

# Online Conversion with Switching Costs: *Robust and Learning-Augmented Algorithms*

ADAM LECHOWICZ, University of Massachusetts Amherst, USA

NICOLAS CHRISTIANSON, California Institute of Technology, USA

BO SUN, University of Waterloo, Canada

NOMAN BASHIR, Massachusetts Institute of Technology, USA

MOHAMMAD HAJIESMAILI, University of Massachusetts Amherst, USA

ADAM WIERMAN, California Institute of Technology, USA

PRASHANT SHENOY, University of Massachusetts Amherst, USA

We introduce and study online conversion with switching costs, a family of online problems that capture emerging problems at the intersection of energy and sustainability. In this problem, an online player attempts to purchase (alternatively, sell) fractional shares of an asset during a fixed time horizon with length  $T$ . At each time step, a cost function (alternatively, price function) is revealed, and the player must irrevocably decide an amount of asset to convert. The player also incurs a *switching cost* whenever their decision changes in consecutive time steps, i.e., when they increase or decrease their purchasing amount. We introduce competitive (robust) threshold-based algorithms for both the minimization and maximization variants of this problem, and show they are optimal among deterministic online algorithms. We then propose learning-augmented algorithms that take advantage of untrusted black-box advice (such as predictions from a machine learning model) to achieve significantly better average-case performance without sacrificing worst-case competitive guarantees. Finally, we empirically evaluate our proposed algorithms using a carbon-aware EV charging case study, showing that our algorithms substantially improve on baseline methods for this problem.

## ACM Reference Format:

Adam Lechowicz, Nicolas Christianson, Bo Sun, Noman Bashir, Mohammad Hajiesmaili, Adam Wierman, and Prashant Shenoy. 2024. Online Conversion with Switching Costs: *Robust and Learning-Augmented Algorithms*. In *Proceedings of the 2024 ACM SIGMETRICS / IFIP Performance Joint International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS / Performance '24)*, June 10–14, 2024, Venice, Italy. ACM, New York, NY, USA, 48 pages. <https://doi.org/XXXXXXXX.XXXXXXX>

## 1 INTRODUCTION

In classic online conversion, an online player buys/sells shares of an item over a fixed time horizon where future prices are unknown, and attempts to minimize their cost or maximize their profit. This paper introduces and studies *online conversion with switching costs* (OCS), a novel class of online problems motivated by emerging control problems in the design of sustainable systems. We consider both minimization (OCS-min) and maximization (OCS-max) variants of the problem. In OCS-min, an online player aims to purchase one item over a sequence of time-varying cost functions and decides the *fractional* amount of item to purchase in each round. The player must purchase the entire item before a deadline  $T$ , and they incur a *movement cost* whenever their decision changes, i.e., whenever they purchase different amounts of the item in consecutive time steps. From the player's perspective, the goal is to minimize their total cost, including the total purchasing cost and any movement cost incurred over the time horizon. In OCS-max, the setting is almost the same,

---

*SIGMETRICS / Performance '24, June 10–14, 2024, Venice, Italy*

© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.

This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in *Proceedings of the 2024 ACM SIGMETRICS / IFIP Performance Joint International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS / Performance '24)*, June 10–14, 2024, Venice, Italy, <https://doi.org/XXXXXXXX.XXXXXXX>.

except the player sells an item fractionally according to time-varying price functions, so the goal is to maximize their total profit, and any movement costs are subtracted from the revenue. In both settings, the cost/price functions are revealed one by one in an online manner, and the player makes an irrevocable decision at each time step without the knowledge of future cost/price functions.

Our motivation behind introducing OCS is an emerging class of *carbon-aware problems* such as carbon-aware electric vehicle (EV) charging [12] and carbon-aware compute shifting [1, 3, 22, 23, 46, 57], which have attracted significant attention in recent years. A common thread is the goal of reducing carbon emissions by temporally shifting flexible workloads to take advantage of times when low-carbon forms of electricity generation are available.

OCS builds on a long line of work in the literature of related online algorithms, which can roughly be classified into two problem types: *online search problems*, which feature deadline constraints but do not consider a penalty for switching decisions [16, 34, 41, 53, 55], and *online metric problems*, which feature switching costs but do not consider a long-term deadline constraint [6, 8–11, 13, 14, 26, 47]. We briefly review the most relevant problems to our work here. In the online search literature, OCS is most closely aligned with the *one-way trading problem* (OWT) introduced by El-Yaniv et al. [16], in which an online player sells fractional shares of an asset over a sequence of time-varying price functions to maximize their profit. In the online metric algorithms literature, OCS is also closely aligned with the problem of *convex function chasing* (CFC) introduced by Friedman and Linial [19], where an online player makes online decisions  $x_t$  in a normed vector space  $(X, \|\cdot\|)$  over a sequence of time-varying cost functions in order to minimize their total cost, which includes the value of the cost function plus a penalty for switching decisions.

Despite extensive prior literature on the above two categories of problems, there is limited work that includes both long-term constraints (as in OWT) and switching costs (as in CFC). To our knowledge, the only prior work is the online pause and resume (OPR) problem [27], which combines switching costs and a long-term deadline constraint. Specifically, OPR is a generalization of  $k$ -search, where the online player’s overarching objective is to accept the  $k$  lowest prices over a time-varying sequence, and the decision space is *binary* (the player either accepts or rejects a price). Similarly to OCS, OPR adds a penalty for switching the decision over consecutive slots. The binary design space in OPR makes it natural to extend the prior single-threshold-based algorithms for  $k$ -search [34]. Therefore, the authors in [27] develop a *double threshold* algorithm for OPR and show that the proposed algorithms can achieve the best competitive ratios possible. In OCS, however, the intuitive idea of the double threshold design fails because the design space of the algorithm is continuous. Therefore, this work aims to answer the following question:

*Is it possible to design competitive algorithms for OCS that simultaneously balance switching costs and deadline constraints in continuous decision spaces?*

Although the previous literature in both OWT and CFC primarily focuses on providing algorithms with near-optimal adversarial competitive guarantees, moving beyond worst-case analysis has received considerable attention in recent years. In particular, because algorithms designed to optimize the worst case may be overly cautious or pessimistic, a line of work focused on using machine-learned advice has emerged. These *learning-augmented algorithms* [35, 45] have, in many settings, been able to take advantage of untrusted machine-learned advice to improve the average-case performance (*consistency*) while still maintaining strong adversarial guarantees when the advice is bad (*robustness*). Recent works proposed learning-augmented algorithms for both OWT [53] and CFC [13, 14]; however, incorporating machine-learned predictions into algorithm design for OCS is uniquely challenging since even the choice of advice model is not clear for OCS. The prior literature for online algorithms with advice for online search and metrics problems uses two different advice models. In Sun et al. [53], based on an advice model for predicting an *input* parameter (i.e., the

*maximum price*), the authors present learning-augmented algorithms for the one-way trading problem that achieves the optimal consistency-robustness trade-off. For metrics problems and, specifically, convex function chasing, Christianson et al. [14] give a learning-augmented algorithm that uses black-box advice of the optimal *decision* at each time step and achieves an optimal consistency-robustness trade-off. Given these challenges, a natural open question is:

*Can predictions be effectively integrated into an online algorithm for OCS to provably improve the average case performance of competitive algorithms while maintaining their worst-case guarantees?*

## 1.1 Contributions and Techniques

We present online algorithms for both variants of OCS and show that they achieve the best possible competitive ratios. We then extend our algorithms to the learning-augmented setting, significantly improving practical performance when predictions are accurate while maintaining competitive guarantees. Our technical contributions are summarized in detail below.

*Competitive results.* The core algorithmic idea is based on a novel optimization-based framework called “Ramp-On Ramp-Off” (RORO), which we instantiate to develop competitive algorithms for OCS-min and OCS-max. RORO extends the framework of online threshold-based algorithms (OTA), a preeminent design paradigm for related online search problems [16, 27, 34, 53]. We give upper bounds on the competitive ratio of RORO in the minimization and maximization settings as a function of the problem’s parameters (see Theorems 3.2 and C.1). Furthermore, we show lower bounds for the competitive ratio of any deterministic online algorithm, which imply that RORO is optimal for this problem (see Theorems 3.3 and C.2).

To achieve the above competitive results for OCS, the RORO framework has to tackle substantially different technical challenges than prior work. Existing work leveraging the OTA framework either does not consider the presence of switching costs [16, 53, 55], or only considers a binary decision space [27]. For the online pause and resume (OPR) problem, authors in [27] show a “double threshold” algorithm achieving the optimal competitive ratio. However, the double threshold technique cannot be generalized to OCS. That is, each of the two defined threshold families corresponds to a single case for the previous decision made by the algorithm – making the decision space continuous in  $[0, 1]$  breaks this relationship between decisions and the threshold. Thus, the primary technical challenge in applying OTA design to OCS is deriving a technique that can specify continuous decisions while incorporating “awareness” of the switching cost. The key enabling result for the RORO framework is a *dynamic threshold* approach (in Section 3.1). Instead of the double threshold idea, the RORO framework uses a single threshold function complemented by an adaptive mechanism that dynamically adjusts admission criteria based on the previous online decision. Compared to existing approaches, this framework is more general and can be applied to the continuous OCS setting. RORO also recovers the double threshold algorithm for OPR [27] as a special case.

*Learning-augmented results.* To go beyond worst-case analysis and attain better *average-case* performance for OCS, we then turn our attention to learning-augmented design. The existing learning-augmented results for related problems can be roughly split into two advice models. Firstly, the advice model in online search (OWT and  $k$ -search) leverages *input predictions*, which, e.g., predicts the “best price” of the instance for OWT. In contrast, prior results for CFC and MTS typically leverage “black-box”-type advice, where the online player receives a prediction of an *optimal decision* at each time step. In learning-augmented design for OCS, we consider both advice models. For the input advice model, we present an impossibility result: our lower bound result proves that “best price” predictions cannot achieve 1-consistency for any value of robustness in OCS. This result implies that switching costs complicates the problem beyond the expressiveness of simple predictions (see Theorem 4.1), and that an advice model providing more information is necessary. Leveraging proof

techniques from the advice complexity literature [4, 5, 20], we further show that any 1-consistent algorithm for OCS must use advice that grows linearly in the sequence length. This motivates our usage of an *untrusted black-box advice model* which predicts the online decision at each time step. Using such advice, we present RO-Advice, a meta-algorithm that uses RORO as a subroutine to achieve  $(1 + \epsilon)$ -consistency and bounded robustness (see Theorems 4.3 and C.3). We expect that our technique of using advice complexity to prove lower bounds for a prediction model will be broadly applicable to other online problems where the choice of a prediction model in learning augmentation is not obvious.

*Case study.* We evaluate the performance of the proposed algorithms using a case study on the carbon-aware electric vehicle (EV) charging problem. In this problem, there is an interruptible and deferrable workload in the form of a charging schedule for an EV. From the user, we receive a requested amount of charge and a deadline of  $T$ , representing the time the vehicle will be plugged into an adaptive charger. Such a charger can change its charging rate based on grid constraints or external signals [29]. Since frequent fluctuations in the charging rate are undesirable from a lithium-ion battery health perspective [59], the operator can define a *switching penalty* to incentivize a “smooth” charging rate. The objective is to minimize the charging carbon footprint while satisfying the charging demand and maintaining a smooth charging rate within the deadline limit. As renewable sources such as solar and wind make up greater fractions of power grids worldwide, their intermittency contributes to time-varying carbon intensities, which can be volatile [27, 36, 37], so future values are unknown. This problem could be exactly modeled by OCS.

In Section 5, we evaluate the performance of our robust RORO and learning-augmented RO-Advice algorithms through the EV charging scheduling case study. We utilize real-world data describing EV charging sessions from the Adaptive Charging Network data set [30], combined with carbon intensity traces for the California ISO [39], and simulated local solar generation computed using the National Solar Radiation Data Base [48]. Our experiments simulate different strategies to charge an EV to the requested capacity before a deadline while dealing with uncertain future carbon intensities. We show that our proposed robust algorithms outperform baseline methods and adapted algorithms for related problems such as OWT. Furthermore, we evaluate our learning-augmented RO-Advice algorithm when advice is provided by an “*off-the-shelf*” ML model providing time-series predictions of carbon intensity [36]. We show that the imperfect predictions provided by this model are powerful enough for RO-Advice to outperform existing algorithms significantly.

Lastly, we note that OCS also captures other interesting applications in the sustainability space. One such example is the shifting of carbon-aware workloads [1, 3, 46, 52, 57], where time-varying resources are allocated to an interruptible and deferrable workload so that total carbon emissions from the workload are minimized. Recent work [23, 49] has additionally considered scalability (i.e., when training an ML model, a parallelizable training job can take more resources to “run harder” when low-carbon energy is available), which is analogous to the continuous decision space we consider in OCS. Progress in this field builds on a long line of work considering the temporal shifting of workloads for more sustainable data centers, e.g., [21, 31–33]. Carbon awareness has also been explored in other contexts, such as HVAC in residential buildings, which considers, e.g., scheduling heating and cooling cycles to coincide with the availability of low carbon energy [7]. For many such applications, the OCS formulation is applicable.

## 2 PROBLEM FORMULATION AND PRELIMINARIES

We first introduce and formulate the OCS problem, then provide brief background on design paradigms which will be useful in the design of our proposed algorithms. Table A1 (in the Appendix) summarizes core OCS notations for convenience.

## 2.1 Problem Formulation

We formulate two variants of the Online Conversion (OCS) problem, primarily considering the minimization variant (OCS-min) in the main body, and deferring the maximization variant (OCS-max) to Appendix C.<sup>1</sup> In OCS-min, an online player must buy an asset with total size  $C$ , while minimizing their total cost. Without loss of generality and for notational simplicity, we assume that  $C = 1$ . At each time step  $t \in [T]$ , a *convex cost function*  $g_t(\cdot)$  arrives online. The player can buy  $x_t \in [0, d_t]$  amount of the asset at a cost of  $g_t(x_t)$ . Following convention, each cost function satisfies  $g_t(0) = 0$ ; i.e., if the player purchases nothing, they pay no cost, and  $g_t(x_t) \geq 0$  for any valid  $x_t$ . Furthermore,  $d_t \leq C$  is a *rate constraint*, which limits the amount of purchases at a given time step  $t$ .<sup>2</sup> Whenever the player's decision changes in consecutive time steps, they incur a *switching cost*, which is proportional to the "distance" between  $x_t$  and  $x_{t-1}$ . For one time step, this is formalized as  $\beta|x_t - x_{t-1}|$ . For modeling purposes, we assume  $x_0 = 0$  and  $x^{T+1} = 0$ , which jointly imply that the player must incur *some* switching costs to "turn on" and "turn off", respectively. Note that OPT must incur at least  $\frac{2}{T}\beta$  in switching cost<sup>3</sup> and that the maximum switching cost is  $2\beta$ . The switching cost parameter  $\beta$  can be considered a linear coefficient that charges the online player proportionally to the movement between consecutive time steps.

For simplicity in modeling, this formulation assumes that the cost to "turn on" is equivalent to the cost to "turn off", and that the magnitude of  $\beta$  is constant across all time steps. Our formulation and algorithms can be extended to cases where the switching cost is time-varying or asymmetric by setting  $\beta$  to be an upper bound on the actual switching cost (e.g.  $\beta|x_t - x_{t-1}| \geq s_t(x_t, x_{t-1})$  for all  $t \in [T]$ , where  $s_t(x, x')$  is the time-varying or asymmetric switching cost).

Importantly, OCS requires the online player to buy/sell the entire demand  $C$  before the end of the sequence (the "deadline"). If the player has bought  $w^{(t)}$  fraction of the utilization at time  $t$ , a *compulsory trade* begins at time step  $t$  as soon as  $\sum_{\tau=t+1}^T d_\tau < 1 - w^{(t)}$  (i.e. when the remaining purchase opportunities in future time steps are not enough to fulfill the demand). During this compulsory trade, a cost-agnostic algorithm takes over and trades maximally at each time step to meet the constraint – intuitively, this algorithm comes with no competitive guarantees. We assume that the rate constraints  $\{d_t\}_{t \in [T]}$  are known in advance, so the player can always identify when the compulsory trade should begin.<sup>4</sup>

To ensure that the problem is non-trivial, we generally assume that the compulsory trade does not make up a large fraction of the sequence. Concretely, the earliest time step  $j$  at which the compulsory trade begins (i.e. the first time step such that  $\sum_{\tau=j+1}^T d_\tau < 1$ ) is  $j \gg 1$ , which implies that  $T$  and the rate constraints  $\{d_t\}_{t \in [T]}$  are sized appropriately for the demand. This assumption is reasonable in practice, since if  $T$  is small or the rate constraints are small, the rate constraints will be binding or near-binding in each time step – i.e., the task's *temporal flexibility* will be low, so that even an omnipotent solution will be unable to take much advantage of ramping up and down to

<sup>1</sup>We use OCS whenever the context applies to both minimization (OCS-min) and maximization (OCS-max) variants of the problem; otherwise, we refer to the specific variant. The same notation style extends to our proposed algorithms (RORO and RO-Advice) throughout the paper.

<sup>2</sup>For motivating applications such as carbon-aware EV charging, this rate constraint primarily models the case where the requested capacity cannot be met in a single time step (e.g., an EV requests 20 kWh of charge from a 10kW source).

<sup>3</sup>This follows by observing that the sequence consists of  $T$  cost functions. If each  $g_t$  is identical, OPT can minimize their switching cost by choosing to purchase  $1/T$  of the asset at every possible time step. OPT then incurs one switching cost to turn on  $\beta|1/T - x_0| = \beta/T$ , and one switching cost to turn off  $\beta|x_{T+1} - 1/T| = \beta/T$ , for a total of  $\frac{2}{T}\beta$ .

<sup>4</sup>We note that if the rate constraints are not known in advance (i.e., revealed online), the formulation still holds as long as the player receives a signal near the end of the sequence indicating that they must begin the compulsory trade.

improve its performance. In summary, the offline version of OCS-min can be formalized as follows:

$$\min_{\{x_t\}_{t \in [T]}} \underbrace{\sum_{t=1}^T g_t(x_t)}_{\text{Purchasing cost}} + \underbrace{\sum_{t=1}^{T+1} \beta |x_t - x_{t-1}|}_{\text{Switching cost}}, \text{ s.t., } \underbrace{\sum_{t=1}^T x_t = 1}_{\text{Deadline constraint}}, \underbrace{x_t \in [0, d_t] \forall t \in [T]}_{\text{Rate constraint}}. \quad (1)$$

Our focus is on the online version of OCS, where the player must make irrevocable decisions  $x_t$  at each time step without the knowledge of future inputs. The most important unknowns are the cost/price functions  $g_t(\cdot)$ , which are revealed online.

*Assumptions and additional notation.* For OCS-min, we assume that cost functions  $\{g_t(\cdot)\}_{t \in [T]}$  have a bounded derivative, i.e.  $L \leq dg_t/dx_t \leq U$ , where  $L$  and  $U$  are known positive constants. Furthermore, we assume that all cost functions  $g_t(\cdot)$  are *convex* – this assumption is important as a way to model *diminishing returns*, and is empirically valid for the applications of interest.

The switching cost coefficient  $\beta$  is assumed to be known to the player, and is bounded within an interval ( $\beta \in (0, U-L/2)$ ). If  $\beta = 0$ , the problem is equivalent to one-way trading, and if  $\beta$  is “too large” (i.e.,  $> U-L/2$ ), we can show that any competitive algorithm should *only* consider the switching cost.<sup>5</sup>

In applications, the deadline  $T$  is generally known in advance, although our formulation and algorithms do not require this to be true. The most important technicality that changes when  $T$  is unknown is the compulsory trading at the end of the period, which is used to ensure that the entire asset is purchased/sold before the deadline. If  $T$  is unknown, the player must receive a signal to indicate that the deadline is close and compulsory trading should begin.

*Competitive analysis.* Our goal is to design an online algorithm that maintains a small *competitive ratio* [6, 38], i.e., performs nearly as well as the offline optimal solution. Given an online algorithm ALG, an offline optimal solver OPT, and a valid input sequence  $\mathcal{I}$ , we denote the optimal cost by  $\text{OPT}(\mathcal{I})$ , and  $\text{ALG}(\mathcal{I})$  is the cost of the solution obtained by running the online algorithm over this input. Then if  $\Omega$  is the set of all feasible input instances, the *competitive ratio* for a minimization problem is defined as:  $\text{CR}(\text{ALG}) = \max_{\mathcal{I} \in \Omega} \text{ALG}(\mathcal{I})/\text{OPT}(\mathcal{I})$ . Under these definitions, the competitive ratio is always greater than or equal to one, and a *lower* competitive ratio implies that the online algorithm is guaranteed to be *closer* to the offline optimal solution.

In the emerging literature on learning-augmented algorithms, the competitive ratio is interpreted via the notions of *consistency* and *robustness*, introduced in [35, 45]. Formally, let  $\text{ALG}(\mathcal{I}, \varepsilon)$  denote the cost of a learning-augmented online algorithm on input sequence  $\mathcal{I}$  when provided predictions with error factor  $\varepsilon$ . Then the notion of consistency describes the competitive ratio of a learning-augmented algorithm when the predictions provided to it are exactly correct, i.e.  $\max_{\mathcal{I} \in \Omega} \text{ALG}(\mathcal{I}, 0)/\text{OPT}(\mathcal{I})$ . Conversely, robustness captures the competitive ratio when predictions are adversarially incorrect, i.e.  $\max_{\mathcal{I} \in \Omega} \text{ALG}(\mathcal{I}, \mathbf{E})/\text{OPT}(\mathcal{I})$ , where  $\mathbf{E}$  is a maximum error factor (or  $\infty$ ).

## 2.2 Background: Threshold-based Algorithms for Online Problems

Existing work on related problems such as online pause and resume [27], online search [16, 28, 34], one-way trading [53, 55], and online knapsack [54, 58, 60] often use of a technique known as *threshold-based algorithm design*. Threshold-based algorithms are a family of online algorithms where a thoughtfully-designed *threshold function* is used to guide an online decision at each time step. In broad terms, this threshold establishes a “minimal acceptable quality” that an incoming

<sup>5</sup>As brief justification for the bounds on  $\beta$ , consider the following solutions for OCS-min. In OCS-min, a feasible solution can have objective value  $L + 2\beta$ . Note that if  $\beta > U-L/2$ ,  $L + 2\beta > U$ , and we argue that the incurred switching cost is more important than the cost values accepted.

input must meet to be accepted by the algorithm. The threshold values are deliberately chosen to ensure that an algorithm consistently accepting/rejecting inputs based on this (potentially changing) “standard” at each step is guaranteed to obtain a certain competitive ratio.

This algorithmic framework has seen success in the related problems mentioned above; many of the threshold-based online algorithms presented in these works are optimal in the sense that the threshold function guarantees a competitive ratio that matches the information-theoretic lower bound for the online problem. Threshold-based design has not historically been applied to problems with switching costs such as OCS. In [27], the authors present the first (to our knowledge) threshold-based algorithm for a problem with switching costs, namely the online pause and resume problem. Although these results and the general framework do not directly translate to the OCS setting, we make use of ideas from these existing works. We briefly detail the most relevant results from prior work before discussing how this framework generalizes to OCS in Section 3.

*One-way trading.* In the canonical one-way trading problem, a sequence of time-varying prices is revealed to an online player one-by-one [16]. The player’s objective is to maximize their profit from selling an asset. When each price  $c_t$  arrives, the player must immediately decide a fraction of their asset  $x_t \in [0, 1]$  to sell at the current price, and they must sell the entire asset by the end of the sequence. A generalization of the problem considers price functions  $g_t(\cdot)$ , which must be *concave*, but are not necessarily linear. Assuming bounded derivative for the price functions, i.e.,  $L \leq dg_t/dx \leq U$ , Sun et al. [55] show a deterministic threshold-based algorithm which achieves the optimal competitive ratio of  $1 + W((U/L - 1)/e)$ , where  $W(\cdot)$  is the Lambert  $W$  function.

