Learning Networks of Stochastic Differential Equations: Supplementary Materials

José Bento
Department of Electrical Engineering
Stanford University
Stanford, CA 94305
jbento@stanford.edu

Morteza Ibrahimi
Department of Electrical Engineering
Stanford University
Stanford, CA 94305
ibrahimi@stanford.edu

Andrea Montanari
Department of Electrical Engineering and Statistics
Stanford University
Stanford, CA 94305
montanari@stanford.edu

In order to prove Proposition 4.1 we first introduce two technical lemmas.

Lemma .5. *For any subset $S \subseteq [p]$ the following decomposition holds,*

$$\widehat{Q}_{S^C,S}\left(\widehat{Q}_{S,S}\right)^{-1} = T_1 + T_2 + T_3 + Q^0_{S^C,S}\left(Q^0_{S,S}\right)^{-1}, \tag{23}$$

where,

$$\begin{aligned} T_1 &= Q^0_{S^C,S}\left(\left(\widehat{Q}_{S,S}\right)^{-1} - \left(Q^0_{S,S}\right)^{-1}\right), && (24)\\ T_2 &= (\widehat{Q}_{S^C,S} - Q^0_{S^C,S})\left(Q^0_{S,S}\right)^{-1}, && (25)\\ T_3 &= (\widehat{Q}_{S^C,S} - Q^0_{S^C,S})\left(\left(\widehat{Q}_{S,S}\right)^{-1} - \left(Q^0_{S,S}\right)^{-1}\right). && (26)\\ & && (27)\end{aligned}$$

In addition, if $\|Q^0_{S^C,S}\left(Q^0_{S,S}\right)^{-1}\|_\infty < 1$ and $\lambda_{\min}(\widehat{Q}_{S,S}) \geq C_{\min}/2 > 0$ the following relations hold,

$$\begin{aligned} \|T_1\|_\infty &\leq \frac{2\sqrt{k}}{C_{\min}}\|\widehat{Q}_{S,S} - Q^0_{S,S}\|_\infty, && (28)\\ \|T_2\|_\infty &\leq \frac{\sqrt{k}}{C_{\min}}\|\widehat{Q}_{S^C,S} - Q^0_{S^C,S}\|_\infty, && (29)\\ \|T_3\|_\infty &\leq \frac{2\sqrt{k}}{C^2_{\min}}\|\widehat{Q}_{S^C,S} - Q^0_{S^C,S}\|_\infty\|\widehat{Q}_{S,S} - Q^0_{S,S}\|_\infty. && (30)\end{aligned}$$

The following lemma taken from the proofs of Proposition 1 in [19] and Proposition 1 in [12] respectively is the crux to guaranteeing correct signed-support reconstruction of A^0_r.

Lemma .6. *If $\widehat{Q}_{S^0,S^0} > 0$, then the dual vector $\hat{z}$ from the KKT conditions of the optimization problem* (8) *satisfies the following inequality,*

$$\|\hat{z}_{(S^0)^C}\|_\infty \leq \|\widehat{Q}_{(S^0)^C,S^0}\left(\widehat{Q}_{S^0,S^0}\right)^{-1}\|_\infty\left(1 + \frac{\|\widehat{G}_{S^0}\|_\infty}{\lambda}\right) + \frac{\|\widehat{G}_{(S^0)^C}\|_\infty}{\lambda}. \tag{31}$$

In addition, if

$$\|\widehat{G}_{S^0}\|_\infty \leq \frac{A_{\min}\lambda_{\min}(\widehat{Q}_{S^0,S^0})}{2k} - \lambda \tag{32}$$

then $\|A^0_r - \hat{A}_r\|_\infty \leq A_{\min}/2$. The same result holds for problem (2).

Proof of Proposition 4.1: To guarantee that our estimated support is at least contained in the true support we need to impose that $\|\hat{z}_{S^C}\|_\infty < 1$. To guarantee that we do not introduce extra elements in estimating the support and also to determine the correct sign of the solution we need to impose that $\|A^0_r - \hat{A}_r\|_\infty \leq A_{\min}/2$. Now notice that since $\lambda_{\min}(Q^0_{S^0,S^0}) = C_{\min}$ the relation $\lambda_{\min}(\widehat{Q}_{S^0,S^0}) \geq C_{\min}/2$ is guaranteed as long as $\|\widehat{Q}_{S^0,S^0} - Q^0_{S^0,S^0}\|_\infty \leq C_{\min}/2$. Using Lemma .5 it is easy to see that the bounds of Proposition 4.1 lead to the conditions of Lemma .6 being verified. Thus, these lead to a correct recovery of the signed structure of A^0_r. $\square$

