J. Appl. Prob. **54,** 1233–1260 (2017)
*doi:*10.1017/jpr.2017.61

OPTIMAL SCALING OF THE RANDOM WALK METROPOLIS ALGORITHM UNDER L^p MEAN DIFFERENTIABILITY

ALAIN DURMUS,* *LTCI, CNRS and Télécom ParisTech*
SYLVAIN LE CORFF,** *Université Paris-Sud, CNRS and Université Paris-Saclay*
ERIC MOULINES,*** *Ecole Polytechnique*
GARETH O. ROBERTS,**** *University of Warwick*

Abstract

In this paper we consider the optimal scaling of high-dimensional random walk Metropolis algorithms for densities differentiable in the L^p mean but which may be irregular at some points (such as the Laplace density, for example) and/or supported on an interval. Our main result is the weak convergence of the Markov chain (appropriately rescaled in time and space) to a Langevin diffusion process as the dimension d goes to ∞. As the log-density might be nondifferentiable, the limiting diffusion could be singular. The scaling limit is established under assumptions which are much weaker than the one used in the original derivation of Roberts *et al.* (1997). This result has important practical implications for the use of random walk Metropolis algorithms in Bayesian frameworks based on sparsity inducing priors.

Keywords: Random walk Metropolis; Markov chain Monte Carlo; optimal scaling; L^p mean differentiability

2010 Mathematics Subject Classification: Primary 65C05
Secondary 60F05

1. Introduction

A wealth of contributions has been devoted to the study of the behaviour of high-dimensional Markov chains. One of the most powerful approaches for that purpose is the scaling analysis, introduced by Roberts *et al.* [15]. Assume that the target distribution has a density with respect to the d-dimensional Lebesgue measure given by

$$\pi^d(x^d) = \prod_{i=1}^{d} \pi(x_i^d). \tag{1}$$

The random walk Metropolis–Hastings (RWM) updating scheme was first applied in [11] and proceeds as follows. Given the current state X_k^d, a new value $Y_{k+1}^d = (Y_{k+1,i}^d)_{i=1}^d$ is

Received 5 April 2016; revision received 18 February 2017.
The supplementary material for this article can be found at http://doi.org/10.1017/jpr.2017.61.
* Postal address: Télécom ParisTech, 46 rue Barrault, 75634 Paris, France.
Email address: alain.durmus@telecom-paristech.fr
** Postal address: Département de Mathématiques, Université Paris-Sud, 91405 Orsay Cedex, France.
Email address: sylvain.lecorff@math.u-psud.fr
*** Postal address: Centre de Mathématiques Appliquées, Ecole Polytechnique, Route de Saclay, 91128 Palaiseau Cedex, France. Email address: eric.moulines@polytechnique.edu
**** Postal address: Department of Statistics, University of Warwick, Coventry CV4 7AL, UK.
Email address: gareth.o.roberts@warwick.ac.uk

obtained by moving independently each coordinate, i.e. $Y^d_{k+1,i} = X^d_{k,i} + \ell d^{-1/2} Z^d_{k+1}$, where $\ell > 0$ is a scaling factor and $(Z_k)_{k \geq 1}$ is a sequence of independent and identically distributed (i.i.d.) Gaussian random variables. Here ℓ governs the overall size of the proposed jump and plays a crucial role in determining the efficiency of the algorithm. The proposal is then accepted or rejected according to the acceptance probability $\alpha(X^d_k, Y^d_{k+1})$, where $\alpha(x^d, y^d) = 1 \wedge \pi^d(y^d)/\pi^d(x^d)$. If the proposed value is accepted it becomes the next current value, otherwise the current value is left unchanged:

$$X^d_{k+1} = X^d_k + \ell d^{-1/2} Z^d_{k+1} \mathbf{1}_{\mathcal{A}^d_{k+1}}, \tag{2}$$

$$\mathcal{A}^d_{k+1} = \left\{ U_{k+1} \leq \prod_{i=1}^{d} \frac{\pi(X^d_{k,i} + \ell d^{-1/2} Z^d_{k+1,i})}{\pi(X^d_{k,i})} \right\}, \tag{3}$$

where $(U_k)_{k \geq 1}$ is a sequence of i.i.d. uniform random variables on $[0, 1]$ independent of $(Z_k)_{k \geq 1}$.

Under some regularity assumptions on π, it has been proved in [15] that if X^d_0 is distributed according to the stationary distribution π^d, then each component of $(X^d_k)_{k \geq 0}$ appropriately rescaled in time converges weakly to a Langevin diffusion process with invariant distribution π as $d \to +\infty$.

This result allows us to compute the asymptotic mean acceptance rate and to derive a practical rule to tune the factor ℓ. It was shown in [15] that the speed of the limiting diffusion as a function of ℓ has a unique maximum. The corresponding mean acceptance rate in stationarity is equal to 0.234.

These results have been derived for target distributions of the form (1), where $\pi(x) \propto \exp(-V(x))$ and where V is three-times continuously differentiable. Therefore, they do not cover the cases where the target density is continuous but not smooth, for example the Laplace distribution which plays a key role as a sparsity-inducing prior in high-dimensional Bayesian inference.

The aim of this paper is to extend the scaling results for the RWM algorithm introduced in the seminal paper [15, Theorem 3] to absolutely continuous densities differentiable in the L^p mean (DLM) for some $p \geq 2$ but which can be either nondifferentiable at some points or supported on an interval. As shown in [10, Section 17.3], differentiability of the square root of the density in the L^2 norm implies a quadratic approximation property for the log-likelihood known as the local asymptotic normality. As shown below, the DLM permits the quadratic expansion of the log-likelihood without paying the twice-differentiability price usually demanded by such a Taylor expansion (such an expansion of the log-likelihood plays a key role in [15]).

The paper is organised as follows. In Section 2 the target density π is assumed to be positive on $\mathbb{R}$. In Theorem 2 we prove that under the DLM assumption of this paper, the average acceptance rate and the expected squared jump distance are the same as in [15]. In Theorem 3 we show that, under the same assumptions, the rescaled in time Markov chain produced by the RWM algorithm converges weakly to a Langevin diffusion. We show that these results may be applied to a density of the form $\pi(x) \propto \exp(-\lambda|x| + U(x))$, where $\lambda \geq 0$ and U is a smooth function. In Section 3 we focus on the case where π is supported only on an open interval of $\mathbb{R}$. Under appropriate assumptions, in Theorem 4 and Theorem 5 we show that the same asymptotic results (limiting average acceptance rate and limiting Langevin diffusion associated with π) hold. We apply our results to gamma and beta distributions. The proofs are postponed to Section 4 and Section 5.

2. Positive target density on $\mathbb{R}$

The key of the proof of our main result is to show that the acceptance ratio and the expected squared jump distance converge to a finite and nontrivial limit. In the original proof of [15], the density of the product form (1) with

$$\pi(x) \propto \exp(-V(x)) \tag{4}$$

is three-times continuously differentiable and the acceptance ratio is expanded using the usual pointwise Taylor formula. More precisely, the log-ratio of the density evaluated at the proposed value and at the current state is given by $\sum_{i=1}^{d} \Delta V_i^d$, where

$$\Delta V_i^d = V(X_i^d) - V(X_i^d + \ell d^{-1/2} Z_i^d), \tag{5}$$

with X^d distributed according to π^d and Z^d a d-dimensional standard Gaussian random variable independent of X. The two leading terms are

$$\ell d^{-1/2} \sum_{i=1}^{d} \dot{V}(X_i^d) Z_i^d \quad \text{and} \quad \frac{\ell^2 d^{-1} \sum_{i=1}^{d} \ddot{V}(X_i^d)(Z_i^d)^2}{2},$$

where $\dot{V}$ and $\ddot{V}$ are the first and second derivatives of V, respectively. By the central limit theorem, the first term converges in distribution to a zero-mean Gaussian random variable with variance $\ell^2 I$, where

$$I = \int_{\mathbb{R}} \dot{V}^2(x)\pi(x)\,\mathrm{d}x. \tag{6}$$

Note that I is the Fisher information associated with the translation model $\theta \mapsto \pi(x+\theta)$ evaluated at $\theta = 0$. Under appropriate technical conditions, and using the dual representation of the Fisher information:

$$-\mathbb{E}[\ddot{V}(X)] = \mathbb{E}[(\dot{V}(X))^2] = I, \tag{7}$$

the second term converges almost surely to $-\ell^2 I/2$. Assuming that these limits exist, the acceptance ratio in the RWM algorithm converges to $\mathbb{E}[1 \wedge \exp(Z)]$, where Z is a Gaussian random variable with mean $-\ell^2 I/2$ and variance $\ell^2 I$; elementary computations show that $\mathbb{E}[1 \wedge \exp(Z)] = 2\Phi(-\ell/2\sqrt{I})$, where Φ denotes the cumulative distribution function of a standard normal distribution.

For $t \geq 0$, denote by Y_t^d the linear interpolation of the Markov chain $(X_k^d)_{k\geq 0}$ after time rescaling:

$$Y_t^d = (\lceil dt \rceil - dt)X_{\lfloor dt \rfloor}^d + (dt - \lfloor dt \rfloor)X_{\lceil dt \rceil}^d = X_{\lfloor dt \rfloor}^d + (dt - \lfloor dt \rfloor)\ell d^{-1/2} Z_{\lceil dt \rceil}^d \mathbf{1}_{\mathcal{A}_{\lceil dt \rceil}^d}, \tag{8}$$

where $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ denote the lower and the upper integer part functions. Note that, for all $k \geq 0$, $Y_{k/d}^d = X_k^d$. Denote by $(B_t,\ t \geq 0)$ the standard Brownian motion.

Theorem 1. (See [15].) *Suppose that the target π^d and the proposal distribution are given by (1)–(4) and (2), respectively. Assume that*

(i) *V is twice continuously differentiable and $\dot{V}$ is Lipschitz continuous;*

(ii) *$\mathbb{E}[(\dot{V}(X))^8] < \infty$ and $\mathbb{E}[(\ddot{V}(X))^4] < \infty$, where X is distributed according to π.*

https://doi.org/10.1017/jpr.2017.61 Published online by Cambridge University Press

Then $(Y^d_{t,1},\ t \geq 0)$, *where* $Y^d_{t,1}$ *is the first component of the vector* Y^d_t *defined in (8), converges weakly in the Wiener space (equipped with the uniform topology) to the Langevin diffusion*

$$\mathrm{d}Y_t = \sqrt{h(\ell)}\,\mathrm{d}B_t - \tfrac{1}{2}h(\ell)\dot{V}(Y_t)\,\mathrm{d}t, \tag{9}$$

where Y_0 *is distributed according to* π, $h(\ell)$ *is given by*

$$h(\ell) = 2\ell^2\Phi\biggl(-\frac{\ell}{2}\sqrt{I}\biggr), \tag{10}$$

and I *is defined in (6).*

Whereas V is assumed to be twice continuously differentiable, the dual representation of the Fisher information (7) allows us to remove, in the statement of the theorem, all mention of the second derivative of V, which hints that two derivatives might not really be required. For all $\theta, x \in \mathbb{R}$, define

$$\xi_\theta(x) = \sqrt{\pi(x+\theta)}. \tag{11}$$

For $p \geq 1$, denote $\|f\|^p_{\pi,p} = \int |f(x)|^p \pi(x)\,\mathrm{d}x$. Consider the following assumptions.

Assumption 1. *There exists a measurable function* $\dot{V} : \mathbb{R} \to \mathbb{R}$ *such that*

(i) *there exist* $p > 4$, $C > 0$, *and* $\beta > 1$ *such that, for all* $\theta \in \mathbb{R}$,

$$\|V(\cdot+\theta) - V(\cdot) - \theta\dot{V}(\cdot)\|_{\pi,p} \leq C|\theta|^\beta;$$

(ii) *the function* $\dot{V}$ *satisfies* $\|\dot{V}\|_{\pi,6} < +\infty$.

In Example 1 we detail how these assumptions might be checked for the Bayesian Lasso with V of the form $V : x \mapsto U(x) + \lambda|x|$, where U is twice continuously differentiable with bounded second derivative.

Lemma 1. *Assume that Assumption 1 holds. Then, the family of densities* $\theta \to \pi(\cdot+\theta)$ *is differentiable in quadratic mean at* $\theta = 0$ *with derivative* $\dot{V}$, *i.e. there exists* $C > 0$ *such that, for all* $\theta \in \mathbb{R}$,

$$\biggl(\int_{\mathbb{R}}\biggl(\xi_\theta(x) - \xi_0(x) + \frac{\theta\dot{V}(x)\xi_0(x)}{2}\biggr)^2 \mathrm{d}x\biggr)^{1/2} \leq C|\theta|^\beta,$$

where ξ_θ *is given by (11).*

Proof. The proof is postponed to Section 4.1. □

The first step in the proof is to show that both the expected squared jump distance $\mathbb{E}[(Z^d_1)^2\{1 \wedge \exp(\sum_{i=1}^d \Delta V^d_i)\}]$, and the acceptance ratio $\mathbb{P}(\mathcal{A}^d_1) = \mathbb{E}(1 \wedge \exp\{\sum_{i=1}^d \Delta V^d_i\})$ converge to a finite value.

Theorem 2. *Assume that Assumption 1 holds. Then* $\lim_{d\to+\infty}\mathbb{P}[\mathcal{A}^d_1] = a(\ell)$, *where* $a(\ell) = 2\Phi(-\sqrt{I}\ell/2)$.

Proof. The proof is postponed to Section 4.2. □

The next step consists in proving that the sequence $\{(Y^d_{t,1})_{t\geq0},\ d \in \mathbb{N}^\star\}$ defined by (8) converges weakly to a Langevin diffusion. Denote by $(\mu_d)_{d\geq1}$ the sequence of distributions of $\{(Y^d_{t,1})_{t\geq0},\ d \in \mathbb{N}^\star\}$. Following the proof of [9], it is shown in Lemma 6 (see Section 4.3)

https://doi.org/10.1017/jpr.2017.61 Published online by Cambridge University Press

that this sequence is tight in the Wiener space $\boldsymbol{W}$. By the Prohorov theorem, the tightness of $(\mu_d)_{d\geq 1}$ implies that this sequence has a weak limit point.

