Queueing Syst
https://doi.org/10.1007/s11134-018-9574-1

On the rate of convergence to equilibrium for reflected Brownian motion

Peter W. Glynn[1] · **Rob J. Wang**[1,2]

Received: 21 September 2017 / Revised: 27 January 2018

Abstract This paper discusses the rate of convergence to equilibrium for one-dimensional reflected Brownian motion with negative drift and lower reflecting boundary at 0. In contrast to prior work on this problem, we focus on studying the rate of convergence for the entire distribution through the total variation norm, rather than just moments of the distribution. In addition, we obtain computable bounds on the total variation distance to equilibrium that can be used to assess the quality of the steady state for queues as an approximation to finite horizon expectations.

Keywords Reflected Brownian motion · Queueing theory · Total variation distance · Rate of convergence to equilibrium · Large deviations · Steady-state simulation · Diffusion processes

Mathematics Subject Classification 60F05 · 60F10 · 60G05 · 60J60 · 60K25

1 Introduction

This paper is concerned with the convergence to equilibrium of the reflected Brownian motion (RBM) process $X = (X(t): t \geq 0)$. This process can be characterized as the unique solution of the stochastic differential equation (SDE)

✉ Rob J. Wang
robjwang@gmail.com

Peter W. Glynn
glynn@stanford.edu

[1] Department of Management Science and Engineering, Stanford University, 475 Via Ortega, Stanford, CA 94305, USA

[2] Present Address: Airbnb, 999 Brannan St, San Francisco, CA 94103, USA

Published online: 23 March 2018

$$\mathrm{d}X(t) = -r\mathrm{d}t + \sigma\mathrm{d}B(t) + \mathrm{d}L(t), \tag{1.1}$$

subject to $X(0) = x \geq 0$, where $r > 0$, $B = (B(t): t \geq 0)$ is standard Brownian motion, and $L = (L(t): t \geq 0)$ is a continuous non-decreasing process for which $\mathbb{I}(X(t) > 0)\mathrm{d}L(t) = 0$ for all $t \geq 0$. The boundary process L is often called the *local time* of X (at the origin). The parameter $-r$ represents the *drift* of X and $\sigma > 0$ its *volatility*.

As is well known, RBM arises naturally as an approximation to the single-server queue in heavy traffic. To illustrate, suppose that $W = (W_n: n \geq 0)$ is the waiting time (exclusive of service) sequence for the first-in first-out (FIFO) single-server queue having inter-arrival times $(\chi_n: n \geq 1)$ and service times $(V_n: n \geq 0)$. If $((\chi_{n+1}, V_n): n \geq 0)$ is a stationary sequence, then heavy-traffic approximation asserts that

$$W_n \overset{\mathscr{D}}{\approx} X(n), \tag{1.2}$$

where $\overset{\mathscr{D}}{\approx}$ denotes *has approximately the same distribution as* (and has no rigorous meaning), and X is an RBM with drift $-r$ and volatility σ given by

$$\begin{aligned} -r &= E(V_0 - \chi_1), \\ \sigma^2 &= \mathrm{Var}(V_0 - \chi_1) + 2\sum_{j=1}^{\infty} \mathrm{Cov}(V_0 - \chi_1, V_j - \chi_{j+1}). \end{aligned} \tag{1.3}$$

A key result giving theoretical support to (1.3) is the limit theorem, due to Iglehart and Whitt [20], that holds as $r \searrow 0$; see also Borovkov [8].

When $r > 0$, both W and X have finite-valued steady states, in the sense that

$$W_n \Rightarrow W_\infty \tag{1.4}$$

and

$$X(t) \Rightarrow X(\infty) \tag{1.5}$$

as $n, t \to \infty$, where $\Rightarrow$ denotes weak convergence. Furthermore, Kingman [21] proved a limit theorem supporting the approximation

$$W_\infty \overset{\mathscr{D}}{\approx} X(\infty)$$

when r is small. The random variable $X(\infty)$ has an exponential distribution:

$$P(X(\infty) \in \mathrm{d}x) \overset{\Delta}{=} \pi(\mathrm{d}x) = 2\eta \mathrm{e}^{-2\eta x}\mathrm{d}x$$

for $x \geq 0$, where $\eta = r/\sigma^2$; see Harrison [19] p.102.

It is common, in many queueing applications, to approximate various transient performance measures by their corresponding more analytically tractable steady-state limits. Steady-state analysis also avoids the need for the modeler to specify a time horizon or initial distribution in pursuing his or her model analysis. Either one of these

modeling choices can be problematic, since there may be no natural candidate for either one.

A key issue that arises in replacing a transient analysis by an equilibrium formulation is the quality of the resulting approximation. In view of (1.2), it is natural to use the rate of convergence to equilibrium for RBM as a vehicle toward approximating the rate of convergence for the pre-limit queueing model. Accordingly, Abate and Whitt [1,2] study the rate of convergence of $E_x X(t)^k$ to $EX(\infty)^k$ as $t \to \infty$ (for $k = 1, 2, \ldots$), where $E_x(\cdot)$ is the expectation operator associated with the probability $P_x(\cdot) \stackrel{\Delta}{=} P(\cdot \mid X(0) = x)$. This was supplemented by an accompanying paper [3] on moment convergence for the $M/M/1$ queue, thereby extending earlier work of Cohen [10]. In addition, Abate and Whitt [4] considered the transient behavior of the workload process of the $M/G/1$ queue, but without a focus on utilizing the results to study the question of the rate of convergence to equilibrium.

This paper provides a further complement to this body of theory on rates of convergence, by comprehensively studying the rate at which the entire distribution of $X(t)$ converges to that of $X(\infty)$ as $t \to \infty$. In Sect. 2, we obtain both asymptotic and finite-time bounds for the rate at which $X(t)$ converges to $X(\infty)$ in the total variation norm; see (2.14) and (2.13). The total variation norm is the most widely used metric within the Markov process context for studying rates of convergence to equilibrium (see, for example, Roberts and Rosenthal [25], Meyn and Tweedie [23], and Diaconis [11]). We further generalize our results to the h-norm distance between $X(t)$ and $X(\infty)$; see Theorem 2. Given (1.3), these results translate immediately into a recommendation for the time $t = t(\epsilon)$ associated with guaranteeing that the "distance to equilibrium" be less than some given $\epsilon > 0$; see (2.21).

In Sect. 3, our focus is on the time t at which a tail probability $P_x(X(t) > y)$ can be well approximated by $P(X(\infty) > y)$. Our theoretical vehicle here is the use of large deviations, so that the suggested approximation (3.1) is likely to be especially relevant when the event $\{X(t) > y\}$ is "rare" under P_x.

While Sect. 2 considers the rate at which instantaneous performance measures of the form $E_x f(X(t))$ converge to $Ef(X(\infty))$ (for a given real-valued performance functional f), Sect. 4 studies their integrated counterpart, namely $E_x \int_0^t f(X(s))\mathrm{d}s$. Such integrated performance measures are especially relevant in cost/reward settings where the costs and rewards are aggregated over a given operational time horizon. Again, we develop approximations and bounds that can be used to assess when such finite horizon integrated performance measures can be replaced by an equilibrium expectation; see (4.9).

In Sect. 5, we change our perspective somewhat, re-interpreting the results of Sects. 2 and 4 to help assess how long a steady-state simulation must be run, in order that the bias introduced by the initial distribution be smaller than some given tolerance. Our recommendations differ, depending on whether the steady state is computed by simulating a single long replicate, or by generating p shorter replicates (as is natural in the parallel computing context).

Finally, Sect. 6 discusses RBM from the standpoint of its spectral representation. This representation is intimately connected to the Hilbert space of functions $L^2(\pi)$ that are square-integrable with respect to RBM's stationary distribution. The fact that

the convergence to equilibrium includes an algebraic non-exponential factor arises as a consequence of the fact that the spectrum has a continuous component. One interesting subtlety has to do with the fact that the spectral representation provides a description of the rate of convergence only for a subset of $L^2(\pi)$; see (6.6) and (6.7).

It should be noted that the companion paper Glynn and Wang [16] considers rates of convergence to equilibrium for two additional diffusion processes that arise naturally in queueing theory: two-sided RBM and the Ornstein–Uhlenbeck process.

2 Total variation convergence to equilibrium for RBM

In this section, we study the rate of convergence for instantaneous performance measures of the form $E_x f(X(t))$ to the equilibrium expectation $Ef(X(\infty))$. We start by recalling that the (conditional) cumulative distribution of function of $X(t)$ is given by

$$P_x(X(t) \le y) = 1 - \Phi\left(\frac{-y + x - rt}{\sigma\sqrt{t}}\right) - \mathrm{e}^{-2\eta y}\Phi\left(\frac{-y - x + rt}{\sigma\sqrt{t}}\right), \tag{2.1}$$

where $P_x(\cdot) \stackrel{\Delta}{=} P(\cdot \mid X(0) = x)$ and $\Phi(\cdot)$ denotes the cumulative distribution function associated with a standard normal random variable (rv) $N(0,\ 1)$; see, for example, Harrison [19], p. 48. It follows immediately that the transition density of X is given by

$$\begin{aligned} p(t,\ x,\ y) = {} & 2\eta\mathrm{e}^{-2\eta y}\Phi\left(\frac{rt - x - y}{\sigma\sqrt{t}}\right) + \frac{1}{\sigma\sqrt{t}}\phi\left(\frac{-rt + x - y}{\sigma\sqrt{t}}\right) \\ & + \frac{1}{\sigma\sqrt{t}}\mathrm{e}^{-2\eta y}\phi\left(\frac{rt - x - y}{\sigma\sqrt{t}}\right), \end{aligned} \tag{2.2}$$

where $\phi(\cdot)$ denotes the standard normal probability density function.

We consider first the most widely adopted "distance measure" used to study convergence rates for Markov processes, namely the *total variation* distance given by

$$\begin{aligned} \mathrm{d}(t,\ x) &\stackrel{\Delta}{=} \|P_x(X(t) \in \cdot) - P(X(\infty) \in \cdot)\| \\ &= \sup_{0 \le f \le 1} |E_x f(X(t)) - Ef(X(\infty))| \\ &= \sup_A |P_x(X(t) \in A) - P(X(\infty) \in A)|, \end{aligned}$$

where the final supremum is taken over all measurable subsets A of $[0,\ \infty)$. This is easily seen to equal

$$\frac{1}{2}\int_0^\infty |p(t,\ x,\ y) - p(y)|\mathrm{d}y. \tag{2.3}$$

A well-known property of the total variation distance for Markov processes, *monotonicity in time* (see, for example, Proposition 3 of Roberts and Rosenthal [25]), makes it especially convenient for the study of convergence rates. Indeed,

$$\begin{aligned}\mathrm{d}(s+t,\,x) &= \frac{1}{2}\int_0^\infty \left|\int_0^\infty (p(t,\,x,\,z)-p(z))p(s,\,z,\,y)\mathrm{d}z\right|\mathrm{d}y\\ &\le \frac{1}{2}\int_0^\infty |p(t,\,x,\,z)-p(z)|\int_0^\infty p(s,\,z,\,y)\mathrm{d}y\mathrm{d}z\\ &= \mathrm{d}(t,\,x)\end{aligned}$$

for $s,\,t \ge 0$.

To explore the rate of convergence of $\mathrm{d}(t,\,x)$ to zero, it is convenient to find an alternative representation to (2.2) for $p(t,\,x,\,y)-p(y)$. In particular, we will now derive the *spectral representation* for the transition density of RBM. Because of the key role it will play in our analysis, we provide a direct proof. (As we shall discuss in Sect. 6, the existing proof of the spectral representation for RBM relies on functional analysis ideas associated with the self-adjointness of the infinitesimal generator associated with X; see Linetsky [22].)

Throughout this paper, we will state our results for general RBM, but often prove the results only for the special case of canonical RBM. The RBM $\tilde{X} = (\tilde{X}(t): t \ge 0)$ is said to be *canonical* if $r = \sigma = 1$. In particular, the self-similarity of Brownian motion ensures that

$$X(\cdot) \stackrel{\mathscr{D}}{=} \frac{1}{\eta}\tilde{X}(\nu\cdot),$$

where $\stackrel{\mathscr{D}}{=}$ denotes equality in distribution and $\nu = r^2/\sigma^2$. It then follows that the transition density p for an RBM with drift $-r$ and volatility σ can be expressed in terms of the transition density $\tilde{p}$ for $\tilde{X}$, namely

$$p(t,\,x,\,y) = \eta\tilde{p}(\nu t,\,\eta x,\,\eta y). \tag{2.4}$$

Consequently, our derivation of the spectral representation will focus on canonical RBM.