*Pause and resume (and  $k$ -search).* In recent work, Lechowicz et al. [27] presents the online pause and resume problem (OPR), which can be summarized as  $k$ -min/ $k$ -max search [34] with an added switching cost. In both OPR and  $k$ -search, a sequence of time-varying prices is revealed to an online player one by one. When each price  $c_t$  arrives, the player must immediately decide whether to reject or accept the price (i.e.,  $x_t \in \{0, 1\}$ ). The player’s objective in both problems is to find the  $k$  lowest (conversely, highest) prices in the sequence to minimize their cost (conversely, maximize their profit), and they must accept  $k$  prices before the end of the sequence. In OPR, the player additionally pays a switching cost  $\beta$  whenever their decision changes in consecutive time steps. Assuming bounded prices in the range  $[L, U]$ , [27] shows deterministic *double threshold* algorithms for OPR, where each threshold separately corresponds to each of the possible prior decisions  $x_{t-1} \in \{0, 1\}$ . These algorithms achieve the optimal competitive ratio for both variants of OPR.

### 3 COMPETITIVE ALGORITHMS

We first introduce a general algorithmic framework that we term Ramp-On, Ramp-Off (RORO). Then, we demonstrate the application of RORO in the context of the pause and resume problem for simplicity before applying it to the OCS-min problem (and OCS-max in Appendix C.2).

#### 3.1 The online Ramp-On, Ramp-Off (RORO) framework

This section presents a general online optimization framework called Ramp-On, Ramp-Off (RORO) (see Algorithm 1). In the RORO framework, whenever an input arrives online, we solve two *pseudo-cost minimization problems*, with a restricted decision space in each. We denote these problems as the **ramping-on** and **ramping-off** problems. In this framework, we employ the *pseudo-cost*, a notion inherited from the online search literature [55]. At a high level, the pseudo-cost generalizes the traditional concept of threshold-based algorithms, where, recall, a threshold value/function establishes the minimum/maximum acceptable inputs for the competitive algorithm to accept, given the current state. Then, at each time step, when input arrives, the pseudo-cost of a particular decision gives the actual cost of the decision at the current time step, *minus* a threshold value term

**Algorithm 1** Online Ramp-On, Ramp-Off (RORO) framework

- 
- 1: **input:** ramping-on problem  $\text{RAMPON}(\cdot)$ , ramping-off problem  $\text{RAMPOFF}(\cdot)$ ,
  - 2: pseudo-cost function  $\text{PCOST}(\cdot)$
  - 3: **initialization:** initial decision  $x_0 = 0$ , initial utilization  $w^{(0)} = 0$ ;
  - 4: **while** cost/price function  $g_t(\cdot)$  is revealed and  $w^{(t-1)} < 1$  **do**
  - 5:     solve the **(ramping-on problem)** to obtain decision  $x_t^+$  and its pseudo cost  $r_t^+$ ,
 
$$x_t^+ = \text{RAMPON}(g_t(\cdot), x_{t-1}), \quad (2)$$

$$r_t^+ = \text{PCOST}(g_t(\cdot), x_t^+, x_{t-1}). \quad (3)$$
  - 6:     solve the **(ramping-off problem)** to obtain decision  $x_t^-$  and its pseudo cost  $r_t^-$ ,
 
$$x_t^- = \text{RAMPOFF}(g_t(\cdot), x_{t-1}), \quad (4)$$

$$r_t^- = \text{PCOST}(g_t(\cdot), x_t^-, x_{t-1}). \quad (5)$$
  - 7:     **if**  $r_t^+ \leq r_t^-$  **then** set  $x_t = x_t^+$  **else** set  $x_t = x_t^-$ ;
  - 8:     update the utilization  $w^{(t)} = w^{(t-1)} + x_t$ ;
- 

describing the maximum/minimum acceptable input that would justify the decision. Intuitively, the pseudo-cost serves as a mechanism preventing an algorithm from “waiting too long” to accept any prices. For example, in OCS, naively minimizing the cost function  $g_t$  at each time step results in decisions  $x_t = 0$  for all  $t \in [T]$ , which incur no cost. However, this sequence of decisions is undesirable since the final compulsory trade needed to satisfy the deadline constraint may force the player to pay a large cost. Thus, the pseudo-cost “balances” between the two extremes of buying immediately and waiting indefinitely by enforcing the acceptance of inputs that are “good enough”.

The namesakes of the RORO framework – the *ramping-on* and *ramping-off* problems – find the *best decision* in either of two restricted decision spaces that minimize this pseudo-cost. In the ramping-on problem, the decision space is restricted to online decisions where  $x_t$  increases or stays the same, i.e.,  $x_t \geq x_{t-1}$ . In the ramping-off problem, the considered decisions are all  $\leq x_{t-1}$  (i.e., the online decision decreases or stays the same). The framework then chooses whichever of the two resulting decisions yields a better pseudo-cost. In what follows, we discuss these as two distinct problems for instructive value and ease of presentation.

*Intuition and a motivating example.* To provide intuition for the design of the RORO framework, we show that the double threshold algorithm shown for online pause and resume in [27] is a special case of this framework. We focus on the minimization version of pause and resume (OPR-min), summarized in Section 2.2. At a high level, OPR-min includes many of the same components as OCS-min, with a few key differences. Most notably, the decision space in OPR-min is binary, i.e.,  $x_t \in \{0, 1\}$ , while decisions in OCS are continuous ( $x_t \in [0, d_t]$ ). Furthermore, the online player in OPR-min is restricted to accept at most  $1/k$  of their total demand in a single time step (recall that  $k$  prices must be accepted before the deadline), while the online player in OCS can choose to purchase their entire demand  $C$  in a single time step (if rate constraint  $d_t$  permits).

Lechowicz et al. [27] present a *double threshold* algorithm for OPR-min which specifies two families of  $k$  threshold values  $\{u_i\}_{i \in [k]}$  and  $\{\ell_i\}_{i \in [k]}$ . The algorithm chooses which family to use based on the previous online decision  $x_{t-1}$ . Then if  $i - 1$  prices have been accepted so far, the  $i$ th price accepted will be the first price which is at most  $\ell_i$  if  $x_{t-1} = 0$ , or at most  $u_i$  if  $x_{t-1} = 1$ .

In [27], the authors present a *double threshold* algorithm for OPR-min. At a high level, the idea behind this algorithm is to change the criteria for accepting a price based on the switching cost that will be incurred with respect to the previous online decision. Intuitively, when a price arrives and a previously inactive player considers whether to accept it, they must add the extra switching



cost of  $\beta$  they will incur to “switch on”. Conversely, if the player has accepted the previous price, once a new price arrives, they should discount the price by  $\beta$ , since *not* accepting this new price will force them to “switch off” and incur an extra switching cost of  $\beta$ .

This idea is formalized in the specification of two threshold families  $\{u_i\}_{i \in [k]}$  and  $\{\ell_i\}_{i \in [k]}$ , each representing  $k$  “minimum acceptable prices”. The algorithm chooses which family to use based on the previous online decision  $x_{t-1}$  as described above. Then if  $i - 1$  prices have been accepted so far, the  $i$ th price accepted will be the first price which is at most  $\ell_i$  if  $x_{t-1} = 0$ , or at most  $u_i$  if  $x_{t-1} = 1$ .

We will illustrate the special case of RORO by first considering the *pseudo-cost* in the context of OPR-min, where inputs are scalar values arriving online, and decisions are binary. When a price  $c_t$  arrives, the algorithm can consider accepting the price ( $x_t = 1$ ) or rejecting it ( $x_t = 0$ ). The pseudo-cost is the actual cost  $c_t x_t + \beta |x_t - x_{t-1}|$ , minus a threshold value determining the maximum acceptable price. Then, if  $x_t = 1$  and the pseudo-cost  $< 0$ , we know that the threshold value is *greater than* the actual cost and the price should be accepted. In the following, we formalize the above intuition by defining the ramping-on problem RAMPON, the ramping-off problem RAMP OFF, and the pseudo-cost function PCOST to obtain a RORO instantiation, which we show to be equivalent to the double threshold algorithm.

**CLAIM 3.1.** *The RORO framework is equivalent to the double threshold algorithm for OPR-min in [27] when the framework is instantiated as follows. Let  $\phi$  be a family of  $k$  thresholds defined as:  $\phi_i = u_i - \beta = \ell_i + \beta, \forall i \in [k]$ , where  $\{u_i\}_{i \in [k]}$  and  $\{\ell_i\}_{i \in [k]}$  are the double thresholds for OPR-min in [27]. Then, define the ramping-on problem, ramping-off problem, and the pseudo-cost function as:*

$$\text{RAMPON}(c_t, x_{t-1}) = \arg \min_{x_t \in \{0,1\}, x_t \geq x_{t-1}} c_t x_t + \beta(x_t - x_{t-1}) - \phi_i x_t, \quad (6)$$

$$\text{RAMP OFF}(c_t, x_{t-1}) = \arg \min_{x_t \in \{0,1\}, x_t \leq x_{t-1}} c_t x_t + \beta(x_{t-1} - x_t) - \phi_i x_t, \quad (7)$$

$$\text{PCOST}(c_t, x_t, x_{t-1}) = c_t x_t + \beta |x_t - x_{t-1}| - \phi_i x_t. \quad (8)$$

We summarize this instantiation in pseudocode form in the appendix, in Algorithm A1. We proceed to justify the equivalence of the RORO instantiation described in Claim 3.1 against the double threshold algorithm shown for OPR-min.

If  $x_{t-1} = 0$ , the ramping-off solution is restricted to  $x_t^- = x_{t-1} = 0$ , and thus the online solution is dominated by the ramping-on solution, i.e.,  $x_t = x_t^+$ . The ramping-on solution is exactly

$$x_t^+ = \begin{cases} 1 & \text{if } c_t + \beta \leq \phi_i \\ 0 & \text{if } c_t + \beta > \phi_i \end{cases}, \quad (9)$$

which is the same as the solution obtained by the double threshold algorithm, i.e.,  $x_t = 1$  if  $c_t \leq \phi_i - \beta = \ell_i$ ; otherwise,  $x_t = 0$ . If  $x_{t-1} = 1$ , the ramping-on is restricted to  $x_t^+ = 1 = x_{t-1}$ , and thus the online solution is dominated by the ramping-off, i.e.,  $x_t = x_t^-$ . The ramping-off solution is

$$x_t^- = \begin{cases} 1 & \text{if } c_t - \phi_i \leq \beta \\ 0 & \text{if } c_t - \phi_i > \beta \end{cases}, \quad (10)$$

which is the same as the solution obtained by the double threshold algorithm, i.e.,  $x_t = 1$  if  $c_t \leq \phi_i + \beta = u_i$  and otherwise  $x_t = 0$ . In what follows, we describe how RORO solves OCS.

### 3.2 Solving OCS-min using the RORO framework

We now move to apply the RORO framework and propose an algorithm for solving OCS-min. Prior work [53] has observed that the limiting case of  $k$ -search (i.e.  $k \rightarrow \infty$ ) can be recast as the one-way trading problem if the formulation is generalized in a few key areas. Leveraging the RORO framework,

we apply a similar technique to recast the limiting case of OPR-min as OCS-min and then generalize the threshold function in Claim 3.1.

Recall that in OPR, the player must accept  $k$  prices total, and at the time that they consider accepting the  $i$ th price, they have already accepted  $i - 1$  prices previously (where  $i \in [k]$ ). In the continuous setting, we let  $w^{(t)} \in [0, 1]$  denote the *utilization* at time  $t$ , i.e.,  $w^{(t)}$  gives the fraction of the total asset which has been purchased up to  $t$ . Equivalently, in the binary setting,  $w^{(t)}$  is defined as  $i^{-1/k}$ . In the binary setting of OPR, the player can only accept a single price at a time, thus satisfying  $1/k$  of their task. This is insufficient for OCS, so we relax this constraint and allow the player to purchase more than  $1/k$  amount in a single time step. In OPR, this is analogous to allowing to purchase more than 1 “item” in a single time step. Finally, instead of a single price  $c_t \in [L, U]$  arriving online, a convex cost function  $g_t(\cdot)$  arrives online and the decision  $x_t$  is in  $[0, d_t]$ .

With these transformations, we can initialize a *dynamic threshold*  $\phi(w)$  for OCS-min that will inform the design of our RORO instantiation for OCS-min. In the following, we slightly abuse notation and use  $w$  instead of  $w^{(t)}$ , wherever dropping  $t$  is appropriate for the context.

*Definition 3.1 (Threshold function  $\phi$  for OCS-min).* For any utilization  $w \in [0, 1]$ ,  $\phi$  is defined as:

$$\phi(w) = U - \beta + (U/\alpha - U + 2\beta) \exp(w/\alpha), \quad (11)$$

where  $\alpha$  is the competitive ratio and is defined in (15).

This threshold function follows by taking the limit (as  $k \rightarrow \infty$ ) of the closed form threshold  $\{\phi_i\}_{i \in [k]}$  defined in Claim 3.1, and observing that  $w = i^{-1/k}$ . We note that  $\phi(1) - \beta = L$ , which corresponds to the case of the final threshold value in the OPR instantiation, i.e.,  $\phi_{k+1} = L + \beta$ . Given  $\phi(\cdot)$ , we fully describe our instantiation of RORO for OCS-min (RORO-min) by defining the ramping-on problem, the ramping-off problem, and the pseudo-cost function in the following.

*Definition 3.2 (RORO instantiation for OCS-min (RORO-min)).* The RORO framework solves OCS-min when instantiated as follows. Let  $\phi(\cdot) : [0, 1] \rightarrow [L, U]$  be the dynamic threshold defined in 3.1. Then, define the ramping-on problem, ramping-off problem, and the pseudo-cost function as:

$$\text{RAMPON}(g_t(\cdot), w^{(t-1)}, x_{t-1}) = \arg \min_{x \in [x_{t-1}, \min(1-w^{(t-1)}, d_t)]} g_t(x) + \beta(x - x_{t-1}) - \int_{w^{(t-1)}}^{w^{(t-1)}+x} \phi(u) du, \quad (12)$$

$$\text{RAMPOFF}(g_t(\cdot), w^{(t-1)}, x_{t-1}) = \arg \min_{x \in [0, \min(x_{t-1}, d_t)]} g_t(x) + \beta(x_{t-1} - x) - \int_{w^{(t-1)}}^{w^{(t-1)}+x} \phi(u) du, \quad (13)$$

$$\text{PCOST}(g_t(\cdot), w^{(t-1)}, x_t, x_{t-1}) = g_t(x_t) + \beta(x_t - x_{t-1}) - \int_{w^{(t-1)}}^{w^{(t-1)}+x_t} \phi(u) du. \quad (14)$$

The optimizations defined above are inserted into the pseudocode defined in Algorithm 1 to create an instance that solves OCS-min (RORO-min). At a glance, it is not obvious that the ramping-on and ramping-off problems given by (13) and (14) can be solved efficiently. In Appendix B.1, we show that both are convex minimization problems, which can be solved efficiently (e.g., using iterative methods). In the following theorem, we state the competitive result for RORO-min. We give a proof sketch of the result and discuss its significance here, and relegate the full proof to Appendix B.2.

**THEOREM 3.2.** *Algorithm 1 for OCS-min (RORO-min) is  $\alpha$ -competitive when the threshold function is given by  $\phi(w)$  (from Def. 3.1), where  $\alpha$  is the solution to  $\frac{U-L-2\beta}{U/\alpha-U-2\beta} = \exp(1/\alpha)$  and is given by*

$$\alpha := \left[ W \left( \left( \frac{2\beta}{U} + \frac{L}{U} - 1 \right) e^{2\beta/U-1} - \frac{2\beta}{U} + 1 \right) \right]^{-1}. \quad (15)$$

In the above,  $W(\cdot)$  is the Lambert  $W$  function, defined as the inverse of  $f(x) = xe^x$ .

PROOF SKETCH FOR THEOREM 3.2. To show this result, we give three lemmas characterizing the optimal solution and the solution obtained by RORO-min. Note that the solution obtained by RORO-min is feasible, as the constraints in the ramping-on problem combined with the compulsory trade always enforce that  $\sum_{t=1}^T x_t = 1$ .

We first consider the case without rate constraints, and derive a lower bound on the cost incurred by OPT. We denote the final utilization on an arbitrary instance  $\mathcal{I} \in \Omega$  attained by RORO-min (before the compulsory trade) by  $w^{(j)}$ . Leveraging the definition of the pseudo-cost and an analysis of the ramping-on solution, we show by contradiction that the *best marginal cost* (i.e., the smallest cost function which arrives at any point in the sequence) is lower bounded by  $\phi(w^{(j)}) - \beta$ . This gives a corresponding lower bound on  $\text{OPT}(\mathcal{I})$  (Lemma B.2).

We continue by showing an upper bound on the cost incurred by RORO-min under the same sequence (denoted by  $\text{RORO-min}(\mathcal{I})$ ). By combining a worst-case analysis of the actual switching cost paid by RORO-min with the definition of the pseudo-cost and its relationship to the cost functions accepted by RORO-min, we show that  $\text{RORO-min}(\mathcal{I})$  is upper bounded by  $\int_0^{w^{(t)}} \phi(u) du + \beta w^{(t)} + (1 - w^{(t)})U$  (Lemma B.3).

Finally, we consider the setting with rate constraints. Through a careful analysis of the different costs paid by OPT and ALG in this setting, we show that the rate constraints do not degrade the worst-case competitive ratio (Lemma B.4). A final analysis of the interaction between these three lemmas yields the result, showing that RORO-min's competitive ratio is bounded by  $\alpha$ . The full proof can be found in Appendix B.2.  $\square$

Theorem 3.2 gives an upper bound on the competitive ratio of RORO-min, showing how the parameters of the problem (e.g.,  $L$ ,  $U$ , and  $\beta$ ) filter through and appear in the competitive results. A reasonable question to ask is whether any other algorithm for OCS-min can achieve a better competitive ratio. In the following result, we highlight that no improvement is possible, i.e., that RORO-min achieves the optimal competitive ratio across all deterministic online algorithms for OCS-min. We give a proof sketch of the result and discuss its significance here, and relegate the full proof to Appendix B.3.

THEOREM 3.3. *No deterministic online algorithm for OCS-min can achieve a competitive ratio better than  $\alpha$ , as defined in (15).*

PROOF SKETCH FOR THEOREM 3.3. We first define a special class of worst-case  $x$ -decreasing instances  $\mathcal{I}_x : x \in [L, U]$  for which the offline optimal solution is  $\text{OPT}(\mathcal{I}_x) = x$ . An  $x$ -decreasing instance (defined formally in Def. B.3) is a sequence of linear cost functions of the form  $g(y) = \kappa y$ , where  $\kappa$  is a linear coefficient. Letting  $n, m \in \mathbb{N}$  be sufficiently large, and  $\delta := (U-L)/n$ , there are  $n_x := 2 \cdot \lceil (x-L)/\delta \rceil + 1$  alternating batches of cost functions. The  $i^{\text{th}}$  batch ( $i \in [n_x - 2]$ ) contains a single cost function with  $\kappa = U - (\lceil i/2 \rceil)\delta$  if  $i$  is even, and these “decreasing” batches are interrupted by blocks of  $m$  cost functions with  $\kappa = U$  (i.e., when  $i$  is odd). See below for an illustration.

$$\underbrace{\{Uy, \dots, Uy\}}_m, \underbrace{(U - \delta)y, Uy, \dots, Uy}_m, \underbrace{(U - 2\delta)y, \dots, Uy, \dots, Uy}_m, \underbrace{(x + \varepsilon)y, \dots, (x + \varepsilon)y}_m, \underbrace{Uy, \dots, Uy}_m$$

As  $n \rightarrow \infty$ , the alternating single cost functions in an  $x$ -decreasing sequence continuously decrease down to  $x$ , and each of these “good cost functions” is interrupted by a section of worst-case  $U$  cost functions. Note that the last few cost functions in an  $x$ -decreasing instance are always  $U$ . We also note that  $\mathcal{I}_U$  is simply a stream of  $m$  cost functions  $U$ .

By defining the special class of instances, we solve a key technical challenge – the actions of any deterministic online algorithm ALG on these instances can then be fully described by an arbitrary

conversion function  $h(x) : [L, U] \rightarrow [0, 1]$ , where  $h(x)$  denotes the fractional amount purchased (before the final compulsory trade) on instance  $\mathcal{I}_x$ .

Using this conversion function, we can construct several necessary conditions that must be satisfied by any  $\alpha^*$ -competitive algorithm, where  $\alpha^*$  is the target (best) competitive ratio. First, we can observe that the cost of the arbitrary algorithm is  $\text{ALG}(\mathcal{I}_x) = h(U/\alpha^*)U/\alpha^* - \int_{U/\alpha^*}^x udh(u) + 2\beta h(x) + (1 - h(x))U$ , where  $udh(u)$  is the cost of purchasing  $dh(u)$  utilization at price  $u$ . Based on this, since ALG's cost must be  $\alpha^*$ -competitive against OPT, we can derive a differential inequality that lower bounds the conversion function  $h(x)$ .

We can then use Grönwall's Inequality [40, Theorem 1, p. 356] to obtain a *necessary condition*, where  $\alpha^*$  and  $h(x)$  must jointly satisfy the following:  $\alpha^* \ln(L + 2\beta - U) - \alpha^* \ln(U/\alpha^* + 2\beta - U) \leq h(L) = 1$ . The optimal  $\alpha^*$  is obtained when the inequality is binding, and the result follows.  $\square$

By combining Theorems 3.2 and 3.3, we conclude that RORO-min is optimal for OCS-min. We note that the definition of  $\alpha$  gives that when the switching cost  $\beta = 0$ , our RORO algorithm recovers the optimal competitive ratio for the minimization variant of one-way trading, which is  $\alpha \sim [W((L/U - 1)e^{-1}) + 1]^{-1}$ , as given by [34, 53]. It is known that  $W(x) \sim \ln(x)$  as  $x \rightarrow \infty$  [24, 50]. As  $\beta$  grows, the competitive ratio  $\alpha$  intuitively degrades linearly in the magnitude of  $\beta$ , although the assumed upper bound on  $\beta$  prevents the competitive ratio from becoming unbounded.

## 4 LEARNING-AUGMENTED ALGORITHMS

In this section, we explore how ML advice can help break past the pessimistic worst-case bounds of competitive algorithms for OCS. We propose a meta-algorithm, RO-Advice, which integrates advice to significantly improve empirical performance and theoretical bounds when the advice is of high quality. Along the way, we discuss and consider two distinct models of advice employed in prior work on learning-augmented algorithms. We present lower bounds that show that “simple predictions” are not powerful enough to allow learning-augmented algorithms to achieve 1-consistency in OCS (Theorem 4.1). In addition, we show that the amount of advice necessary to break through this barrier is at least linear in the length of the instance (Theorem 4.2). These results, combined with the availability of applicable machine learning techniques, heavily influence our choice of advice model, which is detailed below.

### 4.1 Advice Model

In [28, 53], the authors present learning-augmented algorithms for one-way trading and  $k$ -search, respectively. Both works consider a “best price” prediction model, where the algorithm receives a single prediction  $\hat{P} \in [L, U]$  of the best-quality input in the upcoming sequence. For instance,  $\hat{P}$  in the OCS problem would predict the lowest cost function in the minimization setting or the largest price function in the maximization setting. On the other hand, learning-augmented work for convex function chasing (CFC) [13], smoothed online convex optimization (SOCO) [47], and metrical task systems (MTS) [2, 14] use *black-box advice*, which directly predicts the optimal online decision at each time step. In the OCS problem, this advice would involve predicting  $x_t$  for each time step  $t \in [T]$ . We first consider leveraging “best price” predictions since our RORO algorithm design is closer in lineage to the robust one-way trading algorithm used as a baseline in [53]. Notably, however, as we show in the following theorem, these relatively simple predictions are not expressive enough in the OCS setting for learning-augmented algorithms to obtain performance comparable to the offline optimal. The full proof of Theorem 4.1 is deferred to Appendix B.4.

**THEOREM 4.1.** *Any learning-augmented algorithm for OCS which uses a prediction of the “best price” in a sequence cannot achieve 1-consistency, regardless of robustness.*

**PROOF SKETCH FOR THEOREM 4.1.** Assume that an algorithm for OCS receives a prediction of the best price  $\hat{P}$ , and that this prediction is correct. While this information is useful because it enables the algorithm to wait to accept cost/price functions matching  $\hat{P}$ , it does not convey any information about the switching cost, which is a crucial part of making optimal decisions.

