

PARAMETER AND DIMENSION DEPENDENCE OF CONVERGENCE RATES TO STATIONARITY FOR REFLECTING BROWNIAN MOTIONS

BY SAYAN BANERJEE* AND AMARJIT BUDHIRAJA†

Department of Statistics and Operations Research, University of North Carolina, Chapel Hill, *sayan@email.unc.edu;
†amarjit@unc.edu

We obtain rates of convergence to stationarity in L^1 -Wasserstein distance for a d -dimensional reflected Brownian motion (RBM) in the nonnegative orthant that are explicit in the dimension and the system parameters. The results are then applied to a class of RBMs considered in (Blanchet and Xinyun (2016)) and to rank-based diffusions including the Atlas model. In both cases, we obtain explicit rates and bounds on relaxation times. In the first case we improve the relaxation time estimates of $O(d^4(\log d)^2)$ obtained in (Blanchet and Xinyun (2016)) to $O((\log d)^2)$. In the latter case, we give the first results on explicit parameter and dimension dependent rates under the Wasserstein distance. The proofs do not require an explicit form for the stationary measure or reversibility of the process with respect to this measure, and cover settings where these properties are not available. In the special case of the standard Atlas model (In *Stochastic Portfolio Theory* (2002) 1–24 Springer), we obtain a bound on the relaxation time of $O(d^6(\log d)^2)$.

1. Introduction. A d -dimensional obliquely reflected Brownian motion with drift in the nonnegative orthant plays a central role in Queuing Theory where it arises as a diffusion limit of scaled queue length processes when the system is in the *heavy traffic regime* (namely the arrival rate and the service rate are approximately equal) [5, 14, 15, 24, 29]. Such a process is also used to describe the behavior of rank-based diffusions, namely a system of particles whose trajectories are given by Brownian motions with drift, where the drift and diffusion coefficients of a given particle at any given time depend on its relative rank in the system at that time. These models appear frequently in mathematical finance, for example, the Atlas model [2, 18, 23]. There has been extensive work in the study of stability of such reflected Brownian motions (RBM) that gives explicit sufficient conditions for positive recurrence for the RBM and the corresponding queuing systems [1, 8, 9, 16, 27]. In this work, we obtain explicit exponential convergence rates (in Wasserstein distance) to equilibrium for multidimensional reflected Brownian motion (RBM) under a key stability condition identified in [16] (see Assumption (A2)). This assumption is known to be “almost necessary” for stability (see Remark 1 for a precise statement). The convergence rates obtained in this work are *explicit (up to some universal constants) in the dimension and system parameters*. The system parameters are given by the drift vector, the covariance matrix of the Brownian motion, and the reflection matrix. Stationary distributions of RBM are rarely explicit and the convergence rates of the form obtained in this work provide important information for the construction of numerical schemes that sample from these stationary distributions.

There has been some prior work in this area. Exponential ergodicity was proved in [7] for semimartingale reflecting Brownian motions under the stability condition of [9]. This class includes RBM of the form considered in this work. The paper [7] also established exponential ergodicity of certain reflected diffusions with state dependent drift and diffusion coefficients.

Received March 2019; revised August 2019.

MSC2020 subject classifications. Primary 60J60, 60H10; secondary 60J55, 60K25.

Key words and phrases. Reflected Brownian motion, local time, coupling, Wasserstein distance, relaxation time, heavy traffic, Atlas model.

The key ingredient in the proof was the construction of a suitable Lyapunov function along with establishing a minorization condition on a sufficiently large compact set (referred to as a “small set”). The Lyapunov function provides good control on the exponential moments of the return times to the small set while the minorization condition implies the existence of abstract couplings of two copies of the process (via construction of “pseudo-atoms” as described in Chapter 5 of [22]) which have a positive chance of coalescing inside the small set. Together, they furnish exponential rates of convergence (in a weighted total variation distance). However, due to the somewhat implicit treatment of the process inside the small set, the rates obtained by this method shed little light on how they qualitatively depend on the system parameters or the state dimension. The paper [18] obtained explicit convergence rates for a class of reversible rank-based diffusions with explicit stationary measures using Dirichlet form techniques (which crucially use reversibility). See also the discussion in Section 5.2. The convergence considered in [18] corresponds to that of time averages of bounded functionals of the state process to the corresponding stationary values in probability (see Theorem 1 of [18]), which is considerably weaker than the L^1 -Wasserstein distance considered in the current work. The setting of one-dimensional RBM was considered in [28] where (among other results) an estimate on the spectral gap was provided as a function of the drift and the diffusion coefficient. In a recent work, [3] obtained dimension dependent bounds on rates of Wasserstein convergence for a class of RBM. Under conditions on the drift vector, the covariance matrix of the Brownian motion, and the reflection matrix (see Conditions (BC1)–(BC3) in Section 5), [3] analyzed the behavior of the RBM inside the small set explicitly by considering *synchronous couplings* (namely, couplings where the RBM starting from different points are driven by the same Brownian motion). Using explicit couplings to obtain better convergence rate estimates is a relatively recent but developing area. See [4, 10–12] for such results for other classes of diffusions. In this work, we revisit the idea of constructing synchronous couplings for RBM. Under quite general conditions (specifically, the ones introduced in [16] that guarantee the existence of strong solutions and positive recurrence), we construct a suitable Lyapunov function and identify (an appropriate analogue of) a small set that both depend crucially on the process parameters and the state dimension. This, along with a careful treatment of excursions from the small set, enables us to quantify contraction rates in L^1 -distance for synchronous couplings starting from distinct points and thereby obtain rates of Wasserstein convergence that are given explicitly in terms of the system parameters, the state dimension, and some constants (that do not depend on dimension or model parameters). These convergence rates, together with bounds on relaxation times of the RBM that follow from it, are the main results of this work and are given in Theorem 1. In Section 5 we apply these results to the class of RBMs considered in [3] and rank-based diffusions considered in [18]. In the former case, we substantially improve the relaxation time estimates from $O(d^4(\log d)^2)$ obtained in [3] to $O((\log d)^2)$. In the latter case, we give the first results on explicit parameter and dimension dependent rates under the Wasserstein distance. The proofs do not require an explicit form for the stationary measure or reversibility of the process with respect to this measure, and cover settings where these properties are not available. In the special case of the standard Atlas model [13], we give a bound on the relaxation time of $O(d^6(\log d)^2)$ (see Remark 6).

2. Model, notation and assumptions. Let B be a d -dimensional standard Brownian motion and let $\mu \in \mathbb{R}^d$ and $\mathbf{D}, R \in \mathbb{R}^{d \times d}$. Consider for $\mathbf{x} \in \mathbb{R}_+^d := [0, \infty)^d$ the \mathbb{R}_+^d -valued continuous stochastic process given by the equation

$$(1) \quad X(t; \mathbf{x}) = \mathbf{x} + \mathbf{D}B(t) + \mu t + RL(t),$$

where L , referred to as the local time process, is a nondecreasing continuous process satisfying

$$(2) \quad L(0) = 0, \quad \int_0^t X_i(s; \mathbf{x}) dL_i(s) = 0 \quad \text{for all } t > 0 \text{ and } 1 \leq i \leq d.$$

We will make the following basic assumptions.

ASSUMPTIONS.

(A1) The matrix $P := I - R^T$ is substochastic (nonnegative entries and row sums bounded above by 1) and transient ($P^n \rightarrow 0$ as $n \rightarrow \infty$).

(A2) $\mathbf{b} := -R^{-1}\mu > 0$.

(A3) The matrix $\Sigma = \mathbf{D}\mathbf{D}^T$ is positive definite.

The paper [15] shows that under (A1) there is a unique strong solution to (1)–(2), namely for each $\mathbf{x} \in \mathbb{R}_+^d$ there is a unique pair of continuous stochastic processes (X, L) satisfying the above equations. This assumption is satisfied by the routing matrix of any single-class open queueing network [15] and consequently diffusion limits of such networks can be characterized by (1)–(2). The collection $\{X(\cdot; \mathbf{x})\}_{\mathbf{x} \in \mathbb{R}_+^d}$ defines a strong Markov process (see [16]) which we denote as RBM(μ, Σ, R) and refer to simply as the reflected Brownian motion (RBM). The matrix R describes the reflection mechanism, specifically, the i th column of R gives the direction of reflection on the i th face of the orthant. The conditions on P in particular say that its spectral radius is strictly less than 1. The matrix $\Sigma = \mathbf{D}\mathbf{D}^T$ gives the covariance matrix associated with the diffusion term of (1).

NOTATION. Although μ, Σ and R depend on the dimension d , this dependence is suppressed to avoid cumbersome notation. We will write $\mathbf{b} = -R^{-1}\mu$. The entries of \mathbf{b} will be denoted by b_i , and the diagonal entries of Σ will be denoted by σ_i^2 , where $1 \leq i \leq d$. All constants appearing in the statements of lemmas and theorems will be universal in that they do not depend on model parameters or the dimension d , unless noted otherwise.

REMARK 1. Unique strong solutions of the RBM that follow from (A1) imply that any coupling of the driving Brownian motions translate into a coupling of the processes themselves. Throughout this work we will take the family $\{X(\cdot; \mathbf{x})\}_{\mathbf{x} \in \mathbb{R}_+^d}$ to be *notationdriven by the same Brownian motion*, namely we will consider a *synchronous coupling* of the processes starting from different initial conditions. Assumption (A2) is the well-known “stability condition” which is sufficient for the existence of a stationary measure [16]. The condition is almost necessary for stability in that if $\mathbf{b}_i < 0$ for some i then the RBM is transient [6]. Assuming (A3) in addition to (A1)–(A2) gives that the strong Markov process RBM(μ, Σ, R) has a unique stationary probability distribution [16].

3. Main result. Following [3], define the following stopping times: $\eta^0(\mathbf{x}) = 0$ and

$$\eta_i^k(\mathbf{x}) = \inf\{t \geq \eta^{k-1}(\mathbf{x}) + 1 : X_i(t, \mathbf{x}) = 0\}, \quad \eta^k(\mathbf{x}) = \sup\{\eta_i^k(\mathbf{x}) : 1 \leq i \leq d\}.$$

Define

$$\mathcal{N}(t; \mathbf{x}) = \sup\{k \geq 0 : \eta^k(\mathbf{x}) \leq t\}.$$

Also define the *contraction coefficient*

$$(3) \quad n(R) := \inf\{n \geq 1 : \|P^n \mathbf{1}\|_\infty \leq 1/2\},$$

where $\mathbf{1}$ is a d -dimensional vector of ones and for $u \in \mathbb{R}^d$, $\|u\|_\infty := \sup_{1 \leq i \leq d} |u_i|$. By Assumption (A1), $n(R) < \infty$. This quantity plays a key role in quantifying the convergence rate to equilibrium.

We now present the main result of this work. Given probability measures μ and ν on \mathbb{R}_+^d , a probability measure γ on $\mathbb{R}_+^d \times \mathbb{R}_+^d$ is said to be a coupling of μ and ν if $\gamma(\cdot \times \mathbb{R}_+^d) = \mu(\cdot)$ and $\gamma(\mathbb{R}_+^d \times \cdot) = \nu(\cdot)$. The L^1 -Wasserstein distance between two probability measures μ and ν on \mathbb{R}_+^d is given by

$$W_1(\mu, \nu) = \inf \left\{ \int_{\mathbb{R}_+^d \times \mathbb{R}_+^d} \|\mathbf{x} - \mathbf{y}\|_1 \gamma(d\mathbf{x}, d\mathbf{y}) : \gamma \text{ is a coupling of } \mu \text{ and } \nu \right\},$$

where for a vector $z \in \mathbb{R}^d$, $\|z\|_1 = \sum_{i=1}^d |z_i|$. We will denote the law of a random variable X by $\mathcal{L}(X)$. Recall that from [16], under Assumptions (A1)–(A3), there is a unique stationary distribution of the RBM. Denote by $\mathbf{X}(\infty)$ a random vector sampled from this stationary distribution. Define the *relaxation time*, $t_{\text{rel}}(\mathbf{x})$ for the RBM starting from $\mathbf{x} \in \mathbb{R}_+^d$ as

$$t_{\text{rel}}(\mathbf{x}) := \inf \{ t \geq 0 : W_1(\mathcal{L}(X(t; \mathbf{x})), \mathcal{L}(\mathbf{X}(\infty))) \leq 1/2 \}.$$

We will abbreviate the parameters of the RBM as $\Theta := (\mu, \Sigma, R)$. Recall that these parameters are required to satisfy (A1)–(A3). We will quantify rate of convergence to equilibrium in terms of the following functions of Θ, d . Fix $\kappa \in (0, \infty)$. Let

$$a(\Theta) := \sup_{1 \leq i \leq d} \left[\frac{\sum_{j=1}^d (R^{-1})_{ij} \sigma_j}{b_i} \right], \quad b(\Theta) := \sup_{1 \leq i \leq d} \left[\frac{\sum_{j=1}^d (R^{-1})_{ij} \sigma_j}{\sigma_i} \right],$$

$$R_1(\Theta, d) := n(R)(1 + a(\Theta)^2 \log(2d)), \quad R_2(\Theta) := a(\Theta)^2 b(\Theta),$$

$$C_1(\mathbf{x}, \Theta) := 2\|\mathbf{x}\|_1 + a(\Theta) \sum_{i,j} (R^{-1})_{ij} \sigma_j,$$

$$C_2(\mathbf{x}, \Theta, \kappa) := 2\|\mathbf{x}\|_1 e^{3(\kappa a(\Theta) b(\Theta))^{-1} \|\mathbf{x}\|_\infty^*} + a(\Theta) \left[2d(1+d) \left(\sum_{i,j} (R^{-1})_{ij}^2 \right) \left(\sum_{j=1}^d \sigma_j^2 \right) \right]^{1/2}.$$

For any $\mathbf{x} \in \mathbb{R}_+^d$, define $\|\mathbf{x}\|_\infty^* := \sup_{1 \leq i \leq d} \sigma_i^{-1} \mathbf{x}_i$.