The equivalence between the weak formulation of stochastic differential equations and martingale problems is used to prove that any limit point is the law of a solution to (9). The generator L of the Langevin diffusion (9) is given by, for all $\phi \in C_c^2(\mathbb{R}, \mathbb{R})$,

$$L\phi(x) = \frac{h(\ell)}{2}(-\dot{V}(x)\dot{\phi}(x) + \ddot{\phi}(x)), \tag{12}$$

where, for $k \in \mathbb{N}$ and I an open subset of $\mathbb{R}$, $C_c^k(I, \mathbb{R})$ is the space of k-times differentiable functions with compact support, endowed with the topology of uniform convergence of all derivatives up to order k. We set $C_c^\infty(I, \mathbb{R}) = \bigcap_{k=0}^\infty C_c^k(I, \mathbb{R})$ and $\boldsymbol{W} = C(\mathbb{R}_+, \mathbb{R})$. The canonical process is denoted by $(W_t)_{t\geq 0}$ and $(\mathcal{B}_t)_{t\geq 0}$ is the associated filtration. For any probability measure μ on $\boldsymbol{W}$, the expectation with respect to μ is denoted by $\mathbb{E}^\mu$. A probability measure μ on $\boldsymbol{W}$ is said to solve the martingale problem associated with (9) if the pushforward of μ by W_0 is π and if, for all $\phi \in C_c^\infty(\mathbb{R}, \mathbb{R})$, the process

$$\bigg(\phi(W_t) - \phi(W_0) - \int_0^t L\phi(W_u)\,\mathrm{d}u\bigg)_{t\geq 0}$$

is a martingale with respect to μ and the filtration $(\mathcal{B}_t)_{t\geq 0}$, i.e. if, for all $s, t \in \mathbb{R}_+$, $s \leq t$ (almost surely μ),

$$\mathbb{E}^\mu\bigg[\phi(W_t) - \phi(W_0) - \int_0^t L\phi(W_u)\,\mathrm{d}u \;\bigg|\; \mathcal{B}_s\bigg] = \phi(W_s) - \phi(W_0) - \int_0^s L\phi(W_u)\,\mathrm{d}s.$$

Assumption 2. *The function $\dot{V}$ is continuous on $\mathbb{R}$ except on a Lebesgue-negligible set $\mathcal{D}_{\dot{V}}$ and is bounded on all compact sets of $\mathbb{R}$.*

Under Assumption 2, Proposition 2 (see Section 4.4) proves that every limit point of the sequence of probability measures $(\mu_d)_{d\geq 1}$ on $\boldsymbol{W}$ is a solution to the martingale problem associated with (9). In addition, under Assumption 2, in [16, Chapter 5, Lemma 1.9 and Theorem 20.1] it was shown that any solution to the martingale problem associated with (9) coincides with the law of a solution to the stochastic differential equation (SDE) (9), and conversely. Therefore, uniqueness in law of weak solutions to (9) implies uniqueness of the solution to the martingale problem.

Theorem 3. *Assume that Assumptions 1 and 2 hold. Assume also that (9) has a unique weak solution. Then $\{(Y_{t,1}^d)_{t\geq 0},\ d \in \mathbb{N}^*\}$ converges weakly to the solution $(Y_t)_{t\geq 0}$ of the Langevin equation defined by (9). Furthermore, $h(\ell)$ is maximized at the unique value of ℓ for which $a(\ell) = 0.234$, where a is defined in Theorem 2.*

Proof. The proof is postponed to Section 4.5. □

Remark 1. The idea of scaling random walk Metropolis algorithms by maximizing the speed measure $h(\ell)$ was discussed, for instance, in [1]–[6], [12], [14], [15], and [17]. For random walk Metropolis proposals of the form $Y_{k+1,i}^d = X_{k,i}^d + \sqrt{\ell d^{-\vartheta}}\,Z_{k+1}^d$, this choice is closely related to maximizing the expected squared jump distance (ESJD):

$$\mathrm{ESJD}^d(\ell, \vartheta) = \mathbb{E}[\|X_1^d - X_0^d\|^2], \tag{13}$$

where $X_0^d \sim \pi^d$. Note that for one-dimensional distributions, as $\mathrm{ESJD}^1(\ell, \vartheta) = 2(1 - \rho_1)\operatorname{var}[X_0^1]$ with ρ_1 the first-order autocorrelation, maximizing $\mathrm{ESJD}^1(\ell, \vartheta)$ is equivalent to

https://doi.org/10.1017/jpr.2017.61 Published online by Cambridge University Press

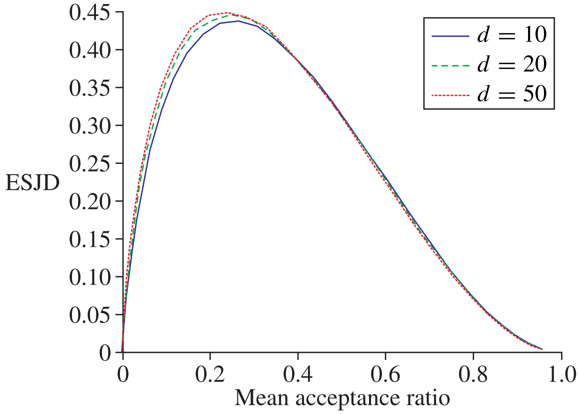

FIGURE 1: Expected squared jump distance for $V(x) = (x-1)^2 + |x|$ as a function of the mean acceptance rate for $d = 10, 20, 50$.

minimizing ρ_1. It is shown in Theorem 2 and Theorem 3 (see also Theorem 4 and Theorem 5) that

$$\lim_{d\to+\infty} \mathrm{ESJD}^d(\ell) = \mathrm{ESJD}^d(\ell, 1) = h(\ell) = \ell^2 a(\ell).$$

Therefore, in the case where $\vartheta = 1$, the optimal value of ℓ obtained in the paper is the value maximizing the limit of the ESJD^d as $d \to \infty$.

Example 1. (*Bayesian Lasso.*) We apply the results obtained above to a target density π on $\mathbb{R}$ given by $x \mapsto \mathrm{e}^{-V(x)} / \int_{\mathbb{R}} \mathrm{e}^{-V(y)}\, \mathrm{d}y$, where V is given by

$$V : x \mapsto U(x) + \lambda |x|,$$

and where $\lambda \geq 0$ and U is twice continuously differentiable with bounded second derivative. Furthermore, $\int_{\mathbb{R}} |x|^6 \mathrm{e}^{-V(x)}\, \mathrm{d}x < +\infty$. Define $\dot{V} : x \mapsto U'(x) + \lambda \operatorname{sign}(x)$, with $\operatorname{sign}(x) = -1$ if $x \leq 0$ and $\operatorname{sign}(x) = 1$ otherwise. We first check that Assumption 1(i) holds. Note that, for all $x, y \in \mathbb{R}$,

$$||x+y| - |x| - \operatorname{sign}(x)y| \leq 2|y|\, \mathbf{1}_{\mathbb{R}_+}(|y| - |x|),$$

which implies that, for any $p \geq 1$, there exists C_p such that

$$\begin{aligned}
&\|V(\cdot + \theta) - V(\cdot) - \theta \dot{V}(\cdot)\|_{\pi,p} \\
&\quad \leq \|U(\cdot + \theta) - U(\cdot) - \theta U'(\cdot)\|_{\pi,p} + \lambda \||\cdot + \theta| - |\cdot| - \theta \operatorname{sign}(\cdot)\|_{\pi,p} \\
&\quad \leq \|U''\|_\infty \theta^2 + 2|\theta| \lambda \{\pi([-\theta, \theta])\}^{1/p} \\
&\quad \leq C|\theta|^{p+1/p} \vee |\theta|^2.
\end{aligned}$$

Assumptions 1(ii) and 2 are easy to check. The uniqueness in law of (9) was established in [7, Theorem 4.5(i)]. Therefore, Theorem 3 can be applied. To numerically illustrate this result, we consider the density π associated with $U(x) = (x-1)^2$ and $\lambda = 1$. In Figure 1 we present an empirical estimation of the ESJD^d defined by (13) for dimensions $d = 10, 20, 50$ as a function of the empirical mean acceptance rate. We can observe that, as expected (see Remark 1), the ESJD^d converges to some limit function as d goes ∞, and this function has a maximum for a mean acceptance probability around 0.23.

3. Target density supported on an interval of $\mathbb{R}$

We now extend our results to densities supported by a open interval $\mathcal{I} \subset \mathbb{R}$:

$$\pi(x) \propto \exp(-V(x))\,\mathbf{1}_{\mathcal{I}}(x),$$

where $V\colon \mathcal{I} \to \mathbb{R}$ is a measurable function. Note that, by convention, $V(x) = -\infty$ for all $x \notin \mathcal{I}$. Denote by $\overline{\mathcal{I}}$ the closure of $\mathcal{I}$ in $\mathbb{R}$. The results of Section 2 cannot be directly used in such a case as π is no longer positive on $\mathbb{R}$. Consider the following assumption.

Assumption 3. *There exists a measurable function $\dot{V}\colon \mathcal{I} \to \mathbb{R}$ and $\mathrm{r} > 1$ such that*

(i) *there exist $p > 4$, $C > 0$, and $\beta > 1$ such that, for all $\theta \in \mathbb{R}$,*

$$\|\{V(\cdot+\theta) - V(\cdot)\}\,\mathbf{1}_{\mathcal{I}}(\cdot+\mathrm{r}\theta)\,\mathbf{1}_{\mathcal{I}}(\cdot+(1-\mathrm{r})\theta) - \theta\dot{V}(\cdot)\|_{\pi,p} \leq C|\theta|^{\beta},$$

with the convention $0 \times \infty = 0$;

(ii) *the function $\dot{V}$ satisfies $\|\dot{V}\|_{\pi,6} < +\infty$;*

(iii) *there exist $\gamma \geq 6$ and $C > 0$ such that, for all $\theta \in \mathbb{R}$,*

$$\int_{\mathbb{R}} \mathbf{1}_{\mathcal{I}^c}(x+\theta)\pi(x)\,\mathrm{d}x \leq C|\theta|^{\gamma}.$$

As an important consequence of Assumption 3(iii), if X is distributed according to π and is independent of the standard random variable Z, there exists a constant C such that

$$\mathbb{P}(X + \ell d^{-1/2} Z \in \mathcal{I}^c) \leq C d^{-\gamma/2}.$$

Theorem 4. *Assume that Assumption 3 holds. Then $\lim_{d\to+\infty} \mathbb{P}[\mathcal{A}_1^d] = a(\ell)$, where $a(\ell) = 2\Phi(-\sqrt{I}\ell/2)$.*

Proof. The proof is similar to the proof of Theorem 2 and can be found in the supplementary material; see [8]. □

We now establish the weak convergence of the sequence $\{(Y_{t,1}^d)_{t\geq 0},\ d \in \mathbb{N}^\star\}$, following the same steps as for the proof of Theorem 3. For all $d \geq 1$, let μ_d be the law of the process $(Y_{t,1}^d)_{t\geq 0}$. Under Assumption 3, Lemma S5 of the supplementary material [8] establishes that the sequence $(\mu_d)_{d\geq 1}$ is tight in $\boldsymbol{W}$ (the proof is similar to the proof of Lemma 6).

Contrary to the case where π is positive on $\mathbb{R}$, we do not assume that $\dot{V}$ is bounded on all compact sets of $\mathbb{R}$. Therefore, we consider the local martingale problem associated with (9): with the notation of Section 2, a probability measure μ on $\boldsymbol{W}$ is said to solve the local martingale problem associated with (9) if the pushforward of μ by W_0 is π and if, for all $\psi \in C^{\infty}(\mathbb{R}, \mathbb{R})$, the process

$$\left(\psi(W_t) - \psi(W_0) - \int_0^t L\psi(W_u)\,\mathrm{d}u\right)_{t\geq 0}, \tag{14}$$

where L is given by (12), is a local martingale with respect to μ and the filtration $(\mathcal{B}_t)_{t\geq 0}$. By [7, Theorem 1.27], any solution to the local martingale problem defined by (14) coincides with the law of a solution to SDE (9) and conversely. If (9) admits a unique solution in law, this law is the unique solution to the local martingale problem defined by (14). We consider the following assumption on $\dot{V}$.

Assumption 4. *The function $\dot V$ is continuous on $\mathcal{I}$ except on a null-set $\mathcal{D}_{\dot V}$, with respect to the Lebesgue measure, and is bounded on all compact sets of $\mathcal{I}$.*

Note that this condition does not preclude that $\dot V$ is unbounded at the boundary of $\mathcal{I}$. In Proposition 4 (see Section 5.2), we prove that any limit point μ of $(\mu_d)_{d\geq 1}$ is a solution to the local martingale problem defined by (14). The key step of the proof is Lemma 2.

Lemma 2. *Assume that Assumptions 3 and 4 hold. Let μ be a limit point of the sequence $(\mu_d)_{d\geq 1}$. If, for all $\phi \in C_c^\infty(\mathcal{I}, \mathbb{R})$, the process $(\phi(W_t) - \phi(W_0) - \int_0^t L\phi(W_u)\,\mathrm{d}u)_{t\geq 0}$ is a martingale with respect to μ and the filtration $(\mathcal{B}_t)_{t\geq 0}$, then μ solves the local martingale problem associated with (9).*

Proof. The proof is postponed to Section 5.1. □

Theorem 5. *Assume that Assumptions 3 and 4 hold. Assume also that (9) has a unique weak solution. Then $\{(Y_{t,1}^d)_{t\geq 0},\ d \in \mathbb{N}^*\}$ converges weakly to the solution $(Y_t)_{t\geq 0}$ of the Langevin equation defined by (9). Furthermore, $h(\ell)$ is maximized at the unique value of ℓ for which $a(\ell) = 0.234$, where a is defined in Theorem 2.*

Proof. The proof follows the same lines as the proof of Theorem 3 and can be found in the supplementary material [8]. □

The conditions for uniqueness in law of singular one-dimensional SDEs are given in [7]. These conditions are rather involved and difficult to summarize in full generality. Rather, we illustrate Theorem 5 by two examples.

Example 2. (*Application to the gamma distribution.*) Define the class of the generalized gamma distributions as the family of densities on $\mathbb{R}$ given by

$$\pi_\gamma : x \mapsto \frac{x^{\mathrm{a}_1 - 1}\exp(-x^{\mathrm{a}_2})\,\mathbf{1}_{\mathbb{R}_+^\star}(x)}{\int_{\mathbb{R}_+^\star} y^{\mathrm{a}_1-1}\exp(-y^{\mathrm{a}_2})\,\mathrm{d}y},$$

with two parameters $\mathrm{a}_1 > 6$ and $\mathrm{a}_2 > 0$. Note that in this case, $\mathcal{I} = \mathbb{R}_+^\star$ for all $x \in \mathcal{I}$, $V_\gamma : x \mapsto x^{\mathrm{a}_2} - (\mathrm{a}_1 - 1)\log x$, and $\dot V_\gamma : x \mapsto \mathrm{a}_2 x^{\mathrm{a}_2 - 1} - (\mathrm{a}_1 - 1)/x$. We check that Assumption 3 holds with $\mathrm{r} = \frac{3}{2}$. First, we show that Assumption 3(i) holds with $p = 5$. Write, for all $\theta \in \mathbb{R}$ and $x \in \mathcal{I}$,

$$\{V_\gamma(x+\theta) - V_\gamma(x)\}\,\mathbf{1}_{\mathcal{I}}(x + (1-r)\theta)\,\mathbf{1}_{\mathcal{I}}(x + r\theta) - \theta \dot V_\gamma(x) = \mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3,$$

where

$$\begin{aligned}
\mathcal{E}_1 &= \theta \dot V_\gamma(x)\left\{\mathbf{1}_{\mathcal{I}}\left(x - \frac{\theta}{2}\right)\mathbf{1}_{\mathcal{I}}\left(x + \frac{3\theta}{2}\right) - 1\right\},\\
\mathcal{E}_2 &= (1-\mathrm{a}_1)\left\{\log\left(1 + \frac{\theta}{x}\right) - \frac{\theta}{x}\right\}\mathbf{1}_{\mathcal{I}}\left(x - \frac{\theta}{2}\right)\mathbf{1}_{\mathcal{I}}\left(x + \frac{3\theta}{2}\right),\\
\mathcal{E}_3 &= ((x+\theta)^{\mathrm{a}_2} - x^{\mathrm{a}_2} - \mathrm{a}_2\theta x^{\mathrm{a}_2 - 1})\,\mathbf{1}_{\mathcal{I}}\left(x - \frac{\theta}{2}\right)\mathbf{1}_{\mathcal{I}}\left(x + \frac{3\theta}{2}\right).
\end{aligned}$$

It is enough to prove that there exists $q > 5$ such that, for all $i \in \{1, 2, 3\}$, $\int_{\mathcal{I}} |\mathcal{E}_i|^5 \pi_\gamma(x)\,\mathrm{d}x \leq C|\theta|^q$. It follows from tedious but straightforward calculations (see the supplementary mate-

