Competitive and Learned Online Algorithms for HVAC Demand Management in Thermal Networks

Adam Lechowicz
University of Massachusetts
Amherst, MA, USA
alechowicz@cs.umass.edu

Nicolas Christianson
Stanford University
Stanford, CA, USA
christianson@stanford.edu

Mohammad Hajiesmaili
University of Massachusetts
Amherst, MA, USA
hajiesmaili@cs.umass.edu

Adam Wierman
California Institute of Technology
Pasadena, CA, USA
adamw@caltech.edu

Prashant Shenoy
University of Massachusetts
Amherst, MA, USA
shenoy@cs.umass.edu

Abstract

The electrification of heating and cooling is critical to global energy transition efforts. In recent years, fifth-generation district heating and cooling (5GDHC) thermal networks have emerged as a promising approach to electrify heating and cooling systems while mitigating grid impacts and repurposing fossil-fuel infrastructure, such as gas distribution lines. 5GDHC systems connect building communities via a shared underground loop of ambient-temperature fluid, using ground-source heat pumps at each building to provide heating and cooling. Existing deployments of these systems operate in a static manner, but dynamic operation can unlock flexibility by leveraging the network's thermal inertia to shift load and reduce costs. In this paper, we model the dynamic operation of a 5GDHC system as an online optimization problem, where an operator makes real-time decisions in the face of volatile electricity prices and uncertain demand. We propose TANDEM, a distributed online algorithm with provable (and optimal) competitive guarantees. We then extend TANDEM using an end-to-end framework that learns an improved online algorithm based on historical data while preserving worst-case guarantees. In a case study of ten real neighborhoods, we demonstrate that TANDEM achieves significant cost savings of 44.1% compared to static operation, 75.4% emissions reductions compared to gas heating, and 71.8% peak demand reductions compared to uncoordinated air-source heat pump adoption.

1 Introduction

Building operations account for 30% of global energy use and 26% of yearly greenhouse gas emissions [27], a large portion is due to heating, ventilation, and air conditioning (HVAC) [56], making HVAC a critical target for decarbonization efforts. Decarbonization strategies for HVAC typically focus on *electrification* (i.e., decommissioning furnaces and boilers that use fossil fuels), combined with a transition to low-carbon forms of electricity generation, such as solar PV, hydro, and wind [6, 24]. To replace fossil-fuel-based heating systems, electric heat pumps are widely considered the most promising technology, owing to their excellent energy efficiency compared to resistive electric heating [28]. The most common type used in residential and commercial buildings are air-source heat pumps (ASHPs), which extract heat from ambient (outdoor) air.

ASHPs have become popular with advances in cold-climate performance that make them viable in regions with harsh winters [35, 71]. However, even with efficiency advantages that ASHPs offer over other forms of electric heating, forecasts indicate widespread ASHP adoption in regions currently reliant on fossil fuels has the potential to increase peak electric demand significantly and shift it to winter periods [30, 31]. This raises concerns about the cost of necessary grid upgrades, as well as the feasibility of meeting wintertime peak demand using low-carbon sources alone [41, 53].

To mitigate the grid impacts of HVAC in general, a litany of *demand response* programs have been proposed, implemented, and studied [26, 40, 66]. These programs often consider direct control of HVAC systems (e.g., via smart thermostats) or price-based incentives that encourage users to shift energy consumption to off-peak periods [26, 63]. However, the realizable flexibility from these demand response programs is often limited—each participant home is an isolated unit with its own HVAC system, and occupant comfort constraints limit the extent to which heating or cooling loads can be shifted without negatively impacting residents [25, 79].

Since 2021, *fifth-generation district heating and cooling* (5GDHC) concepts have emerged as a promising way to decarbonize HVAC while mitigating grid impacts [21]. At the core of 5GDHC systems is a *thermal network* of underground pipes connecting a community of buildings via a shared ambient-temperature loop. 5GDHC networks have gained attention for their potential to replace or repurpose underground infrastructure such as natural gas distribution pipelines [10, 19]; they have gained significant traction in Europe [8, 19], and pilot projects are underway in the U.S. [9, 34].

In contrast to traditional district heating systems that directly deliver high temperature (70-120°C) heat to buildings, 5GDHC networks contain low temperature fluid (e.g., 5-25°C) and rely on *ground-source heat pumps* (GSHPs) to exchange heat with the network, providing heating or cooling as needed. GSHPs are more efficient than ASHPs because they extract heat from the stable and above-freezing temperatures of the underground thermal network—benefits are largest when outdoor air is very cold, corresponding to times when heating demand is highest and ASHP efficiency is lowest [60, 61]. Furthermore, compared to isolated GSHP installations (i.e., using ground loops or boreholes) at each building, 5GDHC networks offer a *shared infrastructure* that provides an additional benefit of *thermal inertia*—the large thermal mass of the fluid in the underground network allows for the storage and shifting of energy over time, providing opportunities for flexibility and load shifting.

In current deployments, 5GDHC thermal networks are operated in a static manner, maintaining, e.g., a constant flow rate through

boreholes at all times [67]. To fully unlock the flexibility offered by the network, however, a 5GDHC system must be *dynamically managed* over time, responding to changing conditions such as building demands and electricity prices. An operator's control levers include managing the influx and outflux of heat at pump stations connecting the network to energy sources or sinks (e.g., waste heat, boreholes, thermal storage, or auxiliary heat), as well as coordinating with flexible building demands (if any) to optimally utilize the thermal inertia of the network while respecting operational constraints.

A few prior works have proposed methods for dynamic control of thermal networks, often using machine learning or model predictive control [59, 73, 80]. However, these methods are seldom used in practice. One challenge for deployment is the inherent online nature of the control problem: the controller must contend with non-stationary intra-day inputs such as volatile electricity prices and uncertain demand timing/amounts while respecting safe operating constraints, posing significant difficulty for developing ML methods or heuristics that can handle distribution shift while satisfying constraints. Furthermore, compute is often limited, preventing the use of computationally intensive algorithms in real-time.

On the other hand, there is a rich body of work in the theoretical literature on online algorithms [7, 15, 43, 68, 78] studying related problems with similar challenges—for instance, [38] propose online algorithms for demand management in grid-integrated data centers with local energy storage, contending with volatile electricity prices and uncertain workloads. However, despite these similarities, there remains a gap between existing online algorithm models and the complexity of the problem faced by an operator in practice. Unlike the "clean" demand modeled by existing work (e.g., independent demands with sizes often known a priori), 5GDHC systems contend with significantly more complexity—demands are unknown a priori, and separate demand requests of heterogeneous types (e.g., heating at house A and cooling at house B) can interact with each other via the shared thermal network, making algorithm design challenging.

Contributions. To approach this problem, we introduce and study **online thermal network demand management** (OTNM), which models several features informed by practice. We make three main contributions, each corresponding to a central research question.

▶ *Can we design online algorithms for* OTNM *that manage the unique challenges of the 5GDHC setting while providing performance guarantees?* We answer this question in the affirmative by proposing TANDEM (Thermal Accounting for Network-Distributed Energy Mediation), a distributed online algorithm for OTNM. TANDEM extends the threshold-based design paradigm from problems such as online search to achieve the best-possible competitive ratio for OTNM under our assumptions, providing a strong worst-case guarantee.

▶ *When historical data is available, can we design data-driven algorithms for* OTNM *that learn from it to improve performance while preserving worst-case guarantees?* We extend TANDEM to incorporate an end-to-end learning framework proposed by recent work [38] that leverages historical data to learn an improved algorithm parameterization (e.g., thresholds governing TANDEM's actions) while preserving a bounded worst-case competitive ratio.

▶ *What practical impacts could demand management have on the operation of a 5GDHC system in realistic settings?* We present a detailed case study of 5GDHC systems in ten real neighborhoods in Massachusetts, using reduced-order thermodynamics to model the network and real-world data for heat demand and grid impacts. Both variants of TANDEM are effective for solving OTNM in practice, and their performance translates into cost savings of 44.1% compared to the 5GDHC status quo of constant operation, emissions reductions of 75.4% compared to natural gas heating, and reductions in daily peak electric load of 71.8% compared to uncoordinated ASHP operation. We also study the potential of data center waste heat integration, finding that sufficiently affordable waste heat is a viable substitute for borehole fields in 5GDHC operations.

2 Background

We start by providing background on heat pumps and the distribution grid, 5GDHC systems, and demand response programs.

Heat Pumps. In contrast to forms of heating that rely on heat generation (e.g., via combustion or resistive electric heating), electric heat pumps transfer heat from a source (e.g., air, ground, water) into an indoor space using a vapor-compression refrigeration cycle [28]. From an efficiency perspective, heat pumps achieve a *coefficient of performance* (COP) greater than 1, meaning they transfer multiple units of heat for each unit of electricity consumed. Heat pumps can also operate in "reverse," providing cooling by transferring heat out of a space. COP depends on the temperature difference between the heat source and output—as a rule of thumb, smaller differences yield better COPs [5]. Since ASHPs extract heat from outdoor air, their efficiency degrades as temperatures drop—units that achieve COPs of 5 or higher in ideal conditions may see COPs drop below 2 in very cold conditions [69]. For a fixed output, electric consumption scales with $1/\text{COP}$, so low COPs lead to extra load when heating demand is highest—prior work has identified that limited capacity in existing distribution grids is a key bottleneck to widespread electrification, including ASHP adoption [12, 16, 76, 77]. On the other hand, GSHPs leverage the observation that temperatures just 1-2 meters underground remain relatively constant year-round (e.g., above 5°C in non-polar climates) [72]—this enables GSHPs to achieve more stable high COPs, especially during cold weather [3].

5GDHC Thermal Networks. Fifth-generation district heating and cooling (5GDHC) systems are an emerging HVAC technology that connects a community of buildings via a shared underground loop of uninsulated pipes containing a fluid (e.g., brine or water) operating at ambient temperatures (e.g., 5-25°C) [21]. Instead of directly delivering high-temperature heat to buildings, GSHPs at each building use the fluid loop as a heat source (for heating) or sink (for cooling). 5GDHCs avoid the need for costly insulated distribution piping, and enable broader deployment of GSHPs without the cost and space requirements of isolated ground loops or boreholes at each building. 5GDHC networks are typically designed to interconnect with *borehole fields* to exchange heat with the ground—a borehole is a vertical loop of pipe, between 100-200m in depth, that leverages extremely stable deep ground temperature to provide a heat source/sink when fluid is pumped through it [8, 18]. However, owing to the ambient temperature of the network, 5GDHC systems can also integrate with other heat sources/sinks, such as waste heat from data centers or water treatment plants [70, 75].

Demand Response. Several demand response programs have been proposed and implemented to manage the electric load of

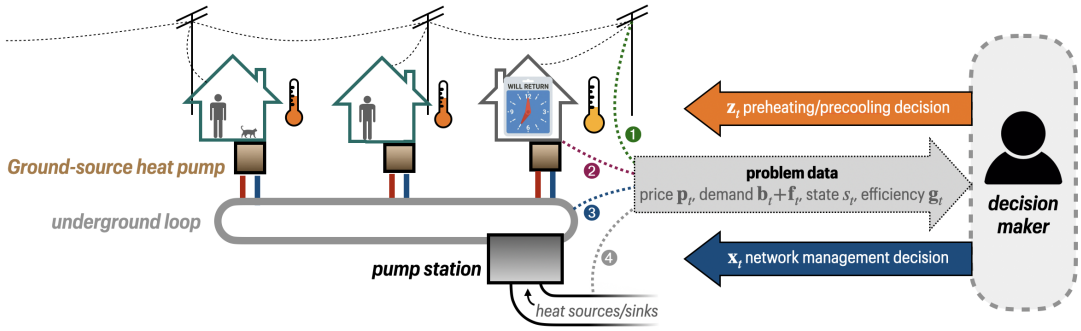


Figure 1: Diagram of the OTNM problem at a single time step t. The decision maker receives input data describing ❶ the electricity price p_t, ❷ system demand, base represented by b_t and flexible represented by f_t, ❸ the state of the underground thermal network s_t, and ❹ the current efficiency of heat sources/sinks attached to the pump station(s). In response, the decision maker must specify a network management decision x_t, which decides where and when to inject/extract heat from the underground thermal loop, and a preheating/precooling decision z_t, which decides when to satisfy flexible demands at customer homes.

HVAC systems, often via direct control or price-based incentives [26, 66]. In the U.S., utilities have deployed direct control programs targeting cooling loads during summer peaks—these programs offer discounted smart thermostats to participants in exchange for the ability to reduce load during a few peak events each year—utilities often *precool* homes before such events to minimize discomfort [63]. Alternatively, time-of-use prices charge customers higher rates during peak hours (and vice-versa) to incentivize self-shifting [52]. In this work, we focus on the former type of control, where a system operator can interface with smart thermostats to slightly shift the timing of heating/cooling loads within occupant comfort constraints. Typically considered in industrial contexts, demand response programs can also adopt the form of regulation signals or *tracking targets* sent by the electric grid to consumers at regular intervals that describe consumption set points or adjustments to help the grid in real-time [57]. In our problem formulation (see Section 3), we model these via a tracking cost, wherein the operator is penalized for deviating from externally-provided targets.

3 Problem and Preliminaries

In this section, we formalize the problem faced by a 5GDHC system operator as online thermal network demand management (OTNM) and state our assumptions. We summarize the key components of OTNM in Fig. 1. Throughout, lowercase bold letters denote vectors, and $\mathbf{1}$ denotes the all-ones vector of appropriate dimension.

3.1 Problem Formulation

Consider a 5GDHC with m pump stations and n customers being served over a time horizon T. At each discrete time step $t \in [T]$, inputs that describe the system's condition are revealed online. The first set of inputs describes *demand* on the system: *base demand* $\mathbf{b}_t \in \mathbb{R}^n$ must be satisfied immediately at time t, while *flexible demand* $\mathbf{f}_t \in \mathbb{R}^n$ must be satisfied at any time step in the interval $[t, \Delta_{i,t}]$ for customer $i \in [n]$, where $\Delta_{i,t} > t$ is a customer-specific deadline. Flexible demands capture requests such as *pre*heating or *pre*cooling of buildings and houses, as reported by, e.g., occupancy-aware smart thermostats. For both demand types, we model positive values as *heat* requests and negative values as *cooling* requests.

The operator must make two decisions at each time step t. First, $\mathbf{x}_t \in \mathbb{R}^m$ is a network management decision that denotes how much heat energy to inject into (resp. extract from) the network at each of the m pump stations. While base demands (e.g., $\mathbf{b}_t$) *must* be satisfied irrespective of the operator's decisions, the operator has the ability to decide when to satisfy flexible demands (e.g., $\mathbf{f}_t$) within their respective deadlines. Thus, the second decision $\mathbf{z}_t \in \mathbb{R}^n$ denotes how much of the flexible demand to satisfy at time t—i.e., how much to preheat or precool at each customer at time t. Both decisions are subject to physical constraints based on the capabilities of the pump stations and the ground-source heat pump (GSHP) at each customer—we model these as feasible sets $\mathbf{x}_t \in \mathcal{X}_t$ and $z_{i,t} \in [-\omega_i, \omega_i]$, where ω_i is the max. heating/cooling rate of customer i's GSHP.

The thermal network has an effective *capacity* of S units of thermal energy based on allowable temperatures for the fluid circulating within.[1] A thermodynamic model $\mathcal{M}$ describes its evolution over time—this model can be simple (e.g., approximating the network's state of charge as if it were a battery) or complex (e.g., capturing the full physics of fluid flow in the network). In practice, a common choice is a linearized nodal model that represents the network as a graph with nodes of thermal capacities and edges representing pipes, capturing basic heat loss and propagation delays [64, 67]. Given $\mathcal{M}$, we describe the time-varying network state as $s_t \in [0, S]$, representing the amount of energy stored in the network at time t.

The operator's goal is to satisfy all demand and manage the thermal network's constraints over time while minimizing an *operational cost* consisting of three components: *(i)* the cost of managing the network, *(ii)* the cost of satisfying demand at customers' homes, and *(iii)* a tracking cost that encourages "grid-healthy" targets for electricity usage. At each time t, the operator receives online inputs that describe conditions affecting these. The first is an *electricity price*, denoted by p_t. Second, a vector $\mathbf{g}_t \in \mathbb{R}^m_+$ describes the time-varying *inverse efficiency* of each pump station in converting electricity (e.g., driving pumps, compressors) into thermal energy (this can vary over time based on, e.g., borehole and outside air temperatures). A base demand efficiency function $c^{(b)}$ describes the energy needed to satisfy base demand at each customer at time t, while a preheating/precooling efficiency function $c_t^{(f)}$ captures the same for flexible demand. The function $c_t^{(f)}$ additionally captures thermal losses—e.g., if a house is preheated 6 hours early, some heat will be lost before the deadline. Finally, the operator receives *tracking targets* (from, e.g., the electric grid) for both decisions, denoted

[1] E.g., 10,000 liters of water allowed to vary between 10°-20° C can store ≈116 kWh.

by $\tilde{\mathbf{x}}_t$ and $\tilde{\mathbf{z}}_t$. These represent desired usage patterns that promote grid stability (such as avoiding peaks) and the operator incurs a cost of η and γ (respectively) for deviating from them [1]. An offline formulation of OTNM illustrating the costs and constraints is below.

$$\min_{\{\mathbf{x}_t,\mathbf{z}_t\}_{t\in[T]}} \sum_{t=1}^{T} p_t\left(\mathbf{g}_t^\top \mathbf{x}_t + c^{(b)}(\mathbf{b}_t) + c_t^{(f)}(\mathbf{z}_t)\right) + \eta\|\mathbf{x}_t - \tilde{\mathbf{x}}_t\|_1 + \gamma\|\mathbf{b}_t + \mathbf{z}_t - \tilde{\mathbf{z}}_t\|_1$$

$$\text{s.t. } s_t \leftarrow \mathcal{M}(s_{t-1}, \mathbf{x}_t, \mathbf{b}_t, \mathbf{z}_t) \text{ with } s_t \geq 0 \quad \forall t \in [T],$$

$$\mathbf{1}^\top \mathbf{x}_t \leq \max\{S - s_{t-1} + \mathbf{1}^\top(\mathbf{b}_t + \mathbf{z}_t), \mathbf{1}^\top(\mathbf{z}_t)\} \quad \forall t \in [T],$$

$$\min\Big\{0, \sum_{\tau\leq t:\Delta_{i,\tau}\geq t} f_{i,\tau}\Big\} \leq z_{i,t} \leq \max\Big\{0, \sum_{\tau\leq t:\Delta_{i,\tau}\geq t} f_{i,\tau}\Big\} \quad \forall t \in [T],$$

$$\sum_{\tau=t}^{\Delta_{i,t}} z_{i,\tau} = f_{i,t} \quad \forall i \in [n], \forall t \in [T],$$

$$\mathbf{x}_t \in \mathcal{X}_t \text{ and } z_{i,t} \in [-\omega_i - b_{i,t}, \omega_i - b_{i,t}] \quad \forall i \in [n], \forall t \in [T],$$

$$\max\{0, f_{i,t}\} \leq \sum_{\tau=t}^{\Delta_{i,t}} \mathbf{1}^\top \mathbf{x}_\tau, \forall i \in [n] \quad \forall t \in [T],$$

$$s_0 \leq s_T \text{ and } \mathbf{0} \preceq \mathbf{x}_t \quad \forall t \in [T].$$

While we focus on a tracking cost model for simplicity, we remark that various other "smoothing costs" could be used in OTNM, such as switching costs that penalize changes in decisions over time [7]—our results can be extended to other such models as well.

As outlined in the formulation, OTNM is subject to several constraints for all $t \in [T]$. In order, they are as follows: First, the network state s_t evolves according to the thermodynamic model based on the previous state, decisions, and demand: $s_t \leftarrow \mathcal{M}(s_{t-1}, \mathbf{x}_t, \mathbf{b}_t, \mathbf{z}_t)$, with the constraint that $s_t \geq 0$ (i.e., minimum charge level) for all $t \in [T]$. Next, the energy injected into (resp. extracted from) the network, i.e., $\mathbf{x}_t$, should reflect the current state of the network—if the network is discharged ($s_{t-1} \approx 0$), then the operator can inject energy up to the shiftable capacity (i.e., $S - s_{t-1}$) plus the amount of energy needed to satisfy demand at time t. Conversely, if the network is already charged ($s_{t-1} \approx S$), then the operator can only inject energy to satisfy flexible demand actively being served (i.e., $\mathbf{z}_t$). Preheating/precooling decisions at the i^{th} customer must satisfy two constraints: first, at time t, we are allowed to satisfy any outstanding flexible demand for that customer that has arrived but is not past due—this is captured by $\min\{0, \sum_{\tau:\Delta_{i,\tau}\geq t} f_{i,\tau}\} \leq z_{i,t} \leq \max\{0, \sum_{\tau:\Delta_{i,\tau}\geq t} f_{i,\tau}\}$. Second, the cumulative preheating and precooling for any i^{th} customer's flexible demand request must satisfy the request before its deadline—this is captured by $\sum_{\tau=t}^{\Delta_{i,t}} z_{i,\tau} = f_{i,t}$. Intuitively, all decisions must lie within their feasible sets.

The final two constraints depend on the sign of the net demand (see Asm. 3.9 for more discussion). If the net demand over the time horizon is positive (meaning the network "requires" heating overall), then the operator must ensure that enough energy is injected during the lifetime of positive flexible demand requests to satisfy it (captured by $\max\{0, f_{i,t}\} \leq \sum_{\tau=t}^{\Delta_{i,t}} \mathbf{1}^\top \mathbf{x}_\tau$), and the final state of the network must be at least as charged as the initial state (i.e., $s_0 \leq s_T$). If the aggregate demand over the time horizon is negative (i.e., the network "requires" cooling overall), then the above are reversed.

While the above captures the offline version of the problem, our goal is to design an online algorithm that makes decisions $\mathbf{x}_t$ and $\mathbf{z}_t$ *without knowledge of future inputs* $\{p_\tau, \mathbf{g}_\tau, c_\tau^{(f)}, \mathbf{b}_\tau, \mathbf{f}_\tau, \tilde{\mathbf{x}}_\tau, \tilde{\mathbf{z}}_\tau\}_{\tau>t}$. We evaluate such an algorithm using competitive ratio [45]:

Definition 3.1 (Competitive Ratio). Let $\texttt{OPT}(\mathcal{I})$ denote the optimal offline cost for an OTNM instance $\mathcal{I}$, and let $\texttt{ALG}(\mathcal{I})$ denote the cost of an online algorithm ALG on the same instance. ALG is α-competitive if for all $\mathcal{I}$, $\texttt{ALG}(\mathcal{I}) \leq \alpha\texttt{OPT}(\mathcal{I}) + C$, where $\alpha \geq 1$ is the competitive ratio and $C \geq 0$ is a constant independent of the instance $\mathcal{I}$.

Note that a smaller competitive ratio is better. This is a *worst-case* metric, since it must hold for all instances, including adversarial ones. As a result, competitive bounds provide strong guarantees against nonstationarity in the environment, which is crucial for reliable operation of real-world systems such as 5GDHCs—as one example, electricity prices are known to be volatile and nonstationary [58]. However, algorithms purely optimized for the worst-case can be overly conservative in practice—to address this, in Section 5, we leverage techniques to learn an improved algorithm from data, minimizing average-case cost while preserving the worst-case guarantees against nonstationarity that are useful in practice.

3.2 Assumptions

To design algorithms for OTNM, we make the following assumptions. We typically assume T (the time horizon) is known a priori.

Assumption 3.2. *Electricity prices are bounded within a known range: $p_{\min} \leq p_t \leq p_{\max}$ for all $t \in [T]$, where $0 < p_{\min} \leq p_{\max} < \infty$.*

This is a common assumption in literature on related online problems such as online search [15, 39] and is necessary for a bounded competitive ratio. Note that these prices correspond to a cost per unit of electrical energy (e.g., kWh) consumed.

Assumption 3.3. *Each element of the efficiency vectors is bounded in a known range: $0 < g_{\min} \leq g_{j,t} \leq g_{\max} \leq 1$ for all $t \in [T]$ and $j \in [m]$.*[2]

Intuitively, $g_{j,t}$ represents the efficiency of pump station j at time t: how much electricity is needed to move one unit of thermal energy. This can vary over time as, e.g., borehole temperatures fluctuate, exterior air temperature changes, etc. Throughout the paper, we let $\lambda_{\min} = p_{\min} g_{\min}$ and $\lambda_{\max} = p_{\max} g_{\max}$ denote *bounds* on the marginal cost of energy injected into (or extracted from) the network.

Assumption 3.4. *While the network model $\mathcal{M}$ is arbitrary, given a fixed time horizon T, we assume there exists a loss factor $L \in [0,1)$ such that evolution satisfies $s_T \geq s_0 + \sum_{t=1}^{T}(\mathbf{1}^\top \mathbf{x}_t - \mathbf{1}^\top(\mathbf{b}_t + \mathbf{f}_t)) - LS$. Also, state $s_t = 0$ corresponds to the case where the network is at or below ground temperature (eliminating further heat loss).*

Asm. 3.4 essentially states that the network's thermodynamics $\mathcal{M}$ must satisfy basic properties for a given time horizon T. In particular, the inequality ($s_T \geq s_0 +$ injected heat $-$ extracted heat $-$ loss) must hold for some loss factor bounded away from 1, meaning that the network can be approximated as a "lossy battery." This simplifies the treatment of the network's otherwise complicated physics from an algorithm perspective, while remaining realistic for 5GDHC systems in practice.

Assumption 3.5. *Given pump station capabilities (feasible sets $\{\mathcal{X}_t\}_{t\in[T]}$) and demand sequences $\{\mathbf{b}_t, \mathbf{f}_t\}_{t\in[T]}$, it is always possible to satisfy network constraints (i.e., $s_t \geq 0\ \forall t \in [T-1]$ and $s_T \geq s_0$). Further, given customer GSHP capabilities (rate constraints given by ω_i), flexible demands are satisfiable at the deadline (i.e., $z_{i,\Delta_{i,t}} = f_{i,t}$).*

Asm. 3.5 is necessary for the feasibility of OTNM and true in practice—networks are sized appropriately to meet demand, and flexible demands can be defined with satisfiability in mind.

[2] Note that $g_{\max} \leq 1$ since "backup" resistive electric heating has an efficiency of 1.

Assumption 3.6. *In* OTNM*, we assume that the coefficient of performance (*COP*) of each home's GSHP is static over time and known. The function $c^{(b)}$ is defined as $c^{(b)}(\mathbf{b}_t) \coloneqq \sum_{i=1}^{n} \frac{1}{\mathsf{COP}_i}|b_{i,t}|$, where $\mathsf{COP}_i \geq 1$ is a conservative estimate of the* COP *of customer i's heat pump.*

Assumption 3.7. *Given a flexible demand $f_{i,t}$ with deadline $\Delta_{i,t}$, $c^{(f)}_{i,t}$ satisfies the following. If the demand is satisfied at its deadline, the electricity needed is equivalent to satisfying base demand (i.e., $c^{(f)}_{i,\Delta_{i,t}}(f_{i,t}) = \frac{1}{\mathsf{COP}_i}|f_{i,t}|$). Further, for any $\tau_1, \tau_2 \in [t, \Delta_{i,t}]$ where $\tau_1 < \tau_2$, we have $c^{(f)}_{i,\tau_1}(f_{i,t}) \geq c^{(f)}_{i,\tau_2}(f_{i,t})$ (i.e., satisfying the demand earlier is at least as difficult as satisfying it later).*

We let r denote the ratio between the efficiency of GSHPs and the efficiency of managing the thermal network—it is defined as $r \coloneqq \frac{c_{\max}}{g_{\min}}$, where $c_{\max} \coloneqq \max_{i\in[n]} \frac{1}{\mathsf{COP}_i}$ is the worst GSHP COP (i.e., at any home), and $g_{\min}$ is the best-case efficiency of any pump station in managing the thermal network (see Asm. 3.2).

Assumption 3.8. *Tracking cost coefficients η and γ must satisfy $\eta \leq (\lambda_{\max} - \lambda_{\min})/2$ and $\gamma \leq (\lambda_{\max} r - \lambda_{\min} r)/2$.*

When η and γ are "too large," tracking costs overwhelm the other corresponding costs in the objective, making algorithm design simple (i.e., one should always match the tracking targets).

3.3 Opposing Demands

In OTNM, a component that has not been addressed in prior literature on related online problems is the presence of *opposing demands*: in any instance, some demands might be for heating (positive values) while others might be for cooling (negative values). This is intuitive in real-world systems—different customers can request heating and cooling at the same time. However, they introduce new challenges for analysis. Since heating demands remove heat from the network while cooling demands add heat, a demand sequence can "trick" an algorithm into making decisions that are difficult to analyze in hindsight—for example, an adversary can trick an algorithm into, e.g., purchasing heat to charge the network, only to follow up with large cooling demands that render these purchases "wasted."

While demand sequences are uncertain, they are not adversarial in practice—in particular, the 5GDHC operator likely has a good prior about what kind of network management is needed over a short time horizon (e.g., if T is a single day)—for example, in the Northern Hemisphere, a day in January will require heat to be added to the network, and a day in July will require the opposite. To formalize this intuition and mitigate the adversarial challenge of opposing demands in our analysis, we assume:

Assumption 3.9. *For any instance $\mathcal{I}$ of* OTNM*, denote the total net demand over the entire time horizon as $D \coloneqq \sum_{t=1}^{T} \mathbf{1}^\top(\mathbf{b}_t + \mathbf{f}_t)$. We assume that an online algorithm for* OTNM *knows whether $D > 0$ (net heating demand) or $D < 0$ (net cooling demand) a priori.*

Using Asm. 3.9, we can focus our algorithmic decisions accordingly: For example, if D reflects net heating demand, then an algorithm should focus on adding heat to the network—this is captured by $\mathbf{x}_t \succeq \mathbf{0}\ \forall t \in [T]$ and $s_0 \leq s_T$. Conversely, if D reflects cooling demand, then an algorithm should focus on extracting heat from the network over time, captured by $\mathbf{x}_t \preceq \mathbf{0}\ \forall t \in [T]$ and $s_0 \geq s_T$.