Lemma .7. *Let $r, j \in [p]$ and let $\rho(\tau)$ represent a $p \times p$ matrix with all rows equal to zero except the r^{th} row which equals the j^{th} row of $(I + \eta A^0)^\tau$ (the τ^{th} power of $I + \eta A^0$). Let $\tilde{R}(j) \in \mathbb{R}^{(n+m+1)\times(n+m+1)}$ be defined as,*

$$\tilde{R} = \begin{pmatrix} 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0\\ \rho(m) & \rho(m-1) & \dots & \rho(1) & \rho(0) & 0 & \dots & 0 & 0\\ \rho(m+1) & \rho(m) & \dots & \rho(2) & \rho(1) & \rho(0) & \dots & 0 & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & 0 & 0\\ \rho(m+n-1) & \rho(m+n-2) & \dots & \rho(n) & \rho(n-1) & \rho(n-2) & \dots & \rho(0) & 0 \end{pmatrix}. \tag{33}$$

Define $R(j) = 1/2(\tilde{R} + \tilde{R}^*)$ and let ν_i denote its i^{th} eigenvalue and assume $\sigma_{\max} \equiv \sigma_{\max}(I + \eta A^0) < 1$. Then,

$$\sum_{i=1}^{p(n+m+1)} \nu_i = 0, \tag{34}$$

$$\max_i |\nu_i| \leq \frac{1}{1-\sigma_{\max}}, \tag{35}$$

$$\sum_{i=1}^{p(n+m+1)} \nu_i^2 \leq \frac{1}{2}\frac{n}{1-\sigma_{\max}}. \tag{36}$$

Proof. First it is immediate to see that $\sum_{i=1}^{p(n+m+1)} \nu_i = Tr(R) = 0$. Let $I_{1\tau}$ represent a $p \times p$ matrix with zeros everywhere and ones in the block-position where $\rho(\tau)$ appears and $I_{2\tau}$ represent a similar matrix but with ones in the block-position where $\rho(\tau)^*$ appears. Then R can be written as,

$$R = \frac{1}{2}\left(\sum_{\tau=0}^{m+n-1} I_{1\tau} \otimes \rho(\tau) + I_{2\tau} \otimes \rho(\tau)^*\right), \tag{37}$$

where $\otimes$ denotes the Kronecker product of matrices. This expression can be used to compute an upper bound on $|\nu_i|$. Namely,

$$\max_i |\nu_i| = \sigma_{\max}(R) \leq \sum_{\tau=0}^{\infty} \sigma_{\max}(I_{1\tau} \otimes \rho(\tau)) \leq \sum_{\tau=0}^{\infty} \sigma_{\max}(I_{1\tau})\sigma_{\max}(\rho(\tau)) \tag{38}$$

$$\leq \sum_{\tau=0}^{\infty} \sigma_{\max}(\rho(\tau)) \leq \sum_{\tau=0}^{\infty} \sigma_{\max}^{\tau} = \frac{1}{1-\sigma_{\max}(\varphi^*)}. \tag{39}$$

For the other bound we do,

$$\sum_{i=1}^{(n+m+1)p} \nu_i^2 = Tr(R^2) \leq \frac{1}{4}\, n\; 2 \sum_{\tau=0}^{\infty} Tr(\rho(\tau)\rho(\tau)^*) \tag{40}$$

$$= \frac{1}{2} n \sum_{\tau=0}^{\infty} \|\rho(\tau)\|_2^2 \tag{41}$$

$$\leq \frac{1}{2} n \sum_{\tau=0}^{\infty} \sigma_{\max}^{2\tau} \leq \frac{1}{2}\frac{n}{1-\sigma_{\max}}, \tag{42}$$

where in the last step we used the fact that $0 \leq \sigma_{\max} < 1$. $\square$

Lemma .8. *Let $j \in [p]$. Define $\rho(\tau) \in \mathbb{R}^{1\times p}$ to be the j^{th} row of $(I + \eta A^0)^\tau$. Let $\Phi_j \in \mathbb{R}^{n\times(n+m)}$ be defined as,*

$$\Phi_j = \begin{pmatrix} \rho(m) & \rho(m-1) & \dots & \rho(1) & \rho(0) & 0 & \dots & 0 \\ \rho(m+1) & \rho(m) & \dots & \rho(2) & \rho(1) & \rho(0) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & 0 \\ \rho(m+n-1) & \rho(m+n-2) & \dots & \rho(n) & \rho(n-1) & \rho(n-2) & \dots & \rho(0) \end{pmatrix}, \tag{43}$$