We construct two toy sequences to illustrate how the best price prediction fails in the presence of switching costs. The first sequence consists of one cost/price function  $g(y) = \hat{P}y$ , followed by any number of significantly worse cost/price functions. The second sequence consists of one cost/price function  $g(y) = \hat{P}y$ , followed by some number of significantly worse cost/price functions, and ending with a block of  $m$  cost/price functions  $\{g(y) = \hat{P}y\} \times m$ .

From the viewpoint of the first time step, both sequences look identical to the learning-augmented algorithm, but the switching cost incurred by the optimal solution is significantly different ( $2\beta$  for the first sequence,  $2\beta/m$  for the second sequence). Therefore, the online algorithm must incur a switching cost of  $2\beta$  (otherwise, it would not be 1-consistent for the first sequence), while the optimal solution in the second sequence incurs a switching cost of  $2\beta/m$ , which approaches 0 as  $m \rightarrow \infty$ . Thus, a learning-augmented algorithm with simple predictions cannot achieve a consistency better than  $\text{ALG}/\text{OPT} \geq L+2\beta/L$  in the minimization or  $\text{OPT}/\text{ALG} \geq U/U-2\beta$  in the maximization setting.  $\square$

The result in Theorem 4.1 motivates considering the question: *how much advice is needed to achieve 1-consistency?* Answering this question provides insight about which advice model is the “right choice” for OCS – although we may not require 1-consistency in all cases, it is valuable to have a parameterized learning-augmented algorithm which *can* achieve 1 consistency if predictions are consistently accurate. In what follows, we leverage techniques from the advice complexity literature [4, 5, 17, 20] to prove a lower bound on the amount of advice necessary for 1-consistency in OCS. We give a proof sketch of the result and discuss its significance here, relegating the full proof to Appendix B.5.

**THEOREM 4.2.** *For any instance of OCS with sequence length  $T$ , at least  $\Omega(T)$  bits of advice are necessary to achieve 1-consistency.*

**PROOF SKETCH FOR THEOREM 4.2.** We first characterize the *number of unique solutions* for OCS sequences with length  $T$ . We leverage the structure of the problem, observing first that any feasible solution satisfies  $\sum_{i=1}^T x_i = 1, 0 \leq x_i \leq 1 \forall i \in [T]$ . We use a careful discretization of the decision space  $[0, 1]$  to the set  $\{1/T, 2/T \dots, 1\}$  to capture two extreme cases for the possible solutions. One feasible solution (rate constraints permitting) is to distribute purchasing decisions across all  $T$  time steps, setting  $x_i = 1/T$  at each step. The other extreme is to group all purchasing decisions into a single time step  $m$ , such that  $x_m = 1$  and  $x_i = 0 : \forall i \neq m$ . This careful discretization allows us to derive a lower bound on the number of unique solutions using combinatorial principles.

Using this lower bound, we can then show that the growth in the number of unique solutions as  $T$  grows is fast enough to overwhelm an advice model that is sublinear in  $T$ . The intuition behind the proof hinges on the fact that as  $T$  grows, any general advice model which uses  $A_T$  bits of advice (where  $A_T = o(T)$ ) can distinguish between  $2^{A_T}$  inputs based on the advice. As  $T$  grows, the number of unique solutions will quickly exceed  $2^{A_T}$ , and by the pigeonhole principle, there must be at least two unique solutions that map to the same advice string. The full proof is given in Appendix B.5.  $\square$

The above result concludes that any algorithm with advice for OCS must read  $\Omega(T)$  advice bits to achieve a competitive ratio of 1; in the *untrusted advice* setting, this corresponds to the advice

necessary for 1-consistency. Consequently, it naturally motivates the usage of black-box advice (as in [2, 13, 14, 47]) which predicts the online decision at each time step. Below we formally define the black-box advice model we use for OCS, and some assumptions.

*Definition 4.1 (Black-box advice model for OCS).* A learning-augmented algorithm ALG for OCS receives advice of the form  $\{\hat{x}_t\}_{t \in [T]}$ . If the advice is correct, a naïve algorithm ADV which chooses  $\hat{x}_t$  at each time step obtains the optimal cost/profit, i.e., for any valid OCS instance  $\mathcal{I}$ ,  $\text{ADV}(\mathcal{I}) = \text{OPT}(\mathcal{I})$ .

We assume that regardless of whether the black-box advice is correct, it is always *feasible*, i.e., it satisfies  $\sum_{t=1}^T \hat{x}_t = 1$  and  $\hat{x}_t \in [0, d_t] \forall t \in [T]$ .

We note here that it is not obvious that this black-box advice model is something that can be directly learned by a machine learning model in practice, and the assumption that the black-box advice is always feasible can similarly be questionable. In practice, however, the model’s generality means that it can accommodate any “black-box predictor”, which can actually be a layer of abstraction over several components, such as an ML model plus a post-processing pipeline to enforce feasibility. In Section 5, we show that this “black-box” advice model is still useful in the real world by leveraging an off-the-shelf ML model for our empirical experiments.

Before we describe our main learning-augmented results, we note that the prediction lower bound shown in Theorem 4.2 can be improved if aspects of the OCS problem are relaxed or simplified. Specifically, if rate constraints are removed (i.e.,  $d_t = 1 \forall t \in [T]$ ) and the cost/price functions are all linear, the problem’s advice complexity seems to decrease substantially. We provide some more discussion on this point in Appendix B.6, although we do not explore this dynamic in detail since these simplifications are not relevant to the real-world applications we consider. It would be very interesting to explore this trade-off between the problem’s complexity and the amount of advice required in future work, particularly since OCS seems to lie on a boundary between the simple advice used in the online search and the decision advice used in the online metric literature.

#### 4.2 R0-Advice-min: robustly incorporating black-box decision advice

We present a learning-augmented algorithm for OCS called Ramp-On-Advice (RO-Advice, pseudocode in Algorithm 2). This algorithm combines the robust decision of RORO ( $\tilde{x}_t$ ) at each time step with the predicted optimal solution ( $\hat{x}_t$ ) obtained from the black-box advice. RO-Advice takes a hyperparameter  $\epsilon$ , which parameterizes a trade-off between consistency and robustness (i.e., as  $\epsilon$  approaches 0, consistency improves and robustness degrades). Based on  $\epsilon$ , RO-Advice sets a combination factor  $\lambda \in [0, 1]$ , which determines the decision portion that it should take from each subroutine (i.e.,  $\lambda$  from the black-box advice and  $(1 - \lambda)$  from the robust solution). We note that the exact definitions of  $\epsilon$  and  $\lambda$  are different in the minimization and maximization settings, and we defer discussion of the maximization setting to Appendix C.3.

---

**Algorithm 2** Ramp-On-Advice meta-algorithm for *learning-augmented* OCS (RO-Advice)

---

- 1: **input:** The black-box advice  $\{\hat{x}_t\}_{t \in [T]}$ , competitive robust advice  $\{\tilde{x}_t\}_{t \in [T]}$  (e.g., given by RORO), combination factor  $\lambda$  (see Definitions 4.2 and C.3).
  - 2: **while** cost/price function  $g_t(\cdot)$  is revealed and  $w^{(t-1)} < 1$  **do**
  - 3:     obtain untrusted advice  $\hat{x}_t$ ;
  - 4:     obtain robust advice  $\tilde{x}_t$ ;
  - 5:     set the online decision as  $x_t = \lambda \hat{x}_t + (1 - \lambda) \tilde{x}_t$ ;
  - 6:     update the utilization  $w^{(t)} = w^{(t-1)} + x_t$ ;
- 

We start by describing our instantiation of RO-Advice for OCS-min (RO-Advice-min), which includes definitions of  $\epsilon$ ,  $\lambda$ , and the robust instantiation of RORO-min as a subroutine.

*Definition 4.2 (RO-Advice instantiation for OCS-min (RO-Advice-min)).* Let  $\epsilon \in [0, \alpha - 1]$ , where  $\alpha$  is the robust competitive ratio defined in (15). RO-Advice-min then sets a combination factor  $\lambda := (\alpha - 1 - \epsilon) \cdot \frac{1}{\alpha - 1}$ , which is in  $[0, 1]$ . The robust advice  $\{\tilde{x}_t\}_{\forall t \in [T]}$  is given by the RORO instantiation for OCS-min (RORO-min), given in Definition 3.2.

In the following theorem, we state the consistency and robustness bounds for the RO-Advice-min meta-algorithm. For brevity, we relegate the full proof to Appendix B.7.

**THEOREM 4.3.** *Given a parameter  $\epsilon \in [0, \alpha - 1]$ , where  $\alpha$  is defined as in (15), RO-Advice-min is  $(1 + \epsilon)$ -consistent and  $\left(\frac{(U+2\beta)(\alpha-1-\epsilon)+\alpha\epsilon}{(\alpha-1)}\right)$ -robust for OCS-min.*

Note that as  $\epsilon \rightarrow \alpha - 1$ , both the consistency and robustness bounds approach  $\alpha$ , which is the optimal competitive bound shown for the setting without advice. When  $\epsilon \rightarrow 0$ , the robustness bound degrades to  $U+2\beta/L$ , but the consistency bound implies that perfect advice allows RO-Advice-min to obtain the optimal solution. Although the result in Theorem 3.3 implies that the consistency-robustness trade off of RO-Advice-min is optimal as  $\epsilon \rightarrow \alpha - 1$ , it is not clear whether the trade off is *Pareto-optimal* for any setting of  $\epsilon$ . It would be very interesting to consider a corresponding lower bound on consistency-robustness in this setting with predictions.

It is notable that a relatively simple linear combination of the untrusted advice and the robust online decision simultaneously achieves bounded consistency and robustness, particularly since prior works have shown that such a technique fails to achieve bounded robustness for other online problems. In OCS, our learning-augmented technique can take advantage of structure in the problem provided by the deadline constraint – namely, since the deadline constraint *must* be satisfied by any valid solution, we can derive upper and lower bounds on the cost of any valid solution. Since any cost function must have derivative  $\geq L$ , the cost of any solution is lower bounded by  $L$ . Conversely, since any cost function must have derivative  $\leq U$  and the worst-case switching cost is exactly  $2\beta$ , the cost of any solution is upper bounded by  $U + 2\beta$  (see Section 2.1). These bounds are exactly reflected in the worst-case robustness ratio.

## 5 CASE STUDY: CARBON-AWARE EV CHARGING

This section uses carbon-aware electric vehicle (EV) charging as a case study and compares our proposed algorithms for OCS-min against existing algorithms from the literature.

### 5.1 Experimental Setup

We consider a carbon-aware EV charging system that supplies a requested amount of electricity to a customer's EV while minimizing the total CO<sub>2</sub> emissions of the energy used. For simplicity, we consider a single EV charger connected to the electric grid and a local solar installation, such as a solar parking canopy [43]. The simulated charger is an *adaptive* Level 2 AC charger with a peak charging rate of 19kW [44] that can modulate its charging rate between 0kW and 19kW. The charger first uses the electricity from the local solar and resorts to the electric grid to meet the deficit. We assume that local solar electricity has zero operational carbon intensity, while grid-supplied electricity has a *time-varying carbon intensity* based on the mix of generation sources for the region [51].

We construct instances of OCS-min as follows: Each EV arrives with a charging demand (in kWh), a connection time, and a disconnection time. The objective of the charging system is to selectively ramp the charging rate up or down to meet the full demand before the deadline (disconnection time) while minimizing total carbon emissions. The switching cost smooths the charging rate and penalizes large fluctuations, which accelerate lithium-ion battery degradation [59] (see Fig. 1 for

“smoothing” effect illustration). The rate constraints  $\{d_t\}_{t \in [T]}$  are set to 1 if the total charging demand is less than 19 kWh, and otherwise it is set to reflect the fraction of the total demand that our simulated Level 2 AC charger can meet during one hour (one time slot) of charging.<sup>6</sup>

*EV charging data.* We use the Adaptive Charging Network Dataset (ACN-Data) [30], which provides real EV charging data for  $\sim 16,000$  EV charging sessions, including arrival time, departure time, and energy demand (in kWh) for each session. We focus on charging sessions that are 5 hours or longer, since the carbon intensity of the electric grid does not change over shorter time periods, and carbon savings can only be realized when charging has temporal flexibility.

*Carbon data traces.* We use historical grid carbon intensity data for the California ISO (CAISO), obtained via Electricity Maps [39], which provides hourly *average carbon emissions* of the CAISO grid, expressed in grams of CO<sub>2</sub> equivalent per kilowatt-hour (gCO<sub>2</sub>eq/kWh).

*Carbon forecasts.* To evaluate our learning-augmented RO-Advice algorithm, we use an “*off-the-shelf*” predictive model called CarbonCast [36]. CarbonCast is a CNN-LSTM-based ML architecture that accurately forecasts the hourly carbon intensity of the grid in up to 96-hour windows. CarbonCast provides pre-trained models and training data for 13 regions, including CAISO.

*Solar data traces.* To simulate the impact of solar installations of varying sizes, we obtain historical solar data from the National Solar Radiation Database (NSRDB) [48]. The NSRDB provides satellite measurements of hourly solar irradiation, including diffuse horizontal irradiance (DHI) and direct normal irradiance (DNI), and solar zenith angle, at a 2 km spatial resolution. Using this data, we can approximate the historical generation of a simulated solar system in a particular location using the following method [15]. We use the optimal panel tilt angle for the location to compute the *global tilted irradiance* (GTI) on the panel’s surface based on the DHI, DNI, and solar elevation angle as  $GTI = DNI \cdot \sin(\text{solar elevation angle}) + DHI$ . The DC rating of a solar cell is customarily given under “ideal” irradiance conditions of 1000 W/m<sup>2</sup>. Given the DC rating of an entire solar system, we can compute the estimated generation as follows:

$$\text{Generation (in kW)} = \text{DC rating} \cdot \frac{\text{GTI}}{1000\text{W/m}^2} \cdot \text{AC Inverter Efficiency} \cdot (1 - \text{System Losses}) .$$

We assume that the AC inverter is 95% efficient and the panel-to-output system losses are  $\sim 14\%$  – default values from the PVWatts solar calculator [42]. In Appendix Fig. A1, we visualize a two-week sample of the grid’s carbon intensity and simulated solar generation for a 15 kW solar system.

*Parameter settings.* Most algorithms we test accept  $L$  and  $U$  as parameters for their threshold functions. To set these parameters, we examine a month’s worth of grid carbon intensity values leading up to the connection time for a given charging session, setting  $L$  and  $U$  to be the minimum and maximum observed carbon intensity values over the past month, respectively.

First, to gauge the impact of local solar generation, we tested differently sized simulated local solar systems with DC-rated nameplate capacities of 0, 5, 10, and 15 kW. A rating of 0 implies that the charger can only draw current from the grid, while a larger DC rating represents greater impact from local solar (during periods when the sun is shining). The switching cost is fixed at  $\beta = 20$ . Second, to gauge the impact of the switching cost magnitude, we run a set of experiments to test different values for  $\beta$  in the range  $[0, 40]$ . Local solar generation is fixed at 0 kW. Note that when  $\beta = 0$ , OCS reduces exactly to the one-way trading problem.

Finally, to measure the impact of prediction error, we generate a spectrum of simulated black-box advice, ranging from perfect advice to completely adversarial advice, and test RO-Advice for different values of  $\epsilon$ . Unlike other experiments, in these experiments, RO-Advice does not use

<sup>6</sup>Our setup assumes a slow AC charger, and the EV can always accept its full charging rate. Considering a DC fast charger, which limits the charging rate when the battery is above a certain charge level [25], is outside the scope of this work.



CarbonCast predictions. Instead, we define a simulated advice vector  $\hat{\mathbf{x}} = (1 - \zeta)\mathbf{x}^* + \zeta\check{\mathbf{x}}$ , where  $\mathbf{x}^*$  represents the decisions made by an optimal solution,  $\check{\mathbf{x}}$  represents the decisions made by a solution which maximizes the objective (rather than minimizing it), and  $\zeta$  is an *adversarial factor* in  $[0, 1]$ . When  $\zeta = 0$ , the simulated prediction is exactly correct ( $\hat{\mathbf{x}} = \mathbf{x}^*$ ), and when  $\zeta = 1$ , the simulated prediction is fully adversarial ( $\hat{\mathbf{x}} = \check{\mathbf{x}}$ ). We note that  $\check{\mathbf{x}}$  is adversarial from the perspective of the objective, although it is still a feasible solution for the problem.

*Benchmark algorithms.* To evaluate the performance of the tested algorithms, we compute the offline optimal solution for each charging session (instance) using a numerical solver [56]. This allows us to report the empirical competitive ratio for each algorithm on each charging session. We compare RORO and RO-Advice against three benchmark algorithms, which are not developed for solving OCS-min. Still, we construct these algorithms to represent simple adaptations of existing ideas for adjacent problems and evaluate how our algorithms measure up.

The first benchmark algorithm is a *carbon-agnostic* approach, which starts charging the EV at the fastest rate as soon as it is plugged in. This carbon-agnostic approach simulates the behavior of a traditional, non-adaptive charger. We also compare RORO against carbon-aware *switching-cost-agnostic* algorithms adapted for OCS. We have two algorithms of this type, each drawing from existing online search methods in the literature. The first algorithm is a *simple threshold algorithm*, which uses the  $\sqrt{UL}$  threshold value first presented for the online search [16]. In our experiments, this algorithm chooses to charge the EV (at the maximum rate) whenever the carbon intensity is less than or equal to  $\sqrt{UL}$ . This approach simulates the behavior of the simplest carbon-aware charger, which is not adaptive but can switch on or off in response to a carbon signal. The other switching-cost-agnostic algorithm tested is the threshold-based one-way trading algorithm, OWT, shown by [55] and described in Section 2.2. This algorithm chooses to ramp to a specific charging rate  $x_t$  based on a threshold function  $\varphi$ . Note that when  $\beta = 0$ , RORO reduces exactly to this one-way trading algorithm. In all experiments, we report the empirical competitive ratio of different algorithms, i.e., closer to 1 is a better competitive ratio, and 1 is the best.

## 5.2 Experimental Results

Table 1 summarizes the most notable overall results, aggregating over the first two experiments. We observe that the average empirical competitive ratio of RORO-min is a 57.3% improvement on the carbon agnostic method, a 52.4% improvement on the simple threshold algorithm, and a 12.1% improvement on the one-way trading algorithm. Across the first two experiments, RO-Advice with  $\epsilon \sim 1.86$  and predictions from CarbonCast shows a 33.4% improvement on RORO-min, a 69.2% improvement on the carbon-agnostic method, a 66.4% improvement on the simple threshold algorithm, and a 41.4% improvement on the one-way trading algorithm.

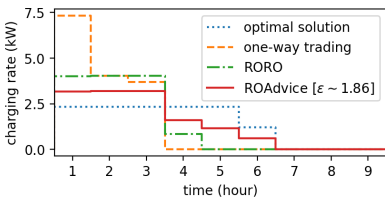


Fig. 1. The decisions made by different algorithms and the optimal solution during a 9-hour EV charging session with 10kW of local solar capacity and  $\beta = 30$ . The total energy requested is  $\sim 12.9$  kWh.

improvement against...	simple threshold		one-way trading		RORO	
	avg.	95 <sup>th</sup> %ile	avg.	95 <sup>th</sup> %ile	avg.	95 <sup>th</sup> %ile
RORO	52.4%	54.1%	12.1%	3.6%	-	-
RO-Advice	66.4%	73.3%	41.4%	46.2%	33.4%	44.9%

Table 1. Overall performance improvement of RORO-min and RO-Advice-min ( $\epsilon \sim 1.86$ ) over baseline threshold  $\sqrt{UL}$  algorithm and one-way trading algorithm. The empirical competitive ratio measures performance, and we report the average and the 95<sup>th</sup> percentile (i.e., worst-case) improvement for all experiments.

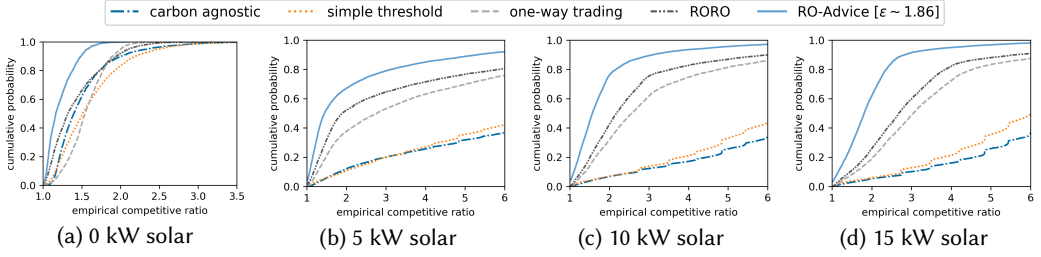


Fig. 2. Cumulative distribution functions (CDFs) of empirical competitive ratios for all tested algorithms, in experiments testing the impact of *differently-sized local solar generation*. Each simulated solar system is described using the DC rating. Switching cost is fixed  $\beta = 20$ , and RO-Advice uses CarbonCast predictions.

To further motivate the inclusion of the switching cost  $\beta$  in this EV charging application, in Fig. 1, we plot a step function of the decisions made by several algorithms on a representative 9-hour charging session from Feb 2020. We note that the switching cost  $\beta$  has a “smoothing” effect on the charging schedule generated by our algorithms, which are *switching-aware*. This is in contrast to the schedule generated by the one-way trading algorithm, which front loads most of the charging into the first hour of the charging session.

We now proceed to explain the details of our experiments. Fig. 2 shows a series of cumulative distribution function (CDF) plots. In this experiment, we test all algorithms for differently sized local solar installations, ranging from 0 kW (no solar) to 15 kW. The switching cost  $\beta$  is fixed to  $\beta = 20$ , and RO-Advice uses CarbonCast predictions with  $\epsilon \sim 1.86$  (this value of  $\epsilon$  gives  $\lambda = 0.5$ ). By testing different simulated solar system ratings, this experiment gauges how the makeup of the grid can affect performance (i.e., more or less renewable volatility). Unsurprisingly, RO-Advice performs well across the board, improving on the next closest algorithm RORO-min by 33.4% in the average case and 52.4% in the 95th percentile, respectively. RORO-min (our robust algorithm without predictions) compares favorably to the other algorithms, achieving an average empirical competitive ratio which is a 57.4% improvement on the carbon-agnostic method, a 52.4% improvement on the simple threshold algorithm, and a 12.1% improvement on the one-way trading algorithm.

We further note that in Fig. 2(a) (experiment with no local solar generation), while RORO-min outperforms the one-way trading algorithm on average, the tail of the CDF implies that RORO-min’s empirical competitive ratio is worse on roughly  $\sim 20\%$  of the instances. Since the one-way trading algorithm is *switching-agnostic*, there are cases where RORO-min is too conservative for a given sequence – in other words, RORO-min avoids ramping up to avoid a large switching penalty and eventually is forced to fulfill the charge at the end of the sequence (a compulsory trade). This dynamic disappears when simulated local solar generation is available because RORO-min’s conservative design matches well with the intermittent presence of local solar.

In the second experiment, reported as a heat map of empirical competitive ratios in Fig. 3, we test our proposed algorithms against the *switching-agnostic* algorithms for different switching costs  $\beta$  in the range from 0 to 40 (see Appendix Fig. A2 for a more detailed version of this plot). By testing different values for  $\beta$ , this experiment tests how an increasing switching cost impacts the performance of RORO-min and RO-Advice-min with respect to other algorithms that do not consider the switching cost in their designs. We observe that as the switching cost  $\beta$  increases, the empirical performance of all algorithms degrades (indicated by the luminance of the plot). Notably, comparing RORO-min and RO-Advice-min against the simple threshold algorithm and the one-way trading algorithm, our proposed algorithms exhibit competitive ratios that degrade *slower* as a function of the increasing competitive ratio, which validates our theoretical results. RO-Advice outperforms across the board, improving on the next closest algorithm RORO-min by 14.3% in the average case and 27.9% in the 95th percentile, respectively. RORO-min also outperforms the other

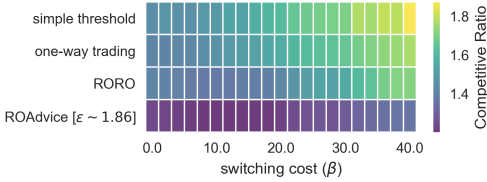


Fig. 3. Heat map of empirical competitive ratios for algorithms in a set of experiments testing the impact of different *switching costs*. Local solar generation is fixed to 0, and RO-Advice uses CarbonCast predictions with  $\epsilon \sim 1.86$ . Note that the simple threshold and one-way trading algorithms are *switching-oblivious*, while our proposed RORO and RO-Advice algorithms are aware of the switching cost.

