Optimal Scaling of Random Walk Metropolis algorithms with Discontinuous target densities

Neal, P and Roberts, G and Yuen, J

2007

MIMS EPrint: **2007.91**

ISSN 1749-9097

The Annals of Applied Probability

OPTIMAL SCALING OF RANDOM WALK METROPOLIS ALGORITHMS WITH DISCONTINUOUS TARGET DENSITIES

By Peter Neal, Gareth Roberts and Wai Kong Yuen

University of Manchester, University of Warwick and Brock University

We consider the optimal scaling problem for high-dimensional Random walk Metropolis (RWM) algorithms where the target distribution has a discontinuous probability density function. All previous analysis has focused upon continuous target densities. The main result is a weak convergence result as the dimensionality d of the target densities converges to ∞. In particular, when the proposal variance is scaled by d^{-2}, the sequence of stochastic processes formed by the first component of each Markov chain converges to an appropriate Langevin diffusion process. Therefore optimising the efficiency of the RWM algorithm is equivalent to maximising the speed of the limiting diffusion. This leads to an asymptotic optimal acceptance rate of $e^{-2}(= 0.1353)$ under quite general conditions. The results have major practical implications for the implementation of RWM algorithms by highlighting the detrimental effect of choosing RWM algorithms over Metropolis-within-Gibbs algorithms.

1. Introduction. Random walk Metropolis (RWM) algorithms are widely used generic Markov chain Monte Carlo (MCMC) algorithms. The ease with which RWM algorithms can be constructed has no doubt played a pivotal role in their popularity. The efficiency of a RWM algorithm depends fundamentally upon the scaling of the proposal density. Choose the variance of the proposal to be too small and the RWM will converge slowly since all its increments are small. Conversely, choose the variance of the proposal to be too large and too high a proportion of proposed moves will be rejected. Of particular interest is how the scaling of the proposal variance depends upon the dimensionality of the target distribution. The target distribution is the distribution of interest and the MCMC algorithm is constructed such that the stationary

AMS 2000 subject classifications: Primary 60F05; secondary 65C05

Keywords and phrases: Random walk Metropolis, Markov chain Monte Carlo, optimal scaling.

distribution of the Markov chain is the target distribution.

In their seminal work, [12] considered the asymptotic problem as the dimensionality d of the target distribution tends to infinity. In particular, they showed that for continuous and suitably smooth IID product densities the (asymptotic) optimal scaling of the proposal density is l^2/d for some $l > 0$. Of great practical importance, it was shown that the optimal scaling of the proposal density leads to an average acceptance probability of 0.234. This has major practical implications for practitioners, in that, to monitor the efficiency of the RWM algorithm it is sufficient to study the proportion of proposed moves accepted. Computer simulations have shown that the asymptotic approximations are very good for $d = 10$. Furthermore the overall efficiency of the algorithm is $O(d)$. A number of extensions of the results in [12] have been considered, see [8], [14], [11] and [4], all with the same conclusions for optimal scaling and the resulting average acceptance probability. Similar results have been obtained for Metropolis adjusted Langevin algorithms (MALA), see [13], where the optimal scaling of the proposal density is $l^2/d^{1/3}$ for some $l > 0$ and the corresponding average acceptance probability is 0.574. Extensions appear in [14], [7] and [11].

In all of the above mentioned examples the target distribution is assumed to have a continuous (and suitably differentiable) probability density function (pdf). The aim of the current work is to investigate the situation where the target distribution has a discontinuous pdf. More specifically, we consider target distributions of the form

$$\pi_d(\mathbf{x}^d) = \prod_{i=1}^{d} f(x_i^d), \tag{1.1}$$

where

$$f(x) \propto \exp(g(x)) \quad (0 < x < 1) \tag{1.2}$$

where $g(\cdot)$ is twice differentiable upon $[0, 1]$ and $f(x) = 0$ otherwise. In particular, we show that the optimal scaling of the proposal variance is $\sigma_d^2 = l^2/d^2$ for some $l > 0$ with a corresponding average acceptance probability of $e^{-2}(= 0.1353)$. Therefore the mixing of the algorithm is $O(d^2)$. This compares unfavourably with Metropolis-within-Gibbs algorithms for which the mixing of the algorithm is $O(d)$. Furthermore, the arguments of [11] can easily be adapted to this situation and it is shown that updating only a proportion $c > 0$ of the components rather than the full

RWM algorithm leads to a gain in efficiency of c^{-1}. Hence lower dimensional updating schema are preferable. This is the main practical implications of the paper, although should a full RWM algorithm be used, the optimal scaling of the proposal variance leads to an average acceptance probability of e^{-2}.

The paper is structured as follows. The algorithm and main Theorem (Theorem 2.1) are given in Section 2. Although the results are similar to [12] and subsequent work the method of proof is markedly different. In particular, the proof of Theorem 2.1 splits into two parts which are proved in Sections 3 and 4, respectively. Finally in Section 5 extensions of Theorem 2.1 are given and the practical conclusions of the work beyond the analytical results are discussed.

2. Algorithm and statement of the main Theorem. For $d \geq 1$, we shall consider a Random walk Metropolis (RWM) algorithm on the d-dimensional hypercube with target density $\pi_d(\mathbf{x}^d)$ of the form given by (1.1) and (1.2).

For $d \geq 1$, let $\mathbf{X}_0^d, \mathbf{X}_1^d, \ldots$ denote the successive states of a Markov chain constructed as follows. Let $\mathbf{X}_0^d \sim \pi_d(\cdot)$. For $t \geq 0$ and $i \geq 1$, let $Z_{t,i}$ be independent and identically distributed according to the uniform random variable $Z \sim U[-1,1]$, with $\mathbf{Z}_t^d = (Z_{t,1}, Z_{t,2}, \ldots, Z_{t,d})$. Then for $d \geq 1$, $t \geq 0$ and $l > 0$, let

$$\mathbf{Y}_{t+1}^d = \mathbf{X}_t^d + \sigma_d \mathbf{Z}_t^d$$

where $\sigma_d = l/d$. Set

$$\mathbf{X}_{t+1}^d = \begin{cases} \mathbf{Y}_{t+1}^d & \text{with probability } \left\{1 \wedge \frac{\pi_d(\mathbf{Y}_{t+1}^d)}{\pi_d(\mathbf{X}_t^d)}\right\}, \\ \mathbf{X}_t^d & \text{otherwise.} \end{cases}$$

Therefore for $d \geq 1$, let

$$h_d(\mathbf{z}^d) = \begin{cases} 2^{-d} & \text{if } \mathbf{z}^d \in (-1,1)^d \\ 0 & \text{otherwise.} \end{cases}$$

Let $J_d(\mathbf{x}^d)$ denote the probability of accepting a move in the RWM process given the current state of the process is $\mathbf{x}^d$. Then

$$J_d(\mathbf{x}^d) = \int h_d(\mathbf{z}^d) \left\{1 \wedge \frac{\pi_d(\mathbf{x}^d + \sigma_d \mathbf{z}^d)}{\pi_d(\mathbf{x}^d)}\right\} \, d\mathbf{z}^d. \tag{2.1}$$

In order to assist in our analysis of the RWM algorithm it is necessary to define a related process which we shall term the pseudo-RWM process. The key difference is that the pseudo-

RWM process moves at each iteration. Furthermore the moves in the pseudo-RWM process are identical to those of the RWM process, conditioned upon a move in the RWM process being accepted. For $d \geq 1$, let $\hat{\mathbf{X}}_0^d, \hat{\mathbf{X}}_1^d, \ldots$ denote the successive states of the pseudo-RWM process, where $\hat{\mathbf{X}}_0^d \sim \pi_d(\cdot)$. The pseudo-RWM process is a Markov process, where for $t \geq 0$, $\hat{\mathbf{X}}_{t+1}^d = \hat{\mathbf{X}}_t^d + \sigma_d \hat{\mathbf{Z}}_t^d$ and given that $\hat{\mathbf{X}}_t^d = \mathbf{x}^d$, $\hat{\mathbf{Z}}_t^d$ has pdf

$$\zeta(\mathbf{z}^d|\mathbf{x}^d) = h_d(\mathbf{z}^d) \left\{ 1 \wedge \frac{\pi_d(\mathbf{x}^d + \sigma_d \mathbf{z}^d)}{\pi_d(\mathbf{x}^d)} \right\} \Big/ J_d(\mathbf{x}^d) \qquad (\mathbf{z}^d \in \mathbb{R}^d).$$

Let $\hat{\pi}_d(\cdot)$ denote the stationary distribution of the pseudo-RWM process. Then for $\mathbf{x}^d \in [0,1]^d$,

$$\begin{aligned} \hat{\pi}_d(\mathbf{x}^d) &= J_d(\mathbf{x}^d)\pi_d(\mathbf{x}^d) \Big/ \int J_d(\mathbf{y}^d)\pi_d(\mathbf{y}^d)\, d\mathbf{y}^d \\ &= J_d(\mathbf{x}^d) \prod_{i=1}^d f(x_i^d) \Big/ \int J_d(\mathbf{y}^d)\pi_d(\mathbf{y}^d)\, d\mathbf{y}^d \\ &= \frac{J_d(\mathbf{x}^d)}{M_d} \prod_{i=1}^d f(x_i^d), \quad \text{say,} \end{aligned} \tag{2.2}$$

where $M_d = \int J_d(\mathbf{y}^d)\pi_d(\mathbf{y}^d)\, d\mathbf{y}^d$. Note that $\mathbf{X}_0^d, \hat{\mathbf{X}}_0^d \sim \pi_d(\cdot)$ and therefore we can couple the two processes to have the same starting value $\mathbf{X}_0^d$. A continued coupling of the two processes is outlined below. For $j \geq 0$ let $G_j(\cdot)$ denote independent Geometric random variables, where for $0 < p \leq 1$, $G_j(p)$ denotes a Geometric random variable with 'success' probability p. For $s \in \mathbb{Z}^+$, let

$$U_t^d = \sup \left\{ s \in \mathbb{N} : \sum_{j=0}^{s-1} G_j(J_d(\hat{\mathbf{X}}_j^d)) \leq t \right\}.$$

The RWM process can then be constructed from the pseudo-RWM process by setting $\mathbf{X}_0^d \equiv \hat{\mathbf{X}}_0^d$, and for all $s \geq 0$, $\mathbf{X}_s^d = \hat{\mathbf{X}}_{U_s^d}^d$. Obviously the above process can be reversed by setting $\hat{\mathbf{X}}_t^d$ equal to the t^{th} accepted move in the RWM process.

For $t \geq 0$, let

$$T_d(t) = \frac{1}{d^2} \sum_{j=0}^{[td^2-1]} G_j(J_d(\hat{\mathbf{X}}_j^d)),$$

and let

$$W_t^d = \sup\{s; T_d(s) \leq t\}.$$

Therefore $W_t^d = \frac{1}{d^2} U_{[d^2 t]}^d$.

For each $d \geq 1$, the components of $\mathbf{X}_0^d$ are independent and identically distributed. Therefore we focus attention on the first component as this is indicative of the behaviour of the whole process. Therefore for $d \geq 1$ and $t \geq 0$, letting $V_t^d = X^d_{[d^2t],1}$ and $\hat{V}_t^d = \hat{X}^d_{[d^2t],1}$, we have that $V_t^d = \hat{V}^d_{W_t^d}$.

THEOREM 2.1. *Fix $l > 0$. For all $d \geq 1$, let $\mathbf{X}_0^d = (X^d_{0,1}, X^d_{0,2}, \ldots, X^d_{0,d})$ be such that all of its components are distributed according to $f(\cdot)$. Then, as $d \to \infty$,*

$$V^d \Rightarrow V$$

where $V_0 \sim f(\cdot)$ and V satisfies the (reflected) Langevin SDE on $[0,1]$

$$dV_t = \sqrt{\nu(l)} dB_t + \frac{1}{2}\nu(l) g'(V_t)\, dt + dL_t^0(V) - dL_t^1(V) \tag{2.3}$$

where

$$\nu(l) = \frac{l^2}{3} \exp\left(-\frac{f^* l}{2}\right)$$

and $f^* = \lim_{x \downarrow 0} \left(\frac{f(x)+f(1-x)}{2}\right)$.

Note that the factor $1/3$ comes from the variance of $U[-1,1]$ being $1/3$.

Proof. Fix $l > 0$ and let $\psi_l(t) = \exp(-f^* l/2)t$ $(t \geq 0)$. We shall prove the theorem under the assumptions that

$$\hat{V}^d_\cdot \quad \Rightarrow \quad \hat{V}_\cdot \quad \text{as } d \to \infty, \tag{2.4}$$

$$W^d_\cdot \quad \xrightarrow{p} \quad \psi_l(\cdot) \quad \text{as } d \to \infty, \tag{2.5}$$

where $\hat{V}_\cdot$ satisfies (2.3) with $\nu(l) = l^2/3$ and $\hat{V}_0 \equiv V_0 \sim f(\cdot)$. The proofs of (2.4) and (2.5) are given in Sections 4 and 3, respectively.

Thus to prove the theorem it is sufficient to show that $V^d_\cdot \Rightarrow \hat{V}_{\psi_l(\cdot)}$ as $d \to \infty$.

By (2.4), (2.5) and [5], Theorem 4.4,

$$(\hat{V}^d_\cdot, W^d_\cdot) \xrightarrow{D} (\hat{V}_\cdot, \psi_l(\cdot)).$$

Since $\hat{V}_\cdot$ is almost surely continuous and $W^d_\cdot$ is non-decreasing in t, it follows by [5], (17.7–17.9) that

$$\hat{V}^d_{W^d_\cdot} \Rightarrow \hat{V}_{\psi_l(\cdot)} \quad \text{as } d \to \infty.$$

Therefore since for all $t \geq 0$ and $d \geq 1$, $V_t^d \equiv \hat{V}^d_{W_t^d}$, we finally have that

$$V_{\cdot}^d \Rightarrow \hat{V}_{\psi_l(\cdot)} \equiv V_{\cdot} \quad \text{as } d \to \infty,$$

as required. □

The speed of the diffusion depends on f^* and therefore depends upon the limiting number of components (under stationarity) close to the boundaries 0 and 1. In fact the speed of the diffusion is not affected at all by the behaviour of $f(\cdot)$ in the interior of $(0,1)$. This contrasts with the continuous density results of [12] where the speed of the limiting diffusion depends upon $\mathbb{E}_f[g'(X)^2]$.