We note that these cases do not preclude providing heating or cooling to customers at any time—ground-source heat pumps at customer homes can be used for either heating or cooling regardless of D. Rather, the purpose of Asm. 3.9 is to smooth over routine intra-day fluctuations in demand for the sake of competitive analysis. As an example, consider a case with a subperiod of heating demand followed by a subperiod of cooling demand—in practice, such a case is actually ideal for 5GDHCs, since the heat injected into the network during the first subperiod can be extracted during the second subperiod, reducing the network management cost to near-zero.[3] In such a case, an algorithm should not "overfit" its decisions to these short-term fluctuations in demand, since these are difficult to analyze in hindsight and not adversarial in practice—Asm. 3.9 ensures that our algorithm analysis can focus on the case where our total instance cost (i.e., for network management) is directionally aligned w.r.t. that of the optimal offline solution.

We also note that the sign of D can typically be estimated in advance with high probability from seasonality and weather data, so algorithms using this assumption remain effective. We formalize this in the following theorem.

Theorem 3.10. *Given a data set $\mathcal{S}$ of N historical instances of* OTNM*, PAC generalization guarantees existence of a classifier predicting the sign of D whose misclassification rate is bounded by δ with probability at least $1-\kappa$ over the random draw of the data set.*

For a new instance $\mathcal{I}$ of OTNM*, an α-competitive* ALG *that uses the predicted sign for Asm. 3.9 satisfies $\mathbb{E}[\mathsf{ALG}(\mathcal{I})] \leq (\alpha + (\zeta - \alpha)\delta)\mathsf{OPT}(\mathcal{I}) + C$ with probability at least $1-\kappa$, where $\zeta = O\left(\frac{\lambda_{\max}+\eta+\gamma}{\lambda_{\min}}\right)$ is the competitive ratio of an algorithm that only ensures feasibility.*

The proof relies on known PAC generalization bounds combined with an analysis of an algorithm that only ensures feasibility without optimizing cost (i.e., a "fallback" used when the sign of D is misclassified). The value of $\kappa > 0$ can be specified, but the prediction error δ depends on the dataset and prediction task—in practice, δ can be empirically estimated for a given κ (e.g., using a held-out test set). We defer more discussion and the full proof to Appendix B.1.

4 Algorithm Design and Guarantees

In this section, we present a distributed online algorithm called **T**hermal **A**ccounting for **N**etwork-**D**istributed **E**nergy **M**ediation (TANDEM) for solving OTNM. We describe TANDEM in Section 4.1 and analyze its competitive ratio in Section 4.2.

4.1 The TANDEM algorithm

The core ideas of TANDEM build on a long line of work on online problems such as load shifting that leverage an algorithmic framework known as *pseudo-cost minimization* [7, 38, 68]. In this framework, an algorithm maintains a state-dependent *threshold function* that guides its decisions by defining the "prices" at which it is willing to take certain actions. This function is designed to capture a *trade-off* between costs now and worst-case future costs, driving decisions that guarantee a bounded competitive ratio.

OTNM's challenges complicate the design of such an algorithm—existing threshold-based techniques are designed for problems with known demand (such as one-way trading [7, 15]) where the algorithm knows how much energy it will need to purchase, and only decides *when* (i.e., at what price) to purchase. Furthermore, the

[3] In seasons where both heating and cooling are needed, 5GDHC thermal inertia allows them to "self-regulate": they can operate with little external energy input [67].

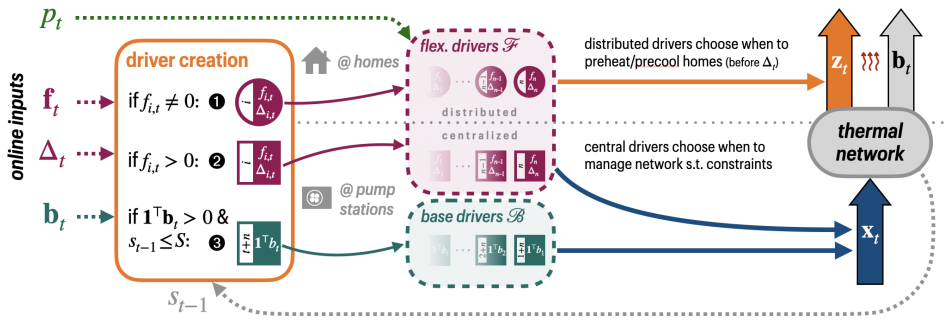


Figure 2: Diagram of the intuition behind the TANDEM algorithm at a single time step. Given the current demand (base $\mathbf{b}_t$ and flexible $\mathbf{f}_t$), TANDEM creates drivers that make decisions as follows: ❶ if customer i has a flexible demand (with deadline $\Delta_{i,t}$), a distributed driver is created that decides when to preheat/precool this home. ❷ if customer i's flexible demand is aligned with the sign of the net demand, a local module is created to manage the thermal network accordingly. Finally, ❸ depending on the current network state (s_{t-1}), a base driver is created to manage the thermal network in response to aggregate demand ($\mathbf{1}^\top \mathbf{b}_t$).

network's state dynamics and the presence of both *positive* and *negative* demands (i.e., heating and cooling loads) are unique to OTNM and require an algorithm to account for situations that impact the total amount of network management needed.

To tackle the first challenge, TANDEM generalizes a partitioning technique introduced by [78] that was extended to pseudo-cost minimization by [38]. Newly arrived units of (base or flexible) demand instantiate "drivers" that perform two important functions: *(i)* they allow TANDEM to account for the total demand that has arrived over time in a granular way, and *(ii)* they generate decisions for subproblems that are aggregated into global decisions. To tackle the second challenge, we leverage two facts—first, Asm. 3.4 guarantees that network state evolution is "well-behaved" in the sense of being approximable as a lossy battery, which significantly simplifies the needed accounting for "charging/discharging" decisions. The presence of both positive and negative demands is a significant departure from prior applications of this partitioning technique, which were applied to unidirectional problems such as load shifting. In those problems, driver-like partitions are instantiated to account for every unit of arriving demand. In contrast, TANDEM implements an "active" driver creation loop that considers the sign of the net demand D and the time-varying network state at each step—this imposes conditions on when drivers are created, which allows TANDEM to selectively ignore certain units of demand (e.g., if the network's state has already been preheated by previous cooling demand, the network is effectively "self-regulating" and additional heating demand does not require further management, hence new drivers are not created). Below, we describe TANDEM in detail.

Driver Types. TANDEM uses two types of drivers to specify network management decisions and preheating/precooling decisions. The first type is *base drivers* (denoted by set $\mathcal{B}$), which the base demand that arrives over time—these drivers focus on managing the network state to replenish thermal energy that has been used by homes. The second type is *flexible drivers* (denoted by set $\mathcal{F}$), which correspond to individual flexible demands at each customer, and are responsible for two actions. Flexible drivers are primarily implemented in a *distributed* manner at the customer level (e.g., at smart thermostats), where they make decisions about how much of the (preheating/precooling) demand to satisfy at each time step, given prices and tracking targets communicated from the central controller. If a given flexible demand is aligned with the sign of the net demand D, (e.g., heating demand when net demand is positive), a local module at the operator also makes network management decisions during its lifetime to satisfy it at low cost. In Fig. 2, we illustrate the main intuition behind TANDEM's state-dependent driver creation process and how demand-specific drivers become decision-making modules that generate TANDEM's aggregate actions.

At initialization, TANDEM creates a *manager driver* (a special base driver with index 0) responsible for managing overall network state, especially preheating the network to anticipate demand that may arrive. This driver is assigned size $d_0 = S - s_0$ (the energy needed to fill the network to capacity). At each time step $t \in [T]$, TANDEM creates new drivers as follows. Base demand is aggregated into a single demand value $\mathbf{1}^\top \mathbf{b}_t$; if this is positive *and* the network is not already preheated (i.e., $s_{t-1} < S$), a new base driver with index $t+n$ and size $d_{t+n} = \mathbf{1}^\top \mathbf{b}_t$ is created to manage it. If any $f_{i,t}$ is nonzero, a new flexible driver with index i, size $d_i = f_{i,t}$ and deadline $\Delta_i = \Delta_{i,t}$ is created to manage this demand in a distributed manner.

Base driver i (and flexible local modules, if applicable) makes decisions governed by a *network management threshold* $\phi_i(w)$ that depends on the amount of energy purchased so far $w_i \in [0, d_i]$:

Definition 4.1. For a driver i with size d_i, ϕ_i is defined as follows, where α is the competitive ratio (defined in (3)):

$$\phi(w) = \lambda_{\max} + 2\eta + \left(\frac{\lambda_{\max} + 2\eta}{\alpha} - \lambda_{\max}\right) \exp\left(\frac{w}{\alpha d_i}\right) : w \in [0, d_i]. \quad (1)$$

Analogously, each distributed flexible driver makes preheating/precooling decisions governed by a *delivery threshold* $\psi_i(v)$, which depends on the demand satisfied by driver i so far, $v_i \in [0, |d_i|]$:

Definition 4.2. For a flexible driver i with size d_i, ψ_i is defined as:

$$\psi(v) = \lambda_{\max} r + 2\gamma + \left(\frac{\lambda_{\max} r + 2\gamma}{\alpha} - \lambda_{\max} r\right) \exp\left(\frac{v}{\alpha d_i}\right) : v \in [0, d_i]. \quad (2)$$

Throughout the rest of the paper, we use the shorthand notation $\Phi(w_1, w_2) \coloneqq \int_{w_1}^{w_2} \phi(u)du : w_1 \leq w_2$ and $\Psi(v_1, v_2) \coloneqq \int_{v_1}^{v_2} \psi(u)du : v_1 \leq v_2$ to denote the definite integrals of threshold functions.

Algorithm Operation. We now describe the operation of the TANDEM algorithm in detail; the main pseudocode is provided in Alg. 2, and a subroutine for (distributed) household-level decisions is provided in Alg. 1. Each driver created by TANDEM makes individual decisions based on its own state and threshold function—these subdecisions are then aggregated into global actions (e.g., $\mathbf{x}_t \leftarrow \sum_{i \in \mathcal{B} \cup \mathcal{F}} \mathbf{x}_{i,t}$ (elementwise), $\mathbf{z}_t \leftarrow \{z_{1,t}, \ldots, z_{n,t}\}$).

TANDEM receives tracking targets that specify desired network management and preheating/precooling actions. In the case of the

Algorithm 1 Pseudo-cost Minimization Subroutine for Flexible Demand at TANDEM Customer (HouseholdDecision)

1: **input:** customer index i, demand size d_i, decision cost $p_t c_{i,t}^{(f)}$, tracking cost coeff. γ, tracking target $\tilde{z}_{i,t}$, threshold Ψ, GSHP capability ω_i.
2: **compute** feasible set $\mathcal{Z}_{i,t}$: if $d_i < 0$, then $\mathcal{Z}_{i,t} \leftarrow [\max\{-\omega_i - b_{i,t}, d_i + v_{i,t-1}\}, 0]$; else $\mathcal{Z}_{i,t} \leftarrow [0, \min\{\omega_i - b_{i,t}, d_i - v_{i,t-1}\}]$
3: **solve** $z_{i,t} \leftarrow \operatorname{argmin}_{z \in \mathcal{Z}_{i,t}} p_t c_{i,t}^{(f)} |z| + \gamma |b_{i,t} + z - \tilde{z}_{i,t}| - \Psi_i(v_{i,t-1}, v_{i,t-1} + |z|)$, **update** delivery state: $v_{i,t} = v_{i,t-1} + |z_{i,t}|$
4: **if** $\Delta_i = t$ **then** $z_{i,t} \leftarrow d_i - \operatorname{sign}(d_i) \cdot v_{i,t-1}$ ▷ satisfy any remaining demand at deadline
5: **output:** i^{th} customer's decision for time step t: $z_{i,t}$

Algorithm 2 **T**hermal **A**ccounting for **N**etwork-**D**istributed **E**nergy **M**ediation (TANDEM) Algorithm for OTNM

1: **input:** Network thermal capacity S, thermodynamic model $\mathcal{M}$, tracking cost coefficients η, γ, threshold functions Φ, Ψ
2: **initialize:** Initial network state $s_0 \in [0, S]$, decisions $\mathbf{x}_0 \leftarrow \mathbf{0}$, $\mathbf{z}_0 \leftarrow \mathbf{0}$, driver sets $\mathcal{B} \leftarrow \emptyset$, $\mathcal{F} \leftarrow \emptyset$
3: **for** $t = 1$ to T observe inputs $p_t, \mathbf{g}_t, c_t^{(f)}, \mathbf{b}_t, \mathbf{f}_t, \Delta_t, \tilde{\mathbf{x}}_t, \tilde{\mathbf{z}}_t, \mathcal{X}_t$ **do**
4: **if** $s_{t-1} = 0$ or $\mathbf{1}^\top \mathbf{b}_t > S$ **then create** manager driver with index 0, size $d_0 = S$; $\mathcal{B} \leftarrow \{0\}$ ▷ refresh base drivers and network manager driver
5: **if** $0 < \mathbf{1}^\top \mathbf{b}_t < S$ and $s_{t-1} \le S$ **then create** base driver with index $t+n$, size $d_{t+n} = \mathbf{1}^\top \mathbf{b}_t$; $\mathcal{B} \leftarrow \mathcal{B} \cup \{t+n\}$
6: **for** each customer $i \in [n]$ **if** $f_{i,t} \neq 0$ **then do** ▷ create driver for each new flexible demand
7: **create** flexible driver i: $d_i \leftarrow f_{i,t}$, $\Delta_i \leftarrow \Delta_{i,t}$; $\mathcal{F} \leftarrow \mathcal{F} \cup \{i\}$
8: **initialize** feasible decision set $\bar{\mathcal{X}}_t \leftarrow \mathcal{X}_t$ ▷ tracks network management decisions to ensure feasibility
9: **for** each active driver $i \in \mathcal{F}$ **do**
10: **send** to customer i: $p_t c_{i,t}^{(f)}$, $\tilde{z}_{i,t}$; receive $z_{i,t} \leftarrow$ HouseholdDecision$(i,\ d_i,\ p_t c_{i,t}^{(f)},\ \gamma,\ \tilde{z}_{i,t},\ \Psi_i)$
11: **if** $\Delta_i = t$ and $d_i > 0$ **then set** $\mathbf{x}_{i,t}$ s.t. $\mathbf{1}^\top \mathbf{x}_{i,t} = d_i - w_{i,t-1}$; **update** $\bar{\mathcal{X}}_t \leftarrow \bar{\mathcal{X}}_t \setminus \{\mathbf{x}_{i,t}\}$ ▷ satisfy any remaining demand at deadline
12: **for** each active driver $i \in \mathcal{B} \cup \mathcal{F}$, $\Delta_i \neq t$, sorted by $w_{i,t-1}/d_i$ **do**
13: **compute** pseudo-target $\tilde{\mathbf{x}}_{i,t} \leftarrow \tilde{\mathbf{x}}_t \cdot \frac{d_i - w_{i,t-1}}{\sum_{j \in \mathcal{B} \cup \mathcal{F}} d_j - w_{j,t-1}}$ ▷ allocate proportional tracking targets
14: **solve** $\mathbf{x}_{i,t} \leftarrow \operatorname{argmin}_{\mathbf{x} \in \bar{\mathcal{X}}_t} p_t \mathbf{g}_t^\top \mathbf{x} + \eta \|\mathbf{x} - \tilde{\mathbf{x}}_{i,t}\|_1 - \Phi_i(w_{i,t-1}, w_{i,t-1} + \mathbf{1}^\top \mathbf{x})$, **update** $\bar{\mathcal{X}}_t \leftarrow \bar{\mathcal{X}}_t \setminus \{\mathbf{x}_{i,t}\}$; $w_{i,t} \leftarrow w_{i,t-1} + \mathbf{1}^\top \mathbf{x}_{i,t}$
15: **aggregate** network management decisions: $\mathbf{x}_t \leftarrow \sum_{i \in \mathcal{B} \cup \mathcal{F}} \mathbf{x}_{i,t}$ and flexible decisions from households $\mathbf{z}_t \leftarrow \{z_{1,t}, z_{2,t}, \ldots, z_{n,t}\}$
16: **if** $\mathbf{1}^\top(\mathbf{b}_t + \mathbf{z}_t) - \mathbf{1}^\top(\mathbf{x}_t) - s_{t-1} < 0$ **then increase** $\mathbf{x}_t$ to meet demand (i.e., until LHS > 0)
17: **if** $t = T$ and $s_{t-1} + \mathbf{1}^\top(\mathbf{x}_t) - \mathbf{1}^\top(\mathbf{b}_t + \mathbf{z}_t) < s_0$ **then increase** $\mathbf{x}_t$ to restore network to initial state s_0
18: **update** network state: $s_t \leftarrow \mathcal{M}(s_{t-1}, \mathbf{x}_t, \mathbf{b}_t, \mathbf{z}_t)$ and **refresh** drivers: remove i from $\mathcal{B}$ if $w_{i,t} = d_i$, from $\mathcal{F}$ if $\Delta_i = t$
19: **output:** Purchase decisions $\{\mathbf{x}_t\}_{t=1}^T$, delivery decisions $\{\mathbf{z}_t\}_{t=1}^T$

network management target $\tilde{\mathbf{x}}_t$, the tracking cost applies to decisions across multiple drivers—to address this, we define *pseudo-targets* that distribute the target (elementwise) across all active drivers based on their outstanding demand (see Line 13 of Alg. 2).

The core of TANDEM's decision-making is the *pseudo-cost minimization* problem (Line 14 in Alg. 2 and Line 3 in Alg. 1). In this problem, the first two terms correspond to the actual cost that will be incurred for an action under consideration (e.g., z or $\mathbf{x}$), while the final term (a definite integral of the relevant threshold) *encourages good decisions*: when the marginal cost of an action is lower than the threshold value at the current state, the final term incentivizes taking that action; otherwise, it waits for better conditions in the future. This enables TANDEM's bounded competitive ratio.

We remark that while TANDEM's definition is involved due to the inherent complexity and requirements of OTNM, its computational burden (i.e., computing solutions to convex programs) is light, making it practical for real systems. In Appendix A.5, we demonstrate that TANDEM's overhead is small—it can be implemented efficiently on low-power devices such as smart thermostats.

4.2 Competitive Guarantees

We now analyze the TANDEM algorithm defined in Section 4.1, providing bounds on its competitive ratio. We show that TANDEM is α-competitive (see Theorem 4.3), which is the best possible for OTNM as shown in Theorem 4.4. Our main result is below—we give a proof sketch here and relegate the full proof to Appendix B.2.

Theorem 4.3. TANDEM *is α-competitive for* OTNM, *where α is defined as follows. W denotes the Lambert W function [11]; μ is defined as $\mu \coloneqq \frac{2\gamma}{r}$ when $\eta \le \frac{\gamma}{r}$ and $\mu \coloneqq \frac{2\eta + 2\gamma}{1+r}$ otherwise.*

$$\alpha \coloneqq \left[\frac{\lambda_{\max}}{\mu + \lambda_{\max}} + W\left(- \frac{\exp\left(- \frac{\lambda_{\max}}{\mu + \lambda_{\max}} \right)(\lambda_{\max} - \lambda_{\min})}{\mu + \lambda_{\max}} \right) \right]^{-1}. \quad (3)$$

Proof Sketch. We start by showing that TANDEM generates a feasible solution (Lemma B.1), which follows by definition. Next, we lower bound the cost of OPT (Lemma B.3) and upper bound the cost of TANDEM (Lemma B.4)—to obtain these, we partition an instance into active and inactive periods (Def. B.2) based on the state of the network in TANDEM's solution—combined with properties of the threshold functions, this gives a granular accounting of costs incurred by both solutions over time. Due to the presence of both positive and negative demands in OTNM, we must then carefully bound how the procurement amount of TANDEM relates to that of OPT over the entire instance (Lemma B.7), which relies on Asm. 3.4. By showing that TANDEM's procurement is an additive constant away from OPT in the worst case, we can use pseudo-cost techniques to compare the costs of both solutions using a relation between the threshold functions and their integrals (Lemmas B.5 & B.6). The rest follows by contradiction—as long as these threshold relations hold, the competitive ratio is at most α. □

We note that although the loss factor L (from Asm. 3.4) does not appear in the above competitive bound, it is charged to a constant in the proof and helps account for TANDEM's total procurement

relative to OPT. Given the results of Theorem 4.3, a natural follow-up question is whether any online algorithm can achieve a better competitive ratio for OTNM—we answer this in the negative, giving a proof sketch below (full proof in Appendix B.3):

Theorem 4.4. *A set of* OTNM *instances exists for which no deterministic online algorithm is better than α-competitive (α as in (3)).*

Proof Sketch. To show this result, we construct a family of adversarial instances (Def. B.10) where the network management cost and preheating/precooling cost decrease over time, down to a "best price" parameterized by $x \in [\lambda_{\min}, \lambda_{\max}]$. These instances consist of one pump station ($m = 1$) and two customers ($n = 2$) with a single flexible demand each. Parameter $P \in (0, 1]$ controls the ratio of net demand to absolute demand (i.e., when $P \to 0$, the two demands nearly cancel each other out, and when $P = 1$, the demands are both positive). Under these special instances, we show that the cost of any deterministic ALG can be described using arbitrary functions $h(x), m(x) : [\lambda_{\min}, \lambda_{\max}] \to [0, 1]$ that capture ALG's network management and delivery decisions at price level x. For ALG to be $\alpha^\star$-competitive, these functions must satisfy an inequality whose solution yields a lower bound on the competitive ratio of α (matching (3)). Two cases for μ follow from solving for the "worst-case" P as a function of the problem's parameters. □

Recall that Asm. 3.6 is a simplifying assumption for analysis of OTNM that models the COP of each customer's GSHP as static over time. In reality, COP varies with the temperature of the heat source/sink, which in turn depends on network state s_t. In the following, we extend our competitive analysis to consider the optimal solution of a more detailed model capturing the state-dependent dynamic COP of each GSHP. Recall $c_{\max}$ is the maximum inverse COP in OTNM (i.e., the worst-case GSHP efficiency in the static model). Letting $\text{COP}_i(s)$ denote the (state-dependent) COP of GSHP i when the network is in state s, we let $c_{\min} \coloneqq \min_{i\in[n], s\in[0,S]} \frac{1}{\text{COP}_i(s)}$ denote the minimum inverse COP across all customers and states, corresponding to the best-case GSHP efficiency in a state-dependent model. Then:

Corollary 4.5. TANDEM *is $\alpha\beta$-competitive against the offline optimal solution for thermal network demand management with a state-dependent model of GSHP* COPs*, where $\beta \coloneqq c_{\max}/c_{\min}$.*

We defer the proof to Appendix B.4. In realistic settings, GSHP COPs fluctuate within a small range (i.e., between 3-6), thus β is usually < 2, and TANDEM's competitive ratio remains strong against a more detailed optimal solution. We verify this empirically in Section 7.

5 Learning from Historical Data

In Section 4, we designed TANDEM to achieve the best (worst-case) competitive ratio for OTNM. However, as discussed in Section 3.1, algorithms optimized for the worst-case can be overly pessimistic in practice. To address this, we extend TANDEM to incorporate a data-driven learning framework. In doing so, we seek a method that is flexible to adapt to historical patterns, yet computationally light enough to be practical, while preserving worst-case guarantees.

Motivation. In recent years, there has been growing interest in *learning-augmented* online algorithms that take predictions about the future as input to improve empirical performance [36, 44]—such algorithms are evaluated using the notions of consistency and robustness, which capture performance when predictions are accurate and adversarially wrong, respectively. These algorithms typically assume access to suitable predictions (e.g., price or demand forecasts) and focus on leveraging such predictions effectively.

Under the pseudo-cost minimization paradigm, recent work [38] has proposed an "end-to-end" methodology that directly *learns a better algorithm* from historical data, while retaining worst-case guarantees by constraining learning to lie within a set that guarantees a desired worst-case robustness. Rather than designing around specific predictions and optimizing for a theoretical notion of consistency, this approach directly finds the best algorithm for the downstream task at hand while retaining guarantees.

Approach. For OTNM, we generalize the end-to-end learning approach of [38] to our TANDEM algorithm as follows. Instead of the analytical threshold functions ϕ, ψ (Defs. 4.1 and 4.2), we define *learned threshold functions* $\hat{\phi}_{\boldsymbol{\theta}}$ and $\hat{\psi}_{\boldsymbol{\theta}}$, whose parameters $\boldsymbol{\theta}$ are learned from data to optimize empirical performance. Given learned thresholds $\hat{\phi}_{\boldsymbol{\theta}}$ and $\hat{\psi}_{\boldsymbol{\theta}}$, the pseudocode of our learned version of TANDEM is identical to that of Alg. 2 and Alg. 1, except that $\hat{\phi}_{\boldsymbol{\theta}}$ and $\hat{\psi}_{\boldsymbol{\theta}}$ are used in place of ϕ and ψ throughout.

We let ρ denote a desired robustness factor for the learned algorithm, where $\rho > \alpha$. To ensure that the learned thresholds yield a worst-case competitive ratio no worse than ρ, we constrain them to lie within a *robust certificate set* $\mathcal{R}_\rho$, defined as follows.

Definition 5.1 (Robust Certificate for Learned Thresholds). Given a desired robustness factor $\rho > \alpha$ (for α defined in (3)), threshold functions ϕ and ψ must lie within the following joint feasible set:

$$\mathcal{R}_\rho \coloneqq \{\phi : [0,1] \to [\lambda_{\min}, \lambda_{\max}] \text{ and } \psi : [0,1] \to [\lambda_{\min} r, \lambda_{\max} r] \;\big|\; \phi, \psi \text{ monotone non-increasing, } \phi(1) \le \lambda_{\min} + 2\eta, \psi(1) \le \lambda_{\min} r + 2\gamma \text{ and } \forall w \in [0,1], v \in [0, w] : \Phi(0, w) + (1-w)(\lambda_{\max} + 2\eta) + \Psi(0, v) + (1-v)(\lambda_{\max} r + 2\gamma) \le \rho\big[\phi(w) + \psi(v) - 2\eta - 2\gamma\big]\}.$$

Assuming our learned thresholds $\hat{\phi}_{\boldsymbol{\theta}}$ and $\hat{\psi}_{\boldsymbol{\theta}}$ lie within this set, we have the necessary conditions to prove that TANDEM using these learned thresholds is at most ρ-competitive for OTNM, as follows:

Theorem 5.2. *If learned thresholds $\hat{\phi}_{\boldsymbol{\theta}}$ and $\hat{\psi}_{\boldsymbol{\theta}}$ lie within the robust certificate set $\mathcal{R}_\rho$ for some desired robustness factor $\rho > \alpha$, then* TANDEM *using these learned thresholds is ρ-competitive for* OTNM*.*

We relegate the full proof to Appendix B.5. The conditions of $\mathcal{R}_\rho$ are critical for the proof—the first three conditions (structure, monotonicity and final value) ensure that thresholds take on each value for the necessary lower bounds on OPT. The final inequality condition is the key expression that relates the cost of TANDEM (using the learned thresholds) to the cost of OPT—the left-hand side forms an upper bound, while the right-hand side is a lower bound on OPT scaled by the robustness factor ρ.

Methodology. Theoretically, the framework described above can accommodate any parameterization of thresholds ϕ and ψ as long as their set membership in $\mathcal{R}_\rho$ can be enforced. One suitable parameterization is a piecewise-affine "knot" representation at a fixed grid of K points $\{0, \frac{1}{K-1}, \frac{2}{K-1}, \ldots, 1\}$: each threshold is defined by a vector $\boldsymbol{\theta}$, where each element corresponds to the value at one point (values in-between are linearly interpolated). We denote $\boldsymbol{\theta}_\phi$ and $\boldsymbol{\theta}_\psi$ as the parameters for $\hat{\phi}_{\boldsymbol{\theta}}$ and $\hat{\psi}_{\boldsymbol{\theta}}$, respectively. We have:

Lemma 5.3. *The robust certificate set $\mathcal{R}_\rho$ is convex when learned thresholds $\hat{\phi}_{\boldsymbol{\theta}}$ and $\hat{\psi}_{\boldsymbol{\theta}}$ are parameterized as piecewise-affine functions.*

The full proof is in Appendix B.6—when thresholds are piecewise-affine, $\mathcal{R}_\rho$ constraints can be expressed as linear inequalities and halfspaces, whose intersection is convex. Using this, we learn a small neural network that maps instance context (e.g., season, weather & price forecasts) to predicted threshold vectors $\boldsymbol{\theta}_\phi$ and $\boldsymbol{\theta}_\psi$. The robust certificate is enforced using a differentiable convex projection layer [2]. Using a training set of historical instances and a differentiable TANDEM implementation, we train the NN *end-to-end* to learn better thresholds with respect to downstream cost. We defer more details of the NN architecture to Appendix A.2.

6 Simulation Environment

In this section, we describe how we simulate the operation of 5GDHC networks in 10 case-study neighborhoods using data from a city in Massachusetts. We implement TANDEM and several baseline strategies to evaluate their performance across multiple metrics, including cost, emissions, and transformer peak load.

Data and Simulator. We combine data sets and use parameters from a pilot program [17] to create a 5GDHC simulator. We model 15-minute time steps, resampling (e.g., via averaging) as needed.

• *Demand & Neighborhood Data.* We use natural gas and electricity consumption data collected from a small city in Massachusetts for 2023-24. Gas data for each meter is reported at an hourly granularity, while electric usage is reported every five-minutes. We identify ten neighborhoods that rely on natural gas for heating (6-44 homes each). The existing gas network topology in each neighborhood is estimated using OpenStreetMap road data, using methodology from prior work [37]. We give a summary of neighborhood statistics in Appendix A.3. To model heating demand for each home, we convert natural gas consumption to thermal demand (in kWh), assuming a gas furnace/boiler efficiency of 90% [76]. We focus on the heating season (defined by local utilities as Nov.-Apr.) [51], as we do not have suitable data to model cooling demand.

