

# Supplementary proofs for: Metric recovery from directed unweighted graphs

January 25, 2015

## Contents

<b>1</b>	<b>Conjecture on uniform equicontinuity of the rescaled stationary distribution</b>	<b>1</b>
<b>2</b>	<b>Full proof of Theorem 2.1</b>	<b>2</b>
2.1	Definition of the objects . . . . .	3
2.2	Quantities used in the Stroock-Varadhan criterion . . . . .	4
2.3	Statement of the Stroock-Varadhan criterion . . . . .	5
2.4	Verification of the Stroock-Varadhan conditions . . . . .	7
2.4.1	Moment conditions . . . . .	7
2.4.2	Boundary conditions . . . . .	8
2.5	Completing the proof of Theorem S2.1 . . . . .	10
<b>3</b>	<b>Generalizing to isotropic graphs</b>	<b>12</b>
<b>4</b>	<b>Recovery of distances via ball-radii</b>	<b>15</b>
4.1	Outline of proof approach . . . . .	15
4.2	The case of exact knowledge of $\varepsilon_n$ . . . . .	15
4.3	The case of stochastic estimates of $\varepsilon_n$ . . . . .	17

## 1 Conjecture on uniform equicontinuity of the rescaled stationary distribution

In the conditions  $(\star)$  we imposed, we required the uniform equicontinuity of  $n\pi_{X_n}$ . Without this condition, our proof technique implies the weak convergence

$$\sum_{x \in \mathcal{X}_n} \pi_{X_n}(x) \delta_x \rightarrow \pi_Y(x) dx$$

of the empirical stationary measures of  $X_n(t)$  to the stationary measure of  $Y(t)$ . The additional imposition of uniform equicontinuity was required solely to upgrade this convergence to a convergence of the rescaled discrete density functions

to the continuous density function. We conjecture that this continuity is true in general.

**Conjecture S1.1.** *Given the other continuity and scaling conditions on  $p(x)$  and  $\varepsilon_n(x)$  in  $(\star)$ ,  $n\pi_{X_n}(x)$  is a.s. uniformly equicontinuous.*

We discuss a few reasons why we might believe this conjecture to hold.

- In the case of constant  $\varepsilon_n(x)$ ,  $n\pi_{X_n}(x)$  is proportional to  $|\mathbf{NB}_n(x)|$ , hence converges to  $p(x)$  uniformly. The conjecture therefore holds in this case.
- Our empirical results produce robust results across a broad range of  $n$ ,  $\bar{\varepsilon}(x)$ , and  $p(x)$ . One possible explanation would be that Conjecture S1.1 holds for all datasets constructed according to  $(\star)$ .
- For  $x, y \in \mathcal{X}_n$ , let  $r_n(x)$  denote the expected first return time to  $x$  and  $c_n(x, y)$  denote the expected commute time from  $x$  to  $y$ . It is known that

$$\pi_{X_n}(x) = \frac{1}{r_n(x)},$$

so to show that  $n\pi_{X_n}(x)$  is uniformly equicontinuous, it suffices to show that  $\frac{n}{r_n(x)}$  is uniformly equicontinuous. Notice that

$$r_n(x) \leq c_n(x, y) + r_n(y) + c_n(y, x)$$

and that

$$r_n(y) \leq c_n(x, y) + r_n(x) + c_n(y, x),$$

which together imply that

$$|r_n(x) - r_n(y)| \leq |c_n(x, y) + c_n(y, x)|.$$

This relates continuity of  $r_n(x)$  and hence  $\pi_{X_n}(x)$  to the commute time  $c_n(x, y)$ . On the other hand, our techniques using the Stroock-Varadhan criterion yield convergence of the simple random walk  $X_n(t)$  to the Itô process  $Y(t)$  in  $\mathcal{D}([0, \infty), \bar{D})$  without assumption of uniform equicontinuity. In a scaling limit, this should lead to a relation between  $c_n(x, y)$  and a rescaling of the commute time of the corresponding Itô process. In future work, we intend to use this result to relate a scaling of  $c_n(x, y)$  to  $|x - y|$  and approach Conjecture S1.1 in conjunction with new methods for metric estimation.

## 2 Full proof of Theorem 2.1

The goal of this section will be to give a fully rigorous proof of Theorem 3.4 from the main text. We first restate the theorem as Theorem S2.1.

**Theorem S2.1.** *Under  $(\star)$ , if  $h_n \rightarrow g_n^2$  as  $n \rightarrow \infty$ , then a.s. in  $\mathcal{X}$ , the process  $X_n(\lfloor t/h_n \rfloor)$  converges in  $\mathcal{D}([0, \infty), \overline{D})$  to the isotropic  $\overline{D}$ -valued Itô process  $Y(t)$  with reflecting boundary condition defined by*

$$dY(t) = \frac{\nabla p(Y(t))}{3p(Y(t))} \bar{\varepsilon}(Y(t))^2 dt + \frac{\bar{\varepsilon}(Y(t))}{\sqrt{3}} dW(t), \quad (1)$$

where the precise meaning of the reflecting boundary condition is given in Subsection 2.1.

Our technique is an application of the Stroock-Varadhan criterion (see [2, Theorem 6.3]) for convergence of discrete time Markov processes in a bounded domain to drift-diffusion processes with reflecting boundary conditions in that domain. In what follows, we preserve the notation used by Stroock-Varadhan in [2] whenever possible.

## 2.1 Definition of the objects

In this subsection, we recall in detail the problem setup. We are given an infinite sequence  $\mathcal{X} = \{x_1, x_2, \dots\}$  of latent coordinate points drawn independently from a distribution with probability density  $p(x)$  in  $\mathbb{R}^d$  supported on a compact domain  $D \subset \mathbb{R}^d$  with smooth boundary  $\partial D$ . We may then find a bounded  $C^2$  function  $\phi(x)$  on  $\mathbb{R}^d$  so that

$$D = \{x \mid \phi(x) > 0\}, \quad \partial D = \{x \mid \phi(x) = 0\}, \text{ and } |\nabla \phi(x)| \geq 1 \text{ on } \partial D.$$

We fix a single random draw of  $\mathcal{X}$  and analyze the quenched setting.

We are then given a radius function  $\varepsilon_n(x_i)$  which may depend on the draw of  $\mathcal{X}$  and a scaling factor  $g_n$  so that

$$\lim_{n \rightarrow \infty} g_n^{-1} \varepsilon_n(x) = \bar{\varepsilon}(x)$$

for some deterministic  $\bar{\varepsilon}(x)$  on  $\overline{D}$ . Let  $G_n = (\mathcal{X}_n, E_n)$  be the unweighted directed neighborhood graph with vertex set  $\mathcal{X}_n = \{x_1, \dots, x_n\}$  and with a directed edge from  $i$  to  $j$  if and only if

$$|x_i - x_j| < \varepsilon_n(x_i).$$

Note that  $G_n$  is stochastic and depends on the specific realization of  $\mathcal{X}_n$  which is drawn.

Let  $X_n(t)$  be the simple random walk on the directed graph  $G_n$  so that  $X_n(t)$  is a discrete-time Markov process with state space  $\mathcal{X}_n$ . We normalize the timestep of  $X_n(t)$  to be  $h_n = g_n^2$  and identify  $X_n(t)$  with the continuous time process given by  $t \mapsto X_n(\lfloor t/h_n \rfloor)$ . From now on, we refer to these two processes interchangeably.

