

Supplementary Materials for “Wavelet Monte Carlo - a principle for sampling from complex distributions”

Walter R. Gilks¹, Lukas Cironis², Stuart Barber^{3*}

¹Unaffiliated, Leeds, UK.

²University of Leeds, Leeds, UK.

^{3*}University of Leeds, Leeds, UK.

*Corresponding author(s). E-mail(s): S.barber@leeds.ac.uk;

Contributing authors: wally.gilks@gmail.com; cironis.lukas@gmail.com;

A Appendix

A.1 Geometric convergence in the Besov-space norm of approximations to h

Throughout this section we assume regularity conditions (i)–(vi) on h and ψ . We provide proofs showing that these regularity conditions are sufficient to ensure that the Besov space norm for h , given in (8), is finite, and that approximation (10) converges geometrically to h in this norm.

We begin by establishing the finiteness of

$$H_j = 2^{-j/2} \sum_{i \in \mathbb{Z}} |\tilde{h}_{ji}|, \quad (45)$$

separately for scales $j \geq 0$ and $j < 0$. For $j \geq 0$, Lemma 1 shows that, as $j \rightarrow \infty$, H_j decays geometrically at a rate r_ψ , the number of vanishing moment of ψ . For $j < 0$, Lemma 2 shows that, as $j \rightarrow -\infty$, H_j decays geometrically at a rate dependent on s_h , the decay rate in the tails of h . Lemma 3 then combines these results to show that restricting scales j to the range $j_0 < j < j_1$, as we do in Section 7.3, leads to discrepancies between h and its WMC approximation which decay geometrically as

$(j_0, j_1) \rightarrow (-\infty, \infty)$. Corollary 4 then shows that the Besov space norm for h , given in (8), is finite.

Lemma 1. For $j \in \mathbb{Z}$, $j \geq 0$,

$$H_j < C_1 2^{-j r_\psi}, \quad (46)$$

where $C_1 = C_\psi C_h a^{r_\psi+1} (1 + a + 2(s_h - 1))^{-1}$.

Proof. From Taylor's Theorem using the Lagrange form of remainder term, for given $x_0, x_1 \in \mathbb{R}$, $-\infty < x_0 < x_1 < \infty$, and for all $x, z \in [x_0, x_1]$,

$$|h(x) - h_z(x)| \leq \frac{1}{r_\psi!} \sup_{\xi \in [x_0, x_1]} |h^{(r_\psi)}(\xi)| \cdot |x - z|^{r_\psi} \quad (47)$$

where

$$h_z(x) = \sum_{k=0}^{r_\psi-1} \frac{1}{k!} h^{(k)}(z) (x - z)^k. \quad (48)$$

From (45), using (4,7),

$$\begin{aligned} H_j &= \sum_{i \in \mathbb{Z}} \left| \int h(x) \psi(2^j x - i) dx \right| = 2^{-j} \sum_{i \in \mathbb{Z}} \left| \int_0^a h(2^{-j}(y + i)) \psi(y) dy \right| \\ &\leq 2^{-j} \sum_{i \in \mathbb{Z}} \left\{ \int_0^a |h(2^{-j}(y + i)) - h_{z_{ji}}(2^{-j}(y + i))| \cdot |\psi(y)| dy \right. \\ &\quad \left. + \left| \int h_{z_{ji}}(2^{-j}(y + i)) \psi(y) dy \right| \right\}, \end{aligned} \quad (49)$$

for an arbitrary choice of z_{ji} . Setting $z_{ji} = 2^{-j}i$ and substituting $x = 2^{-j}(y + i)$; $\xi = 2^{-j}(\eta + i)$; $x_0 = 2^{-j}i$; $x_1 = 2^{-j}(a + i)$; $z = z_{ji}$ into (47) then integrating and summing we obtain

$$\begin{aligned} &\sum_{i \in \mathbb{Z}} \int_0^a |h(2^{-j}(y + i)) - h_{z_{ji}}(2^{-j}(y + i))| |\psi(y)| dy \\ &\leq C_\psi \sum_{i \in \mathbb{Z}} \int_0^a \frac{1}{r_\psi!} 2^{-j r_\psi} \sup_{\eta \in [0, a]} |h^{(r_\psi)}(2^{-j}(\eta + i))| |y|^{r_\psi} dy \\ &< 2^{-j r_\psi} a^{r_\psi+1} C_\psi C_h \sum_{i \in \mathbb{Z}} \sup_{\eta \in [0, a]} (1 + 2^{-j}|\eta + i|)^{-s_h} \\ &< 2^{-j r_\psi} a^{r_\psi+1} C_\psi C_h \left\{ \sum_{i < -a} (1 + 2^{-j}|a + i|)^{-s_h} \right. \end{aligned}$$