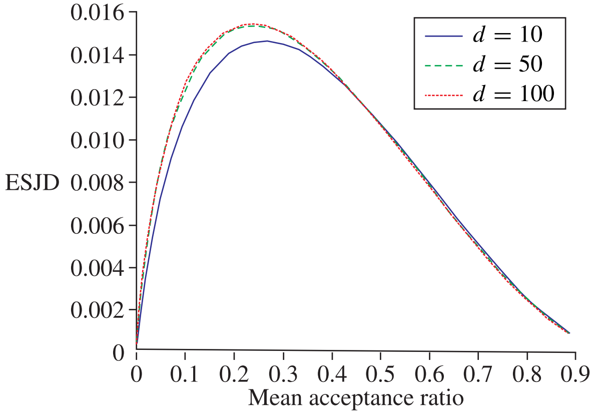

FIGURE 2: The ESJD^d for the beta distribution with parameters $a_1 = 10$ and $a_2 = 10$ as a function of the mean acceptance rate for $d = 10, 50, 100$.

rial [8] for detailed computations) that

$$\int_{\mathbb{R}_+^*} |\mathcal{E}_1(x)|^5 \pi_\gamma(x)\,\mathrm{d}x \le C(|\theta|^{\mathrm{a}_1} + |\theta|^{5\mathrm{a}_2+\mathrm{a}_1}),$$
$$\int_{\mathbb{R}_+^*} |\mathcal{E}_2(x)|^5 \pi_\gamma(x)\,\mathrm{d}x \le C(|\theta|^{\mathrm{a}_1} + |\theta|^{10}),$$
$$\int_{\mathbb{R}_+^*} |\mathcal{E}_3(x)|^5 \pi_\gamma(x)\,\mathrm{d}x \le C(|\theta|^{5\mathrm{a}_2+\mathrm{a}_1} + |\theta|^{10}).$$

The proof of Assumption 3(ii) follows from

$$\int_{\mathbb{R}_+^*} |\dot{V}_\gamma(x)|^6 \pi_\gamma(x)\,\mathrm{d}x \le C\bigg(\int_{\mathbb{R}_+^*} x^{\mathrm{a}_1-1+6(\mathrm{a}_2-1)} \mathrm{e}^{-x^{\mathrm{a}_2}}\,\mathrm{d}x + \int_{\mathbb{R}_+^*} x^{\mathrm{a}_1-7} \mathrm{e}^{-x^{\mathrm{a}_2}}\,\mathrm{d}x\bigg) < \infty,$$

and Assumption 3(iii) follows from $\int_{\mathbb{R}} \mathbf{1}_{\mathcal{I}^c}(x+\theta)\pi_\gamma(x)\,\mathrm{d}x \le C|\theta|^{\mathrm{a}_1}$. Now consider the Langevin equation associated with π_γ given by $\mathrm{d}Y_t = -\dot{V}_\gamma(Y_t)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}B_t$ with initial distribution π_γ. This SDE has 0 as a singular point, which has right type 3 according to the terminology of [7]. On the other hand, ∞ has type A and the existence and uniqueness in law for the SDE follows from [7, Theorem 4.6(viii)]. Since Assumption 4 is straightforward, Theorem 5 can be applied.

Example 3. (*Application to the beta distribution.*) Consider now the case of the beta distributions π_β with density $x \mapsto x^{a_1-1}(1-x)^{a_2-1}\mathbf{1}_{(0,1)}(x)$ with $a_1, a_2 > 6$. Here $\mathcal{I} = (0, 1)$ and the log-density V_β and its derivative on $\mathcal{I}$ are defined by $V_\beta(x) = -(a_1-1)\log x - (a_2-1)\log(1-x)$ and $\dot{V}_\beta(x) = -(a_1-1)/x - (a_2-1)/(1-x)$. Along the same lines as above, π_β satisfies Assumptions 3 and 4. Hence, Theorem 4 can be applied if we establish the uniqueness in law for the Langevin equation associated with π_β defined by $\mathrm{d}Y_t = -\dot{V}_\beta(Y_t)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}B_t$ with initial distribution π_β. In the terminology of [7], 0 has right type 3 and 1 has left type 3. Therefore, by [7, Theorem 2.16(i) and 2.16(ii)], the SDE has a global unique weak solution. To illustrate our findings, we consider the beta distribution with parameters $a_1 = 10$ and $a_2 = 10$. In Figure 2 we present an empirical estimation of the ESJD^d defined by (13) for dimensions $d = 10, 50, 100$

as a function of the empirical mean acceptance rate. We can observe that, as expected (see Remark 1), the ESJD^d converges to some limit function as d goes to ∞, and this function has a maximum for a mean acceptance probability around 0.23.

4. Proofs of Section 2

We first prove differentiability in quadratic mean (Lemma 1) and then the convergence of the acceptance ratio (Theorem 2) in Sections 4.1 and 4.2. Then the proof of Theorem 3 is detailed in the remainder of the section in three main steps. First, Section 4.3 is devoted to the proof of Lemma 6 which establishes the tightness of the sequence $(\mu_d)_{d\geq 1}$ in $\boldsymbol{W}$ which ensures that this sequence has a weak limit point. Then it is proved in Section 4.4 (Proposition 2) that if (29) holds, every limit point of the sequence $(\mu_d)_{d\geq 1}$ on $\boldsymbol{W}$ is a solution to the martingale problem associated with (9) and, therefore, is the law of a solution to SDE (9). The fact that (29) holds is a consequence of Proposition 3.

For any real random variable Y and any $p \geq 1$, let $\|Y\|_p := \mathbb{E}[|Y|^p]^{1/p}$.

4.1. Proof of differentiability in quadratic mean (Lemma 1)

Let $\Delta_\theta V(x) = V(x) - V(x+\theta)$. By definition of ξ_θ and π,

$$\bigg(\xi_\theta(x) - \xi_0(x) + \frac{\theta \dot{V}(x)\xi_0(x)}{2}\bigg)^2 \leq 2\{A_\theta(x) + B_\theta(x)\}\pi(x),$$

where

$$A_\theta(x) = \bigg(\exp\bigg(\frac{\Delta_\theta V(x)}{2}\bigg) - 1 - \frac{\Delta_\theta V(x)}{2}\bigg)^2, \qquad B_\theta(x) = \frac{(\Delta_\theta V(x) + \theta \dot{V}(x))^2}{4}.$$

By Assumption 1(i), $\|B_\theta\|_{\pi,p} \leq C|\theta|^\beta$. For A_θ, note that, for all $x \in \mathbb{R}$, $(\exp(x) - 1 - x)^2 \leq 2x^4(\exp(2x) + 1)$. Then

$$\begin{aligned}\int_{\mathbb{R}} A_\theta(x)\pi(x)\,\mathrm{d}x &\leq C \int_{\mathbb{R}} \Delta_\theta V(x)^4 (1 + \mathrm{e}^{\Delta_\theta V(x)})\pi(x)\,\mathrm{d}x \\ &\leq C \int_{\mathbb{R}} (\Delta_\theta V(x)^4 + \Delta_{-\theta} V(x)^4)\pi(x)\,\mathrm{d}x.\end{aligned}$$

The proof is completed by writing (the same inequality holds for $\Delta_{-\theta} V$):

$$\int_{\mathbb{R}} \Delta_\theta V(x)^4 \pi(x)\,\mathrm{d}x \leq C\bigg[\int_{\mathbb{R}} (\Delta_\theta V(x) - \theta \dot{V}(x))^4 \pi(x)\,\mathrm{d}x + \theta^4 \int_{\mathbb{R}} \dot{V}^4(x)\pi(x)\,\mathrm{d}x\bigg]$$

and using Assumption 1(i) and 1(ii).

4.2. Proof of asymptotic acceptance rate (Theorem 2)

Define

$$\mathbb{E}^d(q) = \mathbb{E}\bigg[(Z_1^d)^q \bigg| 1 \wedge \exp\bigg(\sum_{i=1}^d \Delta V_i^d\bigg) - 1 \wedge \exp(\upsilon^d)\bigg|\bigg],$$

https://doi.org/10.1017/jpr.2017.61 Published online by Cambridge University Press

where ΔV_i^d is given by (5),

$$\upsilon^d = -\ell d^{-1/2} Z_1^d \dot{V}(X_1^d) + \sum_{i=2}^{d} b^d(X_i^d, Z_i^d), \tag{15}$$

$$b^d(x,z) = -\frac{\ell z}{\sqrt{d}}\dot{V}(x) + \mathbb{E}[2\zeta^d(X_1^d, Z_1^d)] - \frac{\ell^2}{4d}\dot{V}^2(x), b^d(x,z) \tag{16}$$

$$\zeta^d(x,z) = \exp\left\{\frac{V(x) - V(x + \ell d^{-1/2} z)}{2}\right\} - 1. \tag{17}$$

The key result to prove Theorem 2 is stated in Proposition 1 which shows that it is enough to consider υ^d to analyse the asymptotic behaviour of the acceptance ratio and the expected squared jump distance as $d \to +\infty$.

Proposition 1. *Assume that Assumption 1 holds. Let X^d be a random variable distributed according to π^d and Z^d be a zero-mean standard Gaussian random variable, independent of X^d. Then, for any $q \geq 0$, $\lim_{d\to+\infty} \mathbb{E}^d(q) = 0$.*

Proof. By the central limit theorem, the term $-\ell \sum_{i=2}^{d} (Z_i^d/\sqrt{d})\dot{V}(X_i^d)$ in (15) converges in distribution to a zero-mean Gaussian random variable with variance $\ell^2 I$, where I is defined in (6). By Lemma 5 (stated and proved below), the second term, which is $d\mathbb{E}[2\zeta^d(X_1^d, Z_1^d)] = -d\mathbb{E}[(\zeta^d(X_1^d, Z_1^d))^2]$, converges to $-\ell^2 I/4$. The last term converges in probability to $-\ell^2 I/4$. Therefore, the two last terms play a similar role in the expansion of the acceptance ratio as the second derivative of V in the regular case. We now present the detailed arguments.

Let $q > 0$ and $\Lambda^d = -\ell d^{-1/2} Z_1^d \dot{V}(X_1^d) + \sum_{i=2}^{d} \Delta V_i^d$. By the triangle inequality, $\mathbb{E}^d(q) \leq \mathbb{E}_1^d(q) + \mathbb{E}_2^d(q)$, where

$$\mathbb{E}_1^d(q) = \mathbb{E}\left[(Z_1^d)^q \left| 1 \wedge \exp\left\{\sum_{i=1}^{d} \Delta V_i^d\right\} - 1 \wedge \exp\{\Lambda^d\}\right|\right],$$

$$\mathbb{E}_2^d(q) = \mathbb{E}[(Z_1^d)^q |1 \wedge \exp\{\Lambda^d\} - 1 \wedge \exp\{\upsilon^d\}|].$$

Since $t \mapsto 1 \wedge \mathrm{e}^t$ is 1-Lipschitz, by the Cauchy–Schwarz inequality, we obtain

$$\mathbb{E}_1^d(q) \leq \|Z_1^d\|_{2q}^q \|\Delta V_1^d + \ell d^{-1/2} Z_1^d \dot{V}(X_1^d)\|_2.$$

By Lemma 3(ii) (stated and proved below), $\mathbb{E}_1^d(q)$ goes to 0 as d goes to $+\infty$. Consider now $\mathbb{E}_2^d(q)$. Using again the fact that $t \mapsto 1 \wedge \mathrm{e}^t$ is 1-Lipschitz and Lemma 4, $\mathbb{E}_2^d(q)$ goes to 0. □

We have now all the necessary ingredients to establish Theorem 2.

Proof of Theorem 2. By the definition of $\mathcal{A}_1^d$, see (3),

$$\mathbb{P}[\mathcal{A}_1^d] = \mathbb{E}\left[1 \wedge \exp\left\{\sum_{i=1}^{d} \Delta V_i^d\right\}\right],$$

where $\Delta V_i^d = V(X_{0,i}^d) - V(X_{0,i}^d + \ell d^{-1/2} Z_{1,i}^d)$ and where X_0^d is distributed according to π^d and independent of the standard d-dimensional Gaussian random variable Z_1^d. Following the same steps as in the proof of Proposition 1 yields

$$\lim_{d\to+\infty} |\mathbb{P}[\mathcal{A}_1^d] - \mathbb{E}[1 \wedge \exp\{\Theta^d\}]| = 0, \tag{18}$$

https://doi.org/10.1017/jpr.2017.61 Published online by Cambridge University Press

where

$$\Theta^d = -\ell d^{-1/2}\sum_{i=1}^{d} Z_{1,i}^d \dot{V}(X_{0,i}^d) - \ell^2\sum_{i=2}^{d}\frac{\dot{V}(X_{0,i}^d)^2}{4d} + 2(d-1)\mathbb{E}[\zeta^d(X_{0,1}^d, Z_{1,1}^d)].$$

Conditional on X_0^d, Θ^d is a one-dimensional Gaussian random variable with mean μ_d and variance σ_d^2 defined by

$$\mu_d = -\ell^2\sum_{i=2}^{d}\frac{\dot{V}(X_{0,i}^d)^2}{4d} + 2(d-1)\mathbb{E}[\zeta^d(X_{0,1}^d, Z_{1,1}^d)], \qquad \sigma_d^2 = \ell^2 d^{-1}\sum_{i=1}^{d}\dot{V}(X_{0,i}^d)^2.$$

Therefore, since, for any $G \sim \mathcal{N}(\mu, \sigma^2)$, $\mathbb{E}[1 \wedge \exp(G)] = \Phi(\mu/\sigma) + \exp(\mu + \sigma^2/2)\Phi(-\sigma - \mu/\sigma)$, taking the expectation conditional on X_0^d, we have

$$\mathbb{E}[1 \wedge \exp\{\Theta^d\}] = \mathbb{E}\bigg[\Phi\bigg(\frac{\mu_d}{\sigma_d}\bigg) + \exp\bigg(\mu_d + \frac{\sigma_d^2}{2}\bigg)\Phi\bigg(-\sigma_d - \frac{\mu_d}{\sigma_d}\bigg)\bigg] = \mathbb{E}[\Gamma(\sigma_d^2, -2\mu_d)],$$

where the function Γ is defined in (23). By Lemma 5 and the law of large numbers, almost surely, $\lim_{d\to+\infty}\mu_d = -\ell^2 I/2$ and $\lim_{d\to+\infty}\sigma_d^2 = \ell^2 I$. Thus, as Γ is bounded, by Lebesgue's dominated convergence theorem,

$$\lim_{d\to+\infty}\mathbb{E}[1 \wedge \exp\{\Theta^d\}] = 2\Phi\bigg(-\frac{\ell\sqrt{I}}{2}\bigg).$$

The proof is then completed by (18). □

We conclude this section by establishing the technical lemmas which are used in the proofs above. Define

$$R(x) = \int_0^x \frac{(x-u)^2}{(1+u)^3}\,\mathrm{d}u, \tag{19}$$

where R is the remainder term of the Taylor expansion of $x \mapsto \log(1+x)$:

$$\log(1+x) = x - \frac{x^2}{2} + R(x). \tag{20}$$

Lemma 3. *Assume that Assumption 1 holds. Then, if X is a random variable distributed according to π and Z is a standard Gaussian random variable independent of X,*

(i) $\lim_{d\to+\infty} d\|\zeta^d(X, Z) + \ell Z\dot{V}(X)/(2\sqrt{d})\|_2^2 = 0$;

(ii) $\lim_{d\to+\infty}\sqrt{d}\|V(X) - V(X + \ell Z/\sqrt{d}) + \ell Z\dot{V}(X)/\sqrt{d}\|_p = 0$;

(iii) $\lim_{d\to\infty} d\|R(\zeta^d(X, Z))\|_1 = 0$,

where ζ^d is given by (17).

