

Report on

# Online Learning for Min Sum Set Cover and Pandora's Box

Kasionis Ioannis  
University of Piraeus  
ioannis.kasionis@gmail.com

Triantafyllos Petros  
University of Piraeus  
petrostriantafyllos@outlook.com

January 2025

## Abstract

This report presents an analysis and summary of the paper *Online Learning for Min Sum Set Cover and Pandora's Box* which provides a simple but effective skeleton for crafting online learning algorithms for the *Pandora's Box Problem*, the *Min Sum Set Cover Problem (MSSC)*, and other related problems in the field of *Stochastic optimization* [1]. Traditionally, these problems assume prior knowledge of value distributions. However, in this work, we explore an online learning approach where we are presented with information over multiple rounds.

Furthermore, the research, extends this framework to a bandit setting where only the values of the opened boxes are shown to the learner after each round. This extension also encompasses other alternatives of the above-mentioned problems [1].

# 1 Introduction

The two essential problems in *Stochastic Optimization* are the *Pandora's Box*, and *MSSC* problems. Introduced, by Weitzman [2] the *Pandora's Box* problem offers a general framework for optimizing decisions under uncertainty, thus the goal is to minimize the cost of uncovering information. *MSSC*, a variation of *Pandora's box*, deals with scenarios where values are either 0 or  $\infty$  [3].

While the traditional formulation of these problems assumes access to known stochastic distributions [1], this study extends these frameworks to *online* settings, where data arrives sequentially without prior distributional knowledge.

The *online learning* perspective brings forth several challenges and opportunities. In this model, adversarial instances or independent scenarios are presented iteratively, and the algorithm must decide on the next action based on the current state and previous observations. The performance metric typically revolves around *regret*, measuring the algorithm's cumulative cost relative to an optimal offline policy [4].

The work's primary contributions are as follows:

1. Propose a computationally efficient online algorithm that achieves constant-competitive guarantees for both MSSC and Pandora's Box generalizations in full-information settings. This includes problems with complex constraints such as matroid bases [5].
2. Extend the analysis to *bandit settings*, where only partial feedback about the explored options is available [6]. Despite these limited observations, our algorithm maintains a logarithmic regret bound, providing robust performance guarantees.
3. Leveraging convex relaxations and online convex optimization techniques, demonstrate practical approximation strategies that bridge theoretical insights with computational efficiency [7].

The implications of these results are twofold. First, they broaden the applicability of MSSC and Pandora's Box to dynamic, real-world contexts where decisions must adapt to unfolding events. Second, the integration of online convex optimization frameworks highlights the role of mathematical tools in addressing combinatorial challenges [1].

The remainder of the paper is organized as follows. Section 2 presents the problem formulations and related work, Section 3 details the proposed methods, and Sections 4 and 5 provide theoretical analyses and experimental evaluations, respectively. It concludes with discussions on potential future extensions.

## 2 Problem Definitions

### 2.1 Pandora’s Box Problem

The Pandora’s Box problem involves a set of  $n$  boxes, each containing an unknown value and an associated cost to open. The objective is to determine an order in which to open the boxes to minimize the combined cost of inspection and the value of the selected item. Specifically, given a distribution of potential values and costs for each box, the goal is to identify an optimal strategy that balances exploration (opening boxes) and exploitation (choosing the best available value).

In the online variant considered in this work, values and costs for the boxes are revealed adversarially over  $T$  rounds, requiring the learner to adopt strategies that adapt to newly presented information while minimizing regret against an optimal offline policy.

### 2.2 Min Sum Set Cover (MSSC)

The MSSC problem is a special case of Pandora’s Box where the values (or costs) inside the boxes are restricted to either 0 or  $\infty$ . The MSSC is particularly relevant in applications where binary outcomes dictate whether an item satisfies certain criteria.

### 3 Methodologies and Results

The authors developed a three-step framework inspired by the scenario-aware relaxations in [?], but adapted to an *online* setting. This framework is applied to two benchmarks—*non-adaptive* (NA) and *partially adaptive* (PA)—under both full-information and bandit feedback.