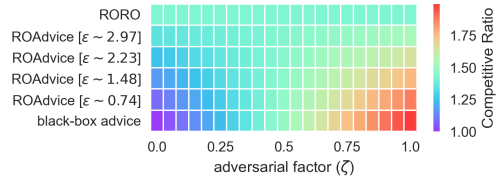


Fig. 4. Heat map of empirical competitive ratios for algorithms in a set of experiments testing the consistency and robustness of RO-Advice-min by simulating advice from perfect ( $\zeta = 0$ ) to adversarial ( $\zeta = 1$ ). No local solar, and four version of RO-Advice-min, with  $\epsilon \in [\sim 2.97, \sim 2.23, \sim 1.48, \sim 0.74]$ . The black-box advice algorithm naïvely plays the simulated advice at each time step.

tested algorithms, achieving an average improvement of 11.4% on the simple threshold algorithm and a 8.5% improvement on the one-way trading algorithm. Note that when  $\beta = 0$ , RORO-min is equivalent to the one-way trading algorithm, which the experimental results confirm.

Finally, in the third experiment, we rigorously test the consistency and robustness of RO-Advice-min and include RORO and the naïve black-box advice as baselines. We test four different variants of RO-Advice-min, parameterized by  $\epsilon \in [\sim 2.97, \sim 2.23, \sim 1.48, \sim 0.74]$  (these values of  $\epsilon$  correspond to  $\lambda$  parameters  $[0.2, 0.4, 0.6, 0.8]$ ). As described in Section 5.1, RO-Advice-min and the naïve black-box advice receive simulated advice as input, which is manipulated according to an adversarial factor  $\zeta$  (where  $\zeta = 0$  is perfect advice and  $\zeta = 1$  is fully adversarial advice). We test different values of  $\zeta$  in the range  $[0, 1]$  and visualize these results as a heat map in Fig. 4 (see Appendix Fig. A3 for a more detailed version). As discussed in the theoretical results, we observe opposing dynamics on either side of the heat map. When  $\zeta = 0$ , setting  $\epsilon \rightarrow 0$  yields increasingly better performance as the (accurate) advice is “trusted” more. On the other hand, when  $\zeta = 1$ , setting  $\epsilon \rightarrow 1$  can quickly increase the competitive ratio well beyond what is achieved by the robust RORO algorithm since trusting a prediction that is actually incorrect or even adversarial can significantly degrade performance. Importantly, Fig. 4 shows that, depending on the desired  $\epsilon$ , RO-Advice-min maintains a competitive ratio bounded away from the naïve worst-case regardless of the prediction quality, while significantly improving the average-case performance when predictions are correct or nearly correct, showing that it can achieve the “best of both worlds”.

## 6 CONCLUSION

Motivated by sustainability applications such as carbon-aware EV charging, we introduce and study online conversion with switching costs (OCS), which bridges gaps between several existing online problems. It is the first online optimization problem to simultaneously capture a continuous decision space, long-term constraints, and switching costs. Our main results introduce and analyze the RORO framework, which achieves the optimal deterministic competitive ratio for the problem. We also consider designing learning-augmented algorithms in this setting, first showing that existing advice models for related online search problems are insufficient to achieve improved consistency. We prove an advice complexity lower bound, which shows that advice growing linearly in the length of the sequence is necessary. This result motivates our design of learning-augmented algorithms that take advantage of untrusted black-box advice (such as ML-based predictions of decisions), which achieve bounded consistency and robustness. Finally, we empirically evaluate our proposed algorithms using a carbon-aware EV charging case study.

There are a number of interesting directions in which to continue the study of OCS. In particular, there are several open theoretical questions. Considering a target application of carbon-aware load shifting, some workloads can be shifted both spatially and temporally [3, 46, 52], which would require a generalization to a multidimensional setting where cost functions and decisions are in  $\mathbb{R}^d$  (e.g., each dimension corresponds to a location). Such a generalization is highly non-trivial, as there are no existing works that consider online search-type deadline constraints in a multidimensional setting, and the threshold-based algorithm design we generalize in this work has not historically been applied to multidimensional problems. As this paper considers a problem that bridges the gap between online search problems and a special case of an online metric problem (convex function chasing in  $\mathbb{R}$ ), any multidimensional generalization naturally prompts further questions about possible extensions of other online metric problems, such as metrical task systems (MTS). These questions are substantial additions that would be very interesting to explore in future work.

## ACKNOWLEDGMENTS

We thank our shepherd Arpan Mukhopadhyay and the anonymous SIGMETRICS / Performance reviewers for their valuable insight and feedback.

This research is supported by National Science Foundation grants CAREER-2045641, CNS-2102963, CNS-2106299, CNS-2146814, CNS-1518941, CPS-2136197, CPS-2136199, NGSDI-2105494, NGSDI-2105648, 1908298, 2020888, 2021693, 2045641, 2213636, and 2211888.

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Department of Energy Computational Science Graduate Fellowship under Award Number DE-SC0024386, and an NSF Graduate Research Fellowship (DGE-1745301).

## DISCLAIMERS

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

## REFERENCES

- [1] Bilge Acun, Benjamin Lee, Fiodar Kazhamiaka, Kiwan Maeng, Udit Gupta, Manoj Chakkaravarthy, David Brooks, and Carole-Jean Wu. 2023. Carbon Explorer: A Holistic Framework for Designing Carbon Aware Datacenters. In *Proceedings of the 28th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 2* (Vancouver, BC, Canada) (ASPLOS 2023). Association for Computing Machinery, New York, NY, USA, 118–132. <https://doi.org/10.1145/3575693.3575754>
- [2] Antonios Antoniadis, Christian Coester, Marek Elias, Adam Polak, and Bertrand Simon. 2020. Online Metric Algorithms with Untrusted Predictions. In *Proceedings of the 37th International Conference on Machine Learning*. PMLR, 345–355.
- [3] Noman Bashir, Tian Guo, Mohammad Hajiesmaili, David Irwin, Prashant Shenoy, Ramesh Sitaraman, Abel Souza, and Adam Wierman. 2021. Enabling Sustainable Clouds: The Case for Virtualizing the Energy System. In *Proceedings of the ACM Symposium on Cloud Computing* (Seattle, WA, USA) (SoCC '21). Association for Computing Machinery, New York, NY, USA, 350–358. <https://doi.org/10.1145/3472883.3487009>
- [4] Hans-Joachim Böckenhauer, Juraj Hromkovič, Dennis Komm, Sacha Krug, Jasmin Smula, and Andreas Sprock. 2014. The string guessing problem as a method to prove lower bounds on the advice complexity. *Theoretical Computer*

- Science* 554 (October 2014), 95–108. <https://doi.org/10.1016/j.tcs.2014.06.006>
- [5] Hans-Joachim Böckenhauer, Dennis Komm, Rastislav Kráľovič, Richard Kráľovič, and Tobias Mömke. 2009. On the advice complexity of online problems. In *Algorithms and Computation: 20th International Symposium, ISAAC 2009, Honolulu, Hawaii, USA, December 16–18, 2009. Proceedings 20*. Springer, 331–340.
- [6] Allan Borodin, Nathan Linial, and Michael E. Saks. 1992. An Optimal On-Line Algorithm for Metrical Task System. *J. ACM* 39, 4 (Oct 1992), 745–763. <https://doi.org/10.1145/146585.146588>
- [7] Phuthipong Bovornkeeratiroj, Noman Bashir, Vivek Deulkar, Bharathan Balaji, Prashant Shenoy, David Irwin, and Mohammad Hajiesmaili. 2023. Quantifying the Decarbonization Potential of Flexible Loads in Residential Buildings. In *Proceedings of the 10th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation (Istanbul, Turkey) (BuildSys '23)*. Association for Computing Machinery, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3563357.3564079>
- [8] Sébastien Bubeck, Christian Coester, and Yuval Rabani. 2023. The Randomized  $k$ -Server Conjecture Is False!. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing (STOC 2023) (Orlando, FL, USA) (STOC 2023)*. Association for Computing Machinery, New York, NY, USA, 581–594. <https://doi.org/10.1145/3564246.3585132>
- [9] Sébastien Bubeck, Michael B. Cohen, James R. Lee, and Yin Tat Lee. 2021. Metrical Task Systems on Trees via Mirror Descent and Unfair Gluing. *SIAM J. Comput.* 50, 3 (Jan. 2021), 909–923. <https://doi.org/10.1137/19M1237879>
- [10] Sébastien Bubeck, Bo'az Klartag, Yin Tat Lee, Yuanzhi Li, and Mark Sellke. 2019. Chasing Nested Convex Bodies Nearly Optimally. In *Proceedings of the 2020 ACM-SIAM Symposium on Discrete Algorithms (SODA)*. Society for Industrial and Applied Mathematics, 1496–1508. <https://doi.org/10.1137/1.9781611975994.91>
- [11] NiangJun Chen, Gautam Goel, and Adam Wierman. 2018. Smoothed Online Convex Optimization in High Dimensions via Online Balanced Descent. In *Proceedings of the 31st Conference On Learning Theory*. PMLR, 1574–1594.
- [12] Kai-Wen Cheng, Yuexin Bian, Yuanyuan Shi, and Yize Chen. 2022. Carbon-Aware EV Charging. In *2022 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*. 186–192. <https://doi.org/10.1109/SmartGridComm52983.2022.9960988>
- [13] Nicolas Christianson, Tinashe Handina, and Adam Wierman. 2022. Chasing Convex Bodies and Functions with Black-Box Advice. In *Proceedings of the 35th Conference on Learning Theory*, Vol. 178. PMLR, 867–908.
- [14] Nicolas Christianson, Junxuan Shen, and Adam Wierman. 2023. Optimal robustness-consistency tradeoffs for learning-augmented metrical task systems. In *International Conference on Artificial Intelligence and Statistics*.
- [15] Aron P Dobos. 2014. *PVWatts Version 5 Manual*. Technical Report. National Renewable Energy Lab.(NREL), Golden, CO (United States).
- [16] Ran El-Yaniv, Amos Fiat, Richard M. Karp, and G. Turpin. 2001. Optimal Search and One-Way Trading Online Algorithms. *Algorithmica* 30, 1 (May 2001), 101–139. <https://doi.org/10.1007/s00453-001-0003-0>
- [17] Yuval Emek, Pierre Fraigniaud, Amos Korman, and Adi Rosén. 2011. Online computation with advice. *Theoretical Computer Science* 412, 24 (2011), 2642–2656. <https://doi.org/10.1016/j.tcs.2010.08.007> Selected Papers from 36th International Colloquium on Automata, Languages and Programming (ICALP 2009).
- [18] Philippe Flajolet and Robert Sedgewick. 2009. *Analytic Combinatorics*. Cambridge University Press, Cambridge, England.
- [19] Joel Friedman and Nathan Linial. 1993. On convex body chasing. *Discrete & Computational Geometry* 9, 3 (March 1993), 293–321. <https://doi.org/10.1007/bf02189324>
- [20] Sushmita Gupta, Shahin Kamali, and Alejandro López-Ortiz. 2013. On advice complexity of the  $k$ -server problem under sparse metrics. In *International Colloquium on Structural Information and Communication Complexity*. Springer, 55–67.
- [21] Vani Gupta, Prashant Shenoy, and Ramesh K Sitaraman. 2019. Combining renewable solar and open air cooling for greening internet-scale distributed networks. In *Proceedings of the Tenth ACM International Conference on Future Energy Systems*. 303–314.
- [22] Walid A. Hanafy, Roozbeh Bostandoost, Noman Bashir, David Irwin, Mohammad Hajiesmaili, and Prashant Shenoy. 2023. The War of the Efficiencies: Understanding the Tension between Carbon and Energy Optimization. In *Proceedings of the 2nd Workshop on Sustainable Computer Systems*. ACM. <https://doi.org/10.1145/3604930.3605709>
- [23] Walid A. Hanafy, Qianlin Liang, Noman Bashir, David Irwin, and Prashant Shenoy. 2023. CarbonScaler: Leveraging Cloud Workload Elasticity for Optimizing Carbon-Efficiency. *Proceedings of the ACM on Measurement and Analysis of Computing Systems* 7, 3 (Dec 2023). arXiv:2302.08681 [cs.DC]
- [24] Abdolhossein Hoorfar and Mehdi Hassani. 2008. Inequalities on the Lambert W function and hyperpower function. *Journal of Inequalities in Pure and Applied Mathematics* 9, 51 (Jan. 2008). Issue 2.
- [25] ChargePoint Inc. 2019. Common DC Fast Charging Curves and How to Find Yours. <https://www.chargepoint.com/blog/common-dc-fast-charging-curves-and-how-find-yours>.
- [26] Elias Koutsoupias. 2009. The  $k$ -server problem. *Computer Science Review* 3, 2 (May 2009), 105–118. <https://doi.org/10.1016/j.cosrev.2009.04.002>

- [27] Adam Lechowicz, Nicolas Christianson, Jinhang Zuo, Noman Bashir, Mohammad Hajiesmaili, Adam Wierman, and Prashant Shenoy. 2023. The Online Pause and Resume Problem: Optimal Algorithms and An Application to Carbon-Aware Load Shifting. *Proceedings of the ACM on Measurement and Analysis of Computing Systems* 7, 3, Article 53 (Dec 2023), 36 pages. arXiv:2303.17551 [cs.DS]
- [28] Russell Lee, Bo Sun, John C. S. Lui, and Mohammad Hajiesmaili. 2022. Pareto-Optimal Learning-Augmented Algorithms for Online k-Search Problems. arXiv:2211.06567 <https://arxiv.org/abs/2211.06567>
- [29] Zachary J. Lee, George Lee, Ted Lee, Cheng Jin, Rand Lee, Zhi Low, Daniel Chang, Christine Ortega, and Steven H. Low. 2021. Adaptive Charging Networks: A Framework for Smart Electric Vehicle Charging. *IEEE Transactions on Smart Grid* 12, 5 (2021), 4339–4350. <https://doi.org/10.1109/TSG.2021.3074437>
- [30] Zachary J. Lee, Tongxin Li, and Steven H. Low. 2019. ACN-Data: Analysis and Applications of an Open EV Charging Dataset. In *Proceedings of the Tenth ACM International Conference on Future Energy Systems* (Phoenix, AZ, USA) (*e-Energy '19*). Association for Computing Machinery, New York, NY, USA, 139–149. <https://doi.org/10.1145/3307772.3328313>
- [31] Minghong Lin, Adam Wierman, Lachlan LH Andrew, and Eno Thereska. 2012. Dynamic right-sizing for power-proportional data centers. *IEEE/ACM Transactions on Networking* 21, 5 (2012), 1378–1391.
- [32] Zhenhua Liu, Yuan Chen, Cullen Bash, Adam Wierman, Daniel Gmach, Zhikui Wang, Manish Marwah, and Chris Hyser. 2012. Renewable and cooling aware workload management for sustainable data centers. In *Proceedings of the 12th ACM SIGMETRICS/PERFORMANCE joint international conference on Measurement and Modeling of Computer Systems*. 175–186.
- [33] Zhenhua Liu, Minghong Lin, Adam Wierman, Steven H Low, and Lachlan LH Andrew. 2011. Greening geographical load balancing. *ACM SIGMETRICS Performance Evaluation Review* 39, 1 (2011), 193–204.
- [34] Julian Lorenz, Konstantinos Panagiotou, and Angelika Steger. 2008. Optimal Algorithms for k-Search with Application in Option Pricing. *Algorithmica* 55, 2 (Aug. 2008), 311–328. <https://doi.org/10.1007/s00453-008-9217-8>
- [35] Thodoris Lykouris and Sergei Vassilvitskii. 2018. Competitive Caching with Machine Learned Advice. In *Proceedings of the 35th International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 80)*, Jennifer Dy and Andreas Krause (Eds.). PMLR, 3296–3305. <https://proceedings.mlr.press/v80/lykouris18a.html>
- [36] Diptyaroop Maji, Prashant Shenoy, and Ramesh K. Sitaraman. 2022. CarbonCast: Multi-Day Forecasting of Grid Carbon Intensity. In *Proceedings of the 9th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation* (Boston, Massachusetts) (*BuildSys '22*). Association for Computing Machinery, New York, NY, USA, 198–207. <https://doi.org/10.1145/3563357.3564079>
- [37] Diptyaroop Maji, Ramesh K. Sitaraman, and Prashant Shenoy. 2022. DACF: Day-Ahead Carbon Intensity Forecasting of Power Grids Using Machine Learning. In *Proceedings of the Thirteenth ACM International Conference on Future Energy Systems (Virtual Event) (e-Energy '22)*. Association for Computing Machinery, New York, NY, USA, 188–192. <https://doi.org/10.1145/3538637.3538849>
- [38] Mark Manasse, Lyle McGeoch, and Daniel Sleator. 1988. Competitive Algorithms for On-Line Problems. In *Proceedings of the Twentieth Annual ACM Symposium on Theory of Computing* (Chicago, Illinois, USA) (*STOC '88*). Association for Computing Machinery, New York, NY, USA, 322–333. <https://doi.org/10.1145/62212.62243>
- [39] Electricity Maps. 2023. Electricity Map. <https://www.electricitymap.org/map>.
- [40] Dragoslav S. Mitrinovic, Josip E. Pečarić, and A. M. Fink. 1991. *Inequalities Involving Functions and Their Integrals and Derivatives*. Vol. 53. Springer Science & Business Media.
- [41] Esther Mohr, Iftikhar Ahmad, and Günter Schmidt. 2014. Online algorithms for conversion problems: A survey. *Surveys in Operations Research and Management Science* 19, 2 (July 2014), 87–104. <https://doi.org/10.1016/j.sorms.2014.08.001>
- [42] National Renewable Energy Laboratory (NREL). 2017. PVWatts Calculator. <https://www.pvwatts.nrel.gov>.
- [43] University of Massachusetts Amherst. 2017. Solar Carports: Turning University Parking Facilities into Renewable Electricity Plants. [https://www.epa.gov/sites/default/files/2017-09/documents/gppwebinar-9-26-17\\_small\\_umass.pdf](https://www.epa.gov/sites/default/files/2017-09/documents/gppwebinar-9-26-17_small_umass.pdf).
- [44] U.S. Department of Transportation. 2023. Rural EV Toolkit: Charger Types and Speeds. <https://www.transportation.gov/rural/ev/toolkit/ev-basics/charging-speeds>.
- [45] Manish Purohit, Zoya Svitkina, and Ravi Kumar. 2018. Improving Online Algorithms via ML Predictions. In *Advances in Neural Information Processing Systems*, S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett (Eds.), Vol. 31. Curran Associates, Inc.
- [46] Ana Radovanovic, Ross Konigstein, Ian Schneider, Bokan Chen, Alexandre Duarte, Binz Roy, Diyue Xiao, Maya Haridasan, Patrick Hung, Nick Care, et al. 2022. Carbon-Aware Computing for Datacenters. *IEEE Transactions on Power Systems* (2022).
- [47] Daan Ruppen, Nicolas Christianson, Debankur Mukherjee, and Adam Wierman. 2023. Smoothed Online Optimization with Unreliable Predictions. *Proceedings of the ACM on Measurement and Analysis of Computing Systems* 7, 1, Article 12 (March 2023), 36 pages. <https://doi.org/10.1145/3579442>
- [48] Manajit Sengupta, Yu Xie, Anthony Lopez, Aron Habte, Galen Maclaurin, and James Shelby. 2018. The National Solar Radiation Data Base (NSRDB). *Renewable and Sustainable Energy Reviews* 89 (2018), 51–60. <https://doi.org/10.1016/j>

rser.2018.03.003

- [49] Abel Souza, Noman Bashir, Jorge Murillo, Walid Hanafy, Qianlin Liang, David Irwin, and Prashant Shenoy. 2023. Ecovisor: A Virtual Energy System for Carbon-Efficient Applications. In *Proceedings of the 28th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 2* (Vancouver, BC, Canada) (ASPLOS 2023). Association for Computing Machinery, New York, NY, USA, 252–265. <https://doi.org/10.1145/3575693.3575709>
- [50] Seán M. Stewart. 2009. On Certain Inequalities Involving the Lambert W function. *Journal of Inequalities in Pure and Applied Mathematics* 10, 96 (Nov. 2009). Issue 4.
- [51] Thanathorn Sukprasert, Abel Souza, Noman Bashir, David Irwin, and Prashant Shenoy. 2023. Quantifying the Benefits of Carbon-Aware Temporal and Spatial Workload Shifting in the Cloud. arXiv:2306.06502 [cs.DC]
- [52] Thanathorn Sukprasert, Abel Souza, Noman Bashir, David Irwin, and Prashant Shenoy. 2023. Spatiotemporal Carbon-Aware Scheduling in the Cloud: Limits and Benefits. In *Companion Proceedings of the 14th ACM International Conference on Future Energy Systems* (Orlando, FL, USA) (*e-Energy '23 Companion*). Association for Computing Machinery, New York, NY, USA. <https://doi.org/10.1145/3599733.3606301>
- [53] Bo Sun, Russell Lee, Mohammad Hajiesmaili, Adam Wierman, and Danny Tsang. 2021. Pareto-Optimal Learning-Augmented Algorithms for Online Conversion Problems. In *Advances in Neural Information Processing Systems*, M. Ranzato, A. Beygelzimer, Y. Dauphin, P.S. Liang, and J. Wortman Vaughan (Eds.), Vol. 34. Curran Associates, Inc., 10339–10350. [https://proceedings.neurips.cc/paper\\_files/paper/2021/file/55a988dfb00a914717b3000a3374694c-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2021/file/55a988dfb00a914717b3000a3374694c-Paper.pdf)
- [54] Bo Sun, Lin Yang, Mohammad Hajiesmaili, Adam Wierman, John CS Lui, Don Towsley, and Danny HK Tsang. 2022. The Online Knapsack Problem with Departures. *Proceedings of the ACM on Measurement and Analysis of Computing Systems* 6, 3 (2022), 1–32.
- [55] Bo Sun, Ali Zeynali, Tongxin Li, Mohammad Hajiesmaili, Adam Wierman, and Danny H.K. Tsang. 2021. Competitive Algorithms for the Online Multiple Knapsack Problem with Application to Electric Vehicle Charging. *Proceedings of the ACM on Measurement and Analysis of Computing Systems* 4, 3, Article 51 (June 2021), 32 pages. <https://doi.org/10.1145/3428336>
- [56] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. 2020. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods* 17 (2020), 261–272. <https://doi.org/10.1038/s41592-019-0686-2>
- [57] Philipp Wiesner, Ija Behnke, Dominik Scheinert, Kordian Gontarska, and Lauritz Thamsen. 2021. Let’s Wait AWhile: How Temporal Workload Shifting Can Reduce Carbon Emissions in the Cloud. In *Proceedings of the 22nd International Middleware Conference*. Association for Computing Machinery, New York, NY, USA, 260–272. <https://doi.org/10.1145/3464298.3493399>
- [58] Lin Yang, Ali Zeynali, Mohammad H. Hajiesmaili, Ramesh K. Sitaraman, and Don Towsley. 2021. Competitive Algorithms for Online Multidimensional Knapsack Problems. *Proceedings of the ACM on Measurement and Analysis of Computing Systems* 5, 3, Article 30 (Dec 2021), 30 pages.
- [59] Sheng Shui Zhang. 2006. The effect of the charging protocol on the cycle life of a Li-ion battery. *Journal of Power Sources* 161, 2 (2006), 1385–1391. <https://doi.org/10.1016/j.jpowsour.2006.06.040>
- [60] Yunhong Zhou, Deeparnab Chakrabarty, and Rajan Lukose. 2008. Budget Constrained Bidding in Keyword Auctions and Online Knapsack Problems. In *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 566–576.

