Dynamic resource allocation: The geometry and robustness of constant regret

Alberto Vera, Alessandro Arlotto , Itai Gurvich*, and Eli Levin

We study a family of dynamic resource allocation problems, wherein requests of different types arrive over time and are accepted or rejected. Each request type is characterized by its reward, arrival probability, and resource consumption.

An upper bound for collected reward is given by a linear optimization problem with a random right-hand side. This type of problem, known as packing LP, is ubiquitous in resource allocation problems. We provide a detailed characterization of the parametric structure of this packing LP. Relying on this geometric understanding, we re-visit and expand on BudgetRatio algorithms that achieve constant regret by re-solving this same packing LP in each period and accepting requests "scored" as sufficiently valuable.

We illustrate the benefits of the geometric view in proving that (i) BudgetRatio achieves constant regret relative to the offline (full information) upper bound in the presence of inventory that is (slowly) restocked (ii) Within explicitly identifiable bounds, the algorithm's regret is robust to misspecification of the model parameters. This gives bounds for the "bandits" version of the problem where the parameters have to be learned. (iii) The algorithm has an equivalent formulation as a generalized bid-price algorithm where the bid prices can be adaptively and efficiently computed.

Our analysis focuses on the evolution of the remaining inventory—in turn of the LP that drives BudgetRatio—as a stochastic process. We prove that it is attracted to "sticky" regions of the state space where the online algorithm takes actions consistent with the optimal basis of the offline upper bound, a basis that is revealed only in hindsight, at the horizon's end.

Key words: multi-secretary problem, online packing, regret, adaptive online policy.

MSC2010 subject classification: Primary: 90C39. Secondary: 60C05, 60J05, 68W27, 68W40, 90C27, 90C40.

OR/MS subject classification: Primary: Decision Analysis: Sequential, Theory; Dynamic programming: Markov. Secondary: Probability: Markov processes; Analysis of Algorithms: Suboptimal algorithms.

History: *This version:* September 20, 2023.

---

* The work of the third author was supported by DoD grant W911NF-20-C-0008

1. Introduction We study a family of dynamic resource allocation problems described as follows. Requests of multiple types arrive over a finite horizon of T discrete periods. If accepted, a request consumes a set of resources (that depends on the request's type) and generates a reward. There is an inventory of resources available at time 0 and additional units of inventory may be re-stocked over time. The controller's objective is to use its resource inventory to maximize the total reward collected over the finite horizon.

The important and well-studied network revenue management problem as well as some assembly, distribution, and matching problems are members of this family.

If the controller could solve the problem in an *offline* fashion, she would wait for the end of the horizon and, given the realization of the random arrivals, choose the best allocation of resources to requests. The reward of the offline controller is an upper bound on any online algorithm.

At each time step $t = 1, \ldots, T$, a request of type $j \in \mathcal{J}$ arrives with probability p_j and simultaneously a unit of resource $i \in \mathcal{R}$ is restocked with probability ϱ_i. The algorithm we study, which we refer to as BudgetRatio, is based on re-solving the following *packing* linear program at each time period:

$$\max_y v'y \quad \text{s.t.} \quad Ay \le \underbrace{\frac{1}{T-t} I^t + \varrho}_{\text{per-period inventory}}, \quad \underbrace{0 \le y \le p}_{\text{per-period demand}}, \tag{1}$$

where $p = (p_j)_{j\in\mathcal{J}}$ are the arrival probabilities of requests, $\varrho = (\varrho_i)_{i\in\mathcal{R}}$ the resource restock probabilities, $v = (v_j)_{j\in\mathcal{J}}$ is the vector of rewards, and A is the $|\mathcal{R}| \times |\mathcal{J}|$ resource-consumption matrix. Finally, $I^t \in \mathbb{N}^{\mathcal{R}}$ is the available inventory of different resources at t. Since there are $(T-t)$ periods to go, the per-period expected available inventory is $\frac{1}{T-t} I^t + \varrho$ and the per-period expected demand is p.

In a solution $\bar{y}^t$ to Eq. (1), $\bar{y}^t_j$ is a proxy for the fraction of type-j requests that we want to accept: an inventory-dependent "score" of type j. BudgetRatio accepts requests with sufficiently large scores, i.e., such that $\bar{y}^t_j \ge \eta_j$ for thresholds η_j that we will explicitly specify. Viewed as a random process, these scores $\bar{y}^t$ depend, through the LP solution, on the random *budget-ratio* process $R^t := \frac{1}{T-t} I^t + \varrho$. This random process, evolving in the space of scaled resources, drives our analysis; see Fig. 1.

Methodology: a geometric view of re-solving policies. The problems we consider cannot be solved optimally due to the so-called "curse of dimensionality". This motivates the pursuit of policies that are simple to implement, adapt, and scale according to the problem instance. Algorithms based on linear programming have been introduced to overcome this challenge.

We uncover fundamental structure of the online stochastic packing problem. We expose the problem's geometric nature and study the budget consumption dynamics as a *stochastic-process* in

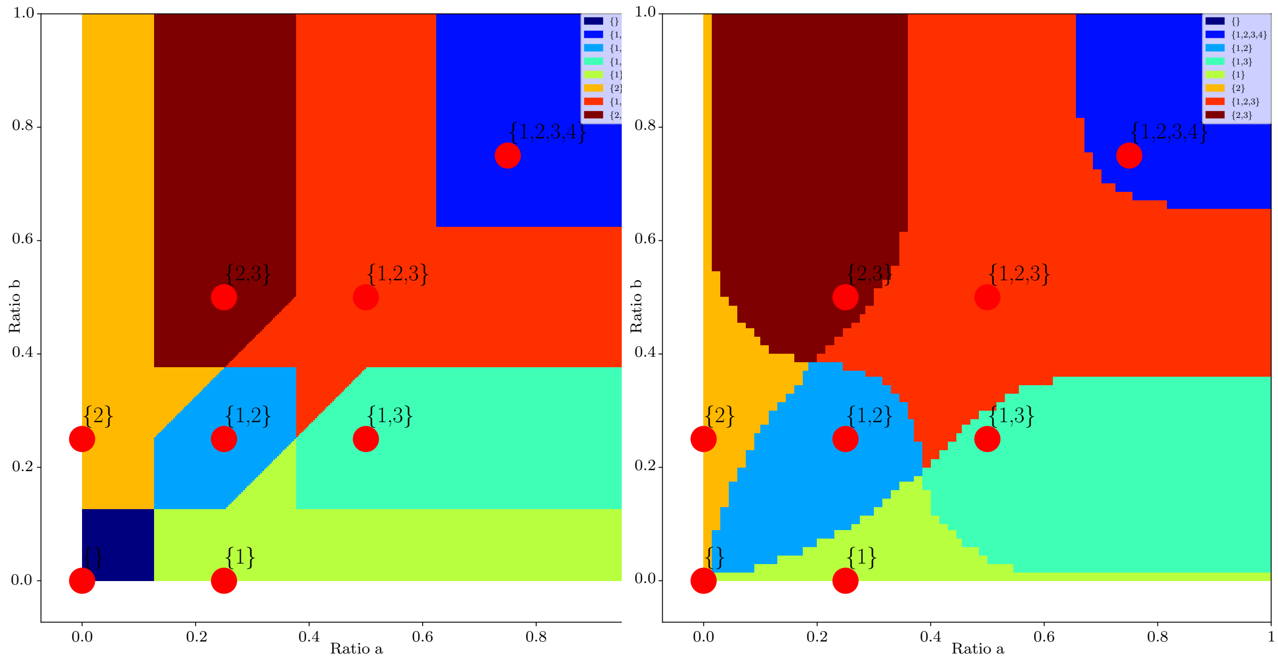

FIGURE 1. Action regions of BudgetRatio for a problem with two resources $i \in \{a,b\}$ and four request types $j \in \{1,2,3,4\}$. The plot is in the space of ratios that represent the per-period resource availability $R_i^t = \frac{1}{T-t} I_i^t + \varrho_i$ for $i \in \{a,b\}$. Each point corresponds to a pair of budget states (R_a, R_b). When solving the LP, we obtain the set of request types $\mathcal{K} = \{j : \bar{y}_j^t \geq \eta_j\}$ that should be accepted at that inventory level. Each color on the plot corresponds to a different such set. The light-blue rhombus-like region, for instance, corresponds to $\mathcal{K} = \{1,2\}$; when R^t is in this region, BudgetRatio accepts only type-1 and type-2 requests. (RIGHT) The action regions of the optimal policy (computed via DP) with 70 periods to go.

the space $\mathbb{R}_+^{\mathcal{R}}$ of budget ratios. The analysis reveals how BudgetRatio interacts with the geometry of the packing LP.

The thresholding of the decision $\bar{y}^t$ divides the space of resource-budgets $\mathbb{R}_+^{\mathcal{R}}$ into mutually exclusive *action* regions. When the ratio is in a given region, all requests j associated to this region (those for which $\bar{y}_j^t \geq \eta_j$) are accepted and all others are rejected; see Fig. 1. In this way, the "location" of R^t determines the actions that the algorithm takes.

The offline problem is a packing LP whose right-hand side corresponds to the (random) realization of total demand and restock over the horizon. To achieve constant regret, an online policy must act in a way that is consistent with the optimal, unknown, basis of the benchmark offline problem. This is made mathematically meaningful in Proposition 3 which relates the regret of any policy to the time in which it stops being consistent with the offline basis. The thresholding of $\bar{y}^t$ guarantees that, notwithstanding the unrevealed offline basis, that stopping time is large: under BudgetRatio the process R^t spends most of its time in the action region (and subset thereof) where it performs *basic allocations*, those that are consistent with the offline basis. To establish this we must (i) develop a *generalizable mathematical description* of Fig. 1 and (ii) study the dynamics of the stochastic process R^t inside and between the action regions in this figure.

While the main contribution is mathematical, the geometric view advances the understanding of practical aspects of BudgetRatio, specifically:

BudgetRatio as a bid-price control. Bid-price heuristics are popular due to their intuitive interpretation: a request should be accepted if its reward exceeds the opportunity cost of the resources it consumes.

In online packing, the standard bid-price algorithm solves the packing LP and accepts a request if its reward exceeds the sum of the shadow prices of all resources it consumes. This is a popular and widely used policy, yet it does not achieve constant regret Jasin and Kumar [2013].

To achieve constant regret, our bid-price version of BudgetRatio is more careful: the bid-price is obtained from a maximum over several shadow prices—the collection of these is determined by the problem's geometry. The generalized bid-prices can be computed adpatively and efficiently.

Robustness to parameter misspecification. Our geometric analysis uncovers the sensitivity to errors in the forecasting of the demand and/or the rewards. We study the case where the true parameters (rewards and probabilities) are (v, p, ϱ), but the algorithm is run with $(\tilde{v}, \tilde{p}, \tilde{\varrho}) \neq (v, p, \varrho)$. We quantify how accurate $(\tilde{v}, \tilde{p}, \tilde{\varrho})$ must be for BudgetRatio to achieve constant regret, despite being executed with incorrect parameters.

We introduce an appealingly simple notion of centroids (see Section 3). As long as the misspecification leaves these centroids unchanged, the collection of action regions in Fig. 1 is stable under perturbations of the parameters. In the one-dimensional case (i.e., with a single resource), $\tilde{v}$ must be accurate enough to deduce the ranking of the requests Vera et al. [2021]. The centroids provide a generalization of the inherently one-dimensional notion of ranking, allowing us to understand the multidimensional problem. These robustness guarantees subsequently yield optimal regret guarantees in the setting where the demand and reward parameters are not apriori known to the controller and must be learned.

The impact of restock on regret. In the baseline setting of online packing (or network revenue management), inventory is not restocked; only the initial inventory is available to the controller.

Generally, restock poses a real challenge: the offline upper bound is too ambitious and constant regret is not attainable. Our geometric view of the problem affords a nuanced consideration of restock. We prove that, under an explicitly identifiable "slow restock" condition, constant regret is attainable in this generally difficult problem, and is achieved by BudgetRatio with suitably tuned thresholds.

2. Model and overview of results A decision maker allocates resources to requests over T periods. There is a set of resources $\mathcal{R} = [d] = \{1, \ldots, d\}$ and, at time $t = 0$, there is an initial inventory I_i^0 for resource $i \in \mathcal{R}$. Additionally, at each time $t \in [T]$, a unit of resource i arrives with probability ϱ_i independently of the past; ϱ denotes the vector of these arrival probabilities and satisfies $\sum_{i\in\mathcal{R}} \varrho_i \leq 1$ (not all resources restock). At most one unit of resource arrives each period. We let $\mathfrak{Z}^t = (\mathfrak{Z}_i^t : i \in \mathcal{R})$ be the accumulated restock over the time interval $[1, t]$. The controller cannot consume more than $I_i^0 + \mathfrak{Z}_i^t$ units of resource i by time t.

There is a set $\mathcal{J} = [n] = \{1, \ldots, n\}$ of possible requests, a request of type $j \in \mathcal{J}$ generates a reward v_j and consumes resources as encoded in a matrix $A \in \{0,1\}^{d\times n}$, where $A_{ij} = 1$ means that type j requires one unit of resource i. At time $t \in [T]$, a request j arrives with probability p_j independently of the past; p denotes the vector of these arrival probabilities and satisfies $\sum_{j\in\mathcal{J}} p_j = 1$. Exactly one request arrives each period. We let $Z^t = (Z_j^t : j \in \mathcal{J})$ be the accumulated arrivals over $[1, t]$. The controller cannot accept more than Z_j^t requests of type j by time t.

We let V^t be the reward brought by the request arriving at time t; the random variables $V^1, V^2, \ldots, V^T$ are assumed to be i.i.d with $\mathbb{P}[V^t = v_j] = p_j$, $j \in \mathcal{J}$. The inventory on hand at time $t \in [T]$ is denoted by $I^t = (I_1^t, \ldots, I_d^t)'$.

The selection process at time t unfolds as follows:

(i) Inventory restock: the inventory is updated to include newly arriving resource units; i.e., $I_i^t \leftarrow I_i^t + 1$ if a resource i arrives.

(ii) Request acceptance and inventory reduction: If the arrival is of type j (i.e., $V^t = v_j$), then the request must be rejected if $I^t \ngeq A_j$. If the request is feasible ($A_j \leq I^t$), then it may be accepted, thereby generating a reward v_j and decreasing the inventory to $I^t - A_j$; or it may be rejected generating zero reward.

Resources do not expire: if not used by time t, they are available at $t + 1$. The decision to accept/reject a request is final: if a type-j request is accepted, reward is collected and the relevant resources are consumed; it if it rejected it is lost forever (requests do not queue).

No online policy can do better than the offline, full information, counterpart in which all rewards are presented in advance. Allowing this offline to use fractional allocations gives a further upper bound. This fractional offline controller is our benchmark; its expected total reward is given by

$$V^*_{\text{off}}(T, I^0) := \mathbb{E}\begin{bmatrix} \max & v'y \\ \text{s.t.} & Ay \leq I^0 + \mathfrak{Z}^T \\ & y \leq Z^T \\ & y \in \mathbb{R}^n_{\geq 0} \end{bmatrix}. \tag{2}$$

Throughout we assume, without loss of generality, that $I_i \leq T$ for all $i \in \mathcal{R}$. If $I_i > T$, resource i is non-binding and we can reduce the problem to one with $d - 1$ resources.

As a pre-processing step, we perturb the rewards: for every $j \in \mathcal{J}$, we take the rewards to be randomly perturbed rewards $v_j \leftarrow v_j + \mathcal{U}_j$ where $\mathcal{U}_j(0, 1/T)$ are n i.i.d uniform $[0, 1/T]$ random variables. Over an horizon of length T, this perturbation introduces at most a $\mathcal{O}(1)$ error. It guarantees that, almost surely, the optimal solutions are uniquely defined at each period of the Algorithm and in the offline problem; see e.g. [Bertsimas and Tsitsiklis, 1997, Exercise 3.15].

2.1. The Primal BudgetRatio Algorithm. The budget is the inventory on hand plus *expected* future restock; the budget *ratio* is the size of the budget relative to the residual horizon.

Definition 1 (Budget Ratio). The budget ratio at time $t \in [1, T]$ is

$$R^t := \frac{1}{T-t}(I^t + \mathbb{E}[\mathfrak{Z}^T - \mathfrak{Z}^t]) = \frac{1}{T-t} I^t + \varrho,$$

where I^t is the inventory on hand at time t. The ratio at $t = 0$ is defined by the random variables (without expectation) $R^0 := \frac{1}{T}(I^0 + \mathfrak{Z}^T)$. The demand at time $t = 0$ is defined by $D^0 := \frac{1}{T} Z^T$.

Define

$$\begin{aligned} \mathrm{LP}(R, D) \qquad & \max \ v'x \\ & \text{s.t.} \quad Ay \le R, \\ & \qquad\quad y \le D, \\ & \qquad\quad y \in \mathbb{R}^n_{\ge 0}. \end{aligned} \tag{3}$$

BudgetRatio re-solves a deterministic relaxation of (2) and thresholds its solution to make acceptance/rejection decisions; see Algorithm 1.

Algorithm 1 Budget Ratio Policy

Input: Aggressiveness parameter $\alpha \in (0, 1)$

Set thresholds: for $j \in \mathcal{J}$, let $\gamma_j := \max_{i: A_{ij} = 1} \varrho_i$ and set $\bar{p}_j \leftarrow p_j + \gamma_j \mathbb{1}_{\{\gamma_j > \alpha p_j\}}$.
for $t = 1, \ldots, T$ **do**
  If a resource $i \in \mathcal{R}$ arrived, $I^t_i \leftarrow I^t_i + 1$.
  Set $R^t \leftarrow \frac{1}{T-t} I^t + \varrho$.
  Solve $\mathrm{LP}(R^t, p)$ to obtain the optimal decision variables $\bar{y}^t$.
  Set j as the type of the arriving request.
  if $I^t \not\ge A_j$ (not feasible to serve j) or $\bar{y}^t_j < \alpha \bar{p}_j$ (no desirable to serve j): reject the request
  else if $\bar{y}^t_j \ge \alpha \bar{p}_j$: accept the request $I^t \leftarrow I^t - A_j$
  Carry over the inventory for the next period: $I^{t+1} \leftarrow I^t$.

LP notation and terminology. Let $\bar{A}$ be the following augmentation of the resource consumption matrix A, where I_n is the identity matrix of dimension $n \times n$

$$\bar{A} = \begin{bmatrix} A & 0 & I_d \\ I_n & I_n & 0 \end{bmatrix}. \tag{4}$$

For any $R \in \mathbb{R}^d_{\geq 0}$ and $D \in \mathbb{R}^n_{\geq 0}$, we re-write the LP relaxation, in standard form, as

$$\max\left\{ v'y : \bar{A}\begin{pmatrix} y \\ u \\ s \end{pmatrix} = \begin{pmatrix} R \\ D \end{pmatrix},\ (y,u,s) \geq 0 \right\}, \qquad \text{(LP}(R,D))$$

where $(y,u,s) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^d$ is the decision vector. The variables $y \in \mathbb{R}^n$ represent the amount of requests accepted, $u \in \mathbb{R}^n$ correspond to the number of unmet (i.e., arrived but not accepted) requests, and $s \in \mathbb{R}^d$ stand for resource surplus. We refer to these henceforth as the *request*, *unmet*, and *surplus* variables.

We use the general notation $\mathcal{B}$ to denote a basis of (LP(R,D)) as well as the $(d+n)\times(d+n)$ sub-matrix of $\bar{A}$ corresponding to the variables in the basis $\mathcal{B}$; $\mathcal{B}^c$ denotes the non-basic columns. Let $\bar{v} = (v,0,0) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^d$ be the extended reward vector, where we assign zero value to the slack variables u and s.

For an optimal basis $\mathcal{B}$ we refer to $\lambda = \lambda(\mathcal{B}, v) = (\mathcal{B}^{-1})'\bar{v}_{\mathcal{B}}$ as the dual variables associated with $\mathcal{B}$.

A useful example. For visualization purposes, we present in detail a two-dimensional example ($d = 2$ resources) that is rich enough to demonstrate key characteristics, yet simple enough to afford a visual representation of the problem's geometry. The example is a traditional packing problem with no restock ($\varrho = 0$).

We denote resources and their initial inventory by a, b and I_a, I_b respectively. There are four customer types $\{1,2,3,4\}$ with the consumption matrix A, reward values, and arrival probabilities

$$A \quad = \quad \begin{array}{c|cccc} & 1 & 2 & 3 & 4 \\ \hline a & 1 & 0 & 1 & 1 \\ b & 0 & 1 & 1 & 1 \end{array}, \quad v = (4,4,5,1), \text{ and } p = (1/4, 1/4, 1/4, 1/4).$$

Type-3 requests bring the highest reward ($v_3 = 5$) but consume both resources a and b. Types 1 and 2 have the highest per-resource-consumption reward. Type-4 consumes both resources but brings little rewards; it is the least desirable. For future reference, we label this as the ***base example***.

The geometry of BudgetRatio. Each point in the colored map in Fig. 1 corresponds to a two-dimensional budget-ratio (R_a, R_b). When solving LP(R,p), we obtain the set of request types $\mathcal{K} = \{j : \bar{y}_j \geq \alpha \bar{p}_j\}$ that BudgetRatio accepts at that inventory level; no other types are accepted. The action region

$$\mathcal{N}_{\mathcal{K}} = \{R \in \mathbb{R}^2_+ : \bar{y}_j \geq \alpha \bar{p}_j, \text{ for } j \in \mathcal{K}, \bar{y}_j < \alpha \bar{p}_j, \text{ for } j \in \mathcal{K}^c\}.$$

is the set of budget ratios R where BudgetRatio accepts exclusively requests from types in the *centroid set* $\mathcal{K}$; each color on the plot corresponds to a different centroid set; the light-blue region, for instance, is the set $\mathcal{N}_{\{1,2\}}$.

The red circle in $\mathcal{N}_{\mathcal{K}}$ represents the *centroid budget*. It is where the budget equals the resource consumption of those request types in $\mathcal{K}$:

$$\sum_{j\in\mathcal{K}} A_{ij}p_j =: r_{\mathcal{K}}. \qquad \text{(Centroid Budget)}$$

For the centroid $\mathcal{K}=\{1,3\}$, the budget is the vector $r_{\{1,3\}}=(0.5,0.25)'$ because $0.5=p_1+p_3$ (both request types consume resource a) and $0.25=p_3$ (only type 3 requires resource b). The centroid budgets anchor the geometry of the action regions.

The LP at a centroid budget, $\mathrm{LP}(r_{\mathcal{K}},p)$, has multiple optimal bases $\mathcal{B}$; these are *the bases associated with the centroid* $\mathcal{K}$. With each of these we have the dual variable $\lambda(\mathcal{B},v)=(\mathcal{B}^{-1})'\bar{v}_{\mathcal{B}}$. This informal description of the geometry suffices for the presentation of our results; formal definitions appear in §3 and §4.

2.2. The max-bid-price BudgetRatio. We present a generalization of bid-price policies that, we will prove, achieves constant regret.

We define the (set of) dual prices at a centroid $\mathcal{K}$ as follows

$$\begin{aligned}&\Lambda_{\mathcal{K}} := \{\lambda : \lambda=\lambda(\mathcal{B},v) \text{ for some optimal basis } \mathcal{B} \text{ associated to } \mathcal{K}\},\\ &\Lambda(R)=\Lambda_{\mathcal{K}} \text{ if } R\in\mathcal{N}_{\mathcal{K}}.\end{aligned} \qquad (5)$$

The map $\Lambda(\cdot)$ identifies which bid-prices are relevant for the budget R. Having identified the centroid $\mathcal{K}$ such that $R\in\mathcal{N}_{\mathcal{K}}$, the set of bid-prices are those associated to its corresponding centroid $\mathcal{K}$.

When close to the origin $R=0$, primal BudgetRatio rejects *all* requests even if $R>0$. To mimic this boundary behavior, we must introduce—through the bid prices—a high shadow price near the boundary of the state space. The centroid $\mathcal{K}$ is near a boundary for type j if $\sum_{l\in\mathcal{K}} A_l \ngeq A_j$, we then write $j\in\partial(\mathcal{K})$. For the centroid $\mathcal{K}=\emptyset$ all types are near the boundary: $\partial(\emptyset)=\mathcal{J}$. Define

$$\lambda^{\partial}(R)=2\sum_j v_j \mathbf{e}_j \mathbb{1}_{\{j\in\partial(\mathcal{K})\}}, \text{ if } R\in\mathcal{N}_{\mathcal{K}},$$

where $\mathbf{e}_j$ is the vector of size $d+n$ that has 1 in entry $d+j$ and 0 elsewhere. We note that $\lambda^{\partial}(R)=0$ if $R\in\mathcal{N}_{\mathcal{K}}$ and $\mathcal{K}$ is such that $r_{\mathcal{K}}>0$; in the base example this is the case for all but $\mathcal{K}=\emptyset,\mathcal{K}=\{1\}$ and $\mathcal{K}=\{2\}$. Below $\bar{A}$ is the augmentation of A as in Eq. (4).

Definition 2 (a max-bid price definition of BudgetRatio). An arriving request of type j is accepted at time t if $I^t\geq A_j$ (there are enough resources), and v_j exceeds the *max-bid price*: $v_j\geq \max_{\lambda\in\Lambda(R^t)} \bar{A}_j'(\lambda+\lambda^{\partial}(R^t))$.

In our base example, the dual vectors at the centroid $\mathcal{K}=\{1\}$ (bottom yellow region in Fig. 1) has the dual vectors $\lambda(\mathcal{B}_1)=(4,4,0,0,0,0)'$, $\lambda(\mathcal{B}_2)=(1,4,3,0,0,0)'$, and $\lambda(\mathcal{B}_3)=(0,5,4,0,0,0)'$ (one vector for each of the optimal bases at the centroid's budget) so that for $R^t\in\mathcal{N}_{\{1\}}$, the decision is to accept a type-j arrival if $v_j\geq\max\{\lambda(\mathcal{B}_1)'\bar{A}_j,\lambda(\mathcal{B}_2)'\bar{A}_j,(\mathcal{B}_3)'\bar{A}_j\}$. Type 1 is accepted here because $\lambda(\mathcal{B}_1)'\bar{A}_1=\lambda(\mathcal{B}_2)'\bar{A}_1=\lambda(\mathcal{B}_3)'\bar{A}_1=4\leq v_1$; type 2 is not accepted because $\lambda(\mathcal{B}_3)'\bar{A}_2=5\geq 4=v_2$ (even without including λ^{∂}).

In Theorem 1, we state an equivalence between the two formulations of BudgetRatio: the one (based on the primal) in Algorithm 1 and the other (based on the dual) in Definition 2. We prove that both algorithms take precisely the same actions at all times: BudgetRatio accepts an arriving request of type j at time t (and facing ratio R^t) if and only if the max-bid price control does so at this time and state.

This bid-price formulation of BudgetRatio could, in some instances, be computationally faster than Algorithm 1. To pre-compute the full map $\Lambda(\cdot)$, we must solve at most $(n+1)!$ parcking LPs where, n, recall, is the number of types; see Remark 4. Algorithm 1, in contrast, requires solving (in real time) T such LPs, one for each period in the horizon. In our base example, there are 4 types so that bid-price BudgetRatio is computationally preferable for $T\gg 24$. Moreover, the map is computed only once and can be subsequently used for multiple runs of the online phase, as is often done in large-scale networks; see Bast et al. [2016]. But pre-computing the full map is not necessary for the bid-price version of BudgetRatio. Instead, bid-prices can be generated adaptively and relatively efficiently; see Remark 7.

2.3. Main results We impose the following requirement throughout.

Assumption 1 **(Slow Restock)**. *For every centroid $\mathcal{K}$ and every resource i used by some $j\in\mathcal{K}$ ($\sum_{j\in\mathcal{K}}A_{ij}\geq 1$) we have $\varrho_i<(r_{\mathcal{K}})_i=\sum_{j\in\mathcal{K}}A_{ij}p_j$.*

The requirement is that a resource restocks at a lower rate than the rate consumed by the centroid set. It is trivially satisfied in the traditional online packing setting where there is initial inventory but no restock ($\varrho=0$); see further discussion of this assumption in Remark 1.

We define two requirements on the primitives (p,ϱ,v,A) that are used to parametrize our robustness statements. They are not needed for constant regret.

Definition 3 (δ-Complementarity). Let $\mathcal{B}$ be a basis (associated with some centroid $\mathcal{K}$) and $\lambda=\lambda(\mathcal{B},v)$ be the dual variables associated to $(\mathcal{B},v)$. We say that $\mathcal{B}$ is δ-complementary if (i) $\lambda_i\geq\delta$ for all resource i whose surplus s_i is not in $\mathcal{B}$, (ii) $\lambda_j\geq\delta$ for all request type j whose slack u_i is not in $\mathcal{B}$, and (iii) $(\bar{A}'\lambda)_j\geq v_j+\delta$ for all request type j whose request variable y_j is not in $\mathcal{B}$.

Our notion of δ-complementarity is a strengthening of the standard notion of complementary slackness in linear programming; the latter is recovered by setting $\delta = 0$ in our definition. Parametrizing strict complementarity by $\delta > 0$ allows us to relate the problem's primitives to allowed perturbation/misspecification of the reward vector v.

Similarly to δ-complementarity, δ-separation will parametrize allowed misspecification of the arrival-probability vector p.

Definition 4 (centroid separation). We say that the centroids are δ-separated if $\min_{\mathcal{K}\neq\mathcal{K}'}\min_{i\in\mathcal{R}}|(r_{\mathcal{K}}(p)-r_{\mathcal{K}'}(p))_i|\geq\delta$.

In the one dimensional case ($d=1$), Definitions 3 and 4 reduce to explicit insightful requirements; see Corollary 1.

Theorem 1 **(constant regret and its robustness)**. *Suppose that slow restock holds. Then,*

1. **Constant Regret:** BudgetRatio *(primal) achieves a uniformly bounded regret: There exists a constant M such that*

$$V^*_{\mathit{off}}(T,I^0)-V_{on}(T,I^0)\leq M, \tag{6}$$

   where $V_{on}(T,I^0)$ is the total reward of BudgetRatio*. The constant M may depend on (p,ϱ,v,A), but not on the horizon T or the initial inventory I^0.*

2. **Robustness with respect to reward:** *The regret remains constant if* BudgetRatio *uses an estimate $\tilde{v}$ of v, as long as*

$$||v-\tilde{v}||_\infty\leq\frac{\delta}{c(d+2)}, \tag{7}$$

   where δ is such that all bases are δ-complementary, and $c\leq\max\{||\mathcal{B}^{-1}||_\infty:\mathcal{B}$ basis $\}$.

3. **Robustness with respect to arrival probabilities:** *The regret remains similarly constant if* BudgetRatio *uses an estimate $(\widetilde{p},\widetilde{\varrho})$ of (p,ϱ), as long as $(\widetilde{p},\widetilde{\varrho})$ satisfy slow-restock and*

$$\max_{\mathcal{K}}||r_{\mathcal{K}}(p)-r_{\mathcal{K}}(\widetilde{p})||_\infty\leq\frac{\delta}{4}, \tag{8}$$

   where δ is such that all centroid budgets are δ-separated.