In Theorem S2.1, we wish to show that  $X_n(t)$  converges weakly in  $\mathcal{D}([0, \infty), \overline{D})$  to the continuous-time continuous-space Itô process  $Y(t)$  defined by (1) with reflecting boundary conditions. We interpret the boundary conditions in terms of

the submartingale condition of [2]. That is, we define the vector function  $\gamma(s, x)$  to be the normal vector to  $\partial D$  at  $x$  whose length is normalized so that

$$\langle \gamma(s, x), \nabla \phi(x) \rangle = 1.$$

Take also the scalar function  $\rho(s, x) = 0$ . Together,  $\gamma$  and  $\rho$  specify the boundary conditions in the following sense.

We say that a process  $Y(t)$  solves the submartingale problem for  $a$ ,  $b$ ,  $\rho$ , and  $\gamma$  if for any function  $f \in C_0^{1,2}([0, \infty) \times \overline{D})$  satisfying

$$\rho(\partial f / \partial t) + \langle \gamma, \nabla f \rangle \geq 0$$

on  $[0, \infty) \times \partial D$ , the random variable

$$f(t, Y(t)) - \int_0^t (f_s + L_s f)(s, Y(s)) 1_D(Y(s)) ds$$

is a submartingale, where

$$L_s f = \frac{1}{2} \sum_{i,j=1}^d a^{ij} \frac{\partial^2 f}{\partial x_i \partial x_j} + \sum_{j=1}^d b^j \frac{\partial f}{\partial x_j}.$$

As explained in [2], when  $\rho = 0$ , this formulation is equivalent to specifying that  $Y(t)$  satisfies (1) on the interior of  $D$  and has reflecting boundary conditions on  $\partial D$ .

## 2.2 Quantities used in the Stroock-Varadhan criterion

We now define the moment and boundary quantities which are used in the Stroock-Varadhan criterion. We follow the notations of [2]. Our discrete time Markov process  $X_n(t)$  has time increment  $h_n = g_n^2$  and transition kernel

$$\Pi_n(x, A) = p(X_n(t+1) \in A | X_n(t) = x) = \frac{|\mathcal{X}_n \cap A \cap B(x, \varepsilon_n(x))|}{|\mathcal{X}_n \cap B(x, \varepsilon_n(x))|}$$

for  $x \in \mathcal{X}_n$ , where we recall that  $\mathcal{X}_n \cap B(x, \varepsilon_n(x)) = \text{NB}_n(x)$ .

The moment quantities in [2] are the discrete time drift  $b_n$ , diffusion  $a_n$ , and tail  $\Delta_{n,\alpha}$  coefficients, defined for  $x \in \mathcal{X}_n$  by

$$\begin{aligned} a_n^{ij}(s, x) &= \frac{1}{h_n} \int (y_i - x_i)(y_j - x_j) \Pi_n(x, dy) = \frac{1}{h_n} \sum_{y \in \text{NB}_n(x)} \frac{(y_i - x_i)(y_j - x_j)}{|\text{NB}_n(x)|} \\ b_n^i(s, x) &= \frac{1}{h_n} \int (y_i - x_i) \Pi_n(x, dy) = \frac{1}{h_n} \sum_{y \in \text{NB}_n(x)} \frac{y_i - x_i}{|\text{NB}_n(x)|} \\ \Delta_{n,\alpha}(s, x) &= \frac{1}{h_n} \int |y - x|^{2+\alpha} \Pi_n(x, dy) = \frac{1}{h_n} \sum_{y \in \text{NB}_n(x)} \frac{|y - x|^{2+\alpha}}{|\text{NB}_n(x)|}. \end{aligned}$$

The boundary conditions are specified by  $\gamma$  and  $\rho$ , where we recall that  $\rho \equiv 0$ . We note that  $\gamma$  has the alternate expression

$$\gamma(s, x) = C_\gamma(x) \lim_{n \rightarrow \infty} \varepsilon_n(x)^{-1} \int_{|y| < \varepsilon_n(x)} y \frac{p(x+y)}{p_{\varepsilon_n(x)}(x)} dy,$$

where  $p_r(x) = \int_{|y| < r} p(x+y) dy$  for a normalization factor  $C_\gamma(x)$ . Define  $J_0 = \{(t, y) : \rho(t, y) = 0\}$  and  $J_1$  as its complement. In our setting,  $J_0 = \partial D$  and  $J_1$  is empty.

**Remark.** In the definitions above, we have included possible time dependence in all quantities to be consistent with the notation of [2]. However, all processes we consider are time-independent, so this dependence will not exist in our case.

### 2.3 Statement of the Stroock-Varadhan criterion

We now state two theorems of Stroock-Varadhan which together imply the convergence of  $X_n(t)$  to  $Y(t)$  in  $D([0, T], \overline{D})$  for any  $T > 0$ . These theorems will depend on several conditions which we label A-E and F1-4 and check in the next subsection.

**Remark.** By [3, Theorem 2.8], convergence in  $D([0, T], \overline{D})$  for all  $T > 0$  implies convergence  $D([0, \infty), \overline{D})$ . Further, by [1, Theorem 4.9.12], this implies weak convergence of the stationary measures of  $X_n(t)$  to the stationary measure of  $Y(t)$ .

The first theorem yields tightness of measures of  $X_n(t)$  on Skorokhod space.

**Theorem S2.2** ([2, Theorem 6.1]). *Suppose a discrete time Markov process  $X_n(t)$  satisfies the following conditions.*

A. (bounded tail mass): *For some  $\alpha > 0$ , as  $n \rightarrow \infty$ , we have*

$$\sup_{0 \leq t \leq T} \sup_{x \in G} \Delta_{n, \alpha}(t, x) \rightarrow 0.$$

B. (all large drifts are reflections): *There exists  $M$  and  $c$  such that for all  $n > n_0$ ,  $|b_n(t, x)| > M$  implies  $\frac{\langle \nabla \phi(x), b_n(t, x) \rangle}{|b_n(t, x)|} \geq c$ .*

C. (bounded drift outside boundary): *For every  $\delta > 0$  there exists some  $M_\delta < \infty$  such that for all  $n > n_0$ ,  $|b_n(t, x)| > M_\delta$  implies  $\phi(x) < \delta$ .*

D. (bounded diffusion): *There exists  $M < \infty$  such that for all  $n > n_0$ ,*

$$\sup_{0 \leq t \leq T} \sup_{x \in G} \|a_n(t, x)\| \leq M,$$

where  $\|\cdot\|$  denotes the Frobenius norm.

Then, the family of distributions  $P_x^n$  induced by  $X_n^x(t)$  over trajectories is conditionally compact in  $D([0, T], \overline{D})$ . Moreover, any weak limit of these is concentrated on the subset  $C([0, T], \overline{D}) \subset D([0, T], \overline{D})$ .

The next theorem yields convergence of  $X_n(t)$  under convergence of the moment quantities and some regularity conditions on the boundary.

