# ON THE GEOMETRIC ERGODICITY OF HYBRID SAMPLERS

G. FORT, E. MOULINES, G. O. ROBERTS, AND J. S. ROSENTHAL

Abstract. In this paper, we consider the random scan symmetric random walk Metropolis algorithm (RSM) on $\mathbb{R}^d$. This algorithm performs a Metropolis step on just one coordinate at a time (as opposed to the full dimensional symmetric Random walk Metropolis algorithm, which proposes a transition on all coordinates at once). We present various sufficient conditions implying $V$-uniform ergodicity of the RSM when the target density decreases either sub-exponentially or exponentially in the tails.

**Keywords:** Markov Chain Monte Carlo; Hybrid sampler; geometric ergodicity.

**AMS 1991 subject classification:** 65C05, 60J10.

G. Fort, (Corresponding author) LMC-IMAG, 51 rue des Mathématiques, B.P. 53, 38041 Grenoble Cedex 9, France *E-mail:* Gersende.Fort@imag.fr

E. Moulines, ENST, 46, rue Barrault, 75634 Paris Cedex 13, France *E-mail:* moulines@tsi.enst.fr

G.O.Roberts, Department of Mathematics and Statistics, Lancaster University, Lancaster LA1 4YF, UK *E-mail:* G.O.Roberts@lancaster.ac.uk

J.S. Rosenthal, Department of Statistics, University of Toronto, 100 St. George Street, Toronto, Ontario, Canada M5S 3G3 *E-mail:* jeff@math.toronto.edu

# 1. Introduction

Markov Chain Monte-Carlo (MCMC) algorithms are well-known schemes to draw sample from an ergodic Markov chain with given stationary distribution $\pi$ on a state space $\mathcal{X}$. Theoretical work on MCMC algorithms has so far mainly concentrated on the properties of simple algorithms, such as the Gibbs sampler (see, e.g. Sahu and Roberts (1999) and Hobert and Geyer (1998)) or the full-dimensional Metropolis algorithm (see, e.g., Mengersen and Tweedie (1996), Roberts and Tweedie (1996), Jarner and Hansen (2000), Fort and Moulines )(2000c). In many practical situations, and in particular when the dimension of the state space is large, these elementary samplers are seldom used as they stand, but are rather used as building blocks for more complex sampling strategies (see e.g. Robert and Casella (1999)).

A rather intuitive idea to deal with large dimensional state space $\mathcal{X}$ is (whenever possible) to write the state space as a product of lower dimensional ones, $\mathcal{X} = \mathcal{X}_1 \times \cdots \times \mathcal{X}_d$, and to construct a Markov transition kernel $P$ on $\mathcal{X}$ having the stationary distribution $\pi$ by combining kernels $P_i$ acting on $\mathcal{X}_i$. The deterministic scan Gibbs sampler is an example of this strategy, where we write $P := Q_d Q_{d-1} \cdots Q_1$, where $Q_k$ is the Markov kernel that replaces the $k$-th coordinate by a draw from $\pi(dx_k | \{x_j\}_{j \neq k})$, leaving $x_j$ fixed for $j \neq k$. The random scan Gibbs sampler, $P := d^{-1} \sum_{i=1}^{d} Q_i$ is sometimes used instead (see Smith and Roberts (1993), Tierney (1994)). When the full conditional distributions $\pi(dx_i | \{x_j\}_{j \neq i})$ are difficult to sample, one can instead define new operators $P_i$ (e.g. one-dimensional Metropolis algorithms) which are easily implemented, such that $P_i^n$ converges to $Q_i$ (in an appropriate sense) as $n$ goes to infinity. This method is referred to as "variable-at-a-time Metropolis-Hastings" or "Metropolis-within-Gibbs" in the terminology of Tierney (1994) and Chan and Geyer (1994).

Let $\mathcal{C} := (P_1, P_2, \cdots, P_d)$ be any collection of Markov kernels on a state space $\mathcal{X} = \mathcal{X}_1 \times \cdots \times \mathcal{X}_d$. The random scan hybrid sampler for $\mathcal{C}$ is the sampler defined by

$$P_{\mathrm{RS}} := d^{-1}(P_1 + \cdots + P_d).$$

In this paper, we focus on the Random-Scan Metropolis (RSM) algorithm, where $\mathcal{X} = \mathbb{R}^d$, and where each operator $P_i$ arises from a symmetric random-walk Metropolis algorithm on the $i$-th coordinate.

This algorithm was studied by Roberts and Rosenthal (1997,1998), and by Jarner and Hansen (2000). One of the assumptions in Roberts and Rosenthal (1998) is expressed in terms of the maximal curvature of all the geodesic curves on the contour manifold $\{y \in \mathbb{R}^d, p(y) = p(x)\}$ as $|x| \to \infty$. This condition is rather difficult to check even when $d = 2$; in addition, as suggested in Jarner and Hansen (2000), it is not clear that this curvature condition should really play a role, since geometric ergodicity can be established fairly easily for densities for which the maximum curvature goes to infinity as $|x| \to \infty$, at least in some directions. In this paper, we shall instead show that geometric ergodicity holds under essentially no condition on the geometry of the contour manifold.

Let the state space $\mathcal{X}$ be equal to $\mathbb{R}^d$, equipped with its Borel $\sigma$-field $\mathcal{B}(\mathbb{R}^d)$. Let $\mu_d$ (resp. $\mu$) be Lebesgue measure on $\mathbb{R}^d$ (resp. $\mathbb{R}$) and $\{e_1, \cdots, e_d\}$ be the coordinate unit vectors. Denote by $|\cdot|$ the Euclidean norm. We shall assume that

**(A1)** the target distribution $\pi$ is absolutely continuous with respect to $\mu_d$, with positive and continuous density $p$ on $\mathbb{R}^d$.

Let $P_i$ be a symmetric random-walk Metropolis (with target density $p$) on the $i$-th coordinate: started from the $d$-vector $x = (x_1, \ldots, x_d)$, the proposal in the $e_i$-direction is given by $x + ye_i$, where $y$ is sampled from a symmetric increment density $q_i$ with respect to the one-dimensional Lebesgue measure $\mu$; this proposal is then accepted with probability $1 \wedge \{p(x+ye_i)/p(x)\}$. We shall assume for simplicity that

**(A2)** $\{q_i\}_{1 \le i \le d}$ is a family of symmetric densities with respect to $\mu$, such that there exist some constants $\eta_i > 0$, $\delta_i < \infty$ (for $i = 1, \ldots, d$) such that $|y| \le \delta_i \Longrightarrow q_i(y) \ge \eta_i$.

Condition (A2) ensures that the resulting Markov chain is $\phi$-irreducible and strongly aperiodic, and allows to identify *small sets* (see Section 2). For $x \in \mathbb{R}^d$ and $i \in \{1, \cdots, d\}$, let $\mathcal{A}(x, i)$ be the acceptance region in the $i$-th direction:

$$\mathcal{A}(x, i) := \{y \in \mathbb{R}, p(x + ye_i) \ge p(x)\}.$$

Similarly, let $\mathcal{R}(x, i)$ be the potential rejection region in the $i$-th direction:

$$\mathcal{R}(x, i) := \{y \in \mathbb{R},\ p(x + ye_i) < p(x)\}. \tag{1}$$

That is, $\mathcal{A}(x,i)$ represents the set of increments which, if proposed as an increment in the $i$-th direction, would always be accepted; $\mathcal{R}(x,i)$ represents those increments which could be rejected with positive probability.

With these notations, the transition kernels $P_i$, $i \in \{1, \cdots, d\}$, on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ are more formally defined as follows ; for $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$, $A = A_1 \times \cdots \times A_d \in \mathcal{B}(\mathbb{R}^d)$,

$$P_i(x, A) := \prod_{k \neq i} \delta_{x_k}(A_k) \int_{A_i - x_i} \alpha(x, x + ye_i) q_i(y) \mu(dy) + \delta_x(A) \int \Big(1 - \alpha(x, x + ye_i)\Big) q_i(y) \mu(dy),$$

where $A_i - x_i := \{y \in \mathbb{R}, x_i + y \in A_i\}$ and $\alpha(x, z) := 1 \wedge p(z)/p(x)$, $(x, z) \in \mathbb{R}^d \times \mathbb{R}^d$, so that for any Borel function $V : \mathbb{R}^d \to \mathbb{R}_+$, $x \in \mathbb{R}^d$,

$$P_i V(x) = \int_{\mathcal{A}(x,i)} V(x + ye_i) q_i(y) \mu(dy) + \int_{\mathcal{R}(x,i)} V(x + ye_i) \frac{p(x + ye_i)}{p(x)} q_i(y) \mu(dy) \\ + V(x) \int_{\mathcal{R}(x,i)} \Big(1 - \frac{p(x + ye_i)}{p(x)}\Big) q_i(y) \mu(dy). \quad (2)$$

The RSM kernel $P_{\mathrm{RS}}$ is the hybrid sampler associated to the collection $\mathcal{C} = (P_1, P_2, \cdots, P_d)$, i.e. $P_{\mathrm{RS}} := \frac{1}{d} \sum_{i=1}^d P_i$. The kernel $P_i$ is reversible with respect to the target distribution $\pi$, and thus $\pi$ is stationary for $P_i$ (and thus also for $P_{\mathrm{RS}}$).

Note finally that

$$P_i(x, A) = \prod_{k \neq i} \delta_{x_k}(A_k) M_i(x_i, A_i; x_{-i}) \quad (3)$$

where $x_{-i} := (x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_d)$ and $M_i(\cdot, \cdot; x_{-i})$ is the kernel of a random-walk Metropolis algorithm admitting the full-conditional distribution $\pi(dx_i | x_{-i})$ as its unique invariant distribution. The RSM is thus a special instance of *Metropolis-within-Gibbs* sampling.

## 2. Geometric ergodicity for sub-exponential densities

In this section we present a sufficient condition for geometric ergodicity of the RSM algorithm on $\mathbb{R}^d$ for sub-exponential densities.

### 2.1. **Background and assumptions.**

The proof of our result below uses the theory of drift and minorisation conditions for general Markov chains. We briefly review the necessary definitions here; see Meyn and Tweedie (1993) for further background.

A transition kernel $P$ (or a Markov Chain $\{X_n\}$) on a state space $\mathcal{X}$ is said to be *$\phi$-irreducible* if there is a non-zero measure $\phi$ on $\mathcal{B}(\mathcal{X})$, such that for all $x \in \mathcal{X}$, and for all measurable subsets $A \subseteq \mathcal{X}$ with $\phi(A) > 0$, the chain has positive probability of hitting $A$ when started at $x$, i.e. $P_x(\tau_A < \infty) > 0$ where $\tau_A = \min\{n \geq 1,\ X_n \in A\}$ is the first return time of $A$. The kernel $P$ is said to be *$V$-uniformly ergodic* for some function $V : \mathcal{X} \to [1, \infty]$ if $P$ is $\phi$-irreducible, with invariant probability measure $\pi$ such that $\pi(V) < \infty$, and there exist constants $r > 1$ and $R < \infty$ such that for $\pi$-almost all $x \in \mathcal{X}$,

$$\|P^n(x, \cdot) - \pi(\cdot)\|_V \leq Rr^{-n} V(x)\,, \qquad n = 1, 2, \ldots,$$

where for any signed measure $\mu$, $\|\mu\|_V := \sup_{|f| \leq V} |\mu(f)|$. (Furthermore, a chain is geometrically ergodic if and only if it is $V$-uniformly ergodic for some such $V$.)