Proof. Using the definitions (11) and (17) of ζ^d and ξ_θ,

$$\zeta^d(x, z) = \frac{\xi_{\ell z d^{-1/2}}(x)}{\xi_0(x)} - 1. \tag{21}$$

https://doi.org/10.1017/jpr.2017.61 Published online by Cambridge University Press

(i) The proof follows from Lemma 1 using the fact that $\beta > 1$:

$$\left\| \zeta^d(X, Z) + \frac{\ell Z \dot{V}(X)}{2\sqrt{d}} \right\|_2^2 \le C\ell^{2\beta} d^{-\beta} \mathbb{E}[|Z|^{2\beta}].$$

(ii) Using Assumption 1(i), we obtain

$$\left\| V(X) - V\left(X + \frac{\ell Z}{\sqrt{d}}\right) + \frac{\ell Z \dot{V}(X)}{\sqrt{d}} \right\|_p^p \le C\ell^{\beta p} d^{-\beta p/2} \mathbb{E}[|Z|^{\beta p}]$$

and the proof follows since $\beta > 1$.

(iii) Note that, for all $x > 0$, $u \in [0, x]$, $|(x-u)(1+u)^{-1}| \le |x|$, and the same inequality holds for $x \in (-1, 0]$ and $u \in [x, 0]$. Then, by (19) and (20), for all $x > -1$, $|R(x)| \le x^2 |\log(1+x)|$.
Then, by (21), setting $\Psi_d(x, z) = R(\zeta^d(x, z))$:

$$|\Psi_d(x, z)| \le \frac{(\xi_{\ell z d^{-1/2}}(x)/\xi_0(x) - 1)^2 |V(x + \ell z d^{-1/2}) - V(x)|}{2}.$$

Since, for all $x \in \mathbb{R}$, $|\exp(x) - 1| \le |x|(\exp(x) + 1)$, this yields

$$|\Psi_d(x, z)| \le 4^{-1} |V(x + \ell z d^{-1/2}) - V(x)|^3 (\exp(V(x) - V(x + \ell z d^{-1/2})) + 1),$$

which implies that

$$\int_{\mathbb{R}} |\Psi_d(x, z)| \pi(x)\, \mathrm{d}x \le 4^{-1} \int_{\mathbb{R}} |V(x + \ell z d^{-1/2}) - V(x)|^3 \{\pi(x) + \pi(x + \ell z d^{-1/2})\}\, \mathrm{d}x.$$

By Hölder's inequality and using Assumption 1(i),

$$\int_{\mathbb{R}} |\Psi_d(x, z)| \pi(x)\, \mathrm{d}x \le C\left(|\ell z d^{-1/2}|^3 \left(\int_{\mathbb{R}} |\dot{V}(x)|^4 \pi(x)\, \mathrm{d}x\right)^{3/4} + |\ell z d^{-1/2}|^{3\beta}\right).$$

The proof follows from Assumption 1(ii) since $\beta > 1$. □

For all $d \ge 1$, let X^d be distributed according to π^d, and Z^d be a d-dimensional Gaussian random variable independent of X^d, set

$$J^d = \left\| \sum_{i=2}^{d} \{\Delta V_i^d - b^d(X_i^d, Z_i^d)\} \right\|_1,$$

where ΔV_i^d and b^d are defined in (5) and (16), respectively.

Lemma 4. *We have* $\lim_{d\to+\infty} J^d = 0$.

Proof. Define

$$J_1^d = \left\| \sum_{i=2}^{d} 2\zeta^d(X_i^d, Z_i^d) + \frac{\ell Z_i^d}{\sqrt{d}} \dot{V}(X_i^d) - \mathbb{E}[2\zeta^d(X_i^d, Z_i^d)] \right\|_1,$$

$$J_2^d = \left\| \sum_{i=2}^{d} \zeta^d(X_i^d, Z_i^d)^2 - \frac{\ell^2}{4d} \dot{V}^2(X_i^d) \right\|_1, \qquad J_3^d = 2\left\| \sum_{i=2}^{d} R(\zeta^d(X_i^d, Z_i^d)) \right\|_1,$$

https://doi.org/10.1017/jpr.2017.61 Published online by Cambridge University Press

where R is defined by (19). As $\Delta V_i^d = 2\log(1+\zeta^d(X_i^d, Z_i^d))$ and using (20), we obtain

$$J^d \le J_1^d + J_2^d + J_3^d.$$

By the Cauchy–Schwarz inequality and as the $(X_i^d, Z_i^d)_{2\le i\le d}$ are independent,

$$J_1^d \le \operatorname{var}\bigg[\sum_{i=2}^{d} 2\zeta^d(X_i^d, Z_i^d) + \frac{\ell Z_i^d}{\sqrt{d}}\dot V(X_i^d)\bigg]^{1/2} \le \sqrt{d}\bigg\|2\zeta^d(X_1^d, Z_1^d) + \frac{\ell Z_1^d}{\sqrt{d}}\dot V(X_1^d)\bigg\|_2.$$

By Lemma 3(i), this term goes to 0 as d goes to $+\infty$. Consider now J_2^d. We use the following decomposition for all $2 \le i \le d$:

$$\begin{aligned}\zeta^d(X_i^d, Z_i^d)^2 - \frac{\ell^2}{4d}\dot V^2(X_i^d) &= \bigg(\zeta^d(X_i^d, Z_i^d) + \frac{\ell}{2\sqrt{d}}Z_i^d\dot V(X_i^d)\bigg)^2 \\ &\quad - \frac{\ell}{\sqrt{d}}Z_i^d\dot V(X_i^d)\bigg(\zeta^d(X_i^d, Z_i^d) + \frac{\ell}{2\sqrt{d}}Z_i^d\dot V(X_i^d)\bigg) \\ &\quad + \frac{\ell^2}{4d}\{(Z_i^d)^2 - 1\}\dot V^2(X_i^d).\end{aligned}$$

Then

$$\begin{aligned}J_2^d &\le d\bigg\|\zeta^d(X_1^d, Z_1^d) + \frac{\ell}{2\sqrt{d}}Z_1^d\dot V(X_1^d)\bigg\|_2^2 + \frac{\ell^2}{4d}\bigg\|\sum_{i=2}^{d}\dot V^2(X_i^d)\{(Z_i^d)^2-1\}\bigg\|_1 \\ &\quad + \ell\sqrt{d}\bigg\|\dot V(X_1^d)Z_1^d\Big(\zeta^d(X_1^d, Z_1^d) + \frac{\ell}{2\sqrt{d}}Z_1^d\dot V(X_1^d)\Big)\bigg\|_1.\end{aligned}$$

Using Assumption 1(ii), Lemma 3(i), and the Cauchy–Schwarz inequality, we see that the first and the last term converge to 0. For the second term, note that $\mathbb{E}[(Z_i^d)^2 - 1] = 0$ so that

$$d^{-1}\bigg\|\sum_{i=2}^{d}\dot V^2(X_i^d)\{(Z_i^d)^2-1\}\bigg\|_1 \le d^{-1/2}\operatorname{var}[\dot V^2(X_1^d)\{(Z_1^d)^2-1\}]^{1/2} \to 0.$$

Finally, $\lim_{d\to\infty} J_3^d = 0$ by (20) and Lemma 3(iii). □

Following [9], we introduce the function $\mathcal{G}$ defined on $\overline{\mathbb{R}}_+ \times \mathbb{R}$ by

$$\mathcal{G}(a,b) = \begin{cases} \exp\Big(\dfrac{a-b}{2}\Big)\Phi\Big(\dfrac{b}{2\sqrt{a}} - \sqrt{a}\Big) & \text{if } a \in (0,+\infty), \\ 0 & \text{if } a = +\infty, \\ \exp\Big(-\dfrac{b}{2}\Big)\mathbf{1}_{\{b>0\}} & \text{if } a = 0, \end{cases} \tag{22}$$

where Φ is the cumulative distribution function of a standard normal variable, and Γ is defined by

$$\Gamma(a,b) = \begin{cases} \Phi\Big(-\dfrac{b}{2\sqrt{a}}\Big) + \exp\Big(\dfrac{a-b}{2}\Big)\Phi\Big(\dfrac{b}{2\sqrt{a}} - \sqrt{a}\Big) & \text{if } a \in (0,+\infty), \\ \dfrac{1}{2} & \text{if } a = +\infty, \\ \exp\Big(-\dfrac{b_+}{2}\Big) & \text{if } a = 0. \end{cases} \tag{23}$$

Note that $\mathcal{G}$ and Γ are bounded on $\overline{\mathbb{R}}_+ \times \mathbb{R}$. We use $\mathcal{G}$ and Γ throughout Section 4.

https://doi.org/10.1017/jpr.2017.61 Published online by Cambridge University Press

Lemma 5. *Assume that Assumption 1 holds. For all $d \in \mathbb{N}^*$, let X^d be a random variable distributed according to π^d and Z^d be a standard Gaussian random variable in $\mathbb{R}^d$, independent of X. Then*

$$\lim_{d\to+\infty} d\mathbb{E}[2\zeta^d(X_1^d, Z_1^d)] = -\frac{\ell^2}{4}I,$$

where I is defined in (6) and ζ^d in (17).

Proof. By (17),

$$\begin{aligned} d\mathbb{E}[2\zeta^d(X_1^d, Z_1^d)] &= 2d\mathbb{E}\bigg[\int_{\mathbb{R}} \sqrt{\pi(x+\ell d^{-1/2}Z_1^d)}\sqrt{\pi(x)}\,\mathrm{d}x - 1\bigg] \\ &= -d\mathbb{E}\bigg[\int_{\mathbb{R}} \Big(\sqrt{\pi(x+\ell d^{-1/2}Z_1^d)} - \sqrt{\pi(x)}\Big)^2 \mathrm{d}x\bigg] \\ &= -d\mathbb{E}[\{\zeta^d(X_1^d, Z_1^d)\}^2]. \end{aligned}$$

The proof is then completed by Lemma 3(i). □

4.3. Proof of tightness (Lemma 6)

Lemma 6. *Assume that Assumption 1 holds. Then the sequence $(\mu_d)_{d\geq 1}$ is tight in $\mathbf{W}$.*

Proof. The proof is adapted from [9]. By Kolmogorov's criterion, it is enough to prove that there exists a nondecreasing function $\gamma : \mathbb{R}_+ \to \mathbb{R}_+$ such that, for all $d \geq 1$ and all $0 \leq s \leq t$,

$$\mathbb{E}[(Y_{t,1}^d - Y_{s,1}^d)^4] \leq \gamma(t)(t-s)^2.$$

The inequality is straightforward for all $0 \leq s \leq t$ such that $\lfloor ds \rfloor = \lfloor dt \rfloor$. For all $0 \leq s \leq t$ such that $\lceil ds \rceil \leq \lfloor dt \rfloor$,

$$Y_{t,1}^d - Y_{s,1}^d = X_{\lfloor dt \rfloor,1}^d - X_{\lceil ds \rceil,1}^d + \frac{dt - \lfloor dt \rfloor}{\sqrt{d}}\ell Z_{\lceil dt \rceil,1}^d \mathbf{1}_{\mathcal{A}_{\lceil dt \rceil}^d} + \frac{\lceil ds \rceil - ds}{\sqrt{d}}\ell Z_{\lceil ds \rceil,1}^d \mathbf{1}_{\mathcal{A}_{\lceil ds \rceil}^d}.$$

Then, by the Jensen inequality,

$$\mathbb{E}[(Y_{t,1}^d - Y_{s,1}^d)^4] \leq C((t-s)^2 + \mathbb{E}[(X_{\lfloor dt \rfloor,1}^d - X_{\lceil ds \rceil,1}^d)^4]),$$

where we have used

$$\frac{(dt - \lfloor dt \rfloor)^2}{d^2} + \frac{(\lceil ds \rceil - ds)^2}{d^2} \leq \frac{(dt-ds)^2 + (\lceil ds \rceil - \lfloor dt \rfloor)^2}{d^2} \leq 2(t-s)^2.$$

The proof is completed using Lemma 7. □

Lemma 7. *Assume that Assumption 1 holds. Then, there exists $C > 0$ such that, for all $0 \leq k_1 < k_2$,*

$$\mathbb{E}[(X_{k_2,1}^d - X_{k_1,1}^d)^4] \leq C\sum_{p=2}^{4} \frac{(k_2-k_1)^p}{d^p}.$$

Proof. For all $0 \leq k_1 < k_2$,

$$\mathbb{E}[(X_{k_2,1}^d - X_{k_1,1}^d)^4] = \frac{\ell^4}{d^2}\mathbb{E}\bigg[\bigg(\sum_{k=k_1+1}^{k_2} Z_{k,1}^d - \sum_{k=k_1+1}^{k_2} Z_{k,1}^d \mathbf{1}_{(\mathcal{A}_k^d)^c}\bigg)^4\bigg].$$

Therefore, by the Hölder inequality,

$$\mathbb{E}[(X^d_{k_2,1} - X^d_{k_1,1})^4] \leq \frac{24\ell^4}{d^2}(k_2 - k_1)^2 + \frac{8\ell^4}{d^2}\mathbb{E}\bigg[\bigg(\sum_{k=k_1+1}^{k_2} Z^d_{k,1}\mathbf{1}_{(\mathcal{A}^d_k)^c}\bigg)^4\bigg]. \tag{24}$$

The second term can be written as

$$\mathbb{E}\bigg[\bigg(\sum_{k=k_1+1}^{k_2} Z^d_{k,1}\mathbf{1}_{(\mathcal{A}^d_k)^c}\bigg)^4\bigg] = \sum \mathbb{E}\bigg[\prod_{i=1}^{4} Z^d_{m_i,1}\mathbf{1}_{(\mathcal{A}^d_{m_i})^c}\bigg],$$

where the sum is over all the quadruplets $(m_i)_{i=1}^4$ satisfying $m_i \in \{k_1+1, \dots, k_2\}$, $i = 1, \dots, 4$. The expectation on the right-hand side can be upper bounded depending on the cardinality of $\{m_1, \dots, m_4\}$. For all $1 \leq j \leq 4$, define

$$\mathcal{I}_j = \{(m_1, \dots, m_4) \in \{k_1 + 1, \dots, k_2\};\ \#\{m_1, \dots, m_4\} = j\}.$$

Let $(m_1, m_2, m_3, m_4) \in \{k_1 + 1, \dots, k_2\}^4$ and $(\tilde{X}^d_k)_{k\geq 0}$ be defined as

$$\tilde{X}^d_0 = X^d_0 \quad \text{and} \quad \tilde{X}^d_{k+1} = \tilde{X}^d_k + \mathbf{1}_{k\notin\{m_1-1,m_2-1,m_3-1,m_4-1\}}\frac{\ell}{\sqrt{d}}Z^d_{k+1}\mathbf{1}_{\tilde{\mathcal{A}}^d_{k+1}},$$

with $\tilde{\mathcal{A}}^d_{k+1} = \{U_{k+1} \leq \exp(\sum_{i=1}^d \Delta\tilde{V}^d_{k,i})\}$, where, for all $k \geq 0$ and all $1 \leq i \leq d$, $\Delta\tilde{V}_{k,i}$ is defined by

$$\Delta\tilde{V}^d_{k,i} = V(\tilde{X}^d_{k,i}) - V\bigg(\tilde{X}^d_{k,i} + \frac{\ell}{\sqrt{d}}Z^d_{k+1,i}\bigg).$$

Note that, on the event $\bigcap_{j=1}^4\{\mathcal{A}^d_{m_j}\}^c$, the two processes $(X_k)_{k\geq 0}$ and $(\tilde{X}_k)_{k\geq 0}$ are equal. Let $\mathcal{F}$ be the σ-field generated by $(\tilde{X}^d_k)_{k\geq 0}$.