The most important consequence of Theorem 2.1 is the following Corollary. First let

$$a_d(l) = \int \pi_d(\mathbf{x}^d) J_d(\mathbf{x}^d)\, d\mathbf{x}^d \tag{2.6}$$

be the average acceptance rate of the RWM algorithm in d dimensions, and let

$$a(l) = \exp(-f^* l/2).$$

Corollary 2.2.

$$\lim_{d \to \infty} a_d(l) = a(l)$$

$\nu(l)$ is maximised by

$$l = \hat{l} = \frac{4}{f^*}.$$

Also

$$a(\hat{l}) = \exp(-2) = 0.1353.$$

Clearly, if $f(\cdot)$ is known $\hat{l}$ can be calculated explicitly. However where MCMC is used $f(\cdot)$ will often only be known up to the constant of proportionality. This is where Corollary 2.2 has major practical implications, in that, to maximise the speed of the limiting diffusion, and hence, the efficiency of the RWM algorithm it is sufficient to monitor the average acceptance rate, and to choose l such that the average acceptance rate is approximately e^{-2}. Therefore there is no need to explicitly calculate or estimate the constant of proportionality.

3. Proof of (2.5). The proof of (2.5) is rather lengthy and is completed in Theorem 3.16. Therefore before jumping into the proof we give an outline of how the argument will proceed. In order to show that $W^d_{\cdot} \stackrel{p}{\longrightarrow} \psi_l(\cdot)$ as $d \to \infty$, we will show that for all $s \geq 0$,

$$T_d(s) \stackrel{p}{\longrightarrow} s \exp(f^* l/2) \quad \text{as } d \to \infty. \tag{3.1}$$

The first step towards (3.1) is to show that under stationarity, the number of components in the rejection region is sufficiently well controlled, where for fixed $l > 0$, we term $R_d^l = (0, \sigma_d) \cup (1 - \sigma_d, 1)$ the rejection region. The term rejection region is used since for a component in R_d^l there is a non-zero probability of proposing a new value outside of the range $(0, 1)$ with such moves rejected automatically. This enables us to consider a simpler quantity $\hat{T}_d(s) = \frac{1}{d^2} \sum_{i=0}^{[sd^2-1]} L_d(\hat{\mathbf{X}}_i^d)$ where $L_d(\mathbf{x}^d) = J_d(\mathbf{x}^d)^{-1}$ $(\mathbf{x}^d \in (0,1)^d)$, in place of $T_d(s)$, see Corollary 3.3.

The second step is to deal with complications caused by the dependence between the components in the pseudo-RWM process. This dependence is very weak and for large d, the pseudo-RWM process behaves in a similar manner to the independent pseudo-RWM (ip-RWM) process defined after Corollary 3.3. In particular, we show that the two processes can be coupled for $[d^\alpha]$ iterations $(0 < \alpha < 1/8)$ with probability tending to 1 as $d \to \infty$. This coupling is sufficient for our needs and is completed in Corollary 3.6.

The above are the preliminary steps towards proving (3.1). The main part of the proof consists in showing that for suitable sequences of positive integers $\{i_d\}$ and $\{k_d\}$,

$$\mathbb{E}[L_d(\hat{\mathbf{X}}^d_{i_d+k_d}) | \hat{\mathbf{X}}^d_{i_d}] \stackrel{p}{\longrightarrow} \exp(f^* l/2) \quad \text{as } d \to \infty. \tag{3.2}$$

In other words, after k_d iterations the 'acceptance probability' of the pseudo-RWM process 'forgets' the initial configuration of the Markov chain. This is done by showing that the position of the components in R_d^l after k_d iterations converge to the points of a suitably defined Poisson point process. By (3.2) and Chebychev's inequality, we will be able to show that

$$\hat{T}_d(s) \stackrel{p}{\longrightarrow} s \exp(f^* l/2) \quad \text{as } d \to \infty.$$

The proof of (3.2) is given in Lemma 3.13 following preliminary analysis of the movements of the components of the pseudo-RWM and ip-RWM processes. The results are then drawn together in Theorem 3.16.

Fix $l > 0$. We shall assume that α, β and γ are arbitrary but satisfy $0 < 20\gamma < \beta < \alpha < \frac{1}{8}$. For $d \geq 1$ and $r \geq 0$, let $R_d^r = (0, r/d) \cup (1 - r/d, 1)$ (this is consistent with the definition of R_d^l as the rejection region) and let $b_d^r(\mathbf{x}^d) = \sum_{j=1}^d 1_{\{x_j^d \in R_d^r\}}$.

Let $F_d^1 = \{\mathbf{x}^d; b_d^l(\mathbf{x}^d) \leq \gamma \ \log d\}$ and let $A_d^1 = \{\hat{\mathbf{X}}_s^d \in F_d^1; 0 \leq s \leq [d^{5/2}]\}$. In order to analyse A_d^1 it is easier to study the RWM process and use the coupling outlined in Section 2 in reverse.

Lemma 3.1. *For any* $\kappa > 0$,

$$d^\kappa \mathbb{P}(\mathbf{X}_0^d \notin F_d^1) \to 0 \qquad \textit{as } d \to \infty.$$

Proof. Note that $\mathbf{X}_0^d \notin F_d^1$ if and only if $b_d^l(\mathbf{X}_0^d) > \gamma \ \log d$. However,

$$b_d^l(\mathbf{X}_0^d) \sim Bin\left(d, \int_0^{l/d} \{f(x) + f(1-x)\}\, dx\right),$$

with

$$d \int_0^{l/d} \{f(x) + f(1-x)\}\, dx \to 2f^* l \qquad \text{as } d \to \infty. \tag{3.3}$$

Fix $\rho > \kappa/\gamma$. By Markov's inequality,

$$\begin{aligned} d^\kappa \mathbb{P}(b_d^l(\mathbf{X}_0^d) > \gamma \ \log d) \ &\leq \ d^\kappa \left(1 + (e^\rho - 1)\int_0^{l/d} \{f(x) + f(1-x)\}\, dx\right)^d \Big/ \exp(\rho\gamma \log d) \\ &\leq \ d^{\kappa - \rho\gamma} \exp\left((e^\rho - 1) d \int_0^{l/d} \{f(x) + f(1-x)\}\, dx\right). \end{aligned} \tag{3.4}$$

The lemma follows since (3.3) implies that the right hand side of (3.4) converges to 0 as $d \to \infty$. □

Lemma 3.2.

$$d^{1/8} \mathbb{P}((A_d^1)^C) \to 0 \qquad \textit{as } d \to \infty. \tag{3.5}$$

Hence for any sequence of sets $\{\tilde{F}_d\}$, *such that*

$$d^3 \mathbb{P}(\mathbf{X}_0^d \notin \tilde{F}_d) \to 0 \qquad \textit{as } d \to \infty, \tag{3.6}$$

we have that

$$d^{1/8} \mathbb{P}\left(\cup_{t=0}^{[d^{5/2}]} \{\hat{\mathbf{X}}_t^d \notin \tilde{F}_d\}\right) \to 0 \qquad \textit{as } d \to \infty. \tag{3.7}$$

Proof. Let $\tilde{A}^1_d = \cap_{t=0}^{[d^{23/8}]} \{\mathbf{X}^d_t \in F^1_d\}$. By stationarity and Lemma 3.1,

$$d^{1/8}\mathbb{P}\left((\tilde{A}^1_d)^C\right) \to 0 \quad \text{as } d \to \infty.$$

From (2.1) we can write

$$J_d(\mathbf{x}^d) = \int h_d(\mathbf{z}^d) 1_{\{\mathbf{x}^d + \sigma_d \mathbf{z}^d \in (0,1)^d\}} \left\{1 \wedge \prod_{i=1}^d \frac{\exp(g(x_i + \sigma_d z_i))}{\exp(g(x_i))}\right\} d\mathbf{z}^d.$$

Let $g^* = \sup_{0 \leq y \leq 1} |g'(y)|$. By Taylor's Theorem

$$\sum_{i=1}^d g(x_i + \sigma_d z_i) - g(x_i) \geq -lg^*.$$

Hence for all $\mathbf{x}^d \in F^1_d$,

$$J_d(\mathbf{x}^d) \geq 2^{-\gamma \log d} e^{-lg^*} \geq d^{-\gamma} e^{-lg^*}.$$

For $d \geq 1$, let $\chi^d_0, \chi^d_1, \ldots$ denote independent and identically distributed random variables with $\mathbb{P}(\chi^d_0 = 1) = \exp(-lg^*) d^{-\gamma}$. Let $\theta_d = \sum_{i=0}^{[d^{23/8}]} 1_{\{\mathbf{X}^d_i \neq \mathbf{X}^d_{i+1}\}}$, the total number of accepted moves in the first $[d^{23/8}]$ iterations of the RWM process. Then $\theta_d \geq [d^{5/2}]$ and $\tilde{A}^1_d$ together imply A^1_d. Therefore

$$d^{1/8}\mathbb{P}((A^1_d)^C) \leq d^{1/8}\mathbb{P}((\tilde{A}^1_d)^C) + d^{1/8}\mathbb{P}(\theta_d < [d^{5/2}] | \tilde{A}^1_d).$$

However, conditional upon $\tilde{A}^1_d$,

$$\sum_{i=0}^{[d^{23/8}]} \chi^d_i \leq_{st} \theta_d,$$

and so,

$$d^{1/8}\mathbb{P}((A^1_d)^C) \quad \leq \quad d^{1/8}\mathbb{P}((\tilde{A}^1_d)^C) + d^{1/8}\mathbb{P}\left(\sum_{i=0}^{[d^{23/8}]} \chi^d_i < [d^{5/2}]\right). \tag{3.8}$$

By Chebychev's inequality the latter term on the right hand side of (3.8) converges to 0 as $d \to \infty$ and hence (3.5) is proved.

For a given sequence of sets $\tilde{F}_d$ satisfying (3.6), let $\tilde{D}_d = \cap_{t=0}^{[d^{23/8}]} \{\mathbf{X}^d_t \in \tilde{F}_d\}$ and $D_d = \cap_{t=0}^{[d^{5/2}]} \{\hat{\mathbf{X}}^d_t \in \tilde{F}_d\}$. Then

$$d^{1/8}\mathbb{P}(D^C_d) \leq d^{1/8}\mathbb{P}(\tilde{D}^C_d) + d^{1/8}\mathbb{P}((\tilde{A}^1_d)^C) + d^{1/8}\mathbb{P}(D^C_d | \tilde{D}_d, \tilde{A}^1_d). \tag{3.9}$$

The first two terms on the right hand side of (3.9) converge to 0 as $d \to \infty$ by (3.6) and Lemma 3.1, respectively. The final term on the right hand side of (3.9) converges to 0 as $d \to \infty$ by identical arguments to (3.8) since $\tilde{D}_d$ and $\theta_d > [d^{5/2}]$ together imply D_d. Therefore (3.7) is proved. □

Lemma 3.2 will prove very useful in proving results for $\{\hat{\mathbf{X}}_i^d\}$ by exploiting the fact that $\hat{\mathbf{X}}_0^d \equiv \mathbf{X}_0^d \sim \pi_d(\cdot)$ and $\{\mathbf{X}^d\}$ is in stationarity.

For $s \geq 0$, let $\hat{T}_d(s) = \frac{1}{d^2} \sum_{i=0}^{[sd^2-1]} L_d(\hat{\mathbf{X}}_i^d)$. An immediate consequence of Lemma 3.2 is the following Corollary.

Corollary 3.3. *For all $s \geq 0$, $T_d(s) \stackrel{p}{\longrightarrow} s\exp(f^* l/2)$ if $\hat{T}_d(s)|A_d^1 \stackrel{p}{\longrightarrow} s\exp(f^* l/2)$ as $d \to \infty$.*

Proof. Fix $s \geq 0$. By Lemma 3.2, $T_d(s) \stackrel{p}{\longrightarrow} s\exp(f^* l/2)$ if $T_d(s)|A_d^1 \stackrel{p}{\longrightarrow} s\exp(f^* l/2)$ as $d \to \infty$. Consider the characteristic function of $T_d(s)|A_d^1$. For all $\tau \in \mathbb{R}$,

$$\begin{aligned} \mathbb{E}[\exp(i\tau T_d(s))|A_d^1] &= \mathbb{E}\left[\prod_{j=0}^{[sd^2-1]} \mathbb{E}\left[\exp\left(\frac{i\tau}{d^2} G_j(J_d(\hat{\mathbf{X}}_j^d))\right)\middle| A_d^1, \{\hat{\mathbf{X}}^d\}\right]\middle| A_d^1\right] \\ &= \mathbb{E}\left[\prod_{j=0}^{[sd^2-1]} \frac{\exp(i\tau/d^2) J_d(\hat{\mathbf{X}}_j^d)}{1-(1-J_d(\hat{\mathbf{X}}_j^d))\exp(i\tau/d^2)}\middle| A_d^1\right]. \end{aligned}$$

Conditional upon A_d^1, $L_d(\hat{\mathbf{X}}_j^d) \leq \exp(lg^*)d^\gamma$. Hence, for all $d > s^2$ and $0 \leq j \leq sd^2 - 1$,

$$\frac{\exp(i\tau/d^2) J_d(\hat{\mathbf{X}}_j^d)}{1-(1-J_d(\hat{\mathbf{X}}_j^d))\exp(i\tau/d^2)} = 1 + \frac{i\tau}{d^2} L_d(\hat{\mathbf{X}}_j^d) + o(d^{-3}).$$

Therefore $\mathbb{E}[\exp(i\tau T_d(s))|A_d^1]$ and $\mathbb{E}[\exp(i\tau \hat{T}_d(s))|A_d^1]$ have the same limit as $d \to \infty$, and the corollary is proved. □

Therefore it is sufficient to study $\hat{T}_d(s)|A_d^1$. For $j \geq 1$, let $\tau_j^{d,\alpha} = \frac{1}{[d^\alpha]} \sum_{i=[(j-1)d^\alpha]}^{[jd^\alpha-1]} L_d(\hat{\mathbf{X}}_i^d)$. Then to prove (3.1) it is sufficient to show that for $j \geq 1$,

$$\tau_j^{d,\alpha}|A_d^1 \stackrel{p}{\longrightarrow} \exp(f^* l/2) \quad \text{as } d \to \infty. \tag{3.10}$$

We introduce the ip-RWM algorithm and define a suitable coupling to overcome problems in analysing the pseudo-RWM algorithm caused by dependence between the components. We shall define the ip-RWM before outlining a useful coupling of the two algorithms. In analysing (3.10), it is sufficient to consider a coupling of the two processes over $[d^{\alpha}]$ iterations. This coupling will allow us to use either of the two processes depending upon which is most convenient.

