Supplementary Material for "*On Achievable Rates of Line Networks with Generalized Batched Network Coding*"

Appendix A
Proofs about Converse

Proof of Lemma 3: Denote by $\mathbf{y}^* = (y^* \cdots y^*)$. We have

$$W(\mathbf{y}|\mathbf{x}) = \begin{cases} \frac{Q^{\otimes N}(\mathbf{y}^*|\mathbf{x}) - p_0}{1-p_0} & \mathbf{y} = \mathbf{y}^*, \\ \frac{Q^{\otimes N}(\mathbf{y}|\mathbf{x})}{1-p_0} & \text{otherwise.} \end{cases} \tag{51}$$

Let $P(\mathbf{y}) = \sum_{\mathbf{x}} Q^{\otimes N}(\mathbf{y}|\mathbf{x}) p(\mathbf{x})$ and $P'(\mathbf{y}) = \sum_{\mathbf{x}} W(\mathbf{y}|\mathbf{x}) p(\mathbf{x})$. We have

$$P'(\mathbf{y}) = \begin{cases} \frac{1}{1-p_0}(P(\mathbf{y}) - p_0) & \mathbf{y} = \mathbf{y}^*, \\ \frac{1}{1-p_0} P(\mathbf{y}) & \text{otherwise.} \end{cases} \tag{52}$$

Substituting (51) and (52) into $I(p, W)$, we get

$$I(p, W) = \sum_{\mathbf{x}} p(\mathbf{x}) \sum_{\mathbf{y}} W(\mathbf{y}|\mathbf{x}) \log \frac{W(\mathbf{y}|\mathbf{x})}{P'(\mathbf{y})} \tag{53}$$

$$= \frac{1}{1-p_0} I(p, Q^{\otimes N}) + \frac{1}{1-p_0} U(\mathbf{y}^*), \tag{54}$$

where

$$U(\mathbf{y}^*) \triangleq \sum_{\mathbf{x}} p(\mathbf{x}) \left((Q^{\otimes N}(\mathbf{y}^*|\mathbf{x}) - p_0) \log \frac{Q^{\otimes N}(\mathbf{y}^*|\mathbf{x}) - p_0}{P(\mathbf{y}^*) - p_0} - Q^{\otimes N}(\mathbf{y}^*|\mathbf{x}) \log \frac{Q^{\otimes N}(\mathbf{y}^*|\mathbf{x})}{P(\mathbf{y}^*)} \right). \tag{55}$$

Using $P(\mathbf{y}^*) = \sum_{\mathbf{x}} Q^{\otimes N}(\mathbf{y}^*|\mathbf{x}) p(\mathbf{x}) \geq \sum_{\mathbf{x}} \epsilon^N p(\mathbf{x}) = \epsilon^N$, we have

$$U(\mathbf{y}^*) = -p_0 \sum_{\mathbf{x}} p(\mathbf{x}) \log(Q^{\otimes N}(\mathbf{y}^*|\mathbf{x}) - p_0) + P(\mathbf{y}^*) \log \frac{P(\mathbf{y}^*)}{P(\mathbf{y}^*) - p_0} \tag{56}$$

$$+ p_0 \log(P(\mathbf{y}^*) - p_0) + \sum_{\mathbf{x}} p(\mathbf{x}) Q^{\otimes N}(\mathbf{y}^*|\mathbf{x}) \log \frac{Q^{\otimes N}(\mathbf{y}^*|\mathbf{x}) - p_0}{Q^{\otimes N}(\mathbf{y}^*|\mathbf{x})} \tag{57}$$

$$\leq -p_0 \log(\epsilon^N - p_0) + q^* \log \frac{\epsilon^N}{\epsilon^N - p_0} + p_0 \log(q^* - p_0) + q^* \log \frac{q^* - p_0}{q^*} \tag{58}$$

$$= (q^* + p_0) \log \frac{q^* - p_0}{\epsilon^N - p_0} + q^* \log \frac{\epsilon^N}{q^*} \tag{59}$$