## APPENDIX

Table A1. A summary of key notations

Notation	Description
$C = 1$	Total item to be bought (or sold)
$T$	Deadline constraint; the player must finish purchasing (or selling) before $T$
$t \in [T]$	Current time step
$d_t : t \in [T]$	Rate constraint at time step $t$
$x_t \in [0, d_t]$	Decision at time $t$ .
$\beta$	Switching cost coefficient of the cost incurred from decision change $ x_t - x_{t-1} $
$U$	Upper bound on any partial derivative of $g_t(\cdot)$ that will be encountered
$L$	Lower bound on any partial derivative of $g_t(\cdot)$ that will be encountered
$g_t(\cdot)$	( <i>Online input</i> ) Cost/price function revealed to the player at time $t$

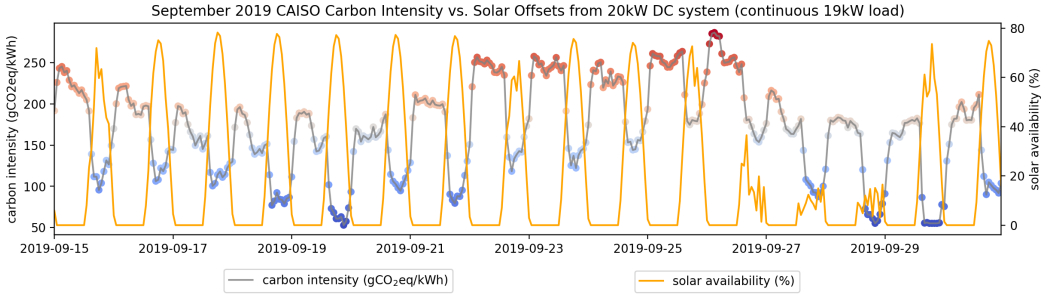


Fig. A1. (see Section 5) Grid carbon intensity (in  $\text{gCO}_2\text{eq/kWh}$ ) and simulated solar generation for the Caltech campus plotted over a two-week period in Sep. 2019, with one-hour granularity. The left-hand  $y$ -axis, scatter dots, and gray line correspond to the carbon intensity on the CAISO grid provided by Electricity Maps [39]. The right-hand  $y$  axis and the orange line give the simulated generation of a 15 kWh (DC nameplate rating) solar system, with irradiation data provided by the NSRDB data set [48]. Solar generation is expressed as a percentage with respect to a 19 kW continuous load, corresponding to the Level 2 AC charger simulated in our case study experiments (see Section 5.1 for more setup details).



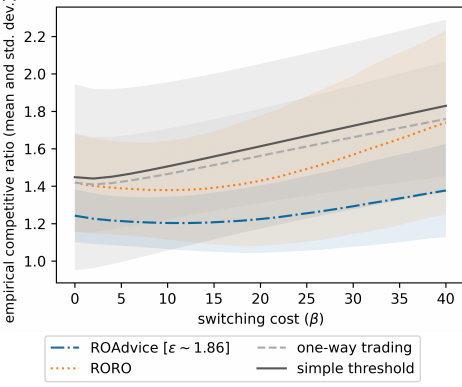


Fig. A2. (see Section 5.2) Average empirical competitive ratios and their standard deviations for experiments testing the impact of *switching cost magnitude*. Local solar generation is fixed to 0, and RO-Advice uses CarbonCast predictions with  $\epsilon \sim 1.86$ . Note that the simple threshold and one-way trading algorithms are *switching-oblivious*, while our proposed RORO and RO-Advice algorithms are aware of the switching cost.

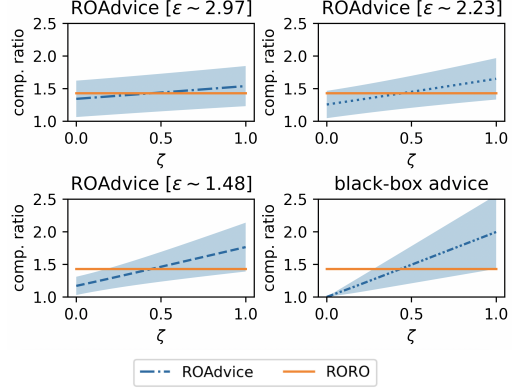


Fig. A3. (see Section 5.2) Average empirical competitive ratios and their standard deviations for experiments testing the consistency and robustness of RO-Advice-min by simulating advice from perfect ( $\zeta = 0$ ) to adversarial ( $\zeta = 1$ ). No local solar, and three versions of RO-Advice-min, with  $\epsilon \in [\sim 2.97, \sim 2.23, \sim 1.48]$ . The black-box advice algorithm naïvely plays the simulated advice at each time step. RORO (orange solid line) is a benchmark.

## A DEFERRED PSEUDOCODE FROM SECTION 3

**Algorithm A1** Online Ramp-On, Ramp-Off (RORO) framework instantiated for the online pause & resume problem (OPR-min)

- 1: **input:** threshold  $\{\phi_i\}_{i \in [k]}$ , switching cost  $\beta$ ;
- 2: **initialization:** initial decision  $x_0 = 0$ ,  $i = 1$ ;
- 3: **while** price  $c_t$  is revealed and  $i < k$  **do**
- 4: solve a (**ramping-on problem**) and obtain decision  $x_t^+$  and its pseudo cost  $r_t^+$ ,

$$x_t^+ = \arg \min_{x_t \in \{0,1\}, x_t \geq x_{t-1}} c_t x_t + \beta(x_t - x_{t-1}) - \phi_i x_t; \quad (16)$$

$$r_t^+ = c_t x_t^+ + \beta(x_t^+ - x_{t-1}) - \phi_i x_t^+; \quad (17)$$

- 5: solve a (**ramping-off problem**) and obtain decision  $x_t^-$  and its pseudo cost  $r_t^-$ ,

$$x_t^- = \arg \min_{x_t \in \{0,1\}, x_t \leq x_{t-1}} c_t x_t + \beta(x_{t-1} - x_t) - \phi_i x_t; \quad (18)$$

$$r_t^- = c_t x_t^- + \beta(x_{t-1} - x_t^-) - \phi_i x_t^-; \quad (19)$$

- 6: **if**  $r_t^+ \leq r_t^-$  **then** set  $x_t = x_t^+$  **else** set  $x_t = x_t^-$ ;
- 7: update  $i \leftarrow i + x_t$ ;

## B PROOFS

We now prove several results described in the main body.

In Appendix B.1, we show that the ramping-on and ramping-off optimization problems used in the RORO instantiations for OCS-min can be efficiently solved using iterative convex optimization techniques.

In Appendix B.2, we prove the competitive upper bound for RORO-min (Theorem 3.2). In Appendix B.3, we provide the proof of the information-theoretic lower bound for OCS-min (Theorem 3.3), which subsequently proves that RORO-min is optimal.

In Appendix B.5, we prove the  $\Omega(T)$  advice complexity lower bound for OCS presented in Theorem 4.2. In Appendix B.7, we prove the consistency and robustness bounds for RO-Advice-min (Theorem 4.3).

### B.1 Efficiently solving the ramping-on and ramping-off problems for RORO-min

Recall the ramping-on and ramping-off problems given by (13) and (14), respectively. First, we note that because the primary difference between the two problems is the restriction on the decision space, if we are only interested in the actual online decision  $x_t$ , it is valid to merge these problems and consider a single optimization problem as follows:

$$\text{RAMPONRAMPOFF}(g_t(\cdot), w^{(t-1)}, x_{t-1}) = \arg \min_{x \in [0, \min(1-w^{(t-1)}, d_t)]} g_t(x) + \beta|x - x_{t-1}| - \int_{w^{(t-1)}}^{w^{(t-1)}+x} \phi(u) du.$$

Let us define  $f_t(x) : t \in [T]$  as the right-hand side of the combined ramping-on and ramping-off minimization problem defined above:

$$f_t(x) = g_t(x) + \beta|x - x_{t-1}| - \int_{w^{(t-1)}}^{w^{(t-1)}+x} \phi(u) du. \quad (20)$$

**THEOREM B.1.** *Under the assumptions of OCS-min,  $f_t(x)$  is convex on the entire interval  $x \in [0, 1]$  for any  $t \in [T]$ .*

**PROOF.** We prove the above statement by contradiction.

By definition, the sum of two convex functions gives a convex function. First, note that the switching cost term can be equivalently defined in terms of the  $\ell_1$  norm as follows:  $\beta\|x - x_{t-1}\|_1$ . By definition and by observing that  $x_{t-1}$  is fixed,  $\beta\|x - x_{t-1}\|_1$  is convex. We have also assumed as part of the OCS-min problem setting that each  $g_t(x)$  is convex. Thus,  $g_t(x) + \beta\|x - x_{t-1}\|_1$  must be convex.

We turn our attention to the term  $-\int_{w^{(t-1)}}^{w^{(t-1)}+x} \phi(u) du$ . Let  $k(x) = \int_{w^{(t-1)}}^{w^{(t-1)}+x} \phi(u) du$ .

By the fundamental theorem of calculus,  $d/dx k(x) = \phi(z^{(t-1)} + x) \cdot d/dx x = \phi(z^{(t-1)} + x)$

Let  $p(x) = \phi(z^{(t-1)} + x)$ . Then  $d^2/d^2x k(x) = p'(x) \cdot d/dxx$ . Since  $\phi$  is monotonically decreasing on the interval  $[0, 1]$ , we know that  $p'(x) < 0$ , and thus  $d/dxx \cdot p'(x)$  is negative. This gives that  $k(x)$  is concave in  $x$ .

Since the negation of a concave function is convex, this causes a contradiction, because the sum of two convex functions is a convex function. Note that  $(g_t(x) + \beta\|x - x_{t-1}\|_1)$  and  $(-\int_{w^{(t-1)}}^{w^{(t-1)}+x} \phi(u) du)$  are both convex.

Thus,  $f_t(\cdot) = g_t(x) + \beta\|x - x_{t-1}\|_1 - \int_{w^{(t-1)}}^{w^{(t-1)}+x} \phi(u) du$  is always convex under the assumptions of the problem setting.  $\square$

Since the right-hand side of the combined ramping-on and ramping-off minimization problem is convex, (13) and (14) can be solved efficiently using an iterative convex optimization method.

## B.2 Competitive results for RORO-min

In the following, we prove Theorem 3.2, which states that the instantiation of RORO for OCS-min is  $\alpha$ -competitive, where  $\alpha$  is as defined in (15).

**PROOF OF THEOREM 3.2.** Let  $\mathcal{I} \in \Omega$  denote any valid OCS-min sequence, and let  $w^{(j)}$  denote RORO-min's final utilization before the compulsory trading, which begins at time step  $j \leq T$ . Note that  $w^{(t)} = \sum_{m \in [t]} x_t$  is non-decreasing over  $t$ .

**LEMMA B.2.** *The offline optimum is lower bounded by  $\text{OPT}(\mathcal{I}) \geq \phi(w^{(j)}) - \beta$ .*

*Proof of Lemma B.2.* We prove this lemma by contradiction. Note that the offline optimum (ignoring additional switching costs) is to trade all items at the best cost function over the sequence  $\{g_t(\cdot)\}_{t \in [T]}$ .

Suppose this best cost function is at an arbitrary step  $m$  ( $m \in [T]$ ,  $m \leq j$ ), denoted by  $g_m(\cdot)$ . Since cost functions are convex and additionally satisfy the conditions  $g_m(0) = 0$  and  $g_m(x) \geq 0 \forall x \in [0, 1]$ , the derivative of  $g_m$  at  $x = 0$  is a lower bound on the best marginal cost (i.e. per unit of capacity) that the optimal solution can obtain. We henceforth denote this derivative by  $\frac{dg_m}{dx}$ . Assume for the sake of contradiction that  $\frac{dg_m}{dx} = \text{OPT}(\mathcal{I}) < \phi(w^{(j)}) - \beta$ .

Next, we consider  $\frac{dg_m}{dx} - \beta < \phi(w^{(j)}) - 2\beta$ . Since  $\phi(z)$  is strictly decreasing on  $z \in [0, 1]$ , it follows that  $\phi(w^{(j)}) - 2\beta \leq \phi(w^{(m)})$ , as  $m \leq j$ . By solving the ramping-off problem, we have  $x_m^- = x_{m-1}$ . Thus, the online solution of step  $m$  is dominated by  $x_m^+$ , i.e.,  $x_m = x_m^+$ .

Furthermore, since  $\frac{dg_m}{dx} < \phi(w^{(j)}) - \beta$  and  $\phi(w^{(j)}) - \beta < \phi(w^{(m)}) - \beta$  as previously, by solving the ramping-on problem, we must have that the resulting decision satisfies  $x_m^+ > w^{(j)} - w^{(m-1)}$ . This implies that  $w^{(m)} > w^{(j)}$ , which, given that  $w^{(t)}$  is non-decreasing in  $t \in [T]$ , contradicts with the assumption that the final utilization before the compulsory trade is  $w^{(j)}$ .

Thus, we conclude that  $\text{OPT}(\mathcal{I}) \geq \phi(w^{(j)}) - \beta$ .

**LEMMA B.3.** *The cost of RORO-min( $\mathcal{I}$ ) is upper bounded by*

$$\text{RORO-min}(\mathcal{I}) \leq \int_0^{w^{(j)}} \phi(u) du + \beta w^{(j)} + (1 - w^{(j)})U. \quad (21)$$

*Proof of Lemma B.3.* By solving the ramping-off problem for any arbitrary time step  $t \in [T]$ , we can observe that  $g_t(x_t^-) - \beta x_t^- \leq \int_{w^{(t-1)}}^{w^{(t-1)} + x_t^-} \phi(u) du, \forall t \in [T]$ . Therefore, we have the following inequality

$$\min\{r_t^+, r_t^-\} \leq r_t^- \leq \beta x_{t-1}, \forall t \in [T]. \quad (22)$$

Thus, we have

$$\beta w^{(j)} \geq \beta w^{(N-1)} = \sum_{t \in [T]} (\beta x_{t-1}) \geq \sum_{t \in [T]} \min\{r_t^+, r_t^-\} \quad (23)$$

$$= \sum_{t \in [T]} \left[ g_t(x_t) + \beta |x_t - x_{t-1}| - \int_{w^{(t-1)}}^{w^{(t-1)} + x_t} \phi(u) du \right] \quad (24)$$

$$= \sum_{t \in [T]} [g_t(x_t) + \beta |x_t - x_{t-1}|] - \int_0^{w^{(j)}} \phi(u) du \quad (25)$$

$$= \text{RORO-min}(\mathcal{I}) - (1 - w^{(j)})U - \int_0^{w^{(j)}} \phi(u) du. \quad (26)$$

Combining Lemma B.2 and Lemma B.3 gives

$$\text{CR} \leq \frac{\text{RORO-min}(\mathcal{I})}{\text{OPT}(\mathcal{I})} \leq \frac{\int_0^{w^{(j)}} \phi(u) du + \beta w^{(j)} + (1 - w^{(j)})U}{\phi(w^{(j)}) - \beta} \leq \alpha, \quad (27)$$

where the last inequality holds since for any  $w \in [0, 1]$

$$\int_0^w \phi(u) du + \beta w + (1 - w)U = \int_0^w [U - \beta + (U/\alpha - U + 2\beta) \exp(w/\alpha)] + \beta w + (1 - w)U \quad (28)$$

$$= (U - \beta)w + \alpha(U/\alpha - U + 2\beta)[\exp(w/\alpha) - 1] + \beta w + (1 - w)U \quad (29)$$

$$= \alpha(U/\alpha - U + 2\beta)[\exp(w/\alpha) - 1] + U \quad (30)$$

$$= \alpha [U - 2\beta + (U/\alpha - U + 2\beta) \exp(w/\alpha)] \quad (31)$$

$$= \alpha[\phi(w) - \beta]. \quad (32)$$

We note that the rate constraints  $\{d_t\}_{t \in [T]}$  surprisingly do not appear in this worst-case analysis. For completeness, we state and prove Lemma B.4.

LEMMA B.4. *If  $d_t < 1 \forall t \in [T]$ , the competitive ratio of RORO-min is still upper bounded by  $\alpha$ .*

PROOF OF LEMMA B.4. Suppose that the presence of a rate constraint  $d_t$  causes RORO-min to make a decision which violates  $\alpha$ -competitiveness. At time  $t$ , the only difference between the setting where  $x_t \in [0, 1]$  and the setting with rate constraints  $< 1$ , where  $x_t \in [0, d_t]$ , is that  $x_t$  cannot be  $> d_t$ .

This implies that a challenging situation for RORO-min under a rate constraint is the case where RORO-min would otherwise accept  $> d_t$  of a good cost function, but it cannot due to the rate constraint.

We can now show that such a situation implies that RORO-min achieves a competitive ratio which is strictly better than  $\alpha$  (in the minimization setting).

From (27), we know that the following holds for any value of  $w \in [0, 1]$ :

$$\int_0^w \phi(u) du + \beta w + (1 - w)U \leq \alpha[\phi(w) - \beta].$$

For an arbitrary instance  $\mathcal{I} \in \Omega$  and an arbitrary time step  $t$ , let  $w^{(t)} = w^{(t-1)} + d_t$ , implying that  $x_t = d_t$ . For the sake of comparison, we first consider this time step with a cost function  $g_t(\cdot)$ , such that  $g_t(x_t) = \phi(w^{(t)}) \cdot x_t$ , implying that without the presence of a rate constraint, RORO-min would set  $x_t = d_t$ . If no more cost functions are accepted by RORO-min after time step  $t$  (excepting the compulsory trade), we have the following:

$$\frac{\text{RORO-min}(\mathcal{I})}{\text{OPT}(\mathcal{I})} \leq \frac{\int_0^{w^{(t)}} \phi(u) du + \beta w^{(t)} + (1 - w^{(t)})U}{\phi(w^{(t)}) - \beta} = \alpha. \quad (33)$$

Now, consider the exact same setting as above, except with a new cost function  $g'_t(\cdot)$ , such that  $g'_t(x_t) < \phi(w^{(t)}) \cdot x_t$ . This implies that without the presence of a rate constraint, RORO-min would set  $x_t > d_t$ . In other words,  $g'_t(\cdot)$  is a good cost function which RORO-min cannot accept more of due to the rate constraint.

Since OPT is subject to the same rate constraint  $d_t$ , we know that  $\text{OPT}(\mathcal{I})$  is lower bounded by  $[\phi(w^{(t)}) - \beta](1 - d_t) + g'_t(d_t)$  – the rest of the optimal solution is bounded by the final threshold value, since we assume that no more prices are accepted by RORO-min after time step  $t$ .

The cost RORO-min incurs is upper bounded by  $\text{RORO-min}(\mathcal{I}) \leq \int_0^{w^{(t)}} \phi(u)du - \int_{w^{(t-1)}}^{w^{(t)}} \phi(u)du + g'_t(d_t) + \beta w^{(t)} + (1 - w^{(t)})U$ .

Observe that compared to the previous setting, the OPT and RORO-min solutions have both decreased – OPT( $\mathcal{I}$ ) has decreased by a subtractive factor of  $-\phi(w^{(t)}) - \beta]d_t + g'_t(d_t)$ , while RORO-min( $\mathcal{I}$ ) has decreased by a subtractive factor of  $-\int_{w^{(t-1)}}^{w^{(t)}} \phi(u)du + g'_t(d_t)$ .

Since  $\phi$  is monotonically decreasing on  $w \in [0, 1]$ , we have that by definition,  $[\phi(w^{(t)}) - \beta]d_t < \int_{w^{(t-1)}}^{w^{(t)}} \phi(u)du$ . Thus, the solution obtained by RORO-min has decreased *more* than the solution obtained by OPT. This then implies the following:

$$\frac{\text{RORO-min}(\mathcal{I})}{\text{OPT}(\mathcal{I})} \leq \frac{\int_0^{w^{(t-1)}} \phi(u)du - \int_{w^{(t-1)}}^{w^{(t)}} \phi(u)du + g'_t(d_t) + \beta w^{(t)} + (1 - w^{(t)})U}{[\phi(w^{(t)}) - \beta](1 - d_t) + g'_t(d_t)} < \alpha,$$

where the final inequality follows from (33).  $\square$

At a high-level, this result shows that even if there is a rate constraint which prevents RORO-min from accepting a good cost function, the worst-case competitive ratio does not change. Combining Lemmas B.2, B.3, and B.4 completes the proof.  $\square$

### B.3 Lower bound for OCS-min

In the following, we prove Theorem 3.3, which states that for any deterministic online algorithm solving OCS-min, the optimal competitive ratio is  $\alpha$ , where  $\alpha$  is as defined in (15).

To show this lower bound, we first define a family of special instances and then show that the competitive ratio for any deterministic algorithm is lower bounded under these instances. Prior work has shown that difficult instances for online search problems with a minimization objective occur when inputs arrive at the algorithm in an decreasing order of cost [16, 27, 34, 55]. For OCS-min, we additionally consider how an adaptive adversary can essentially force an algorithm to incur a large switching cost in the worst-case. We now formalize such a family of instances  $\{\mathcal{I}_x\}_{x \in [L, U]}$ , where  $\mathcal{I}_x$  is called an *x-decreasing instance*.

**DEFINITION B.5 (x-DECREASING INSTANCE FOR OCS-min).** *Let  $n, m \in \mathbb{N}$  be sufficiently large, and  $\delta := (U-L)/n$ . For  $x \in [L, U]$ ,  $\mathcal{I}_x \in \Omega$  is an *x-instance* if it consists of  $n_x := 2 \cdot \lceil (x-L)/\delta \rceil + 1$  alternating batches of linear cost functions. The  $i^{\text{th}}$  batch ( $i \in [n_x - 2]$ ) contains  $m$  cost functions with coefficient  $U$  if  $i$  is odd, and 1 cost function with coefficient  $U - (i/2)\delta$  if  $i$  is even. The last two batches consist of  $m$  cost functions with coefficient  $x + \varepsilon$ , followed by  $m$  cost functions with coefficient  $U$ .*

Note that  $\mathcal{I}_U$  is simply a stream of  $m$  prices  $U$ . See Fig. A4 for an illustration of an *x-decreasing instance*. Since a competitive algorithm ALG should not accept the worst-case cost function with coefficient  $U$  unless it is forced to, this special *x-decreasing instance* can be equivalently interpreted as an *adaptive adversary*, which gives ALG  $m$  worst-case inputs  $U$  whenever any cost function is accepted. As  $n \rightarrow \infty$ , the alternating single cost functions in an *x-decreasing sequence* continuously decrease down to  $x$ , and each of these “good cost functions” is interrupted by a section of worst-case  $U$  cost functions. Note that the last few cost functions in an *x-decreasing instance* are always  $U$ .

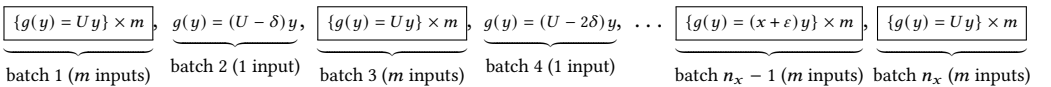


Fig. A4.  $\mathcal{I}_x$  consists of  $n_x$  batches of cost functions, where the alternating single functions are continuously decreasing from  $U$  down to  $x$ .

PROOF OF THEOREM 3.3. Let  $h(x)$  denote a *conversion function*  $[L, U] \rightarrow [0, 1]$ , which fully describes the actions of a deterministic ALG for OCS-min on an instance  $\mathcal{I}_x$ . Note that for large  $n$ , processing the instance  $\mathcal{I}_{x-\delta}$  is equivalent to first processing  $\mathcal{I}_x$  (besides the last two batches), and then processing batches with prices  $x - \delta$  and  $U$ . Since ALG is deterministic and the conversion is unidirectional (irrevocable), we must have that  $h(x - \delta) \geq h(x)$ , i.e.  $h(x)$  is non-increasing in  $[L, U]$ . Intuitively, the entire capacity should be satisfied if the minimum possible price is observed, i.e.  $h(L) = 1$ . For instance  $\mathcal{I}_x$ , the optimal offline solution is  $\text{OPT}(\mathcal{I}_x) = x + 2/m\beta$ . Note that as  $m$  grows large,  $\text{OPT}(\mathcal{I}_x) \rightarrow x$ .

Due to the adaptive nature of each  $x$ -instance, any deterministic ALG incurs a switching cost proportional to  $h(x)$ , which gives the amount of utilization obtained by ALG before the end of the sequence on instance  $\mathcal{I}_x$ .