(i) We have $\#\{m_1, \dots, m_4\} = 4$, as the $\{(U_{m_j}, Z^d_{m_j,1}, \dots, Z^d_{m_j,d})\}_{1\leq j\leq 4}$ are independent conditionally to $\mathcal{F}$,

$$\begin{aligned}\mathbb{E}\bigg[\prod_{j=1}^4 Z^d_{m_j,1}\mathbf{1}_{(\mathcal{A}^d_{m_j})^c}\ \bigg|\ \mathcal{F}\bigg] &= \prod_{j=1}^4 \mathbb{E}[Z^d_{m_j,1}\mathbf{1}_{(\tilde{\mathcal{A}}^d_{m_j})^c}\ |\ \mathcal{F}] \\ &= \prod_{j=1}^4 \mathbb{E}\bigg[Z^d_{m_j,1}\varphi\bigg(\sum_{i=1}^d \Delta\tilde{V}^d_{m_j-1,i}\bigg)\ \bigg|\ \mathcal{F}\bigg],\end{aligned}$$

where $\varphi(x) = (1 - \mathrm{e}^x)_+$. Since the function φ is 1-Lipschitz, we have

$$\begin{aligned}&\bigg|\varphi\bigg(\sum_{i=1}^d \Delta\tilde{V}^d_{m_j-1,i}\bigg) - \varphi\bigg(-\frac{\ell}{\sqrt{d}}\dot{V}(\tilde{X}^d_{m_j-1,1})Z^d_{m_j,1} + \sum_{i=2}^d \Delta\tilde{V}^d_{m_j-1,i}\bigg)\bigg| \\ &\quad\leq \bigg|\Delta\tilde{V}^d_{m_j-1,1} + \frac{\ell}{\sqrt{d}}\dot{V}(\tilde{X}^d_{m_j-1,1})Z^d_{m_j,1}\bigg|.\end{aligned}$$

Then

$$\bigg|\mathbb{E}\bigg[\prod_{j=1}^4 Z^d_{m_j,1}\mathbf{1}_{(\mathcal{A}^d_{m_j})^c}\bigg]\bigg| \leq \mathbb{E}\bigg[\prod_{j=1}^4\{A^d_{m_j} + B^d_{m_j}\}\bigg],$$

where

$$A^d_{m_j} = \mathbb{E}\left[|Z^d_{m_j,1}|\left|\Delta\tilde{V}^d_{m_j-1,1} + \frac{\ell}{\sqrt{d}}\dot{V}(\tilde{X}^d_{m_j-1,1})Z^d_{m_j,1}\right| \;\Big|\; \mathcal{F}\right],$$

$$B^d_{m_j} = \left|\mathbb{E}\left[Z^d_{m_j,1}\left(1-\exp\left\{-\frac{\ell}{\sqrt{d}}\dot{V}(\tilde{X}^d_{m_j-1,1})Z^d_{m_j,1} + \sum_{i=2}^{d}\Delta\tilde{V}^d_{m_j-1,i}\right\}\right)_+ \;\Big|\; \mathcal{F}\right]\right|.$$

By the inequality of arithmetic and geometric means and convex inequalities,

$$\left|\mathbb{E}\left[\prod_{j=1}^{4} Z^d_{m_j,1}\,\mathbf{1}_{(\mathcal{A}^d_{m_j})^c}\right]\right| \le 8\mathbb{E}\left[\sum_{j=1}^{4}(A^d_{m_j})^4 + (B^d_{m_j})^4\right].$$

By Lemma 3(ii) and the Hölder inequality, there exists $C > 0$ such that $\mathbb{E}[(A^d_{m_j})^4] \le Cd^{-2}$. On the other hand, by [9, Lemma 6] since $Z^d_{m_j,1}$ is independent of $\mathcal{F}$,

$$B^d_{m_j} = \left|\mathbb{E}\left[\frac{\ell}{\sqrt{d}}\dot{V}(\tilde{X}^d_{m_j-1,1})\mathcal{G}\left(\frac{\ell^2}{d}\dot{V}(\tilde{X}^d_{m_j-1,1})^2, -2\sum_{i=2}^{d}\Delta\tilde{V}^d_{m_j-1,i}\right) \;\Big|\; \mathcal{F}\right]\right|,$$

where the function $\mathcal{G}$ is defined in (22). By Assumption 1(ii) and since $\mathcal{G}$ is bounded, we have $\mathbb{E}[(B^d_{m_j})^4] \le Cd^{-2}$. Therefore, $|\mathbb{E}[\prod_{j=1}^4 Z^d_{m_j,1}\,\mathbf{1}_{(\mathcal{A}^d_{m_j})^c}]| \le Cd^{-2}$, yielding

$$\sum_{(m_1,m_2,m_3,m_4)\in\mathcal{I}_4}\left|\mathbb{E}\left[\prod_{i=1}^{4} Z^d_{m_i,1}\,\mathbf{1}_{(\mathcal{A}^d_{m_i})^c}\right]\right| \le \frac{C}{d^2}\binom{k_2-k_1}{4}. \tag{25}$$

(ii) We have $\#\{m_1,\dots,m_4\} = 3$, as the $\{(U_{m_j}, Z^d_{m_j,1},\dots,Z^d_{m_j,d})\}_{1\le j\le 3}$ are independent conditionally to $\mathcal{F}$,

$$\left|\mathbb{E}\left[(Z^d_{m_1,1})^2\,\mathbf{1}_{(\mathcal{A}^d_{m_1})^c}\prod_{j=2}^{3} Z^d_{m_j,1}\,\mathbf{1}_{(\mathcal{A}^d_{m_j})^c} \;\Big|\; \mathcal{F}\right]\right| \le \mathbb{E}[(Z^d_{m_1,1})^2 \mid \mathcal{F}]\left|\prod_{j=2}^{3}\mathbb{E}[Z^d_{m_j,1}\,\mathbf{1}_{(\tilde{\mathcal{A}}^d_{m_j})^c} \mid \mathcal{F}]\right|$$
$$\le \left|\prod_{j=2}^{3}\mathbb{E}[Z^d_{m_j,1}\,\mathbf{1}_{(\tilde{\mathcal{A}}^d_{m_j})^c} \mid \mathcal{F}]\right|.$$

Then, following the same steps as above, and using Hölder's inequality yields

$$\left|\mathbb{E}\left[\prod_{j=2}^{3} Z^d_{m_j,1}\,\mathbf{1}_{(\mathcal{A}^d_{m_j})^c}\right]\right| \le C\mathbb{E}\left[\sum_{j=2}^{3}(A^d_{m_j})^2 + (B^d_{m_j})^2\right] \le Cd^{-1}$$

and

$$\sum_{(m_1,m_2,m_3,m_4)\in\mathcal{I}_3}\left|\mathbb{E}\left[\prod_{i=1}^{4} Z^d_{m_i,1}\,\mathbf{1}_{(\mathcal{A}^d_{m_i})^c}\right]\right| \le \frac{C}{d}\binom{k_2-k_1}{3} \le \frac{C}{d}(k_2-k_1)^3. \tag{26}$$

(iii) If $\#\{m_1,\dots,m_4\} = 2$, two cases have to be considered:

$$\mathbb{E}[(Z^d_{m_1,1})^2\,\mathbf{1}_{(\tilde{\mathcal{A}}^d_{m_1})^c}(Z^d_{m_2,1})^2\,\mathbf{1}_{(\mathcal{A}^d_{m_2})^c}] \le \mathbb{E}[(Z^d_{m_1,1})^2]\mathbb{E}[(Z^d_{m_2,1})^2] \le 1,$$

$$\mathbb{E}[(Z^d_{m_1,1})^3\,\mathbf{1}_{(\mathcal{A}^d_{m_1})^c}\,Z^d_{m_2,1}\,\mathbf{1}_{(\mathcal{A}^d_{m_2})^c}] \le \mathbb{E}[|Z^d_{m_1,1}|^3]\mathbb{E}[|Z^d_{m_2,1}|] \le \frac{4}{\pi}.$$

https://doi.org/10.1017/jpr.2017.61 Published online by Cambridge University Press

This yields

$$\sum_{(m_1,m_2,m_3,m_4)\in\mathcal{I}_2}\left|\mathbb{E}\left[\prod_{i=1}^{4} Z^d_{m_i,1}\,\mathbf{1}_{(\mathcal{A}^d_{m_i})^c}\right]\right| \le \left(3+4\cdot\frac{4}{\pi}\right)(k_2-k_1)(k_2-k_1-1)$$
$$\le C(k_2-k_1)^2. \tag{27}$$

(iv) If $\#\{m_1,\dots,m_4\}=1 : \mathbb{E}[(Z^d_{m_i,1}\,\mathbf{1}_{(\mathcal{A}^d_{m_i})^c})^4] \le \mathbb{E}[(Z^d_{m_1,1})^4] \le 3$, then

$$\sum_{(m_1,m_2,m_3,m_4)\in\mathcal{I}_1}\left|\mathbb{E}\left[\prod_{i=1}^{4} Z^d_{m_i,1}\,\mathbf{1}_{(\mathcal{A}^d_{m_i})^c}\right]\right| \le 3(k_2-k_1). \tag{28}$$

The proof is completed by combining (24) with (25)–(28). □

4.4. Proof of reduction to the martingale problem (Proposition 2)

We preface the proof by a preliminary lemma.

Lemma 8. *Assume that Assumption 1 holds. Let μ be a limit point of the sequence of laws $(\mu_d)_{d\ge1}$ of $\{(Y^d_{t,1})_{t\ge0},\ d\in\mathbb{N}^*\}$. Then, for all $t\ge0$, the pushforward measure of μ by W_t is π.*

Proof. By (8),

$$\lim_{d\to+\infty}\mathbb{E}[|Y^d_{t,1}-X^d_{\lfloor dt\rfloor,1}|]=0.$$

Since $(\mu_d)_{d\ge1}$ converges weakly to μ, for all bounded Lipschitz functions $\psi\colon\mathbb{R}\to\mathbb{R}$, $\mathbb{E}^{\mu}[\psi(W_t)] = \lim_{d\to+\infty}\mathbb{E}[\psi(Y^d_{t,1})] = \lim_{d\to+\infty}\mathbb{E}[\psi(X^d_{\lfloor dt\rfloor,1})]$. The proof is completed upon noting that, for all $d\in\mathbb{N}^*$ and all $t\ge0$, $X^d_{\lfloor dt\rfloor,1}$ is distributed according to π . □

Proposition 2. *Assume that Assumptions 1 and 2 hold. Assume also that, for all $\phi\in C^\infty_c(\mathbb{R},\mathbb{R})$, $m\in\mathbb{N}^*$, $g\colon\mathbb{R}^m\to\mathbb{R}$ bounded and continuous, and $0\le t_1\le\cdots\le t_m\le s\le t$,*

$$\lim_{d\to+\infty}\mathbb{E}^{\mu_d}\left[\left(\phi(W_t)-\phi(W_s)-\int_s^t L\phi(W_u)\,\mathrm{d}u\right)g(W_{t_1},\dots,W_{t_m})\right]=0. \tag{29}$$

Then every limit point of the sequence of probability measures $(\mu_d)_{d\ge1}$ on $\boldsymbol{W}$ is a solution to the martingale problem associated with (9).

Proof. Let μ be a limit point of the sequence of laws $(\mu_d)_{d\ge1}$ of $\{(Y^d_{t,1})_{t\ge0},\ d\in\mathbb{N}^*\}$. It is straightforward to show that μ is a solution to the martingale problem associated with L if, for all $\phi\in C^\infty_c(\mathbb{R},\mathbb{R})$, $m\in\mathbb{N}^*$, $g\colon\mathbb{R}^m\to\mathbb{R}$ bounded and continuous, and $0\le t_1\le\cdots\le t_m\le s\le t$,

$$\mathbb{E}^{\mu}\left[\left(\phi(W_t)-\phi(W_s)-\int_s^t L\phi(W_u)\,\mathrm{d}u\right)g(W_{t_1},\dots,W_{t_m})\right]=0. \tag{30}$$

Let $\phi\in C^\infty_c(\mathbb{R},\mathbb{R})$, $m\in\mathbb{N}^*$, $g\colon\mathbb{R}^m\to\mathbb{R}$ continuous and bounded, $0\le t_1\le\cdots\le t_m\le s\le t$, and $\boldsymbol{W}_{\dot V}=\{w\in\boldsymbol{W}\,|\,w_u\notin\mathcal{D}_{\dot V}$ for almost every $u\in[s,t]\}$. Note first that $w\in\boldsymbol{W}^c_{\dot V}$ if and only if $\int_s^t \mathbf{1}_{\mathcal{D}_{\dot V}}(w_u)\,\mathrm{d}u>0$. Therefore, by Assumption 2, Lemma 8, and Fubini's theorem,

$$\mathbb{E}^{\mu}\left[\int_s^t \mathbf{1}_{\mathcal{D}_{\dot V}}(W_u)\,\mathrm{d}u\right]=\int_s^t \mathbb{E}^{\mu}[\mathbf{1}_{\mathcal{D}_{\dot V}}(W_u)]\,\mathrm{d}u=0,$$

https://doi.org/10.1017/jpr.2017.61 Published online by Cambridge University Press

showing that $\mu(\boldsymbol{W}_{\dot V}^c) = 0$. We now prove that on $\boldsymbol{W}_{\dot V}$,

$$\Psi_{s,t} : w \mapsto \left\{\phi(w_t) - \phi(w_s) - \int_s^t L\phi(w_u)\,\mathrm{d}u\right\} g(w_{t_1}, \dots, w_{t_m}) \tag{31}$$

is continuous. It is clear that it is enough to show that $w \mapsto \int_s^t L\phi(w_u)\,\mathrm{d}u$ is continuous on $\boldsymbol{W}_{\dot V}$. So let $w \in \boldsymbol{W}_{\dot V}$ and $(w^n)_{n\geq 0}$ be a sequence in $\boldsymbol{W}$ which converges to w in the uniform topology on compact sets. Then, by Assumption 2, for any u such that $w_u \notin \mathcal{D}_{\dot V}$, $L\phi(w_u^n)$ converges to $L\phi(w_u)$ when n goes to ∞ and $L\phi$ is bounded. Therefore, by Lebesgue's dominated convergence theorem, $\int_s^t L\phi(w_u^n)\,\mathrm{d}u$ converges to $\int_s^t L\phi(w_u)\,\mathrm{d}u$. Hence, the map defined by (31) is continuous on $\boldsymbol{W}_{\dot V}$. Since $(\mu_d)_{d\geq 1}$ converges weakly to μ, by (29),

$$\mu(\Psi_{s,t}) = \lim_{d\to+\infty} \mu^d(\Psi_{s,t}) = 0,$$

which is precisely (30). □

4.5. Proof of Theorem 3

By Proposition 2, it is enough to check (29) to prove that μ is a solution to the martingale problem. The core of the proof of Theorem 3 is Proposition 3, for which we need two technical lemmata.