$$\begin{aligned}
& + \sum_{i=-a}^0 1 + \sum_{i>0} (1 + 2^{-j}|i|)^{-s_h} \Big\} \\
& < 2^{-jr_\psi} a^{r_\psi+1} C_\psi C_h \left\{ a + 1 + 2 \int_0^\infty (1 + 2^{-j}u)^{-s_h} du \right\} \\
& = 2^{-jr_\psi} a^{r_\psi+1} C_\psi C_h \left(a + 1 + 2 \frac{2^j}{s_h - 1} \right), \tag{50}
\end{aligned}$$

using (3,12). Substituting the same ingredients into (48) then integrating we obtain

$$\int h_{z_{ji}}(2^{-j}(y+i)) \psi(y) dy = \int \sum_{k=0}^{r_\psi-1} \frac{1}{k!} h^{(k)}(2^{-j}i) y^k \psi(y) 2^{-jk} dy = 0, \tag{51}$$

using (2). Substituting (50,51) into (49), we obtain

$$\begin{aligned}
H_j & < 2^{-j} 2^{-jr_\psi} a^{r_\psi+1} C_\psi C_h \left(a + 1 + 2 \frac{2^j}{s_h - 1} \right) \\
& < a^{r_\psi+1} C_\psi C_h \left(a + 1 + 2 \frac{1}{s_h - 1} \right) 2^{-jr_\psi} = C_1 2^{-jr_\psi}, \tag{52}
\end{aligned}$$

noting that here we consider only $j \geq 0$. $\square$

Lemma 2. For $j \in \mathbb{Z}$, $j < 0$,

$$H_j < C_2 2^{j(1-s_h^{-1})}, \tag{53}$$

where $C_2 = 6aC_\psi C_h (s_h - 1)^{-1}$.

Proof. For $j, k \in \mathbb{Z}$, $j < 0$, and an arbitrary choice of $\gamma \in (0, 1)$, let

$$x_{jk} = k2^{-\gamma j}, \quad I_{jk} = [x_{jk}, x_{j,k+1}), \quad I_{jk}(x) = \mathbb{I}[x \in I_{jk}], \tag{54}$$

notationally suppressing their dependence on γ , and let

$$J_{jk} = [k2^{-j}, (k+a)2^{-j}).$$

Note that, as $j \rightarrow -\infty$, the length of interval I_{jk} grows more slowly than that of J_{jk} .

From Taylor's theorem using the Lagrange form of remainder term, since $r_\psi \geq 1$, we have for all $x, z \in I_{jk}$ that

$$\begin{aligned}
|\psi(x) - \psi(z)| & \leq \sup_{\xi \in I_{jk}} |\psi^{(1)}(\xi)| \cdot |x - z| \\
& \leq C_\psi |x - z|, \tag{55}
\end{aligned}$$

from (3). From (45), using (4,7) and noting that the support of $\psi(2^j x - i)$ is J_{ji} ,

$$\begin{aligned} H_j &= 2^{-j/2} \sum_{i \in \mathbb{Z}} \left| \int h(x) \psi_{ji}(x) dx \right| = \sum_{i \in \mathbb{Z}} \left| \int_{J_{ji}} h(x) \psi(2^j x - i) dx \right| \\ &= \sum_{i \in \mathbb{Z}} |T_{1ji} + T_{2ji}|, \end{aligned} \quad (56)$$

where

$$\begin{aligned} T_{1ji} &= \int_{J_{ji}} h(x) \left\{ \sum_{k < 0} I_{jk}(x) (\psi(2^j x - i) - \psi(2^j x_{j,k+1} - i)) \right. \\ &\quad \left. + \sum_{k > 0} I_{j,k-1}(x) (\psi(2^j x - i) - \psi(2^j x_{j,k-1} - i)) dx \right\} \end{aligned} \quad (57)$$

$$\begin{aligned} T_{2ji} &= \int_{J_{ji}} h(x) \left\{ \sum_{k < 0} I_{jk}(x) \psi(2^j x_{j,k+1} - i) \right. \\ &\quad \left. + \sum_{k > 0} I_{j,k-1}(x) \psi(2^j x_{j,k-1} - i) dx \right\}. \end{aligned} \quad (58)$$

Then

$$\begin{aligned} \sum_{i \in \mathbb{Z}} |T_{1ji}| &\leq \sum_{i \in \mathbb{Z}} \int_{J_{ji}} |h(x)| \left\{ \sum_{k < 0} I_{jk}(x) |\psi(2^j x - i) - \psi(2^j x_{j,k+1} - i)| \right. \\ &\quad \left. + \sum_{k > 0} I_{j,k-1}(x) |\psi(2^j x - i) - \psi(2^j x_{j,k-1} - i)| \right\} dx \\ &\leq 2^j C_h C_\psi \sum_{i \in \mathbb{Z}} \int_{J_{ji}} (1 + |x|)^{-s_h} \left\{ \sum_{k < 0} I_{jk}(x) |x - x_{j,k+1}| \right. \\ &\quad \left. + \sum_{k > 0} I_{j,k-1}(x) |x - x_{j,k-1}| \right\} dx \\ &\leq 2^j C_h C_\psi \sum_{i \in \mathbb{Z}} \int_{J_{ji}} (1 + |x|)^{-s_h} \sum_k |x_{j,k+1} - x_{j,k}| dx \\ &\leq 2^{j(1-\gamma)} C_h C_\psi 2a \int_0^\infty (1+x)^{-s_h} dx \\ &= 2^{j(1-\gamma)} 2a C_h C_\psi (s_h - 1)^{-1}, \end{aligned} \quad (59)$$