Whenever ALG accepts some price  $U - ([i/2])\delta$ , the adversary presents prices  $U$  starting in the next time step. Any ALG which does not switch away immediately obtains a competitive ratio strictly worse than an algorithm which does switch away (if an algorithm accepts  $c$  fraction of a good price, the switching cost it will pay is  $2\beta c$ . An algorithm may continue accepting  $c$  fraction of prices  $U$  in the subsequent time steps, but a sequence exists where this decision will take up too much utilization to recover when better prices are presented later. In the extreme case, if an algorithm continues accepting  $c$  fraction of these  $U$  prices, it might fill its utilization and then OPT can accept a cost function which is arbitrarily better).

Since accepting any price by a factor of  $c$  incurs a switching cost of  $2\beta c$ , the switching cost paid by ALG on instance  $\mathcal{I}_x$  is  $2\beta h(x)$ . We assume that ALG is notified of the compulsory trade, and does not incur a significant switching cost during the final batch.

Then the total cost incurred by an  $\alpha^*$ -competitive online algorithm ALG on instance  $\mathcal{I}_x$  is  $\text{ALG}(\mathcal{I}_x) = h(U/\alpha^*)U/\alpha^* - \int_{U/\alpha^*}^x udh(u) + 2\beta h(x) + (1 - h(x))U$ , where  $udh(u)$  is the cost of buying  $dh(u)$  utilization at price  $u$ , the last term is from the compulsory trade, and the second to last term is the switching cost incurred by ALG. Note that any deterministic ALG which makes conversions when the price is larger than  $U/\alpha^*$  can be strictly improved by restricting conversions to prices  $\leq U/\alpha^*$ .

For any  $\alpha^*$ -competitive online algorithm, the corresponding conversion function  $h(\cdot)$  must satisfy  $\text{ALG}(\mathcal{I}_x) \leq \alpha^* \text{OPT}(\mathcal{I}_x) = \alpha^* x, \forall x \in [L, U]$ . This gives a necessary condition which the conversion function must satisfy as follows:

$$\text{ALG}(\mathcal{I}_x) = h(U/\alpha^*)U/\alpha^* - \int_{U/\alpha^*}^x udh(u) + 2\beta h(x) + (1 - h(x))U \leq \alpha^* x, \quad \forall x \in [L, U].$$

By integral by parts, the above implies that the conversion function must satisfy  $h(x) \geq \frac{U - \alpha^* x}{U - x - 2\beta} - \frac{1}{U - x - 2\beta} \int_{U/\alpha^*}^x h(u) du$ . By Grönwall's Inequality [40, Theorem 1, p. 356], we have that

$$\begin{aligned} h(x) &\geq \frac{U - \alpha^* x}{U - x - 2\beta} - \frac{1}{U - x - 2\beta} \int_{U/\alpha^*}^x \frac{U - \alpha^* u}{U - u - 2\beta} \cdot \exp\left(\int_u^x \frac{1}{U - r - 2\beta} dr\right) du \\ &\geq \frac{U - \alpha^* x}{U - x - 2\beta} - \int_{U/\alpha^*}^x \frac{U - \alpha^* u}{(U - u - 2\beta)^2} du \\ &\geq \frac{U - \alpha^* x}{U - x - 2\beta} - \left[ \frac{U\alpha^* - U - 2\beta\alpha^*}{u + 2\beta - U} - \alpha^* \ln(u + 2\beta - U) \right]_{U/\alpha^*}^x \\ &\geq \alpha^* \ln(x + 2\beta - U) - \alpha^* \ln(U/\alpha^* + 2\beta - U), \quad \forall x \in [L, U]. \end{aligned}$$

$h(L) = 1$  by the problem definition – we can combine this with the above constraint to give the following condition for an  $\alpha^*$ -competitive online algorithm:

$$\alpha^* \ln(L + 2\beta - U) - \alpha^* \ln(U/\alpha^* + 2\beta - U) \leq h(L) = 1.$$

The optimal  $\alpha^*$  is obtained when the above inequality is binding, so solving for the value of  $\alpha^*$  which solves  $\alpha^* \ln(L + 2\beta - U) - \alpha^* \ln(U/\alpha^* + 2\beta - U) = 1$  yields that the best competitive ratio for any ALG solving OCS-min is  $\alpha^* \geq \left[ W\left(\frac{e^{2\beta/U}(L/U + 2\beta/U - 1)}{e}\right) - \frac{2\beta}{U} + 1 \right]^{-1}$ .  $\square$

#### B.4 Lower bound for the “best-price” predictions in OCS

In the following, we prove Theorem 4.1, which states that any learning-augmented algorithm for OCS which uses a prediction of the “best price” in a sequence cannot achieve 1 consistency.

**PROOF OF THEOREM 4.1.** To prove the result, we state and prove two lemmas, one for each of the minimization (OCS-min) and maximization (OCS-max) settings. These results show that the achievable consistency is lower bounded by  $1 + 2\beta/L$  and  $1/1 - 2\beta/U$  in OCS-min and OCS-max, respectively.

**LEMMA B.6.** *Given a single prediction of the “minimum price” in a sequence, any learning-augmented algorithm for OCS-min which is  $\gamma$ -robust is at least  $\eta$ -consistent, where  $\eta$  is*

$$\eta \geq \gamma - \gamma \left(1 + \frac{2\beta}{L} - \frac{U}{L}\right) \ln \left[ \frac{L + 2\beta - U}{\frac{U}{\gamma} + 2\beta - U} \right] + 1 - \frac{U}{L} + \frac{2\beta}{L}.$$

*Note that for any value of  $\gamma$ ,  $\eta \geq 1 + 2\beta/L$ .*

**PROOF OF LEMMA B.6.** To show this result, we leverage the same special family of  $x$ -decreasing instances defined in Definition B.3.

We let  $h(x)$  denote a *conversion function*  $[L, U] \rightarrow [0, 1]$ , which fully describes the actions of a deterministic ALG for OCS-min on an instance  $\mathcal{I}_x$ , and  $h(x)$  gives the total amount traded under the instance  $\mathcal{I}_x$  before the compulsory trade. Note that for large  $n$ , processing the instance  $\mathcal{I}_{x-\delta}$  is equivalent to first processing  $\mathcal{I}_x$  (besides the last two batches), and then processing batches with prices  $x - \delta$  and  $U$ . Since ALG is deterministic and the conversion is unidirectional (irrevocable), we must have that  $h(x - \delta) \geq h(x)$ , i.e.  $h(x)$  is non-increasing in  $[L, U]$ . Intuitively, the entire capacity should be satisfied if the minimum possible price is observed, i.e.  $h(L) = 1$ .

Recall that for instance  $\mathcal{I}_x$ , the optimal offline solution is  $\text{OPT}(\mathcal{I}_x) = x + 2/m\beta$ , and that as  $m$  grows large,  $\text{OPT}(\mathcal{I}_x) \rightarrow x$ .

For any  $\gamma$ -robust online algorithm ALG given a prediction  $\hat{P} = L$ , the following must hold:

$$\text{ALG}(\mathcal{I}_x) \leq \gamma \text{OPT}(\mathcal{I}_x) = \gamma x, \quad \forall x \in [L, U].$$

The cost of ALG with conversion function  $h$  on an instance  $\mathcal{I}_x$  is  $\text{ALG}(\mathcal{I}_x) = h(U/\gamma)U/\gamma - \int_{U/\gamma}^x udh(u) + 2\beta h(x) + (1 - h(x))U$ , where  $udh(u)$  is the cost of buying  $dh(u)$  utilization at price  $u$ , the last term is from the compulsory conversion, and the second to last term is the switching cost incurred by ALG.

This implies that  $h(x)$  must satisfy the following:

$$h(U/\gamma)U/\gamma - \int_{U/\gamma}^x udh(u) + 2\beta h(x) + (1 - h(x))U \leq \gamma x, \quad \forall x \in [L, U].$$

By integral by parts, the above implies that the conversion function must satisfy  $h(x) \geq \frac{U-\gamma x}{U-x-2\beta} - \frac{1}{U-x-2\beta} \int_{U/\gamma}^x h(u)du$ . By Grönwall's Inequality [40, Theorem 1, p. 356], we have that

$$h(x) \geq \frac{U-\gamma x}{U-x-2\beta} - \frac{1}{U-x-2\beta} \int_{U/\gamma}^x \frac{U-\gamma u}{U-u-2\beta} \cdot \exp\left(\int_u^x \frac{1}{U-r-2\beta} dr\right) du \quad (34)$$

$$\geq \frac{U-\gamma x}{U-x-2\beta} - \int_{U/\gamma}^x \frac{U-\gamma u}{(U-u-2\beta)^2} du \quad (35)$$

$$\geq \frac{U-\gamma x}{U-x-2\beta} - \left[ \frac{U\gamma-U-2\beta\gamma}{u+2\beta-U} - \gamma \ln(u+2\beta-U) \right]_{U/\gamma}^x \quad (36)$$

$$\geq \gamma \ln(x+2\beta-U) - \gamma \ln(U/\gamma+2\beta-U), \quad \forall x \in [L, U]. \quad (37)$$

In addition, to simultaneously satisfy  $\eta$ -consistency when the minimum price prediction is  $\hat{P} = L$ , ALG must satisfy  $\text{ALG}(\mathcal{I}_L) \leq \eta \text{OPT}(\mathcal{I}_L) = \eta L$ . Combining this constraint with  $h(L) = 1$  gives:

$$\int_{U/\gamma}^L h(u)du + 2\beta \leq \eta L - L. \quad (38)$$

By combining equations (37) and (38), the conversion function  $h(x)$  of any  $\gamma$ -robust and  $\eta$ -consistent online algorithm given prediction  $\hat{P} = L$  must satisfy the following:

$$\gamma \int_{U/\gamma}^L \ln\left(\frac{u+2\beta-U}{U/\gamma+2\beta-U}\right) du + 2\beta \leq \eta L - L. \quad (39)$$

When all inequalities are binding, this equivalently gives that

$$\eta \geq \gamma - \gamma(1 + 2\beta/L - U/L) \ln\left(\frac{L+2\beta-U}{U/\gamma+2\beta-U}\right) + 1 - \frac{U}{L} + \frac{2\beta}{L}.$$

For any value of  $\gamma$ ,  $\eta \geq 1 + 2\beta/L$ , and this completes the proof.  $\square$

**LEMMA B.7.** *Given a single prediction of the “maximum price” in a sequence, any learning-augmented algorithm for OCS-max which is  $\gamma$ -robust is at least  $\eta$ -consistent, where  $\eta$  is*

$$\eta \geq \left[ 1 - \frac{U-2\beta-L}{U\gamma} \ln\left(\frac{U-2\beta-L}{\gamma L-2\beta-L}\right) + \frac{1}{\gamma} - \frac{L}{U} - \frac{2\beta}{U} \right]^{-1}.$$

*Note that for any value of  $\gamma$ ,  $\eta \geq 1/(1-2\beta/U)$ .*

**PROOF OF LEMMA B.7.** To show this result, we leverage the same special family of  $x$ -increasing instances defined in Definition D.3.

We let  $h(x)$  denote a *conversion function*  $[L, U] \rightarrow [0, 1]$ , which fully describes the actions of a deterministic ALG for OCS-max on an instance  $\mathcal{I}_x$ , and  $h(x)$  gives the total amount traded under the instance  $\mathcal{I}_x$  before the compulsory trade. Note that for large  $n$ , processing the instance  $\mathcal{I}_{x+\delta}$  is equivalent to first processing  $\mathcal{I}_x$  (besides the last two batches), and then processing batches with prices  $x + \delta$  and  $L$ . Since ALG is deterministic and the conversion is unidirectional (irrevocable), we must have that  $h(x + \delta) \geq h(x)$ , i.e.  $h(x)$  is non-decreasing in  $[L, U]$ . Intuitively, the entire capacity should be satisfied if the minimum possible price is observed, i.e.  $h(U) = 1$ .

Recall that for instance  $\mathcal{I}_x$ , the optimal offline solution is  $\text{OPT}(\mathcal{I}_x) = x - 2/m\beta$ , and that as  $m$  grows large,  $\text{OPT}(\mathcal{I}_x) \rightarrow x$ .

For any  $\gamma$ -robust online algorithm ALG given a prediction  $\hat{P} = U$ , the following must hold:

$$\text{ALG}(\mathcal{I}_x) \leq \text{OPT}(\mathcal{I}_x)/\gamma = x/\gamma, \quad \forall x \in [L, U].$$



The profit of ALG with conversion function  $h$  on an instance  $\mathcal{I}_x$  is  $\text{ALG}(\mathcal{I}_x) = h(\gamma L)\gamma L + \int_{\gamma L}^x udh(u) - 2\beta h(x) + (1 - h(x))L$ , where  $udh(u)$  is the profit of selling  $dh(u)$  utilization at price  $u$ , the last term is from the compulsory conversion, and the second to last term is the switching cost incurred by ALG.

This implies that  $h(x)$  must satisfy the following:

$$h(\gamma L)\gamma L + \int_{\gamma L}^x udh(u) - 2\beta h(x) + (1 - h(x))L \geq x/\gamma, \quad \forall x \in [L, U].$$

By integral by parts, the above implies that the conversion function must satisfy  $h(x) \geq \frac{x/\gamma - L}{x - 2\beta - L} + \frac{1}{x - 2\beta - L} \int_{\gamma L}^x h(u)du$ . By Grönwall's Inequality [40, Theorem 1, p. 356], we have that

$$h(x) \geq \frac{x/\gamma - L}{x - 2\beta - L} + \frac{1}{x - 2\beta - L} \int_{\gamma L}^x \frac{U/\gamma - L}{u - 2\beta - L} \cdot \exp\left(\int_u^x \frac{1}{r - 2\beta - L} dr\right) du \quad (40)$$

$$\geq \frac{x/\gamma - L}{x - 2\beta - L} + \int_{\gamma L}^x \frac{U/\gamma - L}{(u - 2\beta - L)^2} du \quad (41)$$

$$\geq \frac{x/\gamma - L}{x - 2\beta - L} + \left[ \frac{L\gamma - 2\beta - L}{\gamma u - 2\beta\gamma - L\gamma} + \frac{1}{\gamma} \ln(u - 2\beta - L) \right]_{\gamma L}^x \quad (42)$$

$$\geq \frac{1}{\gamma} \ln(x - 2\beta - L) - \frac{1}{\gamma} \ln(\gamma L - 2\beta - L), \quad \forall x \in [L, U]. \quad (43)$$

In addition, to simultaneously satisfy  $\eta$ -consistency when the maximum price prediction is  $\hat{P} = U$ , ALG must satisfy  $\text{ALG}(\mathcal{I}_U) \geq \text{OPT}(\mathcal{I}_U)/\eta = U/\eta$ . Combining this constraint with  $h(U) = 1$  gives:

$$\int_{\gamma L}^U h(u)du + 2\beta \leq \frac{(\eta - 1)U}{\eta}. \quad (44)$$

By combining equations (43) and (44), the conversion function  $h(x)$  of any  $\gamma$ -robust and  $\eta$ -consistent online algorithm given prediction  $\hat{P} = L$  must satisfy the following:

$$\frac{1}{\gamma} \int_{\gamma L}^U \ln\left(\frac{x - 2\beta - L}{\gamma L - 2\beta - L}\right) du + 2\beta \leq \frac{(\eta - 1)U}{\eta}. \quad (45)$$

When all inequalities are binding, this equivalently gives that

$$\eta \geq \left[ 1 - \frac{U - 2\beta - L}{U\gamma} \ln\left(\frac{U - 2\beta - L}{\gamma L - 2\beta - L}\right) + \frac{1}{\gamma} - \frac{L}{U} - \frac{2\beta}{U} \right]^{-1}.$$

For any value of  $\gamma$ ,  $\eta \geq 1/(1 - 2\beta/U)$ , and this completes the proof.  $\square$

By combining Lemmas B.6 and B.7, the result follows.  $\square$

## B.5 Advice complexity lower bound for OCS

In the following, we prove Theorem 4.2, which states that for any instance of OCS with sequence length  $T$ , at least  $\Omega(T)$  bits of advice are necessary to achieve 1-consistency.

**PROOF OF THEOREM 4.2.** Recall that  $T \geq 1$  is the integer length of a valid sequence for OCS. We start by deriving a lower bound on the *number of unique solutions* for sequences of length  $T$ , which

we denote as  $N_T$ . Note that any valid solution satisfies the following, assuming that rate constraints are all 1 (i.e.,  $d_t = 1 \forall t \in [T]$ ):

$$\sum_{i=1}^T x_t = 1, 0 \leq x_t \leq 1 \forall t \in [T].$$

To simplify the problem, we discretize the action space  $x_t \in [0, 1]$  into  $T + 1$  bins, allowing each  $x_t$  to take on values  $\{0/T, 1/T, \dots, T/T\}$ . This simplifies the action space, making it countable while still allowing for solutions that either evenly distribute conversion decisions across all  $T$  time steps or convert the entire asset in one-time step. By the balls-and-bars formula [18, Page 40], the number of solutions which satisfy the above constraints is  $N_T \geq \binom{2T}{T-1}$ .

We prove the theorem statement by contradiction: suppose there exists an algorithm ALG which uses  $A_T := o(T)$  bits of accurate advice (i.e., sublinear in  $T$ ) and achieves a consistency of 1 on any input sequence with length  $T$ . With  $A_T$  bits of advice, ALG can theoretically distinguish between  $2^{A_T}$  different solutions.

Suppose the same ALG processes an arbitrary sequence with length  $T + 1$ . Since the number of advice bits is sublinear in  $T$ , we know that  $A_{T+1} < A_T + 1$ . On the other hand, the number of unique solutions  $N_{T+1}$  satisfies the following:

$$N_{T+1} \geq \binom{2T+2}{T} = \frac{(2T+2)!}{(T)!(T+2)!} = \frac{(2T)!(2T+1)(2T+2)}{(T-1)!(T)(T+1)!(T+2)} = \frac{4T^2 + 6T + 2}{T^2 + 2T} \cdot N_T$$

Note that  $\lim_{T \rightarrow \infty} N_{T+1} \geq 4N_T$ .

By assumption we have that ALG can distinguish between  $2^{A_{T+1}}$  different solutions, where  $2^{A_{T+1}} < 2 \cdot 2^{A_T}$ . Conversely, as the length of the sequence increases by 1, the number of unique solutions grows by at least  $N_{T+1} \geq 4 \cdot N_T$ .

If we assume that the advice is powerful enough to distinguish between the possible solutions with the sequences of length  $T$ , then we have  $2^{A_T} \approx N_T$ . However, this implies that by pigeonhole, because  $2^{A_{T+1}} < 4N_T \leq N_{T+1}$ , there must exist at least two unique solutions for sequences of length  $T + 1$  which map to the same advice string. Then a contradiction follows by constructing two sequences, where each of these two solutions, respectively, is the unique optimal. By definition, such a construction forces ALG to achieve a competitive ratio strictly greater than 1.  $\square$

## B.6 Remarks on the Advice Complexity of OCS

In Section 4.1, we gave a strong advice complexity lower bound (see Theorem 4.2) showing that advice which grows linearly in the length of the input sequence is necessary to achieve 1-consistency for the general formulation of OCS.

In this supplementary section, we pick up from the brief remarks in Section 4.1 and discuss how this lower bound might be improved if certain aspects of the OCS problem are relaxed and simplified. Specifically, we will assume that the rate constraints are effectively “removed”, so they satisfy  $\{d_t = C\}_{\forall t \in [T]}$ , and the cost/price functions are linear. For ease of presentation, we focus on the minimization setting (OCS-min) here. We start with a useful observation about the optimal solution which holds under the above conditions.

**OBSERVATION B.8.** *For any instance of OCS-min where the rate constraints satisfy  $\{d_t = C\}_{\forall t \in [T]}$  and all cost functions are linear, the optimal solution is upper bounded by  $\text{OPT} \leq c_{\min} + 2\beta$ , where  $c_{\min}$  denotes the linear coefficient of the cost function which satisfies  $g_{\min}(1) \leq g_t(1) \forall t \in [T]$ .*

PROOF FOR OBSERVATION B.8. By definition and WLOG, there exists some cost function in the valid sequence which is the lowest, which we denote as  $g_{\min}$ . Since we assume all cost functions are linear, it follows that the linear coefficient  $c_{\min}$  of this cost function satisfies  $c_{\min} \leq c_t \forall t \in [T]$  (i.e. it is the smallest linear coefficient of any cost function in the sequence).

Then by definition, the solution which purchases the full item at cost of  $g_{\min}(1)$  and incurs a switching cost of  $2\beta$  is always an upper bound on the objective value of the optimal solution (i.e., if any solution incurs a cost strictly greater than  $c_{\min} + 2\beta$ , it cannot be optimal).  $\square$

Using this useful observation for the restricted set of OCS instances, we give two conjectures which jointly suggest that within the space of optimal solutions, there exists a “simple” solution which obtains the optimal objective value. First, we suggest that if an optimal solution chooses non-zero decisions across a contiguous sequence of cost functions, there is an equivalent optimal solution which chooses the same non-zero decision across the same contiguous sequence of cost functions.

CONJECTURE B.9. *For any instance of OCS-min where the rate constraints satisfy  $\{d_t = C\}_{\forall t \in [T]}$  and all cost functions are linear, if an optimal solution chooses non-zero actions (i.e.  $x_t > 0$ ) across a contiguous sequence of cost functions with length  $m$ , there is always an equivalent optimal solution which chooses the same non-zero action  $x^*$  for each of the  $m$  cost functions.*

INTUITION FOR CONJECTURE B.9. Consider a contiguous sequence of cost functions with length  $m$ , where (WLOG) the first  $m - 1$  cost functions have linear coefficient  $c_i$ , and the last cost function has linear coefficient  $c_j$ . We abuse notation and generally use subscript  $i$  to refer to the first  $m - 1$  time slots and  $j$  refers to the final time slot, which is of particular interest. Assume that  $c_j > c_i$ .

For the sake of contradiction, assume that there is some optimal solution which assigns non-zero decisions to all of these cost functions, where the decision is denoted by  $x_i$  for the first  $m - 1$  cost functions, and  $x_j$  for the last cost function.

Then we ask the following guiding question: *when should an optimal solution OPT set  $x_j < x_i$ ?* (i.e. when does it make sense for OPT to take less than  $x_i$  of the cost function w/ coefficient  $c_j$ ?)

Intuitively, if  $c_j$  is much larger than  $c_i$ , OPT should not accept it at all. On the other hand, if  $c_j$  is only marginally worse than  $c_i$ , it makes more sense to set  $x'_i = x_j$  to minimize the overall switching cost. Enumerating these cases formally, the following expression describes the case where taking  $x_j < x_i$  gives a lower overall cost than uniformly trading across the whole interval at a rate  $x'_i$ .

$$\sum_{i \neq j} c_i x_i + c_j x_j + 2\beta \max(x_i) \leq \sum_{i \neq j} c_i x'_i + c_j x'_i + 2\beta x'_i.$$

The inequality then implies that if the following inequality holds, setting  $x_j < x_i$  is a better or equal outcome:

$$(c_j - c_i) \left[ |\max(x_i) - x_j| - \frac{|\max(x'_i) - x_j|}{m} \right] \leq 2\beta \left[ \frac{|\max(x'_i) - x_j|}{m} \right].$$

This subsequently gives the condition that when  $c_j = c_i + \frac{2\beta}{m-1}$ ,  $x_j$  can be any value between 0 and  $x'_i$  and the same cost will be attained. However, also this implies that in this edge case, the optimal solution can either decide to distribute the trading equally (i.e.  $x'_i$  for each of the  $m$  cost functions), or completely ignore the last cost function (i.e.  $x_i^*$  for each of the first  $m - 1$  cost functions), obtaining an identical objective value.

It follows, then, that if  $c_j > c_i + \frac{2\beta}{m-1}$ , it should be completely ignored, and if  $c_j < c_i + \frac{2\beta}{m-1}$ , it should be fully accepted (i.e.  $x_j = x_i$ ). Thus, in each of these enumerated cases, there is an optimal solution which assigns a single constant decision to each of the  $m$  cost functions, and the assumption causes a contradiction.  $\square$

This conjecture tells us that within a consecutive interval where the optimal solution decides to trade, there can be a single decision value applied across the entire interval which attains the optimal competitive ratio. Next, we give intuition for a result showing that a solution can achieve the optimal objective by only “ramping on” once during a sequence.