For all $d \geq 1$, define the ip-RWM algorithm as follows. For $j, k \geq 0$, let $\tilde{\mathbf{X}}_k^{d,j}$ denote the position of the ip-RWM process after k iterations with $\tilde{\mathbf{X}}_0^{d,j} = \hat{\mathbf{X}}_j^d$. That is, we start the coupling at the j^{th} iteration of the pseudo-RWM process and if the coupling is maintained for k iterations we have that $\tilde{\mathbf{X}}_k^{d,j} = \hat{\mathbf{X}}_{j+k}^d$. For $j \geq 0$, $k \geq 1$, $1 \leq i \leq d$ and $x \in [\sigma_d, 1-\sigma_d]$, if $\tilde{X}_{k-1,i}^{d,j} = x$, let $\tilde{Z}_{k,i}^{d,j}$ have probability density function (pdf)

$$j_x^d(z) = \begin{cases} \frac{1}{2}\left(1 + \frac{1}{2}\sigma_d g'(x) z\right) & -1 < z < 1, \\ 0 & \text{otherwise.} \end{cases}$$

For $j \geq 0$, $k \geq 1$, $1 \leq i \leq d$ and $x \in R_d^l$, if $\tilde{X}_{k-1,i}^{d,j} = x$, let $\tilde{Z}_{k,i}^{d,j}$ have pdf

$$j_x^d(z) = \begin{cases} (1 + x/\sigma_d)^{-1} & -x/\sigma_d < z < 1,\ x \in (0, \sigma_d) \\ (1 + (1-x)/\sigma_d)^{-1} & -1 < z < (1-x)/\sigma_d,\ x \in (1-\sigma_d, 1) \\ 0 & \text{otherwise.} \end{cases}$$

Finally, set

$$\tilde{X}_{k,i}^{d,j} = \tilde{X}_{k-1,i}^{d,j} + \sigma_d \tilde{Z}_{k,i}^{d,j}.$$

We require the following preliminaries and Lemma 3.4 before coupling the two processes.

We follow [8] and [11] in noting that, for any function h which is a twice differentiable function on $\mathbb{R}$, the function $z \mapsto 1 \wedge e^{h(z)}$ is also twice differentiable, except at a countable number of points, with first derivative given Lebesgue almost everywhere by the function

$$\frac{d}{dz} 1 \wedge e^{h(z)} = \begin{cases} h'(z) e^{h(z)} & \text{if } h(z) < 0 \\ 0 & \text{if } h(z) \geq 0. \end{cases}$$

Therefore letting $J_d^z(\mathbf{x}^d)$ denote the probability of accepting a move in the RWM given that $Z_1^d = z$, we have that

$$J_d^z(\mathbf{x}^d) = \left\{ J_d^0(\mathbf{x}^d) + \sigma_d g'(x_1) z \tilde{J}_d^0(\mathbf{x}^d) + o(\sigma_d^{3/2}) \right\} 1_{\{x_1 + \sigma_d z \in (0,1)\}}, \tag{3.11}$$

where $\mathbf{x}^{d-} = (x_2, x_3, \ldots, x^d)$, $\pi_{d-}(\mathbf{x}^{d-}) = \pi_{d-1}(\mathbf{x}^{d-})$,

$$\begin{aligned}\tilde{J}_d^0(\mathbf{x}^d) &= \mathbb{E}\left[\frac{\pi_{d-}(\mathbf{x}^{d-}+\sigma_d\mathbf{Z}^{d-})}{\pi_{d-}(\mathbf{x}^{d-})}; \frac{\pi_{d-}(\mathbf{x}^{d-}+\sigma_d\mathbf{Z}^{d-})}{\pi_{d-}(\mathbf{x}^{d-})} < 1\right] \\ &= \mathbb{E}\left[A_d(\mathbf{x}^{d-}+\sigma_d\mathbf{Z}^{d-})\exp\left(\sum_{i=2}^d\{g(x_i+\sigma_d Z_i^d)-g(x_i)\}\right); \sum_{i=2}^d\{g(x_i+\sigma_d Z_i^d)-g(x_i)\} < 0\right]\end{aligned} \tag{3.12}$$

and for all $\mathbf{z}^d \in \mathbb{R}^d$,

$$A_d(\mathbf{z}^{d-}) = \begin{cases} 1 & \mathbf{z}^{d-} \in (0,1)^{d-1} \\ 0 & \mathbf{z}^{d-} \notin (0,1)^{d-1}. \end{cases}$$

There are two cases to consider with (3.11) and (3.12). The first case is the uniform density where $g(x) = 0$ for all $0 \leq x \leq 1$. In this case $J_d^z(\mathbf{x}^d) = J_d^0(\mathbf{x}^d)1_{\{x_1+\sigma_d z\in(0,1)\}}$ and $\tilde{J}_d^0(\mathbf{x}^d) = 0$. The second case is where $\mathbb{E}_f[g'(X)^2] = \int g'(x)^2 f(x)\,dx > 0$. We shall focus on the non-uniform case with the uniform case following by similar, but considerably simpler, arguments. For $d \geq 1$, let

$$F_d^2 = \left\{\mathbf{x}^d : \left|\frac{1}{d}\sum_{i=1}^d g'(x_i)^2 - \mathbb{E}[g'(X_1)^2]\right| < d^{-1/4}\right\}.$$

For all $\kappa > 0$, by independence of the components of $\mathbf{X}_0^d$ and Markov's inequality

$$d^\kappa \mathbb{P}(\mathbf{X}_0^d \notin F_d^2) \to 0 \quad \text{as } d \to \infty. \tag{3.13}$$

LEMMA 3.4. *For $\mathbb{E}[g'(X_1)^2] > 0$ and for all $\mathbf{x}^d \in F_d^1 \cap F_d^2$,*

$$d^{1/4}\left|\frac{\tilde{J}_d^0(\mathbf{x}^d)}{J_d^0(\mathbf{x}^d)} - \frac{1}{2}\right| \to 0 \quad \textit{as } d \to \infty.$$

Proof. Fix $\mathbf{x}^d \in F_d^1 \cap F_d^2$. Let

$$\begin{aligned} I_d(\mathbf{x}^d) &= \sqrt{d}\sum_{i=2}^d\{g(x_i+\sigma_d Z_i) - g(x_i)\} \\ &= \frac{l}{\sqrt{d}}\sum_{i=2}^d g'(x_i)Z_i + o(d^{-1/3}). \end{aligned} \tag{3.14}$$

Therefore

$$\begin{aligned} &\exp(-d^{-1/3})\mathbb{E}\left[A_d(\mathbf{x}^{d-}+\sigma_d\mathbf{Z}^{d-})1_{\{I_d(\mathbf{x}^d)/\sqrt{d}<0\}}\right] - \mathbb{E}[1_{\{I_d(\mathbf{x}^d)/\sqrt{d}<-d^{-1/3}\}}] \\ \leq\ & \tilde{J}_d^0(\mathbf{x}^d) \leq \mathbb{E}\left[A_d(\mathbf{x}^{d-}+\sigma_d\mathbf{Z}^{d-})1_{\{I_d(\mathbf{x}^d)/\sqrt{d}<0\}}\right] \end{aligned}$$

and

$$\begin{aligned} & \exp(-d^{-1/3})\mathbb{E}\left[A_d(\mathbf{x}^{d-}+\sigma_d\mathbf{Z}^{d-})\right] - \mathbb{E}[1_{\{I_d(\mathbf{x}^d)/\sqrt{d}<-d^{-1/3}\}}] \\ \leq \; & J_d^0(\mathbf{x}^d) \leq \mathbb{E}\left[A_d(\mathbf{x}^{d-}+\sigma_d\mathbf{Z}^{d-})\right]. \end{aligned}$$

By Markov's inequality,

$$d^{\frac{1}{3}}\mathbb{P}(I_d(\mathbf{x}^d) < -d^{\frac{1}{6}}) \to 0 \quad \text{as } d \to \infty,$$

and so,

$$d^{1/4}\left|\frac{\tilde{J}_d^0(\mathbf{x}^d)}{J_d^0(\mathbf{x}^d)} - \frac{\mathbb{E}[A_d(\mathbf{x}^{d-}+\sigma_d\mathbf{Z}^{d-})1_{\{I_d(\mathbf{x}^d)/\sqrt{d}<0\}}]}{\mathbb{E}[A_d(\mathbf{x}^{d-}+\sigma_d\mathbf{Z}^{d-})]}\right| \to 0 \quad \text{as } d \to \infty, \tag{3.15}$$

since for $\mathbf{x}^d \in F_d^1$, $\mathbb{E}[A_d(\mathbf{x}^{d-}+\sigma_d\mathbf{Z}^{d-})] \geq e^{-lg^*}d^{-\gamma}$.

Note that

$$\frac{\mathbb{E}[A_d(\mathbf{x}^{d-}+\sigma_d\mathbf{Z}^{d-})1_{\{I_d(\mathbf{x}^d)/\sqrt{d}<0\}}]}{\mathbb{E}[A_d(\mathbf{x}^{d-}+\sigma_d\mathbf{Z}^{d-})]} = \mathbb{P}(I_d(\mathbf{x}^d) < 0|A_d(\mathbf{x}^{d-}+\sigma_d\mathbf{Z}^{d-}) = 1).$$

Let $B_d(\mathbf{x}^d) = \{2 \leq i \leq d; \sigma_d < x_i < 1-\sigma_d\}$ and let $\tilde{I}_d(\mathbf{x}^d) = \sqrt{d}\sum_{i\in B_d(\mathbf{x}^d)} lg'(x_i)Z_i$. For $\mathbf{x}^d \in F_d^1$, $|B_d(\mathbf{x}^d)^C| = b_d^l(\mathbf{x}^d) \leq \gamma \log d$, and so, using (3.14), for all sufficiently large d,

$$|I_d(\mathbf{x}^d) - \tilde{I}_d(\mathbf{x}^d)| \leq d^{-1/3}.$$

Thus

$$\begin{aligned} \mathbb{P}\left(\tilde{I}_d(\mathbf{x}^d) < -d^{-1/3}\right) & \leq \mathbb{P}(I_d(\mathbf{x}^d) < 0|A_d(\mathbf{x}^{d-}+\sigma_d\mathbf{Z}^{d-}) = 1) \\ & \leq \mathbb{P}\left(\tilde{I}_d(\mathbf{x}^d) < d^{-1/3}\right), \end{aligned} \tag{3.16}$$

since for $i \in B_d(\mathbf{x}^d)$, Z_i^d is independent of $A_d(\mathbf{x}^{d-}+\sigma_d\mathbf{Z}^{d-})$.

By the Central limit theorem, $\tilde{I}_d(\mathbf{x}^d) \stackrel{D}{\approx} \hat{I}_d(\mathbf{x}^d) \sim N(0,(l^2/d)\sum_{i\in B_d(\mathbf{x}^d)} g'(x_i)^2)$, and so, by the Berry-Esséen inequality, there exists $K_1 < \infty$ such that for any $a \in \mathbb{R}$,

$$|\mathbb{P}(\tilde{I}_d(\mathbf{x}^d) < a) - \mathbb{P}(\hat{I}_d(\mathbf{x}^d) < a)| \leq \frac{K_1}{\sqrt{d}}. \tag{3.17}$$

Since $\mathbf{x}^d \in F_d^2$, $\frac{1}{d}\sum_{i\in B_d(\mathbf{x}^d)} g'(x_i)^2 \geq \frac{1}{2}\mathbb{E}[g'(X)^2]$ for all sufficiently large d. It follows from (3.16) and (3.17) that there exists $K_2 < \infty$ such that

$$\left|\mathbb{P}(I_d(\mathbf{x}^d) < 0|A_d(\mathbf{x}^{d-}+\sigma_d\mathbf{Z}^{d-}) = 1) - \frac{1}{2}\right| \leq K_2 d^{-1/3}. \tag{3.18}$$

The lemma then follows from (3.15) and (3.18). □

LEMMA 3.5. *Suppose that* $\mathbf{x}^d \in F_d^1 \cap F_d^2$. *Then for* $j \geq 0$ *there exists a coupling such that*

$$d^{1/4}\mathbb{P}(\tilde{\mathbf{X}}_1^{d,j} \neq \hat{\mathbf{X}}_{j+1}^d | \tilde{\mathbf{X}}_0^{d,j} \equiv \hat{\mathbf{X}}_j^d = \mathbf{x}^d) \to 0 \quad \textit{as } d \to \infty.$$

Proof. We prove the result for $j = 0$ with the general case following by identical arguments.

Note that

$$\mathbb{P}(\hat{\mathbf{X}}_1^d \neq \tilde{\mathbf{X}}_1^{d,0} | \tilde{\mathbf{X}}_0^{d,0} \equiv \hat{\mathbf{X}}_0^d = \mathbf{x}^d) \leq \sum_{i=1}^d \mathbb{P}(\hat{X}_{1,i}^d \neq \tilde{X}_{1,i}^{d,0} | \tilde{\mathbf{X}}_0^{d,0} \equiv \hat{\mathbf{X}}_0^d = \mathbf{x}^d). \tag{3.19}$$

We shall focus on the first component, $\hat{X}_{0,1}^d \equiv \tilde{X}_{0,1}^{d,0} = x_1$ and consider the cases $x_1 \in [\sigma_d, 1-\sigma_d]$ and $x_1 \in R_d^l$, separately. We have different bounds for the two cases but we can again make use of the fact that, for $\mathbf{x}^d \in F_d^1$, $b_d^l(\mathbf{x}^d) \leq \gamma \log d$.

