten as

(SM63)  
$$x^* = x_0 - \epsilon^2 \left(1 - \frac{\epsilon}{2}\right) \nabla U(x_0) + \epsilon \left(1 - \frac{\epsilon}{2}\right) u_{-\frac{1}{2}} + \frac{\epsilon}{2} \sqrt{c(2 - c)} W;$$

(SM64)  
$$u_{\frac{1}{2}} = (1 - c) u_{-\frac{1}{2}} - \epsilon (1 - c) \nabla U(x_0) + \sqrt{c(2 - c)} W,$$

where $W \sim \mathcal{N}(0, 1)$. With the full-step momentum in Proposition 7, (SM63)–(SM64) leads to

(SM65)  
$$x^* = x_0 - \frac{\epsilon^2}{2} \left(1 - \frac{\epsilon}{2}\right) \nabla U(x_0) + \frac{\epsilon}{2} \sqrt{4 - \epsilon^2} \left(1 - \frac{\epsilon}{2}\right) u_0 + \frac{\epsilon}{2} \sqrt{c(2 - c)} W.$$

(SM66)  
$$u^* = (1 - \frac{\epsilon}{2}) u_0 - \frac{\epsilon(1 - \epsilon)}{\sqrt{4 - \epsilon^2}} \nabla U(x_0) - \frac{\epsilon}{\sqrt{4 - \epsilon^2}} \nabla U(x^*) + \frac{2\epsilon}{\sqrt{4 - \epsilon^2}} \nabla U(x^*) + \frac{2\epsilon}{\sqrt{4 - \epsilon^2}} \nabla U(x^*) W,$$

where $W \sim \mathcal{N}(0, 1)$. By substituting $\tilde{c} = 1 - e^{-\eta}$, we see that (SM65)–(SM66) becomes identical to (SM60) and (SM61). Hence IL with full-step momentum is equivalent to rescaled BAOAB and matches HAMS-A in the same manner.

**Modified OBABO.** The modified OBABO update is

(SM67)  
$$x^* = x_0 - (1 - \sqrt{1 - c^2}) \nabla U(x_0) + \epsilon \sqrt{c} u_0 + \epsilon \sqrt{1 - c} W_1,$$

(SM68)  
$$u^* = cu_0 - \frac{\sqrt{c} \epsilon}{1 + \sqrt{1 - c^2}} \nabla U(x_0) - \frac{\sqrt{c} \epsilon}{1 + \sqrt{1 - c^2}} \nabla U(x^*) + \sqrt{c}(1 - c) W_1 + \sqrt{1 - c} W_2,$$

where $W_1, W_2 \sim \mathcal{N}(0, 1)$ independently. The coefficients of $\nabla U(x_0)$, $\nabla U(x^*)$, and $u_0$ between (SM53)–(SM54) and (SM67)–(SM68) are matched, remarkably, by setting

(SM69)  
$$a_1 = 1 - \sqrt{1 - c^2}, \quad a_2 = \epsilon \sqrt{c}, \quad a_3 = 1 + c \sqrt{1 - c^2}, \quad \phi = \frac{\sqrt{c} \epsilon}{1 + \sqrt{1 - c^2}}.$$
Using the HAMS formulas in (SM55)–(SM56), we find

\[ \text{Var}(x^*) = \epsilon^2(1 - c), \quad \text{Var}(u^*) = 1 - c^2, \quad \text{Cov}(x^*, u^*) = \epsilon(1 - c)\sqrt{c}, \]

which match exactly the variances and covariance of \((x^*, u^*)\) in (SM67) and (SM68). The \(\phi\) choice in (SM69) also agrees with the default value \(\phi = a_2/(2 - a_1)\) for HAMS.

**Modified VEC.** The modified VEC update is

\[
\begin{align*}
SM70 \quad x^* &= x_0 - \frac{c^2}{2} \nabla U(x_0) + \frac{2\epsilon - \eta c^2}{2} u_0 + \frac{\sqrt{2\eta} \epsilon^{3/2}}{2} W_1 + \frac{\sqrt{6\eta} \epsilon^{3/2}}{6} W_2 \\
\quad u^* &= \left(1 - \eta c + \frac{\eta^2 c^2}{2}\right) u_0 + \left(\frac{\eta c^2 - \epsilon^2}{2} - \frac{\epsilon^3}{4}\right) \nabla U(x_0) - \frac{\epsilon}{2} \nabla U(x^*) \\
&\quad + \frac{\sqrt{2\eta}}{2} (2 - \eta c) W_1 - \frac{\sqrt{6}}{6} (\eta \epsilon)^{3/2} W_2,
\end{align*}
\]