4. **Max-bid price Equivalence:** *If all bases are δ-complementary for some $\delta>0$, then the primal and max-bid-price definitions of* BudgetRatio *are equivalent: on any realization of $Z,\mathfrak{Z}$ and at any time t,* BudgetRatio *as specified in Algorithm 1 accepts an arriving request of type j, if and only if the max-bid price algorithm in Definition 2 does.*

In the 2-dimensional base example Fig. 1(RIGHT), the sup norm distance between any two red circles (centroid budgets) equals $1/4$, hence centroid separation (Definition 4) is satisfied with $\delta = 1/4$. δ-complementarity (Definition 3) is satisfied with $\delta = 1$ and $\max\{||\mathcal{B}^{-1}||_\infty : \mathcal{B}\text{ basis }\}\leq 1$; this we found through computational discovery of all the optimal bases. Therefore, Eq. (7) specializes to

$||v - \tilde{v}||_\infty \leq 1/4$. It is important that c, δ depend only on (v, A) and not on p, ϱ or the horizon T. The requirement Eq. (8) on $\widetilde{p}$ imposes 8 constraints, one per centroid.

In the one-dimensional case (7) and (8) simplify to intuitive requirements.

Corollary 1 **(separation conditions for a single resource)**. *With $d = 1$, the centroids are δ-separated, in the sense of Definition 4, with $\delta = \min_j p_j$. Eq. (8) reduces to*

$$\left| \sum_{k \in [j]} (p_k - \widetilde{p}_k) \right| \leq \frac{\min_k p_k}{4} = \frac{\delta}{4} \quad \forall j \in [n], \tag{9}$$

which is, in particular, satisfied if $||p - \widetilde{p}||_\infty \leq \frac{\min_k p_k}{4n}$. The rewards v satisfy δ-complementarity, in the sense of Definition 3, if

$$v_j \geq \delta, \text{ for all } j \in [n], \text{ and } |v_j - v_{j'}| \geq \delta, \text{ for all } j \neq j'. \tag{10}$$

Equation (10) recovers the reward separation requirement in [Vera et al., 2021, Theorem 4]. We conclude this section with a discussion of slow restock and of BudgetRatio parameters.

Remark 1 (slow restock). **Conceptual implications:** In allowing restock in our model, we explore the limits of constant regret and simple re-solving algorithms. The slow-restock requirement in Assumption 1 draws such a limit explicitly: if the condition is met, constant regret is attainable and is achieved by a suitably modified version of BudgetRatio.

Lemma 2 illuminates how the slow restock assumption facilitates the workings of BudgetRatio. If the restock rate is large, much of the *forecasted* inventory at a time t is embedded in future arrivals. This means that although we might want to accept a request at time t, we might not be able to because there is no inventory on hand. With high-restock rates, the system behaves more like a loss-queue than an inventory allocation problem—see further discussion in Section 8.

Assumption 1 can be weakened somewhat: if—in Fig. 1—the initial budget lies in $\mathcal{N}_{\{1,2,3\}}$ (the red action region), it suffices to satisfy slow restock for the centroid $\{1,2,3\}$ and its immediate neighbors $\{1,2\}$, $\{1,3\}$, $\{2,3\}$, and $\{1,2,3,4\}$.

If the parameters satisfy the complete exact of Assumption 1—$\varrho_i > \max_{\mathcal{K}} (r_{\mathcal{K}})_i$ for all resource i and centroids $\mathcal{K}$—the problem trivializes: there is so much capacity that constant regret is achieved by admitting all arriving requests as long as there is available inventory.

The problematic cases are those where $\varrho_i < (r_{\mathcal{K}})_i$ for some centroids but $\varrho_i > (r_{\mathcal{K}})_i$ for others; or those where some of these are held with equality. In Appendix Section A, we provide examples, where a regret proportional to $\sqrt{T}$ is unavoidable.

Algorithmic implications: The rate is accounted for in the definition of $\bar{p}_j$ where γ_j captures the restock of resources used by j; $\bar{p} = p$ if there is no restock. The greater the restock rate, the more opportunities BudgetRatio has to serve type-j requests in the future. Because the resources that j consumes replenish, it is less critical to accept j in the immediate present. The increase in the threshold by γ_j renders BudgetRatio more conservative in accepting j. On the other hand, for a request with low restock of the resources it consumes (hence small γ_j), BudgetRatio might as well accept it now as the opportunities to do so will not increase in the future. ■

Remark 2 (The Aggressiveness Parameter α). The parameter α can be set to any value in $(0,1)$. The closer that the value of α is to 1 (while still away from 1) the greater the restock rate that is allowed without compromising constant regret. Intuitively speaking, as α approaches 1, the algorithm becomes more conservative in accepting requests; it slows down to allow for inventory to accumulate. ■

Final setup details. Let $\mathcal{F}_0$ denote the trivial σ-field and, for $t\in[T]$, let $\mathcal{F}_t=\sigma\{(Z^\tau,\mathfrak{Z}^\tau):\tau=1,\dots,t\}$ be the σ-field generated by the random arrivals of resources and requests. An online policy π can be expressed with binary random variables $(\sigma_j^{\pi,t}: j\in\mathcal{J})$ such that $\sigma_j^{\pi,t}=1$ means that a type-j request is accepted at time t. For adapted online policies, $\sigma^{\pi,t}$ must be $\mathcal{F}_t$-measurable. Let

$$Y_j^{\pi,t}:=\sum_{\tau\in[t]}\sigma_j^{\pi,\tau},$$

be the total number of type-j requests accepted by the policy π over $[1,t]$. A policy is feasible if (1) the total consumption of resource i does not exceed its initial inventory I_i^0 plus its total restock, and (2) the total acceptance does not exceed arrivals:

$$\begin{aligned} AY^{\pi,t} &\le I^0+\mathfrak{Z}^t,\ t\in[T], \\ Y^{\pi,t} &\le Z^t,\ t\in[T], \\ \sigma_j^{\pi,t} &\le \mathbb{1}_{\{V^t=v_j\}},\ t\in[T], j\in\mathcal{J}. \end{aligned} \tag{11}$$

Let Π be the set of feasible online policies, those that are $\mathcal{F}_t$-adapted and satisfying (11). The total reward of an online policy $\pi\in\Pi$ is

$$V^\pi(T,I^0)=\mathbb{E}\left[\sum_{t\in[T]}v'\sigma^{\pi,t}\right].$$

For each (T,I^0), the goal of the decision maker is to maximize the expected value:

$$V^*(T,I^0)=\max_{\pi\in\Pi}V^\pi(T,I^0).$$

To prove optimality guarantees, we compare $V^\pi(T,I^0)$ (with $\pi\leftarrow$ BudgetRatio) against the offline benchmark $V^*_{\text{off}}(T,I^0)$ in (2).

Additional notation. Given a subset $\mathcal{K} \subseteq \mathcal{J}$ we let $A_{\mathcal{K}}$ be the submatrix of A that has only columns in the index set $\mathcal{K}$ (but has all rows). We similarly define sub-vectors: if x is a column vector, $x_{\mathcal{K}}$ is a subvector with the indices in the set $\mathcal{K}$. For real vectors x, y of the same dimension and $\epsilon > 0$, we write $x = y \pm \epsilon$ if $\|x - y\| \leq \epsilon$. Throughout $d(x, y)$ is the Euclidean distance between two points $x, y \in \mathbb{R}^d$. For a subset $\mathcal{C} \subseteq \mathbb{R}^d$ and a point $x \in \mathbb{R}^d$, $d(x, \mathcal{C}) = \inf_{y \in \mathcal{C}} d(x, y)$ is the equclidean distance of x from the set $\mathcal{C}$. We similarly define $d_\infty(x, y)$ and $d_\infty(x, \mathcal{C})$ for the sup-norm. For an integer $m \geq 1$, $[m] = \{1, \ldots, m\}$. We adopt the convention that the maximum over the empty set is zero and the minimum is ∞ ($\max \varnothing = 0, \min \varnothing = \infty$). We use throughout M to be a constant —that can depend on (A, p, ϱ, v), but it is independent of (T, I^0)— whose value can change from one line to the next.

2.4. Related Work

Online packing (network revenue management). The attainability of constant regret has been already established for some online allocation. Arlotto and Gurvich [2019], the first to establish constant regret (regardless of whether the deterministic LP is degenerate or not), takes a geometric stochastic-process view, but it is specific to the one-dimensional (i.e., single resource) case. The geometric analysis of the multidimensional case requires the introduction of generalizable mathematical constructs (centroids, bases, cones, etc). More recently, Vera and Banerjee [2021], Vera et al. [2021] study a large family of resource allocation problems that includes also dynamic posted pricing.

Relative to this earlier work, the geometric view has explanatory power insofar as it provides an alternative and mathematically appealing support for constant regret that is grounded in linear programming and, specifically, in a parametric view of the packing LP. This view provides language through which we can explicitly identify the robustness and flexibility of BudgetRatio.

Bid-price heuristics. Bid-price heuristics are popular due to their intuitive interpretation; see Talluri and Van Ryzin [1998] for asymptotic results and Boyd and Bilegan [2003] for a broader overview of bid-prices. In a setting with multiple resources, a bid-price policy is described as follows: at time t, compute a vector $\lambda^t \in \mathbb{R}^d_{\geq 0}$ of resource prices and reject a type-j arrival if and only if its reward is below the combined price of requested resources, i.e., if and only if $v_j < A_j' \lambda^t$. The standard bid price heuristic sets λ^t to be the dual vector (or "shadow price") of an LP solved at time t. It is known that these bid prices cannot achieve constant regret Jasin and Kumar [2013]. To the best of our knowledge, the strongest available guarantee for bid-price policies is $O(\sqrt{T})$ regret Talluri and Van Ryzin [1998]. We show that BudgetRatio can be interpreted as a bid-price control (Talluri and Van Ryzin [1998] and [Talluri and Van Ryzin, 2004, Chapter 3.2]), albeit a more elaborate one. Our generalized version of bid-price—which we call *max-bid price*—achieves constant regret.

Robustness to parameter misspecification and bandits. In Theorem 1 we identify sufficient conditions on the misspecification of parameters (probabilities p, ϱ and rewards v) under which constant regret persists. When these conditions are met BudgetRatio produces constant regret even if executed under wrong parameters.

Our results have implications for learning and acting in resource allocation. The premise is that the reward of type-j requests is random *with expectation* v_j and neither v nor the arrival probability vectors p, ϱ are initially known to the controller. The empirical frequency of different request types provides the controller with an estimate of p, and the accepted requests allow the controller to estimate v; see Bubeck and Cesa-Bianchi [2012] for more on bandit problems.

Bandit problems known as *bandits with knapsacks* Badanidiyuru et al. [2013] explicitly model budget constraints which, in our setting, correspond to the limited inventory. In *contextual bandits* Agrawal and Devanur [2016], arrivals present a "context" before the controller makes decisions.

The general-purpose results in the literature Agrawal and Devanur [2016], Badanidiyuru et al. [2013] imply $O(\sqrt{T})$ regret bounds for our setting. For our model, where the context is the type $j \in [n]$, we identify the separation condition in Definition 4 that guarantees an optimal regret scaling of $O(\log T)$. This separation condition relies on our notion of centroids: to make good accept/reject decisions, we must learn enough about the primitives to identify the type of instance, i.e., the important centroids. Centroids bring out a natural multi-dimensional notion of separation that is consistent with, yet generalizes, the $\mathcal{O}(\log T)$ regret and the separation condition for the one-dimensional (single resource) case in Vera et al. [2021], Wu et al. [2015].

Two-sided arrivals and assembly. Arrivals of inventory capture assembly networks with fixed production rates. In assembly models, orders arrive to be assembled by using relevant components; see Song and Zipkin [2003] as well as Plambeck and Ward [2006] which gives an asymptotically optimal policy for holding cost minimization under a high demand assumption. We focus on finite-time non-asymptotic guarantees for reward maximization. Our contribution to this literature is in identifying conditions on the restock rate that, when met, render the offline upper bound attainable, and achievable by a simple resolving algorithm that we explicitly construct.

Paremetric linear programming. The objective in this literature is to understand how optimization problems change as the primitives change; see Gal [1984] for a survey. We study the parametric behavior of the packing LP when multiple parameters are perturbed simultaneously. This is in the spirit of multiparametric linear programming Bemporad et al. [2002], Borrelli et al.

[2003] where the parametric analysis is used in support of model predictive control. Our analysis of BudgetRatio requires the characterization of the geometry of the problem. This is made feasible by the special structure of the packing LP.

Drift analysis. Much of our analysis centers on the dynamics of the process R^t. We argue that, when close to the boundary of an action region, the ratio process R^t drifts towards and then "sticks" to this boundary. Such Lyapunov/drift arguments are frequently used in the analysis of stochastic models to establish positive recurrence of Markov processes. In the context of online control, there are similarities to queueing theory where max-weight policies—based on re-solving local optimization problems—lead to the attraction to a subset of the state space; see Eryilmaz and Srikant [2012], Maguluri and Srikant [2016].

3. Overview of our approach.

An online policy builds, in an adapted manner, an approximate solution for a random linear system whose right-hand side is revealed only at the end of the horizon—the offline linear system. The offline optimal decision maker waits until the end of the horizon to solve its LP while the online policy must commit to solutions in a dynamic fashion. Below we make this precise.

Offline representation. Introducing slack variables, we rewrite the constraints of the offline LP (2), $\{Ay \le I^0 + \mathfrak{Z}^T, y \le Z^T\}$ in standard form $\{Ay + s = I^0 + \mathfrak{Z}^T, y + u = Z^T\}$, where $s \in \mathbb{R}^d_{\ge 0}$ is the *surplus* of resource and $u \in \mathbb{R}^n_{\ge 0}$ is the *unmet* demand. Augmenting the matrix A to $\bar{A}$ in Eq. (4), we arrive at the standard form representation of offline's value

$$V^*_{\mathrm{off}}(T, I^0) = \mathbb{E}\left[\max\left\{v'y : \bar{A}\begin{pmatrix} y \\ u \\ s \end{pmatrix} = C\right\}\right], \quad \text{where} \quad C := \begin{pmatrix} I^0 + \mathfrak{Z}^T \\ Z^T \end{pmatrix}. \tag{12}$$

The random vector $C \in \mathbb{R}^{d+n}_{\ge 0}$ is the *maximum consumption* of offline. Given a basis $\mathcal{B}$ (columns of $\bar{A}$) for the LP in Eq. (12), the optimal solution satisfies $\mathcal{B}x_{\mathcal{B}} = C$, where $x = (y, u, s)$ stands for all the variables and $\mathcal{B} = \mathcal{B}(C)$ depends on the right-hand side. The realized (random) value of offline can be written as

$$\sum_{\mathcal{B}} v'_{\mathcal{B}} y_{\mathcal{B}} \mathbb{1}_{\{\mathcal{B} \text{ is optimal}\}} = \sum_{\mathcal{B}} v'_{\mathcal{B}} \mathcal{B}^{-1} C \mathbb{1}_{\{\mathcal{B} \text{ is optimal}\}}. \tag{13}$$

Recall that we use $\mathcal{B}$ for both the indices of basic columns and the sub matrix $\bar{A}_{\mathcal{B}}$.

Online construction of the offline linear system. If the optimal offline basis is $\mathcal{B}$, offline's actions correspond to the unique solution of the system $\mathcal{B}x_{\mathcal{B}} = C$, where $x = (y, u, s)$. The quality of the online approximation to the static offline system depends on how long—out of the total horizon of length T—the policy π takes actions that are consistent with the optimal offline basis $\mathcal{B}$. This consistency is captured in the following definition.

Definition 5 (Basic Allocation). Let π be an online policy and $\mathcal{B}$ be the optimal offline basis (revealed at time T). We say that π performs *basic allocation* at $t \in [T]$ if it only serves requests j such that $y_j \in \mathcal{B}$ (request variable for type j is basic) and it only rejects arriving requests such that $u_j \in \mathcal{B}$ (unmet variable for type j is basic).

Above we refer to *the* basis of the offline problem but there could be multiple optimal bases.[1] In that case, for a given basis $\mathcal{B}$ we would say that π performs a $\mathcal{B}$-*basic allocation* at t. We continue referring to *the* optimal basis for the offline problem on the understanding that the statements apply to any optimal basis if multiple exist.

As long as the policy π performs basic allocations, it is "operating" in an optimal basis. Ifτ^π is the first time where π performs a non-basic allocation, regret is incurred only and at most in the remaining $T - \tau^\pi$ periods; see Proposition 1. That regret, in turn, depends on the amount of resource that remains unused by the online policy relative to the offline solution.

Definition 6 (Wastage). Let π be any online policy and $\mathcal{B}$ the optimal offline basis. Let S_i^t be the surplus of resource $i \in [d]$ at time t when using the policy π, i.e., $S^t = I^0 + \mathfrak{Z}^t - AY^{\pi,t}$. The wastage of π at t is $W^{\pi,t} := \max\{S_i^t : \text{surplus variable } s_i \text{ is non basic}\} = \max\{S_i^t : s_i \notin \mathcal{B}, i \in [d]\}$.

Intuitively, if $s_i \notin \mathcal{B}$, then resource i has no slack: it is completely utilized in the offline solution. The wastage captures the inventory left unused by the online policy that should have been used in its entirety. The quality of the online system, i.e., the approximation to $\mathcal{B}x_{\mathcal{B}} = C$ is determined by this time τ^π and the wastage it induces.

Proposition 1 **(a stopping-time regret criterion)**. *Let $\mathcal{B}$ be the optimal basis for the offline problem (12) and denote $J^t \in \mathcal{J}$ the type of the t-th request. For any online policy π define the time*

$$
\begin{aligned}
\tau^\pi &:= \min\{t \le T : \textit{the policy does not perform a basic allocation at } t\} - 1 \\
&= \min\{t \le T : (\sigma_j^{\pi,t} = 1 \textit{ and } y_j \notin \mathcal{B} \textit{ for some } j) \textit{ or } (\sigma_j^{\pi,t} = 0 \textit{ and } u_j \notin \mathcal{B} \textit{ where } j = J^t)\} - 1.
\end{aligned}
$$

Then for any $\tau \le \tau^\pi$ a.s. the expected regret of π is at most $M\mathbb{E}[T - \tau + W^{\pi,\tau}]$, where M is a constant independent of (T, I^0), but that may depend on (A, v), and $W^{\pi,t}$ is the wastage at time t. In particular, the regret is $\mathcal{O}(1)$ if $\mathbb{E}[T - \tau + W^{\pi,\tau}] = \mathcal{O}(1)$.

Proof. Throughout the proof, the policy π is fixed and omitted from notation. Let Y_j^t, U_j^t be the number of type-j requests accepted and rejected (unmet) by the online policy over the interval

[1] Our perturbation of the rewards guarantees that offline has a unique optimal solution but this solution might be degenerate. Indeed, at centroid budgets, this optimal solution is degenerate.

$[1,t]$. Let $C^t := \begin{pmatrix} I^0 + \mathfrak{Z}^t \\ Z^t \end{pmatrix}$ be the maximal feasible consumption in $[1,t]$, and recall that the surplus is $S^t := I^0 + \mathfrak{Z}^t - AY^t \in \mathbb{R}^d_{\geq 0}$. By definition,

$$\bar{A}X^t = C^t, \text{ where } X^t = (Y^t, S^t, U^t).$$

Let us divide the matrix $\bar{A}$ into basic and non-basic columns as $\bar{A} = [\mathcal{B}, \mathcal{B}^c]$. We note that,

$$\mathcal{B}\begin{pmatrix} Y^t \\ U^t \\ S^t \end{pmatrix}_{\mathcal{B}} + \mathcal{B}^c \begin{pmatrix} 0 \\ 0 \\ S^t \end{pmatrix}_{\mathcal{B}^c} = C^t \quad \text{and} \quad C - C^t = \begin{pmatrix} \mathfrak{Z}^T \\ Z^T \end{pmatrix} - \begin{pmatrix} \mathfrak{Z}^t \\ Z^t \end{pmatrix}, \quad \forall t \leq \tau. \tag{14}$$

The first equation follows from the decomposition $\bar{A} = [\mathcal{B}, \mathcal{B}^c]$ and the fact that, up to time τ, the policy performs basic allocations so that the only non-zero variables Y^t_j, U^t_j are those in the basis $\mathcal{B}$. The second equation follows from the definition of C and C^t. Recall that the offline variables $x_{\mathcal{B}} = (y,u,s)_{\mathcal{B}} = \mathcal{B}^{-1}C$ are the solution to the offline system. Using (14) we then have

$$\begin{aligned} \begin{pmatrix} y \\ u \\ s \end{pmatrix}_{\mathcal{B}} - \begin{pmatrix} Y^t \\ U^t \\ S^t \end{pmatrix}_{\mathcal{B}} &= \mathcal{B}^{-1}C - \mathcal{B}^{-1}\left(C^t - \mathcal{B}^c \begin{pmatrix} 0 \\ 0 \\ S^t \end{pmatrix}_{\mathcal{B}^c}\right) \\ &= \mathcal{B}^{-1}\left(\begin{pmatrix} \mathfrak{Z}^T \\ Z^T \end{pmatrix} - \begin{pmatrix} \mathfrak{Z}^t \\ Z^t \end{pmatrix}\right) + \mathcal{B}^{-1}\mathcal{B}^c \begin{pmatrix} 0 \\ 0 \\ S^t \end{pmatrix}_{\mathcal{B}^c} \qquad \forall t \leq \tau. \end{aligned} \tag{15}$$

The process Y is increasing and non-negative: $Y^T \geq Y^t \geq 0$ for all $t \in [T]$. Consequently,

$$\text{Regret} = (v'_{\mathcal{B}} y_{\mathcal{B}} - v'Y^T) \leq (v'_{\mathcal{B}} y_{\mathcal{B}} - v'Y^t) \leq v'_{\mathcal{B}}(y_{\mathcal{B}} - Y^t_{\mathcal{B}}).$$

We bound the last expression using Eq. (15): since there is at most one arrival per period, $||(\mathfrak{Z}^T, Z^T)' - (\mathfrak{Z}^t, Z^t)'||_\infty \leq T - t$, and the surplus is bounded by definition as $||S^t_{\mathcal{B}^c}||_\infty = W^t$. Finally, we take the worst case over $\mathcal{B}$ in Eq. (15) and conclude the result by setting $t = \tau$. □

To prove item 1. of Theorem 1 (constant regret), it suffices now to find a random time $\tau \leq \tau^\pi$ a.s. and prove that $\mathbb{E}[T - \tau + W^{\pi,\tau}] = \mathcal{O}(1)$. Accordingly, the remainder of our analysis is dedicated to, identifying τ, and then bounding $T - \tau$ and the wastage W for $\pi =$ BudgetRatio.

Analysis overview via the one-dimensional case. Let us consider in some detail the one-dimensional packing problem, also known as the multi-secretary problem [Arlotto and Gurvich, 2019]. There are I^0 positions to be filled and candidates arrive one at a time with abilities (rewards) $V^1, \ldots, V^T$; the goal is to maximize total accumulated reward by selecting at most I^0 candidates.

In our notation, $d = |\mathcal{R}| = 1$ (single resource), $\varrho = 0$ (no-restock so that $\mathbb{E}[R^0] = R^0 = \frac{1}{T}I^0$), and $A = \mathbf{e}'$ (each request consumes one unit of the resource). The deterministic relaxation has $n+1$ constraints, one for each of the demand constraints, and a single budget constraint:

$$\begin{aligned} \text{LP}(R,p) \quad & \max \ v'y \\ & \text{s.t.} \quad \mathbf{e}'y \leq \mathbb{E}[R^0], \\ & \qquad\quad y \leq p, \\ & \qquad\quad y \geq 0. \end{aligned} \tag{16}$$

We assume without loss of generality that types are labelled in decreasing order of rewards, i.e., $v_1 > v_2 > \ldots > v_n$, and let $\bar{F}_i \coloneqq \sum_{j=1}^{i} p_j$ be the survival function at v_i. The deterministic relaxation in Eq. (16) has a simple greedy solution: in increasing order of k, set $\bar{y}_k = p_k$ as long as $\bar{F}_k \leq \mathbb{E}[R^0]$. Letting $i_0 = \max\{k : \bar{F}_k \leq R\}$; finally, set $\bar{y}_{i_0+1} = \mathbb{E}[R^0] - \bar{F}_{i_0}$.

Centroids. If the budget ratio is exactly $\mathbb{E}[R^0] = \bar{F}_j$ (at a jump point of the distribution), the deterministic relaxation (16) takes all types $\mathcal{K} = [j]$, and only those types. In other words, for this choice of right-hand side (budget), the problem $\mathrm{LP}(\bar{F}_j, p)$ has all variables $y_1, \ldots, y_j$ saturated and all other variables equal to zero. The sets $\mathcal{K}$ with this property are *centroids*. The set $\mathcal{K} = [j]$ is optimal when the budget is exactly $r_{\mathcal{K}} = \bar{F}_j$, so we refer to $r_{\mathcal{K}}$ as the *centroid's budget*; see Fig. 2.

The centroids *do not depend on* p. Regardless of the distribution, the LP "takes" all requests $[j]$ before taking any request of type $j+1$. Both the deterministic relaxation $\mathrm{LP}(\mathbb{E}[R^0], p)$ and the offline problem $\mathrm{LP}(R^0, D^0)$—where, we recall, $D^0 = \frac{1}{T} Z^T$—follow the same nested rule. This concept generalizes in multiple dimensions: there are sets of requests $\mathcal{K} \subseteq \mathcal{J}$ that are always prioritized in a subset of the space independent of the demand p. Centroids elicit a useful summary of the matrix A and the reward vector v; see Definition 7.

Action regions: the centroid neighborhood. The thresholding of the algorithm—accepting type j requests only if $\bar{y}_j \geq \frac{1}{2} p_j$—creates a confidence interval (a neighborhood) around the centroid's budget. The *neighborhood* of the centroid $\{1,2\}$ is the interval $\mathcal{N}_{\{1,2\}}(p) = [\bar{F}_2 - \frac{p_2}{2}, \bar{F}_2 + \frac{p_3}{2}]$ centered at exactly the centroid budget $r_{[2]}(p) = \bar{F}_2 = p_1 + p_2$; see Fig. 2. As long as $R^t = I^t/(T-t)$ is in this interval, BudgetRatio accepts only (and all) arriving requests of types $\{1,2\}$; it starts accepting type-3 requests when R^t exceeds the upper threshold $\bar{F}_2 + p_3/2$. It rejects type-2 requests if it goes below the lower threshold $\bar{F}_2 - p_2/2$.

Basic convex subsets. By Proposition 1, for an online policy to be good it must have "almost" oracle access to the offline basis $\mathcal{B}$; its decisions must be consistent with the (apriori unknown) basis $\mathcal{B}$ for much of the horizon.

Consider the centroid $[2] = \{1,2\}$, its budget $r_{[2]}(p) = \bar{F}(v_3) = \bar{F}_2$ and the action region $\mathcal{N}_{[2]}(p) = [\bar{F}_2 - p_2/2, \bar{F}_2 + p_3/2]$: when $R^t \in \mathcal{N}_{[2]}(p)$, BudgetRatio accepts only, and all, arriving requests of types 1 and 2. This action region has two convex subsets, each associated with a specific basis. The basis $\mathcal{B}_2$ that has—in addition to types 1, 2—the request variable y_3 for type 3 is optimal on the set $\mathcal{N}_{[2]}(\mathcal{B}_2, p) = [\bar{F}_2, \bar{F}_2 + p_3/2]$. The basis $\mathcal{B}_1$ that has—in addition to the unmet (slack) variables for types $j > 2$—also the unmet variable u_2 for type 2, is optimal on $\mathcal{N}_{[2]}(\mathcal{B}_1, p) = [\bar{F}_2 - p_2/2, \bar{F}_2]$. When in the proximity of $\mathcal{N}_{[2]}(\mathcal{B}_2, p)$, i.e., on the set $\mathcal{N}^{\epsilon}_{[2]}(\mathcal{B}_2, p) = \{R : d(R, \mathcal{N}_2(\mathcal{B}, p)) \leq \epsilon\}$ (see Fig. 2), BudgetRatio accepts all of type 1, 2 requests and, if R^t crosses into the neighboring centroid [3], also type 3 requests. Similarly, when in the set $\mathcal{N}^{\epsilon}_{[2]}(\mathcal{B}_1, p)$, BudgetRatio accepts all of type-1 requests and, if R^t crosses into the neighboring centroid [1], stops accepting type 2 requests.

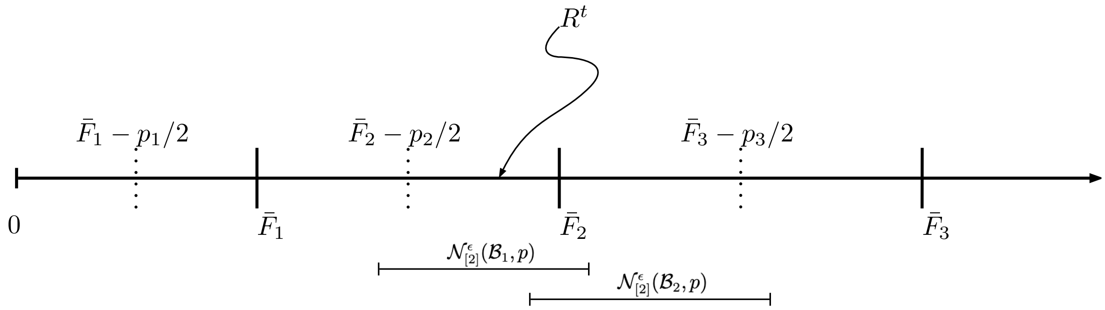

FIGURE 2. The position of the ratio R^t with respect to the centroid budgets $r_{[j]}(p) = \bar{F}_j$ determines the actions of the policy. At time t, the policy accepts a type-j request if and only if $R^t \geq \bar{F}_j - p_j/2$. Oracle containment guarantees that if the realization Z^T is such that offline accepts only types $[2] = \{1, 2\}$, then $R^t \in \mathcal{N}^{\varepsilon}_{[2]}(\mathcal{B}_1, p)$ with high probability for most of the horizon. If, instead, Z^T is such that offline accepts also type-3 requests, then $R^t \in \mathcal{N}^{\varepsilon}_{[2]}(\mathcal{B}_2, p)$ with high probability for most of the horizon. In conclusion, R^t evolves in the "correct" region $\mathcal{N}^{\varepsilon}_{[2]}(\mathcal{B}_1)$ or $\mathcal{N}^{\varepsilon}_{[2]}(\mathcal{B}_2, p)$, this guarantees that the policy accepts only requests in the optimal offline basis.

Oracle containment. By the same arguments as above, but with the probability distribution p replaced by the random realization D^0, offline selects basis $\mathcal{B}_1$ if $R^0 \in \mathcal{N}_{[2]}(\mathcal{B}_1, D^0)$ and accepts only requests of types $\{1, 2\}$. We will prove that BUDGETRATIO—despite not knowing the optimal offline basis—keeps $R^t \in \mathcal{N}^{\varepsilon}_{[2]}(\mathcal{B}_1, p)$ for much of the horizon. Here we have p instead of D^0: if offline has $R^0 \in \mathcal{N}_{[2]}(\mathcal{B}_1, D^0)$ then BUDGETRATIO—acting adaptively in real time—keeps R^t in the proximity of the corresponding "theoretical" set $\mathcal{N}^{\varepsilon}_{[2]}(\mathcal{B}_1, p)$.