Starting with $x_1 \in [\sigma_d, 1-\sigma_d]$. For the ip-RWM, we have that

$$j_{x_1}^d(z) = \frac{1}{2}\left\{1 + \frac{1}{2}g'(x_1)\sigma_d z\right\}. \tag{3.20}$$

For the pseudo-RWM, we have that $\hat{X}_{1,1}^d = \hat{X}_{0,1}^d + \hat{Z}_{1,1}^d$ where $\hat{Z}_{1,1}^d$ has probability density function

$$\hat{j}_{\mathbf{x}^d}^d(z) = \frac{1}{2} \times \frac{J_d^z(\mathbf{x}^d)}{J_d(\mathbf{x}^d)}. \tag{3.21}$$

From (3.20) and (3.21), we can couple $\tilde{Z}_{1,1}^{d,0}$ and $\hat{Z}_{1,1}^d$ so that

$$\mathbb{P}(\tilde{Z}_{1,1}^{d,0} \neq \hat{Z}_{1,1}^d | \hat{\mathbf{X}}_0^d = \mathbf{x}^d) = \int_{-1}^1 |\hat{j}_{\mathbf{x}^d}^d(z) - j_{x_1}^d(z)|\, dz. \tag{3.22}$$

By (3.11), for $x_1 \in [\sigma_d, 1-\sigma_d]$,

$$J_d^z(\mathbf{x}^d) = J_d^0(\mathbf{x}^d) + \sigma_d g'(x_1) z \tilde{J}_d^0(\mathbf{x}^d) + o(\sigma_d^{7/4}). \tag{3.23}$$

Therefore for $\mathbf{x}^d \in F_d^1$ it follows from (3.21) and (3.23) that

$$\begin{aligned} \hat{j}_{\mathbf{x}^d}^d(z) &= \frac{1}{2} \times \frac{J_d^0(\mathbf{x}^d) + \sigma_d g'(x_1) z \tilde{J}_d^0(\mathbf{x}^d) + o(\sigma_d^{3/2})}{J_d^0(\mathbf{x}^d) + o(\sigma_d^{3/2})} \\ &= \frac{1}{2}\left\{1 + \frac{\tilde{J}_d^0(\mathbf{x}^d)}{J_d^0(\mathbf{x}^d)}\sigma_d g'(x_1) z + o(d^\gamma \sigma_d^{7/4})\right\}. \end{aligned} \tag{3.24}$$

It follows from Lemma 3.4 and (3.24) that

$$\hat{j}_{\mathbf{x}^d}^d(z) = j_{x_1}^d(z) + o(d^{-5/4}). \tag{3.25}$$

For $x_1 \in R_d^l$, we have that

$$J_d^z(\mathbf{x}^d) = \begin{cases} J_d^0(\mathbf{x}^d) + \sigma_d g'(x_1) z \tilde{J}_d^0(\mathbf{x}^d) + o(\sigma_d^{3/2}) & x_1 + \sigma_d z \in (0,1) \\ 0 & \text{otherwise.} \end{cases}$$

Focussing upon $0 < x_1 < \sigma_d$ (identical arguments apply for $1 - \sigma_d < x_1 < 1$), we have that

$$\begin{aligned} J_d(\mathbf{x}^d) &= \int_{-x_1/\sigma_d}^{1} \frac{1}{2} J_d^z(\mathbf{x}^d)\, dz \\ &= \frac{1}{2}\left(1 + \frac{x_1}{\sigma_d}\right) J_d^0(\mathbf{x}^d) + \frac{1}{2}\left(1 - \frac{x_1^2}{\sigma_d^2}\right) \sigma_d \tilde{J}_d^0(\mathbf{x}^d) g'(x_1) z + o(\sigma_d^{3/2}). \end{aligned} \tag{3.26}$$

Therefore for $x_1 + \sigma_d z \in (0,1)$, it follows from (3.26) and Lemma 3.4 that

$$\begin{aligned} \hat{j}_{\mathbf{x}^d}^d(z) &= \frac{1}{2} \times \frac{J_d^0(\mathbf{x}^d) + \sigma_d g'(x_1) z \tilde{J}_d^0(\mathbf{x}^d) + o(\sigma_d^{3/2})}{\frac{1}{2}\left(1 + \frac{x_1}{\sigma_d}\right) J_d^0(\mathbf{x}^d) + o(\sigma_d^{7/8})} \\ &= \left(1 + \frac{x_1}{\sigma_d}\right)^{-1} \left\{1 + \frac{\tilde{J}_d^0(\mathbf{x}^d)}{J_d^0(\mathbf{x}^d)} \sigma_d g'(x_1) z + o(d^\gamma \sigma_d^{7/8})\right\} \\ &= j_{x_1}^d(z) + o(d^{-3/4}). \end{aligned} \tag{3.27}$$

For $\mathbf{x}^d \in F_d^1$, $b_d^l(\mathbf{x}^d) \leq \gamma \log d$, and so, the lemma follows from (3.19), (3.22), (3.25) and (3.27).
□

The following Corollary is an immediate consequence of Lemma 3.5.

COROLLARY 3.6. *For* $0 \leq j_d \leq [d^{5/2} - d^\alpha]$, *let* $\tilde{\mathbf{X}}_0^{d,j_d} \equiv \hat{\mathbf{X}}_{j_d}^d$ *and let* $A_d^2(j_d) = \bigcap_{i=0}^{[d^\alpha]-1} \{\hat{\mathbf{X}}_{(j_d+i)}^d = \tilde{\mathbf{X}}_i^{d,j_d}\}$. *Then using the coupling outlined in Lemma 3.5, we have that*

$$d^{1/8} \mathbb{P}((A_d^2(j_d))^C) \to 0 \qquad as\ d \to \infty.$$

Proof. Let $\check{A}_d = \cap_{i=0}^{[d^{5/2}]} \{\hat{\mathbf{X}}_i^d \in F_d^1 \cap F_d^2\}$. By Lemma 3.2 and (3.13),

$$d^{1/8} \mathbb{P}(\check{A}_d^C) \to 0 \quad \text{ as } d \to \infty.$$

Therefore the Corollary follows from Lemma 3.5, since

$$d^{1/8} \mathbb{P}(A_d^2(j_d)^C) \leq d^{1/8} \mathbb{P}(\check{A}_d^C) + d^{1/8} \mathbb{P}(A_d^2(j_d)^C | \check{A}_d). \qquad \square$$

We now focus explicitly upon proving (3.2). This is achieved in Lemma 3.13 and involves studying the (random) configuration of the components in the rejection region. In particular, we show that after sufficiently many steps the configuration of the components is well approximated by a Poisson point process. This is established in Lemma 3.11, via a series of preliminary results. Then (3.2), and hence (3.1), follow straightforwardly.

For $d \geq 1$, $1 \leq i \leq d$, $0 \leq r \leq l$, $k \geq 0$ and $\mathbf{x}^d \in (0,1)^d$, let

$$\tilde{\chi}_i^d(\mathbf{x}^d; r; k) = \begin{cases} 1 & \text{if } \tilde{X}_{k,i}^{d,0} \in R_d^r \text{ given } \tilde{\mathbf{X}}_0^{d,0} = \mathbf{x}^d \\ 0 & \text{otherwise.} \end{cases}$$

The components of the ip-RWM algorithm move independently. Therefore we only need to explicitly state the starting value of component i when considering the i^{th} component. In other words, $\tilde{\chi}_i^d(\mathbf{x}^d; r; k) \equiv \tilde{\chi}_i^d(x_i; r; k)$. For $x \in (0,1)$, let

$$\tilde{q}^d(x; r; k) = \mathbb{E}[\tilde{\chi}_1^d(x; r; k)].$$

For a random variable Y defined upon $(0,1)$, let

$$\tilde{q}^d(Y; r; k) = \mathbb{E}[\tilde{\chi}_1^d(Y; r; k)|Y].$$

For the pseudo-RWM process define $\hat{\chi}_i^d(\mathbf{x}^d; r; k)$ and $\hat{q}^d(\mathbf{x}^d; r; k)$ in the obvious fashion replacing $\tilde{\mathbf{X}}_\cdot^{d,0}$ by $\hat{\mathbf{X}}_\cdot^d$. Finally, for $0 \leq r \leq l$ and $\mathbf{x}^d \in (0,1)^d$, let $\tilde{M}_d(\mathbf{x}^d; r; k) = \sum_{i=1}^d \tilde{\chi}_i^d(\mathbf{x}^d; r; k)$ and $\hat{M}_d(\mathbf{x}^d; r; k) = \sum_{i=1}^d \hat{\chi}_i^d(\mathbf{x}^d; r; k)$. It is worth noting that for $j \geq 0$ and $k \leq [d^\alpha]$, conditional upon $A_d^2(j)$, if $\tilde{\mathbf{X}}_0^{d,j} = \hat{\mathbf{X}}_j^d$, then

$$\tilde{M}_d(\tilde{\mathbf{X}}_0^{d,j}; r; k) \equiv \hat{M}_d(\hat{\mathbf{X}}_j^d; r; k).$$

For $d \geq 1$, $\mathbf{x}^d \in (0,1)^d$, $0 \leq r \leq l$ and $k \geq 0$, let $\tilde{\lambda}_d(\mathbf{x}^d; r; k) = \mathbb{E}[\tilde{M}_d(\mathbf{x}^d; r; k)]$ and let $\hat{\lambda}_d(\mathbf{x}^d; r; k) = \mathbb{E}[\hat{M}_d(\mathbf{x}^d; r; k)]$.

To prove that $\tau_j^{d,\alpha}|A_d^1 \xrightarrow{p} \exp(f^* l/2)$ as $d \to \infty$, it suffices by Chebychev's inequality to show that for any sequences of non-negative integers $\{i_d\}$ and $\{k_d\}$ such that $[d^\beta] \leq k_d \leq [d^\alpha]$ and $0 \leq i_d + k_d \leq [d^{5/2}]$,

$$cov(L_d(\hat{\mathbf{X}}_{i_d}^d), L_d(\hat{\mathbf{X}}_{i_d+k_d}^d)|A_d^1) \to 0 \quad \text{as } d \to \infty. \tag{3.28}$$

We shall prove (3.28) for arbitrary sequences $\{i_d\}$ and $\{k_d\}$ satisfying the above criterion.

For $d \geq 1$, let $a_d = k_d^{3/4}$ and

$$F_d^3 = \{\mathbf{x}^d : |b_d^{a_d}(\mathbf{x}^d) - \mathbb{E}[b_d^{a_d}(\mathbf{X}_0^d)]| \leq a_d^{3/4}\}.$$

Let $A_d^3 = \bigcap_{i=1}^{[d^{5/2}]} \{\hat{\mathbf{X}}_i^d \in F_d^3\}$. In other words, for $\mathbf{x}^d \in F_d^3$ the number of components 'close' to the boundary is well controlled.

LEMMA 3.7. *$d^{1/8}\mathbb{P}((A_d^3)^C) \to 0$ as $d \to \infty$.*

Proof. By stationarity and Markov's inequality, for all $m \in \mathbb{N}$,

$$d^3\mathbb{P}(|b_d^{a_d}(\mathbf{X}_0^d) - \mathbb{E}[b_d^{a_d}(\mathbf{X}_0^d)]| \geq a_d^{3/4}) \leq \frac{d^3}{a_d^{3m/2}}\mathbb{E}\left[(b_d^{a_d}(\mathbf{X}_0^d) - \mathbb{E}[b_d^{a_d}(\mathbf{X}_0^d)])^{2m}\right]. \tag{3.29}$$

However, $b_d^{a_d}(\mathbf{X}_0^d) \sim Bin\left(d, \int_0^{a_d/d}\{f(x) + f(1-x)\}\, dx\right)$, and so, for any $m \in \mathbb{N}$, there exists $K < \infty$ such that

$$\mathbb{E}\left[(b_d^{a_d}(\mathbf{X}_0^d) - \mathbb{E}[b_d^{a_d}(\mathbf{X}_0^d)])^{2m}\right] \leq K a_d^m.$$

Since $a_d \geq d^{3\beta/4}$, the right hand side of (3.29) converges to 0 as $d \to \infty$ by taking $m > 8/\beta$.

The lemma then follows by Lemma 3.2. □

The following Lemma is important for Lemma 3.11 in that it establishes useful bounds for the probability of a component being in the rejection region after $i_d + k_d$ iterations given that $\hat{\mathbf{X}}_{i_d}^d \in F_d^1 \cap F_d^2 \cap F_d^3$.

LEMMA 3.8. *For all $\mathbf{x}^d \in F_d^1 \cap F_d^2 \cap F_d^3$,*

$$d^{2\gamma} \sum_{i=1}^{d} \tilde{q}^d(x_i^d; l; k_d)^2 \to 0 \quad \text{as } d \to \infty. \tag{3.30}$$

Also for all $0 < x < 1$,

$$d^{2\gamma}\tilde{q}^d(x; l; k_d) \to 0 \quad \text{as } d \to \infty. \tag{3.31}$$

Proof. Fix $\mathbf{x}^d \in F_d^1 \cap F_d^2 \cap F_d^3$ and set $\tilde{\mathbf{X}}_0^{d,0} = \mathbf{x}^d$.

Studying the ip-RWM directly is not straightforward. However, by coupling each of the components of the ip-RWM to a reflected random walk (RRW) progress can be made. Explicitly, we couple the ip-RWM process to a random walk which is perturbed at the boundaries (RWB). Since each component of the ip-RWM are independent we focus on the first component with $\tilde{X}_{0,1}^{d,0} = x_1$.

