Learning and Robustifying an Optimal Assortment from Observational Data

Miao Lu

Stanford University

INFORMS 2025

Joint work with

Yuxuan Han
NYU

Han Zhong
PKU

Zhengyuan Zhou
NYU

Jose Blanchet
Stanford

Han, Y., Zhong, H., Lu, M., Blanchet, J., & Zhou, Z. (2025). Learning an Optimal Assortment Policy under Observational Data. arXiv preprint arXiv:2502.06777.

Lu, M., Han, Y., Zhong, H., Zhou, Z., & Blanchet, J. (2025). Robustifying Assortment Optimization from Observational Data with Near-Optimal Sample Complexity. To appear on arXiv.

Work supported by grants NSF 2118199, 2229012, 2312204 and FA9550-20-1-0397.

Assortment Optimization: Math. Setups

Seller (assortment and revenue)

Seller wants to maximize expected revenue.

- Product set: $[N] := \{1, 2, \cdots, N\}$.
- An assortment: $S \subseteq [N]$.
- Revenue of product i: $r_i \in [0, r_{\max}]$.
- Non-purchasing: $i = 0$ and $r_0 = 0$.

Customer (choice model)

Customer buy product based on choice model:

- $\mathbb{P}(i|S) = \mathbb{P}(\text{buying } i \text{ seeing assortment } S)$.
- This work: multinomial logit (MNL),

$$\mathbb{P}_{\boldsymbol{v}}(i|S) := \frac{v_i}{1 + \sum_{j \in S} v_j}, \quad \forall i \in S \cup \{0\}.$$

Expected Revenue of assortment S:

$$R(S; \boldsymbol{v}, \boldsymbol{r}) := \sum_{i \in S} r_i \cdot \mathbb{P}_{\boldsymbol{v}}(i|S) = \sum_{i \in S} \frac{r_i v_i}{1 + \sum_{j \in S} v_j}.$$

Optimal Assortment $S^\star$ that maximizes the expected revenue:

$$S^\star := \arg\max_{|S| \le K} R(S; \boldsymbol{v}, \boldsymbol{r}), \quad K \in \mathbb{N}: \text{ size constraint.}$$

Learning Optimal Assortment $S^\star$ from Data

This Work: Offline Learning of $S^\star$.

- The seller doesn't know the underlying choice pattern $\mathbb{P}_{\boldsymbol{v}}$, but still wants to present good S.
- Learning $S^\star$ purely from offline data of customer choices, denoted by $\mathbb{D} := \{(S_k, i_k)\}_{k=1}^{n}$.
- $S_k \subseteq [N]$ is an observed assortment, $i_k \sim \mathbb{P}_{\boldsymbol{v}}(\cdot|S_k)$ is the observed customer choice.
- Performance metric (sub-optimality gap): $\mathrm{SubOpt}(\widehat{S}; \boldsymbol{v}, \boldsymbol{r}) := R(S^\star; \boldsymbol{v}, \boldsymbol{r}) - R(\widehat{S}; \boldsymbol{v}, \boldsymbol{r})$.

Key Questions and Result Overview:

1. (Lower bound) What is the minimal condition on $\mathbb{D}$ for learning $S^\star$ under the MNL model?
   A new notion of "item-wise partial coverage" condition.
2. (Upper bound) What is the algorithm to learn $S^\star$ requiring only the minimal condition on $\mathbb{D}$?
   Pessimistic Rank-Breaking (PRB), achieving minimax optimal sample complexity.
3. (Robustness) What if the observational data mismatches the actual customer choice pattern?
   Extensions of 1. & 2. to robust assortment optimization.

Learning Optimal Assortment $S^\star$ from Data

This Work: Offline Learning of $S^\star$.