As long as $R^t \in \mathcal{N}^{\varepsilon}_{[2]}(\mathcal{B}_1, p)$, recall, it accepts only requests $\{1, 2\}$ thus performing basic allocations. If, instead, $R^0 \in \mathcal{N}_{[2]}(\mathcal{B}_2, D^0)$ offline selects basis $\mathcal{B}_2$ and accepts type 3. BUDGETRATIO then keeps $R^t \in \mathcal{N}^{\varepsilon}_{[2]}(\mathcal{B}_2, p)$ for much of the horizon.

The overall implication is that BUDGETRATIO performs basic allocations for much of the horizon and, in turn, that τ^π is large. Finally, $R^t = \frac{1}{T-t} I^t \in \mathcal{N}^{\varepsilon}_{[2]}(\mathcal{B}_1, p) \cup \mathcal{N}^{\varepsilon}_{[2]}(\mathcal{B}_2, p) \leq \bar{F}_3$ implies $I^t \leq \bar{F}_3(T - t)$: little inventory (hence little wastage) remains at the end of the horizon. With τ^π large and wastage small, Proposition 1 yields constant regret.

Our analysis consists, then, of two steps: (i) mapping budgets R (a point in the space of budget ratios) to optimal bases of (LP(R, D)); and (ii) showing that, under BUDGETRATIO, R^t remains in the basic subset which is consistent with optimal (unknown to it) offline basis.

The mapping from budgets R to bases (step (i) above) is straightforward in the one dimensional case: On the "right" of the centroid $\mathcal{K} = [2]$ is its *centroid neighbor* $\mathcal{K} = [3]$—request-type 3 is added to the centroid—and the optimal basis $\mathcal{B}_2$ is the one where the request variable y_3 is in the basis. On the left is the neighbor $\mathcal{K} = [1]$—so 2 leaves the centroid—and the optimal basis $\mathcal{B}_1$ has the unmet variable u_2. In this way, there is correspondence between the bases that are optimal at a centroid's budget and the neighbors of the centroid. The fact that BUDGETRATIO drives R^t into the correct basic subset (step (ii) above) is non-trivial already in the one dimensional case; see Arlotto and Gurvich [2019].

In the remainder of the paper we introduce the infrastructure to execute on both of these steps in the multidimensional problem.

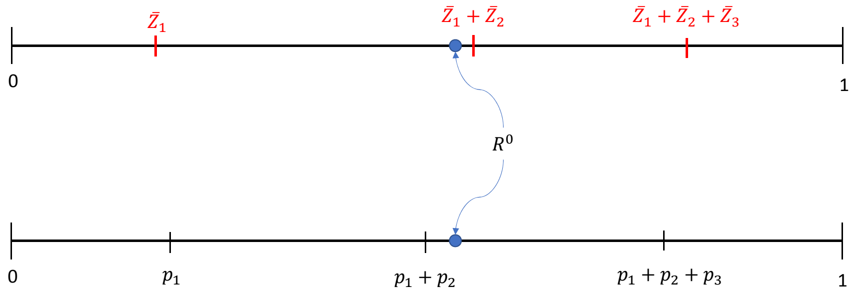

FIGURE 3. Why randomized policies do not maintain basic allocations. An illustration via the one-dimensional ($d = 1$) case. Although offline accepts no type 3 requests because $R^0 \leq \bar{Z}_1^T + \bar{Z}_2^T$ (TOP), the online randomized algorithm accepts type 3 with some probability (BOTTOM).

REMARK 3 (ON RANDOMIZED POLICIES AND BID-PRICE CONTROLS). Our definition of basic allocation (Definition 5) and its associated guarantee in Proposition 1, underscore the relationship between regret and the extent to which an online policy performs allocations that are consistent with the offline basis $\mathcal{B}$. Randomized policies [Jasin and Kumar, 2012] do not satisfy this consistency. Consider the multi-secretary problem and the scenario captured in Fig. 3. With R^0 close enough to $p_1 + p_2$, we can have with non-neglible probability both $\bar{Z}_1^T + \bar{Z}_2^T = \frac{1}{T}Z_1^T + \frac{1}{T}Z_2^T > R^0$ and $p_1 + p_2 < R^0$. In this realization, offline takes all of type-1 and most of type-2 requests but none of the type-3 requests. The standard randomized policy solves $LP(R^t, p)$ and accepts a request of type j with probability y_j/p_j. In this scenario, it accepts at time $t = 1$ an arriving type 3 with probability $(R^0 - (p_1 + p_2))/p_3$ and will continue accepting type-3 requests until $R^t \leq p_1 + p_2$, thus performing multiple *non*-basic allocations. Under this randomized policy, and with this initial budget R^0, the budget fluctuates around $p_1 + p_2$ and performs non-basic allocations (too) frequently. BUDGETRATIO, in contrast, introduces a confidence interval: it does accept type-3 requests unless $R^0 \geq p_1 + p_2 + p_3/2$.

This execution of non-basic allocations is also the shortcoming of the standard bid-price control. Here, a request is accepted if its reward v_j exceeds the sum of shadow prices (shadow price = the dual variables of the resource constraint) of requested resources. In this example, because $R^0 \in (p_1 + p_2, p_1 + p_2 + p_3)$, the shadow price of the (single) resource is v_3 so that type-3 requests are accepted. At the centroid's budget $p_1 + p_2$, there are two optimal dual solutions: in one the shadow price of the capacity constraints is v_3 and in the other it is $v_2 > v_3$. The max-bid equivalent of BudgetRatio in Definition 2 accepts only types 1 and 2 but not type 3. ■

4. Parametric structure of the packing problem. The geometric structure and the stochastic analysis that builds on it (convex subsets, basic cones, etc.) are relatively simple in the one-dimensional case. Formalizing general notions of centroids and action regions requires a parametric analysis of the packing LP.

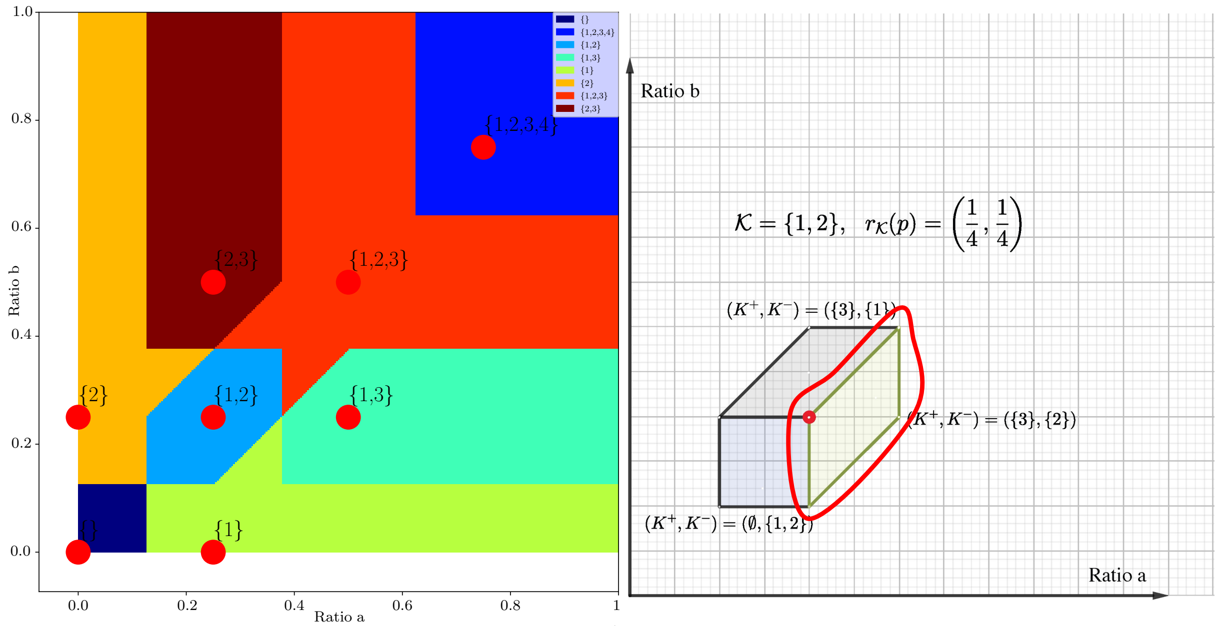

FIGURE 4. The action regions and convex subsets in the base example

The centroid neighborhood $\mathcal{N}_{\{1,2\}}$ in the base example (see Fig. 4(LEFT)) has three *basic convex subsets* as seen on the right of that figure, each of which corresponds to one optimal basis. As long as R^t is in the yellow convex set, a given basis—let us call it $\mathcal{B}^{Yellow}$—is optimal for $LP(R^t, p)$. This basis is fully characterized by its centroid $\{1,2\}$ and the three centroid neighbors $\{1\}, \{1,3\}$ and $\{1,2,3\}$; **see Lemma 4.**

We show that as long as the budget ratio R^t stays "close" to this basic convex subset—e.g. within the red-lined encirclement of the yellow subset—it performs actions consistent with its underlying basis; **see Proposition 2**. This is the multidimensional generalization of the sets $\mathcal{N}^{\epsilon}_{[2]}(\mathcal{B}_1)$ and $\mathcal{N}^{\epsilon}_{[2]}(\mathcal{B}_2)$ in the one-dimensional case; recall Fig. 2. We further show that if $\mathcal{B}^{Yellow}$ is optimal for *offline* then, indeed, R^t remains close to the yellow set under the *online* policy BudgetRatio; **see Proposition 3.**

Below is how this roadmap is divided into sections:

- In §4.1 we define, in optimization terms, the centroid regions and their basic convex subsets. Specifically, we map the the location of the budget ratio R, which features in the right-hand side of (LP(R, D)), to the optimal bases of this optimization problem. We prove that as long as R^t is in the proximity of the basic convex subset corresponding to the offline optimal basis, BudgetRatio is performing basic allocations. In other words, the "escape time" from the convex subset is a lower bound on τ^π in our regret criterion Proposition 1; **see Proposition 2**. Regret would then be small if τ^π is close to the end of the horizon time T.; **See Proposition 3.**
- §4.2 and §5 are dedicated to proving Proposition 2 and 3. To show that the escape time from the basic convex subset is indeed large we must have the language to study movement of R^t as it is driven by BudgetRatio. In §4.2 we characterize the geometry of the basic convex subsets. We show that a basic convex subset is the intersection of a centroid neighborhood and a suitable cone (Lemma 5) that can we characterize in significant detail. This section ends with the proof of Proposition 2.
- With the geometry mapped, §5 is where we analyze the stochastic movement of R^t in space and prove that it stays in proximity of the correct basic convex subset and hence performs actions that are consistent with offline's optimal basis. This includes (1) a sticky boundary property (Theorem 2), that shows that the residual budget process remains close to one centroid neighborhood/action region; and (2) a cone-containment property (Theorem 3) that stipulates that the BudgetRatio controlled budget process remains constrained to the correct basic cone. Combined, Theorem 2 and 3 are the key ingredients in the proof of Proposition 3.
- Items 2. and 3. of Theorem 1 (parameter robustness) are proved in §6. Item 4 (max-bid price control) is proved §7.

4.1. Action regions, exit times and constant regret For ease of exposition, we strengthen Assumption 1 and require that $\varrho_i < \frac{1}{2}(r_{\mathcal{K}})_i$, instead of $\varrho_i < (r_{\mathcal{K}})_i$. This allows us to take the tuning parameter to be $\alpha = \frac{1}{2}$ throughout this section. The analysis remains the same as long as $\alpha \in (0,1)$ is such that $\varrho_i < \alpha(r_{\mathcal{K}})_i$; such $\alpha \in (0,1)$ exists by Assumption 1.

Recall the augmented matrix $\bar{A}$ in (4) and the standard form (LP(R, D)) introduced in Section 2. Our first result focuses attention on a subset of relevant bases; no other bases must be considered.

Lemma 1. *Fix a basis $\mathcal{B}$ and let $\lambda = (\mathcal{B}^{-1})'\bar{v}_{\mathcal{B}}$ be the dual variables associated to $\mathcal{B}$. If (i) $\lambda \geq 0$ and (ii) $\bar{A}'\lambda \geq \bar{v}$, then $\mathcal{B}$ is optimal for* (LP(R, D)) *if $\mathcal{B}^{-1}\begin{pmatrix} R \\ D \end{pmatrix} \geq 0$. Conversely, for any right-hand side (R, D), there is an optimal basis that satisfies (i) and (ii).*

All lemmas are proved in Appendix B.

For a set $\mathcal{K} \subseteq \mathcal{J}$ and a demand vector $D \in \mathbb{R}^n$, the *action region* for $\mathcal{K}$ is the set of ratios $\mathcal{N}_{\mathcal{K}}(D) \subseteq \mathbb{R}^d$ where the algorithm serves *exclusively* requests in $\mathcal{K}$, i.e., all requests $j \in \mathcal{K}$ are accepted and $j \notin \mathcal{K}$ are rejected:

$$\begin{aligned}\mathcal{N}_{\mathcal{K}}(D) &:= \left\{R \in \mathbb{R}^d : \text{ BudgetRatio serves exclusively requests } \mathcal{K} \text{ when } (R^t, p) = (R, D)\right\} \\ &= \bigcup_{\mathcal{B}} \left\{ R \in \mathbb{R}^d : \mathcal{B} \text{ optimal}, \quad y_{\mathcal{K}} = \left(\mathcal{B}^{-1}\begin{pmatrix} R \\ D \end{pmatrix}\right)_{\mathcal{K}} \geq \frac{1}{2} D_{\mathcal{K}}, \quad y_{\mathcal{K}^c} = \left(\mathcal{B}^{-1}\begin{pmatrix} R \\ D \end{pmatrix}\right)_{\mathcal{K}^c} < \frac{1}{2} D_{\mathcal{K}^c} \right\}. \end{aligned} \tag{17}$$

The equality holds because the algorithm serves a request j if and only if $y_j \geq D_j/2$. We use this definition with D taking two possible values: $D = p$ and $D = D^0 = \frac{1}{T} Z^T$. The set $\mathcal{N}_{\mathcal{K}}(D)$ might be empty for some $\mathcal{K} \subseteq \mathcal{J}$ (the algorithm never "prioritizes" the set $\mathcal{K}$ of request types).

Henceforth we use D as a placeholder where the constructs are relevant for both offline and BudgetRatio. As a general rule, D appears for geometric constructions (relevant for both online and offline) while p appears in statements concerning the stochastic process R^t driven by BudgetRatio.

It is in the following lemma, *and only here*, where the slow restock Assumption 1 is used. The assumption guarantees that enough of the total inventory (on-hand plus future restock) is on-hand so that BudgetRatio can accept a type $j \in \mathcal{K}$ request when $R^t \in \mathcal{N}_{\mathcal{K}}(p)$.

Lemma 2. *For $\mathcal{K} \subseteq \mathcal{J}$, BudgetRatio serves exclusively requests in $\mathcal{K}$ if and only if $R^t \in \mathcal{N}_{\mathcal{K}}(p)$. Furthermore, for a constant M that depends only on (A, p, ϱ), whenever $t \leq T - M$ and $R^t \in \mathcal{N}_{\mathcal{K}}(p)$, there is enough inventory to serve any request $j \in \mathcal{K}$, i.e., $I^t \geq A_j$ for all $j \in \mathcal{K}$.*

In view of (17), we can write $\mathcal{N}_{\mathcal{K}}(D) = \cup_{\mathcal{B}} \mathcal{N}_{\mathcal{K}}(D, \mathcal{B})$ where $\mathcal{N}_{\mathcal{K}}(D, \mathcal{B}) \subseteq \mathcal{N}_{\mathcal{K}}(D)$ is the set of ratios where the algorithm serves exclusively requests in $\mathcal{K}$ and the optimal basis for LP(R, D) is $\mathcal{B}$.

$$\begin{aligned}\mathcal{N}_{\mathcal{K}}(D, \mathcal{B}) &:= \left\{R \in \mathbb{R}^d : \text{ BudgetRatio uses } \mathcal{B} \text{ and serves exclusively } \mathcal{K} \text{ when } (R^t, p) = (R, D)\right\} \\ &= \left\{ R \in \mathbb{R}^d : \mathcal{B} \text{ optimal}, \quad y_{\mathcal{K}} = \left(\mathcal{B}^{-1}\begin{pmatrix} R \\ D \end{pmatrix}\right)_{\mathcal{K}} \geq \frac{1}{2} D_{\mathcal{K}}, \quad y_{\mathcal{K}^c} = \left(\mathcal{B}^{-1}\begin{pmatrix} R \\ D \end{pmatrix}\right)_{\mathcal{K}^c} < \frac{1}{2} D_{\mathcal{K}^c} \right\}, \end{aligned} \tag{18}$$

The next result states that the sets $\mathcal{N}_{\mathcal{K}}(D, \mathcal{B})$ are the correct resolution to study the problem. The time of escape from these sets lower bounds τ^π which, per Proposition 1, controls the regret.

Proposition 2. *Let $\mathcal{B}$ be the optimal offline basis and set $\mathcal{K} \subseteq \mathcal{J}$. Given $\epsilon > 0$, define*

$$\tau^{\epsilon, \mathcal{K}} := \min\{t \leq T : d_\infty(R^t, \mathcal{N}_{\mathcal{K}}(p, \mathcal{B})) > \epsilon\}. \tag{19}$$

Then, there exists a choice of $\epsilon > 0$ such that $\tau^{\epsilon, \mathcal{K}} \leq \tau^\pi$, where $\pi =$ BudgetRatio is as spelled out in Algorithm 1 and, as defined in Proposition 1, $\tau^\pi + 1$ is the first time that π does not perform a basic allocation.

Proposition 2 immediately implies that, as long as R^t is close to $\mathcal{N}_{\mathcal{K}}(p,\mathcal{B})$, BudgetRatio is performing basic allocations. The next proposition further guarantees that R^t remains close to $\mathcal{N}_{\mathcal{K}}(D,\mathcal{B})$ for much of the horizon. Recall that $\mathbb{E}[R^0]=\frac{1}{T}I^0+\varrho$ and $\mathbb{E}[D^0]=p$, hence at time $t=0$ we can identify the set $\mathcal{K}$ such that $\mathbb{E}[R^0]\in\mathcal{N}_{\mathcal{K}}(\mathbb{E}[D^0])$; it is obtained the first time we solve the deterministic relaxation.

Proposition 3. *Let $\mathcal{K}$ be such that $\mathbb{E}[R^0]\in\mathcal{N}_{\mathcal{K}}(\mathbb{E}[D^0])=\mathcal{N}_{\mathcal{K}}(p)$, $\epsilon>0$ and $\tau^{\epsilon,\mathcal{K}}$ be as in Proposition 2. Then, there is a constant M such that $\mathbb{E}[T-\tau^{\epsilon,\mathcal{K}}+W^{\tau^{\epsilon,\mathcal{K}}}]\leq M$, where W^t is the wastage at time t (see Definition 6).*

We now have the ingredients to prove part 1. of Theorem 1.

Proof of Theorem 1 (Regret bound). By Proposition 2 we have that $\tau^{\epsilon,\mathcal{K}}\leq\tau^{\pi}$, hence the policy performs only basic allocations over the interval $[1,\tau^{\epsilon,\mathcal{K}}]$. By Proposition 3, the expected wastage and remaining time $T-\tau^{\epsilon,\mathcal{K}}$ are bounded by a constant and so is, by Proposition 1, the regret. □

It remains to prove Propositions 2 and 3. The former is proved at the end of §4.2 and the latter in §5.

4.2. The geometric characterization of action regions: Centroids and basic cones.

Definition 7 (Centroids). A subset $\mathcal{K}\subseteq\mathcal{J}$ is a centroid if, for some $D\in\mathbb{R}^n_{>0}$, there exists a solution (y,u,s) to $\mathrm{LP}(A_{\mathcal{K}}D,D)$ such that $u_{\mathcal{K}}=0$ (no request in $\mathcal{K}$ is unmet) and $y_{\mathcal{K}^c}=0$ (no request in $\mathcal{K}^c$ is accepted). For a centroid $\mathcal{K}$, $r_{\mathcal{K}}(D):=A_{\mathcal{K}}D_{\mathcal{K}}$ is the *centroid budget*.

Intuitively, a set $\mathcal{K}$ is a centroid if, given the exact budget required in expectation for all requests $\mathcal{K}$—this is $r_{\mathcal{K}}(p)=A_{\mathcal{K}}p_{\mathcal{K}}$—it is optimal in the deterministic relaxation to accept all requests $\mathcal{K}$ and no others. In the one-dimensional setting of Section 3, the centroids are the sets $[j]$ for $j=1,2,\ldots,n$ and their corresponding budgets are $r_{\mathcal{K}}(p)=r_{[j]}(p)=\bar{F}(v_j)$. Fig. 1 has (in red) the centroid budgets $r_{\mathcal{K}}(p)$ for our two-dimensional base example.

The optimization problem Eq. (LP(R,D)) has multiple optimal bases at $R=r_{\mathcal{K}}(D)=A_{\mathcal{K}}D_{\mathcal{K}}$ and they are all degenerate: the solution (y,u,s) at $r_{\mathcal{K}}(D)$ is, per Definition 7, $y_{\mathcal{K}}=D_{\mathcal{K}}$, $u_{\mathcal{K}^c}=D_{\mathcal{K}^c}$ with all other variables equal to zero. Because $\mathcal{K}\cup\mathcal{K}^c=\mathcal{J}$, only n of the basic variables are strictly positive, whereas the dimension of the right-hand side is $n+d$; there must then be d zero-valued basic variables.

Definition 8 (Zero Valued Basic Variables). Fix a centroid $\mathcal{K}$ for some $\hat{D}$ as in Definition 7 and let $\mathcal{B}$ be a basis that is optimal at $r_{\mathcal{K}}(\hat{D})$, i.e., optimal for $\mathrm{LP}(r_{\mathcal{K}}(\hat{D}),\hat{D})$, with (y,u,s) the associated solution. Define the sets of basic variables

$$K^+:=\{j\in\mathcal{J}:y_j\in\mathcal{B},y_j=0\},\quad K^-:=\{j\in\mathcal{J}:u_j\in\mathcal{B},u_j=0\},\quad K^0:=\{i\in\mathcal{R}:s_i\in\mathcal{B},s_i=0\}.$$

We sometimes write $K^+(\mathcal{B}), K^-(\mathcal{B}), K^0(\mathcal{B})$ to make explicit the dependence on the basis $\mathcal{B}$.

The characterization of centroids, bases and zero-valued variables associated with them does not depend on the demand distribution D, but only on the matrix A and the rewards v. In particular, $\mathcal{K}$ is a centroid under both the theoretical distribution $(D = p)$ and the empirical distribution $(D^0 = \frac{1}{T} Z^T)$.

LEMMA 3. *Let $\mathcal{K}$ be a centroid for some $\hat{D} \in \mathbb{R}^n_{>0}$ as in Definition 7. Then, the same property holds for any $\tilde{D} \in \mathbb{R}^n_{>0}$, i.e.,* $\mathrm{LP}(A_\mathcal{K} \tilde{D}_\mathcal{K}, \tilde{D})$ *has the solution $u_\mathcal{K} = 0$ and $y_{\mathcal{K}^c} = 0$. Similarly, the bases and the sets of zero-valued basic variables in Definition 8 are the same under $\hat{D}$ and $\tilde{D}$.*

Finally, we define a useful relation between centroids.

DEFINITION 9 (NEIGHBORS). Let $\mathcal{K}$ be a centroid. If the basis $\mathcal{B}$ is optimal at the centroid budget $r_\mathcal{K}(D)$, we say that $\mathcal{B}$ is associated to $\mathcal{K}$. Another centroid $\mathcal{K}'$ is a neighbor of $\mathcal{K}$ if there is a basis $\mathcal{B}$ that is associated to both $\mathcal{K}$ and $\mathcal{K}'$.

Like the centroids themselves, the relation of "neighbor" does not depend on the demand distribution D. Once we fix $\mathcal{K}$ and an associated basis $\mathcal{B}$, we can obtain neighbors of $\mathcal{K}$ based on the zero-valued basic variables; see Definition 8. In Lemma 4 we prove that $\mathcal{N}_\mathcal{K}(D, \mathcal{B})$, which determines the exit time of interest in Proposition 2, can be characterized in terms of the focal centroid $\mathcal{K}$ and its neighbors. The characterization facilitates the analysis of the exit time in Proposition 2.

LEMMA 4 **(characterization of $\mathcal{N}_\mathcal{K}(D, \mathcal{B})$ and neighbors)**. *Fix a centroid $\mathcal{K}$ with associated basis $\mathcal{B}$. Let (K^+, K^-, K^0) be the zero-valued basic variables (Definition 8). Then,*

1. *The basis $\mathcal{B}$ is optimal for any right-hand side (R, D) of the form*

$$R = r_\mathcal{K}(D) + \alpha(A_{\kappa^+} D_{\kappa^+} - A_{\kappa^-} D_{\kappa^-}) + b,$$

where $\kappa^+ \subseteq K^+, \kappa^- \subseteq K^-, \alpha \in [0,1]$, and $b \in \mathbb{R}^d_{\geq 0}$ is zero for components not in K^0, i.e., $b_i = 0$ for $i \notin K^0$. In particular, the set $\mathcal{K} \cup \kappa^+ \backslash \kappa^-$ is a centroid and a neighbor of $\mathcal{K}$.

2. *The basis $\mathcal{B}$ is optimal for (R, D) if and only if R is of the form*

$$R = r_\mathcal{K}(D) + \sum_{\kappa^+ \subseteq K^+, \kappa^- \subseteq K^-} \alpha_{(\kappa^+, \kappa^-)} (A_{\kappa^+} D_{\kappa^+} - A_{\kappa^-} D_{\kappa^-}) + b, \tag{20}$$

where b is as before, $\alpha \geq 0$, and $\sum_{\kappa^+ \subseteq K^+, \kappa^- \subseteq K^-} \alpha_{(\kappa^+, \kappa^-)} = 1$.

3. *$R \in \mathcal{N}_\mathcal{K}(D, \mathcal{B})$ if and only if*

$$R - r_\mathcal{K}(D) = A_{K^+} x_{K^+} - A_{K^-} x_{K^-} + b,$$

where $x_j \in [0, D_j/2]$ for $j \in K^+ \cup K^-$ and b is as in item 1.

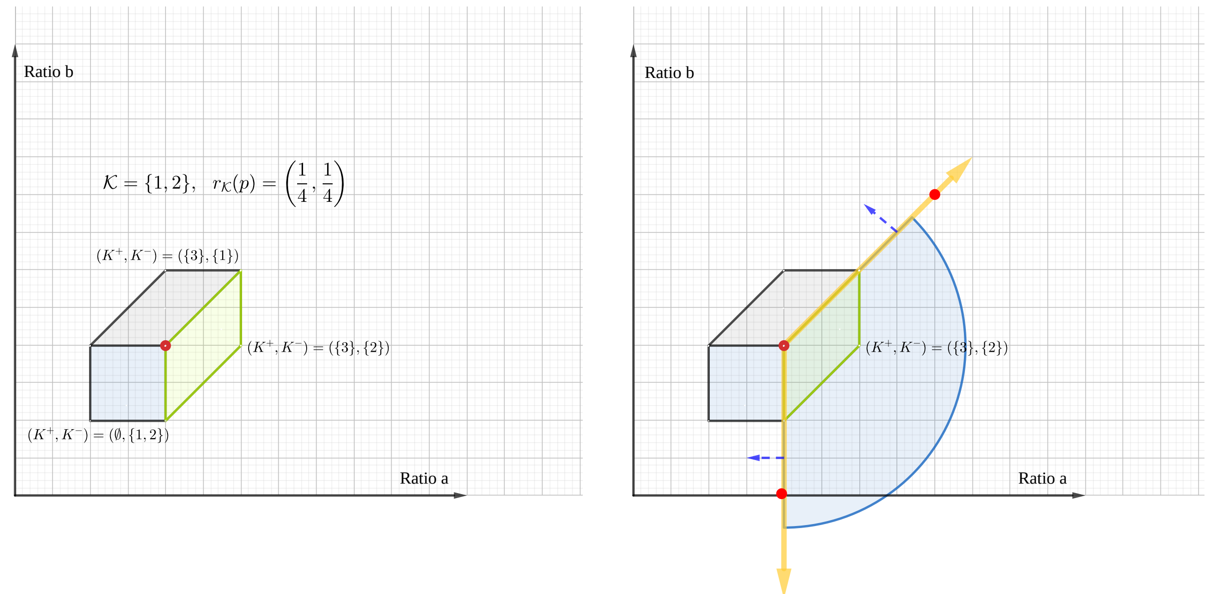

FIGURE 5. Geometric properties in the base example for the centroid $\{1,2\}$ whose budget is $r=(1/4,1/4)$: (LEFT) The extreme points and convex subsets, and (RIGHT) the cone, with orange boundaries, corresponding to $(K^+,K^-,K^0)=(\{3\},\{2\},\varnothing)$. The dashed vectors are the outer normals, $\psi_1=(-1,1)'$ and $\psi_2=(-1,0)'$ that characterize the cone.

In Fig. 5 (RIGHT) we plot three neighbors of the centroid $\mathcal{K}=\{1,2\}$ in our base example. For the direction $(\kappa^+,\kappa^-)=(\{3\},\{2\})$ the neighboring centroid is $\mathcal{K}'=\{1,3\}$. In moving from $\mathcal{K}$ to $\mathcal{K}'$ the request variable y_2 and the unmet variable u_3 leave the basis, and y_3 and u_2 enter the basis.

One optimal basis at the centroid $\mathcal{K}=\{1,2\}$ has $K^+=\{3\}$ and $K^-=\{2\}$. The neighboring centroids with $\kappa^+\in K^+$ and $\kappa^-\subseteq K^-$ are $\{1,3\}$, $\{1\}$, and $\{1,2,3\}$. The set $\mathcal{N}_{\mathcal{K}}(D,\mathcal{B})$ is the convex hull of the mid-points of the lines leading to those neighbors and corresponds to the yellow-colored set; it is the intersection of the action region $\mathcal{N}_{\mathcal{K}}(D)$ with the cone defined by the arrows.

DEFINITION 10 (BASIC CONE). Let $\mathcal{K}$ be a centroid with associated basis $\mathcal{B}$ and (K^+,K^-,K^0) be the zero-valued basic variables (Definition 8). Define

$$\mathrm{cone}(\mathcal{K},\mathcal{B})=\{\xi\in\mathbb{R}^d:\xi=A_{K^+}x_{K^+}-A_{K^-}x_{K^-}+b_{K^0},\ \text{for some } x\in\mathbb{R}^n_+,b\in\mathbb{R}^d_+\}.$$

This definition of the basic cone depends only on $(\mathcal{K},\mathcal{B})$ and not on D.

LEMMA 5. *Let $\mathcal{K}$ be a centroid with basis $\mathcal{B}$. Then, $\mathcal{N}_{\mathcal{K}}(D,\mathcal{B})=\mathcal{N}_{\mathcal{K}}(D)\cap(r_{\mathcal{K}}(D)+\mathrm{cone}(\mathcal{K},\mathcal{B}))$.*

The properties of the outward normals to the cone are central to the proof of oracle containment (see Theorem 3). Fig. 5 (RIGHT) visualizes these vectors.