**Theorem S2.3** ([2, Theorem 6.3]). *Suppose  $X_n(t)$  satisfies the following.*

*E. (convergence of coefficients): Drift and diffusion coefficients  $a_n$  and  $b_n$  converge uniformly on compact subsets  $K \subset [0, T] \times D$  to some  $a$  and  $b$ .*

*F1. (reflectivity at absorbing boundary): Given  $(t, y) \in J_1$  and  $\varepsilon > 0$ , there exists  $n_0 < \infty$ ,  $\delta_0 > 0$  such that if  $|t - s| < \delta_0$ ,  $|x - y| < \delta_0$ ,  $n > n_0$  and  $\langle \nabla \phi(x), a_n(s, x) \nabla \phi(x) \rangle < \delta_0$  the following hold:*

$$|a_n(s, x)| < \varepsilon \quad \text{and} \quad |b_n(s, x) - \rho^{-1}(t, y) \gamma(t, y)| < \varepsilon.$$

*F2. (bounded drift under absorption): Given  $(t, y) \in J_1$  there exist  $M_0 < \infty$  and  $\delta_0 > 0$  such that if  $|s - t| < \delta_0$  and  $|y - x| < \delta_0$ , then*

$$|b_n(s, x)| \leq M_0 \quad \text{for all } n.$$

*F3. (drift dominates diffusion on reflection): Given  $(t, y) \in J_0$  and  $M < \infty$  there exist  $\delta_0 > 0$  and  $n_0 < \infty$  such that if  $|t - s| < \delta_0$ ,  $|x - y| < \delta_0$ ,  $n > n_0$ , and  $\langle \nabla \phi(x), a_n(s, x) \nabla \phi(x) \rangle < \delta_0$ , we have*

$$|b_n(s, x)| \geq M.$$

*F4. (drifts at boundary simulate reflection): Given  $(t, y) \in J_0$  and  $\varepsilon > 0$  there exist  $\delta_0 > 0$ ,  $n < \infty$  and  $M < \infty$  such that if  $|s - t| < \delta_0$ ,  $|x - y| < \delta_0$ ,  $n > n_0$ , and  $|b_n(s, x)| > M$ , then*

$$\left| \frac{b_n(s, x)}{\langle b_n(s, x), \nabla \phi(x) \rangle} - \gamma(t, y) \right| < \varepsilon.$$

Then any weak limit  $Y(t)$  of  $X_n(t)$  in  $D([0, T], \overline{D})$  solves the submartingale problem for  $a$ ,  $b$ ,  $\rho$ , and  $\gamma$ .

Finally, we state a criterion for uniqueness of solution to the submartingale problem for  $a$ ,  $b$ ,  $\rho$ , and  $\gamma$ .

**Theorem S2.4** ([2, Theorem 5.8]). *Suppose  $a$ ,  $b$ ,  $\rho$ , and  $\gamma$  are time independent and satisfy the following conditions.*

1.  *$a$  is continuous, symmetric, and positive definite on  $\overline{D}$ ;*
2.  *$b$  is bounded and measurable;*

3.  $\gamma$  is bounded, locally Lipschitz, and on  $\partial D$  satisfies

$$\langle \gamma(x), \nabla \phi(x) \rangle \geq \beta > 0;$$

4.  $\rho(x)$  is bounded, continuous, and non-negative.

Then the solution to the submartingale problem for  $a$ ,  $b$ ,  $\rho$ , and  $\gamma$  is unique.

Combining Theorem S2.2, Theorem S2.3, and Theorem S2.4 yields the following conclusion.

**Corollary S2.5.** *Suppose that  $X_n(t)$  satisfies the conditions of Theorem S2.2, Theorem S2.3, and Theorem S2.4. Then  $X_n(t)$  converges to  $Y(t)$  in  $D([0, T], \bar{D})$ .*

*Proof.* By Theorem S2.2, some subsequential limit of  $X_n(t)$  exists. Theorem S2.3 implies that any such limit is a solution to the submartingale problem for  $a$ ,  $b$ ,  $\rho$ , and  $\gamma$ , so the uniqueness of Theorem S2.4 yields the desired result.  $\square$

## 2.4 Verification of the Stroock-Varadhan conditions

We now verify each of the nine conditions necessary for weak convergence. Conditions F1 and F2 are vacuous because  $J_1$  is empty for us. We now verify each of the remaining conditions.

### 2.4.1 Moment conditions

**Theorem S2.6** (Condition A). *As  $n \rightarrow 0$ , we have*

$$\sup_{0 \leq t \leq T} \sup_{x \in D} \Delta_{n,1}(t, x) \rightarrow 0.$$

*Specifically, we have*

$$\Delta_{n,1}(s, x) \rightarrow \lim_{n \rightarrow \infty} \frac{1}{h_n} \int_{|y| < \varepsilon_n(x)} |y|^3 \frac{p(x+y)}{p_{\varepsilon_n(x)}(x)} dy = 0.$$

*Proof.* From Lemma 3.2 with  $f(x) = |x|^3$ .  $\square$

**Theorem S2.7** (Condition E). *The sequences of drift and diffusion coefficients  $a_n \rightarrow a$  and  $b_n \rightarrow b$  converge uniformly on compact subsets  $K \subset [0, T] \times G$ . More specifically, the limiting quantities are*

$$\begin{aligned} a_n^{ij}(s, x) &\rightarrow \lim_{n \rightarrow \infty} \frac{1}{h_n} \int_{|y| < \varepsilon_n(x)} y_i y_j \frac{p(x+y)}{p_{\varepsilon_n(x)}(x)} dy \\ b_n^i(s, x) &\rightarrow \lim_{n \rightarrow \infty} \frac{1}{h_n} \int_{|y| < \varepsilon_n(x)} y_i \frac{p(x+y)}{p_{\varepsilon_n(x)}(x)} dy. \end{aligned}$$

*Proof.* From Lemma 3.2 with  $f(x) = x$  and  $f^{ij}(x) = x_i x_j$ .  $\square$

### 2.4.2 Boundary conditions

**Theorem S2.8** (Condition C). *For  $\delta > 0$ , there exists  $M_\delta < \infty$  and  $n_0$  so that for  $n > n_0$ ,  $|b_n(t, x)| > M_\delta$  implies  $\phi(x) < \delta$ .*

*Proof.* On the compact set  $\{\phi(x) \geq \delta\}$ ,  $b_n(t, x)$  converges uniformly by Theorem S2.7 and Theorem S2.14 to  $\frac{1}{3} \frac{\nabla p(x)}{p(x)} \bar{\varepsilon}(x)^2$ , hence is uniformly bounded on this set.  $\square$

**Theorem S2.9** (Condition D). *The diffusion term  $a_n$  is uniformly bounded by some  $M < \infty$  so that*

$$\sup_{s, x, n} \|a_n(s, x)\| \leq M.$$

*Proof.* By definition the diffusion term

$$a_n^{ij}(s, x) = \frac{1}{h_n} \sum_{y \in \text{NB}_n(x)} \frac{1}{|\text{NB}_n(x)|} (y_i - x_i)(y_j - x_j)$$

is an average of numbers bounded by  $\frac{\varepsilon_n(x)^2}{h_n}$ . This quantity converges to the bounded function  $\bar{\varepsilon}(x)$  as  $n \rightarrow \infty$ , yielding the result.  $\square$

**Theorem S2.10** (Condition F3). *Given  $(t, y) \in J_0$  and  $M < \infty$ , there exist  $\delta_0 > 0$  and  $n_0 < \infty$  so that if  $|t - s| < \delta_0$ ,  $|x - y| < \delta_0$ ,  $n > n_0$ , and  $\langle \nabla \phi(x), a_n(s, x) \nabla \phi(x) \rangle < \delta_0$ , then  $|b_n(s, x)| \geq M$ .*

*Proof.* For any  $\delta_1 > 0$ , by Lemma 3.2, we may choose  $n_0$  large enough so that for all  $n > n_0$  and  $x \in \mathcal{X}_n$ , we have

$$\|a_n(s, x) - a(s, x)\| < \delta_1,$$

which implies that

$$\langle \nabla \phi(x), a_n(s, x) \nabla \phi(x) \rangle \geq \left( \frac{1}{3} \bar{\varepsilon}(x)^2 - \delta_1^2 \right) |\nabla \phi(x)|^2.$$