The proof is completed by combining (54) and (59). ■

Proof of Lemma 5: We relax N to a real number and solve $\frac{\mathrm{d}F(N)}{\mathrm{d}N} = 0$, i.e.,

$$1 - \epsilon^N + LN\epsilon^N \ln \epsilon = 0, \tag{60}$$

or

$$\epsilon^{-N} - 1 + LN \ln \epsilon = 0. \tag{61}$$

Let $t = -N \ln \epsilon$, and denote by $t^*(L)$ the solution of $g(t) \triangleq e^t - 1 - Lt = 0, t > 0$. Then the solution of (60) is $N^* = t^*(L)/\ln(1/\epsilon)$.

We know that $g(t) < 0$ for $0 < t < t^*(L)$; and $g(t) > 0$ for $t > t^*(L)$. Since $g(\ln L) = L - 1 - L\ln L < 0$ and $g(2\ln L) = L^2 - 1 - 2L\ln L > 0$ when $L > 1$, we have $\ln l < t^*(L) < 2\ln L$ when $L > 1$. Last, using $\epsilon^{N^*} = e^{-t^*(L)}$,

$$0.25 \leq (1 - 1/L)^L \leq (1 - \epsilon^{N^*})^L \leq \left(1 - 1/L^2\right)^L < 1, \tag{62}$$

and hence $F(N^*) = \frac{(1-\epsilon^{N^*})^L}{N^*} = \frac{\ln\frac{1}{\epsilon}(1-\epsilon^{N^*})^L}{t^*(L)} = \Theta(\frac{\ln\frac{1}{\epsilon}}{\ln L})$. ■

Proof of Lemma 6: We group the elements of $\mathcal{S}_{\mathrm{i}}$ into $\lceil |\mathcal{S}_{\mathrm{i}}|/2 \rceil$ pairs, denoted collectively as $\mathcal{S}_{\mathrm{i}}^{(2)}$, where each element of $\mathcal{S}_{\mathrm{i}}$ appears in exactly one pair. When $|\mathcal{S}_{\mathrm{i}}|$ is even, all pairs have distinct entries. When $|\mathcal{S}_{\mathrm{i}}|$ is odd, exactly one pair has the two entries same and the other pairs have distinct entries.

For each pair $(x, x') \in \mathcal{S}_{\mathrm{i}}^{(2)}$, fix $y_{x,x'}$ such that $Q(y_{x,x'}|x) \geq \varepsilon_Q$ and $Q(y_{x,x'}|x') \geq \varepsilon_Q$. Define $\mathcal{Z}$ as the collection of $z = (z_x, x \in \mathcal{Q}_{\mathrm{i}})$ such that $z_x = y_{x,x'}$ and $z_{x'} = y_{x,x'}$ for all pairs $(x, x') \in \mathcal{S}_{\mathrm{i}}^{(2)}$. Let $\mathcal{S}_{\mathrm{o}} = \{y_{x,x'} : (x, x') \in \mathcal{S}_{\mathrm{i}}^{(2)}\}$. Therefore, $|\mathcal{S}_{\mathrm{o}}| \leq \lceil |\mathcal{S}_{\mathrm{i}}|/2 \rceil$. Hence for any $x \in \mathcal{S}_{\mathrm{i}}$ and $z \in \mathcal{Z}$, $\alpha(x, z) = z_x \in \mathcal{S}_{\mathrm{o}}$. When $\mathcal{A}$ is even,

$$\begin{aligned} P(Z \in \mathcal{Z}) &= \prod_{(x,x') \in \mathcal{S}_{\mathrm{i}}^{(2)}} P(Z[x] = y_{x,x'}) P(Z_{x'} = y_{x,x'}) && (63) \\ &= \prod_{(x,x') \in \mathcal{S}_{\mathrm{i}}^{(2)}} Q(y_{x,x'}|x) Q(y_{x,x'}|x') \geq \prod_{(x,x') \in \mathcal{S}_{\mathrm{i}}^{(2)}} \varepsilon_Q^2 = \varepsilon_Q^{|\mathcal{S}_{\mathrm{i}}|}. && (64) \end{aligned}$$