#### Scenario-Aware Relaxation and Rounding

First, they formulate each round’s problem (selecting one box,  $k$  boxes, or a matroid basis) as a *scenario-aware* convex program whose fractional solution captures the idea of “opening” or “selecting” boxes. Building on [?], this relaxation is crafted so that a single rounding algorithm can be applied *irrespective of the scenario* revealed. In particular, an  $\alpha$ -approximate rounding ensures that converting fractional solutions to integral ones (i.e., actual permutations or subsets of boxes) inflates cost by at most a constant factor  $\alpha$ . This idea eliminates the need to re-optimize for each scenario and is key to preserving no-regret guarantees.

#### Online Convex Optimization for Full Information vs. PA

When facing a *partially adaptive* benchmark in a *full-information* setting, the authors adopt *Follow the Regularized Leader* (FTRL). At each round, since the cost function for the entire convex relaxation is observed, they update their fractional decision by minimizing the sum of past losses plus a strongly convex regularizer. Replacing the gradient-descent approach of earlier works with FTRL allows tighter regret bounds. Crucially, once a fractional solution is chosen, it is rounded using the scenario-independent  $\alpha$ -approximation procedure. In the special case of MSSC, this method recovers the *tight 4-approximation* known offline, showing that their approach effectively matches best-known guarantees while handling adversarial input sequences.

Moreover, against a partially adaptive benchmark, these methods yield  $\alpha$ -approximate no-regret for Pandora’s Box-type objectives under diverse constraints. The paper reports:

$$\begin{cases} 9.22\text{-approx. no-regret for choosing 1 box,} \\ O(1)\text{-approx. no-regret for choosing } k \text{ boxes,} \\ O(\log k)\text{-approx. no-regret for choosing a matroid basis.} \end{cases}$$

These results hold in the full-information model, demonstrating constant or logarithmic blow-ups relative to an optimal partially adaptive strategy in hindsight.

#### Bandit vs. PA via Mixing FTRL and “Open-All”

Under *bandit feedback*, the learner typically only sees  $f^{(t)}(x_t)$  at the chosen point  $x_t$ , and thus must estimate losses without observing the full function each round. To compensate for missing information, the authors’ algorithm randomly alternates between:

1. *Opening all  $n$  boxes* (paying an upfront cost) to learn the entire cost function for that round.
2. *Applying an OCO step* (similar to FTRL) using only partial information from the boxes actually opened in the other rounds.

This mixed strategy balances exploration and exploitation and preserves the same rounding-based  $\alpha$ -approximation factor, ensuring *sublinear* regret compared to a PA benchmark. Intuitively, if we never open all boxes, we risk poor gradient estimates; if we always open them, we overspend on inspection. Balancing these extremes provides robust performance with constant or  $O(\log k)$  approximate ratios mirroring the full-information case. Specifically the same as in the previous full information case.

## Bandit vs. NA and the Role of Non-Lipschitz Costs

A *non-adaptive* benchmark fixes one subset of boxes throughout all rounds. The cost function here is often *non-Lipschitz*, preventing direct application of FTRL. Instead, the paper introduces an *explore-exploit* procedure in which:

1. **Explore:** The algorithm fully opens boxes in certain rounds, discovering a new constraint (e.g., a threshold or cost range). It adds this to a global linear program and uses an ellipsoid subroutine to find a feasible fractional solution.
2. **Exploit:** That fractional solution is then rounded into a non-adaptive set of boxes, used until the next exploration.

**Theorem 5.3** states that if there exists a partially adaptive algorithm that is  $\beta$ -competitive against the NA optimal policy for a given problem (single box,  $k$  boxes, or matroid), then the authors' bandit algorithm matches that  $\beta$  up to a small constant factor plus an  $o(T)$  regret term. By leveraging previously established constants for these problems, they obtain approximate no-regret bounds of:

$$\begin{cases} 3.16\text{-approx. no regret for choosing 1 box,} \\ 12.64\text{-approx. no regret for choosing } k \text{ boxes,} \\ O(\log k)\text{-approx. no regret for choosing a matroid basis.} \end{cases}$$

Hence, even in the more restrictive NA setting and under partial feedback, the approach still achieves near-optimal competitive guarantees.

