Heat-Diffusion: Pareto Optimal Dynamic Routing for Time-Varying Wireless Networks

Reza Banirazi, *Member, IEEE,* Edmond Jonckheere, *Life Fellow, IEEE,*
and Bhaskar Krishnamachari, *Senior Member, IEEE*

This supplementary material carries three sections as appendices to the main paper, with its title and authors being replicated above. The first appendix contains the proof of all theorems and lemmas in the main paper. The second appendix studies the performance of proposed HD routing policy versus V-parameter BP in a real sensor network with forty five wireless sensors. The third appendix is a brief review of graph Laplacian matrix and its properties.

APPENDIX A
PROOF OF THEOREMS AND LEMMAS

Note that in the proofs we often drop timeslot variable (n) for ease of notation and concision.

Proof of Theorem 1 (HD Key Property)

One can verify that

$$D(\boldsymbol{f}, \boldsymbol{q}_\circ, n) = \sum\nolimits_{ij\in\mathcal{E}} 2\,\phi_{ij}(n) q_{ij}(n) f_{ij}(n) - f_{ij}(n)^2.$$

Let us temporarily relax all network constraints. Then each link-related component of $D(\boldsymbol{f}, \boldsymbol{q}_\circ, n)$ turns to be strictly concave. For each link ij, by taking the first derivative with respect to f_{ij}, we find the maximizing link transmission $f_{ij}^{\rm opt} = \phi_{ij} q_{ij}$. Considering the link constraints that f_{ij} must be non-negative and at most equal to the link capacity yields

$$f_{ij}^{\rm opt} = \min\{\phi_{ij}\, {q_{ij}}^+,\, \mu_{ij}\}$$

which follows $\widehat{f_{ij}}$ in (6). Considering the link interference constraint, on the other hand, enforces to activate the links that contribute most to the D maximization. Then assuming that an interference model does not let a node transmit to more than one neighbor at the same time, the latter directly leads to the max-weight scheduling (8) alongside the HD weighting (7), which concludes the proof. ∎

Proof of Lemma 1

Define $\boldsymbol{\Delta} := \boldsymbol{B}_\circ \boldsymbol{B}_\circ^\top$ and $\boldsymbol{\Delta}_\phi := \boldsymbol{B}_\circ \boldsymbol{\Phi} \boldsymbol{B}_\circ^\top$, which are both positive definite matrices. Since $\boldsymbol{\Delta}_\phi^{1/2} \boldsymbol{\Delta}^{-1} \boldsymbol{\Delta}_\phi^{1/2}$ is congruent to $\boldsymbol{\Delta}^{-1}$ which has only positive eigenvalues, by Sylvester's law of inertia, $\boldsymbol{\Delta}_\phi^{1/2} \boldsymbol{\Delta}^{-1} \boldsymbol{\Delta}_\phi^{1/2}$ has only positive eigenvalues too. The latter is similar to $\boldsymbol{M}_\circ$, and so they have the same eigenvalues, proving that $\boldsymbol{M}_\circ$ has only positive eigenvalues.

We now show that $\boldsymbol{x}^\top \boldsymbol{M}_\circ \boldsymbol{x} \geqslant 0$. Letting $\boldsymbol{v} := \boldsymbol{\Delta}^{-1} \boldsymbol{x}$ and substituting for $\boldsymbol{M}_\circ$, it suffices to show that

$$(\boldsymbol{B}_\circ^\top \boldsymbol{v})^\top \big(\boldsymbol{\Phi} \boldsymbol{B}_\circ^\top \boldsymbol{B}_\circ\big)(\boldsymbol{B}_\circ^\top \boldsymbol{v}) \geqslant 0\,. \tag{A1}$$

The authors are with Ming Hsieh Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089.
E-mail: {banirazi, jonckhee, bkrishna}@usc.edu

Doing another change of variable, let $\boldsymbol{f} := \boldsymbol{B}_\circ^\top \boldsymbol{v}$ that represents an edge vector in which $f_{ij} = v_i - v_j$, $\forall ij \in \mathcal{E}$. Recall that $\boldsymbol{B}_\circ$ is a signed node-edge incidence matrix with arbitrarily assigned algebraic topological edge orientations. Let us assign edge orientations such that $f_{ij} \geqslant 0$, $\forall ij \in \mathcal{E}$. Then to fulfill (A1), it suffices to show that $\boldsymbol{f}^\top \boldsymbol{\Phi} \boldsymbol{B}_\circ^\top \boldsymbol{B}_\circ \boldsymbol{f} \geqslant 0$ subject to $\boldsymbol{f} \succcurlyeq \boldsymbol{0}$, which reads $f_{ij} \geqslant 0$, $\forall ij \in \mathcal{E}$. To this end, we equivalently show that minimum cost in the following optimization problem is non-negative:

$$\min_{\boldsymbol{f}\succcurlyeq\boldsymbol{0}} \;\; \boldsymbol{f}^\top \boldsymbol{\Phi} \boldsymbol{B}_\circ^\top \boldsymbol{B}_\circ \boldsymbol{f}\,.$$

Let us construct the Lagrangian dual problem

$$\max_{\boldsymbol{\lambda}\succcurlyeq\boldsymbol{0}} \min_{\boldsymbol{f}} \Big(\mathcal{L}(\boldsymbol{\lambda}, \boldsymbol{f}) := \boldsymbol{f}^\top \boldsymbol{\Phi} \boldsymbol{B}_\circ^\top \boldsymbol{B}_\circ \boldsymbol{f} - \boldsymbol{\lambda}^\top \boldsymbol{f} \Big) \tag{A2}$$

with $\boldsymbol{\lambda}$ being the vector of Lagrange multipliers. Since the primal variable $\boldsymbol{f}$ is continuously differentiable, so the Lagrangian $\mathcal{L}$, and thus the minimum occurs where $\nabla_{\boldsymbol{f}}\, \mathcal{L} = \boldsymbol{0}$, which leads to

$$\boldsymbol{\lambda} = \big(\boldsymbol{B}_\circ^\top \boldsymbol{B}_\circ \boldsymbol{\Phi} + \boldsymbol{\Phi} \boldsymbol{B}_\circ^\top \boldsymbol{B}_\circ\big)\, \boldsymbol{f}^{\rm opt}.$$

Substituting $\boldsymbol{f}^{\rm opt}$ in (A2) and noting that both $\boldsymbol{f}^{\rm opt}$ and $\boldsymbol{\lambda}$ are entrywise non-negative, we obtain

$$\max_{\boldsymbol{\lambda}\succcurlyeq\boldsymbol{0}} \;\; \mathcal{L}(\boldsymbol{\lambda}) = \max_{\boldsymbol{\lambda}\succcurlyeq\boldsymbol{0}} \; -\frac{1}{2}\, \boldsymbol{\lambda}^\top \boldsymbol{f}^{\rm opt} = 0\,. \tag{A3}$$

By the weak duality theorem, the minimum of the primal problem is greater than or equal to the maximum of the dual problem. Thus, (A3) entails $\min_{\boldsymbol{f}\succcurlyeq\boldsymbol{0}} \boldsymbol{f}^\top \boldsymbol{\Phi} \boldsymbol{B}_\circ^\top \boldsymbol{B}_\circ \boldsymbol{f} \geqslant 0$, which equally means $\boldsymbol{x}^\top \boldsymbol{M}_\circ \boldsymbol{x} \geqslant 0$.

It remains to show that $\boldsymbol{x}^\top \boldsymbol{M}_\circ \boldsymbol{x} = 0$ only if $\boldsymbol{x} = 0$, which is equivalent to show that matrix $\boldsymbol{M}_\circ^\top + \boldsymbol{M}_\circ$ is positive definite. Since $\boldsymbol{x}^\top (\boldsymbol{M}_\circ^\top + \boldsymbol{M}_\circ)\, \boldsymbol{x} \geqslant 0$ already guarantees that $\boldsymbol{M}_\circ^\top + \boldsymbol{M}_\circ$ is positive semi-definite, it suffices to show that $\boldsymbol{M}_\circ^\top + \boldsymbol{M}_\circ$ has no zero eigenvalue. Let us assume it does, which implies the existence of an eigenvector $\boldsymbol{\nu} \neq \boldsymbol{0}$ such that

$$(\boldsymbol{M}_\circ^\top + \boldsymbol{M}_\circ)\boldsymbol{\nu} = \boldsymbol{0}\,. \tag{A4}$$

Because $\boldsymbol{M}_\circ$ is the product of two positive definite matrices, $\boldsymbol{\nu} \neq \boldsymbol{0}$ entails $\boldsymbol{M}_\circ \boldsymbol{\nu} \neq \boldsymbol{0}$, which leads to $(\boldsymbol{M}_\circ \boldsymbol{\nu})^\top \boldsymbol{M}_\circ \boldsymbol{\nu} + (\boldsymbol{M}_\circ^\top \boldsymbol{\nu})^\top \boldsymbol{M}_\circ^\top \boldsymbol{\nu} > 0$. Utilizing (A4) in the latter results in

$$\boldsymbol{\nu}^\top \big(\, \boldsymbol{M}_\circ^\top - \boldsymbol{M}_\circ \big)^2\, \boldsymbol{\nu} < 0$$

which is not true as $(\boldsymbol{M}_\circ^\top - \boldsymbol{M}_\circ)^2$ is a symmetric positive semi-definite matrix. Therefore, $\boldsymbol{M}_\circ^\top + \boldsymbol{M}_\circ$ has no zero eigenvalue and so is symmetric positive definite. ∎

Proof of Lemma 2

By definition of $\boldsymbol{M}_\circ$, we already have $\boldsymbol{B}_\circ \boldsymbol{B}_\circ^\top \boldsymbol{M}_\circ \boldsymbol{x} = \boldsymbol{B}_\circ \boldsymbol{\Phi} \boldsymbol{B}_\circ^\top \boldsymbol{x}$, which could easily be seen by substituting $\boldsymbol{M}_\circ$

from (18). Thus, to prove the claim, it suffices to show that for any vectors $\boldsymbol{x}$ and $\boldsymbol{y}$, equality $\boldsymbol{B}_\circ \boldsymbol{y} = \boldsymbol{B}_\circ \boldsymbol{\Phi} \boldsymbol{B}_\circ^\top \boldsymbol{x}$ entails $\boldsymbol{y} = \boldsymbol{\Phi} \boldsymbol{B}_\circ^\top \boldsymbol{x}$. To this end, we utilize the properties of heat equations on undirected graphs (see Sec. VII-B).

Consider a thermal graph with reduced node-edge incidence matrix $\boldsymbol{B}_\circ$ and edge thermal diffusivity matrix $\boldsymbol{\Phi}$ and let the destination node be fixed at zero temperature. As the first scenario, let us envision $\boldsymbol{y}$ as the vector of heat fluxes through the branches, implying that $\boldsymbol{B}_\circ \boldsymbol{y}$ represents the vector of heat sources injected at the nodes (see (32) and (33) under constant heat sources.) As the second scenario, envision $\boldsymbol{x}$ as the vector of temperatures at the nodes, implying that $\boldsymbol{\Phi} \boldsymbol{B}_\circ^\top \boldsymbol{x}$ represents the vector of heat fluxes through the branches and $\boldsymbol{B}_\circ \boldsymbol{\Phi} \boldsymbol{B}_\circ^\top \boldsymbol{x}$ represents the vector of heat sources injected at the nodes.