Let $\overline{\Phi}(\cdot) = 1-\Phi(\cdot)$, and note that (2.1) implies that

$$\begin{aligned}&P_x(X(t)\le y) - P(X(\infty)\le y)\\ &= -\Phi\left(\frac{x-y-t}{\sqrt{t}}\right) + \mathrm{e}^{-2y}\overline{\Phi}\left(\frac{t-x-y}{\sqrt{t}}\right)\\ &= -\int_{-\infty}^{\frac{x-y-t}{\sqrt{t}}} \mathrm{e}^{-\frac{v^2}{2}}\frac{\mathrm{d}v}{\sqrt{2\pi}} + \mathrm{e}^{-2y}\int_{\frac{t-x-y}{\sqrt{t}}}^\infty \mathrm{e}^{-\frac{w^2}{2}}\frac{\mathrm{d}w}{\sqrt{2\pi}}\\ &= \int_t^\infty \frac{y-x-z}{2z^{\frac{3}{2}}}\mathrm{e}^{-\frac{(x-y-z)^2}{2z}}\frac{\mathrm{d}z}{\sqrt{2\pi}} + \mathrm{e}^{-2y}\int_t^\infty \frac{y+x+z}{2z^{\frac{3}{2}}}\mathrm{e}^{-\frac{(z-x-y)^2}{2z}}\frac{\mathrm{d}z}{\sqrt{2\pi}}.\end{aligned} \tag{2.5}$$

We now observe that

$$\begin{aligned}\frac{\mathrm{e}^{-\frac{(x+y)^2}{2z}}}{\sqrt{2\pi z}} &= \frac{E\mathrm{e}^{i(x+y)\frac{N(0,1)}{\sqrt{z}}}}{\sqrt{2\pi z}} \\ &= \int_{-\infty}^{\infty} \mathrm{e}^{-\frac{\xi^2 z}{2}} (\cos(x\xi)\cos(y\xi) - \sin(x\xi)\sin(y\xi))\mathrm{d}\xi \cdot \frac{1}{2\pi} \\ &= \frac{1}{\pi}\int_{0}^{\infty} \mathrm{e}^{-\frac{\xi^2 z}{2}} (\cos(x\xi)\cos(y\xi) - \sin(x\xi)\sin(y\xi))\mathrm{d}\xi, \end{aligned} \tag{2.6}$$

where we use the evenness of the cosine and sine products in the last step.

A similar expression to (2.6) holds for $\mathrm{e}^{-(x-y)^2/(2z)}/\sqrt{2\pi z}$. Taking the difference between these two expressions yields the identity

$$\frac{2}{\pi}\int_{0}^{\infty} \mathrm{e}^{-\frac{\xi^2 z}{2}} \sin(x\xi)\sin(y\xi)\mathrm{d}\xi = \frac{z\mathrm{e}^{-\frac{(x-y)^2}{2z}}}{z^{\frac{3}{2}}\sqrt{2\pi}} - \frac{z\mathrm{e}^{-\frac{(x+y)^2}{2z}}}{z^{\frac{3}{2}}\sqrt{2\pi}}. \tag{2.7}$$

Differentiating (2.7) with respect to x, we find that

$$\frac{2}{\pi}\int_{0}^{\infty} \mathrm{e}^{-\frac{\xi^2 z}{2}} \xi\cos(x\xi)\sin(y\xi)\mathrm{d}\xi = \frac{(y-x)\mathrm{e}^{-\frac{(x-y)^2}{2z}}}{z^{\frac{3}{2}}\sqrt{2\pi}} + \frac{(x+y)\mathrm{e}^{-\frac{(x+y)^2}{2z}}}{z^{\frac{3}{2}}\sqrt{2\pi}}. \tag{2.8}$$

Substituting (2.7) and (2.8) into the right-hand side of (2.5) then shows that

$$\begin{aligned} &P_x(X(t) \le y) - P(X(\infty) \le y) \\ &= \frac{1}{\pi}\mathrm{e}^{x-y}\int_{t}^{\infty}\int_{0}^{\infty} \mathrm{e}^{-\frac{(\xi^2+1)z}{2}} (\xi\cos(x\xi) - \sin(x\xi))\sin(y\xi)\mathrm{d}\xi\mathrm{d}z \\ &= \frac{1}{\pi}\mathrm{e}^{x-y}\int_{0}^{\infty} (\xi\cos(x\xi) - \sin(x\xi))\sin(y\xi)\int_{t}^{\infty} \mathrm{e}^{-\frac{(\xi^2+1)z}{2}}\mathrm{d}z\mathrm{d}\xi \\ &= \frac{2}{\pi}\mathrm{e}^{x-y}\int_{0}^{\infty} (\xi\cos(x\xi) - \sin(x\xi))\sin(y\xi)\frac{\mathrm{e}^{-\frac{(\xi^2+1)t}{2}}}{\xi^2+1}\mathrm{d}\xi. \end{aligned}$$

Differentiating the above expression with respect to y shows that

$$\begin{aligned} &p(t,\, x,\, y) - p(y) \\ &= \frac{2}{\pi}\mathrm{e}^{x-y}\int_{0}^{\infty} (\xi\cos(x\xi) - \sin(x\xi))(\xi\cos(y\xi) - \sin(y\xi))\frac{\mathrm{e}^{-\frac{(\xi^2+1)t}{2}}}{\xi^2+1}\mathrm{d}\xi. \end{aligned} \tag{2.9}$$

(All of the above interchanges of differentiation with integration are easily justified via the dominated convergence theorem.) Substituting $\lambda = -(\xi^2+1)/2$ into (2.9) yields Proposition 1, our desired spectral representation (stated for general RBM).

Proposition 1 *For all $x,\ y \geq 0$ and $t > 0$,*

$$p(t,\ x,\ y) - p(y) = p(y) \int_{-\infty}^{-\frac{\nu}{2}} \mathrm{e}^{\lambda t} u_\lambda(x) u_\lambda(y) \left(-\frac{s(\lambda)}{2\pi\lambda} \right) \mathrm{d}\lambda, \tag{2.10}$$

where $s(\lambda) = \sqrt{-2\lambda - \nu}/\sigma$ for $\lambda \leq -\nu/2$, and

$$u_\lambda(x) = \mathrm{e}^{\eta x} \left(\cos(s(\lambda)x) - \frac{\eta}{s(\lambda)} \sin(s(\lambda)x) \right) \tag{2.11}$$

for $\lambda < -\nu/2$.

Formula (2.10) expresses the difference $p(t,\ x,\ y) - p(y)$ as an integral of decaying exponentials. As we shall see in the proof of Theorem 1, this significantly simplifies our convergence analysis.

Remark 1 We shall also later need the quantity

$$u_{-\frac{\nu}{2}}(x) = \lim_{\lambda \nearrow -\frac{\nu}{2}} u_\lambda(x) = \mathrm{e}^{\eta x}(1 - \eta x). \tag{2.12}$$

We use the notation $o(a(t))$ to denote a function $g(t)$ such that $|g(t)|/a(t) \to 0$ as $t \to \infty$. For $p > 0$, let

$$L^p(\pi) = \left\{ h\colon \int_0^\infty |h(x)|^p \pi(\mathrm{d}x) < \infty \right\}$$

and write

$$\langle f,\ g \rangle = \int_0^\infty f(x) g(x) \pi(\mathrm{d}x)$$

whenever $fg \in L^1(\pi)$.

Theorem 1 *For each $t > 0$ and $x \geq 0$,*

$$\mathrm{d}(t,\ x) \leq \sqrt{\frac{2}{\pi}} (\nu t)^{-\frac{3}{2}} \mathrm{e}^{-\frac{\nu t}{2}} (1 + \eta x) \mathrm{e}^{\eta x} \min(\nu t,\ 1), \tag{2.13}$$

and

$$\mathrm{d}(t,\ x) = \sqrt{\frac{2}{\pi}} (\nu t)^{-\frac{3}{2}} \mathrm{e}^{-\frac{\nu t}{2}} |1 - \eta x| \mathrm{e}^{\eta x} \mathrm{e}^{-1} + o\left(t^{-\frac{3}{2}} \mathrm{e}^{-\frac{\nu t}{2}} \right) \tag{2.14}$$

as $t \to \infty$.

Furthermore, if f is a continuous function for which $f u_{-\nu/2} \in L^1(\pi)$, then

$$E_x f(X(t)) = Ef(X(\infty)) + \frac{1}{\sqrt{2\pi}} (\nu t)^{-\frac{3}{2}} e^{-\frac{\nu t}{2}} u_{-\frac{\nu}{2}}(x) \left\langle f,\ u_{-\frac{\nu}{2}} \right\rangle + o\left(t^{-\frac{3}{2}} \mathrm{e}^{-\frac{\nu t}{2}} \right) \tag{2.15}$$

as $t \to \infty$.

Proof As noted earlier, we can specialize our proof to canonical RBM. Note that the substitution $-z = (\lambda + \frac{1}{2})t$ into (2.10) yields

$$p(t,\ x,\ y) - p(y) = p(y)\frac{\mathrm{e}^{-\frac{t}{2}}}{2\pi t}\int_0^\infty \mathrm{e}^{-z} u_{-\frac{z}{t}-\frac{1}{2}}(x) u_{-\frac{z}{t}-\frac{1}{2}}(y)\frac{\sqrt{\frac{2z}{t}}}{\frac{z}{t}+\frac{1}{2}}\mathrm{d}z$$
$$= \frac{2\sqrt{2}}{\pi} t^{-\frac{3}{2}}\mathrm{e}^{-\frac{t}{2}}\mathrm{e}^{x-y}\int_0^\infty \sqrt{z}\mathrm{e}^{-z}k(t,\ x,\ y,\ z)\mathrm{d}z, \quad (2.16)$$

where

$$k(t,\ x,\ y,\ z) = \left(\frac{1}{1+\frac{2z}{t}}\right)\left(\cos\left(\sqrt{\frac{2z}{t}}x\right) - \sqrt{\frac{t}{2z}}\sin\left(\sqrt{\frac{2z}{t}}x\right)\right)$$
$$\cdot\left(\cos\left(\sqrt{\frac{2z}{t}}y\right) - \sqrt{\frac{t}{2z}}\sin\left(\sqrt{\frac{2z}{t}}y\right)\right).$$

Because $|\sin(w)/w| \le 1$ for all w, it follows that

$$z^{\frac{1}{2}}|k(t,\ x,\ y,\ z)| \le (1+x)(1+y)\frac{z^{\frac{1}{2}}}{\frac{2z}{t}+1} \le (1+x)(1+y)\min\left(\frac{t}{2z^{\frac{1}{2}}},\ z^{\frac{1}{2}}\right). \quad (2.17)$$

Consequently,

$$|p(t,\ x,\ y) - p(y)| \le \frac{2\sqrt{2}}{\pi} t^{-\frac{3}{2}}\mathrm{e}^{-\frac{t}{2}}\mathrm{e}^{x-y}(1+x)(1+y)$$
$$\cdot\int_0^\infty \mathrm{e}^{-z}\min\left(\frac{t}{2}z^{-\frac{1}{2}},\ z^{\frac{1}{2}}\right)\mathrm{d}z$$
$$\le \frac{2\sqrt{2}}{\pi} t^{-\frac{3}{2}}\mathrm{e}^{-\frac{t}{2}}\mathrm{e}^{x-y}(1+x)(1+y)$$
$$\cdot\min\left(\frac{t}{2}\Gamma\left(\frac{1}{2}\right),\ \Gamma\left(\frac{3}{2}\right)\right)$$
$$= \frac{2\sqrt{2}}{\pi} t^{-\frac{3}{2}}\mathrm{e}^{-\frac{t}{2}}\mathrm{e}^{x-y}(1+x)(1+y)$$
$$\cdot\min\left(\frac{t}{2}\sqrt{\pi},\ \frac{\sqrt{\pi}}{2}\right), \quad (2.18)$$

where $\Gamma(\cdot)$ denotes the gamma function; the values for $\Gamma(1/2)$ and $\Gamma(3/2)$ can be found on p. 19 of Artin [5]. Integrating over y, we obtain the bound (2.13) on $\mathrm{d}(t,\ x)$.