Our proof consists in proving a Foster-Lyapunov condition outside a small set. Recall that $C \in \mathcal{B}(\mathbb{R}^d)$ is a small set if there exist an integer $m \geq 1$, a constant $\rho > 0$ and a probability measure $\nu_m$ on $\mathcal{B}(\mathbb{R}^d)$ such that

$$P^m(x, A) \geq \rho\ \nu_m(A) \qquad x \in C, A \in \mathcal{B}(\mathbb{R}^d). \tag{4}$$

From Theorems 15.0.1 and 16.0.1 of Meyn and Tweedie (1993), we have

**Theorem 1.** *Let $P$ be a $\phi$-irreducible aperiodic transition kernel. Assume that there exist some constants $0 < \lambda < 1$, $b < \infty$, some Borel function $V : \mathcal{X} \to [1, \infty]$ with $V(x_0) < \infty$ for some $x_0 \in \mathbb{R}^d$, and a small set $C$ satisfying*

$$PV(x) \leq \lambda V(x) + b\mathbb{1}_C(x)\,, \qquad x \in \mathcal{X}\,. \tag{5}$$

*Then $P$ is $V$-uniformly ergodic.*

**Remark 1.** Conversely, if $P$ is $V$-uniformly ergodic, then there exist $R < \infty$ and $r > 1$ such that for all $n$, $\sup_{x \in \{V < \infty\}} \|P^n(x, \cdot) - \pi(\cdot)\|_V / V(x) \leq Rr^{-n}$. If so, then there exists a function $V_0$ equivalent to $V$, which is a solution of the Foster-Lyapunov drift condition (5) (see Meyn and Tweedie (1993), Theorem 16.1.4).

**Remark 2.** Explicit expressions of the rate $r$ and of the constant $R$ as a function of the terms in (4) and (5) can be found in Meyn and Tweedie (1994), Mengersen and Tweedie (1996), Rosenthal (1995), Roberts and Tweedie (1998), Fort and Moulines (2000a), and Douc et al. (2001).

Under (A2), it is easily shown that $P^d_{\text{RS}}(x, .)$ has a nontrivial continuous component with respect to the Lebesgue measure and that this continuous component is bounded from below on a ball around $x$. From this, the positivity and the continuity of $p$, it is straightforward to prove the following result (Roberts and Rosenthal (1998), Lemma 4 ).

**Proposition 2.** *Assume (A1) and (A2). Then $P_{\mathrm{RS}}$ is Lebesgue-irreducible, aperiodic, with invariant probability measure $\pi(dx) := p(x)\mu_d(dx)$. In addition, any bounded set is small.*

To establish the Foster-Lyapunov condition, we need to find a function $V$ (which will depend on the dimension $d$) such that $\lim_{|x|\to\infty} P_{\mathrm{RS}}V(x)/V(x) < 1$. Consider the drift function $V_s(x) := p(x)^{-s}$, for some $0 < s < 1$. We have (Roberts and Rosenthal (1998), Proposition 3)

**Proposition 3.** *Let $P_i$ be given by (2), and set $V_s(x) := p(x)^{-s}$ for some $0 < s < 1$. For all $x \in \mathbb{R}^d$,*

$$P_i V_s(x) \le r(s) V_s(x) \qquad \textit{where} \qquad r(s) := 1 + s(1-s)^{1/s-1}. \tag{6}$$

*Hence, for all $\eta > 0$, there exists $s$ with $0 < s < \eta$, such that $1 < r(s) < 1 + \eta$.*

*Proof.* We provide a proof for completeness. We have

$$\begin{aligned}\frac{P_i V_s(x)}{V_s(x)} &= \int_{\mathcal{A}(x,i)} \left(\frac{p(x)}{p(x+ye_i)}\right)^s q_i(y)\mu(dy) + \int_{\mathcal{R}(x,i)} \left(1 - \frac{p(x+ye_i)}{p(x)} + \left(\frac{p(x+ye_i)}{p(x)}\right)^{1-s}\right) q_i(y)\mu(dy) \\ &= \int \mathcal{I}(y;x,i,s) q_i(y)\mu(dy)\end{aligned}$$

where

$$\mathcal{I}(y;x,i,s) := \begin{cases} (p(x)/p(x+ye_i))^s & y \in \mathcal{A}(x,i), \\ 1 - \frac{p(x+ye_i)}{p(x)} + \left(\frac{p(x+ye_i)}{p(x)}\right)^{1-s} & y \in \mathcal{R}(x,i). \end{cases} \tag{7}$$

The proof is concluded by noting that $\sup_{u\in[0,1]}(1 - u + u^{1-s}) \le r(s)$. □

Observe that $\lim_{s\to 0} r(s) = 1$, which shows that for any $\eta > 0$, by choosing $s$ small enough, one may find a function $V_s = p^{-s}$ such that for all $i \in \{1, \cdots, d\}$, $x \in \mathbb{R}^d$, $P_i V(x) \le (1+\eta)V(x)$. To prove the geometric ergodicity of the RSM algorithm, we of course need to prove something stronger.

The key assumption may be formulated as follows.

**(A3)** There is $0 \le \delta < \Delta \le +\infty$ such that $\xi := \inf_{1\le i\le d} \int_\delta^\Delta q_i(y)\mu(dy) > 0$, and for any sequence $\mathbf{x} := \{x^j\}$ with $\lim_j |x^j| = +\infty$, one may extract a subsequence $\tilde{\mathbf{x}} := \{\tilde{x}^j\}$, such that, for some $i \in \{1, \ldots, d\}$, and all $y \in [\delta, \Delta]$

$$\lim_j \frac{p(\tilde{x}^j)}{p(\tilde{x}^j - \mathrm{sign}(\tilde{x}^j_i)\ y\ e_i)} = 0, \quad \text{and} \quad \lim_j \frac{p(\tilde{x}^j + \mathrm{sign}(\tilde{x}^j_i)\ y\ e_i)}{p(\tilde{x}^j)} = 0. \tag{8}$$

This condition is somewhat involved. However, we will discuss in section 2.3 a simple criterion to check (A3).

2.2. **Main result.** The key result of Section 2 is the following.

**Theorem 4.** *Assume (A1), (A2), and (A3). Let* $0 < s < 1$ *such that*

$$r(s) < 1 + \frac{\xi}{d - 2\xi} \tag{9}$$

*where* $r(s)$ *and* $\xi$ *are given by (6) and (A3) respectively, and set* $V_s(x) := p(x)^{-s}$*. Then there exist constants* $0 < \lambda < 1$, $b < \infty$ *and a small set* $C \in \mathcal{B}(\mathbb{R}^d)$ *such that*

$$P_{\mathrm{RS}} V_s(x) \leq \lambda V_s(x) + b \mathbb{1}_C(x)\,, \qquad x \in \mathbb{R}^d.$$

*In particular,* $P_{\mathrm{RS}}$ *is* $V$*-uniformly ergodic.*

*Proof.* The proof is by contradiction. Assume that there exists a $\mathbb{R}^d$-valued sequence $\mathbf{x} := \{x^j\}$ such that $\lim_j |x^j| = +\infty$ and $\limsup_j P_{\mathrm{RS}} V_s(x^j)/V_s(x^j) \geq 1$. Then, there exists a subsequence $\hat{\mathbf{x}} := \{\hat{x}^j\}$ such that $\lim_j P_{\mathrm{RS}} V_s(\hat{x}^j)/V_s(\hat{x}^j) \geq 1$. We shall show that there exist a further subsequence $\tilde{\mathbf{x}} := \{\tilde{x}^j\}$ and an integer $i \in \{1, \cdots, d\}$ such that

$$\lim_j \frac{P_i V_s(\tilde{x}^j)}{V_s(\tilde{x}^j)} \leq r(s) - (2r(s) - 1)\,\xi. \tag{10}$$

The contradiction will follow from

$$\lim_j \frac{P_{\mathrm{RS}} V_s(\tilde{x}^j)}{V_s(\tilde{x}^j)} = \lim_j \frac{1}{d} \sum_{k=1}^d \frac{P_k V_s(\tilde{x}^j)}{V_s(\tilde{x}^j)} \leq \frac{1}{d}\left(r(s) - (2r(s)-1)\xi\right) + \frac{1}{d} \lim_j \sum_{k \neq i} \frac{P_k V_s(\tilde{x}^j)}{V_s(\tilde{x}^j)}$$
$$\leq \frac{1}{d}\left(r(s) - (2r(s)-1)\xi\right) + \frac{d-1}{d} r(s) < 1,$$

since, by Proposition 3, $P_k V_s(x)/V_s(x) \leq r(s)$ for all $x \in \mathbb{R}^d$. Under (A3), one may extract from the sequence $\hat{\mathbf{x}}$ a subsequence $\tilde{\mathbf{x}}$ in such a way that, for some $i \in \{1, \cdots, d\}$, (8) is verified and that, for all $j \geq 0$, $\mathrm{sign}(\tilde{x}^j_i) = \epsilon_i \in \{-1, +1\}$; without loss of generality set $\epsilon_i = 1$. We have

$$\frac{P_i V_s(x)}{V_s(x)} = \int_{\delta \leq |y| \leq \Delta} \mathcal{I}(y; x, i, s) q_i(y) \mu(dy) + \int_{\{|y| \leq \delta\} \cup \{|y| \geq \Delta\}} \mathcal{I}(y; x, i, s) q_i(y) \mu(dy),$$

where $\mathcal{I}$ is given by (7). Since $\mathcal{I}(y;x,i,s) \leq r(s)$, the second term on the right hand side of the previous equation is bounded by $r(s)(1-2\int_\delta^\Delta q_i(y)\mu(dy))$. Consider now the first term and set $J(\delta,\Delta) := [-\Delta,-\delta] \cup [\delta,\Delta]$. We first prove that

$$\lim_j \mathcal{R}(\tilde{x}^j, i) \cap J(\delta,\Delta) = [\delta,\Delta], \tag{11}$$

which implies that $\lim_j \mathcal{A}(\tilde{x}^j, i) \cap J(\delta,\Delta) = [-\Delta,-\delta]$. To this end, we show that

$$[\delta,\Delta] \subset \liminf_j \mathcal{R}(\tilde{x}^j, i) \cap J(\delta,\Delta) \subset \limsup_j \mathcal{R}(\tilde{x}^j, i) \cap J(\delta,\Delta) \subset [\delta,\Delta].$$

For $y \in [\delta,\Delta]$, $\lim_j p(\tilde{x}^j + ye_i)/p(\tilde{x}^j) = 0$; hence, $y \in \liminf_j \mathcal{R}(\tilde{x}^j, i) \cap J(\delta,\Delta)$. Assume now that $y \in \limsup_j \mathcal{R}(\tilde{x}^j, i) \cap J(\delta,\Delta)$. Then, $\liminf_j p(\tilde{x}^j)/p(\tilde{x}^j + ye_i) \geq 1$, and since $y \in J(\delta,\Delta)$ the latter relation implies that $y \in [\delta,\Delta]$, showing (11). By definition of the kernel $P_i$, we have

$$\frac{P_i V_s(\tilde{x}^j)}{V_s(\tilde{x}^j)} - \int_{\mathcal{R}(\tilde{x}^j,i)\cap J(\delta,\Delta)} q_i(y)\mu(dy) \leq \int_{\mathcal{A}(\tilde{x}^j,i)\cap J(\delta,\Delta)} \left[\frac{p(\tilde{x}^j + ye_i)}{p(\tilde{x}^j)}\right]^{-s} q_i(y)\mu(dy)$$
$$+\int_{\mathcal{R}(\tilde{x}^j,i)\cap J(\delta,\Delta)} \left\{\left[\frac{p(\tilde{x}^j + ye_i)}{p(\tilde{x}^j)}\right]^{1-s} - \left[\frac{p(\tilde{x}^j + ye_i)}{p(\tilde{x}^j)}\right]\right\} q_i(y)\mu(dy) + r(s)\left(1 - 2\int_\delta^\Delta q_i(y)\mu(dy)\right).$$

Note that for $y \in \mathcal{A}(x,i)$, $p(x+ye_i)/p(x) \geq 1$ and for $y \in \mathcal{R}(x,i)$, $p(x+ye_i)/p(x) \leq 1$. Then by (11), (A3) and the dominated convergence Theorem, we have

$$\lim_j \frac{P_i V_s(\tilde{x}^j)}{V_s(\tilde{x}^j)} \leq \int_\delta^\Delta q_i(y)\mu(dy) + r(s)\left(1 - 2\int_\delta^\Delta q_i(y)\mu(dy)\right)$$
$$\leq r(s) - (2r(s)-1)\int_\delta^\Delta q_i(y)\mu(dy) \leq r(s) - (2r(s)-1)\,\xi$$

which concludes the proof of (10) and thus of the first part of the Theorem.