• *Electricity Prices.* We use locational marginal price (LMP) data for 2023-24 in ISO New England [22]. For the learning implementation of TANDEM, we also use day-ahead LMP forecasts as context.

• *Air and Ground Temperatures.* We obtain 2023-24 air temperature data from a nearby airport [49] and ground temperature data (depth of 2m) from a nearby ecological monitoring site [50].

• *Carbon Intensity.* We use average carbon intensity (in $gCO_2eq.$ / kWh) for ISO New England in 2023-24 from Electricity Maps [46].

• *Flexible Demand.* Most customer demand is treated as *base demand* (recall Section 3), but we model a synthetic *flexible* demand as follows: we set a parameter `flex_prop` $\in [0, 1]$ to define the fraction of homes that have a flexible demand on any given day. If a home is randomly selected, we generate a random period between 3-6 hours and a random start time between 8 AM-2 PM during which the home is "unoccupied." The home's base demand during this period is treated as flexible and due at the end of the period.

To model thermal losses (home → air) that affect shifted flexible demand (i.e., $c_t^{(f)}$ in OTNM), we estimate each home's thermal loss rate (over time) using its size and air temperature data [4, 29].[4] We then increase $c_t^{(f)}$ to charge the operator for additional heating needed if flexible demand is satisfied in advance of the deadline.

• *Tracking Targets.* We aggregate electricity usage across all homes in each neighborhood (plus the 5GDHC system itself) to capture load effects on the grid. Tracking targets are defined as follows: each neighborhood's distribution transformer(s) have an existing load profile (from electric data) *before* 5GDHC is added, with a certain daily peak load. For network management decisions ($\tilde{\mathbf{x}}_t$ in OTNM), we compute the difference between the peak and existing load over time (i.e., targets are 0 at the time of peak load), and scale this difference so $\tilde{\mathbf{x}}_t$ sums to the total thermal demand of the instance. For delivery decisions ($\tilde{\mathbf{z}}_t$ in OTNM), we set targets equal to the base demand, except during periods of flexible demand, when the target is defined as described above, then scaled so that the sum during the flexible period equals the corresponding demand.

• *5GDHC Simulation.* We implement a reduced-order model to capture the thermodynamics of a 5GDHC system constructed based on the (estimated) natural gas distribution topology in each neighborhood. The model is a 1D plug flow system with conductive ground losses, discretized using a finite-difference scheme on a linearized out-and-back (i.e., supply and return loop) pipe network topology [64, 67].[5] We simulate a "pump station" at the neighborhood entrance that circulates fluid through the core neighborhood loop at a constant flow rate[6] and interconnects with (simulated) heat sources such as a borehole field, ASHP, or waste heat (see below). Note that boreholes are only able to supply heat to the network if the network's temperature is below the temperature at the borehole depth of 150m, which is typically about 2−4°C higher than the ground temperature at distribution pipe depth (2m) [48]. Throughout, we use product data sheets to model heat pump COPs as a function of air temperature [54] and fluid temperature [74].

Data Center Waste Heat. While 5GDHC systems are traditionally designed with borehole fields in mind, certain geographies and ground conditions may make borehole installation infeasible or cost-prohibitive [8, 42]. As one alternative, 5GDHCs are well-suited to using "low-quality" waste heat from a data center (DC)—unlike traditional district heating, GSHPs can still effectively scavenge heat from low-temperature sources. Using our simulator, we conduct a case study of using DC waste heat for 5GDHC systems throughout the city. We use time-varying utilization from Google power traces [62] to simulate waste heat availability over time. This is then partitioned to ensure that all homes currently using gas for heating in the case study city can use an equal share of the DC heat.

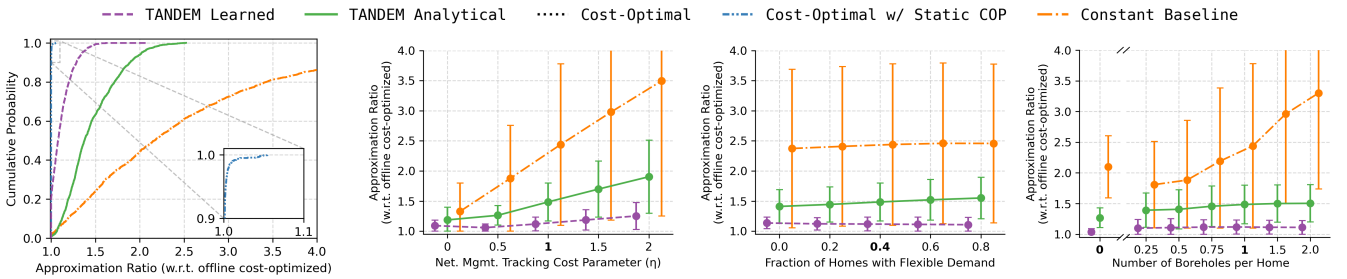
Algorithms and Baselines. We implement and evaluate several strategies for 5GDHC operation and baselines. First, we use Gurobi [23] to compute optimal solutions for OTNM (with full knowledge of prices and demands) under two GSHP efficiency models: *(i) dynamic*, where COP varies with local network temperature, and *(ii) static*, where COP is a conservative estimate. We report *approximation ratio* with respect to the offline cost-optimal dynamic solution.

We implement TANDEM in two variants: one using the ***analytical*** thresholds from Section 4, and another using ***learned*** thresholds via the framework described in Section 5. For the latter, we use 2023 data for training and report results on 2024 data—see Appendix A.2 for more details of the learning implementation. We set $K = 5$ and a robustness factor $\rho = 5\alpha$ for learned thresholds.

[4] For houses, we assume "rule-of-thumb" values of 1.2 W/m^2K for thermal conductivity, 120 kJ/m^2K for thermal mass per area, and an indoor setpoint of 20°C.

[5] We model 150mm diameter uninsulated HDPE pipes buried at 2m with a linear heat loss (soil) coefficient of 1.8 W/mK, and a fluid specific heat of 3.9 kJ/kgK (brine).

[6] We model a flow rate of 10 kg/s, and an allowable temperature range of 5-25°C.



(a) Tracking coefficient η (b) `flex_prop` parameter (c) No. of boreholes

Figure 3: CDF of approx. ratio for all algorithms and all neighborhoods with default parameter settings.

Figure 4: Average approximation ratio for all algorithms in one neighborhood, with varying (a) network management tracking coefficient (η), (b) proportion of homes with flexible demand (`flex_prop`), and (c) number of boreholes.

As a 5GDHC baseline, we implement a ***constant operation*** baseline that always runs the borehole field pumps and uses a simple thresholding rule to fulfill flexible demand when/if electricity prices fall below a fixed threshold ($\sqrt{p_{\min} p_{\max}}$).[7] If boreholes are not available (i.e., if only ASHPs or waste heat are used to supply the network), this strategy injects heat at a rate equivalent to using one borehole per home in the neighborhood. In select experiments, we compare against two *non-5GDHC* baselines: *(i)* the status quo of ***gas heating***, and *(ii)* uncoordinated ***ASHPs*** at each home, with a basic thresholding rule for flexible demand identical to that used in the constant operation strategy above.

Metrics. We evaluate strategies across three metrics: *(i) cost* of energy procurement (the primary objective of `OTNM`), *(ii) carbon emissions* from HVAC (electric or natural gas), and *(iii) peak load increase* (as a percentage relative to status quo) to assess grid impact.

7 Evaluation

We present key results in the main body, deferring additional results (sensitivity analyses, runtime measurements, etc.) to Appendix A.

A summary is in Fig. 3, which plots a cumulative distribution function (CDF) of the approximation ratio for all tested algorithms across all neighborhoods in 2024, with default parameter settings of $\eta = 1, \gamma = 0.1$, `flex_prop` $= 0.4$, and one borehole per home. First, we note that the cost-optimal solution using a static COP model for GSHPs performs well, achieving an approximation ratio ≤ 1.05 in all tested cases, and verifying the validity of Asm. 3.6. Analytical `TANDEM` achieves an approximation ratio of 1.45 on average and 2.02 in the 95$^{\text{th}}$ percentile, improving on the constant baseline by 44.1% and 63.3% (respectively). Learned `TANDEM` further improves performance, achieving an average approximation ratio of 1.11 and a 95$^{\text{th}}$ percentile ratio of 1.35, corresponding to improvements of 23.4% and 33.1% over the analytical version—this highlights the value of adapting to context about instances and price distributions. We explore the impact of several key parameters on performance, varying one parameter at a time while holding others at their default values and evaluating on instances in one neighborhood for 2024. In the following plots, bold labels on the x-axis correspond to a regime that the learned version of `TANDEM` was trained on, while lighter labels correspond to out-of-distribution regimes.

• *Tracking Coefficient η.* To understand the impact of the tracking cost coefficient η on network management decisions, we vary its value from 0 to 2 and plot the resulting approximation ratios for a single neighborhood in Fig. 4(a). We find that the performance of all algorithms generally degrades as η increases—the constant baseline is particularly affected as it does not adapt its decisions based on grid conditions. Learned `TANDEM` algorithm consistently performs well, although its performance degrades slightly for large values of η that are out-of-distribution from its training data.

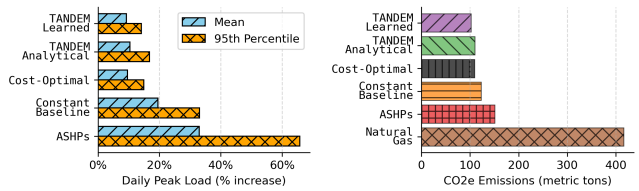
• *Flexible Demand Proportion.* We vary the proportion of homes with a daily flexible demand between 0.0 and 0.8, plotting the resulting approximation ratios in Fig. 4(b). We find that performance is generally stable, although the approximation ratio of analytical `TANDEM` degrades slightly as the number of flexible demands increases. Since the optimal solution has full knowledge, it can leverage flexible demands to substantially reduce cost—in contrast, since analytical `TANDEM` is optimized for the worst-case, it is more conservative and thus less able to exploit this flexibility.

• *Borehole Count.* We vary the number of simulated boreholes in the neighborhood from 0-3 per home and plot the resulting approximation ratios in Fig. 4(c). We retrain learned `TANDEM` for the no-borehole case—in this scenario, the system completely relies on an ASHP at the pump station to bring heat into the network, which significantly changes the problem's cost structure. We find that the constant baseline performs better for low (but non-zero) borehole counts as the pumping overhead of operating boreholes shrinks (vice versa for higher borehole counts). Both variants of `TANDEM` adapt to the amount of demand in the instance, allowing them to achieve good approximation ratios across all borehole counts.

Decarbonization and Grid Impacts. Beyond the objective of `OTNM` (i.e., cost savings), we also evaluate two ancillary benefits of our approach and 5GDHC systems broadly: reductions in distribution transformer peak loads and reductions in carbon emissions.

• *Peak Electric Load.* In Fig. 5(a), we plot the percentage increase in daily peak load on distribution transformers due to 5GDHC operation (compared to a status quo of natural gas heating) for all tested algorithms and all ten neighborhoods in 2024. Compared to deploying ASHPs to replace natural gas, a 5GDHC system under constant operation (constant baseline) reduces the daily peak load on the grid by 40.6% on average (49.7% in the 95$^{\text{th}}$ percentile), primarily due to the efficiency gains of GSHPs over ASHPs. By adaptively managing the thermal network using `TANDEM`, further reductions in peak load are achieved—learned `TANDEM` reduces peak load by an additional 31.2% on average (28.8% in the 95$^{\text{th}}$ percentile), reflecting the ability to shift network management loads to off-peak periods.

[7] This threshold is informed by the classic optimal algorithm for one-min search [15].



(a) Daily peak load increase **(b) CO_2eq. emissions**

Figure 5: For all tested strategies and neighborhoods in 2024, (a) percentage increase in daily peak load on distribution transformers (compared to natural gas status quo), and (b) estimated total yearly CO_2eq. emissions from heating.

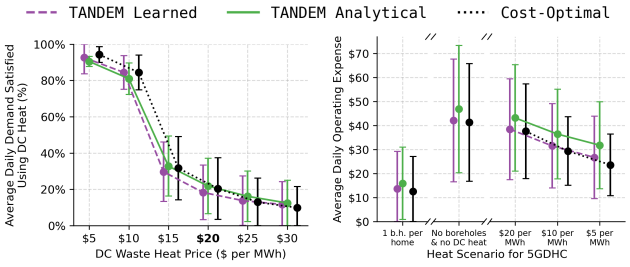
• *Carbon Emissions.* In Fig. 5(b), we estimate the CO_2eq. emissions from heating loads for all tested strategies and all ten neighborhoods in 2024. Using average carbon intensity for the ISO New England grid (36% carbon-free in 2024 [32]), we find that transitioning from natural gas heating to ASHPs already reduces GHG emissions by 63.8%. In comparison, learned TANDEM achieves a total emissions reduction of 75.4% compared to natural gas—the higher reduction is attributable to the efficiency benefits of a 5GDHC system (e.g., constant baseline, 6.7%), and adaptive management using learned TANDEM (4.9%)—the remaining reductions are obtained by switching away from natural gas heating (same as the ASHP baseline).[8]

Data Center Waste Heat. As DC growth has accelerated in recent years, interest has grown in utilizing the "low-quality" waste heat generated by DCs—since 5GDHC systems operate at low temperatures, they are well-suited to this integration [75]. In the case study city, there is an existing DC (15 MW IT load) that we model as a waste heat source, giving results for other DC sizes in Appendix A.7.

Here we assume no boreholes are installed—the 5GDHC system relies on (either) DC waste heat or an ASHP at the pump station to meet demand. We consider a range of *prices* for DC waste heat between \$5 and \$30 per thermal MWh (USD), roughly corresponding to the cost of heating using natural gas or ASHPs (see Appendix A.7). In Fig. 6(a), we plot the amount of daily heat demand met using DC waste heat under three strategies (both TANDEM variants and the cost-optimal offline solution) for a single neighborhood. We train learned TANDEM on data with a waste heat price of \$20 / thermal MWh. We find that for DC heat priced at or below \$10 / thermal MWh, all strategies choose to use DC heat to meet between 80-100% of the daily heat demand on average, demonstrating value in utilizing DC waste heat when it is cost-competitive. Although our analysis simulates a single neighborhood, the amount of available heat from the DC is constrained such that other neighborhoods can access a proportional share of it (see Section 6) When DC waste heat is priced higher, the operator's electricity cost of using an ASHP to inject heat at the pump station is sometimes favorable, leading to reduced utilization of DC heat.

In Fig. 6(b), we plot the average daily operating cost for 2024 data in one neighborhood under different heat source options: 1 borehole per home, no boreholes (i.e., using an ASHP at pump station), and DC waste heat at varying prices. Error bars indicate standard deviation across all days/seasons. Intuitively, borehole fields provide the lowest operating cost because their pumping

[8]Since these estimates use existing average carbon intensity of the grid, they do not account for changes in grid mix due to the added electric load.



(a) Demand met using waste heat **(b) Scenario cost comparison**

Figure 6: Case study of DC waste heat in a single neighborhood for 2024; (a) demand met using waste heat of varying prices, and (b) daily operating cost using boreholes (far left), ASHP-only (2nd from left), and waste heat (remaining).

cost is low compared to both ASHPs and DC waste heat. When no boreholes are available, DC waste heat is used if/when it is *more cost-effective* than using an ASHP at the pump station to inject heat into the network. As such, we find that integration of cheaper waste heat can significantly reduce operating costs—at \$5 / thermal MWh, DC waste heat reduces operating cost by 36.9% (averaged across strategies) compared to using a pump station ASHP alone.

8 Conclusion

We introduce online thermal network demand management (OTNM) to capture the decision-making problem faced in the dynamic operation of fifth-generation district heating and cooling (5GDHC) systems. We propose TANDEM, a distributed online algorithm achieving the best possible competitive ratio under our assumptions. We further extend TANDEM with an end-to-end learning framework that leverages historical data to adapt to patterns in real-world instances while retaining robustness guarantees. We conduct a comprehensive case study of 5GDHC operation in ten real neighborhoods, demonstrating that dynamic management of 5GDHC systems can achieve cost savings of 44.1% compared to the status quo of constant operation, emissions reductions of 75.4% compared to gas heating, and peak load reductions of 71.8% compared to ASHP adoption.

Limitations. Our theoretical study of OTNM and TANDEM assumes a "clean slate" design where the operator has information about and granular control over several components of a 5GDHC system (e.g., controllable smart thermostats, open communication between utility and customers, fine-grained control, etc.). In real-world settings, such a design may not be reasonable due to e.g., technical constraints of legacy infrastructure or regulatory constraints on e.g., data privacy in customer-utility communication [20]. It would be interesting to explore this generalizes to settings with limited input data, additional constraints on control inputs, or where the operator must rely on price signals/incentives over direct control.

Future Work. Several other directions remain for future work. While OTNM is chiefly motivated by the application to 5GDHC systems, a generalization of OTNM may be applicable to several domains involving online resource management in multiple dimensions, such as EV charging networks or distributed BESS systems. It would also be interesting to explore how additional 5GDHC integrations, such as thermal energy storage (TES) tanks or ground-source water heaters, might be incorporated into an extension of OTNM.

References

[1] Fatih Acun, Ioannis Ch Paschalidis, and Ayse K Coskun. 2024. Conductor: A Collaboration Framework for Multi-Data-Center Demand Response. In *2024 IEEE 15th International Green and Sustainable Computing Conference (IGSC)*. IEEE, 22–28.

[2] Akshay Agrawal, Brandon Amos, Shane Barratt, Stephen Boyd, Steven Diamond, and Zico Kolter. 2019. Differentiable Convex Optimization Layers. In *Advances in Neural Information Processing Systems*.

[3] Tine Aprianti, Evan Tan, Chan Diu, Ben Sprivulis, Greg Ryan, Kandadai Srinivasan, and Hui Tong Chua. 2021. A comparison of ground and air source heat pump performance for domestic applications: A case study in Perth, Australia. *International Journal of Energy Research* 45, 15 (2021), 20686–20699. doi:10.1002/er.7133

[4] Peder Bacher and Henrik Madsen. 2011. Identifying suitable models for the heat dynamics of buildings. *Energy and Buildings* 43, 7 (2011), 1511–1522. doi:10.1016/j.enbuild.2011.02.005

[5] Dmitrii Bogdanov, Rasul Satymov, and Christian Breyer. 2024. Impact of temperature dependent coefficient of performance of heat pumps on heating systems in national and regional energy systems modelling. *Applied Energy* 371 (2024), 123647. doi:10.1016/j.apenergy.2024.123647

[6] David Borge-Diez, Daniel Icaza, Diego Francisco Trujillo-Cueva, and Emin Açıkkalp. 2022. Renewable energy driven heat pumps decarbonization potential in existing residential buildings: Roadmap and case study of Spain. *Energy* 247 (2022), 123481. doi:10.1016/j.energy.2022.123481

[7] Roozbeh Bostandoost, Adam Lechowicz, Walid A. Hanafy, Noman Bashir, Prashant Shenoy, and Mohammad Hajiesmaili. 2024. LACS: Learning-Augmented Algorithms for Carbon-Aware Resource Scaling with Uncertain Demand. In *Proceedings of the 15th ACM International Conference on Future and Sustainable Energy Systems* (Singapore, Singapore) *(e-Energy '24)*. Association for Computing Machinery, New York, NY, USA, 27–45. doi:10.1145/3632775.3661942

[8] S. Buffa, Marco Cozzini, Matteo D'Antoni, Marco Baratieri, and Roberto Fedrizzi. 2019. 5th generation district heating and cooling systems: A review of existing cases in Europe. *Renewable and Sustainable Energy Reviews* 104 (2019), 504–522. doi:10.1016/j.rser.2018.12.059

[9] Canary Media. 2024. Massachusetts kicks off first pilot to shift gas utilities to clean heat. https://www.canarymedia.com/articles/geothermal/massachusetts-kicks-off-first-pilot-to-shift-gas-utilities-to-clean-heat

[10] Canary Media. 2024. New York will repurpose gas pipelines to pump clean heat into buildings. https://www.canarymedia.com/articles/carbon-free-buildings/new-york-will-repurpose-gas-pipelines-to-pump-clean-heat-into-buildings

[11] Robert M Corless, Gaston H Gonnet, David EG Hare, David J Jeffrey, and Donald E Knuth. 1996. On the Lambert W function. *Advances in Computational Mathematics* 5 (1996), 329–359.

[12] Nikolaos Damianakis, Gautham Ram Chandra Mouli, Pavol Bauer, and Yunhe Yu. 2023. Assessing the grid impact of Electric Vehicles, Heat Pumps & PV generation in Dutch LV distribution grids. *Applied Energy* 352 (2023), 121878. doi:10.1016/j.apenergy.2023.121878

[13] Luc Devroye, László Györfi, and Gábor Lugosi. 1996. The Bayes Error. In *A Probabilistic Theory of Pattern Recognition*. Stochastic Modelling and Applied Probability, Vol. 31. Springer-Verlag, New York, NY, 9–20. doi:10.1007/978-1-4612-0711-5_2

[14] Steven Diamond and Stephen Boyd. 2016. CVXPY: A Python-embedded modeling language for convex optimization. *Journal of Machine Learning Research* 17, 83 (2016), 1–5.

[15] Ran El-Yaniv, Amos Fiat, Richard M. Karp, and Gordon Turpin. 2001. Optimal Search and One-Way Trading Online Algorithms. *Algorithmica* 30, 1 (May 2001), 101–139. doi:10.1007/s00453-001-0003-0

[16] Salma Elmallah, Anna M Brockway, and Duncan Callaway. 2022. Can distribution grid infrastructure accommodate residential electrification and electric vehicle adoption in Northern California? *Environmental Research: Infrastructure and Sustainability* 2, 4 (Nov. 2022), 045005. doi:10.1088/2634-4505/ac949c

[17] Eversource. 2024. Geothermal Pilot in Framingham. https://www.eversource.com/residential/save-money-energy/clean-energy-options/geothermal-energy/geothermal-pilot-framingham

[18] Eversource. 2025. Geothermal Pilot Reference Guide. https://www.eversource.com/residential/save-money-energy/clean-energy-options/geothermal-energy/geothermal-pilot-reference-guide. Accessed: January 2026.

[19] Peter Fairley. 2025. Geothermal networks let cities warm and cool as one. *Nature* (December 2025). doi:10.1038/d41586-025-03932-6 Nature Outlook: Cities.

[20] Bundesamt für Sicherheit in der Informationstechnik. 2024. Smart Meter Gateway — bsi.bund.de. https://www.bsi.bund.de/EN/Themen/Verbraucherinnen-und-Verbraucher/Informationen-und-Empfehlungen/Internet-der-Dinge-Smart-leben/Smart-Meter-Gateway/smart-meter-gateway_node.html. [Accessed 13-12-2024].

[21] Kristian Gjoka, Behzad Rismanchi, and Robert H. Crawford. 2023. Fifth-generation district heating and cooling systems: A review of recent advancements and implementation barriers. *Renewable and Sustainable Energy Reviews* 171 (2023), 112997. doi:10.1016/j.rser.2022.112997

[22] Grid Status. 2025. Grid Status. https://www.gridstatus.io

[23] Gurobi Optimization, LLC. 2024. Gurobi Optimizer Reference Manual. https://www.gurobi.com

[24] Ian Hamilton. 2023. *Decarbonizing the Global Buildings Sector: Efficiency, Electrification, and Equity*. Technical Report. Center on Global Energy Policy, Columbia University. https://www.energypolicy.columbia.edu/publications/decarbonizing-the-global-buildings-sector-efficiency-electrification-and-equity/

[25] L.A. Hurtado, J.D. Rhodes, P.H. Nguyen, I.G. Kamphuis, and M.E. Webber. 2017. Quantifying demand flexibility based on structural thermal storage and comfort management of non-residential buildings: A comparison between hot and cold climate zones. *Applied Energy* 195 (June 2017), 1047–1054. doi:10.1016/j.apenergy.2017.03.004

[26] Hyunkyeong Hwang, Ahyun Yoon, Yongtae Yoon, and Seungil Moon. 2023. Demand response of HVAC systems for hosting capacity improvement in distribution networks: A comprehensive review and case study. *Renewable and Sustainable Energy Reviews* 187 (2023), 113751. doi:10.1016/j.rser.2023.113751

[27] International Energy Agency. 2024. Buildings: Energy System. https://www.iea.org/energy-system/buildings

[28] International Energy Agency. 2024. Heat Pumps: Energy System. https://www.iea.org/energy-system/buildings/heat-pumps

[29] International Organization for Standardization. 2017. ISO 52016-1:2017 Energy performance of buildings — Energy needs for heating and cooling, internal temperatures and sensible and latent heat loads — Part 1: Calculation procedures. https://www.iso.org/standard/65696.html

[30] ISO New England. 2024. CELT 2024: Heating electrification will drive higher energy use, winter peaks. https://isonewswire.com/2024/05/09/celt-2024-heating-electrification-will-drive-higher-energy-use-winter-peaks/

[31] ISO New England. 2024. *Final 2024 Heating Electrification Forecast*. Technical Report. ISO New England. https://www.iso-ne.com/static-assets/documents/100010/final-2024-heating-electrification-forecast.pdf

[32] ISO New England. 2025. *2024 Annual Markets Report*. Technical Report. ISO New England. https://www.iso-ne.com/static-assets/documents/100023/2024-annual-markets-report.pdf

[33] ISO New England. 2025. Monthly wholesale electricity prices and demand in New England, January 2025. https://isonewswire.com/2025/03/10/monthly-wholesale-electricity-prices-and-demand-in-new-england-january-2025/

[34] June Kim. 2023. Underground thermal energy networks are becoming crucial to the US's energy future. MIT Technology Review. https://www.technologyreview.com/2023/10/04/1080795/us-thermal-energy-networks/

[35] Mary Elizabeth Konrad and Brendan D. MacDonald. 2023. Cold climate air source heat pumps: Industry progress and thermodynamic analysis of market-available residential units. *Renewable and Sustainable Energy Reviews* 188 (2023), 113739. doi:10.1016/j.rser.2023.113739

[36] Ravi Kumar, Manish Purohit, and Zoya Svitkina. 2018. Improving Online Algorithms via ML Predictions. In *Advances in Neural Information Processing Systems*, S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett (Eds.), Vol. 31. Curran Associates, Inc.

[37] Adam Lechowicz, Noman Bashir, John Wamburu, Mohammad Hajiesmaili, and Prashant Shenoy. 2023. Equitable Network-Aware Decarbonization of Residential Heating at City Scale. In *Proceedings of the 14th ACM International Conference on Future Energy Systems* (Orlando, FL, USA) *(e-Energy '23)*. Association for Computing Machinery, New York, NY, USA, 1–13. doi:10.1145/3575813.3576870

[38] Adam Lechowicz, Nicolas Christianson, Mohammad Hajiesmaili, Adam Wierman, and Prashant Shenoy. 2025. Online Smoothed Demand Management. arXiv:2511.18554 [cs.DS] https://arxiv.org/abs/2511.18554

[39] Russell Lee, Bo Sun, Mohammad Hajiesmaili, and John C. S. Lui. 2024. Online Search with Predictions: Pareto-optimal Algorithm and its Applications in Energy Markets. In *Proceedings of the 15th ACM International Conference on Future and Sustainable Energy Systems* (Singapore, Singapore) *(e-Energy '24)*. Association for Computing Machinery, New York, NY, USA, 386–407. doi:10.1145/3632775.3639590

[40] Chang Li, Felix Langner, Kevin Förderer, Jörg Matthes, and Veit Hagenmeyer. 2025. A Survey on the Current State of Demand Response Implementations: Models and Approaches. *ACM SIGEnergy Energy Informatics Review* 5, 1 (2025), 12–24. doi:10.1145/3757888.3757890

[41] Mingquan Li, Qingyuan Ma, Rui Shan, Ahmed Abdulla, Edgar Virguez, Shuo Gao, and Dalia Patiño-Echeverri. 2024. Renewable energy quality trilemma and coincident wind and solar droughts. *Communications Earth & Environment* 5 (2024), 661. doi:10.1038/s43247-024-01850-5

[42] Xiang Li, Selin Yilmaz, Martin K. Patel, and Jonathan Chambers. 2023. Techno-economic analysis of fifth-generation district heating and cooling combined with seasonal borehole thermal energy storage. *Energy* 285 (Dec. 2023), 129382. doi:10.1016/j.energy.2023.129382

[43] Julian Lorenz, Konstantinos Panagiotou, and Angelika Steger. 2008. Optimal Algorithms for k-Search with Application in Option Pricing. *Algorithmica* 55, 2 (Aug. 2008), 311–328. doi:10.1007/s00453-008-9217-8

[44] Thodoris Lykouris and Sergei Vassilvtiskii. 2018. Competitive Caching with Machine Learned Advice. In *Proceedings of the 35th International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 80)*, Jennifer Dy and Andreas Krause (Eds.). PMLR, 3296–3305. https://proceedings.mlr.press/v80/lykouris18a.html

[45] Mark Manasse, Lyle McGeoch, and Daniel Sleator. 1988. Competitive Algorithms for On-Line Problems. In *Proceedings of the 20th Annual ACM Symposium on Theory of Computing* (Chicago, Illinois, USA) *(STOC '88)*. Association for Computing Machinery, New York, NY, USA, 322–333. doi:10.1145/62212.62243

[46] Electricity Maps. 2025. Electricity Map. https://www.electricitymap.org/map.

[47] Dragoslav S. Mitrinovic, Josip E. Pečarić, and A. M. Fink. 1991. *Inequalities Involving Functions and Their Integrals and Derivatives*. Vol. 53. Springer Science & Business Media.

