Wireless Network Thermodynamics: Interfering Stochastic Multiclass Diffusion on Limited Capacity

Reza Banirazi, Edmond Jonckheere, Bhaskar Krishnamachari

Abstract—Concepts of temperature, heat, and entropy are extended to packet routing on multiclass wireless networks with particular challenges including network directionality, channel interference, channel capacity, time-varying topology, and stochastic arrivals. Applying thermodynamic principles in the framework of combinatorial geometry, a novel physics-oriented multiclass diffusion process is developed on directed graphs constrained to edge capacities, leading to a nonlinear graph Laplacian. Taking this as a target model, the multiclass diffusion process on a wireless network is then established using the notion of fluid limit for stochastic queuing networks. Specifically, a family of throughput-optimal control policies is introduced under which the fluid limit of stochastic wireless network comply with the multiclass diffusion on the underlying directed graph with suitably-weighted and capacitated edges. Thus we cross-fertilize wireless networking with classical thermodynamics that opens a door, for the first time, to take advantage of powerful tools from heat calculus in the analysis and optimization of multiclass wireless networks subject to channel interference and capacity. Further, this scheme can be adapted to the general class of stochastic processing networks, which includes many other problems beyond wireless networking.

Index Terms—Wireless network, processing network, control, routing, interference, queuing, diffusion, thermodynamics.

I. INTRODUCTION

Thermodynamics is, without a doubt, one of the most fundamental laws of nature that addresses some intriguing questions in physics, chemistry, mathematics, neuroscience, and even about the origin of the universe. The zeroth law of thermodynamics defines thermal equilibrium based on which neither of two systems with equal temperatures tends to transfer heat to the other. The first law asserts the existence of a state variable, the internal energy, that obeys the principle of energy conservation. By the second law, heat cannot spontaneously flow from cold to hot, observing the fact that temperature tends to even out in an isolated system, where entropy is a measure of how much this process has progressed.

In many areas of science and technology, a process is conceptualized by a graphical model in which a set of subsystems (nodes) are interconnected by a set of interfaces (edges) through which a sort of conserved commodity flows. The commodity can be data or information in a communication system, can be charge in an electrical circuit, can be number of dissolved particles in a particle diffusion process, can be photons or heat in an optical or a thermal system, can be coupling Hamiltonian in a spin network, and so forth.

When the system is deterministic and the graphical model can be represented with a time-invariant, free-capacity, undirected graph, the disciplines of thermodynamics become equivalent to celebrated Kirchhoff’s laws in circuit theory, and thus the notion of network thermodynamics will be well defined [1]. However, there are different sort of problems in which the system is stochastic or the graphical model is time-varying. Further, the graphical model may be restricted to some constraints, e.g. when the edges are of limited capacity, or the graph is directed, or the edges are not independent resources to be used simultaneously. Moreover, there are problems in which multi-type of commodities need to share some limited graphical resources. The goal of this paper is to extend the notion of network thermodynamics to this family of problems by addressing the following questions:

- How to approach stochasticity in network?
- How to lay down multiclass network thermodynamics?
- How to enter edge directionality into the picture?
- How to observe limited capacity on edges?
- How to treat interdependency among edges?
- How to approach stochasticity in network?

Packet routing on a wireless network is a conspicuous example with all of the above-mentioned complexities. Nevertheless, the framework of this paper is applicable to a wider family of stochastic problems with interdependent resources, so-called stochastic processing networks [2], where the resources are a collection of interdependent servers that can only be accessed under certain constraints, and the consumers are of random service time with asynchronous completion. This general model describes a wide variety of problems including queuing networks, product assembly systems, memory or processor managements, call centers, agent allocations, data switches, and healthcare systems, just to name a few.

A. Related Work

The term “network thermodynamics” has been vastly used in biology and chemistry in modeling the kinetics of large chemical reaction networks (see for example [3]–[5] and references therein), where all of them fall into the central pathway of thermodynamic principles on a free-capacity, undirected graph. In data networks, there have been some attempts to design top layer routing strategies inspired by heat propagation for large-scale mesh and sensor networks [6], [7], or by potential field degradation for wireline networks [8], where the concept of temperature field, or potential field for that matter, is used to build a local-maximum free gradient of data towards destination. Those works do not consider any of the main concerns of this paper including network directionality, edge capacity, process stochasticity, and inter-link interference.

This work was supported by NSF grants CNS-1017881 and CNS-1049541. 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
The closest schemes to the work of this paper is the family of Back-Pressure-based policies, also called maximum pressure policies. The original Back-Pressure (BP) was introduced in [11] consisted of a max-weight link scheduling and a congestion gradient packet forwarding, and was shown to be throughput-optimal. During more than two decades from the original paper, BP-based schemes have been a very active research topic in computer networking, product assembly systems, and processing networks [12]–[14].

This paper is the continuation of our research on developing a wireless routing policy inspired by heat diffusion process [15]–[17]. Specifically, this paper extends the results of [16] on "wireless network thermodynamics" in two directions: First, rather than taking a free-capacity directed graph, here we work with a capacity-constrained directed graph. This relaxes the main restriction in [16] that under a natural heat propagation, the flow on each wireless link needed to remain spontaneously lower than the link expected time average capacity (see Theorem 5 and Remark 2 in [16]). Second, here we consider multiclass networks that together with limited edge capacities make the problem far harder.

B. Contribution

As the main contribution, this paper extends the principles of thermodynamics to routing and resource allocation on wireless networks, or more generally on stochastic processing networks. This makes possible of utilizing tools from heat calculus, an active area of pure mathematics, in the analysis and optimization of stochastic multiclass queuing networks with link interference and time-varying topology.

In particular, we introduce a family of multiclass network control policies under which the long-time average flow of packets on the wireless network, called fluid limits, comply with heat propagation process on the underlying directed graph with suitably weighted and capacitated edges. It is further shown that every control policy in this family is throughput-optimal, i.e. it stabilizes (keeps finite) all the queues under any traffic rate matrix within the network capacity region.

As a side contribution, we develop the notion of multiclass heat propagation on the combinatorial domain of directed graphs with limited edge capacities, a novel idea that bridges classical heat diffusion and multiclass network applications. To preserve important characteristics of heat propagation, including the energy minimization principle, we exploit the theory of combinatorial geometry that allows us to directly transfer principles of physics to a combinatorial setting.

C. Organization

The next section formulates the classical thermodynamics on smooth manifolds, and in parallel on uncapacitated undirected graphs. Section III develops the novel idea of multiclass thermodynamics on capacitated directed graphs. The notion of wireless network thermodynamics is established in Sec. IV by introducing a family of routing policies under which the fluid limits of wireless network satisfy the multiclass heat equations developed in Sec. III. We show in Sec. VI that this family of routing policies are throughput-optimal. Simulation results are given in Sec. ?? prior to concluding the paper in Sec. VII.

TERMS AND NOTATION

The paper discusses three different dynamic processes: (i) heat diffusion on graphs, which is continuous and deterministic, (ii) packet routing on wireless networks, which is time-slotted and stochastic, and (iii) fluid limit of packet routing on wireless networks, which is continuous and deterministic. We deliberately use similar notation for these three different systems, emphasizing the similarity in their behavior, but care should be taken to discriminate them.

When the subject is about heat process we use the terms graph, vertices and edges, whereas using the terms network, nodes and links when the subject is about wireless routing. The terms capacity-free and uncapacitated (resp., capacity-constrained and capacitated) are used interchangeably.

We use \( \mathbb{R} \) for the set of real numbers, \( \mathbb{E} \) for expectation, and \( \mathbb{P} \) for probability. For \( S \) as a set, \( |S| \) denotes its cardinality. A superscript \( T \) denotes the transpose operation. For \( v \) as a vector, \( \|v\| := (v^T v)^{1/2} \). For \( w \) as a block vector, \( \text{diag}(w) \) denotes its block diagonal matrix expansion. We denote the zero vector with \( 0 \), the vector of all ones of size \( m \) with \( 1_m \), and the identity matrix of size \( m \) with \( I_m \). Between vectors or matrices, the entrywise (Schur) product is denoted with \( \odot \), and the tensor product with \( \otimes \). Between two vectors \( u \) and \( v \), the operators \( \min\{u, v\} \) and \( \max\{u, v\} \), and also curly inequalities \( \ll \) and \( \gg \), act entrywise. For a vector \( v \), we define \( v^+ := \max\{0, v\} \). On a graph, or a network for that matter, for a value \( x \) corresponding to a directed edge \( f \) from vertex \( i \) to vertex \( j \), we use the notation \( x_f \) and \( x_{ij} \) interchangeably.

II. PRELIM: CLASSICAL THERMODYNAMICS

To bring the concepts of thermodynamics from ordinary smooth geometry to the purely combinatorial domain of a network, we use the theory of combinatorial geometry where the notion of chains-cochains provides a genuine counterpart for differential forms in geometry. To find the details, refer to [18], [19] and references therein.

A. Classical Continuous-Domain Thermodynamics

On a smooth bounded manifold \( M \) charted in local coordinates \( z \), let \( Q(z, t) \) be the scalar field of heat distribution, \( F(z, t) \) the vector field of heat flux through the boundary \( \partial M \), and \( A(z, t) \) the scalar field of heat sources (with minus for sinks) at time \( t \). An isolated system exchanges no energy with other systems, including the environment. However, energy may transfer between two points within the system in the form of mass, heat, or work. Assuming energy is transferred only as heat, the first law of thermodynamics entails that the amount of heat in any region \( M' \subset M \) must either leave through the boundary \( \partial M' \) or have an external source at any time \( t \),

\[
\frac{\partial}{\partial t} \int_{M'} Q(z, t) = -\int_{\partial M'} F(z, t) \cdot N(z) + \int_{M'} A(z, t) \quad (1)
\]

where \( N \) represents the unit outward normal vector field on the boundary, and \( \cdot \) denotes the projection operator.

Let \( U(z, t) \) be the scalar field of temperature distribution, defined as the ability of a spatial point \( z \) at time \( t \) to give up heat to its surroundings. By the second law of thermodynamics, heat spontaneously flows from higher to lower temperature
with a flow proportional to the temperature gradient,
\[ F(z, t) = -\sigma(z) \nabla U(z, t) \] (2)
where \( \sigma \) represents the spatial thermal conductivity that quantifies how fast heat moves through the material.