As for (2.14), note that we can multiply both sides of (2.16) by $\mathrm{e}^{\frac{t}{2}}t^{\frac{3}{2}}$, and use the bound (2.17) to establish the required domination needed for the application of the dominated convergence theorem, thereby obtaining the limit

$$t^{\frac{3}{2}}\mathrm{e}^{\frac{t}{2}}\int_0^\infty |p(t,\,x,\,y)-p(y)|\mathrm{d}y \to \frac{2\sqrt{2}}{\pi}\mathrm{e}^{x}|1-x|\int_0^\infty z^{\frac{1}{2}}\mathrm{e}^{-z}\mathrm{d}z\int_0^\infty |1-y|\mathrm{e}^{-y}\mathrm{d}y$$

as $t \to \infty$. The limit (2.14) then follows from the fact that

$$\int_0^\infty |1-y|\mathrm{e}^{-y}\mathrm{d}y = 2\int_0^1 (1-y)\mathrm{e}^{-y}\mathrm{d}y = 2\mathrm{e}^{-1}.$$

For the final assertion, note that the continuity of f implies that f is bounded over finite intervals, so that the bounded convergence theorem yields the correct limit behavior for the integral (in y) over $[0,\,2/\eta]$. For the integral in y over $[2/\eta,\,\infty)$, we note that $(1+y)/|u_{-\nu/2}(y)|$ is bounded there, and exploit the dominating function (2.17) and the dominated convergence theorem to finish the proof. □

Remark 2 We note that the total variation distance $\mathrm{d}(t,\,x)$ is always guaranteed to be less than or equal to 1, so that the right-hand side of (2.13) is interesting only when the value of t is large enough that the bound is less than 1.

Evidently, for $t \geq 1/\nu$, (2.13) implies that

$$\mathrm{d}(t,\,x) \leq \sqrt{\frac{2}{\pi}}(\nu t)^{-\frac{3}{2}}\mathrm{e}^{-\frac{\nu t}{2}}(1+\eta x)\mathrm{e}^{\eta x}. \tag{2.19}$$

In particular,

$$\mathrm{d}(t,\,0) \leq \sqrt{\frac{2}{\pi}}(\nu t)^{-\frac{3}{2}}\mathrm{e}^{-\frac{\nu t}{2}} \tag{2.20}$$

for $t \geq 1/\nu$. This latter upper bound is within a factor of e of the correct long-term behavior described by (2.14) when $x = 0$.

Theorem 1 provides a solution to the practical question of how large t must be in order that the distance of $X(t)$ be within ϵ in total variation norm from $X(\infty)$. In particular, (2.19) suggests that we choose $t = t(\epsilon)$ according to the formula

$$t(\epsilon) = \frac{2}{\nu}\left(\log\left(\sqrt{\frac{2}{\pi}}\frac{(\eta x+1)}{\epsilon}\right)+\eta x\right), \tag{2.21}$$

provided that $\epsilon \leq \sqrt{\frac{2}{\pi}}(\eta x+1)/e$ (since this choice guarantees that $t(\epsilon) \geq 1/\nu$).

We can now use (2.21) as a guideline for determining when the pre-limit queue has a distribution within ϵ of its equilibrium.

We observe that (2.14) asserts that the total variation convergence rate is minimized asymptotically at $x = 1/\eta = 2EX(\infty)$. Note that (in the case of canonical RBM)

$$
\begin{aligned}
&\mathrm{e}^{-x} u_{-\frac{1}{2}-\frac{z}{t}}(x) \left(\frac{1}{1+\frac{2z}{t}}\right)^{-1} \\
&\quad = \left(1 - \frac{1}{2}\left(\sqrt{\frac{2z}{t}}\right)^2 x^2 + o\left(\frac{1}{t}\right)\right. \\
&\qquad \left. - \sqrt{\frac{t}{2z}}\left(\sqrt{\frac{2z}{t}}x - \frac{1}{6}\left(\sqrt{\frac{2z}{t}}\right)^3 x^3 + o\left(t^{-\frac{3}{2}}\right)\right)\right) \\
&\qquad \cdot \left(1 - \frac{2z}{t} + o\left(\frac{1}{t}\right)\right) \\
&\quad = (1-x)\left(1 - \frac{2z}{t}\right) + \frac{z}{t}\left(-x^2 + \frac{1}{3}x^3\right) + o\left(\frac{1}{t}\right),
\end{aligned}
$$

via a Taylor expansion in $1/t$. Hence

$$
k(t,\ 1,\ y,\ z) = -\frac{2}{3}\frac{z}{t} u_{-\frac{1}{2}-\frac{z}{t}}(y)\mathrm{e}^{-y} + o\left(\frac{1}{t}\right)
$$

as $t \to \infty$. Because

$$
\int_0^\infty z^{\frac{3}{2}} \mathrm{e}^{-z} \mathrm{d}z = \Gamma\left(\frac{5}{2}\right) = \frac{3}{4}\sqrt{\pi},
$$

a dominated convergence argument similar to that used in the proof of Theorem 1 shows that

$$
\mathrm{d}(t,\ 1/\eta) = \sqrt{\frac{2}{\pi}}(\nu t)^{-\frac{5}{2}} \mathrm{e}^{-\frac{\nu t}{2}} + o\left(t^{-\frac{5}{2}} \mathrm{e}^{-\frac{\nu t}{2}}\right)
$$

as $t \to \infty$, thereby yielding the total variation convergence rate associated with the exceptional state $x = 1/\eta = 2EX(\infty)$.

We now generalize our total variation result to the more general weighted variation *h-norm* defined by

$$
\| P_x(X(t) \in \cdot) - P(X(\infty) \in \cdot) \|_h \stackrel{\Delta}{=} \sup_{|f| \le h} |E_x f(X(t)) - Ef(X(\infty))|.
$$

The total variation distance $\mathrm{d}(t,\ x)$ is one half the h-norm corresponding to the envelope function $h = e$, where $e(x) = 1$ for $x \ge 0$. Such h-norms are natural in the queueing context where one is interested in studying convergence rates for unbounded performance measures f (such as $E_x X(t)$). In contrast to the total variation norm, h-norms are typically non-monotone in t.

Theorem 2 *Suppose that $h \ge 1$ is a continuous function for which $hu_{-\frac{\nu}{2}} \in L^1(\pi)$. Then*

$$\|P_x(X(t) \in \cdot) - P(X(\infty) \in \cdot)\|_h$$
$$\leq \sqrt{\frac{2}{\pi}}(\nu t)^{-\frac{3}{2}} e^{-\frac{\nu t}{2}} e^{\eta x}(1+\eta x) \min(\nu t,\ 1) \int_0^\infty \eta e^{-\eta y}(1+\eta y) h(y) dy \quad (2.22)$$

for $t > 0$ *and* $x \geq 0$, *and*

$$\|P_x(X(t) \in \cdot) - P(X(\infty) \in \cdot)\|_h$$
$$= \frac{1}{\sqrt{2\pi}}(\nu t)^{-\frac{3}{2}} e^{-\frac{\nu t}{2}} \left| u_{-\frac{\nu}{2}}(x) \right| \int_0^\infty \left| u_{-\frac{\nu}{2}}(y) \right| h(y) p(y) dy + o\left(t^{-\frac{3}{2}} e^{-\frac{\nu t}{2}}\right)$$

as $t \to \infty$.

The proof of this result is similar to that of Theorem 1 and is therefore omitted.

In many applied settings, modelers choose to simplify a transient analysis by making an equilibrium approximation to the transient performance measure. In particular, if the modeler wishes to be assured that all performance functionals f sitting within an envelope function h have expectation $E_x f(X(t))$ within ϵ of their equilibrium values, the upper bound (2.22) can be used to determine the magnitude of the approximation error.

Example 1 If $h(x) = 1 + x$ for $x \geq 0$ (so that both indicator functions and $E_x X(t)$ are covered), then the upper bound (2.22) is given by

$$\sqrt{\frac{2}{\pi}}(\nu t)^{-\frac{3}{2}} e^{-\frac{\nu t}{2}} e^{\eta x}(1+\eta x) \min(\nu t,\ 1) \frac{3+2\eta}{\eta},$$

which is $1 + \frac{3}{2\eta}$ times as large as that for the total variation distance analogue (in which $h = e$). The time required to make this h-norm less than ϵ is then increased additively by an amount equal to $\log(1 + \frac{3}{2\eta})$ relative to the time $t(\epsilon)$ given by (2.21) for the total variation norm. So, the approximation "cost" to the envelope $h(x) = 1 + x$ relative to the total variation norm is small when $EX(\infty)$ is of moderate size. This suggests that the formula (2.21) is still useful even when dealing with unbounded functionals f.

Example 2 If $h(x) = e^{\theta x}$ for $x \geq 0$, then we need to choose $\theta < \eta$ in order to satisfy the condition of Theorem 2. In this setting, the factor of $1 + \frac{3}{2\eta}$ in Example 1 is replaced by $\eta(2\eta - \theta)/(2(\eta - \theta)^2)$. Again, the time required to make this h-norm less than ϵ relative to the total variation norm is increased additively by the amount $\log(\eta(2\eta-\theta)/(2(\eta-\theta)^2))$, again showing that (2.21) is typically a good convergence guideline even for many unbounded functions.

We conclude this section by noting that an easy consequence of (2.16) is that

$$p(t,x,y) - p(y) = \eta\sqrt{\frac{2}{\pi}}(\nu t)^{-\frac{3}{2}} e^{-\frac{\nu t}{2}} e^{\eta(x-y)}(\eta x - 1)(\eta y - 1) + o\left(t^{-\frac{3}{2}} e^{-\frac{\nu t}{2}}\right) \quad (2.23)$$

as $t \to \infty$, uniformly in x and y on compact intervals. We note that this relation implies that $p(t, x, \cdot)$ crosses $p(\cdot)$ at a point converging (as $t \to \infty$) to $y = \frac{1}{\eta}$, and that the crossing is from above to below when $x < \frac{1}{\eta}$ (and from below to above when $x > \frac{1}{\eta}$). So, the transition density allocates more mass to $[0, 1/\eta]$ for all t sufficiently large when X starts in that interval (with a similar conclusion when X starts in $[1/\eta, \infty)$).

3 Rates of convergence for tail probabilities

We now briefly discuss rates of convergence to equilibrium for tail probabilities of the form $P_x(X(t) > y)$ to $P(X(\infty) > y)$. Such probabilities arise naturally in many application settings, as quality-of-service constraints often involve such probabilities (for example, the requirement that no more than a given proportion of customers experience a delay greater than y).

To provide some additional insight into this question (beyond that of Sect. 2), we consider a "large deviations" setting in which x, y, and t are large. Specifically, we consider an asymptotic regime in which $(x, y, t) = (\tilde{x}, \tilde{y}, \tilde{t})\tau$ with $\tau \to \infty$. Let $a \vee b$ denote $\max(a, b)$ for $a, b \in \mathbb{R}$.

Theorem 3 *If* $(x, y, t) = (\tilde{x}, \tilde{y}, \tilde{t})\tau$, *then*

$$\frac{1}{\tau} \log P_x(X(t) > y) \to \begin{cases} 0, & 0 \le \tilde{t} \le \frac{\tilde{x}-\tilde{y}}{r} \vee 0, \\ -\frac{(\tilde{y}-\tilde{x}+r\tilde{t})^2}{2\sigma^2\tilde{t}}, & \frac{\tilde{x}-\tilde{y}}{r} \vee 0 \le \tilde{t} \le \frac{(\sqrt{\tilde{x}}+\sqrt{\tilde{y}})^2}{r}, \\ -2\eta\tilde{y}, & \tilde{t} \ge \frac{(\sqrt{\tilde{x}}+\sqrt{\tilde{y}})^2}{r}, \end{cases}$$

as $\tau \to \infty$.

Given that $\tau^{-1} \log P(X(\infty) > y) \to -2\eta\tilde{y}$ as $\tau \to \infty$, Theorem 3 suggests that $P_x(X(t) > y)$ equilibrates to $P(X(\infty) > y)$ roughly when

$$t \approx \frac{(\sqrt{x} + \sqrt{y})^2}{r}, \tag{3.1}$$

when x, y, and t are reasonably large. We note that the time to equilibrium is increasing in x. This suggests that when $P_x(X(t) > y)$ is small (and $\{X(t) > y\}$ is a rare event), it can take longer for the probability to reach the equilibrium value when the system is initialized with a large workload.