THEOREM 1. *There exist a $t_0 \in (0, \infty)$ and $D_1, D_2 \in (0, \infty)$ such that for every $d \in \mathbb{N}$, $\mathbf{x} \in \mathbb{R}_+^d$, every parameter choice Θ , and $t \geq t_0(1 + (a(\Theta))^2 \log(2d))$,*

$$\begin{aligned} & W_1(\mathcal{L}(X(t; \mathbf{x})), \mathcal{L}(\mathbf{X}(\infty))) \\ & \leq \mathbb{E}(\|X(t; \mathbf{x}) - X(t; \mathbf{X}(\infty))\|_1) \\ & \leq C_1(\mathbf{x}, \Theta) \left(2e^{-\frac{D_1 t}{R_1(\Theta, d)}} + e^{-\frac{t}{16D_2 R_2(\Theta)}} \right) + C_2(\mathbf{x}, \Theta, D_2) e^{-\frac{t}{8D_2 R_2(\Theta)}}. \end{aligned}$$

In particular, the relaxation time satisfies

$$\begin{aligned} t_{\text{rel}}(\mathbf{x}) & \leq \max \{ D_1^{-1} R_1(\Theta, d) \log(8C_1(\mathbf{x}, \Theta)) \\ & \quad + 16D_2 R_2(\Theta) \log[4(C_1(\mathbf{x}, \Theta) + C_2(\mathbf{x}, \Theta, D_2))] \}, \\ & \quad t_0(1 + (a(\Theta))^2 \log(2d)). \end{aligned}$$

REMARK 2. The universal constants t_0, D_1 and D_2 will be identified in Sections 6 and 7. Specifically, t_0 and D_1 are introduced in Lemma 8 (see (43)) and $D_2 := \max\{A_0, 9\}$, where A_0 is introduced in Lemma 5.

REMARK 3. The proof of Theorem 1 (see Remark 8) will show that one can, in fact, obtain a better bound of the form

$$\begin{aligned} W_1(\mathcal{L}(X(t; \mathbf{x}), \mathcal{L}(\mathbf{X}(\infty))) &\leq \mathbb{E}(\|X(t; \mathbf{x}) - X(t; \mathbf{X}(\infty))\|_1) \\ &\leq C_1(\mathbf{x}, \Theta) \left(e^{-\frac{D_1 n(R)t}{R_1(\Theta, d)}} + e^{-\frac{32(\log 2)D_1 t}{R_1(\Theta, d)}} + e^{-\frac{t}{16D_2 R_2(\Theta)}} \right) \\ &\quad + C_2(\mathbf{x}, \Theta, D_2) e^{-\frac{t}{8D_2 R_2(\Theta)}}. \end{aligned}$$

This bound leads to a better choice of the universal constants appearing in the exponents of the bound when $n(R)$ is large. As a consequence the bounds on relaxation times and the bounds given in the examples of Section 5 can be slightly improved using the above estimate. However, this improved bound leads to cumbersome expressions in the bounds and relaxation time estimates in Section 5. Moreover, our main goal is to highlight the dependence of the convergence rates on system parameters which is completely captured by Theorem 1. Hence, we do not give details on how the improved bound can be obtained, however see Remark 8 for some additional comments.

An important ingredient in the proof is the following analogue of Lemma 3 from [3] which shows that the synchronous coupling gives an a.s. contraction of the L^1 -distance $\|X(t; \mathbf{x}) - X(t; \mathbf{0})\|_1$ which can be quantified as follows. The proof is similar to that in [3] and so only a sketch is provided.

LEMMA 2 (See [3]). For $\mathbf{x} \in \mathbb{R}_+^d$ and $t \geq 0$,

$$\|X(t; \mathbf{x}) - X(t; \mathbf{0})\|_1 \leq 2\|\mathbf{x}\|_1 2^{-\mathcal{N}(t; \mathbf{x})/n(R)}.$$

PROOF. The main idea is to associate the substochastic matrix P with a Markov chain on states $\{0, 1, \dots, d\}$ absorbed at 0 and show that $\|\mathbf{x}\|_1^{-1} \|X(t; \mathbf{x}) - X(t; \mathbf{0})\|_1$ (assuming $\|\mathbf{x}\|_1 \neq 0$) is bounded above by the maximum over the initial state i of the probability that, starting from i , the Markov chain is not absorbed by time $\mathcal{N}(t; \mathbf{x})$. Using this idea, Lemma 2 in [3] and the proof of Lemma 3 in [3] establish

$$(4) \quad \|X(t; \mathbf{x}) - X(t; \mathbf{0})\|_1 \leq \|P^{\mathcal{N}(t; \mathbf{x})} \mathbf{1}\|_\infty \|\mathbf{x}\|_1.$$

The lemma now follows from the definition of $n(R)$ given in (3) above. \square

REMARK 4. The quantity $n(R)^{-1}$ defined in (3) gives an explicit bound on the exponential decay rate of $\|P^n \mathbf{1}\|_\infty$ with n . Note that $n(R)$ possibly depends on the dimension d , but the dependence is solely through R . Sometimes (as we will see in the first example of Section 5) it is possible to get a better bound in the sense that we can obtain positive constants $C(R, d)$ and $n'(R) < n(R)$ such that

$$\|P^n \mathbf{1}\|_\infty \leq C(R, d) 2^{-n/n'(R)}, \quad n \geq 0.$$

In this case, we can replace the bound in Lemma 2 by

$$(5) \quad \|X(t; \mathbf{x}) - X(t; \mathbf{0})\|_1 \leq C(R, d) \|\mathbf{x}\|_1 2^{-\mathcal{N}(t; \mathbf{x})/n'(R)}.$$

REMARK 5. In cases where we can obtain the better bound (5), the constants $R_1(\Theta, d)$ and $C_1(\mathbf{x}, \Theta)$ appearing in the bounds on Wasserstein distance and relaxation time in Theorem 1 can be replaced by $R'_1(\Theta, d)$ and $C'_1(\mathbf{x}, \Theta, d)$ respectively, where

$$\begin{aligned} R'_1(\Theta, d) &:= n'(R)(1 + a(\Theta)^2 \log(2d)), \\ C'_1(\mathbf{x}, \Theta, d) &:= 2\|\mathbf{x}\|_1 + \frac{a(\Theta)C(R, d)}{2} \sum_{i,j} (R^{-1})_{ij} \sigma_j. \end{aligned}$$

4. Outline of approach.

We now give an outline of our approach.

(i) We use a key idea from [3] which shows that, under the synchronous coupling, the L^1 -distance between the two processes $X(\cdot; \mathbf{0})$ and $X(\cdot; \mathbf{x})$ decreases with time. Using this idea, we provide an estimate on the rate of decay of this L^1 -distance in terms of a “contraction coefficient” which quantifies the decay rate of $\|P^n \mathbf{1}\|_\infty$ with n . The precise statement was formulated as Lemma 2 in Section 3.

(ii) We use the fact that for any $v > 0$ in \mathbb{R}^d satisfying $R^{-1}v \leq \mathbf{b}$, one can dominate the process $X(\cdot; \mathbf{x})$ in an appropriate manner by a normally reflected Brownian motion with drift $-v$ in \mathbb{R}_+^d . This process, written as $X_v^+(\cdot; \mathbf{x})$, is technically simpler to analyze. The idea of dominating an RBM(μ, Σ, R) by a normally reflected RBM is due to [16]. Next, we choose an appropriate compact set (which plays a role similar to the “small set” in the terminology of [22]) such that one can obtain a tight control over return times to this set (this is done via Lyapunov function techniques in Lemma 5) and, loosely speaking, is such that the L^1 -distance between the synchronously coupled processes $X(\cdot; \mathbf{0})$ and $X(\cdot; \mathbf{x})$ decreases by a constant factor each time the process $X_v^+(\cdot; \mathbf{x})$ visits this set (this result is formulated in Lemma 7). A crucial ingredient here is the **Introduction** of a suitable weighted norm (see (16)) whose sub-level sets are the appropriate “small sets” with the desired contraction property. The definition of this norm is guided by an analysis of how the maximum process for each coordinate scales with the system parameters. This weighted norm is used to construct the small set and also an appropriate Lyapunov function. These constructions and their properties are studied in Section 6.

(iii) In Section 7, we obtain the rate of decay of $\|X(t; \mathbf{x}) - X(t; \mathbf{0})\|_1$ with time t , in terms of the parameter v of the dominating normally reflected RBM, by decomposing the path of $X_v^+(\cdot; \mathbf{x})$ into excursions from the small set obtained in (ii) and using the estimates from Section 6 for probabilities of certain events associated with these excursions.

(iv) Finally in Section 8 we prove our main result, namely Theorem 1, where we obtain explicit parameter and dimension dependent rates of decay in L^1 -Wasserstein distance between the processes $X(\cdot; \mathbf{0})$ and $X(\cdot; \mathbf{x})$ with time t by optimizing the rates derived in (iii) over the parameter $v > 0$ of the dominating RBM.

Before proceeding to the proof we apply Theorem 1 in two settings, the first is that of RBM satisfying the assumptions of [3] and the second corresponds to that of rank-based diffusions such as the Atlas model.

5. Examples. We will use Theorem 1 (and Remark 5) to obtain bounds on the rate of convergence to equilibrium in two examples that are discussed in Sections 5.1 and 5.2 below.

5.1. Blanchet–Chen RBM. This refers to the class of RBM under the set of assumptions in [3], namely:

(BC1) The matrix P is substochastic and there exist $\kappa > 0$ and $\beta \in (0, 1)$ not depending on the dimension d such that $\|\mathbf{1}^T P^n\|_\infty \leq \kappa(1 - \beta)^n$ for all $n \geq 0$.

(BC2) There exists $\delta > 0$ independent of d such that $R^{-1}\mu < -\delta\mathbf{1}$.

(BC3) There exists $\sigma > 0$ independent of d such that $\sigma_i := \sqrt{\Sigma_{ii}}$ satisfies $\sigma^{-1} \leq \sigma_i \leq \sigma$ for every $1 \leq i \leq d$.

Under the above conditions [3] give a polynomial bound of $O(d^4(\log d)^2)$ on the relaxation time of the RBM. As shown in the following theorem, Theorem 1 gives a substantial improvement by establishing a polylogarithmic relaxation time of $O((\log d)^2)$.

THEOREM 3. *Under Assumptions (BC1), (BC2) and (BC3), there exist positive constants E_1, E_2, E_3, E_4, t_1 such that for any $\mathbf{x} \in \mathbb{R}_+^d, t \geq t_1 \max\{\|\mathbf{x}\|_\infty, \log(2d)\}$,*

$$\begin{aligned} & \mathbb{E}(\|X(t; \mathbf{x}) - X(t; \mathbf{X}(\infty))\|_1) \\ & \leq 2(2\|\mathbf{x}\|_1 + E_1 d^2)e^{-E_2 t / \log(2d)} + (4\|\mathbf{x}\|_1 + E_1 d^2)e^{-E_4 t / 2} + E_3 d^2 e^{-E_4 t}. \end{aligned}$$

In particular, the relaxation time satisfies

$$\begin{aligned} t_{\text{rel}}(\mathbf{x}) & \leq \max\{E_2^{-1} \log[8(2\|\mathbf{x}\|_1 + E_1 d^2)] \log(2d) \\ & \quad + E_4^{-1} [2 \log[8(4\|\mathbf{x}\|_1 + E_1 d^2)] + \log(8E_3 d^2)], \\ & \quad t_1 \max\{\|\mathbf{x}\|_\infty, \log(2d)\}\}. \end{aligned}$$

PROOF. Observe that

$$\|P^n \mathbf{1}\|_\infty \leq \mathbf{1}^T P^n \mathbf{1} \leq d \|\mathbf{1}^T P^n\|_\infty \leq d\kappa(1 - \beta)^n.$$

Thus, the hypothesis of Remark 4 is satisfied with

$$C(R, d) = \kappa d, \quad n'(R) \equiv n' := \frac{\log(2)}{\log(1 - \beta)^{-1}} + 1.$$

Now we will use Theorem 1 in conjunction with Remark 5.

Under Assumptions (BC1), (BC2) and (BC3), we have the following bounds:

$$a(\Theta) = \sup_{1 \leq i \leq d} \left[\frac{\sum_{j=1}^d (R^{-1})_{ij} \sigma_j}{b_i} \right] \leq \frac{\|R^{-1} \mathbf{1}\|_\infty \sigma}{\delta} \leq \frac{\kappa \sigma}{\beta \delta},$$

$$b(\Theta) = \sup_{1 \leq i \leq d} \left[\frac{\sum_{j=1}^d (R^{-1})_{ij} \sigma_j}{\sigma_i} \right] \leq \frac{\|R^{-1} \mathbf{1}\|_\infty \sigma}{\sigma^{-1}} \leq \frac{\kappa \sigma^2}{\beta},$$

$$R'_1(\Theta, d) = n'(R)(1 + a(\Theta)^2 \log(2d)) \leq n' \left(1 + \frac{\kappa^2 \sigma^2}{\beta^2 \delta^2} \log(2d) \right),$$

$$R_2(\Theta) = a(\Theta)^2 b(\Theta) \leq \frac{\kappa^3 \sigma^4}{\beta^3 \delta^2},$$

$$\begin{aligned} C'_1(\mathbf{x}, \Theta, d) & = 2\|\mathbf{x}\|_1 + \frac{a(\Theta)C(R, d)}{2} \sum_{i,j} (R^{-1})_{ij} \sigma_j \\ & \leq 2\|\mathbf{x}\|_1 + \frac{\kappa^2 \sigma d^2}{2\beta \delta} \|R^{-1} \mathbf{1}\|_\infty \sigma \leq 2\|\mathbf{x}\|_1 + \frac{\kappa^3 \sigma^2 d^2}{2\beta^2 \delta}, \end{aligned}$$

where we have used the observation that under Assumption (BC1), one has $\|R^{-1} \mathbf{1}\|_\infty \leq \kappa / \beta$. Next, observe that $\|\mathbf{x}\|_\infty^* \leq \sigma \|\mathbf{x}\|_\infty$. This, along with the bound on $a(\Theta)$ obtained above, implies that for $t \geq \frac{48\kappa\sigma^2}{\beta\delta} \|\mathbf{x}\|_\infty, 3(D_2 a(\Theta) b(\Theta))^{-1} \|\mathbf{x}\|_\infty^* \leq t / (16D_2 R_2(\Theta))$. Hence, for such t ,

$$\begin{aligned} & C_2(\mathbf{x}, \Theta, D_2) e^{-\frac{t}{8D_2 R_2(\Theta)}} \\ & = 2\|\mathbf{x}\|_1 e^{3(D_2 a(\Theta) b(\Theta))^{-1} \|\mathbf{x}\|_\infty^*} e^{-\frac{t}{8D_2 R_2(\Theta)}} \\ & \quad + a(\Theta) \left[2d(1 + d) \left(\sum_{i,j} (R^{-1})_{ij}^2 \right) \left(\sum_{j=1}^d \sigma_j^2 \right) \right]^{1/2} e^{-\frac{t}{8D_2 R_2(\Theta)}} \end{aligned}$$

$$\begin{aligned} &\leq 2\|\mathbf{x}\|_1 e^{-\frac{t}{16D_2R_2(\Theta)}} + \frac{\kappa\sigma}{\beta\delta} [2d(1+d)d(\|R^{-1}\mathbf{1}\|_\infty)^2(d\sigma^2)]^{1/2} e^{-\frac{t}{8D_2R_2(\Theta)}} \\ &\leq 2\|\mathbf{x}\|_1 e^{-\frac{t}{16D_2R_2(\Theta)}} + \frac{2\kappa^2\sigma^2}{\beta^2\delta} d^2 e^{-\frac{t}{8D_2R_2(\Theta)}} \\ &\leq 2\|\mathbf{x}\|_1 e^{-\frac{\beta^3\delta^2 t}{16D_2\kappa^3\sigma^4}} + \frac{2\kappa^2\sigma^2}{\beta^2\delta} d^2 e^{-\frac{\beta^3\delta^2 t}{8D_2\kappa^3\sigma^4}}. \end{aligned}$$

