Dynamic resource allocation: The geometry and robustness of constant regret

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

We study a family of dynamic resource allocation problems, wherein requests of different types arrive over time and are accepted or rejected. Each request type is characterized by its reward and the resources it consumes. We consider an algorithm, called BudgetRatio, that solves an intuitive packing linear program and accepts requests "scored" as sufficiently valuable by the LP. Our analysis method is geometric and focuses on the evolution of the remaining inventory—hence of the LP that drives BudgetRatio—as a stochastic process. The analysis requires a detailed characterization of the parametric structure of the packing LP, which is of independent interest.

We prove that (i) BudgetRatio achieves constant regret relative to the offline (full information) upper bound in the presence of (slow) inventory restock and request queues. (ii) By tuning the algorithm's parameters, we simultaneously achieve near-maximal reward and near-minimal holding cost. (iii) Within explicitly identifiable bounds, the algorithms regret is robust to mis-specification of the model parameters. This gives bounds for the "bandits" version where the parameters have to be learned. (iv) The algorithm has a natural interpretation as a generalized bid-price algorithm.

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

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

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

History: *This version:* October 11, 2020.

---

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

1. Introduction We study a family of dynamic resource allocation problems described as follows. Requests of multiple types arrive over a finite horizon of T discrete periods. If served, a request consumes a set of resources (that depend on the request's type) and generates a reward. There is an inventory of resources available at time 0 and additional units can arrive over time (restock). In general, a request can be either impatient — if not served immediately upon arrival it is lost — or patient, in which case the controller can choose to place it in a queue. The controller's objective is to use "correctly" its resource inventory to maximize the total reward collected over the finite horizon.

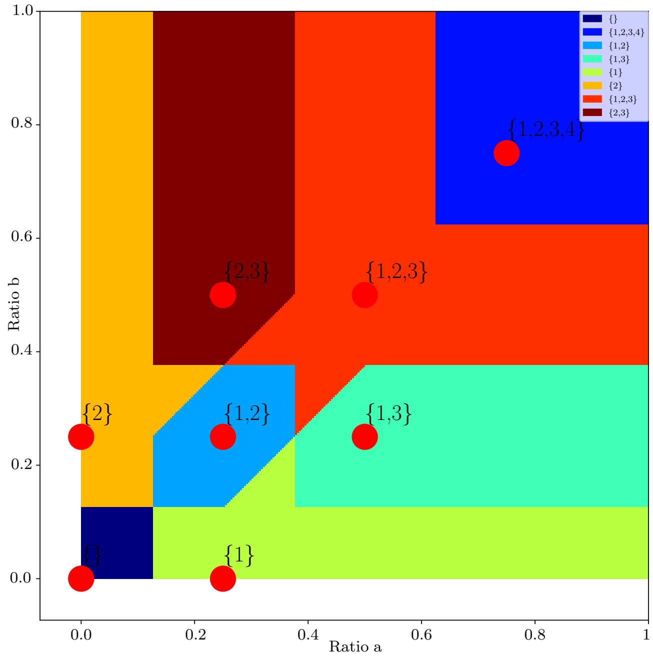
We refer to this large family of problems as *Resource Allocation Networks* or RAN s. If the controller could solve the problem in an *offline* fashion, she would wait till the end of the horizon and, given the realization of the random arrivals, choose the best allocation of resources. The reward of the offline (or hindsight) controller is an upper bound on any online algorithm. *In this paper, we obtain several performance guarantees and robustness results compared to the offline upper bound.*

The algorithm we study, which we refer to as BudgetRatio, is based on re-solving a *packing* linear program at each time. Specifically, let $(p_j)_{j\in\mathcal{J}}$ be the arrival probabilities of requests and $(p_i)_{i\in\mathcal{R}}$ the restock probabilities of resources. If at time t the inventory of different resources is $I^t\in\mathbb{N}^{\mathcal{R}}$, then we obtain $\bar{y}^t$ as the solution to an allocation LP where the capacity is constrained by $\frac{1}{T-t}I^t+p_{\mathcal{R}}$ and the demand by $p_{\mathcal{J}}$; see Fig. 1. Intuitively, the inventory plus expected restock is $I^t+(T-t)p_{\mathcal{R}}$ and we divide it into per-period expected available inventory. Similarly, we divide the expected arrivals of requests into per-period arrivals.

The number $\bar{y}^t_j$ is a proxy for the number of type-j requests that we want to accept: an inventory-dependent "score" of type j. BudgetRatio accepts requests with sufficiently large score, i.e., such that $\bar{y}^t_j\geq\eta_j$ for some thresholds η_j specified later. Observe that the scores $\bar{y}^t$ depend on the random *budget-ratio* process $R^t:=\frac{1}{T-t}I^t+p_{\mathcal{R}}$. This random process, evolving in the space of scaled resources, drives our analysis, see Figure 1.

Methodology: a geometric view of re-solving policies. The problems we consider cannot be solved optimally due to the so-called "curse of dimensionality". This fact alone justifies 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.

Our motivation in this paper is not purely algorithmic. We uncover the fundamental structure of the online stochastic packing problem and its optimal solution by taking a stochastic-process view of the dynamics. Our proof method captures the geometric nature of the problem and the dynamics of the budget consumption as a *stochastic-process* in the space $\mathbb{R}^{\mathcal{R}}_+$ of budget ratios. Our analysis makes explicit how BudgetRatio interacts with the geometry of the packing LP.



$$\begin{aligned} &\max\ v'y \\ &\text{s.t.}\quad Ay \le R^t := \tfrac{1}{T-t} I^t + p_{\mathcal{R}} \\ &\qquad\quad y \le p_{\mathcal{J}} \\ &\qquad\quad y \ge 0. \end{aligned}$$

FIGURE 1. Action regions for a problem with two resources and four request types. The plot is in the space of ratios that represent the per-period resource availability $R_i = \frac{1}{T-t} I^t + p_i$ for $i \in \{a, b\}$. Each point corresponds to a pair of budget states (R_a, R_b). When solving the LP, we obtain the set of request types $\mathcal{K} = \{j : \bar{y}_j \ge \eta_j\}$ that should be accepted at that inventory level. Each color on the plot corresponds to a different such set. For example, the light-blue region shaped like a rhombus corresponds to $\mathcal{K} = \{1, 2\}$. The specific parameters are $p_{\mathcal{J}} = (1/4, 1/4, 1/4, 1/4)$, $p_{\mathcal{R}} = 0$, $v = (4, 4, 5, 1)$, and $A = (1, 0, 1, 1; 0, 1, 1, 1)$.

The thresholding of the decision $\bar{y}^t$ divides the space of resource-budgets $\mathbb{R}_+^{\mathcal{R}}$ into mutually exclusive *action* regions. When the ratio is in a given region, all requests of a subset of types are accepted and all other are rejected, see Fig. 1.

To achieve bounded regret, the online policy must act in a way that is consistent with the optimal, unknown, offline basis. The policy must perform, what we call, *basic allocations* (see Definition 2). Thresholding—which underlies Figure 1— guarantees that, notwithstanding the unrevealed offline basis, the algorithm performs basic allocations. 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.

Since our geometric view exposes the structure of the problem, we can use it to study other popular algorithms. Indeed, we can explain the "failure" of the natural randomized re-solving algorithm and also give an appealing interpretation of BUDGETRATIO as bid-price control. In the context of online packing (a.k.a network revenue management), the standard bid-price algorithm solves the packing LP and accepts a request if its reward exceeds a bid-price that captures the *opportunity cost* of serving the request. The bid-price is the sum of the shadow prices of the requested resources. To achieve bounded regret, BUDGETRATIO is more careful: it is equivalently formulated as a bid-price control where the bid-price is obtained from a maximum over several shadow prices; see Section 8.

We study the robustness of BudgetRatio along multiple dimensions:

Modelling assumptions: the effect of queues and restock. In the baseline setting of online packing (or network revenue management), requests are impatient (leave if not immediately served), and inventory is not restocked, i.e., there is only an initial inventory that is gradually depleted. Our geometric view of the problem uncovers the effect of patient customers and restock on the very feasibility of achieving bounded regret—characterizing when such regret is or is not attainable. Where bounded regret is attainable, the geometric view guides the design of the algorithm.

We show that the offline benchmark is generally too ambitious. Indeed, in the absence of (what we call) a "slow restock" assumption, the offline upper bound is not generally achievable. Under the slow restock assumption, BudgetRatio achieves bounded regret. The thresholds, however, must be suitably perturbed by a quantity that captures the restock probability.

With patient requests, it seems reasonable that utilizing queued requests can increase the rewards. Instead, we prove that—with slow restock—constant regret for total reward can be achieved without interacting with the queues. In particular, this shows the robustness of the algorithm w.r.t. the abandonment model (e.g., infinitely patient, geometric patience, etc).

The performance metric. The algorithm's threshold can be tuned to achieve nested objectives: to minimize—in first order—the holding cost over the horizon subject to near optimal rewards. In other words, we consider the disutility of holding requests or inventory as well as the rewards. We show that, within the family of *all* policies that are nearly optimal for reward, BudgetRatio—with carefully tuned thresholds—achieves the minimal holding costs in first order.

This result hinges to a great extent on the understanding of the geometry of the problem and the dynamics of the remaining inventory as a function of the thresholds that are used.

Parameter misspecification. Our geometric analysis uncovers the sensitivity to errors in the forecasting of the demand and/or the rewards. Specifically, we study the case where the true parameters (values and probabilities) are (v, p) and we run the algorithm with $(\tilde{v}, \tilde{p})$ instead. We quantify how accurate $(\tilde{v}, \tilde{p})$ need to be so that BudgetRatio remains near optimal.

We introduce an appealingly simple notion of centroids (see §3). As long as these remain unchanged, the collection of action regions in Fig. 1 are stable under perturbations of the parameters. It is known that, in one dimension, $\tilde{v}$ needs to be accurate enough to deduce the ranking of the requests [21]. The centroids provide a generalization of the inherently one-dimensional notion of ranking, thus allowing us to understand the multidimensional problem. Our robustness guarantees easily yield the optimal regret guarantees in the learning setting (bandits with knapsacks).

2. Model and overview of results A decision maker must allocate resources to requests over a horizon of T periods. There is a set of resources $\mathcal{R}=\{1,\dots,d\}$ and, at time $t=0$, there is an initial inventory I_i^0 for each $i\in\mathcal{R}$. Additionally, at each time $t\in[T]$, a unit of resource i arrives with probability p_i; we call this the *restock* probability. We let $(Z_i^t : i\in\mathcal{R})$ be the accumulated restock over the time interval $[1,t]$. The controller cannot consume more than $I_i^0+Z_i^t$ units of resource i by time t.

There is a set $\mathcal{J}=\{1,\dots,n\}$ of possible requests, each request $j\in\mathcal{J}$ generates a reward v_j and requires some resource 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. We identify resources $\mathcal{R}$ and requests $\mathcal{J}$ with natural numbers for simplicity. At each time $t\in[T]$, a request j arrives with probability p_j independently of the past with $\sum_{j\in\mathcal{J}}p_j=1$; exactly one request arrives each period. We let $(Z_j^t : j\in\mathcal{J})$ be the accumulated arrivals over $[1,t]$.

Resources accumulate, i.e., if not used by time t, they are available at $t+1$. We allow for two kinds of requests: patient and impatient. If requests are impatient, they must be served when they arrive or not at all. In the classical network revenue managmenet (NRM) setting, all requests are impatient. In contrast, patient requests queue when not served at their arrival time. The controller knows in advance which types $j\in\mathcal{J}$ are patient and which are not. This modelling flexibility allows us to capture, in one framework, both online packing and assembly. Decisions are final: if a request j is served, the resources are consumed. Similarly, an impatient request that is rejected is lost forever.

We let V^t be the reward brought by the request arriving at time t. Thus, $V^1,V^2,\dots,V^T$ are 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 $I^t=(I_1^t,\dots,I_d^t)'$. Say the arrival is of type j, i.e., $V^t=v_j$, then the selection process unfolds as follows. If $I^t\not\geq A_j$ (the available inventory is insufficient), then the request must be queued. On the other hand, if the request is feasible ($A_j\leq I^t$), then it may be served, thereby generating a reward of v_j and decreasing the inventory to I^t-A_j, or it may be queued. Queued requests, corresponding to earlier patient arrivals, may be served at any period after their arrival.

No online policy can do better than the offline, full information, counterpart in which all values are presented in advance. Allowing this offline to use fractional allocations gives a further upper bound. This fractional offline controller is our benchmark. The expected total reward of the offline problem is given by the expectation of the following linear program, where $Z_{\mathcal{R}}$ are resource arrivals (restock) and $Z_{\mathcal{J}}$ are request arrivals:

$$V^*_{off}(T,I^0):=\mathbb{E}\begin{bmatrix}\max & v'y \\ \text{s.t.} & Ay\leq I^0+Z^T_{\mathcal{R}} \\ & y\leq Z^T_{\mathcal{J}} \\ & y\in\mathbb{R}^n_{\geq 0}\end{bmatrix}. \tag{1}$$

We make the following assumption throughout:

ASSUMPTION 1 **(Slow Restock)**. *For each request j and each resource i with $A_{ij} = 1$, we have $p_i < p_j$, i.e., the restock is bounded by the demand.*

Assumption 1 is imposed throughout. If it fails, the regret generally grows proportionally to $\sqrt{T}$. To show this, it suffices to consider a single type of impatient requests arriving at rate p, and a single resource also arriving at rate p. This corresponds to a *critically loaded* queue.

LEMMA 1. *There exists a RAN that violates Assumption 1 and such that, for some $c > 0$,*

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

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

The proof is in Appendix A. A slow restock assumption is thus necessary. When it fails, offline is no longer a useful benchmark for performance measurement.

2.1. Main results

THEOREM 1 **(constant regret)**. *Assuming slow restock,* BUDGETRATIO *(Algorithm 1) has a uniformly bounded regret: there exists a constant M such that*

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

where $V_{on}(T, I^0)$ is the total reward of BUDGETRATIO*. The constant M may depend on the request probabilities p, the reward v, and the resource consumption A, but it is independent of (T, I^0). Furthermore,* BUDGETRATIO *ignores the queued requests, hence it is robust to the modelling of customers' patience.*

Versions of this result have appeared before for the Online Packing problem (a.k.a. Network Revenue Management), i.e., under the assumptions of no restock and no queues (note that our model captures this with $p_{\mathcal{R}} = 0$) [2, 8, 20]. Our result is more general and the different analysis supports the robustness results presented in this paper; for a more detailed discussion of the literature see Section 2.3.

REMARK 1 (MODELLING PATIENCE). We assume that customers are either impatient or have infinite patience. Theorem 1 establishes, however, that this is immaterial: that the regret guarantee does not depend on the patience model. The latter can be taken to be geometric (departure with probability θ in each period), deterministic (departure after $1/\theta$ periods), or other. ■

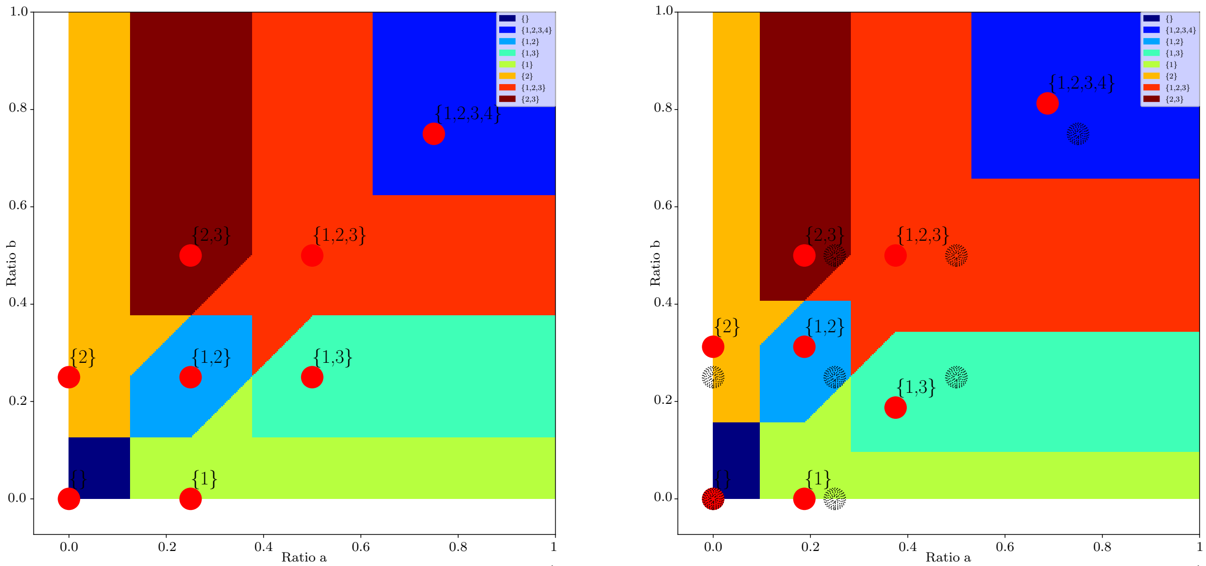


Figure 2. Action regions with true and misspecified probabilities (p and $\tilde{p}$). (LEFT) action regions of BudgetRatio when it is executed relative to p. (RIGHT) action regions when BudgetRatio uses $\tilde{p}_j = p_j - \frac{1}{16}$ for $j = 1, 3$ and $\tilde{p}_j = p_j + \frac{1}{16}$ for $j = 2, 4$. While the centroid sets remain unchanged, the shape of the action regions (the centroid neighborhoods) is slightly different. Crucially, under this particular mis-specification, the true centroid budgets (dashed circles), lie in the interior of the (mis-specified) actions regions. This guarantees that the BudgetRatio, although equipped with wrong probabilities, achieves bounded regret.

Robustness results. In Fig. 1, each action region corresponds to a set of types, or *centroid set*, accepted when the budget ratio is in this region. In this figure, for example, the light-blue region is the subset of $\mathbb{R}^2_+$ (of inventory levels) where BudgetRatio accepts requests of types 1 and 2. The red circle corresponds to the budget (the resource capacity) required to serve in expectation all the arrivals of the corresponding types and nothing else. The point $r_{\{1,2\}}(p) := (0.25, 0.25)$ is then the *centroid budget* for the centroid set $\{1,2\}$; see Section 4 for formal definitions.

As is clear in Figure 1, these red circles anchor the geometry of the action regions. It thus makes sense that robustness can be specified in terms of these points.

Our first robustness result quantifies how accurate our estimation of p must be so that the algorithm maintains constant regret. The key insight, illustrated in Figure 2, is that the true centroid budgets remain in the interior of the action regions of the algorithm with the perturbed "wrong" probabilities.

Proposition 1 **(demand misspecification)**. *Let p be the true underlying distribution and suppose that the centroid budgets are separated by at least $\delta > 0$ (see Definition 8). Then, for any distribution $\widetilde{p}$ that has $||p - \widetilde{p}||_\infty \le \frac{\delta}{4n}$,* BudgetRatio, *with the LP solved each period with $\widetilde{p}$ replacing p, produces bounded regret in the sense of Theorem 1.*

In Section 7 we will provide a refined condition on $\tilde{p}$ that is, in particular, implied by the one in the above proposition.

In the example of Fig. 1, $\delta \equiv \frac{1}{4}$ (by inspection) so that it is enough to have an estimation of the probabilities p with accuracy $\frac{\delta}{4n} = \frac{1}{64}$. We remark that this requires constant accuracy, in contrast to one that improves with T as is often the case in learning algorithms. Furthermore, the weaker condition we give in Section 7 allows us to remove a factor of n, so it suffices to estimate the centroid budgets with accuracy $\frac{\delta}{4} = \frac{1}{16}$.

The next result, concerning the robustness to misspecification of the reward vector v, hinges on a notion of strict complementarity (see Section 7.2 for details). In Figure 1, at each point (R_a, R_b) in the space, the optimal basis for the packing LP is different. In the following proposition, we require strict complementarity for all those optimal bases.

Proposition 2 **(value misspecification).** *Let v be the true values and let δ be such that all the bases are δ-complementary (Definition 9). There exists a constant $c \geq 1$ such that, for any $\tilde{v} \in \mathbb{R}^n_+$ satisfying*

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

BudgetRatio*, with the LP solved each period with $\tilde{v}$ instead of v, produces bounded regret in the sense of Theorem 1. Furthermore $c \leq \max\{||\mathcal{B}^{-1}||_\infty : \mathcal{B}$ basis $\}$.*

In the example of Fig. 1, we can take $\delta = c = 1$ (computed numerically), so that the robustness region is $||v - \tilde{v}||_\infty \leq \frac{1}{4}$. It is important that c, δ depend only on (v, A) and not on p or the horizon T.

Consider next the setting where the holding of inventory (and of customers, if they cannot be rejected) is costly. In that case, there is a tradeoff between maximizing rewards and consuming inventory (or emptying queues) as soon as possible.

We show that, with suitable tuning of the thresholds, among all policies with regret $o(T)$, BudgetRatio has the optimal holding cost scaling. Regardless of whether the holding cost if of requests or resources, we obtain the same result. Specifically, we consider two scenarios for holding cost: (i) we incur a cost $c_j > 0$ per queued request per time period, and (ii) we incur a cost h_i per stocked inventory unit per time period; see Section 6 for details and illustrative graphics. The holding cost of a policy π is

$$\mathcal{C}^\pi(T, I^0) := \mathbb{E}^\pi_{I^0}\left[\sum_{t=1}^T c \cdot Q^t\right] \quad \text{or} \quad \mathcal{C}^\pi(T, I^0) := \mathbb{E}^\pi_{I^0}\left[\sum_{t=1}^T h \cdot I^t\right].$$

PROPOSITION 3 **(dual objectives)**. *Suppose that the deterministic relaxation* $\mathrm{LP}(\mathbb{E}[R^0], D)$ *in Eq.* (2) *has a unique solution* $\bar{y}$ *and that* $\bar{y}_j < p_j$ *for at least one* $j \in \mathcal{J}$*. Then,* $\widehat{\pi} =$BUDGETRATIO *with threshold* $\alpha^T = T^{-1/4}$ *achieves simultaneously (1) constant regret for reward maximization and (2) asymptotic optimality for cost minimization, i.e.,*

$$\liminf_{T\uparrow\infty} \frac{\mathcal{C}^{\pi}(T, I^0)}{\mathcal{C}^{\widehat{\pi}}(T, I^0)} \geq 1 \qquad \textit{for any policy } \pi \textit{ with regret } o(T).$$

Finally, our geometric analysis elicits an interpretation of BUDGETRATIO as a bid-price heuristic: setting the bid price to a maximum of "local" shadow-prices yields constant regret under a complementarity assumption; see Section 8.

PROPOSITION 4 **(bid price heuristic)**. *If all the bases are* δ*-complementary (Definition 9) for some* $\delta > 0$*, then* BUDGETRATIO *is equivalent to a max-bid-price heuristic (Definition 10).*

The meaning of bounded regret. Theorem 1, and its various corollaries, state that the *additive (not multiplicative) gap* in collected reward between our algorithm (BUDGETRATIO) and the optimal algorithm (producecd by the solution of the dynamic program), is bounded by a constant that does not depend on length of the horizon T or on the initial inventory I^0. The proof establishes that BUDGETRATIO makes at most M (constant) mistakes relative to the optimal decision maker. It makes at most M mistakes over an horizon of $T = 10$ periods, and also at most M mistakes if the horizon is $T = 10^6$. Since

$$\text{Regret } = (\text{ Number of errors }) \cdot (\text{ Maximal cost of a single error}),$$

a finite number of mistakes translates into bounded regret. The (multiplicative) approximation factor is then $1 - O(\frac{1}{\min\{T, I_{\min}\}})$, where $I_{\min} = \min\{I_i^0 : i \in \mathcal{R}, p_i = 0\}$ is the minimum initial inventory that does not restock. In percentage terms, the error becomes negligible as the instance becomes "larger"; for example, if all resources restock, then approximation factor is $1 - O(\frac{1}{T})$.