where \(W_1, W_2 \sim \mathcal{N}(0, 1)\), independently. The coefficients of \(\nabla U(x_0), \nabla U(x^*)\), and \(u_0\) between (SM53)–(SM54) and (SM70)–(SM71) are matched, remarkably, by setting

\[
SM72 \quad a_1 = \frac{c^2}{2}, \quad a_2 = \epsilon - \frac{\eta c^2}{2}, \quad a_3 = 2 - \frac{\epsilon}{4}(2 - \eta c)(2\eta + \epsilon), \quad \phi = \frac{\epsilon}{2}.
\]

Using the HAMS formulas in (SM55)–(SM56), we find

\[
\begin{align*}
\text{Var}(x^*) &= \eta \epsilon^3 - \frac{\eta^2 \epsilon^4 + \epsilon^4}{4}, \\
\text{Var}(u^*) &= 2\eta \epsilon - 2\eta^2 \epsilon^2 + \frac{\epsilon}{16} [8\eta^2 (1 + 2\eta^2) - 4\epsilon^4(1 + \eta^2 + \eta^4) + 4\eta^4 - \epsilon^5], \\
\text{Cov}(x^*, u^*) &= \eta \epsilon^2 - \frac{\eta^2 \epsilon^3 + \epsilon^2}{4} \eta (1 + \eta^2) - \frac{1}{8} \epsilon^3,
\end{align*}
\]

whereas according to (SM70) and (SM71) in the modified VEC update,

\[
\begin{align*}
\text{Var}(x^*) &= \frac{2\eta}{3} \epsilon^3, \quad \text{Var}(u^*) = 2\eta \epsilon - 2\eta^2 \epsilon^2 + \frac{2}{3} \eta^3 \epsilon^3, \quad \text{Cov}(x^*, u^*) = \eta \epsilon^2 - \frac{2}{3} \eta^2 \epsilon^3.
\end{align*}
\]

The differences between the corresponding variances and covariances are \(O(\epsilon^3)\). The \(\phi\) choice in (SM72) only differs from the default value \(\phi = a_2/(2 - a_1)\) by \(O(\epsilon^2)\).

**Shifted HAMS.** The update (64) in shifted HAMS can be rewritten as

\[
\begin{align*}
SM73 \quad x^* &= x_0 - a_1 \nabla U(\hat{x}) + (ba_1 + a_2) u_0 + Z_0^{(1)}, \\
SM74 \quad u^* &= (a_3 + ba_2 - 1) u_0 - a_2 \nabla U(\hat{x}) + Z_0^{(2)}.
\end{align*}
\]

Our matching approach using shifted HAMS is similar to using original HAMS. For each method, we first identify \(a_1, a_2, a_3, b\) to match the coefficients for \(\nabla U(\hat{x})\) and \(u_0\) in (SM73)–(SM74) and then we compare \(\text{Var}(x^*), \text{Var}(u^*)\) and \(\text{Cov}(x^*, u^*)\). While the first step is relatively straightforward with four coefficients of \(\nabla U(\hat{x})\) and \(u_0\) and four tuning parameters \(a_1, a_2, a_3, b\), the close matching in the variances and covariance remains nontrivial.