When $\mathcal{A}$ is odd,

$$\begin{aligned} P(Z \in \mathcal{Z}) &= \prod_{(x,x') \in \mathcal{S}_{\mathrm{i}}^{(2)} : x \neq x'} P(Z[x] = y_{x,x'}) P(Z_{x'} = y_{x,x'}) \prod_{(x,x) \in \mathcal{S}_{\mathrm{i}}^{(2)}} P(Z[x] = y_{x,x}) && (65) \\ &= \prod_{(x,x') \in \mathcal{S}_{\mathrm{i}}^{(2)} : x \neq x'} Q(y_{x,x'}|x) Q(y_{x,x'}|x') \prod_{(x,x) \in \mathcal{S}_{\mathrm{i}}^{(2)}} Q(y_{x,x}|x) \geq \varepsilon_Q^{|\mathcal{S}_{\mathrm{i}}|}. && (66) \end{aligned}$$

■

Proof of Theorem 7: Consider a line network of length L of general DMCs Q_ℓ with $\varepsilon_{Q_\ell} \geq \epsilon > 0$ and a GBNC as described in §II. Without loss of optimality, we assume a deterministic recoding scheme, i.e., ϕ_ℓ are deterministic. Channel $Q_\ell^{\otimes N}$ can be modelled by the function α_ℓ^N with the channel status variable $Z_\ell = (Z_\ell[\mathbf{x}], \mathbf{x} \in \mathcal{Q}_\mathrm{i}^N)$ so that

$$\mathbf{Y}_\ell = \alpha_\ell^N(\mathbf{U}_\ell, Z_\ell). \tag{67}$$

As $\varepsilon_{Q_\ell^{\otimes N}} \geq \varepsilon_{Q_\ell}^N > 0$, the condition of applying Lemma 6 on $Q_\ell^{\otimes N}$ is satisfied.

Let $\mathcal{S}_\mathrm{i}^{(1)} = \mathcal{Q}_\mathrm{i}^N$. Applying Lemma 6 on $Q_1^{\otimes N}$ w.r.t. $\mathcal{S}_\mathrm{i}^{(1)}$, there exists subsets $\mathcal{Z}^{(1)}$ and $\mathcal{S}_\mathrm{o}^{(1)} \subseteq \mathcal{Q}_\mathrm{o}^N$ with $|\mathcal{S}_\mathrm{o}^{(1)}| \leq \lceil |\mathcal{S}_\mathrm{i}^{(1)}|/2 \rceil$ such that $\alpha_1^N(\mathbf{x}, z_1) \in \mathcal{S}_\mathrm{o}^{(1)}$ for any $\mathbf{x} \in \mathcal{S}_\mathrm{i}^{(1)}$ and $z_1 \in \mathcal{Z}^{(1)}$, and $P(Z_1 \in \mathcal{Z}^{(1)}) \geq \varepsilon^{N|\mathcal{Q}_\mathrm{i}|^N}$. Fix an integer $K = \lceil N \log |\mathcal{Q}_\mathrm{i}| \rceil$. For $i = 2, 3, \ldots, K$, define recursively $\mathcal{S}_\mathrm{i}^{(i)}$, $\mathcal{S}_\mathrm{o}^{(i)}$ and $\mathcal{Z}^{(i)}$ as follows: $\mathcal{S}_\mathrm{i}^{(i)} = \left\{ \mathbf{x} \in \mathcal{Q}_\mathrm{i}^N : \mathbf{x} = \phi_{i-1}(\mathbf{y}) \text{ for certain } \mathbf{y} \in \mathcal{S}_\mathrm{o}^{(i-1)} \right\}$, and $\mathcal{S}_\mathrm{o}^{(i)}$ and $\mathcal{Z}^{(i)}$ are determined as in the proof of Lemma 6 w.r.t. $Q_i^{\otimes N}$ and $\mathcal{S}_\mathrm{i}^{(i)}$ so that $\alpha_i^{\otimes N}(\mathbf{x}, z) \in \mathcal{S}_\mathrm{o}^{(i)}$ for any $\mathbf{x} \in \mathcal{S}_\mathrm{i}^{(i)}$ and $z \in \mathcal{Z}^{(i)}$, and $P(Z_i \in \mathcal{Z}^{(i)}) \geq \varepsilon^{N|\mathcal{S}_\mathrm{i}^{(i)}|}$.