[48] Manuel Nathenson and Marianne Guffanti. 1988. Geothermal gradients in the conterminous United States. *Journal of Geophysical Research: Solid Earth* 93, B6 (June 1988), 6437–6450. doi:10.1029/jb093ib06p06437

[49] National Centers for Environmental Information (NCEI). 2026. Integrated Surface Global Hourly Data. https://www.ncei.noaa.gov/products/land-based-station/integrated-surface-database.

[50] National Ecological Observatory Network (NEON). 2025. Soil temperature (DP1.00041.001). doi:10.48443/SACF-DC09

[51] National Grid. 2026. *Heat Pump Rate*. https://www.nationalgridus.com/MA-Home/Rates/Heat-Pump-Rate

[52] Guy R. Newsham and Brent G. Bowker. 2010. The effect of utility time-varying pricing and load control strategies on residential summer peak electricity use: A review. *Energy Policy* 38, 7 (2010), 3289–3296. doi:10.1016/j.enpol.2010.01.027

[53] North American Electric Reliability Corporation. 2025. Rising Demand, Evolving Resources Continue to Challenge Winter Grid Reliability. https://www.nerc.com/newsroom/2025-winter-reliability-assessment 2025 Winter Reliability Assessment.

[54] Northeast Energy Efficiency Partnerships. 2023. *Cold Climate Air Source Heat Pump Specification (Version 4.0)*. Technical Report. NEEP. https://neep.org/sites/default/files/media-files/cold_climate_air_source_heat_pump_specification_-_version_4.0_final.pdf Effective January 1, 2023.

[55] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Köpf, Edward Yang, Zach DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. 2019. PyTorch: An Imperative Style, High-Performance Deep Learning Library. arXiv:1912.01703 [cs.LG] https://arxiv.org/abs/1912.01703

[56] Luis Pérez-Lombard, José Ortiz, and Christine Pout. 2008. A review on buildings energy consumption information. *Energy and Buildings* 40, 3 (2008), 394–398. doi:10.1016/j.enbuild.2007.03.007

[57] PJM Interconnection. 2025. *PJM Manual 12: Balancing Operations*. PJM Interconnection. https://www.pjm.com/-/media/DotCom/documents/manuals/m12.pdf Revision 56. Effective Date: October 1, 2025.

[58] Ziyu Qu, Li He, Xinxin Ge, Fei Wang, Fei Xu, and Jinling Lu. 2024. A Two-Stage Forecasting Approach for Day-Ahead Electricity Price Based on Improved Wavelet Neural Network With ELM Initialization. *IEEE Transactions on Industry Applications* 60, 3 (May 2024), 5061–5073. doi:10.1109/tia.2024.3365456

[59] Davide Quaggiotto, Jacopo Vivian, and Angelo Zarrella. 2021. Management of a district heating network using model predictive control with and without thermal storage. *Optimization and Engineering* 22, 3 (May 2021), 1897–1919. doi:10.1007/s11081-021-09644-w

[60] Cristina Sáez Blázquez, Ignacio Martín Nieto, Javier Carrasco García, Pedro Carrasco García, Arturo Farfán Martín, and Diego González-Aguilera. 2023. Comparative Analysis of Ground Source and Air Source Heat Pump Systems under Different Conditions and Scenarios. *Energies* 16, 3 (2023), 1289. doi:10.3390/en16031289

[61] A. A. Safa, A. S. Fung, and R. Kumar. 2015. Comparative thermal performances of a ground source heat pump and a variable capacity air source heat pump systems for sustainable houses. *Applied Thermal Engineering* 81 (2015), 279–287. doi:10.1016/j.applthermaleng.2015.02.039

[62] Varun Sakalkar, Vasileios Kontorinis, David Landhuis, Shaohong Li, Darren De Ronde, Thomas Blooming, Anand Ramesh, James Kennedy, Christopher Malone, Jimmy Clidaras, and Parthasarathy Ranganathan. 2020. Data Center Power Oversubscription with a Medium Voltage Power Plane and Priority-Aware Capping. In *Proceedings of the Twenty-Fifth International Conference on Architectural Support for Programming Languages and Operating Systems*. New York, NY, USA, 497–511. https://dl.acm.org/doi/abs/10.1145/3373376.3378533

[63] Lucile Sarran, H. Burak Gunay, William O'Brien, Christian A. Hviid, and Carsten Rode. 2021. A data-driven study of thermostat overrides during demand response events. *Energy Policy* 153 (2021). doi:10.1016/j.enpol.2021.112290

[64] Dominik Schojda, Jan Scheipers, Jürgen Roes, and Harry Hoster. 2025. Modelling and Transient Simulation of District Heating Networks Based on a Control Theory Approach. *Energies* 18, 3 (Jan. 2025), 658. doi:10.3390/en18030658

[65] Shai Shalev-Shwartz and Shai Ben-David. 2014. *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press, Cambridge, UK.

[66] Pierluigi Siano. 2014. Demand response and smart grids—A survey. *Renewable and Sustainable Energy Reviews* 30 (2014), 461–478. doi:10.1016/j.rser.2013.10.022

[67] Juliet Simpson and Guangdong Zhu. 2024. An Efficient Annual-Performance Model of a Geothermal Network for Improved System Design, Operation, and Control. In *Stanford Geothermal Workshop*. Stanford, CA. https://www.nrel.gov/docs/fy24osti/88488.pdf NREL/CP-5700-88488.

[68] Bo Sun, Ali Zeynali, Tongxin Li, Mohammad Hajiesmaili, Adam Wierman, and Danny H.K. Tsang. 2020. Competitive Algorithms for the Online Multiple Knapsack Problem with Application to Electric Vehicle Charging. *Proc. ACM Meas. Anal. Comput. Syst.* 4, 3, Article 51 (Nov. 2020), 32 pages. doi:10.1145/3428336

[69] Topten. 2025. Energy Efficient Heat Pumps. https://www.topten.eu/private/products/heat_pumps.

[70] United States Environmental Protection Agency. 2011. *Opportunities for Combined Heat and Power at Wastewater Treatment Facilities: Market Analysis and Lessons from the Field*. Technical Report. U.S. Environmental Protection Agency, Combined Heat and Power Partnership. https://www.epa.gov/sites/default/files/2015-07/documents/opportunities_for_combined_heat_and_power_at_wastewater_treatment_facilities_market_analysis_and_lessons_from_the_field.pdf

[71] U.S. Department of Energy. 2017. Cold Climate Air-Source Heat Pumps: An Innovative Technology to Stay Warm in Winter and Cool in Summer. https://www.energy.gov/eere/buildings/articles/cold-climate-air-source-heat-pumps-innovative-technology-stay-warm-winter

[72] U.S. Department of Energy. 2025. Geothermal Heat Pumps. https://www.energy.gov/energysaver/geothermal-heat-pumps. Accessed: 2025.

[73] Sam van der Zwan and Ivo Pothof. 2020. Operational optimization of district heating systems with temperature limited sources. *Energy and Buildings* 226 (Nov. 2020), 110347. doi:10.1016/j.enbuild.2020.110347

[74] Viessmann Limited. 2019. Heating with airborne and geothermal energy: Vitocal Heat Pumps. https://www.viessmann.co.uk/content/dam/public-brands/gb/pdf/technology-brochures/en/Vitocal-Heat-Pumps-Brochure.pdf/_jcr_content/renditions/original./Vitocal-Heat-Pumps-Brochure.pdf.

[75] Mikko Wahlroos, Matti Pärssinen, Jukka Manner, and Sanna Syri. 2017. Utilizing data center waste heat in district heating – Impacts on energy efficiency and prospects for low-temperature district heating networks. *Energy* 140 (Dec. 2017), 1228–1238. doi:10.1016/j.energy.2017.08.078

[76] John Wamburu, Noman Bashir, David Irwin, and Prashant Shenoy. 2023. Analyzing the Impact of Decarbonizing Residential Heating on the Electric Distribution Grid. *SIGENERGY Energy Inform. Rev.* 3, 2 (June 2023), 47–60. doi:10.1145/3607114.3607119

[77] John Wamburu, Stephen Lee, Prashant Shenoy, and David Irwin. 2018. Analyzing Distribution Transformers at City Scale and the Impact of EVs and Storage. In *Proceedings of the Ninth International Conference on Future Energy Systems* (Karlsruhe, Germany) *(e-Energy '18)*. Association for Computing Machinery, New York, NY, USA, 157–167. doi:10.1145/3208903.3208925

[78] Lin Yang, Mohammad H. Hajiesmaili, Ramesh Sitaraman, Adam Wierman, Enrique Mallada, and Wing S. Wong. 2020. Online Linear Optimization with Inventory Management Constraints. *Proceedings of the ACM on Measurement and Analysis of Computing Systems* 4, 1, Article 16 (May 2020), 29 pages. doi:10.1145/3379482

[79] Kun Zhang and Michaël Kummert. 2021. Evaluating the impact of thermostat control strategies on the energy flexibility of residential buildings for space heating. *Building Simulation* 14, 5 (Jan. 2021), 1439–1452. doi:10.1007/s12273-020-0751-x

[80] Xuan Zhang, Wenbo Shi, Bin Yan, Ali Malkawi, and Na Li. 2017. Decentralized and Distributed Temperature Control via HVAC Systems in Energy Efficient Buildings. arXiv:1702.03308 [cs.SY] https://arxiv.org/abs/1702.03308

Appendix

Table 1: Summary of notation used throughout the paper.

Symbol	Meaning
Problem dimensions	
T, n, m	Horizon length; number of households; number of pump stations
Network state	
$s_t \in [0, S]$	Network thermal energy state at time t (S: max. capacity)
$\mathcal{M}$, L	Thermodynamic model of network state evolution, with maximum (T horizon-dependent) thermal loss rate $L \in [0, 1]$
Online inputs	
p_t	Electricity price at time t
$\mathbf{g}_t \in \mathbb{R}_+^m$	Pump station inverse efficiency vector at time t
$b_{i,t}$, $f_{i,t}$	Base and flexible demand of customer i and time t; $\mathbf{b}_t$, $\mathbf{f}_t$ denote vectors that collect demand across customers
$\Delta_{i,t}$	Deadline of flexible demand $f_{i,t}$
$\tilde{\mathbf{x}}_t$, $\tilde{\mathbf{z}}_t$	Tracking targets for network management and preheating/precooling decisions
$\mathcal{X}_t$, $\mathcal{Z}_{i,t}$	Feasible sets for $\mathbf{x}_t$ and $z_{i,t}$; $\mathcal{Z}_{i,t} = [-\omega_i - b_{i,t},\ \omega_i - b_{i,t}]$, where ω_i is the GSHP rate capability of customer i
Decisions	
$\mathbf{x}_t \in \mathcal{X}_t$	Network management (energy purchasing) decision at time t
$z_{i,t}$, $\mathbf{z}_t$	Preheating/precooling decision for customer i at time t; $\mathbf{z}_t = (z_{1,t}, \ldots, z_{n,t})$
Cost functions and parameters	
$c^{(b)}(\cdot)$, $c_t^{(f)}(\cdot)$	Base/flexible demand efficiency (summed over all customers: $c^{(b)}(\mathbf{b}_t) \coloneqq \sum_i \frac{\lvert b_{i,t}\rvert}{\mathtt{COP}_i}$; $c_t^{(f)}(\mathbf{z}_t) \coloneqq \sum_i c_{i,t}^{(f)} \lvert z_{i,t}\rvert$)
$c_{\max}$, $g_{\min}$, r	Worst-case inverse COP of GSHP; best-case pump station efficiency; ratio $r \coloneqq c_{\max}/g_{\min}$
β	Dynamic COP ratio: $\beta \coloneqq c_{\max}/c_{\min}$ (used in Corollary 4.5)
$\lambda_{\min}$, $\lambda_{\max}$	Lower/upper bounds on marginal cost of moving energy in/out of the network
η, γ	Tracking cost coefficients for network management and preheating/precooling
TANDEM *algorithm*	
$\mathcal{B}$, $\mathcal{F}$	Sets of active base and flexible drivers
d_i, Δ_i	Demand size and deadline of driver i
$w_{i,t}$, $v_{i,t}$	Cumulative energy purchased / flexible demand delivered by driver i up to and including time t
$\phi_i(w)$, $\psi_i(v)$	Network management / delivery threshold functions for driver i
$\Phi_i(w_1, w_2)$	Shorthand for $\int_{w_1}^{w_2} \phi_i(u)\, du$
α	Competitive ratio of TANDEM (defined in (3))
Learning framework	
$\hat{\phi}_{\boldsymbol{\theta}}$, $\hat{\psi}_{\boldsymbol{\theta}}$	Learned threshold functions (parameterized by $\boldsymbol{\theta}$)
$\boldsymbol{\theta}_\phi$, $\boldsymbol{\theta}_\psi \in \mathbb{R}_+^K$	Piecewise-affine parameter vectors (K knot points)
$\theta_{j,\phi}$, $\theta_{j,\psi}$; $\mathcal{R}_\rho$	jth knot values of $\boldsymbol{\theta}_\phi$, $\boldsymbol{\theta}_\psi$; robust certificate set (ensures competitive ratio ρ)

A Supplemental Experiments

In this section, we present additional experimental setup details and results to supplement those in Section 7.

A.1 Reduced-Order Thermodynamic Simulation Model

In this section, we describe the reduced-order thermodynamic model that we use in our experiments to model the state evolution of a 5GDHC system in our experiments.

System Topology and Linearization. Informed by prior work [64, 67], the physical network that we simulate consists of a root pump station and a set of bidirectional fluid transport pipes arranged in a tree topology. To bypass the computational overhead of full computational fluid dynamics (CFD) and complex junction mixing equations, the spatial network is linearized into a 1D "folded loop."

This is achieved by computing an Eulerian-like out-and-back traversal (e.g., via a depth-first search preorder traversal) of the tree graph. The resulting path effectively models a single continuous pipe that supplies fluid to all downstream nodes and returns it to the root. The total system length is ℓ, defined by the sum of the edge lengths in the traversal path.

Assumptions. The model relies on the following common reduced-order assumptions to enable fast simulation.

- **Plug Flow:** The fluid exhibits zero axial dispersion. Adjacent control volumes do not mix; heat transfer along the direction of flow occurs strictly via advection.
- **Incompressible Fluid:** The working fluid (water/propylene glycol mixture) has constant density and specific heat capacity (in our simulations, we simulate brine with a density of 1020 kg/m^3 and a specific heat of 3.9 kJ/kgK).

- **Uniform Geometry:** The thermal network is modeled using a constant effective internal pipe diameter D (in our simulations, we simulate pipe diameters of 150mm).
- **Operator Splitting:** The physical processes of advection and localized heat exchange (injection/extraction and ground losses) are decoupled and solved sequentially within each time step.

Spatial and Temporal Discretization. The linearized pipe of length ℓ is discretized into $\mathtt{N}$ equal-length spatial control volumes (or "nodes"), where $\Delta x = \ell/\mathtt{N}$. In our simulations, we set a fixed spatial resolution of $\Delta x = 10$ meters, and calculate $\mathtt{N}$ dynamically based on the total length of the specific neighborhood being simulated. The fluid state is represented by a 1D vector of temperatures $\mathbf{S} \in \mathbb{R}^{\mathtt{N}}$.

For a given time step Δt (in our simulations, the internal physics time step is set to $\Delta t = 60$ seconds), the continuous fluid velocity u is determined by the mass flow rate $\dot{m}$:

$$u = \frac{\dot{m}}{\rho A} \tag{4}$$

where $A = \frac{\pi}{4} D^2$ is the cross-sectional area of the pipe. The discrete spatial shift (the number of control volumes the fluid travels in one time step) is computed as:

$$\Delta k = \left\lfloor \frac{u \Delta t}{\Delta x} \right\rceil \tag{5}$$

State Evolution and Energy Balance. The state of the system $\mathbf{S}(t)$ evolves to $\mathbf{S}(t + \Delta t)$ via a two-step operator splitting method.
Advection: The purely advective transport of the fluid is modeled as a cyclic permutation of the temperature vector (assuming a closed-loop system returning to the pump station):

$$S_i^* = S_{(i-\Delta k) \bmod \mathtt{N}}(t) \quad \forall i \in \{0, 1, \ldots, \mathtt{N}-1\} \tag{6}$$

where S_i^* represents the intermediate advected temperature of the i-th control volume before localized heat transfer is applied.
Heat Exchange: Following advection, the energy balance is computed for each control volume. The mass of fluid in a single control volume is $m_c = \rho A \Delta x$. The change in temperature due to external heat injection/extraction and environmental losses is governed by:

$$m_c C_p \frac{dS_i}{dt} = \dot{Q}_{net,i}(t) - \dot{Q}_{loss,i}(t) \tag{7}$$

Discretizing this via a forward Euler step yields the final nodal temperature update:

$$S_i(t + \Delta t) = S_i^* + \frac{\Delta t}{m_c C_p} \left(\dot{Q}_{net,i}(t) - U \pi D \Delta x \left(S_i^* - S_{\text{ground}} \right) \right) \tag{8}$$

where:

- $\dot{Q}_{net,i}(t)$ is the net thermal power (in Watts) injected into control volume i at time t. For the pump station, $\dot{Q}_{net} > 0$ implies heating the fluid. For end-user nodes, $\dot{Q}_{net} < 0$ implies load extraction (heating the building, cooling the network), and $\dot{Q}_{net} > 0$ implies injection (cooling the building, heating the network).
- U is the overall linear heat transfer coefficient between the pipe and the surrounding earth (in our simulations, we set this to 1.8 W/mK).
- S_{ground} is the uniform ambient soil temperature (in our simulations, we set this based on daily historical data).

Scalar State Representation. To contextualize the total thermal energy stored/available for OTNM, the network's state of charge is normalized against its operational temperature limits as follows. A scalar network state $s_t \in [0, 1]$ is defined as the spatial average of the temperature distribution relative to $S_{\min}$ and $S_{\max}$:

$$s_t = \max\left(0, \min\left(1, \frac{\frac{1}{\mathtt{N}} \sum_{i=0}^{\mathtt{N}-1} S_i(t) - S_{\min}}{S_{\max} - S_{\min}}\right)\right) \tag{9}$$

A state of $s_t = 0$ implies the entire network has been depleted to temperature $S_{\min}$, while $s_t = 1$ implies the network is uniformly saturated at temperature $S_{\max}$. In our simulations, we set $S_{\min} = 5°$ C and $S_{\max} = 25°$ C.

A.2 TANDEM Learning Implementation Details

We provide additional details about our implementation of TANDEM using learned thresholds (piecewise-affine parameterization), as introduced in Section 5. We implement the TANDEM algorithm (particularly the pseudo-cost minimization problems) using CVXPYLayers [2, 14], which allow us to differentiate through convex optimization problems and use gradients to improve the learned threshold parameters with respect to the overall cost objective. To facilitate the flow of gradients, this "training" version of TANDEM uses a "hinge representation" of the integral over the threshold functions which is fully differentiable with respect to the piecewise-affine threshold parameters $\boldsymbol{\theta}$, and skips hard constraint enforcement (e.g., network top-ups, flexible demand delivery at deadlines) in exchange for soft penalties in the end-to-end loss function. We also implement the projection of threshold parameter vectors $\boldsymbol{\theta}$ onto the robust certificate set $\mathcal{R}_\rho$ (see Def. 5.1 and Lemma 5.3) as a CVXPYLayer problem, allowing us to "enforce" robustness during training.

We implement a small neural network in PyTorch [55] to predict two piecewise-affine threshold parameter vectors $\boldsymbol{\theta}$ (one $K+1$-length output for each of $\hat{\phi}$ and $\hat{\psi}$, respectively) given context about the current time step. As context (input to the NN), we use the following

features: time of day (24 hr format, normalized to $[0, 1]$), average daily air temperature (in °C), the season (month number, normalized to $[0, 1]$), the thermal network's current state-of-charge (e.g., s_t/S), the network's current average fluid temperature (in °C), the current electricity price (in USD/kWh), the mean, standard deviation, and min-max of predicted (day-ahead LMP forecast) electricity prices over the rest of the instance (in USD/kWh), the network management decision tracking target $\tilde{\mathbf{x}}_t \in \mathbb{R}^m$, the $\lambda_{\min}$ and $\lambda_{\max}$ marginal cost values used by TANDEM on this instance (based on the previous day's data), and a bias term.

The neural network starts as a fully feed-forward network with two hidden layers of 64 neurons each, using ReLU activations. To enforce monotonicity of the learned thresholds, we encode the constraint architecturally by having the neural network predict an initial value bounded via sigmoid activation, followed by K non-negative decrements computed via softplus. These decrements are scaled to ensure the final value remains above the minimum threshold bound (i.e., $\lambda_{\min}$), then accumulated to produce a guaranteed monotone sequence. This monotone head is applied *twice*, once for each of the two thresholds $\hat{\phi}$ and $\hat{\psi}$. The raw outputs from the monotone head(s) are then projected onto the robust certificate set $\mathcal{R}_\rho$ using the CVXPYLayer projection defined earlier, ensuring that the final learned thresholds used by TANDEM satisfy the robustness conditions.

For training, we generate a data set of historical instances based on a single neighborhood and 2023 data, which is not otherwise used in our evaluation experiments. Each instance corresponds to one day (96 time steps) of heating season operation for a given neighborhood, including historical context data (e.g., weather, price, and demand). We typically set $\eta = 1$, $\gamma = 0.1$, and `flex_prop` $= 0.4$ for these training instances, matching the default parameters used in our main experiments. We train the TANDEM neural-network model in two steps:

(1) **Pretraining**: We first pretrain the NN for 1000 epochs with target thresholds that linearly decrease from $1.2 \times \lambda_{\min}$ to $\lambda_{\min}$ for ϕ and from $1.2 \times \lambda_{\min} r$ to $\lambda_{\min} r$ for ψ, where $\lambda_{\min}$ is the instance-specific minimum marginal cost. We use the Adam optimizer with a learning rate of 10^{-4}. This provides a good initialization point for the learned thresholds and ensures they start within a reasonable range of the intuitive "correct" thresholds, which pick out the best prices/marginal costs in each instance based on context. After epoch 500 (to allow for some initial exploration) we enable the robust certificate projection layer, ensuring that predicted thresholds remain within the robust certificate set $\mathcal{R}_\rho$ for $\rho = 5\alpha$.

(2) **End-to-end Training**: Next, we fine-tune the NN end-to-end using backpropagation through the CVXPYLayer implementation of TANDEM. We use the Adam optimizer with a learning rate of 10^{-5}, training for 1000 epochs with a batch size of 16 instances. The loss function is the total cost incurred by TANDEM over each instance in the batch, plus an L_2 penalty term (with weight 10^{-4}) that "pushes" the learned thresholds up towards $\lambda_{\max}$ if constraints are violated by the differentiable implementation of TANDEM (to encourage feasibility). We again enforce the robust certificate projection layer after epoch 500 to ensure worst-case robustness.

We deploy the trained NN model in our evaluation experiments (using unseen data from 2024) using a "testing" version of TANDEM that *does* implement hard constraint enforcement (e.g., network top-ups, flexible demand delivery at deadlines) rather than soft penalties, ensuring feasibility. In evaluation, we always "enable" the robust certificate projection layer to enforce the learned thresholds remain within $\mathcal{R}_\rho$.

A.3 Neighborhood and Building Details

In Section 6, we describe how we use real natural gas and electricity consumption data for a small city in Massachusetts to simulate neighborhood-level heating demand in our 5GDHC simulation environments. We identified ten neighborhoods containing between 6 and 44 homes each that rely on natural gas for heating—below, in Table 2, we summarize high-level statistics for each of these neighborhoods.[9] In our experiments that consider a single neighborhood (e.g., instead of all ten neighborhoods), we use Neighborhood E (bolded in Table 2) as our "representative neighborhood."

Table 2: Summary statistics for the ten neighborhoods used in our experiments.

Neighborhood	**No. of Homes**	**Avg. Daily Heating Demand** (in kWh, during heating season)	**Thermal Network Length** (in km)	**Avg. Home Size** (in m^2)	**Network Capacity** (in kWh)
A	44	2291	3.86	278	3013.2
B	24	2251	1.58	243	1233.3
C	19	2085	1.50	274	1172.2
D	26	1968	1.65	145	1288.7
E	**9**	**917**	**0.77**	**323**	**603.4**
F	13	863	1.30	178	1015.9
G	6	792	0.39	194	302.7
H	9	731	0.61	198	476.2
I	7	667	0.58	161	450.9
J	6	504	0.64	238	500.4

[9] Capacity estimate is based on an "out-and-back" supply & return topology, 150mm diameter pipe, allowable temperature range of 5-25°C, and a specific heat of 3.9 kJ/kgK (brine).

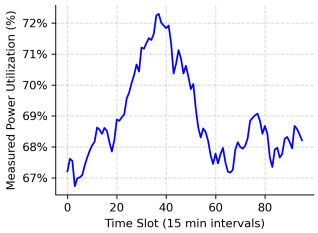


Figure 7: Average power utilization over a representative day obtained from Google power traces [62], used to simulate waste heat availability.

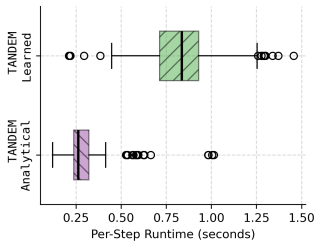
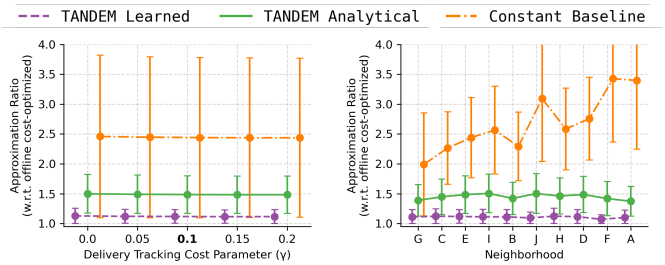


Figure 8: Per time-step runtime of analytical and learned TANDEM algorithm solutions on a Raspberry Pi Zero 2 W.



(a) Tracking coefficient γ **(b) Neighborhoods**

Figure 9: Average approximation ratio for all algorithms, with varying (a) delivery tracking coefficient (γ) in one neighborhood, and (b) results disaggregated for each neighborhood.

A.4 Google Power Trace for Data Center Waste Heat

To simulate the availability of waste heat from a data center (DC) for our case study in Section 7, we use publicly available power traces from Google data centers [62]. These traces provide time-varying power utilization data for multiple cells within a Google DC over several months. We average the utilization across all cells and PDUs to obtain a single time series representing the overall DC power usage—Fig. 7 plots this averaged utilization over a representative day in the trace.

A.5 Runtime Analysis

To demonstrate the feasibility of implementing our proposed TANDEM algorithms in real-world, resource-constrained settings, we conduct a runtime analysis to characterize their computational overhead. We measure runtime by simulating the entire algorithm operation (both network management decisions and customer-level decisions) over a full year of instances (2024) in a single neighborhood with nine homes. We run each algorithm on a Raspberry Pi Zero 2 W with a 1 GHz ARM Cortex-A53 CPU and 512 MB of RAM, representing an IoT device such as a smart thermostat or home energy management system.

In Fig. 8, we plot box plot statistics of the per-time-step runtime for the analytical and learned versions of TANDEM (recall that time steps throughout our experiments are defined as 15-minute intervals). We find that both algorithms have low computational overhead, with average per-time-step runtimes of 0.31 seconds for analytical TANDEM and 0.83 seconds for learned TANDEM. The increased overhead of learned TANDEM is primarily due to the learned threshold functions, which require multiple NN forward passes over the course of an instance and projections to the robust certificate set.

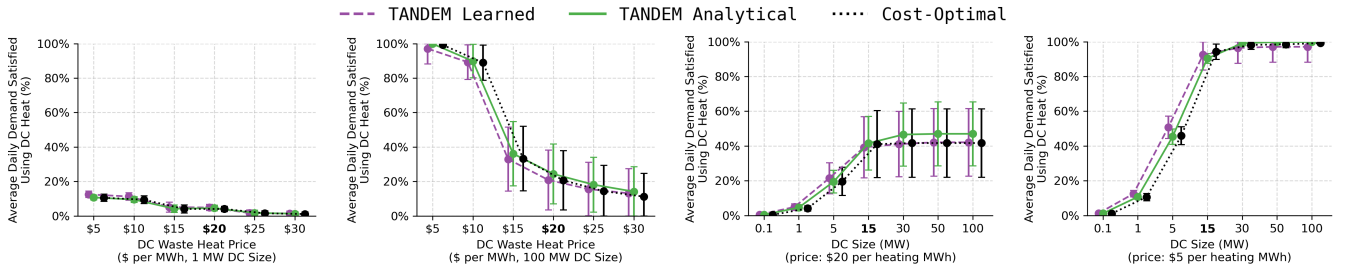
However, even with this additional overhead, both algorithms are able to make fast decisions well within the time step duration, demonstrating their feasibility for real-world deployment on resource-constrained devices. We note that these runtimes also consider the "full" operation of the algorithms, including both network management and customer-level decisions; in practice, customer-level decisions can be made independently and asynchronously, further reducing the effective computational load on any single device.

A.6 Additional Parameter Sensitivity Results

Picking up from Section 7, we explore the impact of several additional parameters on performance, varying one parameter at a time while holding others at their default values.

• *Delivery Tracking Coefficient γ.* To understand the impact of the delivery tracking cost coefficient γ on preheating/precooling decisions, we vary its value from 0 to 0.2 and plot the resulting approximation ratios for a single neighborhood in Fig. 9(a). We find that the performance of all algorithms is generally stable as γ varies—in the setting of our experiments, this makes sense as all algorithms (including OPT) have a limited ability to impact this cost component (e.g., the only lever to reduce this cost is changing preheating/precooling decisions, which are only a portion of demand in 40% of homes).