For $d \geq 1$, let the RWB process, $\{P_t^d; t \geq 0\}$, be defined as follows. Set $P_0^d = x_1$. For $k \geq 1$, $P_k^d \in (\sigma_d, 1-\sigma_d)$, let $P_{k+1}^d = P_k^d + \sigma_d Z_k^d$ where

$$Z_k^d = \begin{cases} U[-P_k^d/\sigma_d, 1] & P_k^d \in (0, \sigma_d] \\ U[-1, 1] & P_k^d \in (\sigma_d, 1-\sigma_d) \\ U[-1, (1-P_k^d)/\sigma_d] & P_k^d \in [1-\sigma_d, 1). \end{cases}$$

For $d \geq 1$, the ip-RWM process, $\{\tilde{X}_{t,1}^d; t \geq 0\}$ can be constructed as follows. For $k \geq 0$, let

$$\tilde{X}_{k+1,1}^{d,0} = \tilde{X}_{k,1}^{d,0} + \sigma_d \tilde{Z}_k^d$$

where for $\tilde{X}_{k,1}^{d,0} \in (\sigma_d, 1-\sigma_d)$ $\tilde{Z}_k^d$ has pdf $j^d_{\tilde{X}_{k,1}^{d,0}}(z)$ $(-1 < z < 1)$ and for $\tilde{X}_{k,1}^{d,0} \in R_d^l$, $\tilde{Z}_k^d \sim U[-\min\{\tilde{X}_{k,1}^{d,0}/\sigma_d, 1\}, \min\{(1-\tilde{X}_{k,1}^{d,0})/\sigma_d, 1\}]$. For all $k \geq 0$, if $\tilde{X}_{k,1}^{d,0} = P_k^d$, with probability $\left(1 - \frac{1}{2} g^* \sigma_d\right)$, set $\tilde{Z}_k^d = Z_k^d$ and with the remaining probability draw $\tilde{Z}_k^d$ so that $\tilde{Z}_k^d$ has the correct distribution. (Remember $g^* = \sup_{0 \leq y \leq 1} |g'(y)|$.) Therefore we have a coupling such that

$$\mathbb{P}(\tilde{X}_{k_d,1}^{d,0} \neq P_{k_d}^d) \leq \frac{k_d l}{2d} g^*.$$

Finally, we define the RRW process $\{\check{P}_t^d; t \geq 0\}$. Set $\check{P}_0^d = x$ for some $x \in (0,1)$. Let $Z_0, Z_1, \ldots$ be independent and identically distributed according to $U \sim U[-1,1]$. For $k \geq 1$, set $\check{P}_{k+1}^d = \check{P}_k^d + \sigma_d Z_k$ with reflection at the boundaries 0 and 1, so that $\check{P}_k^d \in (0,1)$.

For $x \in (0,1)$, $l \geq 0$ and $k \geq 0$, let $p^d(x; l; k) = \mathbb{P}(P_k^d \in R_d^l | P_0^d = x)$ and $\check{p}^d(x; l; k) = \mathbb{P}(\check{P}_k^d \in R_d^l | \check{P}_0^d = x)$. A simple coupling argument shows that, for all $x \in (0,1)$,

$$p^d(x; l; k_d) \leq \check{p}^d(x; l; k_d).$$

This is because it is not possible to move a distance greater than $k_d \sigma_d$ from your starting point in k_d iterations. Therefore without loss of generality, we assume that $0 < x < (k_d+1)\sigma_d$ (symmetry arguments apply for $1-(k_d+1)\sigma_d < x < 1$), and so, by the reflection principle

$$\begin{aligned} \check{p}^d(x; l; k_d) &= \mathbb{P}\left(-\sigma_d < x + \sigma_d \sum_{i=0}^{k_d-1} Z_i < \sigma_d\right) \\ &= \mathbb{P}\left(-1 < x/\sigma_d + \sum_{i=0}^{k_d-1} Z_i < 1\right). \end{aligned} \tag{3.32}$$

By the Berry-Esséen theorem, there exists a positive constant, $K_1 < \infty$ say, such that for all $z \in \mathbb{R}$,

$$\left| \mathbb{P}\left(\sqrt{\frac{3}{k_d}} \sum_{i=0}^{k_d-1} Z_i \leq z \right) - \Phi(z) \right| \leq \frac{K_1}{\sqrt{k_d}} \tag{3.33}$$

where $\Phi(\cdot)$ denotes the cdf of a standard normal. Therefore it follows from (3.32) and (3.33) that there exists a positive constant, $K_2 < \infty$ say, such that for all $x \in (0,1)$,

$$\check{p}^d(x; l; k_d) \leq \frac{K_2}{\sqrt{k_d}}. \tag{3.34}$$

By Markov's inequality, for any $m \in \mathbb{N}$ and $\epsilon > 0$,

$$\begin{aligned} \mathbb{P}\left(\left| \frac{1}{\sqrt{k_d}} \sum_{i=0}^{k_d-1} Z_i \right| > \epsilon \right) &\leq \frac{1}{\epsilon^{2m}} \mathbb{E}\left[\left(\frac{1}{\sqrt{k_d}} \sum_{i=1}^{k_d} Z_i \right)^{2m} \right] \\ &\leq \frac{1}{\epsilon^{2m}} \left(\frac{2m!}{2^m} \right). \end{aligned} \tag{3.35}$$

For $a_d < x < (k_d+1)/d$, we have by (3.35) that, for any $m \in \mathbb{N}$,

$$\begin{aligned} \check{p}^d(x; l; k_d) &= \mathbb{P}\left(\left| x + \sigma_d \sum_{i=0}^{k_d-1} Z_i \right| < \sigma_d \right) \\ &\leq \mathbb{P}\left(\left| \sigma_d \sum_{i=0}^{k_d-1} Z_i \right| > \frac{a_d}{2d} \right) \\ &\leq (2l^2)^m (2m!) \left(\frac{\sqrt{k_d}}{a_d} \right)^{2m} \\ &\leq (2l^2)^m (2m!) k_d^{-m/2}. \end{aligned} \tag{3.36}$$

By taking $m > 2/\beta$ in (3.36), we have that

$$\check{p}^d(x; l; k_d) \leq (2l^2)^m (2m!) d^{-1} \qquad \text{as } d \to \infty. \tag{3.37}$$

Finally, for $(k_d+1)/d < x < 1 - (k_d+1)/d$, $\check{p}^d(x; l; k_d) = 0$.

Therefore we can put all the parts of the proof together. Firstly, by applying the couplings, we have that

$$\begin{aligned} \tilde{q}^d(x; l; k_d) &\leq p^d(x; l; k_d) + \frac{k_d l}{2d} g^* \\ &\leq \check{p}^d(x; l; k_d) + \frac{k_d l}{2d} g^*. \end{aligned} \tag{3.38}$$

By (3.34) and (3.38), (3.31) follows immediately.

Since $\mathbf{x}^d \in F_d^1 \cap F_d^2 \cap F_d^3$ it follows from (3.34), (3.37) and (3.38) that, there exists $K_3, K_4, m < \infty$ such that

$$\begin{aligned} d^{2\gamma} \sum_{i=1}^{d} \tilde{q}^d(x_i^d; l; k_d)^2 &\leq 2d^{2\gamma} \sum_{i=1}^{d} \left\{ \check{p}^d(x_i^d; l; k_d)^2 + \left(\frac{k_d l}{2d} g^* \right)^2 \right\} \\ &\leq 2d^{2\gamma} \left\{ K_3 a_d \left(\frac{K_2}{\sqrt{k_d}} \right)^2 + K_4 k_d (2l^2)^m (2m!) d^{-1} \frac{k_d^2}{d} \left(\frac{lg^*}{2} \right)^2 \right\} \end{aligned} \tag{3.39}$$

with the righthand side of (3.39) converging to 0 as $d \to \infty$. □

COROLLARY 3.9. *For any $m \geq 1$, there exists $K < \infty$ such that for all $d \geq 1$,*

$$\mathbb{E}[\tilde{q}^d(\hat{X}_{0,1}^d; l; k_d)^{2m}] \leq K d^{-(1+(m-3/4)\beta)}. \tag{3.40}$$

Proof Fix $m \geq 1$. From Lemma 3.8, there exists constants $K_1, K_2, K_3 < \infty$ such that for all $d \geq 1$,

$$\begin{aligned} \mathbb{E}[\tilde{q}(\hat{X}_{0,1}^d; l; k_d)^{2m}] &\leq \mathbb{P}(X_{0,1}^d \in R_d^{a_d}) \left(\frac{K_2}{\sqrt{k_d}} \right)^{2m} + \mathbb{P}(X_{0,1}^d \in R_d^{a_d}) \left(\frac{K_1}{d} \right)^{2m} \\ &\leq K_3 \frac{k_d^{3/4}}{d} \left(\frac{K_2}{\sqrt{k_d}} \right)^{2m} + \mathbb{P}(X_{0,1}^d \in R_d^{a_d}) \left(\frac{K_1}{d} \right)^{2m}. \end{aligned}$$

The Corollary follows since $k_d \geq d^\beta$. □

We continue by considering the mean number of components in the rejection region after $i_d + k_d$ iterations given that $\hat{\mathbf{X}}_{i_d}^d \in F_d^1 \cap F_d^2 \cap F_d^3$.

Suppose that $\mathbf{U}^d \sim \hat{\pi}_d(\cdot)$. Then by stationarity and symmetry, for all $0 \leq r \leq l$ and $k \geq 0$,

$$\begin{aligned} \mathbb{E}[\hat{\lambda}_d(\mathbf{U}^d; r; k)] &= \mathbb{E}[\hat{\lambda}_d(\mathbf{U}^d; r; 0)] \\ &= d\mathbb{E}[\hat{q}_1^d(\mathbf{U}^d; r; 0)] \\ &= d\mathbb{P}(U_1^d \in R_d^r). \end{aligned}$$

Note the $U_1^d \sim \hat{f}_d(\cdot)$ where $\hat{f}_d(\cdot)$ denotes the marginal distribution of a component of $\hat{\pi}_d(\cdot)$. For $x_1 \in (\sigma_d, 1 - \sigma_d)$, it follows that

$$\begin{aligned} \hat{f}_d(x_1) &= \frac{f(x_1) \int \{J_d^0(\mathbf{x}^d) + o(\sigma_d^{3/2})\} \pi_{d-1}(\mathbf{x}^{d-}) \, d\mathbf{x}^{d-}}{\int a_d(y_1) \{J_d^0(\mathbf{y}^d) + o(\sigma_d^{3/2})\} \pi_d(\mathbf{y}^d) \, d\mathbf{y}^d} \\ &= f(x_1) + O(d^{-1}) \end{aligned} \tag{3.41}$$

where for $0 < x < 1$, $a_d(x) = \mathbb{E}[1_{\{0<x+\sigma_d Z<1\}}]$. Similarly for $x_1 \in R_d^l$,

$$\hat{f}_d(x_1) = a_d(x_1) f(x_1) + o(d^{-1/2}).$$

Thus if $\hat{\lambda}(r) = f^* r(1 + r/2l)$,

$$\sqrt{d}|d\mathbb{P}(U_1^d \in R_d^r) - \hat{\lambda}(r)| \to 0 \quad \text{as } d \to \infty.$$

Utilising the coupling of the pseudo-RWM and ip-RWM processes

$$d^{1/8}|\mathbb{E}[\tilde{\lambda}_d(\mathbf{U}^d; r; k_d)] - \hat{\lambda}(r)| \to 0 \quad \text{as } d \to \infty. \tag{3.42}$$

For any $0 \leq r \leq l$, let

$$F_d^4(r) = \left\{ \mathbf{x}^d : |\tilde{\lambda}_d(\mathbf{x}^d; r; k_d) - \hat{\lambda}(r)| < d^{-\gamma} \right\}. \tag{3.43}$$

Lemma 3.10. *For any $0 \leq r \leq l$ and for $0 \leq t \leq [d^{5/2}]$,*

$$d^{1/8}\mathbb{P}\left(\cup_{t=0}^{[d^{5/2}]} \{\hat{\mathbf{X}}_t^d \notin F_d^4(r)\} \right) \to 0 \quad \text{as } d \to \infty.$$

Proof. Fix $0 \leq r \leq l$. By Lemma 3.2, it suffices to show that

$$d^3 \mathbb{P}(\mathbf{X}_0^d \notin F_d^4(r)) \to 0 \quad \text{as } d \to \infty.$$

By the triangle inequality,

$$\begin{aligned} & |\tilde{\lambda}_d(\mathbf{X}_0^d; r; k_d) - \hat{\lambda}(r)| \\ \leq \; & |\tilde{\lambda}_d(\mathbf{X}_0^d; r; k_d) - \mathbb{E}[\tilde{\lambda}_d(\mathbf{X}_0^d; r; k_d)]| + |\mathbb{E}[\tilde{\lambda}_d(\mathbf{X}_0^d; r; k_d)] - \mathbb{E}[\tilde{\lambda}_d(\mathbf{U}^d; r; k_d)]| \\ & \quad + |\mathbb{E}[\tilde{\lambda}_d(\mathbf{U}^d; r; k_d)] - \hat{\lambda}(r)|. \end{aligned} \tag{3.44}$$

By (3.42) the third term on the right hand side of (3.44) is less than $d^{-\gamma}/3$ for all sufficiently large d.

From (3.41), there exists a constant $K < \infty$, independent of d, such that for all $d \geq 1$,

$$\int_0^1 |f(x) - \hat{f}_d(x)| \, dx \leq \frac{K}{d}.$$

Therefore it follows from (3.31) that for all sufficiently large d,

$$|\mathbb{E}[\tilde{q}_1^d(X_{0,1}^d; r; k_d)] - \mathbb{E}[\tilde{q}_1^d(U_1^d; r; k_d)]| < d^{-(1+\gamma)}/3$$

and consequently the second term on the right hand side of (3.44) is bounded by $d^{-\gamma}/3$.