According to the construction, $|\mathcal{S}_\mathrm{i}^{(i)}| \leq |\mathcal{S}_\mathrm{o}^{(i-1)}|$ and $|\mathcal{S}_\mathrm{o}^{(i)}| \leq \lceil |\mathcal{S}_\mathrm{i}^{(i)}|/2 \rceil$. Hence $|\mathcal{S}_\mathrm{o}^{(K)}| \leq \lceil |\mathcal{S}_\mathrm{i}^{(1)}|/2^K \rceil = 1$. Since the set $\mathcal{S}_\mathrm{o}^{(K)}$ is non-empty, we have $|\mathcal{S}_\mathrm{o}^{(K)}| = 1$, i.e., there exists an output of $Q_K^{\otimes N}$ that occurs with a positive probability for all inputs of $Q_1^{\otimes N}$. Define the channel $G_1 = Q_1^{\otimes N} \phi_1 Q_2^{\otimes N} \cdots \phi_{K-1} Q_K^{\otimes N}$. Under the condition $Z_i \in \mathcal{Z}^{(i)}, i = 1, \ldots, K$, the output of G_1 must be unique for all possible channel inputs, i.e., G_1 is canonical. Note that

$$P(Z_i \in \mathcal{Z}^{(i)}, i = 1, \ldots, K) \geq \varepsilon^{N \sum_{i=1}^K |\mathcal{A}_i|} \geq \varepsilon^{N(2|\mathcal{Q}_\mathrm{i}|^N + K)}. \tag{68}$$

Let $L' = \lfloor L/K \rfloor$. For $i = 2, \ldots, L'$, define $G_i = Q_{K(i-1)+1}^{\otimes N} \phi_{K(i-1)+1} Q_{K(i-1)+2}^{\otimes N} \cdots \phi_{Ki-1} Q_{Ki}^{\otimes N}$. Similar as G_1, we know that G_i, $i = 2, \ldots, L'$ are all canonical. We see that G_i, $i = 1, \ldots, L'$ forms a length-L' network. Let $\tilde{W}_{L'} = \phi_0 G_1 \phi_K G_2 \phi_{2K} \cdots G_{L'}$, which is the end-to-end transition matrix of a GBNC with inner block length 1 for the length-L' network of canonical channels G_i. By the data processing inequality, $I(p_\mathbf{X}, W_L) \leq I(p_\mathbf{X}, \tilde{W}_{L'})$. Based on this relation, we are ready to prove the theorem, similar to that of Theorem 4. ∎

Appendix B

Proofs about Achievability

Proof of Lemma 9: Suppose that the node $\ell - 1$ transmits $u_\ell(x)$ for N times, where $x \in \mathcal{A}$. We know that the entries of $\mathbf{y}_\ell$ are i.i.d. random variables with distribution $Q_\ell(\cdot \mid u_\ell(x))$. The error probability for ML decoding at the node ℓ satisfies

$$\epsilon_\ell(x) \le P\left(\vee_{\overline{x}\neq x}\mathcal{L}_\ell(\overline{x};\mathbf{y}_\ell) \ge \mathcal{L}_\ell(x;\mathbf{y}_\ell)\right\} \tag{69}$$

$$\le \sum_{\overline{x}\in\mathcal{A}:\ \overline{x}\neq x} P\left(\mathcal{L}_\ell(\overline{x};\mathbf{y}_\ell) \ge \mathcal{L}_\ell(x;\mathbf{y}_\ell)\right), \tag{70}$$

where the second inequality follows from the union bound. For fixed $\overline{x}\in\mathcal{A}$ so that $\overline{x}\neq x$, we bound the probability $P\left(\mathcal{L}_\ell(\overline{x};\mathbf{Y}_\ell) \ge \mathcal{L}_\ell(x;\mathbf{Y}_\ell)\right)$ by considering two cases.