Because  $\bar{\varepsilon}(y) > 0$ , we can choose  $\delta_0 > 0$  so that  $\bar{\varepsilon}(x)^2$  is uniformly bounded away from 0 on  $|x - y| < \delta_0$ , hence choosing  $\delta_1$  small makes the condition vacuous.  $\square$

**Theorem S2.11** (Condition F4). *Given  $(t, y) \in J_0$  and  $\varepsilon > 0$ , there exist  $\delta_0 > 0$ ,  $n_0 < \infty$ , and  $M < \infty$  so that if  $|t - s| < \delta_0$ ,  $|x - y| < \delta_0$ ,  $n > n_0$ , and  $|b_n(s, x)| > M$ , then*

$$\left| \frac{b_n(s, x)}{\langle b_n(s, x), \nabla \phi(x) \rangle} - \gamma(t, y) \right| < \varepsilon.$$

*Proof.* For any  $\varepsilon > 0$ , fix  $M > 0$  to be chosen later. Choose  $\delta_0$  small enough so that if  $|x - y| < 2\delta_0$ , we have

$$\left| p(x) - p(y) - \frac{(y - x) \cdot \nabla p(y)}{p(y)} \right| < C_1$$



for some uniform  $C_1$ . By Lemma 3.2 and the fact that  $|\nabla\phi(x)| \geq 1$  on  $\partial D$  and is continuous, we may choose  $n_0$  large enough so that for all  $n > n_0$  and  $|x - y| < \delta_0$ , we have

- $\varepsilon_n(x) < \delta_0$ ;
- $|\varepsilon_n(x)^2 h_n^{-1} - \bar{\varepsilon}(x)^2| < C_2$  for a uniform  $C_2 > 0$ ;
- $|b_n(s, x) - E[b_n(s, x)]| < M/2$  for  $x \in \mathcal{X}_n$ ;
- $\left| \frac{b_n(s, x)}{\langle b_n(s, x), \nabla\phi(x) \rangle} - \frac{E[b_n(s, x)]}{\langle E[b_n(s, x)], \nabla\phi(x) \rangle} \right| < \varepsilon/2$  for  $x \in \mathcal{X}_n$ .

If  $|b_n(s, x)| > M$  for  $n > n_0$ , then

$$|E[b_n(s, x)]| > M/2.$$

Now, orient the coordinate axes so that the first coordinate axis lies on the normal vector from  $x$  to  $\partial D$ , and let  $\tau$  be the distance from  $x$  to  $\partial D$ . In this case, we compute

$$\begin{aligned} E[b_n^1(s, x)] &= h_n^{-1} \int_{z \in B(x, \varepsilon_n(x)) \cap D} (z_1 - x_1) \frac{p(z)}{p_{\varepsilon_n(x)}(x)} dy \\ &= \frac{\varepsilon_n(x) - \min\{\tau, \varepsilon_n(x)\}}{h_n} + \frac{1}{6} \frac{\partial_1 p(x)}{p(x)} \frac{\varepsilon_n(x)^2 + \tau^2}{h_n} + C_3 \end{aligned}$$

and for  $i > 1$  that

$$E[b_n^i(s, x)] = \frac{1}{6} \frac{\partial_i p(x)}{p(x)} \frac{\varepsilon_n(x)^2}{h_n} + C_4 \quad (2)$$

for error terms  $C_3$  and  $C_4$  independent of  $n$ . Choosing  $M$  large enough, we find

$$\tau < (1 - C_5(M))\varepsilon_n(x)$$

for a constant  $C_5(M) > 0$  independent of  $n$ , which implies that

$$E[b_n^1(s, x)] \geq C_5(M) \frac{\varepsilon_n(x)}{h_n} + \frac{1}{6} \frac{\partial_1 p(x)}{p(x)} \frac{\varepsilon_n(x)^2 + \tau^2}{h_n} + C_3. \quad (3)$$

Now, notice that  $\gamma(s, y)$  is a vector purely in the normal direction to  $\partial D$  at  $y$  normalized so that  $\langle \gamma(s, y), \nabla\phi(y) \rangle = 1$ . Because the constants  $C_3, C_4, C_5(M)$  in (3) and (2) are independent of  $n$ , all terms in these equations aside from  $C_5(M) \frac{\varepsilon_n(x)}{h_n}$  scale to constants as we take  $n_0$  and  $M$  large, so  $\frac{E[b_n(s, x)]}{\langle E[b_n(s, x)], \nabla\phi(x) \rangle}$  becomes arbitrarily close to a vector purely in the normal direction to  $\partial D$  from  $x$ . Choosing  $\delta_0$  small enough makes these vectors coincide up to error  $\varepsilon/2$ , which gives the result when combined with the bound

$$\left| \frac{b_n(s, x)}{\langle b_n(s, x), \nabla\phi(x) \rangle} - \frac{E[b_n(s, x)]}{\langle E[b_n(s, x)], \nabla\phi(x) \rangle} \right| < \varepsilon/2$$

we obtained by taking  $n_0$  large.  $\square$

**Theorem S2.12** (Condition B). *There exist  $M$ ,  $c$ , and  $n_0$  so that for all  $n > n_0$ ,  $|b_n(t, x)| > M$  implies*

$$\frac{\langle \nabla \phi(x), b_n(t, x) \rangle}{|b_n(t, x)|} \geq c.$$

*Proof.* By definition,  $\gamma(t, x)$  is uniformly bounded above by some  $C_0$ . Now, by compactness of  $\partial D = J_0$ , there exists some  $\delta > 0$  so that each  $x \in \{\phi(y) < \delta\}$  has a corresponding  $x' \in \delta D$  so that the conclusion of Theorem S2.11 applies with  $\varepsilon = C_0/2$ . Taking  $M = M_\delta$  and  $n_0$  from Theorem S2.8 for this  $\delta$  and applying Theorem S2.11 yields that

$$\frac{\langle \nabla \phi(x), b_n(t, x) \rangle}{|b_n(t, x)|} \geq \frac{2}{C_0}. \quad \square$$

## 2.5 Completing the proof of Theorem S2.1

By Corollary S2.5, to complete the proof of Theorem S2.1, it suffices for us to compute the limiting terms  $a$  and  $b$  and to verify the conditions of Theorem S2.4 for uniqueness of the submartingale problem. We begin by computing the limiting  $a$  and  $b$ , for which we will need the following lemma.

**Lemma S2.13.** *For  $d \geq 2$ , let  $B_d(r)$  be the  $d$ -dimensional ball of radius  $r$  and  $V_d(r) = V_d r^d$  be its volume. As  $r \rightarrow 0$ , we have*

$$\int_{B_d(r)} x_i^n dx = \begin{cases} 0 & n \text{ odd} \\ \frac{2V_{d-1}}{n+1} r^{n+d} + o(r^{n+d}) & n \text{ even} \end{cases}$$

and

$$\int_{B_d(r)} x_i^n x_j^m dx = 0 \text{ if } n \text{ odd.}$$

*Proof.* If  $n$  is odd, both claims follow because the integrands are odd functions integrated over symmetric domains. If  $n$  is even, for the first claim we compute

$$\int_{B_d(r)} x_i^n dx = \int_{-r}^r V_{d-1}(\sqrt{r^2 - x^2}) x^n dx = \frac{2V_{d-1}}{n+1} r^{n+d} + o(r^{n+d}). \quad \square$$