## Conclusions

Altogether, this methodology—scenario-aware convex relaxations, online convex optimization to handle partial or full feedback, and scenario-independent rounding—provides unified *constant-competitive no-regret* algorithms, even under adversarial instantiations of Pandora's Box and MSSC. While the core scenario-aware approach is similar across PA and NA benchmarks, the *constant factors* differ to reflect the more restrictive nature easier to compete against of non-adaptive policies. Ultimately, their results show that even in adversarial (bandit) contexts, one can achieve performance close to known offline approximations, including the optimal 4-approximation for MSSC, thus reinforcing the power of integrating convex relaxations, online learning, and robust rounding techniques.

## 4 Comparison with Previous Work

This work builds upon the foundational results in stochastic optimization, particularly in the domains of Pandora's Box and MSSC. Previous studies have focused on the stochastic case, where values and costs are drawn from known distributions, enabling efficient offline algorithms and partially adaptive strategies [2, 8].

In contrast, our study addresses the online setting where adversarially chosen values and costs are presented over multiple rounds. Notably, our framework simplifies the implementation of competitive algorithms through convex relaxations and online convex optimization techniques, extending prior work by providing guarantees in bandit settings [1]. Furthermore, our results generalize MSSC algorithms, achieving constant-factor approximations even under matroid constraints, a challenge that previous work struggled to address effectively [3].

## 5 Applications and Future Work

The results of this study have practical implications in various domains, including:

- **Resource Allocation:** Optimizing inspection costs in applications such as manufacturing and quality assurance.
- **Search Algorithms:** Enhancing strategies for finding optimal solutions under uncertain information.
- **Online Auctions:** Adapting bidding strategies in adversarial settings.

Future work could explore extending the framework to dynamic settings where constraints evolve over time or incorporating richer feedback mechanisms in the bandit setting. Additionally, improving the approximation factors under more complex combinatorial constraints remains an open challenge with significant theoretical and practical relevance.

## 6 Conclusion

This work presented a comprehensive framework for addressing the online versions of Pandora’s Box and Min Sum Set Cover (MSSC), two cornerstone problems in stochastic optimization. By leveraging convex relaxations and online convex optimization techniques, the authors developed efficient algorithms that achieve constant competitiveness under full information and bandit settings. These algorithms extend the capabilities of previous approaches, particularly in handling adversarially chosen inputs and accommodating complex constraints such as matroids [1].

## A Appendix

Include any additional proofs, explanations, or extended results here.

## References

- [1] Evangelia Gergatsouli and Christos Tzamos. Online learning for min sum set cover and pandora’s box. In *International Conference on Machine Learning*, pages 7382–7403. PMLR, 2022.
- [2] Martin Weitzman. *Optimal search for the best alternative*, volume 78. Department of Energy, 1978.
- [3] Uriel Feige, László Lovász, and Prasad Tetali. Approximating min sum set cover. *Algorithmica*, 40:219–234, 2004.
- [4] Shai Shalev-Shwartz et al. Online learning and online convex optimization. *Foundations and Trends® in Machine Learning*, 4(2):107–194, 2012.
- [5] Nikhil Bansal, Anupam Gupta, and Ravishankar Krishnaswamy. A constant factor approximation algorithm for generalized min-sum set cover. In *Proceedings of the twenty-first annual ACM-SIAM symposium on Discrete Algorithms*, pages 1539–1545. SIAM, 2010.
- [6] Abraham D Flaxman, Adam Tauman Kalai, and H Brendan McMahan. Online convex optimization in the bandit setting: gradient descent without a gradient. *arXiv preprint cs/0408007*, 2004.
- [7] Shai Shalev-Shwartz and Yoram Singer. A primal-dual perspective of online learning algorithms. *Machine Learning*, 69:115–142, 2007.
- [8] Shuchi Chawla, Evangelia Gergatsouli, Yifeng Teng, Christos Tzamos, and Ruimin Zhang. Pandora’s box with correlations: Learning and approximation. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1214–1225. IEEE, 2020.