**CONJECTURE B.10.** *For any instance of OCS-min where the rate constraints satisfy  $\{d_t = C\}_{\forall t \in [T]}$  and all cost functions are linear, there is an optimal solution where any non-zero actions are clustered into a single contiguous sequence of arbitrary length  $k$ .*

**INTUITION FOR CONJECTURE B.10.** Recall Observation B.8, which states that OPT is upper bounded by  $c_{\min} + 2\beta$ , and there is always a valid solution which simply accepts  $g_{\min}(1)$  (in one time step) and incurs a switching cost of  $2\beta$ .

Now suppose there is a sequence with two contiguous segments of interest. The first is a segment of length 1, containing a single cost function with coefficient  $c_{\min}$ . The other is a segment of arbitrary length  $k$ , containing  $k$  cost functions, where the average linear coefficient is  $c_{\text{avg}}$ . We assume that  $c_{\text{avg}} \geq c_{\min}$ , and that the two segments are disjoint, i.e. there is at least one cost function separating them in time.

Then we ask the following guiding question: *when should an optimal solution OPT consider “transferring” some amount of trading  $x_{\text{amount}}$  from the first segment  $c_{\min}$  to the  $k$ -segment?* Intuitively, if  $c_{\text{avg}} = c_{\min}$ , it makes sense for OPT to move all of the trading onto the  $k$ -segment, since it is strictly better for the switching cost. On the other hand, if  $c_{\text{avg}}$  is too large, the optimal solution will ignore it completely. We are primarily interested in the case where the optimal solution will decide to trade some amount at the first segment, followed by a different amount at the second  $k$ -segment.

Enumerating these cases formally, the following expression describes the case where taking some amount  $x_{\text{amount}}$  in the  $k$ -segment gives a lower overall cost compared to trading completely in the singleton segment.

$$c_{\min}(1 - x_{\text{amount}}) + c_{\text{avg}}x_{\text{amount}} + 2\beta \left( (1 - x_{\text{amount}}) + \frac{x_{\text{amount}}}{k} \right) \leq c_{\min} + 2\beta$$

$$c_{\text{avg}}x_{\text{amount}} + \frac{2\beta x_{\text{amount}}}{k} \leq x_{\text{amount}} [c_{\min} + 2\beta].$$

The above then implies that if the following inequality holds, moving  $x_{\text{amount}}$  to the  $k$ -segment is a better or equal outcome:

$$c_{\text{avg}} \leq c_{\min} + \frac{(k \cdot x_{\text{amount}} - 1)2\beta}{k \cdot x_{\text{amount}}}.$$

Note that  $x_{\text{amount}} \in (0, 1)$  under the assumption that it makes sense for OPT to accept some amount in both segments. However, the above gives the following corollary result:

$$c_{\text{avg}} < \frac{(k - 1)2\beta}{k}.$$

Observe that if  $c_{\text{avg}} = \frac{(k-1)2\beta}{k}$ , the solutions that fully accept in the second segment and fully accept in the first segment, respectively, obtain the same objective value:  $c_{\text{avg}} + 2\beta/k = c_{\min} + 2\beta$ . This causes a contradiction because it implies that if  $c_{\text{avg}} < \frac{(k-1)2\beta}{k}$ , a *strictly better* objective value can be attained by moving *all* of the acceptance from the first to the second segment.

Since the condition for moving some  $x_{\text{amount}} \in (0, 1)$  from the first segment to the second segment implies the above is true as a corollary (i.e.  $c_{\text{avg}} \leq c_{\min} + \frac{(k \cdot x_{\text{amount}} - 1)2\beta}{k \cdot x_{\text{amount}}} < \frac{(k-1)2\beta}{k}$ ), it seems that *as soon as it makes sense to move some acceptance from one segment to another, it equivalently makes sense to move all of the acceptance to the other segment.*

This subsequently gives the condition that when  $c_j = c_i + \frac{2\beta}{m-1}$ ,  $x_j$  can be any value between 0 and  $x'_i$  and the same cost will be attained. However, also this implies that in this edge case, the optimal solution can either decide to distribute the trading equally (i.e.  $x'_i$  for each of the  $m$  cost functions), or completely ignore the last cost function (i.e.  $x_i^*$  for each of the first  $m - 1$  cost functions), obtaining an identical objective value.  $\square$

The significance of these results is primarily in the suggestion that an optimal solution can be fully characterized by *two critical values* in this simplified setting. Namely, if Conjectures B.9 and B.10 theoretically hold, then there is always an optimal solution in the simplified OCS setting which clusters all of its acceptance decisions into a single contiguous interval, and the acceptance rate is uniform.

Such an optimal solution can be recovered by a single start index of the acceptance interval  $i^*$ , and the acceptance rate  $x^*$ . Intuitively, because these parameters do not grow with the length of the input sequence  $T$ , this seems to suggest that the advice complexity lower bound no longer holds in the setting without rate constraints and where cost functions are linear.

Although we do not explore this direction further because rate constraints and convex cost functions are required for most of our motivating applications (particularly carbon-aware EV charging as explored in Section 5) it would be very interesting to explore this dynamic further in future work. It would be particularly interesting to see whether similar or improved consistency-robustness trade-offs can be achieved in this simplified OCS setting where the predictions given to a learning-augmented algorithm are of size sublinear in the length of the input.

### B.7 R0-Advice consistency and robustness for OCS-min

In the following, we prove Theorem 4.3, which states that the instantiation of R0-Advice for OCS-min is  $(1 + \epsilon)$ -consistent and  $\left(\frac{(U+2\beta)/L(\alpha-1-\epsilon)+\alpha\epsilon}{(\alpha-1)}\right)$ -robust, where  $\epsilon$  and  $\alpha$  are defined in Definition 4.2 and (15), respectively. We note that consistency describes the performance of the algorithm when predictions are completely correct, while robustness describes the performance when predictions are adversarially wrong [35, 45].

**PROOF OF THEOREM 4.3.** We start by noting that the online solution given by R0-Advice-min is always feasible (under the assumption that  $\sum_{t=1}^T \hat{x}_t = 1$ ), since

$$\sum_{t=1}^T x_t = \sum_{t=1}^T [\lambda \hat{x}_t + (1 - \lambda) \tilde{x}_t] = \lambda + (1 - \lambda) = 1.$$

Let  $\mathcal{I} \in \Omega$  be an arbitrary valid OCS-min sequence. We denote the *purchasing* and *switching* costs of the robust advice by  $\text{RORO-min}_{\text{bought}}$  and  $\text{RORO-min}_{\text{switch}}$ , respectively. Likewise, the purchasing and switching costs of the black box advice are denoted by  $\text{ADV}_{\text{bought}}$  and  $\text{ADV}_{\text{switch}}$ .

The cost of R0-Advice-min is bounded by

$$\begin{aligned}
\text{R0-Advice-min}(\mathcal{I}) &= \sum_{t=1}^T g_t(x_t) + \sum_{t=1}^{T+1} \beta |x_t - x_{t-1}| \\
&= \sum_{t=1}^T g_t(\lambda \hat{x}_t + (1-\lambda)\tilde{x}_t) + \sum_{t=1}^{T+1} \beta |\lambda \hat{x}_t + (1-\lambda)\tilde{x}_t - \lambda \hat{x}_{t-1} - (1-\lambda)\tilde{x}_{t-1}| \\
&\leq \lambda \sum_{t=1}^T g_t(\hat{x}_t) + (1-\lambda) \sum_{t=1}^T g_t(\tilde{x}_t) \\
&\quad + \sum_{t=1}^{T+1} \beta |\lambda \hat{x}_t - \lambda \hat{x}_{t-1}| + \sum_{t=1}^{T+1} \beta |(1-\lambda)\tilde{x}_t - (1-\lambda)\tilde{x}_{t-1}| \\
&\leq \lambda \text{ADV}_{\text{bought}}(\mathcal{I}) + (1-\lambda) \text{RORO-min}_{\text{bought}}(\mathcal{I}) \\
&\quad + \lambda \sum_{t=1}^{T+1} \beta |\hat{x}_t - \hat{x}_{t-1}| + (1-\lambda) \sum_{t=1}^{T+1} \beta |\tilde{x}_t - \tilde{x}_{t-1}| \\
&\leq \lambda \text{ADV}_{\text{bought}}(\mathcal{I}) + (1-\lambda) \text{RORO-min}_{\text{bought}}(\mathcal{I}) \\
&\quad + \lambda \text{ADV}_{\text{switch}}(\mathcal{I}) + (1-\lambda) \text{RORO-min}_{\text{switch}}(\mathcal{I}) \\
&\leq \lambda \text{ADV}(\mathcal{I}) + (1-\lambda) \text{RORO-min}(\mathcal{I})
\end{aligned}$$

Since  $\text{RORO-min} \leq \alpha \cdot \text{OPT} \leq \alpha \cdot \text{ADV}$ , we have

$$\text{R0-Advice-min}(\mathcal{I}) \leq \lambda \text{ADV}(\mathcal{I}) + (1-\lambda)\alpha \text{ADV}(\mathcal{I}) \quad (46)$$

$$\text{R0-Advice-min}(\mathcal{I}) \leq (\lambda + (1-\lambda)\alpha) \cdot \text{ADV}(\mathcal{I}) \quad (47)$$

$$\text{R0-Advice-min}(\mathcal{I}) \leq (1+\epsilon) \cdot \text{ADV}(\mathcal{I}). \quad (48)$$

Since  $\text{ADV} \leq U + 2\beta \leq \frac{\text{OPT}}{L/(U+2\beta)}$ , we have

$$\text{R0-Advice-min}(\mathcal{I}) \leq \lambda \frac{\text{OPT}(\mathcal{I})}{L/(U+2\beta)} + (1-\lambda)\alpha \text{OPT}(\mathcal{I}) \quad (49)$$

$$\text{R0-Advice-min}(\mathcal{I}) \leq \left[ \frac{\lambda(U+2\beta)}{L} + (1-\lambda)\alpha \right] \cdot \text{OPT}(\mathcal{I}) \quad (50)$$

$$\text{R0-Advice-min}(\mathcal{I}) \leq \left( \frac{(U+2\beta)/L(\alpha-1-\epsilon) + \alpha\epsilon}{(\alpha-1)} \right) \cdot \text{OPT}(\mathcal{I}). \quad (51)$$

By combining (48) and (51), the result follows.  $\square$

## C PROBLEM FORMULATION, RORO INSTANTIATION, AND PROOFS FOR OCS-MAX

In this section, we describe the deferred maximization variant of OCS (OCS-max). In Appendix C.1, we present the offline formulation and relevant assumptions for the OCS-max problem, which builds on the core problem presented in Section 2.1.

### C.1 Formulation and assumptions

In OCS-max, an online player must sell an asset with total capacity  $C$  while maximizing their profit (we again assume for notational simplicity that  $C = 1$ ). In this setting, a *concave price function*  $g_t(\cdot)$  arrives online at each time step  $t \in [T]$ , and the switching cost is subtractive rather than additive. The offline version is formalized below:

$$\max_{\{x_t\}_{t \in [T]}} \underbrace{\sum_{t=1}^T g_t(x_t)}_{\text{Selling profit}} - \underbrace{\sum_{t=1}^{T+1} \beta |x_t - x_{t-1}|}_{\text{Switching cost}}, \text{ s.t., } \underbrace{\sum_{t=1}^T x_t = 1}_{\text{Deadline constraint}}, \underbrace{x_t \in [0, d_t] \forall t \in [T]}_{\text{Rate constraint}}. \quad (52)$$

*Assumptions and additional notation.* For OCS-max, we assume that price functions  $\{g_t(\cdot)\}_{t \in [T]}$  have a bounded derivative, i.e.  $L \leq dg_t/dx_t \leq U$ , where  $L$  and  $U$  are known positive constants. Furthermore, we assume that all cost functions  $g_t(\cdot)$  are *concave* – this assumption is important as a way to model *diminishing returns*, and is empirically valid for the applications of interest.

The switching cost coefficient  $\beta$  is assumed to be known to the player, and is bounded within an interval ( $\beta \in (0, L/2)$ ). If  $\beta = 0$ , the problem is equivalent to one-way trading, and if  $\beta$  is “too large” (i.e.,  $> L/2$ ), we can show that any competitive algorithm should *only* consider the switching cost.<sup>7</sup>

*Competitive analysis.* As mentioned in Section 2.1, our goal is to design an online algorithm that maintains a small *competitive ratio* [6, 38], i.e., performs nearly as well as the offline optimal solution. Here we note that for a problem with a maximization objective (i.e., OCS-max), we define the competitive ratio as  $\max_{\mathcal{I} \in \Omega} \text{OPT}(\mathcal{I})/\text{ALG}(\mathcal{I})$ , where  $\Omega$  is the set of all feasible input instances for the problem.

The notions of *consistency* and *robustness* are similarly redefined for the maximization objective. Letting  $\text{ALG}(\mathcal{I}, \varepsilon)$  denote the cost of a learning-augmented online algorithm on input sequence  $\mathcal{I}$  when provided predictions with error factor  $\varepsilon$ , we have that consistency is defined as  $\max_{\mathcal{I} \in \Omega} \text{OPT}(\mathcal{I})/\text{ALG}(\mathcal{I}, 0)$ , and robustness is defined as  $\max_{\mathcal{I} \in \Omega} \text{OPT}(\mathcal{I})/\text{ALG}(\mathcal{I}, \mathbf{E})$ , where  $\mathbf{E}$  is a maximum error factor (or  $\infty$ ).

## C.2 Solving OCS-max using the RORO framework

We now present our instantiation of RORO for OCS-max (RORO-max), with pseudocode summarized in Algorithm 1. Since this algorithm uses the same RORO framework and shares much of the same conceptual structure as RORO-min (see Section 3.2), here we highlight a few important differences and defer detailed proofs for RORO-max to Appendix D.2 and D.3 for brevity. As previously, RORO-max leverages a dynamic threshold  $\Phi(w)$ , where  $w \in [0, 1]$  denotes the current utilization.

*Definition C.1 (RORO-max threshold function  $\Phi$ ).* For any utilization  $w \in [0, 1]$ ,  $\Phi$  is defined as:

$$\Phi(w) = L + \beta + (\omega L - L - 2\beta) \exp(\omega w), \quad (53)$$

where  $\omega$  is the competitive ratio and is defined in (57).

At each time step, a price function  $g_t$  arrives online and RORO-max solves two *pseudo-profit maximization problems*. Before defining the instantiation, we note here that, because the underlying problem is a maximization problem, the pseudo-cost function we define for the RORO framework has a flipped sign to obtain the correct result on line 6 in Algorithm 1.

*Definition C.2 (RORO instantiation for OCS-max (RORO-max)).* The RORO framework solves OCS-max when instantiated with the following parameters:

Let  $\Phi(\cdot) : [0, 1] \rightarrow [L, U]$  be the dynamic threshold defined in C.1.

<sup>7</sup>As brief justification for the bounds on  $\beta$ , consider the following solutions for OCS-max. In OCS-max, a feasible solution can have objective  $L - 2\beta$ . If  $\beta > L/2$ , our solution may have a negative objective value, and this possibility should be avoided to facilitate competitive analysis.

Then, define the ramping-on problem  $\text{RAMPON}(\cdot)$ , ramping-off problem  $\text{RAMPOFF}(\cdot)$ , and the pseudo-cost function  $\text{PCOST}(\cdot)$  as follows:

$$\text{RAMPON}(g_t(\cdot), w^{(t-1)}, x_{t-1}) = \arg \max_{x \in [x_{t-1}, \min(1-w^{(t-1)}, d_t)]} g_t(x) - \beta(x - x_{t-1}) - \int_{w^{(t-1)}}^{w^{(t-1)}+x} \phi(u) du, \quad (54)$$

$$\text{RAMPOFF}(g_t(\cdot), w^{(t-1)}, x_{t-1}) = \arg \max_{x \in [0, \min(x_{t-1}, d_t)]} g_t(x) - \beta(x_{t-1} - x) - \int_{w^{(t-1)}}^{w^{(t-1)}+x} \phi(u) du, \quad (55)$$

$$\text{PCOST}(g_t(\cdot), w^{(t-1)}, x_t, x_{t-1}) = - \left[ g_t(x_t) + \beta|x_t - x_{t-1}| - \int_{w^{(t-1)}}^{w^{(t-1)}+x_t} \phi(u) du \right]. \quad (56)$$

The optimizations defined above are inserted into the pseudocode defined in Algorithm 1 to create an instance that solves OCS-max (RORO-max). In Appendix D.1, we show that both of these optimizations are concave maximization problems, which can be solved efficiently (e.g., using iterative methods).

As previously, we give upper and lower bounds on the competitive ratio of RORO-max in Theorems C.1 and C.2, respectively. Both proofs are deferred to Appendices D.2 and D.3 for brevity.

**THEOREM C.1.** *Algorithm 1 for OCS-max (RORO-max) is  $\omega$ -competitive when the threshold function is given by  $\Phi(w)$  (from Definition C.1), where  $\omega$  is the solution to  $\frac{U-L-2\beta}{\omega L-L-2\beta} = \exp(\omega)$  and is given by*

$$\omega := W \left( \frac{U/L - 1 - 2\beta/L}{e^{1+2\beta/L}} \right) + 1 + \frac{2\beta}{L}. \quad (57)$$

In the above,  $W(\cdot)$  is the Lambert  $W$  function, defined as the inverse of  $f(x) = xe^x$ .

**THEOREM C.2.** *No deterministic online algorithm for OCS-max can achieve a competitive ratio better than  $\omega$ , as defined in (57).*

By combining Theorems C.1 and C.2, we conclude that RORO-max is optimal for OCS-max. We note that the competitive ratio achievable in the maximization setting (i.e.,  $\omega$  in (57)) is not the same as the competitive ratio achievable in the minimization setting (i.e.,  $\alpha$  in (15)). This distinction aligns with the existing results for online search problems such as one-way trading [16] and  $k$ -search [34], which likewise have different competitive ratios in the maximization and minimization settings. As previously, we note that when  $\beta = 0$ , our RORO algorithm recovers the optimal competitive ratio for the maximization variant of one-way trading, which is  $\omega \sim 1 + W((U/L-1)/e)$ , proven in [34, 53]. The competitive ratio degrades as  $\beta$  grows, but the assumed upper bound on  $\beta$  prevents it from becoming unbounded.

### C.3 RO-Advice-max: robustly incorporating machine-learned advice for OCS-max

We now present our instantiation of the learning-augmented RO-Advice framework for OCS-max (RO-Advice-max), with pseudocode summarized in Algorithm 2. Since this algorithm uses the same RO-Advice framework and shares much of the same conceptual structure as RO-Advice-min (see Section 4.2), here we highlight differences and defer detailed proofs to Appendix D.4 for brevity.

*Definition C.3* (RO-Advice instantiation for OCS-max (RO-Advice-max)). Let  $\epsilon \in [0, \omega - 1]$ , where  $\omega$  is the robust competitive ratio defined in (57). RO-Advice-max then sets a combination factor  $\lambda := \left(\frac{\omega}{1+\epsilon} - 1\right) \cdot \frac{1}{\omega-1}$ , which is in  $[0, 1]$ . The robust advice  $\{\tilde{x}_t\}_{t \in [T]}$  is given by the RORO instantiation for OCS-max (RORO-max), given in Definition C.2.

In the following theorem, we state the consistency and robustness bounds for the RO-Advice-max meta-algorithm. We prove this result and discuss its significance in Appendix D.4.



**THEOREM C.3.** *Given a parameter  $\epsilon \in [0, \omega - 1]$ , where  $\omega$  is defined as in (57), RO-Advice-max is  $(1 + \epsilon)$ -consistent and  $\left(\frac{(\omega-1)(1+\epsilon)}{\epsilon + \frac{(L-2\beta)}{U}(\omega-1-\epsilon)}\right)$ -robust for OCS-max.*

## D OCS-MAX PROOFS

We now prove several of the results described in Appendix C.

In Appendix D.1, we show that the ramping-on and ramping-off optimization problems used in the RORO instantiations for OCS-max can be efficiently solved using iterative convex optimization techniques.

In Appendix D.2, we prove the competitive upper bound for RORO-max (Theorem C.1). In Appendix D.3, we provide the proof of the information-theoretic lower bound for OCS-max (Theorem C.2), which subsequently proves that RORO-max is optimal.

In Appendix D.4, we prove the consistency and robustness bounds for RO-Advice-max (Theorem C.3).

### D.1 Efficiently solving the ramping-on and ramping-off problems for RORO-max

Recall the ramping-on and ramping-off problems given by (55) and (56), respectively. First, we note that because the primary difference between the two problems is the restriction on the decision space, if we are only interested in the actual online decision  $x_t$ , it is valid to merge these problems and consider a single optimization problem as follows:

$$\text{RAMPONRAMPOFF}(g_t(\cdot), w^{(t-1)}, x_{t-1}) = \arg \max_{x \in [0, \min(1 - w^{(t-1)}, d_t)]} g_t(x) - \beta|x - x_{t-1}| - \int_{w^{(t-1)}}^{w^{(t-1)}+x} \Phi(u) du.$$

Let us define  $f_t(x) : t \in [T]$  as the right-hand side of the combined ramping-on and ramping-off maximization problem defined above:

$$f_t(x) = g_t(x) - \beta|x - x_{t-1}| - \int_{w^{(t-1)}}^{w^{(t-1)}+x} \Phi(u) du. \quad (58)$$

**THEOREM D.1.** *Under the assumptions of OCS-max,  $f_t(x)$  is concave on the entire interval  $x \in [0, 1]$  for any  $t \in [T]$ .*

**PROOF.** We prove the above statement by contradiction.

By definition, the sum of two concave functions gives a concave function. First, note that the switching cost term can be equivalently defined in terms of the  $\ell_1$  norm as follows:  $\beta\|x - x_{t-1}\|_1$ . By definition and by observing that  $x_{t-1}$  is fixed,  $\beta\|x - x_{t-1}\|_1$  is convex. We have also assumed as part of the OCS-max problem setting that each  $g_t(x)$  is concave. Since the negation of a convex function is concave, we have that  $g_t(x) - \beta\|x - x_{t-1}\|_1$  must be concave.

We turn our attention to the term  $-\int_{w^{(t-1)}}^{w^{(t-1)}+x} \Phi(u) du$ . Let  $k(x) = \int_{w^{(t-1)}}^{w^{(t-1)}+x} \Phi(u) du$ .

By the fundamental theorem of calculus,  $d/dx k(x) = \Phi(z^{(t-1)} + x) \cdot d/dx x = \Phi(z^{(t-1)} + x)$

Let  $p(x) = \Phi(z^{(t-1)} + x)$ . Then  $d^2/d^2x k(x) = p'(x) \cdot d/dx x$ . Since  $\Phi$  is monotonically increasing on the interval  $[0, 1]$ , we know that  $p'(x) > 0$ , and thus  $d/dx x \cdot p'(x)$  is positive. This gives that  $k(x)$  is convex in  $x$ .

Since the negation of a convex function is concave, this causes a contradiction, because the sum of two concave functions is a concave function. Note that  $(g_t(x) - \beta\|x - x_{t-1}\|_1)$  and  $\left(-\int_{w^{(t-1)}}^{w^{(t-1)}+x} \Phi(u) du\right)$  are both concave.

Thus,  $f_t(\cdot) = g_t(x) - \beta|x - x_{t-1}| - \int_{w^{(t-1)}}^{w^{(t-1)}+x} \Phi(u)du$  is always concave under the assumptions of the problem setting.  $\square$

Since the right-hand side of the combined ramping-on and ramping-off maximization problem is concave, (55) and (56) can be solved efficiently by casting the concave maximization as a convex minimization problem, and using an iterative convex optimization method.

## D.2 Competitive results for RORO-max

In the following, we prove Theorem C.1, which states that the instantiation of RORO for OCS-max is  $\omega$ -competitive, where  $\omega$  is as defined in (57).

**PROOF OF THEOREM C.1.** Let  $\mathcal{I} \in \Omega$  denote any valid OCS-max sequence, and let  $w^{(j)}$  denote the final utilization before the compulsory trading, which begins at time step  $j \leq T$ . Note that  $w^{(t)} = \sum_{m \in [t]} x_m$  is non-decreasing over  $t$ .

