informs
https://pubsonline.informs.org/journal/moor

MATHEMATICS OF OPERATIONS RESEARCH
Articles in Advance, pp. 1–39
ISSN 0364-765X (print), ISSN 1526-5471 (online)

Dynamic Resource Allocation: The Geometry and Robustness of Constant Regret

Alberto Vera,[a] Alessandro Arlotto,[b] Itai Gurvich,[a,*] Eli Levin[c]

[a] School of Operations Research and Information Engineering, Cornell University, Ithaca, New York 14850; [b] The Fuqua School of Business, Duke University, Durham, North Carolina 27708; [c] Smack Technologies, Los Angeles, California 90024

*Corresponding author

Contact: aav39@cornell.edu (AV); aa249@duke.edu, https://orcid.org/0000-0003-2705-9814 (AA); i-gurvich@kellogg.northwestern.edu, https://orcid.org/0000-0001-9746-7755 (IG); elevin@smacktechnologies.com (EL)

Received: October 10, 2020
Revised: September 4, 2022; September 20, 2023
Accepted: April 19, 2024
Published Online in Articles in Advance: November 11, 2024

MSC2020 Subject Classification: Primary: 68W40, 90C39; Secondary: 60G40, 90C40, 90C08

https://doi.org/10.1287/moor.2020.0334

Abstract. 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 the collected reward is given by a linear optimization problem with a random right-hand side. This type of problem, known as packing linear program (LP), is ubiquitous in resource allocation. We provide a detailed characterization of the parametric structure of this packing LP. Relying on this geometric understanding, we revisit and expand on BudgetRatio algorithms that achieve constant regret by resolving 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, and (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 in which the parameters have to be learned. (iii) The algorithm has an equivalent formulation as a generalized bid-price algorithm in which 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 in which 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.

Funding: This work was supported by the U.S. Department of Defense [Grant W911NF-20-C-0008].

Keywords: multisecretary problem • online packing • regret • adaptive online policy

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 restocked 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.

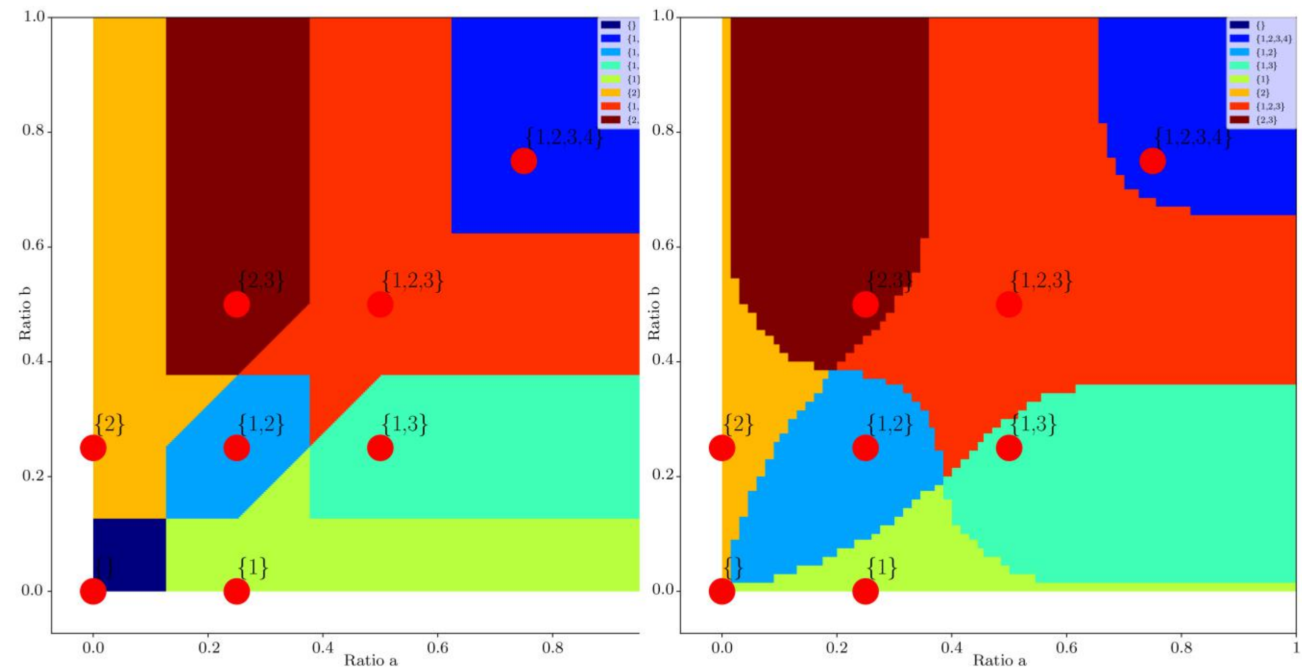
If the controller could solve the problem in an offline fashion, the controller would wait for the end of the horizon and, given the realization of the random arrivals, choose the best allocation of resources to requests. The reward of the offline controller is an upper bound on any online algorithm.

At each time $t = 1, \ldots, T$, 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 resolving the following packing linear program (LP) at each time period t:

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

where $p = (p_j)_{j \in \mathcal{J}}$ are the arrival probabilities of requests, $\varrho = (\varrho_i)_{i \in \mathcal{R}}$ are the resource restock probabilities, $v = (v_j)_{j \in \mathcal{J}}$

Figure 1. (Color online) 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 $\mathcal{K} = \{j : \overline{y}_j^t \geq \eta_j\}$ that should be accepted at that inventory level. Each region on the plot corresponds to a different such set. The 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 are computed via dynamic programming with 70 periods to go.



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. Because 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 $\overline{y}^t$ to Equation (1), $\overline{y}_j^t$ 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, that is, such that $\overline{y}_j^t \geq \eta_j$ for thresholds η_j that we explicitly specify. Viewed as a random process, these scores $\overline{y}^t$ 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 Figure 1.

1.1. Methodology: A Geometric View of Resolving Policies

The problems we consider cannot be solved optimally because of 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 the 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 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 $\overline{y}^t$ 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 $\overline{y}_j^t \geq \eta_j$) are accepted, and all others are rejected; see Figure 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 $\overline{y}^t$ 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) in which it performs basic allocations, those that are consistent with the offline basis. To establish this, we must (i) develop a generalizable mathematical description of Figure 1 and (ii) study the dynamics of the stochastic process R^t inside and between the action regions in this figure.

1.2. Implications

Whereas the main contribution is mathematical, the geometric view advances the understanding of practical aspects of BudgetRatio as follows.

1.2.1. BudgetRatio as a Bid-Price Control. Bid-price heuristics are popular because of 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 [16]).

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 adaptively and efficiently.

1.2.2. 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 in which 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 Figure 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. [25]). 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 in which the demand and reward parameters are not a priori known to the controller and must be learned.

1.2.3. 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 $\mathcal{R} = [d] = \{1, \dots, d\}$, 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 $\mathcal{J} = [n] = \{1, \dots, n\}$ 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 and A_j corresponds to the (column) vector of resources that request j consumes. 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 $V^1, V^2, \dots, V^T$ are assumed to be independent and identically distributed (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 $I^t = (I_1^t, \dots, I_d^t)'$.

The selection process at time t unfolds as follows:

i. Inventory restock: The inventory is updated to include newly arriving resource units; that is, $I_i^t \leftarrow I_i^t + 1$ 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 $I^t \not\geq A_j$. 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 and leaving I^t unchanged.

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; if it is 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 controller 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 nonbinding, and we can reduce the problem to one with $d-1$ resources.

As a preprocessing 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 a 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 [6, 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

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

BudgetRatio resolves 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 $\overline{p}_j \leftarrow p_j + \gamma_j \mathbb{1}_{\{\gamma_j > \alpha p_j\}}$.
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 LP(R^t, p) to obtain the optimal decision variables $\overline{y}^t$.
  Set j as the type of the arriving request.
  if $I^t \not\geq A_j$ (not feasible to serve j) or $\overline{y}^t_j < \alpha\overline{p}_j$ (not desirable to serve j): reject the request
  else if $\overline{y}^t_j \geq \alpha\overline{p}_j$: accept the request $I^t \leftarrow I^t - A_j$
  Carry over the inventory for the next period: $I^{t+1} \leftarrow I^t$.

2.1.1. LP Notation and Terminology. If I_n is the identity matrix of dimension $n \times n$, let $\overline{A}$ be the following augmentation of the resource consumption matrix A:

$$\overline{A} = \begin{bmatrix} A & 0 & \mathbf{I}_d \\ \mathbf{I}_n & \mathbf{I}_n & 0 \end{bmatrix}. \tag{4}$$

For any $R \in \mathbb{R}^d_{\geq 0}$ and $D \in \mathbb{R}^n_{\geq 0}$, we rewrite the LP in (3) in standard form as

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

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 $(d+n) \times (d+n)$ submatrix of $\overline{A}$ corresponding to the variables in the basis $\mathcal{B}$; $\mathcal{B}^c$ denotes the nonbasic columns. Let $\overline{v} = (v,0,0) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^d$ be the extended reward vector, to which we assign zero value to the slack variables u and s. For an optimal basis $\mathcal{B}$, we refer to $\lambda = \lambda(\mathcal{B},v) = (\mathcal{B}^{-1})'\overline{v}_{\mathcal{B}}$ as the dual variables associated with $\mathcal{B}$.

2.1.2. 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

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

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.

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

$$\mathcal{N}_{\mathcal{K}} = \{R \in \mathbb{R}^2_{\geq 0} : \overline{y}_j \geq \alpha \overline{p}_j,\ \text{for}\ j \in \mathcal{K}, \overline{y}_j < \alpha \overline{p}_j,\ \text{for}\ j \in \mathcal{K}^c\}.$$

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

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

$$\sum_{j \in \mathcal{K}} A_{ij} p_j =: r_{\mathcal{K}}. \tag{Centroid Budget}$$

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, 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 $\lambda(\mathcal{B},v) = (\mathcal{B}^{-1})'\overline{v}_{\mathcal{B}}$. This informal description of the geometry suffices for the presentation of our results; formal definitions appear in Sections 3 and 4.

2.2. The Max-Bid-Price BudgetRatio

We present a generalization of bid-price policies that we prove achieves constant regret.

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

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

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, $\Lambda(R)$, 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 $\sum_{l \in \mathcal{K}} A_l \not\geq A_j$, and we then write $j \in \partial(\mathcal{K})$. The centroid $\mathcal{K} = \emptyset$ is near the boundary for all types: $\partial(\emptyset) = \mathcal{J}$. 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 one in entry $d + j$ and zero 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 $\overline{A}$ is the augmentation of A as in Equation (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: $v_j \geq \max_{\lambda \in \Lambda(R^t)} \{\overline{A}_j'(\lambda + \lambda^{\partial}(R^t))\}$.

In our base example, the centroid $\mathcal{K} = \{1, 2\}$ (rombus-like region in Figure 1) has the dual vectors $\lambda(\mathcal{B}_1) = (4, 4, 0, 0, 0, 0)'$, $\lambda(\mathcal{B}_2) = (4, 1, 0, 3, 0, 0)'$, and $\lambda(\mathcal{B}_3) = (1, 4, 3, 0, 0, 0)'$ (one vector for each of the optimal bases at the centroid's budget) so that, for $R^t \in \mathcal{N}_{\{1,2\}}$, the decision is to accept a type-j arrival if $v_j \geq \max\{\lambda(\mathcal{B}_1)'\overline{A}_j, \lambda(\mathcal{B}_2)'\overline{A}_j, (\mathcal{B}_3)'\overline{A}_j\}$. Types 1, 2 are both accepted here because $\lambda(\mathcal{B}_1)'\overline{A}_1 = \lambda(\mathcal{B}_2)'\overline{A}_1 = 4$ and $\lambda(\mathcal{B}_3)'\overline{A}_1 = 1$; type 3 is not accepted because $\lambda(\mathcal{B}_2)'\overline{A}_2 = 8 \geq 5 = v_3$. There are not boundary types for the centroid $\{1, 2\}$.

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 precompute the full map $\Lambda(\cdot)$, we must solve at most $(n + 1)!$ packing LPs in which, 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 four 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. [4]). But precomputing 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 in which 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 parameterize 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 types j whose slack u_i is not in $\mathcal{B}$, and (iii) $(\overline{A}'\lambda)_j \geq v_j + \delta$ for all request types 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. Parameterizing 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 parameterizes 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 simple requirements; see Corollary 1.

Theorem 1 (Constant Regret and Its Robustness). *Suppose that slow restock holds. Then,*

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

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

where $V_{\text{on}}(T,I^0)$ 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.*

ii. *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 $c\leq\max\{|\mathcal{B}^{-1}|_\infty:\mathcal{B}\text{ basis}\}$.

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

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

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

iv. *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 two-dimensional base example Figure 1 (right), the sup norm distance between any two circles (centroid budgets) equals 1/4; hence, centroid separation (Definition 4) is satisfied with $\delta=1/4$. The δ-complementarity of Definition 3 is satisfied with $\delta=1$ and $\max\{|\mathcal{B}^{-1}|_\infty:\mathcal{B}\text{ basis}\}\leq1$; this we found through computational discovery of all the optimal bases. Therefore, Equation (7) specializes to $|v-\tilde{v}|_\infty\leq1/4$. It is important that c,δ depend only on (v,A) and not on p,ϱ or the horizon T. The requirement of Equation (8) on $\tilde{p}$ imposes eight 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 $\delta=\min_j\{p_j\}$. Equation (8) reduces to*

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

which is, in particular, satisfied if $|p-\tilde{p}|_\infty\leq\frac{\min_k\{p_k\}}{4n}$. 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. [25, theorem 4].

We conclude this section with a discussion of slow restock and BudgetRatio parameters.

Remark 1 (Slow Restock, Conceptual Implications). In allowing restock in our model, we explore the limits of constant regret and simple resolving 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 Figure 1 the initial budget lies in $\mathcal{N}_{\{1,2,3\}}$, 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 opposite 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 in which $\varrho_i < (r_{\mathcal{K}})_i$ for some centroids but $\varrho_i > (r_{\mathcal{K}})_i$ for others or those in which some of these are held with equality. In Appendix A, we provide examples in which a regret proportional to $\sqrt{T}$ is unavoidable.

