Moment Ambiguity and the Limits of Robust Stochastic Optimization

Andrés Cristi* Matteo Russo* Jiechen Zhang*

Abstract

We study fundamental information-theoretic limits of robust stochastic optimization when the distribution is known only through its exact moment sequence. We develop a unified framework that produces families of distinct distributions sharing all moments yet inducing radically different optimal decisions, thereby establishing strong impossibility results for a range of decision problems under moment ambiguity.

Our binary and N-way moment-matching constructions separate quantiles, threshold regions, and online selection opportunities while preserving every moment. For the newsvendor problem, moment equivalence causes quantile ambiguity, inducing any fixed or randomized order quantity to fail arbitrarily badly. For revenue maximization, no deterministic or randomized posted pricing scheme can secure a nontrivial approximation relative to the full-information benchmark.

For the secretary problem with cardinal observations, the optimal robust success probability under exact moment disclosure is the classical finite-horizon value s_m, even for i.i.d. observations. This value converges to $1/e$, compared with the full-information limit of approximately 0.5801 due to Gilbert and Mosteller (J. Am. Stat. Assoc., 1966).

We also recover and expand upon the recent impossibility result of Correa et al. (STOC, 2026) for prophet inequalities with moment knowledge. For every fixed $r \geq 1$, exact moment knowledge can yield at best a $\Theta(1/\log m)$ competitive ratio, even when competing against relaxed benchmarks based on expected r-th order statistics or when the algorithm is allowed to select r items.

1 Introduction

Many models in online decision making, resource allocation, inventory control, and mechanism design assume that the relevant distributions are known exactly. This assumption underlies classical prophet inequalities [KS78, SC84], newsvendor prescriptions [Edg88, AHM51], and Bayesian posted pricing [Mye81]. In applications, however, a decision maker may only see summaries of a distribution: means, variances, higher moments, or other aggregate descriptors. Moment ambiguity sets are therefore a central object in distributionally robust optimization and robust mechanism design [DY10, WKS14, RM19, Che22, CLK+18].

We ask how well a decision maker can perform when given every moment of the relevant distributions. The benchmark is the decision that could be made with full distributional knowledge. Even this complete sequence of exact summaries can leave substantial ambiguity: two compatible distributions may have very different quantiles or useful price thresholds, and their online selection problems may offer different opportunities to stop.

To obtain lower bounds, we construct distributions that share their entire moment sequence while controlling the atoms and tails that matter to each objective. Any policy receives the same ex ante

*EPFL, Switzerland, {andres.cristi,matteo.russo,jiechen.zhang}@epfl.ch.

moment information on these instances. The challenge is then to show that a common decision rule cannot exploit their different opportunities, even when it can randomize or observe values online.

1.1 Our Contributions

We construct families of moment-equivalent distributions with explicit bounds on their atoms and tails. A binary pair supplies the quantile and order-statistic separations; an N-way family supplies many alternative thresholds and designated values for pricing and secretary success probability.

Key Constructions. We use two separator constructions.

(A) *Binary separator* (**Theorem 3.1**). We construct two distributions that agree on all moments and have disjoint supports by splitting a moment-annihilating sequence into its positive and negative parts. The construction is guided by Vandermonde null vectors: in finite dimension, the signs alternate (Lemma 3.1); in the infinite construction, rapidly separated nodes make the limiting weights summable and keep all moments finite. Controlled atoms and tail decay then yield the quantile and order-statistic separations in Corollaries 4.1 and 7.1.

(B) *N-way residue-class separator* (**Theorem 3.2**). We construct N moment-equivalent distributions supported on disjoint residue classes. Roots of unity force all residue-class moment sums to agree, while the coefficient estimates in Theorem 3.2 provide the local mass control used for randomized pricing and yield the first-atom concentration in Corollary 6.1 used by the secretary embedding. This N-way structure creates many incompatible decision regions under one common moment disclosure.

Our N-way construction gives a one-sided realization with a common nonnegative coefficient sequence and explicit bounds on its full profile. The disjoint supports in both separator constructions imply total-variation distance 1 between distinct members.

Main Results. These constructions give the following lower bounds.

(1) *Quantiles, risk, and newsvendor* (**Corollary 4.1 and Theorem 4.1**). Exact moment information does not determine quantiles even approximately. This yields separations for Value-at-Risk, lower-tail average quantiles, chance-constrained capacity, and the newsvendor problem, including randomized order quantities; the risk and feasibility consequences are stated in Corollaries 4.2 and 4.3.

(2) *Posted pricing* (**Theorems 5.1 and 5.2**). The binary separator rules out any positive uniform approximation factor for deterministic prices, and the N-way separator does the same for randomized prices. For the latter, disjoint good-price intervals allow a many-scale averaging argument inside one full moment class; we discuss its relation to earlier pricing lower bounds in Section 5.

(3) *Secretary* (**Theorem 6.1**). In the labeled cardinal-value random-order model, the optimal robust success probability under exact moment disclosure is the classical value s_m for every horizon m, both for independent and for i.i.d. values. The classical secretary rule supplies the matching lower bound.

The success-probability upper bound combines first-atom concentration with the finite Ramsey argument in Lemma 6.1. Uniform mixtures of the moment-equivalent laws give i.i.d. instances, and a large homogeneous subgrid makes repeated designated values unlikely.

Order-statistic prophets and relation to prior constructions. The binary separator also gives a self-contained upper-bound proof for the fixed-r order-statistic prophet problem and its r-selection variant (Theorem 7.1 and corollary 7.2). Together with the lower bound of Correa et al. [CCL+26], this yields the logarithmic barrier. Their moment-equivalent hard instances use a related Vandermonde mechanism and also admit an order-statistic extension. Here the atom and tail estimates make the upper bound a direct application of the same pair used for quantiles and newsvendor.

1.2 Related Work

The moment problem and Stieltjes classes. The classical moment problem asks when a measure is determined by its moment sequence [Sti94, Hau21a, Hau21b, ST43, Sod19]. Chihara and Leipnik constructed discrete laws on bilateral geometric orbits that share the lognormal moment sequence [Chi70, Lei82, AGK19]. Suitable choices give disjoint atom sets and can also supply the designated-atom concentration and price separation used in our N-way applications. Our construction realizes these properties on a one-sided grid with a common nonnegative coefficient sequence and explicit bounds. Other discrete Stieltjes classes include the log-Heine constructions of Ostrovska and Turan [OT18].

Distributionally robust optimization. Distributionally robust optimization studies worst-case decisions over ambiguity sets specified by support, moments, or probability metrics [SAK57, DY10, GS10, WKS14, RM19]. We compare a moment-based policy with the distribution-specific benchmark. Our lower bounds show how much this comparison can cost even when the ambiguity set fixes the full moment sequence: quantile, newsvendor, and pricing guarantees can deteriorate arbitrarily, while the online selection objectives exhibit the barriers stated above.

Prophet inequalities and partial information. Classical prophet inequalities assume known distributions [KS78, SC84]; later work studies richer feasibility constraints and posted-price interpretations [HKS07, KW12, RS17, LS18, DFKL20, FSZ21]. Under partial information, sample access and inaccurate-prior models can still yield strong guarantees [CDFS22, DK19]. Correa et al. [CCL+26] prove the tight $\Theta(1/\log m)$ barrier for the expected-maximum benchmark under moment information. Although their theorem is stated for the maximum, their hard instances can also be adapted to every fixed order statistic because each active block contains many copies of its designated value. Section 7 gives an alternative, modular upper-bound derivation from our binary separator.

Newsvendor and pricing. The classical newsvendor solution is determined by a critical quantile [Edg88, AHM51], while posted-price revenue is determined by threshold tail probabilities [Mye81]. Prior work studies sample-based newsvendor policies [LPU15], moment-based robust newsvendor prescriptions [SAK57, DDN21], and robust pricing or mechanism design under coarse information [BBDS17, CLK+18, GPTD22]. The geometric interval lower bound of Babaioff et al. [BBDS17] and the continuum-of-scales, equal-revenue construction of Giannakopoulos, Poças, and Tsigonias-Dimitriadis [GPTD22] are especially close to the final averaging argument in our randomized-pricing lower bound. At a more abstract minimax level, Hartline and Roughgarden use an equal-revenue distribution over value scales to certify the optimal logarithmic price lottery [HR14], and Hartline and Johnsen later formulate the corresponding point-mass/equal-revenue decomposition through equivocal blends [HJ24]. Our results isolate a different information-theoretic limitation: distributions

can require incompatible quantiles and threshold regions even when they agree on their entire moment sequence.

Secretary and best-choice selection. For i.i.d. observations from a known continuous distribution, the full-information best-choice problem has asymptotic value of about 0.5801 [GM66]. Esfandiari et al. [EHLM20] study the random-order model with known, independent, non-identically distributed values, and Nuti [Nut22] shows that its worst-case value coincides with the i.i.d. full-information benchmark. Our success-probability result gives a contrasting information boundary: replacing the distributions by their exact moment sequences reduces the robust value to the classical secretary probability.

1.3 Organization

The remainder of the paper is organized as follows. Section 2 fixes notation, moment classes, and the moment-based policy model. Section 3 develops the two core moment-matching constructions; their quantitative consequences are stated alongside the applications that use them. We treat quantile ambiguity and newsvendor in Section 4, posted-price revenue in Section 5, secretary success probability in Section 6, and order-statistic prophet inequalities in Section 7. Section 8 supplies the deferred proofs of the two separator constructions, followed by concluding remarks in Section 9.

2 Preliminaries

We start by recalling some useful notation defined for a distribution ψ on $\mathbb{R}_+ = [0, \infty)$: namely, let $F_\psi(x) := \mathbb{P}_\psi(X \le x)$ denote its CDF and let $F_\psi(x^-) := \mathbb{P}_\psi(X < x)$ be its "open at x" counterpart. Define the quantile function

$$Q_\psi(u) := \inf\{x \ge 0 : \ F_\psi(x) \ge u\}, \qquad u \in (0, 1).$$

Next, for independent random variables $X_1, \dots, X_m$, we write $X_{(1)} \ge X_{(2)} \ge \cdots \ge X_{(m)}$ for the decreasing order statistics.

We say that a measure is supported on a set when it assigns that set full measure. For discrete laws, we specify the sets of atoms; these need not be closed.

2.1 Background on Moment Ambiguity

The k-th moment of a distribution ψ is

$$m_k(\psi) := \mathbb{E}_\psi[X^k],$$

whenever finite. Distributions agree on all moments if all these finite moments exist and agree for every $k \in \mathbb{N}_0$.