from (3,12,54,55). Also

$$\begin{aligned}
\sum_{i \in \mathbb{Z}} |T_{2ji}| &\leq \sum_{i \in \mathbb{Z}} \left\{ \left| \int_{J_{ji}} h(x) (I_{j,-1}(x) + I_{j,0}(x)) \psi(2^j x_{j,0} - i) dx \right| \right. \\
&\quad + \int_{J_{ji}} |h(x)| \sum_{k \leq -2} I_{jk}(x) |\psi(2^j x_{j,k+1} - i)| dx \\
&\quad \left. + \int_{J_{ji}} |h(x)| \sum_{k \geq 2} I_{j,k-1}(x) |\psi(2^j x_{j,k-1} - i)| dx \right\} \\
&\leq C_\psi \left\{ \sum_{i=1-a}^{-1} \left| \int_{x_{j,-1}}^{x_{j,1}} h(x) dx \right| + \sum_{i \in \mathbb{Z}} \int_{J_{ji}} |h(x)| \sum_{k \leq -2} I_{jk}(x) dx \right. \\
&\quad \left. + \sum_{i \in \mathbb{Z}} \int_{J_{ji}} |h(x)| \sum_{k \geq 2} I_{j,k-1}(x) dx \right\}
\end{aligned}$$

using (3) and that $\psi(2^j x_{j,0} - i) \neq 0$ only for $1-a \leq i \leq -1$, for which $I_{j,-1} \cup I_{j,0} \subseteq J_{ji}$, for $j < 0$,

$$\begin{aligned}
&\leq aC_\psi \left\{ \left| \int_{x_{j,-1}}^{x_{j,1}} h(x) dx \right| + \int_{-\infty}^{x_{j,-1}} |h(x)| dx + \int_{x_{j,1}}^{\infty} |h(x)| dx \right\} \\
&= 2aC_\psi \left\{ \int_{-\infty}^{x_{j,-1}} |h(x)| dx + \int_{x_{j,1}}^{\infty} |h(x)| dx \right\}
\end{aligned}$$

using $\int h(x) dx = 0$, from (11),

$$\begin{aligned}
&< 2aC_\psi C_h \left\{ \int_{-\infty}^{x_{j,-1}} (1 + |x|)^{-s_h} dx + \int_{x_{j,1}}^{\infty} (1 + |x|)^{-s_h} dx \right\} \\
&= 4aC_\psi C_h (s_h - 1)^{-1} (1 + 2^{-j\gamma})^{1-s_h} \\
&< 2^{j\gamma(s_h-1)} 4aC_\psi C_h (s_h - 1)^{-1}, \tag{60}
\end{aligned}$$

using (12,54). Thus, from (56),

$$\begin{aligned}
H_j &\leq \sum_{i \in \mathbb{Z}} |T_{1ji}| + |T_{2ji}| \\
&< 2^{j(1-\gamma)} 2aC_h C_\psi (s_h - 1)^{-1} + 2^{j\gamma(s_h-1)} 4aC_\psi C_h (s_h - 1)^{-1}, \tag{61}
\end{aligned}$$

from (59,60). As $j \rightarrow -\infty$, the rate of decay in (61) is greatest when $1 - \gamma = \gamma(s_h - 1)$, which is when $\gamma = s_h^{-1}$. Substituting this into (61) gives the required result. $\square$

Lemma 3. For $j_0 < 0$ and $j_1 \geq 0$,

$$\sum_{j \geq j_1} H_j < C_1 \frac{2^{-j_1 r_\psi}}{1 - 2^{-r_\psi}} \quad \text{and} \quad \sum_{j \leq j_0} H_j < C_2 \frac{2^{j_0(1-s_h^{-1})}}{1 - 2^{-(1-s_h^{-1})}}. \quad (62)$$

Proof. The results follow by summing the geometric series of the bounds on H_j obtained in Lemmas 1 and 2. $\square$

Corollary 4. The homogeneous Besov space norm $\dot{B}_{1,1}^0$ for h , given in (8), is finite.