**Theorem S2.14** (Drift diffusion coefficients). *The limiting integrals for drift and diffusion are*

$$\begin{aligned} a_n^{ii}(s, x) &= \frac{1}{h_n} \left( \frac{1}{3} \varepsilon_n(x)^2 + o(\varepsilon_n(x)^2) \right) \rightarrow \frac{1}{3} \bar{\varepsilon}(x)^2 \\ a_n^{ij}(s, x) &= \frac{1}{h_n} \frac{o(\varepsilon_n(x)^{d+2})}{2V_{d-1}p(x)\varepsilon_n(x)^d + o(\varepsilon_n(x)^d)} \rightarrow 0 \\ b_n^i(s, x) &= \frac{1}{h_n} \left( \frac{1}{3} \frac{\partial_i p(x)}{p(x)} \varepsilon_n(x)^2 + o(\varepsilon_n(x)^2) \right) \rightarrow \frac{\partial_i p(x)}{3p(x)} \bar{\varepsilon}(x)^2 \\ \Delta_{n,1}(x, s) &= \frac{1}{h_n} \left( \frac{\varepsilon_n(x)^{d+4} p(x) V_{d-1} + o(\varepsilon_n(x)^{d+4})}{2V_{d-1}p(x)\varepsilon_n(x)^d + o(\varepsilon_n(x)^d)} \right) \rightarrow 0. \end{aligned}$$

*Proof.* Because  $p$  is differentiable on  $D$ , for any  $x \in D$  we have the Taylor expansion

$$p(x+y) = p(x) + y \cdot \nabla p(x) + o(|y|^2)$$

of  $p$  at  $x$ , where the convergence is uniform on compact sets. For  $n$  large enough so that the ball of radius  $\varepsilon_n(x)$  centered at  $x$  lies completely inside  $D$ , we can substitute this expansion into the definitions of  $a_n$  and  $b_n$ . Using Lemma S2.13 to estimate the resulting expressions yields

$$\begin{aligned} a_n^{ii}(s, x) &= \frac{1}{h_n} \frac{\int_{|y| < \varepsilon_n(x)} y_i^2 p(x) + y_i^2 y \cdot \nabla p(x) + y_i^2 o(|y|^2) dy}{\int_{|y| < \varepsilon_n(x)} p(x) + y \cdot \nabla p(x) + o(|y|^2) dy} \\ &= \frac{1}{h_n} \frac{\frac{2}{3} V_{d-1} p(x) \varepsilon_n(x)^{d+2} + o(\varepsilon_n(x)^{d+2})}{2 V_{d-1} p(x) \varepsilon_n(x)^d + o(\varepsilon_n(x)^d)} \\ &= \frac{1}{h_n} \left( \frac{1}{3} \varepsilon_n(x)^2 + o(\varepsilon_n(x)^2) \right) \end{aligned}$$

,

$$\begin{aligned} a_n^{ij}(s, x) &= \frac{1}{h_n} \frac{\int_{|y| < \varepsilon_n(x)} y_i y_j p(x) + y_i y_j y \cdot \nabla p(x) + y_i y_j o(|y|^2) dy}{\int_{|y| < \varepsilon_n(x)} p(x) + y \cdot \nabla p(x) + o(|y|^2) dy} \\ &= \frac{1}{h_n} \frac{o(\varepsilon_n(x)^{d+2})}{2 V_{d-1} p(x) \varepsilon_n(x)^d + o(\varepsilon_n(x)^d)} \end{aligned}$$

, and

$$\begin{aligned} b_n^i(s, x) &= \frac{1}{h_n} \frac{\int_{|y| < \varepsilon_n(x)} y_i p(x) + y_i y \cdot \nabla p(x) + y_i o(|y|^2) dy}{\int_{|y| < \varepsilon_n(x)} p(x) + y \cdot \nabla p(x) + o(|y|^2) dy} \\ &= \frac{1}{h_n} \frac{\frac{2}{3} V_{d-1} \frac{\partial_i p(x)}{p(x)} \varepsilon_n(x)^{d+2} + o(\varepsilon_n(x)^{d+2})}{2 V_{d-1} p(x) \varepsilon_n(x)^d + o(\varepsilon_n(x)^d)} \\ &= \frac{1}{h_n} \left( \frac{1}{3} \frac{\partial_i p(x)}{p(x)} \varepsilon_n(x)^2 + o(\varepsilon_n(x)^2) \right). \end{aligned}$$

Defining  $S_d(r)$  to be the surface area of a radius  $r$  ball in  $d$  dimensions, we find

$$\begin{aligned} \Delta_{n,1}(s, x) &= \frac{1}{h_n} \frac{\int_{|y| < \varepsilon_n(x)} |y|^3 p(x) + |y|^3 p(x) + |y|^3 o(|y|^3) dy}{\int_{|y| < \varepsilon_n(x)} p(x) + y \cdot \nabla p(x) + o(|y|^2) dy} \\ &= \frac{1}{h_n} \frac{\int_0^{\varepsilon_n(x)} r^3 S_d(r) p(x) + o(\varepsilon_n(x)^{d+4})}{2 V_{d-1} p(x) \varepsilon_n(x)^d + o(\varepsilon_n(x)^d)} \\ &= \frac{1}{h_n} \frac{o(\varepsilon_n(x)^{d+4})}{2 V_{d-1} p(x) \varepsilon_n(x)^d + o(\varepsilon_n(x)^d)}. \end{aligned}$$

The result follows by taking the  $n \rightarrow \infty$  limit in each estimate and recalling that  $h_n$  was chosen so that  $h_n^{-1} \varepsilon_n(x)^2 \rightarrow \bar{\varepsilon}_n(x)^2$  and  $h_n^{-1} \varepsilon_n(x)^{2+\alpha} \rightarrow 0$ . The final convergence is uniform on compact sets because the convergence of the initial Taylor expansion was, each integration estimate preserves uniformity, and the limit  $h_n^{-1} \varepsilon_n(x)^2 \rightarrow \bar{\varepsilon}(x)^2$  is uniform over all of  $D$ .  $\square$

*Proof of Theorem S2.1.* To prove Theorem S2.1, it remains only to check the conditions of Theorem S2.4. Condition (1) follows because  $a(x) = \frac{1}{3}\varepsilon(x)^2 \cdot I$  is a continuous multiple of the identity. Condition (2) follows because  $b(x) = \frac{1}{3} \frac{\nabla p(x)}{p(x)} \varepsilon(x)^2$  is evidently bounded and measurable. For Condition (3),  $\gamma$  is evidently bounded, locally Lipschitz because it is a normalized vector normal to the smooth  $\partial D$ , and  $\langle \gamma(x), \nabla \phi(x) \rangle = 1$  by definition. Finally, Condition (4) is evident because  $\rho \equiv 0$ .  $\square$

### 3 Generalizing to isotropic graphs

In this section, we give details on how to generalize our results for  $\varepsilon_n(x)$ -ball graphs to isotropic graphs. The approach is exactly parallel; we verify the conditions of the Stroock-Varadhan criterion and consider the limiting rescaled stationary distribution. We give in this section the necessary estimates of the minimal degree and the drift and diffusion terms. We first present a technical lemma.

**Lemma S3.1.** *For  $d \geq 2$ , Let  $S_d(r)$  be the  $d$ -dimensional shell of radius  $r$  and  $V_d(r) = C_d r^d$  be its volume. As  $r \rightarrow 0$ , we have*

$$\int_{S_d(r)} x_i^n dx = \begin{cases} 0 & n \text{ odd} \\ \frac{2C_{d-1}}{n+1} (n+d) r^{n+d-1} + o(r^{n+d-1}) & n \text{ even} \end{cases}$$

and

$$\int_{S_d(r)} x_i^n x_j^m dx = 0 \text{ if } n \text{ odd.}$$

*Proof.* This follows by differentiating Lemma S2.13.  $\square$

Let us now consider an isotropic graph model with kernel function  $h(r)$ . In particular, this implies that there is an edge from  $x_i$  to  $x_j$  with probability  $h(|x_i - x_j| \varepsilon_n(x_i)^{-1})$  and that

$$\int_0^1 h(r) r^{d-1} dr > 0.$$

We characterize the minimal out-degree in this setting.