Let ν_l denote the l^{th} eigenvalue of the matrix $R(i,j) = 1/2(\Phi_j^*\Phi_i + \Phi_i^*\Phi_j) \in \mathbb{R}^{(n+m)\times(n+m)}$ (where $i \in [p]$) and assume $\sigma_{\max} \equiv \sigma_{\max}(I + \eta A^0) < 1$ then,

$$|\nu_l| \leq \frac{1}{(1-\sigma_{\max})^2}, \tag{44}$$

$$\frac{1}{n}\sum_{l=1}^{(n+m)p} \nu_l^2 \leq \frac{2}{(1-\sigma_{\max})^3}\left(1 + \frac{3}{2n}\frac{1}{1-\sigma_{\max}}\right). \tag{45}$$

Proof. The first bound can be proved in a trivial manner. In fact, since for any matrix A and B we have $\sigma_{\max}(A+B) \leq \sigma_{\max}(A) + \sigma_{\max}(B)$ and $\sigma_{\max}(AB) \leq \sigma_{\max}(A)\sigma_{\max}(B)$ we can write

$$\max_l |\nu_l| = \sigma_{\max}(1/2(\Phi_j^*\Phi_i + \Phi_i^*\Phi_j)) \leq 1/2(\sigma_{\max}(\Phi_j^*\Phi_i) + \sigma_{\max}(\Phi_i^*\Phi_j)) \tag{46}$$

$$\leq \sigma_{\max}(\Phi_i^*\Phi_j) \leq \sigma_{\max}(\Phi_i)\sigma_{\max}(\Phi_j) \leq \frac{1}{(1-\sigma_{\max})^2}, \tag{47}$$

where in the last inequality we used the fact $\sigma_{\max}(\Phi_j) \leq 1/(1-\sigma_{\max})$. The proof of this is just a copy of the proof of the bound (35) in Lemma .7.

Before we prove the second bound let us introduce some notation to differentiate $\rho(\tau)$ associated with Φ_j from $\rho(\tau)$ associated with Φ_i. Let us call them $\rho(\tau, j)$ and $\rho(\tau, i)$ respectively. Now notice that $\Phi_i^*\Phi_j$ can be written as a block matrix

$$\begin{pmatrix} \tilde{A} & \tilde{D} \\ \tilde{C} & \tilde{B} \end{pmatrix} \tag{48}$$

where $\tilde{A}, \tilde{B}, \tilde{C}$ and $\tilde{D}$ are matrix blocks where each block is a p by p matrix. $\tilde{A}$ has $p \times p$ blocks, $\tilde{B}$ has $n \times n$ blocks, $\tilde{C}$ has $n \times m$ blocks and $\tilde{D}$ has $m \times n$ blocks. If we index the blocks of each matrix with the indices x, y these can be described in the following way

$$\tilde{A}_{xy} = \sum_{s=1}^{m} \rho(m-x+s, i)^*\rho(m-y+s, j) \tag{49}$$

$$\tilde{B}_{xy} = \sum_{s=0}^{n-x} \rho(s, i)^*\rho(s+x-y, j), x \geq y \tag{50}$$

$$\tilde{B}_{xy} = \sum_{s=0}^{n-y} \rho(s+y-x, i)^*\rho(s, j), x \leq y \tag{51}$$

$$\tilde{C}_{xy} = \sum_{s=0}^{n-x} \rho(s, i)^*\rho(m-y+x+s, j) \tag{52}$$

$$\tilde{D}_{xy} = \sum_{s=0}^{n-y} \rho(m-x+y+s, i)^*\rho(s, j). \tag{53}$$

With this in mind and denoting by A, B, C and D the symmetrized versions of these same matrices (e.g.: $A = 1/2(\tilde{A} + \tilde{A}^*)$) we can write,

$$\sum_{l=1}^{(n+m)p} \nu_l^2 = Tr(R(i,j)^2) = Tr(A^2) + Tr(B^2) + 2Tr(CD). \tag{54}$$

We now compute a bound for each one of the terms. We exemplify in detail the calculation of the first bound only. First write,

$$Tr(A^2) = \sum_{x=1}^{m}\sum_{y=1}^{m} Tr(A_{xy}A_{xy}^*). \tag{55}$$

Now notice that each $Tr(A_{xy}A_{xy}^*)$ is a sum over $\tau_1, \tau_2 \in [p]$ of terms of the type,

$$(\rho(m-x+\tau_1, i)^*\rho(m-y+\tau_1, j) + \rho(m-x+\tau_1, j)^*\rho(m-y+\tau_1, i)) \times \tag{56}$$

$$\times(\rho(m-y+\tau_2, j)^*\rho(m-x+\tau_2, i) + \rho(m-y+\tau_2, i)^*\rho(m-x+\tau_2, j)). \tag{57}$$