For a fixed amount of total energy, in thermal equilibrium, the energy is uniformly distributed throughout the system. Every isolated system spontaneously evolves towards equilibrium, where entropy is a measure of this progress. Thermal equilibrium is the state of maximum entropy; thus the entropy of an isolated system never decreases.

Energy, entropy, and temperature are related by the fundamental thermodynamic relationship. Though the standard symbol for entropy is \( S \), it is denoted by \( H \) in this paper as we will use \( S \) for another purpose. In the absence of work and mass transfer, temperature is expressed as the inverse of the rate of change of entropy \( H \) with internal energy \( Q \),
\[ U(Q) = \left( \frac{\partial H(Q)}{\partial \theta} \right)^{-1}. \]

Defining \( H(Q) := C(z) \ln(Q) \) with \( C \) representing the spatial thermal capacity that quantifies the amount of heat required to change the temperature, the above equation leads to
\[ Q(z, t) = C(z) U(z, t). \] (3)

Using the divergence theorem in (1), and as \( \mathcal{M}' \) is arbitrary and can be chosen infinitesimally small, the first law of thermodynamics can equally be formulated as
\[ \frac{\partial Q(z, t)}{\partial t} = -\text{div} F(z, t) + A(z, t). \]
Substituting (2) and (3) into the above equation leads to the following parabolic partial differential equation that describes the variation in temperature in a given region over time:
\[ C(z) \frac{\partial U(z, t)}{\partial t} = \text{div} (\sigma(z) \nabla U(z, t)) + A(z, t). \] (4)
To solve this equation uniquely, besides time initial condition, one needs to prescribe \( U(t) \) on the boundary \( \partial \mathcal{M} \).

B. Classical Graph Thermodynamics

Combinatorial geometry operates on a cell complex as the discrete domain. In contrast with simplicial electromagnetics that works with 2- or 3-dimensional cell complexes, here the structure of wireless network enforces working only with a 1-dimensional complex, which entails a significant revision of the concepts. A \( p \)-dimensional cell complex is fully described by a collection of \( k = 1, \ldots, p \) incidence matrices. Graph viewed as a 1-dimensional cell complex has only the first incidence relation between 0-cells (vertices) and 1-cells (edges). Denoting the vertex-edge incidence matrix with \( B \), the entry \( B_{ij} \) takes the value 1 if vertex \( i \) is the tail of oriented edge \( \ell \), \(-1 \) if \( i \) is the head, and 0 otherwise.\(^1\)

Transferring the elements of classical thermodynamics to the graph as a 1-complex combinatorial domain leads to the following substitutions: (i) Smooth space \( \mathcal{M} \) is replaced by a 0-chain vector \( \eta_0 \) representing the discrete space. (ii) Scalar fields \( Q(z, t), U(z, t), H(Q) \) and \( A(z, t) \) are respectively replaced by 0-cochain vectors \( q(t), u(t), h(q) \) and \( a(t) \) representing the values of heat, temperature, entropy, and heat generation at vertices. (iii) Vector field \( F(z, t) \) is replaced by a 1-cochain vector \( f(t) \) which is an edge variable representing heat flow on edges. (iv) Spatial thermal conductivity \( \sigma(z) \) is replaced by an edge weight vector \( \sigma \) modifying flow through edges. (v) Spatial thermal capacity \( C(z) \) is replaced by a vertex weight vector \( c \) modifying temperature at vertices.

Now we construct the combinatorial analogues of thermodynamic laws on an undirected graph with free-capacity edges. The counterpart of integral in combinatorial geometry is the pairing between a chain and a cochain, where pairing is simply defined as the inner product of the two vectors. Thus, the combinatorial analogue of the first law in (1) becomes
\[ \frac{d}{dt} \langle \eta_0, q(t) \rangle = -\langle \eta_0, B f(t) \rangle + \langle \eta_0, a(t) \rangle \] (5)
where \( B \) performs as the algebraic codifferential operator acting on the 1-cochain \( f(t) \) to generate a 0-cochain [19].

The gradient operator on a smooth manifold resembles the 0-dimensional exterior derivative in differential geometry that maps 0-forms into 1-forms. The combinatorial equivalent of exterior derivative is the algebraic coboundary operator, where the \( k \)-coboundary moves \( (k-1) \)-cochains into \( k \)-cochains [19]. Graph as a 1-complex has only the first coboundary operator that is simply the adjoint of the \( B \) matrix. Thus the combinatorial analogue of the second law in (2) becomes
\[ f(t) = \text{diag}(\sigma) B^T u(t). \] (6)
The sign change is due to the convention of taking the classical gradient positive when the quantity increases along a direction, which is inverse in differential and combinatorial geometry.

Defining \( h(q) := \text{diag}(c) \ln(q) \) with ln being taken entrywise, the fundamental thermodynamic relationship leads to
\[ q(t) = \text{diag}(c) u(t). \] (7)
Since (5) must be true for an arbitrary chain, \( \eta_0 \) can be canceled from the equation. Then substituting (6) and (7) leads to the combinatorial analogue of heat equation (4) as
\[ \text{diag}(c) \dot{u}(t) = -B \text{diag}(\sigma) B^T u(t) + a(t) \] (8)
with overdot denoting time derivative. Notice the codifferential \( B \) as the combinatorial equivalent of classical divergence.

Matrix \( L := \text{diag}(c)^{-1} B \text{diag}(\sigma) B^T \), called weighted graph Laplacian, is the combinatorial equivalent of 0-dimensional Laplace-deRham operator in differential geometry [19]. The structure of \( L \) clearly shows that it is symmetric and positive semi-definite, whose nullspace is the same as the left nullspace of \( B \). Viewing \( B \) as the algebraic boundary operator, on the other hand, its left nullspace is of rank equal to the number of connected components in the graph [19]. Thus for a connected graph, the nullspace of \( L \) has rank one, and spanned by a vector of equal constant elements, as the rows of incidence matrix \( B \) are summed to \( 0 \).

C. Reduced Graph Thermodynamic Equations

For uniquely solving the combinatorial equation (8), like for its classical counterpart (4), we need to fix \( u(t) \) at a
reference vertex \( d \), which is analogous to a collapse of the boundary \( \partial \mathcal{M} \) to a point on the continuous domain. The thermal interpretation of the reference is the vertex that absorbs the heat generated in all vertices while keeping its own temperature constant, which thermodynamically has to be in the absolute zero. Fixing \( u_d(t) = 0 \), one may eliminate it from (5)–(8) leading to a reduced set of equations that have a unique solution, given time initial condition. Let subscript \( o \) denote a reduced vector or matrix obtained by discarding the entries corresponding to the reference vertex \( d \). Specifically, matrix \( B_o \) denotes a reduction of \( B \) through discarding the row corresponding to \( d \) and is called the basis incidence matrix with respect to \( d \). Then we obtain the reduced set of deterministic graph thermodynamic equations as

\[
\begin{align*}
    f(t) &= \text{diag}(\sigma) B_o^T u_o(t) \\
    q_o(t) &= \text{diag}(c_o) u_o(t) \\
    \text{diag}(c_o) \dot{u}_o(t) &= -L_o u_o(t) + a_o(t)
\end{align*}
\]

where \( L_o := B_o \text{diag}(\sigma) B_o^T \) is called Dirichlet Laplacian with respect to \( d \). For a connected graph, \( L_o \) is positive definite [19] that implies bounded-input bounded-output (BIBO) stability of the dynamical system (11) in general, and its exponential stability when \( a_o(t) \) has a fixed steady-state value.

### III. Multiclass Thermodynamics on Capacity-Constrained Directed Graphs

This section extends the classical graph thermodynamics of previous section along two totally new directions: (i) extension to directed graphs with limited edge capacities, and (ii) extension to multiclass graphs with so-called “different types of calorie” heading to “different heat sinks.” Graph thermodynamics, or more specifically diffusion process, subject to either edge capacity or edge directionality has not been considered in literature, let alone the same problem with the both constraints and at the same time in a multiclass framework.

#### A. Uniclass Diffusion on Capacitated Directed Graphs

Consider a graph whose edges are directed and of limited capacity. On this discrete domain, the first thermodynamic law (5), implying heat conservation, remains unchanged. However, the second law (6) must be modified to allow the flow in only one direction, and to limit it within the edge capacity. With no loss of generality, let the edge orientation concurs with the edge direction. Then the equivalent of (6) on a capacity-constrained directed graph becomes

\[
    f(t) = \min \{ \text{diag}(\sigma) (B^T u(t))^+, \mu \}
\]

where \( \mu \) is the constant vector of edge capacities \( \mu_{ij} \). Like in undirected graphs, one may eliminate the sink \( d \) from the equations by fixing the boundary condition \( u_d(t) = 0 \). Then the reduced set of deterministic thermodynamic equations on a capacitated directed graph is obtained as

\[
\begin{align*}
    \dot{q}_o(t) &= -B_o f(t) + a_o(t) \\
    f(t) &= \min \{ \text{diag}(\sigma) (B_o^T u_o(t))^+, \mu \} \\
    q_o(t) &= \text{diag}(c_o) u_o(t)
\end{align*}
\]

Plugging (13) and (14) in (12) yields

\[
    \text{diag}(c_o) \dot{u}_o(t) = -\bar{L}_o(u_o) + a_o(t)
\]

\[
    \bar{L}_o(u_o) := B_o \min \{ \text{diag}(\sigma) (B_o^T u_o)^+, \mu \}
\]

We refer to \( \bar{L}_o(\cdot) \) as nonlinear Dirichlet Laplacian matrix, which is an operand-dependent nonlinear operator acting on a capacity-constrained directed graph.\(^2\)

Our nonlinear diffusion formulation is in agreement with the recent work in [20] showing that diffusion on Finsler manifolds, the natural counterparts of directed graphs in continuous domain, leads to a nonlinear Laplace operator. In the graph literature, different linear Laplacian matrices have been proposed for directed graphs (see [21, Sec. 3] for a review). While successful to address some purely graphical issues, they do not convey the physics of the diffusion, nor the intrinsic nonlinearity due to the one-way flow restrictions on edges.