**Theorem S3.2** (Minimal out-degree). *For an isotropic graph with kernel  $h(r)$  satisfying  $(\star)$ , we have the almost sure convergence*

$$\varepsilon_n(x)^{-d} \frac{|\mathbf{NB}_n(x)|}{|\mathcal{X}_n \cap B(x, \varepsilon_n(x))|} \rightarrow C(h)p(x)$$

for a constant  $C(h)$  independent of  $x$  and  $n$ , which implies that the minimal degree  $|\mathbf{NB}_n(x)| = \omega(n^{2/(d+2)} \log(n)^{d/(d+2)})$ .

*Proof.* The out-degree of a vertex is the independent sum of binary variables, each with probability  $h(|x_i - x_j|\varepsilon_n(x_i)^{-1})$ , so Kolmogorov's strong law yields

$$\varepsilon_n(x)^{-d} \frac{|\mathbf{NB}_n(x)|}{|\mathcal{X}_n \cap B(x, \varepsilon_n(x))|} \rightarrow E \left[ \varepsilon_n(x)^{-d} \frac{|\mathbf{NB}_n(x)|}{|\mathcal{X}_n \cap B(x, \varepsilon_n(x))|} \right].$$

Let  $y(r, \theta)$  be the radial representation of  $y$  and let  $C = \frac{2C_{d+1}(n+d)}{n+1}$  be the constants in Lemma S3.1. The desired expected value is the integral

$$\begin{aligned} E \left[ \frac{\varepsilon_n(x)^{-d} |\mathbf{NB}_n(x)|}{|\mathcal{X}_n \cap B(x, \varepsilon_n(x))|} \right] &= \int_{y \in B(x, \varepsilon_n(x))} p(x+y) h(|y|\varepsilon_n^{-1}(x)) dy \\ &\sim \varepsilon_n(x)^{-d} \int_{y \in B(x, \varepsilon_n(x))} (p(x) + \nabla p(x) \cdot y) h(|y|\varepsilon_n^{-1}(x)) dy \\ &= \varepsilon_n(x)^{-d} \int_0^{\varepsilon_n(x)} \int_{\theta \in S_d(r)} (p(x) + \nabla p(x) \cdot y(r, \theta)) h(r) dy d\theta \\ &= Cp(x) \varepsilon_n(x)^{-d} \int_0^{\varepsilon_n(x)} h(r) r^{d-1} dr \\ &\quad + \varepsilon_n(x)^{-d} \int_0^{\varepsilon_n(x)} h(r) r^{d-1} \int_{\theta \in S_d(1)} \nabla p(x) \cdot y(1, \theta) dr d\theta. \end{aligned}$$

The latter term is zero by Lemma S3.1 since it is the integral of the odd function  $y(1, \theta)$  over a symmetric domain. Now take the substitution  $s = r/\varepsilon_n(x)$  to obtain

$$E \left[ \frac{\varepsilon_n(x)^{-d} |\mathbf{NB}_n(x)|}{|\mathcal{X}_n \cap B(x, \varepsilon_n(x))|} \right] = Cp(x) \int_0^1 h(s) s^{d-1} ds.$$

The Kolmogorov strong law provides concentration around this value. Noting that  $\varepsilon_n(x)^d = \omega(n^{2/(d+2)} \log(n)^{d/(d+2)})$  gives the asymptotic claim.  $\square$

Since Theorem S3.2 guarantees that asymptotically we achieve the necessary minimal number of points, and  $h(x)$  is zero for  $x > 1$ , Lemma 3.2 applies to show the moment conditions in the Stroock-Varadhan criterion. For the boundary conditions, note that C, D, and F3 only require convergence of coefficients in Lemma S3.3 to those in Theorem S2.14. Conditions F4 and B rely on two facts, the uniform convergence of coefficients given by Lemma S3.3, and the asymmetry induced by the boundary (3), the proof of which is parallel to the one given for  $\varepsilon$ -ball graphs. Therefore, to complete the proof the generalization, it remains only to compute the limiting drift and diffusion coefficients.

**Lemma S3.3** (Polynomial integrals with respect to kernel). *Under the same conditions as Theorem S3.2, for any positive integer  $\alpha$  we have*

$$\int_{y \in B(x, \varepsilon_n(x))} y_i^\alpha p(x+y) h(|y|\varepsilon_n^{-1}(x)) dy \sim V(h, \alpha) \int_{y \in B(x, \varepsilon_n(x))} y_i^\alpha p(x+y) dy$$

as  $n \rightarrow \infty$  for a constant  $V(h, \alpha)$  independent of  $n$  with  $V(h, 1) = V(h, 2)$ .

*Proof.* Perform the same Taylor approximation and radial decomposition as in Theorem S3.2 to obtain

$$\begin{aligned} & \int_{y \in B(x, \varepsilon_n(x))} y_i^\alpha p(x+y) h(|y| \varepsilon_n^{-1}(x)) dy \\ & \sim \int_0^{\varepsilon_n(x)} \int_{\theta \in S_d(r)} y_i(r, \theta)^\alpha (p(x) + \nabla p(x) \cdot y(r, \theta)) h(r \varepsilon_n^{-1}(x)) dr d\theta. \end{aligned}$$

For  $\alpha$  an odd integer, by Lemma S3.1 we have

$$\begin{aligned} & \int_{y \in B(x, \varepsilon_n(x))} y_i^\alpha p(x+y) h(|y| \varepsilon_n^{-1}(x)) dy \\ & \sim \int_0^{\varepsilon_n(x)} h(r \varepsilon_n^{-1}(x)) r^{\alpha+d} \int_{\theta \in S_d(1)} y_i(1, \theta)^\alpha \nabla p(x) \cdot y(1, \theta) dr d\theta \\ & = \partial p_i(x) \int_0^1 h(r) r^{\alpha+d} dr \varepsilon_n(x)^{\alpha+d} \int_{\theta \in S_d(1)} y_i(1, \theta)^{\alpha+1} d\theta \\ & \sim V(h, \alpha) \int_{y \in B(x, \varepsilon_n(x))} y_i^\alpha p(x+y) dy \end{aligned}$$

for

$$V(h, \alpha) = (\alpha + d + 1) \int_0^1 h(r) r^{\alpha+d} dr.$$

If  $\alpha$  is an even integer, we have

$$\begin{aligned} & \int_{y \in B(x, \varepsilon_n(x))} y_i^\alpha p(x+y) h(|y| \varepsilon_n^{-1}(x)) dy \\ & \sim \int_0^{\varepsilon_n(x)} h(r \varepsilon_n^{-1}(x)) r^{\alpha+d-1} \int_{\theta \in S_d(1)} y_i(1, \theta)^\alpha p(x) dr d\theta \\ & = p(x) \int_0^{\varepsilon_n(x)} h(r \varepsilon_n^{-1}(x)) r^{\alpha+d-1} dr \int_{\theta \in S_d(1)} y_i(1, \theta)^\alpha d\theta \\ & = p(x) \varepsilon_n(x)^{\alpha+d} \int_0^1 h(r) r^{\alpha+d-1} dr \int_{\theta \in S_d(1)} y_i(1, \theta)^\alpha d\theta \\ & \sim V(h, \alpha) \int_{y \in B(x, \varepsilon_n(x))} y_i^\alpha p(x+y) dy \end{aligned}$$

for

$$V(h, \alpha) = (\alpha + d) \int_0^1 h(r) r^{\alpha+d-1} dr. \quad \square$$

The limits of drift and diffusion terms in Theorem S2.14 depend only on ratios of these integrals for  $\alpha = 1, 2$ , so applying Lemma S3.3 shows that the limits for isotropic graphs are identical to the ones for  $\varepsilon$ -ball graphs. The remainder of the analysis proceeds unchanged.