**Modified ABOBA.** Consider modified ABOBA update with \(b\) to be determined:

\[
\begin{align*}
SM75 \quad x^* &= x_0 - \frac{b\epsilon}{2}(1 + e^{-\eta c}) \nabla U(\hat{x}) + b(1 + e^{-\eta c}) u_0 + b \sqrt{1 - e^{-2\eta c}W}, \\
SM76 \quad u^* &= e^{-\eta c} u_0 - \frac{\epsilon}{2} (1 + e^{-\eta c}) \nabla U(\hat{x}) + \sqrt{1 - e^{-2\eta c}W},
\end{align*}
\]
where $W \sim \mathcal{N}(0,1)$. Matching the coefficients of $\nabla U(\tilde{x})$ and $u_0$ between (SM73)–(SM74) and (SM75)–(SM76), we obtain
\[
a_1 = \frac{1}{2} (1 + e^{-\eta \epsilon})(1 - \sqrt{1 - \epsilon^2}), \quad a_2 = \frac{1}{2} (1 + e^{-\eta \epsilon})(1 + \sqrt{1 - \epsilon^2}), \quad b = \frac{1}{\sqrt{1 - \epsilon^2}}.
\]
The variances and covariance implied by shifted HAMS are
\[
\text{Var}(x^*) = \frac{1}{2} (1 - \sqrt{1 - \epsilon^2})(1 - e^{-2\eta \epsilon}) = \frac{\eta^3}{2} - \frac{\eta^2 \epsilon^4}{2} + \left( \frac{\eta}{8} + \frac{\eta^3}{3} \right) \epsilon^5 + O(\epsilon^6),
\]
\[
\text{Var}(u^*) = \frac{1}{2} (1 + \sqrt{1 - \epsilon^2})(1 - e^{-2\eta \epsilon}) = 2\eta \epsilon - 2\eta^2 \epsilon^2 + \left( \frac{4}{3} \eta^3 - \eta \right) \epsilon^3 + O(\epsilon^4),
\]
\[
\text{Cov}(x^*, u^*) = \frac{\epsilon}{2} (1 - e^{-2\eta \epsilon}) = \eta^2 \epsilon^2 - \eta^2 \epsilon^3 + \frac{2\eta^3 \epsilon^4}{3} + O(\epsilon^5),
\]
whereas those computed from (SM75)–(SM76) are
\[
\text{Var}(x^*) = \frac{(1 - e^{-2\eta \epsilon})(1 - \sqrt{1 - \epsilon^2})^2}{\epsilon^2} = \frac{\eta^3}{2} - \frac{\eta^2 \epsilon^4}{2} + \left( \frac{\eta}{4} + \frac{\eta^3}{3} \right) \epsilon^5 + O(\epsilon^6),
\]
\[
\text{Var}(u^*) = 1 - e^{-2\eta \epsilon} = 2\eta \epsilon - 2\eta^2 \epsilon^2 + \frac{4}{3} \eta^3 \epsilon^3 + O(\epsilon^4),
\]
\[
\text{Cov}(x^*, u^*) = \frac{1 - \sqrt{1 - \epsilon^2}}{\epsilon} (1 - e^{-2\eta \epsilon}) = \eta^2 \epsilon^2 - \eta^2 \epsilon^3 + \left( \frac{2}{3} \eta^3 + \frac{\eta}{4} \right) \epsilon^4 + O(\epsilon^5).
\]
The differences between the corresponding variances and covariances are $O(\epsilon^3)$.

**Modified SPV.** Consider modified SPV update with $b$ to be determined:
\[
\text{(SM77)} \quad x^* = x_0 - \frac{b(1 - e^{-\eta \epsilon})}{\eta} \nabla U(\tilde{x}) + b(1 + e^{-\eta \epsilon})u_0 + b \sqrt{1 - e^{-2\eta \epsilon}} W_1,
\]
\[
\text{(SM78)} \quad u^* = e^{-\eta \epsilon}u_0 - \frac{1 - e^{-\eta \epsilon}}{\eta} \nabla U(\tilde{x}) + \sqrt{1 - e^{-2\eta \epsilon}} W_1.
\]
where $W \sim \mathcal{N}(0,1)$. Matching the coefficients of $\nabla U(\tilde{x})$ and $u_0$ between (SM73)–(SM74) and (SM77)–(SM78), we obtain
\[
a_1 = \frac{1}{2} \left\{ 1 + e^{-\eta \epsilon} - \sqrt{(1 + e^{-\eta \epsilon})^2 - \frac{4(1 - e^{-\eta \epsilon})^2}{\eta^2}} \right\}, \quad a_2 = \frac{1 - e^{-\eta \epsilon}}{\eta},
\]
\[
a_3 = \frac{1}{2} \left\{ 1 + e^{-\eta \epsilon} + \sqrt{(1 + e^{-\eta \epsilon})^2 - \frac{4(1 - e^{-\eta \epsilon})^2}{\eta^2}} \right\},
\]
\[
b = \frac{\eta \left( 1 + e^{-\eta \epsilon} - \sqrt{(1 + e^{-\eta \epsilon})^2 - \frac{4(1 - e^{-\eta \epsilon})^2}{\eta^2}} \right)}{2(1 - e^{-\eta \epsilon})}.
\]
The variances and covariance implied by shifted HAMS are
\[
\text{Var}(x^*) = \frac{\eta^3 \epsilon^3}{2} - \frac{\eta^2 \epsilon^4}{2} + \left( \frac{\eta}{8} + \frac{\eta^3}{4} \right) \epsilon^5 + O(\epsilon^6),
\]
\[
\text{Var}(u^*) = 2\eta \epsilon - 2\eta^2 \epsilon^2 + \left( \frac{2}{3} \eta^3 + \frac{\eta}{4} \right) \epsilon^3 + O(\epsilon^4),
\]
\[
\text{Cov}(x^*, u^*) = \frac{(1 - e^{-\eta \epsilon})^2}{\eta} = \eta^2 \epsilon^2 - \eta^2 \epsilon^3 + \frac{7\eta^3 \epsilon^4}{12} + O(\epsilon^5),
\]
whereas according to (SM77) and (SM78) in the modified SPV update,
\[
\text{Var}(x^*) = \frac{\eta^3 \epsilon^3}{2} - \frac{\eta^2 \epsilon^4}{2} + \left( \frac{\eta}{4} + \frac{\eta^3}{4} \right) \epsilon^5 + O(\epsilon^6),
\]
\[
\text{Var}(u^*) = 2\eta \epsilon - 2\eta^2 \epsilon^2 + \frac{4\eta^3}{3} \epsilon^3 + O(\epsilon^4),
\]
\[
\text{Cov}(x^*, u^*) = \eta \epsilon^2 - \eta^2 \epsilon^3 + \frac{7\eta^3 \epsilon^4}{12} + \frac{\eta^4}{4} + O(\epsilon^5).
\]
The differences between the corresponding variances and covariances are \(O(\epsilon^3)\).