**Definition 1:** Given a scalar deterministic continuous-time function \( x(t) \), its long-time average value is defined as

\[
    \text{ave}(x) := \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau x(t) dt.
\]

For discrete-time functions, the integral is replaced by the sum.

The definition is extended entrywise to vector functions.

For (15) being BIBO stable, a necessary and sufficient condition is the feasibility of the corresponding network flow problem defined by replacing time-varying heat sources \( a_o(t) \) with their average values \( \text{ave}(a_o) \), which are constant.\(^3\) Also observe that on a directed graph, a reasonable assumption — even though not required for any of our analytical results — is that if a vertex generates heat, then there exists at least one directed path from that vertex to the sink.

**Theorem 1:** Consider a capacitated directed graph with a set of deterministic heat sources \( a_o(t) \). Then the nonlinear diffusion process (15) is BIBO stable if and only if the conventional network flow problem defined by replacing \( a_o(t) \) with \( \text{ave}(a_o) \) is feasible. If, in addition, \( a_o(t) \) is of fixed steady-state value, then the diffusion system has a unique equilibrium point that is exponentially stable.

**Proof:** First, assume that the network flow problem has at least one solution, but the system (15) is not stable; thus there exist some vertices of infinite temperature. Since the heat flows along the negative gradient of temperature, at least one of the vertices with infinite temperature has to be a heat source too, say vertex \( m \) with the heat source \( a_m(t) \). This implies, by the heat flow equation (13), that all possible directed paths from the source \( a_m(t) \) towards the sink have reached their capacity limits, yet unable to transfer the generated heat \( a_m(t) \) to the sink, where the capacity limit of a directed path equals the minimum link capacity on the path. Also observe that replacing \( a_m(t) \) with \( \text{ave}(a_m) \) may only change the time that the temperature at vertex \( m \) blows up to infinity. But this is in contradiction with the initial assumption that the \( \text{ave}(a_o) \) network flow problem has a solution, meaning that the heat diffusion process (15) must be stable.

\(^2\)Note that on a free-capacity directed graph with the unit vertex weights, i.e. when \( \mu_{ij} = \infty \) for all edges and \( c_i = 1 \) for all vertices, \( L_o(\cdot) \) reduces to the nonlinear Dirichlet Laplacian defined in [16].

\(^3\)A network flow problem is feasible if there exists at least one flow configuration holding both vertex flow conservation and edge capacity constraints.
Next, assume the system (15) is stable. Therefore,

\[
\text{ave}(\dot{u}_o) = \limsup_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau \dot{u}_o(t) \, dt = \limsup_{\tau \to \infty} \frac{1}{\tau} u_o(\tau) = 0
\]

where the last equality comes from the system stability. In light of (14), \(\text{ave}(\dot{u}_o) = 0\) also implies \(\dot{q}_o = 0\); thus taking time average from (12) leads to \(\text{ave}(a_o) = B_o \text{ave}(f)\). Due to the structure of matrix \(B_o\), the term \(B_o \text{ave}(f)\) expresses the average net flow of heat drawn away from the vertices through the edges, i.e. the leaving flow through the outgoing edges minus the entering flow through the incoming edges. Thus \(\text{ave}(a_o) = B_o \text{ave}(f)\) indeed reads the average flow conservation at vertices, showing that the \(\text{ave}(a_o)\) network flow problem has at least one solution.

Now we show that for \(\text{ave}(a_o)\) being constant, the stability of the system (14) is exponential. Since \(a_o(t)\) is an independent input into the system — not controlled by the diffusion dynamics — it contributes only to the equilibrium state of the system, which is either an equilibrium point for the case of time-invariant heat sources or an equilibrium trajectory for the case of time-varying heat sources. Thus, the stability of (14) is determined by the stability of the non-excited system

\[
\text{diag}(c_o) \dot{u}_o(t) = - \mathcal{L}_o(u_o) = -B_o f(t).
\]

Consider the Lyapunov function

\[
Y(t) := \frac{1}{2} u_o(t)^\top \text{diag}(c_o) u_o(t).
\]

Exploiting (17) in the time derivative of (18) yields

\[
\dot{Y}(t) = -f(t)^\top B_o^\top u_o(t).
\]

With \(E\) being the set of edges, \(B_o^\top u_o\) is a vector of size \(|\mathcal{E}|\) with entries representing the edge temperature-differences. Then for any given \(u_o\), one can partition \(B_o^\top u_o\) into the three groups: edges with saturated flow where \(\sigma_{ij}u_{ij} > \mu_{ij}\), edges with non-saturated flow where \(0 < \sigma_{ij}u_{ij} \leq \mu_{ij}\), and edges with zero flow where either \(u_{ij} = 0\) or \(u_{ij} < 0\). Accordingly, the incidence matrix \(B_o\) can be partitioned as

\[
B_o = (B_{o\bullet} \mid B_{o\Delta} \mid B_{o\emptyset})
\]

where \(B_{o\bullet}\) denotes the partition of \(B_o\) containing the incidence information of the edges with positive temperature-difference and saturated heat flow, \(B_{o\Delta}\) of the edges with positive temperature-difference but non-saturated heat flow, \(B_{o\emptyset}\) of the edges with either zero or negative temperature-difference and so zero heat flow. Likewise, the edge weights and the edge capacities get partitioned as

\[
\sigma = (\sigma_{\bullet} \mid \sigma_{\Delta} \mid \sigma_{\emptyset}) \quad \text{and} \quad \mu = (\mu_{\bullet} \mid \mu_{\Delta} \mid \mu_{\emptyset}).
\]

Using this partitioning in (19), we get

\[
\dot{Y}(t) = -u_o(t)^\top B_{o\bullet} \mu_{\bullet} - u_o(t)^\top B_{o\Delta} \text{diag}(\sigma_{\Delta}) B_{o\Delta}^\top u_o(t) - \dot{u}_o(t)^\top B_{o\emptyset} \text{diag}(\sigma_{\emptyset}) B_{o\emptyset}^\top u_o(t)
\]

where we have used the fact that

\[
\begin{align*}
\min\{\text{diag}(\sigma_{\bullet})(B_{o\bullet}^\top u_o)^+, \mu_{\bullet}\} &= \mu_{\bullet} \\
\min\{\text{diag}(\sigma_{\Delta})(B_{o\Delta}^\top u_o)^+, \mu_{\Delta}\} &= \text{diag}(\sigma_{\Delta}) B_{o\Delta}^\top u_o \\
\min\{\text{diag}(\sigma_{\emptyset})(B_{o\emptyset}^\top u_o)^+, \mu_{\emptyset}\} &= 0
\end{align*}
\]

For a saturated edge \(ij\) on which \(\sigma_{ij}u_{ij} > \mu_{ij}\) and \(f_{ij} = \mu_{ij}\), one can find a \(\sigma'_{ij}\), say the edge virtual conductivity, such that \(\sigma'_{ij}u_{ij} = \mu_{ij} = f_{ij}\). Let us form the following arrays:

\[
\sigma' := (\sigma'_{\bullet} \mid \sigma'_{\Delta} \mid \sigma'_{\emptyset}) \quad B_o' := (B_{o\bullet} \mid B_{o\Delta} \mid B_{o\emptyset}) \quad \mu' := (\mu_{\bullet} \mid \mu_{\Delta} \mid \mu_{\emptyset}).
\]

Then the Lyapunov derivative (20) can be restated as

\[
\dot{Y}(t) = -u_o(t)^\top B_o' \text{diag}(\sigma') B_o'^\top u_o(t).
\]

As the \(\text{ave}(a_o)\) network flow problem has at least one solution, for every heat source there exists at least one directed path to the heat sink. This implies that the set of edges of positive temperature-difference build a connected graph with the vertex \(d\). For a connected graph, any sub-matrix obtained from its incidence matrix by removing one or more rows has full row rank [19]. Then the matrix \(B_o'^\top \text{diag}(\sigma') B_o'^\top\) has full rank, and so is symmetric positive definite. Let \(\lambda_{\min} > 0\) be the smallest eigenvalue of this matrix, and \(c_{\min}\) and \(c_{\max}\) be respectively the smallest and largest vertex weights in the graph. Then (21) and (18) lead to

\[
\dot{Y}(t) \leq -\lambda_{\min} u_o(t)^\top u_o(t) \leq -2(\lambda_{\min}/c_{\max}) Y(t).
\]

By some matrix manipulation, the latter implies

\[
\lVert u_o(t) \rVert \leq \sqrt{c_{\min}/c_{\max}} \lVert u_o(0) \rVert e^{-(\lambda_{\min}/c_{\max}) t}.
\]

This means that from any initial condition, the system tends to its equilibrium state exponentially fast.

The next theorem formalizes a property of graph thermodynamics that is being used, in the next subsection, to develop the notion of graph multiclass thermal diffusion.

**Theorem 2:** On a uniclass graph with free-capacity undirected edges (resp., capacity-constrained directed edges), the flow assigned by the second thermodynamic law in the equation (9) (resp., in the equation (13)) uniquely minimizes the functional \(\lVert \text{diag}(\sigma') B_o'^\top u_o(t) - f(t) \rVert\) among all admissible flow configurations that respect flow conservation at vertices.

**Proof:** To simplify the notation, we drop the time variable \((t)\). Let \(J(f) := \lVert \text{diag}(\sigma') B_o'^\top u_o - f \rVert\). On uncapacitated undirected graphs, the minimum of \(J(f)\) is obviously zero, resulting in the linear flow equation (9). On a capacitated directed graph, the solution must respect the direction and capacity of edges. For a given configuration of \(q_d\), the constraint on each specific edge influences the flow of only that edge, with no impact on the other edges. Thus this constrained minimization can be solved for each edge independently by pushing each entry of \((\text{diag}(\sigma') B_o'^\top u_o - f)\) towards 0 subject to \(0 \leq f_{ij} \leq \mu_{ij}\). Then it is easily seen that the solution to this problem leads to the nonlinear flow equation (13).

