APPENDIX A
THE DERIVATION OF PROBABILITY EXPRESSIONS

Derivation of q: Recall the definition of q that the average probability of finding at least one source destination pair in the same cell. Without loss of generality, we make a reasonable assumption which will not deteriorate the performance. We assume that N is divisible by $k+1$ with every $k+1$ nodes randomly forming an independent group. Each node i in a specific group acts as a source and the other k nodes are its destinations. Hence, any two nodes within the same group form an S-D pair, and nodes not belonging to this group can only act as relays. This assumption holds in the rest of the proof. Due to the independent location distribution of each node, the probability that no source-destination pair is found in cell c is $(1-\pi_c)^{k+1} + \binom{k+1}{1}\pi_c(1-\pi_c)^k$. Since there are $\frac{N}{k+1}$ groups, the probability of finding at least one S-D pair in the same cell c is $(1-[(1-\pi_c)^{k+1} + \binom{k+1}{1}\pi_c(1-\pi_c)^k]^{\frac{N}{k+1}})$. With this, we obtain the expression of q.

Derivation of p: Recall the definition of p that the average probability of finding at least two nodes in the same cell. Its opposite event is that no node or only one node is found in the cell. Consider a cell c, the probability that no node in it is $(1-\pi_c)^N$, and the probability that only one node in it is $\binom{N}{1}\pi_c(1-\pi_c)^{N-1}$. Thus, the probability that at least two nodes in cell c is $1-(1-\pi_c)^N - \binom{N}{1}\pi_c(1-\pi_c)^{N-1}$. With this, we obtain the expression of p.

Derivation of q': The probability of finding only one user in cell c is given by $\binom{N}{1}\pi_c(1-\pi_c)^{N-1}$. The probability of finding its destination in an adjacent cell of cell c is k times of $\Pi_{adj}(c)$. Using this, we obtain the expression of q'.

Derivation of p': Given that there is exactly 1 user in cell c, the probability that at least one of the other $N-1$ users is in an adjacent cell is given by $1-(1-\Pi_{adj}(c))^{ki}$. Thus, we obtain the expression of p'.

Derivation of q'': We first compute the probability that there are i users in cell c such that there are no S-D pairs. Clearly, $1 \le i \le \frac{N}{i+1}$, otherwise there must be at least one source destination pair. Note that $\binom{N}{i}\pi_c^i(1-\pi_c)^{N-i}$ is the probability of finding i users in cell c. Given that there are i users in a cell, the probability that no S-D pair in the same cell is $\frac{(k+1)^i\binom{\frac{N}{k+1}}{i}}{\binom{N}{i}}$. Furthermore, the probability that there is at least 1 node in an adjacent cell that will make an S-D pair with one of these i nodes given that it is not in cell c is $1-(1-\Pi_{adj}(c))^{ki}$. Combining all the arguments, we obtain the expression of q''.

Derivation of p'': The derivation of p'' is similar to that of p''. The probability of finding i users in cell c such that no S-D pair is found in cell c as well as any adjacent cells is $\frac{(k+1)^i\binom{\frac{N}{k+1}}{i}}{\binom{N}{i}}$. To ensure at least 2 users in cell c, we sum over i from 2 to $\frac{N}{2}$. Combining all these, we obtain the expression of p''.

APPENDIX B
PROOF OF THEOREM 1

Let Ψ denote all the scheduling schemes that can make the network stable. Consider a policy $\psi \in \Psi$. Let $X_{ab}^{\psi}(T)$ be the number of packets successfully transmitted from sources to all the k destinations by exactly "a" same cell transmissions and "b" adjacent cell transmissions during the time interval $(0,T)$. To make the network stable, there should be an arbitrarily large value T which make the total output rate within ϵ ($\epsilon > 0$ is a constant) of the total input rate:

$$\frac{\sum_{a=0}^{\infty}\sum_{b=0}^{\infty} X_{ab}^{\psi}(T)}{T} \ge N\lambda - \epsilon. \qquad (1)$$