Proof. From (8) and (45), $\dot{B}_{1,1}^0(\mathbb{R}) = \sum_{j \in \mathbb{Z}} H_j = \sum_{j \geq 0} H_j + \sum_{j < 0} H_j$, which is finite as can be seen from (62), setting $j_1 = 0$ and $j_0 = -1$. $\square$

Corollary 5. The approximation $\hat{h}_{j_0, j_1}(x)$ given in (10) converges geometrically in L_1 norm to h at rate 2^{-r_ψ} as $j_1 \rightarrow \infty$ and at rate $2^{-(1-s_h^{-1})}$ as $j_0 \rightarrow -\infty$.

Proof. The L_1 discrepancy between h and $\hat{h}_{j_0, j_1}(x)$ is, from (6,10),

$$\begin{aligned} \int |h(x) - \hat{h}_{j_0, j_1}(x)| dx &= \int \left| \sum_{j > j_1} \sum_{i \in \mathbb{Z}} \tilde{h}_{ji} \psi_{ji}(x) + \sum_{j < j_0} \sum_{i \in \mathbb{Z}} \tilde{h}_{ji} \psi_{ji}(x) \right| dx \\ &\leq \int_0^a |\psi(y)| dy \left\{ \sum_{j > j_1} \sum_{i \in \mathbb{Z}} 2^{-j/2} |\tilde{h}_{ji}| + \sum_{j < j_0} \sum_{i \in \mathbb{Z}} 2^{-j/2} |\tilde{h}_{ji}| \right\} \\ &\leq a C_\psi \left\{ \sum_{j \geq j_1} H_j + \sum_{j \leq j_0} H_j \right\} \\ &< a C_\psi C_1 \frac{2^{-j_1 r_\psi}}{1 - 2^{-r_\psi}} + a C_\psi C_2 \frac{2^{j_0(1-s_h^{-1})}}{1 - 2^{-(1-s_h^{-1})}}, \end{aligned} \quad (63)$$

using (3,4,45,62). Geometric convergence to zero of (63) occurs in its first term at rate 2^{-r_ψ} as $j_1 \rightarrow \infty$, and in its second term at rate $2^{-(1-s_h^{-1})}$ as $j_0 \rightarrow -\infty$. $\square$

A.2 Completing the proof of principle of pWMC

The proof of principle of pWMC, given in Section 3.2, requires the following result.

Lemma 6. For given $y \in \mathbb{R}$, $\int_x |q(x, y) - q(y, x)| dx < \infty$, a.e.

Proof. For given y , integrating (23) over $x \neq y \in \mathbb{R}$, using Tonelli's Theorem to exchange the order of integration and summations, we obtain

$$\int_{x \neq y} q(x, y) dx \leq \frac{1}{c_g} \sum_{j \in \mathbb{Z}} \frac{1}{A_j} \sum_{i \in \mathbb{Z}} \left\{ \tilde{h}_{ji}^+ \psi_{ji}^+(y) \int_x \psi_{ji}^-(x) dx + \tilde{h}_{ji}^- \psi_{ji}^-(y) \int_x \psi_{ji}^+(x) dx \right\},$$

$$= \frac{1}{c_g} \sum_{j \in \mathbb{Z}} \sum_{i \in \mathbb{Z}} \left\{ \tilde{h}_{ji}^+ \psi_{ji}^+(y) + \tilde{h}_{ji}^- \psi_{ji}^-(y) \right\}. \quad (64)$$

Therefore

$$\int_y \int_{x \neq y} q(x, y) \, dx \, dy \leq \frac{1}{c_g} \sum_{j \in \mathbb{Z}} A_j \sum_{i \in \mathbb{Z}} \left\{ \tilde{h}_{ji}^+ + \tilde{h}_{ji}^- \right\} = \frac{2A_0}{c_g} \sum_{j \in \mathbb{Z}} 2^{-j/2} \sum_{i \in \mathbb{Z}} |\tilde{h}_{ji}|,$$

using (15). Then, by the finiteness of the Besov space norm for h given in (8) (proved in Appendix A.1), $\int_y \int_{x \neq y} q(x, y) \, dx \, dy < \infty$. This implies that both $\int_{x \neq y} q(x, y) \, dx < \infty$ and $\int_{x \neq y} q(y, x) \, dx < \infty$, a.e. Therefore, noting that $q(x, y) \geq 0$ for all $x, y \in \mathbb{R}$, $\int_x |q(x, y) - q(y, x)| \, dx \leq \int_{x \neq y} q(x, y) + q(y, x) \, dx < \infty$, a.e. $\square$

A.3 Supplementary Tables and Figures

The following tables and figures are referenced in Sections 8.1 and 8.2 of the paper.

Table 5 Additional statistics from the application of Monte-Carlo samplers to the one-dimensional example of Section 8.1 of the paper. See text for further details.