Assuming $\boldsymbol{B}_\circ \boldsymbol{y} = \boldsymbol{B}_\circ \boldsymbol{\Phi} \boldsymbol{B}_\circ^\top \boldsymbol{x}$ means that the thermal graph is charged by the same configuration of heat sources in both scenarios above. It follows that the vector of temperatures at the nodes are also the same as the Dirichlet Laplacian is positive definite in (33). Hence, in both scenarios the vector of heat fluxes through the branches must be equal, because $\boldsymbol{B}_\circ$ has full row rank in (32). This entails $\boldsymbol{y} = \boldsymbol{\Phi} \boldsymbol{B}_\circ^\top \boldsymbol{x}$, which concludes the proof. ∎

Proof of Lemma 3

Let us replace $\boldsymbol{M}_\circ + \boldsymbol{M}_\circ^\top$ by $2\boldsymbol{M}_\circ + (\boldsymbol{M}_\circ^\top - \boldsymbol{M}_\circ)$. Doing some matrix manipulation, we need to show that there exists such $1 \leqslant \eta \leqslant 3$ that for arbitrary vectors $\boldsymbol{x}$ and $\boldsymbol{y}$,

$$\boldsymbol{x}^\top (\boldsymbol{M}_\circ^\top - \boldsymbol{M}_\circ)\, \boldsymbol{y} \leqslant (\eta - 2)\, \boldsymbol{x}^\top \boldsymbol{M}_\circ\, \boldsymbol{y}\,. \tag{A5}$$

To this end, it suffices to show $\left|\boldsymbol{x}^\top (\boldsymbol{M}_\circ^\top - \boldsymbol{M}_\circ)\, \boldsymbol{y}\right| \leqslant \left|\boldsymbol{x}^\top \boldsymbol{M}_\circ\, \boldsymbol{y}\right|$, which then makes the inequality (A5) true for $\eta = 1$ in case of $\boldsymbol{x}^\top \boldsymbol{M}_\circ\, \boldsymbol{y} \leqslant 0$, and for $\eta = 3$ in case of $\boldsymbol{x}^\top \boldsymbol{M}_\circ\, \boldsymbol{y} > 0$. This is equivalent to show that the following inequality holds:

$$\boldsymbol{x}^\top (\boldsymbol{M}_\circ^\top - \boldsymbol{M}_\circ)\, \boldsymbol{y}\boldsymbol{y}^\top (\boldsymbol{M}_\circ - \boldsymbol{M}_\circ^\top)\, \boldsymbol{x} \leqslant \boldsymbol{x}^\top \boldsymbol{M}_\circ\, \boldsymbol{y}\boldsymbol{y}^\top \boldsymbol{M}_\circ^\top \boldsymbol{x}\,.$$

By little algebra, the latter can be rephrased as

$$\boldsymbol{x}^\top (2\,\boldsymbol{M}_\circ - \boldsymbol{M}_\circ^\top)\, \boldsymbol{y}\boldsymbol{y}^\top \boldsymbol{M}_\circ\, \boldsymbol{x} \geqslant 0\,.$$

To prove the above inequality, it suffices to show that the minimum objective value in the following optimization problem is non-negative:

$$\begin{aligned} \min_{\boldsymbol{x},\boldsymbol{y}} \quad & \boldsymbol{x}^\top (2\,\boldsymbol{M}_\circ - \boldsymbol{M}_\circ^\top)\, \boldsymbol{y}\, \boldsymbol{y}^\top \boldsymbol{M}_\circ \boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{x}^\top \boldsymbol{M}_\circ \boldsymbol{x} > 0 \ , \ \boldsymbol{y}^\top \boldsymbol{M}_\circ \boldsymbol{y} > 0 \end{aligned}$$

where the constraints are enforced in light of Lem. 1. The Lagrangian dual problem, with λ_x and λ_y as the Lagrange multipliers, is found as

$$\max_{\lambda_x,\lambda_y \geqslant 0} \min_{\boldsymbol{x},\boldsymbol{y}} \Big(\mathcal{L} := \boldsymbol{x}^\top (2\,\boldsymbol{M}_\circ - \boldsymbol{M}_\circ^\top)\, \boldsymbol{y}\, \boldsymbol{y}^\top \boldsymbol{M}_\circ \boldsymbol{x} - \lambda_x\, \boldsymbol{x}^\top \boldsymbol{M}_\circ \boldsymbol{x} - \lambda_y\, \boldsymbol{y}^\top \boldsymbol{M}_\circ \boldsymbol{y} \Big)$$

Imposing the first order conditions $\nabla_{\boldsymbol{x}}\, \mathcal{L} = \boldsymbol{0}$ and $\nabla_{\boldsymbol{y}}\, \mathcal{L} = \boldsymbol{0}$ leads to

$$\begin{aligned} \lambda_x\, (\boldsymbol{M}_\circ^\top + \boldsymbol{M}_\circ)\, \boldsymbol{x} &= 2\, \boldsymbol{M}_\circ \boldsymbol{y}\, \boldsymbol{y}^\top \boldsymbol{M}_\circ \boldsymbol{x} + 2\, \boldsymbol{M}_\circ^\top \boldsymbol{y}\, \boldsymbol{y}^\top \boldsymbol{M}_\circ^\top \boldsymbol{x} \\ &\quad - 2\, \boldsymbol{M}_\circ^\top \boldsymbol{y}\, \boldsymbol{y}^\top \boldsymbol{M}_\circ \boldsymbol{x} \\ \lambda_y\, (\boldsymbol{M}_\circ^\top + \boldsymbol{M}_\circ)\, \boldsymbol{y} &= 2\, \boldsymbol{M}_\circ \boldsymbol{x}\, \boldsymbol{x}^\top \boldsymbol{M}_\circ \boldsymbol{y} + 2\, \boldsymbol{M}_\circ^\top \boldsymbol{x}\, \boldsymbol{x}^\top \boldsymbol{M}_\circ^\top \boldsymbol{y} \\ &\quad - 2\, \boldsymbol{M}_\circ \boldsymbol{x}\, \boldsymbol{x}^\top \boldsymbol{M}_\circ^\top \boldsymbol{y}\,. \end{aligned}$$

Let us plug these two equations into the Lagrangian $\mathcal{L}$ and utilize the identities $\boldsymbol{x}^\top \boldsymbol{M}_\circ^\top \boldsymbol{y} = \boldsymbol{y}^\top \boldsymbol{M}_\circ \boldsymbol{x} := a$ and $\boldsymbol{x}^\top \boldsymbol{M}_\circ \boldsymbol{y} = \boldsymbol{y}^\top \boldsymbol{M}_\circ^\top \boldsymbol{x} := b$ with a and b being scalars. One can easily confirm the following identities:

$$\begin{aligned} \mathcal{L} &= (2\,a\,b - b^2) - \lambda_x\, \boldsymbol{x}^\top \boldsymbol{M}_\circ \boldsymbol{x} - \lambda_y\, \boldsymbol{y}^\top \boldsymbol{M}_\circ \boldsymbol{y} \\ \lambda_x\, \boldsymbol{x}^\top (\boldsymbol{M}_\circ^\top + \boldsymbol{M}_\circ)\, \boldsymbol{x} &= 2\,(2\,a\,b - b^2) \\ \lambda_y\, \boldsymbol{y}^\top (\boldsymbol{M}_\circ^\top + \boldsymbol{M}_\circ)\, \boldsymbol{y} &= 2\,(2\,a\,b - b^2). \end{aligned}$$

Then by little algebra, the Lagrangian can be transformed to

$$\begin{aligned} \mathcal{L} &= \frac{1}{4}\, \lambda_x\, \boldsymbol{x}^\top \big(\boldsymbol{M}_\circ^\top + \boldsymbol{M}_\circ\big) \boldsymbol{x} + \frac{1}{4}\, \lambda_y\, \boldsymbol{y}^\top \big(\boldsymbol{M}_\circ^\top + \boldsymbol{M}_\circ\big) \boldsymbol{y} \\ &\qquad - \lambda_x\, \boldsymbol{x}^\top \boldsymbol{M}_\circ \boldsymbol{x} - \lambda_y\, \boldsymbol{y}^\top \boldsymbol{M}_\circ \boldsymbol{y} \\ &= \frac{1}{4}\, \lambda_x\, \boldsymbol{x}^\top \big(\boldsymbol{M}_\circ^\top - 3\,\boldsymbol{M}_\circ\big) \boldsymbol{x} + \frac{1}{4}\, \lambda_y\, \boldsymbol{y}^\top \big(\boldsymbol{M}_\circ^\top - 3\,\boldsymbol{M}_\circ\big) \boldsymbol{y}\,. \end{aligned}$$

Since $\boldsymbol{M}_\circ - \boldsymbol{M}_\circ^\top$ is skew-symmetric, both $\boldsymbol{x}^\top (\boldsymbol{M}_\circ - \boldsymbol{M}_\circ^\top) \boldsymbol{x}$ and $\boldsymbol{y}^\top (\boldsymbol{M}_\circ - \boldsymbol{M}_\circ^\top) \boldsymbol{y}$ vanish. In light of $\boldsymbol{x}^\top \boldsymbol{M}_\circ \boldsymbol{x} > 0$ and $\boldsymbol{y}^\top \boldsymbol{M}_\circ \boldsymbol{y} > 0$, the Lagrangian dual problem reads

$$\max_{\lambda_x,\lambda_y \geqslant 0}\ \mathcal{L} = \max_{\lambda_x,\lambda_y \geqslant 0} -\frac{1}{2} \left(\lambda_x\, \boldsymbol{x}^\top \boldsymbol{M}_\circ \boldsymbol{x} + \lambda_y\, \boldsymbol{y}^\top \boldsymbol{M}_\circ \boldsymbol{y}\right) = 0\,.$$

This entails $\boldsymbol{x}^\top (2\,\boldsymbol{M}_\circ - \boldsymbol{M}_\circ^\top)\, \boldsymbol{y}\boldsymbol{y}^\top \boldsymbol{M}_\circ\, \boldsymbol{x} \geqslant 0$ by the weak duality theorem that the maximum of the dual problem provides a lower bound for the minimum of the primal problem. ∎

Proof of Theorem 2 (HD Throughput Optimality)

To simplify the proof, we assume arrivals to be i.i.d. over timeslots. For non-i.i.d. arrivals with stationary ergodic processes of finite mean and variance, a similar analysis can be done using N-slot Lyapunov drift [27], where the queue evolution (10) is modified to

$$\boldsymbol{q}_\circ(n+N) = \boldsymbol{q}_\circ(n) + \sum_{k=n}^{n+N-1} \boldsymbol{a}_\circ(k) - \sum_{k=n}^{n+N-1} \boldsymbol{B}_\circ \boldsymbol{f}(k)\,. \tag{A6}$$

One can view N as the time required for the system to reach "near steady state," noting that in the i.i.d. case, the steady state is reached on each and every timeslot, and so $N = 1$.