• *Neighborhoods.* In Section 7, we present our main results aggregated across all ten neighborhoods. In Fig. 9(b), we disaggregate those results to show performance for each algorithm in each neighborhood. Neighborhoods are ordered from left to right in descending order of the ratio of avg. heating demand to number of homes (e.g., see Table 2). We find that the relative performance of all algorithms is consistent across neighborhoods, with learned TANDEM consistently achieving the best approximation ratios. In this experiment we model one borehole per home in all neighborhoods, but observe trends similar to the number of boreholes sensitivity experiment in Fig. 4(c)—the constant baseline's performance trends worse for neighborhoods with a lower ratio of heating demand to homes, which corresponds to a lower ratio of demand to borehole capacity. As in Section 7, TANDEM variants are better able to adapt to the amount of demand in the instance, allowing them to achieve good approximation ratios across all neighborhoods.



(a) DC size: 1 MW **(b) DC size: 100 MW** **(c) Fixed price: $20/thermal MWh** **(d) Fixed price: $5/thermal MWh**

Figure 10: Demand met using waste heat from data centers of varying size (in MW of peak IT load) under several DC size and waste heat price scenarios. In (a) and (b), we fix the DC size (in terms of peak IT load), and vary the price of waste heat offered to the 5GDHC operator. In (c) and (d), we fix the waste heat price and vary DC size—(c) is a moderate price case, while (d) is a low price case.

A.7 Case Study: Data Center Waste Heat Supplemental Results

In Section 7, we present a case study of using waste heat from a data center (DC) to provide heat for a 5GDHC system—in the main body, we focus on a single "size" of data center and omit some details of our experimental setup for brevity. Here, we provide additional details and results to supplement that case study.

In the main body, we consider a range of prices for DC waste heat between \$5 / thermal MWh to \$30 / thermal MWh (in USD). The rationale for this range is as follows: In 2024-25, wholesale natural gas prices (all USD) in the region ranged between \$3 / MMBtu (average) to \$17 / MMBtu (winter peak)—assuming a boiler efficiency of 90%, this corresponds to a cost of between \$10 / thermal MWh to \$61 / thermal MWh for heating using natural gas. Similarly, wholesale electricity prices ranged between \$40 / MWh (average) to \$135 / MWh (winter spike) over the same period—assuming a winter ASHP COP of 2-2.5, this corresponds to a cost of between \$16 / thermal MWh to \$68 / thermal MWh for heating using ASHPs [32, 33]. Given this, we consider a range of waste heat prices at or below the cost of heating using natural gas or electricity, ranging from \$5 / thermal MWh (to represent very cheap waste heat) to \$30 / thermal MWh (to represent waste heat that is competitive with other heating options, but not always the cheapest option).

• *Varying DC Size.* In these supplemental experiments, we explore the impact of varying the size of the data center (DC) providing waste heat to the 5GDHC system, in addition to varying the price of waste heat as in the main body. In the main body we considered a DC with a peak IT load of 15 MW, based on a model data center in the actual city under study. In what follows, we consider several additional DC sizes, including 0.1 MW (edge / micro DC), 1 MW (small DC), 5 MW (small DC), 30 MW (medium DC), 50 MW (large DC), and 100 MW (hyperscale DC). We apply the same Google power utilization trace [62] to model waste heat availability for each DC size. Note that although the DC changes size, we assume that the available waste heat is still partitioned evenly across homes currently using natural gas in the case study city (see Section 6), which constrains the amount of waste heat that can be used by each neighborhood.

In Fig. 10(a) and Fig. 10(b), we first explore additional variants of Fig. 6(a) from the main body, showing the amount of heating demand met using DC waste heat for varying waste heat prices under two (extreme) scenarios for the DC size: a very small DC with a peak IT load of 1 MW (Fig. 10(a)) and a very large DC with a peak IT load of 100 MW (Fig. 10(b)). In both cases, the amount of demand met using waste heat decreases as the price of waste heat increases, as expected. However, the impact of DC size is significant—in the small DC case (Fig. 10(a)), the maximum amount of demand met using waste heat is limited to around 10% even when waste heat is very cheap, due to the limited amount of waste heat available from the small DC. On the other hand, in the large DC case (Fig. 10(b)), nearly all demand can be met using waste heat when it is cheap (i.e., at or less than \$10 / thermal MWh), as the large DC provides ample waste heat to satisfy the neighborhood's heating demand. Notably though, at moderate prices (e.g., \$15 / thermal MWh and above), the amount of demand met using waste heat is similar to that in the main body for the 15 MW DC case, indicating that at these prices the waste heat price (rather than availability) is the main limiting factor on waste heat usage.

In Fig. 10(c) and Fig. 10(d), we consider varying the data center size while keeping the waste heat price constant, showing results for a moderate waste heat price of \$20 / thermal MWh (Fig. 10(c)) and a cheap waste heat price of \$5 / thermal MWh (Fig. 10(d)). In both cases, the amount of demand met using waste heat increases with DC size on the low end—very small DCs (e.g., 0.1 MW, 1 MW) provide little waste heat relative to the city's heating demand (across neighborhoods), limiting their ability to satisfy demand. As DC size increases, more waste heat becomes available, allowing more demand to be met. However, beyond a certain DC size (e.g., 15 MW and above), the amount of demand met using waste heat plateaus—we find that the main driver of this is the price of the waste heat. Comparing Fig. 10(c) and (d), we see that at the moderate waste heat price of \$20 / thermal MWh, waste heat is cheaper than using an ASHP to inject heat into the network a little less than half the time (based on electricity prices and ASHP COP values in 2024)—as a result, even when very large DCs provide waste heat, the amount of demand met using waste heat plateaus around 40% of daily demand. This is in contrast to the cheap waste heat price of \$5 / thermal MWh, where waste heat is almost always the cheapest heating option—in this case, we see that nearly all demand can be met using waste heat from DCs of size 15 MW and above.

B Deferred Proofs

In this section, we provide deferred proofs for theoretical results stated in Sections 3—5.

B.1 Proof of Theorem 3.10

In this section, we prove Theorem 3.10, which states that given a data set $\mathcal{S}$ of N historical instances of OTNM, there exists a classifier predicting the sign of the net demand D whose misclassification rate is bounded by δ with probability at least $1-\kappa$ over the random draw of the data set, provided that N is sufficiently large. Then, for a new instance $\mathcal{I}$ of OTNM, an α-competitive algorithm ALG that assumes Asm. 3.9 and uses the predicted sign of D satisfies:

$$\mathbb{E}\left[\texttt{ALG}(\mathcal{I})\right] \leq (\alpha + (\zeta - \alpha)\delta) \cdot \texttt{OPT}(\mathcal{I}) + \cdot C,$$

with probability at least $1-\kappa$, where $\zeta = O\left(\frac{\lambda_{\max}+\eta+\gamma}{\lambda_{\min}}\right)$ is the competitive ratio attainable by an algorithm that only ensures feasibility.

Proof. Suppose we are given a data set $\mathcal{S} \coloneqq \{X_i, Y_i\}_{i=1}^N$ of N historical instances of OTNM, where X_i denotes day-ahead features (e.g., weather forecasts, season) and $Y_i \in \{+,-\}$ denotes the sign of the net demand.

PAC learning results guarantee the existence of a classifier where the probability of misclassifying the sign of the net demand on a new instance is *bounded*. Specifically, if a classifier $h(\cdot)$ is learned from a hypothesis class with VC dimension d, we have that for any new daily instance of OTNM (given by (X, Y)) drawn from the same distribution as $\mathcal{S}$, the probability that $h(X) \neq Y$ (i.e., the misclassification probability) is bounded as:

$$\Pr_{(X,Y)}\left[h(X) \neq Y\right] \leq \hat{\text{err}}(h) + O\left(\sqrt{\frac{d + \log(1/\kappa)}{N}}\right),$$

with probability at least $1-\kappa$ over the random draw of the data set $\mathcal{S}$, where $\hat{\text{err}}(h)$ is the empirical error of h on $\mathcal{S}$ [65, Theorem 6.8].

Let $\delta \coloneqq \hat{\text{err}}(h) + O\left(\sqrt{\frac{d+\log(1/\kappa)}{N}}\right)$ denote the upper bound on the misclassification probability. This *sign-prediction error* is task- and feature-dependent: in general it can be no smaller than the Bayes error for the chosen feature set [13], and in degenerate cases (e.g., if Y is independent of the features X, such as if Y is a coin flip), any classifier must err with probability at least 1/2. In practice, $\hat{\text{err}}(h)$ can be empirically assessed from historical data using a held-out test set; the generalization bounds above then convert the observed test error into a high-confidence upper bound (i.e., δ). Note that the confidence parameter κ is user-specified and controls the probability that this estimation/generalization statement fails.

For a random new day (X, Y) drawn from the same distribution as $\mathcal{S}$, we have two cases: if the prediction is correct (with probability at least $1-\delta$), then the α-competitive ALG incurs cost of at most $\alpha\texttt{OPT}(\mathcal{I}) + C$. If the prediction is incorrect (with probability at most δ), then we must consider the cost incurred under a "fallback" case wherein ALG must ensure feasibility without optimizing cost. Consider an online algorithm that ensures feasibility for OTNM. Let $\mathcal{I}$ denote any instance of OTNM, and let $\texttt{ALG}_{\text{feas}}(\mathcal{I})$ denote the cost incurred by the feasibility algorithm on instance $\mathcal{I}$. The objective of OTNM contains three components: the cost of managing the thermal network, the cost of delivering demand to customers, and the tracking cost.

Suppose D_b denotes the net base demand over the entire time horizon, i.e., $D_b = \sum_{t=1}^T \mathbf{1}^\top(\mathbf{b}_t)$, and let $|D_b|$ denote the total absolute base demand over the entire time horizon, i.e., $|D_b| = \sum_{t=1}^T \mathbf{1}^\top|\mathbf{b}_t|$, where the absolute value is taken element-wise w.r.t. $\mathbf{b}_t$. Also, let $|D_f| = \sum_{i=1}^n \sum_{\tau=1}^T |f_{i,\tau}|$ denote the absolute flexible drivers' demand over the entire time horizon (where the absolute value is applied element-wise w.r.t. $\mathbf{f}_\tau$) and let $D_f^{+,-}$ denote the sum of flexible drivers' demand that is aligned with the sign of the net demand D_b over the entire time horizon.

Then, by definition, the cost incurred by the offline optimal solution $\texttt{OPT}(\mathcal{I})$ is at least:

$$\texttt{OPT}(\mathcal{I}) \geq \lambda_{\min} \cdot \left(D_b + D_f^{+,-}\right) + \lambda_{\min} r \left(|D_b| + |D_f|\right),$$

where $\lambda_{\min}$ is the minimum unit cost of network management decisions over the time horizon, and $\lambda_{\min} r$ is the minimum unit cost of delivering demand (i.e., running GSHPs at houses) over the time horizon. Note that OPT cannot do better than incurring 0 tracking cost.

Then, consider the cost incurred by the online algorithm that ensures feasibility:

$$\texttt{ALG}_{\text{feas}}(\mathcal{I}) \leq (\lambda_{\max} + 2\eta) \cdot \left(D_b + D_f^{+,-}\right) + (\lambda_{\max} r + 2\gamma)\left(|D_b| + |D_f|\right) + \lambda_{\max} \cdot S,$$

where $\lambda_{\max}$ is the maximum unit cost of network management decisions over the time horizon, $\lambda_{\max} r$ is an upper bound on the maximum unit cost of delivering demand (i.e. running GSHPs at houses) over the time horizon, and 2η, 2γ are the worst-case tracking cost coefficients. Note that any energy management decisions in "the wrong direction" (i.e., due to the misclassification of the sign of D) have cost bounded by the total storage capacity $\lambda_{\max} S$. E.g., if ALG "thinks" that $\mathcal{I}$ is a cooling instance but it is actually a heating instance, then ALG may expend at most $\lambda_{\max} \cdot S$ cost to fully cool the storage before the minimum temperature constraint is binding—once that is the case, ALG will be forced to only make energy management decisions that ensure feasibility, whose cost is upper bounded by $\lambda_{\max} \cdot \left(D_b + D_f^{+,-}\right)$.

Then, letting $C' = (1-\delta)C + \delta\lambda_{\max}S$ denote a constant independent of the instance $\mathcal{I}$ and probabilistically considering both cases, we have:

$$\mathbb{E}\left[\frac{\texttt{ALG}(\mathcal{I})}{\texttt{OPT}(\mathcal{I})}\right] \leq (1-\delta)\cdot\alpha + \delta\cdot\left(\frac{(\lambda_{\max}+2\eta)\cdot\left(D_b + D_f^{+,-}\right) + (\lambda_{\max}r + 2\gamma)\left(|D_b| + |D_f|\right)}{\lambda_{\min}\cdot\left(D_b + D_f^{+,-}\right) + \lambda_{\min}r\left(|D_b| + |D_f|\right)}\right).$$

Letting $\zeta \coloneqq O\left(\frac{\lambda_{\max}+\eta+\gamma}{\lambda_{\min}}\right)$, we have:

$$\mathbb{E}\left[\frac{\texttt{ALG}(\mathcal{I})}{\texttt{OPT}(\mathcal{I})}\right] \leq \alpha + (\zeta-\alpha)\delta.$$

Thus, we conclude that, given a data set $\mathcal{S}$ of N historical instances of OTNM, there exists a classifier predicting the sign of the net demand D whose misclassification rate is bounded by δ with probability at least $1-\kappa$. Then, for a new instance $\mathcal{I}$ of OTNM, an α-competitive algorithm ALG that assumes Asm. 3.9 and uses the predicted sign of D satisfies:

$$\mathbb{E}\left[\texttt{ALG}(\mathcal{I})\right] \leq (\alpha + (\zeta-\alpha)\delta)\cdot\texttt{OPT}(\mathcal{I}) + C',$$

with probability at least $1-\kappa$, where $\zeta = O\left(\frac{\lambda_{\max}+\eta+\gamma}{\lambda_{\min}}\right)$ is the competitive ratio attainable by an algorithm that only ensures feasibility. □

B.2 Proof of Theorem 4.3

In this section, we prove Theorem 4.3, which states that the TANDEM algorithm is α-competitive for online thermal network demand management (OTNM), where α is defined in (3). We start by proving the feasibility of the solution produced by TANDEM, before proceeding to prove the competitive ratio.

Proof. Before the main competitive proof, we start our analysis by showing that the solution produced by TANDEM is feasible for OTNM.

Lemma B.1. TANDEM *produces a feasible solution for* OTNM*, i.e., it satisfies all demand constraints, storage constraints, and rate constraints.*

Proof. We prove feasibility directly by the definition of the algorithm—there are three main components to check: *(i)* demand satisfaction, *(ii)* network capacity constraints, and *(iii)* rate constraints.

Note that base demand ($\{\mathbf{b}_t\}_{t\in[T]}$) is always satisfied by construction of OTNM. We must show that flexible demands ($\{\mathbf{f}_t\}_{t\in[T]}$) are satisfied before or at their respective deadlines $\{\mathbf{\Delta}_t\}_{t\in[T]}$. We start by considering the delivery (i.e., preheating or precooling) decision at customers—in Algorithm 1, Line 4 ensures that any remaining flexible demand that reaches its deadline at time t is fully satisfied at that time. Note that by Assumption 3.5, this is always a feasible decision. For any flexible demand that is aligned with the sign of the net demand D, Lines 11-12 in Algorithm 2 ensure that the network manager allocates sufficient energy to satisfy the demand during its lifetime (i.e., before its deadline), satisfying the constraint on $\mathbf{x}_t$ w.r.t. flexible demands.

Next, we show that the network capacity constraints are never violated. First, for each time step, Line 18 in Algorithm 2 ensures that sufficient energy is procured to satisfy the net demand (including flexible demand deliveries) and maintain a non-negative storage level. Note that by Assumption 3.5, this is always a feasible decision, and by Assumption 3.4, namely that the state $s_t = 0$ corresponds to a case where the network is at or below ground temperature, it is sufficient to consider the inflow and outflow of energy on line 18 without considering thermal losses when the network must be "topped up" this way. Next, TANDEM ensures that the network is always returned to its initial state at the end of the time horizon (i.e., $s_T \geq s_0$) by Line 20 in Algorithm 2, which again is always feasible by Assumption 3.5.

Finally, we show that TANDEM respects all rate constraints. We start by considering the delivery decisions made at customers. By definition, the base demand sequence respects the rate constraints of customer heat pumps (i.e., $b_{i,t} \in [-\omega_i, \omega_i]$ for all i and t). Further, in Algorithm 1, Line 2 computes a feasible set for each delivery decision $z_{i,t}$ that respects the rate constraints of customer i *after* taking into consideration any base demand that must be simultaneously served. Note that rate constraints are non-binding at the deadline of any flexible demand (e.g., Line 4 in Algorithm 1) by Assumption 3.5. Next, we consider the network management decisions made by the central controller. By construction, in Algorithm 2, Lines 8, 12, and 16 keep track of the remaining capacity (i.e., $\bar{\mathcal{X}}_t$) in the feasible decision set $\mathcal{X}_t$. The remaining capacity is communicated to each driver when solving for their respective decisions–note that Line 11 (increasing $\mathbf{x}_t$ to procure a flexible demand aligned with the sign of the net demand D at its deadline) is always feasible by Assumption 3.5–meaning that the elementwise sum $\sum_{i\in\mathcal{B}\cup\mathcal{F}}\mathbf{x}_{i,t}$ results in a decision $\mathbf{x}_t$ that is always in $\mathcal{X}_t$. Finally, Lines 19 and 20 in Algorithm 2 are always feasible by Assumption 3.5—irrespective of the current state or demand, the pump stations have enough capacity to maintain the network's capacity and/or restore the network to its initial state at the end of the time horizon. □

We now proceed to prove the competitive ratio of TANDEM. We start by considering the competitive ratio in a setting without rate constraints (i.e., where $\mathbf{x}_t$ is unconstrained for all t), and later extend the proof to the general case with rate constraints (see Lemma B.9). To do so, we introduce some notation and a definition that will be useful in our analysis. For notational brevity, the following considers any arbitrary instance of OTNM $\mathcal{I} \in \Omega$ that satisfies Assumptions 3.2 - 3.9, where Ω is the set of all such instances. For ease of presentation, the following generally assumes that the sign of the net demand is positive (i.e., that the instance $\mathcal{I}$ needs heating, see Asm. 3.9). If this is not the case,

swapping the roles of heating and cooling demands throughout yields the result—we point to specific details that require extra consideration using footnotes throughout.

Definition B.2 (Active & Inactive Periods). We partition the time horizon $t \in \{1, \dots, T\}$ into active and inactive periods based on TANDEM*'s solution as follows. An active period is a contiguous interval where the thermal network's effective storage is non-empty (i.e., $s_t > 0$).*[10] *An inactive period is any interval that lies in-between two adjacent active periods.*

We use the following additional notations. Suppose that for an arbitrary instance, TANDEM's actions result in K active periods, indexed by $k \in \{1, \dots, K\}$. During the k^{th} active period, we denote v_k as the number of base demand drivers that are created. Further, we let $\hat{w}^{(b)}_{k,j}$ denote the total amount of energy purchased by TANDEM for the j^{th} base demand driver by the end of the k^{th} active period, where $j \in \{1, \dots, v_k\}$, and let $B^{(b)}_{k,j}$ denote the total demand associated with the same driver.

Flexible demand drivers are indexed by the time at which they were created (i.e., $\tau \in \{1, \dots, T\}$) and customer they belong to (i.e., $i \in \{1, \dots, n\}$). We let $\hat{w}^{(f)}_{i,\tau}$ and $\hat{v}^{(f)}_{i,\tau}$ denote the total amount of energy purchased and delivered before the deadline by TANDEM for the flexible demand driver created at time τ for customer i – note that $f_{i,\tau}$ is the total demand associated with this driver, and if $f_{i,\tau} \le 0$, we have $\hat{w}^{(f)}_{i,\tau} = 0$.

We let $B = \sum_{k=1}^{K} \sum_{j=1}^{v_k} B^{(b)}_{k,j}$ denote the sum of base drivers' demand over the entire time horizon, let $|D_f| = \sum_{i=1}^{n} \sum_{\tau=1}^{T} |f_{i,\tau}|$ denote the absolute flexible drivers' demand over the entire time horizon (where the absolute value is applied element-wise w.r.t. $\mathbf{f}_\tau$) and let $D^+_f = \sum_{i=1}^{n} \sum_{\tau=1}^{T} \max\{0, f_{i,\tau}\}$ denote the sum of positive flexible drivers' demand over the entire time horizon. We let $\hat{s}$ denote the final status of the storage at the end of the time horizon in TANDEM's solution, i.e., $\hat{s} = s_T$.

Finally, we introduce notation used to characterize the optimal offline solution: let $G_k(\beta)$ denote the minimum cost of purchasing β units of energy during the k^{th} active period, let $H_{i,\tau}(\xi)$ denote the minimum cost of purchasing ξ units of asset during the period $[\tau, \Delta_{i,\tau}]$ (i.e., during the lifetime of the τ^{th} flexible demand at customer i), and let $E_{i,\tau}(f)$ denote the minimum cost of delivering f units of asset during the period $[\tau, \Delta_{i,\tau}]$.

We let D denote the net demand over the entire time horizon, i.e., $D = \sum_{t=1}^{T} \mathbf{1}^\top (\mathbf{b}_t + \mathbf{f}_t)$, let D_b denote the net base demand over the time horizon (i.e., less any positive flexible demand), i.e., $D_b = \sum_{t=1}^{T} \mathbf{1}^\top \mathbf{b}_t + \sum_{t=1}^{T} \mathbf{1}^\top \min\{\mathbf{0}, \mathbf{f}_t\}$, where the min is taken element-wise w.r.t. $\mathbf{f}_t$. We let $|D_b|$ denote the total absolute base demand over the entire time horizon, i.e., $|D_b| = \sum_{t=1}^{T} \mathbf{1}^\top |\mathbf{b}_t|$, where the absolute value is taken element-wise w.r.t. $\mathbf{b}_t$. Let $\tilde{pg}$ denote the minimum possible cost of purchasing energy suring the idle periods, and let $\bar{pc}$ denote the weighted average price during periods with non-zero base demand, i.e., $\bar{pc} = \frac{\sum_{\tau=1}^{T} p_t c^{(b)}(\mathbf{b}_t)}{|D_b|}$.

With these preliminaries, we are ready to proceed with the main competitive analysis. We begin by proving a lower bound on the cost incurred by the optimal offline solution OPT.

Lemma B.3. *Given that* TANDEM *produces K active periods, let β_k denote the energy purchased towards the base demand by the offline optimal solution during the k^{th} active period, let $\xi_{i,\tau} = \max\{0, f_{i,\tau}\}$ denote the energy purchased towards the flexible demand driver created at time τ for customer i, and let $\tilde{pg}$ and $\bar{pc}$ be defined as outlined above. Then, $\mathtt{OPT}(\mathcal{I})$ is lower bounded as follows:*

$$\mathtt{OPT}(\mathcal{I}) \ge \sum_{k=1}^{K} G_k(\beta_k) + \left(D_b - \sum_{k=1}^{K} \beta_k\right) \tilde{pg} + \bar{pc}|D_b| + \sum_{i=1}^{n} \sum_{\tau=1}^{T} \left(H_{i,\tau}(\xi_{i,\tau}) + E_{i,\tau}(f_{i,\tau})\right).$$

Proof. The cost of an optimal offline solution, denoted as $\mathtt{OPT}(\mathcal{I})$, can be decomposed into four components. The first and second components are the costs of purchasing energy towards the base demand and delivering the base demand, respectively. The third and fourth components are the costs of purchasing energy towards the flexible demand and delivering the flexible demand, respectively. We analyze each component in turn.

Recall that during the k^{th} active period, the optimal offline solution purchases β_k units of energy towards the base demand. By definition, the minimum cost of purchasing this energy is $G_k(\beta_k)$. Summing over all active periods, we obtain that the total cost of purchasing energy towards the base demand during active periods is at least $\sum_{k=1}^{K} G_k(\beta_k)$. Then, if the optimal solution needs to purchase additional energy towards the base demand during inactive periods, the minimum cost of doing so is $\left(D_b - \sum_{k=1}^{K} \beta_k\right) \tilde{pg}$ (by definition of $\tilde{pg}$), and note that D_b is the minimum possible energy that must be purchased over the entire time horizon (ignoring the flexible demand purchasing constraint). Thus, the total cost of purchasing energy towards the base demand is at least $\sum_{k=1}^{K} G_k(\beta_k) + \left(D_b - \sum_{k=1}^{K} \beta_k\right) \tilde{pg}$. In the objective of OTNM, the cost of delivering the base demand is given by $\sum_{t=1}^{T} p_t c^{(b)}(\mathbf{b}_t)$, which can be rewritten as $\bar{pc}|D_b|$ by definition of $\bar{pc}$.

Next, recall that for each unit of flexible demand (e.g., arriving at time τ at customer i), $\xi_{i,\tau} = \max\{0, f_{i,\tau}\}$ units of energy must be purchased towards this demand by the purchasing constraint, and the least cost of doing so during the interval $[\tau, \Delta_{i,\tau}]$ is $H_{i,\tau}(\xi_{i,\tau})$. The least cost of delivering this flexible demand during the same interval is $E_{i,\tau}(f_{i,\tau})$. Summing over all flexible demand units, we obtain that the total cost towards the flexible demand is at least $\sum_{i=1}^{n} \sum_{\tau=1}^{T} \left(H_{i,\tau}(\xi_{i,\tau}) + E_{i,\tau}(f_{i,\tau})\right)$. This completes the proof. □

After Lemma B.3, we now proceed to upper bound the cost incurred by TANDEM's solution.

[10] If the sign of the net demand is negative (i.e., the instance requires cooling), then a "non-empty" storage is redefined as $s_t < S$. Note that the role of the network is generally reversed in this case, as the network is actually storing a "lack of energy" to accommodate extra heat exchanged from homes into the network.

Lemma B.4. *Given that* TANDEM *produces K active periods, let $\hat{w}^{(b)}_{k,j}$, $B^{(b)}_{k,j}$, B, $\hat{w}^{(f)}_{i,\tau}$, and $\hat{v}^{(f)}_{i,\tau}$ be defined as outlined above. Then, the cost incurred by* TANDEM*'s solution is upper bounded as:*

$$\texttt{TANDEM}(\mathcal{I}) \leq \sum_{k=1}^{K}\sum_{j=1}^{v_k}\int_0^{\hat{w}^{(b)}_{k,j}}\phi^{(b)}_{k,j}(y)dy + \sum_{\tau=1}^{T}\sum_{i=1}^{n}\left(\int_0^{\hat{w}^{(f)}_{i,\tau}}\phi^{(f)}_{i,\tau}(y)dy + \int_0^{\hat{v}^{(f)}_{i,\tau}}\psi_{i,\tau}(y)dy\right) \tag{10}$$

$$+\left(B - \sum_{k=1}^{K}\sum_{j=1}^{v_k}\hat{w}^{(b)}_{k,j}\right)(\lambda_{\max}+2\eta) + \bar{pc}|D_b| + \left(D_f^+ - \sum_{i=1}^{n}\sum_{\tau=1}^{T}\hat{w}^{(f)}_{i,\tau}\right)(\lambda_{\max}+2\eta) + \left(|D_f| - \sum_{i=1}^{n}\sum_{\tau=1}^{T}\hat{v}^{(f)}_{i,\tau}\right)(\lambda_{\max}r+2\gamma) \tag{11}$$

$$+ L\cdot S\,(\lambda_{\max}+2\eta) \tag{12}$$

Proof. This follows by explicitly characterizing the worst-case cost of TANDEM using the definitions in Algorithm 1 and Algorithm 2.

First, (10) corresponds to the worst-case cost that can be charged to each driver's pseudo-cost threshold functions, using the definitions of the pseudo-cost minimization subproblems in Algorithm 1 and Algorithm 2. In short, TANDEM only purchases energy or delivers flexible demand when the marginal cost of doing so (e.g., the market price p_t combined with the current efficiency, either $\mathbf{g}_t$ or $c^{(f)}_{i,t}$) is sufficiently low to make the minimization problems on Line 15 (Algorithm 2) and Line 3 (Algorithm 1) negative. For instance, we can say that the cost of purchasing energy attributable to the j^{th} base driver in the k^{th} active period is upper bounded by $\int_0^{\hat{w}^{(b)}_{k,j}}\phi^{(b)}_{k,j}(y)dy$. Similar reasoning applies to the flexible demand drivers' purchasing and delivery costs, and aggregating over all drivers yields (10).

Next, (11) corresponds to the worst-case cost incurred by TANDEM for purchasing energy, delivering demand, and generally satisfying any constraints that are not already accounted for by integrating over the pseudo-cost thresholds. We start by considering *base demand*. Recall that B is the total size of base demand drivers over the entire time horizon, and that $\sum_{k=1}^{K}\sum_{j=1}^{v_k}\hat{w}^{(b)}_{k,j}$ is the total amount of energy purchased by TANDEM towards base demand drivers – any additional purchasing is captured by the difference, and in the worst case, this additional purchasing is done at the worst price $\lambda_{\max}$ and incurs an additional tracking cost of 2η per unit of energy purchased. Since base demand must be delivered at the time it arrives, the delivery cost is captured by $\bar{pc}|D_b|$ (as in Lemma B.3). Next, we consider *flexible demand*. Recall that D_f^+ is the total size of positive flexible demand drivers over the entire time horizon (these are the units of flexible demand that require energy purchasing), and that $\sum_{i=1}^{n}\sum_{\tau=1}^{T}\hat{w}^{(f)}_{i,\tau}$ is the total amount of energy purchased by TANDEM towards flexible demand drivers – any additional purchasing is captured by the difference, and in the worst case, this additional purchasing is done at the worst price $\lambda_{\max}$ and incurs an additional (worst-case) tracking cost of 2η per unit of energy purchased. Finally, since $|D_f|$ is the total (absolute) size of flexible demand drivers over the entire time horizon, and that $\sum_{i=1}^{n}\sum_{\tau=1}^{T}\hat{v}^{(f)}_{i,\tau}$ is the total amount of flexible demand delivered by TANDEM– any additional delivery is captured by the difference, and in the worst case, this additional delivery is done at the worst marginal cost (which is itself upper bounded by $\lambda_{\max}r$) and incurs an additional (worst-case) tracking cost of 2γ per unit of flexible demand delivered.