<i>Iterations</i>	<i>Scale</i>	<i>sWMC</i>		<i>RS</i>	<i>IS</i>	<i>RWM</i>
		<i>mean jumps</i>	$\hat{\rho}_L(40)$	<i>M</i>	<i>sd(w)</i>	<i>burn-in</i>
1– 20 000	5	8.6	1.06	13240	19.6	0
20 001– 40 000	10	8.7	1.06	698	15.2	20 000
40 001– 60 000	20	12.0	1.02	96	5.3	40 000
60 001– 80 000	40	16.6	1.01	51	4.2	60 000
80 001–100 000	80	7.3	0.93	64	4.9	80 000

Table 6 Additional statistics from the application of Monte-Carlo samplers to the two-dimensional example of Section 8.2 of the paper.

<i>Scale</i>	<i>sWMC</i>		<i>RS</i>	<i>IS</i>	<i>RWM</i>
	<i>mean jumps</i>		<i>M</i>	<i>sd(w)</i>	<i>burn-in</i>
1	4.1		1.8×10^9	9.6	0
2	15.5		21138	7.2	0
4	17.1		3490	18.5	0
6	8.6		4352	1.6	0

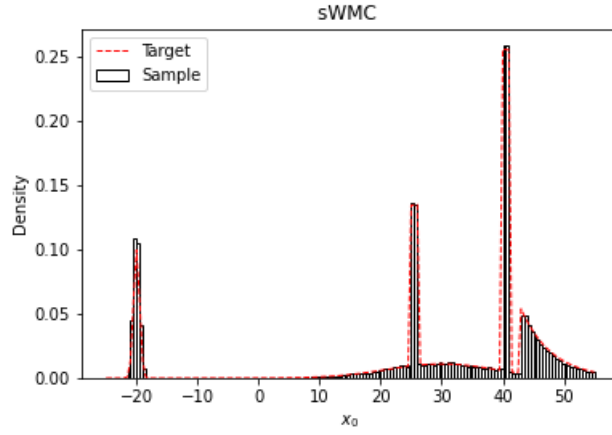


Fig. 4 sWMC applied to the one-dimensional example of Section 8.1 of the paper. The target density g_1 , shown with a broken red line, is described in Figure 2 of the paper. The histogram is of $L = 100\,000$ independent particles generated by sWMC in a single simulation progressing through a sequence of five initial t_5 -distributions g_0 located at -2.0 with increasing scale from 5.0 to 80.0, as described in Table 1 of the paper.

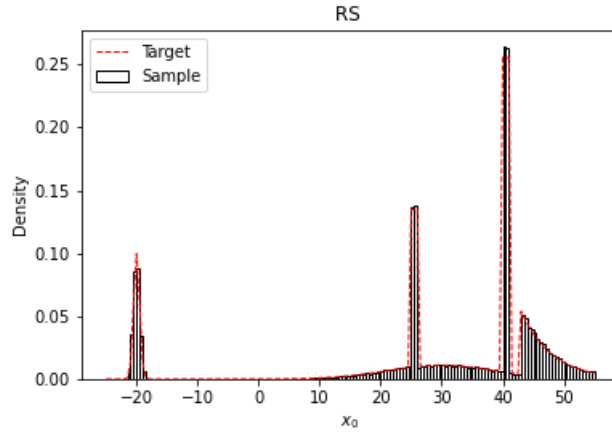


Fig. 5 RS applied to the one-dimensional example of Section 8.1 of the paper. The target density g_1 , shown with a broken red line, is described in Figure 2 of the paper. The histogram is of $L = 100\,000$ independent particles generated by RS in a single simulation progressing through a sequence of five initial t_5 -distributions g_0 located at -2.0 with increasing scale from 5.0 to 80.0, as described in Table 1 of the paper.

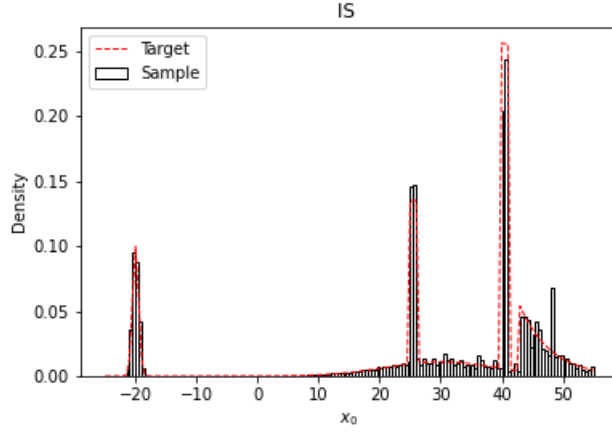


Fig. 6 IS applied to the one-dimensional example of Section 8.1 of the paper. The target density g_1 , shown with a broken red line, is described in Figure 2 of the paper. The histogram is of $L = 100\,000$ independent particles generated by IS in a single simulation progressing through a sequence of five initial t_5 -distributions g_0 located at -2.0 with increasing scale from 5.0 to 80.0, as described in Table 1 of the paper.