Back to the proof for i.i.d. arrivals, suppose that $\overline{\boldsymbol{a}_\circ}$ is interior to the network capacity region $\mathcal{C}$. Thus, there exists an $\epsilon > 0$ such that $\overline{\boldsymbol{a}_\circ} + \epsilon \boldsymbol{1} \in \mathcal{C}$. Since the stationary randomized algorithm that generates $\boldsymbol{f'}(n)$ is throughput optimal [27], it can stabilize the arrival $\overline{\boldsymbol{a}_\circ} + \epsilon \boldsymbol{1}$ at each timeslot. The i.i.d. assumption on arrivals then leads to

$$\mathbb{E}\{\boldsymbol{a}_\circ - \boldsymbol{B}_\circ \boldsymbol{f'}\} = \overline{\boldsymbol{a}_\circ} - (\overline{\boldsymbol{a}_\circ} + \epsilon \boldsymbol{1}) = -\epsilon \boldsymbol{1}$$

implying that both $\boldsymbol{a}_\circ$ and $\boldsymbol{f'}$ reach their steady states on each and every timeslot. Plugging this into the Lyapunov drift inequality (23) yields

$$\mathbb{E}\big\{\Delta W\, |\boldsymbol{q}_\circ\big\} \leqslant -\eta\, \epsilon \boldsymbol{1}^\top \mathbb{E}\{\boldsymbol{M}_\circ\}\, \boldsymbol{q}_\circ + \Gamma_{\max}\,. \tag{A7}$$

Let us assume that there exists a $\mu > 0$, which is explored later, such that $\boldsymbol{1}^\top \mathbb{E}\{\boldsymbol{M}_\circ\}\, \boldsymbol{q}_\circ \geqslant \mu\, \boldsymbol{1}^\top \boldsymbol{q}_\circ$. Using this in the latter drift inequality leads to

$$\mathbb{E}\big\{\Delta W\, |\boldsymbol{q}_\circ\big\} \leqslant -\eta\, \mu\, \epsilon\, \boldsymbol{1}^\top \boldsymbol{q}_\circ + \Gamma_{\max}\,.$$

Thus, $\mathbb{E}\{\Delta W\, |\boldsymbol{q}_\circ\} < 0$ for any $\sum_i q_i > \Gamma_{\max} \big/ (\eta\, \mu\, \epsilon)$. This implies that the queuing system is stable and so $\overline{\boldsymbol{a}_\circ}$ is in the

HD stability region. Thus, any arrival rate $\overline{\boldsymbol{a}_\circ}$ being interior to the network capacity region is stabilized by HD with any $\beta \in [0,1]$, meaning that HD is throughput optimal for all $\beta \in [0,1]$.

We now show that there exists such a $\mu > 0$ that satisfies $\mathbf{1}^\top \mathbb{E}\{\boldsymbol{M}_\circ\}\, \boldsymbol{q}_\circ \geqslant \mu\, \mathbf{1}^\top \boldsymbol{q}_\circ$. Let us temporarily ignore the expectation and solve the problem for $\boldsymbol{M}_\circ$. The claim is trivial for $\boldsymbol{q}_\circ = \mathbf{0}$, and so we assume $\boldsymbol{q}_\circ \neq \mathbf{0}$. Further, $\boldsymbol{q}_\circ$ represents the vector of queue occupancies on the wireless network that are always non-negative. Let $\|\boldsymbol{q}_\circ\|_1$ represent the ℓ^1 norm of $\boldsymbol{q}_\circ$, defined as the sum of all queue occupancies. With no loss of generality, one may normalize $\|\boldsymbol{q}_\circ\|_1$ to one. The problem can then be rephrased as finding a $\mu > 0$ such that

$$\min_{\|\boldsymbol{q}_\circ\|_1 = 1, \boldsymbol{q}_\circ \succcurlyeq \mathbf{0}} \mathbf{1}^\top (\boldsymbol{M}_\circ - \mu \boldsymbol{I})\, \boldsymbol{q}_\circ \geqslant 0\,.$$

The latter is a standard linear programming problem. Using simplex method, the minimum lies on a vertex of the simplex, where the vertices of the simplex are the natural basis elements $\boldsymbol{e}_j : j = 1, ..., |\mathcal{V}|$. Thus, the μ is to be sought such that

$$\mathbf{1}^\top \Big((\boldsymbol{M}_\circ)_{:,j} - \mu\, \boldsymbol{e}_j \Big) \geqslant 0\,.$$

This immediately leads to $\mu = \min_j \mathbf{1}^\top \boldsymbol{M}_\circ\, \boldsymbol{e}_j$.

It remains to show that $\mathbf{1}^\top \boldsymbol{M}_\circ\, \boldsymbol{e}_j > 0$ for every natural basis $\boldsymbol{e}_j$. By Lem. 3, there exists such a $1 \leqslant \eta \leqslant 3$ as

$$\eta\, \mathbf{1}^\top \boldsymbol{M}_\circ\, \boldsymbol{e}_j \geqslant \mathbf{1}^\top \big(\boldsymbol{M}_\circ^\top + \boldsymbol{M}_\circ \big)\, \boldsymbol{e}_j$$

which implies $(\eta - 1)\, \mathbf{1}^\top \boldsymbol{M}_\circ\, \boldsymbol{e}_j \geqslant \mathbf{1}^\top \boldsymbol{M}_\circ^\top\, \boldsymbol{e}_j$. The right hand side is always positive by the next electrical circuit argument, which implies $(\eta - 1)\, \mathbf{1}^\top \boldsymbol{M}_\circ\, \boldsymbol{e}_j > 0$. The latter entails $\eta > 1$ and $\mathbf{1}^\top \boldsymbol{M}_\circ\, \boldsymbol{e}_j > 0$ as we desired.

To argue that $\mathbf{1}^\top \boldsymbol{M}_\circ^\top\, \boldsymbol{e}_j = \boldsymbol{e}_j^\top \boldsymbol{M}_\circ\, \mathbf{1} > 0$, by substituting $\boldsymbol{M}_\circ$ from (18), we need to show

$$\boldsymbol{e}_j^\top (\boldsymbol{B}_\circ \boldsymbol{B}_\circ^\top)^{-1} (\boldsymbol{B}_\circ \boldsymbol{\Phi} \boldsymbol{B}_\circ^\top)\, \mathbf{1} > 0\,.$$

Let us associate node-edge incidence matrix $\boldsymbol{B}_\circ$ with a resistive network and $\boldsymbol{e}_j$ with the vector of independent current sources attached to the nodes. Then the vector $\boldsymbol{v} := (\boldsymbol{B}_\circ \boldsymbol{B}_\circ^\top)^{-1} \boldsymbol{e}_j$ reads the voltages induced at the nodes. Since $\boldsymbol{e}_j$ implies that electrical current is injected into the network by a single current source at the node j, the resulting voltage at each node is non-negative ($\boldsymbol{v} \succcurlyeq \mathbf{0}$). Further, the voltages at the node j and at least at one of the nodes neighbor to ground (destination node) are always positive. On the other hand, the elements of each row of the Dirichlet Laplacian $\boldsymbol{B}_\circ \boldsymbol{\Phi} \boldsymbol{B}_\circ^\top$ sum to zero, except for those rows representing the nodes neighbor to ground, which always sum to a positive value. (Recall that $\boldsymbol{B}_\circ$ is obtained from $\boldsymbol{B}$ by discarding the row related to ground.) This implies that in the vector $\boldsymbol{u} := (\boldsymbol{B}_\circ \boldsymbol{\Phi} \boldsymbol{B}_\circ^\top)\, \mathbf{1}$, the components related to the nodes neighbor to ground are positive, and others are zero. Considering the conditions of $\boldsymbol{u}$ and $\boldsymbol{v}$ together, we get $\boldsymbol{e}_j^\top \boldsymbol{M}_\circ\, \mathbf{1} = \boldsymbol{v}^\top \boldsymbol{u} > 0$.

Replacing $\boldsymbol{M}_\circ$ by $\mathbb{E}\{\boldsymbol{M}_\circ\}$, the same argument leads to $\mu = \min_j \mathbf{1}^\top \mathbb{E}\{\boldsymbol{M}_\circ\}\, \boldsymbol{e}_j > 0$, which concludes the proof. ∎

Proof of Lemma 4

Considering the maximum of $G(\boldsymbol{f})$ subject to $\boldsymbol{f} \succcurlyeq \mathbf{0}$, the Lagrangian dual problem is obtained as

$$\min_{\boldsymbol{\lambda} \succcurlyeq \mathbf{0}}\ \max_{\boldsymbol{f}} \Big(\mathcal{L}(\boldsymbol{\lambda}, \boldsymbol{f}) := 2\, \boldsymbol{f}^\top \boldsymbol{B}_\circ^\top \boldsymbol{q}_\circ - \boldsymbol{f}^\top \boldsymbol{B}_\circ^\top \boldsymbol{B}_\circ \boldsymbol{f} + \boldsymbol{\lambda}^\top \boldsymbol{f} \Big)$$

with $\boldsymbol{\lambda}$ being the vector of Lagrange multipliers. Using the first order condition $\nabla_{\boldsymbol{f}}\, \mathcal{L} = \mathbf{0}$, we get

$$2\, \boldsymbol{B}_\circ^\top \boldsymbol{B}_\circ \boldsymbol{f}^{\mathrm{opt}} = 2\, \boldsymbol{B}_\circ^\top \boldsymbol{q}_\circ + \boldsymbol{\lambda}\,. \tag{A8}$$

Plugging $\boldsymbol{\lambda}$ from (A8) into the Lagrangian $\mathcal{L}$ leads to

$$\min\, \mathcal{L} = \boldsymbol{f}^{\mathrm{opt}\top} \boldsymbol{B}_\circ^\top \boldsymbol{B}_\circ \boldsymbol{f}^{\mathrm{opt}}.$$

By the weak duality theorem, the maximum of the primal problem is smaller than or equal to the minimum of the dual problem. Further, the duality gap is zero as $\mathcal{L}(\boldsymbol{\lambda}, \boldsymbol{f})$ is a convex functional, which leads to

$$\max\, G = \boldsymbol{f}^{\mathrm{opt}\top} \boldsymbol{B}_\circ^\top \boldsymbol{B}_\circ \boldsymbol{f}^{\mathrm{opt}}.$$

Substituting the latter into the G functional entails

$$\boldsymbol{f}^{\mathrm{opt}\top} \boldsymbol{B}_\circ^\top \boldsymbol{q}_\circ - \boldsymbol{f}^{\mathrm{opt}\top} \boldsymbol{B}_\circ^\top \boldsymbol{B}_\circ \boldsymbol{f}^{\mathrm{opt}} = 0\,. \tag{A9}$$

One can verify, on the other hand, that $\boldsymbol{B}_\circ^\top \boldsymbol{B}_\circ \boldsymbol{f}$ is an edge vector, in which the entry corresponding to edge ij reads

$$(\boldsymbol{B}_\circ^\top \boldsymbol{B}_\circ \boldsymbol{f})_{ij} = (\boldsymbol{B}_\circ \boldsymbol{f})_i - (\boldsymbol{B}_\circ \boldsymbol{f})_j\,. \tag{A10}$$