**Modified Mannella’s leapfrog.** Consider modified Mannella’s leapfrog update with \(b\) to be determined:

\[
(x^*) = x_0 - b \frac{2\epsilon}{2 + \eta \epsilon} \nabla U(\tilde{x}) + b \frac{4}{2 + \eta \epsilon} u_0 + b \frac{2\sqrt{2\eta}}{2 + \eta \epsilon} W,
\]
\[
(u^*) = 2 - \frac{\eta \epsilon}{2 + \eta \epsilon} u_0 - \frac{2\epsilon}{2 + \eta \epsilon} \nabla U(\tilde{x}) + \frac{2\sqrt{2\eta}}{2 + \eta \epsilon} W,
\]
where \(W \sim N(0, \epsilon)\). The coefficients of \(\nabla U(\tilde{x})\) and \(u_0\) between (SM73)–(SM74) and (SM79)–(SM80) are matched by setting

\[
a_1 = \frac{2(1 - \sqrt{1 - \epsilon^2})}{2 + \eta \epsilon}, \quad a_2 = \frac{2\epsilon}{2 + \eta \epsilon},
\]
\[
a_3 = \frac{2(1 + \sqrt{1 - \epsilon^2})}{2 + \eta \epsilon}, \quad b = \frac{1 - \sqrt{1 - \epsilon^2}}{\epsilon}.
\]
The variances and covariance given by shifted HAMS are

\[
\text{Var}(x^*) = \frac{4\eta \epsilon (1 - \sqrt{1 - \epsilon^2})}{2 + \eta \epsilon} = \frac{\eta^3 \epsilon^3}{2} - \frac{\eta^2 \epsilon^4}{2} + \frac{\eta}{8} (1 + 3\eta^2) \epsilon^5 + O(\epsilon^6),
\]
\[
\text{Var}(u^*) = \frac{4\eta \epsilon (1 + \sqrt{1 - \epsilon^2})}{2 + \eta \epsilon} = 2\eta \epsilon - 2\eta^2 \epsilon^2 + \frac{\eta}{2} (3\eta^2 - 1) \epsilon^3 + O(\epsilon^4),
\]
\[
\text{Cov}(x^*, u^*) = \frac{4\eta^2 \epsilon^2}{2 + \eta \epsilon} = \eta \epsilon^2 - \eta^2 \epsilon^3 + \frac{3\eta^3 \epsilon^4}{4} + O(\epsilon^5).
\]
The variances and covariance given by (SM79) and (SM80) are

\[
\text{Var}(x^*) = \frac{\eta^3 \epsilon^3}{2} - \frac{\eta^2 \epsilon^4}{2} + \frac{\eta}{8} (2 + 3\eta^2) \epsilon^5 + O(\epsilon^6),
\]
\[
\text{Var}(u^*) = 2\eta \epsilon - 2\eta^2 \epsilon^2 + \frac{3\eta^3 \epsilon^3}{2} + O(\epsilon^4),
\]
\[
\text{Cov}(x^*, u^*) = \eta \epsilon^2 - \eta^2 \epsilon^3 + \frac{\eta}{4} (1 + 3\eta^2) \epsilon^4 + O(\epsilon^5).
\]
The differences between the corresponding variances and covariances are \(O(\epsilon^3)\).