- The seller doesn't know the underlying choice pattern $\mathbb{P}_{\boldsymbol{v}}$, but still wants to present good S.
- Learning $S^\star$ purely from offline data of customer choices, denoted by $\mathbb{D} := \{(S_k, i_k)\}_{k=1}^{n}$.
- $S_k \subseteq [N]$ is an observed assortment, $i_k \sim \mathbb{P}_{\boldsymbol{v}}(\cdot|S_k)$ is the observed customer choice.
- Performance metric (sub-optimality gap): $\mathrm{SubOpt}(\widehat{S}; \boldsymbol{v}, \boldsymbol{r}) := R(S^\star; \boldsymbol{v}, \boldsymbol{r}) - R(\widehat{S}; \boldsymbol{v}, \boldsymbol{r})$.

Key Questions and Result Overview:

1. (Lower bound) What is the minimal condition on $\mathbb{D}$ for learning $S^\star$ under the MNL model?
   A new notion of "item-wise partial coverage" condition.
2. (Upper bound) What is the algorithm to learn $S^\star$ requiring only the minimal condition on $\mathbb{D}$?
   Pessimistic Rank-Breaking (PRB), achieving minimax optimal sample complexity.
3. (Robustness) What if the observational data mismatches the actual customer choice pattern?
   Extensions of 1. & 2. to robust assortment optimization.

What is the Minimal Condition on Data $\mathbb{D}$?

- Arbitrarily bad offline data can not guarantee the seller to learn a good assortment!
- "Partial coverage" is the de facto condition for effective offline data-driven decision making: the data needs and only needs to cover the optimal action.
  - In contrast with "full coverage" style conditions, which require uniform coverage of action space.
- In assortment optimization, the optimal action can be viewed as $S^\star$.
- Prior work Dong et al. [2023] proposes PASTA, which achieves efficient learning if $S^\star$ appears sufficiently often in $\mathbb{D}$: $\mathrm{SubOpt}(\widehat{S}) \lesssim 1/\sqrt{n_{S^\star}}$, where $n_{S^\star} := \sum_{k=1}^{|\mathbb{D}|} \mathbf{1}\{S_k = S^\star\}$.
- However, the vanilla "partial coverage" condition could still be impractical here!
- Combinatorial nature of assortment S: if we randomly sample items to form an assortment of size K, the prob. of observing $S^\star$ once would be of order $\Theta(N^{-K})$.
- Can we further relax the vanilla partial coverage condition?
  - This work: "item-wise partial coverage condition" is both necessary and sufficient!

What is the Minimal Condition on Data $\mathbb{D}$?

- Arbitrarily bad offline data can not guarantee the seller to learn a good assortment!
- "Partial coverage" is the de facto condition for effective offline data-driven decision making: the data needs and only needs to cover the optimal action.
  - In contrast with "full coverage" style conditions, which require uniform coverage of action space.
- In assortment optimization, the optimal action can be viewed as $S^\star$.
- Prior work Dong et al. [2023] proposes PASTA, which achieves efficient learning if $S^\star$ appears sufficiently often in $\mathbb{D}$: $\mathrm{SubOpt}(\widehat{S}) \lesssim 1/\sqrt{n_{S^\star}}$, where $n_{S^\star} := \sum_{k=1}^{|\mathbb{D}|} \mathbf{1}\{S_k = S^\star\}$.
- However, the vanilla "partial coverage" condition could still be impractical here!
- Combinatorial nature of assortment S: if we randomly sample items to form an assortment of size K, the prob. of observing $S^\star$ once would be of order $\Theta(N^{-K})$.
- Can we further relax the vanilla partial coverage condition?
  - This work: "item-wise partial coverage condition" is both necessary and sufficient!

What is the Minimal Condition on Data $\mathbb{D}$?