The existence of a finite family of separating vectors $\Psi(\mathcal{K},\mathcal{B}):=\{\psi_l,l\in\mathcal{L}(\mathcal{K},\mathcal{B})\}$ such that $\xi\in\mathrm{cone}(\mathcal{K},\mathcal{B})$ if and only if $\max_{\psi\in\Psi(\mathcal{K},\mathcal{B})}\psi'\xi\leq 0$ follows from the Minkowski-Weyl thoerem; see e.g. [Bertsimas and Tsitsiklis, 1997, Chapter 4.9]. The next lemma explicitly characterizes $\Psi(\mathcal{K},\mathcal{B})$ in terms of immediate centroid neighbors: those that either add or remove *one* type relative to $\mathcal{K}$.

Lemma 6. *Fix a centroid $\mathcal{K}$ with associated basis $\mathcal{B}$. The set $\Psi(\mathcal{K},\mathcal{B})$ of separating vectors contains one vector $\psi[\kappa]$ for each $\kappa=(\kappa^+,\kappa^-,\kappa^0)\in K^+(\mathcal{B})\times K^-(\mathcal{B})\times K^0(\mathcal{B})$ with $|\kappa^+|+|\kappa^-|+|\kappa^0|=1$. These vectors satisfy the following properties:*

1. $\psi[\kappa]'A_{K^+\backslash\kappa^+}=0$, $\psi[\kappa]'A_{K^-\backslash\kappa^-}=0$, *and* $\psi[\kappa]'e_{K^0\backslash\kappa^0}=0$. *Also,* $\psi[\kappa]'A_{\kappa^+}<0$ *if* $|\kappa^+|=1$, $\psi[\kappa]'A_{\kappa^-}>0$ *if* $|\kappa^-|=1$, *and* $\psi[\kappa]'e_{\kappa^0}<0$ *if* $|\kappa^0|=1$.
2. *For any other basis* $\bar{\mathcal{B}}\neq\mathcal{B}$ *associated to* $\mathcal{K}$*: if* κ *as above is in both* $K^+(\mathcal{B})\times K^-(\mathcal{B})\times K^0(\mathcal{B})$ *and* $K^+(\bar{\mathcal{B}})\times K^-(\bar{\mathcal{B}})\times K^0(\bar{\mathcal{B}})$*, then:* $\psi[\kappa]'A_j\geq 0$ *for* $j\in K^+(\bar{\mathcal{B}})\backslash K^+(\mathcal{B})$, $\psi[\kappa]'A_j\leq 0$ *for* $j\in K^-(\bar{\mathcal{B}})\backslash K^-(\mathcal{B})$*, and* $\psi[\kappa]'e_i\geq 0$ *for* $i\in K^0(\bar{\mathcal{B}})\backslash K^0(\mathcal{B})$.

The separating vectors are defined by $\kappa^+\in K^+,\kappa^-\in K^-,\kappa^0\in K^0$ with $|\kappa^+|+|\kappa^-|+|\kappa^0|=1$. Because $|K^+|+|K^-|+|K^0|=d$, the number of separating vectors for the basic cone of $\mathcal{B}$ is $\mathcal{L}(\mathcal{K},\mathcal{B})=d$.

When the surplus coordinate $\kappa^0=\emptyset$, we write $\psi[\kappa^+,\kappa^-]$ instead of $\psi[\kappa]$. This allows us to focus on the values of κ^+,κ^- which determine the centroid's neighbor.

Remark 4 (generating the geometry of $\mathrm{LP}(R,D)$). Lemma 4 provides a tractable procedure to construct the action map—as in Fig. 1—and to identify the bases associated with each centroid set.

We first identify a single centroid set $\mathcal{K}^0$. Having solved $\mathrm{LP}(r_{\mathcal{K}^0}(D),D)$ (the LP at the centroid's budget) and identified all the optimal bases at this centroid, the sets K^+,K^-,K^0 give us, via Lemma 4, the centroid neighbors of $\mathcal{K}^0$. We repeat the procedure for each of these neighbors.

This requires solving at most n LPs per centroid[2] and produces the following outputs: (i) a map so that, at a time t and with budget-ratio being R^t we can identify $\mathcal{K}$ such that $R^t\in\mathcal{N}_{\mathcal{K}}(D)$, and (ii) the bases associated with a centroid set $\mathcal{K}$ and the set of dual variables $\Lambda_{\mathcal{K}}$ in Eq. (5). In turn, at a time t, we can compute the max-bid prices in Definition 2; see also Remark 7.

Because there are at most $n!$ centroids (as the number of paths on the integer set $[n]$ from the empty centroid to the centroid $\mathcal{K}=[n]$), the computational burden of generating the full map is at most the solution of $(n+1)!$ packing LPs. ■

When $R^t\in\mathcal{N}_{\mathcal{K}}(D,\mathcal{B})\subseteq\mathcal{N}_{\mathcal{K}}(D)$, BudgetRatio accepts only requests in $\mathcal{K}$. Lemma 7 below shows that, as long as R^t is in the proximity of $\mathcal{N}_{\mathcal{K}}(D,\mathcal{B})$, BudgetRatio performs only basic allocations. In Fig. 4(RIGHT): as long as R^t is in the proximity of the yellow subset of $\mathcal{N}_{\{1,2\}}$—it is either in $\mathcal{N}_{\{1,2\}}$ or in one of the neighbors $\mathcal{N}_{\{1\}}$, $\mathcal{N}_{\{1,3\}}$, $\mathcal{N}_{\{1,3\}}$—it accepts only requests in $\{1,2\}\cup\{3\}$.

[2] This is because the cones are characterized by immediate neighbors, for each type $j\in\mathcal{K}$, we only need to find whether $\mathcal{K}\backslash\{j\}$ is a centroid, and for each $j\in\mathcal{J}\backslash\mathcal{K}$ if $\mathcal{K}\cup\{j\}$ is a centroid. Verifying that $\mathcal{K}\subseteq\mathcal{J}$ is a centroid only requires solving $\mathrm{LP}(r_{\mathcal{K}}(D),D)$.

Henceforth, we fix

$$\epsilon^0 := \frac{1}{8}\min\{p_j\} \wedge \min\{\varrho_i : \varrho_i > 0\}. \tag{21}$$

LEMMA 7 **(optimal bases and BudgetRatio actions)**. *There exist constants M_1, M_2 such that, if $d_\infty(R^t, \mathcal{N}_\mathcal{K}(\mathcal{B}, p)) \leq \frac{\epsilon^0}{M_1}$,* BUDGETRATIO *performs basic allocations at t: it serves only (but not necessarily all) requests in $\mathcal{K} \cup K^+(\mathcal{B})$ and it rejects only requests in $\mathcal{K}^c \cup K^-(\mathcal{B})$. Moreover, $I_i^t \leq M_2(T-t)$ for all $i \notin K^0(\mathcal{B})$.*

PROOF OF PROPOSITION 2. Let $\mathcal{B}$ be the optimal offline basis and recall that $\tau^{\epsilon,\mathcal{K}} = \min\{t \leq T : d_\infty(R^t, \mathcal{N}_\mathcal{K}(\mathcal{B}, p)) > \epsilon\}$. Setting $\epsilon = \epsilon^0/M_1$, Lemma 7 guarantees that BUDGETRATIO performs basic allocations at t if $d_\infty(R^t, \mathcal{N}_\mathcal{K}(\mathcal{B}, p)) \leq \epsilon$. In turn, with this choice of ϵ, $\tau^{\epsilon,\mathcal{K}} \leq \tau^\pi$ as stated. □

To complete the proof of Theorem 1 (constant regret), it remains to prove Proposition 3. That is the focus of the next section.

5. Analysis of BudgetRatio's dynamics.

To prove Proposition 3, we must bound the time $\tau^{\epsilon,\mathcal{K}} = \min\{t \leq T : d_\infty(R^t, \mathcal{N}_\mathcal{K}(\mathcal{B}, p)) > \epsilon\}$, where $\mathcal{B}$ is the optimal offline basis and ϵ is as in Lemma 7.

We will lower bound $\tau^{\epsilon,\mathcal{K}}$ by two auxiliary exit times:

$$\tau^{\epsilon',\mathcal{K}}_{\text{region}} := \inf\{t \leq T : d(R^t, \mathcal{N}_\mathcal{K}(p)) > \epsilon'\} \tag{22}$$

$$\tau^{\epsilon',\mathcal{B}}_{\text{cone}} := \inf\{t \leq T : \max_{\psi \in \Psi(\mathcal{K},\mathcal{B})} \psi'(R^t - r_\mathcal{K}(p)) > \epsilon'\}, \tag{23}$$

where the vectors $\psi \in \Psi(\mathcal{K}, \mathcal{B})$ are as in Lemma 6 and $\epsilon' > 0$ depends on ϵ. Recall (Lemma 5) that $\mathcal{N}_\mathcal{K}(\mathcal{B}, p)$ is the intersection of $\mathcal{N}_\mathcal{K}(p)$ and the cone $r_\mathcal{K}(p) + cone(\mathcal{K}, \mathcal{B})$. To exit $\mathcal{N}_\mathcal{K}(\mathcal{B}, p)$ it suffices, then, to exit either of the two; this is formalized in Lemma 8.

LEMMA 8 **(exit times)**. *Let $\mathcal{K}$ be a centroid with associated basis $\mathcal{B}$ and fix $\epsilon > 0$. There exists $\epsilon' > 0$ that depends on (ϵ, A, v) only, such that, for any $R \in \mathbb{R}^d$, if $d(R, \mathcal{N}_\mathcal{K}(D)) \leq \epsilon'$ and $\max_{\psi \in \Psi(\mathcal{K},\mathcal{B})} \psi'(R - r_\mathcal{K}(D)) \leq \epsilon'$, then $d_\infty(R, \mathcal{N}_\mathcal{K}(\mathcal{B}, D)) \leq \epsilon$. Consequently, $\tau^{\epsilon,\mathcal{K}} \geq \tau^{\epsilon',\mathcal{K}}_{region} \wedge \tau^{\epsilon',\mathcal{B}}_{\text{cone}}$.*

Theorem 2 is a bound on $\tau^{\epsilon',\mathcal{K}}_{\text{region}}$ and Theorem 3 is a bound on $\tau^{\epsilon',\mathcal{B}}_{\text{cone}}$. Together, these will provide a lower bound on $\tau^{\epsilon,\mathcal{K}}$ which we use to prove Proposition 3.

THEOREM 2 **(sticky boundaries)**. *Let $\mathcal{K}$ be the centroid such that $\mathbb{E}[R^0] \in \mathcal{N}_\mathcal{K}(\mathbb{E}[D^0]) = \mathcal{N}_\mathcal{K}(p)$ and $\tau^{\epsilon',\mathcal{K}}_{region}$ as in Eq. (22) with ϵ' as in Lemma 8. Then,*

$$\mathbb{P}[T - \tau^{\epsilon',\mathcal{K}}_{region} > \ell] \leq m_1 e^{-m_2 \ell},$$

where $m_1, m_2 > 0$ do not depend on (T, I^0) but possibly depend on p, ϱ, A, v and ϵ'.

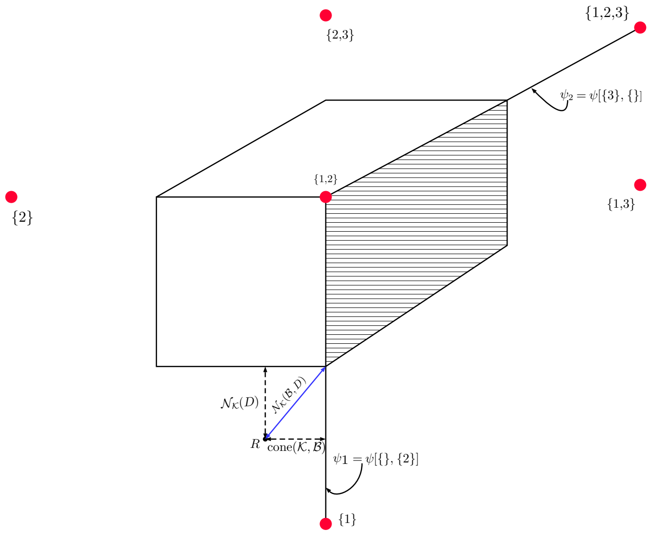

FIGURE 6. Representation of Lemma 8 for the base example. The red circles are the budgets and different centroids. We focus on the action region $\mathcal{N}_{\mathcal{K}}(D)$ for $\mathcal{K} = \{1, 2\}$. The shaded region is $\mathcal{N}_{\mathcal{K}}(\mathcal{B}, D)$ for the basis $\mathcal{B}$ that has $(K^+, K^-) = (\{3\}, \{2\})$. The two rays ψ_1 and ψ_2 define cone$(\mathcal{K}, \mathcal{B})$. At the bottom left we have a ratio $R \notin \mathcal{N}_{\mathcal{K}}(\mathcal{B}, D)$ and the dashed arrows represent the distance from R to $\mathcal{N}_{\mathcal{K}}(D)$ and cone$(\mathcal{K}, \mathcal{B})$, respectively. The solid blue arrow represents the distance from R to $\mathcal{N}_{\mathcal{K}}(\mathcal{B}, D)$ that is bounded by the two previous distances by virtue of Lemma 8. Note that the two extreme rays lead to two different neighboring centroids (second part of Lemma 8): one leads to $\mathcal{K}^0 = \{1\}$ and the other to $\mathcal{K}^0 = \{1, 2, 3\}$. The centroid $\{1, 3\}$ has $\kappa^- = \{2\}, \kappa^+ = \{3\}$, hence $|\kappa^-| + |\kappa^+| > 1$. It is in the interior of the cone.

THEOREM 3 **(cone containment)**. *Let $\mathcal{K}$ be the centroid such that $\mathbb{E}[R^0] \in \mathcal{N}_{\mathcal{K}}(\mathbb{E}[D^0]) = \mathcal{N}_{\mathcal{K}}(p)$, the basis $\mathcal{B}$ be such that $\max_{\psi \in \Psi(\mathcal{K}, \mathcal{B})} \psi'(\mathbb{E}[R^0] - r_{\mathcal{K}}(p)) \leq \epsilon'/2$, and $\tau_{\text{cone}}^{\epsilon', \mathcal{B}}$ be as in Eq. (23). Then, for all $\ell \in [T]$,*

$$\mathbb{P}[T - \tau_{\text{cone}}^{\epsilon', \mathcal{B}} > \ell, R^0 - r_{\mathcal{K}}(D^0) \in \text{cone}(\mathcal{K}, \mathcal{B})] \leq m_1 e^{-m_2 \ell},$$

for constants m_1, m_2 that do not depend on (T, I^0). In particular, letting $\mathcal{B}$ be the (random) optimal offline basis, we have that $\mathbb{P}[T - \tau_{\text{cone}}^{\epsilon', \mathcal{B}} > \ell] \leq m_1 e^{-m_2 \ell}$.

PROOF OF PROPOSITION 3. From Theorems 2 and 3 it follows immediately that $\mathbb{E}[T - \tau_{\text{region}}^{\epsilon', \mathcal{K}} \wedge \tau_{\text{cone}}^{\epsilon', \mathcal{B}}] \leq \mathbb{E}[T - \tau_{\text{region}}^{\epsilon', \mathcal{K}}] + \mathbb{E}[T - \tau_{\text{cone}}^{\epsilon', \mathcal{B}}] \leq M$. By Lemma 8, $\mathbb{E}[T - \tau^{\epsilon, \mathcal{K}}] \leq \mathbb{E}[T - \tau_{\text{region}}^{\epsilon', \mathcal{K}} \wedge \tau_{\text{cone}}^{\epsilon', \mathcal{B}}] \leq M$. By the second claim in Lemma 7, we have that $I_i^t \leq M(T - t)$ for all $i \notin K^0(\mathcal{B})$ and all $t < \tau^{\epsilon, \mathcal{K}}$. In turn, because $I_i^{t+1} \leq I_i^t + 1$, $\mathbb{E}[W^{\tau^{\epsilon, \mathcal{K}}}] \leq M(\mathbb{E}[T - \tau^{\epsilon, \mathcal{K}}] + 1)$. Overall, we have $\mathbb{E}[T - \tau^{\epsilon, \mathcal{K}} + W^{\tau^{\epsilon, \mathcal{K}}}] \leq M$, as stated. □

The two remaining sub-sections include the proofs of Theorems 2 and 3.

5.1. Proof of Theorem 2. We start with some preparatory lemmas. Define the set of request variables consistent with the action region $\mathcal{N}_{\mathcal{K}}(D)$:

$$\mathcal{Y}(\mathcal{K}, D) := \{y : \exists R \in \mathcal{N}_{\mathcal{K}}(D) \text{ s.t. for some } (u, s), (y, u, s) \text{ solves } \text{LP}(R, D)\}.$$

This definition translates, through LP(R,D), centroid neighborhoods—which are functions of budget R and demand D—to decision neighborhoods. In proving Theorem 2—instead of showing directly that R^t remains close to $\mathcal{N}_{\mathcal{K}}(D)$—we show that the solution y *at* R^t remains close to the solution $\theta_{\mathcal{K}}(y,D)$ at a point $R \in \mathcal{N}_{\mathcal{K}}(D)$. Lemma 9 introduces and characterizes this reference point θ.

Lemma 9. *Fix $\mathcal{K}$ and a neighbor $\mathcal{K}^0 = \mathcal{K} \cup \kappa^+ \backslash \kappa^-$. Fix $R \in \mathcal{N}_{\mathcal{K}^0}(D)$ and let (y,u,s) be the solution to* LP(R,D)*. Let*

$$(\theta_{\mathcal{K}}(y,D))_j = \begin{cases} y_j & \text{if } j \notin \kappa^+ \cup \kappa^- \\ D_j/2 & \text{if } j \in \kappa^+ \cup \kappa^-. \end{cases}$$

Then, the following holds:

1. $\theta_{\mathcal{K}}(y,D) \in$ *closure*$(\mathcal{Y}(\mathcal{K},D))$ *and* $(y - \theta_{\mathcal{K}}(y,D))_j = 0$ *for all* $j \notin \kappa^+ \cup \kappa^-$.
2. *If y is the optimal request variable for* LP(R,D) *with optimal basis $\mathcal{B}$ and $\bar{\mathcal{B}}$ is adjacent* $(\kappa^+ \cup \kappa^- \subseteq (K^+(\mathcal{B}) \cup K^-(\mathcal{B})) \cap (K^+(\bar{\mathcal{B}}) \cup K^-(\bar{\mathcal{B}})))$*, then* $\left(\bar{\mathcal{B}}^{-1}\begin{pmatrix} R \\ D \end{pmatrix}\right)_j = y_j$ *for* $j \in \kappa^+ \cup \kappa^-$.

In the one-dimensional case of Fig. 2, take $\mathcal{K} = \{1,2\}$, $\mathcal{K}^0 = \mathcal{K} \cup \{3\} \backslash \{\emptyset\} = \{1,2,3\}$. For $R \in \mathcal{N}_{\mathcal{K}^0}$, $y_1 = p_1, y_2 = p_2$ and $y_3 \geq p_3/2$; the point $\theta_{\{1,2\}}(y) = (p_1,p_2,p_3/2)$ is the closest point to y that is in $\mathcal{Y}(\mathcal{K},p)$.

In the proof, we use a surplus-corrected ratio. Let $y_{R,D}, s_{R,D}$ be the value of the request and surplus variables at LP(R,D). The surplus-corrected budget ratio is given by

$$R_\bullet := R - s_{R,D} = A y_{R,D}. \tag{24}$$

The value of the optimal request variables y are the same for LP(R,D), and LP$(R-s,D)$ for $s = s(R,D)$. In particular, $R \in \mathcal{N}_{\mathcal{K}}(D)$ if and only if $R_\bullet \in \mathcal{N}_{\mathcal{K}}(D)$. Lemma 10 shows that the proximity (not only inclusion) of R to an action region implies that of $R_\bullet$, and vice versa.

Lemma 10. *There exists ϵ' and a constant M so that, for all $\check{\epsilon} \leq \epsilon'$, $d(R,\mathcal{N}_{\mathcal{K}}(D)) \leq M\check{\epsilon}$ if and only if $d(R_\bullet,\mathcal{N}_{\mathcal{K}}(D)) \leq \check{\epsilon}$.*

Finally, we have a simple lemma that shows that the optimal request variables y cannot change too much over one time period.

Lemma 11. *Let R^t be the budget at time t under* BudgetRatio *and (y^t,u^t,s^t) be the solution to $LP(R^t,p)$. Then, there exist M_1, M_2 such that*

$$\|y^{t+1} - y^t\| \leq \frac{M_1}{T-t}, \text{ for all } t \leq T - M_2.$$

Proof of Theorem 2. Take $\check{\epsilon} = \epsilon'/M$ with M as in Lemma 10. We will show that if $\mathbb{E}[R^0] \in \mathcal{N}_{\mathcal{K}}(p)$ as assumed (in particular $\mathbb{E}[R^0]_{\bullet} \in \mathcal{N}_{\mathcal{K}}(p)$) then

$$\mathbb{P}\left[\sup_{t\in[1,T-\ell]} d(R^t_{\bullet}, \mathcal{N}_{\mathcal{K}}(p)) > \check{\epsilon}\right] \le m_1 e^{-m_2 \ell}.$$

This implies the same for R by virtue of Lemma 10.

To simplify notation we write $\theta^t = \theta_{\mathcal{K}}(y^t, p)$, where $\theta_{\mathcal{K}}(y^t, p)$ is as in Lemma 9. We define $\delta^t := y^t - \theta^t$ and the quadratic Lyapunov function

$$g^t := d^2(y^t, \mathcal{Y}(\mathcal{K}, p)) = \|y^t - \theta^t\|^2 = \|\delta^t\|^2.$$

Whenever $g^t \le \check{\epsilon}^2/nd$, we also have $d(R^t_{\bullet}, \mathcal{N}_{\mathcal{K}}(p)) \le \check{\epsilon}$. Indeed, if $g^t \le \check{\epsilon}^2/nd$ then by the Cauchy-Schwarz inequality, $|Ay^t - A\theta^t|_i^2 = (a_i'(y^t - \theta^t))^2 \le \check{\epsilon}^2/d$; we also use here the fact that A is a binary matrix. Since $\theta^t \in \text{closure}(\mathcal{Y}(\mathcal{K}, p))$ (Lemma 9), we have that $A\theta^t \in \mathcal{N}_{\mathcal{K}}(p)$ and $Ay^t = R^t_{\bullet}$, hence $||Ay^t - A\theta^t||^2 \le \check{\epsilon}^2$ implies $d(R^t_{\bullet}, \mathcal{N}_{\mathcal{K}}(p)) \le \check{\epsilon}$.

Setting $\varepsilon^2 := \check{\epsilon}^2/nd$, we conclude that $g^t \le \varepsilon^2$ implies $d(R^t_{\bullet}, \mathcal{N}_{\mathcal{K}}(p)) \le \check{\epsilon}$. This requirement is satisfied at $t = 1$. Indeed, we may assume without loss of generality that $I_i^0 \le 2T$ (otherwise we can remove the i^{th} capacity constraint at both $t = 0$ and $t = 1$). It is then a matter of simple algebra that $|R_i^1 - \mathbb{E}[R_i^0]| = \mathcal{O}\left(\frac{1}{T}\right)$ and by the Lipschitz continuity of the LP [Cook et al., 1986, Mangasarian and Shiau, 1987][3] that $\|R^1_{\bullet} - \mathbb{E}[R^0]_{\bullet}\| = \|Ay^1 - Ay^0\| = \mathcal{O}\left(\frac{1}{T}\right)$. Because $\mathbb{E}[R^0] \in \mathcal{N}_{\mathcal{K}}(p)$ we have, for T large, that $g^1 \le \varepsilon^2/2$.

We next prove the following drift condition: for some constants $M, \bar{M}$, and all $t \le T - \bar{M}$,

$$\mathbb{E}[g^{t+1} - g^t | \mathcal{F}_t] \le -\frac{M}{T-t}, \text{ whenever } g^t \in [\varepsilon^2/2, \varepsilon^2]. \tag{25}$$

Assuming Eq. (25) and using $g^1 \le \varepsilon^2/2$, concentration arguments as in [Arlotto and Gurvich, 2019, Theorem 2] show that $\mathbb{P}[\max_{t\in[1,T-\ell]} g^t > \varepsilon^2] \le m_1 e^{-m_2\ell}$ for constants (m_1, m_2) that depend on M only. This proves the theorem.

The remainder of the proof is thus devoted to Eq. (25). Fix t and $\mathcal{K}^0$ such that $R^t \in \mathcal{N}_{\mathcal{K}^0}(p)$ with $\mathcal{K}^0 = \mathcal{K} \cup \kappa^+ \setminus \kappa^-$, then using Lemma 9 (item 1) we obtain

$$\begin{aligned}\mathbb{E}[g^{t+1} - g^t | \mathcal{F}_t] &= \mathbb{E}[\|\delta^{t+1} - \delta^t\|^2 | \mathcal{F}_t] + 2\mathbb{E}[(\delta^{t+1} - \delta^t)'\delta^t | \mathcal{F}_t] \\ &= \mathbb{E}[\|\delta^{t+1} - \delta^t\|^2 | \mathcal{F}_t] + 2\mathbb{E}[(\delta^{t+1} - \delta^t)'_{\kappa}(\delta^t)_{\kappa} | \mathcal{F}_t],\end{aligned}$$

where $\kappa = \kappa^+ \cup \kappa^-$. Our aim is to prove $\mathbb{E}[\|\delta^{t+1} - \delta^t\|^2 | \mathcal{F}_t] = \mathcal{O}(\frac{1}{(T-t)^2})$ and $\mathbb{E}[(\delta^{t+1} - \delta^t)'_{\kappa}(\delta^t)_{\kappa} | \mathcal{F}_t] \le -\frac{M}{T-t}$, which together would imply Eq. (25). We divide the proof into two parts: the linear bound $\mathbb{E}[(\delta^{t+1} - \delta^t)'_{\kappa}(\delta^t)_{\kappa} | \mathcal{F}_t] \le -\frac{M}{T-t}$ and the quadratic bound $\mathbb{E}[\|\delta^{t+1} - \delta^t\|^2 | \mathcal{F}_t] = \mathcal{O}(\frac{1}{(T-t)^2})$. We start with a fact about the different neighborhoods that the process R^t visits during its evolution.

[3] This is the first of multiple places throughout our proofs where we use the Lipschitz continuity of linear programming in the right hand side. Henceforth, we will do so without citation.

Property of visited neighbors. For a vector $y \in \mathcal{Y}(\mathcal{K}, p)$, by definition, $y_j \geq p_j/2$ for $j \in \mathcal{K}$ and $y_j < p_j/2$ otherwise. Hence, by the Lipschitz continuity of the LP, if $g^t \leq \varepsilon^2$,

$$y_j^t \geq p_j/2 - \varepsilon \quad \forall j \in \mathcal{K} \quad \text{and} \quad y_j^t \leq p_j/2 + \varepsilon \quad \forall j \notin \mathcal{K}. \tag{26}$$

Let $\mathcal{B}$ be the optimal basis of LP(R^t, D) and $\bar{\mathcal{B}}$ be the optimal basis at LP(R^{t+1}, D). Recall that $R^t \in \mathcal{N}_{\mathcal{K}^0}(D)$. We claim that

$$R_{\bullet,i}^{t+1} - R_{\bullet,i}^t = O\left(\frac{1}{T-t}\right), \ \forall i \in [d] \quad \text{and} \quad \kappa^+ \cup \kappa^- \subseteq (K^+(\mathcal{B}) \cup K^-(\mathcal{B})) \cap (K^+(\bar{\mathcal{B}}) \cup K^-(\bar{\mathcal{B}})). \tag{27}$$

The first fact follows directly from Lemma 11 recalling that $R_\bullet^t = Ay^t$. For the second fact, take $j \in \kappa^+ \cup \kappa^-$, then Eq. (26) implies $y_j^t = \frac{1}{2}p_j \pm \varepsilon$ and, in particular, that $j \in K^+(\mathcal{B}) \cup K^-(\mathcal{B})$. Because the solutions to the LP are Lipschitz continuous in the right-hand side, $|y_j^t - y_j^{t+1}| \leq M||R_\bullet^t - R_\bullet^{t+1}|| = \mathcal{O}(\frac{1}{T-t})$, so that for all $t \leq T - \bar{M}$ it must be that $y_j^{t+1} = \frac{1}{2}p_j \pm 2\varepsilon$ and hence (with ε small enough) that $j \in K^+(\bar{\mathcal{B}}) \cup K^-(\bar{\mathcal{B}})$.

Linear bound. We claim that, if $R^t \in \mathcal{N}_{\mathcal{K}^0}(p)$ and $g^t \leq \varepsilon^2/2$, then

$$\begin{aligned} \mathbb{E}[\delta_j^{t+1} - \delta_j^t | \mathcal{F}_t] &\leq -\frac{M}{T-t} \quad j \in \kappa^+, \\ \mathbb{E}[\delta_j^{t+1} - \delta_j^t | \mathcal{F}_t] &\geq \frac{M}{T-t} \quad j \in \kappa^-. \end{aligned} \tag{28}$$

Assuming Eq. (28), if $g^t \geq \varepsilon^2/2$, then there exists $j \in \kappa^+$ s.t. $\delta_j^t = y_j^t - D_j/2 \geq \varepsilon/\sqrt{2}$ or some $j \in \kappa^-$ s.t. $\delta_j^t - y_j^t - D_j/2 \leq -\varepsilon/\sqrt{2}$, hence from Eq. (28) we can bound

$$\mathbb{E}[(\delta^{t+1} - \delta^t)'_\kappa (\delta^t)_\kappa | \mathcal{F}_t] \leq -\frac{M}{T-t}.$$

Recall that $\mathcal{K}^0$ is such that $R^t \in \mathcal{N}_{\mathcal{K}^0}(p)$ and that σ_j^t is the indicator that a request j is accepted at time t. Since only requests in $\mathcal{K}^0$ may be accepted at t, we have the identity $\mathbb{E}[I^{t+1}] = \varrho + I^t - \mathbb{E}[A_{\mathcal{K}^0}\sigma_{\mathcal{K}^0}^t]$, which implies $R^{t+1}(T-t-1) = R^t(T-t) - \mathbb{E}[A_{\mathcal{K}^0}\sigma_{\mathcal{K}^0}^t]$ and, in turn,

$$\mathbb{E}[R^{t+1} - R^t | \mathcal{F}_t] = \frac{1}{(T-t-1)}(R^t - A_{\mathcal{K}^0}\mathbb{E}[\sigma_{\mathcal{K}^0}^t]).$$

Let $\mathcal{B}$ is the optimal basis for LP(R^t, p) and $\bar{\mathcal{B}}$ be the optimal basis for LP(R^{t+1}, p), then $\mathcal{B}y^t + s^t = R^t, \bar{\mathcal{B}}y^{t+1} + s^{t+1} = R^{t+1}$ so that (since $D = p$ does not change with t)

$$\mathbb{E}\left[\bar{\mathcal{B}}\begin{pmatrix} y^{t+1} \\ u^{t+1} \\ s^{t+1} \end{pmatrix} - \mathcal{B}\begin{pmatrix} y^t \\ u^t \\ s^t \end{pmatrix} \middle| \mathcal{F}_t\right] = \frac{1}{T-t-1}\begin{pmatrix} R^t - A_{\mathcal{K}^0}\mathbb{E}[\sigma_{\mathcal{K}^0}^t] \\ 0 \end{pmatrix} = \frac{1}{T-t-1}\left[\begin{pmatrix} R^t \\ p \end{pmatrix} - \begin{pmatrix} A_{\mathcal{K}^0}\mathbb{E}[\sigma_{\mathcal{K}^0}^t] \\ p \end{pmatrix}\right].$$

By (27), $\mathcal{B}$ and $\bar{\mathcal{B}}$ are adjacent in the sense of Lemma 9 (item 2). Multiplying by $\bar{\mathcal{B}}^{-1}$ and using the lemma we have that

$$\mathbb{E}[(y^{t+1} - y^t)_\kappa | \mathcal{F}_t] = \frac{1}{T-t-1}(y^t - \mathbb{E}[\sigma_{\mathcal{K}^0}^t])_\kappa, \tag{29}$$

where $\kappa = \kappa^+ \cup \kappa^-$.