## 4 Recovery of distances via ball-radii

We will prove that given the ball radii  $\varepsilon_n(x_i)$ , we can recover point-to-point distances if  $x_i$  are located in a convex domain. Otherwise, we recover the geodesic distances. Our goal is to show that for any points  $x_i$  and  $x_j$ , the weighted shortest path distance  $d_{ij}$  between the points on the graph  $\overline{G}_n$  where outgoing edges are weighted by  $\varepsilon_n(x_i)$  converges to the distance  $|x_i - x_j|$ .

### 4.1 Outline of proof approach

We proceed in two steps. First, we consider the case when  $\varepsilon_n(x_i)$  is known exactly. In this case, the weighted shortest path is an upper bound on the true distance. We bound its weighted distance  $d_{ij}$  by constructing a path whose weighted distance is close to the geodesic distance.

To control the upper bound, we show that there exists a  $\delta$  that converges to zero faster than  $\min_{x_i} \varepsilon_n(x_i)$  while still guaranteeing that every ball of size  $\delta$  in the domain contains at least one point. Once we find such a  $\delta$ , the upper bound will follow. Indeed, if we are at some  $x$ , we can always find a point that whose distance from our target  $x_j$  is smaller by at least  $\varepsilon_n(x) - \delta$ . This gives an upper bound on the number of steps in our path and therefore the total error.

Second, we assume that we are given noisy estimates of  $\varepsilon(x)$  from our algorithm via the stationary distribution. We use uniform convergence of  $\varepsilon(x)$  to control the overall pathwise error.

We give a detailed analysis of each step in separate subsections below.

### 4.2 The case of exact knowledge of $\varepsilon_n$ .

We begin with two lemmas allowing us to construct for each pair of points  $i, j$  a point  $k$  along which to start a path from  $i$  to  $j$ .

**Lemma S4.1.** *Let  $\delta_n = \Omega(n^{-\frac{1}{d+1}})$ . For any set of  $n^2$  balls with radius  $\delta_n$ , all  $n^2$  balls will have at least one point of  $\mathcal{X}_n$  with high probability.*

*Proof.* The number of points  $N(x)$  in a ball of radius  $\delta_n$  follows a binomial distribution with  $n$  draws and success probability

$$p_{\delta_n}(x) = \int_{|y-x| < \delta(n)} p(y) dy \sim V_d p(x) \delta_n^d.$$

Therefore, the probability that  $N(x) = 0$  is

$$P(N(x) = 0) = (1 - p_{\delta_n}(x))^n = \left( (1 - p_{\delta_n}(x))^{p_{\delta_n}(x)^{-1}} \right)^{np_{\delta_n}(x)} \rightarrow e^{-np_{\delta_n}(x)}$$

if  $n\delta_n^d \rightarrow \infty$ . Recalling that  $\delta_n = \Omega(n^{-\frac{1}{d+1}})$ , this implies that

$$np_{\delta_n}(x) \sim n^{\frac{1}{d+1}}$$

and in particular that  $P(N(x) = 0) = o(n^{-2})$ , so taking the union bound over all  $n^2$  balls yields the result.  $\square$

**Lemma S4.2.** *Let  $\delta_n = \Omega(n^{-\frac{1}{d+1}})$ . For all  $i, j$ , there exists  $x_k \in B(x_i, \varepsilon_n(x_i))$  such that*

$$\left| \left( |x_i - x_j| - |x_k - x_j| \right) - |x_i - x_k| \right| \leq 2\delta_n \text{ and } \left| |x_k - x_i| - \varepsilon_n(x_i) \right| \leq 2\delta_n.$$

*Proof.* Let  $v = \frac{x_j - x_i}{|x_j - x_i|}$  and consider the  $n^2$  balls

$$B_{ij} = B(x_i + v(\varepsilon_n(x_i) - \delta_n), \delta_n).$$

By Lemma S4.1, there must exist with high probability at least one point of  $\mathcal{X}_n$  in each  $B_{ij}$ . Any such  $x_k \in B_{ij}$  verifies the desired conditions.  $\square$

**Theorem S4.3.** *Let  $x_i, x_j \in \mathcal{X}_n$  and  $d_{ij}$  be the weighted shortest path distance over the weighted graph  $\bar{G}_n$  constructed from  $G_n$  by assigning weight  $\varepsilon_n(x_i)$  to all outgoing edges from  $x_i$ . For any  $\varepsilon > 0$ , there exists an  $n$  such that*

$$\left| |x_i - x_j| - d_{ij} \right| < \varepsilon.$$

*Proof.* Take  $\delta_n = \Theta(n^{-\frac{1}{d+1}})$ . We show that with high probability, there exists a path with  $M$  steps whose weighted path distance  $d$  satisfies

$$|x_i - x_j| \leq d \leq |x_i - x_j| + 2M\delta_n + \max_{x \in \mathcal{X}_n} \varepsilon_n(x)$$

and so that  $\lim_{n \rightarrow \infty} M\delta_n = 0$ . The result then follows because  $d_{ij} \leq d$ .

To construct such a path from  $x_i$  to  $x_j$ , we apply the following procedure. Start at the point  $x_i$ . If the current point is  $x_k$  and  $x_j \in B(x_k, \varepsilon_n(x_k))$ , move to it and terminate. Otherwise, pick a point  $x_l \in B_{kj}$  and repeat until  $x_j$  is reached.

The lower bound holds because each edge weight is at least its length. For the upper bound, by Lemma S4.2, moving to  $x_l$  reduces the geodesic distance to  $x_j$  by at least  $|x_k - x_l| - 2\delta_n$  and moves a weighted distance of  $\varepsilon_n(x_k) < |x_k - x_l| + 2\delta_n$ . Thus, if our path has  $M$  steps, the difference between our weighted distance and the geodesic distance is at most  $4M\delta_n + \max_x \varepsilon_n(x)$ , where we add the weighted distance of the last step. This gives the upper bound.

It remains now to bound  $M$ . For this, notice that the geodesic distance to  $x_j$  decreases by at least  $\min_{x \in \mathcal{X}_n} \varepsilon_n(x) - 2\delta_n$  at each step, leading to the bound

$$M \leq \frac{|x_i - x_j|}{\min_{x \in \mathcal{X}_n} \varepsilon_n(x) - 2\delta_n}.$$

Recall now that  $\delta_n = \Theta(n^{-\frac{1}{d+1}})$  so that  $\varepsilon_n(x) = \omega(\delta_n)$  and hence

$$M\delta_n = \frac{|x_i - x_j|}{\min_{x \in \mathcal{X}_n} \frac{\varepsilon_n(x)}{\delta_n} - 1} \rightarrow 0. \quad \square$$



### 4.3 The case of stochastic estimates of $\varepsilon_n$

We now consider the case where we are given only an estimate  $\hat{\varepsilon}_n(x)$  of  $\varepsilon_n$ , obtained by first estimating  $\bar{\varepsilon}(x)$  via the stationary distribution and then applying a normalization to obtain  $\hat{\varepsilon}_n(x)$  on  $\mathcal{X}_n$ . We first control the error in  $\hat{\varepsilon}_n(x)$  along a single path.

