Bernoulli **9**(4), 2003, 559–578

Necessary conditions for geometric and polynomial ergodicity of random-walk-type Markov chains

SØREN F. JARNER[1] and RICHARD L. TWEEDIE*

[1]*Quantitative Research, Fixed Income, The Danish Labour Market Supplementary Pension, Kongens Vænge 8, DK-3400 Hillerød, Denmark*

We give necessary conditions for geometric and polynomial convergence rates of random-walk-type Markov chains to stationarity in terms of existence of exponential and polynomial moments of the invariant distribution and the Markov transition kernel. These results complement the use of Foster–Lyapunov drift conditions for establishing geometric and polynomial ergodicity. For polynomially ergodic Markov chains, the results allow us to derive exact rates of convergence and exact relations between the moments of the invariant distribution and the Markov transition kernel. In an application to Markov chain Monte Carlo we derive tight rates of convergence for symmetric random walk Metropolis algorithms and Langevin algorithms with polynomial target densities.

Keywords: geometric and polynomial moments; Markov chains; Metropolis algorithms

1. Introduction

Let $\mathbf{X} = (X_0, X_1, \ldots)$ be a discrete-time Markov chain on the d-dimensional Euclidean space $E = \mathbb{R}^d$ equipped with its Borel σ-field $\mathcal{B}$. We assume throughout that the chain is ψ-irreducible, aperiodic and positive recurrent (see Meyn and Tweedie 1993). Let π denote the (necessarily unique) invariant distribution. Further, let $P(x, \cdot)$ denote the Markov transition kernel and let $P^n(x, \cdot)$, $n \in \mathbb{N}_0$, denote the n-step kernel,

$$P^n(x, A) = \mathsf{P}_x(X_n \in A) \qquad (x \in E, A \in \mathcal{B}),$$

where P_x is the conditional distribution of the chain given $X_0 \equiv x$. The corresponding expectation operator will be denoted E_x. For any function V we write $PV(x)$ for the function $\int V(y)P(x, \mathrm{d}y)$ and for any signed measure μ we write $\mu(V)$ for $\int V(y)\mu(\mathrm{d}y)$.

Following the terminology of Meyn and Tweedie (1993), a set $C \in \mathcal{B}$ is called small if there exist $n > 0$, $\delta > 0$ and a probability measure ν such that $P^n(x, \cdot) \geqslant \delta\nu(\cdot)$ for all x in C. Under our assumptions of ψ-irreducibility and aperiodicity this is the same as C being petite (see Meyn and Tweedie 1993).

*Richard Tweedie of the University of Minnesota passed away on 7 June 2001, shortly before this paper was submitted. The first author wishes to express his gratitude to the man who has been a great source of inspiration for so many of us.

In this paper we consider geometrically and polynomially ergodic Markov chains, i.e. Markov chains for which there exists a small set C, a function $f : E \to [1, \infty)$ and a geometric or polynomial rate function $r(n)$ such that

$$\sup_{x \in C} \mathsf{E}_x \left[\sum_{k=0}^{\tau_C - 1} r(k) f(X_k) \right] < \infty, \tag{1}$$

where $\tau_C = \inf\{n \geqslant 1 : X_n \in C\}$ is the first return time of the chain to C. In the polynomial case (1) implies that for π-almost all x,

$$r(n) \| P^n(x, \cdot) - \pi \|_f \to 0, \qquad n \to 0, \tag{2}$$

where the f-norm is defined for a signed measure μ as $\|\mu\|_f = \sup_{|g| \leqslant f} |\mu(g)|$. In the geometric case the rate of convergence in (2) is exponential but generally of a lower order than $r(n)$.

The most common way of establishing geometric and subgeometric ergodicity of Markov chains on general state spaces is by verifying an associated Foster–Lyapunov type drift condition (see Nummelin and Tuominen 1982; Tweedie 1983; Nummelin 1984; Meyn and Tweedie 1993; Tuominen and Tweedie 1994; Jarner and Roberts 2002). In the Markov chain Monte Carlo (MCMC) context this approach has been successfully applied a number of times to derive sufficient conditions on π for geometric and subgeometric ergodicity of the Gibbs sampler and other Metropolis–Hastings algorithms (see Chan 1993; Mengersen and Tweedie 1996; Roberts and Tweedie 1996a, 1996b; Fort and Moulines 2000; Jarner and Hansen 2000; Jarner and Roberts 2001). However, it is generally difficult to show that the assumed conditions on π are also necessary or, alternatively, that the rate of convergence is best possible, and there has been only few results in this direction.

In this paper we show that for random-walk-type Markov chains geometric and polynomial ergodicity imply that π and P have certain exponential and polynomial moments. In the polymonially ergodic case, these results can be used in combination with Foster–Lyapunov type drift conditions to derive exact convergence rates and moments of π. We illustrate this by deriving an exact relation between the moments of the invariant distribution and the increment distribution of a random walk on a half line, and to derive tight rates of convergence for two MCMC algorithms, the random walk Metropolis algorithm and the Langevin algorithm with polynomial target densities.

We say that $\mathbf{X}$ is of random walk type if, for every $\epsilon > 0$, there exists $K > 0$ such that $P(x, B(x, K)) > 1 - \epsilon$ for all x, where $B(x, K) = \{y : |y - x| < K\}$ denotes the open ball with centre x and radius K with respect to Euclidean distance $|\cdot|$. This is equivalent to the family of increment distributions $\{P(x, \cdot) - x : x \in E\}$ being tight.

Random-walk-type Markov chains occur frequently in, for example, MCMC methods and queuing theory. Examples of random-walk-type Markov chains include chains of the form

$$X_{n+1} = X_n + f(X_n) + g(X_n) W_{n+1}, \qquad n \in \mathbb{N}_0, \tag{3}$$

where (W_n) is an independent and identically distributed (i.i.d.) sequence of random variables and $\sup_x |f(x)| < \infty$ and $\sup_x |g(x)| < \infty$. Also, the random walk on $[0, \infty)$ given by

$$X_{n+1} = (X_n + W_{n+1})^+, \qquad n \in \mathbb{N}_0, \tag{4}$$

is of random walk type. Examples of Markov chains which are not of random walk type include AR(1) processes,

$$X_{n+1} = \phi X_n + W_{n+1}, \qquad n \in \mathbb{N}_0, \tag{5}$$

and other chains with unbounded drift terms. However, AR(1) processes and, more generally, chains with multiplicative drift structure behave like random-walk-type Markov chains on the log scale, and necessary conditions for geometric and polynomial ergodicity of these chains can be derived using ideas similar to those presented here. This is done in Jarner and Tweedie (2002) for a Markov chain associated with the mean of a Dirichlet process. This chain is essentially an AR(1) process with a stochastic coefficient.

In Section 2 we show that a random-walk-type Markov chain can be geometrically ergodic only if π has exponential moments. In the special case of a symmetric random walk Metropolis algorithm with increment proposal distribution with finite first absolute moment this has previously been proved in Mengersen and Tweedie (1996) and Jarner and Hansen (2000) by using Wald's equation to bound the mean return time to the centre of the space. The approach taken here, however, is more general and involves controlling only a fraction of the probability mass for which more detailed behaviour of the sample paths can be obtained in order to provide sample path bounds on the return times.