The trace of a matrix of this type can be easily upper bounded by

$$(\sigma_{\max})^{m-x+\tau_1+m-y+\tau_1+m-y+\tau_2+m-x+\tau_2} = (\sigma_{\max})^{2(m-x)+2(m-y)+2\tau_1+2\tau_2} \tag{58}$$

which finally leads to

$$Tr(A^2) \leq \frac{1}{(1-\sigma_{\max})^4}. \tag{59}$$

Doing a similar thing to the other terms leads to

$$Tr(B^2) \leq \sum_{x,y}^{n,n} \sum_{\tau_1,\tau_2} \sigma_{\max}^{2\tau_1+2\tau_2+2|x-y|} \leq \frac{2n}{(1-\sigma_{\max})^3} \tag{60}$$

$$Tr(DC) = \sum_{x=1}^{m} \sum_{y=1}^{n} Tr(C_{xy}D_{yx}) \leq \sum_{x,y,\tau_1,\tau_2}^{m,n,n-y,n-y} \sigma_{\max}^{2(m-x)+2y+2\tau_1+2\tau_2} \leq \frac{1}{(1-\sigma_{\max})^4}. \tag{61}$$

Putting all these together leads to the desired bound. □

Proof of Proposition 4.2: We will start by proving that this exact same bound holds when the probability of the event $\{\|\widehat{G}_S\|_\infty > \epsilon\}$ is computed with respect to a trajectory $\{x(t)\}_{t=0}^n$ that is initiated at instant $t = -m$ with the value $w(-m)$. In other words, $x(-m) = w(-m)$. Assume we have done so. Now notice that as $m \to \infty$, X converges in distribution to n consecutive samples from the model (6) when this is initiated from stationary state. Since $\|\widehat{G}_S\|_\infty$ is a continuous function of $X = [x(0), ..., x(n-1)]$, by the Continuous Mapping Theorem, $\|\widehat{G}_S\|_\infty$ converges in distribution to the corresponding random variable in the case when the trajectory $\{x(i)\}_{i=0}^n$ is initiated from stationary state. Since the probability bound does not depend on m we have that this same bound holds for stationary trajectories too.

We now prove our claim. Recall that $\widehat{G}_j = (X_j W_r^*)/(n\eta)$. Since X is a linear function of the independent gaussian random variables W we can write $X_j W_r^* = \eta z^* R(j) z$, where $z \in \mathbb{R}^{p(n+m+1)}$ is a vector of i.i.d. $\mathsf{N}(0,1)$ random variables and $R(j) \in \mathbb{R}^{p(n+m+1)\times p(n+m+1)}$ is the symmetric matrix defined in Lemma .7.

Now apply the standard Bernstein method. First by union bound we have

$$\mathbb{P}\big\{\|\widehat{G}_S\|_\infty > \epsilon\big\} \leq 2|S| \max_{j\in S} \mathbb{P}\big\{z^* R(j) z > n\epsilon\big\}.$$

Next denoting by $\{\nu_i\}_{1\leq i\leq p(n+m+1)}$ the eigenvalues of $R(j)$, we have, for any $\gamma > 0$,

$$\begin{aligned}
\mathbb{P}\big\{z^* R(j) z > n\epsilon\big\} &= \mathbb{P}\Big\{\sum_{i=1}^{p(n+m+1)} \nu_i z_i^2 > n\epsilon\Big\} \\
&\leq e^{-n\gamma\epsilon} \prod_{i=1}^{p(n+m+1)} \mathbb{E}\big\{e^{\gamma\nu_i z_i^2}\big\} \\
&= \exp\left(-n\Big(\gamma\epsilon + \frac{1}{2n}\sum_{i=1}^{(n+m+1)p} \log(1-2\nu_i\gamma)\Big)\right).
\end{aligned}$$