Finally, (12) corresponds to any excess energy that remains stored at the end of the time horizon, as well as a loss term that captures possible losses incurred within the thermal network (e.g., heat lost to surrounding earth). In OTNM, we let $L \in [0,1]$ denote the maximum possible loss rate within the thermal network over a time period of T steps, and thus, in the worst-case, TANDEM must make up for $L\cdot S$ units of lost energy at the worst possible cost of $\lambda_{\max}+2\eta$ per unit of energy lost. This completes the proof. □

In addition to the above lemmas, the following technical lemmas are necessary to prove a relation between the threshold functions and OPT.

Lemma B.5. *By definition of the threshold function $\phi(\cdot)$, the following relation always holds:*

$$\int_0^{w}\phi(y)dy + (1-w)(\lambda_{\max}+2\eta) \leq \alpha\left[\phi(w) - 2\eta\right], \quad \forall w \in [0,1].$$

Lemma B.6. *By definition of the threshold functions $\phi(\cdot)$ and $\psi(\cdot)$, the following relations always hold:*

$$\int_0^{w}\phi(y)dy + (1-w)(\lambda_{\max}+2\eta) + \int_0^{v}\psi(y)dy + (1-v)(\lambda_{\max}r+2\gamma) \leq \alpha\left[\phi(w) - 2\eta + \psi(v) - 2\gamma\right], \quad \forall w,v \in [0,1].$$

$$\int_0^{v}\psi(y)dy + (1-v)(\lambda_{\max}r+2\gamma) \leq \alpha\left[\psi(v) - 2\gamma\right], \quad \forall v \in [0,1].$$

Using the previous lemmas, we now prove the competitive ratio. Using the results in Lemmas B.3, B.4, B.5 and B.6, we claim that the following holds:

$$\frac{\texttt{TANDEM}(\mathcal{I}) - (\lambda_{\max}+2\eta)\cdot(S+LS)}{\texttt{OPT}(\mathcal{I})} \leq \alpha.$$

To show this result, we first substitute the bounds from Lemmas B.3 and B.4 into the left-hand side of the above expression. We define some shorthand notation to facilitate the presentation: let $Q = \sum_{k=1}^{K}\sum_{j=1}^{v_k}\int_0^{\hat{w}^{(b)}_{k,j}}\phi^{(b)}_{k,j}(y)dy + \sum_{\tau=1}^{T}\sum_{i=1}^{n}\left(\int_0^{\hat{w}^{(f)}_{i,\tau}}\phi^{(f)}_{i,\tau}(y)dy + \int_0^{\hat{v}^{(f)}_{i,\tau}}\psi_{i,\tau}(y)dy\right)$ denote the integrals over the thresholds. Let $\hat{W}_b = \sum_{k=1}^{K}\sum_{j=1}^{v_k}\hat{w}^{(b)}_{k,j}$, let $\hat{W}_f = \sum_{i=1}^{n}\sum_{\tau=1}^{T}\hat{w}^{(f)}_{i,\tau}$, and let $\hat{W} = \hat{W}_b + \hat{W}_f$. Further, let $\hat{V} = \sum_{i=1}^{n}\sum_{\tau=1}^{T}\hat{v}^{(f)}_{i,\tau}$.

Let $\boldsymbol{\beta} = \sum_{k=1}^{K} \beta_k$ denote the total amount of asset purchased towards base demand by the optimal offline solution during active periods, noting that $D_b - \boldsymbol{\beta} \geq 0$ by definition. Using this notation, we can rewrite the left-hand side of the above expression as:

$$\frac{\texttt{TANDEM}(\mathcal{I}) - (\lambda_{\max} + 2\eta)\,(S + LS)}{\texttt{OPT}(\mathcal{I})} \leq \frac{Q + (B - \hat{W}_b)(\lambda_{\max} + 2\eta) + \bar{pc}|D_b| + (D_f^+ - \hat{W}_f)(\lambda_{\max} + 2\eta) + (|D_f| - \hat{V})(\lambda_{\max} r + 2\gamma) - (\lambda_{\max} + 2\eta)S}{\sum_{k=1}^{K} G_k(\beta_k) + (D_b - \boldsymbol{\beta})\tilde{pg} + \bar{pc}|D_b| + \sum_{i=1}^{n}\sum_{\tau=1}^{T}\left(H_{i,\tau}(\xi_{i,\tau}) + E_{i,\tau}(f_{i,\tau})\right)}$$

First, note that the total amount of energy purchased by TANDEM towards base demand can be written as follows, where $\{\mathbf{x}_t\}_{t=1}^{T}$ are the purchasing decisions made by TANDEM over the time horizon:

$$\sum_{t=1}^{T} \mathbf{1}^\top \mathbf{x}_t - D_f^+ \leq B,$$

where the last inequality follows by definition of the base demand driver creation process. We have the following lemma:

Lemma B.7. *Under Assumptions 3.2 - 3.9,* TANDEM*'s total procurement towards base demand satisfies the following inequality:*

$$B \leq \sum_{t=1}^{T} \mathbf{1}^\top \mathbf{x}_t^\star - D_f^+ + S \leq D_b + S,$$

where $\{\mathbf{x}_t^\star\}_{t=1}^{T}$ *are the purchasing decisions made by* OPT.

Proof. To show this result, let $u_t = \mathbf{1}^\top \mathbf{x}_t$ and let $u_t^\star = \mathbf{1}^\top \mathbf{x}_t^\star$. The thermal network's state update is governed by an arbitrary thermodynamic model that captures losses within the network. By Assumption 3.4, we know that it can be decomposed into a state update that consists of conservation of energy minus a loss term. Suppose we account for the loss term (for the entire time horizon) separately—then we know that the state update respects the following conservation inequality for both TANDEM and OPT:

$$s_t = s_{t-1} + u_t - \mathbf{1}^\top(\mathbf{b}_t + \mathbf{z}_t),$$

and define cumulative purchased energy as

$$U(t) \coloneqq \sum_{\tau=1}^{t} u_t, \qquad U^\star(t) \coloneqq \sum_{\tau=1}^{t} u_t^\star.$$

Denote the state of the effective storage at time t for TANDEM and OPT as s_t and $s_t^\star$, respectively. By rearranging the conservation equation, we have:

$$s_t - s_t^\star = \left(s_{t-1} - s_{t-1}^\star\right) + \left(u_t - u_t^\star\right).$$

Since both TANDEM and OPT start with the same initial state $s_0 = s_0^\star$, this telescopes to:

$$s_t - s_t^\star = U(t) - U^\star(t).$$

Now, we claim that due to TANDEM's design, we always have $s_t \leq s_t^\star + S$. To see this, suppose for contradiction that there is a first time t such that $s_t - s_t^\star > S$, and let t be minimal with that property. Then $s_{t-1} - s_{t-1}^\star \leq S$, but after the time step it jumps above S. The only way the difference changes is through purchases u_t and $u_t^\star$, so we must have $u_t > u_t^\star$ and $u_t > 0$. Note that $u_t > 0$ implies that TANDEM purchased energy at time t, which means that it must be the case that $s_{t-1} < S$ (otherwise, TANDEM would not have purchased energy due to the storage being full). Then we have that $s_{t-1} < S$ and $s_{t-1} - s_{t-1}^\star \leq S$, which implies that $s_{t-1}^\star \geq 0$. From the update, we have:

$$s_t = s_{t-1} + u_t - \mathbf{1}^\top(\mathbf{b}_t + \mathbf{z}_t).$$

However, purchases (i.e., u_t) can only increase the storage level up to S, and since $s_{t-1}^\star \geq 0$, $s_t - s_t^\star$ cannot exceed S. This contradicts the assumption that $s_t - s_t^\star > S$, and thus, we conclude that $s_t \leq s_t^\star + S$ for all t. Using this result, we have:

$$U(T) - U^\star(T) = s_T - s_T^\star \leq S,$$

which is equivalent (by the definition of $U(\cdot)$) to:

$$\sum_{t=1}^{T} \mathbf{1}^\top \mathbf{x}_t \leq \sum_{t=1}^{T} \mathbf{1}^\top \mathbf{x}_t^\star + S,$$

Since both TANDEM and OPT satisfy purchasing requirements for positive flexible demands exactly, we have:

$$\sum_{t=1}^{T} \mathbf{1}^\top \mathbf{x}_t - D_f^+ \leq \sum_{t=1}^{T} \mathbf{1}^\top \mathbf{x}_t^\star - D_f^+ + S,$$

and furthermore, $\sum_{t=1}^{T} \mathbf{1}^\top \mathbf{x}_t^\star - D_f^+ = D_b$ since OPT is optimal and the base demand purchasing requirement is tight at optimality. Furthermore, since the above holds for any arbitrary TANDEM solution, it holds when $\sum_{t=1}^{T} \mathbf{1}^\top \mathbf{x}_t$ is exactly equal to the amount of base demand drivers created, i.e., B, yielding:

$$B \leq \sum_{t=1}^{T} \mathbf{1}^\top \mathbf{x}_t^\star - D_f^+ + S \leq D_b + S,$$

which completes the proof. □

Using Lemma B.7, we can substitute $B \leq D_b + S$ into the previous expression, yielding:

$$\frac{\texttt{TANDEM}(\mathcal{I}) - (\lambda_{\max} + 2\eta)\,(S + LS)}{\texttt{OPT}(\mathcal{I})} \leq \frac{Q + (D_b - \hat{W}_b)(\lambda_{\max} + 2\eta) + \bar{pc}|D_b| + (D_f^+ - \hat{W}_f)(\lambda_{\max} + 2\eta) + (|D_f| - \hat{V})(\lambda_{\max} r + 2\gamma)}{\sum_{k=1}^{K} G_k(\beta_k) + (D_b - \boldsymbol{\beta})\tilde{pg} + \bar{pc}|D_b| + \sum_{i=1}^{n}\sum_{\tau=1}^{T}\left(H_{i,\tau}(\xi_{i,\tau}) + E_{i,\tau}(f_{i,\tau})\right)} \tag{13}$$

Using the fact that $D_b - \boldsymbol{\beta} \geq 0$ and $|D_b| - \boldsymbol{\beta} \geq 0$, the following is equivalent:

$$= \frac{Q + (\boldsymbol{\beta} - \hat{W}_b)(\lambda_{\max} + 2\eta) + \bar{pc}|D_b| + (D_f^+ - \hat{W}_f)(\lambda_{\max} + 2\eta) + (|D_f| - \hat{V})(\lambda_{\max} r + 2\gamma) + (D_b - \boldsymbol{\beta})(\lambda_{\max} + 2\eta) + (|D_b| - \boldsymbol{\beta})\bar{pc}}{\sum_{k=1}^{K} G_k(\beta_k) + \bar{pc}\boldsymbol{\beta} + \sum_{i=1}^{n}\sum_{\tau=1}^{T}\left(H_{i,\tau}(\xi_{i,\tau}) + E_{i,\tau}(f_{i,\tau})\right) + (D_b - \boldsymbol{\beta})\tilde{pg} + (|D_b| - \boldsymbol{\beta})\bar{pc}}$$

Then, we have the following:

$$\leq \max\left\{ \frac{Q + (\boldsymbol{\beta} - \hat{W}_b)(\lambda_{\max} + 2\eta) + (D_f^+ - \hat{W}_f)(\lambda_{\max} + 2\eta) + (|D_f| - \hat{V})(\lambda_{\max} r + 2\gamma)}{\sum_{k=1}^{K} G_k(\beta_k) + \bar{pc}\boldsymbol{\beta} + \sum_{i=1}^{n}\sum_{\tau=1}^{T}\left(H_{i,\tau}(\xi_{i,\tau}) + E_{i,\tau}(f_{i,\tau})\right)}, \frac{(D_b - \boldsymbol{\beta})(\lambda_{\max} + 2\eta) + (|D_b| - \boldsymbol{\beta})\bar{pc}}{(D_b - \boldsymbol{\beta})\tilde{pg} + (|D_b| - \boldsymbol{\beta})\bar{pc}}, \frac{\bar{pc}\boldsymbol{\beta}}{\bar{pc}\boldsymbol{\beta}} \right\},$$

where the definition of $\tilde{pg}$ ensures that the second term in the max is at most α, and the third term is 1. We now focus on the first term. For the sake of contradiction, suppose that:

$$\frac{Q + (\boldsymbol{\beta} - \hat{W}_b)(\lambda_{\max} + 2\eta) + (D_f^+ - \hat{W}_f)(\lambda_{\max} + 2\eta) + (|D_f| - \hat{V})(\lambda_{\max} r + 2\gamma)}{\sum_{k=1}^{K} G_k(\beta_k) + \sum_{i=1}^{n}\sum_{\tau=1}^{T}\left(H_{i,\tau}(\xi_{i,\tau}) + E_{i,\tau}(f_{i,\tau})\right)} > \alpha. \tag{14}$$

Note that $\boldsymbol{\beta}$ is the total amount of asset purchased towards base demand by the optimal offline solution during active periods, and recall that active periods contain strictly better prices than inactive periods (by definition). We reason about how the cost of OPT and TANDEM relate in terms of $B = \sum_{k=1}^{K}\sum_{j=1}^{v_k} B_{k,j}^{(b)}$, the demand assigned to base drivers. We have the following relation:

$$\frac{Q + (B - \hat{W}_b)(\lambda_{\max} + 2\eta) + (D_f^+ - \hat{W}_f)(\lambda_{\max} + 2\eta) + (|D_f| - \hat{V})(\lambda_{\max} r + 2\gamma)}{\sum_{k=1}^{K} G_k(B_k) + \sum_{i=1}^{n}\sum_{\tau=1}^{T}\left(H_{i,\tau}(\xi_{i,\tau}) + E_{i,\tau}(f_{i,\tau})\right)},$$

$$\geq \frac{Q + (\boldsymbol{\beta} - \hat{W}_b)(\lambda_{\max} + 2\eta) + (D_f^+ - \hat{W}_f)(\lambda_{\max} + 2\eta) + (|D_f| - \hat{V})(\lambda_{\max} r + 2\gamma) + (B - \boldsymbol{\beta})(\lambda_{\max} + 2\eta)}{\sum_{k=1}^{K} G_k(\beta_k) + \sum_{i=1}^{n}\sum_{\tau=1}^{T}\left(H_{i,\tau}(\xi_{i,\tau}) + E_{i,\tau}(f_{i,\tau})\right) + \sum_{k=1}^{K}\left(G_k(B_k) - G_k(\beta_k)\right)} > \alpha,$$

where in the second inequality, we have used the fact that there is a worst-case input instance such that for all $0 < \beta < \beta' < \sum_{j=1}^{v_k} B_{k,j}^{(b)} - \hat{b}_k$, we have $G_k(\beta') - G_k(\beta) < (\beta' - \beta)\tilde{pg}$, where $\hat{b}_k$ is the initial state of the storage at the start of the k^{th} active period in OPT's solution. We have the following lemma to show this:

Lemma B.8. *Defining $\hat{b}_k$ as the initial state of the storage at the start of the k^{th} active period in* OPT*'s solution, there exists a worst-case input instance such that for all $0 < \beta < \beta' < \sum_{j=1}^{v_k} B_{k,j}^{(b)} - \hat{b}_k$, we have:*

(1) $G_k(0) = 0$,
(2) $G_k(\beta') > G_k(\beta)$,
(3) $G_k(\beta') - G_k(\beta) \leq (\beta' - \beta)\tilde{pg}$.

Proof. Statements 1 and 2 follow directly from the definition of $G_k(\cdot)$. Recall that $\tilde{pg}$ is defined as the minimum possible cost of purchasing energy during inactive periods, which corresponds with the worst (highest) marginal cost that TANDEM's base drivers are willing to pay to fill the thermal network's effective storage.

To prove statement 3, consider any input instance for the problem denoted as $\mathcal{I} = [(p_t, \mathbf{g}_t, \mathbf{b}_t, \mathbf{f}_t, \boldsymbol{\Delta}_t)]_{t\in[T]}$. We can construct a modified input instance $\mathcal{I}'$ from $\mathcal{I}$ as follows:

$$\mathcal{I}' = [(\tilde{pg}, \mathbf{0}, \mathbf{0}, \cdot), (p_1, \mathbf{g}_1, \mathbf{b}_1, \mathbf{f}_1, \boldsymbol{\Delta}_1), (\tilde{pg}, \mathbf{0}, \mathbf{0}, \cdot), (p_2, \mathbf{g}_2, \mathbf{b}_2, \mathbf{f}_2, \boldsymbol{\Delta}_2), \ldots, (p_T, \mathbf{g}_T, \mathbf{b}_T, \mathbf{f}_T, \boldsymbol{\Delta}_T)]\,.$$

The purpose of constructing this new input is to guarantee that the purchasing cost of TANDEM does not change, and the purchasing cost of OPT will not increase (for this worst-case logic, we only consider the purchasing cost, ignoring additional e.g., tracking costs that would increase OPT's cost). Note that this is possible because an adversary can set the length of the time horizon. Let $G_k(\beta)$ be the minimum cost of purchasing β units of energy during the k^{th} active period. When OPT purchases another $\beta' - \beta$ units of energy (where $\beta' \leq \sum_{j=1}^{v_k} B_{k,j}^{(b)} - \hat{b}_k$)

during the k^{th} active period, the additional cost will not be larger than $(\beta' - \beta)\tilde{p}\tilde{g}$, since OPT can always choose to purchase at any newly added time slots. Thus, there always exists a worst-case input such that $G_k(\beta') - G_k(\beta) \leq (\beta' - \beta)\tilde{p}\tilde{g}$. □

We proceed to work with the expression in terms of B. During the k^{th} active period and the lifetime of the j^{th} base driver, the minimum marginal purchasing cost observed is given by $\phi_{k,j}^{(b)}(\hat{w}_{k,j}^{(b)}) - 2\eta$ (by definition of the threshold function). Similarly, during the lifetime of the flexible demand driver arriving at time τ for household i, the minimum marginal purchasing and delivery costs observed are given by $\phi_{i,\tau}^{(f)}(\hat{w}_{i,\tau}^{(f)}) - 2\eta$ and $\psi_{i,\tau}(\hat{v}_{i,\tau}^{(f)}) - 2\gamma$, respectively. This gives the following lower bounds on the terms that depend on G_k, $H_{i,\tau}$, and $E_{i,\tau}$, respectively:

$$\sum_{k=1}^{K} G_k\left(\sum_{j=1}^{v_k} B_{k,j}^{(b)}\right) \geq \sum_{j=1}^{v_k}\left(\phi_{k,j}^{(b)}(\hat{w}_{k,j}^{(b)}) - 2\eta\right) \times B_{k,j}^{(b)},$$

$$\sum_{i=1}^{n}\sum_{\tau=1}^{T} H_{i,\tau}(\xi_{i,\tau}) \geq \sum_{i=1}^{n}\sum_{\tau=1}^{T}\left(\phi_{i,\tau}^{(f)}(\hat{w}_{i,\tau}^{(f)}) - 2\eta\right) \times \xi_{i,\tau},$$

$$\sum_{i=1}^{n}\sum_{\tau=1}^{T} E_{i,\tau}(f_{i,\tau}) \geq \sum_{i=1}^{n}\sum_{\tau=1}^{T}\left(\psi_{i,\tau}(\hat{v}_{i,\tau}^{(f)}) - 2\gamma\right) \times f_{i,\tau}.$$

Substituting these bounds into the previous expression, we have that the left-hand side of (14) is less than or equal to:

$$\frac{Q + (B - \hat{W}_b)(\lambda_{\max} + 2\eta) + (D_f^+ - \hat{W}_f)(\lambda_{\max} + 2\eta) + (|D_f| - \hat{V})(\lambda_{\max} r + 2\gamma)}{\sum_{k=1}^{K}\sum_{j=1}^{v_k}\left(\phi_{k,j}^{(b)}(\hat{w}_{k,j}^{(b)}) - 2\eta\right)B_{k,j}^{(b)} + \sum_{i=1}^{n}\sum_{\tau=1}^{T}\left(\phi_{i,\tau}^{(f)}(\hat{w}_{i,\tau}^{(f)}) - 2\eta\right)\xi_{i,\tau} + \sum_{i=1}^{n}\sum_{\tau=1}^{T}\left(\psi_{i,\tau}(\hat{v}_{i,\tau}^{(f)}) - 2\gamma\right)f_{i,\tau}}.$$

By rearranging the terms in the above and substituting for Q, we obtain the following:

$$\leq \frac{\sum_{k=1}^{K}\sum_{j=1}^{v_k} X_{k,j}^{(b)} + \sum_{i=1}^{n}\sum_{\tau=1}^{T} X_{i,\tau}^{(f)}}{\sum_{k=1}^{K}\sum_{j=1}^{v_k}\left(\phi_{k,j}^{(b)}(\hat{w}_{k,j}^{(b)}) - 2\eta\right)B_{k,j}^{(b)} + \sum_{i=1}^{n}\sum_{\tau=1}^{T}\left(\phi_{i,\tau}^{(f)}(\hat{w}_{i,\tau}^{(f)}) - 2\eta\right)\xi_{i,\tau} + \sum_{i=1}^{n}\sum_{\tau=1}^{T}\left(\psi_{i,\tau}(\hat{v}_{i,\tau}^{(f)}) - 2\gamma\right)f_{i,\tau}}. \tag{15}$$

where $X_{k,j}^{(b)}$ and $X_{i,\tau}^{(f)}$ are defined as follows (for each driver):

$$X_{k,j}^{(b)} = \int_0^{\hat{w}_{k,j}^{(b)}} \phi_{k,j}^{(b)}(y)dy + (B_{k,j}^{(b)} - \hat{w}_{k,j}^{(b)})(\lambda_{\max} + 2\eta),$$

$$X_{i,\tau}^{(f)} = \int_0^{\hat{w}_{i,\tau}^{(f)}} \phi_{i,\tau}^{(f)}(y)dy + (\xi_{i,\tau} - \hat{w}_{i,\tau}^{(f)})(\lambda_{\max} + 2\eta) + \int_0^{\hat{v}_{i,\tau}^{(f)}} \psi_{i,\tau}(y)dy + (f_{i,\tau} - \hat{v}_{i,\tau}^{(f)})(\lambda_{\max} r + 2\gamma).$$

Since (14) is $> \alpha$ by assumption, one of the following cases must be true:

Case 1: There exists some base driver with $k \in \{1, \ldots, K\}$ and $j \in \{1, \ldots, v_k\}$ such that:

$$X_{k,j}^{(b)} > \alpha\left(\phi_{k,j}^{(b)}(\hat{w}_{k,j}^{(b)}) - 2\eta\right)B_{k,j}^{(b)}.$$

Case 2: There exists some flexible demand driver with $i \in \{1, \ldots, n\}$ and $\tau \in \{1, \ldots, T\}$ such that:

$$X_{i,\tau}^{(f)} > \alpha\left(\left(\phi_{i,\tau}^{(f)}(\hat{w}_{i,\tau}^{(f)}) - 2\eta\right)\xi_{i,\tau} + \left(\psi_{i,\tau}(\hat{v}_{i,\tau}^{(f)}) - 2\gamma\right)f_{i,\tau}\right).$$

But, if either of these two cases are true (i.e., for any unit of base or flexible demand), they contradict Lemmas B.5 and B.6, respectively. Thus, we conclude that our initial assumption in (14) must be false, completing the proof that $\frac{\text{TANDEM}(\mathcal{I}) - (\lambda_{\max}+2\eta)\cdot(S+LS)}{\text{OPT}(\mathcal{I})} \leq \alpha$.

Using this result, we have:

$$\text{TANDEM}(\mathcal{I}) \leq \alpha\text{OPT}(\mathcal{I}) + (\lambda_{\max} + 2\eta) \cdot (S + LS),$$

where $(\lambda_{\max} + 2\eta) \cdot (S + LS)$ is a constant. This shows that TANDEM is α-competitive under Def. 3.1, which completes the main proof of the competitive ratio. We conclude the proof of Theorem 4.3 with a remark on the impact of rate constraints in the following lemma:

Lemma B.9. TANDEM *remains α-competitive when rate constraints are imposed on the network management decisions of both* TANDEM *and* OPT.

Proof. Denote $\mathbf{x}_t$ and $\mathbf{x}_t^\star$ as the purchasing decisions made by TANDEM and OPT (respectively), and define $\mathcal{T}_r \subset \{1, \ldots, T\}$ as the set of time steps where rate constraints are active for TANDEM (i.e., $\mathbf{x}_t$ is limited by a maximum rate at those time steps).

When TANDEM is *not* rate-constrained, recall that the i^{th} purchasing driver solves the following pseudo-cost minimization problem at each time step: $\mathbf{x}_{i,t} \leftarrow \text{argmin}_{\mathbf{x}\in\mathbb{R}^m}\, p_t\mathbf{g}_t^\top\mathbf{x} + \eta\|\mathbf{x} - \tilde{\mathbf{x}}_{i,t}\| - \Phi_i(w_{i,t-1}, w_{i,t-1} + \mathbf{1}^\top\mathbf{x})$, where $w_{i,t} = w_{i,t-1} + \mathbf{1}^\top\mathbf{x}_{i,t}$ corresponds to the best price observed so far by the i^{th} purchasing driver up to time t, given by $\phi_i(w_{i,t}) - 2\eta$.