Under the K-hop interference model, two wireless links that share a common node cannot be scheduled in the same timeslot. Thus, for each scheduled link $ij \in \mathcal{E}$, we get the net outflow for node i as $(\boldsymbol{B}_\circ \boldsymbol{f})_i = f_{ij}$. When node j is not the final destination, we get $(\boldsymbol{B}_\circ \boldsymbol{f})_j = -f_{ij}$, and when it is, we get $(\boldsymbol{B}_\circ \boldsymbol{f})_j = 0$. (Recall that $\boldsymbol{B}_\circ$ is a reduction of $\boldsymbol{B}$ by discarding the row related to the final destination.) Using these identities in (A10), we get $(\boldsymbol{B}_\circ^\top \boldsymbol{B}_\circ \boldsymbol{f})_{ij} = \vartheta_{ij} f_{ij}$ with ϑ_{ij} defined in (6). Substituting the latter into (A9), we obtain

$$f_{ij}^{\mathrm{opt}} \big(q_{ij} - \vartheta_{ij} f_{ij}^{\mathrm{opt}} \big) = 0\,.$$

Considering the link constraints that f_{ij} must be non-negative and at most equal to the link capacity yields

$$f_{ij}^{\mathrm{opt}} = \min\{\, {q_{ij}}^+ / \vartheta_{ij},\ \mu_{ij} \}$$

which follows $\widehat{f_{ij}}$ in (6) with $\beta = 0$. Next is to activate the links that contribute most to the G maximization that directly leads to the max-weight scheduling (8) alongside the HD weighting (7) with $\beta = 0$, concluding the proof. ∎

Proof of Lemma 5

Consider $W(n) := \boldsymbol{q}_\circ(n)^\top \boldsymbol{q}_\circ(n)$ as the classical quadratic Lyapunov candidate and take expectation from the Lyapunov drift $\Delta W(n) = W(n+1) - W(n)$ to obtain

$$\begin{aligned} \mathbb{E}\{\Delta W\} = {} & \mathbb{E}\{\boldsymbol{a}_\circ - \boldsymbol{B}_\circ \boldsymbol{f}\}^\top \mathbb{E}\{\boldsymbol{a}_\circ - \boldsymbol{B}_\circ \boldsymbol{f} + 2\, \boldsymbol{q}_\circ\} \\ & - 2\, \mathbb{C}\mathrm{ov}\{\boldsymbol{B}_\circ \boldsymbol{f}, \boldsymbol{q}_\circ\} + \mathbb{V}\mathrm{ar}\{\boldsymbol{B}_\circ \boldsymbol{f}\} \\ & + 2\, \mathbb{C}\mathrm{ov}\{\boldsymbol{a}_\circ, \boldsymbol{q}_\circ - \boldsymbol{B}_\circ \boldsymbol{f}\} + \mathbb{V}\mathrm{ar}\{\boldsymbol{a}_\circ\} \end{aligned} \tag{A11}$$

where the equality holds at each timeslot and expectation is with respect to the randomness of arrivals, channel states and (possibly) routing decision. Let $\boldsymbol{g} := \boldsymbol{a}_\circ - \boldsymbol{B}_\circ \boldsymbol{f} + 2\, \boldsymbol{q}_\circ$, sum over timeslots 0 until $\tau - 1$, divide by τ and take a $\limsup$ of $\tau \to \infty$ from both sides of (A11) to obtain the following expected time average equation:

$$\begin{aligned} \limsup_{\tau \to \infty}\ \frac{1}{\tau} \sum_{n=0}^{\tau-1} & \mathbb{E}\{\boldsymbol{a}_\circ(n) - \boldsymbol{B}_\circ \boldsymbol{f}(n)\}^\top \mathbb{E}\{\boldsymbol{g}(n)\} = \\ & + 2\, \overline{\mathbb{C}\mathrm{ov}\{\boldsymbol{B}_\circ \boldsymbol{f}, \boldsymbol{q}_\circ\}} - \overline{\mathbb{V}\mathrm{ar}\{\boldsymbol{B}_\circ \boldsymbol{f}\}} \\ & - 2\, \overline{\mathbb{C}\mathrm{ov}\{\boldsymbol{a}_\circ, \boldsymbol{q}_\circ - \boldsymbol{B}_\circ \boldsymbol{f}\}} - \overline{\mathbb{V}\mathrm{ar}\{\boldsymbol{a}_\circ\}} \end{aligned} \tag{A12}$$

where we utilized $\limsup_{\tau\to\infty}(W(\tau)-W(0))/\tau=0$, as the routing policy stabilizes $\overline{\boldsymbol{a}_\circ}$ and so keeps $W(n)$ finite with probability 1 at each timeslot.

It remains to show that the left-hand side of (A12) vanishes. Observe that $\boldsymbol{g}(n)$ is entrywise non-negative and finite. Thus, there exist constant vectors $\boldsymbol{g}_{\min}$ and $\boldsymbol{g}_{\max}$ such that

$$\boldsymbol{0}\preccurlyeq\boldsymbol{g}_{\min}\preccurlyeq\mathbb{E}\{\boldsymbol{g}(n)\}\preccurlyeq\boldsymbol{g}_{\max}.$$

Hence, the left-hand side of (A12) is bounded from below to $(\,\overline{\boldsymbol{a}_\circ}-\boldsymbol{B}_\circ\overline{\boldsymbol{f}}\,)^\top\boldsymbol{g}_{\min}$ and from above to $(\,\overline{\boldsymbol{a}_\circ}-\boldsymbol{B}_\circ\overline{\boldsymbol{f}}\,)^\top\boldsymbol{g}_{\max}$. Further, as $\overline{\boldsymbol{a}_\circ}$ is stabilized by the routing policy, the feasibility condition in (17) entails $\overline{\boldsymbol{a}_\circ}=\boldsymbol{B}_\circ\overline{\boldsymbol{f}}$, implying that the left-hand side of (A12) vanishes. ∎

Proof of Theorem 3 (HD Minimum Delay)

To simplify the proof, we assume arrivals are i.i.d. over timeslots, with the understanding that it can easily be modified to yield similar result for non-i.i.d. arrivals, using the N-slot analysis derived from (A6).

Consider an arrival rate $\overline{\boldsymbol{a}_\circ}$ interior to the stability region of a $\mathcal{D}$-class routing policy, which we refer to it as "generic". Let the timeslot quantities $\boldsymbol{f}(n)$ and $\boldsymbol{q}_\circ(n)$ be produced by such a generic routing policy. If this generic routing policy also maximizes the G functional (24) at each slot n, by Assum. 1, it will result in the same $\overline{Q}$ as that of HD policy at $\beta=0$. Thus, we assume the G obtained by the generic policy is not maximal. Then for a sufficiently small $\epsilon>0$, there exists a routing algorithm (possibly unfeasible) that can stabilize the arrival $\overline{\boldsymbol{a}_\circ}+\epsilon\boldsymbol{1}$ while making $G(\boldsymbol{f},\boldsymbol{q}_\circ,n)$ not less than that of the generic routing policy at each slot n. Let us refer to this algorithm as "fictitious," as we do not intend to know how it really works. To rest assure that such an algorithm exists, one may endow it with the ability of perfectly predicting all future events with no uncertainty.

Let $\boldsymbol{f'}(n)$ represent the vector of link actual transmissions produced by the fictitious algorithm at slot n given $\boldsymbol{q}_\circ(n)$. Taking expectation from $G(\boldsymbol{f'},\boldsymbol{q}_\circ,n)\geqslant G(\boldsymbol{f},\boldsymbol{q}_\circ,n)$ and considering $\mathbb{E}\{\boldsymbol{B}_\circ\boldsymbol{f'}\}=\mathbb{E}\{\boldsymbol{B}_\circ\boldsymbol{f}\}+\epsilon\boldsymbol{1}$ due to the feasibility condition (17) and the i.i.d. arrivals, we obtain

$$\begin{aligned}2\,\epsilon\,\boldsymbol{1}^\top\mathbb{E}\{\boldsymbol{q}_\circ\}\geqslant\;&2\,\epsilon\,\boldsymbol{1}^\top\mathbb{E}\{\boldsymbol{B}_\circ\boldsymbol{f'}\}-\epsilon^2\boldsymbol{1}^\top\boldsymbol{1}\\&+\big(\,2\,\mathbb{C}\mathrm{ov}\{\boldsymbol{B}_\circ\boldsymbol{f},\boldsymbol{q}_\circ\}-\mathbb{V}\mathrm{ar}\{\boldsymbol{B}_\circ\boldsymbol{f}\}\,\big)\\&-\big(\,2\,\mathbb{C}\mathrm{ov}\{\boldsymbol{B}_\circ\boldsymbol{f'},\boldsymbol{q}_\circ\}-\mathbb{V}\mathrm{ar}\{\boldsymbol{B}_\circ\boldsymbol{f'}\}\,\big)\end{aligned}$$

which holds for each timeslot. Summing over timeslots 0 until $\tau-1$, dividing by τ and taking a $\limsup$ of $\tau\to\infty$ from both sides lead to the following expected time average inequality:

$$\begin{aligned}2\,\epsilon\,\boldsymbol{1}^\top\overline{\boldsymbol{q}_\circ}\geqslant\;&2\,\epsilon\,\boldsymbol{1}^\top(\overline{\boldsymbol{B}_\circ\boldsymbol{f'}})-\epsilon^2\boldsymbol{1}^\top\boldsymbol{1}\\&+\big(\,2\,\overline{\mathbb{C}\mathrm{ov}\{\boldsymbol{B}_\circ\boldsymbol{f},\boldsymbol{q}_\circ\}}-\overline{\mathbb{V}\mathrm{ar}\{\boldsymbol{B}_\circ\boldsymbol{f}\}}\,\big)\\&-\big(\,2\,\overline{\mathbb{C}\mathrm{ov}\{\boldsymbol{B}_\circ\boldsymbol{f'},\boldsymbol{q}_\circ\}}-\overline{\mathbb{V}\mathrm{ar}\{\boldsymbol{B}_\circ\boldsymbol{f'}\}}\,\big).\end{aligned}$$

Let us exploit Lem. 5 in the second and third lines and apply the identities $\mathbb{C}\mathrm{ov}\{\boldsymbol{a}_\circ+\epsilon\boldsymbol{1},\boldsymbol{q}_\circ-\boldsymbol{B}_\circ\boldsymbol{f'}\}=\mathbb{C}\mathrm{ov}\{\boldsymbol{a}_\circ,\boldsymbol{q}_\circ-\boldsymbol{B}_\circ\boldsymbol{f'}\}$ and $\mathbb{V}\mathrm{ar}\{\boldsymbol{a}_\circ+\epsilon\boldsymbol{1}\}=\mathbb{V}\mathrm{ar}\{\boldsymbol{a}_\circ\}$ to obtain

$$\begin{aligned}2\,\epsilon\,\boldsymbol{1}^\top\overline{\boldsymbol{q}_\circ}\geqslant\;&2\,\epsilon\,\boldsymbol{1}^\top(\overline{\boldsymbol{B}_\circ\boldsymbol{f'}})-\epsilon^2\boldsymbol{1}^\top\boldsymbol{1}\\&+2\,\overline{\mathbb{C}\mathrm{ov}\{\boldsymbol{a}_\circ,\boldsymbol{B}_\circ\boldsymbol{f'}\}}-2\,\overline{\mathbb{C}\mathrm{ov}\{\boldsymbol{a}_\circ,\boldsymbol{B}_\circ\boldsymbol{f}\}}.\end{aligned}$$