In Section 3 this idea is taken further to show the existence of polynomial moments of π when $\mathbf{X}$ is polynomially ergodic. We say that $\mathbf{X}$ is *polynomially ergodic of order* (α, β), where $\alpha, \beta \geqslant 0$, if (1) holds with $r(k) = (k+1)^\beta$ and $f(x) = (|x|+1)^\alpha$, and we show that this implies

$$\int_{\mathbb{R}^d} |x|^{\alpha+\eta\beta} \pi(\mathrm{d}x) < \infty, \tag{6}$$

where $0 < \eta \leqslant 2$ depends on the tail behaviour and drift of the family of increment distributions $\{P(x, \cdot) - x\}$. The case $\eta = 1$ corresponds to uniform integrability of the increment distributions, and $\eta = 2$ corresponds to the family of increment distributions having uniformly bounded variance and drift towards the centre of the space of order at most $|x|^{-1}$. As corollaries it is shown that for any $\gamma \geqslant 0$ with $\pi(|x|^\gamma) < \infty$,

$$\int_{\mathbb{R}^d} |x|^{\gamma+\eta} P(z, \mathrm{d}x) < \infty \qquad \text{for } \pi\text{-almost all } z. \tag{7}$$

This is used to derive an exact relation between the moments of the invariant distribution and the increment distribution of the random walk in (4).

Section 4 uses the results above to show that the polynomial rates of convergence of symmetric random walk Metropolis algorithms and Langevin algorithms with polynomial target densities found in Jarner and Roberts (2001) are tight. Knowing the exact rates instead of only lower bounds is quite rare but has obvious advantages when comparing different algorithms.

The technique used in this paper of controlling only part of the probability mass in order to derive lower bounds on expectations of functionals of the Markov chain seems quite general and suitable for providing necessary conditions also for more general (f, r)-ergodic chains than those considered here. Indeed, all chains for which small sets are bounded seem to be subject to the techniques presented here. This includes all chains that make local jumps, while chains that make global jumps, such as uniformly ergodic chains, fall outside of this category.

2. Geometric ergodicity

The Markov chain $\mathbf{X}$ is called geometrically ergodic if (1) holds for some $f \geqslant 1$ and $r(n) = \rho^n$ for some $\rho > 1$. By Theorem 15.0.1 of Meyn and Tweedie (1993) an equivalent condition is that there exists a small set C, constants $\lambda < 1$ and $b < \infty$ and a function $V \geqslant 1$ finite for at least one $x_0 \in E$ satisfying

$$PV \leqslant \lambda V + b 1_C. \tag{8}$$

Showing the Foster–Lyapunov drift condition (8) is often the easiest way to prove geometric ergodicity. Note, that by Theorem 14.3.7 of Meyn and Tweedie (1993) any function V satisfying (8) has finite expectation with respect to π. In particular, V is finite π-almost everywhere. We will use these properties below.

The next theorem shows that for random-walk-type Markov chains as defined in the Introduction geometric ergodicity implies the existence of exponential moments of π. For symmetric random walk Metropolis algorithms this has previously been proved in Mengersen and Tweedie (1996) and Jarner and Hansen (2000) under the additional assumption that the family of increment distributions has finite first absolute moment, but this assumption is not needed for our approach.

We need the following simple lemma from Jarner and Hansen (2000).

Lemma 2.1. *If* $\mathbf{X}$ *is a random-walk-type Markov chain then every small set is bounded.*

Following the terminology of Meyn and Tweedie (1993) a set $A \in \mathcal{B}$ is called an f-Kendall set for a function $f \geqslant 1$ if there exists $\kappa > 1$ such that

$$\sup_{x \in A} \mathsf{E}_x \left[\sum_{k=0}^{\tau_A - 1} \kappa^k f(X_k) \right] < \infty.$$

Theorem 2.2. *Let* $\mathbf{X}$ *be a random-walk-type Markov chain. If* $\mathbf{X}$ *is geometrically ergodic then there exists* $s > 0$ *such that*

$$\int_{\mathbb{R}^d} \mathrm{e}^{s|x|} \pi(\mathrm{d}x) < \infty. \tag{9}$$

Proof. Since $\mathbf{X}$ is geometrically ergodic there exist small set C, constants $\lambda < 1$ and $b < \infty$ and function $V \geqslant 1$ finite π-almost everywhere such that (8) holds. Choose M so large that $A = \{V \leqslant M\}$ has positive π-measure. By Theorems 15.2.6 and 15.2.1 of Meyn and Tweedie

(1993) A is then a small V-Kendall set, and by Theorem 15.2.4 of Meyn and Tweedie (1993) there then exist $\kappa > 1$, $\tilde{\lambda} < 1$ and $\tilde{b} < \infty$ such that

$$P\tilde{V} \leqslant \tilde{\lambda}\tilde{V} + \tilde{b}1_A, \tag{10}$$

where

$$\tilde{V}(x) = \begin{cases} V(x) & \text{for } x \in A, \\ \mathsf{E}_x[\sum_{k=0}^{\tau_A} \kappa^\kappa V(X_k)] & \text{for } x \in A^c. \end{cases}$$

Since $\tilde{V}$ satisfies the drift condition (10) we have $\pi(\tilde{V}) < \infty$, and (9) thus follows if we can find an exponential lower bound on $\tilde{V}(x)$ for $|x|$ sufficiently large.

Choose R so large that $A \subset B(0, R)$; this can be done because A is a small set and hence bounded by Lemma 2.1. Since $V \geqslant 1$ we have the lower bound $\tilde{V}(x) \geqslant \mathsf{E}_x[\kappa^{\tau_A}]$ for $x \in A^c$ and thus in particular for $|x| \geqslant R$.

Choose $\epsilon > 0$ so small that $\kappa(1-\epsilon) > 1$ and then, using the random walk structure, choose K such that $P(x, B(x, K)) > 1 - \epsilon$ for all x. For any real number z let $\lceil z \rceil$ denote the smallest integer equal to or larger than z. For $|x| \geqslant R$ we then have

$$\mathsf{E}_x[\kappa^{\tau_A}] \geqslant (\kappa(1-\epsilon))^w, \tag{11}$$

where $w = \lceil (|x| - R)/K \rceil$, because $(1-\epsilon)^w$ is a lower bound for the probability that the next w jumps are of length at most K and on this event $\tau_A \geqslant w$. It follows that we can find $s > 0$ and $c > 0$ such that $\tilde{V}(x) \geqslant \mathsf{E}_x[\kappa^{\tau_A}] \geqslant c\mathrm{e}^{s|x|}$ for $|x| \geqslant R$, and we are done. $\square$

Note that by stationarity we have for any function $f \geqslant 0$,

$$\int_{\mathbb{R}^d} f(x)\pi(\mathrm{d}x) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(x)P(z, \mathrm{d}x)\pi(\mathrm{d}z).$$

In particular, (9) implies that

$$\int_{\mathbb{R}^d} \mathrm{e}^{s|x|} P(z, \mathrm{d}x) < \infty \qquad \text{for } \pi\text{-almost all } z.$$

3. Polynomial ergodicity

Recall that the Markov chain $\mathbf{X}$ is polynomially ergodic of order (α, β), where $\alpha, \beta \geqslant 0$, if there exists a small set C such that

$$\sup_{x \in C} \mathsf{E}_x\left[\sum_{k=0}^{\tau_C - 1} (k+1)^\beta (|X_k| + 1)^\alpha\right] < \infty. \tag{12}$$

By Theorem 14.0.1 of Meyn and Tweedie (1993) this implies, in particular, that $\mathbf{X}$ is positive recurrent with invariant distribution π and $\pi(|x|^\alpha) < \infty$.

3.1. Polynomial ergodicity and polynomial moments

For random-walk-type Markov chains we show below under varying additional conditions that polynomial ergodicity implies polynomial moments of π. The results take the general form that polynomial ergodicity of order (α, β) implies that

$$\int_{\mathbb{R}^d} |x|^{\alpha+\eta\beta}\pi(\mathrm{d}x) < \infty, \tag{13}$$

where $0 < \eta \leqslant 2$ depends on the heaviness of the tails and the drift of the family of increment distributions $\{P(x, \cdot) - x\}$.