2.3.1. Algorithmic Implications. The restock rate is accounted for in the definition of $\overline{p}_j$, where $\gamma_j = \max_{i:A_{ij}=1}\{\varrho_i\}$ captures the restock of resources used by j; $\overline{p} = p$ 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 $\boldsymbol{\alpha}$). The parameter α can be set to any value in (0, 1). The closer that the value of α is to one (yet still away from one) the greater the restock rate that is allowed without compromising constant regret. Intuitively speaking, as α approaches one, the algorithm becomes more conservative in accepting requests; it slows down to allow for inventory to accumulate.

2.3.2. Final Setup Details. Let $\mathcal{F}_0$ denote the trivial σ-field, and for $t \in [T]$, let $\mathcal{F}_t = \sigma\{(Z^\tau, \mathfrak{Z}^\tau) : \tau = 1, \ldots, t\}$ 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 (i) the total consumption of resource i does not exceed its initial inventory I_i^0 plus its total restock and (ii) the total acceptance does not exceed arrivals:

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

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

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

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

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

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

2.3.3. 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 subvectors: 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, $d(x, y) = \|x - y\|$ 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 Euclidean distance of x from the set $\mathcal{C}$. We similarly define for the sup norm $d_\infty(x, y) = |x - y|_\infty$ and $d_\infty(x, \mathcal{C})$. 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 ∞ ($\max\{\emptyset\} = 0, \min\{\emptyset\} = \infty$). We use throughout M to be a constant that can depend on (A, p, ϱ, v) but is independent of (T, I^0) whose value can change from one line to the next.

2.4. Related Work

2.4.1. Online Packing (Network Revenue Management). The attainability of constant regret has been already established for some online allocation. Arlotto and Gurvich [2], the first to establish constant regret (regardless of whether the deterministic LP is degenerate or not), take 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 et al. [25] and Vera and Banerjee [24] 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.

2.4.2. Bid-Price Heuristics. Bid-price heuristics are popular because of their intuitive interpretation; see Talluri and Van Ryzin [22] for asymptotic results and Boyd and Bilegan [8] 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, that is, 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 [16]). To the best of our knowledge, the strongest available guarantee for bid-price policies is $O(\sqrt{T})$ regret (Talluri and Van Ryzin [22]). We show that BudgetRatio can be interpreted as a bid-price control (see Talluri and Van Ryzin [22, 23, chapter 3.2], albeit a more elaborate one. Our generalized version of bid price—which we call max bid price—achieves constant regret.

2.4.3. 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 the 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 [9] for more on bandit problems.

Bandit problems known as bandits with knapsacks (Badanidiyuru et al. [3]) explicitly model budget constraints that, in our setting, correspond to the limited inventory. In contextual bandits (Agrawal and Devanur [1]), arrivals present a context before the controller makes decisions.

The general-purpose results in the literature (see Agrawal and Devanur [1], Badanidiyuru et al. [3]) imply $\mathcal{O}(\sqrt{T})$ regret bounds for our setting. For our model, in which the context is the type $j \in [n]$, we identify the separation condition in Definition 4 that guarantees an optimal regret scaling of $\mathcal{O}(\log T)$. 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, that is, the important centroids. Centroids bring out a natural multidimensional 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. [25] and Wu et al. [26].

2.4.4. 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 [21] as well as Plambeck and Ward [19], which gives an asymptotically optimal policy for holding cost minimization under a high demand assumption). We focus on finite-time nonasymptotic 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.

2.4.5. Parametric Linear Programming. The objective in this literature is to understand how optimization problems change as the primitives change (see Gal [13] 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. [5], Borrelli et al. [7]) in which 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.

2.4.6. 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 toward 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 queueing control, there are similarities between our arguments and those used to show that max-weight policies—based on resolving local optimization problems—lead to the attraction to a subset of the state space (see Eryilmaz and Srikant [12], Maguluri and Srikant [17]).

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, whereas the online policy must commit to solutions in a dynamic fashion. Below, we make this precise.

3.1. Offline Representation

Introducing slack variables, we rewrite the constraints of the offline LP (2), $\{Ay \leq I^0 + \mathfrak{Z}^T, y \leq Z^T\}$, in standard form as $\{Ay + s = I^0 + \mathfrak{Z}^T, y + u = Z^T\}$, where $s \in \mathbb{R}^d_{\geq 0}$ is the surplus of resource and $u \in \mathbb{R}^n_{\geq 0}$ is the unmet demand. Augmenting the matrix A to $\overline{A}$ as in Equation (4), we arrive at the standard form representation of the offline value:

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

The random vector $C \in \mathbb{R}^{d+n}_{\geq 0}$ is the maximum consumption of offline. Given a basis $\mathcal{B}$ (columns of $\overline{A}$) for the LP in Equation (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 submatrix $\overline{A}_{\mathcal{B}}$.

3.2. Online Construction of the Off-line 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 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. If τ^π is the first time that π performs a nonbasic 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^t_i be the surplus of resource $i \in [d]$ at time t when using the policy π, that is, in vector notation, $S^t = I^0 + \mathfrak{Z}^t - AY^{\pi,t}$. The wastage of π at t is $W^{\pi,t} := \max\{S^t_i : \text{ surplus variable } s_i \text{ is non basic}\} = \max\{S^t_i : s_i \notin \mathcal{B}, i \in [d]\}$.

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, that is, 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 tth request. For any online policy π, define the time*

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

Then, for any $\tau \le \tau^\pi$ almost surely (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 $[1, t]$. 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 $S^t := I^0 + \mathfrak{Z}^t - AY^t \in \mathbb{R}^d_{\ge 0}$. By definition,

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

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

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

The first equation follows from the decomposition $\overline{A} = [\mathcal{B}, \mathcal{B}^c]$ and the fact that, up to time τ, the policy performs basic allocations so that the only nonzero variables Y_j^t, U_j^t 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

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

The process Y is increasing and nonnegative: $Y^T \ge Y^t \ge 0$ for all $t \in [T]$. Consequently,

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

We bound the last expression using Equation (15): because there is at most one arrival per period, $|(\mathfrak{Z}^T, Z^T)' - (\mathfrak{Z}^t, Z^t)'|_\infty \le T - t$, 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 Equation (15) and conclude the result by setting $t = \tau$. □

To prove item (i) of Theorem 1 (constant regret), it suffices now to find a random time $\tau \le \tau^\pi$ 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 $W^{\pi,\tau}$ for $\pi =$BudgetRatio.

3.3. Analysis Overview via the One-Dimensional Case

Let us consider in some detail the one-dimensional packing problem, also known as the multisecretary problem (Arlotto and Gurvich [2]). 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 $A = \mathbf{e}' = (1, 1, \ldots, 1)$ (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:

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

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

3.3.1. Centroids. If the budget ratio is exactly $\mathbb{E}[R^0] = \overline{F}_j$ (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 $\mathrm{LP}(\overline{F}_j, p)$ 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 $r_{\mathcal{K}} = \overline{F}_j$, so we refer to $r_{\mathcal{K}}$ as the centroid's budget; see Figure 2.

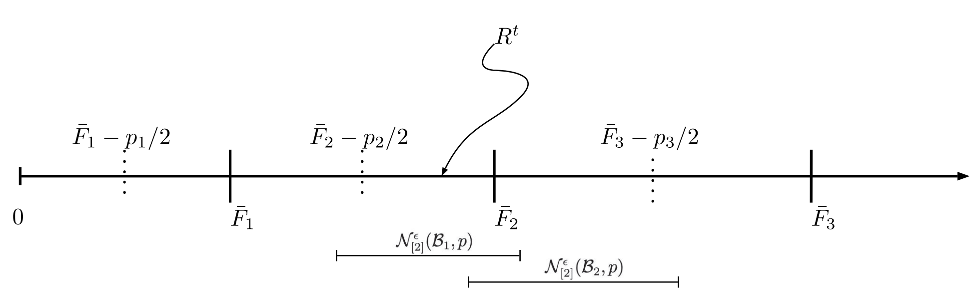
The centroids do not depend on p. Regardless of the distribution, the LP takes all requests $[j]$ before taking any request of type $j + 1$. Both the deterministic relaxation $\mathrm{LP}(\mathbb{E}[R^0], p)$ and the offline problem $\mathrm{LP}(R^0, D^0)$—in which, we recall, $D^0 = \frac{1}{T} Z^T$—follow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

3.3.2. Action Regions: The Centroid Neighborhood. The thresholding of the algorithm when the aggressiveness parameter $\alpha = \frac{1}{2}$ accepts type j requests only if $\overline{y}_j \ge \frac{1}{2} p_j$, as a result, creates an interval (neighborhood) around the centroid's budget. The neighborhood of the centroid {1, 2} is the interval $\mathcal{N}_{\{1,2\}}(p) = \left[\overline{F}_2 - \frac{p_2}{2}, \overline{F}_2 + \frac{p_3}{2}\right)$ centered at exactly the centroid budget $r_{[2]}(p) = \overline{F}_2 = p_1 + p_2$; see Figure 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 is at or exceeds the upper threshold $\overline{F}_2 + p_3/2$. It rejects type-2 requests if it goes below the lower threshold $\overline{F}_2 - p_2/2$.

3.3.3. 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 (a priori unknown) basis $\mathcal{B}$ for much of the horizon.

Consider the centroid $[2] = \{1, 2\}$, its budget $r_{[2]}(p) = \overline{F}_2$, and the action region $\mathcal{N}_{[2]}(p) = [\overline{F}_2 - p_2/2, \overline{F}_2 + p_3/2)$: 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 $\mathcal{N}_{[2]}(\mathcal{B}_2, p) = [\overline{F}_2, \overline{F}_2 + p_3/2)$. The basis $\mathcal{B}_1$ that has—in addition to the

Figure 2. The position of the ratio R^t with respect to the centroid budgets $r_{[j]}(p) = \overline{F}_j$ determines the actions of the policy. At time t, the policy accepts a type-j request if and only if $R^t \ge \overline{F}_j - p_j/2$. 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 $\mathcal{N}^{\varepsilon}_{[2]}(\mathcal{B}_2, p)$, and this guarantees that the policy accepts only requests in the optimal offline basis.



unmet (slack) variables for types $j > 2$—also the unmet variable u_2 for type 2 is optimal on $\mathcal{N}_{[2]}(\mathcal{B}_1, p) = [\overline{F}_2 - p_2/2, \overline{F}_2]$. When in the proximity of $\mathcal{N}_{[2]}(\mathcal{B}_2, p)$, that is, on the set $\mathcal{N}^{\epsilon}_{[2]}(\mathcal{B}_2, p) = \{R : d_\infty(R, \mathcal{N}_2(\mathcal{B}, p)) \le \epsilon\}$ (see Figure 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.

3.3.4. 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 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 $R^t \in \mathcal{N}^{\varepsilon}_{[2]}(\mathcal{B}_1, p)$, BudgetRatio 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, $R^t = \frac{1}{T-t} I^t \in \mathcal{N}^{\varepsilon}_{[2]}(\mathcal{B}_1, p) \cup \mathcal{N}^{\varepsilon}_{[2]}(\mathcal{B}_2, p) \le \overline{F}_3$ implies $I^t \le \overline{F}_3(T-t)$: little inventory (hence, little wastage) remains at the end of the horizon. With τ^π large and wastage small, Proposition 1 yields constant regret.

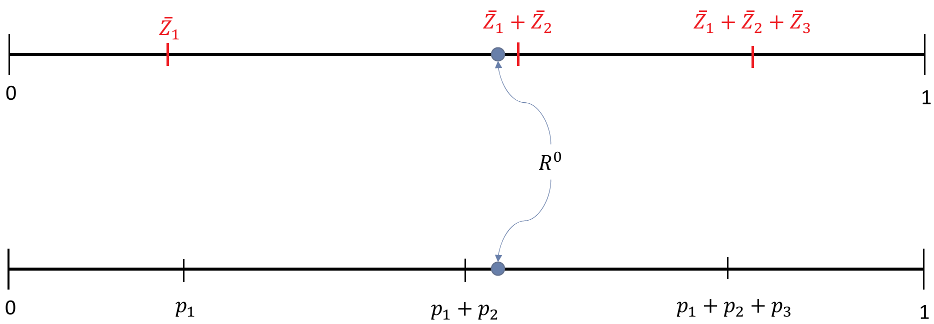
Our analysis consists, then, of two steps: (i) mapping the budget R (a point in the space of budget ratios) to the optimal bases of (LP(R, D)) and (ii) showing that, under BudgetRatio, R^t remains in the basic subset, which is consistent with the optimal (unknown to it) offline basis.

The mapping from budgets R to bases (step (i)) 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 in which the request variable y_3 is in the basis. On the left is the neighbor $\mathcal{K} = [1]$—so two 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)) is nontrivial already in the one-dimensional case (see Arlotto and Gurvich [2]).

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

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 [15]) do not satisfy this consistency. Consider the multisecretary problem and the scenario captured in Figure 3. With R^0 close enough to $p_1 + p_2$, we can have, with nonnegligible probability, both $\overline{Z}_1^T + \overline{Z}_2^T = \frac{1}{T} Z_1^T + \frac{1}{T} Z_2^T > R^0$ and $p_1 + p_2 < R^0$. In this realization, offline takes all type-1 and most 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 continues accepting type-3 requests until $R^t \le p_1 + p_2$, thus performing multiple nonbasic allocations. Under this randomized policy and with this initial budget R^0, the budget fluctuates around $p_1 + p_2$ and performs nonbasic allocations (too) frequently. BudgetRatio, in contrast, introduces a confidence interval: it does accept type-3 requests unless $R^0 \ge p_1 + p_2 + p_3/2$.

Figure 3. (Color online) 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 $R^0 \le \overline{Z}_1^T + \overline{Z}_2^T$ (top), the online randomized algorithm accepts type 3 with some probability (bottom).



This execution of nonbasic 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.