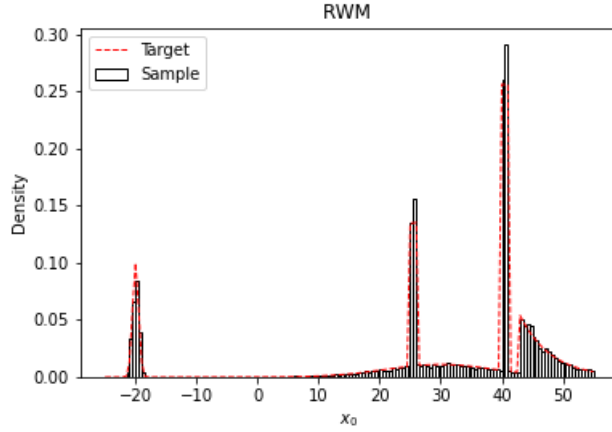


Fig. 7 RWM applied to the one-dimensional example of Section 8.1 of the paper. The target density g_1 , shown with a broken red line, is described in Figure 2 of the paper. The histogram is of $L = 100\,000$ dependent particles generated by RWM in a single simulation progressing through a sequence of five proposal t_5 -distributions q^{prop} centred on the current point with increasing scale from 5.0 to 80.0, as described in Table 1 of the paper.

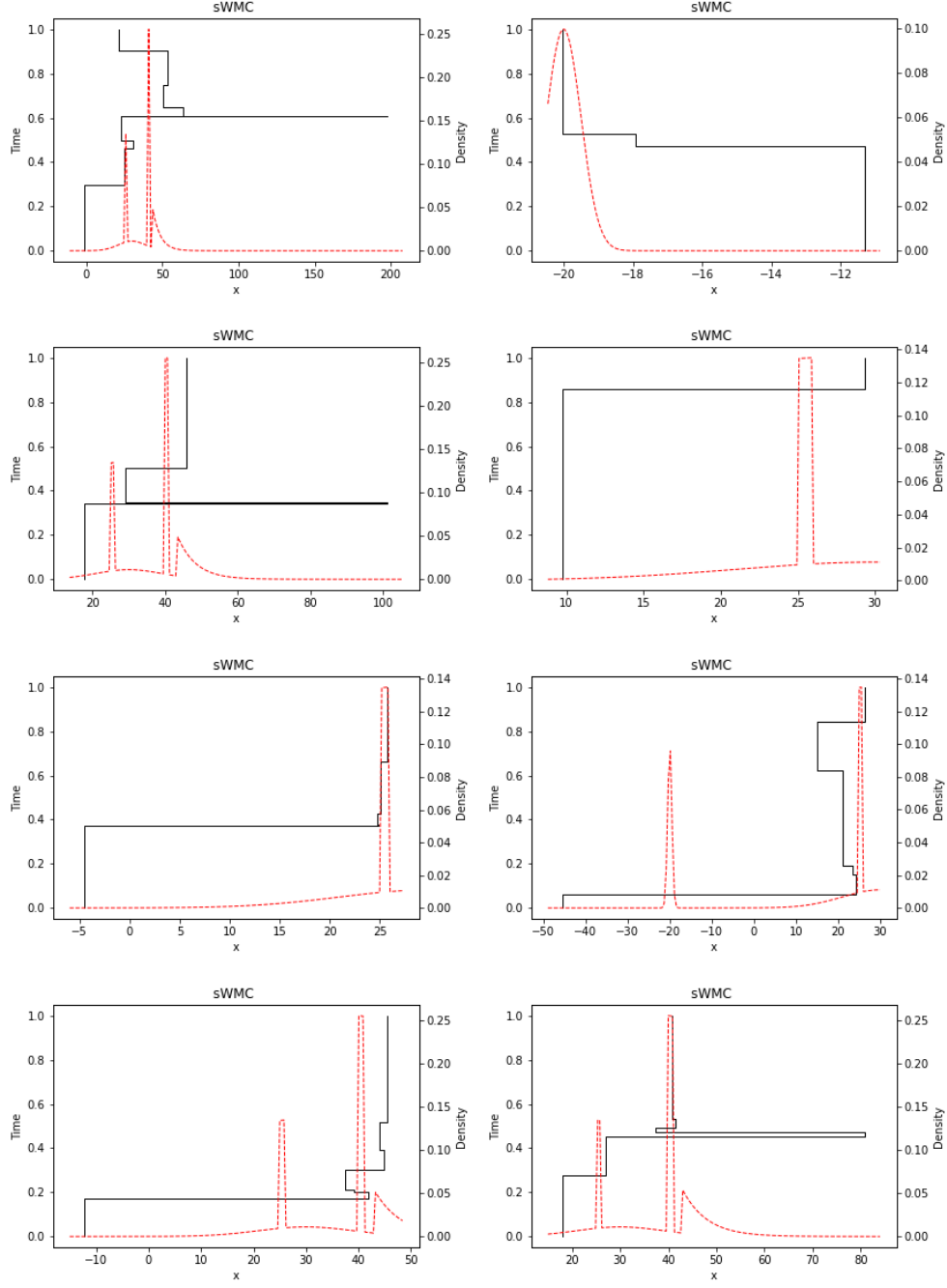


Fig. 8 Position-time tracks of 8 particles from an application of sWMC to the one-dimensional example discussed in Section 8.1 of the paper. The target density g_1 is described in Figure 2 of the paper, and the initial density g_0 is a t_5 -distribution rescaled by a factor of 20.0 and relocated to -2.0. In each panel, the track of one particle is shown with a solid black line, and the corresponding section of the target density is shown with a broken red line.