**B. Multiclass Diffusion on Capacitated Directed Graphs**

In a traditional notion of graph thermodynamics, the heat generated in all vertices is absorbed by one single vertex as the sink. A more complex scenario, however, may be fantasized in parallel with multiclass problems in networking. Specifically, consider a scenario in which different types of heat are generated at vertices, where each type is absorbed by a specific vertex. To make a better connection with multiclass networks, let us refer to each type of heat as a class.

In the absence of edge capacity constraints, the diffusion of different classes happen independently, so that the dynamics of each class are decoupled from the dynamics of other classes.
On a graph of limited edge capacities, however, the dynamics of multiclass diffusion process is no longer the decoupled collection of dynamics of comprising single-sink diffusions. This is simply because the way of allocating edge capacities to each class has a direct impact on the diffusion dynamics of that class, while the sum of the allocated capacities on each edge must be constrained to a bounded value.

Consider the above-described multiclass thermal diffusion on a capacity-constrained directed graph described by set of vertices $V$, set of edges $E$, set of classes $K \subseteq V$, and vector of edge capacities $\mu$. Let superscript $(d)$ denote a quantity related to class $d \in K$, referred to as $d$-quantity. Specifically, we denote the vector of $d$-flows on edges with $f^{(d)}(t) \in \mathbb{R}^{|E|}$, the reduced vector of $d$-heats and $d$-temperatures at vertices respectively with $q^{(d)}_o(t)$ and $u^{(d)}_o(t) \in \mathbb{R}^{|V|-1}$, the reduced vector of $d$-sources at vertices with $a^{(d)}_o(t) \in \mathbb{R}^{|V|-1}$, and the basis incidence matrix with respect to the vertex $d$, i.e. the heat sink of class $d$, with $B^{(d)}_o \in \mathbb{R}^{(|V|-1) \times |E|}$. We also use $0 \leq \theta^{(d)}_{ij}(t) \leq 1$ to represent what portion of total capacity of an edge $ij$ is devoted to the class $d$ at time $t$,

$$\mu^{(d)}_{ij}(t) = \theta^{(d)}_{ij}(t) \mu_{ij} \quad \text{with} \quad \sum_{d \in K} \theta^{(d)}_{ij}(t) \leq 1 \quad (22)$$

and accordingly form $\theta^{(d)}(t) \in \mathbb{R}^{|E|}$ as the vector of edge capacity-factors at time $t$. Note that while the total edge capacity $\mu_{ij}$ is constant, the portion of it devoted to a class $d$ may be a function of time. Further, we let $\sigma^{(d)} \in \mathbb{R}^{|E|}$ denote the vector of $d$-conductivities on edges, giving edges the possibility of having different conductivities for different classes. Also each node may introduce different thermal capacities to different classes, where $c^{(d)}_o \in \mathbb{R}^{|V|-1}$ denotes the reduced vector of $d$-capacities at vertices.

The dynamics of multiclass thermal diffusion is the collection of the diffusion dynamics of its comprising uniclass processes. To formulate the state equations of this multisystem in a compact form, let us cast the following hyper-vectors by placing the corresponding uniclass vectors together:

$$q_o(t) := \begin{bmatrix} q^{(1)}_o(t), \ldots, q^{(|K|)}_o(t) \end{bmatrix}^T \in \mathbb{R}^{|(|V|-1)|K|}$$

$$u_o(t) := \begin{bmatrix} u^{(1)}_o(t), \ldots, u^{(|K|)}_o(t) \end{bmatrix}^T \in \mathbb{R}^{|(|V|-1)|K|}$$

$$a_o(t) := \begin{bmatrix} a^{(1)}_o(t), \ldots, a^{(|K|)}_o(t) \end{bmatrix}^T \in \mathbb{R}^{|(|V|-1)|K|}$$

$$c_o := \begin{bmatrix} c^{(1)}_o, \ldots, c^{(|K|)}_o \end{bmatrix}^T \in \mathbb{R}^{|(|V|-1)|K|}$$

$$f(t) := \begin{bmatrix} f^{(1)}(t), \ldots, f^{(|K|)}(t) \end{bmatrix}^T \in \mathbb{R}^{|E||K|}$$

$$\theta(t) := \begin{bmatrix} \theta^{(1)}(t), \ldots, \theta^{(|K|)}(t) \end{bmatrix}^T \in \mathbb{R}^{|E||K|}$$

$$\sigma := \begin{bmatrix} \sigma^{(1)}, \ldots, \sigma^{(|K|)} \end{bmatrix}^T \in \mathbb{R}^{|E||K|}.$$  

Let us also define the multiclass basis incidence matrix by the following block-diagonal hyper-matrix:

$$B_o := \text{diag}(\begin{bmatrix} B^{(1)}_o, \ldots, B^{(|K|)}_o \end{bmatrix}) \in \mathbb{R}^{(|V|-1)|K| \times |E||K|}. \quad (24)$$

If one can figure out the vector $\theta^{(d)}(t)$ for each class $d$, and so the hyper-vector $\theta(t)$, then thermal diffusion of each class has a direct impact on the diffusion dynamics of its comprising uniclass thermal equations (12)-(15).

In terms of the above-defined hyper-arrays, this leads to

$$f(t) = \min \left\{ \text{diag}(\sigma) \left| B^+_o u_o(t) - Q_o \right| \right\} \quad (25)$$

$$q_o(t) = \text{diag}(c_o) u_o(t) \quad (26)$$

$$\text{diag}(c_o) u_o(t) = -\tilde{L}_o(u_o) + a_o(t) \quad (27)$$

where $\tilde{L}_o(u_o)$ is defined in the similar way as $\tilde{L}_o(u_o)$ in (15) but with the multiclass hyper-arrays and with the $\mu$ replaced by $\theta(t) \otimes (1_{|K|} \otimes \mu)$. The term $(1_{|K|} \otimes \mu)$ extends the vector of edge capacities $\mu \in \mathbb{R}^{|E|}$ to be of size $|E||K|$ and so being used in a multiclass fashion, where its entrywise product with $\theta(t)$ shapes (22) in a hyper-vector form.4

Observe that the $|E|$ diffusion processes are not independent, rather they are coupled together due to the sharing of limited edge capacities. Now, the crucial question is how to divide edge capacities among multiple classes. To this end, we inspire ourselves with the property of uniclass thermal diffusion in Th. 2. Extending this property to a multiclass graph, the multiclass thermal flow $f(t)$ must minimize the multiclass functional $\| \text{diag}(\sigma) B^+_o u_o(t) - f(t) \|$. In the absence of edge capacity constraints, this is readily concluded from Th. 2 together with the flow indepenedency among different classes. Under limited edge capacities, however, the configuration of $f(t)$ depends on the edge capacity-factors $\theta(t)$; thus the minimizing $f(t)$ determines $\theta(t)$,

$$\theta(t) = \arg \min_{\theta} \| \text{diag}(\sigma) B^+_o u_o(t) - f(t) \|$$

subject to: $\sum_{d \in K} \theta^{(d)}(t) \leq 1_{|E|}. \quad (28)$$

Observe that while the optimal flow that solves (28) is unique, the related edge capacity-factors are not necessarily unique, i.e. different $\theta(t)$ may lead to the same optimal flow $f(t)$.

The upshot of this section is the next theorem that formally extends the stability properties of uniclass graph thermodynamics in Th. 1 to the above-developed multiclass case.

**Theorem 3:** Consider a capacitated directed graph with a set of deterministic multiclass heat sources $a_o(t)$. Then the nonlinear multiclass diffusion process (27) under the capacity allocation (28) is BIBO stable if and only if the conventional multiclass network flow problem defined by replacing $a_o(t)$ with $\text{ave}(a_o)$ is feasible.5 If, in addition, $a_o(t)$ is of fixed steady-state value, then the system has a unique equilibrium point that is exponentially stable.

**Proof:** Assume that the multiclass network flow problem has at least one solution, but the system (27) is not stable. Thus there exist some vertices of infinite temperature for some classes, implying that all possible directed paths from some heat sources towards their sinks have reached their assigned capacity limits, with the understanding that the replacement of heat sources with their long-time average values does not change this condition. At the same time, this cannot be the result of our capacity allocation, as (28) entails that a class with the higher gradient in temperature should receive the larger portion of edge capacity, which reaches the full edge capacity when the gradient of class temperature tends to infinity. Therefore, inability of directed paths to convey heats from sources to their related destinations inevitably implies

---

4As an extension of nonlinear Dirichlet Laplacian $\tilde{L}_o(\cdot)$ with the size of $(|V|-1) \times (|V|-1)$ in (15), one may call $\tilde{L}_o(\cdot)$ the multiclass nonlinear Dirichlet Laplacian which is of the size $(|V|-1)|K| \times (|V|-1)|K|$ and acts on a capacity-constrained directed graph with $|K|$ classes.

5A necessary and sufficient condition for the feasibility of a multiclass network flow problem is the existence of a set of multiclass flows that at any time $t$ jointly satisfy flow conservation at vertices and capacity constraints on edges, i.e. $a_o(t) = B_o f(t)$ and $\sum_{d \in K} f^{(d)}_{ij}(t) \leq \mu_{ij}$ for all $ij \in E$. 
the infeasibility of \( \text{ave}(a_0) \) multiclass network flow problem that contradicts the initial assumption.

By a similar reasoning used for Th. 1, on the other hand, stability of (27) leads to \( \text{ave}(a_0) = B_0 \text{ave}(f) \), showing that the \( \text{ave}(a_0) \) multiclass network flow problem is feasible.