Lemma 9. *Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{U}$ be $\mathbb{R}$-valued random variables and $\epsilon > 0$. Assume that $\mathcal{U}$ is nonnegative and bounded by* 1. *Let $g\colon \mathbb{R} \to \mathbb{R}$ be a bounded function on $\mathbb{R}$ such that, for all $(x, y) \in (-\infty, -\epsilon]^2 \cup [\epsilon, +\infty)^2$, $|g(x) - g(y)| \leq C_g|x - y|$.*

(i) *For all $a > 0$,*

$$\begin{aligned}\mathbb{E}[\mathcal{U}|g(\mathcal{X}) - g(Y)|] &\\ \leq C_g\mathbb{E}[\mathcal{U}|\mathcal{X} - Y|] &\\ + \operatorname{osc}(g)\{\mathbb{P}[|\mathcal{X}| \leq \epsilon] + a^{-1}\mathbb{E}[\mathcal{U}|\mathcal{X} - Y|] + \mathbb{P}[\epsilon < |\mathcal{X}| < \epsilon + a]\},&\end{aligned}$$

where $\operatorname{osc}(g) = \sup(g) - \inf(g)$.

(ii) *If there exist $\mu \in \mathbb{R}$ and $\sigma, C_{\mathcal{X}} \in \mathbb{R}_+$ such that*

$$\sup_{x\in\mathbb{R}} \left|\mathbb{P}[\mathcal{X} \leq x] - \Phi\left(\frac{x-\mu}{\sigma}\right)\right| \leq C_{\mathcal{X}},$$

then

$$\begin{aligned}\mathbb{E}[\mathcal{U}|g(\mathcal{X}) - g(Y)|] &\\ \leq C_g\mathbb{E}[\mathcal{U}|\mathcal{X} - Y|] &\\ + 2\operatorname{osc}(g)\{C_{\mathcal{X}} + \sqrt{2\mathbb{E}[\mathcal{U}|\mathcal{X} - Y|](2\pi\sigma^2)^{-1/2}} + \epsilon(2\pi\sigma^2)^{-1/2}\}.&\end{aligned}$$

Proof. (i) Consider the following decomposition:

$$\begin{aligned}\mathbb{E}[\mathcal{U}|g(\mathcal{X}) - g(\mathcal{Y})|] &\\ = \mathbb{E}[\mathcal{U}|(g(\mathcal{X}) - g(\mathcal{Y}))|\,\mathbf{1}_{\{(\mathcal{X},\mathcal{Y})\in(-\infty,-\epsilon]^2\}\cup\{(\mathcal{X},\mathcal{Y})\in[\epsilon,+\infty)^2\}}] &\\ + \mathbb{E}[\mathcal{U}|g(\mathcal{X}) - g(\mathcal{Y})|(\mathbf{1}_{\{\mathcal{X}\in[-\epsilon,\epsilon]\}} + \mathbf{1}_{(\{\mathcal{X}<-\epsilon\}\cap\{\mathcal{Y}\geq-\epsilon\}\cup(\{\mathcal{X}>\epsilon\}\cap\{\mathcal{Y}\leq\epsilon\}))})].&\end{aligned}$$

In addition, for all $a > 0$,

$$(\{\mathcal{X} < -\epsilon\} \cap \{\mathcal{Y} \geq -\epsilon\}) \cup (\{\mathcal{X} > \epsilon\} \cap \{\mathcal{Y} \leq \epsilon\}) \\ \subset \{\epsilon < |\mathcal{X}| < \epsilon + a\} \cup (\{|\mathcal{X}| \geq \epsilon + a\} \cap \{|\mathcal{X} - \mathcal{Y}| \geq a\}).$$

Then using the fact that $\mathcal{U} \in [0, 1)$, we obtain

$$\mathbb{E}[\mathcal{U}|g(\mathcal{X}) - g(\mathcal{Y})|] \leq C_g \mathbb{E}[\mathcal{U}|\mathcal{X} - \mathcal{Y}|] + \operatorname{osc}(g)(\mathbb{P}[|\mathcal{X}| < \epsilon + a] + a^{-1}\mathbb{E}[\mathcal{U}|\mathcal{X} - \mathcal{Y}|]).$$

(ii) The result is straightforward if $\mathbb{E}[\mathcal{U}|\mathcal{X} - \mathcal{Y}|] = 0$. Assume that $\mathbb{E}[\mathcal{U}|\mathcal{X} - \mathcal{Y}|] > 0$. Combining the additional assumption and the previous result,

$$\mathbb{E}[\mathcal{U}|g(\mathcal{X}) - g(\mathcal{Y})|] \\ \leq C_g \mathbb{E}[\mathcal{U}|\mathcal{X} - \mathcal{Y}|] + \operatorname{osc}(g)\{2C_{\mathcal{X}} + 2(\epsilon + a)(2\pi\sigma^2)^{-1/2} + a^{-1}\mathbb{E}[\mathcal{U}|\mathcal{X} - \mathcal{Y}|]\}.$$

As this result holds for all $a > 0$, the proof is concluded by setting

$$a = \sqrt{\frac{\mathbb{E}[\mathcal{U}|\mathcal{X} - \mathcal{Y}|](2\pi\sigma^2)^{1/2}}{2}}. \qquad \square$$

Lemma 10. *Assume that Assumption 1 holds. Let X^d be distributed according to π^d and Z^d be a d-dimensional standard Gaussian random variable, independent of X^d. Then* $\lim_{d \to +\infty} \mathbb{E}^d = 0$*, where*

$$\mathbb{E}^d = \mathbb{E}\bigg[\bigg|\dot{V}(X_1^d)\bigg\{\mathcal{G}\bigg(\frac{\ell^2}{d}\dot{V}(X_1^d)^2, 2\sum_{i=2}^d \Delta V_i^d\bigg) - \mathcal{G}\bigg(\frac{\ell^2}{d}\dot{V}(X_1^d)^2, 2\sum_{i=2}^d b_i^d\bigg)\bigg\}\bigg|\bigg],$$

where ΔV_i^d and b_i^d are given by (5) and (16), respectively.

Proof. Set, for all $d \geq 1$, $\bar{Y}_d = \sum_{i=2}^d \Delta V_i^d$ and $\bar{X}_d = \sum_{i=2}^d b_i^d$. By (22), $\partial_b \mathcal{G}(a, b) = -\mathcal{G}(a, b)/2 + \exp(-b^2/8a)/(2\sqrt{2\pi a})$. As $\mathcal{G}$ is bounded and $x \mapsto x\exp(-x)$ is bounded on $\mathbb{R}_+$, $\sup_{a \in \mathbb{R}_+;\, |b| \geq a^{1/4}} \partial_b \mathcal{G}(a, b) < +\infty$. Therefore, there exists $C \geq 0$ such that, for all $a \in \mathbb{R}_+$ and $(b_1, b_2) \in (-\infty, -a^{1/4})^2 \cup (a^{1/4}, +\infty)^2$,

$$|\mathcal{G}(a, b_1) - \mathcal{G}(a, b_2)| \leq C|b_1 - b_2|. \tag{32}$$

By the definition of b_i^d in (16), $\bar{X}_d$ may be expressed as $\bar{X}_d = \sigma_d \bar{S}_d + \mu_d$, where

$$\mu_d = 2(d-1)\mathbb{E}[\zeta^d(X_1^d, Z_1^d)] - \frac{\ell^2(d-1)}{4d}\mathbb{E}[\dot{V}(X_1^d)^2],$$

$$\sigma_d^2 = \ell^2 \mathbb{E}[\dot{V}(X_1^d)^2] + \frac{\ell^4}{16d}\mathbb{E}[(\dot{V}(X_1^d)^2 - \mathbb{E}[\dot{V}(X_1^d)^2])^2], \qquad \bar{S}_d = (\sqrt{d}\sigma_d)^{-1}\sum_{i=2}^d \beta_i^d,$$

$$\beta_i^d = -\ell Z_i^d \dot{V}(X_i^d) - \frac{\ell^2}{4\sqrt{d}}(\dot{V}(X_i^d)^2 - \mathbb{E}[\dot{V}(X_i^d)^2]).$$

By Assumption 1(ii), the Berry–Essen theorem [13, Theorem 5.7] can be applied to $\bar{S}_d$. Then there exists a universal constant C such that, for all $d > 0$,

$$\sup_{x \in \mathbb{R}} \bigg|\mathbb{P}\bigg[\bigg(\frac{d}{d-1}\bigg)^{1/2} \bar{S}_d \leq x\bigg] - \Phi(x)\bigg| \leq \frac{C}{\sqrt{d}}.$$

https://doi.org/10.1017/jpr.2017.61 Published online by Cambridge University Press

It follows that

$$\sup_{x\in\mathbb{R}}\left|\mathbb{P}[\bar{X}_d \le x] - \Phi\left(\frac{x-\mu_d}{\tilde{\sigma}_d}\right)\right| \le \frac{C}{\sqrt{d}},$$

where $\tilde{\sigma}_d^2 = (d-1)\sigma_d^2/d$. By this result and (32), Lemma 9 can be applied to obtain a constant $C \ge 0$, independent of d, such that

$$\mathbb{E}\left[\left|\mathcal{G}\left(\frac{\ell^2\dot{V}(X_1^d)^2}{d}, 2\bar{Y}_d\right) - \mathcal{G}\left(\frac{\ell^2\dot{V}(X_1^d)^2}{d}, 2\bar{X}_d\right)\right| \,\middle|\, X_1^d\right]$$
$$\le C\left(\varepsilon_d + d^{-1/2} + \sqrt{2\varepsilon_d(2\pi\tilde{\sigma}_d^2)^{-1/2}} + \sqrt{\frac{\ell|\dot{V}(X_1^d)|}{2\pi d^{1/2}\tilde{\sigma}_d^2}}\right),$$

where $\varepsilon_d = \mathbb{E}[|\bar{X}_d - \bar{Y}_d|]$. Using this result, we have

$$\mathbb{E}^d \le C\{(\varepsilon_d + d^{-1/2} + \sqrt{2\varepsilon_d(2\pi\tilde{\sigma}_d^2)^{-1/2}})\mathbb{E}[|\dot{V}(X_1^d)|] + \ell^{1/2}\mathbb{E}[|\dot{V}(X_1^d)|^{3/2}](2\pi d^{1/2}\tilde{\sigma}_d^2)^{-1/2}\}. \tag{33}$$

By Lemma 4, ε_d goes to 0 as d goes to ∞, and, by Assumption 1(ii), $\lim_{d\to+\infty}\sigma_d^2 = \ell^2\mathbb{E}[\dot{V}(X)^2]$. Combining these results with (33), it follows that $\mathbb{E}^d$ goes to 0 when d goes to ∞. □

For all $n \ge 0$, define $\mathcal{F}_n^d = \sigma(\{X_k^d,\ k \le n\})$ and, for all $\phi \in C_c^\infty(\mathbb{R},\mathbb{R})$, let $M_n^d(\phi)$ be the discrete-time martingale

$$M_n^d(\phi) = \frac{\ell}{\sqrt{d}}\sum_{k=0}^{n-1}\phi'(X_{k,1}^d)\{Z_{k+1,1}^d\mathbf{1}_{\mathcal{A}_{k+1}^d} - \mathbb{E}[Z_{k+1,1}^d\mathbf{1}_{\mathcal{A}_{k+1}^d} \mid \mathcal{F}_k^d]\}$$
$$+ \frac{\ell^2}{2d}\sum_{k=0}^{n-1}\phi''(X_{k,1}^d)\{(Z_{k+1,1}^d)^2\mathbf{1}_{\mathcal{A}_{k+1}^d} - \mathbb{E}[(Z_{k+1,1}^d)^2\mathbf{1}_{\mathcal{A}_{k+1}^d} \mid \mathcal{F}_k^d]\}. \tag{34}$$

Proposition 3. *Assume that Assumptions 1 and 2 hold. Then, for all $s \le t$ and all $\phi \in C_c^\infty(\mathbb{R},\mathbb{R})$,*

$$\lim_{d\to+\infty}\mathbb{E}\left[\left|\phi(Y_{t,1}^d) - \phi(Y_{s,1}^d) - \int_s^t L\phi(Y_{r,1}^d)\,\mathrm{d}r - (M_{\lceil dt\rceil}^d(\phi) - M_{\lceil ds\rceil}^d(\phi))\right|\right] = 0.$$

Proof. First, since $\mathrm{d}Y_{r,1}^d = \ell\sqrt{d}Z_{\lceil dr\rceil,1}^d\mathbf{1}_{\mathcal{A}_{\lceil dr\rceil}^d}\,\mathrm{d}r$,

$$\phi(Y_{t,1}^d) - \phi(Y_{s,1}^d) = \ell\sqrt{d}\int_s^t \phi'(Y_{r,1}^d)Z_{\lceil dr\rceil,1}^d\mathbf{1}_{\mathcal{A}_{\lceil dr\rceil}^d}\,\mathrm{d}r. \tag{35}$$

As ϕ is C^3, using (8) and a Taylor expansion, for all $r \in [s,t]$, there exists $\chi_r \in [X_{\lfloor dr\rfloor,1}^d, Y_{r,1}^d]$ such that

$$\phi'(Y_{r,1}^d) = \phi'(X_{\lfloor dr\rfloor,1}^d) + \frac{\ell}{\sqrt{d}}(dr - \lfloor dr\rfloor)\phi''(X_{\lfloor dr\rfloor,1}^d)Z_{\lceil dr\rceil,1}^d\mathbf{1}_{\mathcal{A}_{\lceil dr\rceil}^d}$$
$$+ \frac{\ell^2}{2d}(dr - \lfloor dr\rfloor)^2\phi^{(3)}(\chi_r)(Z_{\lceil dr\rceil,1}^d)^2\mathbf{1}_{\mathcal{A}_{\lceil dr\rceil}^d}.$$

https://doi.org/10.1017/jpr.2017.61 Published online by Cambridge University Press

Substituting this expression into (35) yields

$$\begin{aligned}\phi(Y^d_{t,1})-\phi(Y^d_{s,1}) &= \ell\sqrt{d}\int_s^t \phi'(X^d_{\lfloor dr\rfloor,1})Z^d_{\lceil dr\rceil,1}\,\mathbf{1}_{\mathcal{A}^d_{\lceil dr\rceil}}\,\mathrm{d}r\\ &\quad+\ell^2\int_s^t (dr-\lfloor dr\rfloor)\phi''(X^d_{\lfloor dr\rfloor,1})(Z^d_{\lceil dr\rceil,1})^2\,\mathbf{1}_{\mathcal{A}^d_{\lceil dr\rceil}}\,\mathrm{d}r\\ &\quad+\frac{\ell^3}{2\sqrt{d}}\int_s^t (dr-\lfloor dr\rfloor)^2\phi^{(3)}(\chi_r)(Z^d_{\lceil dr\rceil,1})^3\,\mathbf{1}_{\mathcal{A}^d_{\lceil dr\rceil}}\,\mathrm{d}r.\end{aligned}$$

As $\phi^{(3)}$ is bounded,

$$\lim_{d\to+\infty}\mathbb{E}\bigg[\bigg|d^{-1/2}\int_s^t (dr-\lfloor dr\rfloor)^2\phi^{(3)}(\chi_r)(Z^d_{\lceil dr\rceil,1})^3\,\mathbf{1}_{\mathcal{A}^d_{\lceil dr\rceil}}\,\mathrm{d}r\bigg|\bigg]=0.$$