Finally, we recall that assumptions (A1) and (A2) imply that any compact set is small. Furthermore, the above argument shows that assumption (A3) guarantees that outside a sufficiently large compact set $C$, we have $P_{\mathrm{RS}} V_s / V_s < 1$. Furthermore, $\sup_C V_s < \infty$, and by Proposition 3, $\sup_C P_{\mathrm{RS}} V_s < \infty$. The $V$-uniform ergodicity now follows from Theorem 1. □

**Remark 3.** In fact, it may be deduced from the proof of Theorem 4 that

$$\limsup_{|x|\to+\infty} \frac{P_{\mathrm{RS}} V_s(x)}{V_s(x)} \leq \frac{d-2\xi}{d} r(s) + \frac{\xi}{d}.$$

**Remark 4.** If instead the target density $p$ is positive and continuous and bounded on an unbounded open subset $\mathcal{X} \subset \mathbb{R}^d$, then assumption (A3) can be modified to still imply that the kernel $P_{\mathrm{RS}}$ is $V$-uniformly ergodic. (a) One has to replace "for any sequence $\mathbf{x} := \{x^j\}$ such that $\lim_j |x^j| = +\infty$" by "for any $\mathcal{X}$-valued sequence $\mathbf{x} := \{x^j\}$ such that $x^j \to \partial\mathcal{X}$", where $\partial\mathcal{X}$

is the boundary of $\mathcal{X}$. (b) One has to set, by convention, that for all $y > 0$, the ratio is zero for all $j$ such that $\tilde{x}^j - \mathrm{sign}(\tilde{x}^j_i)\ y\ e_i \notin \mathcal{X}$.

**Remark 5.** In assumption (A3), instead of (8), one could equivalently assume that there exists a function $I : \mathbb{R}^d \to \{1, \ldots, d\}$ such that

$$\lim_j \frac{p(\tilde{x}^j)}{p(\tilde{x}^j - \mathrm{sign}(\tilde{x}^j_{I(\tilde{x}^j)})\ y\ e_{I(\tilde{x}^j)})} = 0, \quad \text{and} \quad \lim_j \frac{p(\tilde{x}^j + \mathrm{sign}(\tilde{x}^j_{I(\tilde{x}^j)})\ y\ e_{I(\tilde{x}^j)})}{p(\tilde{x}^j)} = 0. \tag{12}$$

In fact, since $j \mapsto I(\tilde{x}^j)$ takes at most $d$ different values, one may choose $i \in \{1, \ldots, d\}$ such that $\{j;\ I(\tilde{x}^j) = i\}$ is infinite, and extract a further subsequence $\hat{\mathbf{x}} := \{\hat{x}^j\}$ from $\tilde{\mathbf{x}}$ such that $I(\hat{x}^j) = i$ for all $j \geq 0$. This subsequence satisfies (8).

**Remark 6.** If $p$ is continuously differentiable, the condition (A3) can be rewritten as follows:

- Let $0 < \Delta \leq +\infty$. For any sequence $\mathbf{x} := \{x^j\}$ such that $\lim_j |x^j| = +\infty$, there exists a subsequence $\tilde{\mathbf{x}} := \{\tilde{x}^j\}$ and $i \in \{1, \cdots, d\}$ such that, for all $0 \leq y < \Delta$,

$$\lim_{j \to +\infty} \sup_{\{t, |t| \leq y\}} \mathrm{sign}(\tilde{x}^j_i)\ \nabla_i \log p(\tilde{x}^j + t\, e_i) = -\infty, \tag{13}$$

where $\nabla_i := \partial / \partial x_i$.

2.3. **A criterion to check (A3).** We say that a function $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ is *quasi-monotonic* if the following condition is satisfied:

$$\lim_n \phi(t_n) = +\infty \quad \text{if and only if } \lim_n t_n = +\infty.$$

Non-decreasing functions $\phi$ such that $\lim_{x \to +\infty} \phi(x) = +\infty$ (e.g. the identity function $\phi(x) = x$) are quasi-monotonic, but the set of quasi-monotonic functions is much larger than that. Consider the following assumption:

**(A4)** There is $0 \leq \delta < \Delta \leq +\infty$ such that $\inf_{1 \leq i \leq d} \int_\delta^\Delta q_i(y) \mu(dy) > 0$, and there exist quasi-monotonic functions $\phi_k : \mathbb{R}_+ \to \mathbb{R}_+$ for $k = 1, \ldots, d$, such that for all $i \in \{1, \cdots, d\}$ and all $y \in [\delta, \Delta]$,

$$\lim_{|x_i| \to \infty} \sup_{\{x_{-i} : \phi_j(|x_j|) \leq \phi_i(|x_i|), j \neq i\}} \frac{p(x)}{p(x - \mathrm{sign}(x_i)\ y\ e_i)} = 0$$

$$\text{and} \quad \lim_{|x_i| \to \infty} \sup_{\{x_{-i} : \phi_j(|x_j|) \leq \phi_i(|x_i|), j \neq i\}} \frac{p(x + \mathrm{sign}(x_i)\ y\ e_i)}{p(x)} = 0.$$

**Proposition 5.** *Assume (A4). Then (A3) holds.*

*Proof.* Define

$$I : \mathbb{R}^d \to \{1, \cdots, d\}, \quad (x_1, \ldots, x_d) \mapsto \min \operatorname{argmax}_{k \in \{1, \cdots, d\}} \{\phi_k(|x_k|)\}.$$

Let $\mathbf{x} := \{x^j\}$ be a sequence such that $\lim_{j \to \infty} |x^j| = +\infty$. One may extract a subsequence $\{\tilde{x}^j\}$ and find $i \in \{1, \cdots, d\}$ such that $I(\tilde{x}^j) = i$ for all $j$, and the function $j \mapsto \operatorname{sign}(\tilde{x}_i^j)$ is constant. By construction $\phi_i(|\tilde{x}_i^j|) \geq \phi_k(|\tilde{x}_k^j|)$, $k \in \{1, \cdots, d\}$. In addition, since $|x^j| \to +\infty$, $\lim_j |\tilde{x}_k^j| = +\infty$ for some $k \in \{1, \cdots, d\}$. Thus $\lim_j \phi_k(|\tilde{x}_k^j|) = +\infty$ which implies that $\lim_j \phi_i(|\tilde{x}_i^j|) = +\infty$ and $\lim_j |\tilde{x}_i^j| = +\infty$. (A3) easily follows. □

If $p$ is differentiable, it is convenient to consider the criterion (A5):

**(A5)** The density $p$ is continuously differentiable, and there is $0 < \Delta \leq \infty$ and quasi-monotonic functions $\phi_k : \mathbb{R}_+ \to \mathbb{R}_+$, $k \in \{1, \ldots, d\}$, such that for all $i \in \{1, \cdots, d\}$, all $0 < y < \Delta$, and all $\epsilon \in \{-1, 1\}$,

$$\lim_{x_i \to \epsilon\infty} \sup_{\{x_{-i} : \phi_j(|x_j|) \leq \phi_i(|x_i|), j \neq i\}} \sup_{\{t ; |t| \leq y\}} \epsilon \, \nabla_i \log p(x + t e_i) = -\infty. \tag{14}$$

It is easily checked by standard analytical arguments that (A5) implies (A4).

2.4. **Examples.** We first consider a toy example and prove that if the target density is sub-exponential then $P_{\mathrm{RS}}$ is $V$-uniformly ergodic. We then consider the three examples proposed by Jarner and Hansen (2000) (Examples 2.4.2, 2.4.3 and 2.4.4) and a more realistic example proposed by Zeger (1988) (Example 2.4.5) and deduce from Theorem 4 the $V$-uniform ergodicity of the random scan kernel $P_{\mathrm{RS}}$ for each model. We finally consider an example, for which $P_{\mathrm{RS}}$ is not $V$-uniformly ergodic and (A3) does not hold. This counter-example demonstrates that while (A3) is certainly not a necessary condition for geometric ergodicity, it is far from being redundant.

For the target density $p$ considered in the Examples 2.4.1 to 2.4.5, the condition (A1) trivially holds. In addition, the proposal distributions $q_i$ can always be chosen in order to satisfy Assumption (A2). Assumption (A5) is established in all these examples with $\Delta = +\infty$ and $\phi_k(t) = t$ for all $k \in \{1, \ldots, d\}$ and $t \in \mathbb{R}_+$.

2.4.1. *Example 1.* On $\mathbb{R}^d$, define the density

$$p(x) \propto \exp(-|x|^l), \qquad l > 1, \qquad x = (x_1, \cdots, x_d).$$

Note that $\nabla_i \log p(x) = -l\,|x|^{l-2}\,x_i$ and for all $y > 0$,

$$\forall x_i > y, \forall x_{-i} \in [-x_i, x_i]^{d-1}, \forall |t| \leq y, \qquad \nabla_i \log p(x + te_i) \leq -l \frac{(x_i + y)^l}{\left(dx_i^2 + y^2 + 2yx_i\right)^{1/2}},$$

$$\forall x_i < -y, \forall x_{-i} \in [x_i, -x_i]^{d-1}, \forall |t| \leq y, \qquad \nabla_i \log p(x + te_i) \geq l \frac{|x_i + y|^l}{\left(dx_i^2 + y^2 + 2yx_i\right)^{1/2}}.$$

Since $l > 1$, (A5) easily follows by setting $\Delta = +\infty$ and by choosing $\phi_k$ as the identity function on $\mathbb{R}_+$, $k \in \{1, \cdots, d\}$. Hence, the RSM sampler is $V$-uniformly ergodic for any function $V(x) \propto p(x)^{-s}$ where $s(1-s)^{1/s-1} < (2d-2)^{-1}$.