**LEMMA D.2.** *The offline optimum is upper bounded by  $\text{OPT}(\mathcal{I}) \leq \Phi(w^{(j)}) + \beta$ .*

*Proof of Lemma D.2.* We prove this lemma by contradiction. Note that the offline optimum is to trade all items at the best price function (ignoring additional switching costs) over the sequence  $\{g_t(\cdot)\}_{t \in [T]}$ .

Suppose this maximum price function is at an arbitrary step  $m$  ( $m \in [T]$ ,  $m \leq j$ ), denoted by  $g_m(\cdot)$ . Since cost functions are concave and additionally satisfy the conditions  $g_m(0) = 0$  and  $g_m(x) \geq 0 \forall x \in [0, 1]$ , the derivative of  $g_m$  at  $x = 0$  is an upper bound on the best marginal return (i.e. per unit sale) that the optimal solution can obtain. We henceforth denote this derivative by  $\frac{dg_m}{dx}$ . Assume for the sake of contradiction that  $\frac{dg_m}{dx} = \text{OPT}(\mathcal{I}) > \Phi(w^{(j)}) + \beta$ .

Next, we consider  $\frac{dg_m}{dx} + \beta > \Phi(w^{(j)}) + 2\beta$ . Since  $\Phi(z)$  is strictly increasing on  $z \in [0, 1]$ , it follows that  $\Phi(w^{(j)}) + 2\beta \geq \Phi(w^{(m)})$ , as  $m \leq j$ . By solving the ramping-off problem, we have  $x_m^- = x_{m-1}$ . Thus, the online solution of step  $m$  is dominated by  $x_m^+$ , i.e.,  $x_m = x_m^+$ .

Furthermore, since  $\frac{dg_m}{dx} > \Phi(w^{(j)}) + \beta$  and  $\Phi(w^{(j)}) + \beta > \Phi(w^{(m)}) + \beta$  as previously, by solving the ramping-on problem, we must have that the resulting decision satisfies  $x_m^+ > w^{(j)} - w^{(m-1)}$ . This implies that  $w^{(m)} > w^{(j)}$ , which, given that  $w^{(t)}$  is non-decreasing in  $t \in [T]$ , contradicts with the assumption that the final utilization before the compulsory trade is  $w^{(j)}$ .

Thus, we conclude that  $\text{OPT}(\mathcal{I}) \leq \Phi(w^{(j)}) + \beta$ .

**LEMMA D.3.** *The return of RORO-max( $\mathcal{I}$ ) is lower bounded by*

$$\text{RORO-max}(\mathcal{I}) \geq \int_0^{w^{(j)}} \Phi(u)du - \beta w^{(j)} + (1 - w^{(j)})L. \quad (59)$$

*Proof of Lemma D.3.* By solving the ramping-off problem for any arbitrary time step  $t \in [T]$ , we can observe that  $g_t(x_t^-) + \beta x_t^- \geq \int_{w^{(t-1)}}^{w^{(t-1)}+x_t^-} \Phi(u)du, \forall t \in [T]$ . Therefore, we have the following inequality

$$\max\{r_n^+, r_n^-\} \geq r_n^- \geq -\beta x_{t-1}, \forall t \in [T]. \quad (60)$$

Thus, we have

$$-\beta w^{(j)} \leq -\beta w^{(t-1)} = \sum_{t \in [T]} (-\beta x_{t-1}) \leq \sum_{t \in [T]} \max\{r_n^+, r_n^-\} \quad (61)$$

$$= \sum_{t \in [T]} \left[ g_t(x_t) - \beta |x_t - x_{t-1}| - \int_{w^{(t-1)}}^{w^{(t-1)} + x_t} \Phi(u) du \right] \quad (62)$$

$$= \sum_{t \in [T]} [g_t(x_t) - \beta |x_t - x_{t-1}|] - \int_0^{w^{(j)}} \Phi(u) du \quad (63)$$

$$= \text{RORO-max}(\mathcal{I}) - (1 - w^{(j)})L - \int_0^{w^{(j)}} \Phi(u) du. \quad (64)$$

Combining Lemma D.2 and Lemma D.3 gives

$$\text{CR} \leq \frac{\text{OPT}(\mathcal{I})}{\text{RORO-max}(\mathcal{I})} \leq \frac{\Phi(w^{(j)}) + \beta}{\int_0^{w^{(j)}} \Phi(u) du - \beta w^{(j)} + (1 - w^{(j)})L} \leq \omega, \quad (65)$$

where the last inequality holds since for any  $w \in [0, 1]$

$$\int_0^w \Phi(u) du - \beta w + (1 - w)L = \int_0^w [L + \beta + (\omega L - L - 2\beta) \exp(\omega w)] - \beta w + (1 - w)L \quad (66)$$

$$= (L + \beta)w + \frac{1}{\omega} (\omega L - L - 2\beta) [\exp(\omega w) - 1] - \beta w + (1 - w)L \quad (67)$$

$$= \frac{1}{\omega} [L + 2\beta + (\omega L - L - 2\beta) \exp(\omega w)] \quad (68)$$

$$= \frac{1}{\omega} [\Phi(w) + \beta]. \quad (69)$$

We note that the rate constraints  $\{d_t\}_{t \in [T]}$  surprisingly do not appear in this worst-case analysis. For completeness, we state and prove Lemma D.4.

**LEMMA D.4.** *If  $d_t < 1 \forall t \in [T]$ , the competitive ratio of RORO-max is still upper bounded by  $\omega$ .*

**PROOF OF LEMMA D.4.** Suppose that the presence of a rate constraint  $d_t$  causes RORO-max to make a decision which violates  $\omega$ -competitiveness. At time  $t$ , the only difference between the setting where  $x_t \in [0, 1]$  and the setting with rate constraints  $< 1$ , where  $x_t \in [0, d_t]$ , is that  $x_t$  cannot be  $> d_t$ .

This implies that a challenging situation for RORO-max under a rate constraint is the case where RORO-max would otherwise accept  $> d_t$  of a good price function, but it cannot due to the rate constraint.

We can now show that such a situation implies that RORO-max achieves a competitive ratio which is strictly better than  $\omega$  (in the maximization setting).

From (65), we know that the following holds for any value of  $w \in [0, 1]$ :

$$\int_0^w \Phi(u) du - \beta w + (1 - w)L \geq \frac{1}{\omega} [\Phi(w) + \beta].$$

For an arbitrary instance  $\mathcal{I} \in \Omega$  and an arbitrary time step  $t$ , let  $w^{(t)} = w^{(t-1)} + d_t$ , implying that  $x_t = d_t$ . For the sake of comparison, we first consider this time step with a price function  $g_t(\cdot)$ , such that  $g_t(x_t) = \Phi(w^{(t)}) \cdot x_t$ , implying that without the presence of a rate constraint, RORO-max

would set  $x_t = d_t$ . If no more price functions are accepted by RORO-max after time step  $t$  (excepting the compulsory trade), we have the following:

$$\frac{\text{OPT}(\mathcal{I})}{\text{RORO-max}(\mathcal{I})} \leq \frac{\Phi(w^{(t)}) + \beta}{\int_0^{w^{(t)}} \Phi(u) du - \beta w^{(t)} + (1 - w^{(t)})L} = \omega. \quad (70)$$

Now, consider the exact same setting as above, except with a new price function  $g'_t(\cdot)$ , such that  $g'_t(x_t) > \Phi(w^{(t)}) \cdot x_t$ . This implies that without the presence of a rate constraint, RORO-max would set  $x_t > d_t$ . In other words,  $g'_t(\cdot)$  is a good price function which RORO-max cannot accept more of due to the rate constraint.

Since OPT is subject to the same rate constraint  $d_t$ , we know that  $\text{OPT}(\mathcal{I})$  is upper bounded by  $[\Phi(w^{(t)}) + \beta](1 - d_t) + g'_t(d_t)$  – the rest of the optimal solution is bounded by the final threshold value, since we assume that no more prices are accepted by RORO-max after time step  $t$ .

The profit of RORO-max is lower bounded by  $\text{RORO-max}(\mathcal{I}) \geq \int_0^{w^{(t)}} \Phi(u) du - \int_{w^{(t-1)}}^{w^{(t)}} \Phi(u) du + g'_t(d_t) - \beta w^{(t)} + (1 - w^{(t)})L$ .

Observe that compared to the previous setting, the OPT and RORO-max solutions have both increased – OPT( $\mathcal{I}$ ) has increased by a additive factor of  $g'_t(d_t) - [\Phi(w^{(t)}) + \beta]d_t$ , while  $\text{RORO-max}(\mathcal{I})$  has increased by a additive factor of  $g'_t(d_t) - \int_{w^{(t-1)}}^{w^{(t)}} \Phi(u) du$ .

Since  $\Phi$  is monotonically increasing on  $w \in [0, 1]$ , we have that by definition,  $[\Phi(w^{(t)}) + \beta]d_t > \int_{w^{(t-1)}}^{w^{(t)}} \Phi(u) du$ . Thus, the solution obtained by RORO-max has increased *more* than the solution obtained by OPT. This then implies the following:

$$\frac{\text{OPT}(\mathcal{I})}{\text{RORO-max}(\mathcal{I})} \leq \frac{[\Phi(w^{(t)}) + \beta](1 - d_t) + g'_t(d_t)}{\int_0^{w^{(t)}} \Phi(u) du - \int_{w^{(t-1)}}^{w^{(t)}} \Phi(u) du + g'_t(d_t) - \beta w^{(t)} + (1 - w^{(t)})L} < \omega,$$

where the final inequality follows from (70).  $\square$

At a high-level, this result shows that even if there is a rate constraint which prevents RORO-max from accepting a good price function, the worst-case competitive ratio does not change. Combining Lemmas D.2, D.3, and D.4 completes the proof.  $\square$

### D.3 Lower bound for OCS-max

In the following, we prove Theorem C.2, which states that for any deterministic online algorithm solving OCS-max, the optimal competitive ratio is  $\omega$ , where  $\omega$  is as defined in (57).

As previously, we define a family of special instances and then show that the competitive ratio for any deterministic algorithm is lower bounded under these instances. Prior work has shown that difficult instances for online search problems with a maximization objective occur when inputs arrive at the algorithm in an increasing order of cost [16, 27, 34, 55]. For OCS-max, we additionally consider how an adaptive adversary can essentially force an algorithm to incur a large switching cost in the worst-case. We now formalize such a family of instances  $\{\mathcal{I}_x\}_{x \in [L, U]}$ , where  $\mathcal{I}_x$  is called an  $x$ -increasing instance.

**DEFINITION D.5 ( $x$ -INCREASING INSTANCE FOR OCS-max).** *Let  $n, m \in \mathbb{N}$  be sufficiently large, and  $\delta := (U-L)/n$ . For  $x \in [L, U]$ ,  $\mathcal{I}_x \in \Omega$  is an  $x$ -instance if it consists of  $n_x := 2 \cdot \lceil (x-L)/\delta \rceil + 1$  alternating batches of linear price functions. The  $i$ -th batch ( $i \in [n_x - 2]$ ) contains  $m$  price functions with coefficient  $L$  if  $i$  is odd, and 1 price function with coefficient  $L + (\lceil i/2 \rceil)\delta$  if  $i$  is even. The last two batches of items consist of  $m$  price functions with coefficient  $x - \varepsilon$ , followed by  $m$  price functions with coefficient  $L$ .*

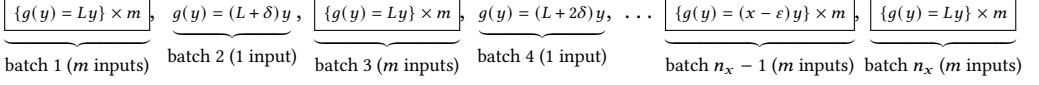


Fig. A5.  $\mathcal{I}_x$  consists of  $n_x$  batches of price functions, where the alternating single functions are continuously increasing from  $L$  up to  $x$ .

Note that  $\mathcal{I}_L$  is simply a stream of  $m$  price functions with coefficient  $L$ . See Fig. A5 for an illustration of an  $x$ -increasing instance. Since a competitive algorithm ALG should not accept the worst-case price function with coefficient  $L$  unless it is forced to, this special  $x$ -increasing instance can be equivalently interpreted as an *adaptive adversary*, which gives ALG  $m$  worst-case inputs  $L$  whenever any price function is accepted. As  $n \rightarrow \infty$ , the alternating single price functions in an  $x$ -increasing sequence continuously increase up to  $x$ , and each of these “good price functions” is interrupted by a section of worst-case  $L$  price functions. Note that the last few price functions in an  $x$ -increasing instance are always  $L$ .

**PROOF OF THEOREM C.2.** Let  $h(x)$  denote a *conversion function*  $[L, U] \rightarrow [0, 1]$ , which fully describes the actions of a deterministic ALG for OCS-max on an  $x$ -increasing instance  $\mathcal{I}_x$ . Note that for large  $n$ , processing the instance  $\mathcal{I}_{x+\delta}$  is equivalent to first processing  $\mathcal{I}_x$  (besides the last two batches), and then processing batches with prices  $x + \delta$  and  $U$ . Since ALG is deterministic and the conversion is unidirectional (irrevocable), we must have that  $h(x + \delta) \geq h(x)$ , i.e.  $h(x)$  is non-decreasing in  $[L, U]$ . Intuitively, the entire capacity should be satisfied if the maximum possible price is observed, i.e.  $h(U) = 1$ .

For instance  $\mathcal{I}_x$ , the optimal offline solution is  $\text{OPT}(\mathcal{I}_x) = x - \frac{2}{m}\beta$ . Note that as  $m$  grows large,  $\text{OPT}(\mathcal{I}_x) \rightarrow x$ .

Due to the adaptive nature of each  $x$ -instance, any deterministic ALG incurs a switching cost proportional to  $h(x)$ , which gives the amount of utilization used by ALG before the end of the sequence on instance  $\mathcal{I}_x$ .

Whenever ALG accepts some price  $L + (\lceil i/2 \rceil)\delta$ , the adversary presents prices  $L$  starting in the next time step. Any ALG which does not switch away immediately obtains a competitive ratio strictly worse than an algorithm which does switch away (if an algorithm accepts  $c$  fraction of a good price, the switching cost it will pay is  $2\beta c$ . An algorithm may continue accepting  $c$  fraction of prices  $L$  in the subsequent time steps, but a sequence exists where this decision will take up too much utilization to recover when better prices are presented later. In the extreme case, if an algorithm continues accepting  $c$  fraction of these  $L$  prices, it might fill its utilization and then OPT can obtain an arbitrarily good profit)

Since accepting any price by a factor of  $c$  incurs a switching cost of  $2\beta c$ , the switching cost (before the final compulsory trade) paid by ALG on instance  $\mathcal{I}_x$  is  $2\beta h(x)$ . We assume that ALG is notified of the compulsory trade, and does not incur a significant switching cost during the final batch.

Then the total profit of an  $\omega^*$ -competitive online algorithm ALG on instance  $\mathcal{I}_x$  is  $\text{ALG}(\mathcal{I}_x) = h(\omega^*L)\omega^*L + \int_{\omega^*L}^x udh(u) - 2\beta h(x) + (1 - h(x))L$ , where  $udh(u)$  is the profit of selling  $dh(u)$  utilization at price  $u$ , the last term is from the compulsory conversion, and the second to last term is the switching cost incurred by ALG. Note that any deterministic ALG which makes conversions when the price is larger than  $\omega^*L$  can be strictly improved by restricting conversions to prices  $\geq \omega^*L$ .

For any  $\omega^*$ -competitive online algorithm, the corresponding conversion function  $h(\cdot)$  must satisfy  $\text{ALG}(\mathcal{I}_x) \geq \text{OPT}(\mathcal{I}_x)/\omega^* = x/\omega^*, \forall x \in [L, U]$ . This gives a necessary condition which the

conversion function must satisfy as follows:

$$\text{ALG}(\mathcal{I}_x) = h(\omega^*L)\omega^*L + \int_{\omega^*L}^x u dh(u) - 2\beta h(x) + (1 - h(x))L \geq x/\omega^*, \quad \forall x \in [L, U].$$

By integral by parts, the above implies that the conversion function must satisfy  $h(x) \geq \frac{x/\omega^* - L}{x - 2\beta - L} + \frac{1}{x - 2\beta - L} \int_{\omega^*L}^x h(u) du$ . By Grönwall's Inequality [40, Theorem 1, p. 356], we have that

$$\begin{aligned} h(x) &\geq \frac{x/\omega^* - L}{x - 2\beta - L} + \frac{1}{x - 2\beta - L} \int_{\omega^*L}^x \frac{u/\omega^* - L}{u - 2\beta - L} \cdot \exp\left(\int_u^x \frac{1}{r - 2\beta - L} dr\right) du \\ &\geq \frac{x/\omega^* - L}{x - 2\beta - L} + \int_{\omega^*L}^x \frac{u/\omega^* - L}{(u - 2\beta - L)^2} du \\ &\geq \frac{x/\omega^* - L}{x - 2\beta - L} + \left[ \frac{L\omega^* - 2\beta - L}{\omega^*u - 2\beta\omega^* - L\omega^*} + \frac{1}{\omega^*} \ln(u - 2\beta - L) \right]_{\omega^*L}^x \\ &\geq \frac{1}{\omega^*} \ln(x - 2\beta - L) - \frac{1}{\omega^*} \ln(\omega^*L - 2\beta - L), \quad \forall x \in [L, U]. \end{aligned}$$

By the problem's definition,  $h(U) = 1$  – we can combine this with the above constraint to give the following condition for an  $\omega^*$ -competitive online algorithm:

$$\frac{1}{\omega^*} \ln(U - 2\beta - L) - \frac{1}{\omega^*} \ln(\omega^*L - 2\beta - L) \leq h(U) = 1.$$

The optimal  $\omega^*$  is obtained when the above inequality is binding, so solving for the value of  $\omega^*$  which solves  $1/\omega^* \ln(U - 2\beta - L) - 1/\omega^* \ln(\omega^*L - 2\beta - L) = 1$  yields that the best competitive ratio for any ALG solving OCS-max is  $\omega^* \geq W\left(\frac{U/L - 1 - 2\beta/L}{e^{1+2\beta/L}}\right) + \frac{2\beta}{L} + 1$ .  $\square$

#### D.4 RO-Advice consistency and robustness for OCS-max

In the following, we prove Theorem C.3, which states that the instantiation of RO-Advice for OCS-max is  $(1 + \epsilon)$ -consistent and  $\left(\frac{(\omega-1)(1+\epsilon)}{\epsilon + \frac{(L-2\beta)}{U}(\omega-1-\epsilon)}\right)$ -robust, where  $\epsilon$  and  $\omega$  are defined in Definition C.3 and (57), respectively.

**PROOF OF THEOREM C.3.** We start by noting that the online solution given by RO-Advice-max is always feasible (under the assumption that  $\sum_{t=1}^T \hat{x}_t = 1$ ), since

$$\sum_{t=1}^T x_t = \sum_{t=1}^T [\lambda \hat{x}_t + (1 - \lambda) \tilde{x}_t] = \lambda + (1 - \lambda) = 1.$$

Let  $\mathcal{I} \in \Omega$  be an arbitrary valid OCS-max sequence. We denote the *selling* and *switching* costs of the robust advice by  $\text{RORO-max}_{\text{sold}}$  and  $\text{RORO-max}_{\text{switch}}$ , respectively. Likewise, the purchasing and switching costs of the black box advice are denoted by  $\text{ADV}_{\text{sold}}$  and  $\text{ADV}_{\text{switch}}$ .

The cost of RO-Advice-max is bounded by

$$\begin{aligned}
\text{RO-Advice-max}(\mathcal{I}) &= \sum_{t=1}^T g_t(x_t) - \sum_{t=1}^{T+1} \beta |x_t - x_{t-1}| \\
&= \sum_{t=1}^T g_t(\lambda \hat{x}_t + (1-\lambda)\tilde{x}_t) - \sum_{t=1}^{T+1} \beta |\lambda \hat{x}_t + (1-\lambda)\tilde{x}_t - \lambda \hat{x}_{t-1} - (1-\lambda)\tilde{x}_{t-1}| \\
&\geq \lambda \sum_{t=1}^T g_t(\hat{x}_t) + (1-\lambda) \sum_{t=1}^T g_t(\tilde{x}_t) \\
&\quad - \sum_{t=1}^{T+1} \beta |\lambda \hat{x}_t - \lambda \hat{x}_{t-1}| - \sum_{t=1}^{T+1} \beta |(1-\lambda)\tilde{x}_t - (1-\lambda)\tilde{x}_{t-1}| \\
&\geq \lambda \text{ADV}_{\text{sold}}(\mathcal{I}) + (1-\lambda) \text{RORO-max}_{\text{sold}}(\mathcal{I}) \\
&\quad - \lambda \sum_{t=1}^{T+1} \beta |\hat{x}_t - \hat{x}_{t-1}| - (1-\lambda) \sum_{t=1}^{T+1} \beta |\tilde{x}_t - \tilde{x}_{t-1}| \\
&\geq \lambda \text{ADV}_{\text{sold}}(\mathcal{I}) + (1-\lambda) \text{RORO-max}_{\text{sold}}(\mathcal{I}) \\
&\quad - \lambda \text{ADV}_{\text{switch}}(\mathcal{I}) - (1-\lambda) \text{RORO-max}_{\text{switch}}(\mathcal{I}) \\
&\geq \lambda \text{ADV}(\mathcal{I}) + (1-\lambda) \text{RORO-max}(\mathcal{I})
\end{aligned}$$

Since  $\text{RORO-max} \geq \text{OPT}/\omega \geq \text{ADV}/\omega$ , we have

$$\text{RO-Advice-max}(\mathcal{I}) \geq \lambda \text{ADV}(\mathcal{I}) + (1-\lambda) (\text{ADV}(\mathcal{I})/\omega) \quad (71)$$

$$\text{RO-Advice-max}(\mathcal{I}) \geq \left( \lambda + \frac{1-\lambda}{\omega} \right) \cdot \text{ADV}(\mathcal{I}) \quad (72)$$

$$\frac{1}{\lambda + \frac{1-\lambda}{\omega}} \cdot \text{RO-Advice-max}(\mathcal{I}) \geq \text{ADV}(\mathcal{I}) \quad (73)$$

$$(1+\epsilon) \cdot \text{RO-Advice-max}(\mathcal{I}) \geq \text{ADV}(\mathcal{I}). \quad (74)$$

Since  $\text{ADV} \geq L - 2\beta \geq ((L-2\beta)/U)\text{OPT}$ , we have

$$\text{RO-Advice-max}(\mathcal{I}) \geq \lambda \frac{(L-2\beta)\text{OPT}(\mathcal{I})}{U} + (1-\lambda) (\text{OPT}(\mathcal{I})/\omega) \quad (75)$$

$$\text{RO-Advice-max}(\mathcal{I}) \geq \left[ \frac{\lambda(L-2\beta)}{U} + \frac{1-\lambda}{\omega} \right] \cdot \text{OPT}(\mathcal{I}) \quad (76)$$

$$\left[ \frac{1}{\frac{\lambda(L-2\beta)}{U} + \frac{1-\lambda}{\omega}} \right] \cdot \text{RO-Advice-max} \geq \text{OPT}(\mathcal{I}) \quad (77)$$

$$\left[ \frac{U\omega}{U - U\lambda + (L-2\beta)\lambda\omega} \right] \cdot \text{RO-Advice-max} \geq \text{OPT}(\mathcal{I}) \quad (78)$$

$$\left[ \frac{(\omega-1)(1+\epsilon)}{\epsilon + (\frac{L}{U} - \frac{2\beta}{U})(\omega-1-\epsilon)} \right] \cdot \text{RO-Advice-max} \geq \text{OPT}(\mathcal{I}). \quad (79)$$

Note that as  $\epsilon \rightarrow \omega - 1$ , both the consistency and robustness bounds approach  $\omega$ , which is the optimal competitive bound shown for the setting without advice. When  $\epsilon \rightarrow 0$ , the robustness bound degrades to  $U/(L-2\beta)$ , but the consistency bound implies that perfect advice allows RO-Advice-max to obtain the optimal solution.

□