When rate constraints are active at time $t \in \mathcal{T}_r$, the purchasing decision is constrained to $\bar{\mathbf{x}}_{i,t}$, which is the solution to the same pseudo-cost minimization problem but with an additional constraint on the maximum allowable purchase rate. Let $\bar{w}_{i,t} = \bar{w}_{i,t-1} + \mathbf{1}^\top\bar{\mathbf{x}}_{i,t}$ denote the (reduced) amount purchased during this step, and note that by the monotonicity of ϕ, we have $\phi_i(\bar{w}_{i,t}) > \phi_i(w_{i,t})$. This has two implications

– first, although TANDEM's total cost is upper-bounded with respect to $\phi_i(\bar{w}_{i,t})$, in a rate-constrained setting, for any $t \in \mathcal{T}_r$, the cost incurred by TANDEM is actually lower since it purchases energy at the lower cost corresponding to $\phi_i(w_{i,t})$. Second, since $\phi_i(\bar{w}_{i,t})$ typically forms a lower bound on OPT's cost (as shown in the main proof), the bound must be adjusted to account for the fact that the actual price paid by OPT is the lower cost corresponding to $\phi_i(w_{i,t})$. Formally, we define a function $\mu(t)$ for each time step $t \in \mathcal{T}_r$ as follows:

$$\mu(t) = \sum_{i \in \mathcal{B}_t \cup \mathcal{F}_t} \left(\phi_i(\bar{w}_{i,t}) - \phi_i(w_{i,t})\right),$$

where $\mathcal{B}_t$ and $\mathcal{F}_t$ denote the sets of active drivers at time t. This function $\chi(\cdot)$ captures the cumulative difference in cost bounds due to rate constraints across all rate-constrained periods $t \in \mathcal{T}_r$. Picking up from (13) in the main proof, we can bound the competitive ratio as follows:

$$\frac{\texttt{TANDEM}(\mathcal{I}) - (\lambda_{\max} + 2\eta)\,(S + LS)}{\texttt{OPT}(\mathcal{I})} \leq \frac{Q + (D_b - \hat{W}_b)(\lambda_{\max} + 2\eta) + \tilde{pc}|D_b| + (D_f^+ - \hat{W}_f)(\lambda_{\max} + 2\eta) + (|D_f| - \hat{V})(\lambda_{\max} r + 2\gamma) - \sum_{t \in \mathcal{T}_r} \mu(t)\mathbf{1}^\top \mathbf{x}_t}{\sum_{k=1}^K G'_k(\beta_k) + (D_b - \boldsymbol{\beta})\tilde{pg} + \tilde{pc}|D_b| + \sum_{i=1}^n \sum_{\tau=1}^T \left(H_\tau^{'(i)}(\xi_{i,\tau}) + E_{i,\tau}(f_{i,\tau})\right)},$$

where $G'_k(\cdot)$ and $H_\tau^{'(i)}(\cdot)$ are defined similarly to $G_k(\cdot)$ but account for rate constraints in OPT's solution. Note that the last term in the numerator captures the reduction in TANDEM's cost due to the "over-estimation" of cost incurred during rate-constrained periods. Then, similarly to the main proof, $G'_k(\cdot)$ and $H_\tau^{'(i)}(\cdot)$ satisfy:

$$\sum_{k=1}^K G_k\left(\sum_{j=1}^{v_k} B_{k,j}^{(b)}\right) + \sum_{i=1}^n \sum_{\tau=1}^T H_{i,\tau}(\xi_{i,\tau}) \geq \sum_{j=1}^{v_k} \left(\phi_{k,j}^{(b)}(\hat{w}_{k,j}^{(b)}) - 2\eta\right) \times B_{k,j}^{(b)} + \sum_{i=1}^n \sum_{\tau=1}^T \left(\phi_{i,\tau}^{(f)}(\hat{w}_{i,\tau}^{(f)}) - 2\eta\right) \times \xi_{i,\tau} - \sum_{t \in \mathcal{T}_r} \mu(t)\mathbf{1}^\top \mathbf{x}_t^\star.$$

Then, using (15) and following the same reasoning as in the main proof, we have:

$$\frac{\texttt{TANDEM}(\mathcal{I}) - (\lambda_{\max} + 2\eta)\,(S + LS)}{\texttt{OPT}(\mathcal{I})} \leq$$
$$\max\left\{\frac{\sum_{k=1}^K \sum_{j=1}^{v_k} X_{k,j}^{(b)} + \sum_{i=1}^n \sum_{\tau=1}^T X_{i,\tau}^{(f)} - \sum_{t \in \mathcal{T}_r} \mu(t)\mathbf{1}^\top \mathbf{x}_t}{\sum_{k=1}^K \sum_{j=1}^{v_k} \left(\phi_{k,j}^{(b)}(\hat{w}_{k,j}^{(b)}) - 2\eta\right) B_{k,j}^{(b)} + \sum_{i=1}^n \sum_{\tau=1}^T \left(\phi_{i,\tau}^{(f)}(\hat{w}_{i,\tau}^{(f)}) - 2\eta\right) \xi_{i,\tau} + \sum_{i=1}^n \sum_{\tau=1}^T \left(\psi_{i,\tau}(\hat{v}_{i,\tau}^{(f)}) - 2\gamma\right) f_{i,\tau} - \sum_{t \in \mathcal{T}_r} \mu(t)\mathbf{1}^\top \mathbf{x}_t^\star}, \alpha\right\}$$

In the worst-case, we have that $\frac{\mathbf{1}^\top \mathbf{x}_t}{\mathbf{1}^\top \mathbf{x}_t^\star} \leq \alpha$ for any $t \in \mathcal{T}_r$—we show this by contradiction. Suppose $\mathcal{I} = [(p_t, \mathbf{g}_t, \mathbf{b}_t, \mathbf{f}_t, \Delta_t)]_{t \in [T]}$ is the worst-case instance and at time slot t, we have $\mathbf{1}^\top \mathbf{x}_t / \mathbf{1}^\top \mathbf{x}_t^\star > \alpha$. We can construct a modified input instance $\mathcal{I}'$ from $\mathcal{I}$ by increasing the price and efficiency at time t by δ_p and δ_g, where $\delta = \delta_p \times \delta_g \leq \chi(t)$. That is:

$$\mathcal{I}' = \left[(p_1, \mathbf{g}_1, \mathbf{b}_1, \mathbf{f}_1, \Delta_1), \ldots, (p_{t-1}, \mathbf{g}_{t-1}, \mathbf{b}_{t-1}, \mathbf{f}_{t-1}, \Delta_{t-1}), (p_t + \delta_p, \mathbf{g}_t + \delta_g, \mathbf{b}_t, \mathbf{f}_t, \Delta_t), (p_{t+1}, \mathbf{g}_{t+1}, \mathbf{b}_{t+1}, \mathbf{f}_{t+1}, \Delta_{t+1}), \ldots, (p_T, \mathbf{g}_T, \mathbf{b}_T, \mathbf{f}_T, \Delta_T)\right].$$

In this modified input instance, the cost of TANDEM increases by $\delta \times \mathbf{1}^\top \mathbf{x}_t$, while the cost of OPT increases by less than $\frac{\delta \times \mathbf{1}^\top \mathbf{x}_t}{\alpha}$, since OPT can always choose to purchase at other time slots. This contradicts the assumption that $\mathcal{I}$ is the worst-case input instance, and thus, we conclude that in the worst-case, $\frac{\mathbf{1}^\top \mathbf{x}_t}{\mathbf{1}^\top \mathbf{x}_t^\star} \leq \alpha$ for any $t \in \mathcal{T}_r$. We have $\frac{\sum_{t \in \mathcal{T}_r} \chi(t)\mathbf{1}^\top \mathbf{x}_t}{\sum_{t \in \mathcal{T}_r} \chi(t)\mathbf{1}^\top \mathbf{x}_t^\star} \leq \alpha$, which yields:

$$\frac{\texttt{TANDEM}(\mathcal{I}) - (\lambda_{\max} + 2\eta)\,(S + LS)}{\texttt{OPT}(\mathcal{I})} \leq$$
$$\max\left\{\frac{\sum_{k=1}^K \sum_{j=1}^{v_k} X_{k,j}^{(b)} + \sum_{i=1}^n \sum_{\tau=1}^T X_{i,\tau}^{(f)}}{\sum_{k=1}^K \sum_{j=1}^{v_k} \left(\phi_{k,j}^{(b)}(\hat{w}_{k,j}^{(b)}) - 2\eta\right) B_{k,j}^{(b)} + \sum_{i=1}^n \sum_{\tau=1}^T \left(\phi_{i,\tau}^{(f)}(\hat{w}_{i,\tau}^{(f)}) - 2\eta\right) \xi_{i,\tau} + \sum_{i=1}^n \sum_{\tau=1}^T \left(\psi_{i,\tau}(\hat{v}_{i,\tau}^{(f)}) - 2\gamma\right) f_{i,\tau}}, \alpha\right\}.$$

Combining the above with the contradiction in (14) from the main proof, we have that TANDEM is α-competitive even when rate constraints are imposed on the network management side, as desired. □

□

B.3 Proof of Theorem 4.4

In this section, we prove Theorem 4.4, which states that α (as defined in (3)) is the best achievable competitive ratio for any deterministic online algorithm solving OTNM.

Proof. To show this result, we define a family of special instances and show that the competitive ratio of any deterministic algorithm is lower bounded under these instances. Consider the following set of instances $\{\mathcal{I}_x\}_{x \in [\lambda_{\min}, \lambda_{\max}]}$, where $\mathcal{I}_x$ is called an *x-decreasing instance*, and $\lambda_{\min}$ and $\lambda_{\max}$ are the best and worst marginal prices for heat exchange at (any) pump station (recall $\lambda_{\min}$ is defined as $p_{\min} \cdot g_{\min}$ and $\lambda_{\max}$ is defined as $p_{\max} \cdot g_{\max}$), where $g_{\max} = 1$. The precise definition of an x-decreasing instance is given below.

Definition B.10 (x-decreasing instance for OTNM*).* Suppose there is a single pump station and two customers. Let $s_0 = 0$ initially. Let $w, q \in \mathbb{N}$ be sufficiently large, and denote $\varrho \coloneqq {}^{\lambda_{\max}-\lambda_{\min}}/_{w}$. Suppose that both customers have identical efficiency factors $c^{(b)} \leq 1$, and let the ratio r be defined as $r \coloneqq \frac{c^{(b)}}{g_{\min}}$.

For $x \in [\lambda_{\min}, \lambda_{\max}]$, we define an x-decreasing instance $\mathcal{I}_x$ of OTNM as follows: The sequence is partitioned into $n_x \coloneqq 2 \cdot \lceil {}^{(x-\lambda_{\min})}/_{\varrho} \rceil + 1$ alternating "batches". If k is odd, the k^{th} batch ($k \in [n_x - 2]$) contains q inputs with $p_t g_t = \lambda_{\max}$ for the pump station (i.e., for decision x_t), and $p_t c_{i,t}^{(f)} = \lambda_{\max} \cdot r$ at both customers (i.e., for decisions $z_{1,t}$ and $z_{2,t}$), where $c_{i,t}^{(f)}$ denotes the coefficient of the function $c_t^{(f)}(\cdot)$ at customer i at time t. If k is even, the k^{th} batch contains 2 inputs with $p_t g_t = \lambda_{\max} - (\lceil k/2 \rceil)\varrho$ for the pump station, and $p_t c_{i,t}^{(f)} = (\lambda_{\max} - (\lceil k/2 \rceil)\varrho) \cdot r$ at both customers. The last two batches consist of q inputs with $p_t g_t = x + \iota$ for the pump station, and $p_t c_{i,t}^{(f)} = (x + \iota) \cdot r$ at both customers, where $\iota > 0$ is an arbitrarily small constant; followed by a final batch of q inputs with $p_t g_t = \lambda_{\max}$ for the pump station, and $p_t c_{i,t}^{(f)} = \lambda_{\max} \cdot r$ at both customers.

Note that when $x \to \lambda_{\min}$, the second-to-last batch of inputs has $p_t c_{i,t}^{(f)} \to \lambda_{\min} \cdot r = p_{\min} \cdot c^{(b)}$.

The demand sequence is defined as follows: At $t = 1$, two flexible demands arrive – a flexible demand $f_{1,1} = 0.5 + P(0.5)$ with deadline $\Delta_{1,1} = T$ at the first customer, and a unit-size flexible demand $f_{2,1} = -0.5 + P(0.5)$ with deadline $\Delta_{2,1} = T$ at the second customer. P is a parameter in $(0, 1]$. All other demands $\mathbf{b}_t$ and $\mathbf{f}_t$ are zero. Note that the total demand is exactly P, and the sequence has positive net demand (i.e., heating mode).

The tracking target is defined separately for pump stations and each customer – we set $\tilde{x}_t = 0$ and $\tilde{z}_{i,t} = 0$ for all batches of prices except the second-to-last batch, where $\tilde{z}_{i,t} = {}^1/_{2q}$ for both customers and $\tilde{x}_t = {}^P/_q$ for all q prices. Note that $\sum_{t=1}^{T} \sum_{i=1}^{2} \tilde{z}_{i,t} = 1$ and $\sum_{t=1}^{T} \tilde{x}_t = P$.

Let $h(x)$ and $m(x)$ denote *conversion functions* that both map $[\lambda_{\min}, \lambda_{\max}] \to [0, 1]$, respectively. Suppose these arbitrary functions fully describe the actions of a deterministic ALG for OTNM on an instance $\mathcal{I}_x$.

Specifically, suppose that $h(x)$ describes ALG's network management decisions, and $m(x)$ describes ALG's fulfillment decisions *before* the arrival of the last batch of prices ($\lambda_{\max}$) in the instance. Note that by definition, $m(x) \leq 1 - P + h(x)P$, since the heating demand fulfillment "matches" up with the cooling demand (up to P). Note that for large w, processing an instance $\mathcal{I}_{x-\varrho}$ is equivalent to first processing $\mathcal{I}_x$ (besides the last two batches), and then processing batches with prices $x - \varrho$ and $\lambda_{\max}$.

Since ALG is deterministic and both conversions (i.e., purchasing and delivery) are both unidirectional (irrevocable), we must have that $h(x - \varrho) \geq h(x)$ and $m(x - \varrho) \geq m(x)$, i.e., $h(x)$ and $m(x)$ are both non-increasing in $[\lambda_{\min}, \lambda_{\max}]$. Intuitively, the entire demand should be fulfilled before the end of the sequence if the lowest price appears, i.e., $h(\lambda_{\min}) = 1, m(\lambda_{\min}) = 1$.

Due to the tracking target, any purchasing or fulfillment decisions before the second-to-last batch incur a tracking penalty of $\eta H(x)P + \gamma M(x)$, since the target is zero during this time. Furthermore, during the last two batches, ALG incurs a tracking penalty of $\eta P + \gamma + \eta(1 - H(x))P + \gamma(1 - M(x))$, since the target is ${}^1/_q$ during the first q prices in the batch, and zero during the second q prices in the batch – note that ALG must purchase $(1 - H(x))$ and fulfill $(1 - M(x))$ at the last price to satisfy the deadlines. Thus, the total tracking penalty incurred by ALG is given by $2\eta P + 2\gamma$. Then the total cost of ALG on instance $\mathcal{I}_x$ is described by:

$$\text{ALG}(\mathcal{I}_x) = Ph(\ell)\ell - P\int_{\ell}^{x} u dh(u) + P(1 - h(x))\lambda_{\max} + 2\eta P + 2\gamma + r\ell m(\ell) - r\int_{\ell}^{x} u dm(u) + (1 - m(x))r\lambda_{\max}.$$

The optimal cost on instance $\mathcal{I}_x$ (if all prices are known in advance) is given by:

$$\text{OPT}(\mathcal{I}_x) = Px + rx,$$

where note that OPT incurs no tracking penalty by perfectly matching the target. If ALG is $\alpha^\star$-competitive, then a necessary condition is that for all $x \in [\lambda_{\min}, \lambda_{\max}]$, $\text{ALG}(\mathcal{I}_x) \leq \alpha^\star \text{OPT}(\mathcal{I}_x)$. This imposes a necessary condition on $h(x)$ and $m(x)$, which can be expressed as the following differential inequality:

$$Ph(\ell)\ell - P\int_{\ell}^{x} u dh(u) + P(1 - h(x))\lambda_{\max} + 2\eta P + 2\gamma + r\ell m(\ell) - r\int_{\ell}^{x} u dm(u) + (1 - m(x))r\lambda_{\max} \leq \alpha^\star (Px + rx).$$

Note that the symmetry of the above implies that $h(\ell) = m(\ell)$ and $h(x) = m(x)$ for all $x \in [\lambda_{\min}, \lambda_{\max}]$. Using integration by parts, this gives the following condition on $h(x)$:

$$\begin{aligned}
(P + r)h(x)x - (P + r)\int_{\ell}^{x} h(u)du + (P + r)(1 - h(x))\lambda_{\max} + 2\eta P + 2\gamma &\leq \alpha^\star (Px + rx), \\
h(x)x - \int_{\ell}^{x} h(u)du + (1 - h(x))\lambda_{\max} + \frac{2\eta P + 2\gamma}{P + r} &\leq \alpha^\star x, \\
\lambda_{\max} + \frac{2\eta P + 2\gamma}{P + r} - \alpha^\star x - \int_{\ell}^{x} h(u)du &\leq (\lambda_{\max} - x)h(x), \\
\frac{\lambda_{\max} + \frac{2\eta P + 2\gamma}{P + r} - \alpha^\star x}{\lambda_{\max} - x} - \frac{1}{\lambda_{\max} - x}\int_{\ell}^{x} h(u)du &\leq h(x).
\end{aligned}$$

Let $Q = \frac{2\eta P + 2\gamma}{P + r}$. To solve for $\alpha^\star$, we can use Grönwall's inequality [47, Theorem 1, p. 356]. First, observe that $H(x)$ must satisfy the following:

$$h(x) \geq \frac{\lambda_{\max} + Q - \alpha^\star x}{\lambda_{\max} - x} - \frac{1}{\lambda_{\max} - x} \int_\ell^x h(u) du.$$

By Grönwall's inequality, it follows that:

$$h(x) \geq \frac{\lambda_{\max} + Q - \alpha^\star x}{\lambda_{\max} - x} - \int_\ell^x \frac{\lambda_{\max} + Q - \alpha^\star u}{(\lambda_{\max} - u)^2} du,$$
$$h(x) \geq \frac{\lambda_{\max} + Q - \alpha^\star x}{\lambda_{\max} - x} - \left[\frac{\lambda_{\max}\alpha^\star - \lambda_{\max} - Q}{u - \lambda_{\max}} - \alpha^\star \ln(\lambda_{\max} - u) \right]_\ell^x$$

Letting $\ell = \frac{\lambda_{\max} + Q}{\alpha^\star}$, we can simplify the above as follows:

$$h(x) \geq \alpha^\star \ln \left[\frac{\lambda_{\max} - x}{\lambda_{\max} - \ell} \right].$$

Recall that we have a boundary condition that $h(\lambda_{\min}) = 1$ – we can combine this with the above to obtain:

$$1 = \alpha^\star \ln \left[\frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} - \frac{\lambda_{\max} + Q}{\alpha^\star}} \right].$$

Solving the above yields that the optimal $\alpha^\star$ for any ALG solving OTNM is lower bounded by:

$$\alpha^\star \geq \left[\frac{\lambda_{\max}}{Q + \lambda_{\max}} + W\left(-\frac{\exp\left(-\frac{\lambda_{\max}}{Q + \lambda_{\max}}\right)(\lambda_{\max} - \lambda_{\min})}{Q + \lambda_{\max}} \right) \right]^{-1},$$

where $Q = \frac{2\eta P + 2\gamma}{P + r}$. Note that Q depends on P, which is the parameter defining the demand sequence in the instance. For given (fixed) values of $\lambda_{\min}, \lambda_{\max}, r, \eta, \gamma, L$, we can optimize over P to obtain the best lower bound on $\alpha^\star$ as follows:

$$\alpha^\star \geq \max_{P \in (0,1]} \left[\frac{\lambda_{\max}}{\frac{2\eta P + 2\gamma}{P + r} + \lambda_{\max}} + W\left(-\frac{\exp\left(-\frac{\lambda_{\max}}{\frac{2\eta P + 2\gamma}{P + r} + \lambda_{\max}}\right)(\lambda_{\max} - \lambda_{\min})}{\frac{2\eta P + 2\gamma}{P + r} + \lambda_{\max}} \right) \right]^{-1}.$$

The above maximization exactly corresponds with maximizing Q over $P \in (0, 1]$, which can be solved analytically as follows. Consider the endpoints (i.e., where $P \to 0^+$ and $P = 1$):

$$\lim_{P \to 0^+} Q = \frac{2\gamma}{r},$$
$$Q|_{P=1} = \frac{2\eta + 2\gamma}{1 + r}.$$

By setting these two expressions equal, we obtain a critical value in terms of η, c, and L as $\eta = \frac{2\gamma}{2r}$— when $\eta < \frac{\gamma}{r}$, Q is maximized at $P \to 0^+$; when $\eta > \frac{\gamma}{r}$, Q is maximized at $P = 1$; and when $\eta = \frac{\gamma}{r}$, Q is constant for all $P \in (0, 1]$. Thus, the optimal lower bound on $\alpha^\star$ is given by:

$$\alpha^\star \geq \begin{cases} \left[\frac{\lambda_{\max}}{\frac{2\gamma}{r} + \lambda_{\max}} + W\left(-\frac{\exp\left(-\frac{\lambda_{\max}}{\frac{2\gamma}{r} + \lambda_{\max}}\right)(\lambda_{\max} - \lambda_{\min})}{\frac{2\gamma}{r} + \lambda_{\max}} \right) \right]^{-1}, & \text{if } \eta \leq \frac{\gamma}{r}, \\ \left[\frac{\lambda_{\max}}{\frac{2\eta + 2\gamma}{1 + r} + \lambda_{\max}} + W\left(-\frac{\exp\left(-\frac{\lambda_{\max}}{\frac{2\eta + 2\gamma}{1 + r} + \lambda_{\max}}\right)(\lambda_{\max} - \lambda_{\min})}{\frac{2\eta + 2\gamma}{1 + r} + \lambda_{\max}} \right) \right]^{-1}, & \text{if } \eta > \frac{\gamma}{r}, \end{cases}$$

letting $\mu \coloneqq \frac{2\gamma}{r}$ when $\eta \leq \frac{\gamma}{r}$, and $\mu \coloneqq \frac{2\eta + 2\gamma}{1 + r}$ when $\eta > \frac{\gamma}{r}$, we exactly obtain the expression for α in (3), which completes the proof. □

B.4 Proof of Corollary 4.5

In this section, we prove Corollary 4.5, which states that TANDEM is $\alpha\beta$-competitive for OTNM with a detailed model of customer COP efficiencies, where $\beta \coloneqq c_{\max}/c_{\min}$.

Consider the offline OTNM problem with feasible set $\mathcal{F}$ over trajectories $\pi := \{(\mathbf{x}_t, \mathbf{z}_t, s_t)\}_{t=1}^{T}$ and note that the feasibility of the problem (constraints, deadlines, and the state evolution $s_t = M(s_{t-1}, \mathbf{x}_t, b_t, \mathbf{z}_t)$) do *not* depend on the heat-pump efficiency model used to define the objective components $c^{(b)}(\cdot)$ and $c_t^{(f)}(\cdot)$. Recall that by Assumption 3.6, in the standard OTNM problem, the efficiency model is static and conservative, i.e., the amount of electricity needed to run ground-source heat pumps at customer homes is modeled via a constant coefficient of performance (COP). We denote $c_{\max} \coloneqq \max_{i\in[n]} \frac{1}{\mathtt{COP}_i}$ as the worst-case electricity-per-heat coefficient across all customers.

Now, note that OTNM objective function can be decomposed as

$$J(\pi; c) \;=\; G(\pi) \;+\; \sum_{t=1}^{T} p_t\, c_t(\pi)\, h_t(\pi),$$

as a function of the trajectory π, where $G(\pi) \geq 0$ captures all costs not related to ground-source heat pumps at households (e.g., fixed costs, demand charges, non-heat-pump energy consumption, etc.), p_t is the electricity price at time t, $h_t(\pi) \coloneqq \mathbf{1}^\top(\mathbf{b}_t + \mathbf{z}_t)$ is the total heating and/or cooling delivered by all heat pumps at time t under trajectory π, and $c_t(\pi) \leq 1$ is the electricity-per-heat coefficient at time t under trajectory π. In the static COP model, we have $c_t(\pi) \sim c_{\max}$ for all π, and it does not vary with time t.

In a more detailed dynamic COP model, the coefficient $c_t(\pi)$ can vary with time t and the trajectory π, reflecting the fact that ground-source heat pump efficiency depends on conditions (e.g., loop temperature, etc.) and the operating point of the heat pump. In practice, these coefficients satisfy the uniform bounds

$$c_{\min} \;\leq\; c_t(\pi) \;\leq\; c_{\max} \qquad \forall t,\ \forall \pi \in \mathcal{F},$$

for some $0 < c_{\min} \leq c_{\max}$. $c_t(\pi) \equiv c_{\max}$ for all t, π. Formally, we denote $c_{\min} \coloneqq \min_{i\in[n]} \frac{1}{\mathtt{COP}_i}$ as the best-case electricity-per-heat coefficient across all customers and operating conditions.

Then, suppose $\mathtt{OPT}_{\mathrm{dyn}}(\mathcal{I})$ is the optimal objective value of the offline OTNM problem under the dynamic COP model for any instance $\mathcal{I}$, and denote $\mathtt{OPT}_{\max}(\mathcal{I})$ as the optimal objective value of the offline OTNM problem under the static conservative COP model (i.e., with $c_t(\pi) \equiv c_{\max}$, under Assumption 3.6).

$$\mathtt{OPT}_{\mathrm{dyn}}(\mathcal{I}) := \min_{\pi\in\mathcal{F}} J(\pi; c) \quad \text{and} \quad \mathtt{OPT}_{\max}(\mathcal{I}) := \min_{\pi\in\mathcal{F}} J(\pi; c_{\max}).$$

Then it follows that $\mathtt{OPT}_{\max}(\mathcal{I}) \;\leq\; \frac{c_{\max}}{c_{\min}}\, \mathtt{OPT}_{\mathrm{dyn}}(\mathcal{I})$. Consequently, if an online algorithm satisfies $\mathtt{ALG}(\mathcal{I}) \;\leq\; \alpha\, \mathtt{OPT}_{\max}(\mathcal{I})$, then it is $\alpha \cdot \frac{c_{\max}}{c_{\min}}$-competitive with respect to the dynamic-COP benchmark:

$$\mathtt{ALG}(\mathcal{I}) \;\leq\; \alpha \frac{c_{\max}}{c_{\min}}\, \mathtt{OPT}_{\mathrm{dyn}}(\mathcal{I}).$$

Proof. Fix any feasible trajectory $\pi \in \mathcal{F}$. Since $c_t(\pi) \geq c_{\min}$ and $p_t, h_t(\pi) \geq 0$,

$$J(\pi; c) = G(\pi) + \sum_{t=1}^{T} p_t\, c_t(\pi)\, h_t(\pi) \;\geq\; \sum_{t=1}^{T} p_t\, c_t(\pi)\, h_t(\pi) \;\geq\; c_{\min} \sum_{t=1}^{T} p_t\, h_t(\pi).$$

Rearranging yields the trajectory-wise bound

$$\sum_{t=1}^{T} p_t\, h_t(\pi) \;\leq\; \frac{1}{c_{\min}}\, J(\pi; c). \tag{16}$$

Next, compare the static and dynamic objectives on the same trajectory π:

$$J(\pi; c_{\max}) = G(\pi) + \sum_{t=1}^{T} p_t\, c_{\max}\, h_t(\pi) = J(\pi; c) + \sum_{t=1}^{T} p_t \big(c_{\max} - c_t(\pi)\big)\, h_t(\pi).$$

Using $c_t(\pi) \geq c_{\min}$ gives $c_{\max} - c_t(\pi) \leq c_{\max} - c_{\min}$, hence

$$J(\pi; c_{\max}) \leq J(\pi; c) + (c_{\max} - c_{\min}) \sum_{t=1}^{T} p_t\, h_t(\pi) \overset{(16)}{\leq} J(\pi; c) + (c_{\max} - c_{\min}) \cdot \frac{1}{c_{\min}}\, J(\pi; c) = \frac{c_{\max}}{c_{\min}}\, J(\pi; c).$$

Let $\pi^\star_{\mathrm{dyn}} \in \arg\min_{\pi\in\mathcal{F}} J(\pi; c)$ be a dynamic-COP optimizer. Since $\mathtt{OPT}_{\max}(\mathcal{I})$ is the minimum of $J(\cdot; c_{\max})$ over the same feasible set $\mathcal{F}$,

$$\mathtt{OPT}_{\max}(\mathcal{I}) = \min_{\pi\in\mathcal{F}} J(\pi; c_{\max}) \leq J(\pi^\star_{\mathrm{dyn}}; c_{\max}) \leq \frac{c_{\max}}{c_{\min}}\, J(\pi^\star_{\mathrm{dyn}}; c) = \frac{c_{\max}}{c_{\min}}\, \mathtt{OPT}_{\mathrm{dyn}}(\mathcal{I}).$$

This completes the proof. □

B.5 Proof of Theorem 5.2

In this section, we prove Theorem 5.2, which states that if learned thresholds $\hat{\phi}_{\boldsymbol{\theta}}$ and $\hat{\psi}_{\boldsymbol{\theta}}$ lie within the robust certificate set $\mathcal{R}_\rho$ for some desired robustness factor $\rho > \rho$, then TANDEM using these learned thresholds is ρ-competitive for OTNM.

Proof. Before the main proof, we note that TANDEM using learned thresholds $\hat{\phi}_{\boldsymbol{\theta}}$ and $\hat{\psi}_{\boldsymbol{\theta}}$ produces a feasible solution to OTNM, following from the same argument as in Lemma B.1, which only relies on the structure of the algorithm, not the thresholds used.

We now proceed to bound the competitive ratio of TANDEM. For notational brevity, the following considers any arbitrary instance of OTNM $\mathcal{I} \in \Omega$ that satisfies Assumptions 3.2 - 3.9, where Ω is the set of all such instances. We start by considering the competitive ratio in a setting without rate constraints (i.e., where $\mathbf{x}_t$ is unconstrained for all t), and later extend the proof to the general case with rate constraints (see Lemma B.9). We recall the definition of active and inactive periods from Definition B.2, and the additional notations used in the proof of Theorem 4.3. We inherent the first result of the existing proof (Lemma B.3), which states a lower bound on OPT for any instance $\mathcal{I} \in \Omega$:

$$\texttt{OPT}(\mathcal{I}) \geq \sum_{k=1}^{K} G_k(\beta_k) + \left(D_b - \sum_{k=1}^{K} \beta_k\right) \hat{p\tilde{g}} + \bar{pc}|D_b| + \sum_{i=1}^{n}\sum_{\tau=1}^{T} \left(H_{i,\tau}(\xi_{i,\tau}) + E_{i,\tau}(f_{i,\tau})\right),$$

where $\hat{p\tilde{g}} = \hat{\phi}_{\boldsymbol{\theta}}(0)$ (i.e., the worst (highest) marginal cost that TANDEM's base drivers are willing to pay to fill the thermal network's effective storage). Next, we state an upper bound on the cost incurred by TANDEM using learned thresholds $\hat{\phi}_{\boldsymbol{\theta}}$ and $\hat{\psi}_{\boldsymbol{\theta}}$ for the same instance $\mathcal{I}$:

Lemma B.11. *Given that* TANDEM *produces K active periods, let $\hat{w}^{(b)}_{k,j}$, $B^{(b)}_{k,j}$, B, $\hat{w}^{(f)}_{i,\tau}$, and $\hat{v}^{(f)}_{i,\tau}$ be defined as outlined above. Then, the cost incurred by* TANDEM*'s solution is upper bounded as:*

$$\texttt{TANDEM}(\mathcal{I}) \leq \sum_{k=1}^{K}\sum_{j=1}^{v_k} \int_0^{\hat{w}^{(b)}_{k,j}} \hat{\phi}^{(b)}_{\boldsymbol{\theta},k,j}(y)dy + \sum_{\tau=1}^{T}\sum_{i=1}^{n} \left(\int_0^{\hat{w}^{(f)}_{i,\tau}} \hat{\phi}^{(f)}_{\boldsymbol{\theta},i,\tau}(y)dy + \int_0^{\hat{v}^{(f)}_{i,\tau}} \hat{\psi}_{\boldsymbol{\theta},i,\tau}(y)dy \right) \tag{17}$$

$$+ \left(B - \sum_{k=1}^{K}\sum_{j=1}^{v_k} \hat{w}^{(b)}_{k,j}\right)(\lambda_{\max} + 2\eta) + \bar{pc}|D_b| + \left(D_f^+ - \sum_{i=1}^{n}\sum_{\tau=1}^{T} \hat{w}^{(f)}_{i,\tau}\right)(\lambda_{\max} + 2\eta) + \left(|D_f| - \sum_{i=1}^{n}\sum_{\tau=1}^{T} \hat{v}^{(f)}_{i,\tau}\right)(\lambda_{\max} r + 2\gamma) \tag{18}$$

$$+ L \cdot S\,(\lambda_{\max} + 2\eta) \tag{19}$$

Proof. As in the proof of Lemma B.4, this follows by explicitly characterizing the worst-case cost of TANDEM using the learned threshold functions in the definitions in Algorithm 1 and Algorithm 2. The primary change with respect to Lemma B.4 is bounding (17)—both (18) and (19) use logic unchanged from the previous proof.