2.4.2. *Example 2.* In this example, we consider the sum of two Gaussian densities on $\mathbb{R}^2$. Define for some $a^2 > 1$

$$p(x) \propto 0.5 \exp\left(-(x_1^2 + a^2 x_2^2)\right) + 0.5 \exp\left(-(a^2 x_1^2 + x_2^2)\right), \qquad x = (x_1, x_2).$$

As shown in Jarner and Hansen (2000), the contour curves have some sharp bends that do not disappear in the limit (even though the contour curves of the two components of the mixture are smooth ellipses). In particular the curvature on the diagonals $(x_1, x_2) = (t, t)$, $t \in \mathbb{R}$, tends to infinity as $t \to \infty$. For such target density, the main result of Roberts and Rosenthal (1998) does not apply (because the curvature does not tend to zero). We now show that nevertheless the RSM is $V$-uniformly ergodic.

We compute that

$$\nabla_1 \log p(x) = -2x_1 \frac{e^{-(x_1^2+a^2x_2^2)} + a^2 e^{-(a^2x_1^2+x_2^2)}}{e^{-(x_1^2+a^2x_2^2)} + e^{-(a^2x_1^2+x_2^2)}},$$
$$\nabla_2 \log p(x) = -2x_2 \frac{a^2 e^{-(x_1^2+a^2x_2^2)} + e^{-(a^2x_1^2+x_2^2)}}{e^{-(x_1^2+a^2x_2^2)} + e^{-(a^2x_1^2+x_2^2)}},$$

from which it easily follows that for all $x \in \mathbb{R}^2$,

$$1 \leq \frac{\nabla_1 \log p(x)}{-2x_1} \leq a^2 \qquad 1 \leq \frac{\nabla_2 \log p(x)}{-2x_2} \leq a^2.$$

(A5) easily follows by setting $\Delta = +\infty$ and by choosing $\phi_k$ as the identity function on $\mathbb{R}_+$, $k \in \{1, \cdots, d\}$. Hence, the RSM sampler is $V$-uniformly ergodic for any function $V(x) \propto p(x)^{-s}$ where $s(1-s)^{1/s-1} < 1/2$.

2.4.3. *Example 3.* Consider the sub-exponential density $p$ on $\mathbb{R}^2$ given by

$$p(x) \propto \exp\left(-(x_1^2 + x_1^2 x_2^2 + x_2^2)\right), \qquad x = (x_1, x_2).$$

A contour plot of the surface is given in Figure 1. This example has also been given in Jarner and Hansen (2000). These authors show that the full-dimensional random walk Metropolis algorithm is *not* geometrically ergodic for this target density. We will nevertheless show that the RSM algorithm *is* geometrically ergodic. This shows, perhaps surprisingly, that: *the RSM algorithm can be geometrically ergodic even in situations where the full-dimensional Metropolis algorithm is not geometrically ergodic.*

Here (A5) easily follows by setting $\Delta = +\infty$ and by choosing $\phi_k$ as the identity function on $\mathbb{R}_+$, $k \in \{1, \cdots, d\}$, and noting that

$$1 \le \frac{\nabla_1 \log p(x)}{-2x_1} = 1 + x_2^2 \qquad \text{and} \qquad 1 \le \frac{\nabla_2 \log p(x)}{-2x_2} = 1 + x_1^2.$$

Hence, the RSM sampler is $V$-uniformly ergodic for any function $V(x) \propto p(x)^{-s}$ where $s$ is chosen to satisfy $s(1-s)^{1/s-1} < 1/2$.

2.4.4. *Example 4.* Consider the sub-exponential density $p$ on $\mathbb{R}^2$

$$p(x) \propto (1 + x_1^2 + x_2^2 + x_1^8 x_2^2) \exp\left(-(x_1^2 + x_2^2)\right), \qquad x = (x_1, x_2),$$

introduced in Jarner and Hansen (2000). Once again, neither the curvature condition nor condition (5) in Roberts and Rosenthal (1998) hold (the curvature tends to infinity along the $x$-axis). Nevertheless, it is once again extremely simple to show that $P_{\mathrm{RS}}$ is $V$-uniformly ergodic. To that purpose, observe that

$$-2x_1 - \frac{2|x_1|}{x_1^2} - \frac{8|x_1|^7}{x_1^8} \le \nabla_1 \log p(x) = -2x_1 + \frac{2x_1 + 8x_1^7 x_2^2}{1 + x_1^2 + x_2^2 + x_1^8 x_2^2} \le -2x_1 + \frac{2|x_1|}{x_1^2} + \frac{8|x_1|^7}{x_1^8},$$

and

$$-2x_2 - \frac{2|x_2|}{x_2^2} - \frac{2|x_2|}{x_2^2} \le \nabla_2 \log p(x) = -2x_2 + \frac{2x_2 + 2x_1^8 x_2}{1 + x_1^2 + x_2^2 + x_1^8 x_2^2} \le -2x_2 + \frac{2|x_2|}{x_2^2} + \frac{2|x_2|}{x_2^2}.$$

(A5) follows by setting $\Delta = +\infty$ and by choosing $\phi_k$ as the identity function on $\mathbb{R}_+$, $k \in \{1, \cdots, d\}$. Hence, the RSM sampler is $V$-uniformly ergodic for any function $V(x) \propto p(x)^{-s}$ where $s(1-s)^{1/s-1} < 1/2$.

2.4.5. *Example 5.* Consider the following density, studied by Zeger (1988) and Chan and Ledolter (1995). Zeger proposed to fit the monthly number of cases of poliomyelitis with a generalised linear model with random effect: it is assumed that the observations $y := (y_1, \cdots, y_d)$ are generated from a Poisson distribution with mean $\lambda_k := \exp(\mu_k + X_k)$, where $\mu := (\mu_1, \cdots, \mu_d) \in \mathbb{R}^d$ is deterministic and $X_k$ is a stationary Gaussian AR(1) latent process $X_k = aX_{k-1} + \epsilon_k$, $\epsilon_k \sim \mathcal{N}(0, \lambda^{-1})$, $|a| < 1$. Chan and Ledolter (1995) considered the estimation of $(\mu, a, \lambda^{-1})$ by maximum likelihood using the Monte-Carlo EM algorithm. To assess the convergence of this algorithm, it is required to show the $V$-uniform ergodicity of the RSM when $p$ is the density of the latent process $X = (X_1, \cdots, X_d)$ given the observations $y$ (see Fort and Moulines )(2000b). Thus, $p$ is given for fixed $y$ by

$$p(x) \propto \exp\Big( \sum_{k=1}^{d} \{y_k(\mu_k + x_k) - \exp(\mu_k + x_k)\} - \lambda/2 \sum_{k=2}^{d} (x_k - a x_{k-1})^2 - \lambda(1 - a^2)x_1^2/2 \Big).$$

Note that

$$\begin{aligned} \nabla_1 \log p(x) &= y_1 - \exp(\mu_1 + x_1) - \lambda x_1 + \lambda a x_2, \\ \nabla_k \log p(x) &= y_k - \exp(\mu_k + x_k) - \lambda(1 + a^2)x_k + \lambda a(x_{k+1} - x_{k-1}), \qquad k \in \{2, \ldots, d-1\}, \\ \nabla_d \log p(x) &= y_d - \exp(\mu_d + x_d) - \lambda x_d + \lambda a x_{d-1}. \end{aligned}$$

It easily holds that for $|x_2| \le |x_1|$, $t \in \mathbb{R}$,

$$y_1 - \lambda t - \exp(\mu_1 + t + x_1) - \lambda x_1 - \lambda|a||x_1| \le \nabla_1 \log p(x + te_1) \le y_1 - \lambda t - \exp(\mu_1 + t + x_1) - \lambda x_1 + \lambda|a||x_1|;$$

for $|x_{d-1}| \le |x_d|$, $t \in \mathbb{R}$,

$$y_d - \lambda t - \exp(\mu_d + t + x_d) - \lambda x_d - \lambda|a||x_d| \le \nabla_d \log p(x + te_d) \le y_d - \lambda t - \exp(\mu_d + t + x_d) - \lambda x_d + \lambda|a||x_d|;$$

and for all $i \in \{2, \ldots, d-1\}$, given $|x_{i-1}| \le |x_i|$ and $|x_{i+1}| \le |x_i|$, $t \in \mathbb{R}$,

$$\begin{aligned} y_i - \lambda(1 + a^2)t - \exp(\mu_i + t + x_i) - \lambda(1 + a^2)x_i - 2\lambda|a||x_i| &\le \nabla_i \log p(x + te_i) \\ &\le y_i - \lambda(1 + a^2)t - \exp(\mu_i + t + x_i) - \lambda(1 + a^2)x_i + 2\lambda|a||x_i|. \end{aligned}$$

As $|a| < 1$, (A5) follows easily by setting $\Delta = +\infty$ and by choosing $\phi_k$ as the identity function on $\mathbb{R}_+$, $k \in \{1, \cdots, d\}$. Hence, the RSM sampler is $V$-uniformly ergodic for any function $V(x) \propto p(x)^{-s}$ where $s(1 - s)^{1/s - 1} < (2d - 2)^{-1}$.

2.4.6. *Example 6.* Consider now the RSM with target density $p$ on $\mathbb{R}^2$ given by

$$p(x) \propto \exp\left(-(x_1^2 + (x_1^2 - x_2^2)^2/4 + x_2^2)\right), \qquad x = (x_1, x_2),$$

which is the density $p$ studied in Example 2.4.3 in the new orthonormal basis $(\bar{e}_1, \bar{e}_2)$, $\bar{e}_1 := \sqrt{2}/2(e_1 + e_2)$ and $\bar{e}_2 := \sqrt{2}/2(e_2 - e_1)$. A contour plot of the surface is given in Figure 2. For this target density, (A3) is not verified (consider for example the sequence $x^j = (j, j)$). We shall show that the RSM algorithm is not geometrically ergodic. To prove this, we use a criterion outlined in Roberts and Tweedie (1996), Theorem 6.1. Let $P$ be a $\phi$-irreducible transition kernel, with invariant measure $\pi$ not concentrated at a single point. Let $h(x) := P(x, \{x\})$, and assume that $x \mapsto h(x)$ is a measurable function. If

$$\operatorname{esssup}_{x\in\mathcal{X}} h(x) := \sup\{h_0;\ \pi(h(x) > h_0) > 0\} = 1,$$

then $P$ is not geometrically ergodic.