Proof (*of Theorem* 3) Recall that when X is a canonical RBM,

$$P_x(X(t) > y) = \Phi\left(\frac{-y + x - t}{\sqrt{t}}\right) + \mathrm{e}^{-2y}\Phi\left(\frac{-y - x + t}{\sqrt{t}}\right).$$

In our large deviations scaling, it follows from the Mill's ratio asymptotic for $\Phi(\cdot)$ (see, for example, Grimmett and Stirzaker [18], p. 98) that in this canonical context,

$$\frac{1}{\tau}\log\Phi\left(\frac{-y+x-t}{\sqrt{t}}\right)\to\begin{cases}0, & 0\le\tilde{t}\le(\tilde{x}-\tilde{y})\vee 0,\\ -\frac{(\tilde{y}+\tilde{t}-\tilde{x})^2}{2\tilde{t}}, & \tilde{t}\ge(\tilde{x}-\tilde{y})\vee 0,\end{cases}\tag{3.2}$$

and

$$\frac{1}{\tau}\log\left(\mathrm{e}^{-2y}\Phi\left(\frac{-y-x+t}{\sqrt{t}}\right)\right)\to\begin{cases}-2\tilde{y}-\frac{(\tilde{t}-\tilde{x}-\tilde{y})^2}{2\tilde{t}}, & \tilde{t}\le\tilde{x}+\tilde{y},\\ -2\tilde{y}, & \tilde{t}\ge\tilde{x}+\tilde{y},\end{cases}\tag{3.3}$$

as $\tau\to\infty$. Observe that $-2\tilde{y}-(\tilde{t}-\tilde{x}-\tilde{y})^2/(2\tilde{t})\le-(\tilde{y}+\tilde{t}-\tilde{x})^2/(2\tilde{t})$ for $\tilde{t}\le\tilde{x}+\tilde{y}$, and hence

$$\frac{1}{\tau}\log P_x(X(t)>y)\to\max\left(-\frac{(\tilde{y}+\tilde{t}-\tilde{x})^2}{2\tilde{t}},\ -2\tilde{y}\right)$$

for $\tilde{t}\ge(\tilde{x}-\tilde{y})\vee 0$. The function $-(\tilde{y}+\tilde{t}-\tilde{x})^2/(2\tilde{t})$, viewed as a function of $\tilde{t}$, crosses below $-2\tilde{y}$ precisely at $\tilde{t}=(\sqrt{\tilde{x}}+\sqrt{\tilde{y}})^2$, thereby proving the theorem in view of (3.2) and (3.3). □

The monotonicity in x of the equilibrium time t given by (3.1) can be explained as follows. When $\{X(t)>y\}$ is "rare," the most likely path (when viewed as a function of time) associated with this event is one in which X roughly follows a straight line path from x to y in which X never touches the reflecting boundary at 0. However, when t is large enough, a competing path becomes dominant. In particular, for t sufficiently large, a more probable trajectory generating $\{X(t)>y\}$ is one in which X drifts down from x to the reflecting barrier at 0 according to its natural drift $-r$ and stays in a vicinity of the origin until time $t-y/r$, after which the process increases linearly from level 0 to level y, crossing level y just before time t. The equilibrium time associated with (3.1) is exactly that time t at which the probability of the second competing path that empties the queue overtakes the probability associated with the straight line path. Since the time required for the system to empty is increasing in x, we find that the equilibrium time (3.1) must be monotone in x.

Our next result makes this explanation rigorous. Let

$$\begin{aligned}\phi_1(\tilde{s})&=\tilde{x}-r\tilde{s},\quad 0\le\tilde{s}\le\tilde{t},\\ \phi_2(\tilde{s})&=\tilde{x}+(\tilde{y}-\tilde{x})\frac{\tilde{s}}{\tilde{t}},\quad 0\le\tilde{s}\le\tilde{t},\\ \phi_3(\tilde{s})&=\begin{cases}\tilde{x}-r\tilde{s}, & 0\le\tilde{s}\le\frac{\tilde{x}}{r},\\ 0, & \frac{\tilde{x}}{r}\le\tilde{s}\le\tilde{t}-\frac{\tilde{y}}{r},\\ \tilde{y}+r(\tilde{s}-\tilde{t}), & \tilde{t}-\frac{\tilde{y}}{r}\le\tilde{s}\le\tilde{t},\end{cases}\end{aligned}$$

and put $I_1=(0,\ (\tilde{x}-\tilde{y})/r\vee 0)$, $I_2=((\tilde{x}-\tilde{y})/r\vee 0,\ (\sqrt{\tilde{x}}+\sqrt{\tilde{y}})^2/r)$, and $I_3=((\sqrt{\tilde{x}}+\sqrt{\tilde{y}})^2/r,\ \infty)$.

Theorem 4 *Suppose that* $(x,\ y,\ s,\ t) = (\tilde{x},\ \tilde{y},\ \tilde{s},\ \tilde{t})\tau$ *with* $0 \le \tilde{s} \le \tilde{t}$. *Then, for all* $\epsilon > 0$, $i = 1,\ 2,\ 3$, *and* $\tilde{t} \in I_i$,

$$\frac{1}{\tau} \log P_x \left(\left| \frac{X(s)}{\tau} - \phi_i(\tilde{s}) \right| < \epsilon \,\middle|\, X(t) > y \right) \to 0$$

as $\tau \to \infty$.

Proof We only outline the argument, since it is straightforward. Consider, for example, the case in which $\tilde{t} \in I_3$. For $0 \le \tilde{s} \le \tilde{x}/r$, note that, for $\epsilon > 0$,

$$\begin{aligned} & P_x \left(\left| \frac{X(s)}{\tau} - \phi_3(\tilde{s}) \right| < \epsilon,\ X(t) > y \right) \\ & = P_x \left(\left| \frac{X(s)}{\tau} - (\tilde{x} - r\tilde{s}) \right| < \epsilon,\ X(t) > y \right) \\ & \ge P_x \left(\left| \frac{X(s)}{\tau} - (\tilde{x} - r\tilde{s}) \right| < \epsilon \right) P_{x - rs - \epsilon\tau}(X(t-s) > y). \end{aligned}$$

For all $\epsilon > 0$ small, $(\sqrt{\tilde{x} - r\tilde{s} - \epsilon} + \sqrt{\tilde{y}})^2/r < \tilde{t} - \tilde{s}$, so Theorem 3 implies that

$$\frac{1}{\tau} \log P_{x - rs - \epsilon\tau}(X(t-s) > y) \to -2\eta\tilde{y}$$

as $\tau \to \infty$. Since $P_x(|X(s)/\tau - (\tilde{x} - r\tilde{s})| < \epsilon) \to 1$ as $\tau \to \infty$, the result follows for this combination of $\tilde{s}$ and $\tilde{t}$. The other cases can be similarly handled. □

4 Rates of convergence for integrated performance measures

Modelers frequently replace the exact integrated performance measure

$$E_x \int_0^t f(X(s))\mathrm{d}s$$

by its steady-state approximation. In particular, it is common to use the approximation

$$E_x \int_0^t f(X(s))\mathrm{d}s \approx t E f(X(\infty))$$

when the time horizon t is large.

While one could approach this problem mathematically by directly integrating the spectral representation (2.10), a more natural vehicle here is to use Poisson's equation to study this problem. Specifically, we can seek a function g so that

$$g(X(t)) + \int_0^t (f(X(s)) - E f(X(\infty)))\mathrm{d}s \tag{4.1}$$

is a P_x-martingale adapted to $(\mathscr{F}_t : t \geq 0)$, where $\mathscr{F}_t = \sigma(X(s) : 0 \leq s \leq t)$ is the σ-algebra generated by X to time t. In the presence of such a function g, it follows that

$$E_x \int_0^t f(X(s))\mathrm{d}s - tEf(X(\infty)) = g(x) - E_x g(X(t)). \tag{4.2}$$

We can then use the results of Sect. 2 to analyze the right-hand side of (4.2).

Indeed, suppose that g is twice continuously differentiable. Itô calculus allows us to express (4.1) as

$$\begin{aligned}&\int_0^t (\mathscr{L}g)(X(s))\mathrm{d}s + \int_0^t g'(X(s))\sigma \mathrm{d}B(s) + \int_0^t g'(X(s))\mathrm{d}L(s)\\ &\quad + \int_0^t (f(X(s)) - Ef(X(\infty)))\mathrm{d}s,\end{aligned} \tag{4.3}$$

where $\mathscr{L}$ is the second-order differential operator given by

$$\mathscr{L} = -r\frac{\mathrm{d}}{\mathrm{d}x} + \frac{\sigma^2}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2}.$$

Recalling that L increases only when X is at the origin, we find that

$$\int_0^t g'(X(s))\mathrm{d}L(s) = g'(0)(L(t) - L(0)).$$

Given that the stochastic integral involving the integrator $\mathrm{d}B(s)$ is always a local martingale, if we choose g to satisfy

$$\begin{aligned}&(\mathscr{L}g)(x) = -(f(x) - Ef(X(\infty)))\\ &\text{s.t. } g'(0) = 0,\end{aligned} \tag{4.4}$$

then (4.3) will be a local martingale; (4.4) is *Poisson's equation* (see Glynn and Meyn [15] and the references therein) corresponding to the right-hand side $-f_c$, where $f_c(x) \stackrel{\Delta}{=} f(x) - Ef(X(\infty))$. Conveniently, Eq. (4.4) can be solved explicitly (for all nonnegative measurable functions f), yielding

$$g(x) = g(0) - \frac{1}{r}\int_0^x f_c(y)(\mathrm{e}^{-2\eta(y-x)} - 1)\mathrm{d}y. \tag{4.5}$$

Because $Ef_c(X(\infty)) = 0$, result (4.5) can further be rewritten as

$$g(x) = g(0) + \frac{\mathrm{e}^{2\eta x}}{r}\int_x^\infty f_c(y)\mathrm{e}^{-2\eta y}\mathrm{d}y + \frac{1}{r}\int_0^x f_c(y)\mathrm{d}y. \tag{4.6}$$

We are now free to choose $g(0) = 0$.

Consider the special case in which $f:[0,\infty)\to[0,1]$ is continuously differentiable. Then, g is twice continuously differentiable and (4.3) is a P_x-martingale (not just local martingale), since the boundedness of f and result (4.6) ensure that

$$\int_0^t E_x g'(X(s))^2 \mathrm{d}s < \infty$$

for all $t > 0$. Moreover, by relation (4.2),

$$\begin{aligned}&\int_0^\infty f(z)\left(\frac{1}{t}\int_0^t P_x(X(s)\in \mathrm{d}z)\mathrm{d}s - P(X(\infty)\in \mathrm{d}z)\right)\\ &\quad=\int_0^\infty g(z)\frac{1}{t}(P_x(X(0)\in \mathrm{d}z) - P_x(X(t)\in \mathrm{d}z)),\end{aligned}$$

where g admits the integral representation in terms of f as specified by (4.5). Since this equality of measures holds for every bounded continuously differentiable f, it holds for all nonnegative measurable functions f. Hence, for all (measurable) indicator functions f,

$$\begin{aligned}&\left|\frac{1}{t}E_x\int_0^t f(X(s))\mathrm{d}s - Ef(X(\infty))\right|\\ &\quad\le\frac{1}{t}(|g(x)| + E|g(X(\infty))| + |E_x g(X(t)) - Eg(X(\infty))|),\end{aligned}\tag{4.7}$$

where g is computed via relation (4.5).

To obtain an upper bound for the right-hand side of (4.7), we note that if f is an indicator function, then $|f_c(x)|\le 1$ and

$$\begin{aligned}|g(x)| &\le \frac{1}{r}\int_0^\infty \mathrm{e}^{-2\eta y+2\eta(x\wedge y)}\mathrm{d}y\\ &=\frac{x}{r}+\frac{1}{2\nu},\end{aligned}$$

where $a\wedge b\stackrel{\Delta}{=}\min(a,\,b)$ for $a,\,b\in\mathbb{R}$. Moreover, (2.18) implies that

$$\begin{aligned}&|E_x g(X(t)) - Eg(X(\infty))|\\ &\quad\le\sqrt{\frac{2}{\pi}}\mathrm{e}^{-\frac{\nu t}{2}}\mathrm{e}^{\eta x}(1+\eta x)\int_0^\infty \eta\mathrm{e}^{-\eta y}(1+\eta y)\left(\frac{y}{r}+\frac{1}{2\nu}\right)\mathrm{d}y\\ &\quad=\frac{4}{\nu}\sqrt{\frac{2}{\pi}}\mathrm{e}^{-\frac{\nu t}{2}}\mathrm{e}^{\eta x}(1+\eta x)\end{aligned}$$

for $\nu t\ge 1$. So, we are led to Proposition 2.

Proposition 2 *For all $x \geq 0$ and $\nu t \geq 1$,*

$$\left\| \frac{1}{t} \int_0^t P_x(X(s) \in \cdot) \mathrm{d}s - P(X(\infty) \in \cdot) \right\| \leq \frac{1}{t} \left(\frac{x}{r} + \frac{3}{2\nu} + \frac{4}{\nu} \sqrt{\frac{2}{\pi}} \mathrm{e}^{-\frac{\nu t}{2}} \mathrm{e}^{\eta x} (1 + \eta x) \right).$$

From Proposition 2, we can easily compute a bound on the time $t = t(\epsilon)$ needed for the total variation distance of $t^{-1} \int_0^t P_x(X(s) \in \cdot) \mathrm{d}s$ to equilibrium to be smaller than any given $\epsilon > 0$.