An alternative notion for algorithm evaluation is that of competitive ratio: an α approximation algorithm is guaranteed to achieve an α fraction of the optimal reward. Bounded regret is not universally a stronger optimality notion than an α competitive ratio. In regimes where the cost of a single mistake is catastrophic, e.g., small T or small I^0, a competitive algorithm may be preferable.

For a variety of operational settings (network revenue management, inventory management etc.), regret is a suitable notion for algorithm evaluation.

2.2. Algorithm Our analysis is driven by the concept of future budget, which corresponds to inventory at hand plus *expected* future restock. Throughout we denote $p_{\mathcal{R}}$ as the vector of restock probabilities and $p_{\mathcal{J}}$ the vector of request arrival probabilities.

Definition 1 (Budget Ratio). If the inventory on hand at time t is I^t, then the ratio is

$$R^t := \frac{1}{T-t}(I^t + \mathbb{E}[Z^T_{\mathcal{R}} - Z^t_{\mathcal{R}}]) = \frac{1}{T-t} I^t + p_{\mathcal{R}}, \qquad t \in [1,T].$$

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

Our policy, presented in Algorithm 1, re-solves a deterministic relaxation and thresholds its solution accordingly. It is useful to define an LP as a function of the right-hand side:

$$\begin{aligned} \mathrm{LP}(R,D) \qquad & \max\ v'x \\ & \text{s.t.} \quad Ay \le R, \\ & \qquad\quad y \le D, \\ & \qquad\quad y \in \mathbb{R}^n_{\ge 0}. \end{aligned} \tag{2}$$

Algorithm 1 Budget Ratio Policy

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

1: Set thresholds: for $j \in \mathcal{J}$, let $\delta_j := \max_{i: A_{ij}=1} p_i$ and set $\bar{p}_j \leftarrow p_j + \delta_j \mathbb{1}_{\{\delta_j > \alpha p_j\}}$.
2: **for** $t = 1, \ldots, T$ **do**
3: If a request $j \in \mathcal{J}$ arrived, $Q^t_j \leftarrow Q^t_j + 1$. If a resource $i \in \mathcal{R}$ arrived, $I^t_i \leftarrow I^t_i + 1$.
4: Set $R^t \leftarrow \frac{1}{T-t} I^t + p_{\mathcal{R}}$.
5: Solve $\mathrm{LP}(R^t, p_{\mathcal{J}})$ to obtain the optimal decision variables $\bar{y}$.
6: **for** requests $j \in \mathcal{J}$ in decreasing order of $\bar{y}_j/\bar{p}_j$ **do**
7: **if** $Q^t_j = 0$ or $I^t \not\ge A_j$: not feasible to serve j
8: **else if** $\bar{y}_j \ge \alpha \bar{p}_j$: accept a request j. Update $Q^t_j \leftarrow Q^t_j - 1$, $I^t \leftarrow I^t - A_j$.
9: Deplete the queues of all impatient requests: $Q^t_j \leftarrow 0$ if $j \in \mathcal{J}$ is impatient.
10: Update for next period $I^{t+1} \leftarrow I^t$ and $Q^{t+1} \leftarrow Q^t$

Remark 2 (The Aggressiveness Parameter α). The parameter α is used to tune the algorithm. Note that the smaller the value of α, the more requests are accepted (see step 8 of Algorithm 1). The main result (Theorem 1) is obtained by setting $\alpha = 1/2$ and we set it to this value up to Section 6, where we will use the flexibility to tackle simultaneously reward maximization and holding-cost minimization. Furthermore, from our constructions it is (a posteriori) clear, that α can be set to any strictly positive constant, hence in practice it can be a useful tuning parameter to improve the performance. Moreover, α can also shrink with T (more details in Section 6). ∎

Remark 3 (the effect of restock: $\bar{p}$ vs. p). In the online packing (no restock), we have $\bar{p} = p$. In the presence of restock, its rate $p_{\mathcal{R}}$ must be taken into account, so that $\bar{p}$ is modified. Notice that we serve requests such that $\bar{y}_j \geq \alpha \bar{p}_j$ (step 8 of Algorithm 1), hence higher $\bar{p}_j$ means that we serve type j requests. Also notice that $\bar{p}_j$ is increased by δ_j (step 1), where δ_j captures the restock of resources used by j. The rationale is that the higher the restock, the more opportunities to serve j in the future because the resources used by j may replenish, hence it is not critical to serve j right now and the threshold increases accordingly. On the other hand, a request with low δ_j is relatively scarce, i.e., the inventory used by j may not replenish, so that it is better to serve it now, hence the threshold decreases. ∎

Final setup details. Let $\mathcal{F}_0$ denote the trivial σ-field. For $t \in [T]$, let $\mathcal{F}_t = \sigma\{Z^\tau : \tau = 1, \ldots, t\}$ 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 request type j is served at time t. Observe that, if $V^t = v_j$ and $\sigma_j^{\pi,t} = 0$, then the arriving request is queued. 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 over $[1,t]$. A policy is feasible if (1) the total consumption of resource i does not exceed its initial inventory I_i^0 plus the restock and (2) the total acceptance does not exceed arrivals, i.e.

$$\begin{aligned} AY^{\pi,t} &\leq I^0 + Z_{\mathcal{R}}^t, \ t \in [T], \\ Y^{\pi,t} &\leq Z_{\mathcal{J}}^t, \ t \in [T] \\ \sigma_j^{\pi,t} &\leq \mathbb{1}_{\{V^t = v_j\}}, \ t \in [T], j \in \mathcal{J}, j \text{ impatient.} \end{aligned} \tag{3}$$

Let Π be the set of feasible online policies, i.e., $\mathcal{F}_t$-adapted and satisfying Eq. (3). For $\pi \in \Pi$, the total reward of an online policy π is then given by

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

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

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

Solving for V^* directly is infeasible for most realistically-sized problems. To prove optimality guarantees, we therefore compare BudgetRatio against the offline benchmark previously defined.

Additional notation. Given a subset $\mathcal{K} \subseteq \mathcal{J}$ we let $A_{\mathcal{K}}$ be the submatrix of A that has only columns in the index set $\mathcal{K}$ (but has all rows). We similarly define sub-vectors: if x is a column vector, $x_{\mathcal{K}}$ is a subvector with the indices in the set $\mathcal{K}$. We use $\mathbf{e}$ for the vector of ones with dimension that will be clear from context. Finally, $\mathbf{e}_{\mathcal{K}}$ is a vector with 1 for $j \in \mathcal{K}$ and 0 otherwise.

For real numbers x, y, ϵ, we write $x = y \pm \epsilon$ if $|x - y| \leq \epsilon$. Following standard notation, for a subset $\mathcal{C} \subseteq \mathbb{R}^d$ and a point $x \in \mathbb{R}^d$ we let $d(x, \mathcal{C}) = \inf_{y \in \mathcal{C}} \|x - y\|$ be the distance of x from the set $\mathcal{C}$. We adopt the convention that the maximum over the empty set is zero, i.e., $\max \varnothing = 0$. We use throughout M to be a constant —that can depend only on (A, p, v), but it is independent of (T, I^0)— whose value can change from one line to the next.

2.3. Related Work

Online packing (network revenue management). Theorem 1 covers, as special cases, results obtained in [2, 20]; . [8, 13] develop, as well, algorithms with constant regret for the online packing problem in restricted settings (where the inventory and horizon scale linearly). We note that [2] also takes a geometric stochastic-process view, but it is specific to the one-dimensional (i.e., single resource) case. The geometric analysis of the multidimensional case requires the introduction of generalizable mathematical constructs (centroids, bases, cones, etc). We study here a larger family of models that includes queues and inventory arrivals.

More recently, [21] studies a large family of resource allocation problems that includes also dynamic posted pricing.

Relative to this earlier work, the geometric view has an explanatory power insofar as it provides an alternative and mathematically appealing support for bounded regret that is grounded in linear programming and, specifically, in a parametric view of the packing LP. This view provides the "language" through which we can explicitly identify the robustness of BudgetRatio to changes in (i) modelling assumptions (queues and inventory arrivals), (ii) parameters (arrival probabilities and rewards)—see further discussion of this immediately below—and (iii) algorithm parameters (the tuning of thresholds to meet multiple objectives)

Robustness to parameter misspecification and bandits. In Section 7 we give explicit conditions on the perturbation of parameters (probabilities p and values v) such that the regret guarantees hold. In simple terms, we identify subsets of $\mathbb{R}^n_+$, "centered" at the true probabilities and values, where BudgetRatio continues to produce bounded regret even if executed under "wrong" parameters.

This is related to, and has implications for, learning. For a clear relative positioning, let us consider the case where the reward of type-j requests is random with expectation v_j and an arrival is, as before, of type j with probability p_j. Suppose neither p nor v are initially known by the controller. The evolution of the arrival process, i.e., observing the frequency of different types, allows the controller to estimate p and the *served* requests allow the controller to estimate v; see [7] for more on bandit problems. Bandits problems known as *bandits with knapsacks* [3] explicitly model budget constraints which, in our setting, correspond to the limited inventory. In *contextual bandits* [1], arrivals present a "context" before the controller makes decisions. In our setting, the "context" is the request's type $j \in [n]$.

The general-purpose results in the literature [1] imply $O(\sqrt{T})$ regret bounds for our specific setting. Our analysis, targeted to resource allocation problems, produces a stronger $O(\log T)$ guarantee, which is the optimal regret scaling.

We obtain this guarantee by appealing to our notion of centroids. Informally speaking, to make good accept/reject decisions one must learn enough about the primitives to distinguish between the customer types. Centroids bring out a natural multi-dimensional notion of separation that is consistent with, yet goes significantly beyond, the condition identified for the one-dimensional (single resource) case in [21].

Two-sided arrivals and assembly. Arrivals of inventory capture assembly networks with fixed production rates. In assembly models, orders arrive (and wait in queue if patient) to be assembled by using relevant components; see [17] as well as [16] which gives an asymptotically optimal policy for holding cost minimization under a high demand assumption (related to our slow restock assumption). We focus on finite-time non-asymptotic guarantees for reward maximization.

Paremetric linear programming. The objective in this literature is to understand how optimization problems changes as the primitives change, see [12] for a survey. We study the parametric behavior of the packing, i.e., multiple parameters are perturbed simultaneously. This is a special case of multiparametric linear programming; see [4, 5] where the parametric analysis is used in support of model predictive control. Instead of developing algorithms for parametric linear programs, our analysis requires the characterization of the geometry of the problem. This is made feasible by the special structure of the packing LP.

Drift analysis. Much of our analysis centers on the dynamics of the process R^t. We argue that, when close to certain boundaries, the ratio process R^t drifts further towards the boundary. Such Lyapunov/drift methods are common in the analysis of stochastic models to establish positive recurrence of underlying Markov processes. In the context of online control, there are similarities to queueing theory where max-weight policies—based on re-solving local optimization problems—lead to the attraction to a subset of the state space; see [11, 14].

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

Offline representation. Introducing slack variables, we rewrite the offline LP Eq. (1), $\{Ay \leq I^0, y \leq Z^T_{\mathcal{J}}\}$ in standard form $\{Ay + s = I^0, y + u = Z^T_{\mathcal{J}}\}$, 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 $\bar{A}$, we have the following standard form LP

$$V^*_{off}(T, I^0) = \mathbb{E}\left[\max\left\{v'y : \bar{A}\begin{pmatrix} y \\ u \\ s\end{pmatrix} = C\right\}\right], \quad \text{where} \quad C := \begin{pmatrix} I^0 + Z^T_{\mathcal{R}} \\ Z^T_{\mathcal{J}}\end{pmatrix}. \tag{4}$$

The random vector $C \in \mathbb{R}^{d+n}_{\geq 0}$ is the *maximum consumption* of offline. Given a basis $\mathcal{B}$ (columns of $\bar{A}$) for the LP in Eq. (4), the optimal solution satisfies $\mathcal{B}x_{\mathcal{B}} = C$, where $x = (y, u, s)$ stands for all the variables. Hence, the realized (random) value of offline can be written as

$$\sum_{\mathcal{B}} v'_{\mathcal{B}} y_{\mathcal{B}} \mathbb{1}_{\{\mathcal{B} \text{ is optimal}\}} = \sum_{\mathcal{B}} v'_{\mathcal{B}} \mathcal{B}^{-1} C \mathbb{1}_{\{\mathcal{B} \text{ is optimal}\}}. \tag{5}$$

We abuse notation in the usual way, where $\mathcal{B}$ denotes both the indices of basic columns and the sub matrix $\bar{A}_{\mathcal{B}}$. Note that the optimality of a basis depends on the random vector C.

Online construction of the offline linear system. As expressed in Eq. (5), if the optimal offline basis is $\mathcal{B}$, then the offline actions correspond to the unique solution of the system $\mathcal{B}x_{\mathcal{B}} = C$, where $x = (y, u, s)$ stands for all the variables. An online policy π builds an approximate solution to the previous offline system. The quality of the approximation depends on how long the policy π "operates" in the basis $\mathcal{B}$. We make this precise in Proposition 5 below.

Definition 2 (Basic Allocation). Let π be any online policy and $\mathcal{B}$ be the optimal offline basis (which is only revealed at 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 is basic) and it only queues arriving requests such that $u_j \in \mathcal{B}$ (unmet variable is basic).

As long as the policy π is performing basic allocations, it is "operating" in the optimal basis, hence, if τ^π is the first time where π performed a non-basic allocation, regret is incurred only and at most in the remaining $T - \tau^\pi$ periods. We need the following additional notion to completely characterize the regret.

Definition 3 (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, i.e., $S^t = I^0 + Z^t_{\mathcal{R}} - AY^{\pi,t}$. 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, i.e., it is completely utilized in the offline solution. The wastage captures the inventory left unused by the online policy that (per the offline solution) should have been used in its entirety. The quality of the online system, i.e., the approximation to $\mathcal{B}x_{\mathcal{B}} = C$ is determined by this time τ^π and the wastage it induces.

PROPOSITION 5. *Let $\mathcal{B}$ be the optimal basis for the offline problem in Eq. (4) and denote $J^t \in \mathcal{J}$ the t-th arriving request. For any online policy π define the time*

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

Then the expected regret of π is at most $M\mathbb{E}[T - \tau^\pi + W^{\pi,\tau^\pi}]$, where M is a constant independent of (T, I^0), but that may depend on (A, v), and $W^{\pi,t}$ is the wastage.

PROOF. Throughout the proof, the policy π is fixed and omitted from the notation. Let Y_j^t, U_j^t be the number of type-j requests accepted and queueed/rejected by the policy over the interval $[1,t]$. Similarly, C^t denotes the maximal feasible consumption in $[1,t]$, i.e., $C^t := \begin{pmatrix} I^0 + Z_{\mathcal{R}}^t \\ Z_{\mathcal{J}}^t \end{pmatrix}$, and the surplus is defined as $S^t := I^0 + Z_{\mathcal{R}}^t - AY^t \in \mathbb{R}_{\ge 0}^d$. By the definition of $\bar{A}$ we then have that $\bar{A}X^t = C^t$, where $X^t = (Y^t, S^t, U^t)$. Let us divide the matrix $\bar{A}$ into basic and non-basic columns as $\bar{A} = [\mathcal{B}, \mathcal{B}^c]$. We claim that,

$$\mathcal{B}\begin{pmatrix} Y^t \\ U^t \\ S^t \end{pmatrix}_{\mathcal{B}} + \mathcal{B}^c \begin{pmatrix} 0 \\ 0 \\ S^t \end{pmatrix}_{\mathcal{B}^c} = C^t \quad \text{and} \quad C - C^t = Z^T - Z^t, \quad \forall t \le \tau^\pi. \tag{6}$$

The first equation follows from the decomposition $\bar{A} = [\mathcal{B}, \mathcal{B}^c]$ and the fact that, up to time τ^π, the only the non-zero variables Y^t, U^t are in the basis $\mathcal{B}$. The second equation follows from the definition of C and C^t. Recall that the offline variables $x = (y, u, s)$ are the solution to the offline system, i.e., $x_{\mathcal{B}} = \mathcal{B}^{-1}C$. From Eq. (6) we have

$$\begin{pmatrix} y \\ u \\ s \end{pmatrix}_{\mathcal{B}} - \begin{pmatrix} Y^t \\ U^t \\ S^t \end{pmatrix}_{\mathcal{B}} = \mathcal{B}^{-1}(Z^T - Z^t) + \mathcal{B}^{-1}\mathcal{B}^c \begin{pmatrix} 0 \\ 0 \\ S^t \end{pmatrix}_{\mathcal{B}^c} \qquad \forall t \le \tau^\pi. \tag{7}$$

Y is an increasing process so that $Y^T \ge Y^t$ for all t and, consequently,

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

We can bound the last expression using Eq. (7): since there is at most one arrival per period, $||Z^T - Z^t||_\infty \le T - t$, and the surplus is bounded by definition as $||S_{\mathcal{B}^c}^t||_\infty = W^{\pi,t}$. Finally, we take the worst case over $\mathcal{B}$ in Eq. (7) and conclude the result by setting $t = \tau^\pi$. $\square$

A consequence of of Proposition 5 is that, to prove Theorem 1, it suffices to prove that $\mathbb{E}[T - \tau^\pi + W^{\pi,\tau^\pi}] = \mathcal{O}(1)$. The remainder of our analysis is dedicated, then, to bounds on τ^π and the wastage W^{π,τ^π} for $\pi =$ BUDGETRATIO.

Remark 4 (On randomized policies and bid-price controls). Our definition of basic allocation (Definition 2) and its associated guarantee in Proposition 5, 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 [13] do not satisfy this consistency and this might explain their sub-optimal performance. The one-dimensional case (see Figure 3) is illustrative here. Suppose that types are ordered in decreasing order of rewards. If the initial ratio R^0 is slightly below $\bar{Z}_1^T + \bar{Z}_2^T$, offline will take all of type 1 and most of type 2 but none of type 3. However, since $\bar{Z}_1^T + \bar{Z}_2^T$ is a (small) perturbation of $p_1 + p_2$ we can have that R^0 is greater than $p_1 + p_2$. In this case, the randomized policy will (with some small probability) accept an arriving type 3 early in the horizon. In doing so, it will perform a *non*-basic allocation. In contrast, BudgetRatio does not take any such type 3, because it requires that $R^0 \geq p_1 + p_2 + p_3/2$ to do so. Informally, the thresholding adds a confidence interval.

The execution of non-basic allocations (too) early in the horizon persists also under the standard bid-price control. Here, one computes the (resource) shadow price vector λ and accepts a request if its reward v_j exceeds the sum of prices of requested resources, i.e., if $v_j \geq \sum_{i\in[d]} a_{ij}\lambda_i$. When $R^0 \in (p_1 + p_2, p_1 + p_2 + p_3)$ the shadow price of the (single) resource is v_3. Thus, we would accept type 3 request although this is not a basic allocation.

We re-visit the question of bid-prices in Section 8 where we show that BudgetRatio can be interpreted as a bid-price control ([18] and [19, Chapter 3.2]), albeit a more elaborate one. ■

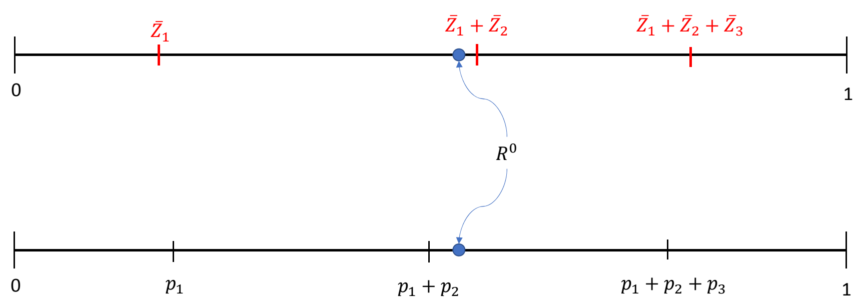


Figure 3. Why randomized policies do not maintain basic allocations. An illustration via the one-dimensional ($d = 1$) case. Although the offline accepts no type 3 requests (TOP), the online randomized algorithm accepts type 3 with some probability (BOTTOM).

Overview of the analysis via the one-dimensional case. Let us consider in some detail the one dimensional packing problem, also known as the multi-secretary problem [2], which can be stated as follows. There are I^0 positions to be filled and candidates arrive one at a time with abilities (values) $V^1, \ldots, V^T$; the goal is to select at most I^0 candidates to maximize the total value.

In our notation, $d = |\mathcal{R}| = 1$ (single resource), $p_{\mathcal{R}} = 0$ (no-restock so that $\mathbb{E}[R^0] = I^0$), and $A = \mathbf{e}'$ (each request consumes one item). All requests are impatient. The deterministic relaxation has $n+1$ constraints, one for each of the demand constraints and a single budget constraint:

$$\begin{array}{lll} \mathrm{LP}(R,p) & \max & v'y \\ & \text{s.t.} & \mathbf{e}'y \leq \mathbb{E}[R^0] \\ & & \quad y \leq p_{\mathcal{J}} \\ & & \quad y \geq 0. \end{array} \tag{8}$$

We assume without loss of generality that types are labelled in decreasing order of rewards, i.e., $v_1 > v_2 > \ldots > v_n$, and let $\bar{F}_i := \sum_{j=1}^{i} p_j$ be the survival function at v_i. The deterministic relaxation in Eq. (8) has a simple greedy solution: in increasing order of k, set $\bar{y}_k = p_k$ as long as $\bar{F}_k \leq \mathbb{E}[R^0]$. Letting $i_0 = \max\{k : \bar{F}_k \leq R\}$. Finally set $\bar{y}_{i_0+1} = \mathbb{E}[R^0] - \bar{F}_{i_0}$.

Centroids. A key construct we introduce here is that of a *centroid* which can be described as follows. If the budget ratio is exactly $\mathbb{E}[R^0] = \bar{F}_j$ (at a break point of the distribution), then the deterministic relaxation (8) takes all types $\mathcal{K} = [j]$ (and only those types). In other words, for this choice of right-hand side (i.e., budget), the problem $\mathrm{LP}(\bar{F}_j, p)$ has all variables $y_1, \ldots, y_j$ saturated and has all other variables equal to zero. The sets $\mathcal{K}$ with this property (and their later generalization to multiple dimensions, see Definition 4) are *centroids*. The set $\mathcal{K} = [j]$ is optimal when the budget is $r_{\mathcal{K}} = \bar{F}_j$, hence we refer to $r_{\mathcal{K}}$ as the *centroid's budget*; see Fig. 4 for an illustration.

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 $\mathrm{LP}(R^0, D^0)$ follow the same nested rule. This concept generalizes in multiple dimensions, i.e., there are sets of requests $\mathcal{K} \subseteq \mathcal{J}$ that are always prioritized in some part of the space independent of the demand p. In conclusion, centroids elicit a geometric view of the problem—a useful summary of the matrix A and the reward vector v—that does not depend on the demand.