Let h be a non-increasing function $h : [0, \infty) \to [0, 1]$ such that, for all $x \in \mathbb{R}^d$ and all $y \geqslant 0$,

$$P(x, B(x, y)^{\mathrm{c}}) \leqslant h(y). \tag{14}$$

Since we can always use $h \equiv 1$, such a function exists for any Markov chain. The Markov chain is of random walk type if and only if there exists h with $h(y) \to 0$ as $y \to \infty$ such that (14) holds.

We first consider conditions in terms of how quickly h tends to zero. Theorems 3.2 and 3.3 show that if h is integrable then (13) holds with $\eta = 1$, while if h tends to zero at a non-integrable polynomial rate we obtain (13) with $0 < \eta < 1$. In Section 3.2 we assume more structure and show in Theorems 3.6 and 3.7 that (13) holds with $1 < \eta \leqslant 2$ for random-walk-type Markov chains where the family of increment distributions has uniformly bounded moments of order η and drift to the centre of the space of order at most $|x|^{1-\eta}$.

For any sequence r we define the sequence Δr by

$$\begin{aligned}\Delta r(0) &= r(0),\\ \Delta r(k) &= r(k) - r(k-1), \qquad \text{for } k = 1, 2, \ldots .\end{aligned}$$

From Theorems 2.1 and 2.3 of Tuominen and Tweedie (1994) and the trivial bound $|x|^\alpha \leqslant (|x|+1)^\alpha$ we obtain the following lemma on which all subsequent results rely.

Lemma 3.1. *If* $\mathbf{X}$ *is polynomially ergodic of order* (α, β), *then there exists a small set* C *such that*

$$\int_{\mathbb{R}^d} \mathsf{E}_x\left[\sum_{k=0}^{\tau_C - 1} \Delta r(k)|X_k|^\alpha\right]\pi(\mathrm{d}x) < \infty, \tag{15}$$

where $r(k) = (k+1)^\beta$.

Note that for $\beta = 0$, (15) reduces to $\pi(|x|^\alpha) < \infty$ since in this case $\Delta r(k) = 0$ for $k > 0$.

Theorem 3.2. *Assume* $\mathbf{X}$ *is of random walk type and that* (14) *holds with* $\int_0^\infty h(y)\mathrm{d}y < \infty$. *If* $\mathbf{X}$ *is polynomially ergodic of order* (α, β), *then*

$$\int_{\mathbb{R}^d} |x|^{\alpha+\beta} \pi(\mathrm{d}x) < \infty. \tag{16}$$

Proof. Assumption (14) with $\int_0^\infty h(y)\mathrm{d}y < \infty$ implies that there exists a sequence of i.i.d. random variables $Y_n > 0$ with finite mean $\mu = \mathsf{E}(Y_n)$ such that, for all $x \in \mathbb{R}^d$ and all $y \geqslant 0$,

$$P(x, B(x, y)^{\mathrm{c}}) \leqslant \mathsf{P}(Y_n \geqslant y). \tag{17}$$

By the weak law of large numbers we have, for any $\epsilon > 0$,

$$\mathsf{P}(S_n < (\mu + \epsilon)n) \to 1, \qquad n \to \infty,$$

where $S_n = Y_1 + \ldots + Y_n$. Hence we can choose N so large that, for $n \geqslant N$,

$$\mathsf{P}(S_n < 2\mu n) \geqslant \tfrac{1}{2}.$$

Using (17), this shows by a stochastic comparison argument that, for all $x \in \mathbb{R}^d$ and all $n \geqslant N$,

$$\mathsf{P}_x(X_k \in B(x, 2\mu n) \text{ for } k = 0, \ldots, n) \geqslant \tfrac{1}{2}. \tag{18}$$

For $|x|$ so large that $|x|/4\mu \geqslant N$ it follows from (18) with $n = \lfloor |x|/4\mu \rfloor \geqslant N$ that

$$\mathsf{P}_x(X_k \in B(x, |x|/2) \text{ for } k = 0, \ldots, \lfloor |x|/4\mu \rfloor) \geqslant \tfrac{1}{2}. \tag{19}$$

By Lemmas 3.1 and 2.1, (15) holds for a small and hence bounded set C. For $|x|$ so large that $B(x, |x|/2) \subset C^{\mathrm{c}}$ we have on the above event

$$\begin{aligned} &\tau_C - 1 \geqslant \lfloor |x|/4\mu \rfloor, \\ &|X_k|^\alpha \geqslant (|x|/2)^\alpha, \qquad \text{for } k = 0, \ldots, \lfloor |x|/4\mu \rfloor, \end{aligned}$$

and thus also

$$\sum_{k=0}^{\tau_C - 1} \Delta r(k)|X_k|^\alpha \geqslant \sum_{k=0}^{\lfloor |x|/4\mu \rfloor} \Delta r(k)|X_k|^\alpha \geqslant \frac{|x|^\alpha}{2^\alpha} r(\lfloor |x|/4\mu \rfloor) \geqslant \frac{|x|^{\alpha+\beta}}{2^\alpha (4\mu)^\beta},$$

where $r(k) = (k+1)^\beta$. For $|x|$ sufficiently large this event has probability at least $\frac{1}{2}$ by (19) and hence, for $|x|$ sufficiently large,

$$\mathsf{E}_x\left[\sum_{k=0}^{\tau_C - 1} \Delta r(k)|X_k|^\alpha\right] \geqslant \frac{|x|^{\alpha+\beta}}{2^{\alpha+1}(4\mu)^\beta},$$

and (16) now follows from (15). □

When the dominating function h decays to zero at a polynomial rate but is not integrable we can use a stable law limit result instead of the weak law of large numbers to get the following result.

Theorem 3.3. *Assume* $\mathbf{X}$ *is of random walk type and that there exists* $0 < \eta < 1$ *and constant*

$c > 0$ such that (14) *holds with $h(y) = cy^{-\eta}$ for y sufficiently large. If* $\mathbf{X}$ *is polynomially ergodic of order* (α, β)*, then*

$$\int_{\mathbb{R}^d} |x|^{\alpha+\eta\beta}\pi(\mathrm{d}x) < \infty. \tag{20}$$

Proof. By (14) and the assumption on h there exists a sequence of i.i.d. random variables $Y_n > 0$ with distribution function F satisfying $1 - F(y) = cy^{-\eta}$ for y sufficiently large such that, for all $x \in \mathbb{R}^d$ and all $y \geqslant 0$,

$$P(x, B(x, y)^{\mathrm{c}}) \leqslant \mathsf{P}(Y_n \geqslant y). \tag{21}$$

From Sections XVII.5 and XIII.6 of Feller (1971) it follows that there exists $\delta > 0$ such that, for all $z > 0$,

$$\mathsf{P}(S_n \leqslant z\delta n^{1/\eta}) \to G_\eta(z), \qquad n \to \infty,$$

where $S_n = Y_1 + \ldots + Y_n$ and G_η is the cdf of a stable law with index η. In particular, we can find $\nu > 0$ and N such that, for $n \geqslant N$,

$$\mathsf{P}(S_n \leqslant \nu n^{1/\eta}) \geqslant \tfrac{1}{2}.$$

Thus by (21) and a stochastic comparison argument we have, for all $x \in \mathbb{R}^d$ and all $n \geqslant N$,

$$\mathsf{P}_x(X_k \in B(x, \nu n^{1/\eta}) \text{ for } k = 0, \ldots, n) \geqslant \tfrac{1}{2}. \tag{22}$$

For $|x|$ so large that $(|x|/2\nu)^\eta \geqslant N$ it follows from (22) with $n = \lfloor (|x|/2\nu)^\eta \rfloor \geqslant N$ that

$$\mathsf{P}_x(X_k \in B(x, |x|/2) \text{ for } k = 0, \ldots, \lfloor (|x|/2\nu)^\eta \rfloor) \geqslant \tfrac{1}{2}, \tag{23}$$

from which (20) follows by the same arguments as in the proof of Theorem 3.2. □

By arguments similar to those used in the proofs of the two preceding theorems we obtain the following corollary which relates the moments of P to those of π.