Since $\boldsymbol{f}$ produced by the generic routing policy is independent of arrival statistics, we get $\mathbb{C}\mathrm{ov}\{\boldsymbol{a}_\circ,\boldsymbol{B}_\circ\boldsymbol{f}\}=0$. Replacing $\boldsymbol{1}^\top\overline{\boldsymbol{q}_\circ}$ by the $\overline{Q}$ expression as defined in (2), we then obtain

$$2\,\epsilon\,\overline{Q}\;\geqslant\;2\,\epsilon\,\boldsymbol{1}^\top(\overline{\boldsymbol{B}_\circ\boldsymbol{f'}})+2\,\overline{\mathbb{C}\mathrm{ov}\{\boldsymbol{a}_\circ,\boldsymbol{B}_\circ\boldsymbol{f'}\}}-\epsilon^2\boldsymbol{1}^\top\boldsymbol{1}\,.\quad\text{(A13)}$$

Consider this time HD policy at $\beta=0$ with the timeslot quantities of $\boldsymbol{q}^\star_\circ(n)$ and $\boldsymbol{f}^\star(n)$. Let again $\boldsymbol{f'}(n)$ be produced by the fictitious algorithm at each slot n to stabilize the arrival $\overline{\boldsymbol{a}_\circ}+\epsilon\boldsymbol{1}$, but this time, given $\boldsymbol{q}^\star_\circ(n)$. In light of Lem. 4, $G(\boldsymbol{f'},\boldsymbol{q}^\star_\circ,n)\leqslant G(\boldsymbol{f}^\star,\boldsymbol{q}^\star_\circ,n)$ at each slot n. Performing the similar steps of taking expectation, exploiting $\mathbb{E}\{\boldsymbol{B}_\circ\boldsymbol{f'}\}=\mathbb{E}\{\boldsymbol{B}_\circ\boldsymbol{f}^\star\}+\epsilon$, translating the results into the expected time average form, using the fact that $\mathbb{C}\mathrm{ov}\{\boldsymbol{a}_\circ,\boldsymbol{B}_\circ\boldsymbol{f}^\star\}=0$ as $\boldsymbol{f}^\star$ is independent of arrival statistics, and applying Lem. 5 by knowing that HD policy is throughput optimal and so stabilizes $\overline{\boldsymbol{a}_\circ}$, we obtain

$$2\,\epsilon\,\overline{Q^\star}\;\leqslant\;2\,\epsilon\,\boldsymbol{1}^\top(\overline{\boldsymbol{B}_\circ\boldsymbol{f'}})+2\,\overline{\mathbb{C}\mathrm{ov}\{\boldsymbol{a}_\circ,\boldsymbol{B}_\circ\boldsymbol{f'}\}}-\epsilon^2\boldsymbol{1}^\top\boldsymbol{1}\,.\quad\text{(A14)}$$

Comparing (A13) and (A14) along with $\epsilon>0$ lead to $\overline{Q^\star}\leqslant\overline{Q}$. This means the average network delay under HD policy with $\beta=0$ remains less than or equal to that under any other $\mathcal{D}$-class routing policy, which was called "generic" here. ∎

Proof of Theorem 4 (HD Fluid Model)

The proof follows the exact same line of argument proposed in [36, Theorem 2.3.2] and [37, Proposition 4.12]. ∎

Proof of Theorem 5 (Wireless Network Thermodynamics)

Let $\boldsymbol{q}^\star_\circ(t)$ and $\boldsymbol{f}^\star(t)$ denote the HD fluid model variables. Consider the continuous-time Lyapunov function

$$Y(t):=\big(\boldsymbol{q}^\star_\circ(t)-\boldsymbol{q}^{\mathrm{opt}}_\circ\big)^\top\,\overline{\boldsymbol{M}}_\circ\,\big(\boldsymbol{q}^\star_\circ(t)-\boldsymbol{q}^{\mathrm{opt}}_\circ\big)$$

where $\overline{\boldsymbol{M}}_\circ=\big(\boldsymbol{B}_\circ\boldsymbol{B}_\circ^\top\big)^{-1}\boldsymbol{B}_\circ\,\overline{\boldsymbol{\Phi}}\,\boldsymbol{B}_\circ^\top$ represents the time average expectation of matrix $\boldsymbol{M}_\circ(n)$ as defined in (18). Taking time derivative from $Y(t)$, we obtain

$$\dot{Y}(t)=\dot{\boldsymbol{q}}^\star_\circ(t)^\top\big(\,\overline{\boldsymbol{M}_\circ}^\top+\overline{\boldsymbol{M}_\circ}\,\big)\big(\boldsymbol{q}^\star_\circ(t)-\boldsymbol{q}^{\mathrm{opt}}_\circ\big).$$

Exploiting Lem. 3 in the latter leads to

$$\dot{Y}(t)\leqslant\eta\;\dot{\boldsymbol{q}}^\star_\circ(t)^\top\,\overline{\boldsymbol{M}_\circ}\,\big(\boldsymbol{q}^\star_\circ(t)-\boldsymbol{q}^{\mathrm{opt}}_\circ\big)\qquad\text{(A15)}$$

for an $1\leqslant\eta\leqslant3$. As a positive coefficient, η has no impact on the Lyapunov argument and can simply be omitted, but for the sake of consistency we prefer to keep it in here.

To find an appropriate expression for $\dot{\boldsymbol{q}}^\star_\circ(t)$, let us begin by plugging (39) in (38) and taking time derivative to obtain

$$\dot{\boldsymbol{q}}^\star_\circ(t)=\overline{\boldsymbol{a}_\circ}-\boldsymbol{B}_\circ\dot{\boldsymbol{f}}^{\star\mathrm{tot}}(t)\,.\qquad\text{(A16)}$$

Note in (43) that the entry of $\widehat{\boldsymbol{f}(t)}$ corresponding to link ij specifies the number of packets the link will send *per unit time* if it is activated at time t. Then $\boldsymbol{f}(t)$ identifies the vector of rate of actual transmissions realized at time t. Assume now that the entry of $\boldsymbol{f}(t)$ corresponding to link ij at time t is equal to $x\geqslant0$, i.e., at time t the link transmits x number of packets per unit time. Then it should be obvious that the same entry of $\dot{\boldsymbol{f}}^{\mathrm{tot}}(t)$ at time t must also be equal to x. In light of $\lim_{\delta\to0}\boldsymbol{f}^{\mathrm{tot}}(t+\delta)=\boldsymbol{f}^{\mathrm{tot}}(t)+\delta\boldsymbol{f}(t)$, this can be explained more formally by the classical definition of limit as

$$\dot{\boldsymbol{f}}^{\mathrm{tot}}(t)=\lim_{\delta\to0}\frac{\boldsymbol{f}^{\mathrm{tot}}(t+\delta)-\boldsymbol{f}^{\mathrm{tot}}(t)}{\delta}=\boldsymbol{f}(t)\,.$$

Further, (48)–(49) imply $\overline{\boldsymbol{a}_\circ} = \vec{\boldsymbol{L}}_\circ^{\,\mathrm{opt}} \boldsymbol{q}_\circ^{\mathrm{opt}} = \boldsymbol{B}_\circ \boldsymbol{f}^{\mathrm{opt}}$. Exploiting these latter identities in (A16) yields

$$\dot{\boldsymbol{q}}_\circ^\star(t) = \boldsymbol{B}_\circ \boldsymbol{f}^{\mathrm{opt}} - \boldsymbol{B}_\circ \boldsymbol{f}^\star(t)\,. \tag{A17}$$

Returning to the Lyapunov argument, let us substitute (A17) in (A15) and utilize equality (19) in Lem. 2 to obtain

$$\eta^{-1}\,\dot{Y}(t) \leqslant \big(\boldsymbol{f}^{\mathrm{opt}} - \boldsymbol{f}^\star(t)\big)^\top \overline{\boldsymbol{\Phi}}\, \boldsymbol{B}_\circ^\top \big(\boldsymbol{q}_\circ^\star(t) - \boldsymbol{q}_\circ^{\mathrm{opt}}\big)\,.$$

Multiplying both sides by two, adding and subtracting the term $\boldsymbol{f}^\star(t)^\top \boldsymbol{f}^\star(t) + \boldsymbol{f}^{\mathrm{opt}\top} \boldsymbol{f}^{\mathrm{opt}}$ on the left-hand side, and recasting the terms lead to

$$\begin{aligned}
2\,\eta^{-1}\,\dot{Y}(t) \leqslant &-\big(2\,\boldsymbol{f}^\star(t)^\top\, \overline{\boldsymbol{\Phi}}\, \boldsymbol{B}_\circ^\top \boldsymbol{q}_\circ^\star(t) - \boldsymbol{f}^\star(t)^\top \boldsymbol{f}^\star(t)\big) && \text{(A18a)}\\
&+ \big(2\,\boldsymbol{f}^{\mathrm{opt}\top}\, \overline{\boldsymbol{\Phi}}\, \boldsymbol{B}_\circ^\top \boldsymbol{q}_\circ^\star(t) - \boldsymbol{f}^{\mathrm{opt}\top} \boldsymbol{f}^{\mathrm{opt}}\big) && \text{(A18b)}\\
&- \big(2\,\boldsymbol{f}^{\mathrm{opt}\top}\, \overline{\boldsymbol{\Phi}}\, \boldsymbol{B}_\circ^\top \boldsymbol{q}_\circ^{\mathrm{opt}} - \boldsymbol{f}^{\mathrm{opt}\top} \boldsymbol{f}^{\mathrm{opt}}\big) && \text{(A18c)}\\
&+ \big(2\,\boldsymbol{f}^\star(t)^\top\, \overline{\boldsymbol{\Phi}}\, \boldsymbol{B}_\circ^\top \boldsymbol{q}_\circ^{\mathrm{opt}} - \boldsymbol{f}^\star(t)^\top \boldsymbol{f}^\star(t)\big). && \text{(A18d)}
\end{aligned}$$

Characterizing (A18a) and (A18b) on the wireless network given $\boldsymbol{q}_\circ^\star(t)$, they respectively read $-D(\boldsymbol{f}^\star, \boldsymbol{q}_\circ^\star, t)$ and $D(\boldsymbol{f}^{\mathrm{opt}}, \boldsymbol{q}_\circ^\star, t)$. Under Assum. 1 and in light of the HD fluid equations (43) and (45), the immediate result of Th. 4 is that at each time t, HD fluid limit maximizes the D functional compared to any alternative forwarding that satisfies wireless network constraints. The $\boldsymbol{f}^{\mathrm{opt}}$ obviously meets the directionality constraints due to the structure of the reference thermal model. It also meets the capacity constraints due to Assum. 2. Hence, (A18a) + (A18b) $\leqslant 0$ which leads to

$$\begin{aligned}
2\,\eta^{-1}\,\dot{Y}(t) \leqslant &-\big(2\,\boldsymbol{f}^{\mathrm{opt}\top}\, \overline{\boldsymbol{\Phi}}\, \boldsymbol{B}_\circ^\top \boldsymbol{q}_\circ^{\mathrm{opt}} - \boldsymbol{f}^{\mathrm{opt}\top} \boldsymbol{f}^{\mathrm{opt}}\big) && \text{(A19a)}\\
&+ \big(2\,\boldsymbol{f}^\star(t)^\top\, \overline{\boldsymbol{\Phi}}\, \boldsymbol{B}_\circ^\top \boldsymbol{q}_\circ^{\mathrm{opt}} - \boldsymbol{f}^\star(t)^\top \boldsymbol{f}^\star(t)\big). && \text{(A19b)}
\end{aligned}$$