Action sets: the centroid neighborhood. The thresholding of the algorithm—that we accept type j request only if $\bar{y}_j/p_j \geq \frac{1}{2}$—creates a confidence interval (a neighborhood) around the centroid's budget. The neighborhood of the centroid $\{1,2\}$ is the interval $[\bar{F}(v_3) - \frac{p_2}{2}, \bar{F}(v_3) + \frac{p_3}{2}]$ "centered" at exactly the centroid budget $\bar{F}(v_3) = p_1 + p_2$. As long as $R^t = I^t/(T-t)$ is in this interval, the algorithm accepts only (and all) arriving requests of types $\{1,2\}$. The algorithm starts accepting type-3 requests if R^t exceeds the right threshold $\bar{F}(v_3) + p_3/2$. It drops some type-2 requests if it goes below the left threshold $\bar{F}(v_3) - p_2/2$.

Oracle containment. Proposition 5 makes clear that a good online policy (with large time τ^π) should have "almost" oracle access to the offline basis $\mathcal{B}$; its decisions must be consistent with the

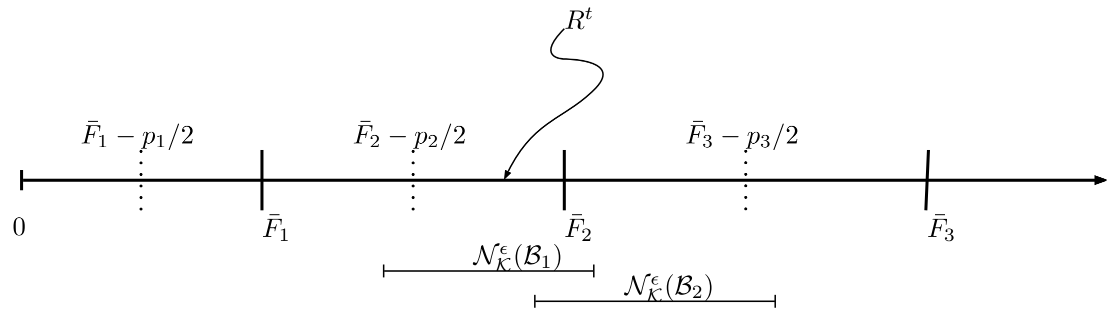


FIGURE 4. The position of the ratio R^t with respect to the centroid budgets $r_{\{1,..,j\}}(p) = \bar{F}_j$ determines the actions of the policy. At time t, the policy accepts a type j if and only if $R^t \geq \bar{F}_j - p_j/2$. The oracle containment property implies that if the realization Z^T is such that offline accepts only types $\{1,2\}$, then $R^t \in \mathcal{N}^\varepsilon(\mathcal{B}_1)$; conversely, if Z^T is such that offline does accept type-3, then $R^t \in \mathcal{N}^\varepsilon(\mathcal{B}_2)$. In conclusion, R^t evolves in the "correct" region $\mathcal{N}^\varepsilon(\mathcal{B}_1)$ or $\mathcal{N}^\varepsilon(\mathcal{B}_2)$, this guarantees that the policy accepts only requests in the optimal offline basis.

(apriori) unknown $\mathcal{B}$. We will formalize this intuition by showing that, if the budget-ratio process R^t is contained in an appropriate region, then we have this oracle access.

Let us suppose that $R^0 = \bar{F}(v_2)$, so that the deterministic relaxation has exactly the amount of budget to take all requests of types 1 and 2 and nothing else (see Fig. 4). There are two possible optimal bases at R^0. One, $\mathcal{B}_1$, has variables $\{y_1, y_2\}$ and the surplus variable s_2 (as a degenerate zero-valued variable) and the other, $\mathcal{B}_2$, has y_3 instead of s_2, but also as a degenerate zero-valued variable.

Consider Fig. 4. We will prove that, if offline takes only types $\{1,2\}$, i.e., if $I^0 - (Z_1^T + Z_2^T) \leq 0$ (offline selects basis $\mathcal{B}_1$), then $R^t \in \mathcal{N}^\varepsilon(\mathcal{B}_1)$ for all $t \in [1,\tau]$ where τ is a large stopping time, i.e., $\mathbb{E}[T - \tau] \leq M$. Observe that, as long as $R^t \in \mathcal{N}^\varepsilon(\mathcal{B}_1)$ (see Fig. 4), the policy only accepts requests $\{1,2\}$ which is consistent with $\mathcal{B}_1$, hence the policy performs basic allocations. On the other hand, if offline does accept type 3, i.e., $I^0 - (Z_1^T + Z_2^T) \geq 0$ (offline selects basis $\mathcal{B}_2$), then (we will prove) $R^t \in \mathcal{N}^\varepsilon(\mathcal{B}_2)$ for $t \in [1,\tau]$ which again implies that the policy performs basic allocation. Finally, because the process $R^t = \frac{1}{T-t} I^t$ remains contained in this bounded region, we have that $R^{\tau^\pi} \leq M$, so that $I^{\tau^\pi} \leq M(T - \tau^\pi)$, which proves a bounded wastage; Proposition 5 then yields bounded regret.

What lies ahead for the multidimensional analysis. The geometric structure, and the stochastic analysis that builds on it (convex sets, basic cones, etc.), is simple in the one dimensional case. Formalizing general notions of centroids and actions sets (in Section 4.2), requires a careful parametric analysis of the packing LP. We will show that the action sets are composed of convex subsets, each corresponding to an optimal basis (see Lemma 5). These convex sets are parametrized by the distribution and hence apply to the deterministic relaxation and to the offline (random) LP.

We will then lay the groundwork for the proof of the oracle containment property in multiple dimensions. The proof builds on (1) a sticky boundary property (Theorem 2), that shows that the residual budget process remains in one action region; and (2) a cone-containment property that stipulates that the BudgetRatio controlled budget process remains constrained to a suitable convex subset of the action region. In this convex subset BudgetRatio's actions are basic allocations, hence consistent with offline's (random) basis.

3.1. Towards the General Analysis: An Example For visualization purposes, a two dimensional example ($d = 2$ resources) that is both rich enough to demonstrate key characteristics and simple enough to afford a visual representation of the problem's geometry. We consider the traditional packing problem, namely one with impatient requests and no resource arrivals.

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 following consumption matrix

$$A \quad = \quad \begin{array}{c|cccc} & 1 & 2 & 3 & 4 \\ \hline a & 1 & 0 & 1 & 1 \\ b & 0 & 1 & 1 & 1 \end{array}$$

We use the reward vector $v = (4, 4, 5, 1)$ and the arrival probability vector $p = (1/4, 1/4, 1/4, 1/4)$. Type-3 requests bring the highest reward ($v_3 = 5$) but consume two resources unit, one of each a and b. Types 1 and 2 have the highest per-resource-consumption reward. Type-4 is clearly the least desirable.

Fig. 5 captures the action-regions of BudgetRatio. For future reference we label this as the *base example.*

Final setup details. Throughout we assume, without loss of generality, that $I_i \le T$ for all $i \in \mathcal{R}$. If $I_i > T$, this resource is non-binding and we can reduce the problem to one with $d - 1$ resources. Additionally, we randomly perturb the rewards as follows. For every $j \in \mathcal{J}$, we add an independent $U(0, \frac{1}{T})$ to v_j. This perturbation induces only a $\mathcal{O}(1)$ error while guaranteeing that all our objects (optimal solutions, optimal bases, etc) are uniquely defined.

4. Parametric structure of packing problems In Proposition 5 we characterize the performance of a policy by the horizon on which its allocations match the optimal offline basis. Since BudgetRatio's actions are dictated by an LP with a changing right hand side we must, to build on this idea, uncover the parametric structure of the packing LP.

The geometric characterization of *static* packing problems in this section is of independent interest. We use these results in Section 5 to analyse the dynamic behavior of BudgetRatio.

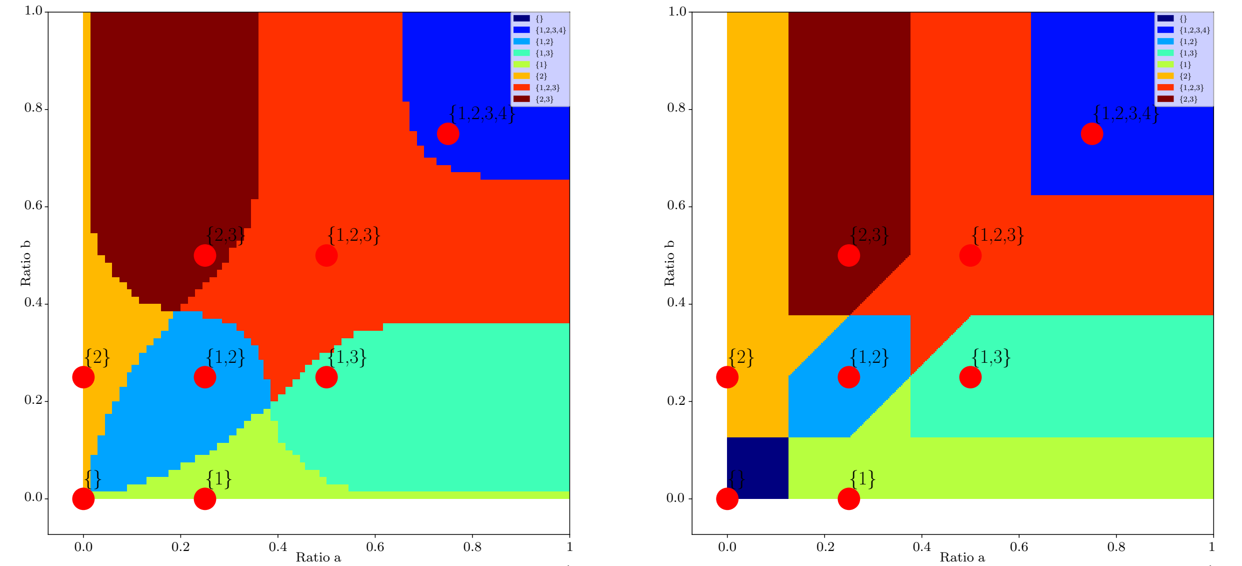


FIGURE 5. (LEFT) The action regions of the optimal policy (computed via DP) with 70 periods to go. When the vector of ratios (R_a, R_b) is in a color coded region, the policy accepts a fixed set of requests and rejects all others. For example, the light blue region in the bottom left corresponds to accepting $\{1,2\}$ and rejecting all other requests. (RIGHT) The action regions for BUDGETRATIO. The color coded regions have the same interpretation.

Let us write the packing LP in standard form. Our key descriptor is the budget ratio $R^t = \frac{1}{T-t} I^t + p_{\mathcal{R}}$. Let $\bar{A}$ be the augmented matrix

$$\bar{A} = \begin{bmatrix} A & 0 & I^d \\ I^n & I^n & 0 \end{bmatrix},$$

where I^n is the identity matrix of dimension $n \times n$. For any $R \in \mathbb{R}^n_{\geq 0}$ and $D \in \mathbb{R}^d_{\geq 0}$, we re-write the LP relaxation as

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

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 served, $u \in \mathbb{R}^n$ correspond to the amount of unmet 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 to denote the $(d+n) \times (d+n)$ sub-matrix of $\bar{A}$ corresponding to the variables in the basis $\mathcal{B}$; $\mathcal{B}^c$ denotes the non-basic columns. Let $\bar{v} = (v,0,0) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^d$ be the extended value vector, where we naturally assign zero value to variables u and s.

We are interested in the parametric structure of the LP (LP(R,D))—in how its solution changes with the right-hand sides (R,D). The first result focuses our attention on a subset of relevant bases; no other bases must be considered.

Lemma 2. *Let $\mathcal{B}$ be a basis and let $\lambda = (\mathcal{B}^{-1})'\bar{v}_{\mathcal{B}}$ be the dual variables associated to $\mathcal{B}$. Assume (i) $\lambda \geq 0$ and (ii) $\bar{A}'\lambda \geq \bar{v}$. Then, for any (R, D), $\mathcal{B}$ 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).*

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

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

For any basis $\mathcal{B}$, the associated dual variables are $\lambda = (\mathcal{B}^{-1})'\bar{v}_{\mathcal{B}}$, hence conditions (i) and (ii) imply that λ is dual-feasible. The associated primal variables are $(y, u, s)'_{\mathcal{B}} = \mathcal{B}^{-1}\begin{pmatrix} R \\ D \end{pmatrix}$. Finally, the primal and dual objectives coincide, so we conclude the optimality of $\mathcal{B}$ by weak duality provided that the solution is primal feasible, i.e., $(y, u, s)'_{\mathcal{B}} = \mathcal{B}^{-1}\begin{pmatrix} R \\ D \end{pmatrix} \geq 0$. Conversely, for any (R, D), if we run the Simplex Algorithm, we can find a basis $\mathcal{B}$ with an associated feasible dual solution. □

4.1. Action regions and exit times For ease of exposition, we strengthen Assumption 1 to require that $p_i < p_j/2$ (rather than $p_i < p_j$) all i, j such that $A_{ij} = 1$. At the end of this section we will comment how the analysis extends to $p_i < p_j$. With this strengthened assumption, $\bar{p}_j = p_j$ (see the first step of Algorithm 1) so that we threshold at at $p_j/2$.

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

$$\begin{aligned}\mathcal{N}_{\mathcal{K}}(D) &:= \left\{R \in \mathbb{R}^d : \text{ BudgetRatio serves exclusively requests } \mathcal{K} \text{ when } (R^t, p) = (R, D)\right\} \\ &= \bigcup_{\mathcal{B}} \left\{R \in \mathbb{R}^d : \mathcal{B} \text{ optimal}, \quad y_{\mathcal{K}} = \left(\mathcal{B}^{-1}\begin{pmatrix} R \\ D \end{pmatrix}\right)_{\mathcal{K}} \geq \frac{1}{2}D_{\mathcal{K}}, \quad y_{\mathcal{K}^c} = \left(\mathcal{B}^{-1}\begin{pmatrix} R \\ D \end{pmatrix}\right)_{\mathcal{K}^c} < \frac{1}{2}D_{\mathcal{K}^c}\right\}.\end{aligned} \tag{9}$$

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

Lemma 3. *For $\mathcal{K} \subseteq \mathcal{J}$,* BudgetRatio *serves exclusively requests in $\mathcal{K}$ if and only if $R^t \in \mathcal{N}_{\mathcal{K}}(p_{\mathcal{J}})$. Furthermore, for a constant M that depends on (A, p) only, whenever $t \leq T - M$ and $R^t \in \mathcal{N}_{\mathcal{K}}(p_{\mathcal{J}})$, there is enough inventory to serve a request from any $j \in \mathcal{K}$, i.e., $I_i^t \geq |\{j \in \mathcal{K} : A_{ij} = 1\}|$ for $i \in [d]$.*

Proof. The first part follows by definition of $\mathcal{N}_{\mathcal{K}}(p_{\mathcal{J}})$. For the second part we claim that, if $R^t \in \mathcal{N}_{\mathcal{K}}(p_{\mathcal{J}})$, then $I_i^t \geq |\{j \in \mathcal{K} : A_{ij} = 1\}|$, which proves that the inventory is enough 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$(R^t, p_{\mathcal{J}})$

implies $Ay \leq \frac{1}{T-t} I^t + p_{\mathcal{R}}$. Since $j \in \mathcal{K}$ it must be that $y_j \geq \frac{1}{2}\bar{p}_j$, hence $I_i^t \geq (T-t)(\frac{1}{2}\bar{p}_j - p_i)$. Since the resources used by j have slow restock (Assumption 1), $\frac{1}{2}\bar{p}_j > p_i$. Taking the constant $M = \frac{|\{j \in \mathcal{K} : A_{ij}=1\}|}{\bar{p}_j/2 - p_i}$ we obtain the claim for all $t \leq T - M$. □

We let $\mathcal{N}_{\mathcal{K}}(D,\mathcal{B}) \subseteq \mathbb{R}^d$ be the set of ratios where the algorithm serves exclusively $\mathcal{K}$ and the optimal basis for the relaxation is $\mathcal{B}$.

$$\begin{aligned}\mathcal{N}_{\mathcal{K}}(D,\mathcal{B}) &:= \left\{ R \in \mathbb{R}^d : \text{ BudgetRatio uses } \mathcal{B} \text{ and serves exclusively } \mathcal{K} \text{ when } (R^t, p) = (R, D) \right\} \\ &= \left\{ R \in \mathbb{R}^d : \mathcal{B} \text{ optimal}, \quad y_{\mathcal{K}} = \left(\mathcal{B}^{-1} \begin{pmatrix} R \\ D \end{pmatrix} \right)_{\mathcal{K}} \geq \frac{1}{2} D_{\mathcal{K}}, \quad y_{\mathcal{K}^c} = \left(\mathcal{B}^{-1} \begin{pmatrix} R \\ D \end{pmatrix} \right)_{\mathcal{K}^c} < \frac{1}{2} D_{\mathcal{K}^c} \right\},\end{aligned} \tag{10}$$

where the condition on $\mathcal{B}$ means that the basis $\mathcal{B}$ is optimal for the relaxation LP(R,D). By definition we have $\mathcal{N}_{\mathcal{K}}(D) = \cup_{\mathcal{B}} \mathcal{N}_{\mathcal{K}}(D,\mathcal{B})$. The next result states that (i) the sets $\mathcal{N}_{\mathcal{K}}(D,\mathcal{B})$ are the "correct resolution" to study the problem and (ii) perturbations of $\mathcal{N}_{\mathcal{K}}(D,\mathcal{B})$ completely characterize the exit time τ^{π} which, per Proposition 5, controls the regret.

Proposition 6. *Let $\mathcal{B}$ be the optimal offline basis and set $\mathcal{K} \subseteq \mathcal{J}$. Given $\epsilon > 0$, define*

$$\tau^{\epsilon,\mathcal{K}} := \min\{t \leq T : d(R^t, \mathcal{N}_{\mathcal{K}}(D,\mathcal{B})) > \epsilon\}.$$

Then, there exists a choice of $\epsilon > 0$ such that $\tau^{\epsilon,\mathcal{K}} \leq \tau^{\pi}$, where π is the policy given by Algorithm 1 and $\tau^{\pi} + 1$ is the first time that π does not perform basic allocation as in Proposition 5.

As long as R^t is close to $\mathcal{N}_{\mathcal{K}}(D,\mathcal{B})$, the algorithm is performing basic allocations. To prove this result, we elicit the structure of the action regions $\mathcal{N}_{\mathcal{K}}(D,\mathcal{B})$. Before turning to this task, we state a lower bound on $\tau^{\epsilon,\mathcal{K}}$ from which follows the proof of Theorem 1 by virtue of Proposition 5. Recall that $\mathbb{E}[R^0] = \frac{1}{T} I^0 + p_{\mathcal{R}}$ and $\mathbb{E}[D^0] = p_{\mathcal{J}}$, hence at time $t = 0$ we can identify the set $\mathcal{K}$ such that $\mathbb{E}[R^0] \in \mathcal{N}_{\mathcal{K}}(\mathbb{E}[D^0])$; indeed, it is obtained the first time we solve the deterministic relaxation.

Proposition 7. *Let $\mathcal{K}$ be such that $\mathbb{E}[R^0] \in \mathcal{N}_{\mathcal{K}}(\mathbb{E}[D^0])$ and let $\epsilon > 0$ and $\tau^{\epsilon,\mathcal{K}}$ be as in Proposition 6. 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 3).*

Proof of Theorem 1. By Proposition 6 we have that $\tau^{\epsilon,\mathcal{K}} \leq \tau^{\pi}$, hence the policy performs basic allocation over the interval $[1, \tau^{\epsilon,\mathcal{K}}]$. By Proposition 7, the expected wastage and remaining time $T - \tau^{\epsilon,\mathcal{K}}$ are bounded by a constant, hence we can apply Proposition 5 and conclude. □

It remains to prove Propositions 6 and 7. Both propositions hinge to a great extent on the action regions $\mathcal{N}_{\mathcal{K}}(D)$ and their subsets $\mathcal{N}_{\mathcal{K}}(D,\mathcal{B})$. The next two subsections provide the geometric characterization of these regions.

4.2. Centroids and neighborhoods In this subsection we formally define the packing *centroids*. In Section 4.3 we give a characterization of $\mathcal{N}_{\mathcal{K}}(D,\mathcal{B})$ based on centroids.

Definition 4 (Centroids). A subset $\mathcal{K}\subseteq\mathcal{J}$ is a centroid if, for some $D\in\mathbb{R}^n_{>0}$, there exists a solution (y,u,s) to $\mathrm{LP}(A_{\mathcal{K}}D,D)$ such that $u_{\mathcal{K}}=0$ (no request in $\mathcal{K}$ is unmet) and $y_{\mathcal{K}^c}=0$ (no request in $\mathcal{K}^c$ is served). If $\mathcal{K}$ is a centroid, we call $r_{\mathcal{K}}(D):=A_{\mathcal{K}}D_{\mathcal{K}}$ 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 $A_{\mathcal{K}}p_{\mathcal{K}}$—it is optimal in the deterministic relaxation to serve all request $\mathcal{K}$ and no others. In the one dimensional setting of Section 3, the centroids are the sets $[j]$ for $j=1,2,\ldots,n$ and their corresponding budgets are $r_{\mathcal{K}}(p)=r_{[j]}(p)=\bar{F}(v_j)$. Surprisingly, the characterization of centroids does not depend on D, but only on the matrix A and the values v. In particular, $\mathcal{K}$ is a centroid under both the theoretical distribution ($D=p$) and the empirical distribution ($D=\frac{1}{T}Z^T$).

Fig. 5 has (in red) the centroids for our base example and the location of the centroid budgets $r_{\mathcal{K}}(p)$ in $\mathbb{R}^2_+$ and their neighborhood.

The optimization problem Eq. (LP(R,D)) has multiple optimal bases at $R=r_{\mathcal{K}}(p)=A_{\mathcal{K}}p_{\mathcal{K}}$ and they are all degenerate: the solution (y,u,s) at $r_{\mathcal{K}}$ is, per Definition 4, $y_{\mathcal{K}}=p_{\mathcal{K}}$, $u_{\mathcal{K}^c}=p_{\mathcal{K}^c}$. All other variables are zero. Thus, only n of the basic variables are strictly positive, whereas the dimension of the right-hand side is $n+d$. There must be d zero-valued basic variables.

Definition 5 (Zero Valued Basic Variables). Fix a centroid $\mathcal{K}$ for some $\hat{p}$ as in Definition 4 and let $\mathcal{B}$ be a basis that is optimal at $r_{\mathcal{K}}(\hat{p})$, i.e., optimal for $\mathrm{LP}(r_{\mathcal{K}}(\hat{p}),\hat{p})$ 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 $K^+(\mathcal{B})$ or $K^+(\mathcal{B},\mathcal{K})$ to make explicit the dependence on the basis $\mathcal{B}$ or the centroid $\mathcal{K}$.

It turns out that the identity of centroids (Definition 4) and the zero-valued variables associated with them (Definition 5) is independent of the distribution $\hat{p}$.

Lemma 4. *Let $\mathcal{K}$ be a centroid for some $\hat{p}\in\mathbb{R}^n_{>0}$ as in Definition 4. Then, for any $\tilde{p}\in\mathbb{R}^n_{>0}$, the same property holds for $\tilde{p}$, i.e., $\mathrm{LP}(A_{\mathcal{K}}\tilde{p}_{\mathcal{K}},\tilde{p})$ has the solution $u_{\mathcal{K}}=0$ and $y_{\mathcal{K}^c}=0$. Similarly, the sets of zero-valued basic variables in Definition 5 are the same under $\hat{p}$ and $\tilde{p}$.*

Proof. Let $\mathcal{B}$ be the optimal basis of $\mathrm{LP}(A_{\mathcal{K}}\hat{p},\hat{p})$. We will prove that $\mathcal{B}$ is also optimal for $\mathrm{LP}(A_{\mathcal{K}}\tilde{p},\tilde{p})$ and has an associated solution $(y,u,s)=(\tilde{p}_{\mathcal{K}},\tilde{p}_{\mathcal{K}^c},0)$, which shows the result.