Definition 2.1. *For a moment sequence $m = (m_k)_{k \ge 0}$, let*

$$\mathcal{M}(m) := \{\psi : \ \mathbb{E}_\psi[X^k] = m_k \text{ for every } k \ge 0\}.$$

We call a subfamily $\mathcal{S} \subseteq \mathcal{M}(m)$ a Stieltjes-type class *if $\mathcal{S}$ contains more than one distribution. In this paper we focus on structured Stieltjes-type subclasses: subfamilies whose supports, atoms, and threshold regions satisfy additional quantitative separation properties.*

Definition 2.2. *A policy is* moment-based *if its rule is measurable with respect to the disclosed moment sequences and the observations revealed online. In particular, if two input distributions have identical moment sequences, the policy receives identical ex ante information under them.*

2.2 Background on Mechanism Design and Stochastic Optimization Problems

Value-at-Risk. Value-at-Risk is the quantile $\mathrm{VaR}_u(X) := Q_X(u)$. We use the lower-tail average quantile

$$\underline{\mathrm{CVaR}}_u(X) := \frac{1}{u}\int_0^u Q_X(t)\,dt, \qquad u \in (0,1).$$

The Newsvendor Problem. In the newsvendor problem, demand $D \sim \psi$ is nonnegative and the normalized profit from ordering $q \geq 0$ units is

$$\mathrm{Profit}_{\psi,\beta}(q) := \mathbb{E}_\psi[\min\{D, q\}] - (1-\beta)q, \qquad \beta \in [0,1].$$

A randomized ordering rule is a probability measure λ on $[0,\infty)$ with finite first moment. Its expected profit is

$$\mathrm{Profit}_{\psi,\beta}(\lambda) := \int_{[0,\infty)} \mathrm{Profit}_{\psi,\beta}(q)\,d\lambda(q).$$

Posted-price revenue. For a buyer value $X \sim \psi$, the revenue from price $v > 0$ is

$$\mathrm{Rev}_\psi(v) := v\mathbb{P}_\psi(X \geq v), \qquad \mathrm{Rev}^*_\psi := \sup_{v>0} \mathrm{Rev}_\psi(v).$$

For a randomized pricing rule λ on $(0,\infty)$, define

$$\mathrm{Rev}_\psi(\lambda) := \int_{(0,\infty)} \mathrm{Rev}_\psi(v)\,d\lambda(v).$$

Randomization is unnecessary under a known distribution, but it matters under ambiguity because the same moment information may be consistent with many distributions.

Secretary success. In the labeled independent random-order cardinal-value secretary problem, the m labels have independent nonnegative values $X_1, \ldots, X_m$ and arrive in a uniformly random order, independently of those values. At each arrival, the policy observes the item's label and realized value and must either reject it irrevocably or stop and select it; at most one item may be selected. Success means selecting any maximizer. Under exact moment disclosure, the policy is given, for each label i, the full sequence $(\mathbb{E}[X_i^k])_{k\geq 0}$ before the arrivals and no other ex ante distributional information. A moment-based secretary policy is a policy in the sense of Definition 2.2 for this disclosure and observation process. Its robust success guarantee is the infimum of its success probability over all independent per-label distributions with finite moments of every order, where the policy observes the resulting moment sequence of each label. Thus the worst case ranges over both the disclosed moment profiles and the compatible distributions. The optimal robust value is the supremum of this guarantee over moment-based secretary policies.

We also consider the restriction to i.i.d. values. In this model, the policy knows that the values are identically distributed and receives the full moment sequence of their common distribution; the robust guarantee ranges over all such disclosures and compatible common distributions.

For each integer horizon $m \geq 1$, the finite-horizon classical secretary value is

$$s_m := \max_{r\in\{1,\ldots,m\}} \frac{r-1}{m}\sum_{t=r}^{m}\frac{1}{t-1},$$

where the expression for $r = 1$ is interpreted as $1/m$. The classical asymptotic is $\lim_{m\to\infty} s_m = 1/e$.

Prophet inequalities. In the prophet inequality problem, values $X_1, \dots, X_m$ are revealed sequentially and the decision maker chooses a stopping time τ. We allow the rule to make no selection, in which case its reward is zero. The classical benchmark is $\mathbb{E}[\max_i X_i]$. In Section 7, we also compare against the weaker benchmark $\mathbb{E}[X_{(r)}]$, the expected r-th largest value, for fixed r.

3 A Residue-Class Moment-Matching Framework

This section develops two constructions for producing distinct distributions that agree on all moments. The binary construction splits an alternating signed identity into two probability measures; the N-way construction uses roots of unity to make the moment sums of N residue classes coincide. The binary pair is enough for two-instance separations, whereas the N-way family is needed when a randomized policy may spread its actions across many disjoint decision regions.

Finite-dimensional cancellation is only the algebraic starting point. On an infinite support, the weights must be summable and admit a common positive normalization, every distribution must have finite moments of every order, and the cancellation identities must remain valid after passing to the limit. Our applications require still more: quantitative control of designated atoms and of the remaining tail. We state these properties here as a toolkit and defer the analytic estimates to Section 8.

3.1 The Binary Construction

We first isolate the finite-dimensional mechanism. On k prescribed nodes, the binary construction is the one-dimensional nullspace of a rectangular Vandermonde matrix.

Lemma 3.1. *Let $k \geq 2$ and let $x_0 < x_1 < \cdots < x_{k-1}$ be distinct real numbers. There exist two distinct probability distributions μ_+ and μ_-, supported on disjoint subsets of $\{x_0, \dots, x_{k-1}\}$, such that*

$$\int x^s \, d\mu_+(x) = \int x^s \, d\mu_-(x), \qquad s = 0, 1, \dots, k-2.$$

Proof. Define the barycentric weights

$$a_i = \frac{1}{\prod_{\ell \neq i}(x_i - x_\ell)}, \qquad i = 0, \dots, k-1.$$

Then

$$\sum_{i=0}^{k-1} a_i x_i^s = 0, \qquad s = 0, 1, \dots, k-2.$$

Indeed, the vector $(a_i)_i$ lies in the nullspace of the $(k-1) \times k$ Vandermonde matrix with rows indexed by $s = 0, \dots, k-2$; since this matrix has rank $k-1$, the nullspace is one-dimensional. Equivalently, the identities follow from the leading-coefficient identity in Lagrange interpolation of every polynomial of degree at most $k-2$. Since the nodes are strictly increasing, the signs of the a_i alternate. Thus the positive and negative parts are both nonempty, and the equation for $s = 0$ gives equal total mass.

Let

$$P = \{i : a_i > 0\}, \qquad M = \{i : a_i < 0\}, \qquad W = \sum_{i \in P} a_i = -\sum_{i \in M} a_i,$$

and define

$$\mu_+(\{x_i\}) = \begin{cases} a_i/W, & i \in P, \\ 0, & i \notin P, \end{cases} \qquad \mu_-(\{x_i\}) = \begin{cases} -a_i/W, & i \in M, \\ 0, & i \notin M. \end{cases}$$

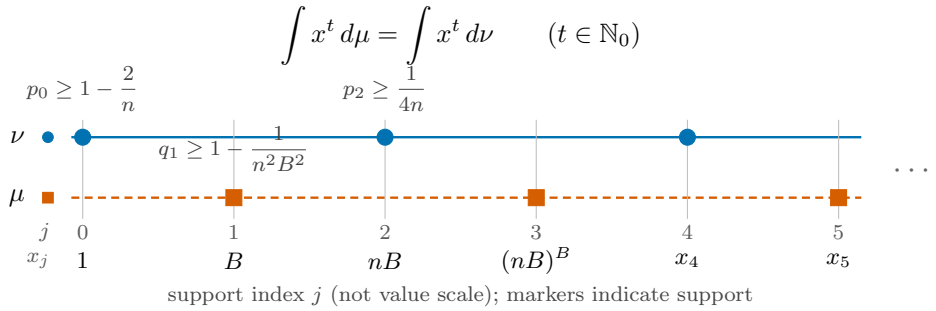


Figure 1: Binary separator: alternating supports and three controlled atoms.

These are probability distributions on disjoint supports. Dividing the identities for $\sum_i a_i x_i^s$ by W gives the claimed moment matching through degree $k-2$. □

Remark 3.1. *If two probability distributions supported on the same k fixed points agree on moments $s = 0, 1, \ldots, k-1$, then they are identical. Their difference vector lies in the kernel of the full $k \times k$ Vandermonde matrix, which is invertible. Thus matching through degree $k-2$ is the strongest possible nontrivial finite matching on k fixed support points.*

To pass from finitely many equations to all moments, we choose a rapidly separated support and take a controlled infinite-support limit of the Vandermonde construction. The decay estimate in the next theorem simultaneously guarantees normalization, finiteness of every moment, and the local mass bounds used later.

Theorem 3.1. *Fix integers $n, B \geq 3$ and define*

$$x_0 = 1, \qquad x_1 = B, \qquad x_j = (nB)^{B^{j-2}} \quad (j \geq 2).$$

There are weights $(w_j)_{j \geq 0}$ and a normalizing constant $W > 0$ such that

$$\nu(\{x_j\}) = \begin{cases} w_j/W, & j \text{ even}, \\ 0, & j \text{ odd}, \end{cases} \qquad \mu(\{x_j\}) = \begin{cases} |w_j|/W, & j \text{ odd}, \\ 0, & j \text{ even}, \end{cases}$$

are probability distributions on disjoint supports. Both distributions have finite moments of every order and agree on all moments. Writing $p_j := \nu(\{x_j\})$ and $q_j := \mu(\{x_j\})$, the following quantitative bounds hold:

$$p_0 \geq 1 - \frac{2}{n}, \qquad q_1 \geq 1 - \frac{1}{n^2B^2}, \qquad p_2 \geq \frac{1}{4n},$$

with $q_0 = p_1 = q_2 = 0$. Moreover, $W \geq n$ and

$$|w_j| \leq \frac{8(nB)^2}{x_j^2} \qquad \text{for all } j \geq 0.$$

The first three bounds expose the local structure: ν places almost all of its mass at 1 and a $1/\Theta(n)$ mass at nB, whereas μ places almost all of its mass at B. The remaining tail is small enough for the quantile and order-statistic applications. Figure 1 summarizes the alternating support pattern and the designated atoms used later.

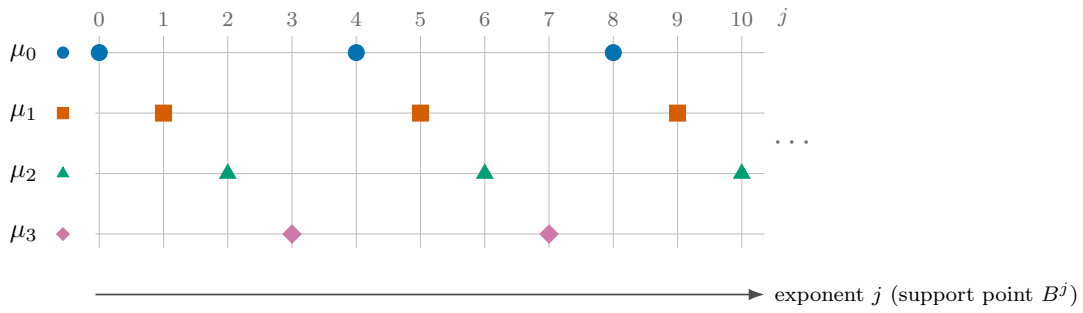


Figure 2: N-way separator for $N = 4$.

3.2 N-Way Residue-Class Construction

A binary pair suffices when a lower bound compares two regimes. To control randomized policies that may distribute their actions over many disjoint regions, we need arbitrarily many distributions with one common moment sequence. The next construction replaces the binary sign split by residue classes and uses roots of unity to force all residue-class moment sums to agree (see Figure 2 for an illustration).

Theorem 3.2. *Fix integers $B, N \geq 2$ and let $(w_j)_{j\in\mathbb{N}_0}$ be the unique coefficients that satisfy*

$$G_N(x) := \prod_{m=0}^{\infty} \sum_{u=0}^{N-1} (xB^{-m})^u = \sum_{j=0}^{\infty} w_j x^j.$$

There is a common normalizing constant $W \in [1, 4)$ such that, for every $h \in \{0, \dots, N-1\}$,

$$\mu_h(\{B^j\}) := \begin{cases} w_j/W, & j \equiv h \pmod N, \\ 0, & \text{otherwise}, \end{cases}$$

is a probability distribution. The distributions $\mu_0, \dots, \mu_{N-1}$ are supported on disjoint residue classes, have finite moments of every order, and agree on all moments.