Let $Y^{\psi}(T)$ be the total number of packet transmissions in interval $(0,T)$ under policy ψ. Since during $(0,T)$, the number of packet transmission is at least $\sum\sum_{a+b\ge k}(a+b)X_{ab}^{\psi}(T)$, we have

$$\begin{aligned}\frac{1}{T}Y^{\psi}(T) \ge& \frac{1}{T}\sum_{a=0}^{\infty}\sum_{b=0}^{\infty}(a+b)X_{ab}^{\psi}(T)\\ \ge& \frac{k}{T}\sum_{a+b=k}X_{ab}^{\psi}(T) + \frac{k+1}{T}\sum_{a+b\ge k}X_{ab}^{\psi}(T)\\ &- \frac{k+1}{T}\sum_{a+b=k}X_{ab}^{\psi}(T)\\ =&(k+1)(N\lambda-\epsilon) - \frac{1}{T}\sum_{a+b=k}X_{ab}^{\psi}(T),\end{aligned}$$

where the last inequality follows from inequality (1). Since ϵ is an arbitrary positive constant, we have,

$$\lambda \le \lim_{T\to\infty}\frac{Y^{\psi}(T) + \sum_{a+b=k}X_{ab}^{\psi}(T)}{(k+1)TN}. \qquad (2)$$

Assume that same cell direct S-D transmission happens with probability p_1 and the adjacent cell direct S-D transmission happens with probability p_2. Thus we have,

$$\begin{aligned}\sum_{a+b=k}X_{ab}^{\psi}(T) =& \quad kp_1^kX_{10}^{\psi}(T) + [(k-1)p_1^{k-1}X_{10}^{\psi}(T)\\ &+p_2X_{01}^{\psi}(T)] + ... + kp_2^kX_{01}^{\psi}(T)\\ =& \quad \sum_{a=0}^{k}ap_1^aX_{10}^{\psi}(T) + \sum_{b=0}^{k}bp_2^bX_{01}^{\psi}(T)\\ \le& \quad X_{10}^{\psi}(T) + X_{01}^{\psi}(T).\end{aligned} \qquad (3)$$

Recall the definition of $X_{ab}^{\psi}(T)$ at the beginning of the proof. We see that $X_{01}^{\psi}(T)$ is the number of packets transmitted from sources to all the k destinations by no same cell transmission and one adjacent cell transmission. Similarly, $X_{10}^{\psi}(T)$ is the number of packets transmitted from sources to all the k destinations by one same cell transmission and no adjacent cell transmission. Thus using inequalities (2) and (3), we have

$$\lambda \le \lim_{T\to\infty}\frac{Y^{\psi}(T) + X_{10}^{\psi}(T) + X_{01}^{\psi}(T)}{(k+1)TN}. \qquad (4)$$

Let $Y_c^{\psi}(\tau)$ be the total number of packet transmissions in cell c under policy ψ at time slot τ. Similarly, $X_{10,c}^{\psi}(\tau)$ and $X_{01,c}^{\psi}(\tau)$ denote the number of direct S-D packets delivered within cell c and from adjacent cell to cell c respectively at slot τ. Hence,

$$\begin{aligned}&Y^{\psi}(T)+X^{\psi}_{10}(T)+X^{\psi}_{01}(T)\\&=\sum_{\tau=0}^{T-1}\sum_{c=1}^{C}\left(Y^{\psi}_{c}(\tau)+X^{\psi}_{10,c}(\tau)+X^{\psi}_{01,c}(\tau)\right).\end{aligned} \tag{5}$$

Let the following four indicator functions denote the packet transmission events under any specific scheduling scheme ψ at slot τ:

$$I^1_c(\tau)=\begin{cases}1, & \text{if a direct S-D transmission happens}\\ & \text{within the same cell } c \text{ at slot } t,\\ 0, & \text{else.}\end{cases} \tag{6}$$

$$I^2_c(\tau)=\begin{cases}1, & \text{if a relay transmission happens}\\ & \text{within the same cell } c \text{ at slot } \tau,\\ 0, & \text{else.}\end{cases} \tag{7}$$

$$I^3_c(\tau)=\begin{cases}1, & \text{if a direct S-D transmission happens}\\ & \text{between cell } c \text{ and its adjacent cell at slot } \tau,\\ 0, & \text{else.}\end{cases} \tag{8}$$

$$I^4_c(\tau)=\begin{cases}1, & \text{if a relay transmission happens between}\\ & \text{cell } c \text{ and its adjacent cell at slot } \tau,\\ 0, & \text{else.}\end{cases} \tag{9}$$