On the other hand, $I=\int_s^t \phi''(X^d_{\lfloor dr\rfloor,1})(dr-\lfloor dr\rfloor)(Z^d_{\lceil dr\rceil,1})^2\,\mathbf{1}_{\mathcal{A}^d_{\lceil dr\rceil}}\,\mathrm{d}r = I_1+I_2$ with

$$\begin{aligned}I_1&=\int_s^{\lceil ds\rceil/d}+\int_{\lfloor dt\rfloor/d}^t \phi''(X^d_{\lfloor dr\rfloor,1})\Big(dr-\lfloor dr\rfloor-\frac12\Big)(Z^d_{\lceil dr\rceil,1})^2\,\mathbf{1}_{\mathcal{A}^d_{\lceil dr\rceil}}\,\mathrm{d}r,\\ I_2&=\frac12\int_s^t \phi''(X^d_{\lfloor dr\rfloor,1})(Z^d_{\lceil dr\rceil,1})^2\,\mathbf{1}_{\mathcal{A}^d_{\lceil dr\rceil}}\,\mathrm{d}r.\end{aligned}$$

Note that

$$\begin{aligned}I_1&=\frac{1}{2d}(\lceil ds\rceil-ds)(ds-\lfloor ds\rfloor)\phi''(X^d_{\lfloor ds\rfloor,1})(Z^d_{\lceil ds\rceil,1})^2\,\mathbf{1}_{\mathcal{A}^d_{\lceil ds\rceil}}\\ &\quad+\frac{1}{2d}(\lceil dt\rceil-dt)(dt-\lfloor dt\rfloor)\phi''(X^d_{\lfloor dt\rfloor,1})(Z^d_{\lceil dt\rceil,1})^2\,\mathbf{1}_{\mathcal{A}^d_{\lceil dt\rceil}},\end{aligned}$$

showing, as ϕ'' is bounded, that $\lim_{d\to+\infty}\mathbb{E}[|I_1|]=0$. Therefore,

$$\lim_{d\to+\infty}\mathbb{E}[|\phi(Y^d_{t,1})-\phi(Y^d_{s,1})-I_{s,t}|]=0,$$

where

$$I_{s,t}=\int_s^t\bigg\{\ell\sqrt{d}\phi'(X^d_{\lfloor dr\rfloor,1})Z^d_{\lceil dr\rceil,1}+\frac{\ell^2\phi''(X^d_{\lfloor dr\rfloor,1})(Z^d_{\lceil dr\rceil,1})^2}{2}\bigg\}\mathbf{1}_{\mathcal{A}^d_{\lceil dr\rceil}}\,\mathrm{d}r.$$

Write

$$I_{s,t}-\int_s^t L\phi(Y^d_{r,1})\,\mathrm{d}r-(M^d_{\lceil dt\rceil}(\phi)-M^d_{\lceil ds\rceil}(\phi))=T^d_1+T^d_2+T^d_3-T^d_4+T^d_5,$$

where

$$\begin{aligned}T^d_1&=\int_s^t\phi'(X^d_{\lfloor dr\rfloor,1})\bigg(\ell\sqrt{d}\,\mathbb{E}\bigg[Z^d_{\lceil dr\rceil,1}\,\mathbf{1}_{\mathcal{A}^d_{\lceil dr\rceil}}\;\bigg|\;\mathcal{F}^d_{\lfloor dr\rfloor}\bigg]+\frac{h(\ell)}{2}\dot V(X^d_{\lfloor dr\rfloor,1})\bigg)\,\mathrm{d}r,\\ T^d_2&=\int_s^t\phi''(X^d_{\lfloor dr\rfloor,1})\bigg(\frac{\ell^2}{2}\mathbb{E}\bigg[(Z^d_{\lceil dr\rceil,1})^2\,\mathbf{1}_{\mathcal{A}^d_{\lceil dr\rceil}}\;\bigg|\;\mathcal{F}^d_{\lfloor dr\rfloor}\bigg]-\frac{h(\ell)}{2}\bigg)\,\mathrm{d}r,\\ T^d_3&=\int_s^t(L\phi(Y^d_{\lfloor dr\rfloor/d,1})-L\phi(Y^d_{r,1}))\,\mathrm{d}r,\\ T^d_4&=\frac{\ell(\lceil dt\rceil-dt)}{\sqrt{d}}\phi'(X^d_{\lfloor dt\rfloor,1})(Z^d_{\lceil dt\rceil,1}\,\mathbf{1}_{\mathcal{A}^d_{\lceil dt\rceil}}-\mathbb{E}[Z^d_{\lceil dt\rceil,1}\,\mathbf{1}_{\mathcal{A}^d_{\lceil dt\rceil}}\mid\mathcal{F}^d_{\lfloor dt\rfloor}])\\ &\quad+\frac{\ell^2(\lceil dt\rceil-dt)}{2d}\phi''(X^d_{\lfloor dt\rfloor,1})((Z^d_{\lceil dt\rceil,1})^2\,\mathbf{1}_{\mathcal{A}^d_{\lceil dt\rceil}}-\mathbb{E}[(Z^d_{\lceil dt\rceil,1})^2\,\mathbf{1}_{\mathcal{A}^d_{\lceil dt\rceil}}\mid\mathcal{F}^d_{\lfloor dt\rfloor}]),\end{aligned}$$

and

$$T_5^d = \frac{\ell(\lceil ds\rceil - ds)}{\sqrt{d}}\phi'(X^d_{\lfloor ds\rfloor,1})(Z^d_{\lceil ds\rceil,1}\mathbf{1}_{\mathcal{A}^d_{\lceil ds\rceil}} - \mathbb{E}[Z^d_{\lceil ds\rceil,1}\mathbf{1}_{\mathcal{A}^d_{\lceil ds\rceil}} \mid \mathcal{F}^d_{\lfloor ds\rfloor}])$$
$$+ \frac{\ell^2(\lceil ds\rceil - ds)}{2d}\phi''(X^d_{\lfloor ds\rfloor,1})((Z^d_{\lceil ds\rceil,1})^2\mathbf{1}_{\mathcal{A}^d_{\lceil ds\rceil}} - \mathbb{E}[(Z^d_{\lceil ds\rceil,1})^2\mathbf{1}_{\mathcal{A}^d_{\lceil ds\rceil}} \mid \mathcal{F}^d_{\lfloor ds\rfloor}]).$$

It is now proved that, for all $1 \le i \le 5$, $\lim_{d\to+\infty}\mathbb{E}[|T_i^d|] = 0$. First, as ϕ' and ϕ'' are bounded,

$$\mathbb{E}[|T_4^d| + |T_5^d|] \le Cd^{-1/2}.$$

Denote for all $r \in [s,t]$ and $d \ge 1$,

$$\Delta V^d_{r,i} = V(X^d_{\lfloor dr\rfloor,i}) - V(X^d_{\lfloor dr\rfloor,i} + \ell d^{-1/2} Z^d_{\lceil dr\rceil,i}),$$
$$\Xi^d_r = 1 \wedge \exp\bigg\{-\frac{\ell Z^d_{\lceil dr\rceil,1}\dot V(X^d_{\lfloor dr\rfloor,1})}{\sqrt{d}} + \sum_{i=2}^d b^d_{\lfloor dr\rfloor,i}\bigg\},$$
$$\Upsilon^d_r = 1 \wedge \exp\bigg\{-\frac{\ell Z^d_{\lceil dr\rceil,1}\dot V(X^d_{\lfloor dr\rfloor,1})}{\sqrt{d}} + \sum_{i=2}^d \Delta V^d_{r,i}\bigg\},$$

where, for all $k, i \ge 0$, $b^d_{k,i} = b^d(X^d_{k,i}, Z^d_{k+1,i})$, and, for all $x, z \in \mathbb{R}$, $b^d(x,y)$ is given by (16). By the triangle inequality,

$$|T_1^d| \le \int_s^t |\phi'(X^d_{\lfloor dr\rfloor,1})|(A_{1,r} + A_{2,r} + A_{3,r})\,\mathrm{d}r, \tag{36}$$

where

$$A_{1,r} = |\ell\sqrt{d}\mathbb{E}[Z^d_{\lceil dr\rceil,1}(\mathbf{1}_{\mathcal{A}^d_{\lceil dr\rceil}} - \Upsilon^d_r) \mid \mathcal{F}^d_{\lfloor dr\rfloor}]|,$$
$$A_{2,r} = |\ell\sqrt{d}\mathbb{E}[Z^d_{\lceil dr\rceil,1}(\Upsilon^d_r - \Xi^d_r) \mid \mathcal{F}^d_{\lfloor dr\rfloor}]|,$$
$$A_{3,r} = \bigg|\ell\sqrt{d}\mathbb{E}[Z^d_{\lceil dr\rceil,1}\Xi^d_r \mid \mathcal{F}^d_{\lfloor dr\rfloor}] + \frac{\dot V(X^d_{\lfloor dr\rfloor,1})h(\ell)}{2}\bigg|.$$

Since $t \mapsto 1 \wedge \exp(t)$ is 1-Lipschitz, by Lemma 3(ii), $\mathbb{E}[|A^d_{1,r}|]$ goes to 0 as d goes to $+\infty$ for almost all r. So by Fubini's theorem, the first term in (36) goes to 0 as d goes to $+\infty$. For $A^d_{2,r}$, by [9, Lemma 6],

$$\mathbb{E}[|A^d_{2,r}|] \le \mathbb{E}\bigg[\bigg|\ell^2\dot V(X^d_{\lfloor dr\rfloor,1})\bigg\{\mathcal{G}\bigg(\frac{\ell^2\dot V(X^d_{\lfloor dr\rfloor,1})^2}{d}, 2\sum_{i=2}^d \Delta V^d_{r,i}\bigg) - \mathcal{G}\bigg(\frac{\ell^2\dot V(X^d_{\lfloor dr\rfloor,1})^2}{d}, 2\sum_{i=2}^d b^d_{\lfloor dr\rfloor,i}\bigg)\bigg\}\bigg|\bigg],$$

where $\mathcal{G}$ is defined in (22). By Lemma 10, this expectation goes to 0 when d goes to ∞. Then by Fubini's theorem and the Lebesgue dominated convergence theorem, the second term of (36) goes to 0 as d goes to $+\infty$. For the last term, by [9, Lemma 6] again:

$$\ell\sqrt{d}\mathbb{E}[Z^d_{\lceil dr\rceil,1}\Xi^d_r \mid \mathcal{F}^d_{\lfloor dr\rfloor}]$$
$$= -\ell^2\dot V(X^d_{\lfloor dr\rfloor,1})\mathcal{G}\bigg(\frac{\ell^2}{d}\sum_{i=1}^d \dot V(X^d_{\lfloor dr\rfloor,i})^2, \frac{\ell^2}{2d}\sum_{i=2}^d \dot V(X^d_{\lfloor dr\rfloor,i})^2 - 4(d-1)\mathbb{E}[\zeta^d(X,Z)]\bigg), \tag{37}$$

https://doi.org/10.1017/jpr.2017.61 Published online by Cambridge University Press

where X is distributed according to π, and Z is a standard Gaussian random variable independent of X. As $\mathcal{G}$ is continuous on $\mathbb{R}_+ \times \mathbb{R} \setminus \{0, 0\}$ (see [9, Lemma 2]), by Assumption 1(ii), Lemma 5, and the law of large numbers, almost surely,

$$\begin{aligned}\lim_{d\to+\infty} \ell^2 \mathcal{G}\bigg(\frac{\ell^2}{d}\sum_{i=1}^{d} \dot V(X^d_{\lfloor dr\rfloor,i})^2, \frac{\ell^2}{2d}\sum_{i=2}^{d} \dot V(X^d_{\lfloor dr\rfloor,i})^2 - 4(d-1)\mathbb{E}[\zeta^d(X,Z)]\bigg) \\ = \ell^2\mathcal{G}(\ell^2\mathbb{E}[\dot V(X)^2], \ell^2\mathbb{E}[\dot V(X)^2]) \\ = \frac{h(\ell)}{2},\end{aligned}$$

where $h(\ell)$ is defined in (10). Therefore, by Fubini's theorem, (37), and Lebesgue's dominated convergence theorem, the last term of (36) goes to 0 as d goes to ∞. The proof for T_2^d follows the same lines. By the triangle inequality,

$$\begin{aligned}|T_2^d| \le \bigg|\int_s^t \phi''(X^d_{\lfloor dr\rfloor,1})\bigg(\frac{\ell^2}{2}\bigg)\mathbb{E}[(Z^d_{\lceil dr\rceil,1})^2(\mathbf{1}_{\mathcal{A}^d_{\lceil dr\rceil}} - \Xi^d_r) \mid \mathcal{F}^d_{\lfloor dr\rfloor}]\,\mathrm{d}r\bigg| \\ + \bigg|\int_s^t \phi''(X^d_{\lfloor dr\rfloor,1})\bigg(\bigg(\frac{\ell^2}{2}\bigg)\mathbb{E}[(Z^d_{\lceil dr\rceil,1})^2\,\Xi^d_r \mid \mathcal{F}^d_{\lfloor dr\rfloor}] - \frac{h(\ell)}{2}\bigg)\,\mathrm{d}r\bigg|. \qquad (38)\end{aligned}$$

By Fubini's theorem, Lebesgue's dominated convergence theorem, and Proposition 1, the expectation of the first term goes to 0 when d goes to ∞. For the second term, by [9, Lemma 6, Equation (A.5)],

$$\begin{aligned}\bigg(\frac{\ell^2}{2}\bigg)\mathbb{E}\bigg[(Z^d_{\lceil dr\rceil,1})^2 1 \wedge \exp\bigg\{-\frac{\ell Z^d_{\lceil dr\rceil,1}}{\sqrt{d}}\dot V(X^d_{\lfloor dr\rfloor,1}) + \sum_{i=2}^{d} b^d_{\lfloor dr\rfloor,i}\bigg\}\ \bigg|\ \mathcal{F}^d_{\lfloor dr\rfloor}\bigg] \\ = \frac{B_1 + B_2 - B_3}{2}, \qquad (39)\end{aligned}$$

where

$$\begin{aligned}B_1 &= \ell^2\Gamma\bigg(\frac{\ell^2}{d}\sum_{i=1}^{d}\dot V(X^d_{\lfloor dr\rfloor,i})^2, \frac{\ell^2}{2d}\sum_{i=2}^{d}\dot V(X^d_{\lfloor dr\rfloor,i})^2 - 4(d-1)\mathbb{E}[\zeta^d(X,Z)]\bigg),\\ B_2 &= \frac{\ell^4\dot V(X^d_{\lfloor dr\rfloor,1})^2}{d}\\ &\quad\times \mathcal{G}\bigg(\frac{\ell^2}{d}\sum_{i=1}^{d}\dot V(X^d_{\lfloor dr\rfloor,i})^2, \frac{\ell^2}{2d}\sum_{i=2}^{d}\dot V(X^d_{\lfloor dr\rfloor,i})^2 - 4(d-1)\mathbb{E}[\zeta^d(X,Z)]\bigg),\\ B_3 &= \frac{\ell^4\dot V(X^d_{\lfloor dr\rfloor,1})^2}{d}\bigg(\frac{2\pi\ell^2\sum_{i=1}^{d}\dot V(X^d_{\lfloor dr\rfloor,i})^2}{d}\bigg)^{-1/2}\\ &\quad\times\exp\bigg\{-\frac{[-(d-1)\mathbb{E}[2\zeta^d(X,Z)] + (\ell^2/(4d))\sum_{i=2}^{d}\dot V(X^d_{\lfloor dr\rfloor,i})^2]^2}{2\ell^2\sum_{i=1}^{d}\dot V(X^d_{\lfloor dr\rfloor,i})^2/d}\bigg\},\end{aligned}$$

where Γ is defined in (23). As Γ is continuous on $\mathbb{R}_+ \times \mathbb{R} \setminus \{0, 0\}$ (see [9, Lemma 2]), by

https://doi.org/10.1017/jpr.2017.61 Published online by Cambridge University Press