**SM4. Details and additional results for numerical experiments.**

**SM4.1. Double well.** In the double well experiment, there is no preconditioning. In Algorithms 1–4 we take \(L = I\). For HAMS-A, we set \(a = 1 - \sqrt{1 - \epsilon^2}, b = e^{-\eta_2 \epsilon/2}(1 + \sqrt{1 - \epsilon^2})\) in Algorithm 1. For HAMS-B, we first set \(\tilde{a} = e^{-\eta_1 \epsilon/2}(1 + \sqrt{1 - \epsilon^2})\) in Algorithm 1.
\sqrt{1-c^2}, \hat{b} = 1 - \sqrt{1-c^2} and then use the transformation \( a = 2 - \hat{a}, b = (\hat{a}\hat{b})/(2 - \hat{a}) \) in Algorithm 1. For HAMS-1/2/3, we set \( c_1 = e^{-k\epsilon^2/2} \) for \( k = 1, 2, 3 \) respectively and \( c_2 = e^{-\eta_2\epsilon^2/2} \), define \( a_1, a_2, a_3 \) by (18), and then apply Algorithm 1. For BAOAB, ABOBA and OBABO we set \( c = e^{-\eta_2\epsilon^2/2} \) in Algorithms 2–4 with \( \eta \) set to \( \eta_2 \) in HAMS.

**Equivalence of temperatures.** We show that \( T_{C1} = T_{C2} = T_K \). By the definition (75), \( x \) and \( u \) are independent and \( u \sim N(0, T) \). Thus \( T_K = E[u^2] = T \). For the configurational temperatures, we use Stein’s identity (Ley et al., 2017), which states that for any differentiable function \( f(x) \) such that \( f(x)\pi(x) \to 0 \) as \( x \to \pm\infty \),

\[
E \left[ f(x) \frac{\nabla \pi(x)}{\pi(x)} \right] = -E[f'(x)].
\]

Notice that \( \nabla U(x) = -T \frac{\nabla \pi(x)}{\pi(x)} \). Taking \( f(x) = x \) shows that

\[
T_{C1} = E[x \cdot \nabla U(x)] = -TE \left[ \frac{x \nabla \pi(x)}{\pi(x)} \right] = T.
\]

Moreover, taking \( f(x) = \nabla U(x) \) shows that

\[
E[(\nabla U(x))^2] = -TE \left[ \nabla U(x) \frac{\nabla \pi(x)}{\pi(x)} \right] = TE[\nabla^2 U(x)].
\]

and hence

\[
T_{C2} = \frac{E[(\nabla U(x))^2]}{E[\nabla^2 U(x)]} = T.
\]

**Density estimation.** In addition to the temperatures, we report the performance of density estimation. Following Leimkuhler and Matthews (2013), the error in density estimation is computed by dividing the interval \([-2, 2]\) into 16 equal sized bins and compare the empirical density with the truth obtained from numerical integration. The left panel of Figure SM1 shows the errors on log scale. Comparison between the methods is consistent with that in temperature estimation. When \( \epsilon \) is small, HAMS-\( k \) has better performance as \( k \) increases (including HAMS-A with \( k = 0 \)). Moreover, HAMS-A, BAOAB, ABOBA and OBABO have comparable performance for small \( \epsilon \). The error of HAMS-B is the smallest for \( \epsilon \leq 0.12 \) but quickly increases afterwards. The overall best performance is achieved by HAMS-1 at \( \epsilon = 0.24 \).

Figure SM1 also shows density plots, produced using \texttt{density()} in R, from an individual run when \( \epsilon = 0.24 \). This confirms that HAMS-1 best tracks the shape of the true density.