Moreover, given $j \in \mathbb{N}_0$, let $L = \lfloor j/(N-1) \rfloor$, and $b = j - L(N-1)$, and define

$$e(j) := \frac{(N-1)L(L-1)}{2} + Lb,$$

the coefficients w_j satisfy

$$B^{-e(j)} \leq w_j \leq 2^j B^{-e(j)}.$$

In addition, $j - e(j) \leq N - 1$, with equality if and only if $j \in \{N-1, \dots, 2N-2\}$.

4 Quantile Ambiguity: VaR, Chance Constraints, and Newsvendor

The binary separator shows that exact moments do not determine quantiles, even up to an arbitrary multiplicative factor. We first state this separation and its consequences for risk and robust capacity, then use the same hidden critical quantile to prove the newsvendor lower bound.

Corollary 4.1. *For every $\beta \in [1/2, 1)$ and every integer $B \geq 3$, there exist two distributions μ, ν on $\mathbb{R}_+$ that agree on all moments and satisfy*

$$Q_\mu(\beta) = B, \qquad Q_\nu(\beta) = 1.$$

Proof. Choose an integer $n > \max\{4, 2/(1-\beta)\}$ and apply Theorem 3.1. Since $\beta \geq 1/2$, the smallest atoms of μ and ν have enough mass to contain level β: $q_1 > \beta$ and $p_0 > \beta$. Thus $Q_\mu(\beta) = B$ and $Q_\nu(\beta) = 1$. □

Corollary 4.2. *Fix integers $B \geq 3$ and $n \geq 4$, and let μ, ν be the binary separator distributions with parameters B, n. Then, for every $u \in [1/2, 1 - 2/n]$,*

$$\mathrm{VaR}_u(\mu) = B, \qquad \mathrm{VaR}_u(\nu) = 1.$$

Moreover, for every $u \in (0, 1 - 2/n]$,

$$\underline{\mathrm{CVaR}}_u(\mu) = B, \qquad \underline{\mathrm{CVaR}}_u(\nu) = 1.$$

Proof. Apply Theorem 3.1. The smallest support point of μ is B and has mass at least $1 - 1/(n^2B^2) > 1 - 2/n$; the smallest support point of ν is 1 and has mass at least $1 - 2/n$. Hence the displayed quantiles are constant over the stated ranges, and integrating the constant quantile functions gives the lower-tail average quantiles. □

Corollary 4.3. *Fix $\beta \in [1/2, 1)$ and $B \geq 3$. There are moment-equivalent distributions μ, ν such that any moment-based capacity rule that is feasible for every distribution in this moment class at service level β must choose capacity at least B. Consequently, on distribution ν its capacity is a factor at least B larger than the distribution-specific minimum feasible capacity.*

Proof. Use the pair from Corollary 4.1. Feasibility at service level β means choosing q with $F(q) \geq \beta$. Under μ, this requires $q \geq Q_\mu(\beta) = B$. Under ν, the distribution-specific minimum feasible capacity is $Q_\nu(\beta) = 1$. □

This corollary concerns rules that require feasibility uniformly over the moment class. It shows that robust feasibility alone can force an arbitrarily large over-capacity relative to the distribution-specific minimum feasible capacity.

4.1 Newsvendor Problem

Recall that in the newsvendor problem, we define the expected profit of ordering q units as

$$\mathrm{Profit}_{\psi,\beta}(q) := \mathbb{E}_\psi[\min\{D, q\}] - (1-\beta)q, \qquad \beta \in [0, 1],$$

where $D \sim \psi$ is the demand, and $(1-\beta)$ is the per-unit cost.

Lemma 4.1. *Fix $\beta \in [0, 1]$ and let ψ be a distribution on $\mathbb{R}_+$ with finite first moment. Then $q \mapsto \mathrm{Profit}_{\psi,\beta}(q)$ is concave on $[0, \infty)$. For $q > 0$, its one-sided derivatives are*

$$\partial_- \mathrm{Profit}_{\psi,\beta}(q) = \beta - F_\psi(q^-), \qquad \partial_+ \mathrm{Profit}_{\psi,\beta}(q) = \beta - F_\psi(q).$$

The right-derivative formula also holds at $q = 0$. Consequently,

$$\arg\max_{q \geq 0} \mathrm{Profit}_{\psi,\beta}(q) = \{q \geq 0 : \ F_\psi(q^-) \leq \beta \leq F_\psi(q)\}.$$

Proof. The tail-integral formula gives

$$\mathrm{Profit}_{\psi,\beta}(q) = \int_0^\infty \mathbb{P}_\psi(\min\{D,q\} \geq t)\, dt - (1-\beta)q = \int_0^q \mathbb{P}_\psi(D \geq t)\, dt - (1-\beta)q,$$

since $\mathbb{P}_\psi(\min\{D,q\} \geq t) = \mathbb{P}_\psi(D \geq t)\mathbf{1}_{\{q \geq t\}}$. The integrand in the display above is nonincreasing, so the profit is concave. The one-sided derivatives are the left and right limits of the integrand minus $1-\beta$, which gives the displayed formulas and the standard first-order optimality condition for a concave function. □

The following lemma states that the highest profit is achieved by a deterministic order amount, and is an immediate consequence of Lemma 4.1 together with Jensen's inequality.

Lemma 4.2. *Fix $\beta \in [0,1]$, let ψ be a distribution on $\mathbb{R}_+$ with finite first moment, and let λ be a randomized ordering rule. Writing $\bar{q} := \int_{[0,\infty)} q\, d\lambda(q)$, we have*

$$\mathrm{Profit}_{\psi,\beta}(\lambda) \leq \mathrm{Profit}_{\psi,\beta}(\bar{q}).$$

Proof. By definition, λ has finite first moment, so $\bar{q} < \infty$. Moreover,

$$\left|\mathrm{Profit}_{\psi,\beta}(q)\right| \leq \mathbb{E}_\psi[D] + q,$$

and hence $q \mapsto \mathrm{Profit}_{\psi,\beta}(q)$ is integrable under λ. The function is concave by Lemma 4.1, so Jensen's inequality gives

$$\mathrm{Profit}_{\psi,\beta}(\lambda) = \int_{[0,\infty)} \mathrm{Profit}_{\psi,\beta}(q)\, d\lambda(q) \leq \mathrm{Profit}_{\psi,\beta}\left(\int_{[0,\infty)} q\, d\lambda(q)\right) = \mathrm{Profit}_{\psi,\beta}(\bar{q}).$$ □

Finally, for the two moment-agreeing distributions μ, ν constructed in Theorem 3.1, the lemmas shown earlier yield that both optimal order amounts and optimal profits differ by an arbitrary (multiplicative) factor.

Lemma 4.3. *Fix $\beta \in [1/2, 1)$ and $\varepsilon \in (0,1)$. Let B, n be integers satisfying*

$$B \geq \frac{4}{\varepsilon(1-\beta)}, \qquad n > \max\left\{2B, \frac{2}{1-\beta}\right\},$$

and let μ, ν be the binary separator distributions with parameters B, n. Then the unique optimal order quantities and profits are

$$q^*_{\mu,\beta} = B, \qquad \mathrm{Profit}^*_{\mu,\beta} = \beta B, \qquad q^*_{\nu,\beta} = 1, \qquad \mathrm{Profit}^*_{\nu,\beta} = \beta.$$

Proof. Since $n > 2/(1-\beta)$, we have $\beta < 1 - 2/n$, so Corollary 4.2 gives $Q_\mu(\beta) = B$ and $Q_\nu(\beta) = 1$. These quantities are optimal by Lemma 4.1.

For uniqueness, $\mathrm{Profit}_{\mu,\beta}(q) = \beta q$ on $[0, B]$ and

$$\partial_+ \mathrm{Profit}_{\mu,\beta}(B) = \beta - F_\mu(B) = \beta - q_1 < 0,$$

using $q_1 > \beta$. Concavity then makes the profit strictly decreasing after B. The proof for ν is identical, using $p_0 > \beta$ at the atom 1. The profit values follow because the supports are contained in $[B, \infty)$ and $[1, \infty)$, respectively. □

For a randomized ordering rule, Lemma 4.2 upper-bounds profit under either demand law by the profit of the deterministic order quantity at its mean $\bar{q}$. Figure 3 visualizes the resulting split at $\bar{q} = \varepsilon B$.

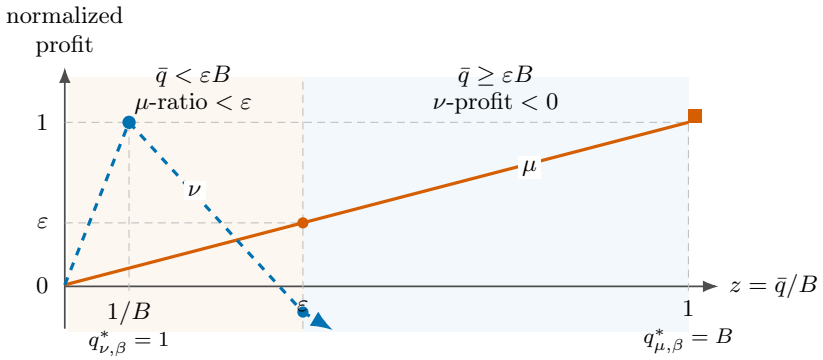


Figure 3: Newsvendor: the two moment-equivalent demand laws require incompatible mean order quantities.

Theorem 4.1. *Fix $\beta \in [1/2, 1)$ and $\varepsilon \in (0, 1)$. Let B, n be integers satisfying*

$$B \geq \frac{4}{\varepsilon(1-\beta)}, \qquad n > \max\left\{2B, \frac{2}{1-\beta}\right\},$$

and let μ, ν be the binary separator distributions with parameters B, n. Then every randomized ordering rule λ satisfies

$$\min\left\{\frac{\mathrm{Profit}_{\mu,\beta}(\lambda)}{\mathrm{Profit}^*_{\mu,\beta}}, \frac{\mathrm{Profit}_{\nu,\beta}(\lambda)}{\mathrm{Profit}^*_{\nu,\beta}}\right\} < \varepsilon.$$

Proof. Let $\bar{q} = \int_{[0,\infty)} q \, d\lambda(q)$, which is finite by the definition of a randomized ordering rule. Lemma 4.2 gives

$$\mathrm{Profit}_{\mu,\beta}(\lambda) \leq \mathrm{Profit}_{\mu,\beta}(\bar{q}), \qquad \mathrm{Profit}_{\nu,\beta}(\lambda) \leq \mathrm{Profit}_{\nu,\beta}(\bar{q}).$$

If $\bar{q} < \varepsilon B$, then $\bar{q} < B$ and

$$\frac{\mathrm{Profit}_{\mu,\beta}(\lambda)}{\mathrm{Profit}^*_{\mu,\beta}} \leq \frac{\beta\bar{q}}{\beta B} < \varepsilon.$$

If $\bar{q} \geq \varepsilon B$, then $\varepsilon B > 1$ and the ν-profit is decreasing on $[1, \infty)$. Since $\varepsilon B < B < nB = x_2$,

$$\mathrm{Profit}_{\nu,\beta}(\lambda) \leq p_0 + (1-p_0)\varepsilon B - (1-\beta)\varepsilon B \leq 1 + \frac{2\varepsilon B}{n} - (1-\beta)\varepsilon B < 0,$$

by $1 - p_0 \leq 2/n$, $n > 2B$, and $B \geq 4/(\varepsilon(1-\beta))$. Thus the ν ratio is below ε. □

The binary separator therefore yields lower bounds for decisions governed by one hidden critical quantile. Posted pricing requires a stronger form of ambiguity: the next section uses the N-way construction to hide many disjoint threshold regions simultaneously.