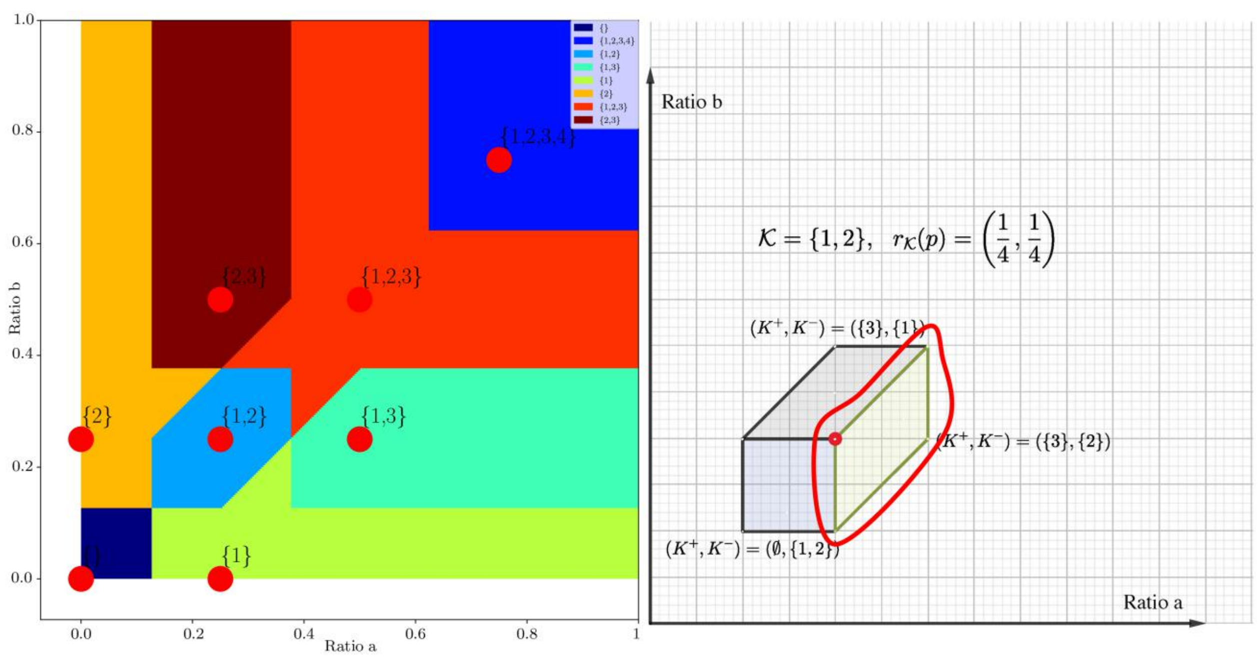
The centroid neighborhood $\mathcal{N}_{\{1,2\}}$ in the base example (see Figure 4 (left)) has three basic convex subsets as seen on the right of that figure, each of which corresponds to one optimal basis. As long as R^t is in the encircled rectangle, a given basis—let us call it $\mathcal{B}^{Encircled}$—is 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—for example, within the encirclement of the rectangle on the right—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 Figure 2. We further show that, if $\mathcal{B}^{Encircled}$ is optimal for offline then, indeed, R^t remains close to the encircled rectangle under the online policy BudgetRatio; see Proposition 3.

Below is how this road map is divided into sections:

- In Section 4.1, we define, in optimization terms, the centroid regions and their basic convex subsets. Specifically, we map 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. The regret would then be small if τ^π is close to the end of the horizon time T; see Proposition 3.
- Sections 4.2 and 5 are dedicated to proving Propositions 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 Section 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 we characterize in significant detail. This section ends with the proof of Proposition 2.

Figure 4. (Color online) The action regions and convex subsets in the base example.



- With the geometry mapped, Section 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 (i) a sticky boundary property (Theorem 2) that shows the residual budget process remains close to one centroid neighborhood/action region and (ii) a cone-containment property (Theorem 3) that stipulates that the BudgetRatio-controlled budget process remains constrained to the correct basic cone. Combined, Theorems 2 and 3 are the key ingredients in the proof of Proposition 3.
- Items (ii) and (iii) of Theorem 1 (parameter robustness) are proved in Section 6. Item (iv) (max bid price control) is proved Section 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 set the aggressiveness parameter $\alpha = \frac{1}{2}$ 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 $\overline{A}$ 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 $\lambda = (\mathcal{B}^{-1})'\overline{v}_{\mathcal{B}}$ be the dual variables associated with $\mathcal{B}$. If* (i) $\lambda \geq 0$ *and* (ii) $\overline{A}'\lambda \geq \overline{v}$, *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 $D \in \mathbb{R}^n_{\geq 0}$, the action region for $\mathcal{K}$ is the set of ratios $\mathcal{N}_{\mathcal{K}}(D) \subseteq \mathbb{R}^d_{\geq 0}$, where the algorithm serves exclusively requests in $\mathcal{K}$; that is, all requests $j \in \mathcal{K}$ are accepted and $j \notin \mathcal{K}$ are rejected:

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

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 when the constructs are relevant for both offline and BudgetRatio. As a general rule, D appears for geometric constructions (relevant for both online and offline), whereas p appears in statements concerning the stochastic process R^t driven by BudgetRatio .

It is in the following lemma, and only here, in which 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 $j \in \mathcal{K}$; that is, $I^t \geq A_j$ 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 in which the algorithm serves exclusively requests in $\mathcal{K}$ and the optimal basis for (LP(R, D)) is $\mathcal{B}$:

$$\begin{aligned}\mathcal{N}_{\mathcal{K}}(D, \mathcal{B}) &:= \{R \in \mathbb{R}^d_{\geq 0} : \text{ BudgetRatio uses } \mathcal{B} \text{ and serves exclusively } \mathcal{K} \text{ when } (R^t, p) = (R, D)\} \\ &= \left\{ R \in \mathbb{R}^d_{\geq 0} : \mathcal{B} \text{ optimal}, \quad y_{\mathcal{K}} = \left(\mathcal{B}^{-1}\begin{pmatrix} R \\ D \end{pmatrix}\right)_{\mathcal{K}} \geq \frac{1}{2}D_{\mathcal{K}}, \quad y_{\mathcal{K}^c} = \left(\mathcal{B}^{-1}\begin{pmatrix} R \\ D \end{pmatrix}\right)_{\mathcal{K}^c} < \frac{1}{2}D_{\mathcal{K}^c} \right\}, \end{aligned} \tag{18}$$

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 (i) 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 Section 4.2 and the latter in Section 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 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 $r_{\mathcal{K}}(p) = r_{[j]}(p) = \overline{F}_j$. Figure 1 has the centroid budgets $r_{\mathcal{K}}(p)$ for our two-dimensional base example.

The optimization problem (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 $r_{\mathcal{K}}(\hat{D})$, that is, optimal for 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 $\left(D^0 = \frac{1}{T}Z^T\right)$.

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 $\tilde{D} \in \mathbb{R}^n_{>0}$; that is, 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 $\tilde{D}$.*

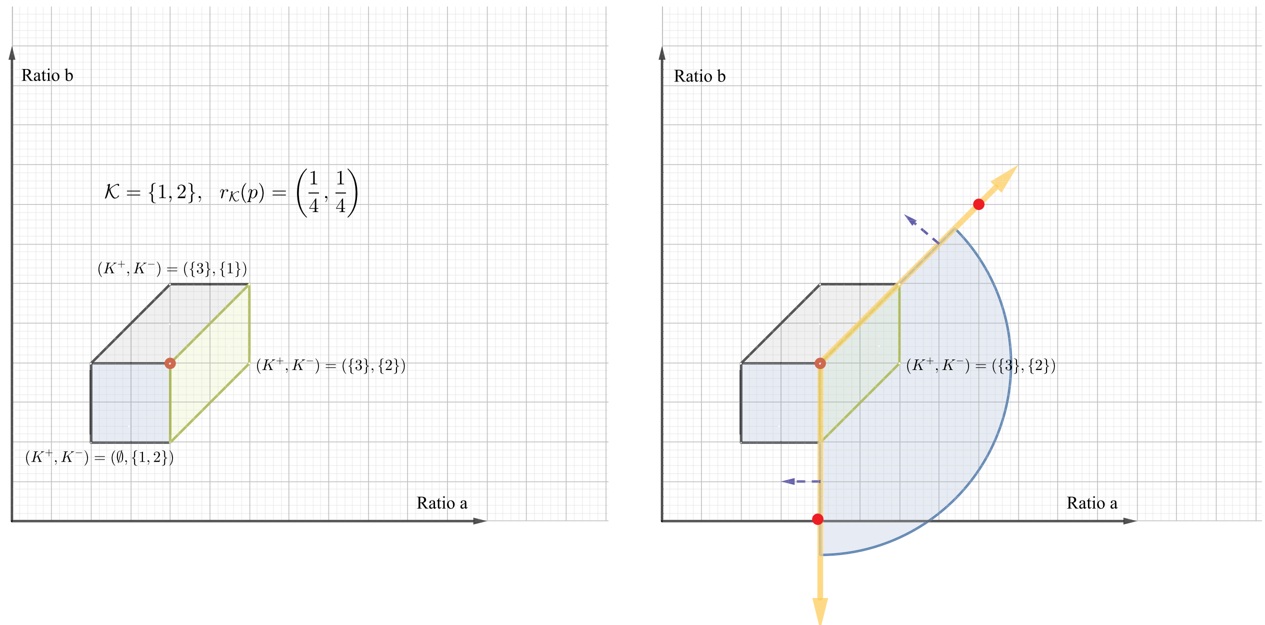
Finally, we define a useful relation between centroids.

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 $\mathcal{K}'$.

As do 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,*

Figure 5. (Color online) 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 corresponding to $(K^+, K^-, K^0) = (\{3\}, \{2\}, \emptyset)$. The dashed vectors are the outer normals, $\psi_1 = (-1, 1)'$ and $\psi_2 = (-1, 0)'$ that characterize the cone.



i. *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, that is, $b_i = 0$ for $i \notin K^0$. In particular, the set $\mathcal{K} \cup \kappa^+ \backslash \kappa^-$ is a centroid and a neighbor of $\mathcal{K}$.

ii. *The basis $\mathcal{B}$ is optimal for (R, D) if and only if $R \in \overline{\mathcal{N}_{\mathcal{K}}(D)}$ 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 $\sum_{\kappa^+ \subseteq K^+, \kappa^- \subseteq K^-} \alpha_{(\kappa^+, \kappa^-)} = 1$.

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

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

where $x_j \in [0, D_j/2), \ j \in K^+, \ x_j \in [0, D_j/2], \ j \in K^-$ and b is as in item (i).

In Figure 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 $\kappa^{\subseteq} \in K^+$ and $\kappa^- \subseteq K^-$ are $\{1,3\}$, $\{1\}$, and $\{1,2,3\}$. The set $\overline{\mathcal{N}_{\mathcal{K}}(D, \mathcal{B})}$ is the convex hull of the midpoints of the lines leading to those neighbors and corresponds to the rectangle on the right; 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

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

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, $\mathcal{N}_{\mathcal{K}}(D, \mathcal{B}) = \mathcal{N}_{\mathcal{K}}(D) \cap (r_{\mathcal{K}}(D) + \text{cone}(\mathcal{K}, \mathcal{B}))$.*

The properties of the outward normals to the cone are central to the proof of oracle containment (see Theorem 3). Figure 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 $\xi \in \text{cone}(\mathcal{K},\mathcal{B})$ if and only if $\max_{\psi\in\Psi(\mathcal{K},\mathcal{B})}\{\psi'\xi\} \le 0$ follows from the Minkowski–Weyl theorem (see, e.g., Bertsimas and Tsitsiklis [6, 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

i. $\psi[\kappa]'A_{K^+\setminus\kappa^+} = 0$, $\psi[\kappa]'A_{K^-\setminus\kappa^-} = 0$, *and* $\psi[\kappa]'e_{K^0\setminus\kappa^0} = 0$. *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* $|\kappa^0| = 1$.

ii. *Given* $\epsilon \ge 0$, $\max_{\psi\in\Psi(\mathcal{K},\mathcal{B})}\{\psi'(R - r_{\mathcal{K}}(D))\} \le \epsilon$ *if and only if* $\mathcal{B}^{-1}\begin{pmatrix} R - r_{\mathcal{K}}(D) \\ 0 \end{pmatrix} \ge -\mathbf{e}\epsilon$.

In the proof, the separating vectors are written explicitly in terms of the basis $\mathcal{B}$ and its inverse. They 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 LP(R, D)). Lemma 4 provides a tractable procedure to construct the action map—as in Figure 1—and to identify the bases associated with each centroid set.

We first identify a single centroid set $\mathcal{K}^0$. Having solved 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 Equation (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 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 Figure 4 (right): as long as R^t is in the proximity of the encircled rectangle in $\mathcal{N}_{\{1,2\}}$—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 $\{1,2\} \cup \{3\}$.

Henceforth, we fix

$$\epsilon^0 := \frac{1}{8}\min\{p_j : j \in [n]\} \wedge \min\{\varrho_i : i \in [d], \varrho_i > 0\}. \tag{21}$$

Lemma 7 (Optimal Bases and BudgetRatio Actions). *There exist constants M_1, M_2 such that, if $d_\infty(R^t,\mathcal{N}_{\mathcal{K}}(\mathcal{B},p)) \le \frac{\epsilon^0}{M_1}$, 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, $I_i^t \le M_2(T-t)$ for all $i \notin K^0(\mathcal{B})$.*

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

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

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

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

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 $r_{\mathcal{K}}(p) + \text{cone}(\mathcal{K},\mathcal{B})$. To exit $\mathcal{N}_{\mathcal{K}}(\mathcal{B},p)$, it suffices then to exit either of the two; this is formalized in Lemma 8 (see also Figure 6).