**Error calculation.** We describe how the errors are calculated in Figure 1 and Figure SM1. Let \( \{x_{ij}\} \) and \( \{u_{ij}\} \) be the samples collected, indexed by \( i = 1, \ldots, N \) draws and \( j = 1, \ldots, J \) repetitions. Let \( T \) be the true temperature. Then

\[
T_{C1}^{(j)} = \frac{1}{N} \sum_{i=1}^{N} x_{ij} \cdot \nabla U(x_{ij}), \quad \text{Error in } T_{C1} = \sqrt{\frac{1}{J} \sum_{j=1}^{J} (T_{C1}^{(j)} - T)^2},
\]

\[
T_{C2}^{(j)} = \frac{\sum_{i=1}^{N} (\nabla U(x_{ij}))^2}{\sum_{i=1}^{N} \nabla^2 U(x_{ij})}, \quad \text{Error in } T_{C2} = \sqrt{\frac{1}{J} \sum_{j=1}^{J} (T_{C2}^{(j)} - T)^2},
\]

\[
T_K^{(j)} = \frac{1}{N} \sum_{i=1}^{N} u_{ij}^2, \quad \text{Error in } T_K = \sqrt{\frac{1}{J} \sum_{j=1}^{J} (T_K^{(j)} - T)^2}.
\]
For errors in densities, let $\omega_k^*$ be the area under the true density curve in the $k$th bin, $k = 1, \ldots, M (= 16)$. Then

$$
\omega_{kj} = \frac{1}{N} \sum_{i=1}^{N} I\{x_{ij} \in k\text{th bin}\}, \quad e_j = \frac{1}{M} \sum_{k=1}^{M} |\omega_{kj} - \omega_k^*|, \quad \text{Error in density} = \sqrt{\frac{1}{J} \sum_{j=1}^{J} e_j^2}.
$$

The normalizing constants needed to evaluate $\omega_k^*$s are obtained from Mathematica.

**SM4.2. Stochastic volatility. Detailed expressions.** We apply preconditioning and use the default parameter choices implied by Proposition 5 and 6. For HAMS-A, we set $a = 1 - \sqrt{1 - \epsilon^2}, b = (\sqrt{2} - \sqrt{\alpha})^2$ in Algorithm 1. For HAMS-B, we first set $\tilde{b} = 1 - \sqrt{1 - \epsilon^2}, \tilde{a} = (\sqrt{2} - \sqrt{\tilde{b}})^2$ and then use the transformation $a = 2 - \tilde{a}, b = (\tilde{a}\tilde{b})/(2 - \tilde{a})$ in Algorithm 1. For HAMS-1/2/3, we set for $k = 1, 2, 3$

$$
c_1 = \exp(-k\epsilon^2/2), \quad c_2 = \max \left\{ \frac{1}{2}, \left\{ \frac{3 - \sqrt{1 - \epsilon^2}}{1 + \sqrt{1 - \epsilon^2}} - 2\sqrt{2}\epsilon(1 + \sqrt{1 - \epsilon^2})^{-3/2} \right\} c_1 \right\},
$$

define $a_1, a_2, a_3$ by (18), and then apply Algorithm 1. Here we restrict $c_2 \leq \frac{1}{2}$ to ensure the condition $a_1 \leq 1 + \nu$ in Proposition 5 is satisfied. For BAOAB, ABOBA and OBABO, we set $c$ as below to be consistent with the choice in HAMS-A,

$$
c = \left\{ \frac{3 - \sqrt{1 - \epsilon^2}}{1 + \sqrt{1 - \epsilon^2}} - 2\sqrt{2}\epsilon(1 + \sqrt{1 - \epsilon^2})^{-3/2} \right\}.
$$

Next, we give the detailed expressions for stochastic volatility model. Using (80) and (81), we derive the conditional density of $x$ given $y$.
\begin{equation}
\begin{aligned}
p(x|y, \beta, \sigma, \phi) &\propto p(x_1) \prod_{t=2}^{T} p(x_t|x_{t-1}, \varphi, \sigma) \prod_{t=1}^{T} p(y_t|x_t, \beta) \\
&\quad \cdot \mathcal{N}(y|0, \sigma^2) \\
&\quad \cdot \mathcal{N}(x|0, C) \\
&\propto \exp \left\{ -\frac{1}{2} x^T C^{-1} x \right\} \beta^{-T} \exp \left\{ -\frac{1}{2} \sum_{t=1}^{T} (x_t + \beta^{-2} y_t^2 \exp(-x_t)) \right\}.
\end{aligned}
\end{equation}

where the \((i, j)\) position of \(C\) matrix is given by \(\phi_{i-j} \sigma^2 / (1 - \varphi^2)\). After suppressing the dependency on \(y\) and the parameters, the negative log density or potential function of \(x\) is

\[U(x) = \frac{1}{2} x^T C^{-1} x + \frac{1}{2} \sum_{t=1}^{T} (x_t + \beta^{-2} y_t^2 \exp(-x_t)).\]