The first line worst-case cost that can be charged to each driver's pseudo-cost threshold functions, using the definitions of the pseudo-cost minimization subproblems in Algorithm 1 and Algorithm 2. In this case, TANDEM uses the learned thresholds $\hat{\phi}_{\boldsymbol{\theta}}$ and $\hat{\psi}_{\boldsymbol{\theta}}$ in place of the analytical thresholds. Even when learned thresholds are used, TANDEM only purchases energy or delivers flexible demand when the marginal cost of doing so (e.g., the market price p_t combined with the current efficiency, either $\mathbf{g}_t$ or $c^{(f)}_{i,t}$) is sufficiently low to make the minimization problems on Line 15 (Algorithm 2) and Line 3 (Algorithm 1) negative. Thus, we can say that the cost of purchasing energy attributable to the j^{th} base driver in the k^{th} active period is upper bounded by $\int_0^{\hat{w}^{(b)}_{k,j}} \hat{\phi}^{(b)}_{\boldsymbol{\theta},k,j}(y)dy$, analogous reasoning applies to the flexible demand drivers' purchasing and delivery costs, and aggregating over all drivers yields (17). □

Leveraging the definition of the robust certificate set $\mathcal{R}_\rho$ in Def. 5.1, we have the following technical lemmas to provide a relation between the learned thresholds and the optimal cost.

Lemma B.12. *By definition of the threshold function $\hat{\phi}_{\boldsymbol{\theta}}(\cdot)$ within the robust certificate set $\mathcal{R}_\rho$, the following relation always holds:*

$$\int_0^w \hat{\phi}_{\boldsymbol{\theta}}(y)dy + (1-w)(\lambda_{\max} + 2\eta) \leq \rho\left[\hat{\phi}_{\boldsymbol{\theta}}(w) - 2\eta\right], \quad \forall w \in [0,1].$$

Lemma B.13. *By definition of the threshold functions $\hat{\phi}_{\boldsymbol{\theta}}(\cdot)$ and $\hat{\psi}_{\boldsymbol{\theta}}(\cdot)$ within the robust certificate set $\mathcal{R}_\rho$, the following relation always holds:*

$$\int_0^w \hat{\phi}_{\boldsymbol{\theta}}(y)dy + (1-w)(\lambda_{\max} + 2\eta) + \int_0^v \hat{\psi}_{\boldsymbol{\theta}}(y)dy + (1-v)(\lambda_{\max} r + 2\gamma) \leq \rho\left[\hat{\phi}_{\boldsymbol{\theta}}(w) - 2\eta + \hat{\psi}_{\boldsymbol{\theta}}(v) - 2\gamma\right], \quad \forall w, v \in [0,1].$$

$$\int_0^v \hat{\psi}_{\boldsymbol{\theta}}(y)dy + (1-v)(\lambda_{\max} r + 2\gamma) \leq \rho\left[\hat{\psi}_{\boldsymbol{\theta}}(v) - 2\gamma\right], \quad \forall v \in [0,1].$$

The proofs of both lemmas follow directly from the definition of the robust certificate set $\mathcal{R}_\rho$ in Def. 5.1. Then, combining Lemmas B.3, B.11, B.12, and B.13, we claim the following holds:

$$\frac{\texttt{TANDEM}(\mathcal{I}) - (\lambda_{\max} + 2\eta)\cdot(S + LS)}{\texttt{OPT}(\mathcal{I})} \leq \rho.$$

Letting $Q = \sum_{k=1}^{K}\sum_{j=1}^{v_k}\int_0^{\hat{w}_{k,j}^{(b)}} \hat{\phi}_{\boldsymbol{\theta},k,j}^{(b)}(y)dy + \sum_{\tau=1}^{T}\sum_{i=1}^{n}\left(\int_0^{\hat{w}_{i,\tau}^{(f)}} \hat{\phi}_{\boldsymbol{\theta},i,\tau}^{(f)}(y)dy + \int_0^{\hat{v}_{i,\tau}^{(f)}} \hat{\psi}_{\boldsymbol{\theta},i,\tau}(y)dy\right)$ denote the integrals over the *learned* thresholds, we have the following using the same logic as in the proof of Theorem 4.3—we can rewrite the left-hand side of the above expression as:

$$\frac{\texttt{TANDEM}(\mathcal{I}) - (\lambda_{\max}+2\eta)\,(S+LS)}{\texttt{OPT}(\mathcal{I})} \leq \frac{Q + (B-\hat{W}_b)(\lambda_{\max}+2\eta) + \bar{pc}|D_b| + (D_f^+ - \hat{W}_f)(\lambda_{\max}+2\eta) + (|D_f| - \hat{V})(\lambda_{\max}r+2\gamma) - (\lambda_{\max}+2\eta)S}{\sum_{k=1}^{K} G_k(\beta_k) + (D_b - \boldsymbol{\beta})\hat{\bar{pg}} + \bar{pc}|D_b| + \sum_{i=1}^{n}\sum_{\tau=1}^{T}\left(H_{i,\tau}(\xi_{i,\tau}) + E_{i,\tau}(f_{i,\tau})\right)},$$

By Lemma B.7, we can write:

$$\leq \max\left\{\frac{Q + (\boldsymbol{\beta}-\hat{W}_b)(\lambda_{\max}+2\eta) + (D_f^+ - \hat{W}_f)(\lambda_{\max}+2\eta) + (|D_f| - \hat{V})(\lambda_{\max}r+2\gamma)}{\sum_{k=1}^{K} G_k(\beta_k) + \bar{pc}\boldsymbol{\beta} + \sum_{i=1}^{n}\sum_{\tau=1}^{T}\left(H_{i,\tau}(\xi_{i,\tau}) + E_{i,\tau}(f_{i,\tau})\right)}, \frac{(D_b-\boldsymbol{\beta})(\lambda_{\max}+2\eta) + (|D_b|-\boldsymbol{\beta})\bar{pc}}{(D_b-\boldsymbol{\beta})\hat{\bar{pg}} + (|D_b|-\boldsymbol{\beta})\bar{pc}}, \frac{\bar{pc}\boldsymbol{\beta}}{\bar{pc}\boldsymbol{\beta}}\right\}, \quad (20)$$

where the definition of $\hat{\bar{pg}}$ (using Def. 5.1) ensures that the second term in the max is at most ρ, and the third term is 1. We now focus on the first term. For the sake of contradiction, suppose that:

$$\frac{Q + (\boldsymbol{\beta}-\hat{W}_b)(\lambda_{\max}+2\eta) + (D_f^+ - \hat{W}_f)(\lambda_{\max}+2\eta) + (|D_f| - \hat{V})(\lambda_{\max}r+2\gamma)}{\sum_{k=1}^{K} G_k(\beta_k) + \sum_{i=1}^{n}\sum_{\tau=1}^{T}\left(H_{i,\tau}(\xi_{i,\tau}) + E_{i,\tau}(f_{i,\tau})\right)} > \rho. \quad (21)$$

Then, continuing with the same logic as in the proof of Theorem 4.3 and using Lemma B.8, we have the following relation:

$$\begin{aligned}
&\frac{Q + (B-\hat{W}_b)(\lambda_{\max}+2\eta) + (D_f^+ - \hat{W}_f)(\lambda_{\max}+2\eta) + (|D_f| - \hat{V})(\lambda_{\max}r+2\gamma)}{\sum_{k=1}^{K} G_k(B_k) + \sum_{i=1}^{n}\sum_{\tau=1}^{T}\left(H_{i,\tau}(\xi_{i,\tau}) + E_{i,\tau}(f_{i,\tau})\right)},\\
&\geq \frac{Q + (\boldsymbol{\beta}-\hat{W}_b)(\lambda_{\max}+2\eta) + (D_f^+ - \hat{W}_f)(\lambda_{\max}+2\eta) + (|D_f| - \hat{V})(\lambda_{\max}r+2\gamma) + (B-\boldsymbol{\beta})(\lambda_{\max}+2\eta)}{\sum_{k=1}^{K} G_k(\beta_k) + \sum_{i=1}^{n}\sum_{\tau=1}^{T}\left(H_{i,\tau}(\xi_{i,\tau}) + E_{i,\tau}(f_{i,\tau})\right) + \sum_{k=1}^{K}\left(G_k(B_k) - G_k(\beta_k)\right)} > \rho,
\end{aligned}$$

Using the learned threshold functions, we know that during the k^{th} active period and the lifetime of the j^{th} base driver, the minimum marginal purchasing cost observed is given by $\hat{\phi}_{\boldsymbol{\theta},k,j}^{(b)}(\hat{w}_{k,j}^{(b)}) - 2\eta$ (by definition of the threshold function). Similarly, during the lifetime of the flexible demand driver arriving at time τ for customer i, the minimum marginal purchasing and delivery costs observed are given by $\hat{\phi}_{\boldsymbol{\theta},i,\tau}^{(f)}(\hat{w}_{i,\tau}^{(f)}) - 2\eta$ and $\hat{\psi}_{\boldsymbol{\theta},i,\tau}(\hat{v}_{i,\tau}^{(f)}) - 2\gamma$, respectively. This gives the following lower bounds on the terms that depend on G_k, $H_{i,\tau}$, and $E_{i,\tau}$, respectively:

$$\begin{aligned}
\sum_{k=1}^{K} G_k\left(\sum_{j=1}^{v_k} B_{k,j}^{(b)}\right) &\geq \sum_{j=1}^{v_k}\left(\hat{\phi}_{\boldsymbol{\theta},k,j}^{(b)}(\hat{w}_{k,j}^{(b)}) - 2\eta\right) \times B_{k,j}^{(b)},\\
\sum_{i=1}^{n}\sum_{\tau=1}^{T} H_{i,\tau}(\xi_{i,\tau}) &\geq \sum_{i=1}^{n}\sum_{\tau=1}^{T}\left(\hat{\phi}_{\boldsymbol{\theta},i,\tau}^{(f)}(\hat{w}_{i,\tau}^{(f)}) - 2\eta\right) \times \xi_{i,\tau},\\
\sum_{i=1}^{n}\sum_{\tau=1}^{T} E_{i,\tau}(f_{i,\tau}) &\geq \sum_{i=1}^{n}\sum_{\tau=1}^{T}\left(\hat{\psi}_{\boldsymbol{\theta},i,\tau}(\hat{v}_{i,\tau}^{(f)}) - 2\gamma\right) \times f_{i,\tau}.
\end{aligned}$$

Substituting these bounds into the previous expression, we have that the left-hand side of (21) is less than or equal to:

$$\frac{Q + (B-\hat{W}_b)(\lambda_{\max}+2\eta) + (D_f^+ - \hat{W}_f)(\lambda_{\max}+2\eta) + (|D_f| - \hat{V})(\lambda_{\max}r+2\gamma)}{\sum_{k=1}^{K}\sum_{j=1}^{v_k}\left(\hat{\phi}_{\boldsymbol{\theta},k,j}^{(b)}(\hat{w}_{k,j}^{(b)}) - 2\eta\right) B_{k,j}^{(b)} + \sum_{i=1}^{n}\sum_{\tau=1}^{T}\left(\hat{\phi}_{\boldsymbol{\theta},i,\tau}^{(f)}(\hat{w}_{i,\tau}^{(f)}) - 2\eta\right)\xi_{i,\tau} + \sum_{i=1}^{n}\sum_{\tau=1}^{T}\left(\hat{\psi}_{\boldsymbol{\theta},i,\tau}(\hat{v}_{i,\tau}^{(f)}) - 2\gamma\right) f_{i,\tau}}.$$

By rearranging the terms in the above and substituting for Q, we obtain the following:

$$\leq \frac{\sum_{k=1}^{K}\sum_{j=1}^{v_k} X_{k,j}^{(b)} + \sum_{i=1}^{n}\sum_{\tau=1}^{T} X_{i,\tau}^{(f)}}{\sum_{k=1}^{K}\sum_{j=1}^{v_k}\left(\hat{\phi}_{\boldsymbol{\theta},k,j}^{(b)}(\hat{w}_{k,j}^{(b)}) - 2\eta\right) B_{k,j}^{(b)} + \sum_{i=1}^{n}\sum_{\tau=1}^{T}\left(\hat{\phi}_{\boldsymbol{\theta},i,\tau}^{(f)}(\hat{w}_{i,\tau}^{(f)}) - 2\eta\right)\xi_{i,\tau} + \sum_{i=1}^{n}\sum_{\tau=1}^{T}\left(\hat{\psi}_{\boldsymbol{\theta},i,\tau}(\hat{v}_{i,\tau}^{(f)}) - 2\gamma\right) f_{i,\tau}}. \quad (22)$$

where $X_{k,j}^{(b)}$ and $X_{i,\tau}^{(f)}$ are defined as follows (for each driver):

$$\begin{aligned}
X_{k,j}^{(b)} &= \int_0^{\hat{w}_{k,j}^{(b)}} \hat{\phi}_{\boldsymbol{\theta},k,j}^{(b)}(y)dy + (B_{k,j}^{(b)} - \hat{w}_{k,j}^{(b)})(\lambda_{\max}+2\eta),\\
X_{i,\tau}^{(f)} &= \int_0^{\hat{w}_{i,\tau}^{(f)}} \hat{\phi}_{\boldsymbol{\theta},i,\tau}^{(f)}(y)dy + (\xi_{i,\tau} - \hat{w}_{i,\tau}^{(f)})(\lambda_{\max}+2\eta) + \int_0^{\hat{v}_{i,\tau}^{(f)}} \hat{\psi}_{\boldsymbol{\theta},i,\tau}(y)dy + (f_{i,\tau} - \hat{v}_{i,\tau}^{(f)})(\lambda_{\max}r+2\gamma).
\end{aligned}$$

Since (21) is $> \rho$ by assumption, one of the following cases must be true:

Case 1: There exists some base driver with $k \in \{1, \ldots, K\}$ and $j \in \{1, \ldots, v_k\}$ such that:

$$X_{k,j}^{(b)} > \rho \left(\hat{\phi}_{\boldsymbol{\theta},k,j}^{(b)}(\hat{w}_{k,j}^{(b)}) - 2\eta \right) B_{k,j}^{(b)}.$$

Case 2: There exists some flexible demand driver with $i \in \{1, \ldots, n\}$ and $\tau \in \{1, \ldots, T\}$ such that:

$$X_{i,\tau}^{(f)} > \rho \left(\left(\hat{\phi}_{\boldsymbol{\theta},i,\tau}^{(f)}(\hat{w}_{i,\tau}^{(f)}) - 2\eta \right) \xi_{i,\tau} + \left(\hat{\psi}_{\boldsymbol{\theta},i,\tau}(\hat{v}_{i,\tau}^{(f)}) - 2\gamma \right) f_{i,\tau} \right).$$

But, if either of these two cases are true (i.e., for any unit of base or flexible demand), they contradict Lemmas B.12 and B.13, respectively. Thus, we conclude that our initial assumption in (21) must be false, completing the proof that $\frac{\texttt{TANDEM}(\mathcal{I}) - (\lambda_{\max} + 2\eta) \cdot (S + LS)}{\texttt{OPT}(\mathcal{I})} \leq \rho$.

Using this result, we have:

$$\texttt{TANDEM}(\mathcal{I}) \leq \rho \texttt{OPT}(\mathcal{I}) + (\lambda_{\max} + 2\eta) \cdot (S + LS),$$

where $(\lambda_{\max} + 2\eta) \cdot (S + LS)$ is a constant. This shows that TANDEM using learned threshold functions is ρ-competitive under Def. 3.1.

Finally, we remark on the impact of rate constraints in the following lemma:

Lemma B.14. TANDEM *using learned threshold functions* $\hat{\phi}_{\boldsymbol{\theta}}$ *and* $\hat{\psi}_{\boldsymbol{\theta}}$ *lying within the robust certificate set* $\mathcal{R}_\rho$ *for some desired robustness factor* $\rho > \rho$ *remains* ρ*-competitive when rate constraints are imposed on the network management decisions of both* TANDEM *and* OPT.

Proof. Denote $\mathbf{x}_t$ and $\mathbf{x}_t^\star$ as the purchasing decisions made by TANDEM and OPT (respectively), and define $\mathcal{T}_r \subset \{1, \ldots, T\}$ as the set of time steps where rate constraints are active for TANDEM (i.e., $\mathbf{x}_t$ is limited by a maximum rate at those time steps).

When TANDEM is *not* rate-constrained, recall that the i^{th} purchasing driver solves the following pseudo-cost minimization problem at each time step: $\mathbf{x}_{i,t} \leftarrow \text{argmin}_{\mathbf{x} \in \mathbb{R}^m} \, p_t \mathbf{g}_t^\top \mathbf{x} + \eta \|\mathbf{x} - \tilde{\mathbf{x}}_{i,t}\| - \hat{\Phi}_{\boldsymbol{\theta},i}(w_{i,t-1}, w_{i,t-1} + \mathbf{1}^\top \mathbf{x})$, where $w_{i,t} = w_{i,t-1} + \mathbf{1}^\top \mathbf{x}_{i,t}$ corresponds to the best price observed so far by the i^{th} purchasing driver up to time t, given by $\hat{\phi}_{\boldsymbol{\theta},i}(w_{i,t}) - 2\eta$.

When rate constraints are active at time $t \in \mathcal{T}_r$, the purchasing decision is constrained to $\bar{\mathbf{x}}_{i,t}$, which is the solution to the same pseudo-cost minimization problem but with an additional constraint on the maximum allowable purchase rate. Let $\bar{w}_{i,t} = \bar{w}_{i,t-1} + \mathbf{1}^\top \bar{\mathbf{x}}_{i,t}$ denote the (reduced) amount purchased during this step, and note that by the monotonicity of ϕ (see Def. 5.1), we have $\hat{\phi}_{\boldsymbol{\theta},i}(\bar{w}_{i,t}) > \hat{\phi}_{\boldsymbol{\theta},i}(w_{i,t})$. This means that although TANDEM's total cost is upper-bounded with respect to $\hat{\phi}_{\boldsymbol{\theta},i}(\bar{w}_{i,t})$, in a rate-constrained setting, for any $t \in \mathcal{T}_r$, the cost incurred by TANDEM is actually lower since it purchases energy at the lower cost corresponding to $\hat{\phi}_{\boldsymbol{\theta},i}(w_{i,t})$. Second, since $\hat{\phi}_{\boldsymbol{\theta},i}(\bar{w}_{i,t})$ typically forms a lower bound on OPT's cost (as shown in the main proof), the bound must be adjusted to account for the fact that the actual price paid by OPT is the lower cost corresponding to $\hat{\phi}_{\boldsymbol{\theta},i}(w_{i,t})$. Formally, we define a function $\mu(t)$ for each time step $t \in \mathcal{T}_r$ as follows:

$$\mu(t) = \sum_{i \in \mathcal{B}_t \cup \mathcal{F}_t} \left(\hat{\phi}_{\boldsymbol{\theta},i}(\bar{w}_{i,t}) - \hat{\phi}_{\boldsymbol{\theta},i}(w_{i,t}) \right),$$

where $\mathcal{B}_t$ and $\mathcal{F}_t$ denote the sets of active drivers at time t. This function $\mu(\cdot)$ captures the cumulative difference in cost bounds due to rate constraints across all rate-constrained periods $t \in \mathcal{T}_r$. Picking up from (20) in the main proof, we can bound the competitive ratio as follows:

$$\frac{\texttt{TANDEM}(\mathcal{I}) - (\lambda_{\max} + 2\eta)(S + LS)}{\texttt{OPT}(\mathcal{I})} \leq \frac{Q + (D_b - \hat{W}_b)(\lambda_{\max} + 2\eta) + \bar{pc}|D_b| + (D_f^+ - \hat{W}_f)(\lambda_{\max} + 2\eta) + (|D_f| - \hat{V})(\lambda_{\max} r + 2\gamma) - \sum_{t \in \mathcal{T}_r} \mu(t) \mathbf{1}^\top \mathbf{x}_t}{\sum_{k=1}^K G_k'(\beta_k) + (D_b - \boldsymbol{\beta})\hat{\bar{pg}} + \bar{pc}|D_b| + \sum_{i=1}^n \sum_{\tau=1}^T \left(H_\tau^{'(i)}(\xi_{i,\tau}) + E_{i,\tau}(f_{i,\tau}) \right)},$$

where $G_k'(\cdot)$ and $H_\tau^{'(i)}(\cdot)$ are defined similarly to $G_k(\cdot)$ but accounts for rate constraints in OPT's solution. Note that the last term in the numerator captures the reduction in TANDEM's cost due to the "over-estimation" of cost incurred during rate-constrained periods. Then, similarly to the main proof, $G_k'(\cdot)$ and $H_\tau^{'(i)}(\cdot)$ satisfy:

$$\sum_{k=1}^K G_k \left(\sum_{j=1}^{v_k} B_{k,j}^{(b)} \right) + \sum_{i=1}^n \sum_{\tau=1}^T H_{i,\tau}(\xi_{i,\tau}) \geq \sum_{j=1}^{v_k} \left(\phi_{k,j}^{(b)}(\hat{w}_{k,j}^{(b)}) - 2\eta \right) \times B_{k,j}^{(b)} + \sum_{i=1}^n \sum_{\tau=1}^T \left(\phi_{i,\tau}^{(f)}(\hat{w}_{i,\tau}^{(f)}) - 2\eta \right) \times \xi_{i,\tau} - \sum_{t \in \mathcal{T}_r} \mu(t) \mathbf{1}^\top \mathbf{x}_t^\star.$$

Then, using (15) and following the same reasoning as in the main proof, we have:

$$\frac{\texttt{TANDEM}(\mathcal{I}) - (\lambda_{\max} + 2\eta)(S + LS)}{\texttt{OPT}(\mathcal{I})} \leq$$

$$\max \left\{ \frac{\sum_{k=1}^K \sum_{j=1}^{v_k} X_{k,j}^{(b)} + \sum_{i=1}^n \sum_{\tau=1}^T X_{i,\tau}^{(f)} - \sum_{t \in \mathcal{T}_r} \mu(t) \mathbf{1}^\top \mathbf{x}_t}{\sum_{k=1}^K \sum_{j=1}^{v_k} \left(\phi_{k,j}^{(b)}(\hat{w}_{k,j}^{(b)}) - 2\eta \right) B_{k,j}^{(b)} + \sum_{i=1}^n \sum_{\tau=1}^T \left(\phi_{i,\tau}^{(f)}(\hat{w}_{i,\tau}^{(f)}) - 2\eta \right) \xi_{i,\tau} + \sum_{i=1}^n \sum_{\tau=1}^T \left(\psi_{i,\tau}(\hat{v}_{i,\tau}^{(f)}) - 2\gamma \right) f_{i,\tau} - \sum_{t \in \mathcal{T}_r} \mu(t) \mathbf{1}^\top \mathbf{x}_t^\star}, \rho \right\}$$

In the worst-case, we have that $\frac{\mathbf{1}^\top \mathbf{x}_t}{\mathbf{1}^\top \mathbf{x}_t^\star} \leq \rho$ for any $t \in \mathcal{T}_r$ – we show this by contradiction. Suppose $\mathcal{I} = [(p_t, \mathbf{g}_t, \mathbf{b}_t, \mathbf{f}_t, \Delta_t)]_{t \in [T]}$ is the worst-case instance and at time slot t, we have $\frac{\mathbf{1}^\top \mathbf{x}_t}{\mathbf{1}^\top \mathbf{x}_t^\star} > \rho$. We can construct a modified input instance $\mathcal{I}'$ from $\mathcal{I}$ by increasing the price and efficiency at time t by δ_p and δ_g, where $\delta = \delta_p \times \delta_g \leq \mu(t)$. That is:

$$\mathcal{I}' = \left[(p_1, \mathbf{g}_1, \mathbf{b}_1, \mathbf{f}_1, \Delta_1), \ldots, (p_{t-1}, \mathbf{g}_{t-1}, \mathbf{b}_{t-1}, \mathbf{f}_{t-1}, \Delta_{t-1}), (p_t + \delta_p, \mathbf{g}_t + \delta_g, \mathbf{b}_t, \mathbf{f}_t, \Delta_t), (p_{t+1}, \mathbf{g}_{t+1}, \mathbf{b}_{t+1}, \mathbf{f}_{t+1}, \Delta_{t+1}), \ldots, (p_T, \mathbf{g}_T, \mathbf{b}_T, \mathbf{f}_T, \Delta_T) \right].$$

In this modified input instance, the cost of TANDEM increases by $\delta \times \mathbf{1}^\top \mathbf{x}_t$, while the cost of OPT increases by less than $\frac{\delta \times \mathbf{1}^\top \mathbf{x}_t}{\rho}$, since OPT can always choose to purchase at other time slots. This contradicts the assumption that $\mathcal{I}$ is the worst-case input instance, and thus, we conclude that in the worst-case, $\mathbf{1}^\top \mathbf{x}_t / \mathbf{1}^\top \mathbf{x}_t^\star \le \rho$ for any $t \in \mathcal{T}_r$. By the above we have that $\frac{\sum_{t \in \mathcal{T}_r} \mu(t) \mathbf{1}^\top \mathbf{x}_t}{\sum_{t \in \mathcal{T}_r} \mu(t) \mathbf{1}^\top \mathbf{x}_t^\star} \le \rho$, which yields:

$$\frac{\texttt{TANDEM}(\mathcal{I}) - (\lambda_{\max} + 2\eta)\,(S + LS)}{\texttt{OPT}(\mathcal{I})} \le$$

$$\max\left\{ \frac{\sum_{k=1}^{K} \sum_{j=1}^{v_k} X_{k,j}^{(b)} + \sum_{i=1}^{n} \sum_{\tau=1}^{T} X_{i,\tau}^{(f)}}{\sum_{k=1}^{K} \sum_{j=1}^{v_k} \left(\phi_{k,j}^{(b)}(\hat{w}_{k,j}^{(b)}) - 2\eta \right) B_{k,j}^{(b)} + \sum_{i=1}^{n} \sum_{\tau=1}^{T} \left(\phi_{i,\tau}^{(f)}(\hat{w}_{i,\tau}^{(f)}) - 2\eta \right) \xi_{i,\tau} + \sum_{i=1}^{n} \sum_{\tau=1}^{T} \left(\psi_{i,\tau}(\hat{v}_{i,\tau}^{(f)}) - 2\gamma \right) f_{i,\tau}},\ \rho \right\}.$$

Combining the above with the contradiction in (21) from the main proof, we have that TANDEM using learned threshold functions is ρ-competitive even when rate constraints are imposed on the network management side, as desired. □

□

B.6 Proof of Lemma 5.3

In this section, we prove Lemma 5.3, which states that the robust certificate set $\mathcal{R}_\rho$ is a convex set when learned thresholds $\hat{\phi}_{\boldsymbol{\theta}}$ and $\hat{\psi}_{\boldsymbol{\theta}}$ are parameterized as piecewise-affine functions defined by $\boldsymbol{\theta}_\phi, \boldsymbol{\theta}_\psi \in \mathbb{R}_+^K$.

Proof. Denote the grid discretization points by $0 = w_1 < w_2 < \cdots < w_K = 1$, with $w_j = \frac{j-1}{K-1}$. Each threshold function $\phi_{\boldsymbol{\theta}} : [0,1] \to \mathbb{R}$ or $\psi_{\boldsymbol{\theta}} : [0,1] \to \mathbb{R}$ is parameterized by its knot values $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_K)$ and linear interpolation on each interval $[w_j, w_{j+1}]$. We want to show that $\mathcal{R}_\rho$ is a convex set in $\boldsymbol{\theta}_\phi, \boldsymbol{\theta}_\psi \in \mathbb{R}_+^K$.

First, we introduce the hat-basis representation of piecewise-affine functions, which will be useful in the proof. Let $\{h_j(\cdot)\}_{j=1}^K$ be the standard PWA "tent" basis: each h_j is piecewise-linear, $h_j(w_i) = \mathbf{1}\{i = j\}$, supported on $[w_{j-1}, w_{j+1}]$ (with boundary conventions for $j = 1, K$), and $\sum_j h_j(u) = 1$ for all $u \in [0,1]$. Under linear interpolation, we have:

$$\phi_{\boldsymbol{\theta}}(u) = \sum_{j=1}^{K} \theta_{j,\phi} h_j(u), \qquad \Phi_{\boldsymbol{\theta}}(0,w) = \int_0^w \phi_{\boldsymbol{\theta}}(u)\,du = \sum_{j=1}^{K} \theta_{j,\phi} H_j(w),$$

so for fixed u, w, both $\phi_{\boldsymbol{\theta}}(u)$ and $\Phi_{\boldsymbol{\theta}}(0,w)$ are affine in the parameter vector $\boldsymbol{\theta}_\phi$. The same applies to $\psi_{\boldsymbol{\theta}}$ and $\Psi_{\boldsymbol{\theta}}(0,v)$.

All defining constraints of $\mathcal{R}_\rho$ are affine in the product space $(\boldsymbol{\theta}_\phi, \boldsymbol{\theta}_\psi)$. (i) range and endpoint: $\lambda_{\min} \le \theta_{j,\phi} \le \lambda_{\max}$, $\lambda_{\min} r \le \theta_{j,\psi} \le \lambda_{\max} r$, $\theta_{K,\phi} \le \lambda_{\min} + 2\eta$, and $\theta_{K,\psi} \le \lambda_{\min} r + 2\gamma$ (box and halfspace); (ii) monotone non-increasing: $\theta_{j+1,\phi} - \theta_{j,\phi} \le 0$ and $\theta_{j+1,\psi} - \theta_{j,\psi} \le 0$ for $j = 1, \ldots, K-1$ (segment slopes); (iii) for $\forall w \in [0,1], v \in [0,w]$, the robustness inequality

$$\sum_{j=1}^{K} \theta_{j,\phi} H_j(w) + (1-w)(\lambda_{\max} + 2\eta) + \sum_{j=1}^{K} \theta_{j,\psi} H_j(v) + (1-v)(\lambda_{\max} r + 2\gamma) \le \rho \left[\sum_{j=1}^{K} \theta_{j,\phi} h_j(w) + \sum_{j=1}^{K} \theta_{j,\psi} h_j(v) - 2\eta - 2\gamma \right].$$

is an affine halfspace in $(\boldsymbol{\theta}_\phi, \boldsymbol{\theta}_\psi)$. Intersecting these halfspaces over all (w,v) and with the linear constraints yields a convex set. Hence $\mathcal{R}_\rho$ is convex. Therefore, the robust certificate set is convex in the piecewise-affine parameter space.

□