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 $d(R,\mathcal{N}_{\mathcal{K}}(D)) \le \epsilon'$ and $\max_{\psi\in\Psi(\mathcal{K},\mathcal{B})}\{\psi'(R - r_{\mathcal{K}}(D))\} \le \epsilon'$, then $d_\infty(R,\mathcal{N}_{\mathcal{K}}(\mathcal{B},D)) \le \epsilon$. Consequently, $\tau^{\epsilon,\mathcal{K}} \ge \tau^{\epsilon',\mathcal{K}}_{\text{region}} \wedge \tau^{\epsilon',\mathcal{B}}_{\text{cone}}$.*

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 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 $\tau^{\epsilon',\mathcal{K}}_{\text{region}}$ as in* Equation (22) *with ϵ' as in Lemma 8. Then,*

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

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

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 $\max_{\psi\in\Psi(\mathcal{K},\mathcal{B})}\{\psi'(\mathbb{E}[R^0] - r_{\mathcal{K}}(p))\} \le \epsilon'/2$, and $\tau^{\epsilon',\mathcal{B}}_{\text{cone}}$ be as in Equation (23). Then, for all $\ell \in [T]$,*

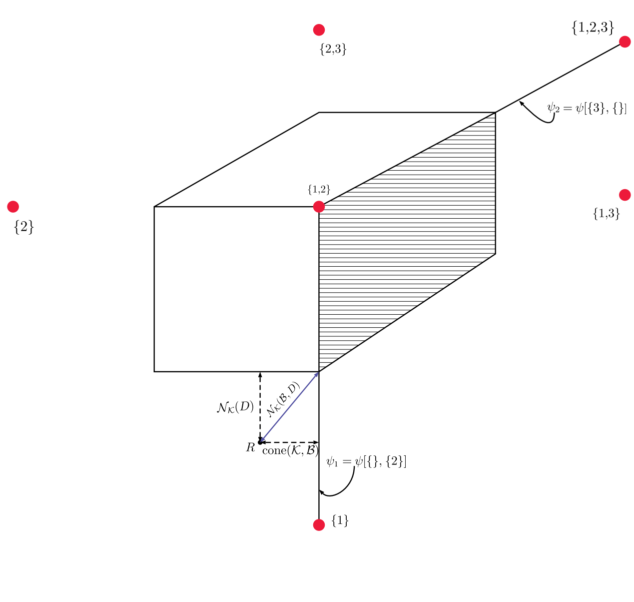
$$\mathbb{P}[T - \tau^{\epsilon',\mathcal{B}}_{\text{cone}} > \ell, R^0 - r_{\mathcal{K}}(D^0) \in \text{cone}(\mathcal{K},\mathcal{B})] \le m_1 e^{-m_2\ell},$$

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 $\mathbb{P}[T - \tau^{\epsilon',\mathcal{B}}_{\text{cone}} > \ell] \le m_1 e^{-m_2\ell}$.

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

The two remaining sections include the proofs of Theorems 2 and 3. For simplicity of exposition, these proofs are written for the case of no restock ($\rho = 0$). The changes are obvious if, instead, there is slow (but positive) restock; the restock rate is relevant only in the application of Lemma 2.

Figure 6. (Color online) Representation of Lemma 8 for the base example. The 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 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 $\kappa^- = \{2\}, \kappa^+ = \{3\}$, and hence, $|\kappa^-| + |\kappa^+| > 1$. It is in the interior of the cone.



5.1. Proof of Theorem 2

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

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

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 a solution $\theta_{\mathcal{K}}(y,D)$ at a point $R \in \mathcal{N}_{\mathcal{K}}(D)$. Lemma 9 introduces and characterizes this reference point $\theta_{\mathcal{K}}$.

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:

i. *$\theta_{\mathcal{K}}(y,D) \in$ closure($\mathcal{Y}(\mathcal{K},D)$) and $(y - \theta_{\mathcal{K}}(y,D))_j = 0$ for all $j \notin \kappa^+ \cup \kappa^-$.*

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

In the one-dimensional case of Figure 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

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

The values 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. It also states that centroids whose neighborhoods are suitably close must be neighbors.

Lemma 10. *There exists ϵ' and a constant M so that*

i. *For all $\check{\epsilon} \leq \epsilon'$, $d(R, \mathcal{N}_{\mathcal{K}}(D)) \leq M\check{\epsilon}$ if and only if $d(R_\bullet, \mathcal{N}_{\mathcal{K}}(D)) \leq \check{\epsilon}$.*

ii. *Fix two centroids $\mathcal{K}$ and $\mathcal{K}^0 = \mathcal{K} \cup \kappa^+ \backslash \kappa^0$. If both $d(R^t, \mathcal{N}_{\mathcal{K}}(p)) \leq \epsilon'$, $d(R^t, \mathcal{N}_{\mathcal{K}^0}(p)) \leq \epsilon'$. Then, $\mathcal{K}$ and $\mathcal{K}^0$ are neighbors in the sense of Definition 9.*

Finally, we have a simple lemma that shows 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, M_3 such that*

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

In turn, if $R^t \in \mathcal{N}_{\mathcal{K}}(p)$, then $d_\infty(R^t_\bullet, \mathcal{N}_{\mathcal{K}}(p)) \leq \frac{M_3}{T-t}$ for all $t \leq T - M_2$.

Proof of Theorem 2. Take $\check{\epsilon} = \epsilon'/M$ as in Lemma 10. We 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

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

By virtue of Lemma 10, this implies the same for R (with $\check{\epsilon}$ replaced by $M\check{\epsilon}$).

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

$$g^t := \|y^t - \theta^t\|^2 = \|\delta^t\|^2.$$

Whenever $g^t \leq \check{\epsilon}^2/nd$, we also have $d(R^t_\bullet, \mathcal{N}_{\mathcal{K}}(p)) \leq \check{\epsilon}$. Indeed, if $g^t \leq \check{\epsilon}^2/nd$, then by the Cauchy–Schwarz inequality and noting that A has binary entries, we have $|Ay^t - A\theta^t|_i^2 = (a_i'(y^t - \theta^t))^2 \leq \check{\epsilon}^2/d$. Because $\theta^t \in$ closure($\mathcal{Y}(\mathcal{K}, p)$) (Lemma 9), we have that $A\theta^t \in \overline{\mathcal{N}_{\mathcal{K}}(p)}$, and because $R^t_\bullet = Ay^t$, $\|Ay^t - A\theta^t\|^2 \leq \check{\epsilon}^2$ implies $d(R^t_\bullet, \mathcal{N}_{\mathcal{K}}(p)) \leq \check{\epsilon}$.

Setting $\varepsilon^2 := \check{\varepsilon}^2/nd$, we conclude that $g^t \le \varepsilon^2$ implies $d(R^t_\bullet, \mathcal{N}_\mathcal{K}(p)) \le \check{\varepsilon}$. The requirement $g^t \le \varepsilon^2$ is satisfied at $t = 1$ (and for T large) because $\mathbb{E}[R_0]_\bullet \in \mathcal{N}_\mathcal{K}(p)$ and using the second claim in Lemma 11.

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

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

Assuming Equation (25) and using $g^1 \le \varepsilon^2/2$, concentration arguments as in Arlotto and Gurvich [2, 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 Equation (25). Fix t with $g^t \le \varepsilon^2$ and let $\mathcal{K}^0 = \mathcal{K} \cup \kappa^+ \setminus \kappa^-$ be the centroid for which $R^t \in \mathcal{N}_{\mathcal{K}^0}(p)$. Because $d(R^t_\bullet, \mathcal{N}_\mathcal{K}(p)) \le \check{\varepsilon}$, the second item of Lemma 10 guarantees that $\mathcal{K}^0$ is a neighbor of $\mathcal{K}$.

Using Lemma 9 (item (i)), we obtain

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

where $\kappa = \kappa^+ \cup \kappa^-$. Our aim is to prove the quadratic bound $\mathbb{E}[\|\delta^{t+1} - \delta^t\|^2 \mid \mathcal{F}_t] = \mathcal{O}\left(\frac{1}{(T-t)^2}\right)$ and the linear bound $\mathbb{E}[(\delta^{t+1} - \delta^t)'_\kappa (\delta^t)_\kappa \mid \mathcal{F}_t] \le -\frac{M}{T-t}$, which together imply Equation (25).

We start with a fact about the different neighborhoods that the process R^t visits during its evolution.

5.1.1. Property of Visited Neighbors. For a vector $y \in \mathcal{Y}(\mathcal{K}, p)$, by definition, $y_j \ge p_j/2$ for $j \in \mathcal{K}$, and $y_j < p_j/2$ otherwise. Hence, if $g^t \le \varepsilon^2$ (in particular, $|y^t - \theta^t|_\infty \le \varepsilon$),

$$y^t_j \ge p_j/2 - \varepsilon \quad \forall j \in \mathcal{K} \quad \text{and} \quad y^t_j \le p_j/2 + \varepsilon \quad \forall j \notin \mathcal{K}. \tag{26}$$

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

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

The first fact follows directly from Lemma 11. For the second fact, take $j \in \kappa^+ \cup \kappa^-$. Using Equation (26) and recalling that $R^t \in \mathcal{N}_{\mathcal{K}^0}(p)$—in particular, $y^t_j \ge p_j/2, j \in \kappa^+$—we have $y^t_j = \frac{1}{2}p_j \pm \varepsilon$ and, in turn, that $j \in K^+(\mathcal{B}) \cup K^-(\mathcal{B})$. Because the solutions to the LP are Lipschitz continuous in the right-hand side (Cook et al. [11], Mangasarian and Shiau [18]),[3] $|y^t_j - y^{t+1}_j| \le M|R^t_\bullet - R^{t+1}_\bullet|_\infty = \mathcal{O}(\frac{1}{T-t})$. Thus, for a constant M and all $t \le T - M$, it must be that $y^{t+1}_j = \frac{1}{2}p_j \pm 2\varepsilon$ and, hence, (with ε small enough) that $j \in K^+(\overline{\mathcal{B}}) \cup K^-(\overline{\mathcal{B}})$.

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

$$\begin{aligned}\mathbb{E}[\delta^{t+1}_j - \delta^t_j \mid \mathcal{F}_t] &\le -\frac{M}{T-t} \quad j \in \kappa^+, \\ \mathbb{E}[\delta^{t+1}_j - \delta^t_j \mid \mathcal{F}_t] &\ge \frac{M}{T-t} \quad j \in \kappa^-.\end{aligned} \tag{28}$$

If $g^t \ge \varepsilon^2/2$, then there exists $j \in \kappa^+$ such that $\delta^t_j = y^t_j - D_j/2 \ge \varepsilon/\sqrt{2d}$ or some $j \in \kappa^-$ such that $\delta^t_j = y^t_j - D_j/2 \le -\varepsilon/\sqrt{2d}$ (recall that $\delta^t_j = 0$ for $j \notin \kappa^+ \cup \kappa^-$). Using Equation (28), we have, as desired, that

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

We turn to prove Equation (28). Recall that $\mathcal{K}^0$ is such that $R^t \in \mathcal{N}_{\mathcal{K}^0}(p)$ and that σ^t_j is the indicator that a request j is accepted at time t. Because only requests in $\mathcal{K}^0$ may be accepted at t, we have the identity $\mathbb{E}[I^{t+1}] = \varrho + I^t - \mathbb{E}[A_{\mathcal{K}^0}\sigma^t_{\mathcal{K}^0}]$, which implies $R^{t+1}(T-t-1) = R^t(T-t) - \mathbb{E}[A_{\mathcal{K}^0}\sigma^t_{\mathcal{K}^0}]$ and, in turn,

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

Let $\mathcal{B}$ be the optimal basis for LP(R^t, p) and $\overline{\mathcal{B}}$ be the optimal basis for LP(R^{t+1}, p); then, $\begin{pmatrix} R^t \\ p \end{pmatrix} = \mathcal{B}\begin{pmatrix} y^t \\ u^t \\ s^t \end{pmatrix}$ and $\begin{pmatrix} R^{t+1} \\ p \end{pmatrix} = \overline{\mathcal{B}}\begin{pmatrix} y^{t+1} \\ u^{t+1} \\ s^{t+1} \end{pmatrix}$ so that

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

By (27), $\mathcal{B}$ and $\overline{\mathcal{B}}$ are adjacent in the sense of Lemma 9 (item (ii)). Multiplying by $\overline{\mathcal{B}}^{-1}$, using the lemma, and recalling that $\delta_j = p_j/2$ for $j \in \kappa$, we have that

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

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, $\mathbb{E}[\sigma^t_j] = \mathbb{E}[\mathbb{1}_{\{J^t = j\}}] = p_j$, where, we recall, $J^t = j$ means that the arrival at time t is of type j. Because $y^t_j \le p_j/2 + \varepsilon$ for $j \in \kappa^+$ (see Equation (26)), we have using Equation (29) that (recall $\delta^t_j = y^t_j - D_j/2, j \in \kappa^+ \cup \kappa^-$)

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

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

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

here, we used $y^t_j \ge p_j/2 - \varepsilon$ (see Equation (26)). This concludes the proof of Equation (28).

5.1.3. Quadratic Bound. Finally, we prove that $\mathbb{E}[\|\delta^{t+1} - \delta^t\|^2 | \mathcal{F}_t] = \mathcal{O}\left(\frac{1}{(T-t)^2}\right)$. By Lemma 9 (item (i)), we have $\delta^t_\kappa = (y^t - p/2)_\kappa$ and $\delta^{t+1}_{\kappa} = (y^{t+1} - p/2)_{\tilde\kappa}$, where $\tilde\kappa = \tilde\kappa^+ \cup \tilde\kappa^-$ defines the neighbor $\tilde{\mathcal{K}}$ visited at time $t+1$; recall that $\delta^t_j = 0, j \notin \kappa$ and $\delta^{t+1}_j = 0, j \in \tilde\kappa$. If $j \in \kappa \cap \tilde\kappa$, $(\theta_{\mathcal{K}}(y^{t+1}, p))_j = (\theta_{\mathcal{K}}(y^t, p))_j = p_j/2$ so that $\delta^{t+1}_j - \delta^t_j = y^{t+1}_j - y^t_j$. If $j \notin \kappa \cup \tilde\kappa$, $(\theta_{\mathcal{K}}(y^{t+1}, p))_j = y^{t+1}_j, (\theta_{\mathcal{K}}(y^t, p))_j = y^t_j$ so that $\delta^{t+1}_j - \delta^t_j = 0$.

Consider the remaining coordinates. If $j \in \tilde\kappa^+$ (in particular $j \notin \mathcal{K}$) but $j \notin \kappa^+$, then $\delta^t_j = 0$ and $y^{t+1}_j \ge p_j/2 \ge y^t_j$ so that $|\delta^{t+1}_j - \delta^t_j| = |\delta^{t+1}_j| = |y^{t+1}_j - p_j/2| \le |y^{t+1}_j - y^t_j|$. 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

$$\|\delta^{t+1} - \delta^t\|^2 \le \|(y^{t+1} - y^t)_{\kappa \cup \tilde\kappa}\|^2 \le \|y^{t+1} - y^t\|^2.$$

The bound on the conditional expectation now follows from 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

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

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

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

Indeed, by the continuity of LP in the right-hand side, $y^t_j \le M\epsilon'$ for all $j \notin \mathcal{K} \cup (\cup_{\mathcal{B}} K^+(\mathcal{B}))$ and all $t < \tau^{\epsilon', \mathcal{K}}_{\text{region}}$. Similarly, $y^t_j \ge p_j - M\epsilon'$ for all $j \in \mathcal{K} \backslash (\cup_{\mathcal{B}} K^-(\mathcal{B}))$ and all such t.