To proceed, let $\lambda \geq 2$, and for all $j \geq \sqrt{\lambda}$, let $\mathcal{X}_j = [\sqrt{j^2 - \lambda}, j] \times [\sqrt{j^2 - \lambda}, j]$. Note that $\pi(\mathcal{X}_j) > 0$. We shall prove that

$$\lim_j \inf_{x\in\mathcal{X}_j} h(x) = 1, \tag{15}$$

which shows that the RSM is not geometrically ergodic for this target density. Towards that goal, we write

$$\begin{aligned} P_{\mathrm{RS}}(x, \{x\}) &\geq \frac{1}{2}\int (1 - \alpha(x, x + ye_1))\, q_1(y)\mu_2(dy) + \frac{1}{2}\int (1 - \alpha(x, x + ye_2))\, q_2(y)\mu_2(dy) \\ &\geq 1 - \frac{1}{2}\sum_{i=1}^{2}\int \alpha(x, x + ye_i) q_i(y)\mu_2(dy), \end{aligned}$$

and show that $\lim_j \sup_{x\in\mathcal{X}_j} \int \alpha(x, x + ye_i) q_i(y)\mu_2(dy) = 0$, $i = \{1, 2\}$. We consider only the case $i = 1$; the case $i = 2$ is similar. Observe that for all $x = (x_1, x_2) \in \mathbb{R}_+ \times \mathbb{R}$ such that $2x_2^2 - x_1^2 - 4 \geq 0$, the acceptance regions $\mathcal{A}(x, i)$ satisfy

$$\mathcal{A}(x,1) = \begin{cases} \left[-2x_1, -x_1 - \sqrt{2x_2^2 - x_1^2 - 4}\right] \cup \left[-x_1 + \sqrt{2x_2^2 - x_1^2 - 4}, 0\right], & \text{if } x_1^2 \geq x_2^2 - 2, \\ \left[-x_1 - \sqrt{2x_2^2 - x_1^2 - 4}, -2x_1\right] \cup \left[0, -x_1 + \sqrt{2x_2^2 - x_1^2 - 4}\right] & \text{otherwise} \end{cases}$$

(see Figure 3). For sufficiently large $j$, and $x \in \mathcal{X}_j$

$$\mathcal{A}(x,1) \subset \begin{cases} \left[-2j, -\sqrt{j^2 - \lambda} - \sqrt{j^2 - 2(\lambda + 2)}\right] \cup \left[\sqrt{j^2 - 2(\lambda+2)} - j, 0\right], & \text{if } x_1^2 \geq x_2^2 - 2, \\ \left[-\sqrt{j^2 + \lambda - 4} - j, -2\sqrt{j^2 - \lambda}\right] \cup \left[0, \sqrt{j^2 + \lambda - 4} - \sqrt{j^2 - \lambda}\right] & \text{otherwise} \end{cases}$$

Hence $\lim_j \sup_{x\in\mathcal{X}_j} \int_{\mathcal{A}(x,1)} q_1(y)\mu_2(dy) = 0$ (see Figure 4). Note that, for $y \neq 0$,

$$\log p(x+ye_1) - \log p(x) = -y(2x_1+y)\left(1 + x_1^2/2 - x_2^2/2 + y^2/4 + yx_1/2\right).$$

Then, for large enough $j$, $x \in \mathcal{X}_j$, and $y > 0$,

$$\log \frac{p(x+ye_1)}{p(x)} \leq -y\left(2\sqrt{j^2-\lambda}+y\right)\left(1 - \lambda/2 + y^2/4 + y\sqrt{j^2-\lambda}/2\right)$$

whereas, for $y < 0$,

$$\log \frac{p(x+ye_1)}{p(x)} \leq -y\left(2j+y\right)\left(1 + \lambda/2 + y^2/4 + y\sqrt{j^2-\lambda}/2\right).$$

In both cases, for all $y \neq 0$, $\lim_{j\to+\infty} \sup_{x\in\mathcal{X}_j} p(x+ye_1)/p(x) = 0$. Hence, by applying Lebesgue's dominated convergence theorem, $\lim_j \sup_{x\in\mathcal{X}_j} \int_{\mathcal{R}(x,1)} \frac{p(x+ye_1)}{p(x)} q_1(y)\mu_2(dy) = 0$. It follows that

$$\lim_j \sup_{x\in\mathcal{X}_j} \int \alpha(x, x+ye_1) q_1(y)\mu_2(dy) = 0,$$

and the RSM algorithm cannot be geometrically ergodic.

## 3. Geometric ergodicity for densities which are log-concave in the tails

Condition (A3) does not cover target densities which are log-concave in the tails. When $d = 1$, a target density is said to be *log-concave in the tails*, if there exist $\alpha > 0$ and some $x_1 > 0$ such that for all $x \geq x_1$ and $h \geq 0$,

$$\log p(x) - \log p(x+h) \geq \alpha h \tag{16}$$

and similarly for all $x \leq -x_1$,

$$\log p(x) - \log p(x-h) \geq \alpha h. \tag{17}$$

It has been shown in Mengersen and Tweedie, Theorem 3.2 (1996), that if (a) the target density $p$ satisfies (A1) and is log-concave in the tails, and (b) the proposal density satisfies (A2) and has a bounded exponential moment (a condition which can be avoided by adapting the proof), then the full dimensional random walk Metropolis algorithm is $V$-uniformly ergodic, and the drift function $V$ may be chosen as $V(x) = e^{s|x|}$, for any $s < \alpha$.

The main purpose of this section is to adapt these results to the RSM algorithm. Such extensions are also considered in Roberts and Rosenthal (1998), under the additional conditions that the target density has smooth contours in an explicit sense. We show that these conditions are in fact not needed. The first step in the construction is to extend the notion of log-concavity to distributions over $\mathbb{R}^d$. Following the approach of the previous section, this is most conveniently expressed in terms of limits of sub-sequences.

3.1. **Main result.** We shall replace (A3) with the following.

(**A3′**) There is $\alpha > 0$ and $1/\alpha \le \delta < \Delta \le \infty$ such that for any sequence $\mathbf{x} := \{x^j\}$ with $\lim_j |x^j| = +\infty$, one may extract a subsequence $\tilde{\mathbf{x}} := \{\tilde{x}^j\}$, such that for some $i \in \{1, \cdots, d\}$, we have, for all $y \in [\delta, \Delta]$,

$$\lim_j \frac{p(\tilde{x}^j)}{p(\tilde{x}^j - \mathrm{sign}(x_i^j)\ y\ e_i)} \le \exp(-\alpha y) \quad \text{and} \quad \lim_j \frac{p(\tilde{x}^j + \mathrm{sign}(\tilde{x}_i^j)\ y\ e_i)}{p(\tilde{x}^j)} \le \exp(-\alpha y). \tag{18}$$

It is easily seen that this notion of log-concavity in the tails generalises (16) and (17).

**Theorem 6.** *Assume (A1), (A2), and (A3'). Assume in addition that*

$$\inf_{i \in \{1,\dots,d\}} \int_\delta^\Delta y\ q_i(y)\mu(dy) \ge \frac{d}{\alpha(e-1)}. \tag{19}$$

*Then there exist* $0 < s < 1$*, some constants* $0 < \lambda < 1$*,* $b < \infty$ *and a small set* $C \in \mathcal{B}(\mathbb{R}^d)$ *such that*

$$P_{\mathrm{RS}} V_s(x) \le \lambda V_s(x) + b \mathbb{1}_C(x), \qquad x \in \mathcal{X},$$

*where* $V_s(x) := p^{-s}(x)$*. In particular,* $P_{\mathrm{RS}}$ *is* $V$*-uniformly ergodic.*

*Proof.* The proof is along the same lines as the proof of Theorem 4. Assume that there exists a sequence $\mathbf{x} := \{x^j\}$ such that $\lim_j |x^j| = +\infty$ and $\limsup_j P_{\mathrm{RS}} V_s(x^j)/V_s(x^j) \ge 1$. We shall show that there exists a further subsequence $\tilde{\mathbf{x}} := \{\tilde{x}^j\}$ such that $\lim_j \frac{P_{\mathrm{RS}} V_s(\tilde{x}^j)}{V_s(\tilde{x}^j)} < 1$, thereby obtaining a contradiction.

We first show that there exists a further subsequence $\tilde{\mathbf{x}} := \{\tilde{x}^j\}$ and $i \in \{1, \cdots, d\}$ such that

$$\lim_j \frac{P_i V_s(\tilde{x}^j)}{V_s(\tilde{x}^j)} \le r(s) - (2r(s) - 1)\, \mathcal{J}_i(0) + \mathcal{J}_i(\alpha s) + \mathcal{J}_i(\alpha(1-s)) - \mathcal{J}_i(\alpha)\,, \tag{20}$$

where for $b \ge 0$,

$$\mathcal{J}_i(b) := \int_{\delta \vee 1/\alpha}^{\Delta} e^{-by} q_i(y)\mu(dy).$$

Indeed, let $\tilde{\mathbf{x}}$ and $i$ be given by (A3′). We assume without loss of generality that $\mathrm{sign}(\tilde{x}_i^j) = 1$. Observe that

$$\lim_j \mathcal{R}(\tilde{x}^j, i) \cap J(\delta, \Delta) = [\delta, \Delta] \quad \text{and} \quad \lim_j \mathcal{A}(\tilde{x}^j, i) \cap J(\delta, \Delta) = [-\Delta, -\delta],$$

where $J(\delta,\Delta) := [-\Delta,-\delta] \cup [\delta,\Delta]$. (The proof of this assertion is along the same lines than the proof of (11), and is omitted here.) Proposition 3 implies

$$\frac{P_i V_s(\tilde{x}^j)}{V_s(\tilde{x}^j)} = \int_{\delta \le |y| \le \Delta} \mathcal{I}(y;\tilde{x}^j,i,s) q_i(y)\mu(dy) + \int_{\{|y|\le\delta\}\cup\{|y|\ge\Delta\}} \mathcal{I}(y;\tilde{x}^j,i,s) q_i(y)\mu(dy)$$
$$\le \int_{\delta \le |y| \le \Delta} \mathcal{I}(y;\tilde{x}^j,i,s) q_i(y)\mu(dy) + r(s)\left(1 - 2\mathcal{J}_i(0)\right), \quad (21)$$

where $\mathcal{I}$ is given by (7). In addition,

$$\int_{\delta \le |y| \le \Delta} \mathcal{I}(y;\tilde{x}^j,i,s) q_i(y)\mu(dy) = 2\mathcal{J}_i(0) + \int_{\mathcal{A}(\tilde{x}^j,i)\cap J(\delta,\Delta)} \left( \left[ \frac{p(\tilde{x}^j + y e_i)}{p(\tilde{x}^j)} \right]^{-s} - 1 \right) q_i(y)\mu(dy)$$
$$+ \int_{\mathcal{R}(\tilde{x}^j,i)\cap J(\delta,\Delta)} \left( \left[ \frac{p(\tilde{x}^j + y e_i)}{p(\tilde{x}^j)} \right]^{1-s} - \left[ \frac{p(\tilde{x}^j + y e_i)}{p(\tilde{x}^j)} \right] \right) q_i(y)\mu(dy).$$

Recall now that if $y \in \mathcal{A}(\tilde{x}^j,i)$, $p(\tilde{x}^j + ye_i) \ge p(\tilde{x}^j)$ whereas if $y \in \mathcal{R}(\tilde{x}^j,i)$, $p(\tilde{x}^j + ye_i) \le p(\tilde{x}^j)$. Hence, using Lebesgue's dominated convergence theorem,

$$\lim_j \int_{\delta \le |y| \le \Delta} \mathcal{I}(y;\tilde{x}^j,i,s) q_i(y)\mu(dy) = 2\mathcal{J}_i(0) + \int_{[-\Delta,-\delta]} \left( \left[ \lim_j \frac{p(\tilde{x}^j + y e_i)}{p(\tilde{x}^j)} \right]^{-s} - 1 \right) q_i(y)\mu(dy)$$
$$+ \int_{[\delta,\Delta]} \left( \left[ \lim_j \frac{p(\tilde{x}^j + y e_i)}{p(\tilde{x}^j)} \right]^{1-s} - \left[ \lim_j \frac{p(\tilde{x}^j + y e_i)}{p(\tilde{x}^j)} \right] \right) q_i(y)\mu(dy).$$

Since $q_i d\mu$ is a symmetric distribution and $u \mapsto u^{1-s} - u$ is non-decreasing on $[0, e^{-1}]$ for all $0 < s < 1$, we have

$$\lim_j \int_{\delta \le |y| \le \Delta} \mathcal{I}(y;\tilde{x}^j,i,s) q_i(y)\mu(dy) \le \mathcal{J}_i(0) + \mathcal{J}_i(\alpha s) + \mathcal{J}_i(\alpha(1-s)) - \mathcal{J}_i(\alpha). \quad (22)$$

Combining (21) and (22) yields (20).