Corollary 3.4. *Assume* $\mathbf{X}$ *is of random walk-type with invariant distribution* π.

(i) *Assume* (14) *holds with* $\int_0^\infty h(y)\mathrm{d}y < \infty$*. Then, for any* $\gamma \geqslant 0$ *with* $\pi(|x|^\gamma) < \infty$,

$$\int_{\mathbb{R}^d} |x|^{\gamma+1} P(z, \mathrm{d}x) < \infty \qquad \textit{for } \pi\textit{-almost all } z. \tag{24}$$

(ii) *Assume there exist* $0 < \eta < 1$ *and a constant* $c > 0$ *such that* (14) *holds with* $h(y) = cy^{-\eta}$*, for* y *sufficiently large. Then, for any* $\gamma \geqslant 0$ *with* $\pi(|x|^\gamma) < \infty$,

$$\int_{\mathbb{R}^d} |x|^{\gamma+\eta} P(z, \mathrm{d}x) < \infty \qquad \textit{for } \pi\textit{-almost all } z. \tag{25}$$

Proof. (i) Assume (14) holds with $\int_0^\infty h(y)\mathrm{d}y < \infty$. As shown in the proof of Theorem 3.2, there exist $\mu > 0$ and $R > 0$ such that, for $|x| \geqslant R$,

$$\mathsf{P}_x(X_k \in B(x, |x|/2) \text{ for } k = 0, \ldots, \lfloor |x|/4\mu \rfloor) \geqslant \tfrac{1}{2}. \tag{26}$$

Choose $K > 0$ so large that $C = B(0, K)$ has strictly positive π-measure. For any $\gamma \geqslant 0$ and any $|x| \geqslant 2K$, we have on the above event.

$$\tau_C - 1 \geqslant \lfloor |x|/4\mu \rfloor,$$

$$|X_k|^\gamma \geqslant (|x|/2)^\gamma, \qquad \text{for } k = 0, \ldots, \lfloor |x|/4\mu \rfloor,$$

and thus also

$$\sum_{k=0}^{\tau_C - 1} |X_k|^\gamma \geqslant \sum_{k=0}^{\lfloor |x|/4\mu \rfloor} |X_k|^\gamma \geqslant \frac{|x|^\gamma}{2^\gamma}(\lfloor |x|/4\mu \rfloor + 1) \geqslant \frac{|x|^{\gamma+1}}{2^\gamma (4\mu)}.$$

For $|x|$ sufficiently large this event has probability at least $\frac{1}{2}$ by (26). Hence, for any $\gamma \geqslant 0$, we have

$$\mathsf{E}_x\left[\sum_{k=0}^{\tau_C - 1} |X_k|^\gamma\right] \geqslant \frac{|x|^{\gamma+1}}{2^{\gamma+1}(4\mu)}, \tag{27}$$

for $|x| \geqslant H$, where $H = \max\{2K, R\}$.

Assume now that $\pi(|x|^\gamma) < \infty$. Since $\pi(C) > 0$ we have, by Theorem 10.4.9 of Meyn and Tweedie (1993),

$$\infty > \pi(|x|^\gamma) = \int_C \mathsf{E}_z\left[\sum_{k=0}^{\tau_C - 1} |X_k|^\gamma\right] \pi(\mathrm{d}z).$$

Thus, for π-almost all z in C, the integrand is finite and for such z we have

$$\begin{aligned}
\infty > \mathsf{E}_z\left[\sum_{k=0}^{\tau_C - 1} |X_k|^\gamma\right] &= |z|^\gamma + \int_{C^c} \mathsf{E}_x\left[\sum_{k=0}^{\tau_C - 1} |X_k|^\gamma\right] P(z, \mathrm{d}x) \\
&\geqslant \int_{B(0,H)^c} \mathsf{E}_x\left[\sum_{k=0}^{\tau_C - 1} |X_k|^\gamma\right] P(z, \mathrm{d}x) \\
&\geqslant \int_{B(0,H)^c} \frac{|x|^{\gamma+1}}{2^{\gamma+1}(4\mu)} P(z, \mathrm{d}x),
\end{aligned}$$

where the last inequality uses (27). This shows that (24) holds for π-almost all z in $C = B(0, K)$. Since K can be chosen arbitrarily large, we conclude that (24) holds for π-almost all z, and we are done.

(ii) Replace (26) by (23) from the proof of Theorem 3.3 and proceed as in (i). □

Example. Let P be the Markov transition kernel for the random walk on $[0, \infty)$ given by

$$X_{n+1} = (X_n + W_{n+1})^+, \qquad n \in \mathbb{N}_0, \tag{28}$$

where (W_n) is an i.i.d. sequence of real-valued random variables with common law Γ. Clearly,

$\mathbf{X}$ is of random walk type. It is easy to show that if $\mathsf{E}[W] < 0$, then $\mathbf{X}$ is δ_0-irreducible, aperiodic and positive recurrent with invariant distribution π.

Proposition 3.5. *Assume* $\mathsf{E}[W] < 0$. *Then, for any* $\gamma \geqslant 2$,

$$\mathsf{E}[(W^+)^\gamma] < \infty \text{ if and only if } \pi(x^{\gamma-1}) < \infty. \tag{29}$$

Proof. From Proposition 5.1 of Jarner and Roberts (2002) it follows that if $\mathsf{E}[(W^+)^\gamma] < \infty$ then $\pi(x^{\gamma-1}) < \infty$. Strictly speaking, Proposition 5.1 of Jarner and Roberts (2002) assumes that γ is an integer, but the proof of the proposition is valid for any real $\gamma \geqslant 2$.

Assume instead that $\pi(x^{\gamma-1}) < \infty$. By assumption, $\mathsf{E}[W] < 0$. In particular, $\mathsf{E}[|W|] < \infty$ and we can therefore find h with $\int_0^\infty h(y)\mathrm{d}y < \infty$ satisfying (14). It then follows from Corollary 3.4(i) that, for π-almost all z in $[0, \infty)$,

$$\infty > \int_0^\infty x^\gamma P(z, \mathrm{d}x) = \int_{-z}^\infty (z+w)^\gamma \Gamma(\mathrm{d}w),$$

and we conclude that $\mathsf{E}[(W^+)^\gamma] < \infty$. $\square$

In fact, Proposition 3.5 holds for any $\gamma \geqslant 1$. The 'if' part of (29) follows as in the proof above by use of Corollary 3.4, while the 'only if' part can be proved by a suitable modification of the proof of Proposition 5.1 in Jarner and Roberts (2002), but we omit the details.

This means that there exist Markov chains of random walk type with invariant distributions without any (polynomial) moments. For example, let the density of W in (28) be given by $f(w) \propto [(w+c)\log(w+c)]^{-2}$ for $w \geqslant e - c$, where $c > 0$ is chosen such that $\mathsf{E}[W] < 0$. Then the invariant distribution π exists, but $\pi(x^s) = \infty$ for all $s > 0$ since $\mathsf{E}[(W^+)^\gamma] = \infty$ for all $\gamma > 1$.