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. Recall that the empirical demand distribution is $D^0 = \frac{1}{T} Z^T =: \overline{Z}^T$ and the empirical budget ratio is $R^0 = \frac{1}{T}(I^0 + \mathfrak{Z}^t) = \frac{1}{T} I^0 + \overline{\mathfrak{Z}}^t$, where $\overline{\mathfrak{Z}}^t = \frac{1}{T}\mathfrak{Z}^t$. By Lemma 5, $\mathcal{B}$ is optimal for offline if (i) $R^0 \in \mathcal{N}_{\mathcal{K}}(D^0)$ and (ii) $R^0 - r_{\mathcal{K}}(D^0) \in \text{cone}(\mathcal{K}, \mathcal{B})$. For ϵ^0 as in (21), the next lemma shows that, on the high-probability event

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

we can replace the requirement that $R^0 \in \mathcal{N}_{\mathcal{K}}(D^0)$ with one in which 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*

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

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

Lemma 13 is used for a subcase in the proof of Theorem 3. It stipulates that, if R is in the proximity of $d+1$ centroid neighborhoods, 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 $\kappa_i \in [n], i \in [d]$ be such that $|\kappa_i| = 1$ and $\mathcal{K}_i = \mathcal{K} \cup \kappa_i^+ \backslash \kappa_i^-$, $i \in [d]$ be d centroids. There then exists $\epsilon'', \delta > 0$ such that, if $d(R, \mathcal{N}_{\mathcal{K}}(D)), d(R, \mathcal{N}_{\mathcal{K}_i}(D)) \le \epsilon'', i \in [d]$, then*

i. *$\mathcal{K}$, $\mathcal{K}_i, i \in [d]$ 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 $\cup_{i\in[d]}\kappa_i^0 \subseteq K^0(\mathcal{B})$.*
ii. *R is in the strict interior of $\text{cone}(\mathcal{K}, \mathcal{B})$, that is, 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 $d+1$ centroids $\mathcal{K}_i$ other than $\mathcal{K}$ (and with $|\kappa_i| = 1$), then there would be $|\mathcal{K}| + |\mathcal{K}^c| + |\cup_i \kappa_i^+| + |\cup_i \kappa^-| = n + d + 1$ strictly positive variables at the solution to LP(R, D), whereas an optimal solution has at most $n + d$ basic variables.

The next lemma is the last ingredient for the proof of Theorem 3.

Let $\mathcal{K}$ be a centroid such that $\mathbb{E}[R^0] \in \mathcal{N}_{\mathcal{K}}(p)$. 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 also define $\mathcal{V} = \mathcal{V}[\tau_{\text{cone}}^{\epsilon',\mathcal{B}}]$ to be the random variable that counts how many centroid neighbors of $\mathcal{K}$, that is, $\mathcal{K}_i = \mathcal{K} \cup \kappa_i^+ \backslash \kappa_i^-$ with $|\kappa| = 1$, are visited by time $\tau_{\text{cone}}^{\epsilon',\mathcal{B}}$; $\mathcal{V} = 0$ means that no other such centroid was visited.

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 $\mathcal{K}_i = \mathcal{K} \cup \kappa_i^+ \backslash \kappa_i^-$ for $i \in [\mathcal{V}]$ be the ith 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

i. *For each of the centroids $i \in \mathcal{V}$, $\kappa_i^+ \subseteq K^+(\mathcal{B})$, $\kappa_i^- \subseteq K^-(\mathcal{B})$.*
ii. *Ψ_0 does not contain any of the vectors $\psi[\kappa^+, \kappa^-] \in \Psi(\mathcal{K}, \mathcal{B})$ for $\kappa^+ \in \cup_{i\in[d^0]}\kappa_i^+$, $\kappa^- \in \cup_{i\in[d^0]}\kappa_i^-$.*

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 $\mathcal{D}(\mathcal{B}) := \{T - \tau_{\text{cone}}^{\epsilon',\mathcal{B}} > \ell, \mathcal{M}(\mathcal{B})\}$.

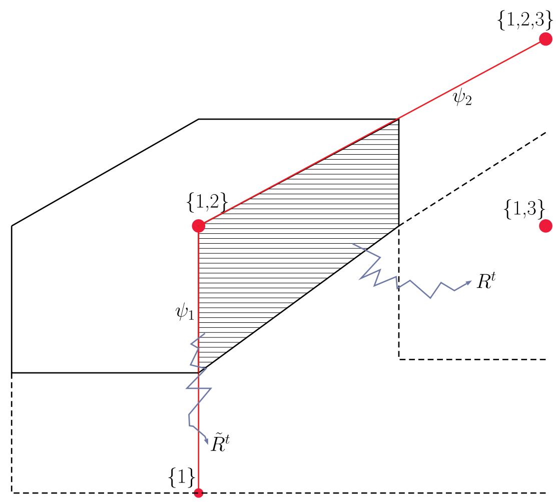
5.2.1. Outline of the Proof. To bound the measure of $\mathcal{D}(\mathcal{B})$, we consider two cases. In the first, at most $d-1$ action regions other than $\mathcal{N}_{\mathcal{K}}(p)$ are visited during the horizon ($\mathcal{V} < d$); this is the challenging case. The second case, in which d or 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 in which analysis is simpler; see Figure 7.

5.2.1.1. First Case (Boundary): $\mathcal{V} < d$. Assume that, over the interval $[1, \tau_{\text{cone}}^{\epsilon',\mathcal{B}}]$, at most $d^0 < d$ centroids other than $\mathcal{K}$ are visited. 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.

Let Ψ^0 be as in (32). We introduce processes $G_\psi^t, \psi \in \Psi^0$ with zero-mean increments that have the following properties on the event $\mathcal{M}(\mathcal{B})$:

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

Figure 7. (Color online) Two random walks for our base example. Solid lines enclose the region of interest $\mathcal{N}_{\{1,2\}}(D)$, and dashed lines enclose neighboring 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, that is, 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}.



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

By Lemma 14, we have the existence of an event $\mathcal{C}_\ell$ with $\mathbb{P}[\mathcal{V} = d^0, (\mathcal{C}_\ell)^c] \le m_1 e^{-m_2 \ell}$ such that if the ith 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 $\mathcal{K}_i = \mathcal{K} \cup \kappa_i^+ \setminus \kappa_i^-$.

We next study sample paths of R^t on the event $\{\mathcal{V} = d^0\} \cap \mathcal{C}_\ell$. We have the inventory equation $I^s = I^0 + \mathfrak{Z}^s - AY^s$, where Y^t_j is the total number of type-j requests accepted over $[1,t]$. Because $(T-s)R^s = I^s + (T-s)\varrho$ and $(T-s)r_\mathcal{K}(p) = (T-s)A_\mathcal{K} p_\mathcal{K}$,

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

and, after basic algebraic manipulations, that

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

where we define the centered process $\hat{\mathfrak{Z}}^t := \mathfrak{Z}^t - t\varrho$. Over the interval $[1, \tau^{\epsilon,\mathcal{B}}_{\text{cone}})$ 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 κ_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 centered process $\hat{Z}^t := Z^t - tp$ and using these identities together with Equation (34), we have

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

Define the process

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

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

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

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

$$\begin{aligned} G^T_\psi &= T\psi'(\mathbb{E}[R^0] - r_\mathcal{K}(p)) + \psi'(\hat{\mathfrak{Z}}^T - A_\mathcal{K} \hat{Z}^T_\mathcal{K}) \\ &= \psi'(I^0 + \mathfrak{Z}^T - A_\mathcal{K} Z^T_\mathcal{K}). \end{aligned}$$

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

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

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

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; see Lemma B.1.

5.2.1.2. Second Case (Strict Interior): $\mathcal{V} \geq d$. Let $\mathcal{K}_i$ be the ith 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 $\Omega_\ell = \{T - \tau_{\text{region}}^{\epsilon',\mathcal{K}} < \ell, T - \tau_{\partial,i}^1 < \ell, i \in [\mathcal{V}]\}$.

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

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

Combining the two cases (boundary and strict interior), we conclude that $\mathbb{P}[\mathcal{D}(\mathcal{B})] = \sum_{d^0=0}^{d-1}\mathbb{P}[\mathcal{D}(\mathcal{B}), \mathcal{V} = d^0] + \mathbb{P}[\mathcal{D}(\mathcal{B}), \mathcal{V} \geq d] \leq m_1 e^{-m_2\ell}$. The last implication in the theorem follows from Lemma 12 that guarantees, if $\mathbb{E}[R^0] \in \mathcal{N}_{\mathcal{K}}(D)$, then, on the event $\mathcal{A}^{\epsilon^0}$, $\mathcal{B}$ is offline optimal if and only if $\mathcal{M}(\mathcal{B})$ holds. Recalling that $\mathcal{D}(\mathcal{B}) := \{T - \tau_{\text{cone}}^{\epsilon',\mathcal{B}} > \ell, \mathcal{M}(\mathcal{B})\}$, we have that

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

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

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

and 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.

6.1. Demand Perturbation

Figure 8 illustrates a key idea behind Equation (8). When p is replaced with an estimate $\tilde{p}$ that satisfies (8), the centroids remain unchanged, but the shape of the centroid action regions is affected. Crucially, under Equation (8), the true centroid budgets (patterned circles) lie in the interior of the (misspecified) action regions. This guarantees that BudgetRatio, although fed with wrong probabilities, achieves constant regret.

Proof of Theorem 1 (Equation (8): Demand Robustness). Because $(\tilde{p}, \tilde{\varrho})$ 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 $\tilde{p}$; so does Lemma 8. Bounded regret then 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 $r_{\mathcal{K}}(p) \in \mathcal{N}_{\mathcal{K}}(\tilde{p})$ so that the initial condition is preserved and the proof of Theorem 2 uses the action region $\mathcal{N}_{\mathcal{K}}(\tilde{p})$ (instead of $\mathcal{N}_{\mathcal{K}}(p)$). The true probability p appears only in the expectation $\mathbb{E}[\sigma_{\mathcal{K}^0}^t]$ there. 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 of centroid neighbors and basic cones to D. □

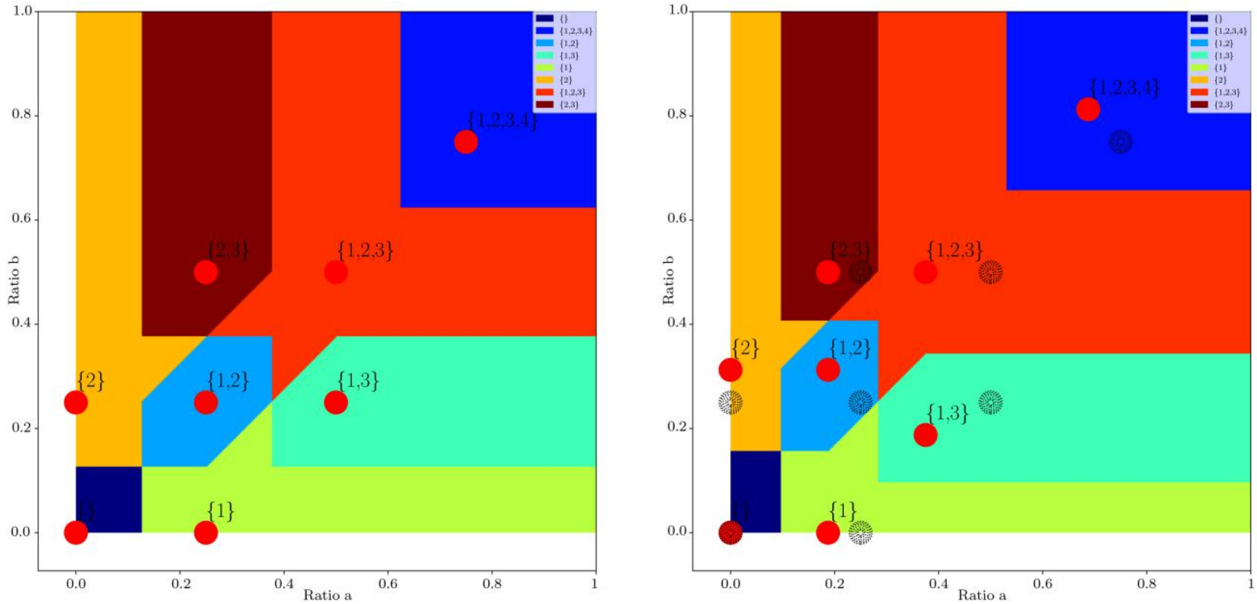


Figure 8. (Color online) Action regions with true and misspecified probabilities (p and $\tilde{p}$). (Left) Action regions of BudgetRatio when it is executed relative to p. (Right) Action regions when BudgetRatio uses $\tilde{p}_j = p_j - \frac{1}{16}$ for $j = 1, 3$ and $\tilde{p}_j = p_j + \frac{1}{16}$ for $j = 2, 4$.

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

$$\delta = \min\{e'_{[j+1]}p_{[j+1]} - e'_{[j]}p_{[j]},\ j \in [n-1]\} = \min\{p_j, j \in [n]\},$$

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

6.2. 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; $\hat{\varrho}^t_i$ 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 $\mathbb{P}[|(p, \varrho)' - (\hat{p}^{c\log T}, \hat{\varrho}^{c\log T})'|_\infty > \epsilon] \le 1/T$. 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 $|(p, \varrho)' - (\hat{p}^{c\log T}, \hat{\varrho}^{c\log T})'|_\infty > \epsilon$, the regret is at most $T \max_{j \in [n]}\{v_j\}$; this event's expected contribution to regret is at most $T \max_{j\in[n]}\{v_j\} \times 1/T = \max_{j\in[n]}\{v_j\} = \mathcal{O}(1)$. □

6.3. Reward Perturbation

We show that, under the δ-complementarity condition (Definition 3), the centroids are stable to local perturbations of v satisfying (7) 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 $(x, \lambda) \in \mathbb{R}^{d+2n} \times \mathbb{R}^{d+n}$ ($x = (y, u, s)$) 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 nonbasic (hence, $y_j = 0$), we have $([A'|I^n]\lambda)_j = v_j$, that is, complementarity is not strict. Definition 3 requires strict complementarity.

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

Proof of Theorem 1 (Equation (7): Reward Robustness). Fix a basis $\mathcal{B}$ associated to some centroid $\mathcal{K}$ under rewards v and let λ be the dual variable associated to $(\mathcal{B}, v)$. By Lemma 1, $\lambda \geq 0$ and $\overline{A}'\lambda \geq \overline{v}$ so that—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. $\tilde{\lambda} \geq 0$.
ii. $\overline{A}'\tilde{\lambda} \geq \overline{\tilde{v}}$.