Now assume that \( a_0(t) \) is of fixed steady-state value, meaning that \( \text{ave}(a_0) \) is constant. By the same way as for Th. 1, the stability of (27) is determined by the stability of the non-excited system \( \dot{u}_0(t) = -\text{diag}(e_0)^{-1}B_0 f(t) \). Using the Lyapunov function \( Y(t) := \frac{1}{2} u_0(t)^T \text{diag}(e_0) u_0(t) \) leads to
\[
\dot{Y}(t) = -\sum_{d \in K} f_d(t)^T B_0^{d\top} \mu_d(t)
- \sum_{d \in K} u_d(t)^T B_0^{d\top} \text{diag}(\sigma_d(t)) B_0^{d\top} u_d(t)
\]
where subscripts \( \Delta \) and \( \triangle \) respectively show the partition related to the edges of positive temperature-difference with saturated and non-saturated heat flows, and \( \mu_d(t) \) denotes the vector of edge capacities devoted to the class \( d \) at time \( t \) as defined in (22). Then for an edge \( ij \) with a saturated \( d \)-flow, i.e. \( \sigma_i^{(d)}(t) > \mu_{ij}^{(d)} \) and \( f_{ij}^{(d)} = \mu_{ij}^{(d)} \), one can find \( d_{ij}^{(d)} \) such that \( \sigma_i^{(d)}(t) = \mu_{ij}^{(d)} = f_{ij}^{(d)} \). Forming \( \sigma_{ij}^{(d)} := (\sigma_i^{(d)} \mid \sigma_j^{(d)}) \), \( B_0^{(d)} := (B_0^{(d)} \mid B_0^{(d)}) \), and \( \mu_{ij}^{(d)} := (\mu_i^{(d)} \mid \mu_j^{(d)}) \) results in
\[
\dot{Y}(t) = -\sum_{d \in K} u_d(t)^T B_0^{d\top} \text{diag}(\sigma_{ij}^{(d)}) B_0^{d\top} u_d(t).
\]
As the \( \text{ave}(a_0) \) multiclass network flow problem has at least one solution, for every heat source of each class \( d \) there exists at least one directed path to the heat sink \( d \). Thus the reduced matrix \( B_0^{d\top} \text{diag}(\sigma_{ij}^{(d)}) B_0^{d\top} \) is positive definite with the smallest eigenvalue \( \lambda_{\text{min}}(d) > 0 \). Let \( c_{\text{min}} \) and \( c_{\text{max}} \) be respectively the smallest and largest vertex \( d \)-weights. Defining
\[
\lambda_{\text{min}} := \min_d \lambda_{\text{min}}(d), \quad c_{\text{min}} := \min \{ c_{\text{min}} \}, \quad c_{\text{max}} := \max_d c_{\text{max}}(d)
\]
leads to \( \dot{Y}(t) \leq -2 (\lambda_{\text{min}} / c_{\text{max}}) Y(t) \). This implies
\[
\| u_0(t) \| \leq c_{\text{max}} / \lambda_{\text{min}} \| u_0(0) \| e^{-2(\lambda_{\text{min}} / c_{\text{max}}) t}
\]
meaning that from any initial condition, the system tends to its equilibrium state exponentially fast.

IV. FLUID APPROXIMATION

Consider a wireless network in which new packets of the same size may randomly arrive at different nodes, destined for any other node, potentially several hops away. The network operates in slotted time \( n \), and is described by a simple directed connectivity graph \( (\mathcal{V}, \mathcal{E}) \). Packets of the same destination form a class (regardless of their sources) and every node holds a separate queue for each class \( d \in K \subseteq \mathcal{V} \).

To make a one-to-one correspondence between wireless networking and graph thermodynamics, we endow each wireless link \( ij \) with a profit-factor \( c_{ij}^{d}(n) \) that represents the profit — reciprocal of the cost — of sending one \( d \)-packet over the link at slot \( n \). It may reflect different topology-based routing penalties, e.g. channel quality, routing distance, relaying cost, power usage, etc. We also endow each node \( i \) with a capacity \( c_i^{d}(n) \) for the class \( d \) at slot \( n \), which may reflect buffer capacity, weights for prioritizing classes, etc.

Due to interference, not all wireless links can transmit at the same time, where the restriction on concurrent transmissions is defined by an interference model. Define a maximal schedule as a set of links such that no two links interfere with each other, and no more link can be added to it without violating the constraints of the interference model. We describe a maximal schedule with a scheduling vector \( \pi \in \{0,1\}^{\mathcal{E}} \) where \( \pi_{ij} \) takes the value 1 if the link \( ij \) is included, and 0 otherwise. Given a connectivity graph \( (\mathcal{V}, \mathcal{E}) \), we also define the scheduling set \( \mathcal{I} \) as the collection of all maximal scheduling vectors.

Due to environmental factors and user mobility, topology of the wireless network may randomly change in time. We assume that the sets \( \mathcal{V} \) and \( \mathcal{E} \) change far slower than link states, thus we can take them fixed during the time of our interest. Then, a temporarily unavailable link is characterized by zero link capacity. Persistent variations, due to e.g. non-local mobility, can be caught in a long scale regime that updates connectivity graph \( (\mathcal{V}, \mathcal{E}) \). We also assume that link states remain fixed during a timeslot, while they may change across slots according to some (unknown) probability laws.

Let a stochastic process \( S(n) = (S_1(n), \ldots, S_{|\mathcal{E}|}(n)) \) represent link states at slot \( n \), describing uncontrollable properties of wireless channels that affect channel capacities, link profit-factors, and node capacities. We assume that \( S(n) \) evolves according to an ergodic stationary process and takes values in a finite, but arbitrarily large, set \( S \). Thus by Birkhoff’s ergodic theorem, each state \( S \in S \) has a probability given by
\[
s := P\{S(n) = S\} = \lim_{n \to \infty} \frac{1}{n} \sum_{n=0}^{n-1} I_{S(n) = S} \tag{29}
\]
where \( \sum_{S \in S} I_S = 1 \) and \( I_x \) is the indicator function taking the value 1 when the statement \( x \) is true, and 0 otherwise.

Definition 2: Given a stochastic time-slotted process \( x(n) \), its expected time average value is defined as
\[
\tau := \lim_{n \to \infty} \frac{1}{n} \sum_{n=0}^{n-1} E\{x(n)\} \tag{30}
\]
where the definition is extended entrywise to vector functions.

A. State Space Model of Multiclass Networks

Let \( q_i^{d}(n) \) represent the integer number of \( d \)-packets in the node \( i \) at slot \( n \), and assume the backlog of \( d \)-packets in the destination node \( d \) is 0 for all \( d \in K \). The network state variables are given by a time-slotted hyper-vector \( q_0(n) \) that is conformably structured as the continuous-time hyper-vector of heats \( q_0(t) \) in (23). Specifically, \( q_i^{d}(n) \equiv 0 \) is dropped from the set of state variables. Also let a stochastic process \( a_i^{d}(n) \) represent the integer number of exogenous \( d \)-packets arriving into the node \( i \) at slot \( n \). Discarding \( a_i^{d}(n) \equiv 0 \), we form the time-slotted hyper-vector of node arrivals \( a_0(n) \) with a structure conformable with the continuous-time hyper-vector of heat sources \( a_0(t) \) in (23).

For a link \( ij \in \mathcal{E} \), we define actual-transmission \( f_{ij}^{d}(n) \) as the integer number of \( d \)-packets genuinely sent over the link at slot \( n \). Notice the contrast with the link capacity \( \mu_{ij}(n) \) that counts the maximum integer number of packets the link can transmit at slot \( n \), which is fre-
quently called link transmission rate in literature. Specifically, while link capacities vary with channel states, link actual-transmissions are assigned by a network control policy subject to $0 \leq f_{ij}(n) \leq \min\{q_{ij}(n), \mu_{ij}(n)\}$. Let us form the time-slotted hyper-vector of link actual-transmissions $f(n)$ with a structure conformable with the continuous-time hyper-vector of heat flows $f(t)$ in (23).

Using the above notation, the $f$-controlled state dynamics of a stochastic time-slotted wireless network is captured by

$$q_{c}(n+1) = q_{c}(n) + a_{c}(n) - B_{c}f(n) \tag{31}$$

where $B_{c}$ is the multiclass basis incidence matrix on the connectivity graph $(\mathcal{V}, \mathcal{E})$ as defined in (24).

B. Fluid Limit versus Fluid Model

For a continuous-time stochastic process, its fluid limit is the limit dynamics of the process obtained by scaling in time and amplitude, called fluid scaling. Under very mild conditions, it is shown that these scaled trajectories converge to a set of deterministic dynamical equations called fluid model. Using this deterministic model, one can analyze the rate-level, rather than packet-level, behavior of the original stochastic process. For the details, refer to [22], [23] and references therein.

Fluid limit: Let $X(\omega, t)$ be a realization of continuous-time stochastic process $X$ along a sample path $\omega$. Define the scaled process $X^{\tau}(\omega, t) := X(\omega, rt)/r$ for any $r > 0$. A deterministic function $X(t)$ is a fluid limit if there exist a sequence $r$ and a sample path $\omega$ such that $\lim_{r \to \infty} X^{\tau}(\omega, t) \to X(t)$ uniformly on compact sets. For a stable flow network, the existence of fluid limits is guaranteed if exogenous arrivals are of finite variance. It is further shown that each fluid limit is Lipschitz-continuous, and so differentiable, almost everywhere with respect to Lebesgue measure on $[0, \infty)$ [22], [23].

Cumulative process: To develop a continuous-time approximation, we model the wireless network by its cumulative processes. Let $a^{\text{tot}}(n)$ and $f^{\text{tot}}(n)$ be the hyper-vector of cumulative node arrivals and cumulative link actual-transmissions up to the slot index $n$. Assuming $a^{\text{tot}}(0) = 0$ and $f^{\text{tot}}(0) = 0$,

$$q_{c}(n) = q_{c}(0) + a_{c}(n) - B_{c}f^{\text{tot}}(n).$$

Let $f_{ij}(n)$ represent the predicted number of d-packets the link $ij$ would transmit if it were activated in the slot $n$, and form the hyper-vector $\overline{f}(n)$ conformably structured as $f(n)$. Also let $T_{\pi}(n)$ represent the cumulative number of timeslots in which the scheduling vector $\pi \in \Pi$ has been selected. Assuming $T_{\pi}(0) = 0$, it is easy to verify that $f^{\text{tot}}(n)$

$$= \sum_{\pi \in \Pi} \sum_{k=1}^{n} (T_{\pi}(k) - T_{\pi}(k-1)) \left(1_{[k]} \otimes \pi \right) \otimes \overline{f}(k).$$

The first parenthesis in the sum is equal to 1 if the scheduling vector $\pi$ is selected at slot $k$, and 0 otherwise. The term $(1_{[k]} \otimes \pi)$ extends the scheduling vector $\pi \in \mathbb{R}^{[k]}$ to be used in a multiclass fashion. Then its entrywise product with $\overline{f}(k)$ reads the number of packets that could be sent from each class over each link if the scheduling vector $\pi$ was selected.