Remark 3 We derive our theory here for general RBM (as opposed to the canonical case), because the additional notational burden is light, and the theory here does not deal with density-related issues where the scaling relationship (2.4) applies. (Of course, scaling relationships for Sect. 4's quantities could also be derived if we wished, again reducing the calculations to the canonical setting.)

We can similarly compute the h-norm distance between $t^{-1} \int_0^t P_x(X(s) \in \cdot) \mathrm{d}s$ and the stationary distribution π, allowing us to develop uniform bounds on the rate of convergence to equilibrium for time averages of all unbounded performance functionals f bounded by a suitable continuous envelope function h for which

$$\int_0^\infty y h(y) \mathrm{e}^{-\eta y} \mathrm{d}y < \infty. \tag{4.8}$$

In particular, for any continuously differentiable f for which $|f| \leq h$, (4.6) implies that if we set $g(0) = 0$, then

$$|g(x)| \leq \frac{2}{r} \int_0^\infty h(y) \mathrm{e}^{-2\eta y + 2\eta(x \wedge y)} \mathrm{d}y.$$

We now apply (2.18) to obtain the bound

$$\begin{aligned} |E_x g(X(t)) - E g(X(\infty))| &\leq \sqrt{\frac{2}{\pi}} \mathrm{e}^{-\frac{\nu t}{2}} \mathrm{e}^{\eta x} (1 + \eta x) \int_0^\infty \eta \mathrm{e}^{-\eta y} (1 + \eta y) \frac{2}{r} \\ &\quad \cdot \int_0^\infty h(z) \mathrm{e}^{-2\eta z + 2\eta(y \wedge z)} \mathrm{d}z \mathrm{d}y \\ &= \frac{4}{r} \sqrt{\frac{2}{\pi}} \mathrm{e}^{-\frac{\nu t}{2}} \mathrm{e}^{\eta x} (1 + \eta x) \int_0^\infty h(z)(1 + \eta z) \mathrm{e}^{-\eta z} \mathrm{d}z. \end{aligned}$$

Finally, in view of the fact that

$$E|g(X(\infty))| \leq \frac{2}{r} \int_0^\infty (1 + 2\eta y) \mathrm{e}^{-2\eta y} h(y) \mathrm{d}y,$$

we obtain Proposition 3. (As in the proof of Proposition 2, although the upper bound in Proposition 3 is derived for continuously differentiable functions, it in fact holds for general nonnegative functions.)

Proposition 3 *For all $x \geq 0$ and $\nu t \geq 1$,*

$$\begin{aligned}
&\left\| \frac{1}{t} \int_0^t P_x(X(s) \in \cdot) \mathrm{d}s - P(X(\infty) \in \cdot) \right\|_h \\
&\quad \leq \frac{1}{t} \left(\frac{2}{r} \int_0^\infty h(y) \mathrm{e}^{-2\eta y + 2\eta(x \wedge y)} \mathrm{d}y + \frac{2}{r} \int_0^\infty (1 + 2\eta y) \mathrm{e}^{-2\eta y} h(y) \mathrm{d}y \right. \\
&\qquad \left. + \frac{4}{r} \sqrt{\frac{2}{\pi}} \mathrm{e}^{-\frac{\nu t}{2}} \mathrm{e}^{\eta x} (1 + \eta x) \int_0^\infty h(z)(1 + \eta z) \mathrm{e}^{-\eta z} \mathrm{d}z \right).
\end{aligned} \tag{4.9}$$

As in the total variation context, we can use (4.9) to obtain an upper bound on the time $t = t(\epsilon)$ required for the h-norm to be less than any given $\epsilon > 0$.

Of course, the dominant terms in the bound (4.9) are the first two terms on the right-hand side. Below we compute these two terms for several important envelope functions.

Example 3 When $h(x) = 1 + x^p$ for $p > 0$ (as occurs with performance measures that correspond to moments),

$$\begin{aligned}
&|g(x)| + E|g(X(\infty))| \\
&\quad \leq \frac{2}{r} \int_0^x (1 + y^p) \mathrm{d}y + \frac{\mathrm{e}^{2\eta x}}{\eta r} \int_{2\eta x}^\infty \left(1 + \frac{y^p}{(2\eta)^p} \right) \mathrm{e}^{-y} \mathrm{d}y \\
&\qquad + \frac{2}{r} \int_0^\infty (1 + 2\eta y) \mathrm{e}^{-2\eta y} (1 + y^p) \mathrm{d}y \\
&\quad = \frac{2x^{1+p}}{(1+p)r} + \frac{2x}{r} + \frac{1}{\nu} \mathrm{e}^{2\eta x} (\mathrm{e}^{-2\eta x} + (2\eta)^{-p}) \Gamma(1 + p,\ 2\eta x) \\
&\qquad + \frac{1}{\nu} \left(2 + \frac{(2\eta)^{-p} \Gamma(3 + p)}{1 + p} \right),
\end{aligned}$$

where $\Gamma(s,\ x) = \int_x^\infty y^{s-1} \mathrm{e}^{-y} \mathrm{d}y$ denotes the *incomplete Gamma function*. In particular, if $p = 1$, then

$$|g(x)| + E|g(X(\infty))| \leq \frac{x^2}{r} + \frac{(2r + \sigma^2)x}{r^2} + \frac{3r\sigma^2 + 2\sigma^4}{r^3}.$$

For $p = 2$,

$$|g(x)| + E|g(X(\infty))| \leq \frac{2x^3}{3r} + \frac{x^2}{\nu} + \frac{(2r^2 + \sigma^4)x}{r^3} + \frac{6r^2\sigma^2 + 5\sigma^6}{2r^4}.$$

Example 4 For exponential moments, an appropriate envelope function is one of the form $h(x) = e^{\theta x}$ with $\theta < 2\eta$. (If we were imposing the conditions needed for $E_x g(X(t))$ to converge to $Eg(X(\infty))$ exponentially fast at rate $\nu/2$, then we would require the stronger condition $\theta < \eta$.) In this case,

$$|g(x)| + E|g(X(\infty))| = \frac{2(2\eta(e^{\theta x} - 1) + \theta)}{\theta r(2\eta - \theta)} + \frac{2(4\eta - \theta)}{r(2\eta - \theta)^2}.$$

We conclude this section by noting that the distance to equilibrium for integrated performance measures of RBM decreases in proportion to t^{-1}, whereas for instantaneous performance measures, it tends to zero exponentially fast. Related results can be found in Thorisson [26]. So, when ϵ is small, the time horizon at which an integrated performance measure can be replaced by its equilibrium analogue is much larger than that required for an instantaneous performance measure.

5 Simulation implications

In Sects. 2, 3 and 4, we have discussed approximations for the distance to equilibrium for instantaneous performance measures, tail probabilities and integrated performance measures. These results were motivated by operations management and queueing applications in which the natural decision horizon leads to consideration of transient quantities, and the intent is to determine the approximation error induced when the transient quantity is replaced by its corresponding steady-state quantity.

In this section, we change our perspective somewhat. In particular, computing transient performance measures via simulation is often easier than computing steady-state quantities via simulation, because transient quantities can easily be generated in finite time. On the other hand, generating rv's from the steady-state distribution is impossible in general (see Asmussen et al. [7]), so that in the simulation setting, the natural question is how long a transient simulation must be run, in order that a good approximation to the steady state be obtained. This question is variously known, in the simulation literature, as the *warm-up problem*, the *start-up problem*, or the *initial transient problem* (see Wang and Glynn [28] and the references therein for additional discussion).

This problem also arises in the setting of Markov chain Monte Carlo (MCMC). In that context, the goal is to sample from a Bayesian posterior distribution, and the computational approach involves simulating a Markov process chosen in such a way that its stationary distribution coincides with the desired posterior. In the MCMC setting, one can always construct the sampler so as to guarantee that it is a reversible Markov process. This ensures a spectral representation for the transition density and allows for the development of special-purpose mathematical tools (such as Cheeger's inequality) for bounding the "spectral gap" that separates the dominant eigenvalue 0 (associated with the stationary distribution) from the rest of the spectrum. With the use of such tools, one can now sometimes obtain useful bounds on the rate of convergence to equilibrium.

In simulating queueing models, the typical model is not a reversible Markov process. Thus, it is unclear how one can use Markov process theory to usefully generate insights into the initial transient problem in this simulation context. The approach we take in this paper is to use the RBM as a mathematical vehicle for generating practically useful guidelines that can be used for such steady-state simulations. Specifically, we take the view that if our underlying queue can be weakly approximated via RBM, then one can use the time to equilibrium for the RBM to generate a guideline for the underlying queueing model. This perspective is very similar in philosophy to the approach followed by Whitt [29] and Asmussen [6] in analyzing the interaction between heavy-traffic effects and steady-state simulation run-length determination in queues.

In computing $\alpha \stackrel{\Delta}{=} Ef(X(\infty))$, the single replication strategy involves generating $X(0)$ from a distribution μ (say), simulating X to time t, and utilizing the time average $t^{-1}\int_0^t f(X(s))\mathrm{d}s$ as an estimator for α. An alternative is to delete the first $s(t)$ time units from the observed time series, and to instead use the modified quantity $(t - s(t))^{-1}\int_{s(t)}^t f(X(s))\mathrm{d}s$ as an estimator for α. Not surprisingly, unless $X(0)$ is chosen badly (for example, much larger than $EX(\infty)$), very little is gained through the deletion of the observations associated with the time interval $[0, s(t)]$, because the biggest issue governing the quality of the estimator is its variability, rather than its bias. From a variability viewpoint, we wish to average over as much data as possible, leading to the conclusion that $s(t)$ should be small or even zero. Mean square error calculations supporting this "no deletion" conclusion can be found in Section 6 of Wang and Glynn [27]; see Grassmann [17] for further discussion of this point.

However, the situation changes in the multiple replication setting, because bias due to initialization effects can then have a substantial effect on estimator quality. Such multiple replication strategies are becoming increasingly attractive, because of the emerging prevalence of multi-processor computing platforms on which independent replications can potentially be run on each of the available processors. Assuming that one estimates a steady-state expectation $\alpha = Ef(X(\infty))$ by simulating independent and identically distributed (iid) replicates $X_1, \ldots, X_p$ of X to time t on each of p processors and retains all the simulated data, the corresponding estimator for α is then given by

$$\alpha(p, t) \stackrel{\Delta}{=} \frac{1}{pt}\sum_{i=1}^{p}\int_0^t f(X_i(s))\mathrm{d}s.$$

Suppose that $|f| \leq h$, where h satisfies (4.8). If all p simulations are initialized with $X_i(0) = x$, then the bias of $\alpha(p, t)$ is

$$E\alpha(1, t) - \alpha = \frac{1}{t}g_c(x) + O\left(\mathrm{e}^{-\frac{\nu t}{2}}\right) \tag{5.1}$$

as $t \to \infty$, uniformly in p, where $O(c(t))$ is a function such that $O(c(t))/c(t)$ is bounded as $t \to \infty$, and $g_c(\cdot)$ is defined as in Sect. 4.

Furthermore, if we strengthen (4.8) so that

$$h(y) = O(\mathrm{e}^{\theta y}) \tag{5.2}$$

as $y \to \infty$ for some $\theta < \eta$, then we are in a position to analyze the variability of $\alpha(p,\ t)$ via the central limit theorem (CLT). Our proof uses a mixing argument rather than the martingale representation (4.1) (followed by an application of the martingale CLT) so as to avoid a need to assume that f_c is sufficiently smooth so that Itô's formula can be applied directly to g_c. (Note that (4.4) means that g_c's second derivative is essentially $-f_c$.)

Let $P_\pi(\cdot) = \int_0^\infty P_x(\cdot)\pi(\mathrm{d}x)$ and let $E_\pi(\cdot)$ be the corresponding expectation operator.

Theorem 5 *Suppose that f satisfies (5.2). Then,*

$$\sqrt{t}\left(\frac{1}{t}\int_0^t f(X(s))\mathrm{d}s - \alpha\right) \Rightarrow aN(0,\ 1) \tag{5.3}$$

as $t \to \infty$, where

$$a^2 = 2\int_0^\infty E_\pi\, f_c(X(0))\, f_c(X(t))\mathrm{d}t.$$

Furthermore, if p is such that $\sqrt{pt}(E_x\alpha(1,\ t) - \alpha) \Rightarrow 0$ as $t \to \infty$, then

$$\sqrt{pt}(\alpha(p,\ t) - E\alpha(1,\ t)) \Rightarrow aN(0,\ 1) \tag{5.4}$$

as $t \to \infty$.