We now characterize (A19a) and (A19b) on the reference thermal model and let

$$H(\boldsymbol{f}) := 2\,\boldsymbol{f}^\top\, \overline{\boldsymbol{\Phi}}\, \boldsymbol{B}_\circ^\top \boldsymbol{q}_\circ^{\mathrm{opt}} - \boldsymbol{f}^\top \boldsymbol{f}\,.$$

One can show that given $\boldsymbol{q}_\circ^{\mathrm{opt}}$, the maximum of H occurs at $\boldsymbol{f} = \boldsymbol{f}^{\mathrm{opt}}$ produced by heat flow. To this end, rephrase H as

$$H(\boldsymbol{f}) = \sum\nolimits_{ij\in\mathcal{E}} 2\,\overline{\phi_{ij}}\, q_{ij}^{\mathrm{opt}} f_{ij} - \big(f_{ij}\big)^2$$

where directionality constraints entail $f_{ij} \geqslant 0$. To maximize H, one then needs to assign $f_{ij} = 0$ if $q_{ij}^{\mathrm{opt}} \leqslant 0$, and $f_{ij} = \overline{\phi_{ij}}\, q_{ij}^{\mathrm{opt}}$ otherwise. Putting this back in a matrix form, we arrive at the same expression as $\boldsymbol{f}^{\mathrm{opt}}$ in (48). Further, the maximizing $\boldsymbol{f}$ is unique by the reason that a given $\boldsymbol{q}_\circ^{\mathrm{opt}}$ leads to a unique $\boldsymbol{B}_\circ^\top \boldsymbol{q}_\circ^{\mathrm{opt}}$, and so to unique q_{ij}^{opt} components, as the matrix $\boldsymbol{B}_\circ$ has full row rank. From $H(\boldsymbol{f}^{\mathrm{opt}}) \geqslant H(\boldsymbol{f}^\star(t))$, we then obtain (A19a) + (A19b) $\leqslant 0$, which by $1 \leqslant \eta \leqslant 3$ yields $\dot{Y}(t) \leqslant 0$.

Let Ω be the largest invariant set in the set of all $\boldsymbol{q}_\circ^\star(t)$ trajectories for which $\dot{Y}(t) = 0$. Since $Y(t)$ is a nonnegative and radially unbounded function with $\dot{Y}(t) \leqslant 0$, LaSalle's invariance principle states that every trajectory $\boldsymbol{q}_\circ^\star(t)$ asymptotically converges to Ω. It remains to show that Ω contains only the trivial trajectory of $\boldsymbol{q}_\circ^\star = \boldsymbol{q}_\circ^{\mathrm{opt}}$. If $\dot{Y} = 0$, then (2) entails $H(\boldsymbol{f}^\star(t)) = H(\boldsymbol{f}^{\mathrm{opt}})$. We previously showed as well that $\boldsymbol{f}^{\mathrm{opt}}$ maximizes H and is unique, which implies

$$\boldsymbol{f}^\star = \boldsymbol{f}^{\mathrm{opt}}. \tag{A20}$$

The intentionally dropped time variable (t) in (A20) emphasizes that $\boldsymbol{f}^\star(t)$ turns to be stationary by converging to $\boldsymbol{f}^{\mathrm{opt}}$, meaning that $\boldsymbol{q}^\star(t)$ converges to a stationary $\boldsymbol{q}^\star$ too.

Given $\boldsymbol{q}^\star$, the equality (A20) entails that $\boldsymbol{f}^{\mathrm{opt}}$ must maximize D, which implies $f_{ij}^{\mathrm{opt}} = 0$ if $q_{ij}^\star \leqslant 0$, and $f_{ij}^{\mathrm{opt}} = \overline{\phi_{ij}}\, q_{ij}^{\mathrm{opt}}$ otherwise. In a matrix form, this is equivalent to $\boldsymbol{f}^{\mathrm{opt}} = \overline{\boldsymbol{\Phi}}\, \max\big\{\boldsymbol{0}, \boldsymbol{B}_\circ^\top \boldsymbol{q}_\circ^\star\big\}$. Putting the latter against the equation (48) leads to

$$\max\big\{\boldsymbol{0},\, \boldsymbol{B}_\circ^\top \boldsymbol{q}_\circ^\star\big\} = \max\big\{\boldsymbol{0},\, \boldsymbol{B}_\circ^\top \boldsymbol{q}_\circ^{\mathrm{opt}}\big\}. \tag{A21}$$

Consider a directed edge ad with its head at the destination node, which has zero queue on the wireless network and zero temperature on the reference thermal model. By (A21), $q_a^\star$ and $\boldsymbol{q}_a^{\mathrm{opt}}$ must be equal. Repeating this argument eventually yields $(\boldsymbol{q}_\circ^\star)^+ = (\boldsymbol{q}_\circ^{\mathrm{opt}})^+$, as any node with positive queue (resp. positive temperature) on the wireless network (resp. on the reference thermal model) must be connected to the destination node d through a directed path. Further, observe that $\boldsymbol{q}_\circ^\star \succcurlyeq 0$, as queues cannot be negative in a wireless network, and $\boldsymbol{q}_\circ^{\mathrm{opt}} \succcurlyeq 0$, as temperatures cannot fall below zero in a thermal system with no negative heat source. Thus, $\boldsymbol{q}_\circ^\star = \boldsymbol{q}_\circ^{\mathrm{opt}}$, which together with (A20) conclude the proof. ∎

Proof of Theorem 6 (Nonlinear Dirichlet Principle)

One can verify, by the $\vec{\boldsymbol{L}}_\circ$ structure in (35), that

$$\vec{E}_D(\boldsymbol{q}_\circ) = \frac{1}{2}\,\big(\boldsymbol{q}_\circ^\top \boldsymbol{B}_\circ\big)^+ \operatorname{diag}(\boldsymbol{\sigma})\big(\boldsymbol{B}_\circ^\top \boldsymbol{q}_\circ\big)^+ - \boldsymbol{q}_\circ^\top \boldsymbol{a}_\circ$$

where each entry of $\boldsymbol{B}_\circ^\top \boldsymbol{q}_\circ$ represents temperature-difference along the corresponding edge. Let $\boldsymbol{q}_\circ^*$ be the $\vec{E}_D(\boldsymbol{q}_\circ)$ minimizing solution and let us rearrange and partition $\boldsymbol{B}_\circ^\top \boldsymbol{q}_\circ^*$ into positive, zero and negative components. Accordingly, $\boldsymbol{B}_\circ$ gets partitioned into $\boldsymbol{B}_\oplus$, $\boldsymbol{B}_\varnothing$ and $\boldsymbol{B}_\ominus$, which respectively contain the incidence information of edges with positive, zero and negative values in $\boldsymbol{B}_\circ^\top \boldsymbol{q}_\circ^*$. Likewise, $\boldsymbol{\sigma}$ gets partitioned into $\boldsymbol{\sigma}_\oplus$, $\boldsymbol{\sigma}_\varnothing$ and $\boldsymbol{\sigma}_\ominus$. Then at $\boldsymbol{q}_\circ = \boldsymbol{q}_\circ^*$, we obtain

$$\begin{aligned}
\vec{E}_D(\boldsymbol{q}_\circ^*) = &-\boldsymbol{q}_\circ^{*\top} \boldsymbol{a}_\circ + \frac{1}{2}\,\big(\boldsymbol{q}_\circ^{*\top} \boldsymbol{B}_\oplus\big)^+ \operatorname{diag}(\boldsymbol{\sigma}_\oplus)\big(\boldsymbol{B}_\oplus^\top \boldsymbol{q}_\circ^*\big)^+\\
&+ \frac{1}{2}\,\big(\boldsymbol{q}_\circ^{*\top} \boldsymbol{B}_\varnothing\big)^+ \operatorname{diag}(\boldsymbol{\sigma}_\varnothing)\big(\boldsymbol{B}_\varnothing^\top \boldsymbol{q}_\circ^*\big)^+ && \text{(A22a)}\\
&+ \frac{1}{2}\,\big(\boldsymbol{q}_\circ^{*\top} \boldsymbol{B}_\ominus\big)^+ \operatorname{diag}(\boldsymbol{\sigma}_\ominus)\big(\boldsymbol{B}_\ominus^\top \boldsymbol{q}_\circ^*\big)^+. && \text{(A22b)}
\end{aligned}$$

Observe that (A22a) is strongly zero due to the $(\cdot)^+$ operation. On the other hand, (A22b) vanishes since $\boldsymbol{B}_\varnothing^\top \boldsymbol{q}_\circ^* = \boldsymbol{0}$. In light of $(\boldsymbol{B}_\oplus^\top \boldsymbol{q}_\circ^*)^+ = \boldsymbol{B}_\oplus^\top \boldsymbol{q}_\circ^*$, we then obtain

$$\vec{E}_D(\boldsymbol{q}_\circ^*) = \frac{1}{2}\,\boldsymbol{q}_\circ^{*\top} \boldsymbol{B}_\oplus \operatorname{diag}(\boldsymbol{\sigma}_\oplus) \boldsymbol{B}_\oplus^\top \boldsymbol{q}_\circ^* - \boldsymbol{q}_\circ^{*\top} \boldsymbol{a}_\circ\,. \tag{A23}$$

Since $\boldsymbol{a}_\circ$ is feasible, each nonzero heat source connects to the sink through at least one directed path. Thus, under any flow that keeps $\boldsymbol{q}_\circ$ entrywise finite, the edges with positive temperature-difference build a connected graph with the node d. On the other hand, $\boldsymbol{q}_\circ^*$ is entrywise finite as it minimizes $\vec{E}_D(\boldsymbol{q}_\circ)$, and so the corresponding edges in $\boldsymbol{B}_\oplus$ build a connected graph with the node d. This implies that $\boldsymbol{B}_\oplus \operatorname{diag}(\boldsymbol{\sigma}_\oplus) \boldsymbol{B}_\oplus^\top$ is a positive definite matrix. Thus, the functional $\frac{1}{2}\,\boldsymbol{q}_\circ^\top \boldsymbol{B}_\oplus \operatorname{diag}(\boldsymbol{\sigma}_\oplus) \boldsymbol{B}_\oplus^\top \boldsymbol{q}_\circ - \boldsymbol{q}_\circ^\top \boldsymbol{a}_\circ$ is strictly convex in $\boldsymbol{q}_\circ$ and so finds its minimum at the critical point, where its first order variation with respect to $\boldsymbol{q}_\circ$ vanishes. Comparing this with (A23), it turns out that the minimizing $\boldsymbol{q}_\circ^*$ must satisfy

$$\boldsymbol{a}_\circ = \boldsymbol{B}_\oplus \operatorname{diag}(\boldsymbol{\sigma}_\oplus)\, \boldsymbol{B}_\oplus^\top \boldsymbol{q}_\circ^*\,. \tag{A24}$$