Because $R^t \in \mathcal{K}^0$, a request of type $j \in \kappa^+$ that arrives at time t is accepted; Lemma 2 guarantees that it is feasible to do so. Hence $\mathbb{E}[\sigma_j^t] = \mathbb{E}[\mathbb{1}_{\{J^t=j\}}] = p_j$; where, we recall, $J^t = j$ means that the arrival at time t is of type j. Because $y_j^t \le p_j/2 + \varepsilon$ for $j \in \kappa^+$ (see Eq. (26)), we then have using Eq. (29) that

$$\mathbb{E}[\delta_j^{t+1} - \delta_j^t | \mathcal{F}_t] = \frac{1}{T-t-1}(y_j^t - \mathbb{E}[\sigma_j^t]) \le \frac{1}{T-t-1}(-p_j/2 + \varepsilon), \text{ for } j \in \kappa^+, \tag{30}$$

This establishes the first row of Eq. (28). For $j \in \kappa^-$, since $j \notin \mathcal{K}^0$, from Eq. (29) we have

$$\mathbb{E}[\delta_j^{t+1} - \delta_j^t | \mathcal{F}_t] = \frac{1}{T-t-1} y_j^t \ge \frac{1}{T-t-1}(p_j/2 - \varepsilon), \text{ for } j \in \kappa^- \tag{31}$$

where we used $y_j^t \ge p_j/2 - \varepsilon$ (see Eq. (26)). This concludes the proof of Eq. (28).

Quadratic bound. Finally, we prove that $||\delta^{t+1} - \delta^t|| = \mathcal{O}(\frac{1}{T-t})$. By Lemma 9 (item 1) we have $\delta^t = (y^t - \frac{1}{2}p)_\kappa$ and $\delta^{t+1} = (y^{t+1} - \frac{1}{2}p)_{\tilde{\kappa}}$, where $\tilde{\kappa} = \tilde{\kappa}^+ \cup \tilde{\kappa}^-$ defines the neighbor $\tilde{\mathcal{K}}$ visited at time $t+1$. If either $j \in \kappa \cap \tilde{\kappa}$ or $j \notin \kappa \cup \tilde{\kappa}$, $\delta_j^{t+1} - \delta_j^t = y_j^{t+1} - y_j^t$; in the first case $(\theta_{\mathcal{K}}(y^{t+1}, p))_j = (\theta_{\mathcal{K}}(y^t, p))_j = D_j/2$ and in the latter $(\theta_{\mathcal{K}}(y^{t+1}, p))_j = y_j^{t+1}, (\theta_{\mathcal{K}}(y^t, p))_j = y_j^t$.

Now consider what happens in the remaining coordinates. If $j \in \tilde{\kappa}^+$ (in particular $j \notin \mathcal{K}$) but $j \notin \kappa^+$, then $y_j^{t+1} \ge p_j/2 \ge y_j^t$ so that $\delta_j^t = 0$, $|\delta_j^{t+1} - \delta_j^t| = |\delta_j^{t+1}| = |y_j^{t+1} - p_j/2| \le |y_j^{t+1} - y_j^t|$. An identical argument applies to the case that $j \in \tilde{\kappa}^-$ (in particular $j \in \mathcal{K}$) but $j \notin \kappa^-$. It follows that

$$||\delta^{t+1} - \delta^t|| \le ||(y^t - y^{t+1})_{\kappa \cup \tilde{\kappa}}|| \le ||y^t - y^{t+1}||.$$

The norm on the right-hand side is bounded using Lemma 11. □

Remark 5 (sticky boundaries). The arguments in the proof of Theorem 2 imply that, once close to the boundary, the process R^t stays there. Formally, let

$$\tau_\partial^0 = \inf\{t \le T : d(R^t, \partial \mathcal{N}_{\mathcal{K}}(p)) \le \epsilon'\}, \text{ and } \tau_\partial^1 = \inf\{t \ge \tau_\partial^0 : d(R^t, \partial \mathcal{N}_{\mathcal{K}}(p)) \ge 2\epsilon'\}.$$

Then, $\mathbb{P}\{T - \tau_\partial^1 \ge \ell\} \le m_1 e^{-m_2 \ell}$. ■

Remark 6 (centroids visited). Because $d(R^t, \mathcal{N}_{\mathcal{K}}(p)) \le \epsilon'$ up to $\tau_{\text{region}}^{\epsilon', \mathcal{K}}$, centroids visited must be of the form $\mathcal{K} \cup \bar{K}^+ \backslash \bar{K}^-$ where $\bar{K}^+ \subseteq \cup_{\mathcal{B}} K^+(\mathcal{B})$ (with the union taken over bases associated to $\mathcal{K}$) and $\bar{K}^- \subseteq \cup_{\mathcal{B}} K^-(\mathcal{B})$.

Indeed, by the continuity of LP in the right-hand side, we have up to $\tau_{\text{region}}^{\epsilon', \mathcal{K}}$ that $y_j^t \le M\epsilon'$ for all $j \notin \mathcal{K} \cup (\cup_{\mathcal{B}} K^+(\mathcal{B}))$. Similarly, $y_j^t \ge p_j/2 - M\epsilon'$ for all $j \in \mathcal{K} \backslash (\cup_{\mathcal{B}} K^+(\mathcal{B}))$. ■

5.2. Proof of Theorem 3. We start with some preparatory lemmas. We first relate basic cones to the optimality of a basis $\mathcal{B}$ for offline, The empirical demand distribution, recall, is $D^0 = \frac{1}{T}Z^T =: \bar{Z}^T$ and the empirical budget ratio is $R^0 = \frac{1}{T}(I^0 + \mathfrak{Z}^t) = \frac{1}{T}I^0 + \bar{\mathfrak{Z}}^t$ where $\bar{\mathfrak{Z}}^t = \frac{1}{T}\mathfrak{Z}^t$. By Lemma 5, $\mathcal{B}$ is optimal for offline if (i) $R^0 \in \mathcal{N}_{\mathcal{K}}(D^0)$, and (ii) $R^0 - r_{\mathcal{K}}(D^0) \in \text{cone}(\mathcal{K},\mathcal{B})$. The next lemma shows that, on the high probability event

$$\mathcal{A}^{\epsilon^0} := \left\{\omega \in \Omega : \|(\bar{Z}^T, \bar{\mathfrak{Z}}^t)' - (p, \varrho)'\|_\infty \le \epsilon^0\right\},$$

where ϵ^0 is as in (21), we can replace the requirement that $R^0 \in \mathcal{N}_{\mathcal{K}}(D^0)$ with one where R^0 and D^0 are replaced with their expectations, $\mathbb{E}[R^0] \in \mathcal{N}_{\mathcal{K}}(\mathbb{E}[D^0])$.

Lemma 12. *Let $\mathcal{K}$ be the centroid such that $\mathbb{E}[R^0] \in \mathcal{N}_{\mathcal{K}}(\mathbb{E}[D^0]) = \mathcal{N}_{\mathcal{K}}(p)$ and let*

$$\mathcal{M}(\mathcal{B}) := \left\{\omega \in \Omega : R^0 - r_{\mathcal{K}}(D^0) \in \text{cone}(\mathcal{K},\mathcal{B})\right\}.$$

Then, on the event $\mathcal{A}^{\epsilon^0}$, $\mathcal{M}(\mathcal{B}) = \{\mathcal{B}$ is offline optimal$\}$: $\mathcal{B}$ is an optimal offline basis if and only if $R^0 - r_{\mathcal{K}}(D^0) \in cone(\mathcal{K},\mathcal{B})$.

Lemma 13 below is used for a sub-case in the proof of Theorem 3. It says that if R is in the proximity of d+1 centroid neighbrhoods, it must in the strict interior of an explicitly identifiable basic cone and hence will take actions consistent with that basis.

Lemma 13. *Fix $\mathcal{K}$. Let $\mathcal{K}_i = \mathcal{K} \cup \kappa_i^+ \backslash \kappa_i^-$, $i \in [d]$ be d other centroids. There exists $\epsilon'', \delta > 0$ such that: if $\exists R$ such that $d(R, \mathcal{N}_{\mathcal{K}_i}(D)) \le \epsilon'', i \in [d]$, then*

1. *$\mathcal{K}$,$\mathcal{K}_i, i \in [d]$ share a basis $\mathcal{B}$ associated to $\mathcal{K}$: $\cup_{i\in[d]}\kappa_i^+ \in K^+(\mathcal{B})$, $\cup_{i\in[d]}\kappa_i^- \subseteq K^-(\mathcal{B})$, and $\cup_{i\in[d]}\kappa_i^0 \subseteq K^0(\mathcal{B})$.*
2. *R is in the strict interior of $\text{cone}(\mathcal{K},\mathcal{B})$, i.e., for the vectors $\Psi(\mathcal{K},\mathcal{B})$ characterizing $\text{cone}(\mathcal{K},\mathcal{B})$ in Lemma 6, we have $\max_{\psi\in\Psi(\mathcal{K},\mathcal{B})} \psi'(R - r_{\mathcal{K}}(D)) \le -\delta$.*

We note, in passing, that one cannot have the intersection of strictly more than $d+1$ different action regions. This is because if $R \in \mathcal{N}_{\mathcal{K}}(D)$ and $d(R, \mathcal{N}_{\mathcal{K}_i}(D)) \le \epsilon''$ for some R and $d+1$ other centroids $\mathcal{K}_i$, then there would $|\mathcal{K}| + |\mathcal{K}^c| + |\cup_i \kappa_i^+| + |\cup_i \kappa^-| = n + d + 1$ strictly positive variables at the solution to $LP(R, D)$, while an optimal solution has at most $n + d$ basic variables.

The following lemma is the last ingredient for the proof of Theorem 3. We define the random set

$$\Psi_0 = \{\psi \in \Psi(\mathcal{K},\mathcal{B}) : \psi'(R^{\tau_{\text{cone}}^{\epsilon',\mathcal{B}}} - r_{\mathcal{K}}(p)) > \epsilon'\}; \tag{32}$$

if $\tau_{\text{cone}}^{\epsilon',\mathcal{B}} = \infty$, we set $\Psi_0 = \emptyset$. These are the separation conditions that are violated at the exit from the cone. We define also $\mathcal{V} = \mathcal{V}[\tau_{\text{cone}}^{\epsilon',\mathcal{B}}]$ the be the random variable that count how many centroid neighborhoods—in addition to $\mathcal{K}$ for which $\mathbb{E}[R^0] \in \mathcal{N}_{\mathcal{K}}(p)$—-are visited by time $\tau_{\text{cone}}^{\epsilon',\mathcal{B}}$.

LEMMA 14. *Let $\mathcal{K}$ be the centroid such that $\mathbb{E}[R^0] \in \mathcal{N}_{\mathcal{K}}(\mathbb{E}[D^0]) = \mathcal{N}_{\mathcal{K}}(p)$ and let $\mathcal{K}_i = \mathcal{K} \cup \bar{K}_i^+ \backslash \bar{K}_i^-$ for $i \in [\mathcal{V}]$ be the i^{th} centroid visited after $\mathcal{K}$.*

Given $d^0 < d$, there exist $m_1, m_2 > 0$ and an event $\mathcal{C}_\ell$, with $\mathbb{P}[\mathcal{V} = d^0, (\mathcal{C}_\ell)^c] \le m_1 e^{-m_2 \ell}$ on which the following holds: if $T - \tau_{\text{cone}}^{\epsilon',\mathcal{B}} > \ell$ and $\mathcal{V} = d^0$, then Ψ_0 does not contain any of the vectors $\psi[\kappa^+, \kappa^-] \in \Psi(\mathcal{K}, \mathcal{B})$ for $\kappa^+ \in \cup_{i \in [d]} \bar{K}_i^+, \kappa^- \in \cup_{i \in [d]} \bar{K}_i^-$.

If $\mathcal{V} = 0$, Ψ^0 might contain any of the vectors in $\Psi(\mathcal{K}, \mathcal{B})$.

PROOF OF THEOREM 3. Throughout, $\mathcal{K}$ and the basis $\mathcal{B}$ are fixed. Because $\mathbb{E}[R^0] \in \mathcal{N}_{\mathcal{K}}(p)$, we have by Theorem 2 that

$$\mathbb{P}[T - \tau_{\text{region}}^{\epsilon',\mathcal{K}} > \ell] \le m_1 e^{-m_2 \ell} \quad \text{where} \quad \tau_{\text{region}}^{\epsilon',\mathcal{K}} = \inf\{t \le T : d(R^t, \mathcal{N}_{\mathcal{K}}) \ge \epsilon'\}. \tag{33}$$

Recall that $\mathcal{M}(\mathcal{B}) := \{\omega \in \Omega : R^0 - r_{\mathcal{K}}(D^0) \in \text{cone}(\mathcal{K}, \mathcal{B})\}$. Define the events $\mathcal{D} := \{T - \tau_{\text{cone}}^{\epsilon',\mathcal{B}} > \ell, \mathcal{M}(\mathcal{B})\}$.

Outline of the proof. To bound the measure of $\mathcal{D} \cap \Omega^\ell$, we consider two cases. In the first, at most $d-1$ action region other than $\mathcal{N}_{\mathcal{K}}(p)$ are visited during the horizon ($\mathcal{V} < d$); this corresponds to the case where the process starts close to the boundary of the cone and it is therefore the challenging case. The second case, where d or more other action regions are visited turns out to be easier. This is because, by Lemma 13, this only happens when the process moves in the strict interior of the cone where analysis is simpler; see Fig. 7.

First case (boundary):$\mathcal{V} < d$. Assume that over the interval $[1, \tau_{\text{cone}}^{\epsilon',\mathcal{B}}]$ at most d^0 centroids other $\mathcal{K}$ are visited: $\mathcal{N}_{\mathcal{K}}(p)$ and neighbors $\mathcal{N}_{\mathcal{K}^i}(p), i \in [d^0]$. The case that exactly one centroid neighborhood is visited during the horizon—namely that $R^t \in \mathcal{N}_{\mathcal{K}}(p)$ for all $t \le \tau_{\text{cone}}^{\epsilon',\mathcal{B}}$—is a simplified version of the argument for $d^0 \ge 1$, so we focus on the latter.

We will introduce a processes G^t with zero-mean increments that has the following properties on the event $\mathcal{M}(\mathcal{B})$:

$$G^1 \le T\epsilon'/2 \text{ and } G^T \le 0 \text{ a.s.} \quad \text{and} \quad \tau_{\text{cone}}^{\epsilon',\mathcal{B}} \le T - \ell \Longleftrightarrow G^t > (T-t)\epsilon' \text{ for some } t \in [1, T-\ell].$$

The event $\tau_{\text{cone}}^{\epsilon',\mathcal{B}} < T - \ell$, requires the process G^t to grow faster than the linear target $(T-t)\epsilon'$; an event that, we will prove, has an exponentially small probability.

Let Ψ^0 be as in (32). By Lemma 14 we have the existence of an event $\mathcal{C}_\ell$ with probability $\mathbb{P}[\mathcal{V} = d^0, (\mathcal{C}_\ell)^c] \le m_1 e^{-m_2 \ell}$ such that if the i^{th} centroid to be visited corresponds to $\mathcal{K}^0 = \mathcal{K} \cup \kappa^+ \setminus \kappa^-$ then Ψ_0 does not contain $\psi[\kappa_i]$ where $\mathcal{K}_i = \mathcal{K} \cup \kappa_i^+ \backslash \kappa_i^-$.

We next study sample paths of R^t on the event $\{\mathcal{V} = d^0\} \cap \mathcal{C}_\ell$. We have the inventory equation $I^s = I^0 + \mathfrak{Z}^s - AY^s$. Since $(T-s)R^s = I^s + (T-s)\varrho$ and $(T-s) r_{\mathcal{K}}(p) = (T-s) A_{\mathcal{K}} p_{\mathcal{K}}$,

$$I^s - (T-s) r_{\mathcal{K}}(p) = I^0 + \mathfrak{Z}^s - AY^s - T r_{\mathcal{K}}(p) + s A_{\mathcal{K}} p_{\mathcal{K}},$$

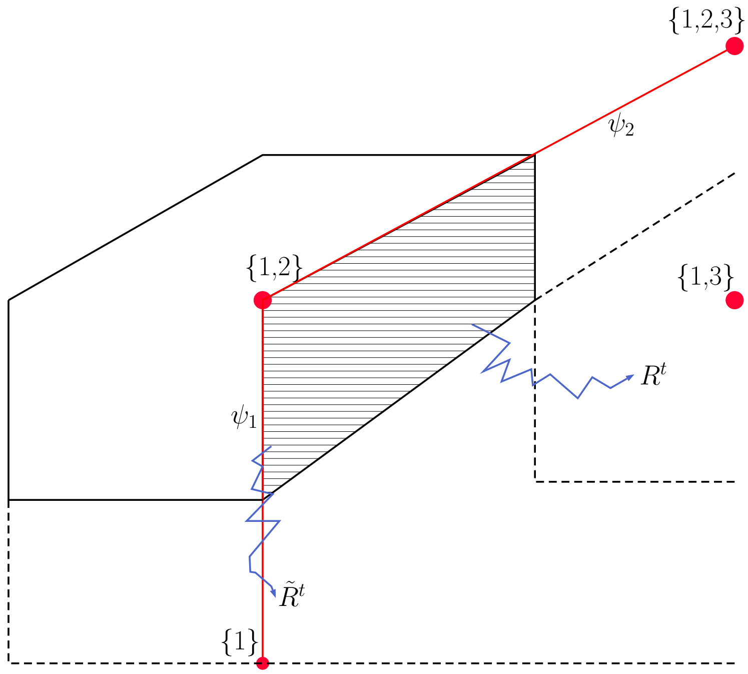

FIGURE 7. Two random walks for our base example. Solid lines enclose the region of interest $\mathcal{N}_{\{1,2\}}(D)$ and dashed lines enclose neighbouring action regions corresponding to $\{1\}$ and $\{1,3\}$. The random walk R^t visits three regions, namely $\{1,2\}$, $\{1\}$, and $\{1,3\}$, it is thus constrained to be in the interior of the cone, i.e., far away from the normals ψ_1, ψ_2. On the other hand, $\tilde{R}^t$ evolves close to the boundary of the cone, but in doing so visits only two regions, namely $\{1,2\}$ and $\{1\}$.

and, after basic algebraic manipulations, that

$$(T-s)(R^s - r_{\mathcal{K}}(p)) = T(\mathbb{E}[R^0] - r_{\mathcal{K}}(p)) + \widehat{\mathfrak{Z}}^s - AY^s + sA_{\mathcal{K}}p_{\mathcal{K}}, \tag{34}$$

where we define the centered process $\widehat{\mathfrak{Z}}^t := \mathfrak{Z}^t - t\varrho$. Over the interval $[1, \tau_{cone}^{\epsilon,\mathcal{B}})$ and on the event $\{\mathcal{V} = d^0\} \cap \mathcal{C}_\ell$, the only requests accepted correspond to $\mathcal{K} \cup \kappa_0^+$ where $\kappa_0^+ = \cup_{i\in[d^0]}\kappa_i^+$, hence $Y^s = Y^s_{\mathcal{K}} + Y^s_{\kappa_0^+}$. Additionally, all of the requests in $\mathcal{K} \setminus \kappa_0^-$, where $\kappa_0^- = \cup_{i\in[d^0]}\kappa_i^-$, are accepted. Thus, we have $Y^s_{\mathcal{K}\setminus\kappa_0^-} = Z^s_{\mathcal{K}\setminus\kappa_0^-}$. By Lemma 6 any $\psi \in \Psi_0$ is orthogonal to the columns of A corresponding to $\kappa_0 -^+$ and κ_0^-, so we arrive at the identities

$$\psi' AY^s = \psi' A_{\mathcal{K}\setminus\kappa_0^-} Z^s_{\mathcal{K}\setminus\kappa_0^-} \quad \text{and} \quad \psi' A_{\mathcal{K}} p_{\mathcal{K}} = \psi' A_{\mathcal{K}\setminus\kappa_0^-} p_{\mathcal{K}\setminus\kappa_0^-}.$$

Defining the centred process $\widehat{Z}^t := Z^t - tp$ and using these identities together with Eq. (34) we have

$$(T-s)\psi'(R^s - r_{\mathcal{K}}(p)) = T\psi'(\mathbb{E}[R^0] - r_{\mathcal{K}}(p)) + \psi'(\widehat{\mathfrak{Z}}^s - A_{\mathcal{K}\setminus\kappa_0^-}\widehat{Z}^s_{\mathcal{K}\setminus\kappa_0^-}) \quad \forall s < \tau_{\text{cone}}^{\epsilon',\mathcal{B}}.$$

Define the process

$$G^t_\psi := T\psi'(\mathbb{E}[R^0] - r_{\mathcal{K}}(p)) + \psi'(\widehat{\mathfrak{Z}}^t - A_{\mathcal{K}\setminus\kappa_0^-}\widehat{Z}^t_{\mathcal{K}\setminus\kappa_0^-}), \quad t \in [1, T].$$

Then, $G_\psi^t = (T-s)\psi'(R^s - r_{\mathcal{K}})$ for all $t < \tau_{\text{cone}}^{\epsilon',\mathcal{B}}$ and

$$\tau_{\text{cone}}^{\epsilon',\mathcal{B}} \le T-\ell \Longleftrightarrow \max_{\psi\in\Psi_0} G_\psi^t > (T-t)\epsilon' \quad \text{for some } t\in[0,T-\ell]. \tag{35}$$

The process G_ψ^t has zero-mean increments and $G_\psi^0 = T\psi'(\mathbb{E}[R^0] - r_{\mathcal{K}}(p)) \le T\epsilon'/2$ by assumption. Finally, $G_\psi^T \le 0$. This is because ψ is orthogonal to the columns $A_{\kappa_0^-}$, so that

$$\begin{aligned} G_\psi^T &= T\psi'(\mathbb{E}[R^0] - r_{\mathcal{K}}(p)) + \psi'(\widehat{\mathfrak{Z}}^T - A_{\mathcal{K}}\widehat{Z}_{\mathcal{K}}^T) \\ &= \psi'(I^0 + \mathfrak{Z}^T - A_{\mathcal{K}} Z_{\mathcal{K}}^T), \end{aligned}$$

In the event $\mathcal{M}(\mathcal{B})$ we have, by definition, $\frac{1}{T}(I^0 + \mathfrak{Z}^T - A_{\mathcal{K}} Z_{\mathcal{K}}^T) = R^0 - r_{\mathcal{K}}(D^0) \in \text{cone}(\mathcal{K},\mathcal{B})$ so that $G_\psi^T = T\psi'(R^0 - r_{\mathcal{K}}(D^0)) \le 0$.

We conclude that $\mathcal{D} \subseteq \cup_{\psi\in\Psi_0}\{G_\psi^T \le 0, \exists t\in T-\ell : G_\psi^t > (T-t)\epsilon'\}$. From Eq. (35), we deduce

$$\begin{aligned} \mathbb{P}[\mathcal{D}, \{\mathcal{V} = d^0\}] &\le \mathbb{P}[\mathcal{V} = d^0, (\mathcal{C}_\ell)^c] + \mathbb{P}\big[\mathcal{D}, \mathcal{C}_\ell, \{\mathcal{V} = d^0\}\big] + \\ &\le \mathbb{P}[\mathcal{V} = d^0, (\mathcal{C}_\ell)^c] + \sum_{\psi\in\Psi(\mathcal{K},\mathcal{B})} \mathbb{P}\left[\bigcup_{t\in[T-\ell]} \{G_\psi^t \ge (T-t)\epsilon'\}, G_\psi^T \le 0\right] \le m_1 e^{-m_2\ell}, \end{aligned}$$

for some $m_1, m_2 > 0$. The final bound follows from the analysis of a random walk crossing a positive moving threshold conditional on being negative at the end of the horizon. This is formally proved in Lemma 16 in the appendix.

Second case (strict interior): $\mathcal{V} \ge d$**.** Let $\mathcal{K}_i$ be the i^{th} centroid visited after $\mathcal{K}$. Let $\tau_{\text{region}}^{\epsilon',\mathcal{K}}$ be the exit time from $\mathcal{N}_{\mathcal{K}}(p)$ as in Theorem 2. Let $\tau_{\partial,i}^0, \tau_{\partial,i}^1$ be as in Remark 5 for $\mathcal{K}_i$. Define the event $\Omega_\ell = \{T - \tau_{\text{region}}^{\epsilon',\mathcal{K}} < \ell, T - \tau_{\partial,i}^1 < \ell, i\in[d]\}$, where we set $\tau_{\partial,i}^1 = \infty$ if $\mathcal{V} < i$ (fewer than i centroids other than $\mathcal{K}$ are visited).

From Theorem 2 and Remark 5 we have $\mathbb{P}[\mathcal{V} \ge d, (\Omega_\ell)^c] \le m_1 e^{-m_2\ell}$. On the event $\Omega_\ell \cap \{\mathcal{V} \ge d\}$ we have by Lemma 13 that $\max_{\psi\in\Psi(\mathcal{K},\mathcal{B})} \psi'(R^t - r_{\mathcal{K}}) \le -\delta$ for all $t < T-\ell$ and, in particular, $\mathbb{P}[\Omega_l, \mathcal{V} \ge d, T - \tau_{\text{cone}}^{\epsilon',\mathcal{B}} > \ell] = 0$.

$$\mathbb{P}[\mathcal{D}, \mathcal{V} \ge d] \le \mathbb{P}[T - \tau_{\text{cone}}^{\epsilon',\mathcal{B}} > \ell, \mathcal{V} \ge d, \Omega_\ell] + \mathbb{P}[\mathcal{V} \ge d, (\Omega_\ell)^c] \le \mathbb{P}[\mathcal{V} \ge d, (\Omega_\ell)^c] \le m_1 e^{-m_2\ell}.$$

Combining the two cases, we conclude that $\mathbb{P}[\mathcal{D}] = \sum_{d^0=0}^{d-1} \mathbb{P}[\mathcal{D}, \mathcal{V} = d^0] + \mathbb{P}[\mathcal{D}, \mathcal{V} \ge d] \le m_1 e^{-m_2\ell}$. The last implication in the theorem follows from Lemma 12 that guarantees that if $\mathbb{E}[R^0] \in \mathcal{N}_{\mathcal{K}}(D)$ then, on the event $\mathcal{A}^{\epsilon^0}$, $\mathcal{B}$ is the offline optimal if and only if $\mathcal{M}(\mathcal{B})$ holds. In turn,

$$\begin{aligned} \mathbb{P}\Big[T - \tau_{\text{cone}}^{\epsilon',\mathcal{B}} > \ell, \mathcal{B} \text{ is optimal}\Big] &= \mathbb{P}\Big[T - \tau_{\text{cone}}^{\epsilon',\mathcal{B}} > \ell, \mathcal{B} \text{ is optimal}, \mathcal{A}^{\epsilon^0}\Big] + \mathbb{P}\Big[T - \tau_{\text{cone}}^{\epsilon',\mathcal{B}} > \ell, \mathcal{B} \text{ is optimal}, (\mathcal{A}^{\epsilon^0})^c\Big] \\ &= \mathbb{P}\Big[\mathcal{D}, \mathcal{A}^{\epsilon^0}\Big] + \mathbb{P}\Big[T - \tau_{\text{cone}}^{\epsilon',\mathcal{B}} > \ell, \mathcal{B} \text{ is optimal}, (\mathcal{A}^{\epsilon^0})^c\Big] \\ &\le \mathbb{P}[\mathcal{D}] + \mathbb{P}\Big[\mathcal{B} \text{ is optimal}, (\mathcal{A}^{\epsilon^0})^c\Big]. \end{aligned} \tag{36}$$

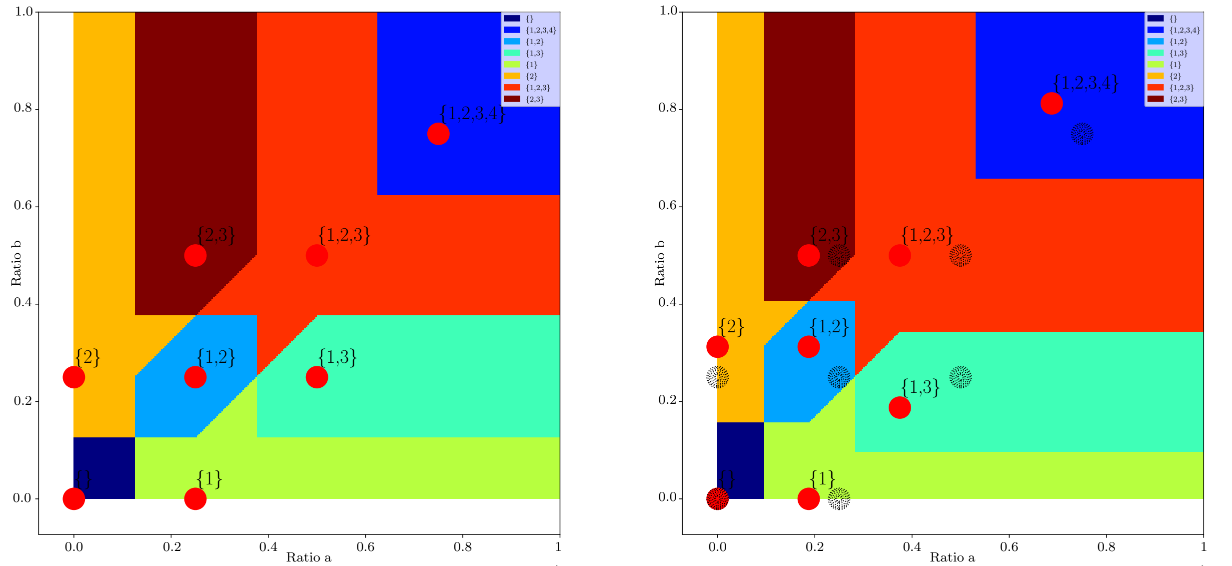

FIGURE 8. Action regions with true and misspecified probabilities (p and $\tilde{p}$). (LEFT) Action regions of BUDGETRATIO when it is executed relative to p. (RIGHT) Action regions when BUDGETRATIO uses $\tilde{p}_j = p_j - \frac{1}{16}$ for $j = 1,3$ and $\tilde{p}_j = p_j + \frac{1}{16}$ for $j = 2,4$.