5 Threshold Ambiguity: Revenue Maximization and Randomized Pricing

In this section, we show that the constructions in Section 3 yield strong impossibility results for revenue maximization. To be more precise, posted-price revenue is a threshold objective since a price performs well only if it lies in the right tail region of the value distribution. The constructions in Section 3 then create moment-equivalent distributions whose good threshold regions are far apart.

5.1 Deterministic Posted Prices

Theorem 5.1. *For every $\varepsilon \in (0,1)$, there exist two distributions μ, ν that agree on all moments such that every deterministic price $v > 0$ satisfies*

$$\min\left\{\frac{\mathrm{Rev}_\mu(v)}{\mathrm{Rev}^*_\mu}, \frac{\mathrm{Rev}_\nu(v)}{\mathrm{Rev}^*_\nu}\right\} < \varepsilon.$$

Proof. Choose integers B, n with $B \geq \lceil 4/\varepsilon \rceil + 1$ and $n \geq 2B$, and let μ, ν be the binary separator distributions. Then

$$\mathrm{Rev}^*_\mu \geq \mathrm{Rev}_\mu(B) = B, \qquad \mathrm{Rev}^*_\nu \geq \mathrm{Rev}_\nu(nB) = nB\, p_2 \geq B/4.$$

Fix a price $v > 0$. If $v \leq 1$, then $\mathrm{Rev}_\nu(v) \leq 1$, so $\mathrm{Rev}_\nu(v)/\mathrm{Rev}^*_\nu < \varepsilon$. If $1 < v \leq B$, the support of ν has no atom in $(1, nB)$ and

$$\mathrm{Rev}_\nu(v) = v(1 - p_0) \leq \frac{2B}{n},$$

so $\mathrm{Rev}_\nu(v)/\mathrm{Rev}^*_\nu \leq 8/n < \varepsilon$. If $v > B$, then the atom B under μ does not buy, and

$$\mathrm{Rev}_\mu(v) \leq \sum_{\substack{j \geq 3 \\ j \text{ odd}}} x_j q_j \leq \frac{1}{W} \sum_{j \geq 3} x_j |w_j| \leq \frac{16nB^2}{x_3}.$$

Since $x_3 = (nB)^B$, this is an ε fraction of $\mathrm{Rev}^*_\mu \geq B$ for the chosen parameters. □

Remark 5.1. *The binary construction only rules out a single deterministic price. A lottery that mixes the two distribution-specific good prices can obtain a constant factor for this pair because revenue is nonnegative. The N-way construction below therefore addresses randomized rules by creating many disjoint good price regions with identical moment disclosures.*

5.2 Randomized Posted Prices

For integers $N, B \geq 2$, let n_h be the unique integer in $\{N-1, \ldots, 2N-2\}$ with $n_h \equiv h \pmod N$, and define

$$I_h := (B^{n_h - 1}, B^{n_h}].$$

Also set

$$C_N := \sum_{j \notin \{N-1, \ldots, 2N-2\}} 2^{2j - e(j) - (N-2)},$$

where $e(j)$ is defined in Theorem 3.2. This constant is finite; we establish convergence in the proof of the following localization bound.

Lemma 5.1. *Fix integers $N, B \geq 2$, and let $\mu_0, \ldots, \mu_{N-1}$ be the distributions from Theorem 3.2. For every $h \in \{0, \ldots, N-1\}$:*

$$\mathrm{Rev}^*_{\mu_h} \geq \frac{B^{N-1}}{4},$$

and every $v \notin I_h$ satisfies

$$\mathrm{Rev}_{\mu_h}(v) \leq (2^{2N-2} + C_N) B^{N-2}.$$

Proof. First verify that C_N is finite. Write $j = L(N-1) + b$, with $L \in \mathbb{N}_0$ and $b \in \{0, \ldots, N-2\}$. The exponent in its defining series is $A_L + (2-L)b$, where $A_L := (N-1)(5L - L^2)/2 - (N-2)$. For $L \geq 3$, the L-th block is bounded by $(N-1)2^{A_L}$, and $A_{L+1} - A_L = (N-1)(2-L) \leq -1$. These bounds form a geometric tail:

$$\sum_{L=3}^{\infty} \sum_{b=0}^{N-2} 2^{A_L + (2-L)b} \leq (N-1) 2^{A_3} \sum_{k=0}^{\infty} 2^{-k} < \infty.$$

The finitely many terms with $L \leq 2$ cause no difficulty, so $C_N < \infty$.

For the lower bound, Theorem 3.2 gives $w_{n_h} \geq B^{-e(n_h)}$ and $n_h - e(n_h) = N - 1$, while $W < 4$. Hence

$$\mathrm{Rev}^*_{\mu_h} \geq B^{n_h} \mu_h(\{B^{n_h}\}) \geq \frac{B^{N-1}}{4}.$$

For the upper bound, first sum the contribution of all exponents outside the special block. By Theorem 3.2,

$$\sum_{j \notin \{N-1, \ldots, 2N-2\}} B^j \mu_h(\{B^j\}) \leq C_N B^{N-2}.$$

Now fix $v \notin I_h$. If $v \leq B^{N-2}$, then $\mathrm{Rev}_{\mu_h}(v) \leq B^{N-2}$. If $B^{N-2} < v \leq B^{n_h - 1}$, the only special-block atom that can contribute is B^{n_h}, and

$$v \mu_h(\{B^{n_h}\}) \leq B^{n_h - 1} w_{n_h} \leq 2^{2N-2} B^{N-2}.$$

If $v > B^{n_h}$, no special-block atom contributes. Combining these cases gives the claim. □

Figure 4 illustrates how the disjoint intervals I_h convert revenue localization into a lower bound against every randomized price.

We now prove this section's main result:

Theorem 5.2. *For every $\varepsilon > 0$, there exist an integer $N \geq 2$ and moment-equivalent distributions $\mu_0, \ldots, \mu_{N-1}$ such that, for every randomized pricing strategy λ, there exists an index $h \in \{0, \ldots, N-1\}$ satisfying*

$$\frac{\mathrm{Rev}_{\mu_h}(\lambda)}{\mathrm{Rev}^*_{\mu_h}} < \varepsilon.$$

Proof. Choose an integer $N \geq 2$ so that $1/N < \varepsilon/2$, and then choose an integer $B \geq 2$ so large that $4(2^{2N-2} + C_N)/B \leq \varepsilon/2$. Let $\mu_0, \ldots, \mu_{N-1}$ be the distributions from Theorem 3.2 with these parameters N and B. Consider now any randomized pricing strategy λ. Since the intervals $I_0, \ldots, I_{N-1}$ are pairwise disjoint, $\sum_{h=0}^{N-1} \lambda(I_h) \leq 1$, and so there exists some $h_* \in \{0, \ldots, N-1\}$ such that $\lambda(I_{h_*}) \leq 1/N$. We next split the expected revenue of λ against μ_{h_*} into prices inside and outside I_{h_*}:

$$\begin{aligned}
\mathrm{Rev}_{\mu_{h_*}}(\lambda) &= \int_{I_{h_*}} \mathrm{Rev}_{\mu_{h_*}}(v) d\lambda(v) + \int_{(0,\infty) \setminus I_{h_*}} \mathrm{Rev}_{\mu_{h_*}}(v) d\lambda(v) \\
&\leq \lambda(I_{h_*}) \mathrm{Rev}^*_{\mu_{h_*}} + (2^{2N-2} + C_N) B^{N-2},
\end{aligned}$$

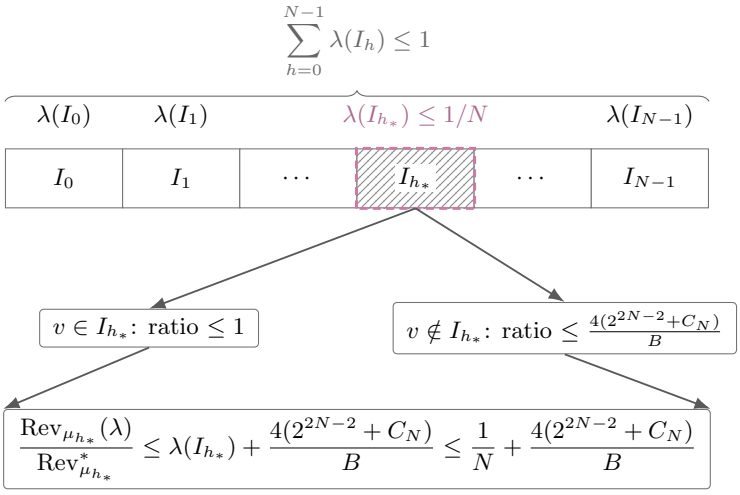


Figure 4: Randomized pricing: one disjoint region receives lottery mass at most $1/N$.

where the last inequality holds by using the trivial bound $\mathrm{Rev}_{\mu_{h_*}}(v) \le \mathrm{Rev}^*_{\mu_{h_*}}$ on I_{h_*} and by Lemma 5.1 outside I_{h_*}. We conclude by using again Lemma 5.1 to lower bound $\mathrm{Rev}^*_{\mu_{h_*}}$,

$$\frac{\mathrm{Rev}_{\mu_{h_*}}(\lambda)}{\mathrm{Rev}^*_{\mu_{h_*}}} \le \lambda(I_{h_*}) + \frac{(2^{2N-2}+C_N)B^{N-2}}{B^{N-1}/4} \le \frac{1}{N} + \frac{4(2^{2N-2}+C_N)}{B} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

This proves the claim. □

Relation to earlier many-scale pricing lower bounds. Once Lemma 5.1 has localized the revenue of each μ_h, the final averaging step is very close to Claim A.1 of Babaioff et al. [BBDS17]: prices are divided into geometrically separated intervals, one selects an interval receiving little probability under the price lottery, and prices outside that interval are charged to a multiplicative revenue loss. The difference lies in the hard instances. Babaioff et al. may use unrelated value distributions—in the relevant claim, deterministic values at the selected scales—whereas our distributions $\mu_0, \ldots, \mu_{N-1}$ are fixed members of one common full moment class. The role of Theorem 3.2 and Lemma 5.1 is precisely to preserve all moments while retaining the off-interval revenue bound needed by that argument. A continuous version of the same many-scale obstruction appears in [GPTD22], who mix two-point laws with a common mean but different monopoly prices so that the resulting mixture is a truncated equal-revenue distribution. In contrast, their higher moments vary across the hard family, whereas all members of our finite family agree on every moment.

The N-way construction yields the randomized-pricing lower bound by hiding the appropriate threshold region among many disjoint alternatives. In the next section, its first-atom concentration property lets us mix laws concentrated at different grid points to obtain hard i.i.d. instances for secretary success probability.

6 Selection Ambiguity: The Secretary Problem

We study the probability of selecting a maximum in the labeled cardinal-value model defined in Section 2. When the independent distributions are known in full, the worst-case finite-horizon success probability equals the full-information i.i.d. value [EHLM20, Nut22]; the latter converges to

roughly 0.5801 [GM66]. Replacing full distributional knowledge by exact moment sequences lowers the robust value to s_m, even when the values are i.i.d. This value converges to $1/e$ (Corollary 6.2).

The upper bound uses the following consequence of the N-way construction: each residue-class distribution can be concentrated near a different designated value while all distributions agree on all moments.