Utilizing $\boldsymbol{B}_{\oplus}^{\top}\boldsymbol{q}_{\circ}^{*}=(\boldsymbol{B}_{\oplus}^{\top}\boldsymbol{q}_{\circ}^{*})^{+}$ and $(\boldsymbol{B}_{\ominus}^{\top}\boldsymbol{q}_{\circ}^{*})^{+}=(\boldsymbol{B}_{\varnothing}^{\top}\boldsymbol{q}_{\circ}^{*})^{+}=\boldsymbol{0}$, one can rephrase (A24) as

$$\boldsymbol{a}_{\circ}=\boldsymbol{B}_{\circ}\,\mathrm{diag}(\boldsymbol{\sigma})\big(\boldsymbol{B}_{\circ}^{\top}\boldsymbol{q}_{\circ}^{*}\big)^{+}=\vec{\boldsymbol{L}}_{\circ}\boldsymbol{q}_{\circ}^{*}$$

that recovers the nonlinear Poisson equation (52) at $\boldsymbol{q}_{\circ}^{*}$. Further, $\boldsymbol{q}_{\circ}^{*}$ is unique as it minimizes the strictly convex functional $\frac{1}{2}\,\boldsymbol{q}_{\circ}^{\top}\boldsymbol{B}_{\oplus}\,\mathrm{diag}(\boldsymbol{\sigma}_{\oplus})\boldsymbol{B}_{\oplus}^{\top}\boldsymbol{q}_{\circ}-\boldsymbol{q}_{\circ}^{\top}\boldsymbol{a}_{\circ}$. ■

Proof of Theorem 7 (Nonlinear Thomson Principle)

Consider (54) as the primal optimization problem and let us construct its Lagrangian dual problem as

$$\max_{\boldsymbol{\lambda}}\ \min_{\boldsymbol{f}\succcurlyeq\boldsymbol{0}}\Big(\mathcal{L}(\boldsymbol{\lambda},\boldsymbol{f}):=\boldsymbol{f}^{\top}\,\mathrm{diag}(\boldsymbol{\sigma})^{-1}\boldsymbol{f}+2\,\boldsymbol{\lambda}^{\top}\big(\boldsymbol{a}_{\circ}-\boldsymbol{B}_{\circ}\boldsymbol{f}\big)\Big)$$

where $\boldsymbol{\lambda}\succcurlyeq 0$ is the vector of Lagrange multipliers. From the first order condition $\nabla_{\boldsymbol{f}}\,\mathcal{L}=\boldsymbol{0}$, we get $\boldsymbol{f}^{\mathrm{opt}}=\mathrm{diag}(\boldsymbol{\sigma})\,\boldsymbol{B}_{\circ}^{\top}\boldsymbol{\lambda}$. Then enforcing the constraint $\boldsymbol{f}^{\mathrm{opt}}\succcurlyeq\boldsymbol{0}$ leads to $\boldsymbol{B}_{\circ}^{\top}\boldsymbol{\lambda}\succcurlyeq\boldsymbol{0}$, which is equivalent to $\boldsymbol{B}_{\circ}^{\top}\boldsymbol{\lambda}=(\boldsymbol{B}_{\circ}^{\top}\boldsymbol{\lambda})^{+}$. Thus, we obtain

$$\boldsymbol{f}^{\mathrm{opt}}=\mathrm{diag}(\boldsymbol{\sigma})\,(\boldsymbol{B}_{\circ}^{\top}\boldsymbol{\lambda})^{+}. \tag{A25}$$

Plugging this $\boldsymbol{f}^{\mathrm{opt}}$ into the Lagrangian $\mathcal{L}$ and utilizing the structure of $\vec{\boldsymbol{L}}_{\circ}$ in (35), we obtain

$$\mathcal{L}(\boldsymbol{\lambda})=-\boldsymbol{\lambda}^{\top}\vec{\boldsymbol{L}}_{\circ}\,\boldsymbol{\lambda}+2\,\boldsymbol{\lambda}^{\top}\boldsymbol{a}_{\circ}\,.$$

Then the dual problem reads $\max_{\boldsymbol{\lambda}}\,\mathcal{L}(\boldsymbol{\lambda})$, which is equivalent to the following minimization problem:

$$\min_{\boldsymbol{\lambda}}\ \frac{1}{2}\,\boldsymbol{\lambda}^{\top}\vec{\boldsymbol{L}}_{\circ}\,\boldsymbol{\lambda}-\boldsymbol{\lambda}^{\top}\boldsymbol{a}_{\circ}\,. \tag{A26}$$

Further, as $\boldsymbol{f}\succcurlyeq\boldsymbol{0}$ makes a convex set and $\mathcal{L}(\boldsymbol{\lambda},\boldsymbol{f})$ is a convex function, the duality gap is zero, and so both the primal and dual problems result in the same optimal solution.

Comparing (A26) with the nonlinear Dirichlet equation (53), it remains to show that the Lagrangian multipliers $\boldsymbol{\lambda}$ are identical to the node temperatures $\boldsymbol{q}_{\circ}$. In (A25), multiplying both sides by $\boldsymbol{B}_{\circ}$ and using the $\vec{\boldsymbol{L}}_{\circ}$ expression, we obtain

$$\vec{\boldsymbol{L}}_{\circ}\boldsymbol{\lambda}=\boldsymbol{B}_{\circ}\boldsymbol{f}^{\mathrm{opt}}=\boldsymbol{a}_{\circ} \tag{A27}$$

where the second equality comes from the constraint in the primal problem (54). Further, by Th. 6, the nonlinear Poisson equation $\vec{\boldsymbol{L}}_{\circ}\boldsymbol{q}_{\circ}=\boldsymbol{a}_{\circ}$ has a unique solution. Putting this against (A27) leads to $\boldsymbol{\lambda}=\boldsymbol{q}_{\circ}$, which concludes the proof. ■

Proof of Theorem 8 (HD Minimum Routing Cost)

It was shown by Th. 6 that if $\overline{\boldsymbol{a}_{\circ}}$ is feasible, then under the nonlinear heat equations (48)–(49), the stationary value of the nonlinear Dirichlet energy $\vec{E}_D(\boldsymbol{q}_{\circ})$ is strictly minimized. It was shown by Th. 7, on the other hand, that minimizing $\vec{E}_D(\boldsymbol{q}_{\circ})$ is equivalent to minimizing the stationary value of total energy dissipation $\vec{E}_R(\boldsymbol{f})$ on the graph. Then the proof immediately follows from Th. 5 which states that under a stabilizable arrival rate $\overline{\boldsymbol{a}_{\circ}}$, HD fluid model complies with the nonlinear heat equations (48)–(49). Note that if $\overline{\boldsymbol{a}_{\circ}}$ is stabilizable, i.e., it satisfies condition (17), then its feasibility is trivial in the sense of Def. 10. ■

Proof of Theorem 9 (HD Pareto Optimality)

Observe that HD policy minimizes $\overline{Q}$ at $\beta=0$, minimizes $\overline{R}$ at $\beta=1$, and changes weight on these two objectives by altering β between 0 and 1. In fact, HD transforms the two objectives of minimizing $\overline{Q}$ and $\overline{R}$ into an aggregated objective function by multiplying each objective function by a weighting factor and summing up the two weighted objective functions. Further, the weighted sum is a convex combination of objectives as the sum of weighting factors β and $1-\beta$ equals 1. Under the assumption that the region $(\overline{Q}\,,\overline{R})$ has a convex Pareto boundary, the proof then follows by the fact that the entire boundary can be reached using the weighted-sum method [38]–[40] that changes the weight on the weighted convex combination of the two objective functions. ■

APPENDIX B
REAL EXPERIMENT RESULTS AND ANALYSIS

To analyze the performance of HD in a real sensor network, we have implemented the HD algorithm, referred to as Heat-Diffusion Collection Protocol (HDCP) here, and the V-parameter BP algorithm, referred to as Backpressure Collection Protocol (BCP) here, on Contiki OS and used the CTP implementation available with the Contiki OS. We perform a set of evaluation experiments on an indoor wireless network testbed called Tutornet [41] with forty five IEEE 802.15.4-based Tmote-sky nodes distributed over a floor with roughly $80,000$ sq.ft of area. The network topology is presented in Fig. A1 where the marked node is the sink, the rest of the nodes are the source nodes, and the furthest node is three hops away from the sink. We use the channel number 26 with Tmote sky power level 31 for this experiment. The number of neighbors to each node varies from 19 to 35 with an average of 29. However, only about 7-8 of the neighbors are typically connected via good links ($ETX\approx 1$). Thus, the topology is very diverse with a considerable number of different paths between any two nodes in the network, which well represents a multihop wireless network. On the negative side, a considerable amount of interference exists among the nodes, limiting the bandwidth, which well represents a time-varying wireless network. The data packets in our experiments are all 26 Bytes in size.

All the experiments are performed on weekdays during daytime with lots of moving people and physical objects around. Each experiment is performed for 35 min: the network settles down during the first 5 min and the data is collected

Fig. A1. Real experiment testbed with forty five IEEE 802.15.4-based Tmote-sky nodes distributed over a floor with roughly $80,000$ sq.ft of area.

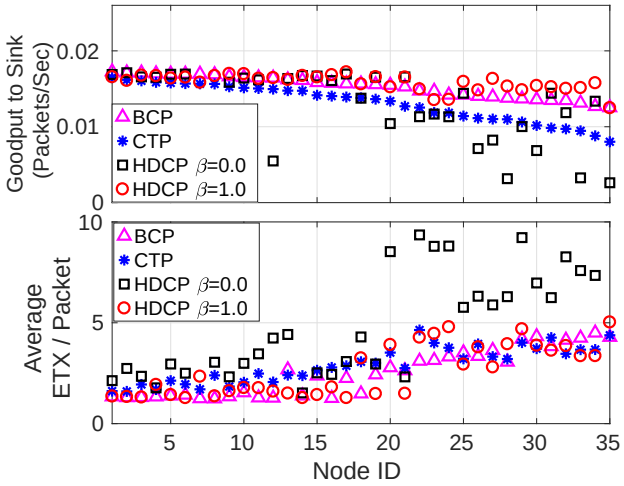

Fig. A2. Performance of HDCP versus BCP and CTP for a low power communication stack: (Top) Goodput to the sink for each node, (Bottom) Average ETX path costs to the sink for each node.

during the next 30 min. Each experiment is repeated at least 10 times to improve the confidence levels.

The goodput of each source node is defined as the number of packets received by the sink over a one second interval. Note the direct correlation between cumulative goodput of all nodes and network throughput. The end to end delay calculation for each packet is performed by adding up all the queuing and processing delays in all intermediate nodes, which can be determined by locally time stamping packets during reception and departure. This ignores the propagation times which in any case are negligible compared to the processing delays.

We evaluate HDCP performance in terms of different values of β and different packet generation rates. We select the value of V to be 2 for BCP, which has empirically shown to be an efficient operating point for Tutornet [10].

Low Power Communication Stack Based Experiments: In order to verify the performance of HDCP on a low power communication stack, we performed a set of experiments with 35 sources and one sink (first 36 nodes of the testbed). For these experiments, we used CX-MAC protocol, a version of X-MAC that is provided in Contiki, with duty cycle of 5% for HDCP, BCP and CTP. However, the choice of CX-MAC protocol over the other protocols is just a matter of the availability of Contiki implementation. Further, since we are using a duty cycle, we also need to cut-back our source rates to a very low rate. For the presented set of experiments, we used a packet generation rate of 1 packet per 60 seconds (i.e., 1/60 PPS). The results in Fig. A2 show that HDCP with $\beta = 1$ performs well in a low power communication stack, at a very low duty cycle setting where even CTP shows some deterioration in fairness of goodput.