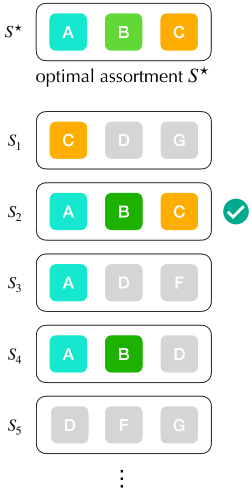
- Arbitrarily bad offline data can not guarantee the seller to learn a good assortment!
- "Partial coverage" is the de facto condition for effective offline data-driven decision making: the data needs and only needs to cover the optimal action.
  - In contrast with "full coverage" style conditions, which require uniform coverage of action space.
- In assortment optimization, the optimal action can be viewed as $S^\star$.
- Prior work Dong et al. [2023] proposes PASTA, which achieves efficient learning if $S^\star$ appears sufficiently often in $\mathbb{D}$: $\mathrm{SubOpt}(\widehat{S}) \lesssim 1/\sqrt{n_{S^\star}}$, where $n_{S^\star} := \sum_{k=1}^{|\mathbb{D}|} \mathbf{1}\{S_k = S^\star\}$.
- However, the vanilla "partial coverage" condition could still be impractical here!
- Combinatorial nature of assortment S: if we randomly sample items to form an assortment of size K, the prob. of observing $S^\star$ once would be of order $\Theta(N^{-K})$.
- Can we further relax the vanilla partial coverage condition?
  - This work: "item-wise partial coverage condition" is both necessary and sufficient!

What is the Minimal Condition on Data $\mathbb{D}$?

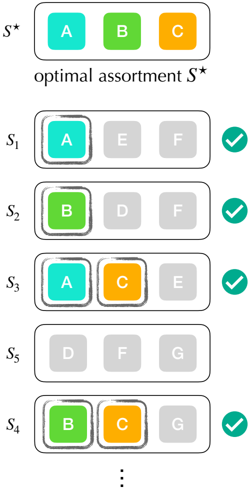
- Arbitrarily bad offline data can not guarantee the seller to learn a good assortment!
- "Partial coverage" is the de facto condition for effective offline data-driven decision making: the data needs and only needs to cover the optimal action.
  - In contrast with "full coverage" style conditions, which require uniform coverage of action space.
- In assortment optimization, the optimal action can be viewed as $S^\star$.
- Prior work Dong et al. [2023] proposes PASTA, which achieves efficient learning if $S^\star$ appears sufficiently often in $\mathbb{D}$: $\mathrm{SubOpt}(\widehat{S}) \lesssim 1/\sqrt{n_{S^\star}}$, where $n_{S^\star} := \sum_{k=1}^{|\mathbb{D}|} \mathbf{1}\{S_k = S^\star\}$.
- However, the vanilla "partial coverage" condition could still be impractical here!
- Combinatorial nature of assortment S: if we randomly sample items to form an assortment of size K, the prob. of observing $S^\star$ once would be of order $\Theta(N^{-K})$.
- Can we further relax the vanilla partial coverage condition?
  - This work: "item-wise partial coverage condition" is both necessary and sufficient!

What is the Minimal Condition on Data $\mathbb{D}$?

- Arbitrarily bad offline data can not guarantee the seller to learn a good assortment!
- "Partial coverage" is the de facto condition for effective offline data-driven decision making: the data needs and only needs to cover the optimal action.
  - In contrast with "full coverage" style conditions, which require uniform coverage of action space.
- In assortment optimization, the optimal action can be viewed as $S^\star$.
- Prior work Dong et al. [2023] proposes PASTA, which achieves efficient learning if $S^\star$ appears sufficiently often in $\mathbb{D}$: $\mathrm{SubOpt}(\widehat{S}) \lesssim 1/\sqrt{n_{S^\star}}$, where $n_{S^\star} := \sum_{k=1}^{|\mathbb{D}|} \mathbf{1}\{S_k = S^\star\}$.
- However, the vanilla "partial coverage" condition could still be impractical here!
- Combinatorial nature of assortment S: if we randomly sample items to form an assortment of size K, the prob. of observing $S^\star$ once would be of order $\Theta(N^{-K})$.
- Can we further relax the vanilla partial coverage condition?
  - This work: "item-wise partial coverage condition" is both necessary and sufficient!