First, we claim that $\lambda = (\mathcal{B}^{-1})'\overline{v}$ must have $\lambda_j = 0$ for $\{j : u_j \in \mathcal{B}\}$ and $\lambda_i = 0$ 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 nondegenerate ($x_{K^+}, x_{K^-} > 0$ and $b_{K^0} > 0$). 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:

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

The dual vector $\lambda = (\mathcal{B}^{-1})'\overline{v}$ is the unique solution to this linear system. Define now $\tilde{\lambda} = (\mathcal{B}^{-1})'\overline{\tilde{v}}$, where $\overline{\tilde{v}} = (\tilde{v}, 0, 0)'$. 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,

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

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

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

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

where the last inequality follows from Equation (7), noting that $\|A_j\|_1 \leq d + 1$ because all entries of A are in $\{0, 1\}$.

Lemma 1 allows us to conclude that all centroids and their bases remain the same under $\tilde{v}$. We finally observe that v is used only in the identification of the centroids; thus, none of our proofs or results change. □

Proof of Corollary 1 (Equation (10)). The primal and dual problems with a single resource ($d = 1$) are

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

Here, λ_0 denotes the resource multiplier and λ_j the demand multipliers for $j \in [n]$. Consider a basis of the form $\mathcal{B} = \{y_j : j = 1, \ldots, k+1\} \cup \{u_j : j = k+1, \ldots, n\}$, that is, all requests $j \leq k$ have $y_j = p_j$ and all requests $j > k + 1$ have $y_j = 0$; request $k + 1$ has both its request and slack variables in the basis. 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 holds for all bases if and only if $v_j \geq \delta$ for all $j \in [n]$ and $|v_j - v_{j'}| \geq \delta$ for all $j \neq j'$. □

6.4. 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 $V_j \sim \Phi_j$, where Φ_j is some unknown distribution, and $v_j = \mathbb{E}[V_j]$ represents the true expectation. At each time, the controller 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 Φ_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 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 $\mathbb{P}[|v - \hat{v}^{c \log T}|_\infty > \epsilon] \le 1/T$. Starting at $t = c' \log T$, we run BudgetRatio. Constant regret is guaranteed by Theorem 1, specifically Equation (7) there. On the event that $|v - \hat{v}^{c \log T}|_\infty > \epsilon$, the regret is at most $T \max_{j \in [n]}\{v_j\}$; this event's expected 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 (iv) 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 δ > 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 because, at the centroid budget $r_{\mathcal{K}}(p)$, $y_j = p_j > 0$ for all bases $\mathcal{B}$ associated to $\mathcal{K}$ so that $y_j \ge p_j/2$ for any $R \in \mathcal{N}_{\mathcal{K}}(p)$, and complementary slackness guarantees that $v_j \ge \overline{A}_j' \lambda$ 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 $\overline{A}_j' \lambda_{\mathcal{K}}^{\partial} = 0$. Overall, $v_j \ge \max_{\lambda \in \Lambda(R^t)}\{\overline{A}_j'(\lambda + \lambda^{\partial}(R^t))\}$ 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 $(\lambda_{\mathcal{K}}^{\partial})_j \ge 2 v_j \mathbf{e}_j$ so that $v_j < \max_{\lambda \in \Lambda(R^t)}\{\overline{A}_j'(\lambda + \lambda^{\partial}(R^t))\}$ and j is rejected by max bid price as required, or (ii) there is a basis $\mathcal{B}$ associated with $\mathcal{K}$ such that $y_j \notin \mathcal{B}$. In that case, δ-complementarity yields $v_j \le \overline{A}_j' \lambda - \delta$ so that j fails the acceptance condition of the policy with λ being the dual vector associated to this basis $\mathcal{B}$. In turn, $v_j < \max_{\lambda \in \Lambda(R^t)}\{\overline{A}_j' \lambda\}$ and max bid price rejects request j as required.

It remains to prove for $j \notin \mathcal{K}$ that, if $j \notin \partial(\mathcal{K})$, there must exist a basis associated to $\mathcal{K}$ for which $y_j \notin \mathcal{B}$.

Suppose that $y_j \in \mathcal{B}$ for all bases associated to $\mathcal{K}$. 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 $r_{\mathcal{K}}(p) = \sum_{j \in \mathcal{K}} A_j p_j \ge \zeta A_j$. By the Lipschitz continuity of LPs, $R = r_{\mathcal{K}}(p) \pm \zeta A_j \in \mathcal{N}_{\mathcal{K}}(p)$ for all ζ sufficiently small (that is, $y_l \ge p_l - M\zeta \ge p_l/2$ for all $l \in \mathcal{K}$). Let $\mathcal{B}$ be the optimal basis at $R = r_{\mathcal{K}}(p) - \frac{1}{2}\zeta A_j$ and let (y, u, s) be the optimal solution. Because $y_j \in K^+(\mathcal{B})$, we have by Lemma 4 that

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

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

Remark 7 (Adaptively Generating Bid Prices). In Remark 4, we discuss how the geometry—the centroids and associated bases—can be precomputed and, in turn, so can the max bid prices. There is an adaptive alternative to this, possibly expensive, precomputation.

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 $y_j \ge \overline{p}_j/2$ 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 not-yet-visited centroid neighborhood[4] that we must recompute the collection of bases. Theorem 2 implies that R^t spends most of its horizon in $\mathcal{N}_{\mathcal{K}}$ and its neighbors; see Remark 6. By Lemma 13, at most $d + 1$ centroid neighborhoods are 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 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 identify robustness boundaries of BudgetRatio .

i. Modeling 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.

ii. Implementation: We prove that, under a suitable complementarity 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.

iii. Model parameters: We consider the effect of running BudgetRatio with misspecified arrival probabilities (p, ϱ) and rewards v. We prove 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 generalize, to multiple dimensions, separation conditions that were previously known only in the one-dimensional (single-resource) case.

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

i. Perturbation conditions and learning. The separation conditions in Definitions 4 and 3, whereas sufficient, may not be necessary for the robustness of BudgetRatio . For instance, whereas the δ-complementarity is necessary for the one-dimensional case ($d = 1$), it is unclear whether our generalization to $d > 1$ is necessary.

ii. 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 restocking 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 the relative simplicity of inventory allocation problems is maintained. A more complete characterization of restock rates that permit this simplicity would be illuminating.

Appendix A. Restock and Feasibility of Constant Regret

If Assumption 1 fails, it is not generally possible to achieve constant regret relative to the offline (2). Consider, for illustration, 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 $T = 1, \ldots, 2{,}000$. For each horizon, we ran 1,000 replications of both the optimal policy and the offline upper bound in (2). In Figure A.1 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 $\Omega(\sqrt{T})$.

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 and achieved by the policy that accepts requests of types 1 and 2 whenever possible. With further simulations, it can be verified that any $\varrho \in [0.4, 0.6]$ does not produce constant regret. This suggests that 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 in which we can prove that the regret exhibits $\sqrt{T}$.

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

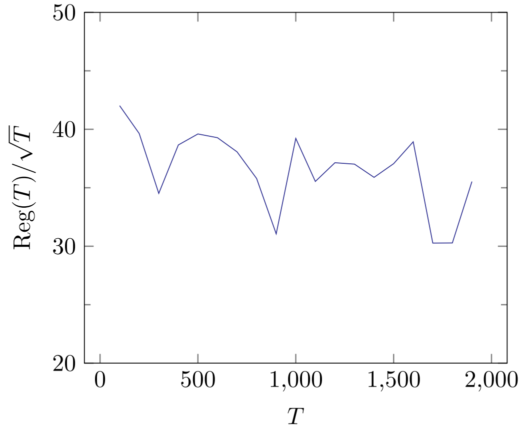
$$V^*_{\text{off}}(T) - V^*_{\text{on}}(T) = c\sqrt{T} + o(\sqrt{T}), \ \text{as } T \to \infty,$$

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

Proof. This counterexample is a network with one request type and one resource, both arriving with probability p. It is a so-called two-sided queue in which one side balks immediately if not served upon arrival. Figure A.2 is the numerical illustration of the $\sqrt{T}$ regret.

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 the offline (online)

Figure A.1. (Color online) 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 $\Omega(\sqrt{T})$.



solution. Then, $V^*_{\text{off}}(T) = \mathbb{E}[\min\{\mathfrak{Z}^t, Z^t\}] = \mathbb{E}[\mathfrak{Z}^T - I^T_{\text{off}}]$, where

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

The optimal online policy serves 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.

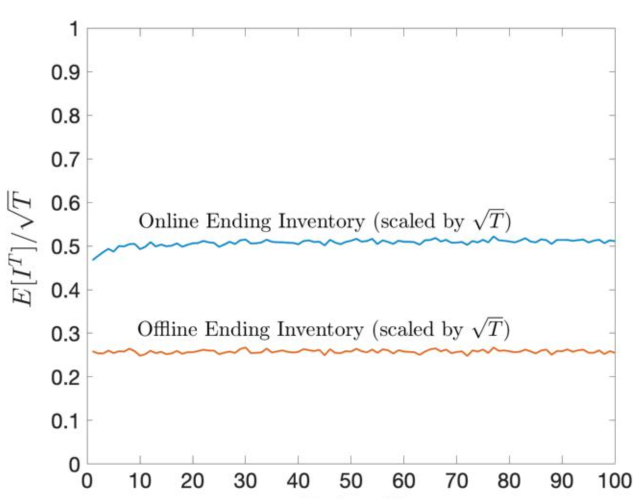
The process $\mathfrak{Z}^t - Z^t$ is a random walk starting at zero and with i.i.d. zero-mean increments $X_1, \dots, X_T$ taking values $\{-1, 0, 1\}$ with probabilities $p(1-p), 1-2p(1-p), p(1-p)$; X_t is the difference between the restock at t (0 or 1) and the request arrival (0 or 1). Write $G^T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

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

where $\sigma^2 := 2p(1-p)$. Because $I^T_{\text{off}} = (G^T)^+$, we have, by the continuous mapping theorem, that $\frac{1}{\sqrt{T}} I^T_{\text{off}} \Rightarrow (\mathcal{N}(0, \sigma^2))^+$, and the convergence here also holds in expectation (see Gut [14, theorem 4.2]). On the other hand, the online inventory satisfies the queueing recursion $I^{t+1}_{\text{on}} = [I^t_{\text{on}} + X_t]^+$ so that

$$I^T_{\text{on}} = \sup_{t \le T}\{G^T - G^t\}.$$

Figure A.2. (Color online) A single requires a type with a single restocking resource. 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 10,000 replications. The y-axis is the (expected) ending inventory divided by $\sqrt{T}$.



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

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

where $\mathcal{Z}\overset{d}{=}|\mathcal{N}(0,\sigma^2)|$ (see Gut [14, theorem 12.2]). This convergence also holds in expectation by Doob's moment inequality and standard results for moment convergence (see Chung [10, theorem 4.5.2]). 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

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

and, in turn, that $V^*_{\text{off}}(T) - V^*_{\text{on}}(T) = \Omega(\sqrt{T})$. □

Appendix B. Proofs of Lemmas

Proof of Lemma 1. The dual problem of (LP(R, D)) is

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

For any basis $\mathcal{B}$, our defined vector $\lambda = (\mathcal{B}^{-1})'\overline{v}_{\mathcal{B}}$ 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 $(R,D)'\lambda = x'_{\mathcal{B}}\overline{v}$ so that, by weak duality, $\mathcal{B}$ is optimal provided that $x_{\mathcal{B}}$ is primal feasible, that is, $x'_{\mathcal{B}} = \mathcal{B}^{-1}\begin{pmatrix}R\\D\end{pmatrix}\ge 0$ (see Bertsimas and Tsitsiklis [6, corollary 4.2]). Conversely, for any (R, D), when the simplex algorithm terminates, it produces an optimal basis $\mathcal{B}$, where $\lambda = (\mathcal{B}^{-1})'\overline{v}_{\mathcal{B}}$ is a feasible solution (it has nonpositive reduced costs), and it is optimal because $(R,D)'\lambda = x'_{\mathcal{B}}\overline{v}$ (see Bertsimas and Tsitsiklis [6, 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 $I^t_i \ge |\{j\in\mathcal{K}: A_{ij}=1\}|$ for some M and all $t\le T-M$, 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 LP(R^t, p) implies $Ay \le \frac{1}{T-t}I^t + \varrho$. Because $j\in\mathcal{K}$, it must be that $y_j\ge\frac{1}{2}\overline{p}_j$ 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

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

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

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

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

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

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

Because $\mathcal{B}$ satisfies properties (i) and (ii) in Lemma 1, $\mathcal{B}$ 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 at the centroid budget. □

Proof of Lemma 4

Item (i). 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 $\mathcal{B}^{-1}\begin{pmatrix}R\\D\end{pmatrix}\ge 0$.

Recall the augmented matrix $\overline{A}$ given by

$$\overline{A} = \begin{bmatrix}A & 0 & \mathbf{I}_d\\ \mathbf{I}_n & \mathbf{I}_n & 0\end{bmatrix},$$

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 submatrix $\mathcal{B}$ has a subset of these columns and can be written as

$$\mathcal{B} = \begin{bmatrix} A_{\mathcal{K}\cup K^+} & 0 & \mathbf{I}^d_{K^0} \\ \mathbf{I}^n_{\mathcal{K}\cup K^+} & \mathbf{I}^n_{\mathcal{K}^c\cup K^-} & 0 \end{bmatrix}, \tag{B.1}$$

where $\mathbf{I}^n_{\mathcal{K}\cup K^+}$ has the columns of $\mathbf{I}_n$ 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 $r_{\mathcal{K}}(D)$, and 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 deduce the inclusions $\kappa^+ \subseteq \mathcal{K}^c \cap K^+$ and $\kappa^- \subseteq \mathcal{K}\cap K^-$. By inspection, we have the identities

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

Premultiplying 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 nonnegative.