Corollary 6.1. *Fix an integer $N \geq 2$, and for each integer $B \geq 2$ let $\mu_0, \ldots, \mu_{N-1}$ be the distributions from Theorem 3.2. Then*

$$\min_{0 \leq h \leq N-1} \mu_h(\{B^h\}) \longrightarrow 1 \qquad \text{as } B \to \infty.$$

Proof. Fix N. For each $h \in \{0, \ldots, N-1\}$, choosing the monomial x^h from the $m = 0$ factor and the constant term from every other factor shows that $w_h \geq 1$. It therefore suffices to show that the common normalizer tends to one.

Use the zero residue class to write $W = \sum_{k \geq 0} w_{Nk}$. The exponent bounds in Theorem 3.2 give $e(Nk) \geq k$: this is immediate for $k = 0$, and for $k \geq 1$ it follows from $Nk - e(Nk) \leq N - 1$. Thus, for $B \geq 2^{N+1}$,

$$1 \leq W = \sum_{k \geq 0} w_{Nk} \leq \sum_{k \geq 0} \left(\frac{2^N}{B}\right)^k = \frac{1}{1 - 2^N/B} \longrightarrow 1.$$

Since μ_h is a probability measure, we have, uniformly in h,

$$\frac{1}{W} \leq \mu_h(\{B^h\}) = \frac{w_h}{W} \leq 1.$$

Taking the minimum over h proves the claim.

□

The next lemma finds a large subgrid on which a cardinal-value policy behaves almost like a rule based on relative ranks. Uniform i.i.d. sampling from that subgrid gives a hard instance; the additional error accounts for repeated values.

Lemma 6.1. *Fix integers $k \geq m \geq 1$ and $\varepsilon > 0$. There exists an integer $N_0 = N_0(m, k, \varepsilon) \geq k$ such that, for every integer $N \geq N_0$, every ordered grid*

$$v_0 < v_1 < \cdots < v_{N-1},$$

and every randomized cardinal-value secretary policy π with horizon m, there is a set $H \subseteq \{0, \ldots, N-1\}$ of size k such that, when the m labeled values are drawn independently and uniformly from $\{v_h : h \in H\}$ and arrive in uniformly random order,

$$\mathbb{P}_\pi(\textit{selects a maximum}) \leq s_m + \varepsilon + \frac{\binom{m}{2}}{k}.$$

Proof. Represent π by its conditional stopping probability after each history of observed labels and values. Choose a mesh size $\delta > 0$ with $m\delta \leq \varepsilon$, and round each stopping probability down to a multiple of δ. For each $t = 1, \ldots, m$, color every t-element subset of the grid by the finite vector of rounded stopping probabilities over all orderings of those values and all ordered t-tuples of distinct labels.

Successive applications of the finite Ramsey theorem give a k-element set H that is homogeneous for every history length, provided N is sufficiently large as a function of m, k, δ. On histories of

distinct grid values indexed by H, the rounded rule $\widetilde{\pi}$ depends only on time, the ordered label history, and the relative order of the observed values. Couple π and $\widetilde{\pi}$ using the same uniform random choices. For any fixed sequence of distinct values and labels, their decisions differ with probability at most $m\delta$.

Let $\mathcal{D}$ be the event that the m independent uniform draws are distinct. Conditional on any unordered set of m distinct draws, their assignment to labels is uniform and their arrival order is independent and uniform. The observed label order is therefore independent of the value ranks. On this event, $\widetilde{\pi}$ is a randomized classical secretary rule, with labels supplying only independent auxiliary randomness. Its success probability is at most s_m [GM66, Fer89], so π succeeds with conditional probability at most $s_m + m\delta$.

Each pair of draws agrees with probability $1/k$, and hence $\mathbb{P}(\mathcal{D}^c) \le \binom{m}{2}/k$. It follows that

$$\mathbb{P}_\pi(\text{success}) \le s_m + m\delta + \mathbb{P}(\mathcal{D}^c) \le s_m + \varepsilon + \frac{\binom{m}{2}}{k}. \qquad \square$$

Theorem 6.1. *For every integer horizon $m \ge 1$, the optimal robust success probability under exact moment disclosure in the labeled cardinal-value random-order secretary problem is s_m, both for independent values and for i.i.d. values.*

Proof. For the lower bound, use a classical optimal skip-then-record rule, breaking value ties by independent continuous tie-breakers. Conditional on any realized values, this rule selects the lexicographic maximum with probability s_m; that item is one of the value maximizers. Hence the robust success probability is at least s_m in both models.

For the upper bound, we construct an i.i.d. instance by mixing moment-equivalent laws concentrated at the grid points of Lemma 6.1. Fix $\varepsilon > 0$ and choose an integer $k \ge m$ such that $\binom{m}{2}/k \le \varepsilon/3$. Take an integer $N \ge \max\{2, N_0(m, k, \varepsilon/3)\}$. By Corollary 6.1, choose an integer $B \ge 2$ large enough that the N-way construction satisfies

$$\mu_h(\{B^h\}) \ge 1 - \frac{\varepsilon}{3m} \qquad \text{for every } h = 0, \dots, N-1.$$

All uniform mixtures of k members of $\{\mu_0, \dots, \mu_{N-1}\}$ share the same full moment sequence. This disclosure is fixed before the policy is chosen, since N and B depend only on m and ε. Fix a moment-based policy π under this disclosure, allowing it to use the fact that the values are i.i.d.

Apply Lemma 6.1 to π on the ordered grid $B^0 < B^1 < \cdots < B^{N-1}$. It gives a set H of size k such that π succeeds with probability at most $s_m + 2\varepsilon/3$ on independent uniform draws from $\{B^h : h \in H\}$. Define

$$\psi_H := \frac{1}{k} \sum_{h \in H} \mu_h.$$

This distribution is compatible with the disclosure and has finite moments of every order.

To compare the two instances, draw $h_1, \dots, h_m$ independently and uniformly from H. Set $Y_i = B^{h_i}$ and, conditional on these indices, draw $X_i \sim \mu_{h_i}$ independently. Then the X_i are i.i.d. with law ψ_H, and

$$\mathbb{P}(\text{some } X_i \ne Y_i) \le \sum_{i=1}^{m} \mathbb{P}(X_i \ne B^{h_i}) \le \frac{\varepsilon}{3}.$$

Run π on both vectors with the same disclosure, arrival order, and private randomization. Whenever the vectors agree, the histories and the success events agree. Therefore

$$\mathbb{P}_\pi(\text{success on } X) \le \mathbb{P}_\pi(\text{success on } Y) + \frac{\varepsilon}{3} \le s_m + \varepsilon.$$

Since π and ε were arbitrary, the robust value in the i.i.d. model is at most s_m. The same upper bound holds in the independent model, whose admissible instances include the i.i.d. ones. Together with the lower bound, this proves the theorem. □

Corollary 6.2. *In both the independent and the i.i.d. models, the finite-horizon robust success values satisfy*

$$\lim_{m\to\infty} s_m = \frac{1}{e}.$$

Proof. By Theorem 6.1, the robust value at every finite horizon $m \geq 1$ is s_m in both models. The classical limit recalled in Section 2.2 gives the displayed convergence. □

The next section treats prophet inequalities, where the binary construction separates the probabilities of seeing enough large observations to change an order-statistic benchmark.

7 Order-Statistic Ambiguity: Prophet Inequalities

The hard instances of Correa et al. [CCL+26] are stated for the expected-maximum benchmark, but their construction can also be adapted to every fixed order statistic. Each active block contains, with high probability, many copies of its designated value. The fixed-r logarithmic barrier can therefore also be obtained by adapting their construction.

The binary separator yields the required order-statistic gap directly from its designated-atom and tail estimates. The resulting proof is modular and uses the same construction that drives the quantile and newsvendor lower bounds.

For integers $m \geq r \geq 1$, let $\rho^{\mathrm{mom}}_{m,r}$ denote the supremum of all $\alpha \geq 0$ for which there is a randomized moment-based single-choice policy such that, for every collection $Z_1, \ldots, Z_m$ of independent nonnegative random variables with finite moments of every order, the policy's stopping time τ satisfies

$$\mathbb{E}[Z_\tau] \geq \alpha\, \mathbb{E}[Z_{(r)}]$$

under the disclosed moment sequences. This guarantee-based definition also covers instances with zero benchmark, without forming an indeterminate ratio. In this section, log denotes the natural logarithm.

The following consequence of the binary separator provides exactly the order-statistic gap needed for the upper bound.

Corollary 7.1. *For every fixed integer $r \geq 1$ there is a constant $\kappa_r > 0$ such that the following holds. Let B, n be integers with $B, n \geq 3$ and $n \geq \max\{2B, 2r\}$, and let μ, ν be the distributions from Theorem 3.1. If $X_1, \ldots, X_n$ are i.i.d. from ν and $Y_1, \ldots, Y_n$ are i.i.d. from μ, then*

$$\mathbb{E}_\nu[X_{(r)}] \geq \kappa_r n B, \qquad \mathbb{E}_\mu[Y_{(1)}] \leq 2B.$$

In addition,

$$\mathbb{E}_\nu\left[\sum_{i=1}^n X_i \mathbf{1}_{\{X_i \geq x_3\}}\right] \leq \frac{6}{n}, \qquad \mathbb{E}_\mu\left[\sum_{i=1}^n Y_i \mathbf{1}_{\{Y_i \geq x_3\}}\right] \leq \frac{6}{n}.$$

Proof. Let $H := \sum_{i=1}^n \mathbf{1}_{\{X_i = x_2\}}$, and let $H_0 \sim \mathrm{Bin}(n, 1/(4n))$. Since $p_2 \geq 1/(4n)$, the variable H stochastically dominates H_0, and hence

$$\mathbb{P}(H \geq r) \geq \mathbb{P}(H_0 = r) = \binom{n}{r} \frac{1}{(4n)^r} \left(1 - \frac{1}{4n}\right)^{n-r} \geq \frac{e^{-1/4}}{8^r r!} =: \kappa_r,$$

using $n \geq 2r$. Here $\binom{n}{r} \geq (n/2)^r/r!$, while $\log(1-x) \geq -x/(1-x)$ with $x = 1/(4n)$ gives $(1-1/(4n))^{n-r} \geq e^{-1/4}$. On the event $H \geq r$, $X_{(r)} \geq x_2 = nB$.

For the tails, Theorem 3.1 gives

$$n \sum_{j \geq 3} x_j \frac{|w_j|}{W} \leq 8(nB)^2 \sum_{j \geq 3} \frac{1}{x_j} \leq \frac{16(nB)^2}{x_3} \leq \frac{16}{nB} < \frac{6}{n},$$

because $x_3 = (nB)^B$ and $B \geq 3$. This proves both displayed tail bounds. Finally, every non-tail sample from μ equals B, so

$$\mathbb{E}_\mu[Y_{(1)}] \leq B + \mathbb{E}_\mu\left[\sum_{i=1}^{n} Y_i \mathbf{1}_{\{Y_i \geq x_3\}}\right] \leq 2B$$

for the stated parameter range.

□

Theorem 7.1. *For every integer $r \geq 1$, there exist constants $a_r, b_r > 0$ and an integer $m_0(r) \geq \max\{r, 2\}$ such that, for every integer $m \geq m_0(r)$,*

$$\frac{a_r}{\log m} \leq \rho_{m,r}^{\mathrm{mom}} \leq \frac{b_r}{\log m}.$$

For the lower-bound direction we use the following theorem of Correa et al. [CCL+26], stated directly in terms of the guarantee above.