Therefore to prove the lemma it is sufficient to show that

$$d^3\mathbb{P}(|\tilde{\lambda}_d(\mathbf{X}_0^d; r; k_d) - \mathbb{E}[\tilde{\lambda}_d(\mathbf{X}_0^d; r; k_d)]| > d^{-\gamma}/3) \to 0 \quad \text{as } d \to \infty.$$

By Markov's inequality, for all $m \in \mathbb{N}$,

$$\begin{aligned} & d^3\mathbb{P}(|\tilde{\lambda}_d(\mathbf{X}_0^d; r; k_d) - \mathbb{E}[\tilde{\lambda}_d(\mathbf{X}_0^d; r; k_d)]| > d^{-\gamma}/3) \\ \leq\ & 3^{2m} d^{3+2m\gamma} \mathbb{E}[(\tilde{\lambda}_d(\mathbf{X}_0^d; r; k_d) - \mathbb{E}[\tilde{\lambda}_d(\mathbf{X}_0^d; r; k_d)])^{2m}]. \end{aligned}$$

Since the components of $\mathbf{X}_0^d$ are independent, we have that

$$\begin{aligned} & d^3\mathbb{P}(|\tilde{\lambda}_d(\mathbf{X}_0^d; r; k_d) - \mathbb{E}[\tilde{\lambda}_d(\mathbf{X}_0^d; r; k_d)]| > d^{-\gamma}/3) \\ \leq\ & 3^{2m} d^{3+2m\gamma} \sum_{i_1=1}^d \sum_{i_2=1}^d \cdots \sum_{i_m=1}^d \mathbb{E}\left[\prod_{j=1}^m (\tilde{q}_d(X_{0,i_j}^d; r; k_d) - \mathbb{E}[\tilde{q}_d(X_{0,i_j}^d; r; k_d)])^2\right] \\ \leq\ & 3^{2m} d^{3+2m\gamma} \sum_{i_1=1}^d \sum_{i_2=1}^d \cdots \sum_{i_m=1}^d \mathbb{E}\left[\prod_{j=1}^m \tilde{q}_d(X_{0,i_j}^d; r; k_d)^2\right] \\ \leq\ & 3^{2m} d^{3+2m\gamma} \sum_{i_1=1}^d \sum_{i_2=1}^d \cdots \sum_{i_m=1}^d \mathbb{E}\left[\prod_{j=1}^m \tilde{q}_d(X_{0,i_j}^d; l; k_d)^2\right]. \end{aligned} \tag{3.45}$$

By Corollary 3.9, (3.40), there exists $K < \infty$ such that the right hand side of (3.45) is less than $Kd^{3+2m\gamma}d^{-\beta m/4} \leq Kd^{3-3m\gamma}$ (since $20\gamma < \beta$).

The lemma follows by choosing $m > 1/\gamma$. □

For $n \geq 1$, let

$$F_d(n) = F_d^1 \cap F_d^2 \cap F_d^3 \cap \left\{\bigcap_{i=1}^n F_d^4\left(\frac{il}{n}\right)\right\}.$$

It follows immediately from the previous results that for any sequence $\{i_d\}$

$$d^{1/8}\mathbb{P}(\hat{\mathbf{X}}_{i_d}^d \notin F_d(n)) \to 0 \quad \text{as } d \to \infty. \tag{3.46}$$

For any $n \in \mathbb{N}$ and $1 \leq i \leq n$, let

$$\tilde{S}_n^d(\mathbf{x}^d; i; k) = \sum_{j=1}^d \left\{\tilde{\chi}_i^d(x_j; il/n; k) - \tilde{\chi}_j^d(x_j; (i-1)l/n; k)\right\}$$

with

$$\tilde{\mathbf{S}}_n^d(\mathbf{x}^d;k) = (S_n^d(\mathbf{x}^d;1;k), S_n^d(\mathbf{x}^d;2;k), \ldots, S_n^d(\mathbf{x}^d;n;k)).$$

Let $\mathbf{S}_n = (S_n^1, S_n^2, \ldots, S_n^n)$ where the components of $\mathbf{S}_n$ are independent Poisson random variables with $S_n^i \sim Po(\lambda_{n,i})$ and

$$\lambda_{n,i} = \hat{\lambda}(il/n) - \hat{\lambda}((i-1)l/n) \quad (1 \leq i \leq n).$$

We are now in position to prove the following key result.

LEMMA 3.11. *For any $n \in \mathbb{N}$ and $\mathbf{x}^d \in F_d(n)$,*

$$\tilde{\mathbf{S}}_n^d(\mathbf{x}^d;k_d) \xrightarrow{D} \mathbf{S}_n \quad \textit{as } d \to \infty.$$

Proof. Fix $n \in \mathbb{N}$ and $\mathbf{x}^d \in F_d(n)$. Let

$$\check{\mathbf{S}}_n^d(\mathbf{x}^d;k_d) = (\check{S}_n^d(\mathbf{x}^d;1;k_d), \check{S}_n^d(\mathbf{x}^d;2;k_d), \ldots, \check{S}_n^d(\mathbf{x}^d;n;k_d))$$

where for $1 \leq i \leq n$, $\check{S}_n^d(\mathbf{x}^d;i;k_d)$ are independent Poisson random variables with means

$$\lambda_{n,i}^d(\mathbf{x}^d;k_d) = \tilde{\lambda}_d(\mathbf{x}^d;il/n;k_d) - \tilde{\lambda}_d(\mathbf{x}^d;(i-1)l/n;k_d).$$

The lemma is proved by showing that

$$\begin{aligned} d_{TV}(\tilde{\mathbf{S}}_n^d(\mathbf{x}^d;k_d), \mathbf{S}_n) &\leq d_{TV}(\tilde{\mathbf{S}}_n^d(\mathbf{x}^d;k_d), \check{\mathbf{S}}_n^d(\mathbf{x}^d;k_d)) + d_{TV}(\check{\mathbf{S}}_n^d(\mathbf{x}^d;k_d), \mathbf{S}_n) \\ &\to 0 \quad \text{as } d \to \infty. \end{aligned} \tag{3.47}$$

By [2], Theorem 1,

$$d_{TV}(\tilde{\mathbf{S}}_n^d(\mathbf{x}^d;k_d), \check{\mathbf{S}}_n^d(\mathbf{x}^d;k_d)) \leq \sum_{i=1}^d \tilde{q}^d(x_i;l;k_d)^2. \tag{3.48}$$

By Lemma 3.8, (3.30) the right hand side of (3.48) converges to 0 as $d \to \infty$.

For the second term on the right hand side of (3.47), it suffices to show that

$$\check{\mathbf{S}}_n^d(\mathbf{x}^d;k_d) \xrightarrow{D} \mathbf{S}_n \quad \text{as } d \to \infty.$$

(For discrete random variables convergence in distribution and convergence in total variation distance are equivalent, see [3], page 254.)

The components of $\check{\mathbf{S}}_n^d(\hat{\mathbf{X}}^d; k_d)$ and $\mathbf{S}_n$ are independent, and therefore it is sufficient to show that, for all $1 \leq i \leq n$,

$$\check{S}_n^d(\mathbf{x}^d; i; k_d) \xrightarrow{D} S_{n,i} \quad \text{as } d \to \infty. \tag{3.49}$$

For all $1 \leq i \leq n$, (3.49) holds, if

$$\lambda_{n,i}^d(\mathbf{x}^d; k_d) \to \lambda_{n,i} \quad \text{as } d \to \infty. \tag{3.50}$$

Therefore the lemma follows from (3.50) since $\mathbf{x}^d \in \bigcap_{i=1}^n F_d^4(il/n)$. □

We set about utilising Lemma 3.11 to prove (3.2).

For $n \in \mathbb{N}$, $1 \leq i \leq n$ and $\mathbf{x}^d \in (0,1)^d$, let $\tilde{b}_d^{n,i}(\mathbf{x}^d) = b_d^{il/n}(\mathbf{x}^d) - b_d^{(i-1)l/n}(\mathbf{x}^d)$ with $\tilde{\mathbf{b}}_d^n(\mathbf{x}^d) = (\tilde{b}_d^{n,1}(\mathbf{x}^d), \tilde{b}_d^{n,2}(\mathbf{x}^d), \ldots, \tilde{b}_d^{n,n}(\mathbf{x}^d))$. For $n \in \mathbb{N}$ and $\mathbf{s} = (s_1, s_2, \ldots, s_n) \in \mathbb{R}^n$, let

$$\begin{aligned} \check{a}_n(\mathbf{s}) &= \prod_{i=1}^n \left(\frac{1}{2} + \frac{(i-1)}{2n} \right)^{-s_i} \\ \hat{a}_n(\mathbf{s}) &= \prod_{i=1}^n \left(\frac{1}{2} + \frac{i}{2n} \right)^{-s_i}. \end{aligned}$$

The key point to note is, that for all $n \in \mathbb{N}$ and for all $\mathbf{x}^d \in (0,1)^d$,

$$\hat{a}_n(\tilde{\mathbf{b}}_d^n(\mathbf{x}^d)) \leq L_d(\mathbf{x}^d) \leq \check{a}_n(\tilde{\mathbf{b}}_d^n(\mathbf{x}^d)).$$

Therefore, conditional upon $A_d^2(i_d)$,

$$\hat{a}_n(\tilde{\mathbf{S}}_n^d(\mathbf{x}^d; k_d)) \leq_{st} \{L_d(\hat{\mathbf{X}}_{i_d+k_d}^d) | \hat{\mathbf{X}}_{i_d}^d = \mathbf{x}^d\} \leq_{st} \check{a}_n(\tilde{\mathbf{S}}_n^d(\mathbf{x}^d; k_d)). \tag{3.51}$$

Corollary 3.12. *Fix $n \in \mathbb{N}$. For $d \geq 1$ let $\mathbf{x}^d \in F_d(n)$. Then*

$$\begin{aligned} \mathbb{E}[\hat{a}_n(\tilde{\mathbf{S}}_n^d(\mathbf{x}^d; k_d))] &\to \mathbb{E}[\hat{a}_n(\mathbf{S}_n)] \quad \text{as } d \to \infty, \\ \mathbb{E}[\check{a}_n(\tilde{\mathbf{S}}_n^d(\mathbf{x}^d; k_d))] &\to \mathbb{E}[\check{a}_n(\mathbf{S}_n)] \quad \text{as } d \to \infty. \end{aligned}$$

Proof. By [6], Theorem 29.2, and Lemma 3.11

$$\hat{a}_n(\tilde{\mathbf{S}}_n^d(\mathbf{x}^d; k_d)) \xrightarrow{D} \hat{a}_n(\mathbf{S}_n) \quad \text{as } d \to \infty, \tag{3.52}$$

$$\check{a}_n(\tilde{\mathbf{S}}_n^d(\mathbf{x}^d; k_d)) \xrightarrow{D} \check{a}_n(\mathbf{S}_n) \quad \text{as } d \to \infty. \tag{3.53}$$

For all $n \in \mathbb{N}$,

$$\hat{a}_n(\tilde{\mathbf{S}}_n^d(\mathbf{x}^d; k_d)) \leq \check{a}_n(\tilde{\mathbf{S}}_n^d(\mathbf{x}^d; k_d)) \leq 2^{\tilde{S}_1^d(\mathbf{x}^d;1;k_d)}. \tag{3.54}$$

The lemma follows immediately from (3.52), (3.53) and (3.54) if there exists $t > 2$ such that

$$\sup_d \mathbb{E}[t^{\tilde{S}_1^d(\mathbf{x}^d;1;k_d)}] < \infty.$$

Since the components of $\tilde{S}_1^d(\mathbf{x}^d; 1; k_d)$ are independent Bernoulli random variables, for all $t \in \mathbb{R}$,

$$\mathbb{E}[t^{\tilde{S}_1^d(\mathbf{x}^d;1;k_d)}] \quad \leq \quad \exp\left((t-1)\tilde{\lambda}_d(\mathbf{x}^d; l; k_d)\right). \tag{3.55}$$

Finally for $\mathbf{x}^d \in F_d(n)$, the right hand side of (3.55) is less than $\exp((t-1)2\hat{\lambda}(l))$ for all sufficiently large d and the lemma is proved. □

Note that $\mathbb{E}[\hat{a}_n(\mathbf{S}_n)], \mathbb{E}[\check{a}_n(\mathbf{S}_n)] \to \exp(f^* l/2)$ as $n \to \infty$.

Lemma 3.13.

$$\mathbb{E}[L_d(\hat{\mathbf{X}}_{i_d+k_d}^d)|\hat{\mathbf{X}}_{i_d}^d] \xrightarrow{p} \exp(f^* l/2) \qquad \textit{as } d \to \infty.$$

Proof. Fix $\epsilon > 0$. Choose $n \geq 1$, such that

$$|\mathbb{E}[\hat{a}_n(\mathbf{S}_n)] - \exp(f^* l/2)|, \; |\mathbb{E}[\check{a}_n(\mathbf{S}_n)] - \exp(f^* l/2)| \leq \epsilon/2.$$

For all $\mathbf{x}^d \in (0,1)^d$, by the triangle inequality,

$$\begin{aligned} |\mathbb{E}[\hat{a}_n(\tilde{\mathbf{S}}_n(\mathbf{x}^d; k_d)] - \exp(f^* l/2)| \quad &\leq \quad |\mathbb{E}[\hat{a}_n(\tilde{\mathbf{S}}_n(\mathbf{x}^d; k_d)] - \mathbb{E}[\hat{a}_n(\mathbf{S}_n)]| + |\mathbb{E}[\hat{a}_n(\mathbf{S}_n)] - \exp(f^* l/2)| \\ |\mathbb{E}[\check{a}_n(\tilde{\mathbf{S}}_n(\mathbf{x}^d; k_d)] - \exp(f^* l/2)| \quad &\leq \quad |\mathbb{E}[\check{a}_n(\tilde{\mathbf{S}}_n(\mathbf{x}^d; k_d)] - \mathbb{E}[\check{a}_n(\mathbf{S}_n)]| + |\mathbb{E}[\check{a}_n(\mathbf{S}_n)] - \exp(f^* l/2)|. \end{aligned} \tag{3.56}$$

For $\mathbf{x}^d \in F_d(n)$, the right hand sides of the equations (3.56) are less than ϵ for all sufficiently large d. Therefore, for $\hat{\mathbf{X}}_{i_d}^d \in F_d(n)$, it follows from (3.51) and (3.56), that conditional upon $A_d^2(i_d)$,

$$|\mathbb{E}[L_d(\tilde{\mathbf{X}}_{i_d+k_d}^d)|\hat{\mathbf{X}}_{i_d}^d = \mathbf{x}^d] - \exp(f^* l/2)| < \epsilon,$$

for all sufficiently large d. The lemma follows by Corollary 3.6 and (3.46). □

LEMMA 3.14. *For any sequence of positive integers $\{i_d\}$ such that $0 \leq i_d \leq [d^{5/2}]$,*

$$\sup_d \mathbb{E}[L_d(\hat{\mathbf{X}}^d_{i_d})^4 | A^1_d] < \infty.$$

Proof. We shall begin by assuming that for all $d \geq 1$, $i_d \geq [d^\beta]$.