Proof We first apply (2.13), thereby allowing us to conclude that for each (measurable) subset A and $n \geq 0$,

$$\left|P_x\left(\int_n^{n+1} f_c(X(s))\mathrm{d}s \in A\right) - P_\pi\left(\int_0^1 f_c(X(s))\mathrm{d}s \in A\right)\right| \leq c_1\mathrm{e}^{\eta x}(1+\eta x)\mathrm{e}^{-\frac{\nu n}{2}}$$

for some constant $c_1 > 0$. So, for $1 \leq q < 2$,

$$\begin{aligned}&\phi_q(n)\\ &\quad\stackrel{\Delta}{=} \sup_A E_\pi^{\frac{1}{q}}\left|P_\pi\left(\int_n^{n+1} f_c(X(s))\mathrm{d}s \in A \,\middle|\, X(0)\right) - P_\pi\left(\int_0^1 f_c(X(s))\mathrm{d}s \in A\right)\right|^q\\ &\quad\leq c_2\mathrm{e}^{-\frac{\nu n}{2}} E_\pi^{\frac{1}{q}}\mathrm{e}^{\eta q X(0)}(1+\eta X(0))\end{aligned}$$

for some constant $c_2 > 0$. We now choose $\delta > 0$ sufficiently small so that $\theta(2+\delta) < 2\eta$. If we set $q = (2+\delta)/(1+\delta)$, we find that

$$\sum_{n=1}^{\infty} \phi_q(n)^{\frac{\delta}{1+\delta}} < \infty.$$

In addition,

$$\left| \int_0^1 f_c(X(s))\mathrm{d}s \right|^{2+\delta} \leq \int_0^1 |f_c(X(s))|^{2+\delta}\mathrm{d}s,$$

so

$$E_\pi \left| \int_0^1 f_c(X(s))\mathrm{d}s \right|^{2+\delta} \leq E_\pi |f_c(X(0))|^{2+\delta} < \infty.$$

Hence, Theorem 3.1, p. 351 of Ethier and Kurtz [13] can be applied, yielding the CLT (5.3) under P_π. Because $X(t)$ converges in total variation to $X(\infty)$ under P_x (see Sect. 2), X can be coupled to a stationary version of X; see Thorisson [26]. The coupling then easily leads to (5.3) under P_x.

To prove (5.4), we first develop an additional mixing result to complement those of Sect. 2. Let w satisfy $\theta + \eta < w < 2\eta$, where θ is given as in (5.2). Set $v(x) = 2\mathrm{e}^{wx} - 2wx + 1$, and note that we have constructed v so that $v'(0) = 0$ and it is positive over $[0, \infty)$. In fact, $v(x) \geq \mathrm{e}^{wx}$ for $x \geq 0$. Furthermore,

$$(\mathscr{L}v)(x) \leq -cv(x) + d$$

for some $c, d > 0$ and all $x \geq 0$. Consequently, Theorem 6.1 of Meyn and Tweedie [24] can be applied, so that there exist constants $\kappa > 0$ and $c_3 > 0$ such that

$$\|P_x(X(t) \in \cdot) - P(X(\infty) \in \cdot)\|_v \leq c_3(v(x)+1)\mathrm{e}^{-\kappa t} \tag{5.5}$$

for all $t, x \geq 0$.

To now verify (5.4), we apply the Lindeberg–Feller theorem; see p.214-215 of Chung [9]. This theorem can be applied to obtain (5.4), provided that $(t(\alpha(1,t) - E\alpha(1,t))^2 : t > 0)$ is P_x-uniformly integrable. In view of our assumption on p, this is equivalent to proving that $(t(\alpha(1,t) - \alpha)^2 : t > 0)$ is P_x-uniformly integrable, which is, in turn, equivalent to showing that

$$\frac{1}{t} E_x \left(\int_0^t f_c(X(s))\mathrm{d}s \right)^2 \to a^2 \tag{5.6}$$

as $t \to \infty$; see p. 101 of Chung [9].

Note that (5.2) allows us to apply Theorem 2, thereby yielding

$$|E_x f_c(X(t))| \leq c_4 \mathrm{e}^{-\frac{vt}{2}} \mathrm{e}^{\eta x}(1+\eta x)$$

for some constant $c_4 > 0$. Hence,

$$\begin{aligned}\left|e^{\frac{vt}{2}} f_c(x) E_x f_c(X(t))\right| &\le c_5 e^{(\theta+\eta)x}(1+\eta x)\\ &\le c_6 e^{wx}\end{aligned}$$

for suitable constants $c_5,\ c_6 > 0$. The bound (5.5) therefore implies that, for $0 \le s \le u$,

$$\begin{aligned}&\left|e^{\frac{v(u-s)}{2}} E_x f_c(X(s)) f_c(X(u)) - e^{\frac{v(u-s)}{2}} E_\pi f_c(X(0)) f_c(X(u-s))\right|\\ &\le c_3(v(x)+1)e^{-\kappa s}\\ &\le c_7 e^{-\kappa s} e^{wx}\end{aligned}$$

for some constants $c_7 > 0$. It follows that there exists $\beta > 0$ for which

$$|E_x f_c(X(s)) f_c(X(u)) - E_\pi f_c(X(0)) f_c(X(u-s)) \le c_7 e^{-\beta u} e^{wx}. \tag{5.7}$$

An immediate consequence of (5.7) is that

$$\frac{1}{t} E_x \left(\int_0^t f_c(X(s))\mathrm{d}s\right)^2 - \frac{1}{t} E_\pi \left(\int_0^t f_c(X(s))\mathrm{d}s\right)^2 \to 0$$

as $t \to \infty$. Finally, it is straightforward to show that

$$\frac{1}{t} E_\pi \left(\int_0^t f_c(X(s))\mathrm{d}s\right)^2 \to a^2$$

as $t \to \infty$, proving (5.6) and completing the argument to justify (5.4). □

Related CLTs are discussed in Whitt [29] and Asmussen [6], but their discussion focuses exclusively on $f(x) = x$ (and on interchanging the heavy-traffic limit and $t \to \infty$).

Note that (5.1) and (5.4) together imply that when the number of processors p is of the order t or larger, the bias becomes a dominant source of error in the estimator $\alpha(p,\ t)$'s ability to accurately compute α. In this case, choosing a good starting state so as to minimize $|g_c(\cdot)|$ becomes important. The examples below show that the best possible starting state x depends heavily on the choice of f.

Example 5 When $f(x) = x$,

$$g_c(x) = \frac{x^2}{2r} - \frac{1}{4\eta^2 r}$$

and the unique zero of $g_c(\cdot)$ occurs at $x = \frac{\sigma}{\sqrt{2r}}$ ($x \approx 0.7071$ when $r = \sigma = 1$).

Example 6 Let $f(x) = x^2$. Then,

$$g_c(x) = \frac{x^3}{3r} + \frac{x^2}{2\eta r} - \frac{1}{2\eta^3 r}$$

and the unique zero of $g_c(\cdot)$ occurs at $x \approx 0.8064$ (when $r = \sigma = 1$).

Example 7 Suppose that $f(x) = \mathrm{e}^{\theta x}$ for $\theta < \eta$. Then,

$$g_c(x) = \frac{-\theta^2 + (4\eta^2 - 2\theta\eta)(\mathrm{e}^{\theta x} - \theta x - 1)}{\theta r(2\eta - \theta)^2}.$$

If $r = \sigma = 1$ and $\theta = 1/2$, then the unique zero of $g_c(\cdot)$ occurs at $x \approx 0.7645$.

Example 8 Let $f(x) = \mathbb{I}(x > b)$ for $b \geq 0$. In this setting,

$$g_c(x) = \begin{cases} \frac{\mathrm{e}^{-2\eta b}(\mathrm{e}^{2\eta x} - 1 - 2\eta(b+x))}{2\eta r}, & x < b, \\ \frac{\mathrm{e}^{-2\eta b}(\mathrm{e}^{2\eta b}(-2\eta b + 2\eta x + 1) - 2\eta(b+x) - 1)}{2\eta r}, & x \geq b, \end{cases}$$

The unique zero of $g_c(\cdot)$ occurs at $x \approx 0.7526$ when $r = \sigma = 1$ and $b = 1$. When $b = 2$ (and $r = \sigma = 1$), $x \approx 0.9684$. When $b = 10$ (and $r = \sigma = 1$), $x \approx 1.5929$.

Of course, in implementing these recommended starting values in a queue that can be well approximated by RBM, one substitutes the values of r and σ given by (1.3) into the above formulae.

An alternative strategy is to delete the initial time segment $[0,\ s(t)]$ from each of the independent simulations generated by the p processors, thereby yielding

$$\tilde{\alpha}(p,\ t) = \frac{1}{p(t - s(t))} \sum_{i=1}^{p} \int_{s(t)}^{t} f(X_i(s))\mathrm{d}s.$$

If $s(t)/t \to 0$ as $t \to \infty$, then (5.4) holds with $\tilde{\alpha}(p,\ t)$ substituting for $\alpha(p,\ t)$ (and $E\tilde{\alpha}(1,\ t)$ substituting for $E\alpha(1,\ t)$). So, this deletion has no asymptotic effect on the variability of $\tilde{\alpha}(p,\ t) - E\tilde{\alpha}(1,\ t)$ (which is of order $(pt)^{-\frac{1}{2}}$). However, the bias is significantly reduced. In particular,

$$E_x\tilde{\alpha}(1,\ t) - \alpha = \frac{1}{t - s(t)} \int_{s(t)}^{t} E_x f_c(X(s))\mathrm{d}s. \tag{5.8}$$

Under condition (5.2), (2.22) applies, so that

$$\begin{aligned} E_x\tilde{\alpha}(1,\ t) - \alpha &= \frac{1}{t - s(t)} \int_{s(t)}^{t} O\left(s^{-\frac{3}{2}}\mathrm{e}^{-\frac{vs}{2}}\right)\mathrm{d}s \\ &= O\left((s(t))^{-\frac{3}{2}}\mathrm{e}^{-\frac{vs(t)}{2}}\right) \end{aligned}$$

as $t \to \infty$.

Suppose now that the number of processors p is large and that $t = p^w$ for $w > 0$. If we choose $s(t) = k \log t$ with $k > (1 + w)/(\nu w)$, then the bias of $\tilde{\alpha}(p, t)$ is of smaller asymptotic order than its variability $(pt)^{-\frac{1}{2}}$. Because we can now choose w as small as we wish (by choosing k correspondingly large), we note that the completion time t for such a parallel computation can be made small, provided that one uses such an initial bias deletion strategy.

So, in the multiple replication setting, deletion of the segment $[0, s(t)]$ is a sensible strategy (unlike the single replication context). For queues that are well approximated by RBM, we note that the bias (5.8) can be made uniformly small over all functions f_c satisfying (5.2); see Theorem 2. (This uniformity is important when we wish to estimate multiple expectations $Ef(X(\infty))$ from the same simulation runs.) Theorem 2 further suggests that an especially good choice of starting state is then given by $x = 1/\eta = 2EX(\infty)$ and that this choice will work uniformly over all functions f dominated by an envelope function satisfying (5.2). This is in sharp contrast to our earlier analysis of $\alpha(p, t)$, where the choice of starting state depends heavily on the specific choice of f.

In implementing these ideas in a queueing context, one could either approximate $2EX(\infty)$ via σ^2/r or estimate $2EX(\infty)$ (crudely) by running a few preliminary simulation runs to obtain a rough estimate of $2EX(\infty)$, followed by simulating the "production runs" described above.

6 Spectral theory for RBM

The spectral representation of Sect. 2 is a special case of the more general spectral representations that are typically available for one-dimensional reversible Markov processes. However, as we shall see below, there are several subtle issues that arise in the RBM setting that are worth illuminating.