Applying Proposition 3 again, we conclude that $P_{\mathrm{RS}} V_s(\tilde{x}^j)/V_s(\tilde{x}^j) \le \mathcal{H}_i(\alpha,s)$, where

$$\mathcal{H}_i(\alpha,s) := r(s) - \frac{1}{d}\left( (2r(s)-1)\,\mathcal{J}_i(0) + \mathcal{J}_i(\alpha) - \mathcal{J}_i(\alpha s) - \mathcal{J}_i(\alpha(1-s)) \right). \quad (23)$$

The result will follow if we can find $s > 0$ such that $\mathcal{H}_i(\alpha,s) < 1$. Since $\mathcal{H}_i(\alpha,0) = 1$ and $s \mapsto \mathcal{H}_i(\alpha,s)$ is differentiable at 0, it suffices to show $\frac{\partial}{\partial s}\mathcal{H}_i(\alpha,0) < 0$. This condition is fulfilled under (19), since

$$\frac{\partial}{\partial s}\mathcal{H}_i(\alpha,0) = (d - 2\mathcal{J}_i(0))e^{-1} - \alpha \int_\delta^\Delta y\ q_i(y)\mu(dy) + \alpha \int_\delta^\Delta y\ e^{-\alpha y} q_i(y)\mu(dy),$$
$$\le de^{-1} - \alpha(1 - e^{-1}) \int_\delta^\Delta y\ q_i(y)\mu(dy) < 0.$$

□

**Remark 7.** When $\Delta = +\infty$, the condition (19) is satisfied by choosing for example proposal distributions which are uniform on $[-Q, Q]$ for $Q$ large enough, or centered gaussian distributions with large enough variances.

**Remark 8.** As shown by the proof, the Foster-Lyapunov drift condition in Theorem 6 holds for any function $V_s := p^{-s}$ where $0 < s < 1$ is chosen such that $\sup_{i\in\{1,\dots,d\}} \mathcal{H}_i(\alpha, s) < 1$. If $\alpha = +\infty$, then (a) this condition essentially becomes $r(s) < 1 + \xi/(d-2\xi)$ where $\xi := \inf_{1\le i\le d} \int_\delta^\Delta q_i(y)\mu(dy)$; (b) the assumptions (A3) and (A3′) become similar; and (c) the condition (19) is always verified. Thus Theorem 4 and Theorem 6 essentially coincide in this case.

3.2. **Examples.** We first consider a toy-example and prove that if the target density is exponential then $P_{\mathrm{RS}}$ is $V$-uniformly ergodic. We then consider two examples adapted from examples 2.4.2 and 2.4.3. In example 3.2.2, the target density $p$ is a mixture of exponential densities on $\mathbb{R}^2$. For that density, the curvature condition of Roberts and Rosenthal (1998) fails to hold and their result does not apply. In example 3.2.3, we prove that the full dimensional random walk Metropolis algorithm can not be geometrically ergodic for the given target density whereas, as it is verified by application of our result, the random scan $P_{\mathrm{RS}}$ is $V$-uniformly ergodic. In these examples, assumption (A1) trivially holds; in addition, one can always choose the proposal densities $\{q_k\}$, $k \in \{1, \cdots, d\}$, in such a way that (A2) and (19) hold. Hence, in Paragraphs 3.2.1 to 3.2.3, if the condition (A3′) is proved, then it may be deduced from Theorem 6 that $P_{\mathrm{RS}}$ is $V$-uniformly ergodic for some function $V(x) \propto p(x)^{-s}$, $0 < s < 1$.

3.2.1. *Example 7.* On $\mathbb{R}^d$, define for $\lambda > 0$ the density

$$p(x) \propto \exp(-\lambda|x|), \qquad x = (x_1, \cdots, x_d).$$

Let $\mathbf{x} := \{x^j\}$ be a sequence such that $\lim_j |x^j| = +\infty$. Let $\tilde{\mathbf{x}} := \{\tilde{x}^j\}$ be a subsequence such that there exist $i \in \{1, \cdots, d\}$ and $\epsilon_i \in \{-1, 1\}$ and (a) for all $j \ge 0$ and $k \in \{1, \cdots, d\}$, we have $|\tilde{x}^j_k| \le |\tilde{x}^j_i|$; (b) for all $j \ge 0$, $\mathrm{sign}(\tilde{x}^j_i) = \epsilon_i$ and (c) $\lim_j |\tilde{x}^j_i|/|\tilde{x}^j|$ exists. Then, for all $y \ge 0$,

$$\log \frac{p\left(\tilde{x}^j + \epsilon_i y e_i\right)}{p(\tilde{x}^j)} = -\lambda\,|\tilde{x}^j|\,\left[\left(1 + \frac{y^2}{|\tilde{x}^j|^2} + 2y\frac{|\tilde{x}^j_i|}{|\tilde{x}^j|^2}\right)^{1/2} - 1\right].$$

Now, as $j \to +\infty$, $|\tilde{x}^j| \to +\infty$, so that

$$\lim_j \log \frac{p\left(\tilde{x}^j + \epsilon_i y e_i\right)}{p(\tilde{x}^j)} = -\lambda\, y \lim_j \frac{|\tilde{x}^j_i|}{|\tilde{x}^j|}.$$

Similarly,

$$\lim_j \log \frac{p(\tilde{x}^j)}{p\left(\tilde{x}^j - \epsilon_i y e_i\right)} = -\lambda\, y \lim_j \frac{|\tilde{x}^j_i|}{|\tilde{x}^j|}.$$

Hence, for all $y \geq 0$,

$$\lim_j \frac{p\left(\tilde{x}^j + \epsilon_i y e_i\right)}{p(\tilde{x}^j)} = \lim_j \frac{p(\tilde{x}^j)}{p\left(\tilde{x}^j - \epsilon_i y e_i\right)} \leq \exp(-\lambda y/\sqrt{d})$$

since $1/\sqrt{d} \leq \lim_j |\tilde{x}^j_i|/|\tilde{x}^j| \leq 1$. (A3$'$) is thus verified with $\alpha = 1/\sqrt{d}$, $\delta = \sqrt{d}$ and $\Delta = +\infty$. For proposal distributions satisfying (A2) and (19), Theorem 6 asserts that there exists $0 < s < 1$ such that $P_{\mathrm{RS}}$ is $V$-uniformly ergodic with $V(x) \propto p(x)^{-s}$.

3.2.2. *Example 8.* In this example, we consider the sum of two exponential densities on $\mathbb{R}^2$. Define for some $a > 1$

$$p(x) \propto 0.5 \exp\left(-\left(|x_1| + a|x_2|\right)\right) + 0.5 \exp\left(-\left(a|x_1| + |x_2|\right)\right), \qquad x = (x_1, x_2).$$

A contour plot is given in Figure 5 when $a = 4$. Similarly to what is done in Jarner and Hansen (2000) for a mixture of Gaussian densities, it may be proved that the curvature on the diagonal $(x_1, x_2) = (t, t)$, $t \in \mathbb{R}_+$, is a positive constant. Let the contour curve corresponding to a given level be given by $(x_1, h(x_1))$ in the first quadrant (that is $x_1 > 0$ and $h(x_1) > 0$). Then the curvature $\mathcal{K}((x_1, h(x_1)))$ of the curve at $(x_1, h(x_1))$ is given by

$$\mathcal{K}((x_1, h(x_1))) := \frac{|h''(x_1)|}{\left(1 + h'^2(x_1)\right)^{3/2}},$$

(see (46) in Jarner and Hansen (2000) or (1.11) in Laugwitz )(1965). In the present case, we find by implicit differentiation that for $(x_1, h(x_1)) = (t, t)$, $h'(x_1) = -1$ and $h''(x_1) = 2(1-a)^2/(1+a)$ so that

$$\mathcal{K}((t, t)) = \frac{1}{\sqrt{2}} \frac{(1-a)^2}{1+a} > 0.$$

Consequently, the result by Roberts and Rosenthal (1998) does not apply. We now show that the RSM is $V$-uniformly ergodic by application of Theorem 6. Let $\mathbf{x} := \{x^j\}$ be a sequence such that, without loss of generality, $\lim_j x^j_1 = +\infty$. We may extract a subsequence $\tilde{\mathbf{x}} := \{\tilde{x}^j\}$ such that (a) $|\tilde{x}_2^j| \leq \tilde{x}_1^j$; (b) for all $j \geq 0$, $\tilde{x}_1^j > 0$; and (c) $\lim_j(\tilde{x}_1^j - |\tilde{x}_2^j|) = L$, $L \in [0, +\infty]$. For all $y \geq 0$, we have

$$\lim_j \frac{p\left(\tilde{x}^j + ye_1\right)}{p(\tilde{x}^j)} = \begin{cases} \exp(-y) & L = +\infty, \\ \frac{\exp(-y)+\exp((1-aL))\exp(-ay)}{1+\exp((1-aL))} \leq \exp(-y) & L < +\infty \end{cases}$$

and similarly

$$\lim_j \frac{p\left(\tilde{x}^j\right)}{p(\tilde{x}^j - ye_1)} = \begin{cases} \exp(-y) & L = +\infty, \\ \frac{1+\exp((1-aL))}{\exp(y)+\exp((1-aL))\exp(ay)} \exp(-y) \leq \exp(-y) & L < +\infty \end{cases}$$

Hence (A3′) is verified by setting $\alpha = \delta = 1$ and $\Delta = +\infty$, and the RSM is $V$-uniformly ergodic.

3.2.3. *Example 9.* Consider the density $p$ on $\mathbb{R}^2$ given by

$$p(x) \propto \exp\left(-\left(|x_1| + |x_1||x_2| + |x_2|\right)\right) \qquad x = (x_1, x_2).$$

The contour plot of the surface is given in Figure 6. The full-dimensional random walk Metropolis kernel $P_{\mathrm{HM}}$ is not geometrically ergodic for this target density. As above, we may find a sequence of sets $\mathcal{X}_j \subset \mathbb{R}^2$, such that $\pi(\mathcal{X}_j) > 0$ and $\lim_j \inf_{x \in \mathcal{X}_j} P_{\mathrm{HM}}(x, \{x\}) = 1$; the criterion (15) shows that $P_{\mathrm{HM}}$ cannot be geometrically ergodic. Consider for example $\mathcal{X}_j := [j-1, j] \times [0, 1/j]$. Using straightforward calculations, it is shown that for all $y \in \mathbb{R}^2$, there exists $j$ large enough such that

$$\sup_{x \in \mathcal{X}_j} \alpha(x, x+y) \leq \exp\left(-j(|y_2| - 2/j) + |y_1|(1 + 1/j + |y_2|)\right).$$

Hence, if $y_2 \neq 0$, $\lim_j \sup_{x \in \mathcal{X}_j} \alpha(x, x+y) = 0$ and Lebesgue's dominated convergence theorem thus shows that

$$\liminf_j \inf_{x \in \mathcal{X}_j} P_{HM}(x, \{x\}) \geq 1 - \limsup_j \int \sup_{x \in \mathcal{X}_j} \alpha(x, x+y) q(y) \mu_2(dy) = 1.$$

For this target density, the RSM is $V$-uniformly ergodic. Indeed, let $\mathbf{x} := \{x^j\}$ be a sequence such that, without loss of generality, $\lim_j x_1^j = +\infty$. Let $\tilde{\mathbf{x}} := \{\tilde{x}^j\}$ be a subsequence such that (a) for all $j \geq 0$, $|\tilde{x}_2^j| \leq \tilde{x}_1^j$; (b) for all $j \geq 0$, $\tilde{x}_1^j > 0$; and (c) $\lim_j |\tilde{x}_2^j| = L \in [0, +\infty]$. We have for all $y \geq 0$,

$$\log \frac{p\left(\tilde{x}^j + ye_1\right)}{p(\tilde{x}^j)} = -\left(\tilde{x}_1^j + y - \tilde{x}_1^j\right)\left(1 + |\tilde{x}_2^j|\right) = -y\left(1 + |\tilde{x}_2^j|\right),$$

so that

$$\lim_j \frac{p\left(\tilde{x}^j + ye_1\right)}{p(\tilde{x}^j)} = \exp(-y(1+L)).$$

Similarly,

$$\lim_j \frac{p(\tilde{x}^j)}{p\left(\tilde{x}^j - ye_1\right)} = \exp(-y(1+L)).$$

Hence,

$$\lim_j \frac{p\left(\tilde{x}^j + ye_1\right)}{p(\tilde{x}^j)} = \lim_j \frac{p(\tilde{x}^j)}{p\left(\tilde{x}^j - ye_1\right)} = \exp(-y(1+L)) \le \exp(-y),$$

and (A3$'$) is verified with $\alpha = \delta = 1$ and $\Delta = +\infty$. For proposal distributions satisfying (A2) and (19), Theorem 6 asserts that there exists $0 < s < 1$ such that $P_{\mathrm{RS}}$ is $V$-uniformly ergodic with $V(x) \propto p(x)^{-s}$.