Let $Z^{\psi}_c(\tau)=Y^{\psi}_c(\tau)+X^{\psi}_{10,c}(\tau)+X^{\psi}_{01,c}(\tau)$, and note that the transmission rates within the same cell and between adjacent cells are R_1 and R_2 respectively. Hence, we have

$$\begin{aligned}Z^{\psi}_c(\tau)=&Y^{\psi}_c(\tau)+X^{\psi}_{10,c}(\tau)+X^{\psi}_{01,c}(\tau)\\=&R_1(I^1_c(\tau)+I^2_c(\tau))+R_2(I^3_c(\tau)+I^4_c(\tau))\\&+R_1I^1_c(\tau)+R_2I^3_c(\tau)\\=&2R_1I^1_c(\tau)+R_1I^2_c(\tau)+2R_2I^3_c(\tau)+R_2I^4_c(\tau).\end{aligned} \tag{10}$$

Note that only one of these four indicator functions is 1 at slot τ. According to the priority order, we schedule the packet transmission events as follows: if $R_1\geq 2R_2$, the optimal order is $I^1_c(\tau)\succ I^2_c(\tau)\succ I^3_c(\tau)\succ I^4_c(\tau)$; if $2R_2>R_1\geq R_2$, the optimal order is $I^1_c(\tau)\succ I^3_c(\tau)\succ I^2_c(\tau)\succ I^4_c(\tau)$. Let $Z_c(\tau)=max_{\psi\in\Psi}Z^{\psi}_c(\tau)$. Following the optimal order and combining Eq. (4), (5) and (10), we have

$$\begin{aligned}\lambda\leq&\lim_{T\leftarrow\infty}\frac{\sum_{\tau=0}^{T-1}\sum_{c=1}^{C}Z_c(\tau)}{(k+1)TN}=\frac{1}{(k+1)N}\sum_{c=1}^{C}\mathbb{E}\{Z_c(\tau)\}\\\leq&\begin{cases}\frac{2R_1q+R_1(p-q)+2R_2q'+R_2(p'-q')}{(k+1)\theta}, & \text{if}\quad R_1\geq 2R_2,\\ \frac{2R_1q+2R_2q''+R_1p''+R_2(p'-q')}{(k+1)\theta}, & \text{if}\quad 2R_2\geq R_1\geq R_2.\end{cases}\end{aligned}$$

Appendix C
Proof of Theorem 2

A packet is called type c if its destination is node c. Let $U^{(c)}_i(t)$ be the number of type c packets queued in the buffer of node i at time t. For all node $i\neq c$, the d-step dynamics of unserved packets satisfy:

$$\begin{aligned}U^{(c)}_i(t+d)\leq&\max\left[U^{(c)}_i(t)-\sum_{\tau=t}^{t+d-1}\sum_b\mu^{(c)}_{ib}(\tau),0\right]\\&+\sum_{\tau=t}^{t+d-1}\sum_a\mu^{(c)}_{ai}(\tau)+\sum_{\tau=t}^{t+d-1}A^{(c)}_i(\tau),\end{aligned} \tag{11}$$

where $A^{(c)}_i(\tau)$ is the number of type c packets arriving at source node i at the beginning of slot τ and $\mu^{(c)}_{ab}(\tau)$ is the serving rate for type c packets from transmitter a to receiver b in slot τ. Note that the above formula is not an equality but an inequality. Because the actual rate of type c packets from other nodes may be smaller than $\sum_{\tau=t}^{t+d-1}\sum_a\mu^{(c)}_{ai}(\tau)$ if they do not have enough packets at this moment.

Now we define the Lyapunov function as $L(\overrightarrow{U}(t))=\sum_{i=1}^{N}\sum_{c\neq i}(U^{(c)}_i(t))^2$. By Eq. (11), we can obtain the following inequality for the d-step Lyapunov drift:

$$\begin{aligned}&\mathbb{E}\{L(\overrightarrow{U}(t+d))-L(\overrightarrow{U}(t))|\overrightarrow{U}(t)\}\\\leq&d^2BN-2d\sum_{i\neq c}U^{(c)}_i(t)\frac{1}{d}\sum_{\tau=t}^{t+d-1}\mathbb{E}\{\sum_b\mu^{(c)}_{ib}(\tau)\sum_a\mu^{(c)}_{ai}(\tau)\\&-A^{(c)}_i(\tau)|\overrightarrow{U}(t)\},\end{aligned} \tag{12}$$

where $B=(A_{max}+\mu^{in}_{max})^2+(\mu^{out}_{max})^2$. And A_{max} is the maximum extraneous arrival rate of any node. μ^{in}_{max} is the maximum transmission rate into any node, which equals R_1+JR_2. μ^{out}_{max} is the maximum transmission rate out of any node, which equals R_1.