Take $E_1 = \frac{\kappa^3\sigma^2}{2\beta^2\delta}$, $E_2 = D_1[n'(2 + \frac{\kappa^2\sigma^2}{\beta^2\delta^2})]^{-1}$, $E_3 = \frac{2\kappa^2\sigma^2}{\beta^2\delta}$, $E_4 = \frac{\beta^3\delta^2}{8D_2\kappa^3\sigma^4}$. Using the above bounds in Theorem 1 (modified as in Remark 5), for any $\mathbf{x} \in \mathbb{R}_+^d$, $t \geq \max\{t_0(1 + \frac{\kappa^2\sigma^2}{\beta^2\delta^2} \log(2d)), \frac{48\kappa\sigma^2}{\beta\delta} \|\mathbf{x}\|_\infty\}$,

$$\begin{aligned} &\mathbb{E}(\|X(t; \mathbf{x}) - X(t; \mathbf{X}(\infty))\|_1) \\ &\leq C'_1(\mathbf{x}, \Theta, d)(2e^{-\frac{D_1 t}{R_1^{\prime}(\Theta, d)}} + e^{-\frac{t}{16D_2R_2(\Theta)}}) + C_2(\mathbf{x}, \Theta, D_2) e^{-\frac{t}{8D_2R_2(\Theta)}} \\ &\leq (2\|\mathbf{x}\|_1 + E_1 d^2)(2e^{-E_2 t/\log(2d)} + e^{-E_4 t/2}) + 2\|\mathbf{x}\|_1 e^{-E_4 t/2} + E_3 d^2 e^{-E_4 t}. \end{aligned}$$

This proves the first part of the theorem upon taking $t_1 = \max\{t_0(1 + \frac{\kappa^2\sigma^2}{\beta^2\delta^2}), \frac{48\kappa\sigma^2}{\beta\delta}\}$. The bound on the relaxation time follows immediately from the first part. \square

5.2. *Gap process of rank-based diffusions.* Rank based diffusions are interacting particle systems where the drift and diffusion coefficient of each particle depends on its rank. Mathematically, they are represented by the SDE

$$(6) \quad dX_i(t) = \left(\sum_{j=1}^{d+1} \delta_j \mathbf{1}_{[X_i(t)=X_{(j)}(t)]} \right) dt + \left(\sum_{j=1}^{d+1} \sigma_j \mathbf{1}_{[X_i(t)=X_{(j)}(t)]} \right) dW_i(t)$$

for $1 \leq i \leq d + 1$, where $\{X_{(j)}(t) : t \geq 0\}$ denotes the trajectory of the rank j particle as a function of time t ($X_{(1)}(t) \leq \dots \leq X_{(d+1)}(t)$ for all $t \geq 0$), δ_j, σ_j denote the drift and diffusion coefficients of the rank j particle, and $W_i, 1 \leq i \leq d$, are mutually independent standard one-dimensional Brownian motions. We will assume throughout that $\sigma_i > 0$ for all $1 \leq i \leq d + 1$. Rank-based diffusions have been proposed and extensively studied as models for problems in finance and economics. A special case is the Atlas model [13] where the minimum particle (i.e., the particle with rank 1) is a Brownian motion with positive drift and the remaining particles are Brownian motions without drift (i.e. $\delta_i = 0$ for all $i > 1$). The general setting considered in (6) was introduced in [2]. In order to study the long time behavior, it is convenient to consider the gap process $Y = (Y_1, \dots, Y_d)$, given by $Y_i = X_{(i+1)} - X_{(i)}$ for $1 \leq i \leq d$. The process $Y \equiv Y(t; \mathbf{y})$ is a RBM in \mathbb{R}_+^d given as

$$Y(t; \mathbf{y}) = \mathbf{y} + \mathbf{D}B(t) + \mu t + RL(t),$$

where \mathbf{y} is the initial gap sequence, B is a standard d -dimensional Brownian motion, $\mu_i = \delta_{i+1} - \delta_i$ for $1 \leq i \leq d$, $\mathbf{D} \in \mathbb{R}^{d \times d}$, L is the local time process associated with Y and R satisfies Assumption (A1). The covariance matrix $\Sigma = \mathbf{D}\mathbf{D}^T$ has entries $\Sigma_{ii} = \sigma_i^2 + \sigma_{i+1}^2$ when $1 \leq i \leq d$, $\Sigma_{i(i-1)} = -\sigma_i^2$ for $2 \leq i \leq d$, $\Sigma_{i(i+1)} = -\sigma_{i+1}^2$ for $1 \leq i \leq d - 1$ and $\Sigma_{ij} = 0$ otherwise. In particular, (A3) is satisfied, namely Σ is positive definite. Moreover, R is given explicitly as $R = I - P^T$, where P is the substochastic matrix given by $P_{i(i+1)} = P_{i(i-1)} = 1/2$ for all $2 \leq i \leq d - 1$, $P_{12} = P_{d(d-1)} = 1/2$ and $P_{ij} = 0$ if $|i - j| \geq 2$. From [16] the

process is positive recurrent and has a unique stationary distribution if Assumption (A2) is satisfied, namely $\mathbf{b} = -R^{-1}\mu > 0$, which is same as the following condition:

$$(7) \quad b_k = \sum_{i=1}^k (\delta_i - \bar{\delta}) > 0 \quad \text{for } 1 \leq k \leq d, \quad \text{where } \bar{\delta} = (d+1)^{-1} \sum_{j=1}^{d+1} \delta_j.$$

In the special case where

$$(8) \quad \sigma_{i+1}^2 - \sigma_i^2 = \sigma_2^2 - \sigma_1^2 \quad \text{for all } 1 \leq i \leq d,$$

the stationary distribution is explicit and takes the form $\mathcal{L}(\mathbf{Y}(\infty)) = \otimes_{k=1}^d \text{Exp}(2b_k(\sigma_k^2 + \sigma_{k+1}^2)^{-1})$ (see Section 5 of [19]). For the general case (i.e., σ_i are strictly positive and (7) is satisfied) explicit formulas for stationary distribution are not available, however from [7], the law of $Y(t; \mathbf{y})$ converges to the unique stationary distribution in (weighted) total variation distance at an exponential rate. As noted previously, this result does not provide information on parameter or dimension dependence of the rate of convergence. The paper [18] provides explicit rate of convergence to stationarity, that shows a clear parameter dependence, under the stability condition (7) and the assumption that $\sigma_i = 1$ for all $1 \leq i \leq d$. In this case the stationary measure takes an explicit form and the process is reversible with respect to the stationary measure. The proofs in [18], which are based on Dirichlet form techniques, crucially make use of these properties. The explicit representation of the stationary measure is available only under the skew-symmetry condition (see [17]) guaranteed by (8) and the reversibility of the process with respect to this measure is not available if the σ_i are not all equal. The convergence considered in [18] corresponds to that of time averages of bounded functionals of the state process to the corresponding stationary values in probability (see Theorem 1 of [18]), which is considerably weaker than the L^1 -Wasserstein distance or total variation convergence.

From Theorem 1 we have the following bound on the rate of L^1 -Wasserstein convergence of the gap process to $\mathbf{Y}(\infty)$. Note that we do not require reversibility or an explicit expression for the stationary measure.

Two key quantities appearing in the rate of convergence are

$$(9) \quad a^* := \sup_{1 \leq i \leq d} \frac{i(d+1-i)}{b_i}, \quad \sigma = \left(\sup_{1 \leq i \leq d} \sigma_i \right) \vee \left(\sup_{1 \leq i \leq d} \sigma_i^{-1} \right),$$

where b_i are defined in (7) and σ_i is the standard deviation of the rank i particle (see (6)).

THEOREM 4. *There exist positive constants F_1, F_2, F_3, F_4, t_2 such that for any $\mathbf{y} \in \mathbb{R}_+^d$ and any $t \geq t_2 \max\{\sigma^2 a^* \|\mathbf{y}\|_\infty, 1 + \sigma^2 a^{*2} \log(2d)\}$,*

$$\begin{aligned} \mathbb{E}(\|Y(t; \mathbf{y}) - Y(t; \mathbf{Y}(\infty))\|_1) &\leq 2(2\|\mathbf{y}\|_1 + F_1 \sigma^2 a^* d^3) e^{-F_2 t / [d^2(1 + \sigma^2 a^{*2} \log(2d))]} \\ &\quad + (4\|\mathbf{y}\|_1 + F_1 \sigma^2 a^* d^3) e^{-F_4 t / [2\sigma^4 a^{*2} (d+1)^2]} \\ &\quad + F_3 \sigma^2 a^* d^{7/2} e^{-F_4 t / [\sigma^4 a^{*2} (d+1)^2]}. \end{aligned}$$

In particular, the relaxation time satisfies

$$\begin{aligned} t_{\text{rel}}(\mathbf{y}) &\leq \max \{ F_2^{-1} [d^2(1 + \sigma^2 a^{*2} \log(2d))] \log[8(2\|\mathbf{y}\|_1 + F_1 \sigma^2 a^* d^3)] \\ &\quad + F_4^{-1} \sigma^4 a^{*2} (d+1)^2 [2 \log[8(4\|\mathbf{y}\|_1 + F_1 \sigma^2 a^* d^3)] + \log(8F_3 \sigma^2 a^* d^{7/2})], \\ &\quad t_2 \max\{\sigma^2 a^* \|\mathbf{y}\|_\infty, 1 + \sigma^2 a^{*2} \log(2d)\} \}. \end{aligned}$$

PROOF. Direct calculation shows that R^{-1} takes the form

$$(R^{-1})_{ij} = \begin{cases} \frac{2j(d+1-i)}{(d+1)} & \text{if } j \leq i, \\ \frac{2i(d+1-j)}{(d+1)} & \text{if } j > i. \end{cases}$$

Therefore,

$$(10) \quad \sum_{j=1}^d (R^{-1})_{ij} = (R^{-1}\mathbf{1})_i = i(d+1-i).$$

Using (10) and recalling (9), we obtain

$$(11) \quad \begin{aligned} a(\Theta) &= \sup_{1 \leq i \leq d} \left[\frac{\sum_{j=1}^d (R^{-1})_{ij} \sigma_j}{b_i} \right] \leq \sigma \sup_{1 \leq i \leq d} \frac{i(d+1-i)}{b_i} = \sigma a^*, \\ b(\Theta) &= \sup_{1 \leq i \leq d} \left[\frac{\sum_{j=1}^d (R^{-1})_{ij} \sigma_j}{\sigma_i} \right] \leq \sigma^2 \sup_{1 \leq i \leq d} i(d+1-i) \leq \sigma^2 \frac{(d+1)^2}{4}, \\ R_2(\Theta) &= a(\Theta)^2 b(\Theta) \leq \frac{\sigma^4 a^{*2} (d+1)^2}{4}, \\ C_1(\mathbf{y}, \Theta) &= 2\|\mathbf{y}\|_1 + a(\Theta) \sum_{i,j} (R^{-1})_{ij} \sigma_j \\ &\leq 2\|\mathbf{y}\|_1 + \sigma^2 a^* \sum_{i=1}^d i(d+1-i) \\ &= 2\|\mathbf{y}\|_1 + \sigma^2 a^* \frac{d(d+1)(d+2)}{6}. \end{aligned}$$

To compute $R_1(\Theta, d)$, we need to estimate $n(R)$. To do this, let $\{S_n^*\}_{n \geq 0}$ denote a simple, symmetric random walk on \mathbb{Z} starting from $S_0^* \in \{1, 2, \dots, d\}$ and absorbed when it hits 0 or $d+1$. Then, for any $n \geq 0$,

$$(P^n \mathbf{1})_i = \mathbb{P}(S_n^* \notin \{0, d+1\} \mid S_0^* = i) = \mathbb{P}(S_k^* \notin \{0, d+1\} \text{ for any } 1 \leq k \leq n \mid S_0^* = i).$$

For $j \in \{0, 1, \dots, d+1\}$, define $\tau_j^* := \inf\{k \geq 0 : S_k^* = j\}$. By Chapter 10, Example 10.17 of [21], for any $i \in \{1, \dots, d\}$, $\mathbb{E}(\tau_{d+1}^{S_0^*} \wedge \tau_0^{S_0^*} \mid S_0^* = i) = i(d+1-i)$. Using this observation and Markov's inequality, for all $d \in \mathbb{N}$ and $i \in \{1, \dots, d\}$,

$$\begin{aligned} &\mathbb{P}(S_k^* \notin \{0, d+1\} \text{ for any } 1 \leq k \leq 2d^2 \mid S_0^* = i) \\ &\leq \mathbb{P}(\tau_{d+1}^{S_0^*} \wedge \tau_0^{S_0^*} > 2d^2 \mid S_0^* = i) \\ &\leq \frac{\mathbb{E}(\tau_{d+1}^{S_0^*} \wedge \tau_0^{S_0^*} \mid S_0^* = i)}{2d^2} = \frac{i(d+1-i)}{2d^2} \leq \frac{(d+1)^2}{8d^2} \leq 1/2 \end{aligned}$$

and consequently,

$$(12) \quad n(R) \leq 2d^2.$$

Using (11) and (12), we obtain

$$(13) \quad R_1(\Theta, d) = n(R)(1 + a(\Theta)^2 \log(2d)) \leq 2d^2(1 + \sigma^2 a^{*2} \log(2d)).$$