Data Pattern needed for previous SOTA
(Require sufficient $S^\star$ in data)



Our algorithm learns $S^\star$ from this data pattern
(No $S^\star$ appearing in the data)

Our Theoretical Results: An Overview

Item-wise Partial Coverage (Ours). Effective size of $\mathbb{D}$:

$$n_{\text{effect}} := \min_{i \in S^\star} n_i, \quad n_i := \sum_{k=1}^{|\mathbb{D}|} \mathbf{1}\{i \in S_k\}.$$

Vanilla Partial Coverage [Dong et al., 2023].

$$n_{S^\star} := \sum_{k=1}^{|\mathbb{D}|} \mathbf{1}\{S_k = S^\star\} \leq n_{\text{effect}}.$$

	Setup	Upper Bound	Lower Bound
[Dong et al., 2023]	general revenue	$\widetilde{\mathcal{O}}\left(\sqrt{\frac{NK}{n_{S^\star}}}\right)$	N/A
This Work	general revenue	$\widetilde{\mathcal{O}}\left(\frac{K}{\sqrt{n_{\text{effect}}}}\right)$	$\Omega\left(\frac{K}{\sqrt{n_{\text{effect}}}}\right)$
This Work	uniform revenue	$\widetilde{\mathcal{O}}\left(\frac{\sqrt{K}}{\sqrt{n_{\text{effect}}}}\right)$	$\Omega\left(\frac{\sqrt{K}}{\sqrt{n_{\text{effect}}}}\right)$

Our Theoretical Results: An Overview

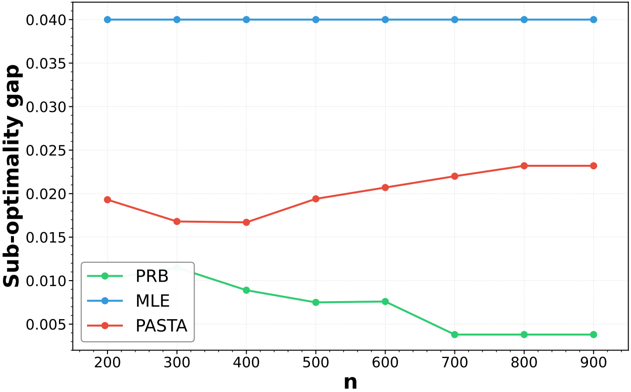
Item-wise Partial Coverage (Ours). Effective size of $\mathbb{D}$:

$$n_{\text{effect}} := \min_{i \in S^\star} n_i, \quad n_i := \sum_{k=1}^{|\mathbb{D}|} \mathbf{1}\{i \in S_k\}.$$

Vanilla Partial Coverage [Dong et al., 2023].

$$n_{S^\star} := \sum_{k=1}^{|\mathbb{D}|} \mathbf{1}\{S_k = S^\star\} \leq n_{\text{effect}}.$$

	Setup	Upper Bound	Lower Bound
[Dong et al., 2023]	general revenue	$\widetilde{\mathcal{O}}\left(\sqrt{\frac{NK}{n_{S^\star}}}\right)$	N/A
This Work	general revenue	$\widetilde{\mathcal{O}}\left(\frac{K}{\sqrt{n_{\text{effect}}}}\right)$	$\Omega\left(\frac{K}{\sqrt{n_{\text{effect}}}}\right)$
This Work	uniform revenue	$\widetilde{\mathcal{O}}\left(\frac{\sqrt{K}}{\sqrt{n_{\text{effect}}}}\right)$	$\Omega\left(\frac{\sqrt{K}}{\sqrt{n_{\text{effect}}}}\right)$



- PRB: Our algorithm
- PASTA [Dong et al., 2023], MLE (naive baseline).

Our Algorithm: Pessimistic Rank-Breaking (PRB)

Takeaways

- For offline A.O. with MNL model, "item-wise partial coverage" is the minimal data condition.
- PRB achieves minimax-optimal sample complexity, requiring only item-wise coverage.
- There is a $\sqrt{K}$-statistical gap in general revenue case and the uniform revenue case.

Robust Assortment Optimization

- So far we assumed that the choice pattern $\mathbb{P}$ generating $\mathbb{D}$ matches the future customer choice distribution.
- What if the learning from $\mathbb{P}$ generates a wrong choice model for the future?
- Our solution: robust assortment optimization from observational data.