3.3. **A necessary condition for $V$-uniform ergodicity of the RSM algorithm.** It is of interest to find necessary conditions for $V$-uniform ergodicity. Theorem 3.3 of Mengersen and Tweedie (1996) states that, if (a) the target density is positive and continuous on $\mathbb{R}$, (b) the full dimensional random walk Metropolis algorithm is $V$-uniformly ergodic (here the state space is $\mathbb{R}$), and (c) the proposal distribution is symmetric, bounded away from zero in a neighborhood of zero and has a bounded mean $\int |y|\; q(y)\mu(dy) < \infty$, then the target density has an exponential moment, i.e. $\int e^{s|x|} p(x)\mu(dx) < \infty$ for some $s > 0$. This result has been extended to the multidimensional case (*i.e.* to the full dimensional random walk Metropolis algorithm) by Jarner and Hansen (2000), Theorem 3.3. A similar result holds for the RSM. The proof below is adapted from the proof of Theorem 3.3. of Jarner and Hansen .

**Proposition 7.** *Assume (A1) and (A2). Assume in addition that for all $i \in \{1, \dots, d\}$, $\int |y|\; q_i(y)\mu(dy) < \infty$. Then, if the RSM is $V$-uniformly ergodic, there exists $s_i > 0$, $i \in \{1, \dots, d\}$ such that*

$$\int_{\mathbb{R}^d} e^{\sum_{i=1}^d s_i |x_i|} p(x)\mu_d(dx) < \infty.$$

*Proof.* By Theorem 16.3.2 in Meyn and Tweedie (1993), it is known that if $P_{\mathrm{RS}}$ is $V$-uniformly ergodic, there exists $\beta > 1$ such that

$$\int \beta^{\mathbb{E}_x[\sigma_C]} p(x)\mu_d(dx) < \infty,$$

where $\sigma_C$ is the hitting time on some small set $C$ such that $\pi(C) > 0$, and $\mathbb{E}_x$ is the expectation with respect to the chain $\{X_k\}$ starting from $X_0 = x$ with transition kernel $P_{\mathrm{RS}}$. In addition, $C$ can be assumed to be on the form $C := [-c, c]^d$.

Denote by $\{I_{n+1}\}$ the i.i.d. sequence of proposed increments and define, for $k \in \{1, \dots, d\}$,

$J^{(k)}_{n+1} := \mathrm{sign}(X_n \cdot e_k) \quad I_{n+1} \cdot e_k \quad \mathbb{1}_{\mathrm{sign}(X_n \cdot e_k) \quad I_{n+1} \cdot e_k} < 0$, $n \geq 0$. Observe that $P(J^{(k)}_{n+1} \leq v | X_n = x) = 1$ for any $v \geq 0$ whereas for any $v < 0$,

$$\mathbb{P}(J^{(k)}_{n+1} \leq v | X_n = x) = \mathbb{P}(\mathrm{sign}(x \cdot e_k) \quad I_{n+1} \cdot e_k \leq v).$$

Since $q_k d\mu$ is a symmetric distribution, it follows that $J^{(k)}_{n+1}$ and $X_n$ are independant, $\{J^{(k)}_{n+1}\}$ is an i.i.d. sequence and

$$\mathbb{E}[|J^{(k)}_{n+1}|] = \int_0^\infty t \ q_k(t) \ dt =: \gamma_k. \tag{24}$$

Define, for $k \in \{1, \ldots, d\}$, the $\mathbb{R}$-valued random walk $W^{(k)}$ by

$$W^{(k)}_0 := |X_0 \cdot e_k|, \qquad W^{(k)}_{n+1} = W^{(k)}_n + J_{n+1}, \quad n \geq 0.$$

We prove by induction that $W^{(k)}_n \leq |X_n \cdot e_k|$. This is true for $n = 0$. Assume that the property holds for $n$. Then

$$|X_{n+1} \cdot e_k| = ||X_n \cdot e_k| + \mathrm{sign}(X_n \cdot e_k) \quad I_{n+1} \cdot e_k| \geq |X_n \cdot e_k| + J^{(k)}_{n+1}.$$

Using the induction assumption, it follows that $|X_{n+1} \cdot e_k| \geq W^{(k)}_{n+1}$, which concludes the proof. Thus for all $x = (x_1, \ldots, x_d)$, $k \in \{1, \ldots, d\}$,

$$\mathbb{E}_{|x_k|}\left[\sigma^{(k)}_c\right] \leq \mathbb{E}_x\left[\sigma_C\right] \tag{25}$$

where $\mathbb{E}_u\left[\sigma^{(k)}_c\right]$ is the mean of the first hitting time on $(-\infty, c]$ of the random walk $W^{(k)}$ started from $u$. Finally, since $\mathbb{E}_x\left[\sigma_C\right] < \infty$ for $p\mu_d$-a.a. $x$, the optional stoppping theorem for martingale and the monotone convergence theorem imply that for all $u > c$,

$$\mathbb{E}_u\left[\sigma^{(k)}_c\right] \quad \gamma_k = u - \mathbb{E}_u\left[W^{(k)}_{\sigma^{(k)}_c}\right] \geq u - c, \tag{26}$$

where $\gamma_k$ is given by (24). Combining (25) and (26) gives

$$\mathbb{E}_x\left[\sigma_C\right] \geq \sup_{k \in \{1, \ldots, d\}} \gamma_k^{-1} \ (|x_k| - c) \vee 0 \geq \frac{1}{d} \sum_{k=1}^{d} \gamma_k^{-1} \ (|x_k| - c) \vee 0.$$

The result follows. □

This result is of interest because it is not straightforward to show that (A3′) implies the existence of exponential moments.

**Remark 9.** The previous proposition still holds if instead of (A1), it is assumed that the target density $p$ is positive and continuous on an unbounded subset $\mathcal{X} \subset \mathbb{R}^d$.

3.3.1. *Example 10.* The so-called Normal-Inverse Gamma model appears as the posterior distribution in one of the simplest two-dimensional Bayesian analyses of an i.i.d. Gaussian model. Although there are many other ways to simulate from this distribution without having to resort to MCMC, it provides a fruitful testing ground for simple algorithms (see Roberts and Tweedie (2001)). The model assumes an i.i.d. collection of data $\{y_1, \ldots, y_n\}$ from the $N(\mu, \tau^{-1})$ distribution, with unknown mean $\mu$ and precision $\tau$ (so that the variance of the Gaussian is just $\tau^{-1}$). The distribution $p$ that we will consider is the joint posterior density for these parameters represented by

$$p(\mu, \tau) \propto \tau^{(n+1)/2} \prod_{i=1}^{n} \exp\{-\tau(y_i - \mu)^2/2\} \qquad \mu \in \mathbb{R}, \tau > 0. \tag{27}$$

This posterior is obtained if we assume a flat prior on $\mu$ (that is, the prior is an improper distribution with constant density on $\mathbb{R}$) and the prior density $\tau^{-1/2}$ on $\mathbb{R}_+$ on the precision. The contours of an example of this distribution is given in Figure 7. Notice how the contours are stretched into long thin ridges for small values of the precision parameter $\tau$.

For this target density, the random scan Metropolis kernel is not be geometrically ergodic. Indeed, set $\bar{y} := n^{-1} \sum_{k=1}^{n} y_k$ and $S^2 := \sum_{k=1}^{n} (y_k - \bar{y})^2$. Then

$$p(\mu, \tau) \propto \tau^{(n+1)/2} \exp\left(-\tau \left(S^2 + n(\mu - \bar{y})^2\right)/2\right).$$

Using the equality $b^a \int_{\mathbb{R}_+} x^{a-1} \exp(-bx) dx = \int_{\mathbb{R}_+} x^{a-1} \exp(-x) dx$, $a, b > 0$, we have, for $s_1 > 0, 0 < s_2 < S^2/2$ that

$$\int_{\mathbb{R}} d\mu \int_0^\infty d\tau \ \exp(s_1 \mu + s_2 \tau) p(\mu, \tau) \propto \int_{\mathbb{R}} \exp(s_1 \mu) \left(S^2 + n(\mu - \bar{y})^2 - 2 s_2\right)^{-(n+3)/2} d\mu,$$

which shows that the target density does not have exponential moments. Hence, Proposition 7 is not satisfied and the RSM algorithm cannot be geometrically ergodic.

The same conclusion can also be reached by using the notion of capacitance of a Markov chain. Recall that for a given Markov chain $P$ with stationary distribution $\pi$, the conductance $c(A)$ of a measurable set $A$ is given by

$$c(A) := \int_A \frac{\pi(dx)}{\pi(A)} P(x, A^c),$$

and the capacitance of the Markov chain is defined as

$$\kappa := \inf_{A, \pi(A) \le 1/2} c(A).$$

The following result, which applies generally to a large number of applications of the Metropolis-within-Gibbs algorithm, is proved in Roberts and Tweedie (2001), Theorem 9.7.1 (see also Roberts and Rosenthal (1998), Lemma 11).

**Theorem 8.** *Suppose that $\pi$ is a $d-$dimensional distribution, and for each $i$, $P_i$ is a Markov chain which is reversible with respect to $\pi$, and updates just the $i$-th coordinate. Consider running a random scan of the $P_i$'s, that is a chain $P$ with*

$$P = \frac{P_1 + P_2 + \ldots + P_d}{d} .$$

*Suppose that for some component $i$, $P_i$ is a random walk Metropolis algorithm with fixed increment proposal density $q$, and that*

$$\lim_{K\to\infty} \frac{\log \pi(X_i \in (K,\infty))}{K} = 0 . \tag{28}$$

*Then*

$$\liminf_{K\to\infty} c(\{X_i \in (K,\infty\}) = 0 \ , \tag{29}$$

*and consequently $\kappa = 0$, so that $P$ is not geometrically ergodic.*

We prove that for the present model, $\mu$ has an heavy tailed distribution (decreasing as $|\mu|^{-(n+3)}$ in the tails) so that (28) is verified with $i = 1$. Indeed

$$\pi(\{\mu \geq K\}) = \int_{[K,\infty)} d\mu \int_{\mathbb{R}_+} d\tau \ p(\mu,\tau) \propto \int_{[K,\infty)} d\mu \left(S^2/2 + n(\mu - \bar{y})^2/2\right)^{-(n+3)/2} .$$

Hence, $\lim_{K\to\infty} \log \pi(\mu \geq K)/K \propto \lim_{K\to\infty} \log K/K = 0$, and (28) is verified.