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

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

By Lemma 2 it follows that $\mathcal{B}$ is optimal for the right-hand side $(A_{\mathcal{K}}\tilde{p}_{\mathcal{K}},\tilde{p})$ and the associated solution is indeed $(y,u,s)=(\tilde{p}_{\mathcal{K}},\tilde{p}_{\mathcal{K}^c},0)$. Finally, it is clear from the structure of the solution (y,u,s) that the set of zero-valued basic variables is the same under $\hat{p}$ and $\tilde{p}$. □

Finally, we define a useful relation between centroids.

Definition 6 (Neighbors). Let $\mathcal{K}$ be a centroid. If the basis $\mathcal{B}$ is optimal at the centroid budget $r_{\mathcal{K}}$, 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 $\mathcal{K}'$, i.e., both centroids share an optimal basis.

Like the centroids themselves, the relation of "neighborhood" does not depend on the distribution $\hat{p}$. Once we fix $\mathcal{K}$ with associated basis $\mathcal{B}$, we can obtain neighbors of $\mathcal{K}$ based on the zero-valued basic variables (see Definition 5). In Section 4.3 we prove that $\mathcal{N}_{\mathcal{K}}(D,\mathcal{B})$, which determines the exit time of interest, can be characterized in terms of the focal centroid $\mathcal{K}$ and its neighbors.

4.3. Cones and characterization of action sets

The exit time in Proposition 6 depends on the distance of the budget ratio from the action set $\mathcal{N}_{\mathcal{K}}(D,\mathcal{B})$. In the study of this distance, the following representation of $\mathcal{N}_{\mathcal{K}}(D,\mathcal{B})$ is useful.

Lemma 5 **(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 5). Then,*

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

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

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

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

$$R=r_{\mathcal{K}}(D)+\sum_{\kappa^+\subseteq K^+,\kappa^-\subseteq K^-}\alpha_{(\kappa^+,\kappa^-)}(A_{\kappa^+}D_{\kappa^+}-A_{\kappa^-}D_{\kappa^-})+b, \tag{11}$$

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

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

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

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

We present the proof of the following result in the appendix. In Figure 6 (RIGHT) we plot three of the neighbors of the centroid $\mathcal{K}=\{1,2\}$. For the direction $(\kappa^+,\kappa^-)=(\{3\},\{2\})$ the neighboring centroid is $\{1,3\}$. In moving from $\mathcal{K}=\{1,2\}$ to $\mathcal{K}'=\{1,3\}$ the request variable y_2 and the slack u_3 leave the basis, while y_3 and u_2 enter the basis.

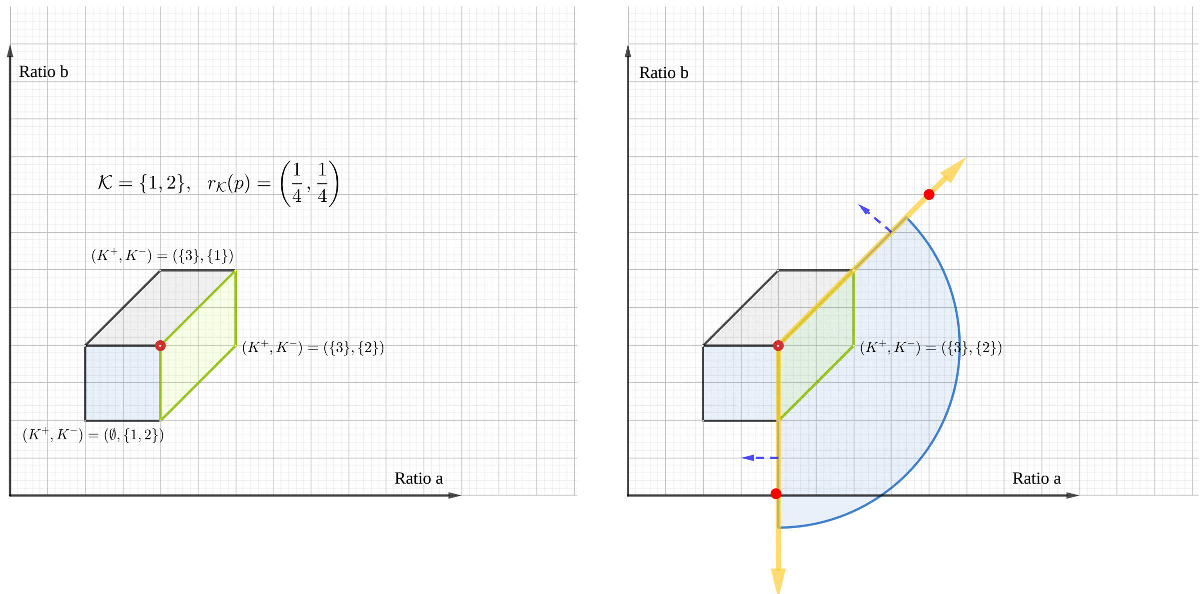


FIGURE 6. Geometric properties in the base example for the centroid $\{1,2\}$ whose budget is $r=(1/4,1/4)$: (LEFT) The extreme points and convex subsets, and (RIGHT) the cone, with orange boundaries, corresponding to $(K^+,K^-)=(\{3\},\{2\})$. The dashed vectors are the outer normals, $\psi_1=(-1,1)'$ and $\psi_2=(-1,0)'$.

In Fig. 6, we zoom-in on the centroid $\mathcal{K}=\{1,2\}$. One optimal basis at this centroid has $K^+=\{3\}$ and $K^-=\{2\}$. The neighboring centroids with $\kappa^+\in K^+$ and $\kappa^-\subseteq K^-$ are $\{1,3\}$, $\{1\}$, and $\{1,2,3\}$. The set $\mathcal{N}_{\mathcal{K}}(D,\mathcal{B})$ is the convex hull of the mid-points of the lines leading to those neighbors and corresponds to the yellow region.

The following definition is central to the analysis.

DEFINITION 7 (BASIC CONE). Let $\mathcal{K}$ be a centroid with associated basis $\mathcal{B}$ and (K^+,K^-) be the zero-valued basic variables (Definition 5). We define

$$\operatorname{cone}(\mathcal{K},\mathcal{B})=\{y\in\mathbb{R}^n: y=A_{K^+}x_{K^+}-A_{K^-}x_{K^-}, \text{ for some } x\geq 0\}.$$

From Lemma 5 (item 3) we know that $R\in\mathcal{N}_{\mathcal{K}}(D,\mathcal{B})$ if and only if $R-r_{\mathcal{K}}(D)=A_{K^+}x_{K^+}-A_{K^-}x_{K^-}+b$ with $x_j\in[0,D_j/2]$. The cone definition depends only on $(\mathcal{K},\mathcal{B})$ and not on D.

Instead of bounding $\tau^{\epsilon,\mathcal{K}}$ (see Proposition 6) directly, we consider the minimum of two times: the exit time from $\operatorname{cone}(\mathcal{K},\mathcal{B})$ and the exit time from $\mathcal{N}_{\mathcal{K}}(D)$; $\mathcal{N}_{\mathcal{K}}(D,\mathcal{B})$ is the intersection of these two sets.

LEMMA 6. *Let $\mathcal{K}$ be a centroid with basis $\mathcal{B}$. Then,* $\mathcal{N}_{\mathcal{K}}(D,\mathcal{B})=\mathcal{N}_{\mathcal{K}}(D)\cap(r_{\mathcal{K}}(D)+\operatorname{cone}(\mathcal{K},\mathcal{B}))$.

PROOF. 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 5 (item 3) we have $R-r_{\mathcal{K}}(D)=A_{K^+}x_{K^+}-A_{K^-}x_{K^-}+b$, hence $R\in r_{\mathcal{K}}(D)+\operatorname{cone}(\mathcal{K},\mathcal{B})$.

If $R^t \in \mathcal{N}_{\mathcal{K}}(D) \cap (r_{\mathcal{K}}(D) + \text{cone}(\mathcal{K},\mathcal{B}))$, then necessarily $R = r_{\mathcal{K}}(D) + A_{K^+}x_{K^+} - A_{K^-}x_{K^-}$ (since it is in the cone) and $x_j \le D_j/2$, since otherwise, by Lemma 5 (item 3), the associated solution would not have $y_{\mathcal{K}} > \frac{1}{2}D_{\mathcal{K}}$ and $y_{\mathcal{K}^c} < \frac{1}{2}D_{\mathcal{K}^c}$. □

The properties of the outwards normals to the cone are central to the proof of oracle containment (see Theorem 3). Fig. 6 (RIGHT) illustrates these vectors.

Lemma 7. *Fix a centroid $\mathcal{K}$ with associated basis $\mathcal{B}$. The vectors characterizing cone$(\mathcal{K},\mathcal{B})$ (i.e., such that $\max_l \psi_l' x \le 0$) have the following properties: For each $\kappa^+ \subseteq K^+$ and $\kappa^- \subseteq K^-$ with $|\kappa^+| + |\kappa^-| = 1$, $\psi[\kappa^+,\kappa^-]' A_{\kappa^+} = 0$ or $\psi[\kappa^+,\kappa^-]' A_{\kappa^-} = 0$. Also, $\psi[\kappa^+,\kappa^-]' A_j < 0$ for all $j \in K^+(\mathcal{B}), j \notin \kappa^+$ and $\psi[\kappa^+,\kappa^-]' A_j > 0$ for all $j \in K^-(\mathcal{B}), j \notin \kappa^-$.*

Additionally, for any other basis $\bar{\mathcal{B}} \ne \mathcal{B}$ associated to $\mathcal{K}$, $\psi[\kappa^+,\kappa^-]' A_j > 0$ for all $j \in (\kappa^+)^c \cup K^+(\bar{\mathcal{B}})$ and $\psi[\kappa^+,\kappa^-]' A_j < 0$ for any $j \in (\kappa^-)^c \cup K^-(\bar{\mathcal{B}})$.

When in $\mathcal{N}_{\mathcal{K}}(D,\mathcal{B}) \subseteq \mathcal{N}_{\mathcal{K}}(D)$, BudgetRatio accept only requests of types in $\mathcal{K}$. The next lemma further shows that BudgetRatio accepts only requests of types in set $\mathcal{K} \cup K^+$ (i.e., K^+ is added) when in a suitably small neighborhood of the subset $\mathcal{N}_{\mathcal{K}}(D,\mathcal{B})$. That is, it performs only basic allocations for the basis $\mathcal{B}$.

Henceforth, we fix

$$\epsilon := \frac{1}{4} \min\{p_k : k \in \mathcal{R} \cup \mathcal{J}, p_k > 0\}. \tag{12}$$

Lemma 8 **(optimal bases and BudgetRatio actions)**. *There exist constants M_1, M_2 such that, if $d(R^t, \mathcal{N}_{\mathcal{K}}(\mathcal{B},D)) \le \frac{\epsilon}{M_1}$, then the policy performs basic allocations at t: it serves only (but not necessarily all) requests in $\mathcal{K} \cup K^+$ and it queues only requests in $\mathcal{K}^c \cup K^-$. Moreover, $I_i^t \le M_2(T-t)$ for all $i \notin K^0(\mathcal{B})$.*

Proof. Let $(\bar{y},\bar{u},\bar{s})$ be the solution to LP(R^t,D) and define

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

By assumption $d_\infty(R^t, \mathcal{N}_{\mathcal{K}}(\mathcal{B},D)) \le \frac{\epsilon}{M_1}$. By the Lipschitz continuity of the LP solution [9, Theorem 5], we can choose M_1 large enough (depending on A) such that $d_\infty((\bar{y},\bar{u},\bar{s}),\mathcal{Y}) \le \epsilon$. Let $(y^0,u^0,s^0) \in \mathcal{Y}$ be such that $d_\infty((\bar{y},\bar{u},\bar{s}),(y^0,u^0,s^0)) \le \epsilon$. Since for all $j \in \mathcal{K} \backslash K^-$ we have that $y_j^0 = D_j$ then we also have that $\bar{y}_j \ge D_j - \epsilon \ge D_j/2$ so that all these items are taken. Also, for any $j \notin \mathcal{K} \cup K^+$, we have that $u_j^0 = D_j$ so that $\bar{u}_j^0 \ge D_j - \epsilon$ and hence $\bar{y}_j^0 < D_j/2$ so these requests are queued. Finally, $s_i^0 = 0$ for all $i \notin K^0$, hence $\bar{s}_i \le \epsilon$ for all such i, which implies $R_i^t \le Ay + \epsilon$ and using $y \le p$ we get the result. □

To relate online to offline, we need to characterize in the same geometric terms the optimality of a basis $\mathcal{B}$ for offline. Recall that the empirical demand distribution is $D^0 = \frac{1}{T}Z_{\mathcal{J}}^T$ and the empirical budget ratio is $R^0 = \frac{1}{T}(I^0 + Z^T_{\mathcal{R}})$. Considering the one-dimensional case can be instructive here. The optimal offline basis has type variables $1,2$ (and not 3) if $R^0 \in (\bar{Z}^T_1, \bar{Z}^T_1 + \bar{Z}^T_2]$. This is the intersection of the cone $r_{\{1\}}(\bar{Z}^T) + [0,\infty)$ (i.e., $R^0 - r_{\{1\}} \in [0,\infty)$) with the set $(0, \bar{Z}^T_1 + \bar{Z}^T_2]$. Define the following event

$$\mathcal{A}^\epsilon := \left\{\omega \in \Omega : \|\bar{Z}^T - p\|_\infty \le \epsilon\right\} \quad \text{where} \quad \epsilon := \frac{1}{4}\min\{p_1, \dots, p_{\mathcal{J}}\}.$$

Notice that we can write $R^0 = \frac{1}{T}I^0 + \bar{Z}^T_{\mathcal{R}} = \mathbb{E}[R^0] + \bar{Z}^T_{\mathcal{R}} - p_{\mathcal{R}}$, hence, if $\mathbb{E}[R^0] \in [p_1/2, p_1 + p_2/2]$, then, on the event $\mathcal{A}^\epsilon$, $R^0 \in [p_1/2, p_1 + p_2/2] \pm \epsilon$ so that $R^0 \in (0, \bar{Z}^T_1 + \bar{Z}^T_2]$. We generalize this reasoning in the following result. We present the proof in the appendix.

Lemma 9. *Let $\mathcal{K}$ be the centroid such that $\mathbb{E}[R^0] \in \mathcal{N}_{\mathcal{K}}(\mathbb{E}[D^0])$ and let*

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

Then, on the event $\mathcal{M}(\mathcal{B}) \cap \mathcal{A}^\epsilon$, $\mathcal{B}$ is an optimal offline basis. In particular, $\mathcal{B}$ has the following variables: $\{j : y_j \in \mathcal{B}\} \subseteq \mathcal{K} \cup K^+(\mathcal{B})$, $\{j : u_j \in \mathcal{B}\} \subseteq \mathcal{K}^c \cup K^-$, and $\{i : s_i \in \mathcal{B}\} \subseteq K^0$.

Recall that $\tau^{\epsilon,\mathcal{K}} = \min\{t \le T : d_\infty(R^t, \mathcal{N}_{\mathcal{K}}(\mathcal{B}, D)) > \epsilon\}$ is the first time the process gets too far from the action region. We are now in the position to prove Proposition 6.

Proof of Proposition 6. By Lemma 8, if $\mathcal{B}$ is the offline basis and $d(R^t, \mathcal{N}_{\mathcal{K}}(\mathcal{B}, D)) \le \epsilon$, the policy performs basic allocation at t. By Lemma 9, on this event the basis $\mathcal{B}$ is offline optimal. $\square$

To complete the proof of Theorem 1 it remains now to prove Proposition 7. That is the focus of the next section.

5. Analysis of BudgetRatio's dynamics Recall that we need to bound the time $\tau^{\epsilon,\mathcal{K}} = \min\{t \le T : d_\infty(R^t, \mathcal{N}_{\mathcal{K}}(\mathcal{B}, D)) > \epsilon\}$. We introduce the two auxiliary exit times:

$$\tau^{\epsilon,\mathcal{K}}_{\text{region}} := \inf\left\{t \le T : d(R^t, \mathcal{N}_{\mathcal{K}}(D)) > \epsilon\right\} \tag{13}$$

$$\tau^{\epsilon,\mathcal{B}}_{\text{cone}} := \inf\{t \le T : \max_l \psi_l'(R^t - r_{\mathcal{K}}(D)) > \epsilon\}, \tag{14}$$

where the vectors $\psi_l \equiv \psi_l[\mathcal{B}]$ are as in Lemma 7. By definition, the action region $\mathcal{N}_{\mathcal{K}}(\mathcal{B}, D)$ is the set where (i) the basis $\mathcal{B}$ is optimal and (ii) requests $\mathcal{K}$ are served exclusively (see Lemma 6). These auxiliary exit times relate to the aforementioned conditions (i) and (ii). Indeed, the time in Eq. (13) relates to condition (i) and the time in Eq. (14) relates to condition (ii) by virtue of Lemma 7. This is formalized in the following result, which we illustrate in Fig. 7 and formally prove in the appendix.

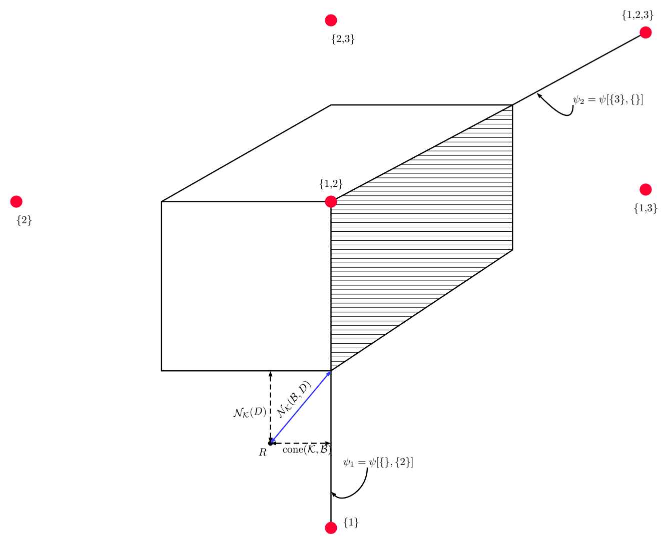


FIGURE 7. Representation of Lemma 10 for our base example. As before, centroids correspond to red circles and we focus on the action region of $\mathcal{K}=\{1,2\}$. The whole depicted area is $\mathcal{N}_\mathcal{K}(D)$, whereas the shaded region is $\mathcal{N}_\mathcal{K}(\mathcal{B},D)$ for a specific basis $\mathcal{B}$. Two rays define cone$(\mathcal{K},\mathcal{B})$, ψ_1 and ψ_2. At the bottom left we have some ratio R and the dashed arrows represent the distance from R to $\mathcal{N}_\mathcal{K}(D)$ and cone$(\mathcal{K},\mathcal{B})$, respectively. The solid blue arrow represents the distance from R to $\mathcal{N}_\mathcal{K}(\mathcal{B},D)$ that is bounded by the two previous distances in virtue of Lemma 10. Finally, notice that the two rays ψ_1 and ψ_2 lead to two different centroids (second part of Lemma 10); ψ_1 leads to $\mathcal{K}^0=\{1\}$ and ψ_2 to $\mathcal{K}^0=\{1,2,3\}$. In contrast, note that $\{1,3\}$ cannot be reached with $|\kappa^+|+|\kappa^-|=1$ as in the lemma because it is in the interior of the cone.

LEMMA 10 **(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 $\max_l \psi_l'(R-r_\mathcal{K}(D))\le\epsilon'$, then $d(R,\mathcal{N}_\mathcal{K}(\mathcal{B},D))\le\epsilon$. Consequently, $\tau^{\epsilon,\mathcal{K}}\ge\tau^{\epsilon',\mathcal{K}}_{region}\wedge\tau^{\epsilon',\mathcal{B}}_{\mathrm{cone}}$.*

Additionally, if for some κ^+,κ^- with $|\kappa^+|+|\kappa^-|=1$ we have $\psi[\kappa^+,\kappa^-]'(R-r_\mathcal{K}(D))\le\epsilon$ and $d(R,\mathcal{N}_\mathcal{K}(D))\le\epsilon$, then $R\in\mathcal{N}_{\mathcal{K}^0}(D)$ where $\mathcal{K}^0=\mathcal{K}\cup\kappa^+\backslash\kappa^-$.

THEOREM 2 **(sticky boundaries)**. *Let $\mathcal{K}$ be the centroid such that $\mathbb{E}[R^0]\in\mathcal{N}_\mathcal{K}(\mathbb{E}[D^0])$ and $\tau^{\epsilon,\mathcal{K}}_{region}$ as in Eq. (13). Then,*

$$\mathbb{P}[T-\tau^{\epsilon,\mathcal{K}}_{region}>\ell]\le\theta_1 e^{-\theta_2\ell},$$

where $\theta_1,\theta_2>0$ do not depend on (T,I^0) but could possibly depend on p,A,v.

Define the set of request variables consistent with the action region $\mathcal{N}_\mathcal{K}(D)$:

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

LEMMA 11. *Fix $\mathcal{K}$ and a neighbor $\mathcal{K}^0=\mathcal{K}\cup\kappa^+\backslash\kappa^-$. Let $R\in\mathcal{N}_{\mathcal{K}^0}(D)$ and (y,u,s) be the solution to* LP(R,D). *Let*

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

Then, the following holds:

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

Proof of Theorem 2. Let $S^t \in \mathbb{R}^d$ be the surplus of LP(R^t,D) at time t, i.e., the value of the surplus variable s. Observe that the request variable y is exactly the same for both problems LP(R^t,D) and LP$(R^t - S^t, D)$, hence, for any centroid $\mathcal{K}$, $R^t \in \mathcal{N}_{\mathcal{K}}(D)$ if and only if $R^t - S^t \in \mathcal{N}_{\mathcal{K}}(D)$. We can assume, then, that R^t has zero surplus. We will show that

$$\mathbb{P}\left[\sup_{t\in[1,T-\ell]} d(R^t,\mathcal{N}_{\mathcal{K}}(D)) > \epsilon\right] \le \theta_1 e^{-\theta_2 \ell},$$

as stated in the theorem.

To simplify notation we will write $\theta^t = \theta_{\mathcal{K}}(y^t,D)$, where $\theta_{\mathcal{K}}(y^t,D)$ is as in Lemma 11 and $\delta^t := y^t - \theta^t$. We define the following Lyapunov function

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

Whenever $g^t \le \epsilon^2/nd$, we also have $d(R^t,\mathcal{N}_{\mathcal{K}}(D)) \le \epsilon^2$. Indeed, if $g^t \le \epsilon^2/nd$ then, by Cauchy-Schwarz, $|Ay^t - A\theta^t|_i^2 = (a_i'(y^t-\theta^t))^2 \le \epsilon^2/d$. Finally, since $\theta^t \in \text{closure}(\mathcal{Y}(\mathcal{K},D))$ (Lemma 11), we have that $A\theta^t \in \mathcal{N}_{\mathcal{K}}(D)$ and $Ay^t = R^t$ (since R^t has zero surplus), hence $||Ay^t - A\theta^t||^2 \le \epsilon^2$ implies $d(R^t,\mathcal{N}_{\mathcal{K}}(D)) \le \epsilon^2$.