Note that the *Two hop relay algorithm* makes decisions only relying on the current distribution of nodes. Besides, the Markovian mobility model has the property of ergodicity and stability. Thus applying *Lemma 1* in [4], we obtain

$$\mathbb{E}\{\sum_b\mu^{(c)}_{ib}(t+d)|\overrightarrow{U}(t)\}\geq(\sum_b\overline{\mu}^{(c)}_{ib})(1-2N\alpha\gamma^d). \tag{13}$$

Taking the average of an interval of length d, we have

$$\begin{aligned}&\frac{1}{d}\sum_{\tau=t}^{t+d-1}\mathbb{E}\{\sum_b\mu^{(c)}_{ib}(\tau)|\overrightarrow{U(t)}\}\\\geq&\frac{1}{d}\sum_{\tau=t}^{t+d-1}(\sum_b\overline{\mu}^{(c)}_{ib})(1-2N\alpha\gamma^{\tau-t})\\=&(\sum_b\overline{\mu}^{(c)}_{ib})(1-\frac{2N\alpha(1-\gamma^d)}{d(1-\gamma)}),\end{aligned} \tag{14}$$

$$\frac{1}{d}\sum_{\tau=t}^{t+d-1}\mathbb{E}\{\sum_a\mu^{(c)}_{ai}(\tau)|\overrightarrow{U(t)}\}\leq(\sum_a\overline{\mu}^{(c)}_{ai})(1+\frac{2N\alpha(1-\gamma^d)}{d(1-\gamma)}), \tag{15}$$

where $\frac{\alpha(1-\gamma^d)}{d(1-\gamma)}=\frac{\delta}{2N^2}$.

By applying inequalities (14) and (15), we can derive a lower bound on the last term of (12) and further bound the d-step Lyapunov drift. The derivation is divided into two cases.

1) In an odd subslot, node i operates in the *source-to-relay mode* and is the source of packet c: $A^{(c)}_i(\tau)=\lambda$ and $\sum_a\mu^{(c)}_{ai}(\tau)=0$. So we just need to calculate $\sum_b\overline{\mu}^{(c)}_{ib}$. Let r_1 denote the transmission rate in this mode. According to our algorithm, we have $Nr_1=C(R_1p+R_2p')$, i.e., $r_1=\frac{R_1p+R_2p'}{\theta}$. Let $\varphi\triangleq\frac{R_1p+R_2p'}{R_1p+R_2p'+R_1q+R_2q'}(0<\varphi<1)$, then $r_1=(k+1)\mu\varphi$. Let $\delta=\frac{1-\rho}{2}$ and $\frac{\alpha(1-\gamma^d)}{d(1-\gamma)}=\frac{\delta}{2N^2}$, then

the time average summation of the last term of (12) can be expressed as follows:

$$\begin{aligned}
&\frac{1}{d}\sum_{\tau=t}^{t+d-1}\mathbb{E}\{\sum_b \mu_{ib}^{(c)}(\tau)-\sum_a \mu_{ai}^{(c)}(\tau)-A_i^{(c)}(\tau)|\overrightarrow{U}(t)\}\\
\geq\ & (\sum_b \overline{\mu}_{ib}^{(c)})(1-\frac{2N\alpha(1-\gamma^d)}{d(1-\gamma)})-\lambda \qquad (16)\\
=\ & (k+1)\mu\varphi(1-\frac{\delta}{N})-\rho\mu\\
\geq\ & (k+1)\mu[(1-\frac{\delta}{N})-\frac{\rho}{k+1}]\geq\frac{(k+1)\mu(1-\rho)}{2}.
\end{aligned}$$