Let $\gamma = \frac{1}{2}(1-\sigma_{\max})\epsilon$. Using the bound obtained for $|\max_i \nu_i|$ in Eq. (35), Lemma .7, $|2\nu_i\gamma| \leq \epsilon$. Now notice that if $|x| < 1/2$ then $\log(1-x) > -x - x^2$. Thus, if we assume $\epsilon < 1/2$ and given that $\sum_{i=1}^{(n+m+1)p} \nu_i = 0$ (see Eq. (34)) we can continue the chain of inequalities,

$$\mathbb{P}(\|\widehat{G}_S\|_\infty > \epsilon) \leq 2|S| \max_j \exp\left(-n(\gamma\epsilon - 2\gamma^2 \frac{1}{n}\sum_{i=1}^{(n+m+1)p} \nu_i^2)\right) \tag{62}$$

$$\leq 2|S| \exp\left(-n(\frac{1}{2}(1-\sigma_{\max})\epsilon^2 - \frac{1}{4}(1-\sigma_{\max})^2\epsilon^2(1-\sigma_{\max})^{-1})\right) \tag{63}$$

$$\leq 2|S| \exp\left(-\frac{n}{4}(1-\sigma_{\max})\epsilon^2\right). \tag{64}$$

where the second inequality is obtained using the bound in Eq. (36). □

Proof of Proposition 4.3: The proof is very similar to that of proposition 4.2. We will first show that the bound

$$\mathbb{P}(|\widehat{Q}_{ij} - \mathbb{E}(\widehat{Q}_{ij})| > \epsilon) \leq 2e^{-\frac{n}{32\eta^2}(1-\sigma_{max})^3\epsilon^2}, \tag{65}$$

holds in the case where the probability measure and expectation are taken with respect to trajectories $\{x(i)\}_{i=0}^{n}$ that started at time instant $t = -m$ with $x(-m) = w(-m)$. Assume we have done so. Now notice that as $m \to \infty$, X converges in distribution to n consecutive samples from the model 6 when this is initiated from stationary state. In addition, as $m \to \infty$, we have from lemma .9 that $\mathbb{E}(\widehat{Q}_{ij}) \to Q^0_{ij}$. Since $\widehat{Q}_{ij}$ is a continuous function of $X = [x(0), ..., x(n-1)]$, a simple application of the Continuous Mapping Theorem plus the fact that the upper bound is continuous in ϵ leads us to conclude that the bound also holds when the system is initiated from stationary state.

To prove our previous statement first recall the definition of $\widehat{Q}$ and notice that we can write,

$$\widehat{Q}_{ij} = \frac{\eta}{n} z^* R(i,j) z, \tag{66}$$

where $z \in \mathbb{R}^{m+n}$ is a vector of i.i.d. $\mathsf{N}(0,1)$ and $R(i,j) \in \mathbb{R}^{(n+m)\times(n+m)}$ is defined has in lemma .8. Letting ν_l denote the l^{th} eigenvalue of the symmetric matrix $R(i,j)$ we can further write,

$$\widehat{Q}_{ij} - \mathbb{E}(\widehat{Q}_{ij}) = \frac{\eta}{n} \sum_{l=1}^{(n+m)p} \nu_l (z_l^2 - 1). \tag{67}$$

By Lemma .8 we know that,

$$|\nu_l| \leq \frac{1}{(1-\sigma_{\max})^2}, \tag{68}$$

$$\frac{1}{n} \sum_{l=1}^{(n+m)p} \nu_l^2 \leq \frac{2}{(1-\sigma_{\max})^3} \left(1 + \frac{3}{2n} \frac{1}{1-\sigma_{\max}}\right) \leq \frac{3}{(1-\sigma_{\max})^3}, \tag{69}$$

where we applied $T > 3/D$ in the last line.

Now we are done since applying Bernstein trick, this time with $\gamma = 1/8\,(1-\sigma_{\max})^3 \epsilon/\eta$, and making again use of the fact that $\log(1-x) > -x - x^2$ for $|x| < 1/2$ we get,

$$\mathbb{P}(\widehat{Q}_{ij} - \mathbb{E}(\widehat{Q}_{ij}) > \epsilon) = \mathbb{P}(\sum_{l=1}^{(n+m)p} \nu_l (z_l^2 - 1) > \epsilon n/\eta) \tag{70}$$

$$\leq e^{-\frac{\gamma \epsilon n}{\eta}} e^{-\gamma \sum_{l=1}^{(n+m)p} \nu_l} + e^{-1/2 \sum_{l=1}^{(m+n)p} \log(1-2\gamma\nu_l)} \tag{71}$$

$$\leq e^{-\frac{\gamma \epsilon n}{\eta} - \gamma \sum_{l=1}^{(n+m)p} \nu_l + \gamma \sum_{l=1}^{(n+m)p} \nu_l + 2\gamma^2 \sum_{l=1}^{(n+m)p} \nu_l^2} \tag{72}$$

$$\leq e^{-\frac{n}{32\eta^2}(1-\sigma_{\max})^3 \epsilon^2}, \tag{73}$$

where had to assume that $\epsilon < 2/D$ in order to apply the bound on $\log(1-x)$. An analogous reasoning leads us to,

$$\mathbb{P}(\widehat{Q}_{ij} - \mathbb{E}(\widehat{Q}_{ij}) < -\epsilon) \leq e^{-\frac{n}{32\eta^2}(1-\sigma_{\max})^3 \epsilon^2} \tag{74}$$

and the results follows.