3.2. Random-walk-type Markov chains with moments and drift

In this section we obtain higher polynomial moments of π for random-walk-type Markov chains with increment distributions having uniformly bounded moments and polynomial drift. For the Markov kernel $P(x, \cdot)$ let $D(x, \cdot)$ be the distribution of $|X_1| - |x|$, that is,

$$D(x, (-\infty, z)) = P(x, B(0, |x|+z)) \qquad (x \in \mathbb{R}^d,\ z \in \mathbb{R}).$$

For $z \leqslant -|x|$ the ball $B(0, |x|+z)$ is the empty set and the probability on the right-hand side is zero. We assume that there exists a family of distributions $(H(r))_{r\geqslant 0}$ on $\mathbb{R}$ such that, for all $r \geqslant 0$ and all $|x| \geqslant r$,

$$H(r) \overset{\text{st}}{\leqslant} D(x, \cdot), \tag{30}$$

where for two distributions Q_1 and Q_2 on $\mathbb{R}$ we write $Q_1 \overset{\text{st}}{\leqslant} Q_2$ when the corresponding distribution functions F_1 and F_2 satisfy $F_1(y) \geqslant F_2(y)$ for all y.

Let $Y(r)$ be a random variable with distribution $H(r)$. The assumption is then that whenever the Markov chain is at distance at least r from the origin the amount by which the distance increases after one iteration is stochastically larger than $Y(r)$. If $\mathbf{X}$ is of random walk type this assumption is always satisfied, and in fact we can choose $H(r) = H$ independent of r such that (30) holds. In general, however, we want to choose $H(r)$ depending on r and 'as large as possible' for the given r.

In Theorems 3.6 and 3.7 below we assume that, for some $1 < \eta \leqslant 2$ and r sufficiently large, $Y(r)$ has uniformly bounded moments of order η and the drift $\mathsf{E}[Y(r)]$ is bounded from below by $-cr^{1-\eta}$, $c > 0$. We show that under these assumptions polynomial ergodicity of order (α, β) implies (13). These assumptions imply that the drift of the Markov chain towards the centre of the space becomes smaller and smaller such that it looks more and more like an unbiased random walk the further away it gets from the origin. As shown in Section 4, Langevin and symmetric random walk Metropolis algorithms with polynomial target densities behave like this. Since we only make assumptions about the limit behaviour of $Y(r)$ it is enough that (30) holds for r sufficiently large.

Theorem 3.6. *Assume that* $\mathbf{X}$ *is of random walk type and that there exists a constant* $c > 0$ *such that*

$$\mathsf{E}[Y(r)] \geqslant -\frac{c}{r} \qquad \textit{for } r \textit{ sufficiently large}, \tag{31}$$

and that

$$\limsup_{r\to\infty} \operatorname{var}(Y(r)) < \infty. \tag{32}$$

If $\mathbf{X}$ *is polynomially ergodic of order* (α, β), *then*

$$\int_{\mathbb{R}^d} |x|^{\alpha+2\beta}\pi(\mathrm{d}x) < \infty. \tag{33}$$

Proof. For any x, we have by (30) the stochastic ordering $H(|x|/2) \overset{\text{st}}{\leqslant} D(y, \cdot)$ for all $|y| \geqslant |x|/2$, and by a stochastic comparison argument we then also have

$$\hat{\tau}_{(-\infty,|x|/2)} \overset{\text{st}}{\leqslant} \tau_{B(0,|x|/2)}, \tag{34}$$

where $\tau_{B(0,|x|/2)}$ is the first return time to $B(0, |x|/2)$ of the Markov chain $\mathbf{X}$ started at x and $\hat{\tau}_{(-\infty,|x|/2)}$ is the first return time to the interval $(-\infty, |x|/2)$ of the random walk (W_i) on $\mathbb{R}$ given by

$$W_0 = |x|,$$

$$W_i = W_{i-1} + Z_i \qquad (i \geqslant 1),$$

where (Z_i) is an i.i.d. sequence of random variables with distribution $H(|x|/2)$. To simplify the notation we are suppressing the dependence of W_i and Z_i on x.

Let $R > 0$ be so large that (31) holds for $r \geqslant R$ and so large that $K = \sup_{r\geqslant R} \operatorname{var}(Y(r)) < \infty$. For $|x| \geqslant 2R$ we then have by Kolmogorov's inequality, for any $a > 0$ and any $n > 0$,

$$\mathsf{P}\left(\max_{1 \leqslant k \leqslant n} |S_k - m_k| > a \right) \leqslant \frac{\mathrm{var}(S_n)}{a^2} \leqslant \frac{nK}{a^2}, \tag{35}$$

where $S_k = Z_1 + \ldots + Z_k$ and $m_k = \mathsf{E}[S_k] = k\mathsf{E}[Y(|x|/2)]$.

Let $\delta = \min\{1/32K, 1/8c\}$. From (35) with $n = \lfloor \delta |x|^2 \rfloor$ and $a = |x|/4$, it follows that

$$\mathsf{P}\left(\max_{1 \leqslant k \leqslant \lfloor \delta |x|^2 \rfloor} |S_k - m_k| > |x|/4 \right) \leqslant \frac{\lfloor \delta |x|^2 \rfloor 16K}{|x|^2} \leqslant \frac{1}{2}. \tag{36}$$

Since $m_k \geqslant -2ck/|x| \geqslant -2c\lfloor \delta |x|^2 \rfloor / |x| \geqslant -|x|/4$ for $k = 1, \ldots, \lfloor \delta |x|^2 \rfloor$, it follows that

$$\mathsf{P}(\hat{\tau}_{(-\infty, |x|/2)} > \lfloor \delta |x|^2 \rfloor) \geqslant \tfrac{1}{2}, \tag{37}$$

and then by (34) also that

$$\mathsf{P}_x(\tau_{B(0,|x|/2)} > \lfloor \delta |x|^2 \rfloor) \geqslant \tfrac{1}{2}. \tag{38}$$

By Lemmas 3.1 and 2.1, (15) holds for a small and hence bounded set C. For $|x|$ so large that $C \subset B(0, |x|/2)$ we have on the above event

$$\tau_C - 1 \geqslant \lfloor \delta |x|^2 \rfloor$$

$$|X_k|^\alpha \geqslant (|x|/2)^\alpha, \qquad \text{for } k = 0, \ldots, \lfloor \delta |x|^2 \rfloor,$$

and thus also

$$\sum_{k=0}^{\tau_C - 1} \Delta r(k) |X_k|^\alpha \geqslant \sum_{k=0}^{\lfloor \delta |x|^2 \rfloor} \Delta r(k) |X_k|^\alpha \geqslant \frac{|x|^\alpha}{2^\alpha} r(\lfloor \delta |x|^2 \rfloor) \geqslant \frac{|x|^{\alpha + 2\beta} \delta^\beta}{2^\alpha},$$

where $r(k) = (k+1)^\beta$. For $|x|$ sufficiently large this event has probability at least $\frac{1}{2}$ by (38) and hence, for $|x|$ sufficiently large,

$$\mathsf{E}_x \left[\sum_{k=0}^{\tau_C - 1} \Delta r(k) |X_k|^\alpha \right] \geqslant \frac{|x|^{\alpha + 2\beta} \delta^\beta}{2^{\alpha + 1}},$$

and (33) now follows from (15). □

It is well known that for a one-dimensional symmetric random walk with finite variance the return time to the centre increases as $|x|^2$ (see Chapter III of Feller 1968). The theorem above says that this is still the case if the random walk is biased of order $|x|^{-1}$.