Item (ii). 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 nonnegative combination of the right-hand sides has the same optimal basis. By item (i) of the lemma, we are taking nonnegative combinations of right-hand sides that 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, and 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 $y_j, u_j \geq 0$. Because $R \in \overline{\mathcal{N}_{\mathcal{K}}(D)}$, 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 $(\overline{y}, \overline{u}, \overline{s})$ given by

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

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

For convenience, let us relabel (and reorder) 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

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

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

We apply the following transformation:

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

Because $y_{K^+} \le D_{K^+}/2$ and $y_{K^-} \ge D_{k^-}/2$ for $R \in \overline{\mathcal{N}_{\mathcal{K}}(D)}$, 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 $z = (y - \overline{y}, u - \overline{u})$ 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 a vector in the polyhedron 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 $\overline{y}, \overline{u}$), 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

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

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 (iii). We proved that $\mathcal{B}$ is optimal for (R, D) with $R \in \overline{\mathcal{N}_{\mathcal{K}}(D)}$ if and only if it can be written as

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

Define

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

Putting $x_j := \alpha_j D_j$, we can write

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

We claim that, for R as above, the solution of LP(R, D) is

$$\begin{pmatrix} y \\ u \\ s \end{pmatrix} = \mathcal{B}^{-1}\begin{pmatrix} R \\ D \end{pmatrix} = \begin{pmatrix} D_{\mathcal{K}} \\ D_{\mathcal{K}^c} \\ 0 \end{pmatrix} + \begin{pmatrix} x_{K^+} - x_{K^-} \\ -x_{K^+} + x_{K^-} \\ b \end{pmatrix}.$$

By definition, $R \in \mathcal{N}_{\mathcal{K}}(\mathcal{B}, D)$ if and only if (i) $R \in \mathcal{N}_{\mathcal{K}}(D)$ and (ii) the basis is optimal. Recalling that $K^+ \subseteq \mathcal{K}^c, K^- \subseteq \mathcal{K}$, the claim above guarantees that $y_{\mathcal{K}} \geq \frac{1}{2}D_{\mathcal{K}}$ and $y_{\mathcal{K}^c} < \frac{1}{2}D_{\mathcal{K}^c}$—in turn $R \in \mathcal{N}_{\mathcal{K}}(D)$—if and only if x_j are as assumed.

From all the above, we immediately have that, if $R \in \mathcal{N}_{\mathcal{K}}(\mathcal{B}, D)$, then R has the stated representation. For the other direction, if R has the stated representation with x_j as assumed, then the claim guarantees that both $R \in \mathcal{N}_{\mathcal{K}}(D)$ and $\mathcal{B}^{-1}\begin{pmatrix} R \\ D \end{pmatrix} \geq 0$. Because $\mathcal{B}$ is associated to $\mathcal{K}$, this $\mathcal{B}$ is then optimal at R by Lemma 1 so that $R \in \mathcal{N}_{\mathcal{K}}(\mathcal{B}, D)$.

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

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

and by the definition of centroid, it follows that

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

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 (iii)), 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 $x_j \leq D_j/2, j \in K^+$ and $x_j < D_j/2, j \in K^+ \cup K^-$. We can now use Lemma 4 (item (iii)) to conclude that $R \in \mathcal{N}_{\mathcal{K}}(D, \mathcal{B})$. □

Proof of Lemma 6. We construct the separating vectors 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, that is, with $|\kappa^+| + |\kappa^-| + |\kappa^0| = 1$.

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, $\psi[\kappa]'A_j = 0$, $j \in K^+(\mathcal{B})\backslash\kappa^+$, $\psi[\kappa]'A_j = 0$, $j \in K^-(\mathcal{B})\backslash\kappa^-$, and $\psi[\kappa]'e_i = 0$ for $i \in K^0(\mathcal{B})\backslash\kappa^0$. 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 $|\kappa^0| = 1$.

For item (ii), we claim that for $j \in \mathcal{K}\backslash K^-$ or $l \in \mathcal{K}^c\backslash K^+$ $\left(\mathcal{B}^{-1}\begin{pmatrix} A_l \\ 0 \end{pmatrix}\right)_j$ for all $l \in [n]$. Recall that, for $R \in \mathcal{N}_{\mathcal{K}}(B, D)$, $y_j = D_j$ for all $j \in \mathcal{K}\backslash K^-$ and $y_j = 0$ for all $j \in \mathcal{K}^c\backslash K^+$. Take a point R in $\mathcal{N}_{\mathcal{K}}(\mathcal{B}, D)$ by setting, for each κ with $|\kappa| = 1$, either $x_{\kappa^+} > 0$, $x_{\kappa^-} > 0$ or $b_{\kappa^0} > 0$. For $l \in [n]$ and for sufficiently small δ, the continuity of the LP in the right-hand side guarantees that $\tilde{R} = R + A_l\delta \in \mathcal{N}_{\mathcal{K}}(\mathcal{B}, D)$. Take now $j \in \mathcal{K}\backslash K^-$ or $l \in \mathcal{K}^c\backslash K^+$. It must be that $\left(\mathcal{B}^{-1}\begin{pmatrix} A_l \\ 0 \end{pmatrix}\right)_j = 0$, or otherwise, y_j would obtain a value other than D_j for $j \in \mathcal{K}\backslash K^-$ or other than zero for $l \in \mathcal{K}^c\backslash K^+$.

In particular, $\max_{\psi \in \Psi(\mathcal{K},\mathcal{B})}\{\psi'(R - r_{\mathcal{K}}(D))\} \leq \epsilon$ if and only if $-\mathcal{B}^{-1}\begin{pmatrix} R - r_{\mathcal{K}}(D) \\ 0 \end{pmatrix} \leq \epsilon\mathbf{e}$. □

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

Finally, $s^0_i = 0$ for all $i\notin K^0$; hence, $\overline{s}_i = R^t_i - (A\overline{y})_i \le \epsilon^0$ for all such i, which implies $R^t_i = \frac{1}{T-t}I^t_i + \varrho_i \le (Ay)_i + \epsilon$, and using $y\le p$, we get as required that $I^t_i \le (T-t)((Ap)_i + \epsilon^0)$ 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. Also recall the surplus-corrected budget ratio $R_\bullet$ as defined in (24).

We argue that, for any ϵ'', we can choose ϵ' small enough so that, under the assumptions of the lemma,

i. $y_j \ge D_j/2 - \epsilon''$ for all $j\in\mathcal{K}$.

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

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

Item (i): Because $d_\infty(R,\mathcal{N}_\mathcal{K}(D)) \le d(R,\mathcal{N}_\mathcal{K}(D)) \le \epsilon'$, the Lipschitz continuity of LPs implies that $y_j \ge D_j/2 - M\epsilon'$ 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): By assumption and using item (ii) of Lemma 6, we have that $\mathcal{B}^{-1}\begin{pmatrix} R - r_\mathcal{K}(D) \\ 0\end{pmatrix} \ge -\mathbf{e}\epsilon'$.

Take the vector $\tilde{R} = R + \mathcal{B}\mathbf{e}\epsilon'$. Then, $\mathcal{B}^{-1}\begin{pmatrix} \tilde{R} - r_\mathcal{K}(D) \\ 0\end{pmatrix} = \mathcal{B}^{-1}\begin{pmatrix} R - r_\mathcal{K}(D) \\ 0\end{pmatrix} + \mathbf{e}\epsilon' \ge 0$ so that $\tilde{R} - r_\mathcal{K}(D)\in\text{cone}(\mathcal{K},\mathcal{B})$, and

$$d_\infty(R,\tilde{R}) = |R-\tilde{R}|_\infty \le \epsilon'|\mathcal{B}|_\infty.$$

Because $d(R,\mathcal{N}_\mathcal{K}(d)) \le \epsilon'$, we also have that $d_\infty(\tilde{R},\mathcal{N}_\mathcal{K}(D)) \le M\epsilon'$, for a suitable constant M.

Recalling that $\mathcal{N}_\mathcal{K}(\mathcal{B},D) = \mathcal{N}_\mathcal{K}(D)\cap(r_\mathcal{K}(D) + \text{cone}(\mathcal{K},\mathcal{B}))$,

$$\begin{aligned} d_\infty(R,\mathcal{N}_\mathcal{K}(\mathcal{B},D)) &\le d_\infty(R,\tilde{R}) + d_\infty(\tilde{R},\mathcal{N}_\mathcal{K}(\mathcal{B},D)) \\ &= d_\infty(R,\tilde{R}) + d_\infty(\tilde{R},\mathcal{N}_\mathcal{K}(D)) \le M\epsilon' =: \epsilon. \end{aligned}$$

for a suitably modified constant M. □

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

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

This implies $\theta_j \ge \frac{D_j}{2}$ for $j\in\mathcal{K}$ because, by definition, $\theta_j = \frac{D_j}{2}$ for $j\in\kappa^-$; also, $\theta_j \le \frac{D_j}{2}$ for $j\in\mathcal{K}^c$ because $\theta_j = \frac{D_j}{2}$ for $j\in\kappa^+$. That $(y - \theta_\mathcal{K}(y,D))_j = 0$ for all $j\notin\kappa^+\cup\kappa^-$ follows immediately from our definition of θ.

Furthermore, θ is optimal at the ratio $R^\theta := A\theta$. Indeed, if we take $\mathcal{B}$ as a basis that $\mathcal{K}$ and $\mathcal{K}^0$ share, then it has all of the y variables in $\mathcal{K}$ and $\kappa^+\cup\kappa^-$, in particular, all the variables for which $\theta_j > 0$. Thus, the support of θ is all basic variables, and we have

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

which proves the optimality of θ at R^θ by Lemma 1. In turn, $\theta_\mathcal{K}(y,D)\in\text{closure}(\mathcal{Y}(\mathcal{K},D))$.

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

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

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

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

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

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

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

where the union is over bases associated to $\mathcal{K}$.

We use the following claim: (*) If $i \notin \cup_\mathcal{B}K^0(\mathcal{B})$, then there exists $\kappa^+ \subseteq \cup_\mathcal{B}K^+(\mathcal{B})$ such that $\mathbf{e}_i \ge A_{\kappa^+}$.

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

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

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

Because $\overline{R} \in \mathcal{N}_\mathcal{K}(D)$, $\overline{y}_j \le D_j/2, j \in \cup_\mathcal{B}K^+(\mathcal{B})$. By the Lipschitz continuity of LPs, $y_j \le D_j/2 + M\check{\epsilon}, j \in \cup_\mathcal{B}K^+(\mathcal{B})$. If $s_i > 0$ for some $i \notin \cup_\mathcal{B}K^0(\mathcal{B})$, then by the claim (*), for small enough δ, $y + \mathbf{e}_\kappa\delta$ 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

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

as required.

It remains to prove the claim (*). Suppose that there exists no $\kappa^+ \in \cup_\mathcal{B}K^+(\mathcal{B})$ with $\mathbf{e}_i \ge A_{\kappa^+}$. Take small enough δ so that $r_\mathcal{K}(D) + \mathbf{e}_i\delta \in \mathcal{N}_\mathcal{K}(D)$; such δ always exists by the definition $\mathcal{N}_\mathcal{K}(D)$ and the Lipschitz continuity of the LP solution. Because there exists no κ^+ with $\mathbf{e}_i \ge A_{\kappa^+}$ 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 second item of the lemma, suppose that $d(R^t,\mathcal{N}_\mathcal{K}(p)) \le \epsilon', d(R^t,\mathcal{N}_{\mathcal{K}^0}(p)) \le \epsilon'$; in particular, $d_\infty(R^t,\mathcal{N}_\mathcal{K}(p)) \le \epsilon'$, $d_\infty(R^t,\mathcal{N}_{\mathcal{K}^0}(p)) \le \epsilon'$. Here, $\mathcal{K}^0 = \mathcal{K} \cup \kappa^+ \backslash \kappa^-$ for some $\kappa^+, \kappa^- \subseteq [n]$. Take the point $\overline{R}^t \in \overline{\mathcal{N}_\mathcal{K}(p)}$ such that $d(R^t,\mathcal{N}_\mathcal{K}(p))) = d(R^t,\overline{R}^t)$; let $\overline{x} = (\overline{x},\overline{u},\overline{s})$ be the solution to LP$(\overline{R}^t, p)$.

Then, LP continuity guarantees that $\overline{y}^t_j \ge p_j/2 - \epsilon', j \in \mathcal{K} \cup \kappa^+$. It also guarantees that $\overline{y}^t_j \ge p_j/2 - \epsilon'$ for $j \in \kappa^-$. Thus, the optimal basis at $\overline{R}^t$ must have $\kappa^+ \subseteq K^+(\mathcal{B})$ and $\kappa^- \subseteq K^-(\mathcal{B})$.

By the first item of Lemma 4, $\mathcal{B}$ is optimal at $R = r_{\mathcal{K}^0}(D) = r_\mathcal{K}(D) + A_{\kappa^+}D_{\kappa^+} - A_{\kappa^-}D_{\kappa^-}$, and hence, it is associated to $\mathcal{K}^0$. □

Proof of Lemma 11. Define $\overline{R}^t_i = \frac{I^t_i \wedge M_0(T-t)}{T-t} + \varrho_i$, where $M_0 = 2\sum_{j\in[n]} p_j$. Because $y^t_j \le p_j, j \in [n]$, it is immediate that $LP(R^t, D)$ has the same optimal request variable values as $LP(\overline{R}^t, D)$:

$$\overline{R}^{t+1}_i - \overline{R}^t_i = \frac{I^{t+1}_i \wedge M_0(T-t-1)}{T-t-1} - \frac{I^t_i \wedge M_0(T-t)}{T-t}.$$

Because $I^{t+1}_i \le I^t_i + 1$, $I^{t+1}_i \wedge M_0(T-t-1) \le I^t_i \wedge M_0(T-t) + 1$ so that

$$|\overline{R}^{t+1}_i - \overline{R}^t_i| \le \frac{1}{T-t-1} + \frac{I^t_i \wedge M_0(T-t)}{(T-t)(T-t-1)} \le \frac{M_0+1}{T-t-1}.$$

Finally, the Lipschitz continuity of LPs guarantees that $|y^{t+1} - y^t|_\infty \le \tilde{M}_0|\overline{R}^{t+1} - \overline{R}^t|_\infty$. □

Proof of Lemma 12. Let $\mathcal{N}^+_\mathcal{K}(D,\mathcal{B}) \supset \mathcal{N}_\mathcal{K}(D,\mathcal{B})$ have $x_j \in [0, D_j), j \in K^+$ and $x_j \in [0, D_j], j \in K^-$ (instead of $x_j \in [0, D_j/2)$ and $x_j \in [0, D_j/2]$ as in Lemma 4, item (iii)).