□

Lemma .9. *As before, assume* $\sigma_{\max} \equiv \sigma_{\max}(I + \eta A^0) < 1$ *and consider that model* (6) *was initiated at time* $-m$ *with* $w(-m)$*, that is,* $x(-m) = w(-m)$ *then*

$$|\mathbb{E}(\widehat{Q}_{ij}) - Q^0_{ij}| \leq \frac{1}{n+m} \frac{\eta}{(1-\sigma_{\max})^2}. \tag{75}$$

Proof. Let $\rho = I + \eta A^0$. Since,

$$Q^0_{ij} = \eta \sum_{l=0}^{\infty} (\rho^l \rho^{*l})_{ij}, \tag{76}$$

and

$$\mathbb{E}(\widehat{Q}_{ij}) = \eta \sum_{l=0}^{n+m-1} \frac{m+n-l}{n+m} (\rho^l \rho^{*l})_{ij}, \tag{77}$$

we can write,

$$Q^0_{ij} - \mathbb{E}(\widehat{Q}_{ij}) = \eta \left(\sum_{l=m+n}^{\infty} (\rho^l {\rho^*}^l)_{ij} + \sum_{l=1}^{n+m-1} \frac{l}{m+n} (\rho^l {\rho^*}^l)_{ij} \right). \quad (78)$$

Using the fact that for any matrix A and B $\max_{ij}(A_{ij}) \leq \sigma_{\max}(A)$, $\sigma_{\max}(AB) \leq \sigma_{\max}(A)\sigma_{\max}(B)$ and $\sigma_{\max}(A+B) \leq \sigma_{\max}(A) + \sigma_{\max}(B)$ and introducing the notation $\zeta = \rho^2$ we can write,

$$|\mathbb{E}(\widehat{Q}_{ij}) - Q^0_{ij}| \leq \eta \left(\frac{\zeta^{n+m}}{1-\zeta} + \frac{\zeta}{n+m} \sum_{l=0}^{m+n-2} \zeta^l \right) = \frac{\eta(\zeta^2 + \zeta^{n+m} - 2\zeta^{m+n+1})}{(m+n)(1-\zeta)^2} \quad (79)$$

$$\leq \frac{\eta}{(m+n)(1-\sigma_{max})^2}, \quad (80)$$

where we used the fact that for $\zeta \in [0,1]$ and $n \in \mathbb{N}$ we have $1-\zeta \geq 1-\sqrt{\zeta}$ and $\zeta^2 + \zeta^n - 2\zeta^{1+n} \leq 1$. □

Proof of Theorem 3.1:

In order to prove Theorem 3.1 we need to compute the probability that the conditions given by Proposition 4.1 hold. From the statement of the theorem we have that the first two conditions ($\alpha, C_{\min} > 0$) of Proposition 4.1 hold. In order to make the first condition on $\widehat{G}$ imply the second condition on $\widehat{G}$ we assume that

$$\frac{\lambda\alpha}{3} \leq \frac{A_{\min}C_{\min}}{4k} - \lambda \quad (81)$$

which is guaranteed to hold if

$$\lambda \leq A_{\min}C_{\min}/8k. \quad (82)$$

We also combine the two last conditions on $\widehat{Q}$ to

$$\|\widehat{Q}_{[p],S^0} - Q^0_{[p],S^0}\|_\infty \leq \frac{\alpha}{12} \frac{C_{\min}}{\sqrt{k}}. \quad (83)$$

Where $[p] = S^0 \cup (S^0)^c$. We then impose that both the probability of the condition on $\widehat{Q}$ failing and the probability of the condition on $\widehat{G}$ failing are upper bounded by $\delta/2$. Using Proposition 4.2 we see that the condition on $\widehat{G}$ fails with probability smaller than $\delta/2$ given that the following is satisfied

$$\lambda^2 = 36\alpha^{-2}(n\eta D)^{-1}\log(4p/\delta). \quad (84)$$

But we also want (82) to be satisfied and so substituting λ from the previous expression in (82) we conclude that n must satisfy

$$n \geq 2304 k^2 {C_{\min}}^{-2} {A_{\min}}^{-2} \alpha^{-2} (D\eta)^{-1} \log(4p/\delta). \quad (85)$$