Therefore $i_d = j_d + [d^\beta]$ where $j_d \geq 0$. By (3.51), we have that

$$\begin{aligned} \mathbb{E}[L_d(\hat{\mathbf{X}}^d_{i_d})^4 | A^1_d] \quad \leq \quad & \left(e^{lg^*} d^\gamma\right)^4 \{\mathbb{P}(A^2_d(j_d)^C | A^1_d) + \mathbb{P}(\hat{\mathbf{X}}^d_{j_d} \notin F_d(1) | A^1_d)\} \\ & + \mathbb{E}\left[(2^4)^{\tilde{S}^d_1(\hat{\mathbf{X}}^d_{j_d};1;[d^\beta])} | \hat{\mathbf{X}}^d_{j_d} \in F_d(1), A^2_d(j_d), A^1_d\right]. \end{aligned} \tag{3.57}$$

Since $4\gamma < 1/8$, the first two terms on the right hand side of (3.57) converge to 0 as $d \to \infty$, by Corollary 3.6 and (3.46), respectively. The latter term is bounded for all $d \geq 1$ since

$$\begin{aligned} \sup_{\mathbf{x}^d \in F_d(1)} \mathbb{E}\left[16^{\tilde{S}^d_1(\mathbf{x}^d;1;[d^\beta])}\right] \quad & \leq \quad \exp(15\tilde{\lambda}(\mathbf{x}^d; l; [d^\beta])) \\ & \leq \quad \exp(30\hat{\lambda}(l)), \end{aligned}$$

for all sufficiently large d, *c.f.* Corollary 3.12.

The proof for $i_d \leq [d^\beta]$ is similar making use of $A^2_d(0)$ and the fact that $\hat{\mathbf{X}}^d_0 \sim \pi_d(\cdot)$. □

COROLLARY 3.15.

$$cov(L_d(\hat{\mathbf{X}}^d_{i_d}), L_d(\hat{\mathbf{X}}^d_{i_d+k_d}) | A^1_d) \to 0 \quad \textit{as } d \to \infty.$$

Proof. Since $\mathbb{E}[L_d(\hat{\mathbf{X}}^d_{i_d+k_d}) | A^1_d] \to \exp(f^* l/2)$ as $d \to \infty$, an immediate consequence of Lemmas 3.13 and 3.14 is that

$$L_d(\hat{\mathbf{X}}^d_{i_d}) \left\{\mathbb{E}[L_d(\hat{\mathbf{X}}^d_{i_d+k_d}) | \hat{\mathbf{X}}^d_0] - \mathbb{E}[L_d(\hat{\mathbf{X}}^d_{i_d+k_d})]\right\} | A^1_d \xrightarrow{p} 0 \quad \text{as } d \to \infty. \tag{3.58}$$

The lemma follows from (3.58) by proving that the sequence

$$\left\{L_d(\hat{\mathbf{X}}^d_{i_d}) \left(\mathbb{E}[L_d(\hat{\mathbf{X}}^d_{i_d+k_d}) | \hat{\mathbf{X}}^d_{i_d}] - \mathbb{E}[L_d(\hat{\mathbf{X}}^d_{i_d+k_d})]\right) | A^1_d; d \geq 1\right\}$$

is uniformly integrable which is an immediate consequence of Lemma 3.14. □

We can finally prove (2.5).

THEOREM 3.16. *For any $T > 0$,*

$$\sup_{0 \leq t \leq T} |W^d_t - \psi_l(t)| \xrightarrow{p} 0 \quad \textit{as } d \to \infty. \tag{3.59}$$

Proof. Fix $T > 0$. By Chebychev's inequality, (3.10) follows immediately from Corollary 3.15. However (3.10) implies (3.1). Since $W^d_\cdot$ and $T_d(\cdot)$ are non-decreasing in t, it follows from (3.1), for all $t \geq 0$,

$$W_t^d \xrightarrow{p} t\exp(-f^*l/2) = \psi_l(t) \quad \text{as } d \to \infty. \tag{3.60}$$

Since W_t^d is bounded and non-decreasing in t with $\psi_l(\cdot)$ continuous, (3.59) follows trivially from (3.60), see, for example, [1], Lemma 5.1. □

4. Convergence of the pseudo RWM process. The aim of this section is to formally state and prove the main weak convergence result of the paper, Theorem 4.1.

THEOREM 4.1.

$$\hat{V}^d \Rightarrow \hat{V} \quad \textit{as } d \to \infty.$$

We begin with some background before embarking on the proof itself. Let G_d be the (discrete-time) generator of $\hat{\mathbf{X}}^d$, and let H be an arbitrary test function of the first component only. Thus

$$G_d H(\mathbf{x}^d) \quad = \quad d^2 \mathbb{E}\left[(H(\hat{\mathbf{X}}_1^d) - H(\hat{\mathbf{X}}_0^d)) | \hat{\mathbf{X}}_0^d = \mathbf{x}^d\right]$$

The generator G of the (limiting) one-dimensional diffusion for an arbitrary test function H, is given by

$$GH(x) = \frac{l^2}{3}\left\{\frac{1}{2}g'(x)H'(x) + \frac{1}{2}H''(x)\right\}, \tag{4.1}$$

at least for $x \in (0,1)$.

Firstly note that the diffusion defined by (4.1) is regular, see [9], page 366. Therefore by [9], Chapter 8, Corollary 1.2, it is sufficient to restrict attention to functions

$$H \in \mathcal{D} \equiv \{h : h \in \hat{C}([0,1]) \cap C^2((0,1)) \cap \mathcal{D}^*, Gh \in \hat{C}([0,1])\}$$

where $\mathcal{D}^*$ is obtained by setting $q_i = 0$ $(i = 0, 1)$ in [9], page 367, (1.11) and is given by

$$\mathcal{D}^* = \{h : h'(0) = h'(1) = 0\}. \tag{4.2}$$

For $d \geq 1$, let F_d^1 and A_d^1 be defined as in Section 3. Since $\mathbb{P}((A_d^1)^C) \to 0$ as $d \to \infty$, by [9], Chapter 4, Corollary 8.7, $\hat{V}^d \Rightarrow \hat{V}$ would follow from

$$\sup_{\mathbf{x}^d \in F_d^1} |G_d H(\mathbf{x}^d) - GH(x_1)| \to 0 \quad \text{as } d \to \infty. \tag{4.3}$$

This would be the standard method of proof and would follow from that of [12] (with some differences in dealing with boundary behaviour) and $H \in C_c^\infty$ ([9], Chapter 8, Theorem 2.1).

Unfortunately (4.3) does *not* hold for our problem, the uniformity of convergence breaking down at the boundaries. This is more than just a technical issue, since it merely reflects the fact that the one step transition dynamics of the first component is not asymptotically Markov, transitions depending on $b_d^r(\mathbf{x}_d)$ in particular.

Instead we need a homogenisation argument which involves looking at the generator of $[d^\delta]$ steps of $\hat{\mathbf{X}}^d$, sufficient time to allow the rapidly mixing $b_d^r(\mathbf{x}_d)$ to average over its stationary measure.

We shall require some constants throughout the proof and so throughout this section, fix $0 < \gamma < \beta < \delta < \alpha < \frac{1}{4}$.

For $s \in \mathbb{Z}^+$, let $\hat{\mathbf{P}}_s^d = \hat{\mathbf{X}}_{[d^\delta]s}^d$. Let G_d^δ be the (discrete-time) generator of $\hat{\mathbf{P}}_s^d$, and so, for $H \in \mathcal{D}$

$$\begin{aligned} G_d^\delta H(\mathbf{x}^d) &= \frac{d^2}{[d^\delta]} \mathbb{E}\left[(H(\hat{\mathbf{P}}_1^d) - H(\hat{\mathbf{P}}_0^d)) | \hat{\mathbf{P}}_0^d = \mathbf{x}^d \right] \\ &= \frac{d^2}{[d^\delta]} \sum_{j=0}^{[d^\delta - 1]} \mathbb{E}\left[H(\hat{\mathbf{X}}_{j+1}^d) - H(\hat{\mathbf{X}}_j^d) | \hat{\mathbf{X}}_0^d = \mathbf{x}^d \right] \\ &= \frac{1}{[d^\delta]} \sum_{j=0}^{[d^\delta - 1]} \mathbb{E}[G_d H(\hat{\mathbf{X}}_j^d) | \hat{\mathbf{X}}_0^d = \mathbf{x}^d]. \end{aligned} \tag{4.4}$$

For all $t \geq 0$, let $\check{V}_t^d = \hat{P}_{[d^2t/[d^\delta]],1} = \hat{X}_{[d^\delta] \times [d^2t/[d^\delta]],1}$. Note that for all $t \geq 0$,

$$|[d^2t] - [d^\delta] \times [d^2t/[d^\delta]]| \leq [d^\delta]$$

and

$$|\hat{X}_{t+1}^d - \hat{X}_t^d| \leq \sigma_d.$$

Hence for all $T > 0$,

$$\sup_{0 \leq s \leq T} |\hat{V}_s^d - \check{V}_s^d| \leq [d^\delta] \sigma_d,$$

and so, $\hat{V}^d \Rightarrow \hat{V}$ if $\check{V}^d \Rightarrow \hat{V}$ as $d \to \infty$.

We proceed by analysing $G_dH(\mathbf{x}^d)$. Then using (4.4) we can study the generator $G_d^\delta H(\mathbf{x}^d)$.

LEMMA 4.2.

$$G_dH(\mathbf{x}^d) = \frac{J_d^0(\mathbf{x}^d)}{J_d(\mathbf{x}^d)}\left\{GH(x_1) + K_dH(x_1) + \frac{l^2}{2}\mathbb{E}[Z_1^2(H''(x_1+\psi_1^d) - H''(x_1))] + o(d^{-1/5})\right\} \tag{4.5}$$

where ψ_1^d *lies between* 0 *and* $\sigma_d Z_1$ *and for* $y \in (\sigma_d, 1-\sigma_d)$, $K_d(y) = 0$ *and for* $y \in R_d^l$,

$$K_dH(y) = ld\mathbb{E}[Z_1 1_{\{0<y+\sigma_d Z_1<1\}}]H'(y) + \frac{l^2}{2}\{\mathbb{E}[Z_1^2 1_{0<y+\sigma_d Z_1<1}] - 1/3\}(H''(y) + g'(y)H'(y)). \tag{4.6}$$

Proof. For $d \geq 1$, fix $\mathbf{x}^d \in F_d^1$. Then

$$\begin{aligned} G_dH(\mathbf{x}^d) &= d^2\mathbb{E}\left[H(\hat{\mathbf{X}}_1^d) - H(\hat{\mathbf{X}}_0^d)|\hat{\mathbf{X}}_0^d = \mathbf{x}^d\right] \\ &= \frac{d^2}{J_d(\mathbf{x}^d)}\mathbb{E}\left[(H(\mathbf{x}^d + \sigma_d\mathbf{Z}^d) - H(\mathbf{x}^d))\left\{1 \wedge \frac{\pi_d(\mathbf{x}^d+\sigma_d\mathbf{Z}^d)}{\pi_d(\mathbf{x}^d)}\right\}\right]. \end{aligned} \tag{4.7}$$

The right hand side of (4.7) is familiar in that it is the generator of the RWM-algorithm divided by the acceptance probability, see for example, [12] page 113.

Firstly, note that

$$H(x_1+\sigma_d Z_1) - H(x_1) = \sigma_d Z_1 H'(x_1) + \frac{\sigma_d^2}{2}Z_1^2 H''(x_1) + \frac{\sigma_d^2}{2}Z_1^2\{H''(x_1+\psi_1^d) - H''(x_1)\},$$

where the random variable ψ_1^d lies between 0 and $\sigma_d Z_1$.

An immediate consequence of (3.11) and Lemma 3.4 is that

$$J_d^z(\mathbf{x}^d) = 1_{\{0<x_1+\sigma_d z<1\}}\left\{J_d^0(\mathbf{x}^d)\left\{1 + \frac{1}{2}\sigma_d g'(x_1)z\right\} + o(d^{-6/5})\right\}, \tag{4.8}$$

since $J_d^0(\mathbf{x}^d) \geq e^{-lg^*}d^{-\gamma}$ $(\mathbf{x}^d \in F_d^1)$.

Thus

$$\begin{aligned} &G_dH(\mathbf{x}^d) \\ = &\frac{d^2}{J_d(\mathbf{x}^d)}\mathbb{E}_{Z_1}\left[(H(x_1+\sigma_d Z_1) - H(x_1))\mathbb{E}_{\mathbf{Z}^{d-}}\left[1 \wedge \frac{\pi_d(\mathbf{x}^d+\sigma_d\mathbf{Z}^d)}{\pi_d(\mathbf{x}^d)}\right]\right] \text{ by (4.8)} \\ = &\frac{d^2}{J_d(\mathbf{x}^d)}\mathbb{E}\left[\left\{\sigma_d Z_1 H'(x_1) + \frac{\sigma_d^2}{2}Z_1^2 H''(x_1) + \frac{\sigma_d^2}{2}Z_1^2\{H''(x_1+\psi_1^d) - H''(x_1)\}\right\}\right. \\ &\left.\times\left\{J_d^0(\mathbf{x}^d)\left\{1+\frac{1}{2}\sigma_d g'(x_1)Z_1\right\} + o(d^{-6/5})\right\}1_{\{0<x_1+\sigma_d Z_1<1\}}\right] \end{aligned}$$

which is equal to (4.5) as required. □

LEMMA 4.3. *For all* $\mathbf{x}^d \in F_d^1$ *with* $x_1 \in (\sigma_d, 1-\sigma_d)$,

$$\sup_{x_1 \in (\sigma_d, 1-\sigma_d)} |G_d H(\mathbf{x}^d) - GH(x_1)| \to 0 \quad \text{as } d \to \infty. \tag{4.9}$$

Furthermore, there exists $0 < K^* < \infty$ *such that for all* $d \geq 1$, $\mathbf{x}^d \in F_d$ *and* $x_1 \notin (\sigma_d, 1-\sigma_d)$,

$$|G_d H(\mathbf{x}^d)| \leq K^*. \tag{4.10}$$

Proof. Since H'' is continuous on $[0,1]$,

$$\mathbb{E}[Z_1^2 (H''(x_1 + \psi_1^d) - H''(x_1))] \to 0 \quad \text{as } d \to \infty.$$

For $\sigma_d < x_1 < 1 - \sigma_d$, we have that

$$J_d(\mathbf{x}^d) = \int_{-1}^{1} \frac{1}{2} J_d^z(\mathbf{x}^d)\, dz = J_d^0(\mathbf{x}^d) + o(\sigma_d^{3/2}). \tag{4.11}$$

Therefore (4.9) follows immediately from (4.5) and (4.11).