If there exists a non-empty subset $\mathcal{Y}_0 \subseteq \mathcal{Q}_\mathrm{o}$ so that for any $y_0 \in \mathcal{Y}_0$, $Q_\ell(y_0 \mid u_\ell(x)) > 0$ but $Q_\ell(y_0 \mid u_\ell(\overline{x})) = 0$, as long as $\mathbf{y}_\ell[i]\in\mathcal{Y}_0$ for some i, we can assert that $\mathcal{L}_\ell(\overline{x};\mathbf{y}_\ell) < \mathcal{L}_\ell(x;\mathbf{y}_\ell)$. Therefore,

$$P\left(\mathcal{L}_\ell(\overline{x};\mathbf{y}_\ell) \ge \mathcal{L}_\ell(x;\mathbf{y}_\ell)\right) \le P\left(\mathbf{Y}_\ell[i]\notin\mathcal{Y}_0, i=1,\dots,N\right) \tag{71}$$

$$= \left[\sum_{y\notin\mathcal{Y}_0} Q_\ell(y\mid u_\ell(x))\right]^N = \exp\left(-N\log\frac{1}{\sum_{y\notin\mathcal{Y}_0} Q_\ell(y\mid u_\ell(x))}\right), \tag{72}$$

where $\sum_{y\notin\mathcal{Y}_0} Q_\ell(y\mid u_\ell(x)) = 1-\sum_{y\in\mathcal{Y}_0} Q_\ell(y\mid u_\ell(x)) < 1$.

Otherwise, consider that the support of $Q_\ell(\cdot\mid u_\ell(x))$ belongs to the support of $Q_\ell(\cdot\mid u_\ell(\overline{x}))$. For $i=1,\dots,N$, define the random variable $D_i = \log\frac{Q_\ell(\mathbf{Y}_\ell[i]|u_\ell(\overline{x}))}{Q_\ell(\mathbf{Y}_\ell[i]|u_\ell(x))}$. We see that D_i are i.i.d., and satisfy

$$\log\varrho_\ell \le D_i \le -\log\varrho_\ell, \tag{73}$$

where $\varrho_\ell = \min_{x\in\mathcal{Q}_\mathrm{i}, y\in\mathcal{Q}_\mathrm{o}: Q_\ell(y|x)>0} Q_\ell(y|x)$, and

$$\mathbb{E}[D_i] = E'_\ell \triangleq -\mathcal{D}_\mathrm{KL}\left(Q_\ell(\cdot\mid u_\ell(x))\|Q_\ell(\cdot\mid u_\ell(\overline{x}))\right), \tag{74}$$

where $\mathcal{D}_\mathrm{KL}$ denotes the Kullback-Leibler divergence. We see that $E'_\ell > -\infty$. Moreover, as $u_\ell(x)\neq u_\ell(\bar{x})\in\mathcal{Q}_\mathrm{i}^\ell$, $Q_\ell(\cdot\mid u_\ell(x)) \neq Q_\ell(\cdot\mid u_\ell(\overline{x}))$ and hence $E'_\ell \neq 0$. Applying Hoeffding's inequality, we obtain

$$P\left(\mathcal{L}_\ell(\overline{x};\mathbf{y}_\ell) \ge \mathcal{L}_\ell(x;\mathbf{y}_\ell)\right) = P\left(\sum_{i=1}^N D_i \ge 0\right) \tag{75}$$

$$= P\left(\sum_{i=1}^N \left(D_i - E'_\ell\right) \ge -NE'_\ell\right) \tag{76}$$