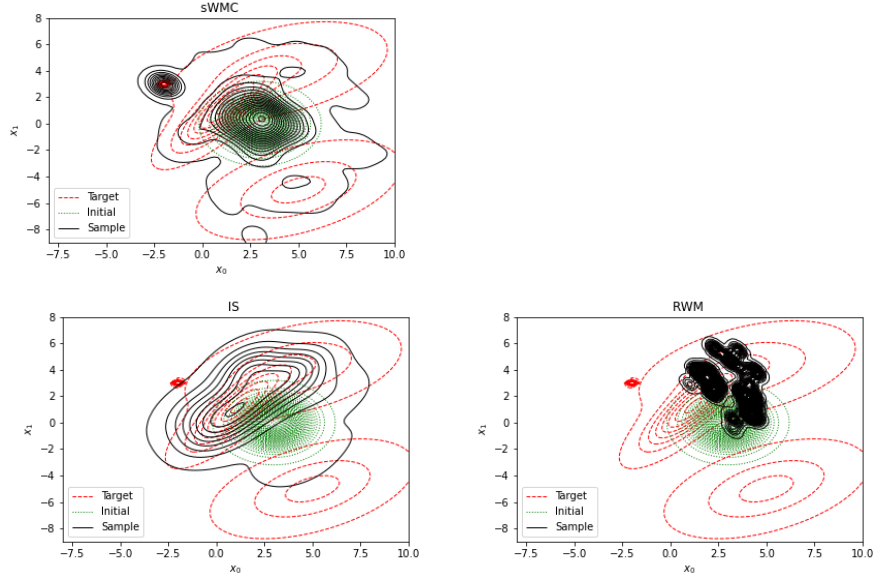


Fig. 9 Monte-Carlo samplers applied to the two-dimensional example of Section 8.2 of the paper (top-left: *sWMC*; bottom-left: *IS*; bottom-right: *RWM*). In each panel, the target density g_1 , described in Figure 3 of the paper, is shown in broken red contours; the initial distribution g_0 is a bivariate Normal distribution $\mathcal{N}\left(\begin{bmatrix} 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)$ (scale 1 in Table 3 of the paper) and is shown in dotted green contours; and a kernel-density estimate based on $L = 10000$ particles generated by one of the samplers, is shown in continuous black contours.

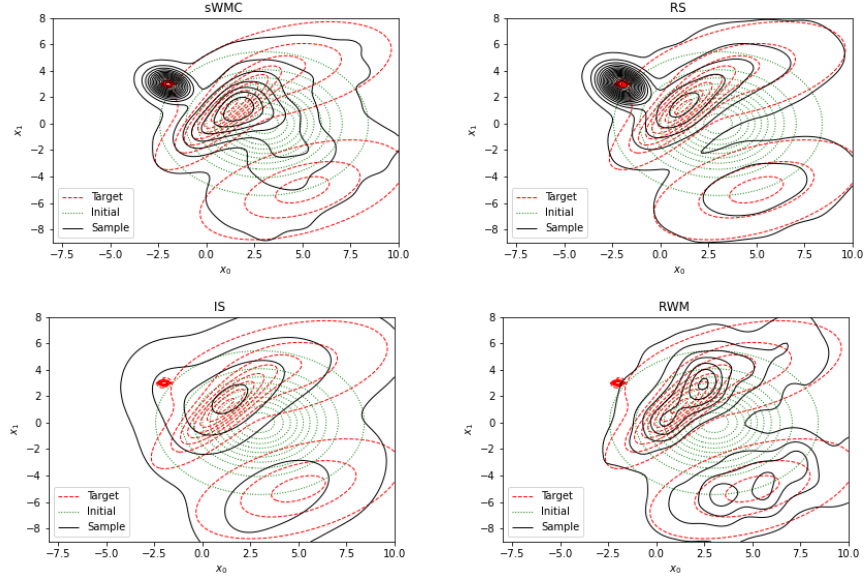


Fig. 10 Monte-Carlo samplers applied to the two-dimensional example of Section 8.2 of the paper (top-left: sWMC; top-right: RS; bottom-left: IS; bottom-right: RWM). In each panel, the target density g_1 , described in Figure 3 of the paper, is shown in broken red contours; the initial distribution g_0 is a bivariate Normal distribution $N\left(\begin{bmatrix} 3 \\ 0 \end{bmatrix}, 2^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)$ (scale 2 in Table 3 of the paper) and is shown in dotted green contours; and a kernel-density estimate based on $L = 10\,000$ particles generated by one of the samplers, is shown in continuous black contours.