For $x_1 \notin (\sigma_d, 1-\sigma_d)$, there exists $0 < y < l$ such that $x_1 = y/d$ or $x_1 = 1 - y/d$. For $x_1 = y/d$, by Taylor's Theorem

$$ld\mathbb{E}[Z_1 1_{\{0 < y + \sigma_d Z_1 < 1\}}] H'(x_1) \;=\; \frac{ly}{4}\left(1 - \left(\frac{y}{l}\right)^2\right) H''(\tilde{\psi}_1^d)$$

where $\tilde{\psi}_i^d \in (0, \sigma_d)$ since $H'(0) = 0$. Therefore there exists $K_0 < \infty$ such that the right hand side of (4.6) is less than K_0 for all $d \geq 1$. Similarly, there exists $K_1 < \infty$ such that for all $d \geq 1$ and $x_1 \in (1-\sigma_d, 1)$, $|K_d H(\mathbf{x}^d)| \leq K_1$. For all $\mathbf{x}^d$, $J_d^0(\mathbf{x}^d)/J_d(\mathbf{x}^d) \leq 2$, and so, (4.10) follows from (4.5) and (4.6). □

We are now in a position to complete the proof of Theorem 4.1.

Proof of Theorem 4.1. By [9], Chapter 4, Corollary 8.7 (f) and Lemma 3.2, the result $\check{V}^d \Rightarrow \hat{V}$ follows if we can demonstrate

$$\sup_{\mathbf{x}^d \in F_d^1} |G_d^\delta H(\mathbf{x}^d) - GH(x_1)| \to 0 \quad \text{as } d \to \infty. \tag{4.12}$$

Let $\{j_d\}$ denote any sequence of positive integers such that $[d^\beta] \leq j_d \leq [d^\delta]$. By Lemma 3.8, (3.31) for any $0 < x_1 < 1$ and for all sufficiently large $d \geq 1$, $\mathbb{P}(\hat{X}^d_{j_d,1} \notin (\sigma_d, 1-\sigma_d) | \hat{\mathbf{X}}^d_0 \in F^1_d, \hat{X}^d_{0,1} = x_1) \leq d^{-\gamma}$. Furthermore, note that for $\hat{X}^d_{0,1} = x_1$,

$$|\hat{X}^d_{j_d,1} - x_1| \leq j_d \sigma_d,$$

and so,

$$|GH(\hat{X}^d_{j_d,1}) - GH(x_1)| \stackrel{p}{\longrightarrow} 0 \quad \text{as } d \to \infty. \tag{4.13}$$

For any random variable Y, let

$$\mathbb{E}^{j,d}_x[Y] = \mathbb{E}[Y | \hat{X}^d_{0,1} = x, \hat{X}^d_{j,1} \in (\sigma_d, 1-\sigma_d)].$$

Therefore by (4.5)

$$\begin{aligned} & |\mathbb{E}[G_d H(\hat{\mathbf{X}}^d_{j_d}) | \hat{X}^d_{0,1} = x_1] - GH(x_1)| \\ \leq\; & (K^* + \sup_{0 \leq x \leq 1} |GH(x)|) \mathbb{P}(\hat{X}^d_{j_d,1} \notin R^l_d | \hat{\mathbf{X}}^d_0 \in F^1_d, \hat{X}^d_{0,1} = x_1) + |\mathbb{E}^{j_d,d}_{x_1}[G_d H(\hat{X}^d_{j_d,1})] - GH(x_1)| \\ \leq\; & (K^* + \sup_{0 \leq x \leq 1} |GH(x)|) d^{-\gamma} + \left| \mathbb{E}^{j_d,d}_{x_1} \left[G_d H(\hat{\mathbf{X}}^d_{j_d}) \right] - GH(x_1) \right|. \end{aligned} \tag{4.14}$$

By (4.5),

$$\mathbb{E}^{j_d,d}_{x_1} \left[G_d H(\hat{\mathbf{X}}^d_{j_d}) \right] = \mathbb{E}^{j_d,d}_{x_1} \left[\frac{J^0_d(\hat{\mathbf{X}}^d_{j_d})}{J_d(\hat{\mathbf{X}}^d_{j_d})} GH(\hat{X}^d_{j_d,1}) \right] + \epsilon_d$$

where $\epsilon_d \to 0$ as $d \to \infty$.

By the triangle inequality,

$$\begin{aligned} & \left| \mathbb{E}^{j_d,d}_{x_1} \left[\frac{J^0_d(\hat{\mathbf{X}}^d_{j_d})}{J_d(\hat{\mathbf{X}}^d_{j_d})} GH(\hat{X}^d_{j_d,1}) \right] - GH(x_1) \right| \\ \leq & \left| \mathbb{E}^{j_d,d}_{x_1} \left[\frac{J^0_d(\hat{\mathbf{X}}^d_{j_d})}{J_d(\hat{\mathbf{X}}^d_{j_d})} \{GH(\hat{X}^d_{j_d,1}) - GH(x_1)\} \right] \right| + \left| \mathbb{E}^{j_d,d}_{x_1} \left[\frac{J^0_d(\hat{\mathbf{X}}^d_{j_d})}{J_d(\hat{\mathbf{X}}^d_{j_d})} GH(x_1) \right] - GH(x_1) \right|. \end{aligned} \tag{4.15}$$

The first term on the right hand side of (4.15) converges to 0 as $d \to \infty$ by (4.13). The latter term on the right hand side of (4.15) converges to 0 as $d \to \infty$ since $X^d_{j_d,1} \notin R^d_d$. Therefore the right hand side of (4.14) converges to 0 as $d \to \infty$, and so, for all $\mathbf{x}^d \in F^1_d$,

$$\left| \frac{1}{[d^\delta]} \sum_{j=[d^\beta]}^{[d^\delta - 1]} \mathbb{E}[G_d H(\hat{\mathbf{X}}^d_j) | \hat{\mathbf{X}}^d_0 = \mathbf{x}^d] - GH(x_1) \right| \to 0 \quad \text{as } d \to \infty. \tag{4.16}$$

Therefore (4.12) follows from (4.4), (4.5) and (4.16). □

5. Extensions. We consider two extensions of Theorem 2.1. Firstly, rather than having $f(\cdot)$ non-zero on a bounded interval, we could have $f(\cdot)$ non-zero on the positive half-line. That is,

$$f(x)\exp(g(x)) \qquad (x > 0)$$

and $f(x) = 0$ otherwise.

THEOREM 5.1. *Fix $l > 0$. For all $d \geq 1$, let $\mathbf{X}_0^d = (X_{0,1}^d, X_{0,2}^d, \ldots, X_{0,d}^d)$ be such that all of its components are distributed according to $f(\cdot)$ with $\sup_{x \geq 0} |g'(x)| = g^* < \infty$. Then, as $d \to \infty$,*

$$V^d \Rightarrow V$$

where $V_0 \sim f(\cdot)$ and V satisfies the (reflected) Langevin SDE on $[0, \infty)$

$$dV_t = \sqrt{\nu(l)} dB_t + \frac{1}{2}\nu(l) g'(V_t)\, dt + dL_t^0(V)$$

where

$$\nu(l) = \frac{l^2}{3} \exp\left(-\frac{f^\star l}{4}\right)$$

and $f^\star = \lim_{x \downarrow 0} f(x)$.

Proof. The proof of the theorem is virtually identical to the proof of Theorem 2.1, and so, the details are omitted. □

Note that we have assumed that $g'(\cdot)$ is bounded on $[0, \infty)$. This assumption is almost certainly stronger than necessary with $g'(\cdot)$ being Lipschitz and/or satisfying certain moment conditions probably being sufficient, *c.f.* [12].

Theorem 5.1 is unsurprising with the speed of the diffusion depending upon the number of components close to the discontinuity at 0. Let $a_d(l)$ be defined as in (2.6), the average acceptance rate of the random walk Metropolis algorithm in d dimensions, and let

$$a(l) = \exp(-f^\star l/4).$$

COROLLARY 5.2.

$$\lim_{d \to \infty} a_d(l) = a(l)$$

$\nu(l)$ is maximised by

$$l = \hat{l} = \frac{8}{f^\star}.$$

Also

$$a(\hat{l}) = \exp(-2) = 0.1353.$$

Therefore the conclusions are identical to Corollary 2.2 that in order to maximise the speed of the limiting diffusion it is sufficient to choose l such that the average acceptance rate is e^{-2}.

The second and more important extension of Theorem 2.1 follows on from [11]. In [11], the Metropolis-within-Gibbs algorithm was considered, where only a fixed proportion c ($0 < c \leq 1$) of the components are updated at each iteration. For given $d \geq 1$ at each iteration $c_d d$ of the components are randomly chosen and new values for these components are proposed using random walk Metropolis with proposal variance $\sigma^2_{d,c_d} = (l/d)^2$. The remaining $(1-c_d)d$ components remain fixed at their current values. Finally it is assumed that $c_d \to c$ as $d \to \infty$.

The following result assumes that $f(\cdot)$ is nonzero on $(0,1)$ only. The extension to the positive half-line is trivial.

THEOREM 5.3. *Fix $0 < c \leq 1$ and $l > 0$. For all $d \geq 1$, let $\mathbf{X}^d_0 = (X^d_{0,1}, X^d_{0,2}, \ldots, X^d_{0,d})$ be such that all of its components are distributed according to $f(\cdot)$. Then, as $d \to \infty$,*

$$V^d \Rightarrow V$$

where $V_0 \sim f(\cdot)$ and V satisfies the (reflected) Langevin SDE on $[0,1]$

$$dV_t = \sqrt{\nu_c(l)} dB_t + \frac{1}{2}\nu_c(l) g'(V_t)\, dt + dL^0_t(V) - dL^1_t(V)$$

where

$$\nu_c(l) = \frac{cl^2}{3} \exp\left(-\frac{cf^* l}{2}\right)$$

and $f^* = \lim_{x \downarrow 0} \frac{f(x)+f(1-x)}{2}$.

Let $a^{c_d}_d(l)$ denote the average acceptance rate of the RWM algorithm in d dimensions where a proportion c_d of the components are updated at each iteration. Let

$$a^c(l) = \exp(-cf^* l/2).$$

We then have the following Corollary which mirrors Corollaries (2.2) and (5.2).

Corollary 5.4. *Let $c_d \to c$ as $d \to \infty$.*

$$\lim_{d\to\infty} a_d^{c_d}(l) = a^c(l)$$

For fixed $0 < c \leq 1$ $\nu_c(l)$ is maximised by

$$l = \hat{l}_c = \frac{4}{cf^*},$$

and

$$\nu_c(\hat{l}_c) = \frac{1}{c}\nu_1(\hat{l}_1).$$

Also

$$a(\hat{l}_c) = \exp(-2) = 0.1353.$$

Corollary 5.4 is of fundamental importance from a practical point of view, in that, it shows that the optimal speed of the limiting diffusion is inversely proportional to c. Therefore the optimal action is to choose c as close to 0 as possible. Furthermore we have shown that not only is full-dimensional RWM bad for discontinuous target densities but it is the worst algorithm of all the Metropolis-within-Gibbs RWM algorithms.

REFERENCES

[1] Ball, F. and Lyne, O. (2001) Stochastic multitype SIR epidemics among a population partitioned into households. *Adv. Appl. Prob.* **33** 99–123.

[2] Barbour, A. D. (1988) Stein's Method and Poisson Process Convergence. *J. Appl. Prob.* **25 A** 175–184

[3] Barbour, A. D., Holst, L. and Janson, S. (1992). *Poisson Approximation*. Oxford University Press, Oxford.

[4] Bédard, M. (2006) Weak convergence of Metropolis algorithms for non-IID target distributions. *Submitted to Ann. Appl. Prob.*

[5] Billingsley, P. (1968) *Convergence of Probability Measures*. Wiley, New York.

[6] Billingsley, P. (1979) *Probability and Measure*. Wiley, New York.

[7] Breyer, L., Piccioni, M. and Scarlatti, S. (2004) Optimal scaling of MALA for nonlinear regression. *Ann. Appl. Prob.* **14** 1479–1505.

[8] Breyer, L. and Roberts, G. O. (2000) From Metropolis to diffusions: Gibbs states and optimal scaling. *Stochastic Process. Appl.* **90** 181–206.

[9] Ethier, S. N. and Kurtz, T. G. (1986) *Markov processes, characterization and convergence.* Wiley, New York.

[10] Lefèvre, C. and Utev, S. (1999). Branching Approximation for the Collective Epidemic Model. *Method. and Comp. Appl. Prob.* **1** 211–228.

[11] Neal, P. J. and Roberts, G. O. (2006) Optimal Scaling for partially updating MCMC algorithms. *Ann. Appl. Prob.* **16** 475–515.

[12] Roberts, G. O., Gelman, A. and Gilks, W. R. (1997) Weak convergence and optimal scaling of Random walk Metropolis algorithms. *Ann. Appl. Prob.* **7** 110–120.

[13] Roberts, G. O. and Rosenthal, J. S. (1998) Optimal scaling of discrete approximations to Langevin diffusions. *J. Roy. Statist. Soc. Ser. B* **60** 255–268.

[14] Roberts, G. O. and Rosenthal, J. S. (2001) Optimal scaling for various Metropolis-Hastings algorithms. *Statist. Science* **16** 351–367.

School of Mathematics
University of Manchester
Sackville Street
Manchester
M60 1QD
UK
e-mail: P.Neal-2@manchester.ac.uk

Dept of Statistics
University of Warwick
Coventry
CV4 7AL
UK
e-mail: Gareth.O.Roberts@warwick.ac.uk

Department of Mathematics
Brock University
St. Catharines
ON L2S 3A1
Canada
e-mail: jyuen@brocku.ca