Theorem 3.7. *Assume that* $\mathbf{X}$ *is of random walk type and that there exist* $1 < \eta < 2$ *and a constant* $c > 0$ *such that*

$$\mathsf{E}[Y(r)] \geqslant -\frac{c}{r^{\eta - 1}} \qquad \textit{for } r \textit{ sufficiently large}, \tag{39}$$

and that

$$\limsup_{r \to \infty} \mathsf{E}|Y(r)|^\eta < \infty. \tag{40}$$

If $\mathbf{X}$ *is polynomially ergodic of order* (α, β)*, then*

$$\int_{\mathbb{R}^d} |x|^{\alpha+\eta\beta} \pi(\mathrm{d}x) < \infty. \tag{41}$$

Proof. As in the proof of Theorem 3.6, we seek to bound the return time $\hat{\tau}_{(-\infty,|x|/2)}$ of the random walk

$$W_0 = |x|,$$

$$W_i = W_{i-1} + Z_i \qquad (i \geqslant 1),$$

where (Z_i) is an i.i.d. sequence of random variables with distribution $H(|x|/2)$. Let $\tilde{Z}_i = Z_i 1_{(|Z_i| \leqslant |x|)}$ be the random variable Z_i truncated at $|x|$, and let $\tilde{\tau}_{(-\infty,|x|/2)}$ be the return time to $(-\infty, |x|/2)$ of the random walk $(\tilde{W}_i)$ given by

$$\tilde{W}_0 = |x|,$$

$$\tilde{W}_i = \tilde{W}_{i-1} + \tilde{Z}_i \qquad (i \geqslant 1).$$

Let $R > 0$ be so large that (39) holds for $r \geqslant R$ and so large that $K = \sup_{r \geqslant R} \mathsf{E}|Y(r)|^\eta < \infty$. For $|x| \geqslant 2R$ we then have the bounds

$$\mathsf{P}(\tilde{Z}_i \neq Z_i) = \mathsf{P}(|Z_i| > |x|) \leqslant \frac{\mathsf{E}|Z_i|^\eta}{|x|^\eta} \leqslant \frac{K}{|x|^\eta},$$

$$|\mathsf{E}[\tilde{Z}_i] - \mathsf{E}[Z_i]| \leqslant \mathsf{E}|\tilde{Z}_i - Z_i| = \mathsf{E}[|Z_i| 1_{(|Z_i| > |x|)}] \leqslant \frac{\mathsf{E}|Z_i|^\eta}{|x|^{\eta-1}} \leqslant \frac{K}{|x|^{\eta-1}},$$

$$\mathsf{E}[\tilde{Z}_i] \geqslant \mathsf{E}[Z_i] - |\mathsf{E}[\tilde{Z}_i] - \mathsf{E}[Z_i]| \geqslant -\frac{2^{\eta-1}c + K}{|x|^{\eta-1}} = -\frac{d}{|x|^{\eta-1}},$$

where $d = 2^{\eta-1}c + K$, and

$$\mathrm{var}(\tilde{Z}_i) \leqslant \mathsf{E}\tilde{Z}_i^2 = \mathsf{E}[|Z_i|^2 1_{(|Z_i| \leqslant |x|)}] \leqslant |x|^{2-\eta} \mathsf{E}|Z_i|^\eta \leqslant K|x|^{2-\eta}.$$

For $|x| \geqslant 2R$ we have by Kolmogorov's inequality, for any $a > 0$ and any $n > 0$,

$$\mathsf{P}\left(\max_{1 \leqslant k \leqslant n} |\tilde{S}_k - \tilde{m}_k| > a \right) \leqslant \frac{\mathrm{var}(\tilde{S}_n)}{a^2} \leqslant \frac{nK|x|^{2-\eta}}{a^2}, \tag{42}$$

where $\tilde{S}_k = \tilde{Z}_1 + \ldots + \tilde{Z}_k$ and $\tilde{m}_k = \mathsf{E}[\tilde{S}_k] = k\mathsf{E}[\tilde{Z}_1]$.

Let $\delta = \min\{1/32K, 1/4d\}$. From (42) with $n = \lfloor \delta |x|^\eta \rfloor$ and $a = |x|/4$, we obtain

$$\mathsf{P}\left(\max_{1 \leqslant k \leqslant \lfloor \delta |x|^\eta \rfloor} |\tilde{S}_k - \tilde{m}_k| > |x|/4 \right) \leqslant \frac{\lfloor \delta |x|^\eta \rfloor |x|^{2-\eta} 16K}{|x|^2} \leqslant \frac{1}{2}. \tag{43}$$

Since $\tilde{m}_k \geqslant -kd/|x|^{\eta-1} \geqslant -\lfloor \delta |x|^\eta \rfloor d/|x|^{\eta-1} \geqslant -|x|/4$ for $k = 1, \ldots, \lfloor \delta |x|^\eta \rfloor$, it follows that

$$\mathsf{P}(\tilde{\tau}_{(-\infty,|x|/2)} > \lfloor \delta |x|^\eta \rfloor) \geqslant \tfrac{1}{2}. \tag{44}$$

Further,

$$\mathsf{P}(\tilde{Z}_k = Z_k \text{ for } k = 1, \ldots, \lfloor \delta|x|^\eta \rfloor) = 1 - \mathsf{P}\left(\bigcup_{k=1}^{\lfloor \delta|x|^\eta \rfloor} (\tilde{Z}_k \neq Z_k) \right) \geqslant 1 - \frac{\lfloor \delta|x|^\eta \rfloor K}{|x|^\eta} \geqslant \frac{31}{32},$$

and then

$$\begin{aligned}\mathsf{P}\hat{\tau}_{(-\infty,|x|/2)} > \lfloor \delta|x|^\eta \rfloor) &\geqslant \mathsf{P}(\tilde{\tau}_{(-\infty,|x|/2)} > \lfloor \delta|x|^\eta \rfloor \text{ and } \tilde{Z}_k = Z_k \text{ for } k = 1, \ldots, \lfloor \delta|x|^\eta \rfloor) \\ &\geqslant \frac{15}{32}.\end{aligned}$$

As in the proof of Theorem 3.6 we then also have

$$\mathsf{P}_x(\tau_{B(0,|x|/2)} > \lfloor \delta|x|^\eta \rfloor) \geqslant \frac{15}{32},$$

from which the conclusion (41) follows as in the proof of Theorem 3.6. □

In Theorems 3.6 and 3.7 the drift is assumed to be bounded from below by $-cr^{1-\eta}$. This assumption is only made to match the random fluctuations such that after n iterations the distance travelled due to drift and that due to random fluctuations are of the same order. However, the arguments used in the proofs of Theorems 3.6 and 3.7 can also be used if the drift is allowed to be larger, but in this case the drift will dominate the random fluctuations and the inferred polynomial moments of π will be smaller.

As in the previous section we have the following corollary which we state without proof.

Corollary 3.8. *Assume that* **X** *is of random walk type with invariant distribution* π.

(i) *Assume that conditions* (31) *and* (32) *of Theorem 3.6 are satisfied. Then, for any* $\gamma \geqslant 0$ *with* $\pi(|x|^\gamma) < \infty$,

$$\int_{\mathbb{R}^d} |x|^{\gamma+2} P(z, \mathrm{d}x) < \infty \qquad \textit{for } \pi\textit{-almost all } z. \tag{45}$$

(ii) *Assume that there exists* $1 < \eta < 2$ *such that conditions* (39) *and* (40) *of Theorem 3.7 are satisfied. Then, for any* $\gamma \geqslant 0$ *with* $\pi(|x|^\gamma) < \infty$,

$$\int_{\mathbb{R}^d} |x|^{\gamma+\eta} P(z, \mathrm{d}x) < \infty \qquad \textit{for } \pi\textit{-almost all } z. \tag{46}$$

4. Applications to Markov chain Monte Carlo

We will give two MCMC applications of the results in the previous sections: one concerning the symmetric random walk Metropolis algorithm, and one concerning the Langevin algorithm, both of which are special cases of the Metropolis–Hastings algorithm (Hastings 1970). In both cases we will show that the polynomial ergodicity results obtained in Jarner and Roberts (2001) are tight.