Definition (Optimal Robust Assortment)

Given nominal choice model $\mathbb{P}$, robust level function ρ, we define optimal robust assortment as

$$S^\star_{\rho(\cdot;\cdot)} = \arg\max_{|S|\leq K} R_{\rho(\cdot;\cdot)}(S;\mathbb{P},\boldsymbol{r}), \tag{1}$$

$$R_{\rho(\cdot;\cdot)}(S;\mathbb{P},\boldsymbol{r}) := \inf_{\substack{\mathbb{Q}_{S_+}\in\mathcal{P}(S_+)\\ D_{\mathrm{KL}}(\mathbb{Q}_{S_+}\|\mathbb{P}(\cdot|S))\leq\rho(S;\mathbb{P}(\cdot|S))}} \sum_{j\in S_+} \mathbb{Q}_{S_+}(j)\cdot r_j. \tag{2}$$

Robust Assortment Optimization

- So far we assumed that the choice pattern $\mathbb{P}$ generating $\mathbb{D}$ matches the future customer choice distribution.
- What if the learning from $\mathbb{P}$ generates a wrong choice model for the future?
- Our solution: robust assortment optimization from observational data.

Definition (Optimal Robust Assortment)

Given nominal choice model $\mathbb{P}$, robust level function ρ, we define optimal robust assortment as

$$S^{\star}_{\rho(\cdot;\cdot)} = \arg\max_{|S| \le K} R_{\rho(\cdot;\cdot)}(S; \mathbb{P}, \boldsymbol{r}), \tag{1}$$

$$R_{\rho(\cdot;\cdot)}(S; \mathbb{P}, \boldsymbol{r}) := \inf_{\substack{\mathbb{Q}_{S_+} \in \mathcal{P}(S_+) \\ D_{\mathrm{KL}}(\mathbb{Q}_{S_+} \| \mathbb{P}(\cdot|S)) \le \rho(S; \mathbb{P}(\cdot|S))}} \sum_{j \in S_+} \mathbb{Q}_{S_+}(j) \cdot r_j. \tag{2}$$

Robust Item-wise Partial Coverage. Effective size of $\mathbb{D}$:

$$n_{\text{effect};\rho(\cdot;\cdot)} := \min_{i \in S^\star_{\rho(\cdot;\cdot)}} n_i, \quad n_i := \sum_{k=1}^{|\mathbb{D}|} \mathbf{1}\{i \in S_k\}.$$

	Setup	Upper Bound	Lower Bound
[Dong et al., 2023]	general revenue	$\widetilde{\mathcal{O}}\left(\sqrt{\frac{NK}{n_{S^\star}}}\right)$	N/A
Non-robust A.O.	general revenue	$\widetilde{\mathcal{O}}\left(\frac{K}{\sqrt{n_{\text{effect}}}}\right)$	$\Omega\left(\frac{K}{\sqrt{n_{\text{effect}}}}\right)$
Non-robust A.O.	uniform revenue	$\widetilde{\mathcal{O}}\left(\frac{\sqrt{K}}{\sqrt{n_{\text{effect}}}}\right)$	$\Omega\left(\frac{\sqrt{K}}{\sqrt{n_{\text{effect}}}}\right)$
Robust A.O.	general revenue	$\widetilde{\mathcal{O}}\left(\frac{K}{\sqrt{n_{\text{effect},\rho(\cdot;\cdot)}}}\right)$	$\Omega\left(\frac{K}{\sqrt{n_{\text{effect},\rho(\cdot;\cdot)}}}\right)$
Robust A.O.	uniform revenue	$\widetilde{\mathcal{O}}\left(\frac{\sqrt{K}}{\sqrt{n_{\text{effect},\rho(\cdot;\cdot)}}}\right)$	$\Omega\left(\frac{\sqrt{K}}{\sqrt{n_{\text{effect},\rho(\cdot;\cdot)}}}\right)$

Takeaways

- For offline robust A.O. with MNL as nominal choice model, "robust item-wise partial coverage" is the minimal data condition.
- New algorithms to achieve minimax-optimal sample complexity.
- There keeps to be a $\sqrt{K}$-statistical gap in general revenue case and the uniform revenue case.