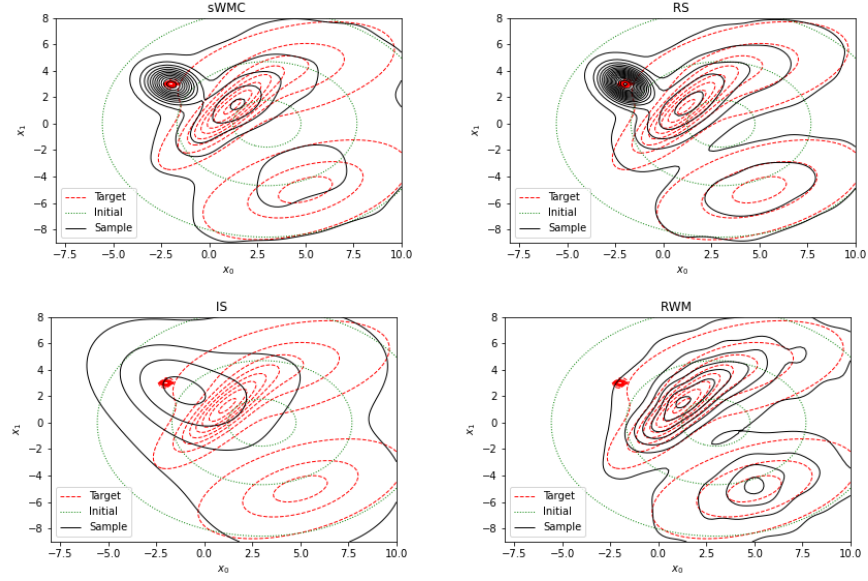


Fig. 11 Monte-Carlo samplers applied to the two-dimensional example of Section 8.2 of the paper (top-left: sWMC; top-right: RS; bottom-left: IS; bottom-right: RWM). In each panel, the target density g_1 , described in Figure 3 of the paper, is shown in broken red contours; the initial distribution g_0 is a bivariate Normal distribution $N\left(\begin{bmatrix} 3 \\ 0 \end{bmatrix}, 4^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)$ (scale 4 in Table 3 of the paper) and is shown in dotted green contours; and a kernel-density estimate based on $L = 10\,000$ particles generated by one of the samplers, is shown in continuous black contours.

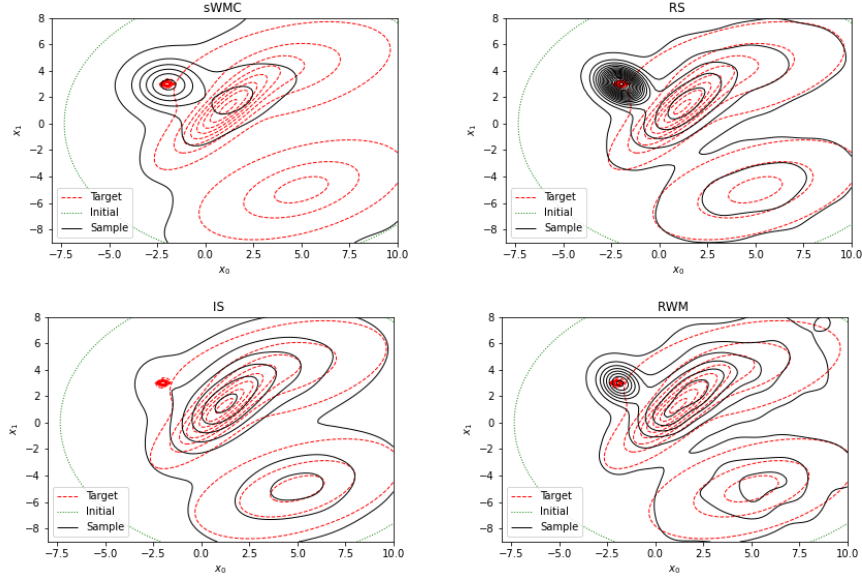


Fig. 12 Monte-Carlo samplers applied to the two-dimensional example of Section 8.2 of the paper (top-left: sWMC; top-right: RS; bottom-left: IS; bottom-right: RWM). In each panel, the target density g_1 , described in Figure 3 of the paper, is shown in broken red contours; the initial distribution g_0 is a bivariate Normal distribution $N\left(\begin{bmatrix} 3 \\ 0 \end{bmatrix}, 6^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)$ (scale 6 in Table 3 of the paper) and is shown in dotted green contours; and a kernel-density estimate based on $L = 10\,000$ particles generated by one of the samplers, is shown in continuous black contours.