Now, for this target density, one of the conditions (A3), (A3') must fail. Consider the sequence $x^j = (a^2, 1/j)$, which tends to $\partial\mathcal{X}$ as $j$ tends to infinity. For any $z > 0$,

$$\log p(x^j) - \log p(x^j - ze_1) = -z \sum_{k=1}^{n} (2y_k - 2a^2 + z)/(2j) \to 0 \quad \text{as } j \to \infty.$$

In addition, for large $j$, $p(x^j)/p(x^j - ze_2) = 0$ (see Remark 4) and

$$\log p(x^j + ze_2) - \log p(x^j) = \frac{n+1}{2} \log(1 + jz) - z \sum_{k=1}^{n} (y_k - a^2)^2/2 \to +\infty \quad \text{as } j \to \infty.$$

Hence, neither (A3) nor (A3') can hold.

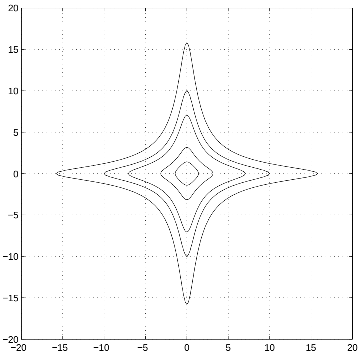


FIGURE 1. A contour plot of the surface of the density $p(x_1, x_2) = \exp\left(-(x_1^2 + x_1^2 x_2^2 + x_2^2)\right)$.

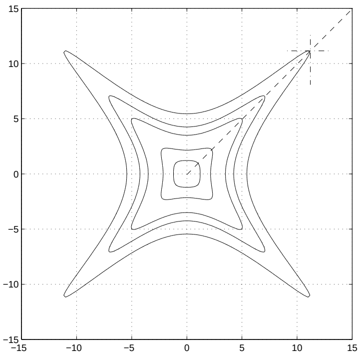


FIGURE 2. A contour plot of the surface of the density $p(x_1, x_2) = \exp\left(-(x_1^2 + (x_1^2 - x_2^2)^2/4 + x_2^2)\right)$.

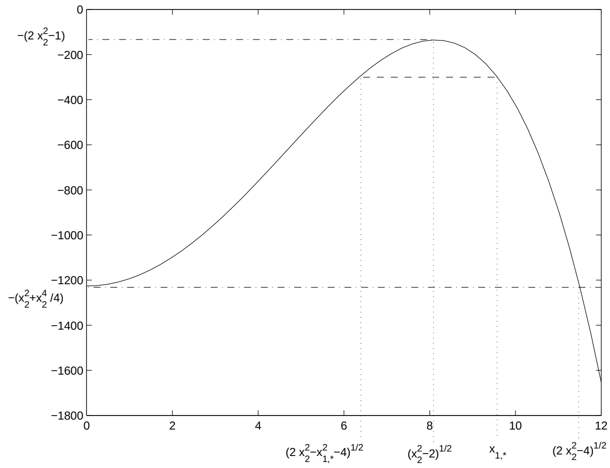


FIGURE 3. For some fixed $|x_2| > \sqrt{2}$, this is the plot of the function $L_{x_2}$ : $\mathbb{R}_+ \to \mathbb{R}_-, x_1 \mapsto -(x_1^2 + (x_1^2 - x_2{}^2)^2/4 + x_2{}^2)$. For all $x_{1,*} > 0$, the set $\{x_1 \in \mathbb{R}, L_{x_2}(x_1) \geq L_{x_2}(x_{1,*})\}$ is the acceptance region on $\mathbb{R}_+$ in the $e_1$-direction at the point $(x_{1,*}, x_2)$, that is $\mathcal{A}((x_{1,*}, x_2), 1) \cap \mathbb{R}_+$. The set $\mathcal{A}((x_{1,*}, x_2), 1) \cap \mathbb{R}_-$ is found easily by symmetry.

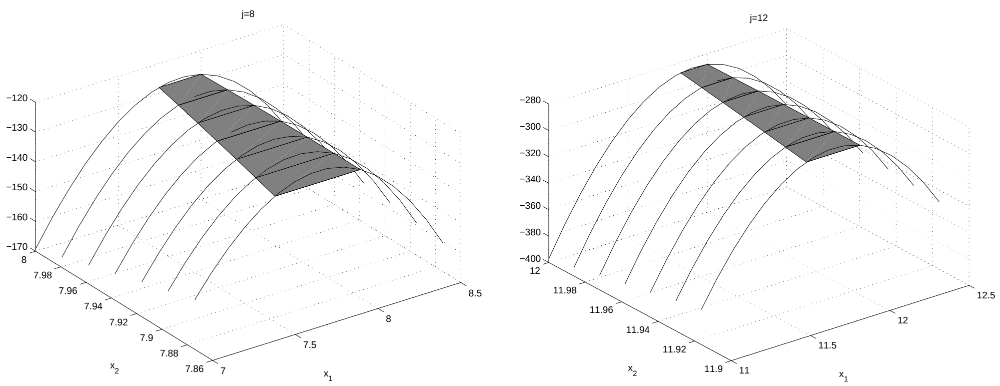


FIGURE 4. We illustrate that for all $x \in \mathcal{X}_j$, the Lebesgue measure of the acceptance region in the $e_i$-direction $\mathcal{A}(x, i)$ goes to zero as $j$ goes to infinity. As the target density $p$ is symmetric $(p(-x_1, x_2) = p(x_1, x_2) = p(x_2, x_1))$, we consider the case $i = 1$ and the figures on $x_1 > 0$. For $\lambda = 2$ and two different values of $j$, we plot **(a)** $\log p(x)$ for different values of $x \in \mathcal{X}_j$ [*solid lines*], and **(b)** the translated acceptance region $\mathcal{A}((j, x_2), 1) + j$ [*grey surface*].

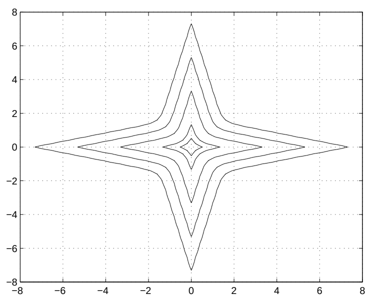

FIGURE 5. A contour plot of the surface of the density $p(x_1, x_2) = 0.5\exp(-(|x_1| + 4|x_2|)) + 0.5\exp(-(4|x_1| + |x_2|))$.

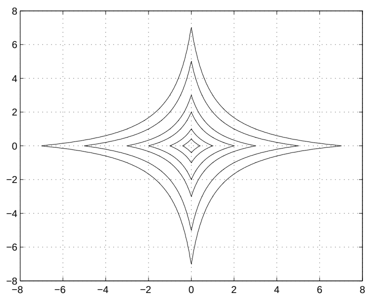

FIGURE 6. A contour plot of the surface of the density $p(x_1, x_2) = \exp(-(|x_1| + |x_1||x_2| + |x_2|))$.

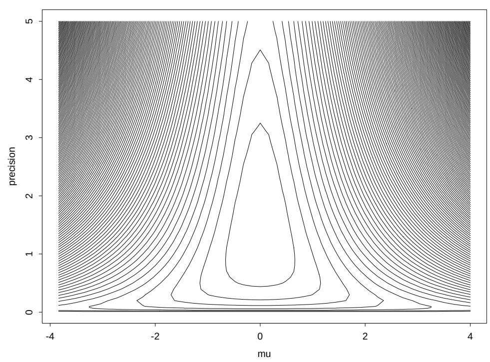


FIGURE 7. A contour plot of the surface of a Normal-Inverse Gamma density function.

## References


Chan, K. and Geyer, C. (1994). Discussion of Tierney. *Ann. Statist.*, 22:1747–1758.

Chan, K. and Ledolter, J. (1995). Monte Carlo EM estimation for time series models involving counts. *J. Amer. Statist. Assoc.*, 90(429):242–252.

Douc, R. Moulines, E. and Rosenthal, J.S. (2001) Quantitative convergence rates for inhomogeneous Markov chains. *In preparation.*

Fort, G. and Moulines, E. (2000a). Computable bounds for subgeometrical and geometrical ergodicity. Submitted for publication in Stochastic Processes Appl.

Fort, G. and Moulines, E. (2000b). Convergence of the Monte-Carlo EM for curved exponential families. Submitted for publication in Ann. Statist.

Fort, G. and Moulines, E. (2000c). V-subgeometric ergodicity for a Hastings-Metropolis algorithm. *Statist. Probab. Lett.*, 49(4):401–410.

Hobert, J. and Geyer, C. (1998). Geometric ergodicity of Gibbs and Block Gibbs Samplers for a Hierarchical random Effects Models. *J. Multivariate Anal.*, 67(2):414–430.

Jarner, S. and Hansen, E. (2000). Geometric ergodicity of Metropolis algorithms. *Stochastic Processes Appl.*, 85(2):341–361.

Laugwitz, D. (1965). *Differential and Riemannian geometry.* Academic Press.

Mengersen, K. and Tweedie, R. (1996). Rates of convergence of the Hastings and Metropolis algorithms. *Ann. Statist.*, 24(1):101–121.

Meyn, S. and Tweedie, R. (1994). Computable bounds for geometric convergence rates of Markov chains. *Ann. Appl. Probab.*, 4(4):981–1011.

Meyn, S. P. and Tweedie, R. L. (1993). *Markov chains and stochastic stability.* Springer-Verlag London Ltd., London.

Robert, C. and Casella, G. (1999). *Monte Carlo statistical methods.* Springer-Verlag, New York.

Roberts, G. and Rosenthal, J. (1997). Geometric ergodicity and hybrid Markov chains. *Electronic Comm. Prob.*, 2:13–25.

Roberts, G. and Rosenthal, J. (1998). Two convergence properties of hybrid samplers. *Ann. Appl. Probab.*, 8(2):397–407.

Roberts, G. and Tweedie, R. Understanding MCMC. To be published.

Roberts, G. and Tweedie, R. (1996). Geometric convergence and central limit theorems for multidimensional Hastings and Metropolis algorithms. *Biometrika*, 83(1):95–110.

Roberts, G. and Tweedie, R. (1999). Bounds on regeneration times and convergence rates for Markov Chains. *Stochastic Processes Appl.*, 80(2):211–229.

Rosenthal, J. S. (1995). Minorization conditions and convergence rates for Markov chain Monte Carlo. *J. Amer. Statist. Assoc.*, 90(430):558–566.

Sahu, S. and Roberts, G. (1999). On Convergence of the EM algorithm and the Gibbs sampler. *Statistics and Computing*, 9:55–64.

Smith, A. and Roberts, G. (1993). Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods (with discussion). *J. R. Stat. Soc. Ser. B*, 55(1):3–24.

Tierney, L. (1994). Markov chains for exploring posterior distributions. *Ann. Statist.*, 22(4):1701–1762. With discussion and a rejoinder by the author.

Zeger, S. L. (1988). A regression model for time series of counts. *Biometrika*, 75(4):621–629.