Setting $\varepsilon^2 := \epsilon^2/nd$, we conclude that $g^t \le \varepsilon^2$ implies $d(R^t,\mathcal{N}_{\mathcal{K}}(D)) \le \epsilon^2$. We argue the following drift condition: for some constant M,

$$\mathbb{E}[g^{t+1} - g^t|\mathcal{F}_t] \le -\frac{M}{T-t}, \text{ whenever } g^t \in [\varepsilon/2, \varepsilon]. \tag{15}$$

Assuming Eq. (15), concentration arguments as in [2, Theorem 3] show that $\mathbb{P}[\max_{t\in[1,T-\ell]} g^t > \varepsilon^2] \le \theta_1 e^{-\theta_2(T-\ell)}$ for some constants (θ_1,θ_2) that depend on M only, which proves the theorem.

The remainder of the proof is devoted to Eq. (15). If $R^t \in \mathcal{N}_{\mathcal{K}^0}(D)$ with $\mathcal{K}^0 = \mathcal{K} \cup \kappa^+ \setminus \kappa^-$, then using Lemma 11 (item 1) we obtain

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

where $\kappa = \kappa^+ \cup \kappa^-$. Our aim is to prove $\mathbb{E}[\|\delta^{t+1}-\delta^t\|^2|\mathcal{F}_t] = O(\frac{1}{(T-t)^2})$ and $\mathbb{E}[(\delta^{t+1}-\delta^t)_\kappa'(\delta^t)_\kappa|\mathcal{F}_t] \le \frac{-M}{T-t}$, which together would imply Eq. (15). We divide the proof in two parts: obtaining the linear bound $\mathbb{E}[(\delta^{t+1}-\delta^t)_\kappa'(\delta^t)_\kappa|\mathcal{F}_t] \le \frac{-M}{T-t}$ and the quadratic bound $\mathbb{E}[\|\delta^{t+1}-\delta^t\|^2|\mathcal{F}_t] = O(\frac{1}{(T-t)^2})$. We remark that the linear bound is the challenging part and the quadratic is purely algebraic. We start with a fact about the different neighborhoods that the process R^t visits during its evolution.

Main property of visited neighbors. For a vector $y \in \mathcal{Y}(\mathcal{K}, D)$, by definition, $y_j \geq D_j/2$ for $j \in \mathcal{K}$ and $y_j < D_j/2$ otherwise. Hence, if $g^t \leq \varepsilon^2$,

$$y_j^t \geq D_j/2 - \varepsilon \; \forall j \in \mathcal{K} \quad \text{and} \quad y_j^t \leq D_j/2 + \varepsilon \; \forall j \notin \mathcal{K}. \tag{16}$$

Let $\tilde{\mathcal{K}}$ be such that $R^{t+1} \in \mathcal{N}_{\tilde{\mathcal{K}}}(D)$, where $\tilde{\mathcal{K}} = \mathcal{K} \cup \tilde{\kappa}^+ \setminus \tilde{\kappa}^-$. We claim that

$$R_i^{t+1} - R_i^t = O\left(\frac{1}{T-t}\right), \; \forall i \in [d] \quad \text{and} \quad \kappa^+ \cup \kappa^- = \tilde{\kappa}^+ \cup \tilde{\kappa}^-. \tag{17}$$

The first fact in Eq. (17) is straightforward since we are taking ratios and since, as discussed before, we can assume zero surplus. Assume towards contradiction that the second fact in Eq. (17) fails. Take $j \in \kappa^+ \cup \kappa^-$, then Eq. (16) implies $y_j^t = \frac{1}{2} D_j \pm \varepsilon$. If $j \notin \tilde{\kappa}^+ \cup \tilde{\kappa}^-$, then $y_j^{t+1} \in \{0, D_j\}$. We conclude $|y_j^t - y_j^{t+1}| > M$. At the same time, solutions to the LP are Lipschitz continuous (see [9, Theorem 5]), so that $|y_j^t - y_j^{t+1}| \leq M ||R^t - R^{t+1}|| = O(\frac{1}{T-t})$, and we arrive at a contradiction. The case $j \notin \kappa^+ \cup \kappa^-$ and $j \in \tilde{\kappa}^+ \cup \tilde{\kappa}^-$ is completely analogous.

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

$$\begin{aligned} \mathbb{E}[\delta_j^{t+1} - \delta_j^t | \mathcal{F}_t] &\leq \frac{-M}{T-t} \quad j \in \kappa^+, \\ \mathbb{E}[\delta_j^{t+1} - \delta_j^t | \mathcal{F}_t] &\geq \frac{M}{T-t} \quad j \in \kappa^-. \end{aligned} \tag{18}$$

Assuming Eq. (18), if $g^t \geq \varepsilon^2/2$, then there exists $j \in \kappa^+$ s.t. $\delta_j \geq \varepsilon/\sqrt{2}$ or some $j \in \kappa^-$ s.t. $\delta_j \leq -\varepsilon/\sqrt{2}$, hence from Eq. (18) we can conclude our desired linear bound

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

Now let us prove Eq. (18). Let $\mathcal{K}^0$ be such that $R^t \in \mathcal{N}_{\mathcal{K}^0}(D^t)$. Recall that σ_j^t is the indicator that a request j is served at time t. Since only requests in $\mathcal{K}^0$ may be served, we have the identity $\mathbb{E}[I^{t+1}] = p_{\mathcal{R}} + I^t - \mathbb{E}[A_{\mathcal{K}^0} \sigma_{\mathcal{K}^0}^t]$, which implies

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

If $\mathcal{B}$ is the optimal basis for LP(R^t, D) and $\bar{\mathcal{B}}$ is optimal for LP(R^{t+1}, D^{t+1}), then we can write the previous equation in vector form (adding n components)

$$\mathbb{E}\left[\bar{\mathcal{B}} \begin{pmatrix} y^{t+1} \\ 0 \\ u^{t+1} \end{pmatrix} - \mathcal{B} \begin{pmatrix} y^t \\ 0 \\ u^t \end{pmatrix} \Big| \mathcal{F}_t \right] = \frac{1}{T-t-1} \mathcal{B} \begin{pmatrix} y^t - \mathbb{E}[\sigma_{\mathcal{K}^0}^t] \\ 0 \\ x \end{pmatrix},$$

where x is set so that the equation is satisfied in the n new components. Now we can multiply this equation by $\mathcal{B}^{-1}$ and use Lemma 11 (item 2) to conclude that, in components $j \in \kappa^+ \cup \kappa^-$ we have

$$\mathbb{E}[(y^{t+1} - y^t)_\kappa | \mathcal{F}_t] = \frac{1}{T-t-1} (y^t - \mathbb{E}[\sigma_{\mathcal{K}^0}^t])_\kappa. \tag{19}$$

For $j \in \kappa^+$ we use Lemma 3 to conclude that j is served whenever possible (the request does not use resources low on inventory), hence $\mathbb{E}[\sigma_j^t] = \mathbb{E}[\mathbb{1}_{\{J^t=j \text{ or } Q_j^t>0\}}] \geq D_j$. We know $\theta_j^t = \frac{1}{2}D_j$ and $\theta_j^{t+1} = \frac{1}{2}D_j$ (since $\tilde{\kappa} = \kappa$ according to Eq. (17)). From Eq. (19),

$$\mathbb{E}[\delta_j^{t+1} - \delta_j^t | \mathcal{F}_t] = \frac{1}{T-t-1}(y_j^t - \mathbb{E}[\sigma_j^t]) \leq \frac{1}{T-t-1}(-D_j/2 + \varepsilon),$$

where we used $y_j^t \leq D_j^t/2 + \varepsilon$ for $j \in \kappa^+$ (see Eq. (16)), and the fact $\mathbb{E}[\sigma_j^t] \geq D_j^t$. This gives Eq. (18) in the case for $j \in \kappa^+$.

On the other hand, for $j \in \kappa^-$, since $j \notin \mathcal{K}^0$, from Eq. (19) we have

$$\mathbb{E}[\delta_j^{t+1} - \delta_j^t | \mathcal{F}_t] = \frac{1}{T-t-1} y_j^t \geq \frac{1}{T-t-1}(D_j/2 - \varepsilon),$$

where we used $y_j^t \geq D_j/2 - \varepsilon$ (see Eq. (16)). This concludes the proof of Eq. (18).

Quadratic bound. Finally, we prove that $||\delta^{t+1} - \delta^t|| = O(\frac{1}{T-t})$. By Lemma 11 (item 1) we have $\delta^t = (y^t - \frac{1}{2}D)_\kappa$ and $\delta^{t+1} = (y^{t+1} - \frac{1}{2}D)_{\tilde{\kappa}}$. Using Eq. (17), we have $\delta^{t+1} = (y^{t+1} - \frac{1}{2}D)_\kappa$. It follows

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

The term $||(y^t - y^{t+1})_\kappa||$ is bounded by LP's Lipschitz continuity as in the proof of Eq. (17). □

REMARK 5 (STICKY BOUNDARIES). The proof of Theorem 2 implies that, once close to the boundary, the process R^t stays there. Formally, let

$$\tau_\partial^0 = \inf\{t \leq T : d(R^t, \partial \mathcal{N}_\mathcal{K}(D)) \leq \epsilon\}, \text{ and } \tau_\partial^1 = \inf\{T \leq t \geq \tau_\partial^0 : d(R^t, \partial \mathcal{N}_\mathcal{K}(D) \geq 2\epsilon\}.$$

Hence, $\mathbb{P}\{T - \tau_\partial^1 \geq \ell\} \leq \theta_1 e^{-\theta \ell}$. ▪

THEOREM 3. *Let $\mathcal{K}$ be the centroid such that $\mathbb{E}[R^0] \in \mathcal{N}_\mathcal{K}$ is such that $\max_l \psi_l'(\mathbb{E}[R^0] - r_\mathcal{K}) \leq \epsilon/2$ and let $\tau_{\text{cone}}^{\epsilon',\mathcal{B}}$ be as in Eq. (14). Then,*

$$\mathbb{P}[T - \tau_{\text{cone}}^{\epsilon',\mathcal{B}} > \ell, R^0 - A_\mathcal{K} \bar{Z}_\mathcal{K}^T \in \text{cone}(\mathcal{N}_\mathcal{K}(\mathcal{B}, D))] \leq \theta_1 e^{-\theta_2 \ell},$$

for some constants θ_1, θ_2 that do not depend on (T, I^0).

LEMMA 12. *Fix $\mathcal{K}$. Let $\mathcal{K}_0$ and $\mathcal{K}_1$ be two centroid neighbors of $\mathcal{K}$. There exists $\epsilon', \delta > 0$ such that: if $\exists R$ such that $d(R, \mathcal{N}_\mathcal{K}(D)), d(R, \mathcal{N}_{\mathcal{K}_0}(D)), d(R, \mathcal{N}_{\mathcal{K}_1}(D)) \leq \epsilon'$ then*

1. *$\mathcal{K}, \mathcal{K}_0$, and $\mathcal{K}_1$ share a basis $\mathcal{B}$ such that: $\mathcal{K}_0 = \mathcal{K} \cup \kappa_0^+ \backslash \kappa_0^-$ and $\mathcal{K}_1 = \mathcal{K} \cup \kappa_1^+ \backslash \kappa_1^-$ with $\kappa_0^+, \kappa_1^+ \subseteq K^+(\mathcal{B})$ and $\kappa_0^-, \kappa_1^- \subseteq K^-(\mathcal{B})$.*
2. *R is in the strict interior of $\text{cone}(\mathcal{K}, \mathcal{B})$, i.e., for the vectors $\{\psi_l\}$ characterizing $\text{cone}(\mathcal{K}, \mathcal{B})$ as in Lemma 7, we have $\max_\ell \psi_l'(R - r_\mathcal{K}(D)) \leq -\delta$.*

Proof. Let us write $\mathcal{K}_0 = \mathcal{K} \cup \kappa_0^+ \backslash \kappa_0^-$ and $\mathcal{K}_1 = \mathcal{K} \cup \kappa_1^+ \backslash \kappa_1^-$. Then, by the Lipschitz continuity of the LP, if y solves LP(R,D), it must be that $y_j \geq D_j/2 - \epsilon$ for all $j \in \kappa_0^+ \cup \kappa_1^0$ as well as $y_j \geq D_j/2 - \epsilon$ for all $j \in \kappa_0^- \cup \kappa_1^-$. Additionally, it must be that $y_j \leq \epsilon$ for all j outside these sets. Thus, R must be of the form

$$R = r_{\mathcal{K}}(D) + \frac{1}{2} A_{\kappa_0^+ \cup \kappa_1^+} D_{\kappa_0^+ \cup \kappa_1^+} - \frac{1}{2} A_{\kappa_0^- \cup \kappa_1^-} D_{\kappa_0^- \cup \kappa_1^-} \pm M\epsilon.$$

For the first item, assume towards contradiction that the bases are different. That is, $\mathcal{K}_0$ is generated (from $\mathcal{K}$) by a basis $\mathcal{B}_0$ and $\mathcal{K}_1$ is generated by a basis $\mathcal{B}_1 \neq \mathcal{B}_0$. Let $\{\psi_l^0\}$ be the vectors characterizing cone$(\mathcal{K}, \mathcal{B}_0)$. Then we must have that $\max_l (\psi_l^0)'(\frac{1}{2} A_{\kappa_0^+ \cup \kappa_1^+} D_{\kappa_0^+ \cup \kappa_1^+} - \frac{1}{2} A_{\kappa_0^- \cup \kappa_1^-} D_{\kappa_0^- \cup \kappa_1^-}) \leq M\epsilon$. On the other hand, we claim that, if there is $j \in \kappa_1^+$ or $j \in \kappa_1^-$ such that $j \notin K^+(\mathcal{B}_0) \cup K^-(\mathcal{B}_0)$, we would have $\max_l (\psi_l^0)'(\frac{1}{2} A_{\kappa_0^+ \cup \kappa_1^+} D_{\kappa_0^+ \cup \kappa_1^+} - \frac{1}{2} A_{\kappa_0^- \cup \kappa_1^-} D_{\kappa_0^- \cup \kappa_1^-}) \geq \zeta \min_j D_j$. The two previous inequalities contradict each other.

To see the claim, recall from Lemma 7 that the vectors defining cone$(\mathcal{K}, \mathcal{B})$, for some generic $\mathcal{B}$, correspond to neighboring centroids with $|\kappa^+| + |\kappa^-| = 1$ and are such that

$$\psi[\kappa^+, \kappa^-]' A_j < 0, \; j \in K^+(\mathcal{B}), j \neq \kappa^+ \quad \text{and} \quad \psi[\kappa^+, \kappa^-]' A_j > 0, \; j \in K^-(\mathcal{B}), j \neq \kappa^-. \tag{20}$$

Fix a vector $\psi[\kappa^+, \kappa^-]$ for cone$(\mathcal{K}, \mathcal{B}^0)$. From Lemma 7 we have $\psi[\kappa^+, \kappa^-]' A_j > 0$ for all $j \in (\kappa^+)^c \cup \kappa_1^+$ and $\psi[\kappa^+, \kappa^-]' A_j < 0$ for any $j \in (\kappa^-)^c \cup \kappa_1^-$. This last fact together with Eq. (20) imply $\psi[\kappa^+, \kappa^-]'(\frac{1}{2} A_{\kappa_0^+ \cup \kappa_1^+} D_{\kappa_0^+ \cup \kappa_1^+}) > \zeta \min_j D_j$ and $\psi[\kappa^+, \kappa^-]'(\frac{1}{2} A_{\kappa_0^- \cup \kappa_1^-} D_{\kappa_0^- \cup \kappa_1^-}) < -\zeta \min_j D_j$. The claim follows from here.

For the second item we can assume $\mathcal{B}_0 = \mathcal{B}_1 = \mathcal{B}$. Fix a vector $\psi[\kappa^+, \kappa^-]$ defining cone$(\mathcal{K}, \mathcal{B})$. Then it must be that either $\kappa^+ \neq \kappa_0^+ \cup \kappa_1^+$ or $\kappa^- \neq \kappa_0^- \cup \kappa_1^-$ because $\mathcal{K}_0 \neq \mathcal{K}_1$. Thus, from Eq. (20), $\psi[\kappa^+, \kappa^-]' A_{\kappa_0^+ \cup \kappa_1^+} D_{\kappa_0^+ \cup \kappa_1^+} \leq -\zeta \min_j D_j$. Similarly, $\psi[\kappa^+, \kappa^-]'(-A_{\kappa_0^- \cup \kappa_1^-} D_{\kappa_0^- \cup \kappa_1^-}) \leq -\zeta \min_j D_j$. In turn, we have $\psi[\kappa^+, \kappa^-]'(R - r_{\mathcal{K}}(D)) \leq M\epsilon - 2\zeta \min_j D_j$. This is negative for small ϵ. $\square$

Proof of Theorem 3. Throughout, $\mathcal{K}$ and the basis $\mathcal{B}$ are fixed. Also, we write $r_{\mathcal{K}} = r_{\mathcal{K}}(p)$. Recall that, whenever $\mathbb{E}[R^0] \in \mathcal{N}_{\mathcal{K}}(p)$, the following bound follows from Theorem 2

$$\mathbb{P}[T - \tau_{\text{region}}^{\epsilon, \mathcal{K}} > \ell] \leq \theta_1 e^{-\theta_2 \ell} \quad \text{where} \quad \tau_{\text{region}}^{\epsilon, \mathcal{K}} = \inf\{t \leq T : d(R^t, \mathcal{N}_{\mathcal{K}}(p)) \geq \epsilon\}. \tag{21}$$

For each ℓ, define the event $\Omega^\ell := \{T - \tau_{\text{region}}^{\epsilon, \mathcal{K}} \leq \ell\}$ and observe that, by Eq. (21), $\mathbb{P}[(\Omega^\ell)^c]$ is exponentially small. As the event of interest is $\mathcal{D} := \{T - \tau_{\text{cone}}^{\epsilon, \mathcal{B}} > \ell, \mathcal{M}(\mathcal{B})\}$, it suffices to bound $\mathbb{P}[\mathcal{D}, \Omega^\ell]$.

Outline of the proof. We are left to bound the measure of $\mathcal{D} \cap \Omega^\ell$. To do so, we consider two cases. First, we assume that at most one action region other than $\mathcal{N}_{\mathcal{K}}(D)$ was visited during the horizon; this corresponds to the case where the process starts close to the boundary of the cone

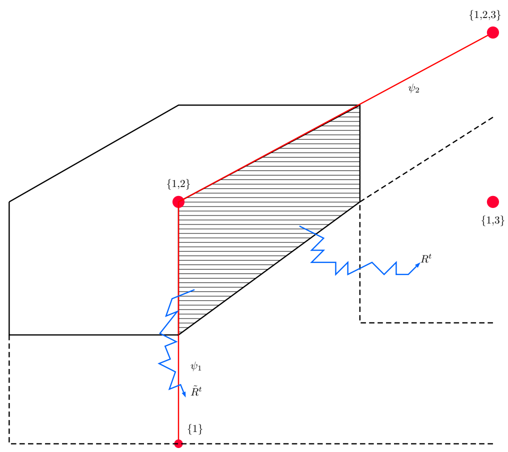


FIGURE 8. Two random walks for our base example. Solid lines enclose the region of interest $\mathcal{N}_{\{1,2\}}(D)$ and dashed lines enclose neighbouring action regions corresponding to $\{1\}$ and $\{1,3\}$. The random walk R^t visits three regions, namely $\{1,2\}$, $\{1\}$, and $\{1,3\}$, it is thus constrained to be in the interior of the cone, i.e., far away from the normals ψ_1, ψ_2. On the other hand, $\tilde{R}^t$ evolves close to the boundary of the cone, but in doing so visits only two regions, namely $\{1,2\}$ and $\{1\}$.

and it is therefore the challenging case. The second case, where more action regions were visited turns out to be easier. This is because, we show, this can only happen when the process moves in the strict interior of the cone where analysis is simpler as depicted in Fig. 8.

First case (boundary). Assume that over the interval $[1, \tau^{\epsilon,\mathcal{B}}_{\text{cone}}]$ only two action regions are visited, $\mathcal{N}_{\mathcal{K}}(p)$ and $\mathcal{N}_{\mathcal{K}^0}(p)$ for some neighbor $\mathcal{K}^0$. We bound the measure of the desired event as follows. We will introduce a process G^t with zero-mean increments and the following properties:

$$G^1 \le T\epsilon/2 \text{ and } G^T \le 0 \text{ a.s.} \quad \text{and} \quad \tau^{\epsilon,\mathcal{B}}_{\text{cone}} < T - \ell \Longrightarrow G^t > \epsilon(T-t) \text{ for some } t \in [1, T-\ell].$$

The event $\tau^{\epsilon,\mathcal{B}}_{\text{cone}} < T - \ell$, requires this process to grow faster than the linear target $\epsilon(T-t)$; an event that has an exponentially small probability. We make this precise after we explicitly define the process G^t.

Let l^0 be such that $\psi'_{l^0}(R^t - r_{\mathcal{K}}) > \epsilon$ at time $t = \tau^{\epsilon,\mathcal{B}}_{\text{cone}}$. In other words, the normal vector ψ_{l^0} is the one that makes the condition $\max_l \psi'_l(R^t - r_{\mathcal{K}}) \le \epsilon$ fail. The normal ψ_{l^0} is defined by some κ^+, κ^- such that $|\kappa^+ \cup \kappa^-| = 1$ (see Lemma 7). Since only two action regions are visited, we claim that the other action region corresponds to the neighbor $\mathcal{K}^0 = \mathcal{K} \cup \kappa^+ \setminus \kappa^-$. Indeed, the segment $L = \{\alpha r_{\mathcal{K}} + (1-\alpha) r_{\mathcal{K}^0} : \alpha \in [0,1]\}$ is completely contained in $\mathcal{N}_{\mathcal{K}}(p) \cap \mathcal{N}_{\mathcal{K}^0}(p)$ whenever $|\kappa^+| + |\kappa^-| = 1$, see Lemma 5. Furthermore, ϵ is small enough so that $L \pm \epsilon \subseteq \mathcal{N}_{\mathcal{K}}(p) \cap \mathcal{N}_{\mathcal{K}^0}(p)$, which proves the claim.

We have the inventory equation $I^s = I^0 + Z^s_{\mathcal{R}} - AY^s$. Since $(T-s)R^s = I^s + (T-s)p_{\mathcal{R}}$ and $(T-s)r_{\mathcal{K}} = (T-s)A_{\mathcal{K}} p_{\mathcal{K}}$,

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

$$(T-s)(R^s - r_{\mathcal{K}}) = T(\mathbb{E}[R^0] - r_{\mathcal{K}}) + \widehat{Z}^s_{\mathcal{R}} - AY^s + sA_{\mathcal{K}}p_{\mathcal{K}} \tag{22}$$

where we have defined the centered process $\widehat{Z}^t_{\mathcal{R}} := Z^t_{\mathcal{R}} - tp_{\mathcal{R}}$. Over the interval $[1, \tau^{\epsilon,\mathcal{B}}_{cone}]$ the only requests accepted correspond to $\mathcal{K} \cup \kappa^+$, hence $Y^s = Y^s_{\mathcal{K}} + Y^s_{\kappa^+}$. Additionally, all of the requests in $\mathcal{K} \setminus \kappa^-$ are accepted, hence $Y^s_{\mathcal{K}\setminus\kappa^-} = Z^s_{\mathcal{K}\setminus\kappa^-}$. Finally, by Lemma 7 we know that ψ_{l_0} is orthogonal to the columns of A corresponding to κ^+ and κ^-, hence we have the following identities

$$\psi'_{l_0} AY^s = \psi'_{l_0} A_{\mathcal{K}\setminus\kappa^-} Z^s_{\mathcal{K}\setminus\kappa^-} \quad \text{and} \quad \psi'_{l_0} A_{\mathcal{K}} p_{\mathcal{K}} = \psi'_{l_0} A_{\mathcal{K}\setminus\kappa^-} p_{\mathcal{K}\setminus\kappa^-}.$$

Defining the centred process $\widehat{Z}^t_{\mathcal{J}} := Z^t_{\mathcal{J}} - tp_{\mathcal{J}}$ and using the previous identities together with Eq. (22) we have

$$(T-s)\psi'_{l^0}(R^s - r_{\mathcal{K}}) = T\psi'_{l^0}(\mathbb{E}[R^0] - r_{\mathcal{K}}) + \psi'_{l^0}(\widehat{Z}^s_{\mathcal{R}} - A_{\mathcal{K}\setminus\kappa^-}\widehat{Z}^s_{\mathcal{K}\setminus\kappa^-}) \quad \forall s \le \tau^{\epsilon,\mathcal{B}}_{cone}.$$

From the previous equation, we can define the following process

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

We have proved the following:

$$\tau^{\epsilon,\mathcal{B}}_{\text{cone}} < T - \ell \iff G^t > (T-t)\epsilon \quad \text{for some } t \in [0, T-\ell]. \tag{23}$$

The process G^t has zero-mean increments. Furthermore, $G^0 = T\psi'_{l^0}(R^{\tau^\epsilon} - r_{\mathcal{K}}) \le T\epsilon/2$ by assumption. The remainder of the proof demonstrates $G^T \le 0$. Using that ψ_{l^0} is orthogonal to the columns A_{κ^-},

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

In the event $\mathcal{M}(\mathcal{B})$ we have $I^0 + Z^T_{\mathcal{R}} - A_{\mathcal{K}} Z^T_{\mathcal{K}} \in \text{cone}(\mathcal{K}, \mathcal{B})$, hence we conclude $G^T \le 0$.