For $t \geq 48\sigma^2 a^* \|\mathbf{y}\|_\infty$, using the bound on $a(\Theta)$ obtained in (11), and noting $b(\Theta) \geq 1$ and $\|\mathbf{y}\|_\infty^* \leq \sigma \|\mathbf{y}\|_\infty$,

$$3(D_2 a(\Theta) b(\Theta))^{-1} \|\mathbf{y}\|_\infty^* \leq 3(D_2 R_2(\Theta))^{-1} a(\Theta) \sigma \|\mathbf{y}\|_\infty \leq \frac{t}{16D_2 R_2(\Theta)}.$$

Moreover, from the explicit form of R^{-1} , $\sum_{i,j} (R^{-1})_{ij}^2 \leq 4d^4$. Using the above two bounds along with (11), for $t \geq 6a^*(d+1)^2 \|\mathbf{y}\|_\infty$,

$$\begin{aligned} & C_2(\mathbf{y}, \Theta) e^{-\frac{t}{8D_2 R_2(\Theta)}} \\ &= 2\|\mathbf{y}\|_1 e^{3(D_2 a(\Theta) b(\Theta))^{-1} \|\mathbf{y}\|_\infty^*} e^{-\frac{t}{8D_2 R_2(\Theta)}} \\ &+ a(\Theta) \left[2d(1+d) \left(\sum_{i,j} (R^{-1})_{ij}^2 \right) \left(\sum_{j=1}^d \sigma_j^2 \right) \right]^{1/2} e^{-\frac{t}{8D_2 R_2(\Theta)}} \\ (14) \quad &\leq 2\|\mathbf{y}\|_1 e^{-\frac{t}{16D_2 R_2(\Theta)}} + \sigma^2 a^* \left[2d^2(1+d) \left(\sum_{i,j} (R^{-1})_{ij}^2 \right) \right]^{1/2} e^{-\frac{t}{8D_2 R_2(\Theta)}} \\ &\leq 2\|\mathbf{y}\|_1 e^{-\frac{t}{4D_2 \sigma^4 a^{*2} (d+1)^2}} + \sigma^2 a^* [8d^6(1+d)]^{1/2} e^{-\frac{t}{2D_2 \sigma^4 a^{*2} (d+1)^2}}. \end{aligned}$$

Take $F_1 = 1$, $F_2 = D_1/2$, $F_3 = 4$, $F_4 = (2D_2)^{-1}$. Using the bounds obtained in (11), (13) and (14) in Theorem 1, for any $\mathbf{y} \in \mathbb{R}_+^d$, $t \geq \max\{t_0(1 + \sigma^2 a^{*2} \log(2d)), 48\sigma^2 a^* \|\mathbf{y}\|_\infty\}$,

$$\begin{aligned} & \mathbb{E}(\|Y(t; \mathbf{y}) - Y(t; \mathbf{Y}(\infty))\|_1) \\ &\leq C_1(\mathbf{y}, \Theta) (2e^{-\frac{D_1 t}{R_1(\Theta, d)}} + e^{-\frac{t}{16D_2 R_2(\Theta)}}) + C_2(\mathbf{y}, \Theta, D_2) e^{-\frac{t}{8D_2 R_2(\Theta)}} \\ &\leq (2\|\mathbf{y}\|_1 + F_1 \sigma^2 a^* d^3) (2e^{-F_2 t/[d^2(1+\sigma^2 a^{*2} \log(2d))]} + e^{-F_4 t/[2\sigma^4 a^{*2} (d+1)^2]}) \\ &\quad + 2\|\mathbf{y}\|_1 e^{-F_4 t/[2\sigma^4 a^{*2} (d+1)^2]} + F_3 \sigma^2 a^* d^{7/2} e^{-F_4 t/[2\sigma^4 a^{*2} (d+1)^2]}. \end{aligned}$$

This proves the first part of the theorem upon taking $t_2 = \max\{t_0, 48\}$. The bound on the relaxation time follows from the first part. \square

REMARK 6. The standard Atlas model [13] is a special case of (6) with $\delta_1 = 1$, $\delta_i = 0$ for all $i \geq 2$ and $\sigma_i = 1$ for all i . For this model, using (7), for any $k \geq 1$,

$$b_k = \sum_{i=1}^k (\delta_i - \bar{\delta}) = \frac{(d+1-k)}{d+1}$$

and

$$a^* := \sup_{1 \leq i \leq d} \frac{i(d+1-i)}{b_i} = \sup_{1 \leq i \leq d} i(d+1) = d(d+1), \quad \sigma = 1.$$

Using these in Theorem 4, we obtain positive constants G_1, G_2, G_3, G_4, t_3 such that for any $\mathbf{y} \in \mathbb{R}_+^d$ and any $t \geq t_3\{d^2 \|\mathbf{y}\|_\infty, 1 + d^2 \log(2d)\}$,

$$\mathbb{E}(\|Y(t; \mathbf{y}) - Y(t; \mathbf{Y}(\infty))\|_1) \leq G_1(\|\mathbf{y}\|_1 + d^5) e^{-G_2 t/d^6 \log(2d)} + G_3 d^{11/2} e^{-G_4 t/d^6}.$$

In particular, the relaxation time for the standard Atlas model is $O(d^6(\log d)^2)$ as $d \rightarrow \infty$.

6. Bounding processes, small sets and return times. Fix a vector $v > 0$ satisfying $R^{-1}v \leq \mathbf{b}$ and consider the collection $\{X_v^+(\cdot; \mathbf{x})\}_{\mathbf{x} \in \mathbb{R}_+^d} = \text{RBM}(-v, \Sigma, I)$, where I is the identity matrix, given as

$$(15) \quad X_v^+(t; \mathbf{x}) = \mathbf{x} + \mathbf{D}B(t) - vt + L^+(t),$$

where B is the same Brownian motion as used in the synchronous coupling of $\{X(\cdot; \mathbf{x})\}_{\mathbf{x} \in \mathbb{R}_+^d}$, and L^+ is the local time process associated with X_v^+ . Observe that $X_v^+(\cdot; \mathbf{x})$ can be written as

$$X_v^+(t; \mathbf{x}) = \mathbf{x} + \mathbf{D}B(t) + \mu t + RL^*(t),$$

where $L^*(t) = R^{-1}L^+(t) + (\mathbf{b} - R^{-1}v)t$ is a nondecreasing process. By minimality of the local time process (see [25], Appendix), $L^*(t) \geq L(t)$ for all $t \geq 0$ implying $R^{-1}X(t; \mathbf{x}) \leq R^{-1}X_v^+(t; \mathbf{x})$ for every $t \geq 0$. Since in this section v will be fixed, we abbreviate $X_v^+(\cdot; \mathbf{x})$ as $X^+(\cdot; \mathbf{x})$. An optimal choice of v will be made later in Section 8. We will hereby refer to $X^+(\cdot; \mathbf{x})$ as the bounding process.

We now introduce an appropriate compact set that depends on system parameters and for which one can obtain useful bounds on exponential moments of return times to the set. In order to motivate the choice of the set consider a one-dimensional Brownian motion $W_{a,b}(t) = bW(t) - at$ with variance b^2 and drift $-a$ (here W is a standard one-dimensional Brownian motion). Standard techniques using scale functions (see [26], V.46) show that for any $a > 0$ and $b \in \mathbb{R}$, $ab^{-2} \sup_{t < \infty} W_{a,b}(t)$ has an Exponential distribution with mean $1/2$. This result says that the maximum of the i th co-ordinate of $X^+(\cdot, \mathbf{x})$ scales like $v_i \sigma_i^{-2}$. This scaling property suggests considering the following *weighted supremum norm*:

$$(16) \quad \|\mathbf{x}\|_{\infty,v} = \sup_{1 \leq i \leq d} v_i \sigma_i^{-2} x_i, \quad \mathbf{x} \in \mathbb{R}_+^d.$$

This weighted norm will play a central role in our analysis. Also define

$$(17) \quad \phi(v) = 2 \sum_{i=1}^d v_i^2 \sigma_i^{-2} / \inf_{1 \leq i \leq d} v_i^2 \sigma_i^{-2}.$$

Note that $\phi(v) \geq 2d$. For $A > 0$, consider the compact set

$$K_A := \{\mathbf{x} \in \mathbb{R}_+^d : \|\mathbf{x}\|_{\infty,v} \leq A \log \phi(v)\}$$

and define the following stopping time for the process $X^+(\cdot; \mathbf{x})$:

$$(18) \quad \tau_A^+(\mathbf{x}) := \inf\{t \geq 0 : X^+(t; \mathbf{x}) \in K_A\} = \inf\{t \geq 0 : \|X^+(t; \mathbf{x})\|_{\infty,v} \leq A \log \phi(v)\}.$$

The following lemma gives bounds on the exponential moments of the hitting time of the compact set K_A , namely $\tau_A^+(\mathbf{x})$.

LEMMA 5. *There exists $A_0 > 0$ such that for any $A \geq A_0$ and any $\mathbf{x} \in \mathbb{R}_+^d$,*

$$\mathbb{E}(e^{\frac{\Lambda(v)}{2A} \tau_A^+(\mathbf{x})}) \leq e^{3A^{-1} \|\mathbf{x}\|_{\infty,v}},$$

where $\Lambda(v) := \inf_{1 \leq i \leq d} \frac{v_i^2}{\sigma_i^2}$.

PROOF. Fix $A > 0$ and without loss of generality assume that $\|\mathbf{x}\|_{\infty,v} > A \log \phi(v)$. Consider the ‘‘Lyapunov function’’

$$V(\mathbf{y}) = \log \left(\sum_{i=1}^d e^{g(2A^{-1} v_i \sigma_i^{-2} y_i)} \right),$$

where g is any nonnegative, nondecreasing C^2 function defined on \mathbb{R}_+ such that $g'(0) = 0$, $g(u) \leq u$, $g'(u) \leq 2$, $g''(u) \leq 9$ for all $u \geq 0$ and $g(u) = u$ for all $u \geq \log 2$. An example of such a function is $g(u) = (\log 2)h(u/\log 2)\mathbf{1}_{[u \leq \log 2]} + u\mathbf{1}_{[u > \log 2]}$ where $h(u) = u^4 - 3u^3 + 3u^2$. The definition of the Lyapunov function is motivated by a similar function introduced in [3]. The main difference is that here different coordinates are weighted differently depending on system parameters. We will prove that for sufficiently large A ,

$$(19) \quad -v^T \nabla V(\mathbf{y}) + \frac{1}{2} \text{Tr}(\Sigma \nabla^2 V(\mathbf{y})) + \frac{1}{2} (\nabla V(\mathbf{y}))^T \Sigma (\nabla V(\mathbf{y})) \leq -\frac{\Lambda(v)}{2A}, \quad \mathbf{y} \in \mathbb{R}_+^d,$$

where ∇ denotes the gradient and ∇^2 denotes the Hessian. By Itô's formula, this will imply that $M(t) := \exp(V(X^+(t; \mathbf{x})) + \frac{\Lambda(v)}{2A}t)$ is a positive supermartingale and therefore, by the optional sampling theorem, for such A ,

$$\mathbb{E}(e^{\frac{\Lambda(v)}{2A}\tau_A^+(\mathbf{x})}) \leq \mathbb{E}(e^{V(X^+(\tau_A^+(\mathbf{x}); \mathbf{x})) + \frac{\Lambda(v)}{2A}\tau_A^+(\mathbf{x})}) \leq e^{V(\mathbf{x})}.$$

Since $\|\mathbf{x}\|_{\infty, v} > A \log \phi(v) > A \log d$, we have

$$V(\mathbf{x}) \leq 2A^{-1}\|\mathbf{x}\|_{\infty, v} + \log d \leq 3A^{-1}\|\mathbf{x}\|_{\infty, v}.$$

Combining the two displays we have that for A that satisfy (19)

$$(20) \quad \mathbb{E}(e^{\frac{\Lambda(v)}{2A}\tau_A^+(\mathbf{x})}) \leq e^{3A^{-1}\|\mathbf{x}\|_{\infty, v}}.$$

Thus, in order to prove the lemma, it suffices to establish (19) for sufficiently large A . Let $w_i(\mathbf{y}, A) = \frac{e^{g(2A^{-1}v_i\sigma_i^{-2}y_i)}}{\sum_{k=1}^d e^{g(2A^{-1}v_k\sigma_k^{-2}y_k)}}$. By similar calculations as in the proof of Lemma 4 of [3], it follows that

$$\begin{aligned} \text{Tr}(\Sigma \nabla^2 V(\mathbf{y})) &\leq 4A^{-2} \sum_{i=1}^d (v_i^2 \sigma_i^{-4} w_i(\mathbf{y}, A) \sigma_i^2 (g''(2A^{-1}v_i\sigma_i^{-2}y_i) + g'(2A^{-1}v_i\sigma_i^{-2}y_i)^2)) \\ &\leq 52A^{-2} \sum_{i=1}^d v_i^2 \sigma_i^{-2} w_i(\mathbf{y}, A) \end{aligned}$$

using $g''(u) \leq 9$ and $g'(u) \leq 2$ for all $u \geq 0$. Moreover,

$$\begin{aligned} &(\nabla V(\mathbf{y}))^T \Sigma (\nabla V(\mathbf{y})) \\ &= 4A^{-2} \sum_{1 \leq i, j \leq d} v_i \sigma_i^{-2} w_i(\mathbf{y}, A) g'(2A^{-1}v_i\sigma_i^{-2}y_i) \Sigma_{ij} w_j(\mathbf{y}, A) v_j \sigma_j^{-2} g'(2A^{-1}v_j\sigma_j^{-2}y_j) \\ &\leq 4A^{-2} \left(\sum_{i=1}^d v_i \sigma_i^{-1} w_i(\mathbf{y}, A) g'(2A^{-1}v_i\sigma_i^{-2}y_i) \right)^2 \\ &\leq 16A^{-2} \sum_{i=1}^d v_i^2 \sigma_i^{-2} w_i(\mathbf{y}, A), \end{aligned}$$

where we have used $\Sigma_{ij} \leq \sigma_i \sigma_j$ in the first inequality on the second line and the Cauchy–Schwarz inequality, the fact that $g'(u) \leq 2$ for all $u \geq 0$, and the fact that $\sum_{j=1}^d w_j(\mathbf{y}, A) = 1$ in the last inequality. From the above bounds, we obtain

$$(21) \quad \text{Tr}(\Sigma \nabla^2 V(\mathbf{y})) + (\nabla V(\mathbf{y}))^T \Sigma (\nabla V(\mathbf{y})) \leq 68A^{-2} \sum_{i=1}^d v_i^2 \sigma_i^{-2} w_i(\mathbf{y}, A).$$