2) In an even subslot, node i operates in the *relay-to-destination mode* and is the relay of packet c: $A_i^{(c)}(\tau)=0$, $\mu_{ai}^{(c)}(\tau)>0$ only when node a is the source node of packet c; $\mu_{ib}^{(c)}(\tau)>0$ only when node b is node c. According to our algorithm, with probability $\frac{1-\delta}{2}$ the source-to-relay transmission happens and with probability $\frac{1+\delta}{2}$ the relay-to-destination transmission happens. Thus the total relay-to-destination transmission rate $r_2=\frac{1+\delta}{1-\delta}r_1$. Moreover, in our algorithm, a relay can receive packets from $N-1$ source nodes with equal probability (except for itself). Likewise, it can relay packets to $N-2$ nodes with equal probability (except for the source and itself). Thus,

$$\sum_a \overline{\mu}_{ai}^{(c)}=\frac{r_1}{N-1},\quad \sum_b \overline{\mu}_{ib}^{(c)}=\frac{r_2}{N-2}.$$

Let $\delta=\frac{1-\rho}{2}$ and $\frac{\alpha(1-\gamma^d)}{d(1-\gamma)}=\frac{\delta}{2N^2}$, we derive the time average summation in (12) as follows:

$$\begin{aligned}
&\frac{1}{d}\sum_{\tau=t}^{t+d-1}\mathbb{E}\{\sum_b \mu_{ib}^{(c)}(\tau)-\sum_a \mu_{ai}^{(c)}(\tau)-A_i^{(c)}(\tau)|\overrightarrow{U}(t)\}\\
\geq\ & (\sum_b \overline{\mu}_{ib}^{(c)})(1-\frac{2N\alpha(1-\gamma^d)}{d(1-\gamma)})\\
&-(\sum_a \overline{\mu}_{ai}^{(c)})(1+\frac{2N\alpha(1-\gamma^d)}{d(1-\gamma)})\\
=\ & (\sum_b \overline{\mu}_{ib}^{(c)}-\sum_a \overline{\mu}_{ai}^{(c)})-(\sum_b \overline{\mu}_{ib}^{(c)}+\sum_a \overline{\mu}_{ai}^{(c)})\frac{\delta}{N}\\
\geq\ & \frac{(r_2-r_1)-\frac{(r_2+r_1)\delta}{N}}{N-2}\geq\frac{(k+1)\mu\varphi(1-\rho)}{N(1-\delta)}. \qquad (17)
\end{aligned}$$

Applying (16) and (17) into (12), we have

$$\begin{aligned}
&\mathbb{E}\{L(\overrightarrow{U}(t+d))-L(\overrightarrow{U}(t))|\overrightarrow{U}(t)\}\\
&\leq d^2BN-d(\frac{1+\delta}{2}\frac{(k+1)\mu(1-\rho)}{2}\\
&\quad+\frac{1-\delta}{2}\frac{(k+1)\mu\varphi(1-\rho)}{N(1-\delta)})\sum_{i\neq c}U_i^{(c)}(t)\\
&\leq d^2BN-d\frac{(k+1)\mu\varphi(1-\rho)}{2N})\sum_{i\neq c}U_i^{(c)}(t). \qquad (18)
\end{aligned}$$

Note that the inequality (18) satisfies the condition in *Lemma 1*, then by *Lemma 1*, we have

$$\limsup_{t\to\infty}\frac{1}{t}\sum_{\tau=0}^{t-1}\sum_{i\neq c}\mathbb{E}\{U_i^{(c)}(\tau)\}\leq\frac{2dBN^2}{(k+1)\mu\varphi(1-\rho)}. \qquad (19)$$

Since the total input rate is $N\lambda$, by Little's Law, the average delay for a source destination pair is $\frac{2dBN}{(k+1)\lambda\mu\varphi(1-\rho)}$. Recall the definition of network delay that the time needed for a packet to be transmitted from the source to its k destinations. To analyze the network delay, we apply the cover time of the Markov chain.

We formulate the multicast transmission session as an ergodic Markov chain $S=\{s_0,s_1,s_2,...,s_k\}$ with $k+1$ states. The state $S_i=1$ denotes that the ith destination has obtained the packet. At the very beginning of the transmission, no destination holds the packet, thus all of the states in the Markov chain is 0. The first hitting time T_i of the state i ($1\leq i\leq k$) is defined as the time needed for the ith destination to obtain the packet. Thus the expected time for a source sending the packet to all its k destinations equals to the time needed to cover all the $k+1$ states of the Markov chain.