From Eq. (23), we deduce the inequality

$$\mathbb{P}[\mathcal{D}, \Omega^\ell] = \mathbb{P}\big[T - \tau^{\epsilon,\mathcal{B}}_{\text{cone}} > \ell, \mathcal{M}(\mathcal{B}), \Omega^\ell\big] \le \mathbb{P}\left[\bigcup_{t\in[T-\ell]} \big\{G^t \ge \epsilon(T-t)\big\}, \max_l S^T_l \le 0\right] \le \theta_1 e^{-\theta_2 \ell},$$

for some $\theta_1, \theta_2 > 0$. The final bound follows from the analysis of a random walk crossing a positive moving threshold conditional on being negative at the end of the horizon. This is formally proved in Lemma 14 in the appendix.

Second case (strict interior). Assume that over the interval $[1,\tau_{\text{cone}}^{\epsilon,\mathcal{B}}]$ the process R^t visits three or more action regions. This case corresponds to where the process is in the strict interior of the cone. Indeed, by Theorem 2, the process R^t remains close to each action region visited. Formally, $d(R^t,\mathcal{N}_{\mathcal{K}^k}(p)) \leq \epsilon$ for each $\mathcal{K}^k$ visited over the interval $[1,T-\ell]$ with probability $\theta_1 e^{-\theta_2(T-\ell)}$. By Lemma 12, it must be that R^t is far from the boundary of the cone over the interval $[1,T-\ell]$, which completes the proof. □

Proof of Proposition 7. By Lemma 10, we have that the policy performs basic allocation up to time $\tau_{\text{region}}^{\epsilon',\mathcal{K}} \wedge \tau_{\text{cone}}^{\epsilon',\mathcal{B}}$. From Lemma 8, the wastage is bounded by $M(T-\tau_{\text{region}}^{\epsilon',\mathcal{K}} \wedge \tau_{\text{cone}}^{\epsilon',\mathcal{B}})$. Finally, from Theorems 2 and 3 we obtain that $\mathbb{E}[T-\tau_{\text{region}}^{\epsilon',\mathcal{K}} \wedge \tau_{\text{cone}}^{\epsilon',\mathcal{B}}] \leq M$. □

6. A Tale of two objectives We focused, thus, on reward maximization. BudgetRatio is a family of algorithms, instances of which are defined by the thresholds that map LP solutions to actions. Those thresholds are not. There is a range of thresholds that achieve bounded regret and this flexibility is useful; we can tailor our choices to achieve secondary objectives.

Theorem 1 proves that BudgetRatio is near-optimal for reward maximization. In this section we will show that BudgetRatio is simultaneously holding-cost minimizing; that with suitably tuned thresholds, BudgetRatio maintains bounded regret while achieving—in a competitive ratio sense—the least holding cost among *all policies* π that have regret $o(T)$. Equivalently, in order to obtain a lower holding cost than BudgetRatio, an algorithm would need to incur a significant $\Omega(T)$ regret for reward maximality.

We will focus on queue (request/customer) holding cost in most of our exposition. In Lemma 13 we show that inventory holding cost can be transformed into queue holding costs.

Let $c \in \mathbb{R}^n_{>0}$ be a vector of holding-cost coefficients: each period that a type-j request spends in queue, we incur a cost c_j. The horizon holding cost of a policy is then $\mathcal{C}_c^\pi(T,I^0) := \mathbb{E}_{I^0}^\pi\left[\sum_{t=1}^T c\cdot Q^t\right]$.

In this section we assume that there is no restock and adopt the following standard asymptotic framework [13]: given a fixed inventory $\bar{I}^0$, we consider a sequence of instances indexed by the horizon T such that the T-th instance has initial inventory $I_T^0 = T\bar{I}^0$.

To prove the main result we use the flexibility of BudgetRatio. Specifically, in step 8 of Algorithm 1, a request is accepted whenever $y_j \geq \alpha p_j$, where we treat $\alpha = \alpha^T$ as a tuning parameter.

Proposition 3. *Suppose that the deterministic relaxation* $\mathrm{LP}(\mathbb{E}[R^0],D)$ *has a unique solution* $\bar{y}$ *and that* $\bar{y}_j < p_j$ *for at least one* $j\in\mathcal{J}$. *Then, Algorithm 1 with tuning parameter* $\alpha^T = T^{-1/4}$ *achieves simultaneously (1) constant regret for reward maximization and (2) asymptotic optimality for cost minimization, i.e.,*

$$\liminf_{T\uparrow\infty} \frac{\mathbb{E}^\pi[\sum_{t=1}^T c\cdot Q^t]}{\mathbb{E}^{\hat{\pi}}[\sum_{t=1}^T c\cdot Q^t]} \geq 1 \qquad \textit{for any policy } \pi \textit{ with regret } o(T).$$

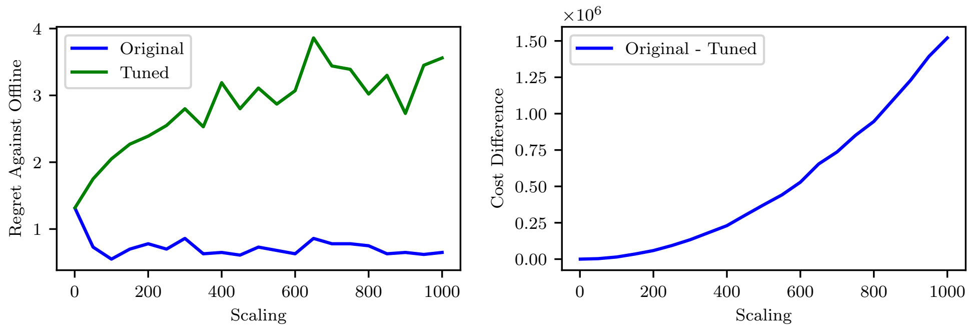


FIGURE 9. (LEFT) Regret, in terms of reward, of BUDGETRATIO with different thresholds. The tuned version has only slightly higher reward regret (still a constant) due to its greater aggressiveness in accepting requests. (RIGHT) Difference in holding cost of the two algorithms. The difference diverges proportional to T^2, i.e., the tuned algorithm's performance is superior by orders of magnitude. Indeed, notice that the scale of the y-axis is 10^6.

The next lemma justifies our focus on queue holding costs.

LEMMA 13. *Minimizing the inventory holding cost* $\mathbb{E}[\sum_{t=1}^T h\cdot I^t]$ *is equivalent to minimizing the request holding cost* $\mathbb{E}[\sum_{t=1}^T c\cdot Q^t]$ *with* $c_j := A_j' h$.

PROOF. We have the inventory equation $I^t = I^0 - AY^t$, where $Y^t \in \mathbb{N}^n$ represents the total amount of each request served over the interval $[1,t]$. Furthermore, we have $Q^t = Z^t - Y^t$. These two equations together imply $I^t = I^0 - A(Z^t - Q^t)$, thus

$$\sum_{t=1}^T h\cdot I^t = Th\cdot I^0 - \sum_{t=1}^T h' A Z^t + \sum_{t=1}^T h' A Q^t.$$

In terms of optimization, the first two terms are constant (cannot be affected by the controller), hence minimizing inventory cost is equivalent to minimizing queue cost. $\square$

Numerical demonstration. Before proving Proposition 3, we present some simulation results that underscore the effect of the tuning parameter.

Let us consider the base example in Section 3.1, which has $n=6$ types and $d=2$ resources. We assign costs $c_j = 1$ and initial inventories $\bar{I}_a^0 = 9, \bar{I}_b^0 = 11$. We then consider a sequence of problems with increasing horizon-length T and with initial inventory equal to $T\bar{I}^0$.

In Fig. 9 we present the performance of the two algorithms, original (with threshold $\alpha = 1/2$) and tuned. We observe that the tuned algorithm has slightly higher regret, but it has a cost that scales at a slower rate. Indeed, the difference in cost between the algorithms is $\Omega(T^2)$. In Fig. 10 we present the queues maintained by the two algorithms, where we can see that, for type $j=1$, they grow at different rates, thus explaining the difference in cost.

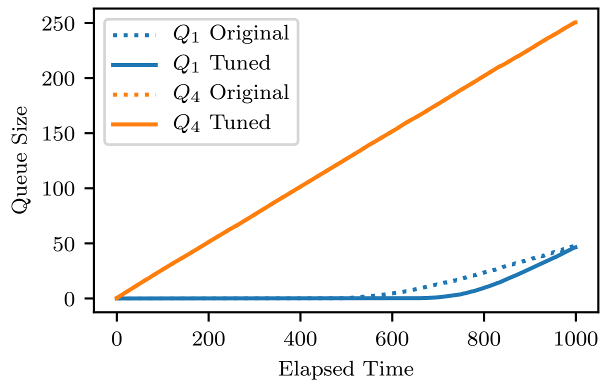


FIGURE 10. Queues maintained by BUDGETRATIO with different thresholds. The queues for types $j = 2$ and $j = 3$ (not plotted) are maintained empty with either the threshold of $1/2$ or α^T. The queue for type $j = 4$ (least desirable request type) grows at the same linear rate for both algorithms. This is because the reward maximizing policy rejects all requests of this type. The tuned threshold affects, however, the growth of the type-1 queue. This queue grows at a slower rate for the tuned (α^T) algorithm, hence accumulating lower holding cost.

Analysis of the tuned algorithm. To prove Proposition 3, we define a lower bound $C(c, \bar{y}, T)$ for the holding cost of any policy with regret $o(T)$, then show that, for any policy π with regret $o(T)$, we have:

$$\liminf_{T\uparrow\infty} \frac{\mathbb{E}^{\pi}[\sum_{t=1}^{T} c\cdot Q^t]}{C(c,\bar{y},T)} \geq 1. \tag{24}$$

Finally, we show that BUDGETRATIO achieves this bound:

$$\lim_{T\uparrow\infty} \frac{\mathbb{E}^{\hat{\pi}}[\sum_{t=1}^{T} c\cdot Q^t]}{C(c,\bar{y},T)} = 1. \tag{25}$$

The lower bound. We now proceed to construct the bound $C(c, \bar{y}, T)$. Let us recall the deterministic linear relaxation at time 0

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

Let $\bar{y}$ be the LP's solution. Recall that $\bar{y}_j$ represents the fraction of type-j requests accepted, hence $\bar{y}_j T$ is a first-order proxy for number of type-j requests that the optimal policy should accept. Indeed, let us denote by $\tilde{Y}_j^T$ the number of type-j requests accepted by the offline policy. It follows from the Lipschitz continuity of LPs that $\tilde{Y}_j^T = \bar{y}_j T + o(T)$, hence any policy π that has sub-linear

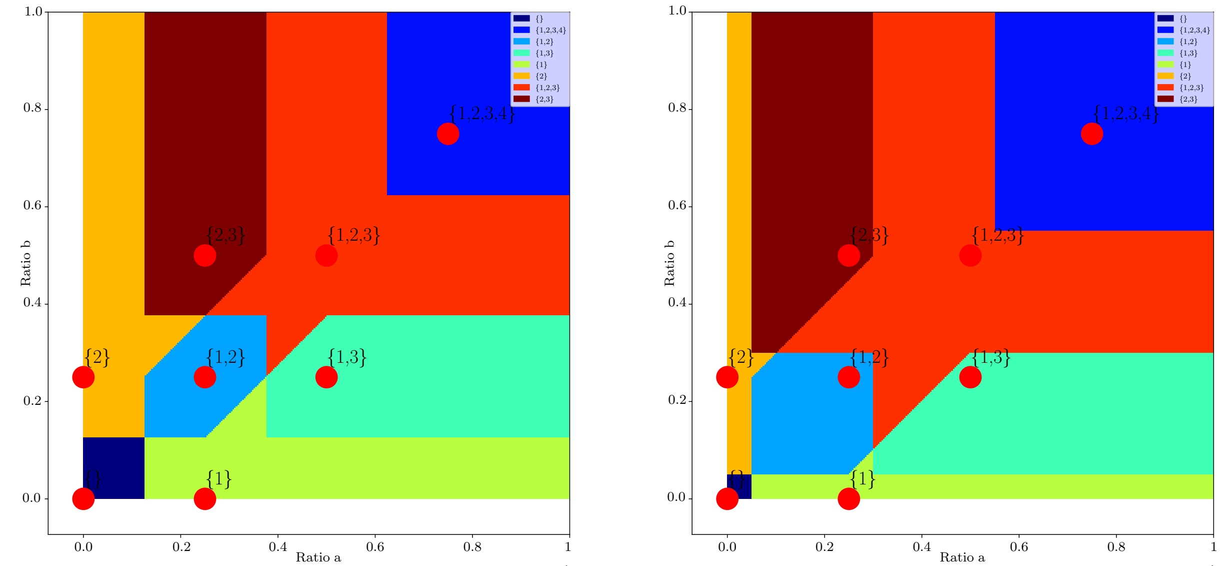


FIGURE 11. Action regions with different thresholds: (LEFT) 0.5 and, (RIGHT) 0.2. We observe that the structure of centroids and neighbourhoods is maintained. The only difference introduced by the threshold is in the shape of the regions. Recall that the smaller the threshold, the more aggressive the policy (serves more frequently). Observe that, for example, the region corresponding to $\{1,2,3,4\}$ (top-right) is larger under the threshold 0.2, i.e., request 4 is served in a larger part of the space, which is explained by the extra aggressiveness of a smaller threshold.

regret has $\mathbb{E}^\pi[Y_j^T] = \bar{y}_j T + o(T)$. Let $\delta^s_{j,\pi} = \mathbb{E}^\pi[Y_j^s]$ and $q^s_{j,\pi} = \mathbb{E}^\pi[Q_j^s]$ then $q^s_{j,\pi} = p_j s - \delta^s_{j,\pi}$. The following LP is our lower bound for the (constrained) holding-cost problem

$$
\begin{aligned}
C(c,\bar{y},T) := \min\ & \textstyle\sum_{t=1}^T c\cdot q^t \\
\text{s.t.}\ & q_j^t = p_j t - \textstyle\sum_{s=1}^t x_j^s && t\in[T],\ j\in\mathcal{J} \\
& \textstyle\sum_{s=1}^t x_j^s \le p_j t && t\in[T],\ j\in\mathcal{J} \\
& \textstyle\sum_{t=1}^T x_j^t = \bar{y}_j T && j\in\mathcal{J} \\
& x_j^t \ge 0 && t\in[T],\ j\in\mathcal{J}.
\end{aligned}
\tag{26}
$$

This problem has the following interpretation: the variables x_j^t represent how much of type-j is accepted at time t, hence the first constraint is the fluid queue (expected arrivals minus acceptance). The second constraint guarantees that the queues remain positive. Finally, the third constraint imposes that the overall acceptance matches the deterministic relaxation.

We are now ready to prove Proposition 3.

Changes in the geometry. The centroid sets remain unchanged as they do not depend on the choice of the threshold. Similarly, the optimal bases at a centroid and the centroid's neighbors also remain the same. By using the threshold α^T instead of 1/2, the action regions $\mathcal{N}_\mathcal{K}(D)$ and $\mathcal{N}_\mathcal{K}(D,\mathcal{B})$ get shifted, see Fig. 11. Specifically, in Lemma 5 (item 3), we replace the statement with:

> $R\in\mathcal{N}_\mathcal{K}(\mathcal{B},D)$ if and only if $R = r_\mathcal{K}(D) + A_{K^+}x_{K^+} - A_{K^-}x_{K^-} + b$ where $x_j\in[0,\alpha^T D_j]$ for all $j\in K^+$ and $x_j\in[0,(1-\alpha^T)D_j]$ for all $j\in K^-$.

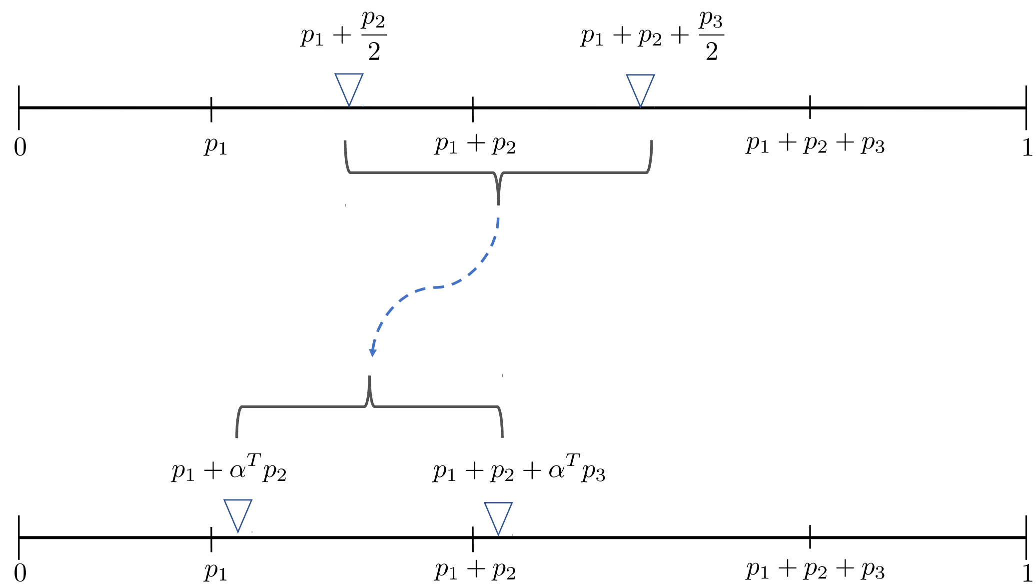


FIGURE 12. Changing the threshold from $1/2$ to α^T shifts the centroid neighborhood. For example, the set $\mathcal{N}_{\mathcal{K}}(\mathcal{B},D)$ corresponding to $\mathcal{K}=\{1,2\}$ changes from $[p_1+p_2, p_1+p_2+p_3/2]$ to $[p_1+p_2, p_1+p_2+\alpha^T p_3]$. Observe that making the parameter α smaller makes the algorithm more aggressive. Indeed, the centroids remain the same, but the boundary from $[k]$ to $[k+1]$ is now easier to cross (to start accepting more requests), because all the midpoints shift to left.

As before, $\mathcal{N}_{\mathcal{K}}(D)=\cup_{\mathcal{B}}\mathcal{N}_{\mathcal{K}}(\mathcal{B},D)$, where the union is over all bases optimal at the centroid $\mathcal{K}$. See Fig. 12 for a more detailed illustration of this "shifting".

PROOF OF PROPOSITION 3 (ITEM 1, REWARD MAXIMIZATION). The proof of bounded regret remains mostly unchanged. We limit the argument here to identifying steps in the proof of Theorem 2 and Theorem 3 where the threshold α^T plays a role and prove that those steps remain unchanged. For the remaining proofs we replace the small constants ϵ,ϵ' with their analogues $\epsilon_T=\alpha^T\epsilon$ and $\epsilon'_T=\alpha^T\epsilon'$. In particular, we replace the set $\mathcal{A}^{\epsilon}$ with the set $\mathcal{A}^{\epsilon_T}$. Notice that concentration inequalities still guarantee that $T\mathbb{P}[(A^{\epsilon_T})^c]\to 0$ as $T\to\infty$.

Changes to the proof of Theorem 2. In the statement of Lemma 11, that is used in the proof of the theorem, the vector $\theta_{\mathcal{K}}(y,D)$ is now defined by $(\theta_{\mathcal{K}}(y,D))_j=\alpha^T D_j$ for $j\in\kappa^+$ and $(\theta_{\mathcal{K}}(y,D))_j=(1-\alpha^T)D_j$ (instead of $(\theta_{\mathcal{K}}(y,D))_j=D_j/2$) for $j\in\kappa^+\cup\kappa^-$.

The drift bound in Eq. (18) changes to

$$\mathbb{E}[\delta_j^{t+1}-\delta_j^t|\mathcal{F}_t]\le\frac{-\alpha^T M}{T-t}\quad j\in\kappa^+,$$
$$\mathbb{E}[\delta_j^{t+1}-\delta_j^t|\mathcal{F}_t]\ge\frac{\alpha^T M}{T-t}\quad j\in\kappa^-.$$

which then implies an updated version of the Lyapunov drift in Eq. (15):

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

where $\xi^T = M\alpha^T$ for some constant M. The concentration argument in [2, Theorem 3] continues to hold for $\epsilon_T = \epsilon\alpha^T$ and ξ^T whenever ϵ^T, ξ^T shrink slower than $1/\sqrt{T}$.

Changes to the proof of Theorem 3. This proof remains mostly unchanged with the exception of replacing ϵ with ϵ_T as discussed before. It suffices to observe that the random walk in Lemma 14 continues to produce an exponentially decaying tail as long as ϵ_T shrinks slower than $1/\sqrt{T}$. $\square$

Proof of Proposition 3 (Item 2, Cost Minimization). We start with the lower bound in Eq. (24). Fix a policy π and recall $q^s := \mathbb{E}^\pi[Q^s]$, $\delta_j^s = \mathbb{E}^\pi[Y_j^s]$, and $x_j^t = \mathbb{E}[Y_j^t - Y_j^{t-1}] \geq 0$ for $t \in [T]$ (with $Y^0 = 0$). The expected queue length satisfies $q_j^s = p_j s - \delta_j^s$. Also, $q_j^t, x_j^t \geq 0$ for all $t \in [T]$. Because the queue must be positive we also have $\sum_{s=1}^t x_j^s \leq p_j t$.