In addition, the application of the probability bound in Proposition 4.2 requires that

$$\frac{\lambda^2\alpha^2}{9} < 1/4 \quad (86)$$

so we need to impose further that,

$$n \geq 16(D\eta)^{-1}\log(4p/\delta). \quad (87)$$

To use Corollary 4.4 for computing the probability that the condition on $\widehat{Q}$ holds we need,

$$n\eta > 3/D, \quad (88)$$

and

$$\frac{\alpha C_{\min}}{12\sqrt{k}} < 2kD^{-1}. \quad (89)$$

The last expression imposes the following conditions on k,

$$k^{3/2} > 24^{-1}\alpha C_{\min} D. \quad (90)$$

The probability of the condition on $\widehat{Q}$ will be upper bounded by $\delta/2$ if

$$n > 4608\eta^{-1}k^3\alpha^{-2}{C_{\min}}^{-2}D^{-3}\log 4pk/\delta. \tag{91}$$

The restriction (90) on k looks unfortunate but since $k \geq 1$ we can actually show it always holds. Just notice $\alpha < 1$ and that

$$\sigma_{\max}(Q^0_{S^0,S^0}) \leq \sigma_{\max}(Q^0) \leq \frac{\eta}{1-\sigma_{\max}} \Leftrightarrow D \leq \sigma^{-1}_{\max}(Q^0_{S^0,S^0}) \tag{92}$$

therefore $C_{\min}D \leq \sigma_{\min}(Q^0_{S^0,S^0})/\sigma_{\max}(Q^0_{S^0,S^0}) \leq 1$. This last expression also allows us to simplify the four restrictions on n into a single one that dominates them. In fact, since $C_{\min}D \leq 1$ we also have $C^{-2}_{\min}D^{-2} \geq C^{-1}_{\min}D^{-1} \geq 1$ and this allows us to conclude that the only two conditions on n that we actually need to impose are the one at Equations (85), and (91). A little more of algebra shows that these two inequalities are satisfied if

$$n\eta > \frac{10^4k^2(kD^{-2} + A^{-2}_{\min})}{\alpha^2 DC^2_{\min}}\log(4pk/\delta). \tag{93}$$

This conclude the proof of Theorem 3.1.

□

Lemma .10. *Let $\sigma_{\max} \equiv \sigma_{\max}(I + \eta A^0)$ and $\rho_{\min}(A^0) = -\lambda_{\max}((A^0 + (A^0)^*)/2) > 0$ then,*

$$-\lambda_{\min}\left(\frac{A^0 + (A^0)^*}{2}\right) \geq \limsup_{\eta\to 0}\frac{1-\sigma_{\max}}{\eta}, \tag{94}$$

$$\liminf_{\eta\to 0}\frac{1-\sigma_{\max}}{\eta} \geq -\lambda_{\max}\left(\frac{A^0 + (A^0)^*}{2}\right). \tag{95}$$

Proof.

$$\frac{1-\sigma_{\max}}{\eta} = \frac{1-\lambda^{1/2}_{\max}((I+\eta A^0)^*(I+\eta A^0))}{\eta} \tag{96}$$

$$= \frac{1-\lambda^{1/2}_{max}(I+\eta(A^0+(A^0)^*)+\eta^2(A^0)^*A^0)}{\eta} \tag{97}$$

$$= \frac{1-(1+\eta u^*(A^0+(A^0)^*+\eta(A^0)^*A^0)u)^{1/2}}{\eta}, \tag{98}$$

where u is some unit vector that depends on η. Thus, since $\sqrt{1+x} = 1 + x/2 + O(x^2)$,

$$\liminf_{\eta\to 0}\frac{1-\sigma_{\max}}{\eta} = -\limsup_{\eta\to 0} u^*\left(\frac{A^0+(A^0)^*}{2}\right)u \geq -\lambda_{\max}\left(\frac{A^0+(A^0)^*}{2}\right). \tag{99}$$

The other inequality is proved in a similar way. □

Proof of Theorem 2.1:

In order to prove Theorem 2.1 we first state and prove the following lemma,

Lemma .11. *Let G be a simple connected graph of vertex degree bounded above by k. Let $\tilde{A}$ be its adjacency matrix and $A^0 = -hI + \tilde{A}$ with $h > k$ then for this A^0 the system in* (1) *has $Q^0 = -(1/2)(A^0)^{-1}$ and,*

$$|||Q^0_{(S^0)^C,S^0}(Q^0_{S^0,S^0})^{-1}|||_\infty = |||(A^0_{(S^0)^C,(S^0)^C})^{-1}A^0_{(S^0)^C,S^0}|||_\infty \leq k/h. \tag{100}$$

Proof. $\tilde{A}$ is symmetric so A^0 is symmetric. Since $\tilde{A}$ is irreducible and non-negative, Perron-Frobenious theorem tells that $\lambda_{\max}(\tilde{A}) \leq k$ and consequently $\lambda_{\max}(A^0) \leq -h + \lambda_{\max}(\tilde{A}) \leq$

$-h+k$. Thus $h>k$ implies that A^0 is negative definite and using equation (4) we can compute $Q^0 = -(1/2)(A^0)^{-1}$. Now notice that, by the block matrix inverse formula, we have

$$(Q^0_{S^0,S^0})^{-1} = -2C^{-1}, \tag{101}$$

$$Q^0_{(S^0)^C,S^0} = \frac{1}{2}((A^0_{(S^0)^C,(S^0)^C})^{-1}A^0_{(S^0)^C,S^0}C), \tag{102}$$

where $C = A^0_{S^0,S^0} - A^0_{S^0,(S^0)^C}(A^0_{(S^0)^C,(S^0)^C})^{-1}A^0_{(S^0)^C,S^0}$ and thus

$$\|Q^0_{(S^0)^C,S^0}(Q^0_{S^0,S^0})^{-1}\|_\infty = \|(A^0_{(S^0)^C,(S^0)^C})^{-1}A^0_{(S^0)^C,S^0}\|_\infty. \tag{103}$$

Recall the definition of $\|B\|_\infty$,

$$\|B\|_\infty = \max_i \sum_j |B_{ij}|. \tag{104}$$

Let $z = h^{-1}$ and write,

$$(A^0_{(S^0)^C,(S^0)^C})^{-1} = -z(I - z\tilde{A}_{(S^0)^C,(S^0)^C})^{-1} = -z\sum_{n=0}^{\infty}(z\tilde{A}_{(S^0)^C,(S^0)^C})^n, \tag{105}$$

$$A^0_{(S^0)^C,S^0} = z^{-1}z\tilde{A}_{(S^0)^C,S^0}. \tag{106}$$

This allows us to conclude that $\|(A^0_{(S^0)^C,(S^0)^C})^{-1}A^0_{(S^0)^C,S^0}\|_\infty$ is in fact the maximum over all path generating functions of paths starting from a node $i \in (S^0)^C$ and hitting S^0 for a first time. Let Ω_i denote this set of paths, ω a general path in G and $|\omega|$ its length. Let $k_1, ..., k_{|\omega|}$ denote the degree of each vertex visited by ω and note that $k_m \le k, \forall m$. Then each of these path generating functions can be written in the following form,

$$\sum_{\omega\in\Omega_i} z^{|\omega|} \le \sum_{\omega\in\Omega_i} \frac{1}{k_1 ... k_{|\omega|}}(kz)^{|\omega|} = \mathbb{E}_G((kz)^{T_{i,S^0}}), \tag{107}$$

where T_{i,S^0} is the first hitting time of the set S^0 by a random walk that starts at node $i \in {S^0}^C$ and moves with equal probability to each neighboring node. But $T_{i,S^0} \ge 1$ and $kz < 1$ so the previous expression is upper bounded by kz. □

Now what remains to complete the proof of Theorem 2.1 is to compute the quantities α, $A_{\min}$, $\rho_{\min}(A^0)$ and $C_{\min}$ in Theorem 1.1 . From Lemma .11 we know that $\alpha = 1 - k/(k+m)$. Clearly, $A_{\min} = 1$. We also have that $\rho_{\min}(A^0) = \sigma_{\min}(A^0) \ge k + m - \sigma_{\max}(\tilde{A}) \ge m + k - k = m$. Finally,

$$\lambda_{\min}(Q^0_{S^0,S^0}) = \frac{1}{2}\lambda_{\min}(-(A^0)^{-1}) = \frac{1}{2}\frac{1}{\lambda_{\max}(-A^0)} \ge \frac{1}{2}\frac{1}{m+k+k} \ge \frac{1}{4(m+k)} \tag{108}$$

where in the last step we made use of the fact that $m+k > k$. Substituting these values in the inequality from Theorem 1.1 gives the desired result.

□