We next introduce a following lemma to compute the cover time in such case.

Lemma 1. *For the Markov chain with $k+1$ states defined as before with initial state 0, the cover time $D=max_{1\leq i\leq k}T_i$ for all the states $\{1,2,...,K\}$ satisfies*

$$\mathbb{E}[D]=(\sum_{m=1}^{k}\frac{1}{m})\mathbb{E}[T_1].$$

Proof. Let $T_{0,m}(m\leq k)$ denote the time when all the states $\{0,1,2,...,m\}$ have been hit in some sequence and thus $D=T_{0,k}$. Due to the i.i.d. property, we have:

$$\mathbb{E}[T_{0,m}-T_{0,m-1}|X_t,t\leq T_{0,m-1}]=\mathbb{E}[T_m]Pr(X_{T_{0,m}}=m).$$

Due to the finite static ergodic Markov chain mobility model, we have

$$\mathbb{E}[T_1]=\mathbb{E}[T_2]=...\mathbb{E}[T_k].$$

Since each of the k states is hit with equal probability, we can obtain that

$$Pr(X_{T_{0,m}}=m)=\frac{1}{m}.$$

Thus, the mean cover time $\mathbb{E}[D]$ is as follows:

$$\begin{aligned}
\mathbb{E}[D]&=\sum_{m=1}^{k}\mathbb{E}[T_{0,m}-T_{0,m-1}|X_t,t\leq T_{0,m-1}]\\
&=\sum_{m=1}^{k}\mathbb{E}[T_m]Pr(X_{T_{1,m}}=m) \qquad (20)\\
&=(\sum_{m=1}^{k}\frac{1}{m})\mathbb{E}[T_1].
\end{aligned}$$

□

In the network, with the *Two hop relay algorithm*, $\mathbb{E}[T_i]$ is $\frac{2dBN}{(k+1)\lambda\mu\varphi(1-\rho)}$ for all $1\leq i\leq k$. Also notice the summation of the harmonic series $\sum_{m=1}^{k}\frac{1}{m}\sim\log k$, by *Lemma 2*, we have that if the input rate $\lambda<\frac{R_1q+R_1P+R_2q'+R_2p'}{(k+1)\theta}$ in the case $R_1\geq 2R_2$, then the delay of the network is $\overline{D}\leq\frac{2dBN\log k}{(k+1)\lambda\mu\varphi(1-\rho)}$. In order sense, when $R_1\geq 2R_2$, the network capacity is $\Theta(\frac{1}{k})$ [1] and accordingly the delay is $\Theta(\frac{N\log k}{k})$. Therefore the tradeoff between delay and capacity in the network is $\Theta(N\log k)$, which is a better result than [6].

Appendix D
Derivation of the Minimum Energy Function

Here, we present the derivation of the minimum energy function under four different functions. Recall that the minimum energy function $\Phi(\lambda)$ is derived by solving an optimization function: $\overline{e} \geq \inf_{x\in\tilde{\Omega}} g(x)$, where $\Omega \subset \tilde{\Omega} = \tilde{\Omega}_0 \cap \tilde{\Omega}_1 \cap \tilde{\Omega}_2 \cap \tilde{\Omega}_3$ and $g(x) \leq f(x)$. The definitions of new constraint sets $\tilde{\Omega}_0$, $\tilde{\Omega}_1$, $\tilde{\Omega}_2$, $\tilde{\Omega}_3$ are as follows:

$$\tilde{\Omega}_0 \triangleq \Omega_0, \quad \tilde{\Omega}_1 \triangleq \Omega_1,$$

$$\tilde{\Omega}_2 \triangleq \{x | \frac{kx_{k0}}{R_1} + \frac{(k+1)}{R_1}\sum_{a\geq k+1} x_{a0} \leq c_1 + c_2\},$$

$$\tilde{\Omega}_3 \triangleq \{x | \frac{kx_{k0}}{R_1} + \frac{(k+1)}{R_1}\sum_{a\geq k+1} x_{a0} + \frac{kx_{0k}}{R_2} \leq c_1 + c_2 + c_3\}.$$

Now we analyze the four different bounds for $\overline{e}$, with each in the linear form of $\overline{e} \geq \alpha\lambda + \beta$. These bounds define the four piecewise linear regions of $\Phi(\lambda)$.