Using the definition of w_i and the monotonicity of g ,

$$\begin{aligned}
-v^T \nabla V(\mathbf{y}) &= -2A^{-1} \sum_{i=1}^d g'(2A^{-1} v_i \sigma_i^{-2} y_i) v_i^2 \sigma_i^{-2} w_i(\mathbf{y}, A) \\
&\leq -2A^{-1} \sum_{i=1}^d v_i^2 \sigma_i^{-2} w_i(\mathbf{y}, A) \mathbf{1}_{[2A^{-1} v_i \sigma_i^{-2} y_i \geq \log 2]} \\
&= -2A^{-1} \sum_{i=1}^d v_i^2 \sigma_i^{-2} w_i(\mathbf{y}, A) + 2A^{-1} \sum_{i=1}^d v_i^2 \sigma_i^{-2} w_i(\mathbf{y}, A) \mathbf{1}_{[2A^{-1} v_i \sigma_i^{-2} y_i < \log 2]} \\
&\leq -2A^{-1} \sum_{i=1}^d v_i^2 \sigma_i^{-2} w_i(\mathbf{y}, A) + 2A^{-1} \frac{e^{g(\log 2)} \sum_{i=1}^d v_i^2 \sigma_i^{-2}}{\sum_{k=1}^d e^{g(2A^{-1} v_k \sigma_k^{-2} y_k)}}.
\end{aligned}$$

Next, note for any $\|\mathbf{y}\|_{\infty, v} > A \log \phi(v)$, there is $1 \leq j \leq d$ such that $v_j \sigma_j^{-2} y_j > A \log(\frac{2 \sum_{i=1}^d v_i^2 \sigma_i^{-2}}{\inf_{1 \leq i \leq d} v_i^2 \sigma_i^{-2}})$. Hence, since $\frac{2 \sum_{i=1}^d v_i^2 \sigma_i^{-2}}{\inf_{1 \leq i \leq d} v_i^2 \sigma_i^{-2}} \geq 2d \geq 2$, we obtain

$$\begin{aligned}
(22) \quad -v^T \nabla V(\mathbf{y}) &\leq -2A^{-1} \sum_{i=1}^d v_i^2 \sigma_i^{-2} w_i(\mathbf{y}, A) + 2A^{-1} \frac{2 \sum_{i=1}^d v_i^2 \sigma_i^{-2}}{e^{g(2 \log(\frac{2 \sum_{i=1}^d v_i^2 \sigma_i^{-2}}{\inf_{1 \leq i \leq d} v_i^2 \sigma_i^{-2}}))}} \\
&= -2A^{-1} \sum_{i=1}^d v_i^2 \sigma_i^{-2} w_i(\mathbf{y}, A) + A^{-1} \frac{(\inf_{1 \leq i \leq d} v_i^2 \sigma_i^{-2})^2}{\sum_{i=1}^d v_i^2 \sigma_i^{-2}} \\
&\leq -2A^{-1} \sum_{i=1}^d v_i^2 \sigma_i^{-2} w_i(\mathbf{y}, A) + A^{-1} \inf_{1 \leq i \leq d} v_i^2 \sigma_i^{-2} \\
&\leq -2A^{-1} \sum_{i=1}^d v_i^2 \sigma_i^{-2} w_i(\mathbf{y}, A) + A^{-1} \sum_{i=1}^d v_i^2 \sigma_i^{-2} w_i(\mathbf{y}, A) \\
&= -A^{-1} \sum_{i=1}^d v_i^2 \sigma_i^{-2} w_i(\mathbf{y}, A).
\end{aligned}$$

From (21) and (22),

$$\begin{aligned}
&-v^T \nabla V(\mathbf{y}) + \frac{1}{2} \text{Tr}(\Sigma \nabla^2 V(\mathbf{y})) + \frac{1}{2} (\nabla V(\mathbf{y}))^T \Sigma (\nabla V(\mathbf{y})) \\
&\leq -A^{-1} \sum_{i=1}^d v_i^2 \sigma_i^{-2} w_i(\mathbf{y}, A) + 34A^{-2} \sum_{i=1}^d v_i^2 \sigma_i^{-2} w_i(\mathbf{y}, A).
\end{aligned}$$

Hence, for any $A \geq 68$, we obtain

$$\begin{aligned}
&-v^T \nabla V(\mathbf{y}) + \frac{1}{2} \text{Tr}(\Sigma \nabla^2 V(\mathbf{y})) + \frac{1}{2} (\nabla V(\mathbf{y}))^T \Sigma (\nabla V(\mathbf{y})) \\
&\leq -\frac{1}{2A} \sum_{i=1}^d v_i^2 \sigma_i^{-2} w_i(\mathbf{y}, A) \leq -\frac{1}{2A} \inf_{1 \leq i \leq d} v_i^2 \sigma_i^{-2} = -\frac{\Lambda(v)}{2A}
\end{aligned}$$

proving (19) and hence the lemma holds with $A_0 = 68$. \square

The next lemma gives an estimate of the running maximum of a reflected Brownian motion with drift.

LEMMA 6. Let $X_t = x + \sigma' B_t - \mu' t - \min\{\inf_{s \leq t} (x + \sigma' B_s - \mu' s), 0\}$, where $x \geq 0$, $\sigma', \mu' > 0$ and B is a one-dimensional standard Brownian motion. Then for any $A, T > 0$ and any $x \in [0, A/2]$,

$$\mathbb{P}\left(\sup_{0 \leq t \leq T} X_t \geq A\right) \leq e^{-\frac{\mu'^2 T}{2\sigma'^2}} + (4\mu' T A^{-1} + 2)e^{-A\mu'/\sigma'^2}.$$

PROOF. Fix $A, T > 0$ and $x \in [0, A/2]$. We define the following stopping times: $\tau_0 = 0$, and for $k \geq 0$,

$$\begin{aligned} \tau_{2k+1} &:= \inf\{t \geq \tau_{2k} : X_t = 0\}, \\ \tau_{2k+2} &:= \inf\{t \geq \tau_{2k+1} : X_t = A/2\}. \end{aligned}$$

Let $\mathcal{N} := \inf\{k \geq 0 : \sup_{t \in [\tau_{2k}, \tau_{2k+1}]} X_t \geq A\}$. By the strong Markov property, $\{\tau_{2k+1} - \tau_{2k}\}_{k \geq 1}$ are i.i.d., each being distributed as the hitting time of the level $-A/2$ by the process $\sigma' B_t - \mu' t$. By [20], Exercise 5.10, for any $\alpha > 0, k \geq 1$,

$$\mathbb{E}(e^{-\alpha(\tau_{2k+1} - \tau_{2k})}) = e^{\frac{\mu' A}{2\sigma'^2} - \frac{A}{2\sigma'} \sqrt{\frac{\mu'^2}{\sigma'^2} + 2\alpha}}.$$

Thus, for any $n \geq 0$,

$$\begin{aligned} \mathbb{P}\left(\sum_{k=0}^n (\tau_{2k+1} - \tau_{2k}) < \frac{nA}{4\mu'}\right) &\leq \mathbb{P}(e^{-\alpha \sum_{k=1}^n (\tau_{2k+1} - \tau_{2k})} > e^{-\alpha n A / (4\mu')}) \\ &\leq e^{\alpha n A / (4\mu')} \mathbb{E}(e^{-\alpha \sum_{k=1}^n (\tau_{2k+1} - \tau_{2k})}) \\ &= e^{\frac{\alpha n A}{4\mu'} + \frac{n\mu' A}{2\sigma'^2} - \frac{nA}{2\sigma'} \sqrt{\frac{\mu'^2}{\sigma'^2} + 2\alpha}}. \end{aligned}$$

Optimizing the above bound in α yields the following bound:

$$(23) \quad \mathbb{P}\left(\sum_{k=0}^n (\tau_{2k+1} - \tau_{2k}) < \frac{nA}{4\mu'}\right) \leq e^{-\frac{nA\mu'}{8\sigma'^2}}.$$

Moreover, recalling that the scale function for the process $t \mapsto \sigma' B_t - \mu' t$ is given by $s(z) = e^{2\mu' z / \sigma'^2}$,

$$\mathbb{P}(\sigma' B_t - \mu' t \text{ hits } A/2 \text{ before } -A/2) = \frac{1 - e^{-A\mu'/\sigma'^2}}{e^{A\mu'/\sigma'^2} - e^{-A\mu'/\sigma'^2}} \leq e^{-A\mu'/\sigma'^2}$$

and hence, for $n \geq 1$,

$$(24) \quad \mathbb{P}(\mathcal{N} \leq n) \leq (n+1)e^{-A\mu'/\sigma'^2}.$$

From (23) and (24), for any $n \in \mathbb{N}$,

$$(25) \quad \begin{aligned} \mathbb{P}\left(\sup_{0 \leq t \leq nA/(4\mu')} X_t \geq A \mid X_0 = x\right) &\leq \mathbb{P}\left(\sum_{k=0}^n (\tau_{2k+1} - \tau_{2k}) < \frac{nA}{4\mu'}, \mathcal{N} > n\right) + \mathbb{P}(\mathcal{N} \leq n) \\ &\leq e^{-\frac{nA\mu'}{8\sigma'^2}} + (n+1)e^{-A\mu'/\sigma'^2}. \end{aligned}$$

The result follows on taking $n = \lfloor 4\mu' T A^{-1} \rfloor + 1$ in (25). \square

Recall the quantities $\Lambda(v)$ defined in the statement of Lemma 5 and $\phi(v)$ defined in (17). Define

$$M(v) := \Lambda(v) + \log \phi(v), \quad T(v) := M(v) / \Lambda(v).$$

The next lemma shows that for any $C_0 \in (0, \infty)$ there are positive constants C_1, C_2 such that whenever $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}_+^d$ satisfy $\|\mathbf{x}_1\|_{\infty, v} \leq C_0 M(v)$ and $R^{-1}\mathbf{x}_2 \leq R^{-1}\mathbf{x}_1$, with (uniform) positive probability, all the coordinates of $X(\cdot; \mathbf{x}_2)$ hit zero by time $C_2 T(v)$ and the weighted supremum norm $\|\cdot\|_{\infty, v}$ of $X^+(\cdot; \mathbf{x}_1)$ is bounded by $C_1 M(v)$ over the time interval $[0, C_2 T(v)]$.

LEMMA 7. *For any $C_0 > 0$, there exists $C_1 > C_0$ and $C_2 > 0$ such that for any $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}_+^d$ satisfying $\|\mathbf{x}_1\|_{\infty, v} \leq C_0 M(v)$ and $R^{-1}\mathbf{x}_2 \leq R^{-1}\mathbf{x}_1$,*

$$\mathbb{P}\left(\sup_{t \in [0, C_2 T(v)]} \|X^+(t; \mathbf{x}_1)\|_{\infty, v} \leq C_1 M(v), \eta^1(\mathbf{x}_2) \leq C_2 T(v)\right) \geq \frac{1}{2}.$$

PROOF. Let $\{e_i\}_{1 \leq i \leq d}$ denote the unit coordinate vectors in \mathbb{R}^d and let $S_i := \{R^{-1}\mathbf{y} : \mathbf{y} \geq 0, y_i = 0\}$ for $1 \leq i \leq d$. Let $U(t; \mathbf{x}) = R^{-1}\mathbf{x} + R^{-1}\mathbf{D}B(t) + R^{-1}\mu t$. We first claim that for any $1 \leq i \leq d$ and $T > 0$,

$$(26) \quad \{U_i(t; \mathbf{x}) = 0 \text{ for some } 0 \leq t \leq T\} \subseteq \{R^{-1}X(t; \mathbf{x}) \in S_i \text{ for some } 0 \leq t \leq T\}.$$

To see this, suppose $R^{-1}X(t; \mathbf{x}) \notin S_i$ for all $0 \leq t \leq T$. Then $X(t; \mathbf{x})$ is strictly positive over $[0, T]$. Since $(R^{-1}X)_i(t; \mathbf{x}) = U_i(t, \mathbf{x}) + L_i(t)$, we have from (2) that $(R^{-1}X)_i(t; \mathbf{x}) = U_i(t, \mathbf{x})$ for all $0 \leq t \leq T$. For any $\mathbf{y} \in S \setminus S_i$, there exists $\mathbf{z} \in \mathbb{R}_+^d$ with $z_i > 0$ such that $\mathbf{y} = R^{-1}\mathbf{z}$. Hence,

$$y_i = \sum_{j=1}^d (R^{-1})_{ij} z_j \geq (R^{-1})_{ii} z_i > 0$$

as $(R^{-1})_{ii} = (I + P^T + (P^T)^2 + \dots)_{ii} \geq 1$. Therefore, $U_i(t; \mathbf{x}) = (R^{-1}X)_i(t; \mathbf{x}) > 0$ for all $0 \leq t \leq T$. This proves (26).

Note that $U(\cdot; \mathbf{x})$ is a Brownian motion with drift in \mathbb{R}^d with covariance matrix $R^{-1}\Sigma(R^{-1})^T$ and drift vector $-b$. Write $\hat{\sigma}_i^2 := (R^{-1}\Sigma(R^{-1})^T)_{ii}$ for the variance of the i th coordinate process U_i of U . Define for each $1 \leq i \leq d$, $\tau_i^U(\mathbf{x}) = \inf\{t \geq 0 : U_i(t; \mathbf{x}) = 0\}$. Also define the vector \hat{w} given by $\hat{w}_i = \sigma_i^2 v_i^{-1}$ for $1 \leq i \leq d$.