$$\le \exp\left(-\frac{NE'^2_\ell}{2\log^2\varrho_\ell}\right). \tag{77}$$

The proof is completed by combining both cases. ■

Proof of Lemma 11: Suppose Q has size $m \times n$. As $C(Q) > \epsilon > 0$, $m \geq 2$. Let $\mathbf{a} = (a_1, \ldots, a_n)$ be a row of Q, and construct a new $m \times n$ stochastic matrix $\tilde{Q}$ with all the rows $\mathbf{a}$. We have $C(\tilde{Q}) = 0$ and hence $|C(Q) - C(\tilde{Q})| > \epsilon$. Since channel capacity as a function of stochastic matrices is uniformly continuous [10, Lemma I.1], there exists a constant $\delta > 0$ depending on ϵ such that $\|\tilde{Q} - Q\|_\infty > \delta$. As a consequence, there exists another row $\mathbf{a}' = (a'_1, \ldots, a'_n)$ of Q such that $\|\mathbf{a} - \mathbf{a}'\|_\infty > \delta$. Denote by j the index such that $|a_j - a'_j| > \delta$.

Using the example of uniform reduction with $s = 2$, we can choose R so that RQ is formed by $\mathbf{a}$ and $\mathbf{a}'$. Then we can find W so that $RQW = U_2(\rho_1)$, where

$$\rho_1 = \sum_{k: a_k + a'_k > 0} \frac{a_k^2}{a_k + a'_k} = 1 - \sum_{k: a_k + a'_k > 0} \frac{a_k a'_k}{a_k + a'_k}. \tag{78}$$

Based on the relation that

$$\frac{1}{2} - \sum_{k: a_k + a'_k > 0} \frac{a_k a'_k}{a_k + a'_k} = \frac{1}{4} \sum_{k: a_k + a'_k > 0} \frac{(a_k - a'_k)^2}{a_k + a'_k} \geq \frac{1}{4} \frac{(a_j - a'_j)^2}{a_j + a'_j} \geq \frac{\delta^2}{8}, \tag{79}$$

we have the lower bound $\rho_1 \geq B$ with $B = \frac{1}{2} + \frac{\delta^2}{8} > 1/2$. For any ϱ such that $1/2 < \varrho \leq B$, we have $U_2(\varrho) = U_2(\rho_1) U_2(\frac{\rho_1 + \varrho - 1}{2\rho_1 - 1})$, and hence $RQWU_2(\frac{\rho_1 + \varrho - 1}{2\rho_1 - 1}) = U_2(\varrho)$. ■

Proof of Lemma 13: As $\operatorname{rank}(Q) = r \geq s$, we can find stochastic matrices R and W such that $\min \operatorname{inv}(RQW) = \kappa_s(Q)$. Let $B = (RQW)^{-1}$, and $K = BU_s(\varrho)$. As $RQWK = U_s(\varrho)$, we only need to show that for $1/s < \varrho \leq \rho_s(Q)$, K is a stochastic matrix. Let $\mathbf{1}$ be the all-one vector of certain length. We see that $K\mathbf{1} = BU_s(\rho)\mathbf{1} = B\mathbf{1} = \mathbf{1}$, where the last equality follows because $RQW\mathbf{1} = \mathbf{1}$ and RQW is invertible.

It remains to show that all the entries of K are nonnegative. Let b_{ij} be the (i, j) entry of B. The (i, j) entry of K is $k_{ij} = \frac{1}{s-1}\left[(1 - \varrho) + b_{ij}(s\varrho - 1)\right] \geq \frac{1}{s-1}\left[(1 - \varrho) + \kappa_s(Q)(s\varrho - 1)\right]$. When $\kappa_s(Q) \geq 0$, we have $k_{ij} \geq 0$ for any $\varrho \in (1/s, 1]$. When $\kappa_s(Q) < 0$, we have $k_{ij} \geq 0$ for any $\varrho \in (1/s, \frac{\kappa_s(Q) - 1}{s\kappa_s(Q) - 1}]$. ■