By standard concentration results, there exist $\bar{m}_1, \bar{m}_2 > 0$ such that $\mathbb{P}[(\mathcal{A}^{\epsilon^0})^c] \leq \bar{m}_1 e^{-\bar{m}_2 T}$. Summing up the right hand side of (36) over bases $\mathcal{B}$ we get that

$$\mathbb{P}\left[T - \tau_{\text{cone}}^{\epsilon',\mathcal{B}} > \ell\right] \leq m_1 e^{-m_2 l} + \bar{m}_1 e^{-\bar{m}_2 T},$$

Because $\ell \leq T$, we have the statement of the theorem with modified constants m_1, m_2. □

6. Parameter misspecification. In this section we prove equations (7) and (8) of Theorem 1, and further discuss and illustrate their implications.

Demand perturbation. Fig. 8 illustrates a key idea behind equation (8). When p is replaced with an estimate $\widetilde{p}$ that satisfies (8), the centroids remain unchanged, but the shape of the centroid action regions is affected. Crucially, under Eq. (8), the true centroid budgets (dashed circles) lie in the interior of the (misspecified) actions regions. This guarantees that BUDGETRATIO, although equipped with wrong probabilities, achieves constant regret.

PROOF OF THEOREM 1(EQUATION (8): DEMAND ROBUSTNESS). Since $(\widetilde{p}, \widetilde{\varrho})$ satisfies slow restock, Lemma 2 does not change. Our constructions of the action regions $\mathcal{N}_{\mathcal{K}}(D)$ and their subsets $\mathcal{N}_{\mathcal{K}}(D, \mathcal{B})$ are for arbitrary D. Proposition 2 holds with D there set to $\widetilde{p}$; so does Lemma 8. Hence, bounded regret depends on whether Theorems 2 and 3 hold with the perturbed probabilities.

Importantly, the centroid sets, the bases associated with them and centroid neighbors are all invariant to D. For Theorem 2, by assumption (8), we have that $r_{\mathcal{K}}(p) \in \mathcal{N}_{\mathcal{K}}(\widetilde{p})$ so that the initial condition is preserved and the proof of Theorem 2 uses the action region $\mathcal{N}_{\mathcal{K}}(\widetilde{p})$ (instead of $\mathcal{N}_{\mathcal{K}}(p)$). The true probability p appears only in the expectation $\mathbb{E}[\sigma^t_{\mathcal{K}0}]$ there. The condition (8) guarantees that the drift is still negative in (30) and positive in (31) with $D_j/2$ replaced there with $D_j/4$.

The proof of Theorem 3 does not change because of the invariance centroid neighbors and basic cones to D. □

Proof of Corollary 1(Equation (10)). In the one-dimensional case, the centroids are of the form $[j]$ so that

$$\delta = \min_{j=1,\ldots,n-1} e'_{[j+1]}p_{[j+1]} - e'_{[j]}p_{[j]} = \min_j p_j,$$

where we define $[0] := \varnothing$. With this, (8) immediately reduces to (9). □

Learning the demand distribution. When p, ϱ are not known a priori, the controller must make decisions while learning the correct type probabilities p and ϱ. The controller observes the type of the request $j \in [n]$ and that of the restocked resource $i \in [d]$ at each period and builds the empirical estimates $\hat{p}^t_j = \frac{1}{t}\sum_{\tau=1}^t \mathbb{1}_{\{J^\tau = j\}}$, where J^τ is the type of the request arrival at τ, and $\hat{\varrho}^t_i$ is similarly defined.

Corollary 2 **(regret with demand learning)**. *Assume that the centroids are δ-separated (see Definition 4). Then, without prior knowledge of p, a modification of* BudgetRatio *achieves $\mathcal{O}(\log T)$ regret.*

Proof. Fix the constant $\epsilon = \frac{\delta}{4n}$. We build a simple policy of the form "learn, then act". We take an initial exploration phase of length $c\log T$ (for some $c = c(\epsilon) > 0$) during which all requests are rejected but the revealed types are used to build the empirical estimates $\hat{p}^t$ and $\hat{\varrho}^t$. By standard concentration results, we can choose c large enough to guarantee that $\mathbb{P}[||(p,\varrho)' - (\hat{p}^{c\log T}, \hat{\varrho}^{c\log T})|| > \epsilon] \leq 1/T$. After time $c\log T$, BudgetRatio is executed with the estimates $\hat{p}^{c\log T}, \hat{\varrho}^{c\log T}$ and achieves constant regret in the remaining periods by virtue of Theorem 1 and equation (8) there. On the event that $||(p,\varrho)' - (\hat{p}^{c\log T}, \hat{\varrho}^{c\log T})|| > \epsilon$ the regret is at most $T\max_{j\in[n]} v_j$; this event's contribution to regret is at most $T\max_{j\in[n]} v_j \times 1/T = \max_{j\in[n]} v_j = \mathcal{O}(1)$. □

Reward perturbation. To prove (7), we show that, under the δ-complementarity condition (see Definition 3), the centroids are stable to local perturbation of v and so is, in turn, the regret.

In the standard form (LP(R,D)) there are d resource-consumption constraints of the form $Ay + s = R$ and n demand constraints of the form $y + u = D$; a total of $d+n$ dual variables. An optimal primal-dual pair $(y,\lambda) \in \mathbb{R}^n \times \mathbb{R}^{d+n}$ must satisfy the complementarity properties: $s_i > 0 \Rightarrow \lambda_i = 0$, $u_j > 0 \Rightarrow \lambda_j = 0$, and, from the dual constraints, $([A'|I^n]\lambda)_j > v_j \Rightarrow y_j = 0$. There is the possibility that, for y_j non-basic (hence $y_j = 0$), we have $([A'|I^n]\lambda)_j = v_j$, i.e., complementarity is not strict. Definition 3 requires strict complementarity.

Recall that we associate to v the extended reward vector $\bar{v} := (v, 0, 0) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^d$, where the zeros correspond to unmet and surplus variables. Also, recall that the dual variable λ associated with $(\mathcal{B}, v)$ is $\lambda = (\mathcal{B}^{-1})'\bar{v}_{\mathcal{B}}$.

Proof of Theorem 1(Equation (7): reward robustness). Fix a basis $\mathcal{B}$ associated to some centroid $\mathcal{K}$ under rewards v. Let λ be the dual variables associated to $(\mathcal{B}, v)$. By Lemma 1, $\lambda \geq 0$ and $\bar{A}'\lambda \geq \bar{v}$. By virtue of Lemma 1, to prove that $\mathcal{B}$ is also an optimal basis under the rewards $\tilde{v}$, it suffices to show that the dual variables $\tilde{\lambda}$ associated to $(\mathcal{B}, \tilde{v})$ satisfy

(i) $\tilde{\lambda} \geq 0$, and

(ii) $\bar{A}'\tilde{\lambda} \geq \bar{\tilde{v}}$.

First, we claim that $\lambda = (\mathcal{B}^{-1})'\bar{v}$ must have $\lambda_j = 0$ for $j : u_j \in \mathcal{B}$, $\lambda_i = 0$ for $i : s_i \in \mathcal{B}$. Indeed, for a basis associated with $\mathcal{K}$ we have by Lemma 4 that there exists $R \in \mathcal{N}_{\mathcal{K}}(D, \mathcal{B})$ in which the solution is non-degenerate ($x_{K^+}, x_{K^-} > 0$ and $b_i > 0, i \in K^0$). At such R, complementary slackness implies that the $d+n$ dual variables associated with this basis are the unique solution to the following $d+n$ independent linear equations:

$$
\begin{aligned}
([A'|I^n]\lambda)_j &= v_j && j \text{ s.t. } y_j \in \mathcal{B} \\
\lambda_j &= 0 && j \text{ s.t. } u_j \in \mathcal{B} \\
\lambda_i &= 0 && i \text{ s.t. } s_i \in \mathcal{B}.
\end{aligned}
$$

The dual vector $\lambda = (\mathcal{B}^{-1})'\bar{v}$ is the unique solution to this linear system. Define now $\tilde{\lambda} = (\mathcal{B}^{-1})'\bar{\tilde{v}}$ where $\bar{\tilde{v}} = (\tilde{v}, 0, 0)'$. Then $\tilde{\lambda}_j = 0$ for $j : u_j \in \mathcal{B}$ and $\tilde{\lambda}_i = 0$ for $i : s_i \in \mathcal{B}$. Also,

$$||\tilde{\lambda} - \lambda|| = ||(\mathcal{B}^{-1})'\bar{\tilde{v}} - (\mathcal{B}^{-1})'\bar{v}|| \leq c||\bar{\tilde{v}} - \bar{v}||,$$

where $c = \{||\mathcal{B}^{-1}||_\infty : \mathcal{B} \text{ basis }\}$.

For (i) ($\tilde{\lambda} \geq 0$), because $\tilde{\lambda}_j = 0, u_j \in \mathcal{B}$ and $\tilde{\lambda}_i = 0, s_i \in \mathcal{B}$ we only need to study the case $u_j \notin \mathcal{B}$ or $s_i \notin \mathcal{B}$. For $\tilde{\lambda}_j, u_j \notin \mathcal{B}$, δ-complementarity and the requirement in (7) guarantee that $\tilde{\lambda}_j \geq \delta - c||\bar{v} - \bar{\tilde{v}}||_\infty \geq 0$, as desired. The same argument applies to $\tilde{\lambda}_i, s_i \notin \mathcal{B}$. For (ii) ($\bar{A}'\tilde{\lambda} \geq \tilde{v}$), by the same reasoning, we need to study only the dual constraints $([A'|I^n]\tilde{\lambda})_j \geq \bar{\tilde{v}}_j$ for $y_j \notin \mathcal{B}$. Let us denote the j-th row of $[A'|I^n]$ by η. By δ-complementarity, we have

$$([A'|I^n]\tilde{\lambda})_j = \eta\tilde{\lambda} \geq \delta + v_j + \eta(\tilde{\lambda} - \lambda) \geq v_j + \delta - (||A_j||_1 + 1)c||\bar{v} - \bar{\tilde{v}}||_\infty \geq v_j,$$

where the last inequality follows from Eq. (7) noting that $||A_j||_1 \leq d+1$.

Lemma 1 allows us to conclude that all centroids and their bases remain the same under $\tilde{v}$. We conclude by observing that all our constructions depend on identifying the centroids only and the rewards v are used nowhere else, thus none of our proofs change. □

PROOF OF COROLLARY 1(EQUATION (10)). The primal and dual problems with a single resource ($d=1$) are

$$\begin{array}{lc} \max & v'y \\ \text{s.t.} & \sum_j y_j + s = R \\ & y + u = D \\ & y_j, u_j, s \geq 0 \end{array} \qquad \begin{array}{lr} \min & R\lambda_0 + \sum_j D_j \lambda_j \\ \text{s.t.} & \lambda_0 + \lambda_j \geq v_j \quad \forall j \in [n] \\ & \lambda_0, \lambda_j \geq 0. \end{array}$$

Here λ_0 denotes the resource multiplier and λ_j the demand multipliers for $j \in [n]$. Consider a basis of the form $\mathcal{B} = \{y_j : j = 1, \ldots, k+1\} \cup \{u_j : j = k+1, \ldots, n\}$, i.e., all requests $j \in [k]$ are completely accepted, $j = k+1$ is partially accepted, and all other requests are completely rejected. By inspection, the dual variables associated to this basis are as follows: $\lambda_0 = v_{k+1}$, $\lambda_j = v_j - v_{k+1}$ for $j \in [k]$ and $\lambda_j = 0$ for $j > k$.

We now verify conditions (i)-(iii) of Definition 3. For (i), we need $v_{k+1} \geq \delta$. For (ii), we need $v_j - v_{k+1} \geq \delta$ for $j \in [k]$. Finally, for (iii) $v_{k+1} \geq v_j + \delta$ for $j = k+2, \ldots, n$. This establishes that δ-complementarity reduces in this case to $v_j \geq \delta$ for all $j \in [n]$ and $|v_j - v_{j'}| \geq \delta$ for all $j \neq j'$. □

Learning the rewards. If v is not known, the controller must make decisions while learning the correct rewards v from their random realizations. In this setting, type-j requests draw a reward $V_j \sim F_j$, where F_j is some unknown distribution, and $v_j = \mathbb{E}[V_j]$ represents the true expectation. At each time, she observes the type $j \in [n]$ and, if the request is accepted, the controller observes a realization of V_j, and uses it to estimate v through its empirical average.

COROLLARY 3 **(learning the reward distribution)**. *Assume that all the bases are δ-complementary for some $\delta > 0$. Further assume that the distributions F_j are sub-Gaussian. Then, a modification of* BUDGETRATIO *achieves $\mathcal{O}(\log T)$ regret.*

PROOF. Fix the constant $\epsilon = \frac{\delta}{c(d+2)}$. We again use a "learn, then act" policy. We set an an initial exploration phase of length $c' \log T$, for some $c' = c'(\epsilon) > 0$, where all requests are accepted and we build the empirical estimates $\hat{v}_j^t = \frac{\sum_{\tau=1}^t V^\tau \mathbb{1}_{\{J^\tau = j\}}}{\sum_{\tau=1}^t \mathbb{1}_{\{J^\tau = j\}}}$, where V^τ is the reward observed at time τ and J^τ is the type of the request at time τ. By standard concentration results, we can choose c so that $\mathbb{P}[||v - \hat{v}^{c \log T}|| > \epsilon] \leq 1/T$. Starting at $t = c' \log T$ we run BUDGETRATIO. Constant regret is guaranteed by Theorem 1, specifically equation (7) there. On the event that $||v - \hat{v}^{c \log T}|| > \epsilon$ the regret is at most $T \max_{j \in [n]} v_j$; this event's contribution to regret is at most $T \max_{j \in [n]} v_j \times 1/T = \max_{j \in [n]} v_j = \mathcal{O}(1)$. □

7. BudgetRatio as a max-bid-price. The proposition below concerns item (4) of Theorem 1. It formalizes the equivalence between the primal and the bid-price versions of BudgetRatio.

Proposition 4. *Assume that all the bases are δ-complementary (Definition 3) for some $\delta > 0$. Then, primal* BudgetRatio *is equivalent to the max-bid price* BudgetRatio *in Definition 2: on any realization of $Z, \mathfrak{Z}$ and at any time t,* BudgetRatio *as specified in Algorithm 1 accepts an arriving request of type j, if and only if the max-bid price algorithm in Definition 2 does.*

Proof. Suppose that $R^t \in \mathcal{N}_{\mathcal{K}}(p)$. We divide the analysis into $j \in \mathcal{K}$ (acceptance) and $j \notin \mathcal{K}$ (rejection). The case $j \in \mathcal{K} \neq \emptyset$ follows easily since, at the centroid budget $r_{\mathcal{K}}(p)$, $y_j = p_j > 0$ for all bases $\mathcal{B}$ associated to $\mathcal{K}$—so that $y_j \geq p_j/2$ for any $R \in \mathcal{N}_{\mathcal{K}}(p)$—we have by complementary slackness that $v_j \geq \bar{A}_j'\lambda$ for the dual variable associated with any of these bases. We note that because $j \in \mathcal{K}$, then $j \notin \partial(\mathcal{K})$ so that $\bar{A}_j'\lambda_{\mathcal{K}}^{\partial} = 0$. Overall, $v_j \geq \max_{\lambda \in \Lambda(R^t)} \bar{A}_j'(\lambda + \lambda^{\partial}(R^t))$ so that max-bid price accepts request j.

We are left to prove the case $j \notin \mathcal{K}$ and hence not accepted by primal BudgetRatio.

We claim that if $j \notin \mathcal{K}$, either (i) $j \in \partial(\mathcal{K})$ in which case $(\lambda_{\mathcal{K}}^{\partial})_j \geq 2v_j\mathbf{e}_j$ so that $v_j < \max_{\lambda \in \Lambda(R^t)} \bar{A}_j'(\lambda + \lambda^{\partial}(R^t))$ and j is rejected by max-bid price, or (ii) there is a basis $\mathcal{B}$ associated with $\mathcal{K}$ such that $y_j \notin \mathcal{B}$. If that is the case, δ-complementarity yields $v_j \leq \bar{A}_j'\lambda - \delta$, so that j fails acceptance condition of the policy with λ being the dual vector associated to this basis $\mathcal{B}$. In turn, $v_j < \max_{\lambda \in \Lambda(R^t)} \bar{A}_j'\lambda$ and max-bid price, as well, rejects request j.

It only remains to prove, then, that if $j \notin \mathcal{K}$ and $j \notin \partial(\mathcal{K})$ there exists an optimal basis associated with $\mathcal{K}$ for which $y_j \notin \mathcal{B}$.

Suppose that $y_j \in \mathcal{B}$ for all bases associated with $\mathcal{K}$. Because $j \notin \mathcal{K}$, it must be that $j \in K^+(\mathcal{B})$ for all bases associated with $\mathcal{K}$. Because $j \notin \partial(\mathcal{K})$, there exists $\zeta > 0$ such that $r_{\mathcal{K}} \geq \zeta A_j$. By the Lipschitz continuity of LPs $R = r_{\mathcal{K}} \pm \zeta A_j \in \mathcal{N}_{\mathcal{K}}(p)$ for all ζ sufficiently small ($y_j \geq D_j - M\zeta \geq D_j/2$ for all $j \in \mathcal{K}$). Let $\mathcal{B}$ be the optimal basis at $R = r_{\mathcal{K}} - \frac{1}{2}\zeta A_j$ and let (y, u, s) be the optimal solution. Because $y_j \in K^+(\mathcal{B})$ we have by Lemma 4 that

$$(\mathcal{B}^{-1} r_{\mathcal{K}})_j = \left(\mathcal{B}^{-1}\begin{pmatrix} R - \frac{1}{2}\zeta A_j \\ p \end{pmatrix} + \mathcal{B}^{-1}\begin{pmatrix} \frac{1}{2}\zeta A_j \\ 0 \end{pmatrix}\right)_j = y_j + \frac{1}{2}\zeta > 0;$$

a contradiction to the fact that $j \notin \mathcal{K}$. □

Remark 7 (Adaptively generating bid-prices). In Remark 4 we discussed how the geometry—the centroids and associated bases—can be pre-computed and, in turn, so can the max-bid prices. There is an adaptive alternative to this, possibly expensive, pre-computation.

The initial centroid $\mathcal{K}$ (the one for which $\mathbb{E}[R^0] \in \mathcal{K}$) is easily identifiable. It contains the requests that have $y_j \geq \bar{p}_j/2$ in the solution to $LP(\mathbb{E}[R^0], p)$. We need to solve at most n LPs (as the number of request types) to identify all the centroid neighbors of (hence all the bases associated to) $\mathcal{K}$. It is only upon entry to a new centroid, that we must compute the new collection of bases.

Proposition 3 (and Theorems 2 and 3 that support it) imply that R^t spends most of its horizon in $\mathcal{N}_{\mathcal{K}}$ and its neighbors. By the proof of Theorem 3 at most $d+1$ centroid action regions will be visited for most of the horizon. With high likelihood, then, at most $(d+1)n$ LP computations would be required over the horizon. ■

8. Concluding remarks We consider a family of resource allocation problems and, extending existing results, show that a simple resolving algorithm achieves constant regret in terms of the total rewards collected. We provide a new proof that is geometric and based on a parametric characterization of the packing LP.

Our fundamental definition is that of centroids, which correspond to subsets of requests that should be fully accepted. For each demand D, a centroid $\mathcal{K}$ is associated with an action region $\mathcal{N}_{\mathcal{K}}(D)$. The key to our analysis is to understand how the budget-ratio process $(R^t, t \in [T])$ evolves relative to these regions $\mathcal{N}_{\mathcal{K}}(D)$ and their basic subsets $\mathcal{N}_{\mathcal{K}}(\mathcal{B}, D)$.

This geometric stochastic-process view has appealing explanatory power. By showing how the process is attracted to the the "basic" subset $\mathcal{N}_{\mathcal{K}}(\mathcal{B}, D)$ consistent with the offline basis, we uncover the mechanism used by the online policy to dynamically build a nearly optimal solution to the offline problem.

Relying on this infrastructure, we were able to identify the "robustness boundaries" of BudgetRatio.

1. *Modelling assumptions*: Inventory arrivals are allowed but they must be slow. In the absence of such an assumption, no algorithm can achieve constant regret. In the presence of slow-restock, a suitably tuned BudgetRatio achieves constant regret. Within slow restock, the aggressiveness parameter $\alpha \in (0,1)$ can be tuned to the rate of restock.
2. *Implementation:* We proved that—under a suitable complementarty assumption, BudgetRatio is equivalent to a max-bid price control. A request is accepted if its reward exceeds a maximum of several shadow prices; these correspond to all the bases associated with the centroid. The max-bid prices can be adaptively and efficiently computed.
3. *Model parameters*: We considered the effect of running BudgetRatio with misspecified arrival probabilities (p, ϱ) and rewards v. We proved that, if the parameters (p, ϱ) and v are estimated within a constant error, BudgetRatio still achieves constant regret. Crucially, both robustness results hinge on the notion of centroids. These provide a language through which we can generalize to multiple dimensions separation conditions that were previously provided only in the single-dimensional case.

We introduced sufficient conditions that guarantee the robustness of BudgetRatio to both parameter perturbation, and restock rates. It is possible that tighter characterization are possible on both fronts:

1. **Perturbation conditions and learning.** The separation conditions in Definitions 4 and Definition 3, while sufficient, may not be necessary for the robustness of BudgetRatio. For instance, while the δ-complementarity is necessary for the one-dimensional case ($d = 1$), it is unclear whether our generalization for $d > 1$ is necessary.
2. **Restock rate.** Our analysis of restock reveals that as the rate of restock increases, the problem "transitions" from one concerning the allocation of finite inventory, to one concerning the control of so-called *loss networks* (or loss queues); the restock of inventory in our problem corresponds to the release of servers in the corresponding loss-network. In loss queues, customers that arrive and find all servers taken must be rejected. When restock rates are high, much of the *forecasted* inventory is embedded in future arrivals. Requests that we might ideally want to accept cannot be accepted because there is no on-hand inventory. Loss networks are difficult and the regret is generally of the order of $\sqrt{T}$; see the examples in Appendix A. The slow restock requirement is, then, one that guarantees that the relative simplicity of inventory-allocation problems is maintained. A more complete characterization of restock levels that permit this simplicity is interesting to pursue.

References

Shipra Agrawal and Nikhil Devanur. Linear contextual bandits with knapsacks. In *Advances in Neural Information Processing Systems*, pages 3450–3458, 2016.

Alessandro Arlotto and Itai Gurvich. Uniformly bounded regret in the multi-secretary problem. *Stochastic Systems*, 9(3):231–260, 2019.

Ashwinkumar Badanidiyuru, Robert Kleinberg, and Aleksandrs Slivkins. Bandits with knapsacks. In *2013 IEEE 54th Annual Symposium on Foundations of Computer Science*, pages 207–216, 2013.

Hannah Bast, Daniel Delling, Andrew Goldberg, Matthias Müller-Hannemann, Thomas Pajor, Peter Sanders, Dorothea Wagner, and Renato F Werneck. Route planning in transportation networks. In *Algorithm engineering*, pages 19–80. Springer, 2016.

Alberto Bemporad, Francesco Borrelli, and Manfred Morari. Model predictive control based on linear programming – the explicit solution. *IEEE transactions on automatic control*, 47(12): 1974–1985, 2002.

Dimitries Bertsimas and John N. Tsitsiklis. *Introduction to linear optimization.* Athena Scientific Belmont, MA, 1997.

Patrick Billingsley. *Convergence of probability measures*. John Wiley & Sons, 2013.

Francesco Borrelli, Alberto Bemporad, and Manfred Morari. Geometric algorithm for multiparametric linear programming. *Journal of optimization theory and applications*, 118(3):515–540, 2003.

Andrew Boyd and Ioana Bilegan. Revenue management and e-commerce. *Management science*, 49(10):1363–1386, 2003.

Sébastien Bubeck and Nicolo Cesa-Bianchi. Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *Foundations and Trends® in Machine Learning*, 5(1):1–122, 2012.

William Cook, Albertus Gerards, Alexander Schrijver, and Éva Tardos. Sensitivity theorems in integer linear programming. *Mathematical Programming*, 34(3):251–264, 1986.

Atilla Eryilmaz and Rayadurgam Srikant. Asymptotically tight steady-state queue length bounds implied by drift conditions. *Queueing Systems*, 72(3-4):311–359, 2012.

Tomáš Gal. Linear parametric programming – a brief survey. In *Sensitivity, Stability and Parametric Analysis*, pages 43–68. Springer, 1984.

Stefanus Jasin and Sunil Kumar. A re-solving heuristic with bounded revenue loss for network revenue management with customer choice. *Mathematics of Operations Research*, 37(2):313–345, 2012.

Stefanus Jasin and Sunil Kumar. Analysis of deterministic lp-based booking limit and bid price controls for revenue management. *Operations Research*, 61(6):1312–1320, 2013.

Siva Maguluri and R Srikant. Heavy traffic queue length behavior in a switch under the maxweight algorithm. *Stochastic Systems*, 6(1):211–250, 2016.

Olvi Mangasarian and Tzong Shiau. Lipschitz continuity of solutions of linear inequalities, programs and complementarity problems. *Siam J. Control and Optimization*, 25(3), 1987.

Erica Plambeck and Amy Ward. Optimal control of a high-volume assemble-to-order system. *Mathematics of Operations Research*, 31(3):453–477, 2006. doi: 10.1287/moor.1060.0196.

Sheldon M Ross. *Stochastic processes*. John Wiley & Sons, 1995.

Jing-Sheng Song and Paul Zipkin. Supply chain operations: Assemble-to-order systems. *Handbooks in operations research and management science*, 11:561–596, 2003.

Kalyan Talluri and Garrett Van Ryzin. An analysis of bid-price controls for network revenue management. *Management science*, 44(11-part-1):1577–1593, 1998.

Kalyan Talluri and Garrett Van Ryzin. *The theory and practice of revenue management*. International Series in Operations Research & Management Science, 68. Kluwer Academic Publishers, Boston, MA, 2004. ISBN 1-4020-7701-7.

Alberto Vera and Siddhartha Banerjee. The bayesian prophet: A low-regret framework for online decision making. *Management Science*, 67(3):1368–1391, 2021.

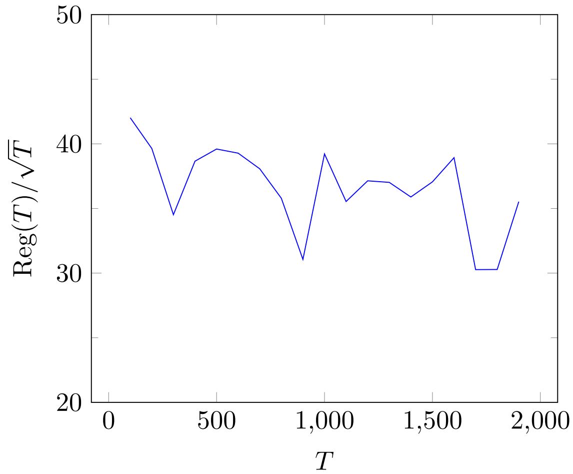

FIGURE 9. Regret of an instance with $n = 3$ and fast restock for increasing time horizon T. The regret is $V^*_{\text{off}}(T) - V^*_{\text{on}}(T)$, where we compute $V^*_{\text{on}}(T)$ via dynamic programming. We conclude that the regret scales as $\Omega(\sqrt{T})$

Alberto Vera, Siddhartha Banerjee, and Itai Gurvich. Online allocation and pricing: Constant regret via bellman inequalities. *Operations Research*, 69(3):821–840, 2021.

Huasen Wu, R Srikant, Xin Liu, and Chong Jiang. Algorithms with logarithmic or sublinear regret for constrained contextual bandits. In *Advances in Neural Information Processing Systems*, pages 433–441, 2015.

Appendix A: Restock and feasibilty of constant regret. If Assumption 1 fails, it is not generally possible to achieve constant regret relative to the offline (2). Consider, as an example, a problem with a single resource and three customer types with rewards $(v_1, v_2, v_3) = (200, 100, 0)$, arrival probabilities $p = (0.4, 0.2, 0.4)$ and restock probability $\varrho = 0.41$. There is no initial budget ($I^0 = 0$). This example violates Assumption 1 because $\varrho > r_{\{1\}} = 0.4$.

We computed the *optimal* dynamic-programming policy for horizons $T = 1, \ldots, 2000$. For each horizon, we ran 1000 replication of both the optimal policy and the offline upper bound in (2). In Fig. 9 we display, for each T, the average (over replications) gap between the two. Evidently, the regret of the optimal policy—in turn, of any online policy—is proportional to $\sqrt{T}$.

If $\varrho < 0.4$, the slow-restock assumption is satisfied and constant regret is guaranteed by Theorem 1. With $\varrho > 0.6$ there are enough resources for both types 1 and 2 and it is easy to show that the regret is constant. With further simulations it can be verified that any $\varrho \in [0.4, 0.6]$ does not produce constant regret. In other words, the set of restock probabilities that imply constant regret is the *disconnected* set $[0, 0.4) \cup (0.6, 1]$.

Below is another example with a single type of requests where we can prove that the regret exhibits $\sqrt{T}$ grows. A slow restock assumption is necessary here; when it fails, offline is no longer a useful benchmark for performance measurement.

LEMMA 15. *There exists a resource allocation network that violates Assumption 1 and such that, for some* $c>0$,

$$V^*_{off}(T) - V^*_{on}(T) = c\sqrt{T} + o(\sqrt{T}), \text{ as } T \to \infty,$$

where $V^*_{on}(T)$ *is the value of the optimal policy for the horizon* $[T]$. *Hence, no policy can achieve* $o(\sqrt{T})$ *regret.*

This counter example is a network with one request type and one resource, both arriving with probability p. It is a so-called two-sided queue where one side balks immediately if not served upon arrival. Fig. 10 is the numerical illustration of the $\sqrt{T}$ regret.

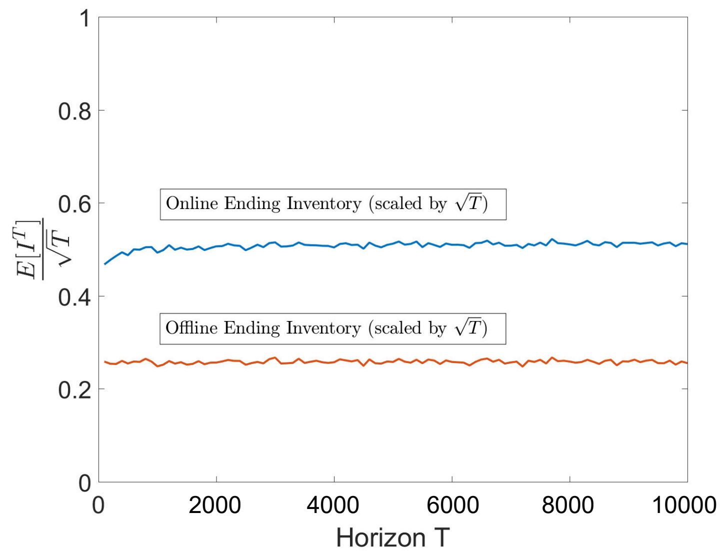

FIGURE 10. Single RAN with restock. The first line depicts the expected remaining inventory of offline. The second line depicts the expected remaining inventory of the optimal online policy. The difference is the regret. The expectation is computed as an average over 10000 replications. The y-axis is the (expected) ending inventory divided by $\sqrt{T}$.