In turn, $\mathbb{E}^\pi[\sum_{t=1}^T c\cdot Q^t] \geq \sum_{t=1}^T c\cdot \bar{q}^t$, for any pair $\bar{q}, \bar{x} \geq 0$ that satisfies these constraints and has $\bar{\delta}^T = \sum_{s=1}^T \bar{x}_j^s = \delta^T$ (same total number of acceptances). Thus, we must have that

$$\mathbb{E}^\pi\left[\sum_{t=1}^T c\cdot Q^t\right] \geq C(c, \delta^T/T, T).$$

We make two claims: (i) $||\tilde{Y}^T - \bar{y}T||_\infty = o(T)$, and (ii) $||\delta^T - \bar{y}T||_\infty = o(T)$. Recalling that $\tilde{Y}^T$ is the expected offline allocation, the first claim is that offline is $o(T)$ consistent with the deterministic relaxation. Recalling that $\delta^T = \mathbb{E}^\pi[Y^T]$, the second claim is that the online policy is, as well, $o(T)$ consistent with that relaxation and, in turn, with offline. The first claim follows directly from the Lipschitz continuity of LPs [15]. For claim (ii), from (i) and the fact that π has $o(T)$ regret, we conclude $v'\delta^T = v'\bar{y}T + o(T)$. Take a convergent subsequence of δ^T/T with limit $\tilde{y}$. By our previous equation, $v'\tilde{y} = v'\bar{y}$ and $\tilde{y}$ is feasible for $LP(\mathbb{E}[R^0], D)$, hence $\tilde{y} = \bar{y}$ by our uniqueness assumption, proving claim (ii).

Under the assumption that $\bar{y}_j < p_j$ for at least one j, we trivially have that $C(c, \bar{y}, T) \geq \Gamma T^2$ for some $\Gamma > 0$. By the Lipschitz continuity of LP in the right hand side, we have that $||q - \bar{q}|| = o(T)$ so that $C(c, \bar{y}, T) = C(c, \delta^T/T, T) + o(T^2)$. In particular,

$$\liminf_{T\uparrow\infty} \frac{\mathbb{E}^\pi[\sum_{t=1}^T c\cdot Q^t]}{C(c, \bar{y}, T)} = \liminf_{T\uparrow\infty} \frac{\mathbb{E}^\pi[\sum_{t=1}^T c\cdot Q^t]}{C(c, \delta_j^T/T, T)} \frac{C(c, \delta_j^T/T, T)}{C(c, \bar{y}, T)} \geq 1,$$

which corresponds to the lower bound Eq. (24).

We now prove that BudgetRatio achieves the lower bound, i.e., Eq. (25). We start with a simple observation about the structure of optimal solution to $C(c, \bar{y}, T)$. For every j such that $\bar{y}_j = p_j$, we have $\bar{q}_j^t = 0$ for all $t \in [T]$. For all j with $\bar{y}_j < p_j$ we let $t_j^0 := \bar{y}_j T/p_j < T$. Then, $x_j^s = p_j$ for all $s \leq \lfloor t_j^0 \rfloor$; $x_j^{\lfloor t_j^0 + 1\rfloor} = \bar{y}_j T - p_j t_j^0$ and $x_j^t = 0$ for all $t > \lfloor t_j^0 \rfloor + 1$. That is, a type j is fully served up to some time t_j^0 and then not served at all.

We claim that for each j, there exists a time τ_j such that:

(I) $\mathbb{E}[Y_j^{\tau_j}] = p_j \mathbb{E}[\tau_j] + o(T)$: most requests j are taken over the interval $[1, \tau_j]$.

(II) $\mathbb{E}[|\tau_j - t_j^0|] = o(T)$: τ_j is close to the deterministic time t_j^0

(III) $\mathbb{E}[Y_j^T - Y_j^{\tau_j}] = o(T)$: few requests j are taken after τ_j.

These properties are sufficient to prove Eq. (25). Indeed, (I) and (II) combined guarantee that $\mathbb{E}[Y_j^{t_j^0}] = p_j t_j^0 + o(T)$. Since $\mathbb{E}[Y_j^t] \leq p_j t$ for all t, we must have that $\mathbb{E}[Y_j^t] = p_j t + o(T)$ for all $t \leq t_j^0$ so that $\mathbb{E}[Q_j^t] = o(T)$ for all $t \leq t_j^0$. By (III), $\mathbb{E}[Y_j^T - Y_j^{t_j^0}] = o(T)$ so that $\mathbb{E}[Q_j^t] = p_j(t - t_j^0) + o(T)$ for all $t \geq t_j^0$. In conclusion, for all $t \geq 0$, $\mathbb{E}[Q^t] - \bar{q}^t = o(T)$ so that

$$\mathbb{E}^{\widehat{\pi}}\left[\sum_{t=1}^{T} c \cdot Q^t\right] = C(c, \bar{y}, T) + o(T^2).$$

This then guarantees that

$$\lim_{T \uparrow \infty} \frac{\mathbb{E}^{\widehat{\pi}}[\sum_{t=1}^{T} c \cdot Q^t]}{C(c, \bar{y}, T)} = 1.$$

Let us turn, then, to establishing properties (I)-(III) for BudgetRatio. We divide the argument into two cases: $j \in \mathcal{K}$ and $j \notin \mathcal{K}$.

Case 1: $j \in \mathcal{K}$**.** Let us define τ_j as the first time when the algorithm rejects j, i.e., the first time the solution y to $\mathrm{LP}(R^t, p)$ has $y_j < \alpha^T p_j$. Let us also define $\tau = \tau^{\epsilon_T', \mathcal{K}} \wedge \tau^{\epsilon_T', \mathcal{B}}$ to be the time when the process completely escapes the action region and the cone. From Theorems 2 and 3 we have the properties $\mathbb{E}[T - \tau] = O(1)$ and $\mathbb{E}[Y^T - Y^\tau] = O(1)$. On the other hand, from our main result Theorem 1 we have $\mathbb{E}[\tilde{Y}^T - Y^T] = O(1)$. Recall that up to time τ we perform basic allocation, hence if we denote $\mathcal{B}$ the optimal offline basis, we have the inventory equation $I^t = I^0 - \mathcal{B} Y^t$ for all $t \leq \tau$, so that

$$\mathcal{B}(Y^\tau - Y^{\tau_j}) = (T - \tau)(R^{\tau_j} - R^\tau) + (\tau - \tau_j) R^{\tau_j}.$$

From this equation we claim $Y_j^\tau - Y_j^{\tau_j} = o(T)$. To see this, recall that the solution y to the LP is $\mathcal{B}^{-1} R^t$ and by definition of τ_j we have $y_j \leq \alpha^T p_j$, hence the second term is $(\tau - \tau_j) R^{\tau_j} = o(T)$. The first term is bounded by Theorem 2 (see also Remark 5), where we proved that the process R^t remains close to any boundary it hits, i.e., from Theorem 2 we deduce $||R^{\tau_j} - R^\tau|| \leq \epsilon^T = \epsilon \alpha^T$, finishing the proof of our claim.

From here properties (I)-(III) are straightforward algebraic checks. For (I), since all requests j are accepted up to τ_j, $\mathbb{E}[Y^{\tau_j}] = \mathbb{E}[Z^{\tau_j}] = p_j \mathbb{E}[\tau_j]$. For (III), since $\mathbb{E}[Y^T - Y^\tau] = O(1)$ (Theorems 2 and 3) and we just proved the claim $Y_j^\tau - Y_j^{\tau_j} = o(T)$, the property holds. For (II), since $\tilde{Y}^T - Y^\tau = O(1)$ and $\tilde{Y}_j^T - T\bar{y}_j = o(T)$, from our claim we obtain $Y_j^{\tau_j} - T\bar{y}_j = o(T)$. Since, by definition, $t_j^0 = \bar{y}_j T / p_j$ and $\mathbb{E}[Y_j^{\tau_j}] = p_j \mathbb{E}[\tau_j]$, we get the result.

Case 2: $j \notin \mathcal{K}$**.** Set $\tau_j = 0$ and notice that, in this case, $\bar{y}_j < \alpha^T p_j$, which implies $\tilde{Y}_j^T = T\bar{y}_j + o(T) = o(T)$. By our main result $Y^T - \tilde{Y}^T = O(1)$, hence $Y^T = o(T)$. Finally, since $t_j^0 = \bar{y}_j T / p_j = o(T)$, all properties (I)-(III) follow directly. □

7. Robustness with respect to primitives We focus on the case where either the arrival and restock probabilities p, or the rewards v, are misspecified. We will identify conditions that guarantee the robustness of BudgetRatio to misspecification. Under these conditions BudgetRatio's constant regret persists under perturbed primitives. We will subsequently use these conditions to state a regret guarantee for the case where the primitives are learned rather than known in advance.

7.1. Demand perturbations Suppose that, instead of knowing the distribution p exactly, we have an estimate $\widetilde{p} \neq p$. In this section, we quantify how close $\widetilde{p}$ needs to be to p so that the algorithm using the estimate $\widetilde{p}$ still has constant regret. Crucially, our "measure of closeness" arises naturally from our geometric framework and is expressed in terms of centroid budgets.

Definition 8 (centroid separation). We say that the centroids are δ-separated if

$$\min_{\mathcal{K}\neq\mathcal{K}'} \min_{i\in[d]} |(r_{\mathcal{K}}(p) - r_{\mathcal{K}'}(p))_i| \geq \delta.$$

Note that in Definition 8 we focus on the "true centroid budgets", i.e., centroid budgets under the true distribution p. We present an illustration of the action regions for p and $\widetilde{p}$ in Fig. 2.

Proposition 1. *Let p be the true underlying distribution and assume the centroids are δ-separated. Then, for any distribution $\widetilde{p}$ such that*

$$\max_{\mathcal{K}} ||r_{\mathcal{K}}(p) - r_{\mathcal{K}}(\widetilde{p})||_\infty \leq \frac{\delta}{4} \quad \text{and} \quad \widetilde{p} \text{ satisfies Assumption 1,} \tag{27}$$

BudgetRatio*, with the LP solved each period with $\widetilde{p}$ replacing p, produces bounded regret in the sense of Theorem 1.*

Proof. The proof of our main result does not change, hence we limit to point the places where p appears in the proof and explain why the same steps continue to hold. Since $\widetilde{p}$ satisfies slow restock, Lemma 3 does not change. Observe that all of our constructions of the action regions $\mathcal{N}_{\mathcal{K}}(D)$ and $\mathcal{N}_{\mathcal{K}}(D,\mathcal{B})$ are for arbitrary D. Additionally, cones do not depend on the probability distribution, hence Lemma 2 continues to hold. For Theorem 2, the main thing to prove is that the initial condition implies (through the drift) that the process R^t remains close to the action region; we have that $r_{\mathcal{K}}(p) \in \mathcal{N}_{\mathcal{K}}(\widetilde{p})$ so that the initial condition is preserved and the proof Theorem 2 uses the region $\mathcal{N}_{\mathcal{K}}(\widetilde{p})$ (instead of $\mathcal{N}_{\mathcal{K}}(p)$). Note that the process R^t is shifted by a constant, namely $\widetilde{p}_{\mathcal{R}} - p_{\mathcal{R}}$ and all the drift analysis is on differences $R^{t+1} - R^t$, so the perturbed restock plays no role. Finally, for Theorem 3, there is an offset $(T-s)(\widetilde{p}-p)$ in Eq. (22), but this does not affect the concentration arguments because we still have $G^1 \leq \epsilon T$ and $G^T \leq 0$. $\square$

Example 1 (One Dimension). The separation condition in Proposition 1 reduces into an intuitive requirement in the one dimensional case. In this case, centroids are of the form $[j]$ so that

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

where we define $[0] := \varnothing$. Then, the separation condition (27) stipulates that

$$\left|\sum_{k=1}^{j}(p_k - \widetilde{p}_k)\right| \le \frac{\min_k p_k}{4} \quad \forall j \in [n].$$

It is, in particular, sufficient that $||p - \widetilde{p}||_\infty \le \frac{\min_k p_k}{4n}$. ∎

Learning. Proposition 1 implies immediately an $O(\log T)$ regret for the setting where p is not known and hence to be learned. The controller must make decisions while learning the correct type probabilities p. The controller observes the type (or "context") $j \in [n]$ at each period and builds the estimate $\hat{p}_j^t = \frac{1}{t}\sum_{\tau=1}^{t} \mathbb{1}_{\{J^\tau = j\}}$, where J^τ is the random arrival at τ.

Corollary 1 **(learning the demand distribution)**. *Assume the separation condition* (27)*. Then, without prior knowledge of* p*, a modification of* BudgetRatio *achieves* $O(\log T)$ *regret.*

Proof. Fix the constant $\epsilon = \frac{\delta}{4n}$. If we have an initial exploration phase of length $c\log T$, for some $c = c(\epsilon) > 0$, where all requests are rejected, then the empirical $\hat{p}^t$ is s.t. $||p - \hat{p}^t|| \le \epsilon$ for all $t \ge c\log T$ with probability $\frac{1}{T}$. After time $c\log T$ we run BudgetRatio and obtain constant regret in the remaining periods in virtue of Proposition 1. □

7.2. Reward perturbations

Suppose now that the values v are not known a priori and we have, instead, an estimate $\tilde{v}$. How close does $\tilde{v}$ have to be to v, for BudgetRatio (run with $\tilde{v}$) to achieve constant regret?

A "separation" requirement is established for the one dimensional case in [21]. There, the requirement is based on ranking the requests by reward. With multiple resources one must take into account the resource consumed by each request type. We show that, under a complementary slackness condition, the centroids are stable to local perturbation of v and so is, in turn, the regret. In Example 3 we show that the our general condition reduces to the separation requirement in [21].

Reminder of complementary slackness. In the standard form (LP(R,D)) we have d resource consumption constraints of the form $Ay + s = R$ and n demand constraints of the form $y + u = D$. Therefore, we have $d+n$ dual variables and an optimal primal-dual pair $(y,\lambda) \in \mathbb{R}^n \times \mathbb{R}^{d+n}$ must satisfy the following complementarity: $s_i > 0 \Rightarrow \lambda_i = 0$, $u_j > 0 \Rightarrow \lambda_j = 0$, and, from the dual constraints, $([A'|I^n]\lambda)_j > v_j \Rightarrow y_j = 0$. There is the possibility that, for y_j non-basic (hence $y_j = 0$), we have $([A'|I^n]\lambda)_j = v_j$, i.e., complementarity is not strict. In our separation condition, we require strict complementarity.

Recall that we associate to v the extended value vector $\bar{v} := (v, 0, 0) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^d$, where the zeros correspond to unmet and surplus variables.

Definition 9 (δ-Complementarity). Let $\mathcal{B}$ be a basis and λ be the dual variables associated to $(\mathcal{B}, v)$, i.e., $\lambda = (\mathcal{B}^{-1})'\bar{v}_{\mathcal{B}}$. We say that $\mathcal{B}$ is δ-complementary if (i) $\lambda_i \geq \delta$ for i s.t. $s_i \notin \mathcal{B}$, (ii) $\lambda_j \geq \delta$ for j s.t. $u_j \notin \mathcal{B}$, and (iii) $(\bar{A}'\lambda)_j \geq v_j + \delta$ for j s.t. $y_j \notin \mathcal{B}$.

Proposition 2. *Let v be the true values and let δ be such that all the bases are δ-complementary. Then, there exists a constant $c \geq 1$ such that, for any $\tilde{v}$ that satisfies*

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

BudgetRatio*, with the LP solved each period with $\tilde{v}$ replacing v, produces bounded regret in the sense of Theorem 1. Furthermore $c \leq \max_{\mathcal{B}} ||\mathcal{B}^{-1}||_\infty$.*

Proof. Fix a basis $\mathcal{B}$ associated to some centroid $\mathcal{K}$ under values v. Let λ be the dual variables associated to $(\mathcal{B}, v)$. Recall that, by Lemma 2, we know that $\lambda \geq 0$ and $\bar{A}'\lambda \geq v$. We first prove that $\mathcal{B}$ is also an optimal basis associated to $\mathcal{K}$ under values $\tilde{v}$. In virtue of Lemma 2, it suffices to show that the dual variables $\tilde{\lambda}$ associated to $(\mathcal{B}, \tilde{v})$ satisfy (i) $\tilde{\lambda} \geq 0$ and (ii) $\bar{A}'\tilde{\lambda} \geq \tilde{v}$.

Using complementary slackness, the $d+n$ dual variables are the solution to the following $d+n$ linear equations:

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

Observe that the left-hand side of the equations is invariant (depends on $A, \mathcal{B}$ only). We thus can write $H\tilde{\lambda} = \tilde{h}$ and note that the same system is satisfied with λ, i.e., $H\lambda = h$, with $(h - \tilde{h})_j = (v_j - \tilde{v}_j)\mathbb{1}_{\{y_j \in \mathcal{B}\}}$. Finally, by the Lipschitz continuity of linear systems [15], we have that $||\lambda - \tilde{\lambda}|| \leq c||h - \tilde{h}||$. It remains to conclude the properties (i) $\tilde{\lambda} \geq 0$ and (ii) $\bar{A}'\tilde{\lambda} \geq \tilde{v}$.

For (i), from the system $H\tilde{\lambda} = h$ it follows that we only need to study the case $u_j \notin \mathcal{B}$ or $s_i \notin \mathcal{B}$. In this case, by δ-complementarity, we are guaranteed $\tilde{\lambda} \geq \delta - c||h - \tilde{h}||_\infty \geq 0$, as desired. For (ii), by the same reasoning, we need to study only the dual constraints $([A'|I^n]\tilde{\lambda})_j \geq \tilde{v}_j$ for $y_j \notin \mathcal{B}$. Let us denote η the j-th row of $[A'|I^n]$. By δ-complementarity, we have

$$([A'|I^n]\tilde{\lambda})_j = \eta\tilde{\lambda} \geq \delta + v_j + \eta(\tilde{\lambda} - \lambda) \geq v_j + \delta - (||A_j||_1 + 1)c||v - \tilde{v}||_\infty \geq 0.$$

We have proved that all centroids and their bases remain the same under $\tilde{v}$. We conclude now the proof by observing that all of our constructions depend on identifying the centroids only and the values v are used nowhere else, thus none of our proofs change. $\square$

Example 2 (One Dimension). With $d=1$, δ-complementarity reduces to the following natural condition: $v_j \geq \delta$ for all $j \in [n]$ and $|v_j - v_{j'}| \geq \delta$ for all $j \neq j'$. In other words, δ-complementarity captures exactly the condition needed to distinguish among requests.

The primal and dual problems are as follows, where λ_0 denotes the resource multiplier and λ_j the demand multipliers for $j \in [n]$:

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

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

Finally, we can verify conditions (i)-(iii) of Definition 9. 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 shows that indeed the proposed condition guarantees δ-complementarity. ■

Example 3 (Necessity of Separation). Consider the case $d = 1$ resource with $n = 2$ infinitely patient request types. We focus on the robust version (no learning/feedback), where all we have is access to an estimate $\tilde{v}$ and the true rewards are never observed. If the inventory is small compared to T, say $I^0 = T/100$, the optimal policy waits till the end and chooses just one type—the one with maximum reward—and allocates all the inventory to that type. In the absence of separation, we cannot distinguish between $v_1 > v_2$ and $v_1 < v_2$ with $\tilde{v}$, hence no policy can obtain constant regret in this context. ■

Learning. Consider the setting where the controller must make decisions while learning the correct values v. Formally, type-j requests draw a reward $V_j \sim F_j$, where F_j is some unknown distribution, and $v_j = \mathbb{E}[V_j]$ represents the true expectation. At each time, the controller observes the type $j \in [n]$ and, if the request is accepted, the controller observes a realization of V_j, that is used to estimate v through its empirical average $\hat{v}_j^t$.

Corollary 2 **(learning the reward distribution)**. *Assume the separation condition in Proposition 2. Further, assume that the distributions F_j are sub-Gaussian, but otherwise unknown and arbitrary. Then, a modification of* BudgetRatio *achieves $O(\log T)$ regret.*

Proof. Fix the constant $\epsilon = \frac{\delta}{c(d+2)}$. If we have an initial exploration phase of length $c' \log T$, for some $c' = c'(\epsilon) > 0$, where all requests are accepted, then the empirical $\hat{v}^t$ is s.t. $||v - \hat{v}^t|| \leq \epsilon$ for all $t \geq c \log T$ with probability $\frac{1}{T}$. After time $c' \log T$ we run BudgetRatio and obtain constant regret in the remaining periods in virtue of Proposition 2. □

8. Interpretation of BudgetRatio as a max-bid-price control Bid-price heuristics are popular due to their intuitive interpretation. There are several examples where these policies achieve provably good results; see [18] for asymptotic results and [6] for a broader overview of bid-prices in the context of revenue management.

In a setting with multiple resources, a bid-price policy is described as follows: at time t compute a vector $\lambda^t \in \mathbb{R}^d_{\geq 0}$ of resource prices, and reject a type-j arrival if and only if its reward is below the combined price of requested resources, i.e., if and only if $v_j < A_j' \lambda^t$. To the best of our knowledge, there are no bid-price policies with constant regret in the literature and the strongest available guarantee is $O(\sqrt{T})$ regret in asymptotic regimes [18].

An enhanced version of bid-price achieves constant regret. This version *considers the maximum of several prices.*

Definition 10. A policy is max-bid-price if it can be described by a (possibly adaptive) set of prices $\Lambda^t \subseteq \mathbb{R}^d_{\geq 0}$ for each t and such that, at time t, a request type j is accepted $I_i^t > 0$ for all i with $A_{ij} = 1$ (there are enough resources) and $v_j \geq \max_{\lambda \in \Lambda^t} A_j' \lambda$.

Recall that a basis $\mathcal{B}$ is associated to a centroid $\mathcal{K}$ if $\mathcal{B}$ is optimal at the centroid budget (Definition 6) and the dual variables associated to a basis $\mathcal{B}$ are $\lambda = (\mathcal{B}^{-1})' \bar{v}_{\mathcal{B}}$ (Definition 9). To define bid-prices, we identify the centroid $\mathcal{K}$ such that $R^t \in \mathcal{N}_{\mathcal{K}}(p)$, and subsequently define

$$\Lambda^t(R^t) := \{\lambda : \lambda \text{ are duals associated to } \mathcal{B} \text{ for some basis } \mathcal{B} \text{ associated to } \mathcal{K}\}.$$

At each time t, we compute this set of prices $\Lambda^t(R^t) \subseteq \mathbb{R}^d_{\geq 0}$ and use the decision rule of Definition 10. In Fig. 13 we show the action regions and their interpretation as max-bid-price. Notice that *the "traditional" bid-price heuristic is recovered in this definition when the set $\Lambda^t(R^t)$ is a singleton for each t.*

Proposition 4. *Assume that all the bases are δ-complementary (Definition 9) for some $\delta > 0$. Then,* BudgetRatio *is equivalent to max-bid-price.*

Proof. We prove that the decision rule with this set of prices is equivalent to BudgetRatio. We divide the analysis into $j \in \mathcal{K}$ (acceptance) and $j \notin \mathcal{K}$ (rejection). The case $j \in \mathcal{K}$ follows easily since, at the centroid budget $r_{\mathcal{K}}$, $y_j = p_j > 0$ for all bases $\mathcal{B}$ associated to $\mathcal{K}$, so by complementary slackness $v_j \geq A_j' \lambda$ for any such associated dual variable.

We are left to prove the case $j \notin \mathcal{K}$. First, we claim that some basis $\mathcal{B}$ is such that $y_j \notin \mathcal{B}$. Assuming this claim, δ-complementarity yields $v_j < A_j' \lambda - \delta$, so that j fails the test with the λ associated to this basis.