Proof of Theorem 14: Recall the Markov chain relation in (45), where the transition matrix $\mathbf{P}$ is an $(M + 1) \times (M + 1)$ matrix with the (i, j) entry ($0 \leq i, j \leq M$):

$$p_{i,j} = \begin{cases} 0 & i < j, \\ \sum_{k=j}^{N} f(k; N, \epsilon) \zeta_j^{i,k} & i \geq j, \end{cases} \tag{80}$$

where $f(k; N, \epsilon) = \binom{N}{k}(1 - \epsilon)^k \epsilon^{N-k}$ is the probability mass function (PMF) of the binomial distribution with parameters N and $1 - \epsilon$, and $\zeta_j^{i,k}$ is the probability that the $i \times k$ matrix with independent entries uniformly distributed over the field $\mathbb{F}_q$ has rank j. We know that (ref. [27, (2.4)]) $\zeta_j^{i,k} = \frac{\zeta_j^i \zeta_j^k}{\zeta_j^j q^{(i-j)(k-j)}}$, where

$$\zeta_r^m = \begin{cases} 1 & r = 0, \\ (1-q^{-m})(1-q^{-m+1})\cdots(1-q^{-m+r-1}) & 1 \le r \le m. \end{cases} \tag{81}$$

As shown in [29], the matrix $\mathbf{P}$ admits the eigendecomposition $\mathbf{P} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$, where $\mathbf{V} = (v_{i,j})_{0 \le i,j \le M}$ and $\mathbf{\Lambda} = \mathrm{diag}(\lambda_0, \lambda_1, \ldots, \lambda_M)$. Here $\lambda_j = \sum_{k=j}^{N} f(k; N, \epsilon)\zeta_j^k$, $v_{i,j} = \zeta_j^i$ for $i \ge j$ and otherwise $v_{i,j} = 0$. It can be checked that $\lambda_0 > \lambda_1 > \cdots > \lambda_M$. Denote the (i,j) entry $0 \le i,j \le M$ of V^{-1} by $u_{i,j}$. We know that $u_{i,j} = 0$ for $i < j$ and $u_{i,i} = 1/\zeta_i^i$. Based on the formulation above, we have

$$\mathbf{E}[\pi_L] = \pi_0 \mathbf{V}\mathbf{\Lambda}^L\mathbf{V}^{-1} \begin{bmatrix} 0 & 1 & \cdots & M \end{bmatrix}^\top = \sum_{i=1}^{M} \lambda_i^L v_{M,i} \sum_{j=1}^{i} j u_{i,j} \tag{82}$$

$$= \lambda_1^L v_{M,1} u_{1,1} \left(1 + \sum_{i=2}^{M} \frac{\lambda_i^L v_{M,i}}{\lambda_1^L v_{M,1} u_{1,1}} \sum_{j=1}^{i} j u_{i,j} \right) \tag{83}$$

$$= \Theta(\lambda_1^L), \tag{84}$$

where (83) follows from the fact that $v_{M,1}u_{1,1} > 0$, and (84) is obtained by noting that

$$\sum_{i=2}^{M} \frac{\lambda_i^L v_{M,i}}{\lambda_1^L v_{M,1} u_{1,1}} \sum_{j=1}^{i} j u_{i,j} = o(1) \tag{85}$$

as $\lambda_i \le \lambda_1$ for $i \ge 2$. By (81), we further have

$$\lambda_1 = \sum_{k=1}^{N} f(k; N, \epsilon)(1 - q^{-k}) = \sum_{k=1}^{N} f(k; N, \epsilon) - \sum_{k=1}^{N} f(k; N, \epsilon) q^{-k} \tag{86}$$

$$= 1 - f(0; N, \epsilon) - \sum_{k=1}^{N} \binom{N}{k} (1-\epsilon)^k \epsilon^{N-k} q^{-k} = 1 - (\epsilon + (1-\epsilon)/q)^N. \tag{87}$$

The proof is completed. ■