For any i , recalling $R^{-1}v \leq \mathbf{b}$, note that

$$(R^{-1}\hat{w})_i = \sum_{j=1}^d (R^{-1})_{ij} \frac{\sigma_j^2}{v_j} \leq \sup_{1 \leq k \leq d} \frac{\sigma_k^2}{v_k} \sum_{j=1}^d (R^{-1})_{ij} v_j \leq \left(\sup_{1 \leq k \leq d} \frac{\sigma_k^2}{v_k}\right) b_i.$$

Moreover, using $\Sigma_{jk} \leq \sigma_j \sigma_k$ for all $1 \leq j, k \leq d$,

$$\begin{aligned} \hat{\sigma}_i^2 &= \sum_{j,k} (R^{-1})_{ij} \Sigma_{jk} (R^{-1})_{ik} \\ &\leq \left(\sum_{j=1}^d (R^{-1})_{ij} \sigma_j\right)^2 \leq \sup_{1 \leq k \leq d} \frac{\sigma_k^2}{v_k} \left(\sum_{j=1}^d (R^{-1})_{ij} v_j\right)^2 \leq \left(\sup_{1 \leq k \leq d} \frac{\sigma_k^2}{v_k}\right) b_i^2. \end{aligned}$$

From the above two bounds, we conclude from the definition of $T(v)$ that for any i ,

$$(27) \quad T(v) \geq \left(\frac{(R^{-1}\hat{w})_i}{b_i} \vee \frac{\hat{\sigma}_i^2}{b_i^2}\right) M(v).$$

Fix $C' > 0$ and take $\mathbf{y} \in \mathbb{R}_+^d$ satisfying $R^{-1}\mathbf{y} \leq C'(R^{-1}\hat{w})M(v)$. Using (27) and writing $N(0, 1)$ for a standard normal random variable, we obtain that for any $C'' > \max\{2C', 1\}$,

$$\begin{aligned}
 \mathbb{P}(\tau_i^U(\mathbf{y}) > C''T(v)) &\leq \mathbb{P}((R^{-1}\mathbf{y})_i + (R^{-1}\mathbf{D}B(C''T(v)) + R^{-1}\mu(C''T(v)))_i > 0) \\
 &\leq \mathbb{P}(C'(R^{-1}\hat{w})_i M(v) + \hat{\sigma}_i B_i(C''T(v)) - b_i C''T(v) > 0) \\
 &= \mathbb{P}(\hat{\sigma}_i B_i(C''T(v)) > (b_i C''T(v) - C'(R^{-1}\hat{w})_i M(v))) \\
 &\leq \mathbb{P}\left(\hat{\sigma}_i B_i(C''T(v)) > \frac{b_i C''}{2} T(v)\right) \\
 (28) \quad &= \mathbb{P}\left(N(0, 1) > \frac{b_i}{2\hat{\sigma}_i} \sqrt{C''T(v)}\right) \\
 &\leq \mathbb{P}\left(N(0, 1) > \frac{1}{2} \sqrt{C'' \log \phi(v)}\right) \\
 &\leq \mathbb{P}\left(N(0, 1) > \frac{1}{2} \sqrt{C'' \log(2d)}\right) \leq \frac{1}{(2d)^{C''/8}},
 \end{aligned}$$

where on the last line we have used (27) in the first inequality and $\phi(v) \geq 2d$ in the second inequality.

Recall the upper bounding process X^+ from (15). Note that the i th coordinate process X_i^+ is a one-dimensional reflected Brownian motion with variance σ_i^2 and drift $-v_i$. Now let $C_0 > 0$ be arbitrary and consider any $C' > \max\{2C_0, 1\}$, any $C'' > \max\{C', 2\}$. Then, from Lemma 6, for any $\mathbf{x} \in \mathbb{R}_+^d$ satisfying $\|\mathbf{x}\|_{\infty, v} \leq C_0 M(v)$,

$$\begin{aligned}
 &\mathbb{P}\left(\sup_{t \in [0, C''T(v)]} \|X^+(t; \mathbf{x})\|_{\infty, v} > C' M(v)\right) \\
 &\leq \sum_{i=1}^d \mathbb{P}\left(\sup_{t \in [0, C''T(v)]} X_i^+(t; \mathbf{x}) > C' \frac{\sigma_i^2}{v_i} (\Lambda(v) + \log \phi(v))\right) \\
 (29) \quad &\leq \sum_{i=1}^d \left(e^{-\frac{C'' v_i^2 T(v)}{2\sigma_i^2}} + \left(\frac{4C'' v_i^2 T(v)}{C' \sigma_i^2 (\Lambda(v) + \log \phi(v))} + 2 \right) e^{-C' (\Lambda(v) + \log \phi(v))} \right) \\
 &\leq d e^{-2^{-1} C'' \log \phi(v)} + \left(\frac{4C'' T(v)}{C' M(v)} \sum_{i=1}^d \frac{v_i^2}{\sigma_i^2} + 2d \right) e^{-C' \log \phi(v)},
 \end{aligned}$$

where we have used $T(v) \geq (\sigma_i^2/v_i^2)M(v) \geq (\sigma_i^2/v_i^2) \log \phi(v)$ in the last step. From the definition of $\phi(v)$ and $T(v)$ respectively,

$$\sum_{i=1}^d \frac{v_i^2}{\sigma_i^2} = \frac{\phi(v)}{2} \Lambda(v) \quad \text{and} \quad \frac{T(v)}{M(v)} = 1/\Lambda(v).$$

Using these observations in (29), we obtain

$$\begin{aligned}
 &\mathbb{P}\left(\sup_{t \in [0, C''T(v)]} \|X^+(t; \mathbf{x})\|_{\infty, v} > C' M(v)\right) \\
 (30) \quad &\leq d e^{-C'' \log \phi(v)/2} + \left(\frac{2C'' \phi(v)}{C'} + 2d \right) e^{-C' \log \phi(v)} \\
 &\leq \frac{1}{(2d)^{\frac{C''}{2}-1}} + \left(\frac{2C''}{C'} + 1 \right) \frac{1}{(2d)^{C'-1}},
 \end{aligned}$$

where once more we have used $\phi(v) \geq 2d$ to obtain the last bound.

Note that for any $C', C'' > 0$, for any $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}_+^d$ satisfying $\|\mathbf{x}_1\|_{\infty, v} \leq C_0 M(v)$ and $R^{-1}\mathbf{x}_2 \leq R^{-1}\mathbf{x}_1$,

$$\begin{aligned}
 & \mathbb{P}\left(\sup_{t \in [0, C''T(v)]} \|X^+(t; \mathbf{x}_1)\|_{\infty, v} > C'M(v) \text{ or } \eta^1(\mathbf{x}_2) > C''T(v)\right) \\
 (31) \quad &= \mathbb{P}\left(\sup_{t \in [0, C''T(v)]} \|X^+(t; \mathbf{x}_1)\|_{\infty, v} > C'M(v)\right) \\
 &+ \mathbb{P}\left(\eta^1(\mathbf{x}_2) > C''T(v), \sup_{t \in [0, C''T(v)]} \|X^+(t; \mathbf{x}_1)\|_{\infty, v} \leq C'M(v)\right).
 \end{aligned}$$

Note that for $\mathbf{z}_1, \mathbf{z}_2 \in \mathbb{R}_+^d$, if $\|\mathbf{z}_1\|_{\infty, v} \leq CM(v)$ for some $C > 0$, and $R^{-1}\mathbf{z}_2 \leq R^{-1}\mathbf{z}_1$, then $R^{-1}\mathbf{z}_2 \leq C(R^{-1}\hat{w})M(v)$. Using (26) and (28), choosing any $C'' > 1 + \max\{2C', 1\}$, we obtain by the Markov property applied at time 1,

$$\begin{aligned}
 & \mathbb{P}\left(\eta^1(\mathbf{x}_2) > C''T(v), \sup_{t \in [0, C''T(v)]} \|X^+(t; \mathbf{x}_1)\|_{\infty, v} \leq C'M(v)\right) \\
 (32) \quad & \leq \sum_{i=1}^d \sup_{\mathbf{y} \in \mathbb{R}_+^d: R^{-1}\mathbf{y} \leq C'(R^{-1}w)M(v)} \mathbb{P}(\tau_i^U(\mathbf{y}) > (C'' - 1)T(v)) \\
 & \leq \frac{d}{(2d)^{(C''-1)/8}} = \frac{1}{2(2d)^{(C''-9)/8}}.
 \end{aligned}$$

Using the estimates (30) and (32) in (31), we obtain for $C' > \max\{2C_0, 8\}$, $C'' = 2 + \max\{2C', 33\}$, and any $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}_+^d$ satisfying $\|\mathbf{x}_1\|_{\infty, v} \leq C_0 M(v)$ and $R^{-1}\mathbf{x}_2 \leq R^{-1}\mathbf{x}_1$,

$$\begin{aligned}
 & \mathbb{P}\left(\sup_{t \in [0, C''T(v)]} \|X^+(t; \mathbf{x}_1)\|_{\infty, v} > C'M(v) \text{ or } \eta^1(\mathbf{x}_2) > C''T(v)\right) \\
 & \leq \frac{1}{(2d)^{\frac{C''}{2}-1}} + \left(\frac{2C''}{C'} + 1\right) \frac{1}{(2d)^{C'-1}} + \frac{1}{2(2d)^{(C''-9)/8}} < \frac{1}{2}.
 \end{aligned}$$

The lemma follows on taking $C_1 = C', C_2 = C''$. \square

REMARK 7. Recall the quantity A_0 from Lemma 5 and consider $C_0 \geq A_0$. Let C_1 be as in Lemma 7 associated with this C_0 . Then the set $S := \{\mathbf{y} : \|\mathbf{y}\|_{\infty, v} \leq C_1 M(v)\}$ plays a role similar to that of a ‘small set’ in the theory developed in [22], in the following sense. For any $\mathbf{x} \geq 0$, (i) by Lemma 5, we have tight control over return times of the bounding process $X^+(\cdot; \mathbf{x})$ to the set $S' = \{\mathbf{y} : \|\mathbf{y}\|_{\infty, v} \leq C_0 M(v)\} \subset S$, and (ii) by Lemma 7, given that the bounding process $X^+(t; \mathbf{x})$ lies in S' for some $t \geq 0$, then with probability at least a half, all the co-ordinates of $X(\cdot; \mathbf{x})$ hit zero at least once in the time interval $[t, t + T(v)]$ without $X^+(\cdot; \mathbf{x})$ leaving S . This, in view of Lemma 2, says that $\|X(\cdot; \mathbf{x}) - X(\cdot; \mathbf{0})\|_1$ is reduced by a factor $2^{-1/n(R)}$ over this time interval.

7. Excursions from the small set. In the following lemma, we combine the estimates from Sections 3 and 6 to decompose the path of $X^+(\cdot; \mathbf{x})$ into excursions from the small set (described in Remark 7) and quantify the rate of decay of $\|X(t; \mathbf{x}) - X(t; \mathbf{0})\|_1$ as t increases.

LEMMA 8. For any $A \geq A_0$, where A_0 is the constant appearing in Lemma 5, there exist positive constants t_0, D_1 such that for any $\mathbf{x} \in \mathbb{R}_+^d$, any $v > 0$ with $R^{-1}v \leq \mathbf{b}$, and any $t \geq t_0 T(v)$,

$$\mathbb{E}(\|X(t; \mathbf{x}) - X(t; \mathbf{0})\|_1) \leq 2\|\mathbf{x}\|_1 \left(2e^{-\frac{D_1 t}{n(R)T(v)}} + e^{-\frac{\Lambda(v)t}{16A}}\right) + 2\|\mathbf{x}\|_1 e^{3A^{-1}\|\mathbf{x}\|_{\infty, v}} e^{-\frac{\Lambda(v)t}{8A}}.$$

PROOF. Fix $A \geq A_0$ and consider constants C_1, C_2 from Lemma 7 that are associated with $C_0 = A$. We also consider the following stopping times. Let $\tau_0 = \inf\{t \geq 0 : \|X^+(t; \mathbf{x})\|_{\infty, v} \leq C_1 M(v)\}$. For $k \geq 0$, having defined the stopping times τ_0, \dots, τ_{2k} , define

$$\begin{aligned} \tau_{2k+1} &:= \inf\{t \geq \tau_{2k} : \|X^+(t; \mathbf{x})\|_{\infty, v} \leq C_0 M(v)\}, \\ \tau_{2k+2} &:= \inf\{t \geq \tau_{2k+1} : \|X^+(t; \mathbf{x})\|_{\infty, v} = C_1 M(v)\} \wedge (\tau_{2k+1} + C_2 T(v)). \end{aligned}$$

Define $\mathbf{N}_t := \inf\{k \geq 0 : \tau_{2k} \leq t\}$. For $k \geq 0$, define the event

$$E_k := \{\tau_{2k+2} \geq \tau_{2k+1} + 1, \text{ and all the co-ordinates of } \{X(t; \mathbf{x})\}_{t \geq 0} \text{ hit zero in the time interval } [\tau_{2k+1} + 1, \tau_{2k+2}]\}.$$

On the event E_k , all the coordinates of $X(\cdot; \mathbf{x})$ hit zero in the time interval $[\eta^{\mathcal{N}(\tau_{2k}; \mathbf{x})} + 1, \tau_{2k+2}]$ as it contains the interval $[\tau_{2k+1} + 1, \tau_{2k+2}]$. Consequently, $\mathcal{N}(\tau_{2k+2}; \mathbf{x}) - \mathcal{N}(\tau_{2k}; \mathbf{x}) \geq 1$. Thus, for any $k \geq 0$,

$$\mathcal{N}(\tau_{2k+2}; \mathbf{x}) - \mathcal{N}(\tau_k; \mathbf{x}) \geq \mathbf{1}_{E_k}.$$

Hence,

$$(33) \quad \mathcal{N}(t; \mathbf{x}) \geq \sum_{k=0}^{\mathbf{N}_t-1} \mathbf{1}_{E_k}.$$

Let $\mathcal{F}_t := \sigma\{B(s) : 0 \leq s \leq t\}$ be the filtration generated by the Brownian motion. For $k \geq 0$, let

$$M_n := \sum_{k=0}^{n-1} (\mathbf{1}_{E_k} - \mathbb{E}(\mathbf{1}_{E_k} | \mathcal{F}_{\tau_{2k}})).$$

Then $(M_n, \mathcal{F}_{\tau_{2n}})_{n \geq 1}$ is a martingale with increments bounded by 1. By Lemma 7, for every $k \geq 0$, $\mathbb{E}(\mathbf{1}_{E_k} | \mathcal{F}_{\tau_{2k}}) \geq 1/2$. Thus, for any $\delta' > 0$, using the Azuma–Hoeffding inequality with $t \geq 4T(v)/\delta'$,

$$\begin{aligned} &\mathbb{P}\left(\sum_{k=0}^{\mathbf{N}_t-1} \mathbf{1}_{E_k} < \delta' t / (4T(v)), \mathbf{N}_t \geq \frac{\delta' t}{T(v)}\right) \\ &\leq \mathbb{P}\left(\sum_{k=0}^{\lfloor \frac{\delta' t}{T(v)} \rfloor - 1} \mathbf{1}_{E_k} < \delta' t / (4T(v))\right) \\ (34) \quad &= \mathbb{P}\left(M_{\lfloor \frac{\delta' t}{T(v)} \rfloor} < \frac{\delta' t}{4T(v)} - \sum_{k=0}^{\lfloor \frac{\delta' t}{T(v)} \rfloor - 1} \mathbb{E}(\mathbf{1}_{E_k} | \mathcal{F}_{\tau_{2k}})\right) \\ &\leq \mathbb{P}\left(M_{\lfloor \frac{\delta' t}{T(v)} \rfloor} < \frac{\delta' t}{4T(v)} - \frac{1}{2} \lfloor \frac{\delta' t}{T(v)} \rfloor\right) \\ &\leq \mathbb{P}\left(M_{\lfloor \frac{\delta' t}{T(v)} \rfloor} < -\frac{\delta' t}{8T(v)}\right) \leq e^{-\delta' t / (128T(v))}. \end{aligned}$$