Assumption 1(ii), Lemma 5, and the law of large numbers, almost surely,

$$\lim_{d\to+\infty} \ell^2\Gamma\bigg(\frac{\ell^2}{d}\sum_{i=1}^{d}\dot{V}(X^d_{\lfloor dr\rfloor,i})^2, \frac{\ell^2}{2d}\sum_{i=2}^{d}\dot{V}(X^d_{\lfloor dr\rfloor,i})^2 - 4(d-1)\mathbb{E}[\zeta^d(X,Z)]\bigg)$$
$$= \ell^2\Gamma(\ell^2\mathbb{E}[\dot{V}(X)^2], \ell^2\mathbb{E}[\dot{V}(X)^2])$$
$$= h(\ell). \tag{40}$$

By Lemma 5, Assumption 1(ii), and the law of large numbers, almost surely,

$$\lim_{d\to+\infty} \exp\bigg\{-\frac{[-(d-1)\mathbb{E}[2\zeta^d(X,Z)] + (\ell^2/(4d))\sum_{i=2}^{d}\dot{V}(X^d_{\lfloor dr\rfloor,i})^2]^2}{2\ell^2\sum_{i=1}^{d}\dot{V}(X^d_{\lfloor dr\rfloor,i})^2/d}\bigg\}$$
$$= \exp\bigg\{-\frac{\ell^2}{8}\mathbb{E}[\dot{V}(X)^2]\bigg\}.$$

Then, as $\mathcal{G}$ is bounded on $\mathbb{R}_+ \times \mathbb{R}$,

$$\lim_{d\to+\infty} \mathbb{E}\bigg[\bigg|\int_s^t \phi''(X^d_{\lfloor dr\rfloor,1})(B_2 - B_3)\,\mathrm{d}r\bigg|\bigg] = 0. \tag{41}$$

Therefore, by Fubini's theorem, (39)–(41), and Lebesgue's dominated convergence theorem, the second term of (38) goes to 0 as d goes to ∞. Write $T^d_3 = (h(\ell)/2)(T^d_{3,1} - T^d_{3,2})$, where

$$T^d_{3,1} = \int_s^t \{\phi''(X^d_{\lfloor dr\rfloor,1}) - \phi''(Y^d_{r,1})\}\,\mathrm{d}r,$$
$$T^d_{3,2} = \int_s^t \{\dot{V}(X^d_{\lfloor dr\rfloor,1})\phi'(X^d_{\lfloor dr\rfloor,1}) - \dot{V}(Y^d_{r,1})\phi'(Y^d_{r,1})\}\,\mathrm{d}r.$$

It is enough to show that $\mathbb{E}[|T^d_{3,1}|]$ and $\mathbb{E}[|T^d_{3,2}|]$ go to 0 when d goes to ∞ to conclude the proof. By (8) and the mean value theorem, for all $r \in [s,t]$, there exists $\chi_r \in [X^d_{\lfloor dr\rfloor,1}, Y^d_{r,1}]$ such that

$$\phi''(X^d_{\lfloor dr\rfloor,1}) - \phi''(Y^d_{r,1}) = \phi^{(3)}(\chi_r)(dr - \lfloor dr\rfloor)\bigg(\frac{\ell}{\sqrt{d}}\bigg)Z^d_{\lceil dr\rceil,1}\,\mathbf{1}_{\mathcal{A}^d_{\lceil dr\rceil}}.$$

Since $\phi^{(3)}$ is bounded, it follows that $\lim_{d\to+\infty}\mathbb{E}[|T^d_{3,1}|] = 0$. On the other hand,

$$T^d_{3,2} = \int_s^t \{\dot{V}(X^d_{\lfloor dr\rfloor,1}) - \dot{V}(Y^d_{r,1})\}\phi'(X^d_{\lfloor dr\rfloor,1})\,\mathrm{d}r + \int_s^t \{\phi'(X^d_{\lfloor dr\rfloor,1}) - \phi'(Y^d_{r,1})\}\dot{V}(Y^d_{r,1})\,\mathrm{d}r.$$

Since ϕ' has a bounded support, by Assumption 2, Fubini's theorem, and Lebesgue's dominated convergence theorem, the expectation of the absolute value of the first term goes to 0 as d goes to ∞. The second term is dealt with by following the same steps as for $T^d_{3,1}$ and using Assumption 1(ii). □

Proof of Theorem 3. By Lemma 6, Proposition 2, and Proposition 3, it is enough to prove that, for all $\phi \in C^\infty_c(\mathbb{R},\mathbb{R})$, $p \geq 1$, all $0 \leq t_1 \leq \cdots \leq t_p \leq s \leq t$, and $g\colon \mathbb{R}^p \to \mathbb{R}$ bounded and continuous function,

$$\lim_{d\to+\infty} \mathbb{E}[(M^d_{\lceil dt\rceil}(\phi) - M^d_{\lceil ds\rceil}(\phi))g(Y^d_{t_1},\ldots,Y^d_{t_p})] = 0,$$

where, for $n \geq 1$, $M^d_n(\phi)$ is defined in (34). But this result is straightforward taking successively the conditional expectations with respect to $\mathcal{F}_k$ for $k = \lceil dt\rceil,\ldots,\lceil ds\rceil$. □

https://doi.org/10.1017/jpr.2017.61 Published online by Cambridge University Press

5. Proofs of Section 3

5.1. Proof of Lemma 2

Lemma 11. *Assume that Assumption 3 holds. Let μ be a limit point of the sequence of laws $(\mu_d)_{d\geq 0}$ of $\{(Y^d_{t,1})_{t\geq 0},\ d\in\mathbb{N}^*\}$. Then, for all $t\geq 0$, the pushforward measure of μ by W_t is π.*

Proof. The proof is the same as in Lemma 8 and is omitted. □

Proof of Lemma 2. As, for all $t\geq 0$ and $d\geq 1$, $Y^d_{t,1}\in\mathcal{I}$, for all $d\geq 1$, $\mu^d(C(\mathbb{R}_+,\overline{\mathcal{I}}))=1$. Since $C(\mathbb{R}_+,\overline{\mathcal{I}})$ is closed in $\boldsymbol{W}$, we have, by the Portmanteau theorem, $\mu(C(\mathbb{R}_+,\overline{\mathcal{I}}))=1$. Therefore, we only need to prove that, for all $\psi\in C^\infty(\overline{\mathcal{I}},\mathbb{R})$, the process $(\psi(W_t)-\psi(W_0)-\int_0^t L\psi(W_u)\,\mathrm{d}u)_{t\geq 0}$ is a local martingale with respect to μ and the filtration $(\mathcal{B}_t)_{t\geq 0}$. Let $\psi\in C^\infty(\overline{\mathcal{I}},\mathbb{R})$.

Suppose first that, for all $\varpi\in C^\infty_c(\overline{\mathcal{I}},\mathbb{R})$, $(\varpi(W_t)-\varpi(W_t)-\int_0^t L\varpi(W_u)\,\mathrm{d}u)_{t\geq 0}$ is a martingale. Then consider the sequence of stopping times defined for $k\geq 1$ by $\tau_k=\inf\{t\geq 0\mid |W_t|\geq k\}$ and a sequence $(\varpi_k)_{k\geq 0}$ in $C^\infty_c(\overline{\mathcal{I}},\mathbb{R})$ satisfying,

(i) for all $k\geq 1$ and all $x\in\overline{\mathcal{I}}\cap[-k,k]$, $\varpi_k(x)=\psi(x)$;

(ii) $\lim_{k\to+\infty}\varpi_k=\psi$ in $C^\infty(\overline{\mathcal{I}},\mathbb{R})$.

Since, for all $k\geq 1$,

$$\begin{aligned}\left(\psi(W_{t\wedge\tau_k})-\psi(W_0)-\int_0^{t\wedge\tau_k}L\psi(W_u)\,\mathrm{d}u\right)_{t\geq 0}\\ =\left(\varpi_k(W_{t\wedge\tau_k})-\varpi_k(W_0)-\int_0^{t\wedge\tau_k}L\varpi_k(W_u)\,\mathrm{d}u\right)_{t\geq 0}\end{aligned}$$

and the sequence $(\tau_k)_{k\geq 1}$ goes to $+\infty$ as k goes to $+\infty$ almost surely, it follows that $(\psi(W_t)-\psi(W_0)-\int_0^t L\psi(W_u)\,\mathrm{d}u)_{t\geq 0}$ is a local martingale with respect to μ and the filtration $(\mathcal{B}_t)_{t\geq 0}$. It remains to show that, for all $\varpi\in C^\infty_c(\overline{\mathcal{I}},\mathbb{R})$, $(\varpi(W_t)-\varpi(W_0)-\int_0^t L\varpi(W_u)\,\mathrm{d}u)_{t\geq 0}$ is a martingale under the assumption of the proposition. We only need to prove that, for all $\varpi\in C^\infty_c(\overline{\mathcal{I}},\mathbb{R})$, $0\leq s\leq t$, $m\in\mathbb{N}^*$, $g\colon\mathbb{R}^m\to\mathbb{R}$ bounded and continuous, and $0\leq t_1\leq\cdots\leq t_m\leq s\leq t$,

$$\mathbb{E}^\mu\left[\left(\varpi(W_t)-\varpi(W_s)-\int_s^t L\varpi(W_u)\,\mathrm{d}u\right)g(W_{t_1},\ldots,W_{t_m})\right]=0. \tag{42}$$

Let $(\phi_k)_{k\geq 0}$ be a sequence of functions in $C^\infty_c(\mathcal{I},\mathbb{R})$ and converging to ϖ in $C^\infty_c(\overline{\mathcal{I}},\mathbb{R})$. First note that, for all $u\in[s,t]$, μ-almost everywhere,

$$\lim_{k\to+\infty}\phi_k(W_u)=\varpi(W_u). \tag{43}$$

By Lemma 11, for all $u\in[s,t]$, the pushforward measure of μ by W_u has density π with respect to the Lebesgue measure and μ-almost everywhere, $\lim_{k\to+\infty}L\phi_k(W_u)=L\varpi(W_u)$. On the other hand, there exists $C\geq 0$ such that, for all $k\geq 0$, $|L\phi_k(W_u)|\leq C(1+|\dot V(W_u)|)$. Then

$$\begin{aligned}\mathbb{E}^\mu\left[\int_s^t(1+|\dot V(W_u)|)\,\mathrm{d}u\right]&\leq(t-s)+\int_s^t\mathbb{E}^\mu[|\dot V(W_u)|]\,\mathrm{d}u\\&\leq(t-s)\left(1+\int_{\mathcal{I}}|\dot V(x)|\pi(x)\,\mathrm{d}x\right).\end{aligned}$$

https://doi.org/10.1017/jpr.2017.61 Published online by Cambridge University Press

Therefore, μ-almost everywhere, by Assumption 3(ii) and the Lebesgue dominated convergence theorem, we obtain

$$\lim_{k\to+\infty}\int_s^t L\phi_k(W_u)\,\mathrm{d}u = \int_s^t L\varpi(W_u)\,\mathrm{d}u. \tag{44}$$

Therefore, (42) follows from (43) and (44), using again the Lebesgue dominated convergence theorem and Assumption 3(ii) □

5.2. Proof of reduction to the martingale problem (Proposition 4)

Proposition 4. *We assume that Assumptions 3 and 4 hold. Assume also that, for all* $\phi \in C_c^\infty(\mathcal{I},\mathbb{R})$, $m \in \mathbb{N}^*$, $g\colon \mathbb{R}^m \to \mathbb{R}$ *bounded and continuous, and* $0 \le t_1 \le \cdots \le t_m \le s \le t$,

$$\lim_{d\to+\infty}\mathbb{E}^{\mu_d}\left[\left(\phi(W_t)-\phi(W_s)-\int_s^t L\phi(W_u)\,\mathrm{d}u\right)g(W_{t_1},\ldots,W_{t_m})\right]=0.$$

Then every limit point of the sequence of probability measures $(\mu_d)_{d\ge1}$ *on* $\mathbf{W}$ *is a solution to the local martingale problem associated with (9).*

Proof. Let μ be a limit point of $(\mu_d)_{d\ge1}$. First, following the proof of Proposition 2, it is straightforward to show that, for all $t \ge 0$, the pushforward measure of μ by W_t is π. By Lemma 2, we only need to prove that, for all $\phi \in C_c^\infty(\mathcal{I},\mathbb{R})$, the process $(\phi(W_t) - \phi(W_0) - \int_0^t L\phi(W_u)\,\mathrm{d}u)_{t\ge0}$ is a martingale with respect to μ and the filtration $(\mathcal{B}_t)_{t\ge0}$. Then the proof follows the same line as the proof of Proposition 2 and is omitted. □

Acknowledgement

The work of A. D. and E. M. is supported by the Agence Nationale de la Recherche, under grant ANR-14-CE23-0012 (COSMOS).

References

[1] Bédard, M. (2007). Weak convergence of Metropolis algorithms for non-i.i.d. target distributions. *Ann. Appl. Prob.* **17,** 1222–1244.

[2] Bédard, M. (2008). Efficient sampling using Metropolis algorithms: applications of optimal scaling results. *J. Comput. Graph. Statist.* **17,** 312–332.

[3] Bédard, M. (2008). Optimal acceptance rates for Metropolis algorithms: moving beyond 0.234. *Stoch. Process. Appl.* **118,** 2198–2222.

[4] Bédard, M. and Rosenthal, J. S. (2008). Optimal scaling of Metropolis algorithms: heading toward general target distributions. *Canad. J. Statist.* **36,** 483–503.

[5] Beskos, A., Roberts, G. and Stuart, A. (2009). Optimal scalings for local Metropolis–Hastings chains on nonproduct targets in high dimensions. *Ann. Appl. Prob.* **19,** 863–898.

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

[7] Cherny, A. S. and Engelbert, H.-J. (2005). *Singular Stochastic Differential Equations* (Lecture Notes Math. **1858**). Springer, Berlin.

[8] Durmus, A., Le Corff, S., Moulines, E. and Roberts, G. O. (2017). Optimal scaling of the random walk Metropolis algorithm under L^p mean differentiability. Supplementary material. Available at http://doi.org/10.1017/jpr.2017.61.

[9] Jourdain, B., Lelièvre, T. and Miasojedow, B. (2015). Optimal scaling for the transient phase of the random walk Metropolis algorithm: the mean-field limit. *Ann. Appl. Prob.* **25,** 2263–2300.

[10] Le Cam, L. (1986). *Asymptotic Methods in Statistical Decision Theory*. Springer, New York.

[11] Metropolis, N. *et al.* (1953). Equations of state calculations by fast computing machines. *J. Chem. Phys.* **21,** 1087–1091.

[12] Neal, P., Roberts, G. and Yuen, W. K. (2012). Optimal scaling of random walk Metropolis algorithms with discontinuous target densities. *Ann. Appl. Prob.* **22,** 1880–1927.

[13] PETROV, V. V. (1995). *Limit Theorems of Probability Theory* (Oxford Stud. Prob. **4**). Oxford University Press.
[14] ROBERTS, G. O. AND ROSENTHAL, J. S. (2014). Minimising MCMC variance via diffusion limits, with an application to simulated tempering. *Ann. Appl. Prob.* **24,** 131–149.
[15] 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.
[16] ROGERS, L. C. G. AND WILLIAMS, D. (2000). *Diffusions, Markov Processes and Martingales,* Vol 2, *Itô Calculus*. Cambridge University Press.
[17] SHERLOCK, C. AND ROBERTS, G. (2009). Optimal scaling of the random walk Metropolis on elliptically symmetric unimodal targets. *Bernoulli* **15,** 774–798.