Proposition 7.1. *There are a constant $a > 0$ and an integer $m_{\mathrm{lb}} \geq 2$ such that, for every integer $m \geq m_{\mathrm{lb}}$,*

$$\rho_{m,1}^{\mathrm{mom}} \geq \frac{a}{\log m}.$$

Proof of Theorem 7.1. The lower bound follows from Proposition 7.1: since $Z_{(1)} \geq Z_{(r)}$ pointwise, a guarantee against the maximum is also one against the r-th largest value. We may therefore take $a_r := a$.

For the upper bound, we arrange the binary separator pairs in blocks of doubling size. At base $B = 3$, a block of n draws from ν_n has an order-statistic benchmark proportional to n, whereas the values below the tail threshold x_3 in its moment-equivalent μ_n block are all 3. The hard instances switch from ν-blocks to μ-blocks at different points. A policy must allocate its stopping probability across the growing blocks without learning the switch point from the disclosed moments.

Let κ_r be the constant in Corollary 7.1 and take $N_r := \lceil \max\{6, 2r, 1/\kappa_r\} \rceil$, so that the separator estimates apply to every block size $n \geq N_r$. Write (μ_n, ν_n) for the corresponding pair with base 3. Fix $m \geq 4N_r^2$ and let L be the largest integer with $N_r(2^L - 1) \leq m$. We can then fit L blocks of sizes $n_j = N_r 2^{j-1}$ into the horizon. In instance $\mathcal{I}_t$, blocks $1, \ldots, t$ contain independent draws from ν_{n_j}, blocks $t+1, \ldots, L$ contain independent draws from μ_{n_j}, and any remaining coordinates are zero. All coordinates are independent. Every instance discloses the same moment sequence at each coordinate, as illustrated in Figure 5.

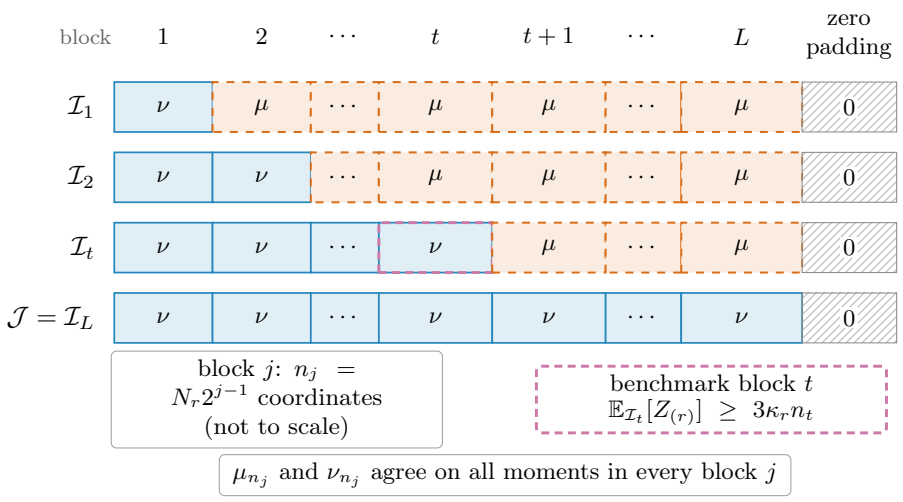


Figure 5: Prophet hard instances: the switch point varies while block sizes double.

Fix any randomized moment-based stopping rule τ. On the all-ν instance $\mathcal{J} := \mathcal{I}_L$, let $a_j := \mathbb{P}_{\mathcal{J}}(\tau \text{ stops in block } j)$. These probabilities satisfy $\sum_{j=1}^{L} a_j \leq 1$. Through any block $j \leq t$, the history has the same law under $\mathcal{I}_t$ as under $\mathcal{J}$, so the stopping probability in that block is still a_j. Ordinary rewards in a ν-block are at most $3n_j$; after the switch, they are at most 3. Adding the tail bounds from Corollary 7.1 gives

$$\mathbb{E}_{\mathcal{I}_t}[Z_\tau] \leq 3\sum_{j\leq t} n_j a_j + 3 + 6\sum_{j=1}^{L} \frac{1}{n_j} \leq 3\sum_{j\leq t} n_j a_j + 3 + \frac{12}{N_r}.$$

Here the last inequality uses the geometric sum $\sum_{j=1}^{L} 2^{1-j} < 2$.

To compare these rewards across scales, average the switch points with weights $\pi_t = 2^{-t}/Z_L$, where $Z_L := \sum_{t=1}^{L} 2^{-t} \in [1/2, 1)$. The decreasing weights offset the growing block sizes: $\sum_{t=j}^{L} \pi_t \leq 2^{2-j}$, while $n_j = N_r 2^{j-1}$. The total stopping probability is at most one, so the weighted policy reward stays bounded. In contrast, block t alone gives $\mathbb{E}_{\mathcal{I}_t}[Z_{(r)}] \geq 3\kappa_r n_t$, and each switch point contributes the same amount to the weighted benchmark. Thus

$$\sum_{t=1}^{L} \pi_t \, \mathbb{E}_{\mathcal{I}_t}[Z_\tau] \leq 6N_r + 3 + \frac{12}{N_r} =: A_r,$$
$$\sum_{t=1}^{L} \pi_t \, \mathbb{E}_{\mathcal{I}_t}[Z_{(r)}] \geq \frac{3\kappa_r N_r L}{2Z_L} \geq \frac{3\kappa_r N_r L}{2}.$$

Consequently, at least one switch point t satisfies

$$\mathbb{E}_{\mathcal{I}_t}[Z_\tau] \leq \frac{2A_r}{3\kappa_r N_r L} \, \mathbb{E}_{\mathcal{I}_t}[Z_{(r)}].$$

It remains to relate the number of blocks to the horizon. Maximality of L gives $N_r(2^{L+1} - 1) > m$, and $m \geq 4N_r^2$ then implies

$$L > \log_2(m/N_r) - 1 \geq \tfrac{1}{2}\log_2 m = \frac{\log m}{2\log 2}.$$

The reward on the selected instance is therefore at most $b_r/\log m$ times its benchmark, with the explicit constants

$$b_r := \frac{4A_r \log 2}{3\kappa_r N_r}, \qquad m_0(r) := \max\{r, m_{\mathrm{lb}}, 4N_r^2\}.$$

This proves the upper bound for every $m \geq m_0(r)$ and completes the theorem.

□

For the multi-selection variant, we take the reward to be the total value of the selected items. This is the stronger reward model for an upper bound; the same conclusion then also applies when the reward is only the maximum selected value.

Corollary 7.2. *Fix an integer $r \geq 1$, and let b_r and $m_0(r)$ be upper-bound constants from Theorem 7.1. For every integer $m \geq m_0(r)$ and every moment-based algorithm that selects at most r values and receives their total value, there is a length-m instance $Z_1, \ldots, Z_m$ of independent nonnegative random variables with finite moments of every order on which the algorithm's expected reward is at most*

$$\frac{rb_r}{\log m}\,\mathbb{E}[Z_{(r)}].$$

Proof. Given an algorithm that selects at most r values, sample J uniformly from $\{1, \ldots, r\}$ before the arrivals and stop when the algorithm makes its J-th selection, receiving value 0 if it makes fewer than J selections. Conditional on the algorithm's selected values, the resulting single-choice algorithm obtains exactly $1/r$ of their total value in expectation over J. Apply the explicit upper-bound construction in the proof of Theorem 7.1 to this induced policy. Its witness instance, after multiplying the inequality by r, gives the claim.

□

This completes the lower-bound applications of the construction toolkit.

8 Proofs of the Constructions

This section supplies the deferred proofs of the two separators in Section 3. We first establish the limiting construction and decay estimates for the binary separator, followed by the analytic estimates for the N-way residue-class separator.

8.1 Proof of the Binary Separator

The goal of this section is to prove Theorem 3.1. We restate an extended version first.

Proposition 8.1. *Fix integers $n, B \geq 3$. Let*

$$x_0 = 1, \qquad x_1 = B, \qquad x_j = (nB)^{B^{j-2}} \quad (j \geq 2).$$

For each $K \geq 0$, let $(w_j^{(n,K)})_{j=0}^{K+2}$ be the unique solution of

$$\sum_{j=0}^{K+2} w_j^{(n,K)} x_j^s = 0 \quad (s = 0, \ldots, K+1), \qquad w_2^{(n,K)} = 1.$$

For each fixed $j \geq 0$, the quantity $w_j^{(n,K)}$ is defined for all $K \geq \max\{0, j-2\}$. Define

$$w_j := \lim_{K \to \infty} w_j^{(n,K)},$$

where the limit is taken over integers $K \geq \max\{0, j-2\}$.

Set

$$W^+ := \sum_{\substack{j\geq 0\\ j \text{ even}}} w_j, \qquad W^- := \sum_{\substack{j\geq 0\\ j \text{ odd}}} |w_j|.$$

Then $W^+ = W^- =: W \in (0,\infty)$. *Define probability measures* ν *and* μ *on* $\{x_j : j \geq 0\}$ *by*

$$\nu(\{x_j\}) := \begin{cases} \frac{w_j}{W}, & j \text{ even}, \\ 0, & j \text{ odd}, \end{cases} \qquad \mu(\{x_j\}) := \begin{cases} \frac{|w_j|}{W}, & j \text{ odd}, \\ 0, & j \text{ even}. \end{cases}$$

Then μ *and* ν *agree on all moments, and all of these moments are finite. Moreover, these distributions satisfy:*

(i) $p_0 := \mathbb{P}_\nu(X = 1) \geq 1 - \frac{2}{n}$ *and* $q_0 := \mathbb{P}_\mu(Y = 1) = 0$;

(ii) $p_1 := \mathbb{P}_\nu(X = B) = 0$ *and* $q_1 := \mathbb{P}_\mu(Y = B) \geq 1 - \frac{1}{n^2B^2}$;

(iii) $p_2 := \mathbb{P}_\nu(X = nB) \geq \frac{1}{4n}$ *and* $q_2 := \mathbb{P}_\mu(Y = nB) = 0$;

(iv) $W \geq n$ *and* $|w_j| \leq \frac{8(nB)^2}{x_j^2}$ *for all* $j \in \mathbb{N}$.

In accordance with the statement, let us fix integers $B, n \geq 3$, and define

$$x_0 = 1, \qquad x_1 = B, \qquad x_j = (nB)^{B^{j-2}} \quad \forall j \geq 2.$$

Note that $x_{j+1} = x_j^B \geq x_j^3$ for all $j \geq 3$. In the remainder, we first construct finite-K barycentric weights $w_j^{(n,K)}$, then pass to the limit $K \to \infty$. After proving existence of the limit in Lemma 8.1, we will abbreviate $w_j := w_j^{(n,\infty)}$ for $j \in \mathbb{N}$, which is the notation used in the statement of Theorem 3.1.

8.1.1 Constructing Distributions Agreeing on Finitely Many Moments

We now construct, for each integer $K \geq 0$, a pair of distributions μ, ν matching (finite) moments through order $K+1$. To this end, consider the linear system

$$\sum_{j=0}^{K+2} w_j^{(n,K)} x_j^s = 0, \qquad s \in \{0, 1, \ldots, K+1\},$$

together with the normalization $w_2^{(n,K)} = 1$. Since the nodes $x_0, \ldots, x_{K+2}$ are distinct, the matrix

$$V_K = (x_j^s)_{\substack{0\leq s\leq K+1\\ 0\leq j\leq K+2}} \in \mathbb{R}^{(K+2)\times(K+3)}$$

has rank $K+2$. Hence $\ker(V_K)$ is one-dimensional. The vector

$$\lambda^{(n,K)} := \left(\lambda_j^{(n,K)}\right)_{j=0}^{K+2}, \qquad \lambda_j^{(n,K)} = \left(\prod_{\substack{0\leq i\leq K+2\\ i\neq j}} (x_j - x_i)\right)^{-1},$$

lies in $\ker(V_K)$ and satisfies $\lambda_2^{(n,K)} \neq 0$ by the Lagrange interpolation identity below.