In order to estimate the actual energy consumptions, we record the different components of energy consumption using the Contiki PowerTrace tool (in terms of the percentage of time spent in different radio phases: transmit, listen/receive). Based on our traces, in HDCP with 5% duty cycle, the radio of each node is on for 5.92% of the total execution time, out of which the node is transmitting and receiving approximately 0.65% and 5.27% of the total execution time, respectively. To get the actual energy consumption, one can use the current and voltage ratings from the specifications of the devices used in the experiment. For example, in Tmote-sky, the rated voltage of operation is approx $3.3V$ and the average current consumptions are $17.4mA$ and $19.7mA$ for radio transmission and radio reception, respectively. This results in approximately $113.78mJ$ energy consumption in each Tmote-Sky for the experiment period of 30 minutes.

External Interference: We evaluate the performance of HDCP with the optimized $\beta = 1$ in the presence of external interference and compare it with both BCP and CTP. This is necessary because the 802.15.4 radios share frequency band with WiFi, Bluetooth and other Zigbee radios, which as a result their performance often suffers from severe interference. To emulate such scenarios, we performed a set of experiments with 40 sources and a single sink (Node 1) while four nodes are used as interference sources on Channel 26. (All nodes are transmitting in Channel 26 but on different Contiki channels.) The interfering nodes are inactive for the first five minutes of the experiment, periodically transmit for next 15 minutes, and become inactive again for the last five minutes of the experiment. During the on period, each of the interfering nodes transmits 110 Byte packets at a rate of 100PPS for 15 seconds and then does not transmit anything for the next 15 seconds, and so on. Furthermore, we reduced the power level of all 41 nodes from level 31 to level 15 whereas the interfering

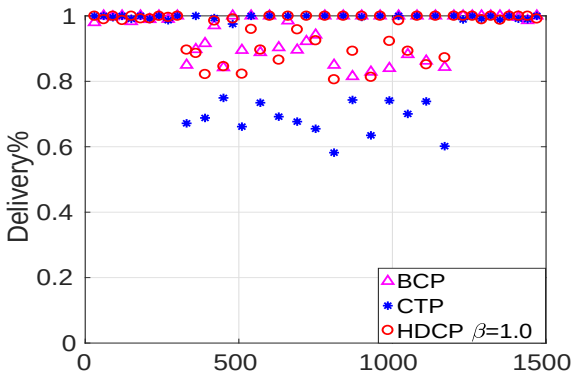

Fig. A3. Thirty second windowed average sourced packet delivery ratio with synthetically generated interfering 802.15.4 Channel 26 traffic.

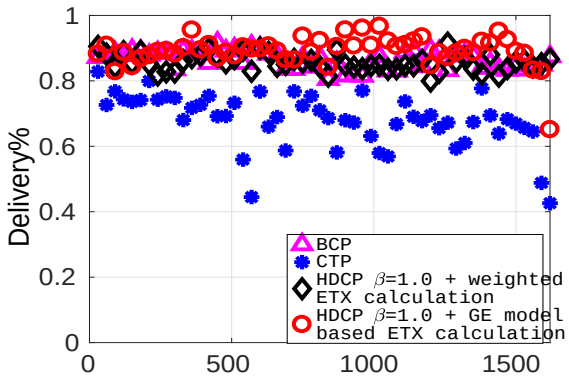

Fig. A4. Thirty second windowed average sourced packet delivery ratio with real interference scenario on 802.15.4 Channel 13.

nodes were kept at level 31, which aims to intensify the effect of interference. The outcome of this set of experiments is presented in Fig. A3 that plots the delivery percentage of the packets over a series of 30 seconds time window for CTP, BCP with $V = 2$ and HDCP with $\beta = 1$. While CTP performance significantly suffers from the interference, both HDCP and BCP maintain their efficient ratio of packet delivery.

The above interference settings are used to stay consistent with those in the original BCP paper [10]. However, it is known that the simple Gilbert-Eliot model used for ETX estimation might work perfectly with some specific synthetic interference models, while failing in real interference scenarios. In order to explore the performance of HDCP algorithm in presence of real interference, we performed a set of experiments with 44 source nodes and one sink node, running on Channel 13 of the 802.15.4 standard, known as one of the most interfered channels. We also compare the performance of HDCP based on the Gilbert-Eliot (GE) ETX model with the performance of BCP with the GE model, as well as with the HDCP configured by the ETX model used in [10].

For this set of experiments, we do not use the link switching method as we empirically found that interference itself causes sufficient link switching, thereby, adding extra link switching negatively affects performance. The results in Fig. A4 demonstrate that even in the presence of constant real interference, the HDCP algorithm with Gilbert-Eliot ETX model performs comparable to the BCP and the HDCP algorithms with the basic ETX model [10], while outperforms the CTP algorithm. Further, both figures show that HDCP and BCP can achieve approx 85% delivery ratio in the presence of interference, while CTP achieves approx 70%.

APPENDIX C
GRAPH LAPLACIAN

Consider a *connected,* weighted graph with set of nodes $\mathcal{V}$, set of undirected edges $\mathcal{E}$, node-edge incidence matrix $\boldsymbol{B}$, edge weight vector $\boldsymbol{\sigma}$ and the Laplacian $\boldsymbol{L} := \boldsymbol{B}\,\mathrm{diag}(\boldsymbol{\sigma})\,\boldsymbol{B}^\top$. One can verify that $\boldsymbol{x}^\top \boldsymbol{L}\boldsymbol{x} = \sum_{ij\in\mathcal{E}} \sigma_{ij}(x_i - x_j)^2 \geqslant 0$, which means $\boldsymbol{L}$ is *positive semi-definite.*

In vector calculus, the gradient of a scalar field is positive in the direction of increase of the field. On a graph, on the other hand, we take the gradient of a node variable positive in the direction of decrease of the variable. By the same reason, the classical Laplace operator is a negative semi-definite operator, while the graph Laplacian is a positive semi-definite matrix.

Observe that $\boldsymbol{L}\mathbf{1} = \mathbf{0}$, which implies that $\mathbf{1}$ is an eigenvector corresponding to the smallest eigenvalue $\lambda_1 = 0$. For the second smallest eigenvalue λ_2, let $\boldsymbol{\nu}$ be the eigenvector orthogonal to $\mathbf{1}$. Thus, $\boldsymbol{\nu}^\top\mathbf{1} = 0$ and $\lambda_2 = \boldsymbol{\nu}^\top\boldsymbol{L}\boldsymbol{\nu} = \sum_{ij\in\mathcal{E}} \sigma_{ij}(\nu_i - \nu_j)^2$. Assume $\lambda_2 = 0$. On a connected graph, there exists a path between every two nodes, which enforces $\boldsymbol{\nu} = c\mathbf{1}$ for a constant c. This contradicts $\boldsymbol{\nu}^\top\mathbf{1} = 0$, and so λ_2 is positive.

Let $\boldsymbol{L}^\dagger$ be the Moore-Penrose pseudoinverse of $\boldsymbol{L}$. Both $\boldsymbol{L}$ and $\boldsymbol{L}^\dagger$ have the same eigenvectors, while two corresponding eigenvalues are reciprocals of each other, except that 1 replaces the zero eigenvalue of $\boldsymbol{L}$. One can verify that

$$\boldsymbol{L}^\dagger = \Big(\boldsymbol{L} + \frac{1}{n}\,\mathbf{1}^\top\mathbf{1}\Big)^{-1} - \frac{1}{n}\,\mathbf{1}^\top\mathbf{1} \quad \text{with} \quad n := |\mathcal{V}|.$$

Also $\boldsymbol{L}^\dagger$ enjoys the structural property of $\boldsymbol{L}^\dagger\mathbf{1} = \mathbf{0}$.

The Dirichlet Laplacian $\boldsymbol{L}_\circ := \boldsymbol{B}_\circ\,\mathrm{diag}(\boldsymbol{\sigma})\,\boldsymbol{B}_\circ^\top$ is made from $\boldsymbol{L}$ by discarding the entries corresponding to a *reference* node d. Let $\mathcal{E} = \mathcal{E}_1 + \mathcal{E}_\circ$, where $\mathcal{E}_\circ$ is the set of edges with one end connected to node d. One can verify that $\forall\,\boldsymbol{x} \neq \mathbf{0}$,

$$\boldsymbol{x}^\top\boldsymbol{L}_\circ\boldsymbol{x} = \textstyle\sum_{ij\in\mathcal{E}_1}\sigma_{ij}(x_i - x_j)^2 + \sum_{id\in\mathcal{E}_\circ}\sigma_{id}\,x_i^2 > 0$$

which means $\boldsymbol{L}_\circ$ is *positive definite* and so *invertible.* Further, $\boldsymbol{L}_\circ\mathbf{1} = \boldsymbol{y}$ is a nonzero vector with non-negative coordinates, where $y_i = 0$ if $id \in \mathcal{E}_1$ and $y_i > 0$ if $id \in \mathcal{E}_\circ$.

Sometimes, it is misunderstood that $\boldsymbol{L}_\circ$ carries the same eigenvalues as $\boldsymbol{L}$ but the zero, which is not true.

In combinatorial geometry, one can view graph as a 1-complex, where $\boldsymbol{B}$ is its 1-incidence matrix that describes the correlation between all oriented 1-cells (edges) and 0-cells (nodes) in the complex.

The structure of incidence matrix $\boldsymbol{B}$ is the same for both directed and undirected graphs except that the edge directions in directed graphs are substituted for the arbitrarily assigned algebraic topological edge orientations in undirected graphs.

Reza Banirazi received his B.S. from Tehran Polytechnic, and his M.S. and Ph.D. from the University of Southern California, all in Electrical Engineering. He currently works on the design and implementation of data-driven algorithms and applications for stochastic optimization and dynamic resource allocation in high dimensions. He has served as a reviewer for several high-impact journals in systems, control and networking, including the IEEE Transactions on Neural Networks and Learning Systems, the IEEE Transactions on Automatic Control, and the IEEE Transactions on Mobile Computing.

Edmond Jonckheere received his Ph.D. in Electrical Engineering from the University of Southern California in 1978. In 1979, he was with the Philips Research Laboratory, Brussels, Belgium. Since 1980, he has been with the University of Southern California, Los Angeles, where he is currently a (full) Professor of Electrical Engineering and Mathematics, a member of the Center for Applied Mathematical Sciences, and a member of the Center for Quantum Information Science and Technology. His main research interests are geometric topology of networks, including power grid, and quantum information technology. He is a Fellow of the Institute of Electrical and Electronics Engineers for his contribution to robust control.

Bhaskar Krishnamachari is a Professor of Electrical and Computer Engineering at the Viterbi School of Engineering, University of Southern California. He received his B.E. from The Cooper Union for the Advancement of Science and Art, and his M.S. and Ph.D. from Cornell University, all in Electrical Engineering. He works on the design and analysis of algorithms, protocols and applications for the internet of things and other distributed systems.