Thank you for your attention!

References I

J. Dong, W. Mo, Z. Qi, C. Shi, E. X. Fang, and V. Tarokh. Pasta: pessimistic assortment optimization. In International Conference on Machine Learning, pages 8276–8295. PMLR, 2023.

Y. Han, H. Zhong, M. Lu, J. Blanchet, and Z. Zhou. Learning an optimal assortment policy under observational data. arXiv preprint arXiv:2502.06777, 2025.

A. Saha and P. Gaillard. Stop relying on no-choice and do not repeat the moves: Optimal, efficient and practical algorithms for assortment optimization. arXiv preprint arXiv:2402.18917, 2024.

A. Saha and A. Gopalan. Active ranking with subset-wise preferences. In The 22nd International Conference on Artificial Intelligence and Statistics, pages 3312–3321. PMLR, 2019.

Lower Bound

Theorem (Minimax Lower Bound 1 [Han et al., 2025])

For sufficiently large N, K with $N \geq 5K$, there is an $\mathbb{D}$ of K-sized assort. $\{S_k\}_{k=1}^n$ and a class of problem instances $\mathcal{V}$ s.t.,

$$\inf_{\pi(\cdot):\mathbb{D}\mapsto S} \sup_{(\boldsymbol{v},\boldsymbol{r})\in\mathcal{V}} \mathbb{E}_{\mathbb{D}\sim\otimes_{k=1}^n \mathbb{P}_{\boldsymbol{v}}(\cdot|S_k)} \left[\mathrm{SupOpt}(\pi(\mathbb{D}); \boldsymbol{v}, \boldsymbol{r})\right] \geq \Omega\left(\frac{r_{\max}\cdot K}{\sqrt{n_{\mathrm{effect}}}}\right). \tag{3}$$

Corollary (noframenumbering)

[Minimal Data Condition: Item-wise Partial Coverage] The minimal condition on the observational data $\mathbb{D}$ is to have a good coverage of each single product i in the optimal assortment $S^\star$.

Lower Bound

Theorem (Minimax Lower Bound 1 [Han et al., 2025])

For sufficiently large N, K with $N \geq 5K$, there is an $\mathbb{D}$ of K-sized assort. $\{S_k\}_{k=1}^n$ and a class of problem instances $\mathcal{V}$ s.t.,

$$\inf_{\pi(\cdot):\mathbb{D}\mapsto S} \sup_{(\boldsymbol{v},\boldsymbol{r})\in\mathcal{V}} \mathbb{E}_{\mathbb{D}\sim\otimes_{k=1}^n \mathbb{P}_{\boldsymbol{v}}(\cdot|S_k)} \left[\mathrm{SupOpt}(\pi(\mathbb{D});\boldsymbol{v},\boldsymbol{r})\right] \geq \Omega\left(\frac{r_{\max}\cdot K}{\sqrt{n_{\text{effect}}}}\right). \tag{3}$$

Corollary (noframenumbering)

[Minimal Data Condition: Item-wise Partial Coverage] The minimal condition on the observational data $\mathbb{D}$ is to have a good coverage of each single product i in the optimal assortment $S^\star$.

Our Algorithm: Pessimistic Rank-Breaking (PRB)

Two key steps: (i) rank-breaking based estimation; (ii) pessimistic assortment optimization.

Step 1: Rank-breaking Estimation [Saha and Gopalan, 2019, Saha and Gaillard, 2024].

$$\widehat{p}_j = \frac{\tau_j}{\tau_{j0}}, \quad \text{where} \quad \tau_{j0} = \underbrace{\sum_{k=1}^{|\mathbb{D}|} \mathbf{1}\{i_k \in \{0, j\}, j \in S_k\}}_{\#\text{ Choosing item } j \text{ or leave}}, \quad \tau_j = \underbrace{\sum_{k=1}^{|\mathbb{D}|} \mathbf{1}\{i_k = j\}}_{\#\text{ Choosing item } j}, \tag{4}$$

- Insight: pairwise choice probabilities $p_j := \mathbb{P}(j \mid S = \{0, j\}) = v_j/(1 + v_j)$ for different $j \in [N]$ can be estimated separately, which can then be mapped to v_j, i.e., $\widehat{v}_j = \widehat{p}_j/(1 - \widehat{p}_j)$.