General equations: Given a sample path $\omega$, we extend a time-slotted process to be continuous-time via linear interpolation in each interval $(n, n+1)$. Let exogenous arrivals occur at the beginning of each timeslot, so that $a^{\text{tot}}(t)$ represents the cumulative arrivals by the time $t$. Assuming normalized timeslots with the period of time unit, we obtain a set of approximate continuous-time stochastic basic equations as

$$q_{c}(t) = q_{c}(0) + a^{\text{tot}}(t) - B_{c}f^{\text{tot}}(t) \tag{32}$$

$$f^{\text{tot}}(t) = \sum_{\pi \in \Pi} T_{\pi}(t) \left(1_{[k]} \otimes \pi \right) \otimes \overline{f}(t) \tag{33}$$

$$T_{\pi}(t) = \begin{cases} 1 & \text{if } \pi \text{ is chosen at time } t \\ 0 & \text{otherwise} \end{cases} \tag{34}$$

$$\sum_{\pi \in \Pi} T_{\pi}(t) = 1 \text{ with } T_{\pi}(t) \text{ nondecreasing.} \tag{35}$$

The equality (33) entails the existence of a $\delta > 0$ such that $f^{\text{tot}}(t') - f^{\text{tot}}(t) = \sum_{\pi \in \Pi} T_{\pi}(t')(T_{\pi}(t') - T_{\pi}(t))$ for any $t' \in [t, t + \delta]$. This states the fact that if a link has a positive flow of d-packets at time $t$, the number of d-packets transmitted by the link in the interval $[t, t'] \subset [t, t + \delta]$ is equal to amount of time it has been activated during $[t, t']$ multiplied by its transmission rate prediction at time $t$. To obtain a deterministic first-order approximation, all stochastic inputs and parameters are replaced by their long-time average values, which equal the expected time average values for ergodic stationary processes. Therefore,

$$a^{\text{tot}}(t) = \overline{a}_{c} t. \tag{36}$$

Likewise, one replaces the vector of link capacities $\mu(n)$ with $\overline{\mu}$, the hyper-vector of link profit-factors $\overline{\sigma}(n)$ with $\overline{\sigma}$, and the hyper-vector of node capacities $\overline{c}(n)$ with $\overline{c}$. The existence of $\overline{\mu}$, $\overline{\sigma}$, and $\overline{c}$ is assured by (29), precisely as

$$\overline{\mu} = \lim_{\tau \to \infty} \sup_{\tau} \frac{1}{\tau} \sum_{n=0}^{\tau-1} \sum_{s \in S} E \{ \mu(n) | S(n) = S \}$$

$$\overline{\sigma} = \lim_{\tau \to \infty} \sup_{\tau} \frac{1}{\tau} \sum_{n=0}^{\tau-1} \sum_{s \in S} E \{ \sigma(n) | S(n) = S \}$$

$$\overline{c} = \lim_{\tau \to \infty} \sup_{\tau} \frac{1}{\tau} \sum_{n=0}^{\tau-1} \sum_{s \in S} E \{ c(n) | S(n) = S \}.$$

Note that the existence of these average values does not imply that they are known to the network controller.

Particular equations: At every timeslot, a network controller makes a routing decision consists of activating a particular set of non-interfering links and assigning a specific number of packets that each activated link should transmit from different classes. Translating into our general model, this control action is equivalent to modifying $T_{\pi}(n)$ and assigning $f(n)$ at each slot $n$. Therefore, while equations (32)–(36) always hold for any finite-queue network operating under an arbitrary non-idling control policy, each policy takes its own particular way to determine timeslot control values $\overline{f}(n)$ and $T_{\pi}(n)$ leading to link actual-transmissions $\overline{f}(n)$. Obviously, this in turn shapes the forms of continuous-time approximations $\overline{f}(t)$ and $T_{\pi}(t)$ that lead to $f(t)$ for a specific control policy.

Theorem 4: Consider a wireless network stabilized under a control policy that assigns $\overline{f}(n)$ and $T_{\pi}(n)$ at each slot $n$. Define network fluid model as the collection of deterministic general equations (32)–(36) together with a set of particular equations that determine $\overline{f}(t)$ and $T_{\pi}(t)$ as the linear approximations of time-slotted variables $\overline{f}(n)$ and $T_{\pi}(n)$. Then every network fluid limit satisfies the network fluid model.

Proof: The proof follows the same line of argument proposed by [22, Theorem 2.3.2] or [23, Proposition 4.12] and is omitted for brevity.
Note that for a time-slotted stochastic process, the fluid limit is the scaled process of the first-order continuous-time approximation for an arbitrary realization, while the fluid model is a set of fully deterministic continuous-time equations. Now consider a wireless network under a stabilizing control policy, where the time-slotted stochastic processes $q_\sigma(n)$, $f^{\text{tot}}(n)$ and $T_\pi(t)$ have their corresponding continuous-time fluid limits. The theorem claims that for large scaling factors, the fluid limit of every realization converges to a set of deterministic continuous-time variables $q_\sigma(t)$, $f^{\text{tot}}(t)$ and $T_\pi(t)$ that solve the set of deterministic fluid model equations. Thus the fluid limit is independent of sample path.

V. THERMODYNAMIC MULTICLASS ROUTING POLICIES

Consider a multihop multiclass wireless network subject to an arrival rate stabilized by a control policy. In steady-state, there exist fluid limits for the stochastic variables of each class $d$. This section introduces a family of network control policies under which the fluid limits of every realization take the form of nonlinear thermal diffusion on the underlying directed graph with suitably-weighted and -capacitated edges. Hence, for the first time in literature, we propose a genuine diffusion process on stochastic multiclass wireless networks, or processing networks for that matter, subject to interdependent resources (link interference) and time-varying topology.

Definition 3: Given a stochastic multiclass wireless network, we define its deterministic thermal model with a capacitated directed graph charged by a set of multiclass heat sources as follows. The thermal graph has the same incidence matrix as that of wireless network. Edge capacities on the thermal graph equal the expected time average capacity of links on the wireless network. For each class, heat sink on the thermal graph is the destination node on the wireless network. For each class, intensity of heat sources on the thermal graph equal the expected time average rate of packet arrivals on the wireless network. For each class, thermal conductivity of edges on the thermal graph equal the expected time average capacity of nodes on the wireless network.

Theorem 5: Consider a stochastic multiclass wireless network with a deterministic thermal model in the sense of Def. 3. Define the target trajectory as the multiclass diffusion dynamics on the thermal graph, captured by (25)–(28). Suppose that at every timeslot, subject to wireless network constraints, a control policy maximizes the $f$-controlled functional
\[
D(f, u_0, n) := 2 f(n)^T \text{diag}(\sigma(n)) B_o^T u_o(n) - f(n)^T f(n)
\]
where $u_0 := \text{diag}(c_0(n))^{-1} q_0(n)$. Then every fluid limit of wireless network under such a control policy asymptotically converges to the above-defined target trajectory.

Proof: It is given in the next subsection.

In Th. 5, though the functional $D(f, u_0, n)$ has a unique maximum value, the control policy that leads to this maximum is not necessarily unique. Therefore, rather than having a unique optimal control policy, there indeed exists a family of optimal control policies. Specifically, in [24] we have developed a control policy that solves $D$ maximization problem using a local link weighing followed by a global max-weight scheduling. One important merit of such a policy is that it has the same algorithmic structure, complexity, and overhead as BP policy, giving them the same wide-reaching impact.

To conceptualize Th. 5, consider the simpler case of a uniclass wireless network under a $D$-maximizing control policy. From a packet-level point of view, at every timeslot, the policy activates a particular set of links and transmits a specific number of packets over them. Thus each link may transmit packets at some slots while being switched off at some other slots. Now consider the expected time average flow of packets on links and the expected time average length of queues at nodes. The theorem claims that looking at this rate-level behavior, it complies with the thermal diffusion process on the underlying directed graph. In a multiclass case, on the other hand, each class has its own average characteristics. Under a $D$-maximizing control policy, the theorem claims that looking at the rate-level behavior of each individual class, it still follows the thermal diffusion process, where the edge capacities are divided among different classes according to (28). This concept is formalized by the next corollary.

Corollary 1: Consider a stochastic multiclass wireless network with the vector of link capacities $\mu(n)$ and the hyper-vectors of link profit-factors $\sigma(n)$ at timeslot $n$. Suppose that at every timeslot, a control policy maximizes the $D(f, u_0, n)$ functional as defined in Th. 5, and let the hyper-vector $\theta$ with $\sum_{d \in K} \theta^{(d)} \leq 1 |\varepsilon|$ represent the portion of expected time average capacity of each link devoted to each class. Then the rate-level behavior of the wireless network is described by the steady-state thermal diffusion equations
\[
\bar{f} = \min \left\{ \text{diag}(\sigma)(B_o^T \bar{u}_o)^+, \theta \odot (1_{|K|} \otimes \mu) \right\}
\]
\[
\theta = \arg \min_\theta \| \text{diag}(\sigma) B_o \bar{u}_o - f \|_F
\]
\[
\bar{u}_o = B_o \min \left\{ \text{diag}(\sigma)(B_o^T \bar{u}_o)^+, \theta \odot (1_{|K|} \otimes \mu) \right\}.
\]

A. Proof of Theorem 5

To prove Th. 5, we exploit the Lyapunov argument using
\[
M_\sigma(n) := (B_o B_o^T)^{-1} B_o \text{diag}(\sigma(n)) B_o^T 
\]
which, contrary to tradition, is a non-symmetric matrix. When all links are of the same profit-factor, i.e. $\sigma(n) = \alpha I_{|\varepsilon|} |K|$ for a scalar $\alpha > 0$, then $M_\sigma(n)$ reduces to $\sigma(n)$.