Therefore there is a unique null vector normalized by $w_2^{(n,K)} = 1$, namely $w_j^{(n,K)} = \frac{\lambda_j^{(n,K)}}{\lambda_2^{(n,K)}}$. With this notation,

$$w_j^{(n,K)} = \frac{\lambda_j^{(n,K)}}{\lambda_2^{(n,K)}} = \frac{\prod_{i\neq 2}(x_2 - x_i)}{\prod_{i\neq j}(x_j - x_i)}.$$

Indeed, for every polynomial p of degree at most $K+1$, the Lagrange interpolation formula

$$p(t) = \sum_{j=0}^{K+2} p(x_j) \prod_{\substack{0\leq i\leq K+2 \\ i\neq j}} \frac{t - x_i}{x_j - x_i}$$

shows, by comparing the coefficient of t^{K+2}, that $\sum_{j=0}^{K+2} \lambda_j^{(n,K)} p(x_j) = 0$. Applying this with $p(t) = t^s$, $s = 0, \ldots, K+1$, gives the annihilation identities, and normalizing by $\lambda_2^{(n,K)}$ enforces $w_2^{(n,K)} = 1$.

The sign pattern is explicit. The numerator $\prod_{i\neq 2}(x_2 - x_i)$ has exactly K negative factors, while the denominator $\prod_{i\neq j}(x_j - x_i)$ has exactly $K+2-j$ negative factors. Hence $\operatorname{sgn}(w_j^{(n,K)}) = (-1)^{K-(K+2-j)} = (-1)^{j-2} = (-1)^j$. Accordingly, define

$$I_+ := \{0 \leq j \leq K+2 : \ j \text{ even}\}, \qquad I_- := \{0 \leq j \leq K+2 : \ j \text{ odd}\}.$$

Then, for every $K \geq 0$,

$$\operatorname{sgn}(w_j^{(n,K)}) > 0 \iff j \in I_+, \qquad \operatorname{sgn}(w_j^{(n,K)}) < 0 \iff j \in I_-, \qquad 0 \leq j \leq K+2.$$

The relation for $s = 0$ gives $\sum_{j=0}^{K+2} w_j^{(n,K)} = 0$, so the positive and negative parts agree:

$$W^{(n,K)} := \sum_{\substack{0\leq j\leq K+2 \\ j\in I_+}} w_j^{(n,K)} = \sum_{\substack{0\leq j\leq K+2 \\ j\in I_-}} |w_j^{(n,K)}| > 0.$$

We define two probability measures on $\{x_0, \ldots, x_{K+2}\}$ by

$$\nu^{(n,K)}(\{x_j\}) := \frac{w_j^{(n,K)}}{W^{(n,K)}} \mathbf{1}_{\{j\in I_+\}}, \qquad \mu^{(n,K)}(\{x_j\}) := \frac{|w_j^{(n,K)}|}{W^{(n,K)}} \mathbf{1}_{\{j\in I_-\}}.$$

These measures have disjoint support, and for every $s \in \{0, 1, \ldots, K+1\}$, $\sum_j x_j^s \nu^{(n,K)}(\{x_j\}) = \sum_j x_j^s \mu^{(n,K)}(\{x_j\})$. Thus $\mu^{(n,K)}$ and $\nu^{(n,K)}$ match moments through order $K+1$.

8.1.2 Extending to Infinity

We now pass to the limit $K \to \infty$ and show that the resulting limiting measures $\mu^{(n,\infty)}$ and $\nu^{(n,\infty)}$ match moments of every order.

Lemma 8.1. *For each fixed $j \geq 0$, the limit $w_j^{(n,\infty)} := \lim_{K\to\infty} w_j^{(n,K)}$ exists.*

Proof. Fix $j \geq 0$, and set $K_0 := \max\{0, j-2\}$. For $K \geq K_0$, when the new node x_{K+3} is added, the normalized barycentric weights satisfy

$$w_j^{(n,K+1)} = w_j^{(n,K)} \frac{x_2 - x_{K+3}}{x_j - x_{K+3}}. \tag{1}$$

Indeed, if

$$\widetilde{w}_j^{(n,K)} := \left(\prod_{\substack{0 \leq i \leq K+2 \\ i \neq j}} (x_j - x_i) \right)^{-1},$$

then

$$\widetilde{w}_j^{(n,K+1)} = \frac{\widetilde{w}_j^{(n,K)}}{x_j - x_{K+3}}, \qquad \widetilde{w}_2^{(n,K+1)} = \frac{\widetilde{w}_2^{(n,K)}}{x_2 - x_{K+3}},$$

and dividing gives (1).

Iterating (1) from K_0 to K, we obtain

$$w_j^{(n,K)} = w_j^{(n,K_0)} \prod_{\ell=K_0+3}^{K+2} \frac{x_2 - x_\ell}{x_j - x_\ell}.$$

Since $\ell \geq K_0 + 3 \geq \max\{j+1, 3\}$, we have $x_\ell > x_j$, and therefore $\frac{x_2 - x_\ell}{x_j - x_\ell} = \frac{x_\ell - x_2}{x_\ell - x_j} = 1 + \frac{x_j - x_2}{x_\ell - x_j}$.

We claim that $x_\ell \geq 2\max\{x_j, x_2\}$ for all $\ell \geq K_0 + 3$. Indeed, if $j \leq 2$, then $\ell \geq 3$, so $x_\ell \geq x_3 = (nB)^B \geq (nB)^3 \geq 2nB = 2x_2 \geq 2x_j$. If $j \geq 3$, then $\ell \geq j+1$, so $x_\ell \geq x_{j+1} = x_j^B \geq x_j^3 \geq 2x_j$, and also $x_j \geq x_3 \geq 2x_2$, hence $x_\ell \geq 2x_2$. Thus $x_\ell - x_j \geq \frac{x_\ell}{2}$, and therefore $\left| \frac{x_j - x_2}{x_\ell - x_j} \right| \leq \frac{2(x_j + x_2)}{x_\ell}$.

Because $x_{\ell+1} = x_\ell^B \geq x_\ell^3$, the series $\sum_{\ell \geq 3} x_\ell^{-1}$ converges. Hence $\sum_{\ell=K_0+3}^{\infty} \left| \frac{x_j - x_2}{x_\ell - x_j} \right| < \infty$. Therefore the infinite product

$$\prod_{\ell=K_0+3}^{\infty} \left(1 + \frac{x_j - x_2}{x_\ell - x_j} \right)$$

converges to a finite, nonzero limit. Hence $w_j^{(n,K)}$ converges as $K \to \infty$, and $\operatorname{sgn}(w_j^{(n,\infty)}) = \operatorname{sgn}(w_j^{(n,K_0)}) = (-1)^j$, as desired. □

From now on, we write $w_j := w_j^{(n,\infty)}$ for $j \geq 0$ in agreement with the notation of Theorem 3.1.

8.1.3 Estimating Decay Rates for Limiting Weights

We next derive quantitative decay estimates for the limiting weights. The normalization into limiting probability measures will be carried out only after absolute summability has been established. We first record a convenient tail estimate for the barycentric weights, to then prove that the distributions defined above match in all (infinite) moments:

Lemma 8.2. *For every integer $M \geq 1$, define*

$$L_M := \max\left\{0, \left\lceil \frac{M-2}{1 - 1/B} \right\rceil \right\}.$$

Then, with

$$C_M^{(0,1,2)} := \max\{2n, 2nB^M, (nB)^M\}, \qquad C_M^{(\geq 3)} := \begin{cases} 8(nB)^2 & M = 1, 2, \\ 8(nB)^2 \max\{x_{L_M+2}^{M-2}, 2^{L_M}\} & M \geq 3, \end{cases}$$

the constant $C_M := \max\{C_M^{(0,1,2)}, C_M^{(\geq 3)}\}$ *satisfies*

$$|w_j| \leq \frac{C_M}{x_j^M} \qquad \forall j \geq 0.$$

Proof. Fix $M \geq 1$ and write $L := L_M$. We begin with the cases $j = 0, 1, 2$. For $j = 0$, using the explicit formula,

$$|w_0^{(n,\infty)}| = \left| \frac{x_2 - x_1}{x_1 - x_0} \prod_{\ell=3}^{\infty} \frac{x_\ell - x_2}{x_\ell - x_0} \right| \leq \frac{x_2 - x_1}{x_1 - x_0} = \frac{nB - B}{B - 1} \leq 2n = \frac{2n}{x_0^M},$$

since each factor in the product is at most 1, and simply by observing that $x_0 = 1$. Next, for $j = 1$, we similarly have

$$|w_1^{(n,\infty)}| = \frac{x_2 - x_0}{x_1 - x_0} \prod_{\ell=3}^{\infty} \frac{x_\ell - x_2}{x_\ell - x_1} \leq \frac{x_2 - x_0}{x_1 - x_0} = \frac{nB - 1}{B - 1} \leq 2n = \frac{2nB^M}{x_1^M},$$

since $x_1 = B$. For $j = 2$, recall that $w_2^{(n,\infty)} = 1$. Since $x_2 = nB$, we get $|w_2^{(n,\infty)}| = 1 = \frac{(nB)^M}{x_2^M}$. In the remainder, we let $C_M^{(0,1,2)} := \max\{2n, 2nB^M, (nB)^M\}$. We now turn to the case $j \geq 3$. Using the explicit barycentric formula and cancelling the common factor $|x_j - x_2|$, we obtain

$$|w_j| = \frac{(x_2 - x_0)(x_2 - x_1)}{(x_j - x_0)(x_j - x_1)} \prod_{m=3}^{j-1} \frac{x_m - x_2}{x_j - x_m} \prod_{\ell=j+1}^{\infty} \frac{x_\ell - x_2}{x_\ell - x_j}.$$

We first bound the first and third factors. Since $x_j \geq x_3 = (nB)^B \geq (nB)^3$, we have $x_j \geq 2$ and $x_j \geq 2B$, hence $x_j - x_0 \geq \frac{x_j}{2}, x_j - x_1 \geq \frac{x_j}{2}$. Therefore, $\frac{(x_2 - x_0)(x_2 - x_1)}{(x_j - x_0)(x_j - x_1)} \leq \frac{(nB)^2}{(x_j/2)^2} = \frac{4(nB)^2}{x_j^2}$. For the third factor, we have that,

$$\prod_{\ell=j+1}^{\infty} \frac{x_\ell - x_2}{x_\ell - x_j} \leq \exp\left(\sum_{\ell=j+1}^{\infty} \frac{2x_j}{x_\ell} \right) \leq \exp\left(\frac{4}{x_j^2} \right) \leq \exp\left(\frac{4}{x_3^2} \right) < 2.$$

The first inequality above holds using $\ln(1 + z) \leq z$ for $z \geq 0$ and since, for every $\ell \geq j + 1$, $x_\ell \geq x_{j+1} \geq x_j^3$, and so $x_\ell - x_j \geq x_\ell/2$. Thus, $\frac{x_\ell - x_2}{x_\ell - x_j} = 1 + \frac{x_j - x_2}{x_\ell - x_j} \leq 1 + \frac{2x_j}{x_\ell}$. The second inequality follows since $x_{\ell+1} \geq x_\ell^3 \geq 2x_\ell$ for $\ell \geq j + 1$, which means that $\sum_{\ell=j+1}^{\infty} \frac{1}{x_\ell} \leq \frac{2}{x_{j+1}} \leq \frac{2}{x_j^3}$. The third follows by recalling that $x_j \geq x_3$, and the final bound uses $x_3 = (nB)^B \geq 6^3 = 216$. Combining these two bounds, we obtain for every $j \geq 3$,

$$|w_j| \leq \frac{8(nB)^2}{x_j^2} \prod_{m=3}^{j-1} \frac{x_m - x_2}{x_j - x_m}.$$