We label the request by 1, the resource by a, and set $v_1 = 1$, $p_1 = \varrho_a = p$. Let Z^t be the number of request arrivals by time t and let $\mathfrak{z}^t$ be the restock by time t. Let I^T_{off} (respectively I^T_{on}) be the end-of-horizon residual inventory under offline (respectively online) policies. Then, $V^*_{\text{off}}(T) = \mathbb{E}[\min\{\mathfrak{z}^t, Z^t\}] = \mathbb{E}[\mathfrak{z}^T - I^T_{\text{off}}]$.

The optimal online policy servs any arriving request if there is inventory available. Let Y^t be the number of requests accepted by the online policy by (and including) time t. Then, $V^*_{\text{on}} = \mathbb{E}[Y^T] = \mathbb{E}[\mathfrak{z}^T - I^T_{\text{on}}]$. Thus, $V^*_{\text{off}}(T) - V^*_{\text{on}}(T) = \mathbb{E}[I^T_{\text{on}}] - \mathbb{E}[I^T_{\text{off}}]$.

We study the two (end-of-horizon) inventory levels, starting with offline, which satisfies:

$$I^T_{\text{off}} = (\mathfrak{z}^T - Z^T)^+.$$

The process $\mathfrak{Z}^t - Z^t$ is a random walk starting at 0 and with i.i.d zero-mean increments $X_1, \ldots, X_T$ taking values $\{-1,0,1\}$ with probabilities $p(1-p), 1-2p(1-p), p(1-p)$; X_t is the difference between the restock at t (0 or 1) and the request arrival (0 or 1). Write $G^T = \sum_{t=1}^T X_t$. By the central limit theorem

$$\frac{1}{\sqrt{T}} G^T \Rightarrow \mathcal{N}(0,\sigma^2), \text{ as } T \uparrow \infty,$$

where $\sigma^2 := 2p(1-p)$. Since $I_{\text{off}}^T = (G^T)^+$, we have by the continuous mapping theorem that $\frac{1}{\sqrt{T}} I_{\text{off}}^T \Rightarrow (\mathcal{N}(0,\sigma^2))^+$, and the convergence here also holds in expectation. On the other hand, the online inventory satisfies the queueing recursion $I_{\text{on}}^{t+1} = [I_{\text{on}}^t + X_t]^+$ so that

$$I_{\text{on}}^T = \sup_{t \le T} (G^T - G^t).$$

The so-called reflection principle implies the equivalence in law

$$\sup_{t \le T} (G^T - G^t) \overset{\mathcal{L}}{=} \sup_{t \le T} G^t.$$

We also have that

$$\frac{1}{\sqrt{T}} \sup_{t \le T} G^t \Rightarrow \mathcal{Z}, \text{ as } T \uparrow \infty,$$

where $\mathcal{Z}$ is distributed as the supremum of a Brownian motion over $[0,1]$. This standard result follows from Donsker's theorem [Billingsley, 2013] and the continuity of the supremum map in the space of continuous functions. This convergence also holds in expectation. The reflection principle for Brownian motion then guarantees that, for $a \ge 0$, $\mathbb{P}\{\mathcal{Z}' \ge a\} = 2\mathbb{P}\{\mathcal{N}(0,\sigma^2) \ge a\}$ so that $\mathbb{E}[\mathcal{Z}'] = 2\mathbb{E}[(\mathcal{N}(0,\sigma^2))^+]$ and this allows us to conclude that

$$\mathbb{E}[I_{\text{on}}^T]/\mathbb{E}[I_{\text{off}}^T] \to 2 \quad \text{ and } \quad \frac{1}{\sqrt{T}} (\mathbb{E}[I_{\text{on}}^T] - \mathbb{E}[I_{\text{off}}^T]) \to \mathbb{E}[\mathcal{N}(0,\sigma^2)^+] > 0, \quad \text{ as } T \to \infty.$$

Appendix B: Proofs of Lemmas.

PROOF OF LEMMA 1. The dual problem of LP(R,D) is

$$\min\{(R,D)'\lambda : \bar{A}'\lambda \ge \bar{v}, \lambda \ge 0\}.$$

For any basis $\mathcal{B}$, our defined vector $\lambda = (\mathcal{B}^{-1})'\bar{v}_{\mathcal{B}}$ is dual-feasible if it satisfies conditions (i) and (ii). The associated primal variables are $x'_{\mathcal{B}} = (y,u,s)'_{\mathcal{B}} = \mathcal{B}^{-1}\begin{pmatrix} R \\ D \end{pmatrix}$. By construction $(R,D)'\lambda = x'_{\mathcal{B}}\bar{v}$ so that, by weak duality, $\mathcal{B}$ is optimal provided that $x_{\mathcal{B}}$ is primal feasible, i.e., $x'_{\mathcal{B}} = \mathcal{B}^{-1}\begin{pmatrix} R \\ D \end{pmatrix} \ge 0$; see [Bertsimas and Tsitsiklis, 1997, Corollary 4.2]. Conversely, for any (R,D), when the simplex algorithm terminates it produces an optimal basis $\mathcal{B}$ where $\lambda = (\mathcal{B}^{-1})'\bar{v}_{\mathcal{B}}$ is an optimal dual solution (the reduced costs are non-negative); see [Bertsimas and Tsitsiklis, 1997, Chapter 3]. Notice that the packing problem is always primal feasible because $x = (0,D,R)$ is feasible. □

Proof of Lemma 2. The first part follows by definition of $\mathcal{N}_{\mathcal{K}}(p)$. For the second part we claim that if $R^t \in \mathcal{N}_{\mathcal{K}}(p)$, then $I_i^t \geq |\{j \in \mathcal{K} : A_{ij} = 1\}|$ for some M and all $t \leq T - M$ which proves that there is sufficient inventory to serve an arriving request in $\mathcal{K}$.

Fix $j \in \mathcal{K}$ and $i \in [d]$ such that $A_{ij} = 1$. The fact that (y, u, s) solves $\mathrm{LP}(R^t, p)$ implies $Ay \leq \frac{1}{T-t} I^t + \varrho$. Since $j \in \mathcal{K}$ it must be that $y_j \geq \frac{1}{2}\bar{p}_j$, for all $j \in \mathcal{K}$. In turn, for each $i \in \mathcal{R}$ such that $A_{ij} = 1$ for some $j \in \mathcal{K}$, we have that

$$I_i^t \geq (T-t)[(Ay)_i - \varrho_i] \geq (T-t)\left[\sum_{j \in \mathcal{K}} A_{ij} \frac{1}{2}\bar{p}_j - \varrho_i\right].$$

By Assumption 1, $\sum_{j \in \mathcal{K}} A_{ij} \frac{1}{2}\bar{p}_j - \varrho_i > 0$ so that taking $M = \max_{i : \sum_{j \in \mathcal{K}} A_{ij} \geq 1} \frac{|\{j \in \mathcal{K} : A_{ij} = 1\}|}{\sum_{j \in \mathcal{K}} A_{ij} \frac{1}{2}\bar{p}_j - \varrho_i}$ we obtain the claim for all $t \leq T - M$.

We note, finally, that $1/2$ can be replaced everywhere with $\alpha \in (0,1)$ that matches the slow restock condition; i.e., such that $\rho_i < \alpha (r_{\mathcal{K}})_i$ □

Proof of Lemma 3. Let $\mathcal{B}$ be an optimal basis of $\mathrm{LP}(A_{\mathcal{K}}\hat{D}, \hat{D})$. In particular, $\mathcal{B}$ is invertible and $\lambda = (\mathcal{B}^{-1})' \bar{v}_{\mathcal{B}}$ satisfies properties (i) and (ii) in Lemma 1. We will prove that $\mathcal{B}$ is also optimal for $\mathrm{LP}(A_{\mathcal{K}}\tilde{D}, \tilde{D})$ and has an associated solution $(y, u, s) = (\tilde{D}_{\mathcal{K}}, \tilde{D}_{\mathcal{K}^c}, 0)$.

Since $\mathcal{B}$ has the basic variables $y_{\mathcal{K}}$ and $u_{\mathcal{K}^c}$, by inspection we have the following:

$$\mathcal{B}\begin{pmatrix} \tilde{D}_{\mathcal{K}} \\ \tilde{D}_{\mathcal{K}^c} \\ 0 \end{pmatrix} = \begin{pmatrix} A_{\mathcal{K}}\tilde{D}_{\mathcal{K}} \\ \tilde{D} \end{pmatrix} \Longrightarrow \mathcal{B}^{-1}\begin{pmatrix} A_{\mathcal{K}}\tilde{D}_{\mathcal{K}} \\ \tilde{D} \end{pmatrix} \geq 0.$$

Since $\mathcal{B}$ satisfies properties (i) and (ii) in Lemma 1, we have by that lemma that $\mathcal{B}$ is optimal for the right-hand side $(A_{\mathcal{K}}\tilde{D}_{\mathcal{K}}, \tilde{D})$. Also, per our derivation above, the associated solution is indeed $(y, u, s) = (\tilde{D}_{\mathcal{K}}, \tilde{D}_{\mathcal{K}^c}, 0)$. Because the set of optimal bases, as we have now shown, is identical under $\hat{D}$ and $\tilde{D}$, so are the sets of zero-valued basic variables. □

Proof of Lemma 4.

Item 1. Because $\mathcal{B}$ is optimal at $(r_{\mathcal{K}}(D), D)$ it is invertible and satisfies properties (i) and (ii) in Lemma 1. To prove that it is optimal also at (R, D), with R of the stated form, it suffices by Lemma 1 to show that $\mathcal{B}^{-1}\begin{pmatrix} R \\ D \end{pmatrix} \geq 0$.

Recall the augmented matrix $\bar{A}$ is given by

$$\bar{A} = \begin{bmatrix} A & 0 & I_d \\ I_n & I_n & 0 \end{bmatrix},$$

where the columns are associated, from left to right, to request variables $y \in \mathbb{R}^n$, unmet variables $u \in \mathbb{R}^n$ and surplus variables $s \in \mathbb{R}^d$. The basic sub-matrix $\mathcal{B}$ has a subset of these columns and can be written as

$$\mathcal{B} = \begin{bmatrix} A_{\mathcal{K} \cup K+} & 0 & I^d_{K^0} \\ I^n_{\mathcal{K} \cup K+} & I^n_{\mathcal{K}^c \cup K-} & 0 \end{bmatrix}, \tag{37}$$

where $I^n_{\mathcal{K}\cup K^+}$ has the columns of I_n corresponding to the request variables in $\mathcal{K}\cup K^+$. The matrix $\mathcal{B}$ is of dimension $(n+d)\times(n+d)$, and each column is associated to either a variable y_j, u_j, or s_i.

We write vectors of dimension $n+d$ in this same order, specifying the components associated to y, u, and s respectively from top to bottom.

By the definition of centroid, all request variables $\mathcal{K}$ are saturated at $r_{\mathcal{K}}(D)$, hence $K^+ = K^+(\mathcal{B}) \subseteq \mathcal{K}^c$; in other words, zero-valued requests cannot come from $\mathcal{K}$. Similarly, unmet variables $\mathcal{K}^c$ are saturated at $r_{\mathcal{K}}(D)$, therefore $K^- = K^-(\mathcal{B}) \subseteq \mathcal{K}$. We deduced the inclusions $\kappa^+ \subseteq \mathcal{K}^c \cap K^+$ and $\kappa^- \subseteq \mathcal{K}\cap K^-$. By inspection we then have the identities

$$\mathcal{B}\begin{pmatrix} D_{\kappa^-} \\ -D_{\kappa^-} \\ 0\end{pmatrix} = \begin{pmatrix} A_{\kappa^-}D_{\kappa^-} \\ 0\end{pmatrix}, \qquad \mathcal{B}\begin{pmatrix} D_{\kappa^+} \\ -D_{\kappa^+} \\ 0\end{pmatrix} = \begin{pmatrix} A_{\kappa^+}D_{\kappa^+} \\ 0\end{pmatrix}, \qquad \mathcal{B}\begin{pmatrix} 0 \\ 0 \\ b_{K^0}\end{pmatrix} = \begin{pmatrix} b_{K^0} \\ 0\end{pmatrix}.$$

Pre-multiplying these identities by $\mathcal{B}^{-1}$, and taking R of the stated form, we have

$$\begin{aligned}\mathcal{B}^{-1}\begin{pmatrix} R \\ D\end{pmatrix} &= \mathcal{B}^{-1}\left[\begin{pmatrix} A_{\mathcal{K}}D_{\mathcal{K}} \\ D\end{pmatrix} + \alpha\begin{pmatrix} A^+_\kappa D^+_\kappa \\ 0\end{pmatrix} - \alpha\begin{pmatrix} A^-_\kappa D^-_\kappa \\ 0\end{pmatrix} + \begin{pmatrix} b \\ 0\end{pmatrix}\right] \\ &= \begin{pmatrix} D_{\mathcal{K}} \\ D_{\mathcal{K}^c} \\ 0\end{pmatrix} + \alpha\begin{pmatrix} D_{\kappa^+} \\ -D_{\kappa^+} \\ 0\end{pmatrix} - \alpha\begin{pmatrix} D_{\kappa^-} \\ -D_{\kappa^-} \\ 0\end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ b_{K^0}\end{pmatrix}.\end{aligned}$$

Because $\kappa^+ \subseteq \mathcal{K}^c\cap K^+$ and $\kappa^- \subseteq \mathcal{K}\cap K^-$, the right-hand side above is non-negative.

Item 2. Assume R has the stated form and let us prove that $\mathcal{B}$ is optimal. If two right-hand sides have the same optimal candidate basis, then, by virtue of Lemma 1, any non-negative combination of the right-hand sides has the same optimal basis. By the first item of the lemma, we are taking non-negative combinations of right-hand sides which have $\mathcal{B}$ as optimal basis, so we conclude optimality.

We turn to prove that, if $\mathcal{B}$ is optimal, then R has the stated representation. Let $(y,\ u,\ s)' = \mathcal{B}^{-1}\begin{pmatrix} R \\ D\end{pmatrix}$, where y are the request variables, u are unmet variables s are surplus variables. Let $K^+ = K^+(\mathcal{B})$, $K^- = K^-(\mathcal{B})$ and $K^0 = K^0(\mathcal{B})$ be as in Definition 8. By the definition of centroid and the optimality of $\mathcal{B}$, we have the following:

$$\begin{aligned} &y_j = D_j \text{ and } u_j = 0 \quad \forall j\in\mathcal{K}\backslash K^-, \\ &y_j = 0 \text{ and } u_j = D_j \quad \forall j\in\mathcal{K}^c\backslash K^+.\end{aligned}$$

For all other indices j, $u_j = D_j - y_j$ and $y_j, u_j \geq 0$. Since $R\in\mathcal{N}_{\mathcal{K}}$ we also have that $y_j \geq D_j/2$ for all $j\in K^-$ and $y_j < D_j/2$ for all $j\in K^+$. From the vector (y,u,s) we subtract the vector $(\bar{y},\bar{u},\bar{s})$ given by

$$\bar{y}_j = \begin{cases} D_j & \text{for } j\in\mathcal{K}\backslash K^-, \\ 0 & \text{for } j\in\mathcal{K}^c\backslash K^+, \\ D_j/2 & \text{for } j\in K^- \\ 0 & \text{otherwise.}\end{cases} \qquad \bar{u}_j = \begin{cases} D_j/2 & \text{for } j\in K^-, \\ D_j & \text{for } j\in K^+\end{cases},$$

and $\bar{s} = s$ (subtracting fully the budget slack variables). Since $R \in \mathcal{N}_{\mathcal{K}}(D)$ we have, by definition, that $y \geq \bar{y}$ and $u \leq \bar{u}$. Thus, $(y-\bar{y}, u-\bar{u}, s-\bar{s}) = (y-\bar{y}, u-\bar{u}, 0) \geq 0$. We will study next the vector $z = (y-\bar{y}, u-\bar{u})$.

For convenience, let us re-label (and re-order) the indices so that indices in $K^+ \cup K^-$ are at the top of the vector (y, u, s). The vector z then has the form

$$z = \begin{pmatrix} y_{K^+} \\ y_{K^-} - D_{K^-}/2 \\ u_{K^+} - D_{K^+} \\ u_{K^-} - D_{K^-}/2 \end{pmatrix} = \begin{pmatrix} y_{K^+} \\ y_{K^-} - D_{K^-}/2 \\ -y_{K^+} \\ D_{K^-}/2 - y_{K^-} \end{pmatrix}.$$

We will identify a representation for z that will help us show that R has the desired form. Since all other entries of (y, u, s) have fixed values, we will then append those to all vectors in the resulting combination.

We apply the following transformation

$$x = Pz \text{ where } P = 2\,\mathrm{diag}(1/D_{K^+}, 1/D_{K^-}, 1/D_{K^+}, 1/D_{K^-}).$$

By definition, all request elements of Pz are in $[0,1]$ and unmet elements are in $[-1,0]$. If x can be written as a convex combination of vectors $x_1, \ldots, x_m$ then $z = (y-\bar{y}, u-\bar{u})$ can be written as a convex combination of $P^{-1}x_1, \ldots, P^{-1}x_m$.

Vectors $x = Pz$ are elements in the polyhedron

$$\left\{ \begin{pmatrix} x_{K^+} \\ x_{K^-} \\ \mathfrak{s}_{K^+} \\ \mathfrak{s}_{K^-} \end{pmatrix} : x_j + \mathfrak{s}_j = 0, x_j \in [0,1], \mathfrak{s}_j \in [0,1] \right\}.$$

This polyhedron is integral because the constraint matrix is totally unimodular consisting, as it does, of only $\{0,1\}$ entries and having a single 1 per column. In turn, we can write each such vector as a convex combination of *binary* vectors of the form

$$\begin{pmatrix} x_{K^+} \\ x_{K^-} \\ -x_{K^+} \\ -x_{K^-} \end{pmatrix},$$

where $x_j \in \{0,1\}$. For such a vector x we have a set $\kappa^+ \subseteq K^+$ of entries such that $x_j = 1$ for $j \in \kappa^+$ and a set $\kappa^- \subseteq K^-$ with $x_j = 0$ for $j \in \kappa^-$. Thus, each of these binary vectors can be written as

$$\begin{pmatrix} e_{\kappa^+} \\ 0_{K^+ \setminus \kappa^+} \\ 0_{\kappa^-} \\ e_{K^- \setminus \kappa^-} \\ -e_{\kappa^+} \\ 0_{K^+ \setminus \kappa^-} \\ 0_{\kappa^-} \\ -e_{K^- \setminus \kappa^-} \end{pmatrix},$$

for some subsets $\kappa^+ \subseteq K^+$ and $\kappa^- \subseteq K^-$. Transforming back (multiplying by D^{-1} and adding $\bar{y}, \bar{u}$), we have that we can write (y,u,s) as a convex combination of vectors of the form

$$\begin{pmatrix} D_{\kappa^+}/2 \\ 0_{K^+\setminus\kappa^+} \\ D_{\kappa^-}/2 \\ D_{\mathcal{K}\setminus\kappa^-} \\ D_{\kappa^+}/2 \\ D_{K^+\setminus\kappa^+} \\ D_{\kappa^-}/2 \\ 0_{\mathcal{K}\setminus\kappa^-} \\ s \end{pmatrix}.$$

Notice that multiplying this vector by $\mathcal{B}$ we get a vector of the form

$$r_{\kappa^+,\kappa^-,u} = r_{\mathcal{K}} + A_{\kappa^+}D_{\kappa^-}/2 - A_{\kappa^-}D_{\kappa^+}/2 + s$$

where we use the fact that $\mathcal{B}s$ (multiplying by vector of surplus) gives back the surplus. We conclude that we can write the top elements of y as a sum of a vector s and a convex combination of vectors (y,u) of the desired form.

Item 3. We just proved that $\mathcal{B}$ is optimal for (R,D) if and only if it can be written as

$$R = r_{\mathcal{K}}(D) + \sum_{\kappa^+\subseteq K^+,\kappa^-\subseteq K^-} \alpha_{(\kappa^+,\kappa^-)}(A_{\kappa^+}D_{\kappa^+} - A_{\kappa^-}D_{\kappa^-}) + b.$$

Observe that the sum ranges over subsets of K^+, K^-. Let us group it instead for each $j \in K^+ \cup K^-$. With this end, define

$$\alpha_j := \sum_{(\kappa^+,\kappa^-):j\in\kappa^+} \alpha(\kappa^+,\kappa^-) \text{ for } j \in K^+ \quad \text{and} \quad \alpha_j := \sum_{(\kappa^+,\kappa^-):j\in\kappa^-} \alpha(\kappa^+,\kappa^-) \text{ for } j \in K^-.$$

Now, if we put $x_j := \alpha_j D_j$, we can write $R = r_{\mathcal{K}}(D) + A_{K^+}x_{K^+} - A_{K^-}x_{K^-}$. We *claim* that the solution (y,u,s) associated to the right-hand side (R,D) is

$$\begin{pmatrix} y \\ u \\ s \end{pmatrix} = \begin{pmatrix} D_{\mathcal{K}} \\ D_{\mathcal{K}^c} \\ 0 \end{pmatrix} + \begin{pmatrix} x_{K^+} - x_{K^-} \\ -x_{K^+} + x_{K^-} \\ b \end{pmatrix}.$$

Assuming this claim, we can conclude since, by definition, $R \in \mathcal{N}_{\mathcal{K}}(D,\mathcal{B})$ if (1) the basis $\mathcal{B}$ is optimal and (2) we have $y_{\mathcal{K}} \geq \frac{1}{2}D_{\mathcal{K}}$ and $y_{\mathcal{K}^c} < \frac{1}{2}D_{\mathcal{K}^c}$. Indeed, condition (2) follows by recalling $K^+ \subseteq \mathcal{K}^c$, $K^- \subseteq \mathcal{K}$ and our definition $x_j = \alpha_j D_j$.

We are left to prove the claim. Using that the variables $\{y_j, u_j : j \in K^- \cup K^+\}$ and s_{K^0} are in the basis $\mathcal{B}$ we have

$$\mathcal{B}\begin{pmatrix} x_{K^+} - x_{K^-} \\ -x_{K^+} + x_{K^-} \\ b_{K^0} \end{pmatrix} = \begin{pmatrix} A_{K^+}x_{K^+} - A_{K^-}x_{K^-} + b \\ 0 \end{pmatrix}.$$

Finally, by the definition of centroid

$$\mathcal{B}\begin{pmatrix} D_{\mathcal{K}} \\ D_{\mathcal{K}^c} \\ 0 \end{pmatrix} = \begin{pmatrix} r_{\mathcal{K}}(D) \\ D \end{pmatrix}.$$

The last two equations together prove the claim by virtue of Lemma 1. □

Proof of Lemma 5. If $R \in \mathcal{N}_{\mathcal{K}}(D,\mathcal{B})$ then in particular $R \in \mathcal{N}_{\mathcal{K}}(D) = \cup_{\mathcal{B}}\mathcal{N}_{\mathcal{K}}(D,\mathcal{B})$ and from Lemma 4 (item 3) we have $R - r_{\mathcal{K}}(D) = A_{K^+}x_{K^+} - A_{K^-}x_{K^-} + b$, hence $R \in r_{\mathcal{K}}(D) + \text{cone}(\mathcal{K},\mathcal{B})$.

If $R \in \mathcal{N}_{\mathcal{K}}(D) \cap (r_{\mathcal{K}}(D) + \text{cone}(\mathcal{K},\mathcal{B}))$, then necessarily $R = r_{\mathcal{K}}(D) + A_{K^+}x_{K^+} - A_{K^-}x_{K^-} + b_{\mathcal{K}^0}$ because it is in the cone. Also, as in the proof of Lemma 4 we have that $(\mathcal{B}^{-1}R)_j = x_j$ for $j \in K^+$ and $(\mathcal{B}^{-1}R)_j = D_j - x_j$ for $j \in K^-$. Because $R \in \mathcal{N}_{\mathcal{K}}(D)$ it must then be that $x_j \leq D_j/2$ for all $j \in K^+ \cup K^-$. We can then use Lemma 4 (item 3) to conclude that $R \in \mathcal{N}_{\mathcal{K}}(D,\mathcal{B})$. □

Proof of Lemma 6. The existence of a finite family of separating vectors such that $\xi \in cone(\mathcal{K},\mathcal{B})$ if and only if $\max_{\psi\in\Psi(\mathcal{K},\mathcal{B})} \psi'\xi \leq 0$ follows from the Minkowski-Weyl theorem; see e.g. [Bertsimas and Tsitsiklis, 1997, Chapter 4.9]. We construct these explicitly for $cone(\mathcal{K},\mathcal{B})$.

Per our construction of the set $\mathcal{N}_{\mathcal{K}}(\mathcal{B},D)$ a vector ξ is in the cone if and only if ξ can be written as

$$\xi = \sum_{(\kappa^+\subseteq K^+,\kappa^-\subseteq K^-)} \alpha[\kappa^+,\kappa^-,\kappa^0](A_{\kappa^+}z_{\kappa^+} - A_{\kappa^-}z_{\kappa^-}) + b_{\kappa^0},$$

where $b, z \geq 0$. It is immediate that in the convex combination it suffices to include $\kappa^+,\kappa^-,\kappa^0$ that are minimal, i.e., with $|\kappa^+| + |\kappa^-| + |\kappa^0| = 1$. These are the extreme rays of the cone; by Lemma 4 we have, for instance, $\left(\mathcal{B}^{-1}\begin{pmatrix} A_{k^+} \\ 0\end{pmatrix}\right)_j = 0$ for all $j \neq \kappa^+$.

Take $\Psi := -(\mathcal{B}^{-1})_{K^+\cup K^-\cup K^0,d}$ (the resource columns for the rows corresponding to $K^+ \cup K^- \cup K^0$). By Lemma 4, $\psi[\kappa]'A_j = 0,\ j \in K^+(\mathcal{B})\backslash\kappa^+$ and $\psi[\kappa]'A_j = 0,\ j \in K^-(\mathcal{B})\backslash\kappa^-$, and $\psi[\kappa]'e_i = 0$, for $i \in K^0(\mathcal{B})\backslash\kappa^-$. Similarly, $\psi[\kappa]'A_{\kappa^+} = -1 < 0$ if $|\kappa^+| = 1$, $\psi[\kappa]'A_{\kappa^-} = -1 < 0$ if $|\kappa^-| = 1$, and $\psi[\kappa]'e_{\kappa^0} < 0$ if $|\kappa^0| = 1$.

We turn to the second item in the lemma. Take $j \in K^+(\bar{\mathcal{B}})$ and $R = r_{\mathcal{K}}(D) + A_j$. Then, $\bar{\mathcal{B}}^{-1}\begin{pmatrix} R \\ 0\end{pmatrix} \geq 0$ by Lemma 4. As in the proof of Lemma 9 for $j \in K^+(\mathcal{B}) \cap K^+(\bar{\mathcal{B}})$, we have

$$\left(\mathcal{B}^{-1}\begin{pmatrix} R \\ 0\end{pmatrix}\right)_j = \left(\bar{\mathcal{B}}^{-1}\begin{pmatrix} R \\ 0\end{pmatrix}\right)_j \geq 0,$$

so recalling that $\Psi := -(\mathcal{B}^{-1})_{K^+\cup K^-\cup K^0,d}$ we have that $\psi[\kappa]'A_j \leq 0$ as required. The argument is the same for $j \in K^-(\bar{\mathcal{B}})$ or $j \in K^0(\bar{\mathcal{B}})$. □

Proof of Lemma 7. The basis $\mathcal{B}$ is fixed for the proof and we write K^+, K^-, K^0 for the corresponding sets in Definition 8. Let $(\bar{y},\bar{u},\bar{s})$ be the solution to LP(R^t,p) and define

$$\mathcal{Y} = \{(y,u,s) : \exists R \in \mathcal{N}_{\mathcal{K}}(\mathcal{B},p) \text{ s.t. } (y,u,s) \text{ solves LP}(R,p)\}.$$

By assumption $d_\infty(R^t, \mathcal{N}_\mathcal{K}(\mathcal{B},p)) \leq \frac{\epsilon^0}{M}$. By the Lipschitz continuity of the LP solution, we can choose M large enough (depending on A) such that $d_\infty((\bar{y},\bar{u},\bar{s}),\mathcal{Y}) \leq \epsilon$. Let $(y^0,u^0,s^0) \in \mathcal{Y}$ be such that $d_\infty((\bar{y},\bar{u},\bar{s}),(y^0,u^0,s^0)) \leq \epsilon^0$. Since for all $j \in \mathcal{K}\backslash K^-$ we have that $y^0_j = p_j$ then we also have that $\bar{y}_j \geq p_j - \epsilon^0 \geq p_j/2$ so that all these requests are accepted. Also, for any $j \notin \mathcal{K} \cup K^+$, we have that $u^0_j = p_j$ so that $\bar{u}^0_j \geq p_j - \epsilon^0$ and hence $\bar{y}^0_j < p_j/2$ so these requests are rejected. Hence, the policy is making basic allocations at t.

Finally, $s^0_i = 0$ for all $i \notin K^0$, hence $\bar{s}_i = R^t_i - (A\bar{y})_i \leq \epsilon^0$ for all such i, which implies $R^t_i = \frac{1}{T-t} I^t_i + \varrho_i \leq (Ay)_i + \epsilon$ and using $y \leq p$ we get, as required, that $I^t_i \leq (T-t)((Ap)_i + \epsilon^0)$ for $i \notin K^0$. □

Proof of Lemma 8. The basis $\mathcal{B}$ is fixed for the proof and we write K^+, K^-, K^0 for the corresponding sets in Definition 8. We will argue that that for any ϵ'' we can choose ϵ' small enough so that under the assumptions of the lemma:

(i) $y_j \geq D_j/2 - \epsilon''$ for all $j \in \mathcal{K}$, and

(ii) $y_j \leq \epsilon''$ for all $j \notin \mathcal{K} \cup K^+$, and $u_j \leq \epsilon''$ for all $j \in \mathcal{K}\backslash K^-$.

We would then have that $R_\bullet := R - s = \bar{R} \pm \|A\|_\infty \epsilon''$ where $\bar{R} = r_\mathcal{K}(D) + A_{K^+}x_{K^+} - A_{K^-}x_{K^-}$ for some x with $x_j \in [0, D_j/2]$ for all $j \in K^+ \cap K^-$; hence $\bar{R} \in \mathcal{N}_\mathcal{K}(\mathcal{B},D)$. Choosing ϵ' (and subsequently ϵ'') so that $\epsilon'' \leq \frac{\epsilon}{\|A\|_\infty}$, we will conclude that $d_\infty(R_\bullet, \mathcal{N}_\mathcal{K}(\mathcal{B},D)) \leq \epsilon$. Finally, we have that $d_\infty(R_\bullet, \mathcal{N}_\mathcal{K}(\mathcal{B},D)) \leq \epsilon$, if and only if $d_\infty(R, \mathcal{N}_\mathcal{K}(\mathcal{B},D)) \leq M\epsilon$; see the proof of Lemma 10.

Item (i): Because, $d_\infty(R, \mathcal{N}_\mathcal{K}(D)) \leq d(R, \mathcal{N}_\mathcal{K}(D)) \leq \epsilon'$, the Lipschitz continuity of LPs implies that $y_j \geq D_j/2 - M\epsilon'$ for all $j \in K^-$ and some constant M that depends on (p, ϱ, A, v). Taking $\epsilon' = \epsilon''/M$ proves this item.