Note that for any $\delta' \in (0, C_2^{-1}/2]$,

$$(35) \quad \begin{aligned} \mathbb{P}(\mathbf{N}_t < \delta' t / T(v)) &\leq \mathbb{P}\left(\tau_0 + \sum_{k=0}^{\lfloor \frac{\delta' t}{T(v)} \rfloor} (\tau_{2k+2} - \tau_{2k+1}) + \sum_{k=0}^{\lfloor \frac{\delta' t}{T(v)} \rfloor} (\tau_{2k+1} - \tau_{2k}) > t\right) \\ &\leq \mathbb{P}\left(\tau_0 + \sum_{k=0}^{\lfloor \frac{\delta' t}{T(v)} \rfloor} (\tau_{2k+1} - \tau_{2k}) > t/2\right), \end{aligned}$$

where the last inequality follows because $\tau_{2k+2} - \tau_{2k+1} \leq C_2 T(v)$ and hence, as $\delta' \in (0, C_2^{-1}/2]$,

$$(36) \quad \sum_{k=0}^{\lfloor \frac{\delta' t}{T(v)} \rfloor} (\tau_{2k+2} - \tau_{2k+1}) \leq t/2.$$

Since $C_0 = A$ and $M(v) > \log \phi(v)$, for any $k \geq 0$, conditionally on $\mathcal{F}_{\tau_{2k}}$, $\tau_{2k+1} - \tau_{2k}$ is stochastically dominated by $\tau_A^+(X^+(\tau_{2k}; \mathbf{x}))$ where $\tau_A^+(\cdot)$ is defined in (18).

For any $n \geq 0$ $s > 0$, using Lemma 5, we obtain

$$\begin{aligned} \mathbb{P}\left(\sum_{k=0}^n (\tau_{2k+1} - \tau_{2k}) > s\right) &\leq e^{-\frac{\Lambda(v)s}{2A}} \mathbb{E}\left(e^{\frac{\Lambda(v)}{2A} (\sum_{k=0}^n (\tau_{2k+1} - \tau_{2k}))}\right) \\ &= e^{-\frac{\Lambda(v)s}{2A}} \mathbb{E}\left(e^{\frac{\Lambda(v)}{2A} (\sum_{k=0}^{n-1} (\tau_{2k+1} - \tau_{2k}))}\right) \mathbb{E}\left(e^{\frac{\Lambda(v)}{2A} (\tau_{2n+1} - \tau_{2n})} | \mathcal{F}_{\tau_{2n}}\right) \\ &\leq e^{-\frac{\Lambda(v)s}{2A}} \mathbb{E}\left(e^{\frac{\Lambda(v)}{2A} (\sum_{k=0}^{n-1} (\tau_{2k+1} - \tau_{2k}))}\right) e^{3A^{-1} \|X^+(\tau_{2n}; \mathbf{x})\|_{\infty, v}} \\ &\leq e^{-\frac{\Lambda(v)s}{2A}} e^{3A^{-1} C_1 M(v)} \mathbb{E}\left(e^{\frac{\Lambda(v)}{2A} (\sum_{k=0}^{n-1} (\tau_{2k+1} - \tau_{2k}))}\right), \end{aligned}$$

where we have used $\|X^+(\tau_{2n}; \mathbf{x})\|_{\infty, v} \leq C_1 M(v)$ by definition of τ_{2n} , and we take $\sum_{k=0}^{n-1} (\tau_{2k+1} - \tau_{2k}) = 0$ when $n = 0$. Iteratively using the same argument, we obtain

$$(37) \quad \mathbb{P}\left(\sum_{k=0}^n (\tau_{2k+1} - \tau_{2k}) > s\right) \leq e^{-\frac{\Lambda(v)s}{2A}} e^{3(n+1)A^{-1} C_1 M(v)}.$$

From (37), for any positive $\delta' \leq \min\{C_2^{-1}/2, (64C_1)^{-1}\}$ and $t \geq 3T(v)/\delta'$, taking $n = \lfloor \frac{\delta' t}{T(v)} \rfloor$ and $s = t/4$,

$$(38) \quad \begin{aligned} \mathbb{P}\left(\sum_{k=0}^{\lfloor \frac{\delta' t}{T(v)} \rfloor} (\tau_{2k+1} - \tau_{2k}) > t/4\right) &\leq e^{-\frac{\Lambda(v)t}{8A}} e^{3\delta' t A^{-1} C_1 M(v)/T(v)} e^{3M(v)A^{-1} C_1} \\ &\leq e^{-\frac{\Lambda(v)t}{8A}} e^{4\delta' t A^{-1} C_1 M(v)/T(v)} \\ &= e^{-\frac{\Lambda(v)t}{8A}} e^{4\delta' t A^{-1} C_1 \Lambda(v)} \leq e^{-\frac{\Lambda(v)t}{16A}}, \end{aligned}$$

where we have used $T(v) = M(v)/\Lambda(v)$ in the equality above. Moreover, as $\tau_0 \leq \tau_A^+(\mathbf{x})$, by Lemma 5,

$$(39) \quad \mathbb{P}(\tau_0 > t/4) \leq e^{-\frac{\Lambda(v)t}{8A}} \mathbb{E}\left(e^{\frac{\Lambda(v)}{2A} \tau_A^+(\mathbf{x})}\right) \leq e^{-\frac{\Lambda(v)t}{8A}} e^{3A^{-1} \|\mathbf{x}\|_{\infty, v}}.$$

Using (36), (38) and (39) in (35), for any δ' and t as above,

$$(40) \quad \begin{aligned} \mathbb{P}(\mathbf{N}_t < \delta' t / T(v)) &\leq \mathbb{P}\left(\sum_{k=0}^{\lfloor \frac{\delta' t}{T(v)} \rfloor} (\tau_{2k+1} - \tau_{2k}) > t/4\right) + \mathbb{P}(\tau_0 > t/4) \\ &\leq e^{-\frac{\Lambda(v)t}{16A}} + e^{-\frac{\Lambda(v)t}{8A}} e^{3A^{-1} \|\mathbf{x}\|_{\infty, v}}. \end{aligned}$$

From (34) and (40), for positive $\delta' \leq \min\{C_2^{-1}/2, (64C_1)^{-1}\}$ and $t \geq 4T(v)/\delta'$,

$$\begin{aligned}
 & \mathbb{P}\left(\sum_{k=0}^{N_t-1} \mathbf{1}_{E_k} < \delta' t / (4T(v))\right) \\
 (41) \quad & \leq \mathbb{P}\left(\sum_{k=0}^{N_t-1} \mathbf{1}_{E_k} < \delta' t / (4T(v)), N_t \geq \frac{\delta' t}{T(v)}\right) + \mathbb{P}(N_t < \delta' t / T(v)) \\
 & \leq e^{-\delta' t / (128T(v))} + e^{-\frac{\Lambda(v)t}{16A}} + e^{-\frac{\Lambda(v)t}{8A}} e^{3A^{-1} \|\mathbf{x}\|_{\infty, v}}.
 \end{aligned}$$

Now, using Lemma 2, (33) and (41), for positive $\delta' \leq \min\{C_2^{-1}/2, (64C_1)^{-1}\}$ and $t \geq 4T(v)/\delta'$,

$$\begin{aligned}
 & \mathbb{E}(\|X(t; \mathbf{x}) - X(t; \mathbf{0})\|_1) \\
 & \leq 2\|\mathbf{x}\|_1 \mathbb{E}(2^{-\mathcal{N}(t; \mathbf{x})/n(R)}) \\
 & \leq 2\|\mathbf{x}\|_1 \mathbb{E}(2^{-\frac{1}{n(R)} \sum_{k=0}^{N_t-1} \mathbf{1}_{E_k}}) \\
 (42) \quad & \leq 2\|\mathbf{x}\|_1 \mathbb{P}\left(\sum_{k=0}^{N_t-1} \mathbf{1}_{E_k} < \delta' t / (4T(v))\right) + 2\|\mathbf{x}\|_1 2^{-\frac{\delta' t}{4n(R)T(v)}} \\
 & \leq 2\|\mathbf{x}\|_1 (e^{-\delta' t / (128T(v))} + e^{-\frac{\Lambda(v)t}{16A}} + e^{-\frac{\Lambda(v)t}{8A}} e^{3A^{-1} \|\mathbf{x}\|_{\infty, v}} + 2^{-\frac{\delta' t}{4n(R)T(v)}}) \\
 & \leq 2\|\mathbf{x}\|_1 (2e^{-\frac{\delta' t}{128n(R)T(v)}} + e^{-\frac{\Lambda(v)t}{16A}}) + 2\|\mathbf{x}\|_1 e^{3A^{-1} \|\mathbf{x}\|_{\infty, v}} e^{-\frac{\Lambda(v)t}{8A}}.
 \end{aligned}$$

This proves the lemma with

$$(43) \quad t_0 = 4/\delta', \quad D_1 = \delta'/128. \quad \square$$

8. Main result: Optimizing over v . In this section, we state and prove our main theorem. This will involve optimizing the bound obtained in Lemma 8 over all possible choices of v along with making an appropriate choice of A .

PROOF OF THEOREM 1. Fix any $A \geq A_0$ whose value will be appropriately chosen later. Recall from Lemma 8 that the quantities $\frac{D_1}{n(R)T(v)}$ and $\frac{\Lambda(v)}{A}$ govern the rate of decay of $\mathbb{E}(\|X(t; \mathbf{x}) - X(t; \mathbf{0})\|_1)$ for any $\mathbf{x} \in \mathbb{R}_+^d$. To obtain the result in the theorem, we first obtain a value $v(\Theta)$ of v which simultaneously maximizes $\frac{D_1}{n(R)T(v)}$ and $\frac{\Lambda(v)}{A}$ over all vectors $v > 0$ satisfying $R^{-1}v \leq \mathbf{b}$. For any such v , define the vector $s(v) > 0$ by

$$(s(v))_i = \sigma_i \inf_{1 \leq k \leq d} \frac{v_k}{\sigma_k}, \quad 1 \leq i \leq d.$$

Then, from the definition of Λ we see that $\Lambda(s(v)) = \Lambda(v)$. Moreover, $\phi(s(v)) = 2d \leq \phi(v)$ from definition of ϕ . Therefore,

$$T(s(v)) = \frac{\Lambda(s(v)) + \log(2d)}{\Lambda(s(v))} \leq \frac{\Lambda(v) + \log(\phi(v))}{\Lambda(v)} = T(v).$$

Thus, for maximizing the rate, it suffices to restrict attention to vectors v of the form $v_i = v_* \sigma_i$ for $1 \leq i \leq d$. From the constraint $R^{-1}v \leq \mathbf{b}$, we obtain

$$b_i \geq (R^{-1}v)_i = v_* \sum_{j=1}^d (R^{-1})_{ij} \sigma_j, \quad \text{for all } i,$$

and hence,

$$v_* \leq \inf_{1 \leq i \leq d} \left[\frac{b_i}{\sum_{j=1}^d (R^{-1})_{ij} \sigma_j} \right] = \frac{1}{a(\Theta)}.$$

From this observation, it follows that for any such vector v , $\Lambda(v) = v_*^2 \leq (a(\Theta))^{-2}$ and $T(v) \geq 1 + (a(\Theta))^2 \log(2d)$. The vector \tilde{v} given by $\tilde{v}_i := (a(\Theta))^{-1} \sigma_i$ for each i satisfies $R^{-1} \tilde{v} \leq \mathbf{b}$ and hence, simultaneously maximizes $\frac{D_1}{n(R)T(v)}$ and $\frac{\Lambda(v)}{A}$. Also note that $T(\tilde{v}) = 1 + (a(\Theta))^2 \log(2d)$ and $\Lambda(\tilde{v}) = (a(\Theta))^{-2}$.