Our Algorithm: Pessimistic Rank-Breaking (PRB)

Two key steps: (i) rank-breaking based estimation; (ii) pessimistic assortment optimization.

Step 1: Rank-breaking Estimation [Saha and Gopalan, 2019, Saha and Gaillard, 2024].

$$\widehat{p}_j = \frac{\tau_j}{\tau_{j0}}, \quad \text{where} \quad \tau_{j0} = \underbrace{\sum_{k=1}^{|\mathbb{D}|} \mathbf{1}\{i_k \in \{0, j\}, j \in S_k\}}_{\#\text{ Choosing item } j \text{ or leave}}, \quad \tau_j = \underbrace{\sum_{k=1}^{|\mathbb{D}|} \mathbf{1}\{i_k = j\}}_{\#\text{ Choosing item } j}, \qquad (4)$$

- Insight: pairwise choice probabilities $p_j := \mathbb{P}(j \mid S = \{0, j\}) = v_j/(1 + v_j)$ for different $j \in [N]$ can be estimated separately, which can then be mapped to v_j, i.e., $\widehat{v}_j = \widehat{p}_j/(1 - \widehat{p}_j)$.

Our Algorithm: Pessimistic Rank-Breaking (PRB)

Step 2: Pessimistic Assortment Optimization.

To hedge the potential large uncertainty for those items not well covered by $\mathbb{D}$, we optimize the assortment against a conservatively estimated revenue. Firstly,

$$p_j^{\mathsf{LCB}} = \left(\widehat{p}_j - \sqrt{\frac{2\widehat{p}_j(1-\widehat{p}_j)\log(1/\delta)}{\tau_{j0}}} - \frac{\log(1/\delta)}{\tau_{j0}}\right)_+ \leq p_j \tag{5}$$

Then we optimize the assortment to maximize the pessimistic estimation of the expected revenue:

$$\widehat{S} := \arg\max_{|S|\leq K} R(S; \boldsymbol{v}^{\mathsf{LCB}}, \boldsymbol{r}), \quad v_j^{\mathsf{LCB}} = p_j^{\mathsf{LCB}}/(1-p_j^{\mathsf{LCB}}) \leq v_j, \quad \forall j \in [N], \tag{6}$$

Lemma (Monotonicity under Optimal Assortment [Han et al., 2025])

Given $\{\underline{v}_j\}_{j=1}^N$ *s.t.* $v_j \geq \underline{v}_j, \forall j \in [N]$. *Then for the maximizer* $\underline{S} := \arg\max_{|S|\leq K} R(S; \underline{\boldsymbol{v}}, \boldsymbol{r})$,

$$R(S^\star; \boldsymbol{v}, \boldsymbol{r}) - R(\underline{S}; \boldsymbol{v}, \boldsymbol{r}) \leq \sum_{j\in S^\star} \frac{r_j(v_j - \underline{v}_j)}{1 + \sum_{j\in S^\star} v_j}. \tag{7}$$

Our Algorithm: Pessimistic Rank-Breaking (PRB)

Theorem (Sub-Optimality Gap of PRB [Han et al., 2025])

Assume that $n_i \geq 256(1 + KV)\log(N/\delta)$ *for all* $i \in S^\star$*. Then, with probability at least* $1 - 2\delta$*, the output* $\widehat{S}$ *of PRB satisfies that*

$$\mathrm{SubOpt}(\widehat{S}; \boldsymbol{v}, \boldsymbol{r}) \leq 2 \cdot r_{\max} \cdot (1 + v_{\max}) \cdot K \cdot \sqrt{\frac{\log(N/\delta)}{n_{\text{effect}}}} + \text{higher order terms.} \tag{8}$$