Item (ii): First, by the continuity of LPs, we have that $y_j \leq \epsilon''$ for all $j \notin \mathcal{K} \cup (\cup_{\mathcal{B}'} K^+(\mathcal{B}'))$ where the union is overall bases $\mathcal{B}'$ associated with $\mathcal{K}$. For the same reason $u_j \leq \epsilon''$ for all $j \in \mathcal{K}\backslash(\cup_{\mathcal{B}'} K^-(\mathcal{B}'))$. It remains only to consider bases that are associated to $\mathcal{K}$.

Take $R \notin \mathcal{N}_\mathcal{K}(\mathcal{B},D)$ that satisfies the assumptions and consider two cases:

First case ($\max_{\psi \in \Psi(\mathcal{K},\mathcal{B})} \psi'(R - r_\mathcal{K}(D)) \leq 0$): In this case $R - r_\mathcal{K}(D) \in cone(\mathcal{K},\mathcal{B})$. Recall the equivalent definition of $\mathcal{N}_\mathcal{K}(D,\mathcal{B})$ in item 3 of Lemma 4 and let $\mathcal{N}^+_\mathcal{K}(D,\mathcal{B}) \supset \mathcal{N}_\mathcal{K}(D,\mathcal{B})$ has $x_j \in [0, D_j]$ (instead of $x_j \in [0, D_j/2]$) for $j \in K^+ \cup K^-$. Let

$$\mathcal{N}^+_\mathcal{K}(D) = \cup_\mathcal{B} \mathcal{N}^+_\mathcal{K}(D,\mathcal{B}). \tag{38}$$

In words, this is the neighborhood obtained taking convex combination of the full lines connecting centroids rather than mid-points. With ϵ' small, $d(R, \mathcal{N}_\mathcal{K}(D)) \leq \epsilon'$ implies that $R \in \mathcal{N}^+_\mathcal{K}(D)$. Per Lemma 4, we then have that $y_j = 0$ for all $j \notin K \cup K^+$ and $u_j = 0$ for all $j \in \mathcal{K}\backslash K^-$ as required.

Second case ($\max_{\psi\in\Psi(\mathcal{K},\mathcal{B})}\psi'(R-r_{\mathcal{K}}(D))>0$): Here $0<\max_{\psi\in\Psi(\mathcal{K},\mathcal{B})}\psi'(R-r_{\mathcal{K}}(D))\leq\epsilon'$. Let $\bar{\mathcal{B}}$ be such that $R\in\text{cone}(\mathcal{K},\bar{\mathcal{B}})$ for $\bar{\mathcal{B}}\neq\mathcal{B}$. We can write R as

$$R=\bar{R}+\sum_{j\in K^+(\bar{\mathcal{B}})\setminus K^+(\mathcal{B})}A_jy_j-\sum_{j\in K^-(\bar{\mathcal{B}})\setminus K^-(\mathcal{B})}A_jy_j+\sum_{i\in K^0(\bar{\mathcal{B}})\setminus K^0(\mathcal{B})}b_i,$$

where

$$\bar{R}=r_{\mathcal{K}}(D)+\sum_{j\in K^+(\mathcal{B})\cap K^+(\bar{\mathcal{B}})}A_jy_j-\sum_{j\in K^-(\mathcal{B})\cap K^-(\bar{\mathcal{B}})}A_jy_j+\sum_{i\in K^0(\mathcal{B})\cap K^0(\bar{\mathcal{B}})}b_i.$$

First, we claim that the maximizer in $\max_{\psi\in\Psi(\mathcal{K},\mathcal{B})}\psi'(R-r_{\mathcal{K}}(D))>0$ cannot be $\psi[\kappa_0]$ for κ_0 that is in both $K^+(\mathcal{B})\times K^-(\mathcal{B})\times K^0(\mathcal{B})$ and $K^+(\bar{\mathcal{B}})\times K^-(\bar{\mathcal{B}})\times K^0(\bar{\mathcal{B}})$. Let us assume that κ_0 has $|\kappa_0^+|=1$ (the other cases are treated identically). By the second part of Lemma 6, we have $\psi[\kappa_0]'(\bar{R}-r_{\mathcal{K}}(D))\leq 0$ as well as $\psi'(R-\bar{R})\leq 0$. In turn, $\psi[\kappa^0]'(R-r_{\mathcal{K}}(D))\leq 0$, contradicting the assumption of this case.

Hence, the maximum cannot be attained at any κ that is on the face between the two bases; it must be attained at $\psi[\kappa_*]$ such that $y_{\kappa_*}=0$. Such $\psi[\kappa_*]$ notice has, by Lemma 6, $\psi[\kappa_*]'A_j=0$ for all $j\in K^+(\bar{\mathcal{B}})\cap K^+(\mathcal{B})$ or $j\in K^-(\bar{\mathcal{B}})\cap K^-(\mathcal{B})$ as well as $\psi[\kappa_*]'e_i=0$ for $i\in K^0(\bar{\mathcal{B}})\cap K^0(\mathcal{B})$. In particular, $\psi[\kappa_*]'(\bar{R}-r_{\mathcal{K}}(D))=0$.

Overall, we have that

$$\begin{aligned}\epsilon'\geq\max_{\psi\in\Psi(\mathcal{K},\mathcal{B})}\psi'(R-r_{\mathcal{K}}(D))&=\max_{\psi\in\Psi(\mathcal{K},\mathcal{B})}\psi'(R-\bar{R}) \\ &=\max_{\psi\in\Psi(\mathcal{K},\mathcal{B})}\psi'\left(\sum_{j\in K^+(\bar{\mathcal{B}})\setminus K^+(\mathcal{B})}A_jy_j-\sum_{j\in K^-(\bar{\mathcal{B}})\setminus K^-(\mathcal{B})}A_jy_j+\sum_{i\in K^0(\bar{\mathcal{B}})\setminus K^0(\mathcal{B})}b_i\right) \\ &\geq\zeta\left[\left(\min_{j\in K^+(\bar{\mathcal{B}})\setminus K^+(\mathcal{B})}y_j\right)\wedge\left(\min_{j\in K^-(\bar{\mathcal{B}})\setminus K^-(\mathcal{B})}y_j\right)\wedge\left(\min_{i\in K^0(\bar{\mathcal{B}})\setminus K^0(\mathcal{B})}b_i\right)\right],\end{aligned} \tag{39}$$

where $\zeta>0$ is a constant. Before proving the last inequality, we note that all the elements on the right-hand side must now be smaller than ϵ'/ζ and get the desired result by taking ϵ' small enough.

Finally, the inequality in (39), is argued as follows. Suppose that $j_0\in K^+(\bar{\mathcal{B}})\setminus K^+(\mathcal{B})$ is such that y_{j_0} is the minimizer on the right hand side of the last row of (39). We then have

$$\begin{aligned}&\max_{\psi\in\Psi(\mathcal{K},\mathcal{B})}\psi'\left(\sum_{j\in K^+(\bar{\mathcal{B}})\setminus K^+(\mathcal{B})}A_jy_j-\sum_{j\in K^-(\bar{\mathcal{B}})\setminus K^-(\mathcal{B})}A_jy_j+\sum_{i\in K^0(\bar{\mathcal{B}})\setminus K^0(\mathcal{B})}b_i\right) \\ &=y_{j_0}\max_{\psi\in\Psi(\mathcal{K},\mathcal{B})}\psi'\left(A_{j_0}+\sum_{j\in K^+(\bar{\mathcal{B}})\setminus K^+(\mathcal{B}),j\neq j_0}A_jx_j-\sum_{j\in K^-(\bar{\mathcal{B}})\setminus K^-(\mathcal{B})}A_jx_j+\sum_{i\in K^0(\bar{\mathcal{B}})\setminus K^0(\mathcal{B})}\beta_i\right) \\ &\geq y_{j_0}\min_{x,\beta\geq 0}\max_{\psi\in\Psi(\mathcal{K},\mathcal{B})}\psi'\left(A_{j_0}+\sum_{j\in K^+(\bar{\mathcal{B}})\setminus K^+(\mathcal{B}),j\neq j_0}A_jx_j-\sum_{j\in K^-(\bar{\mathcal{B}})\setminus K^-(\mathcal{B})}A_jx_j+\sum_{i\in K^0(\bar{\mathcal{B}})\setminus K^0(\mathcal{B})}\beta_i\right).\end{aligned}$$

The fact that the minimum in the last row is strictly positive follows then from the fact that $A_{j_0} \notin \text{cone}(\mathcal{K}, \mathcal{B})$. □

Proof of Lemma 9. First we argue $\theta = \theta_{\mathcal{K}}(y, D) \in \text{closure}(\mathcal{Y}(\mathcal{K}, D))$. Since y solves $\text{LP}(R, D)$ and $R \in \mathcal{N}_{\mathcal{K}^0}(D)$:

$$y_j \geq \frac{D_j}{2},\ j \in \mathcal{K} \setminus \kappa^-, \qquad y_j \geq \frac{D_j}{2},\ j \in \kappa^+, \quad \text{and} \quad y_j < \frac{D_j}{2},\ j \in \kappa^-.$$

This implies $\theta_j \geq \frac{D_j}{2}$ for $j \in \mathcal{K}$ and $\theta_j \leq \frac{D_j}{2}$ for $j \in \mathcal{K}^c$; in turn, $\theta_{\mathcal{K}}(y, D) \in \text{closure}(\mathcal{Y}(\mathcal{K}, D))$. That $(y - \theta_{\mathcal{K}}(y, D))_j = 0$ for all $j \notin \kappa^+ \cup \kappa^-$ follows immediately from our definition of θ.

In fact, θ is optimal at the ratio $R^\theta := A\theta$. Indeed, if we take $\mathcal{B}$ as the basis that $\mathcal{K}$ and $\mathcal{K}^0$ share, then the support of θ is all basic variables and we have

$$\mathcal{B}\begin{pmatrix}\theta \\ D-\theta \\ 0\end{pmatrix}_{\mathcal{B}} = \begin{pmatrix}R^\theta \\ D\end{pmatrix} \Rightarrow \mathcal{B}^{-1}\begin{pmatrix}R^\theta \\ D\end{pmatrix} = \begin{pmatrix}\theta \\ D-\theta \\ 0\end{pmatrix}_{\mathcal{B}} \geq 0,$$

which proves the optimality of θ at R^θ by Lemma 1.

For the second item, let u and s be the unmet and surplus variables for $\text{LP}(R, D)$. Since both bases share the request variables $\kappa := \kappa^+ \cup \kappa^-$, we can write

$$\mathcal{B}\begin{pmatrix}y_\kappa \\ 0 \\ 0\end{pmatrix} = \bar{\mathcal{B}}\begin{pmatrix}y_\kappa \\ 0 \\ 0\end{pmatrix}, \text{ and } \mathcal{B}\begin{pmatrix}y_\kappa \\ 0 \\ 0\end{pmatrix} + \mathcal{B}\begin{pmatrix}y_{\mathcal{B}\setminus\kappa} \\ u_{\mathcal{B}} \\ s_{\mathcal{B}}\end{pmatrix} = \begin{pmatrix}R \\ D\end{pmatrix} \Longrightarrow \bar{\mathcal{B}}^{-1}\begin{pmatrix}R \\ D\end{pmatrix} = \bar{\mathcal{B}}^{-1}\mathcal{B}\begin{pmatrix}y_\kappa \\ 0 \\ 0\end{pmatrix} + \bar{\mathcal{B}}^{-1}\mathcal{B}\begin{pmatrix}y_{\mathcal{B}\setminus\kappa} \\ u_{\mathcal{B}} \\ s_{\mathcal{B}}\end{pmatrix}$$

$$= \begin{pmatrix}y_\kappa \\ 0 \\ 0\end{pmatrix} + \bar{\mathcal{B}}^{-1}\mathcal{B}\begin{pmatrix}y_{\mathcal{B}\setminus\kappa} \\ u_{\mathcal{B}} \\ s_{\mathcal{B}}\end{pmatrix},$$

where the second equation is from the optimality of $\mathcal{B}$. Finally, we claim that $\bar{\mathcal{B}}^{-1}\mathcal{B}$ has an identity in the columns corresponding to κ, which proves the result. To see this, note that by assumption $\bar{\mathcal{B}}_\kappa = \mathcal{B}_\kappa$ and we can separate by columns $\mathcal{B} = [\bar{\mathcal{B}}_\kappa | 0_{\kappa^c}] + [0_\kappa | \mathcal{B}_{\kappa^c}] = \bar{\mathcal{B}} + [0_\kappa | \mathcal{B}_{\kappa^c} - \bar{\mathcal{B}}_{\kappa^c}]$. □

Proof of Lemma 10. We prove a slightly stronger result than stated; that for all $\check{\epsilon} \leq \epsilon'$, $d_\infty(R, \mathcal{N}_{\mathcal{K}}(\mathcal{B}, D)) \leq M\check{\epsilon}$ if and only if $d_\infty(R_\bullet, \mathcal{N}_{\mathcal{K}}(\mathcal{B}, D)) \leq \check{\epsilon}$. This immediately implies the stated result recalling that $\mathcal{N}_{\mathcal{K}}(D) = \cup_{\mathcal{B}}\mathcal{N}_{\mathcal{K}}(\mathcal{B}, D)$ and that $d_\infty(x, y) \leq d(x, y) \leq \sqrt{d} \times d_\infty(x, y)$.

In the first direction, suppose that $d_\infty(R, \mathcal{N}_{\mathcal{K}}(\mathcal{B}, D)) \leq \check{\epsilon}$ and let (y, u, s) be the solution to $\text{LP}(R, D)$. Let $\bar{R} \in \mathcal{N}_{\mathcal{K}}(\mathcal{B}, D)$ be such that $d_\infty(R, \bar{R}) = d_\infty(R, \mathcal{N}_{\mathcal{K}}(\mathcal{B}, D)$. By the Lipschitz continuity of LPs, we must have $|y_j - \bar{y}_j| \leq M\check{\epsilon}$ for all $j \in [n]$. In turn, $\|R_\bullet - \bar{R}_\bullet\| = \|Ay - A\bar{y}\| \leq M\check{\epsilon}$. By Lemma 4, $\bar{R}_\bullet \in \mathcal{N}_{\mathcal{K}}(\mathcal{B}, D)$ if $\bar{R} \in \mathcal{N}_{\mathcal{K}}(\mathcal{B}, D)$, so that we can conclude that $d_\infty(R_\bullet, \mathcal{N}_{\mathcal{K}}(\mathcal{B}, D)) \leq M\check{\epsilon}$.

For the other direction, assume that $d_\infty(R_\bullet, \mathcal{N}_{\mathcal{K}}(\mathcal{B}, D)) \leq \check{\epsilon}$. Let (y, u, s) be the solution at $\text{LP}(R, D)$, and $\bar{R} \in \mathcal{N}_{\mathcal{K}}(\mathcal{B}, D)$ be such that $d_\infty(R_\bullet, \bar{R}) = d_\infty(R_\bullet, \mathcal{N}_{\mathcal{K}}(\mathcal{B}, D)) \leq \check{\epsilon}$. Notice that

$$R = Ay + s = R_\bullet + \sum_{i \in K^0(\mathcal{B})} s_i + \sum_{i \notin K^0(\mathcal{B})} s_i.$$

We use the following claim:

(*) $i \notin K^0(\mathcal{B})$, if and only if there exists $\kappa^+ \subseteq K^+(\mathcal{B})$ such that $\mathbf{e}_i \geq A_{\kappa^+}$.

By item (3) of Lemma 4, $\bar{R} + \sum_{i \in K^0(\mathcal{B})} s_i \in \mathcal{N}_{\mathcal{K}}(\mathcal{B}, D)$. Because

$$d_\infty\left(R_\bullet + \sum_{i \in K^0(\mathcal{B})} s_i, \bar{R} + \sum_{i \in K^0(\mathcal{B})} s_i\right) = d_\infty\left(R_\bullet, \bar{R}\right) = \check{\epsilon}.$$

it suffices to show that $s_i = 0$ for all $i \notin K^0(\mathcal{B})$, to conclude that $d_\infty(R, \bar{R} + \sum_{i \in K^0(\mathcal{B})} s_i) = \check{\epsilon}$ and, in turn, $d_\infty(R, \mathcal{N}_{\mathcal{K}}(\mathcal{B}, D)) = \check{\epsilon}$.

Because $\bar{R} \in \mathcal{N}_{\mathcal{K}}(\mathcal{B}, D)$, $\bar{y}_j \leq D_j/2, j \in K^+(\mathcal{B})$. By the Lipschitz continity of LPs, $y_j \leq D_j/2 + M\check{\epsilon}, j \in K^+(\mathcal{B})$ (these variables are not saturated under the optimal solution at LP(R, D)). If $s_i > 0$ for some $i \notin K^0(\mathcal{B})$, then by the claim (*), for small enough δ, $y + A_\kappa \delta$ is feasible at R and, because $v_j > 0$ for all $j \in [n]$, has higher objective function value than y, contradicting the optimality of y. It must be, then, that $s_i = 0$ for all $i \notin \cup_{\mathcal{B}} K^0(\mathcal{B})$. We may conclude then that

$$d_\infty\left(R, \bar{R} + \sum_{i \in K^0(\mathcal{B})} s_i\right) = d_\infty\left(R_\bullet, \bar{R}\right) = \check{\epsilon},$$

as required.

It remains to prove the claim (*). Suppose that there exists no $\kappa^+ \in K^+(\mathcal{B})$ with $\mathbf{e}_i \geq A_{\kappa^+}$. Take R in the interior where the slack is 0 (the centroid budget is one such choice). Increase $R_i \to R_i + \delta$ for small enough δ so that $R_i + \delta \in \mathcal{N}_{\mathcal{K}}(\mathcal{B}, D)$. Because there exists no κ^+ as desired it means that the slack i must enter the basis. In turn, s_i must be in $\mathcal{B}$ for some $\mathcal{B}$ associated with $\mathcal{K}$.

For the other direction, suppose that there exists $\kappa^+ \in K^+(\mathcal{B})$ with $\mathbf{e}_i \geq A_{\kappa^+}$ then because $y_{\kappa^+} \leq D_{\kappa^+}/2$ we can increase y_{κ^+}. This increases the objective function value contradicting the optimality of R. In turn, s_i cannot be in $\mathcal{B}$. □

Proof of Lemma 11. Define $\bar{R}_i^t = \frac{I_i^t \wedge M(T-t)}{T-t} + \varrho_i$, where $M = 2\sum_{j \in [n]} p_j$. Because $y_j^t \leq p_j, j \in [n]$, it is immediate that $LP(R^t, D)$ has the same optimal request-variable values as $LP(\bar{R}^t, D)$.

$$\bar{R}_i^{t+1} - \bar{R}_i^t = \frac{I_i^{t+1} \wedge M(T-t-1)}{T-t-1} - \frac{I_i^t \wedge M(T-t)}{T-t}.$$

Because $I_i^{t+1} \leq I_i^t + 1$, $I_i^{t+1} \wedge M(T-t-1) \leq I_i^t \wedge M(T-t) + 1$ so that

$$|\bar{R}_i^{t+1} - \bar{R}_i^t| \leq \frac{1}{T-t-1} + \frac{I_i^t \wedge M(T-t)}{(T-t)(T-t-1)} \leq \frac{M+1}{T-t-1}.$$

Finally, the Lipschitz continuity of LPs guarantees that $\|y^{t+1} - y^t\|_\infty \leq \|\bar{R}^{t+1} - \bar{R}^t\|_\infty$. □

Proof of Lemma 12. Let $\mathcal{N}_{\mathcal{K}}^+(D^0)$ be as in (38) with D^0 replacing D there. By assumption $\mathbb{E}[R^0] = \frac{1}{T} I^0 + \varrho \in \mathcal{N}_{\mathcal{K}}(p)$ and $(\bar{Z}^T, \bar{\mathfrak{Z}}^T) \in \mathcal{A}^{\epsilon^0}$. Because $R^0 = \frac{1}{T} I^0 + \bar{\mathfrak{Z}}^T$, we then have $d(R^0, \mathcal{N}_{\mathcal{K}}(p)) \leq$

$2\epsilon^0 = \frac{1}{4}\min\{p_j\} \wedge \min\{\varrho_i : \varrho_i > 0\}$. Because, on $\mathcal{A}^{\epsilon^0}$, $d(\partial\mathcal{N}^+_{\mathcal{K}}(D^0), \mathcal{N}_{\mathcal{K}}(p)) \geq \min_j(D^0_j - p_j/2) \geq min_j p_j/2 - \epsilon^0 \geq 3\epsilon^0$, we have that $R^0 \in \mathcal{N}^+_{\mathcal{K}}(D^0)$.

It then follows from item 2 of Lemma 4 (with D replaced with D^0 there) that $R^0 - A_{\mathcal{K}}\bar{Z}^0_{\mathcal{K}} = R^0 - r_{\mathcal{K}}(D^0) \in \text{cone}(\mathcal{K}, \mathcal{B})$ if and only if $\mathcal{B}$ is offline optimal. □

Proof of Lemma 13. Let us write $\mathcal{K}_i = \mathcal{K} \cup \kappa_i^+ \backslash \kappa_i^-$. By assumption $d(R, \mathcal{N}_{\mathcal{K}_i}(D)) \leq \epsilon''$ for all $i \in [d]$. By the Lipschitz continuity of the LP, y that solves $\text{LP}(R, D)$, must have for some $\check{\epsilon}(\epsilon'')$ that $y_j \geq D_j/2 - \check{\epsilon}$ for all $j \in \cup_{i\in[d]}\kappa_i^+$ (because at centroid i, $y_j \geq D_j/2$ for $j \in \kappa_i^+$) as well as $y_j \geq D_j/2 - \check{\epsilon}$ for all $j \in \cup_{i\in[d]}\kappa_i^-$ (because $j \in \mathcal{K}$). Additionally, $y_j \leq \check{\epsilon}$ for all $j \notin \mathcal{K} \cup (\cup_{i\in[d]}\kappa_i^+) \backslash (\cup_{i\in[d]}\kappa_i^-)$. We choose ϵ'' (in turn $\check{\epsilon}$) so that $\check{\epsilon} < \min_j D_j/2$. Thus, because we are considering d neighbors $|\cup_{i\in[d]} \kappa_i^+| + |\cup_{i\in[d]} \kappa_i^-| = d$. Overall, we have $|\mathcal{K}| + |\mathcal{K}^c| + |\cup_{i\in[d]} \kappa_i^+| + |\cup_{i\in[d]} \kappa_i^-| = n + d$ strictly positive, hence basic, variables. The optimal basis then must be the one that has these variables in the basis.

For the second item of the lemma we can assume $\mathcal{B}_i = \mathcal{B}_0$. Because there are d different centroids $\mathcal{K}_i$ considered and they are different, the set $(\cup_i \kappa_i^+) \cup (\cup_i \kappa_i^-)$ contains at least d different j. Because of the construction of the d vectors ψ, there exists no $\kappa \in (\cup_i \kappa_i^+) \cup (\cup_i \kappa_i^-)$ such that $\psi[\kappa]' A_{\tilde{\kappa}} = 0$ for all $\tilde{\kappa} \in (\cup_i \kappa_i^+) \cup (\cup_i \kappa_i^-)$. Thus, for any such ψ we have, by Lemma 6, that for all κ $\psi[\kappa]'(A_{\cup_i \kappa_i^+} D_{\cup_i \kappa_i^+} - A_{\cup_i \kappa_i^-} D_{\cup_i \kappa_i^-}) \leq -\zeta \min_j D_j$. In turn, we have $\max'_{\psi \in \Psi(\mathcal{K},\mathcal{B})}(R - r_{\mathcal{K}}(D)) \leq M\check{\epsilon} - \zeta \min_j D_j$. The right-hand side is negative for small $\check{\epsilon}$. □

Proof of Lemma 14. We prove this for $d = 2$, the proof for $d > 2$ is identical. Let

$$\tau^0_\partial = \inf\{t \leq T : d(R^t, \partial\mathcal{N}_{\mathcal{K}}(p)) \leq \epsilon'\}, \text{ and } \tau^1_\partial = \inf\{t \geq \tau^0_\partial : d(R^t, \partial\mathcal{N}_{\mathcal{K}}(p)) \geq 2\epsilon'\},$$

and

$$\mathcal{C}_\ell := \{T - \tau^1_\partial \leq \ell\},$$

where we define $\tau^1_\partial = -\infty$ if $\tau^0_\partial \not< T$. On the event $\mathcal{V} := \mathcal{V}[\tau^{\epsilon',\mathcal{B}}_{\text{cone}}] = 1$, $\tau^\partial_0 < T$ and, by Remark 5, $\mathbb{P}[\mathcal{V} = 1, (\mathcal{C}_\ell)^c] \leq m_1 e^{-m_2\ell}$.

On the event $\{\mathcal{V} = 1\} \cap \mathcal{C}_\ell \cap \{T - \tau^{\epsilon',\mathcal{B}}_{\text{cone}} > \ell\}$, $\tau^1_\partial \geq \tau^{\epsilon',\mathcal{B}}_{\text{cone}}$ and there exist $t_0 < \tau^{\epsilon',\mathcal{B}}_{\text{cone}}$ such that $d(R^t, \partial\mathcal{N}_{\mathcal{K}}(p)) \leq 2\epsilon'$ for all $t \in [t_0, \tau^{\epsilon',\mathcal{B}}_{\text{cone}}]$.

Fix such t, and let y be the optimal solution to $LP(R^t, p)$. Because $d(R^t, \partial\mathcal{N}_{\mathcal{K}}) \leq 2\epsilon'$ and $t \leq \tau^{\epsilon',\mathcal{B}}_{\text{cone}}$, there exists M such that $|y_j| \leq M\epsilon'$ for all $j \notin \mathcal{K} \cup K^+(\mathcal{B}) \backslash K^-(\mathcal{B})$ and $y_j \geq p_j/2 - M\epsilon'$ for all $j \in \mathcal{K} \cup \kappa^+$ where $\mathcal{K}^0 = \mathcal{K} \cup \kappa^+ \backslash \kappa^-$ is the second centroid visited.

Notice that for $i \notin K^0(\mathcal{B})$, and $t < \tau^{\epsilon',\mathcal{B}}_{\text{cone}}$, $s_i = 0$ so that $R^t_i = (Ay^t)_i + s_i \leq M$ for some $M > 0$, so that $I^t_i \leq M(T - t)$. Simple algebra then establishes that $R^{t+1}_i - R^t_i = \mathcal{O}\left(\frac{1}{T-t}\right)$ for such i; see e.g. the proof of Lemma 11.

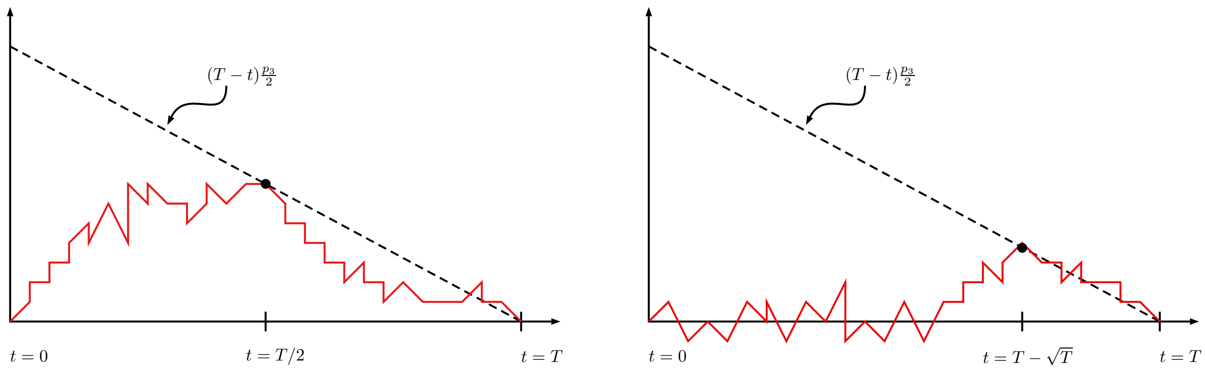

FIGURE 11. Paths on $\mathcal{M}=\{R-e_{\mathcal{K}}\bar{Z}_{\mathcal{K}}^T \le 0\}$. The two paths illustrate unlikely events. On the left: because the random walk has variation $\mathcal{O}(\sqrt{t})$ it cannot hit the target line by a time t of the form $t=T-\Omega(T)$. On the right: the path could hit the target by a time of the form $t=T-\mathcal{O}(\sqrt{T})$. In this case, however, it does not have enough time to get back to 0 which it must on the event $\mathcal{M}$

Suppose that $|\kappa^+|=1$ (the proof is the same if $|\kappa^-|=1$). Take $\psi=\psi[\kappa]$. Using Lemma 6, we then have, for all $t\le \tau_{\text{cone}}^{\epsilon',\mathcal{B}}$, that

$$\psi'(R^t - r_{\mathcal{K}}(p)) = \psi'(A_{K^+}y_{K^+} - A_{K^-}y_{K^-})$$
$$= \psi'(A_{\kappa^+}y_{\kappa^+} - A_{\kappa^-}y_{\kappa^-}) \le \psi'(A_{\kappa^+}p_{\kappa^+}/2 - A_{\kappa^-}p_{\kappa^-}/2) + M\epsilon' \le 0.$$

The first equality and first inequality follow from Lemma 6. The last inequality follows by choosing ϵ' small enough. Thus, it must be that a vector $\psi\in\Psi$ for which $\psi'(R^{\tau_{\text{cone}}^{\epsilon',\mathcal{B}}} - r_{\mathcal{K}}(p)) > \epsilon'$ is distinct from $\psi[\kappa]$. □

Lemma 16 below was used in the proof of Theorem 3.

LEMMA 16. *Fix $\epsilon>0$. Consider a random walk G of the form $G_t = G_0 + \sum_{s=1}^t X_t$ where the increments $X_t, t\in[T]$ are zero-mean i.i.d and bounded ($\mathbb{E}[X_t]=0$ and $\mathbb{P}[|X_t|\le b]=1$ for some $b>0$) and independent of G_0 which satisfies $|G_0|\le \epsilon/2$. Let*

$$\tau = \inf\{t\ge 0 : (T-t)\epsilon \le G_t\} \wedge T.$$

Then, for all $t\in[T]$

$$\mathbb{P}[T-\tau > t, G_T\le 0] \le m_1 e^{-m_2 t},$$

for constants $m_1, m_2 > 0$ that may depend on b,ϵ.

PROOF. Fig. 11 is a graphic illustration of the event whose probability we wish to bound. Notice $\{\tau\le t\} = \{\exists s\le t : (T-s)\epsilon - G_s \le 0\}$. In turn, $\{\tau\le t, G_T\le 0\}\subseteq\{\inf_{s\le t}(G_T - G_s + \epsilon(T-s))\le 0\}$ so that

$$\mathbb{P}[\tau\le t, G_T\le 0] \le \mathbb{P}[\inf_{s\le t} G_T - G_s + \epsilon(T-s)\le 0]$$
$$= \mathbb{P}[\inf_{u\ge T-t} G_u + \epsilon u \le 0] \le m_1 e^{-m_2(T-t)},$$

where the last inequality is a generalized version of Azuma's inequality; see, e.g., [Ross, 1995, Theorem 6.5.2]. Replacing $t \leftarrow T - \ell$ gives the result. $\square$