From Lemma 8 with \tilde{v} in place of v , we obtain for any $\mathbf{x} \in \mathbb{R}_+^d$ and $t \geq t_0(1 + (a(\Theta))^2 \log(2d))$,

$$(44) \quad \begin{aligned} & \mathbb{E}(\|X(t; \mathbf{x}) - X(t; \mathbf{0})\|_1) \\ & \leq 2\|\mathbf{x}\|_1 \left(2e^{-\frac{D_1 t}{R_1(\Theta, d)}} + e^{-\frac{t}{16A(a(\Theta))^2}} \right) + 2\|\mathbf{x}\|_1 e^{3A^{-1}\|\mathbf{x}\|_\infty, \tilde{v}} e^{-\frac{t}{8A(a(\Theta))^2}}. \end{aligned}$$

Consider the dominating process $X^+(\cdot; \mathbf{x}) = X_v^+(\cdot; \mathbf{x})$ with $v = \tilde{v}$. Since $R^{-1}X(t; \mathbf{x}) \leq R^{-1}X^+(t; \mathbf{x})$, for each i ,

$$\begin{aligned} X_i(t; \mathbf{x}) & \leq (R^{-1}X(t; \mathbf{x}))_i \\ & \leq (R^{-1}X^+(t; \mathbf{x}))_i \\ & = \sum_{j=1}^d (R^{-1})_{ij} X_j^+(t; \mathbf{x}) \\ & \leq \left(\sup_{1 \leq k \leq d} \frac{X_k^+(t; \mathbf{x})}{\sigma_k} \right) \sum_{j=1}^d (R^{-1})_{ij} \sigma_j \end{aligned}$$

and hence,

$$\|X(t; \mathbf{x})\|_{\infty, \tilde{v}} \leq b(\Theta) \|X^+(t; \mathbf{x})\|_{\infty, \tilde{v}}, \quad t \geq 0.$$

Moreover,

$$\|X(t; \mathbf{x})\|_1 \leq \sum_{i,j} (R^{-1})_{ij} X_j(t; \mathbf{x}) \leq \sum_{i,j} (R^{-1})_{ij} X_j^+(t; \mathbf{x}), \quad t \geq 0.$$

Denote by $X(\infty)$ and $X^+(\infty)$ the random vectors sampled from the stationary distribution of $X(\cdot; \mathbf{x})$ and $X^+(\cdot; \mathbf{x})$ respectively. By [7], the laws of $X(t; \mathbf{x})$ and $X^+(t; \mathbf{x})$ converge in total variation to those of $X(\infty)$ and $X^+(\infty)$ respectively. Consequently, $\|X(\infty)\|_1$ is stochastically dominated by $\sum_{i,j} (R^{-1})_{ij} X_j^+(\infty)$ and $\|X(\infty)\|_{\infty, \tilde{v}}$ is stochastically dominated by $b(\Theta) \|X^+(\infty)\|_{\infty, \tilde{v}}$. As $X_i^+(\infty)$ is the stationary distribution of a one-dimensional reflected Brownian motion with drift $-\tilde{v}_i$ and variance σ_i^2 , it is a standard fact that $X_i^+(\infty)$ follows an exponential distribution with mean $\sigma_i^2 / (2\tilde{v}_i) = \sigma_i a(\Theta) / 2$. This implies

$$\begin{aligned} \mathbb{E}(\|X(\infty)\|_1) & \leq \mathbb{E} \left(\sum_{i,j} (R^{-1})_{ij} X_j^+(\infty) \right) = \frac{a(\Theta)}{2} \sum_{i,j} (R^{-1})_{ij} \sigma_j, \\ \mathbb{E}(\|X(\infty)\|_1)^2 & \leq \mathbb{E} \left(\sum_{i,j} (R^{-1})_{ij} X_j^+(\infty) \right)^2 \leq \frac{d(a(\Theta))^2}{2} \left(\sum_{i,j} (R^{-1})_{ij}^2 \right) \left(\sum_{j=1}^d \sigma_j^2 \right) \end{aligned}$$

and

$$\mathbb{P}(\|X^+(\infty)\|_{\infty, \tilde{v}} > t) \leq \sum_{i=1}^d \mathbb{P}(X_i^+(\infty) > a(\Theta) \sigma_i t) = d e^{-2t}.$$

Using the above estimates we have for any $D_2 \geq 9$,

$$\begin{aligned} \mathbb{E}(e^{9(D_2b(\Theta))^{-1}\|X(\infty)\|_{\infty,\tilde{v}}}) &\leq \mathbb{E}(e^{9D_2^{-1}\|X^+(\infty)\|_{\infty,\tilde{v}}}) \\ &\leq \int_0^\infty \mathbb{P}(\|X^+(\infty)\|_{\infty,\tilde{v}} > D_2 \log t/9) dt \\ &\leq 1 + d \int_1^\infty e^{-2D_2 \log t/9} dt \leq 1 + d \int_1^\infty t^{-2} dt = 1 + d. \end{aligned}$$

Now fix $D_2 = \max\{A_0, 9\}$. Using the above estimates and (44) with $A = D_2b(\Theta)$ (noting that $b(\Theta) \geq 1$), we obtain for any $\mathbf{x} \in \mathbb{R}_+^d$ and $t \geq t_0(1 + (a(\Theta))^2 \log(2d))$,

$$\begin{aligned} &\mathbb{E}(\|X(t; \mathbf{x}) - X(t; \mathbf{X}(\infty))\|_1) \\ &\leq \mathbb{E}(\|X(t; \mathbf{x}) - X(t; \mathbf{0})\|_1) + \mathbb{E}(\|X(t; \mathbf{X}(\infty)) - X(t; \mathbf{0})\|_1) \\ &\leq 2\|\mathbf{x}\|_1(2e^{-\frac{D_1 t}{R_1(\Theta,d)}} + e^{-\frac{t}{16D_2 R_2(\Theta)}}) + 2\|\mathbf{x}\|_1 e^{3(D_2b(\Theta))^{-1}\|\mathbf{x}\|_{\infty,\tilde{v}}} e^{-\frac{t}{8D_2 R_2(\Theta)}} \\ &\quad + 2\mathbb{E}(\|X(\infty)\|_1)(2e^{-\frac{D_1 t}{R_1(\Theta,d)}} + e^{-\frac{t}{16D_2 R_2(\Theta)}}) \\ &\quad + 2\mathbb{E}(\|X(\infty)\|_1 e^{3(D_2b(\Theta))^{-1}\|X(\infty)\|_{\infty,\tilde{v}}}) e^{-\frac{t}{8D_2 R_2(\Theta)}} \\ &\leq 2\|\mathbf{x}\|_1(2e^{-\frac{D_1 t}{R_1(\Theta,d)}} + e^{-\frac{t}{16D_2 R_2(\Theta)}}) + 2\|\mathbf{x}\|_1 e^{3(D_2b(\Theta))^{-1}\|\mathbf{x}\|_{\infty,\tilde{v}}} e^{-\frac{t}{8D_2 R_2(\Theta)}} \\ &\quad + 2\mathbb{E}(\|X(\infty)\|_1)(2e^{-\frac{D_1 t}{R_1(\Theta,d)}} + e^{-\frac{t}{16D_2 R_2(\Theta)}}) \\ &\quad + 2(\mathbb{E}(\|X(\infty)\|_1^2))^{1/2} (\mathbb{E}(e^{9(D_2b(\Theta))^{-1}\|X(\infty)\|_{\infty,\tilde{v}}}))^{1/2} e^{-\frac{t}{8D_2 R_2(\Theta)}} \\ &\leq 2\|\mathbf{x}\|_1(2e^{-\frac{D_1 t}{R_1(\Theta,d)}} + e^{-\frac{t}{16D_2 R_2(\Theta)}}) + 2\|\mathbf{x}\|_1 e^{3(D_2b(\Theta))^{-1}\|\mathbf{x}\|_{\infty,\tilde{v}}} e^{-\frac{t}{8D_2 R_2(\Theta)}} \\ &\quad + a(\Theta) \sum_{i,j} (R^{-1})_{ij} \sigma_j (2e^{-\frac{D_1 t}{R_1(\Theta,d)}} + e^{-\frac{t}{16D_2 R_2(\Theta)}}) \\ &\quad + a(\Theta) \left[2d(1+d) \left(\sum_{i,j} (R^{-1})_{ij}^2 \right) \left(\sum_{j=1}^d \sigma_j^2 \right) \right]^{1/2} e^{-\frac{t}{8D_2 R_2(\Theta)}} \end{aligned}$$

which proves the Wasserstein bound in the theorem upon noting that $\|\mathbf{x}\|_{\infty,\tilde{v}} = (a(\Theta))^{-1}\|\mathbf{x}\|_{\infty}^*$. This, in turn, implies the stated bound on the relaxation time. \square

REMARK 8. To obtain the better bound displayed in Remark 3, note that we can replace the bound in Lemma 8 by

$$\mathbb{E}(\|X(t; \mathbf{x}) - X(t; \mathbf{0})\|_1) \leq 2\|\mathbf{x}\|_1 (e^{-\frac{\delta' t}{128T(v)}} + e^{-\frac{\Lambda(v)t}{16A}} + e^{-\frac{\Lambda(v)t}{8A}} e^{3A^{-1}\|\mathbf{x}\|_{\infty,v}} + 2^{-\frac{\delta' t}{4n(R)T(v)}})$$

which follows from the second-to-last inequality in the calculation (42). The vector \tilde{v} in the proof of Theorem 1 still optimizes the above bound over all v and leads to the bound displayed in Remark 3.

Acknowledgements. S. Banerjee was partially supported by a Junior Faculty Development Award made by UNC, Chapel Hill. Research of A. Budhiraja was partially supported by the National Science Foundation Grants DMS-1814894 and DMS-1853968. We thank Kavita Ramanan and Andrey Sarantsev for helpful discussions. We also thank two anonymous referees whose comments and suggestions helped to improve the article.

REFERENCES

- [1] ATAR, R., BUDHIRAJA, A. and DUPUIS, P. (2001). On positive recurrence of constrained diffusion processes. *Ann. Probab.* **29** 979–1000. MR1849184 <https://doi.org/10.1214/aop/1008956699>
- [2] BANNER, A. D., FERNHOLZ, R. and KARATZAS, I. (2005). Atlas models of equity markets. *Ann. Appl. Probab.* **15** 2296–2330. MR2187296 <https://doi.org/10.1214/105051605000000449>
- [3] BLANCHET, J. and CHEN, X. (2016). Rates of convergence to stationarity for multidimensional RBM. ArXiv preprint. Available at arXiv:1601.04111.
- [4] BOLLEY, F., GUILLIN, A. and MALRIEU, F. (2010). Trend to equilibrium and particle approximation for a weakly selfconsistent Vlasov–Fokker–Planck equation. *ESAIM Math. Model. Numer. Anal.* **44** 867–884. MR2731396 <https://doi.org/10.1051/m2an/2010045>
- [5] BRAMSON, M. and DAI, J. G. (2001). Heavy traffic limits for some queueing networks. *Ann. Appl. Probab.* **11** 49–90. MR1825460 <https://doi.org/10.1214/aoap/998926987>
- [6] BUDHIRAJA, A. and DUPUIS, P. (1999). Simple necessary and sufficient conditions for the stability of constrained processes. *SIAM J. Appl. Math.* **59** 1686–1700. MR1699034 <https://doi.org/10.1137/S0036139997330222>
- [7] BUDHIRAJA, A. and LEE, C. (2007). Long time asymptotics for constrained diffusions in polyhedral domains. *Stochastic Process. Appl.* **117** 1014–1036. MR2340877 <https://doi.org/10.1016/j.spa.2006.11.007>
- [8] DAI, J. G. (1995). On positive Harris recurrence of multiclass queueing networks: A unified approach via fluid limit models. *Ann. Appl. Probab.* **5** 49–77. MR1325041
- [9] DUPUIS, P. and WILLIAMS, R. J. (1994). Lyapunov functions for semimartingale reflecting Brownian motions. *Ann. Probab.* **22** 680–702. MR1288127
- [10] EBERLE, A. (2016). Reflection couplings and contraction rates for diffusions. *Probab. Theory Related Fields* **166** 851–886. MR3568041 <https://doi.org/10.1007/s00440-015-0673-1>
- [11] EBERLE, A., GUILLIN, A. and ZIMMER, R. (2019). Quantitative Harris-type theorems for diffusions and McKean–Vlasov processes. *Trans. Amer. Math. Soc.* **371** 7135–7173. MR3939573 <https://doi.org/10.1090/tran/7576>
- [12] EBERLE, A., GUILLIN, A. and ZIMMER, R. (2019). Couplings and quantitative contraction rates for Langevin dynamics. *Ann. Probab.* **47** 1982–2010. MR3980913 <https://doi.org/10.1214/18-AOP1299>
- [13] FERNHOLZ, E. R. (2002). Stochastic portfolio theory. In *Stochastic Portfolio Theory* 1–24. Springer, Berlin.
- [14] HALFIN, S. and WHITT, W. (1981). Heavy-traffic limits for queues with many exponential servers. *Oper. Res.* **29** 567–588. MR0629195 <https://doi.org/10.1287/opre.29.3.567>
- [15] HARRISON, J. M. and REIMAN, M. I. (1981). Reflected Brownian motion on an orthant. *Ann. Probab.* **9** 302–308. MR0606992
- [16] HARRISON, J. M. and WILLIAMS, R. J. (1987). Brownian models of open queueing networks with homogeneous customer populations. *Stochastics* **22** 77–115. MR0912049 <https://doi.org/10.1080/17442508708833469>
- [17] HARRISON, J. M. and WILLIAMS, R. J. (1987). Multidimensional reflected Brownian motions having exponential stationary distributions. *Ann. Probab.* **15** 115–137. MR0877593
- [18] ICHIBA, T., PAL, S. and SHKOLNIKOV, M. (2013). Convergence rates for rank-based models with applications to portfolio theory. *Probab. Theory Related Fields* **156** 415–448. MR3055264 <https://doi.org/10.1007/s00440-012-0432-5>
- [19] ICHIBA, T., PAPATHANAKOS, V., BANNER, A., KARATZAS, I. and FERNHOLZ, R. (2011). Hybrid atlas models. *Ann. Appl. Probab.* **21** 609–644. MR2807968 <https://doi.org/10.1214/10-AAP706>
- [20] KARATZAS, I. and SHREVE, S. E. (2012). *Brownian Motion and Stochastic Calculus, Graduate Texts in Mathematics* **113**. Springer, New York.
- [21] KLENKE, A. (2013). *Probability Theory: A Comprehensive Course*, Springer, Berlin.
- [22] MEYN, S. and TWEEDIE, R. L. (2009). *Markov Chains and Stochastic Stability*, 2nd ed. Cambridge Univ. Press, Cambridge. MR2509253 <https://doi.org/10.1017/CBO9780511626630>
- [23] PAL, S. and PITMAN, J. (2008). One-dimensional Brownian particle systems with rank-dependent drifts. *Ann. Appl. Probab.* **18** 2179–2207. MR2473654 <https://doi.org/10.1214/08-AAP516>
- [24] REIMAN, M. I. (1977). Queueing networks in heavy traffic. Technical report, Stanford Univ. Calif., Dept. of Operations Research.
- [25] REIMAN, M. I. (1984). Open queueing networks in heavy traffic. *Math. Oper. Res.* **9** 441–458. MR0757317 <https://doi.org/10.1287/moor.9.3.441>
- [26] ROGERS, L. C. G. and WILLIAMS, D. (2000). *Diffusions, Markov Processes, and Martingales. Vol. 2. Cambridge Mathematical Library*. Cambridge Univ. Press, Cambridge. MR1780932 <https://doi.org/10.1017/CBO9781107590120>

- [27] STOLYAR, A. L. (1995). On the stability of multiclass queueing networks: A relaxed sufficient condition via limiting fluid processes. *Markov Process. Related Fields* **1** 491–512. [MR1403094](#)
- [28] WANG, R. J. and GLYNN, P. W. (2014). Measuring the initial transient: Reflected Brownian motion. In *Simulation Conference (WSC), 2014 Winter* 652–661. IEEE.
- [29] WILLIAMS, R. J. (1998). Diffusion approximations for open multiclass queueing networks: Sufficient conditions involving state space collapse. *Queueing Syst.* **30** 27–88. [MR1663759](#) <https://doi.org/10.1023/A:1019108819713>