(1) First note that $f(x) \geq \frac{k}{R_1}\sum_{a,b|a+b\geq k}\frac{x_{ab}}{N}$. Therefore taking $g(x) = \frac{k}{R_1}\sum_{a,b|a+b\geq k}\frac{x_{ab}}{N}$, we have:

$$\overline{e} \geq \inf_{x\in\tilde{\Omega}_0} \frac{k}{R_1}\sum_{a,b|a+b\geq k}\frac{x_{ab}}{N},$$

where $\tilde{\Omega}_0 \triangleq \{x | \sum_{a,b|a+b\geq k} x_{ab} = N\lambda\}$. Hence, we have the first linear function: $\overline{e} \geq \frac{k\lambda}{R_1}$.

(2) For the second case, we note that $f(x) \geq \frac{kx_{k0}}{R_1 N} + \frac{\sum_{(a,b)\neq(k,0)} a x_{ab}}{R_1 N} \geq \frac{(k+1)x_{k0}}{R_1 N} + \frac{k\sum_{(a,b)\neq(k,0)} x_{ab}}{R_1 N}$. Thus taking this lower bound of $f(x)$ as $g(x)$, we have:

$$\overline{e} \geq \inf_{x\in\tilde{\Omega}_0\bigcap\tilde{\Omega}_1} \frac{kx_{k0}}{R_1 N} + \frac{(k+1)\sum_{(a,b)\neq(k,0)} x_{ab}}{R_1 N},$$

where $\tilde{\Omega}_1 \triangleq \{x | \frac{kX_{k0}}{R_1} \leq c_1\}$. Thus solving this optimization problem we have:

$$\begin{aligned}
\overline{e} &\geq \frac{kx_{k0}}{R_1 N} + \frac{(k+1)\sum_{(a,b)\neq(k,0)} x_{ab}}{R_1 N} \\
&\geq \frac{kx_{k0}}{R_1 N} + \frac{(k+1)(\sum_{a+b\geq k} x_{ab} - x_{k0})}{R_1 N} \\
&\geq \frac{kx_{k0}}{R_1 N} + \frac{(k+1)(N\lambda - x_{k0})}{R_1 N} \\
&\geq \frac{(k+1)\lambda}{R_1} - \frac{c_1}{Nk} \quad (\tilde{\Omega}_1) \\
&= \frac{q}{\theta} + \frac{k+1}{R_1}[\lambda - \frac{R_1 q}{k\theta}] \quad (c_1 = Cq).
\end{aligned}$$

(3) For the third case, we have:

$$f(x) \geq \frac{kx_{k0}}{R_1 N} + \frac{(k+1)\sum_{a\geq k+1} x_{a0}}{R_1 N} + \frac{k\sum_{a+b\geq k+1\&b\neq 0} x_{ab}}{R_2 N}.$$

Taking the lower bound of $f(x)$ as $g(x)$, we have:

$$\begin{aligned}
\overline{e} \geq \inf_{x\in\tilde{\Omega}_0\bigcap\tilde{\Omega}_1\bigcap\tilde{\Omega}_2} &\frac{kx_{k0}}{R_1 N} + \frac{(k+1)\sum_{a\geq k+1} x_{a0}}{R_1 N} \\
&+ \frac{k\sum_{a+b\geq k+1\&b\neq 0} x_{ab}}{R_2 N},
\end{aligned}$$

where $\tilde{\Omega}_2 \triangleq \{x | \frac{kx_{k0}}{R_1} + \frac{(k+1)}{R_1}\sum_{a\geq k+1} x_{a0} \leq c_1 + c_2\}$. Thus solving this optimization problem we have:

$$\begin{aligned}
&\overline{e} + \frac{k}{R_2 N}[N\lambda - x_{k0} - \sum_{a\geq k+1} x_{a0}] \\
\geq&\frac{c_1+c_2}{N} + \frac{k\lambda}{R_2} - \frac{k}{R_2 N}[x_{k0} + \sum_{a\geq k+1} x_{a0}] \quad (\tilde{\Omega}_2) \\
\geq&\frac{c_1+c_2}{N} + \frac{k\lambda}{R_2} - \frac{kR_1}{R_2 N}[\frac{c_1+c_2}{k+1} + \frac{c_1}{k(k+1)}] \quad (\tilde{\Omega}_2) \\
=&\frac{p}{\theta} + \frac{k}{R_2}[\lambda - \frac{R_1 p}{(k+1)\theta} - \frac{R_1 q}{(k+1)k\theta}] \\
&(c_1 = Cq, c_2 = C(p-q)).
\end{aligned}$$