4.1. The symmetric random walk Metropolis algorithm

The Metropolis–Hastings algorithm is an algorithm for constructing a Markov chain with a prescribed invariant distribution π referred to as the target distribution. We assume that the state space is $\mathbb{R}^d$ equipped with its Borel σ-field, and that the target distribution π has density, also denoted by π, with respect to Lebesgue measure μ^{Leb}. The algorithm is based on a *candidate transition kernel* $Q(x, \cdot)$ which generates proposed moves for the Markov chain $\mathbf{X}$. We assume that $Q(x, \cdot)$ has density $q(x, y)$ with respect to μ^{Leb}. If the current state is x, a proposed move to y, generated according to the density $q(x, y)$, is then accepted with probability

$$\alpha(x, y) = \begin{cases} \min\left\{\dfrac{\pi(y)q(y, x)}{\pi(x)q(x, y)}, 1\right\}, & \text{when } \pi(x)q(x, y) > 0, \\ 1, & \text{when } \pi(x)q(x, y) = 0. \end{cases} \tag{47}$$

Thus the Markov transition kernel P for the Markov chain $\mathbf{X}$ is given by

$$P(x, \mathrm{d}y) = p(x, y)\mu^{\text{Leb}}(\mathrm{d}y) + r(x)\delta_x(\mathrm{d}y), \tag{48}$$

where $p(x, y) = \alpha(x, y)q(x, y)$ for $x \neq y$ and 0 otherwise, δ_x is the point mass at x and

$$r(x) = \int (1 - \alpha(x, y))q(x, y)\mu^{\text{Leb}}(\mathrm{d}y) \tag{49}$$

is the probability of staying at x. The kernel P is reversible with respect to π and hence has π as its invariant distribution.

We first consider the special case of this algorithm known as the symmetric random walk Metropolis algorithm (Metropolis *et al.* 1953), in which q has the form

$$q(x, y) = q(|x - y|), \tag{50}$$

that is, the proposed increments are generated according to the same symmetric distribution $Q(\mathrm{d}x) = q(x)\mu^{\text{Leb}}(\mathrm{d}x)$. In this case the acceptance probability simplifies to

$$\alpha(x, y) = \min\left\{\frac{\pi(y)}{\pi(x)}, 1\right\}. \tag{51}$$

The symmetric random walk Metropolis algorithm is clearly of random walk type as defined in the Introduction.

We assume that π is bounded away from zero and infinity on bounded sets and that q is bounded away from zero in some region around zero, that is, there exist $\delta_q > 0$ and $\epsilon_q > 0$ such that $q(x) \geqslant \epsilon_q$ for $|x| \leqslant \delta_q$. Under these assumptions P is μ^{Leb}-irreducible and aperiodic by Theorem 2.2 of Roberts and Tweedie (1996b).

By Theorem 2.2 in the present paper it follows that exponential or lighter tails of π are a necessary condition for geometric ergodicity of the symmetric random walk Metropolis algorithm irrespective of the proposal distribution Q. In one dimension this is essentially also a sufficient condition, while in higher dimensions additional assumptions on the contour manifolds are needed (see Mengersen and Tweedie 1996; Roberts and Tweedie 1996b; Jarner and Hansen 2000).

If π has polynomial tails the algorithm will be only polynomially ergodic. It is somewhat surprising, however, that the order of polynomial ergodicity depends on the tails of both π and Q. For ease of exposition we consider only a stylized one-dimensional case.

Assume that π is a continuous, strictly positive density on the half-line $[0, \infty)$ and that there exists $r > 0$ such that

$$\pi(x) \propto \frac{1}{x^{1+r}} \qquad \text{for } x \text{ sufficiently large.} \tag{52}$$

Proposition 4.1. *Assume that π takes the form of* (52) *and that Q has finite variance. For $\alpha, \beta \geqslant 0$ the symmetric random walk Metropolis algorithm is polynomially ergodic of order (α, β) if and only if $\alpha + 2\beta < r$.*

Proof. The 'if' part follows from Proposition 3.1 of Jarner and Roberts (2001). The 'only if' part will follow from Theorem 3.6 above if we can show that (31) and (32) are satisfied.

By the assumptions on π we have that, for x sufficiently large, all proposed moves to the left are accepted. Also, by (52) we have that, for x sufficiently large, the acceptance probability of any positive increment y is an increasing function in x, that is, $\alpha(x, x + y)$ is increasing in x for any $y \geqslant 0$. Thus for z sufficiently large we have, for all $x \geqslant z$,

$$H(z) \overset{\text{st}}{\leqslant} P(x, \cdot) - x, \tag{53}$$

where $H(z, \mathrm{d}y) = h(z, y)\mu^{\text{Leb}}(\mathrm{d}y)$ for $y \neq 0$, with $h(z, y)$ given by

$$h(z, y) = \begin{cases} q(y), & \text{for } y < 0, \\ q(y)\left(\dfrac{z}{z + y}\right)^{1+r}, & \text{for } y > 0, \end{cases} \tag{54}$$

and $H(z, \{0\}) = 1 - \int h(z, y)\mathrm{d}y$. Let $Y(z)$ be a random variable with distribution $H(z)$. Since Q is assumed to have variance, (32) is satisfied and it only remains to show (31). Now using the fact that, for any $u \geqslant 0$,

$$\left(\frac{1}{1 + u}\right)^{1+r} - 1 \geqslant -(1 + r)u,$$

we find

$$\mathsf{E}[Y(z)] = \int_0^\infty q(y)y\left[\left(\frac{z}{z + y}\right)^{1+r} - 1\right]\mathrm{d}y \geqslant -\frac{(1 + r)}{z}\int_0^\infty q(y)y^2\mathrm{d}y,$$

and since $\int_0^\infty q(y)y^2\,\mathrm{d}y < \infty$ this shows that (31) holds and we are done. $\square$

For symmetric random walk Metropolis algorithms with proposal distribution without variance we have the following result. Recall that a function l is normalized slowly varying if, for all $a > 0$, $x^a l(x)$ is eventually increasing and $x^{-a} l(x)$ is eventually decreasing.

Proposition 4.2. *Assume that π takes the form of* (52) *and that there exists $0 < \eta \leqslant 2$ such that, for $|x|$ sufficiently large, $q(x)$ can be written*

$$q(x) = \frac{l(|x|)}{|x|^{1+\eta}}, \tag{55}$$

where l is a normalized slowly varying function.

The symmetric random walk Metropolis algorithm is polynomially ergodic of order (α, β) for all $\alpha, \beta \geqslant 0$ with $\alpha + \eta\beta < r$, and not polynomially ergodic of order (α, β) for any $\alpha, \beta \geqslant 0$ with $\alpha + \eta\beta > r$.

Proof. That the symmetric random walk Metropolis algorithm is polynomially ergodic of order (α, β) when $\alpha + \eta\beta < r$ follows from Proposition 3.2 of Jarner and Roberts (2001) and the remarks following it. For the second part of the statement, assume by way of contradiction that the algorithm is polynomially ergodic of order (α, β) with $\alpha + \eta\beta > r$.

Consider first the case where $0 < \eta \leqslant 1$. Choose $0 < \eta' < \eta$ such that $\alpha + \eta'\beta > r$. By (55) we have, for $|x|$ sufficiently large,

$$q(x) \leqslant \frac{1}{|x|^{1+\eta'}}, \tag{56}$$

and it then follows from Theorem 3.3 in the present paper that $\pi(|x|^{\alpha+\eta'\beta}) < \infty$, which contradicts $\alpha + \eta'\beta > r$.