The simplest setting in which spectral representations arise is that of an irreducible, aperiodic and reversible finite state Markov chain $Y = (Y_n : n \geq 0)$ having one step transition matrix $R = (R(x, y) : x, y \in \mathbb{S})$, where $\mathbb{S}$ is the state space of Y. The reversibility implies the existence of real eigenvalues $\lambda_1, \ldots, \lambda_d$ of R (with $d = |\mathbb{S}|$), and corresponding real column eigenvectors $u_1, \ldots, u_d$, such that the u_i are orthonormal with respect to the inner product induced by the stationary distribution $(\pi(x) : x \in \mathbb{S})$ associated with R. In other words, $\langle u_i, u_j \rangle = 0$ if $i \neq j$ and $\langle u_i, u_j \rangle = 1$ if $i = j$, where

$$\langle f, g \rangle \stackrel{\Delta}{=} \sum_{x \in \mathbb{S}} \pi(x) f(x) g(x).$$

Because R is stochastic, one of the eigenvalues, say λ_1, must be equal to one, and its corresponding orthonormalized eigenvector is $u_1(x) \equiv 1$ for $x \in \mathbb{S}$. We can then assume that the eigenvalues have been indexed so that $1 = \lambda_1 > |\lambda_2| \geq \cdots \geq |\lambda_d|$. It follows that, for any f,

$$R^n f = \sum_{i=1}^{d} \langle f,\ u_i \rangle R^n u_i$$
$$= \sum_{i=1}^{d} \langle f,\ u_i \rangle \lambda_i^n u_i,$$

and hence (by specializing f to an indicator on the state y), we find that

$$R^n(x,\ y) - \pi(y) = \pi(y) \sum_{i=2}^{d} \lambda_i^n u_i(x) u_i(y). \tag{6.1}$$

The spectral representation (6.1) makes the computation of the rate of convergence to equilibrium relatively straightforward. Much of the literature on Markov chain Monte Carlo rests on this observation; see, for example, the Cheeger bound of Diaconis and Stroock [12].

The finite state perspective generalizes, in a natural way, to one-dimensional diffusions (which are an especially tractable class of reversible Markov processes). In this context, the transition matrix R is replaced by the infinitesimal generator of the diffusion, together with whatever restrictions must be placed on the domain of the generator. In the RBM setting, $\mathscr{L}$ replaces R, and the domain consists of twice continuously differentiable functions having a vanishing derivative at the origin.

In view of our finite state discussion, we are therefore led to the consideration of the eigenvalues of $\mathscr{L}$, so that we wish to find values of λ for which there exists a corresponding twice continuously differentiable eigenfunction u_λ for which

$$\begin{aligned} &\mathscr{L} u_\lambda = \lambda u_\lambda \\ &\text{subject to} \quad u_\lambda'(0) = 0. \end{aligned} \tag{6.2}$$

If there exists a solution to the eigenvalue problem (6.2), then we say that λ is a *formal eigenvalue* of $\mathscr{L}$.

Remark 4 As in Sect. 4, we derive the theory directly for the general case, because the additional notational burden is light.

Proposition 4 *Every $\lambda \in \mathbb{R}$ is a formal eigenvalue of $\mathscr{L}$.*

Proof The functions u_λ defined in (2.11) and (2.12) solve (6.2) for $\lambda \leq -\nu/2$. For $\lambda > -\nu/2$, the corresponding eigenfunction

$$u_\lambda(x) = \frac{(\eta + \beta(\lambda))\mathrm{e}^{x(\eta - \beta(\lambda))} - (\eta - \beta(\lambda))\mathrm{e}^{x(\eta + \beta(\lambda))}}{2\beta(\lambda)}$$

solves (6.2), where $\beta(\lambda) = \sqrt{2\lambda + \nu}/\sigma$. □

If λ is such that $\mathrm{e}^{-\lambda t} u_\lambda(X(t))$ is a P_x-martingale, then we call λ a *probabilistic eigenvalue* of $\mathscr{L}$.

Proposition 5 *Every $\lambda \in \mathbb{R}$ is a probabilistic eigenvalue of $\mathscr{L}$.*

Proof According to Itô's formula,

$$\begin{aligned} \mathrm{d}(\mathrm{e}^{-\lambda t} u_\lambda(X(t))) = \mathrm{e}^{-\lambda t}(\mathscr{L}u_\lambda - \lambda u_\lambda)(X(t))\mathrm{d}t + \mathrm{e}^{-\lambda t} u'_\lambda(X(t))\sigma \mathrm{d}B(t) \\ + \mathrm{e}^{-\lambda t} u'_\lambda(X(t))\mathrm{d}L(t), \end{aligned}$$

so that (6.2) implies that

$$\mathrm{e}^{-\lambda t} u_\lambda(X(t)) = u_\lambda(X(0)) + \int_0^t \mathrm{e}^{-\lambda s} u'_\lambda(X(s))\sigma \mathrm{d}B(s).$$

Since $u_\lambda(\cdot)$ grows (at most) exponentially, and $X(\cdot)$ has tails that are uniformly dominated on compact time intervals by Gaussian tails, it follows that

$$\int_0^t \mathrm{e}^{-2\lambda s} E_x u'_\lambda(X(s))^2 \mathrm{d}s < \infty$$

for each $t \geq 0$, so that $\mathrm{e}^{-\lambda t} u_\lambda(X(t))$ is a square-integrable martingale with respect to P_x for each $x \geq 0$. □

It turns out that these approaches to defining the spectrum of $\mathscr{L}$ are not appropriate from a functional analysis perspective. Recall that our finite state development rests on the inner product $(f,\ g)$, so that the appropriate generalization to the RBM setting naturally relies upon the Hilbert space $L^2(\pi)$ and its associated inner product $\langle f,\ g\rangle$. Hence, a better definitional approach takes the view that λ lies in the *spectrum* of $\mathscr{L}$ whenever $\mathscr{L} - \lambda I$ fails to have a bounded inverse on $L^2(\pi)$, where I is the identity operator. Put $\|f\| = \sqrt{\langle f,\ f\rangle}$ for $f \in L^2(\pi)$.

Theorem 6 *Let $\mathscr{S} = \{0\} \cup \left(-\infty,\ -\frac{\nu}{2}\right]$. For $\lambda \notin \mathscr{S}$, there exists a twice continuously differentiable function $g = g_{\lambda,\, f}$ solving*

$$\begin{aligned} &\mathscr{L}g - \lambda g = -f \\ &\textit{subject to} \quad g'(0) = 0, \end{aligned} \tag{6.3}$$

whenever $f \in L^2(\pi)$ is a continuous function. Furthermore,

$$\sup_{\|f\| \neq 0} \frac{\|g_{\lambda,\, f}\|}{\|f\|} < \infty. \tag{6.4}$$

On the other hand, when $\lambda \in \mathscr{S}$, there exist continuous functions $f \in L^2(\pi)$ for which no $g \in L^2(\pi)$ solves (6.3).

Proof For $\lambda \notin \mathscr{S}$, put

$$G(x,\ y,\ \lambda) = \frac{1}{2r\beta(\lambda)} \mathrm{e}^{\eta(x+y) - \beta(\lambda)|x-y|} + \frac{\lambda + r\beta(\lambda) + \nu}{2\lambda r\beta(\lambda)} \mathrm{e}^{(x+y)(\eta - \beta(\lambda))}$$

and set

$$g_{\lambda,\,f}(x) = \int_0^\infty G(x,\ y,\ \lambda) f(y)\pi(\mathrm{d}y). \tag{6.5}$$

Then, $g_{\lambda,\,f} = g$ satisfies (6.3). Furthermore, for $\lambda > 0$, Itô's formula and (6.3) imply that

$$g(x) = \int_0^\infty \mathrm{e}^{-\lambda t} E_x f(X(t))\mathrm{d}t,$$

and hence

$$g(x) = \frac{1}{\lambda} E_x f(X(\xi)),$$

where ξ is an exponential rv with rate parameter λ independent of X. The Cauchy–Schwarz inequality therefore implies that

$$g^2(x) \le \frac{1}{\lambda^2} E_x f^2(X(\xi))$$

and hence

$$\begin{aligned} \|g\|^2 &\le \frac{1}{\lambda^2} \int_0^\infty \int_0^\infty \lambda \mathrm{e}^{-\lambda t} E_x f^2(X(t))\mathrm{d}t\pi(\mathrm{d}x) \\ &= \frac{1}{\lambda^2}\|f\|^2; \end{aligned}$$

in particular, $G(\cdot,\ \cdot,\ \lambda)$ induces a bounded linear operator on $L^2(\pi)$ for $\lambda > 0$.

On the other hand, for $\lambda \in (-\nu/2,\ 0)$, $0 < \beta(\lambda) < \eta$ and

$$\frac{\mathrm{e}^{(x+y)(\eta-\beta(\lambda))}}{\mathrm{e}^{\eta(x+y)-\beta(\lambda)|x-y|}} = \mathrm{e}^{-2\beta(\lambda)(x\wedge y)} \le 1,$$

so there exists a constant c_1 such that

$$|G(x,\ y,\ \lambda)| \le c_1 \mathrm{e}^{\eta(x+y)-\beta(\lambda)|x-y|}.$$

It follows that, when $f \in L^2(\pi)$,

$$\|g\|^2 \le c_1^2 \int_0^\infty \left(\int_0^\infty \mathrm{e}^{\eta x-\beta(\lambda)|x-y|} \mathrm{e}^{-\eta y} f(y)\mathrm{d}y \right)^2 \pi(\mathrm{d}x).$$

Again, the Cauchy–Schwarz inequality can be applied, thereby yielding

$$\left(\frac{\int_0^\infty \mathrm{e}^{-\beta(\lambda)|x-y|} \mathrm{e}^{-\eta y} f(y)\mathrm{d}y}{\int_0^\infty \mathrm{e}^{-\beta(\lambda)|x-y|}\mathrm{d}y} \right)^2 \le \frac{\int_0^\infty \mathrm{e}^{-\beta(\lambda)|x-y|} \mathrm{e}^{-2\eta y} f^2(y)\mathrm{d}y}{\int_0^\infty \mathrm{e}^{-\beta(\lambda)|x-y|}\mathrm{d}y},$$

and consequently

$$\|g\|^2 \le c_1^2 \int_0^\infty \int_0^\infty \mathrm{e}^{-\beta(\lambda)|x-y|}\mathrm{e}^{-2\eta y} f^2(y)\mathrm{d}y \cdot \int_0^\infty \mathrm{e}^{-\beta(\lambda)|x-z|}\mathrm{d}z\mathrm{e}^{2\eta x}\pi(\mathrm{d}x).$$

But, for $x \ge 0$,

$$\int_0^\infty \mathrm{e}^{-\beta(\lambda)|x-z|}\mathrm{d}z \le \int_{-\infty}^\infty \mathrm{e}^{-\beta(\lambda)|w|}\mathrm{d}w < \infty,$$

so there exists c_2 for which

$$\begin{aligned}\|g\|^2 &\le c_2 \int_0^\infty \int_0^\infty \mathrm{e}^{-\beta(\lambda)|x-y|}2\eta\mathrm{e}^{-2\eta y} f^2(y)\mathrm{d}y\mathrm{d}x \\ &= c_2 \int_0^\infty f^2(y)2\eta\mathrm{e}^{-2\eta y} \int_0^\infty \mathrm{e}^{-\beta(\lambda)|x-y|}\mathrm{d}x\mathrm{d}y \\ &\le c_2 \int_{-\infty}^\infty \mathrm{e}^{-\beta(\lambda)|w|}\mathrm{d}w \cdot \|f\|^2,\end{aligned}$$

proving (6.4) for $-\nu/2 < \lambda < 0$. (We have given here a direct proof of a result that can also be established via a convolution inequality due to Young; see, for example, Folland [14], pp. 240–241.)

When $\lambda = 0 \in \mathscr{S}$, the general solution to (6.3) for $f(x) = -x \in L^2(\pi)$ is given by

$$\frac{\mathrm{e}^{2\eta x} - 2\eta x(\eta x + 1)}{4r\eta^2} + c$$

for $c \in \mathbb{R}$, so that $g \notin L^1(\pi)$ (and hence is not in $L^2(\pi)$). Furthermore, for $\lambda = -\nu/2$, the general solution to (6.3) with $f(x) = -x$ takes the form

$$\frac{2\eta x(1 - \mathrm{e}^{\eta x}) + 4}{r\eta^2} + c\mathrm{e}^{\eta x}(1 - \eta x)$$

for $c \in \mathbb{R}$, so that g is never in $L^2(\pi)$. Finally, for $\lambda < -\nu/2$, (6.3) with $f(x) = -x$ admits the family of solutions

$$\frac{\mathrm{e}^{\eta x}\cos(s(\lambda)x)}{\lambda} + \frac{r - \lambda x}{\lambda^2} + c\left(\mathrm{e}^{\eta x}\left(\cos(s(\lambda)x) - \frac{\eta}{s(\lambda)}\sin(s(\lambda)x)\right)\right)$$

for $c \in \mathbb{R}$. Again, there is no choice of c for which $g \in L^2(\pi)$, proving our final assertion. □

We note that $\mathscr{S}$ is a mixed spectrum that has both a continuous part $(-\infty, -\nu/2]$ and a discrete component $\{0\}$. Furthermore, even for $\lambda \in \mathscr{S}$, $u_\lambda \notin L^2(\pi)$. In addition,

since $u_\lambda \to u_\gamma$ as $\lambda \to \gamma$, the u_λ cannot be orthonormal. Nevertheless, despite these complications, an analogue to (6.1) holds, namely (2.10).