The gradient is

\[\nabla U(x) = C^{-1} x - \frac{1}{2} \beta^{-2} y \exp(-x) + \frac{1}{2} 1,\]

where \(1\) is a vector of all 1’s. The Hessian is

\[\nabla^2 U(x) = C^{-1} + \frac{1}{2} \text{diag}[\beta^{-2} y^2 \exp(-x)].\]

The square \(y^2\) is taken component-wise. Using the relation between \(y\) and \(x\), the diagonal elements in the second term can be expressed as

\[\beta^{-2} y^2 \exp(-x) = \beta^{-2} \exp(-x) z^2 \beta^2 \exp(x) = z^2.\]

Hence \(E[\nabla^2 U(x)] = C^{-1} + \frac{1}{2} I\). For preconditioning, we use \(\hat{\Sigma} = (C^{-1} + \frac{1}{2} I)^{-1}\).

**Means and variance of sample means.** In the left panel of Figure SM2, we plot the average of sample means of all the 1000 latent coordinates across 50 repetitions. The curves are shifted by a constant to be aligned with zero. We see that the overall shapes of average sample means are similar among all methods. However, the variances of sample means as shown in the right panel of Figure SM2 indicate that HAMS methods are more consistent than the remaining methods.

**SM4.3. Log-Gaussian Cox model. Detailed expressions.** The tuning parameters \((a, b), (c_1, c_2),\) and \(c\) are set in the same manner as in Section SM4.2. We now provide the necessary expressions for the log-Gaussian Cox model. From (82), we obtain the potential function of \(x\) (given \(y\)),

\[U(x) = \frac{1}{2} x^T C^{-1} x - \sum_{i,j} (y_{ij} x_{ij} - n^{-1} \exp(x_{ij} + \mu)).\]

The gradient is

\[\nabla U(x) = C^{-1} x - y + n^{-1} \exp(x + \mu).\]

The Hessian is

\[\nabla^2 U(x) = C^{-1} + n^{-1} \text{diag}[x + \mu].\]

Because marginally \(x \sim \mathcal{N}(0, C)\), we take the expectation \(E[\nabla^2 U(x)] = C^{-1} + n^{-1} \text{diag}[\sigma^2 / 2 + \mu]\). We take \(\hat{\Sigma} = \{C^{-1} + n^{-1} \text{diag}[\sigma^2 / 2 + \mu]\}^{-1}\) for preconditioning.
Means and variance of sample means. The average sample means for the Cox model line up consistently among different methods as shown by Figure SM3. But the variances of sample means can be differentiated in the right panel of Figure SM3, in spite of the overlaps. We see that HAMS-A has the smallest variation across repeated simulations, followed by HAMS-1, HAMS-2, HAMS-3, HAMS-B, OBABO, ABOBA and BAOAB.
REFERENCES

N. Bou-Rabee and A. Eberle, *Markov Chain Monte Carlo Methods*, Lecture Notes. Hausdorff School on MCMC: Recent developments and new connections, 2020.

K. Burrage, I. Lenane, and G. Lythe, *Numerical methods for second-order stochastic differential equations*, SIAM Journal on Scientific Computing, 29 (2007), pp. 245–264.

R. Kan, *From moments of sum to moments of product*, Journal of Multivariate Analysis, 99 (2008), pp. 542–554.

B. Leimkuhler and C. Matthews, *Robust and efficient configurational molecular sampling via Langevin dynamics*, The Journal of Chemical Physics, 138 (2013), p. 174102.

C. Ley, G. Reinert, and Y. Swan, *Stein’s method for comparison of univariate distributions*, Probability Surveys, 14 (2017), pp. 1–52.

Z. Song and Z. Tan, *Hamiltonian assisted Metropolis sampling*, Journal of the American Statistical Association. to appear (arXiv preprint:2005.08159), 2021.