Consider next the case where $1 < \eta \leqslant 2$. As in the proof of Proposition 4.1, we have that (53) holds with $H(z, \mathrm{d}y) = h(z, y)\mu^{\mathrm{Leb}}(\mathrm{d}y)$ for $y \neq 0$, where $h(x, y)$ is given by (54). Let $Y(z)$ be a random variable with distribution $H(z)$. By the same argument as above, it is clear that, for any $1 < \eta' < \eta$,

$$\limsup_{z\to\infty} \mathsf{E}|Y(z)|^{\eta'} < \infty.$$

Now choose $1 < \eta' < \eta$ such that $\alpha + \eta'\beta > r$, and let K be so large that (56) holds for $|x| \geqslant K$. We then have

$$\mathsf{E}[Y(z)] = I_1(z) + I_2(z), \tag{57}$$

where

$$I_1(z) = \int_0^K q(y)y\left[\left(\frac{z}{z+y}\right)^{1+r} - 1\right]\mathrm{d}y, \qquad I_2(z) = \int_K^\infty q(y)y\left[\left(\frac{z}{z+y}\right)^{1+r} - 1\right]\mathrm{d}y.$$

As in the proof of Proposition 4.1, $I_1(z) \geqslant -c_1/z$ for some constant $c_1 > 0$. For $I_2(z)$ we obtain, using (56) and the transformation $u = y/z$,

$$I_2(z) \geqslant \int_K^\infty \frac{1}{y^{\eta'}}\left[\left(\frac{z}{z+y}\right)^{1+r} - 1\right]\mathrm{d}y = \frac{1}{z^{\eta'-1}}\int_{K/z}^\infty \frac{1}{u^{\eta'}}\left[\left(\frac{1}{1+u}\right)^{1+r} - 1\right]\mathrm{d}y \geqslant -\frac{c_2}{z^{\eta'-1}},$$

where

$$c_2 = \int_0^\infty \frac{1}{u^{\eta'}}\left[1 - \left(\frac{1}{1+u}\right)^{1+r}\right]\mathrm{d}y < \infty.$$

The integral is finite since the integrand looks like $u^{-\eta'}$ at infinity and $\eta' > 1$, and like $u^{1-\eta'}$ at the origin and $1-\eta' > -1$. Hence by (57) there exists constant $c > 0$ such that $\mathsf{E}[Y(z)] \geqslant -c/z^{\eta'-1}$ for z sufficiently large. By Theorem 3.7 it then follows that $\pi(|x|^{\alpha+\eta'\beta}) < \infty$, which again contradicts $\alpha + \eta'\beta > r$. $\square$

Whereas the sufficiency part of, in particular, the last proposition is not straightforward to extend to higher dimensions, the necessity part can easily be extended to $\mathbb{R}^d$ and to more general polynomially decaying target densities.

4.2. The Langevin algorithm

When π is a positive, twice differentiable density, Besag (1994), for example, has proposed using the candidate transition kernel

$$Q(x, \cdot) = N\left(x + \frac{h}{2}\nabla \log \pi(x),\ h\right), \tag{58}$$

where $h > 0$, N denotes the normal distribution and ∇ is the differential operator. As shown by Roberts and Tweedie (1996a), this choice performs well when π has exponentially decaying or Gaussian tails, giving geometrically ergodic algorithms in both these cases. They also show that the algorithm fails to be geometrically ergodic when $\nabla \log \pi(x) \to 0$ as $|x| \to \infty$, for example when π has polynomial tails. In fact, when π has polynomial tails the results of this paper are particularly easy to apply and we obtain the following result which shows that in this case the convergence properties of the Langevin algorithm are exactly the same as for a symmetric random walk Metropolis algorithm with finite variance.

Let P denote the Markov transition kernel for the Langevin algorithm, i.e. the Metropolis–Hastings algorithm with Q given by (58).

Proposition 4.3. *Assume that π is a strictly positive, twice differentiable density on* $[0, \infty)$ *which takes the form of* (52). *For any $h > 0$, the Langevin algorithm is polynomially ergodic of order (α, β) if and only if $\alpha + 2\beta < r$.*

Proof. The 'if' part follows from Proposition 4.1 of Jarner and Roberts (2001). For the 'only if' part, first note that P is μ^{Leb}-irreducible and aperiodic and of random walk type since $|\nabla \log \pi(x)|$ is bounded away from infinity. Further, it is shown in the proof of Proposition 4.1 of Jarner and Roberts (2001) that, for x sufficiently large, all positive increments are accepted while negative increments are possibly rejected. Thus for x sufficiently large we have, for all $y \geqslant x$,

$$N\left(-\frac{h(1+r)}{2x},\ h\right) \overset{\mathrm{st}}{\leqslant} P(y, \cdot) - y.$$

It then follows from Theorem 3.6 above that P cannot be polynomially ergodic of order (α, β) for any $\alpha, \beta \geqslant 0$ with $\alpha + 2\beta \geqslant r$. □

Acknowledgements

We are grateful to Gareth Roberts for useful discussions and to the referee for inspiring comments. The work was supported in part by National Science Foundation Grant DMS-9803682 and the EU TMR network ERB-FMRX-CT96-0095 on 'Computational and statistical methods for the analysis of spatial data'. The first author would also like to thank the Danish Research Training Council for financial support, including a visit to the University of Minnesota.

References

Besag, J. (1994) Comments on 'Representations of knowledge in complex systems' by U. Grenander and M. Miller. *Roy. Statist. Soc. Ser. B*, **56**, 591–592.

Chan, K.S. (1993) Asymptotic behaviour of the Gibbs sampler. *J. Amer. Statist. Assoc.*, **88**, 320–326.

Feller, W. (1968) *An Introduction to Probability Theory and Its Applications, Vol. I*, 3rd edition. New York: Wiley.

Feller, W. (1971) *An Introduction to Probability Theory and Its Applications, Vol. II*, 2nd edition. New York: Wiley.

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

Hastings, W. K. (1970) Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, **57**, 97–109.

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

Jarner, S.F. and Roberts, G.O. (2001) Convergence of heavy tailed MCMC algorithms. http://www.statslab.cam.ac.uk/~mcmc. Submitted for publication.

Jarner, S.F. and Roberts, G.O. (2002) Polynomial convergence rates of Markov chains. *Ann. Appl. Probab.*, **12**, 224–247.

Jarner, S.F. and Tweedie, R.L. (2002) Convergence rates and moments of Markov chains associated with the mean of Dirichlet processes. *Stochastic Process. Appl.*, **101**, 257–271.

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

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. and Teller, E. (1953) Equation of state calculations by fast computing machines. *J. Chem. Phys.*, **21**, 1087–1091.

Meyn, S.P. and Tweedie, R.L. (1993) *Markov Chains and Stochastic Stability.* London: Springer-Verlag.

Nummelin, E. (1984) *General Irreducible Markov Chains and Non-negative Operators.* Cambridge: Cambridge University Press.

Nummelin, E. and Tuominen, P. (1982) Geometric ergodicity of Harris recurrent Markov chains with applications to renewal theory. *Stochastic Process. Appl*., **12**, 187–202.

Roberts, G.O. and Tweedie, R.L. (1996a) Exponential convergence of Langevin diffusions and their discrete approximations. *Bernoulli*, **2**, 341–364.

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

Tuominen, P. and Tweedie, R.L. (1994) Subgeometric rates of convergence of f-ergodic Markov chains. *Adv. in Appl. Probab.*, **26**, 775–798.

Tweedie, R.L. (1983) Criteria for rates of convergence of Markov chains with application to queueing and storage theory. In J.F.C. Kingman and G.E.H. Reuter (eds), *Probability, Statistics and Analysis*, London Math. Soc. Lecture Note Ser. Cambridge: Cambridge University Press.

Received June 2001 and revised March 2003