Lemma 1: On a connected wireless network, $M_\sigma(n)$ is quasi-positive in the sense that $x^T M_\sigma(n) x \geq 0$ for any vector $x$ and for all $n$, with equality if and only if $x = 0$. Further,
\[
B_o^T M_\sigma(n) x = \text{diag}(\sigma(n)) B_o^T x, \forall x \in \mathbb{R}^{(|\varepsilon| - 1) |K|}.
\]

Proof: Let $\Sigma(n) := \text{diag}(\sigma(n))$. First we show that the matrix $B_o \Sigma(n) B_o^T$ is symmetric positive definite. Since the network is connected, any sub-matrix obtained from its incidence matrix by removing one or more rows has full row rank [19], i.e. $B_o^{(d)}$ has full row rank for any $d \in K$. Accordingly, $B_o$ has full row rank, implying that $B_o \Sigma(n) B_o^T$ is full rank, and so symmetric positive definite.

As $B_o B_o^T$ is symmetric positive definite, so is its inverse. Thus $M_\sigma(n)$ is a product of two positive definite matrices, and thus is similar to $\left( B_o B_o^T \right)^{-1/2} M_\sigma(n) \left( B_o B_o^T \right)^{-1/2}$, which is equal to $\left( B_o B_o^T \right)^{-1/2} B_o \Sigma(n) B_o^T \left( B_o B_o^T \right)^{-1/2}$.
latter is congruent to $B_x\Sigma(\eta)B_x^\top$, and so positive definite. Thus $M_\eta(\eta)$ has positive eigenvalues for all $\eta$, meaning that $x^\top M_\eta(n)x > 0$ for any nonzero vector $x$.

Now we show that $B_x^\top B_x(\eta)M_\eta(n)x = \Sigma(\eta)B_x^\top x$ for any vector $x$. Starting with $M_\eta(\eta) = (B_x^\top B_x)^{-1}B_x\Sigma(\eta)B_x^\top$, it implies $B_x^\top B_x(\eta)M_\eta(n)x = B_x\Sigma(\eta)B_x^\top x$ as $B_x^\top B_x$ is full rank. Let us assume $B_x^\top B_x(\eta)M_\eta(n)x \neq \Sigma(\eta)B_x^\top x$. Since $B_x^\top B_x(\eta)M_\eta(n)x = B_x\Sigma(\eta)B_x^\top x$ must be true for any $x$, under our assumption both $B_x^\top B_x(\eta)M_\eta(n)x$ and $\Sigma(\eta)B_x^\top x$ must belong to the null space of $B_x$, leading to $B_x^\top B_x(\eta)M_\eta(n)x = B_x\Sigma(\eta)B_x^\top x = 0$. But this is in contradiction with the positive definiteness of $B_x\Sigma(\eta)B_x^\top$, implying that the assumption of $B_x^\top B_x(\eta)M_\eta(n)x \neq \Sigma(\eta)B_x^\top x$ cannot be true.

**Lemma 2:** For arbitrary vectors $x, y \in \mathbb{R}^{[\nu]-1|\xi|}$,

$$x^\top (M_\eta(n) + M_\eta(\eta)) y \leq \eta x^\top M_\eta(\eta) y$$

for a scalar $\eta$ that takes the value 1 if $x^\top M_\eta(\eta) y \leq 0$ and the value 3 if $x^\top M_\eta(\eta) y > 0$.

**Proof:** Replacing $M_\eta + M_\eta$ with $2M_\eta + (M_\eta - M_\eta)$, we need to prove that for arbitrary vectors $x, y \in \mathbb{R}^{[\nu]-1|\xi|}$,

$$2x^\top M_\eta(n)y + x^\top (M_\eta(n) - M_\eta(n)) y \leq \eta x^\top M_\eta(\eta) y$$

for $\eta$ being either 1 or 3 depending on if $x^\top M_\eta(\eta) y$ is either negative or positive. This is equivalent to prove that

$$x^\top (M_\eta(n) - M_\eta(n)) y \leq (\eta - 2) x^\top M_\eta(\eta) y.$$

Considering that $\eta$ switches between 1 and 3 corresponding to the sign of $x^\top M_\eta(\eta)x$, it is easy to see that to prove the above inequality, it is sufficient to prove that

$$|x^\top (M_\eta(n) + M_\eta(\eta)) y| \leq |x^\top M_\eta(\eta) y|.$$

Dropping $(\eta)$ for brevity, the latter is equivalent to

$$x^\top (M_\eta(n) - M_\eta(n)) y y^\top (M_\eta(n) - M_\eta(n)) x \leq x^\top M_\eta(\eta) y y^\top M_\eta(\eta) x$$

which by some matrix algebra can be transformed to

$$x^\top (2M_\eta(n) - M_\eta(n)) y y^\top M_\eta(\eta) x \geq 0. \tag{39}$$

First observe that as $M_\eta - M_\eta$ is a skew-symmetric matrix, $z^\top (2M_\eta(n) - M_\eta(n)) z = 0$ for an arbitrary vector $z$. Therefore, $z^\top (2M_\eta(n) - M_\eta(n)) z = z^\top M_\eta(\eta) z$ which shows that $2M_\eta(n) - M_\eta(n)$ is a quasi-positive matrix in the sense of Lem. 1.

Now we show that for any quasi-positive matrix $A$, there exist two symmetric positive definite matrices $L$ and $R$ such that $A = LR$. As a square matrix, $A$ has a Jordan decomposition as $A = VJW^{-1}$ where $J$ is a Jordan form with positive diagonal entries, as $A$ has all eigenvalues positive. Then choose $L = VEV^\top$ and $R = (V^{-1})^\top E^{-1}JV^{-1}$ where $E$ is a symmetric positive definite matrix commuting with $J$. Since $L$ is congruent to $E$, it has positive eigenvalues, and so it is a symmetric positive definite matrix. Likewise, since $R$ is congruent to $E^{-1}J$, it has positive eigenvalues, and so it is a symmetric positive definite matrix. To see why $E^{-1}J$ has positive eigenvalues, observe that it is similar to

$$E^{-1/2}E^{-1/2}J = E^{-1/2}E^{-1/2}J$$

and so congruent to $J$ that has positive eigenvalues.

Turing back to (39), let $D := (2M_\eta(n) - M_\eta(n)) y y^\top M_\eta$. Since both $2M_\eta - M_\eta$ and $M_\eta$ are quasi-positive, they can be decomposed into the product of two symmetric positive definite matrices as $2M_\eta - M_\eta = L_1R_1$ and $M_\eta = L_2R_2$ that leads to $D = L_1R_1 y y^\top L_2$.

Applying this rule iteratively, it is seen that $D$ is congruent to $y y^\top$, and therefore both of them has the same number of negative eigenvalues. But $y y^\top$ is a symmetric positive semi-definite matrix for any vector $y$ and so has no negative eigenvalue, meaning that $D$ also has no negative eigenvalue. Then it is obvious that $x^\top D x \geq 0$, concluding (39).

Getting back to the proof of Th. 5, first observe that in the thermal model of Def. 3, the heat sources are of fixed values. Thus by Th. 3, the thermal model reaches its equilibrium point exponentially fast. By Th. 4, on the other hand, every fluid limit satisfies the fluid model. Thus to prove Th. 5, it is sufficient to show that both of the thermal model and the fluid model have the same behavior in steady-state condition, while they can possibly be different in transients. Let $q^*_t$ and $f^*$ represent the equilibrium point for the thermal model. Also let $q_0(t)$ and $f(t)$ represent the fluid model variables of a $D$-maximizing control policy. In the sequel, we show that $q_0(t)$ and $f(t)$ respectively converge to $q^*_t$ and $f^*$ in limit.

Consider the continuous-time Lyapunov function

$$Y(t) := \langle q_0(t) - q^*_t \rangle^\top \overline{M}_0 \text{diag}(\overline{c})^{-1} (q_0(t) - q^*_t) \tag{40}$$

where overbar notation denotes the expected time average value. Note that all eigenvalues of $\overline{M}_0$ are strictly positive by Lem. 1 and so $Y(t)$ is positive-definite, and radially unbounded with respect to $\langle q_0(t) - q^*_t \rangle$. Taking time derivative from (40), using the equalities $q_0(t) = \text{diag}(\overline{c})u_0(t)$ and $q^*_t = \text{diag}(\overline{c})u^*_t$, and noting that $q^*_t = 0$ as the equilibrium point of equations (25)–(28) is a fixed point, we obtain

$$\dot{Y}(t) = \dot{q}_0(t)^\top (\overline{M}_0^\top + \overline{M}_0) (u_0(t) - u^*_t). \tag{41}$$

Exploiting Lem. 2 in the latter leads to

$$\dot{Y}(t) \leq \eta \dot{q}_0(t)^\top \overline{M}_0 (u_0(t) - u^*_t) \tag{42}$$

where $\eta$ takes the value either 1 or 3 depending on if the functional $\dot{q}_0(t)^\top \overline{M}_0 (u_0(t) - u^*_t)$ is either negative or positive. Since $\eta$ is a positive coefficient, it has no impact on the Lyapunov argument and can simply be omitted, but for the sake of completeness we will keep it in our statements.

Plugging (36) in (32) and taking time derivative lead to

$$\dot{q}_0(t) = \overline{w}_0 - B_x f^{\text{tot}}(t). \tag{43}$$

Note in (33) that the entry of $f(t)$ corresponding to link $ij$ and class $d$ specifies the number of $d$-packets the link will send per unit time if it is activated at time $t$. Then $f(t)$ identifies the vector of rate of actual-transmissions realized at time $t$. Now assume that the entry of $f(t)$ corresponding to link $ij$ and class $d$ at time $t$ is equal to $x \geq 0$, meaning that at time $t$ the link transmits $x$ number of $d$-packets per unit time. Then it should be obvious that the same entry of $f^{\text{tot}}(t)$ at time $t$ must be also equal to $x$. This is seen more formally from

$$f^{\text{tot}}(t) = \lim_{\delta \to 0} \frac{f^{\text{tot}}(t + \delta) - f^{\text{tot}}(t)}{\delta} = f(t) \tag{44}$$

in light of $\lim_{\delta \to 0} f^{\text{tot}}(t + \delta) = f^{\text{tot}}(t)$ and $f(t)$.