**Lemma S4.4.** *For  $k_1 = i$  and  $k_{l_n} = j$ , let  $x_{k_1}, \dots, x_{k_{l_n}}$  be a path between  $i$  and  $j$  in  $\bar{G}_n$ . If  $l_n = O(g_n^{-1})$ , we have*

$$\sum_{i=1}^{l_n} |\hat{\varepsilon}_n(x_{k_i}) - \varepsilon_n(x_{k_i})| \rightarrow 0$$

in probability.

*Proof.* By uniform convergence of the stationary distribution and continuity of the out degree estimate  $p(x)\varepsilon_n(x)^d V_d = k/n$ , for all  $\gamma$  and  $\delta$ , we have

$$P\left(\sup_{x \in \mathcal{X}_n} \left| \frac{\hat{\varepsilon}_n(x)}{g_n} - \bar{\varepsilon}(x) \right| > \gamma\right) < \delta$$

for large enough  $n$ . This implies that

$$P\left(\sup_{x \in \mathcal{X}_n} |\hat{\varepsilon}_n(x) - \varepsilon_n(x)| > \gamma g_n\right) < \delta.$$

Now notice that

$$P\left(\sum_{i=1}^{l_n} |\hat{\varepsilon}_n(x_{k_i}) - \varepsilon_n(x_{k_i})| > \gamma\right) < P\left(l_n \sup_x |\hat{\varepsilon}_n(x) - \varepsilon_n(x)| > \gamma\right).$$

By assumption, the number of steps in the path is  $l_n = O(g_n^{-1})$ . Therefore, there exists a constant  $M > 0$  such that

$$P\left(\sum_{i=1}^{l_n} |\hat{\varepsilon}_n(x_{k_i}) - \varepsilon_n(x_{k_i})| > \gamma\right) < P\left(\sup_x |\hat{\varepsilon}_n(x) - \varepsilon_n(x)| > M\gamma g_n\right) < \delta,$$

from which the claim follows by choosing  $n$  large enough.  $\square$

We now show that the shortest weighted distance path recovers the geodesic distance with stochastic estimates  $\hat{\varepsilon}_n(x)$  instead of the true values. Our approach is the same as in the deterministic case; we will construct a weighted path and show that its weighted distance converges to the geodesic distance and is close to the weighted distance of the shortest weighted path. Let  $\hat{d}_{ij}$  denote the weighted distance of the shortest weighted distance path from  $x_i$  to  $x_j$ .

**Theorem S4.5.** *For any  $\varepsilon > 0$ , there exists  $n$  such that*

$$\left| |x_i - x_j| - \hat{d}_{ij} \right| < \varepsilon$$

with high probability.

*Proof.* Let  $\delta_n = \Theta(n^{-\frac{1}{d+2}})$ . For any  $\gamma > 0$ , we show that for large enough  $n$ , with high probability there exists a path from  $x_i$  to  $x_j$  with  $M$  steps whose weighted path distance  $\widehat{d}$  satisfies

$$\widehat{d} \leq |x_i - x_j| + 4M\delta_n + \gamma + \max_{x \in \mathcal{X}_n} \widehat{\varepsilon}_n(x). \quad (4)$$

Construct the path as in Theorem S4.3 with  $\varepsilon_n(x)$  replaced by  $\widehat{\varepsilon}_n(x)$ .

We now analyze its weighted distance. Arguing as in Lemma S4.2, a step from  $x_k$  to  $x_l$  which is not the last step in this path reduces the geodesic distance to  $x_j$  by between  $|x_k - x_l| - 2\delta_n$  and  $|x_k - x_l|$ . On the other hand, this step has a weighted distance of  $\widehat{\varepsilon}_n(x_k)$ , which satisfies

$$|x_k - x_l| - 2\delta_n - |\widehat{\varepsilon}_n(x_k) - \varepsilon_n(x_k)| \leq \widehat{\varepsilon}_n(x_k) \leq |x_k - x_l| + |\widehat{\varepsilon}_n(x_k) - \varepsilon_n(x_k)|.$$

Therefore, the geodesic distance traveled and weighted distance  $\widehat{d}$  along our constructed path differ by at most

$$\sum_{i=1}^{M-1} |\widehat{\varepsilon}_n(x_{k_i}) - \varepsilon_n(x_{k_{i+1}})| + 4M\delta_n$$

in the first  $M - 1$  steps. By arguing as in the proof of Theorem S4.3 with  $\varepsilon_n(x)$  replaced by  $\widehat{\varepsilon}_n(x)$  and noting that  $\widehat{\varepsilon}_n(x)$  converges uniformly to  $\varepsilon_n(x)$ , the number of steps in the constructed path satisfies

$$\frac{|x_i - x_j|}{\max_x \varepsilon_n(x)} \leq M \leq \frac{|x_i - x_j|}{\min_x \varepsilon_n(x) - 2\delta_n}. \quad (5)$$

In particular, we note that  $M = O(g_n^{-1})$ . Applying Lemma S4.4 to choose  $n$  large enough so that

$$\sum_{i=1}^{M-1} |\widehat{\varepsilon}_n(x_{k_i}) - \varepsilon_n(x_{k_{i+1}})| < \gamma$$

and adding  $\max_{x \in \mathcal{X}_n} \widehat{\varepsilon}_n(x)$  for the last step yields (4). Noting by (5) that  $M\delta_n \rightarrow 0$ , taking large enough  $n$  in (4) shows that  $\widehat{d}_{ij} \leq \widehat{d} \leq |x_i - x_j|$ .

We now show that  $\widehat{d}_{ij} \geq |x_i - x_j|$ . It suffices to show that the length  $L$  of the shortest weighted distance path must be  $L = O(g_n^{-1})$ , as Lemma S4.4 would then imply that its weighted distance with respect to  $\widehat{\varepsilon}_n(x)$  converges to its weighted distance with respect to  $\varepsilon_n(x)$ , which is bounded below by  $|x_i - x_j|$ .

To bound  $L$ , note that the minimum weighted distance at each step is  $\min_{x \in \mathcal{X}_n} \widehat{\varepsilon}_n(x)$ , while the total weighed distance is at most  $\widehat{d}$ . Therefore, by (4), we obtain that for any  $\gamma > 0$  we have

$$L \min_{x \in \mathcal{X}_n} \widehat{\varepsilon}_n(x) \leq |x_i - x_j| + 4M\delta_n + \gamma + \max_{x \in \mathcal{X}_n} \widehat{\varepsilon}_n(x)$$

for large enough  $n$ . By uniform convergence of  $\widehat{\varepsilon}_n(x)$  to  $\varepsilon_n(x)$ , this shows that for any  $\gamma > 0$  we have

$$L \leq \frac{|x_i - x_j| + 4M\delta_n + \gamma + \max_{x \in \mathcal{X}_n} \widehat{\varepsilon}_n(x)}{\min_{x \in \mathcal{X}_n} \varepsilon_n(x)} = O(g_n^{-1})$$

for large enough  $n$ , yielding the desired.  $\square$

## References

- [1] S. N. Ethier and T. G. Kurtz. *Markov processes: characterization and convergence*. John Wiley & Sons, 1986.
- [2] D. Stroock and S. Varadhan. Diffusion processes with boundary conditions. *Communications on Pure and Applied Mathematics*, 24:147–225, 1971.
- [3] W. Whitt. Some useful functions for functional limit theorems. *Math. Oper. Res.*, 5(1):67–85, 1980.