(4) For the forth case, we have:

$$\begin{aligned}
f(x) \geq &\frac{kx_{k0}}{R_1 N} + \frac{(k+1)\sum_{a\geq k+1} x_{a0}}{R_1 N} + \frac{kx_{0k}}{R_2 N} \\
&+ \frac{(k+1)\sum_{b\geq k+1} x_{ab}}{R_2 N}.
\end{aligned}$$

Taking the lower bound of $f(x)$ as $g(x)$, we have:

$$\begin{aligned}
\overline{e} \geq \inf_{x\in\tilde{\Omega}_0\bigcap\tilde{\Omega}_1\bigcap\tilde{\Omega}_2\bigcap\tilde{\Omega}_3} &\frac{kx_{k0}}{R_1 N} + \frac{(k+1)\sum_{a\geq k+1} x_{a0}}{R_1 N} \\
&+ \frac{kx_{0k}}{R_2 N} + \frac{(k+1)\sum_{b\geq k+1} x_{ab}}{R_2 N},
\end{aligned}$$

where $\tilde{\Omega}_3 \triangleq \{x | \frac{kx_{k0}}{R_1} + \frac{(k+1)}{R_1}\sum_{a\geq k+1} x_{a0} + \frac{kx_{0k}}{R_2} \leq c_1 + c_2 + c_3\}$.

1. We use the standard asymptotic notations throughput this paper. For two functions $f(n)$ and $g(n)$, the notations are as follows: $f(n) = o(g(n)) \Leftrightarrow \lim_{n\to\infty}\frac{f(n)}{g(n)} = 0$; $f(n) = \omega(g(n)) \Leftrightarrow \lim_{n\to\infty}\frac{g(n)}{f(n)} = 0$; $f(n) = O(g(n)) \Leftrightarrow \lim_{n\to\infty}\frac{f(n)}{g(n)} < \infty$; $f(n) = \Omega(g(n)) \Leftrightarrow \lim_{n\to\infty}\frac{g(n)}{f(n)} < \infty$; $f(n) = \Theta(g(n)) \Leftrightarrow f(n) = \Omega(g(n))$ and $g(n) = O(f(n))$.

Thus solving this optimization problem we have:

$$\begin{aligned}
\overline{e} \geq & \frac{kx_{k0}}{R_1 N} + \frac{(k+1)\sum\limits_{a\geq k+1} x_{a0}}{R_1 N} + \frac{kx_{0k}}{R_2 N} \\
& + \frac{k+1}{R_2 N}\big[N\lambda - x_{k0} - \sum_{a\geq k+1} x_{a0} - x_{0k}\big] \\
= & \frac{x_{k0}}{N}\big[\frac{k}{R_1} - \frac{k+1}{R_2}\big] + \frac{x_{a0}}{N}\big[\frac{k+1}{R_1} - \frac{k+1}{R_2}\big] \\
& + \frac{(k+1)N\lambda}{R_2 N} - \frac{x_{0k}}{R_2 N} \\
\geq & \frac{1}{N}\big[\frac{1}{R_1} - \frac{1}{R_2}\big]\big[kx_{k0} - (k+1)x_{a0}\big] + \frac{(k+1)\lambda}{R_2} \\
& - \frac{c_3}{kN} - \frac{x_{k0}}{R_2 N} \quad (\tilde{\Omega}_3) \\
\geq & \big[\frac{1}{R_1} - \frac{1}{R_2}\big]\frac{R_1(c_1+c_2)}{N} + \frac{(k+1)\lambda}{R_2} - \frac{c_3}{kN} \\
& - \frac{c_1 R_1}{kR_2 N} \quad (\tilde{\Omega}_1, \tilde{\Omega}_2) \\
= & \frac{p+q'}{\theta} + \frac{k+1}{R_2}\big[\lambda - \frac{R_2 q' + R_1(kp+q)}{(k+1)k\theta}\big] \\
& (c_1 = Cq, c_2 = C(p-q), c_3 = Cq').
\end{aligned}$$