At the equilibrium point of the thermal model, we have $\overline{w}_0 = L_0(u^*_t) = B_x f^*$, since the derivatives of temperatures
vanish. Plugging this and (44) into (43), we get
\[ \dot{q}_c(t) = B_c f^* - B_c f(t). \]
Then substituting (45) for \( \dot{q}_c(t) \) in (42) yields
\[ \eta^{-1} \dot{Y}(t) \leq (B_c f^* - B_c f(t))^\top \Sigma \Sigma (u_c(t) - u_c^*). \]
Next, using the equality (38) with \( \Sigma := \text{diag}(\sigma) \) lead to
\[ \eta^{-1} \dot{Y}(t) \leq (f^* - f(t))^\top \Sigma B_c^\top u_c(t) - u_c^*. \]
Multiply two sides by 2, and on the right-hand side add and subtract \( f(t)^\top f(t) \) to get
\[ 2 \eta^{-1} \dot{Y}(t) \leq \left[ (2 f(t)^\top \Sigma B_c^\top u_c(t) - f(t)^\top f(t)) + \| \Sigma B_c^\top u_c^* - f^* \|^2 \right]. \]

We interpret (46a) and (46b) through the fluid model with the queues \( \dot{q}_c(t) = \text{diag}(\Sigma) u_c(t) \), while (46c) is interpreted through the thermal model with the temperatures \( u_c^* \).

Observe that \( 2 f(t)^\top \Sigma B_c^\top u_c(t) - f(t)^\top f(t) \) is the fluid model counterpart of \( D \) function defined in Th. 5. This means that under a D-maximizing routing policy, the control values \( f(t) \) and \( T_c(t) \), which lead to \( f(t) \), are designed to maximize \( 2 f(t)^\top \Sigma B_c^\top u_c(t) - f(t)^\top f(t) \) at any time \( t \) given queue backlogs \( q_c(t) \). Thus, on the right-hand side of (46), the summation of (46a) and (46b) is nonpositive. On the thermal graph, on the other hand, (28) ensures that the first norm in (46a) is not greater than the second norm; thus (46a) is also nonpositive. This leads to \( 2 \eta^{-1} \dot{Y}(t) \leq 0 \) that, as \( \eta \) takes the value either 1 or 3, is equivalent to \( \dot{Y}(t) \leq 0 \).

Now consider a joint solution \( (q_c(t), q_c^*) \) to the fluid model of a D-maximizing control policy and to the thermal model of Def. 3. Let \( \Omega \) be the largest invariant set in the set of all points for which \( \dot{Y}(t) = 0 \). Since \( \dot{Y}(t) \) is a radially unbounded positive definite function with \( \dot{Y}(t) \leq 0 \), LaSalle’s Invariance Principle implies that every joint solution \( (q_c(t), q_c^*) \) asymptotically converges to \( \Omega \). Next we show that the invariant set \( \Omega \) contains only the point \( (q_c(t) = q_c^*, q_c^*) \).

If \( \dot{Y}(t) = 0 \), then since \( \Sigma B_c^\top + \Sigma I_c^\top \) is symmetric positive-definite, (41) entails \( u_c(t) = u_c^* \) and/or \( \dot{q}_c(t) = 0 \), where the former is equivalent to \( q_c(t) = q_c^* \). We also have \( q_c^* = 0 \), since \( q_c^* \) is a fixed equilibrium point, which means that \( q_c(t) = q_c^* \) leads to \( \dot{q}_c(t) = 0 \) too. Now if \( \dot{q}_c(t) = 0 \), then (45) entails \( B_c f(t) = B_c f^* \) which implies \( B_c f(t) = B_c f^* \) as \( t \to \infty \). Noting that \( f^* \) cannot be in the null space of \( B_c \), since \( B_c f^* = \Sigma f^* \neq 0 \), it equally implies that
\[ \lim_{t \to \infty} f(t) = f^*. \]
It also implies \( B_c f(t)^\top = B_c f^\top + k \) for a constant vector \( k \), as the cumulative flows are the integral of instant flows. Now let us take a limit from (32) to get
\[ \lim_{t \to \infty} q_c(t) = q_c(0) + \lim_{t \to \infty} (a_c(t) - B_c f(t)) \]
Since both the fluid model and the thermal model have the same \( \Sigma \) as the input, we have \( a_c(t) = a_c(t) \). Therefore,
\[ \lim_{t \to \infty} q_c(t) = q_c(0) + \lim_{t \to \infty} (a_c(t) - B_c f(t)) + k. \]
The rightmost limit above is found on the thermal graph by taking integral from 0 to \( \infty \) from (27), which transfers the equality to \( \lim_{t \to \infty} q_c(t) = q_c^* + (q_c^*(0) - q_c^*(0) + k) \) with \( q_c^*(0) \) the initial heat quantities. This means that \( q_c(\infty) \) equals \( q_c^* \) up to a bias \( k' \). Next we show that \( k' \) has to be 0.

Note that given \( u_c(t) \), a D-maximizing control policy chooses \( f(t) \) such that to maximize \( 2 f(t)^\top \Sigma B_c^\top u_c(t) - f(t)^\top f(t) \) at any time \( t \). At the same time, (47) implies that the \( f(t) \) in a D-maximizing policy approaches \( f^* \) when \( t \to \infty \), meaning that \( f^* \) maximizes \( 2 f^*^\top \Sigma B_c^\top u_c(\infty) - f^*^\top f^* \) for a given \( u_c(\infty) \). This equally means that \( f^* \) maximizes \( -u_c(\infty)^\top B_c \Sigma^2 B_c^\top u_c(\infty) + 2 f^*^\top \Sigma B_c^\top u_c(\infty) - f^*^\top f^* \) for a given \( u_c(\infty) \), because adding an \( f^* \)-independent term such as \( -u_c(\infty)^\top B_c \Sigma^2 B_c^\top u_c(\infty) \) does not change the maximization problem. The latter is equivalent to the minimization of \( \| \Sigma B_c^\top u_c(\infty) - f^* \| \). On the thermal graph, on the other hand, (28) entails that \( f^* \) also minimizes \( \| \Sigma B_c^\top u_c^* - f^* \| \). Considering these two together, we get \( B_c^\top u_c(\infty) = B_c^\top u_c^* \) or equally \( B_c^\top u_c(\infty) = B_c^\top q_c^* \). Multiplying both sides by \( B_c \) and considering that \( B_c B_c^\top \) is positive definite and so invertible, we obtain
\[ \lim_{t \to \infty} q_c(t) \to q_c^*. \]
Therefore, if \( \dot{Y}(t) = 0 \), then (48). This entails that the invariant set \( \Omega \) has only one element \( (q_c(\infty) = q_c^*, q_c^*) \), which ensures that the limits (47) and (48) are globally valid.

VI. STABILITY OF THERMODYNAMIC ROUTING POLICIES

In queuing systems, rate stability and strong stability are two common types of stability in literature. A stochastic discrete-time process \( x(n) \) is rate stable if \( \lim_{n \to \infty} x(n)/n = 0 \) with probability 1, and is strong stable if its expected time average value is finite, \( \bar{x} < \infty \). It is shown that under mild assumptions, strong stability implies rate stability [14].

A queuing network is stable if all its queues are stable. An arrival rate is called stabilizable if there exists a control policy — possibly unfeasible — that can make the network stable under that arrival rate. The network capacity region \( C \) is defined as the closure of the set of all stabilizable arrival rates, which is invariant — as it is a closure — under either rate or strong stability definition. It is shown that if arrival rates are outside of the set \( C \) then no control policy can stabilize the network under either rate or strong stability definition; and conversely, if they are interior to \( C \) then there always exists a control policy to strongly stabilize the network [14].

For a stochastic arrival vector \( a_c(n) \) to be in the network capacity region \( C \), the necessary and sufficient condition is the feasibility of the conventional multiclass network flow problem defined by replacing \( a_c(n) \) with \( \text{ave}(a_c) \) and \( \mu(n) \) with \( \text{ave}(\mu) \). Under ergodic stationary assumption for stochastic processes, this is equivalent to the existence of a set of link actual-transmissions \( f(n) \) such that
\[ \bar{a}_c = B_c f \]
\[ \sum_{d \in K} f_{ij}(d) \leq \bar{m}_{ij} \]
for all \( ij \in E \).

A control policy is called throughput-optimal if it secures network stability under all stabilizable arrival rates — interior to the capacity region. For the family of control policies in Th. 5, throughput optimality under rate stability definition is the immediate result of Th. 3. A stronger result, however, is formalized by the next theorem that guarantees strong stability
over the entire capacity region. To simplify the proof, we assume stochastic processes to be independently and identically distributed (i.i.d.) with the understanding that the result can easily be extended to non-i.i.d. systems with stationary ergodic processes of finite mean and variance.

**Theorem 6:** Consider a wireless network with arrivals and channel states being i.i.d. random variables over timeslots and with respect to each other. Then the family of network control policies in Th. 5 are strong throughput-optimal.

**Proof:** It follows the same line of argument proposed for the proof of Theorem 2 in [16] and is omitted for brevity.

**VII. CONCLUSION**

A recently-introduced paradigm in [16], named “wireless network thermodynamics,” was enhanced not to need a restrictive assumption in [16] on the one hand, and to work with multiclass networks on the other hand. The enhancement was established on two pillars: (i) developing the notion of multiclass thermodynamics on capacitated directed graphs, and (ii) introducing a family of control policies under which the long-time average flow of packets on wireless network concurs with the multiclass graph thermodynamics. A natural impact is that under this paradigm, one can employ tools from heat calculus, a very active area of pure mathematics, in the analysis and optimization of interfering multiclass queuing networks. Besides this, we showed that the family of control policies in (ii) are throughput optimal. Also worth to mention that they have several other prominent qualities such as minimizing average network delay in a class of routing algorithms, minimizing “Dirichlet routing cost,” providing fast transient performance, and having an electrical circuit equivalent [24].

**REFERENCES**