To prove this claim, assume towards contradiction that $j \notin \mathcal{K}$ is such that $y_j \in \mathcal{B}$ for all bases associated to $\mathcal{K}$. Since $\mathcal{K} \neq \varnothing$, we have $r_{\mathcal{K}} \neq 0$. Recall that $\mathcal{N}_{\mathcal{K}}(D) = \cup_{\mathcal{B}} \mathcal{N}_{\mathcal{K}}(D, \mathcal{B})$, i.e., $\mathcal{N}_{\mathcal{K}}(D)$ is

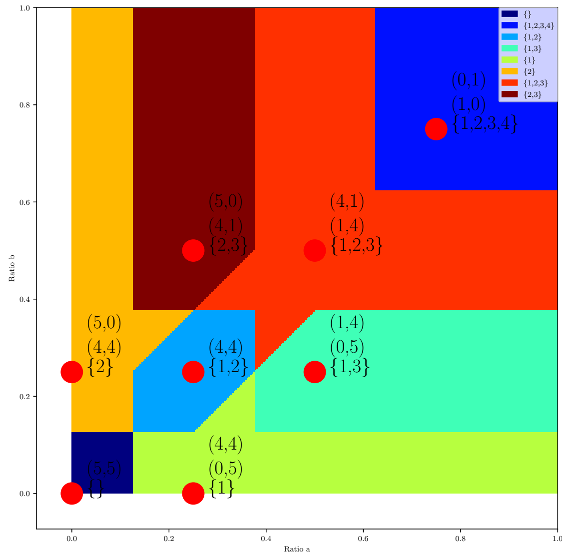


FIGURE 13. Action regions with the associated bid-prices. Every centroid $\mathcal{K} \neq \varnothing, \{1,2\}$ has two vectors of shadow prices with components (λ_a, λ_b). For example, if $R^t \in \mathcal{N}_{\{1\}}$ (bottom region), then the decision rule is to reject if $v_j < (4,4)' A_j$ or $v_j < (0,5)' A_j$.

divided into finitely many regions. Therefore, we can take a direction η such that $r_{\mathcal{K}} \pm \eta$ is in the interior of these regions (if $r_{\mathcal{K}} + \eta$ lies in a boundary, a perturbation will be in the interior because there are finitely many boundaries).

In conclusion, for some bases $\mathcal{B}, \mathcal{B}'$ we have $r_{\mathcal{K}} + \eta \in \text{interior}(\mathcal{N}_{\mathcal{K}}(\mathcal{B}, D))$ and $r_{\mathcal{K}} - \eta \in \text{interior}(\mathcal{N}_{\mathcal{K}}(\mathcal{B}', D))$. Observe that, in the interior of a region, all the basic variables are strictly positive. Since $y_j \in \mathcal{B}, \mathcal{B}'$, it must be that the solutions y^+ and y^- at ratios $r_{\mathcal{K}} + \eta$ and $r_{\mathcal{K}} - \eta$ have $y_j^+, y_j^- > 0$. This is a contradiction with the fact $y_j = 0$ at the centroid budget $r_{\mathcal{K}}$. $\square$

The way we construct Λ^t may suggest that we need to pre-compute all bases associated to all centroids, but this is not the case. Indeed, it suffices to pre-compute the duals for the initial centroid (i.e., such that $R^0 \in \mathcal{N}_{\mathcal{K}}(p)$ and) and its immediate neighbors; by our analysis, up to $\tau_{region}^{\epsilon',\mathcal{K}} \wedge \tau_{cone}^{\epsilon',\mathcal{B}}$ the process R^t remains in these sets.

9. Concluding remarks We consider a family of resource allocation problems and show that a simple resolving algorithm achieves bounded regret in terms of the total rewards collected. Our approach is geometric and based on a parametric characterization of the packing LP. This approach allows us not only to expand the coverage in terms of network primitives (adding, for example, inventory arrivals), but also formally pose and resolve some open problems that are not obviously amenable to other methods.

Our fundamental definition is that of centroids, which correspond to subsets of requests that should be fully served. For each demand D, a centroid $\mathcal{K}$ has associated a region $\mathcal{N}_{\mathcal{K}}(D)$. The key to our analysis is to understand how the process of ratios R^t evolves relative to these regions $\mathcal{N}_{\mathcal{K}}(D)$ and show that it is attracted to the correct subset (that used by the offline controller).

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

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

1. *Modelling assumptions*: we showed that (i) queues (even if allowed) can be ignored and (ii) inventory arrivals are allowed but one must impose a slow-restock assumption. In the absence of such an assumption, no algorithm can achieve bounded regret.
2. *Optimization objectives*: we showed that the algorithm is sufficiently flexible to allow for nested objectives. With suitable tuning of the thresholds, BudgetRatio achieves bounded regret on the reward while minimizing (in the sense of competitive ratio) the inventory holding cost incurred over the horizon.
3. *Model parameters*: We considered the effect of running BudgetRatio with misspecified arrival probabilities p and rewards v. We proved that, if the parameters p and v are estimated within a constant error, BudgetRatio still achieves constant regret. Crucially, both robustness results hinge on the the notion of centroids. These provide a "language" through which can generalize to multiple dimensions separation conditions that were previously provided only in the single dimensional case.

It remains interesting to uncover the extent to which some of our assumptions and conditions are necessary. Although a slow restock assumption is generally necessary, it is evident that Assumption 1 is not the weakest possible. Lemma 3—the only place where this assumption is used—still holds if we assume instead $p_i < \sum_{j \in \mathcal{K}} a_{ij} p_j$ for every centroid $\mathcal{K}$ and every resource i used by some $j \in \mathcal{K}$.

Possibly more interesting is the necessity of the separation conditions we derived in Section 7, i.e., if tighter conditions can be obtained. We know that the separation requirement (based on δ-complementarity; see Proposition 2) is necessary for the one-dimensional case. It is unclear whether it is necessary for $d > 1$.

References

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

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

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

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

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

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

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

[8] Pornpawee Bumpensanti and He Wang. A re-solving heuristic with uniformly bounded loss for network revenue management. *Management Science*, 66(7), 2020.

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

[10] Devdatt Dubhashi and Alessandro Panconesi. *Concentration of measure for the analysis of randomized algorithms.* Cambridge University Press, 2009.

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

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

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

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

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

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

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

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

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

[20] Alberto Vera and Siddhartha Banerjee. The bayesian prophet: A low-regret framework for online decision making. *Management Science*, 2020. Forthcoming.

[21] Alberto Vera, Siddhartha Banerjee, and Itai Gurvich. Online allocation and pricing: Constant regret via bellman inequalities. *Operations Research*, 2020. Forthcoming.

Appendix A: Fast restock In this section we prove Lemma 1 showing that, when Assumption 1 (slow restock) is violated, it is not possible to obtain constant regret against the offline benchmark. We consider as a counter example a network with one request type and one resource, both arriving with probability p. Figure 14 is the numerical illustration of the $\sqrt{T}$ regret.

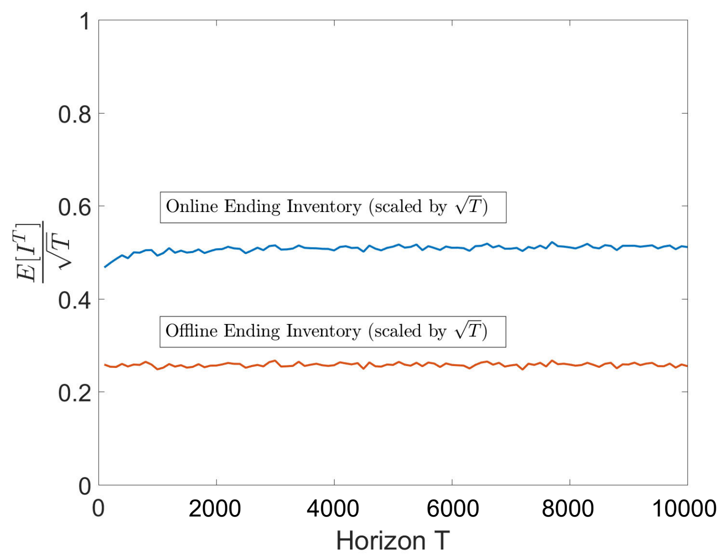


FIGURE 14. Single RAN with impatient requests and restock. The first line depicts the expected remaining inventory of offline. The second line depicts the expected remaining inventory of the optimal online policy. The difference is the regret. The expectation is computed as an average over 10000 replications. The y-axis is the (expected) ending inventory divided by $\sqrt{T}$.

We label the request by 1, the resource by a, and set $v_1 = 1$, $p_1 = p_a = p$. Let Z_1^t be the number of request arrivals by time t and let Z_a^t be the restock by time t. Let I_{off}^T (respectively I_{on}^T) be the end-of-horizon residual inventory under offline (respectively online) policies. Then, $V_{off}^*(T) = \mathbb{E}[\min\{Z_a^t, Z_1^t\}] = \mathbb{E}[Z_a^T - I_{off}^T]$.

The optimal online policy is to always serve if possible. Let Y^t be the number of requests accepted by the online policy by time t. Then, $V^*_{on} = \mathbb{E}[Y^T] = \mathbb{E}[Z^T_a - I^T_{on}]$. Thus, $V^*_{off}(T) - V^*_{on}(T) = \mathbb{E}[I^T_{on}] - \mathbb{E}[I^T_{off}]$. We analyze now the two (end-of-horizon) inventory levels, starting with offline, which satisfies:

$$I^T_{off} = (Z^T_a - Z^T_1)^+.$$

The process $Z^t_a - Z^t_1$ is a random walk starting at 0 and with i.i.d zero-mean increments $X_1, \ldots, X_T$ taking values $\{-1, 0, 1\}$ with probabilities $p(1-p), 1-2p(1-p), p(1-p)$; X_t is the difference between the restock at t (0 or 1) and the request arrival (0 or 1). Write $G^T = \sum_{t=1}^T X_t$. By the central limit theorem

$$\frac{1}{\sqrt{T}} \sum_{i=1}^T X_i \Rightarrow \mathcal{N}(0, \sigma^2),$$

where $\sigma^2 := 2p(1-p)$. Since $I^T_{off} = (G^T)^+$, by the continuous mapping theorem we have $\frac{1}{\sqrt{T}}\mathbb{E}[I^T_{off}] \to \mathbb{E}[(\mathcal{N}(0,\sigma^2))^+]$. On the other hand, the online inventory satisfies the queueing recursion $I^{t+1}_{on} = [I^t_{on} + X_t]^+$ so that

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

It is well-known (reflection principle) that the following equivalence in law holds

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

It is also well-known that

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

where $\mathcal{Z}$ is distributed as the supremum of a Brownian motion. This convergence also holds in expectation. The reflection principle for Brownian motion then guarantees that $\mathbb{P}\{\mathcal{Z}' > a\} = 2\mathbb{P}\{\mathcal{N}(0,\sigma^2) > a\}$ and this allows us to conclude that

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

Appendix B: Proofs of auxiliary lemmas

Proof of Lemma 5.

Item 1. In virtue of Lemma 2, it suffices to show that $\mathcal{B}^{-1}\begin{pmatrix} R \\ D \end{pmatrix} \ge 0$ to conclude the optimality of $\mathcal{B}$. Recall that the matrix $\bar{A}$ is given by

$$\bar{A} = \begin{bmatrix} A & 0 & I^d \\ I^n & I^n & 0 \end{bmatrix},$$

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

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

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

By the definition of centroid, all request variables $\mathcal{K}$ are saturated at $r_{\mathcal{K}}$, hence $K^+ \subseteq \mathcal{K}^c$; in other words, zero-valued requests cannot come from $\mathcal{K}$. Similarly, unfulfilled variables $\mathcal{K}^c$ are saturated at $r_{\mathcal{K}}$, therefore $K^- \subseteq \mathcal{K}$. We deduced the inclusions $\kappa^+ \subseteq \mathcal{K}^c \cap \mathcal{K}^+$ and $\kappa^- \subseteq \mathcal{K} \cap \mathcal{K}^-$. Using the previous fact, by inspection we have the following identities

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

We are now ready to prove $\mathcal{B}^{-1}\begin{pmatrix} R \\ D \end{pmatrix} \geq 0$. Indeed, pre-multiplying the previous identities by $\mathcal{B}^{-1}$,

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

Since we have the inclusions $\kappa^+ \subseteq \mathcal{K}^c \cap \mathcal{K}^+$ and $\kappa^- \subseteq \mathcal{K} \cap \mathcal{K}^-$, it can be verified that the last expression is non-negative, completing the proof of the first item.

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

We turn to prove that, if $\mathcal{B}$ is optimal, then R has the stated representation. Let $(y,\ u,\ s)' = \mathcal{B}^{-1}\begin{pmatrix} R \\ D \end{pmatrix}$, where y are the request variables, u are unmet variables s are surplus variables. Let K^+, K^- and K^0 be as in Definition 5. 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, we know that $u_j = D_j - y_j$ and $y_j, u_j \geq 0$. Since $R \in \mathcal{N}_{\mathcal{K}}$ we also know that $y_j \geq D_j/2$ for all $j \in K^-$ and $y_j < D_j/2$ for all $j \in K^+$. From the vector (y, u, s) we subtract the vector $(\bar{y}, \bar{u}, \bar{s})$ given by

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

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

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

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

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

Let us apply the following transformation

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

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

Vectors $x = Pz$ are elements in the polyhedron

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Finally, by the definition of centroid

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

The last two equations together prove the claim in virtue of Lemma 2. □

Proof of Lemma 7. The existence of separating vectors that have $\max_l \psi_l' x \leq 0$ for any $x \in \text{cone}(\mathcal{K}, \mathcal{B})$ follows from the standard results.

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

$$x - r_{\mathcal{K}}(D) = \sum_{(\kappa^+, \kappa^-)} \alpha[\kappa^+, \kappa^-](A_{\kappa^+} z_{\kappa^+} - A_{\kappa^-} z_{\kappa^-}),$$

where $z \geq 0$. It follows immediately that in the convex combination it suffices to include κ^+ and κ^- that are minimal. That is such that $|\kappa^+| = 1$ or that $|\kappa^-| = 1$. We refer to these as *extreme neighbors*.

In turn, the separating vector must have $\psi_l = \psi[\kappa^+, \kappa^-]' A_{\kappa^+} = 0$ for those extreme neighbors. By the requirement that $\max_l \psi_l' x \leq 0$ it must also be that for other vectors A_j—that do not define this cone but do define the cone of $\mathcal{B}'$ (for some other basis $\mathcal{B}'$ that is optimal at $\mathcal{K}$)—we have $\psi'[\kappa^+, \kappa^-]' A_j = 1$ if $j \in \cup_{\mathcal{B}' \neq \mathcal{B}} K^+(\mathcal{B}')$ and $\psi'[\kappa^+, \kappa^-]' A_j = -1$ if $j \in \cup_{\mathcal{B}' \neq \mathcal{B}} K^-(\mathcal{B}')$. □

Proof of Lemma 9. Note that, if $\mathbb{E}[R^0_{\mathcal{K}}] = \frac{1}{T}(I^0 + p_{\mathcal{R}} T - A_{\mathcal{K}} Q^0_{\mathcal{K}}) \in \mathcal{N}_{\mathcal{K}}(p)$, then since $R^0 = \frac{1}{T}(I^0 + Z^T_{\mathcal{R}} - A_{\mathcal{K}} Q^0_{\mathcal{K}})$ we have for $\bar{Z}^T \in \mathcal{A}^\epsilon$, $d(R^0_{\mathcal{K}}, \mathcal{N}_{\mathcal{K}}(p)) \leq 2\epsilon$. Let $\widehat{\mathcal{N}}_{\mathcal{K}}$ be the $\mathcal{K}$ centroid neighborhood under the empirical distribution. Suppose that $R^0_{\mathcal{K}} \in \widehat{\mathcal{N}}^+_{\mathcal{K}}$ where the $+$ stands for the "bigger" neighborhood. (The one obtained by taking convex combination of the full lines rather than mid-points). If, in addition $R^0_{\mathcal{K}} - A_{\mathcal{K}} \bar{Z}^0_{\mathcal{K}} \in \text{cone}(\mathcal{N}_{\mathcal{K}}(\mathcal{B}))$. Then, it follows from our main geometric lemma that $\mathcal{B}$ is the optimal basis and from here that the solution to the LP has precisely the properties we need. To conclude the lemma notice that if $\mathbb{E}[R^0_{\mathcal{K}}] \in \mathcal{N}_{\mathcal{K}}$ then, on $\mathcal{A}^\epsilon$, $R^0_{\mathcal{K}} \in \widehat{\mathcal{N}}^+_{\mathcal{K}}$. □

Proof of Lemma 10. Let (y, u, s) be the solution to $\text{LP}(R, D)$. Defining (y', u', s') to be the solution to $\text{LP}(R - s, D)$, we have that $y' = y$ and $u' = u$.

We will argue that we can choose ϵ' so that the following properties hold:

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

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

We would then have that $R - s = \bar{R} \pm \|A\|_\infty \epsilon''$ where $\bar{R} = r_{\mathcal{K}}(D) + A_{K^+} x_{K^+} - A_{K^-} x_{K^-}$ for some x with $x_j \in [0, D_j/2]$ for all $j \in K^+ \cap K^-$. Choosing ϵ' (and subsequently ϵ'') so that $\epsilon'' \leq \frac{\epsilon}{\sqrt{d}\|A\|_\infty}$, we will conclude that $d(R-s, \mathcal{N}_{\mathcal{K}}(\mathcal{B}, D)) \leq \sqrt{d} \times d_\infty(R-s, \mathcal{N}_{\mathcal{K}}(\mathcal{B}, D)) \leq \epsilon$.

Item (i): Because, $d_\infty(R-s, \mathcal{N}_{\mathcal{K}}(D)) \leq d(R-s, \mathcal{N}_{\mathcal{K}}(D)) \leq \epsilon'$, the Lipschitz continuity of LP (see [9]) implies that $y_j \geq D_j/2 - M\epsilon'$ for all $j \in K^-$ and some constant M that depends on (v, A, p). Taking $\epsilon' = \epsilon''/M$ covers item (i) of the requirements.

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

First case ($\max_l \psi_l'(R - r_{\mathcal{K}}(D)) \leq 0$): In this case, in particular, $R - r_{\mathcal{K}}(D) \in cone(\mathcal{K}, \mathcal{B})$. With ϵ' small, $d(R-s, \mathcal{N}_{\mathcal{K}}(D)) \leq \epsilon'$ implies that $R - s \in \mathcal{N}_{\mathcal{K}}^+(D)$. Per Lemma 5, we then have that $y_j = 0$ for all $j \notin K \cup K^+$ and $u_j = 0$ for all $j \in \mathcal{K} \backslash K^-$ as required.

Second case ($\max_l \psi_l'(R - r_{\mathcal{K}}(D)) > 0$): Take l such that $0 < \psi_l'(R - r_{\mathcal{K}}(D)) \leq \epsilon'$. Let κ^+, κ^- be such that $\psi_l = \psi[\kappa^+, \kappa^-]$. By Lemma 7, ψ is orthogonal to A_j for all $j \in \kappa^+ \cup \kappa^-$. Furthermore, defining $\delta_{lj} := \psi_l' A_j$, we have $\delta_{lj} > 0$ for $j \notin \kappa^+$ but $j \in \cup_{\bar{\mathcal{B}} \neq \mathcal{B}} K^+(\bar{\mathcal{B}})$ and $\delta_{lj} < 0$ for each $j \notin \kappa^-$ but $j \in \cup_{\bar{\mathcal{B}} \neq \mathcal{B}} K^-(\bar{\mathcal{B}})$. We thus have that $\psi'(R - r_{\mathcal{K}}(D)) = \psi'(\sum_{j \notin \mathcal{K} \cup \kappa^+} A_j y_j + \sum_{j \in \mathcal{K} \backslash \kappa^-} A_j u_j) = \sum_{j \notin \mathcal{K} \cup \kappa^+} \delta_{lj} y_j + \sum_{j \in \mathcal{K} \backslash \kappa^-} |\delta_{lj}| u_j$. The right-hand side is, by assumption, bounded by ϵ' and we get the desired result by taking again $\epsilon' \leq \epsilon''$ for a suitable constant M.

The proof of the lemma's second part is very similar and omitted. ☐

Proof of Lemma 11. First we argue $\theta = \theta_{\mathcal{K}}(y, D) \in \text{closure}(\mathcal{Y}(\mathcal{K}, D))$. Since y solves $\text{LP}(R, D)$ and $R \in \mathcal{N}_{\mathcal{K}^0}(D)$:

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

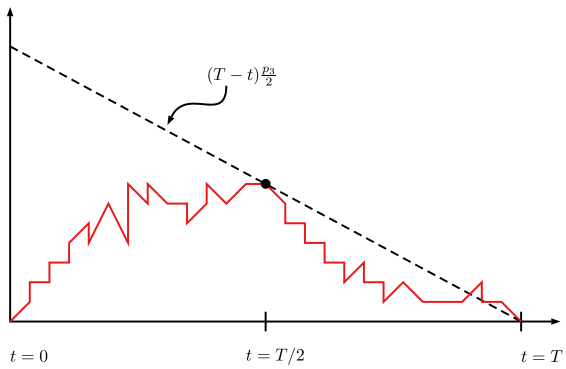
This implies $\theta_j \geq \frac{D_j}{2}$ for $j \in \mathcal{K}$ and $\theta_j \leq \frac{D_j}{2}$ for $j \in \mathcal{K}^c$. Finally, we claim that θ is optimal for the ratio $R^\theta := A\theta$. Indeed, if we take $\mathcal{B}$ as the basis that $\mathcal{K}$ and $\mathcal{K}^0$ share, then the support of θ is all basic variables and we have

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

which proves the optimality of θ at R^θ by Lemma 2 and concludes the first item.

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

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



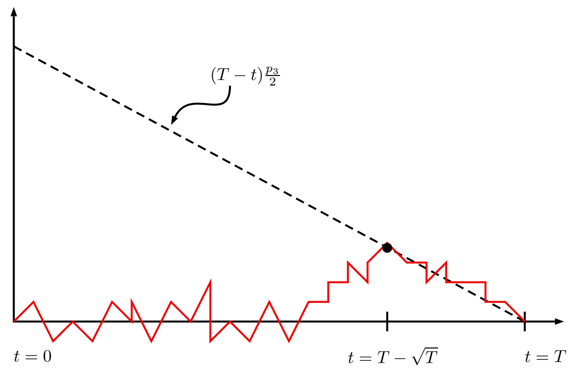


FIGURE 15. Paths on $\mathcal{M}=\{R-e_{\mathcal{K}}\bar{Z}^T_{\mathcal{K}}\leq 0\}$. The two paths illustrate unlikely events. On the left: because the random walk has variation $\mathcal{O}(\sqrt{t})$ it cannot hit the target line by a time t of the form $t=T-\Omega(T)$. On the right: the path could hit the target by a time of the form $t=T-\mathcal{O}(\sqrt{T})$. In this case, however, it does not have enough time to get back to 0 which it must on the event $\mathcal{M}$

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

LEMMA 14. *Consider m random walks $G^1,\dots,G^m$ of the form $G^i_t=G^i_0+\sum_{t=1}^n X^i_t$ where, for each i, the increments $X^i_t, t\in[T]$ are zero-mean i.i.d and bounded ($\mathbb{E}[X^i_t]=0$ and $\mathbb{P}[|X^i_t|\leq b]=1$ for some $b>0$) and independent of G^i_0 which satisfies $|G^i_0|\leq\epsilon/2$. Let*

$$\tau=\inf\{t\geq 0:(T-t)\epsilon\leq\max_i G^i_t\}\wedge T.$$

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

$$\mathbb{P}[T-\tau>t,\max_i G^i_T\leq 0]\leq\theta_1 e^{-\theta_2 t},$$

for constants $\theta_1,\theta_2>0$ that may depend on b,ϵ.

PROOF. Figure 15 is a graphic illustration of the event we want to bound for the case of $m=1$. Fix M and notice that, for $t\leq T-M$,

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

where the last inequality follows from standard results; see e.g. [10]. We conclude that

$$\mathbb{E}[\tau\mathbb{1}_{\{G_T\leq 0\}}]=\sum_{t=0}^{T-1}\mathbb{P}[\tau>t,G_T\leq 0]\geq T-M$$

for some constant M. □