Because spectral theory is built upon $L^2(\pi)$, we have no reason to expect that the convergence rate theory of Sect. 2 generalizes to convergence rates to equilibrium for functions $f \in L^1(\pi)$ that are not in $L^2(\pi)$. Indeed, for $\lambda \in (-\nu/2, 0)$, we have already established that $e^{-\lambda t} u_\lambda(X(t))$ is a martingale, so that

$$E_x u_\lambda(X(t)) = e^{\lambda t} u_\lambda(x) \to 0 = E u_\lambda(X(\infty)) \tag{6.6}$$

as $t \to \infty$. Hence, $u_\lambda \in L^1(\pi)$ and the rate of convergence of $E_x u_\lambda(X(t))$ to its equilibrium value is precisely exponential with rate parameter λ. Thus, depending on how close λ is to 0, the exponential rate can be arbitrarily close to 0.

Note that $u_\lambda(x)$ is of order $e^{x(\eta+\beta(\lambda))}$ as $x \to \infty$. When $\lambda \nearrow 0$, $\eta + \beta(\lambda) \nearrow 2\eta$. So, an alternative interpretation is that exponential moments in which $f(x) = e^{\theta x}$, with $\eta < \theta < 2\eta$, have exponential convergence rates to equilibrium with rate

$$\beta^{-1}(\theta - \eta) = \frac{\sigma^2(\theta - \eta)^2 - \nu}{2}.$$

On the other hand, when $\theta < \eta$, the theory of Sect. 2 applies, and $E_x f(X(t))$ converges to $Ef(X(\infty))$ at the rate $t^{-3/2} e^{-\nu t/2}$ specified there.

We have just argued that (2.15) does not describe the rate of convergence to equilibrium for functions $f \in L^1(\pi)$ that are not in $L^2(\pi)$. The rate of convergence for such functions must be analyzed on a case-by-case basis. But, despite the fact that spectral theory is built upon $L^2(\pi)$, there are also some functions $f \in L^2(\pi)$ for which (2.15) is not descriptive of the rate of convergence. In particular, the asymptotic (2.15) includes the limiting quantity $\langle f, u_{-\frac{\nu}{2}} \rangle$. Note that there exist functions $f \in L^2(\pi)$ for which $\langle f, u_{-\frac{\nu}{2}} \rangle = \infty$. This suggests that the rate of convergence for such functions can be slower than the order $t^{-\frac{3}{2}} e^{-\frac{\nu t}{2}}$ described in Sect. 2. Proposition 6 makes this rigorous.

Proposition 6 *Let $f(x) = e^{\eta x}(1 + \eta x)^{-2}$ for $x \geq 0$. Then, $f \in L^2(\pi)$ and*

$$\liminf_{t \to \infty} t^{\frac{3}{2}} e^{\frac{\nu t}{2}} |E_x f(X(t)) - Ef(X(\infty))| = \infty. \tag{6.7}$$

Proof We focus on the case of canonical RBM, in which $f(x) = e^x(1 + x)^{-2}$. Obviously,

$$Ef^2(X(\infty)) = 2 \int_0^\infty (1 + y)^{-4} dy < \infty,$$

so $f \in L^2(\pi)$. Also, since $E_x|f(X(t))| < \infty$ and $f \in L^1(\pi)$, (2.16) implies that

$$\begin{aligned}
&\frac{\pi}{2\sqrt{2}} t^{\frac{3}{2}} \mathrm{e}^{\frac{t}{2}} \mathrm{e}^{-x} (E_x f(X(t)) - Ef(X(\infty))) \\
&= \int_0^\infty \int_0^\infty \mathrm{e}^{-z} z^{\frac{1}{2}} \left(\cos\left(\sqrt{\frac{2z}{t}} x \right) - \sqrt{\frac{t}{2z}} \sin\left(\sqrt{\frac{2z}{t}} x \right) \right) \\
&\quad \cdot \left(\cos\left(\sqrt{\frac{2z}{t}} y \right) - \sqrt{\frac{t}{2z}} \sin\left(\sqrt{\frac{2z}{t}} y \right) \right) \left(1 + \frac{2z}{t} \right)^{-1} \mathrm{d}z \\
&\quad \cdot \frac{\mathrm{d}y}{(1+y)^2}.
\end{aligned}$$

Because $|\cos(w)| \le 1$, $\left|\frac{1}{w}\sin(wx)\right| \le |x|$, and

$$\int_0^\infty z^\gamma \mathrm{e}^{-z} \mathrm{d}z < \infty$$

for $\gamma > -1$ and $w > 0$, (6.7) follows if we establish that

$$\liminf_{t\to\infty} \int_0^\infty \int_0^\infty \mathrm{e}^{-z} z^{\frac{1}{2}} \sqrt{\frac{t}{2z}} \sin\left(\sqrt{\frac{2z}{t}} y \right) \left(1 + \frac{2z}{t} \right)^{-1} \mathrm{d}z \frac{\mathrm{d}y}{(1+y)^2} = \infty. \quad (6.8)$$

For $\epsilon > 0$, choose $\delta > 0$ so that $\frac{1}{w}\sin(w) \ge 1 - \epsilon$ for $0 \le w \le \sqrt{2}\delta$. Then,

$$\sqrt{\frac{t}{2z}} \sin\left(\sqrt{\frac{2z}{t}} y \right) \ge (1-\epsilon) y$$

for $y \le \delta\sqrt{\frac{t}{z}}$. Hence, the double integral in (6.8) can be lower bounded by $m(t, x)$, where

$$\begin{aligned}
m(t, x) \triangleq &\int_0^\infty \int_0^{\delta\sqrt{\frac{t}{z}}} \mathrm{e}^{-z} z^{\frac{1}{2}} (1-\epsilon) y \left(1 + \frac{2z}{t} \right)^{-1} \frac{\mathrm{d}y}{(1+y)^2} \mathrm{d}z \\
&- \int_0^\infty \int_{\delta\sqrt{\frac{t}{z}}}^\infty \mathrm{e}^{-z} z^{\frac{1}{2}} \sqrt{\frac{t}{2z}} \frac{\mathrm{d}y}{(1+y)^2} \mathrm{d}z. \quad (6.9)
\end{aligned}$$

Moreover,

$$\begin{aligned}
m(t, x) &\ge \frac{1-\epsilon}{2} \int_0^t \mathrm{e}^{-z} z^{\frac{1}{2}} \int_0^{\delta\sqrt{\frac{t}{z}}} \frac{y}{(1+y)^2} \mathrm{d}y \mathrm{d}z - \sqrt{\frac{t}{2}} \int_0^\infty \mathrm{e}^{-z} \frac{\sqrt{z}}{\sqrt{z} + \delta\sqrt{t}} \mathrm{d}z \\
&\to \frac{1-\epsilon}{2} \int_0^\infty \mathrm{e}^{-z} z^{\frac{1}{2}} \int_0^\infty \frac{y \mathrm{d}y}{(1+y)^2} \mathrm{d}z - \frac{1}{\delta\sqrt{2}} \int_0^\infty \sqrt{z} \mathrm{e}^{-z} \mathrm{d}z \\
&= \infty
\end{aligned}$$

as $t \to \infty$, where we used the monotone convergence theorem for simplifying the first integral at the last step. This proves (6.8). □

As a consequence, we see that some "extra condition" (such as $fu_{-\frac{\nu}{2}} \in L^1(\pi)$) beyond just requiring $f \in L^2(\pi)$ is indeed needed for the results of Sect. 2 to hold.

Acknowledgements This paper honors the fundamental contributions of Ward Whitt to the applied probability community, particularly to queueing theory, weak convergence, and diffusion approximations. The authors would also like to thank Vadim Linetsky for a helpful personal communication, regarding how to directly obtain the spectral representation for RBM, and the referee for a careful reading of this paper and for related comments on improving the exposition. Rob J. Wang is grateful to have been supported by an Arvanitidis Stanford Graduate Fellowship in memory of William K. Linvill, the Thomas Ford Fellowship, as well as NSERC Postgraduate Scholarships.

References

1. Abate, J., Whitt, W.: Transient behavior of regulated Brownian motion, I: starting at the origin. Adv. Appl. Probab. **19**(3), 560–598 (1987a)
2. Abate, J., Whitt, W.: Transient behavior of regulated Brownian motion, II: non-zero initial conditions. Adv. Appl. Probab. **19**(3), 599–631 (1987b)
3. Abate, J., Whitt, W.: Transient behavior of the M/M/1 queue: starting at the origin. Queueing Syst. **2**, 41–65 (1987c)
4. Abate, J., Whitt, W.: Transient behavior of the M/G/1 workload process. Oper. Res. **42**(4), 750–764 (1994)
5. Artin, E.: The Gamma Function. Holt, Rinchart, and Winston Inc, New York (1964)
6. Asmussen, S.: Queueing simulation in heavy traffic. Math. Oper. Res. **17**(1), 84–111 (1992)
7. Asmussen, S., Glynn, P.W., Thorisson, H.: Stationarity detection in the initial transient problem. ACM Trans. Model. Comput. Simul. **2**(2), 130–157 (1992)
8. Borovkov, A.A.: Some limit theorems in the theory of mass service II. Theor. Probab. Appl. **10**, 375–400 (1965)
9. Chung, K.L.: A Course in Probability Theory, 3rd edn. Academic Press, San Diego (2001)
10. Cohen, J.W.: The Single Server Queue, 2nd. Revised edn. Elsevier, Amsterdam (1982)
11. Diaconis, P.: The Markov chain Monte Carlo revolution. Bull. Am. Math. Soc. **46**(2), 179–1205 (2009)
12. Diaconis, P., Stroock, D.: Geometric bounds for eigenvalues of Markov chains. Ann. Appl. Probab. **1**(1), 36–61 (1991)
13. Ethier, S.N., Kurtz, T.G.: Markov Processes: Characterization and Convergence, 2nd edn. Wiley, New York (2005)
14. Folland, G.B.: Real Analysis: Modern Techniques and Their Applications, 2nd edn. Wiley, New York (1999)
15. Glynn, P.W., Meyn, S.P.: A Liapounov bound for solutions of the Poisson equation. Ann. Probab. **2**(2), 916–931 (1996)
16. Glynn, P.W., Wang, R.J.: On the rate of convergence to equilibrium for two-sided reflected Brownian motion and for the Ornstein–Uhlenbeck process, pp. 1–10 (2018) (Submitted for Publication)
17. Grassmann, W.K.: Factors affecting warm-up periods in discrete event simulation. Simulation **90**(1), 11–23 (2014)
18. Grimmett, G.R., Stirzaker, D.R.: Probability and Random Processes, 3rd edn. Oxford University Press, Oxford (2001)
19. Harrison, J.M.: Brownian Models of Performance and Control. Cambridge University Press, Cambridge (2013)
20. Iglehart, D.L., Whitt, W.: Multiple channel queues in heavy traffic. I. Adv. Appl. Probab. **2**(1), 150–177 (1970)
21. Kingman, J.F.C.: The single server queue in heavy traffic. Proc. Camb. Philos. Soc. **57**, 902–904 (1961)
22. Linetsky, V.: On the transition densities for reflected diffusions. Adv. Appl. Probab. **37**, 435–460 (2005)
23. Meyn, S., Tweedie, R.L.: Markov Chains and Stochastic Stability, 2nd edn. Cambridge University Press, Cambridge (2009)

24. Meyn, S.P., Tweedie, R.L.: Stability of Markovian processes III: Foster–Lyapunov criteria for continuous-time processes. Adv. Appl. Probab. **25**(3), 518–548 (1993)
25. Roberts, G.O., Rosenthal, J.S.: General state space Markov chains and MCMC algorithms. Probab. Surv. **1**, 20–71 (2004)
26. Thorisson, H.: Coupling, Stationarity, and Regeneration. Springer, New York (2000)
27. Wang, R.J., Glynn, P.W.: Measuring the initial transient: reflected Brownian motion. In: Tolk, A., Diallo, S.Y., Ryzhov, I.O., Yilmaz, L., Buckley, S., Miller, J.A. (eds.) Proceedings of the 2014 Winter Simulation Conference, pp. 652–661 (2014)
28. Wang, R.J., Glynn, P.W.: On the marginal standard error rule and the testing of initial transient deletion methods. ACM Trans. Model. Comput. Simul. **27**(1), 1–30 (2016)
29. Whitt, W.: Planning queueing simulations. Manag. Sci. **35**(11), 1341–1366 (1989)