Let

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

and let $\mathcal{N}^+_\mathcal{K}(D^0)$ be as above with D^0 replacing D. By assumption, $\mathbb{E}[R^0] = \frac{1}{T}I^0 + \varrho \in \mathcal{N}_\mathcal{K}(p)$ and $(\overline{Z}^T, \overline{\mathfrak{Z}}^T) \in \mathcal{A}^{\epsilon^0}$. Because $R^0 = \frac{1}{T}I^0 + \overline{\mathfrak{Z}}^T$, we have $d(R^0,\mathcal{N}_\mathcal{K}(p)) \le 2\epsilon^0 = \frac{1}{4}\min\{p_j : j \in [n]\} \wedge \min\{\varrho_i : i \in [d], \varrho_i > 0\}$. Because, on $\mathcal{A}^{\epsilon^0}$, $d(\partial\mathcal{N}^+_\mathcal{K}(D^0),\mathcal{N}_\mathcal{K}(p)) \ge \min_j\{D^0_j - p_j/2\} \ge \min_j\{p_j/2\} - \epsilon^0 \ge 3\epsilon^0$, we have that $R^0 \in \mathcal{N}^+_\mathcal{K}(D^0)$.

It then follows from item (i) of Lemma 4 (with D replaced with D^0 there) that $R^0 - A_\mathcal{K}\overline{Z}^0_\mathcal{K} = R^0 - r_\mathcal{K}(D^0) \in \text{cone}(\mathcal{K},\mathcal{B})$ 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, $d(R,\mathcal{N}_\mathcal{K}(D)) \le \epsilon''$ and $d(R,\mathcal{N}_{\mathcal{K}_i}(D)) \le \epsilon''$ for all $i \in [d]$. By the Lipschitz continuity of the LP, y that solves LP(R, D) must have, for some $\check{\epsilon}(\epsilon'')$, that $y_j \ge D_j/2 - \check{\epsilon}$ for all $j \in \cup_{i\in[d]}\kappa^+_i$ (because, at centroid i, $y_j \ge D_j/2$ for $j \in \kappa^+_i$) as well as $y_j \ge D_j/2 - \check{\epsilon}$ for all $j \in \cup_{i\in[d]}\kappa^-_i$ (because $j \in \mathcal{K}_i$ for some i). Additionally, $y_j \le \check{\epsilon}$ 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 distinct 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.

Denote this basis by $\mathcal{B}_0$. 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 there are d separating vectors $\psi\in\Psi(\mathcal{K},\mathcal{B})$, 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, $\max_{\psi\in\Psi(\mathcal{K},\mathcal{B})}\{\psi'(R - r_{\mathcal{K}}(D))\} \leq M\check{\epsilon} - \zeta\min_j\{D_j\}$. 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

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

and

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

where we define $\tau_\partial^1=-\infty$ if $\tau_\partial^0<T$. On the event $\mathcal{V}:=\mathcal{V}[\tau_{\text{cone}}^{\epsilon',\mathcal{B}}]=1$, $\tau_0^\partial<T$ and, by Remark 5, $\mathbb{P}[\mathcal{V}=1,(\mathcal{C}_\ell)^c]\leq m_1e^{-m_2\ell}$.

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

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 $t\leq\tau_{\text{cone}}^{\epsilon',\mathcal{B}}$, $d_\infty(R^t,\mathcal{N}_{\mathcal{K}}(\mathcal{B},D))\leq\epsilon$ (see Lemma 8) and there exists M such that

$$y_j^t\leq M\epsilon,\ j\notin\mathcal{K}\cup K^+(\mathcal{B}),\ \text{and}\ y_j^t\geq p_j-M\epsilon,\ j\in\mathcal{K}\backslash K^-(\mathcal{B}). \tag{B.2}$$

If t is a time when $R^t\in\mathcal{N}_{\mathcal{K}^0}(p)$ ($\mathcal{K}^0=\mathcal{K}\cup\kappa^+\backslash\kappa^-$ is the second centroid visited by $\tau_{\text{cone}}^{\epsilon',\mathcal{B}}$), then $y_j^t\geq p_j/2$ for all $j\in\mathcal{K}\cup\kappa^+$ and $y_j^t\leq p_j/2$ for all $j\in\kappa^-$. Choosing ϵ' (subsequently ϵ) as small as needed, it must be that $\kappa^+\subseteq K^+(\mathcal{B})$ and $\kappa^-\subseteq K^-(\mathcal{B})$ as stated in item (i).

We turn to item (ii). By Lemma 11, $|y^{t+1}-y^t|_\infty\leq M/(T-t)$ so that the properties in (B.2) apply also to y^{t+1} with suitably modified constants. In particular, (B.2) holds up to and including $\tau_{cone}^{\epsilon',\mathcal{B}}$.

Take $\psi=\psi[\kappa]$. Using Lemma 6, we then have, for all $t\leq\tau_{\text{cone}}^{\epsilon',\mathcal{B}}$ and a redefined constant M, that

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

The first equality and first inequality follow from Lemma 6, which guarantees, for instance, that $\psi'A_{K^+\backslash\kappa^+}=0$. The last inequality follows by choosing ϵ' (subsequently 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 B.1 is used in the proof of Theorem 3.

Lemma B.1. *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 $\mathbb{P}[|X_t|_\infty\leq b]=1$ for some $b>0$) and independent of G_0, which satisfies $|G_0|\leq\epsilon/2$. Let*

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

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

$$\mathbb{P}[T-\tau>\ell, G_T\leq 0]\leq m_1e^{-m_2\ell},$$

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

Proof. Figure B.1 is a graphic illustration of the event whose probability we wish to bound. Notice that $\{\tau\leq t\}=\{\exists s\leq t:(T-s)\epsilon-G_s\leq 0\}$. In turn,

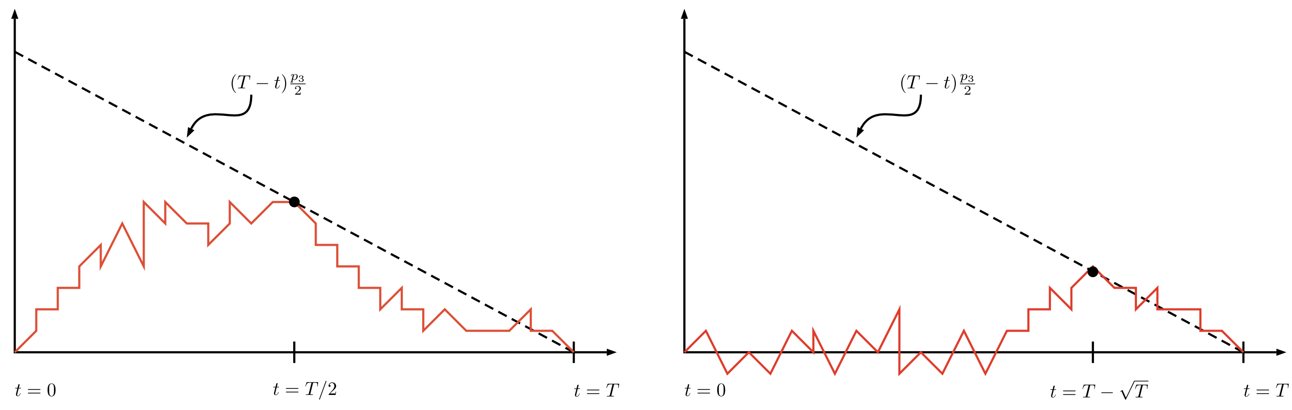
$$\{\tau\leq t, G_T\leq 0\}\subseteq\left\{\inf_{s\leq t}(-G_s+\epsilon(T-s))\leq -G_T, G_T\leq 0\right\}\subseteq\left\{\inf_{s\leq t}(G_T-G_s+\epsilon(T-s))\leq 0\right\},$$

so that

$$\begin{aligned}\mathbb{P}[\tau\leq t, G_T\leq 0]&\leq\mathbb{P}\left[\inf_{s\leq t}\{G_T-G_s+\epsilon(T-s)\}\leq 0\right]\\ &=\mathbb{P}\left[\inf_{u\geq T-t}\{G_u+\epsilon u\}\leq 0\right]\leq m_1e^{-m_2(T-t)},\end{aligned}$$

where the equality follows because the random vector $(G_T-G_t, G_T-G_{t-1},\ldots,G_T-G_0)$ has the same distribution as $(G_{T-t},G_{T-t+1},\ldots,G_T)$ (see, e.g., Gut [14, equation 11.5]), and the last inequality is a generalized version of Azuma's inequality (see, e.g., Ross [20, theorem 6.5.2]). Replacing $t\leftarrow T-\ell$ gives the result. □

Figure B.1. (Color online) Paths on $\mathcal{M} = \{R - e_{\mathcal{K}}\overline{Z}_{\mathcal{K}}^T \le 0\}$. 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 zero, which it must on the event $\mathcal{M}$.



Endnotes

[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.

[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. One verifies that $\mathcal{K} \subseteq \mathcal{J}$ by solving LP$(r_{\mathcal{K}}(D), D)$.

[3] This is the first of multiple places throughout our proofs in which we use the Lipschitz continuity of the LP in its right-hand side. Henceforth, we do so without citation.

[4] Because all bases associated with the initial centroid are computed, we know that $R^t \in \mathcal{N}_{\mathcal{K}}(D)$ as long as $(\mathcal{B}^{-1}R^t)_j \ge \overline{p}_j/2$ for $j \in \mathcal{K}$ and all bases $\mathcal{B}$ associated to $\mathcal{K}$.

References

[1] Agrawal S, Devanur N (2016) Linear contextual bandits with knapsacks. Lee DD, von Luxburg U, Garnett R, Sugiyama M, Guyon I, eds. *Proc. 30th Internat. Conf. Neural Inform. Processing Systems*, Advances in Neural Information Processing Systems, vol. 29 (Curran Associates Inc., Red Hook, NY), 3450–3458.
[2] Arlotto A, Gurvich I (2019) Uniformly bounded regret in the multi-secretary problem. *Stochastic Systems* 9(3):231–260.
[3] Badanidiyuru A, Kleinberg R, Slivkins A (2013) Bandits with knapsacks. *Proc. 2013 IEEE 54th Annual Sympos. Foundations Comput. Sci.* (IEEE Computer Society, Washington, DC), 207–216.
[4] Bast H, Delling D, Goldberg A, Müller-Hannemann M, Pajor T, Sanders P, Wagner D, Werneck RF (2016) Route planning in transportation networks. Kliemann L, Sanders P, eds. *Algorithm Engineering*, Lecture Notes in Computer Science, vol. 9220 (Springer, Cham, Switzerland), 19–80.
[5] Bemporad A, Borrelli F, Morari M (2002) Model predictive control based on linear programming—The explicit solution. *IEEE Trans. Automatic Control* 47(12):1974–1985.
[6] Bertsimas D, Tsitsiklis JN (1997) *Introduction to Linear Optimization* (Athena Scientific, Belmont, MA).
[7] Borrelli F, Bemporad A, Morari M (2003) Geometric algorithm for multiparametric linear programming. *J. Optim. Theory Appl.* 118(3):515–540.
[8] Boyd A, Bilegan I (2003) Revenue management and e-commerce. *Management Sci.* 49(10):1363–1386.
[9] Bubeck S, Cesa-Bianchi N (2012) Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *Foundations Trends Machine Learn.* 5(1):1–122.
[10] Chung KL (2000) *A Course in Probability Theory*, 3rd ed. (Academic Press, Inc., San Diego), xviii+419.
[11] Cook W, Gerards A, Schrijver A, Tardos É (1986) Sensitivity theorems in integer linear programming. *Math. Programming* 34(3):251–264.
[12] Eryilmaz A, Srikant R (2012) Asymptotically tight steady-state queue length bounds implied by drift conditions. *Queueing Systems* 72(3–4):311–359.
[13] Gal T (1984) Linear parametric programming—A brief survey. *Sensitivity, Stability and Parametric Analysis*, Mathematical Programming Studies (North-Holland Publishing Co., Amsterdam), 43–68.
[14] Gut A (2009) *Stopped Random Walks*, Springer Series in Operations Research and Financial Engineering, 2nd ed. (Springer, New York), xiv+263.
[15] Jasin S, Kumar S (2012) A re-solving heuristic with bounded revenue loss for network revenue management with customer choice. *Math. Oper. Res.* 37(2):313–345.
[16] Jasin S, Kumar S (2013) Analysis of deterministic LP-based booking limit and bid price controls for revenue management. *Oper. Res.* 61(6):1312–1320.
[17] Maguluri S, Srikant R (2016) Heavy traffic queue length behavior in a switch under the maxweight algorithm. *Stochastic Systems* 6(1):211–250.

[18] Mangasarian O, Shiau T (1987) Lipschitz continuity of solutions of linear inequalities, programs and complementarity problems. *SIAM J. Control Optim.* 25(3):583–595.
[19] Plambeck E, Ward A (2006) Optimal control of a high-volume assemble-to-order system. *Math. Oper. Res.* 31(3):453–477.
[20] Ross SM (1996) *Stochastic Processes*, Wiley Series in Probability and Statistics: Probability and Statistics, 2nd ed. (John Wiley & Sons, Inc., New York).
[21] Song J-S, Zipkin P (2003) Supply chain operations: Assemble-to-order systems. *Supply Chain Management: Design, Coordination and Operation*, Handbooks in Operations Research and Management Science, vol. 11 (Elsevier, Amsterdam), 561–596.
[22] Talluri K, Van Ryzin G (1998) An analysis of bid-price controls for network revenue management. *Management Sci.* 44(11):1577–1593.
[23] Talluri K, Van Ryzin G (2004) *The Theory and Practice of Revenue Management*, International Series in Operations Research & Management Science, vol. 68 (Kluwer Academic Publishers, Boston).
[24] Vera A, Banerjee S (2021) The Bayesian prophet: A low-regret framework for online decision making. *Management Sci.* 67(3):1368–1391.
[25] Vera A, Banerjee S, Gurvich I (2021) Online allocation and pricing: Constant regret via bellman inequalities. *Oper. Res.* 69(3):821–840.
[26] Wu H, Srikant R, Liu X, Jiang C (2015) Algorithms with logarithmic or sublinear regret for constrained contextual bandits. Cortes C, Lee DD, Sugiyama M, Garnett R, eds. *Proc. 28th Internat. Conf. Neural Inform. Processing Systems – Volume 1* (MIT Press, Cambridge, MA), 433–441.