If $M = 1$ or $M = 2$, then $L = 0$, and since the middle product is at most 1, we immediately get

$$|w_j| \le \frac{8(nB)^2}{x_j^2} \le \frac{8(nB)^2}{x_j^M} \qquad \forall j \ge 3.$$

Thus, assume now that $M \ge 3$, so $L \ge 1$. First consider the range $3 \le j \le L+2$. Since $M - 2 \ge 1$ and $x_j \le x_{L+2}$, the previous $M = 2$ estimate yields

$$|w_j| \le \frac{8(nB)^2}{x_j^2} = \frac{8(nB)^2 x_j^{M-2}}{x_j^M} \le \frac{8(nB)^2 x_{L+2}^{M-2}}{x_j^M}.$$

Now consider $j \ge L+3$. For each $m \in \{3, \dots, j-1\}$, we have $x_m > x_2$, so the factor is positive. Moreover, since $j \ge m+1$, we have $x_j \ge x_{m+1} = x_m^B$, and hence $x_j - x_m \ge x_m^B - x_m$. Also, because $m \ge 3$, we have $x_m \ge x_2 > 0$, and since $B \ge 3$, it follows that $x_m^B \ge x_m^2 \ge 2x_m$. Therefore $x_m^B - x_m \ge x_m \ge x_m - x_2$, and so $0 < \frac{x_m - x_2}{x_j - x_m} \le 1$. Therefore

$$\prod_{m=3}^{j-1} \frac{x_m - x_2}{x_j - x_m} \le \prod_{m=j-L}^{j-1} \frac{x_m - x_2}{x_j - x_m}.$$

Fix $m \in \{j-L, \dots, j-1\}$. Since $m \ge 3$, we have $x_m \ge x_2$, and since $j > m$, $x_j \ge x_{m+1} = x_m^B \ge x_m^3 \ge 2x_m$. Therefore, $0 < \frac{x_m - x_2}{x_j - x_m} \le \frac{x_m}{x_j/2} = \frac{2x_m}{x_j}$. Also $m \le j-1$, and hence $x_m \le x_{j-1} = x_j^{1/B}$. This means that

$$\frac{x_m - x_2}{x_j - x_m} \le \frac{2}{x_j^{1-1/B}}.$$

Multiplying over $m = j-L, \dots, j-1$, we obtain $\prod_{m=3}^{j-1} \frac{x_m - x_2}{x_j - x_m} \le \left(\frac{2}{x_j^{1-1/B}}\right)^L$. Overall, for $j \ge L+3$,

$$|w_j| \le \frac{8(nB)^2}{x_j^2} \left(\frac{2}{x_j^{1-1/B}}\right)^L = 8(nB)^2 2^L x_j^{-2-L(1-1/B)} \le \frac{8(nB)^2 2^L}{x_j^M},$$

since, by definition, $2 + L(1-1/B) \ge M$. Combining the bounds for $3 \le j \le L+2$ and $j \ge L+3$, we conclude that for all $j \ge 3$, $|w_j| \le \frac{C_M^{(\ge 3)}}{x_j^M}$, where $C_M^{(\ge 3)} := 8(nB)^2 \max\{x_{L+2}^{M-2}, 2^L\}$. Hence, the constant $C_M := \max\{C_M^{(0,1,2)}, C_M^{(\ge 3)}\}$ satisfies

$$|w_j| \le \frac{C_M}{x_j^M} \qquad \forall j \ge 0,$$

which proves the lemma. □

We also stress that the same decay bound holds uniformly in K:

Lemma 8.3. *For every integer $M \ge 1$, the constant C_M from Lemma 8.2 also satisfies*

$$|w_j^{(n,K)}| \le \frac{C_M}{x_j^M} \qquad \forall K \ge 0,\ \forall j \in \{0, \dots, K+2\}.$$

Proof. For $j = 0, 1, 2$, the same estimates as in the proof of Lemma 8.2 give $|w_0^{(n,K)}| \le 2n$, $|w_1^{(n,K)}| \le 2n$, $|w_2^{(n,K)}| = 1$. For $j \ge 3$, the finite-K barycentric formula reads

$$|w_j^{(n,K)}| = \frac{(x_2 - x_0)(x_2 - x_1)}{(x_j - x_0)(x_j - x_1)} \prod_{m=3}^{j-1} \frac{x_m - x_2}{x_j - x_m} \prod_{\ell=j+1}^{K+2} \frac{x_\ell - x_2}{x_\ell - x_j}.$$

The proof of Lemma 8.2 uses only upper bounds on the factors, and the final product over ℓ is smaller here than in the limiting case, because all its factors are > 1. Therefore the same argument yields $|w_j^{(n,K)}| \le \frac{C_M}{x_j^M}$ $(j \ge 3)$. Combining the cases $j = 0, 1, 2$ and $j \ge 3$ proves the claim. □

Lemma 8.4. *For every integer $s \ge 0$,*

$$\sum_{j \ge 0} w_j x_j^s = 0.$$

Proof. Fix $s \ge 0$, and set $M := s + 2$. For each $K \ge 0$ and $j \ge 0$, define

$$a_{K,j} := \begin{cases} w_j^{(n,K)} x_j^s, & 0 \le j \le K + 2, \\ 0, & j > K + 2. \end{cases}$$

For each fixed j, Lemma 8.1 gives $a_{K,j} \to w_j x_j^s$ for $K \to \infty$. By Lemma 8.3, $|a_{K,j}| \le \frac{C_M x_j^s}{x_j^M} = \frac{C_M}{x_j^2}$ for all K, j. Since $x_{j+1} \ge x_j^3 \ge 2x_j$ for all $j \ge 3$, the series $\sum_{j \ge 0} x_j^{-2}$ converges. Hence dominated convergence for the counting measure on $\mathbb{N}$ yields

$$\sum_{j \ge 0} w_j x_j^s = \lim_{K \to \infty} \sum_{j \ge 0} a_{K,j} = \lim_{K \to \infty} \sum_{j=0}^{K+2} w_j^{(n,K)} x_j^s.$$

For every $K \ge \max\{0, s - 1\}$, the last sum is 0 by the defining moment-annihilation identities. Therefore the limit is 0. □

This allows us to conclude that the limiting probabilities are well-defined:

Corollary 8.1. *For the limiting weights, now define*

$$I_+ := \{j \ge 0 : \ j \text{ even}\}, \qquad I_- := \{j \ge 0 : \ j \text{ odd}\}.$$

Then the series $\sum_{j \ge 0} |w_j|$ converges, and

$$W := \sum_{j \in I_+} w_j = \sum_{j \in I_-} |w_j| \in (0, \infty).$$

Consequently,

$$\nu^{(n,\infty)}(\{x_j\}) =: p_j := \frac{w_j}{W} \mathbf{1}_{\{j \in I_+\}}, \qquad \mu^{(n,\infty)}(\{x_j\}) =: q_j := \frac{|w_j|}{W} \mathbf{1}_{\{j \in I_-\}}$$

define two probability measures on $\{x_j : j \ge 0\}$. Moreover, for every $M \ge 1$,

$$p_j + q_j = \frac{|w_j|}{W} \le \frac{C_M}{W x_j^M} \qquad \forall j \ge 0.$$

Proof. Taking $M = 2$ in Lemma 8.2, we get $\sum_{j\geq 0} |w_j| \leq C_2 \sum_{j\geq 0} \frac{1}{x_j^2} < \infty$. By Lemma 8.4 with $s = 0$, $\sum_{j\geq 0} w_j = 0$. Also $w_2^{(n,\infty)} = \lim_{K\to\infty} w_2^{(n,K)} = 1 > 0$, so the positive part is nontrivial; hence the negative part is also nontrivial because the total signed sum is 0. Therefore $\sum_{j\in I_+} w_j = \sum_{j\in I_-} |w_j| =: W \in (0,\infty)$. The stated formulas for p_j, q_j are nonnegative and have total mass 1, so they define probability measures. The final bound follows immediately from Lemma 8.2. □

We are now ready to show that the limiting measures have all moments matching:

Lemma 8.5. *For every integer $s \geq 0$,*

$$\sum_{j\geq 0} x_j^s p_j < \infty, \qquad \sum_{j\geq 0} x_j^s q_j < \infty,$$

and

$$\sum_{j\geq 0} x_j^s p_j = \sum_{j\geq 0} x_j^s q_j.$$

Therefore, $\mu^{(n,\infty)}$ and $\nu^{(n,\infty)}$ match moments of every order.

Proof. Fix $s \geq 0$. By Corollary 8.1 with $M = s + 2$, $x_j^s(p_j + q_j) \leq \frac{C_{s+2}}{W x_j^2}$ $(j \geq 0)$. Since $\sum_{j\geq 0} x_j^{-2} < \infty$, both series

$$\sum_{j\geq 0} x_j^s p_j, \qquad \sum_{j\geq 0} x_j^s q_j$$

converge absolutely. Thus the s-th moments are finite. Moreover, by the definitions of p_j and q_j, $p_j - q_j = \frac{w_j}{W}$ for all $j \geq 0$. Therefore,

$$\sum_{j\geq 0} x_j^s p_j - \sum_{j\geq 0} x_j^s q_j = \frac{1}{W} \sum_{j\geq 0} w_j x_j^s = 0$$

by Lemma 8.4. Hence the s-th moments are equal for every $s \geq 0$. □

Having established all moments matching, we can now derive some useful bounds on the probabilities:

Lemma 8.6. *It holds that*

$$W \geq n, \qquad p_2 \geq \frac{1}{4n}, \qquad p_0 \geq 1 - \frac{2}{n}, \qquad q_1 \geq 1 - \frac{1}{n^2 B^2}.$$

Proof. We begin with the lower bound on p_2: for $w_0^{(n,\infty)}$, the explicit formula gives

$$w_0^{(n,\infty)} = \frac{x_2 - x_1}{x_1 - x_0} \prod_{\ell=3}^{\infty} \frac{x_\ell - x_2}{x_\ell - x_0} \leq \frac{nB - B}{B - 1} \leq 2n,$$

since each factor in the product is at most 1. Observe that Lemma 8.2 applied with $M = 2$ gives

$$|w_j| \leq \frac{8(nB)^2}{x_j^2} \qquad \forall j \geq 3.$$

Therefore, we also have

$$\sum_{\substack{j\geq 4\\ j\in I_+}} w_j \leq \sum_{j=4}^{\infty} |w_j| \leq 8(nB)^2 \sum_{j=4}^{\infty} \frac{1}{x_j^2} \leq \frac{16(nB)^2}{x_4^2} < 1,$$

where the third inequality holds because, as $x_{j+1} \geq x_j^3 \geq 2x_j$, the sequence $1/x_j^2$ is dominated by a geometric progression with ratio at most $1/2$, i.e., $\sum_{j=4}^{\infty} \frac{1}{x_j^2} \leq \frac{2}{x_4^2}$. Hence, recalling that $w_2^{(n,\infty)} = 1$,

$$W = w_0^{(n,\infty)} + w_2^{(n,\infty)} + \sum_{\substack{j\geq 4\\ j\in I_+}} w_j \leq 2n + 1 + 1 \leq 4n \implies p_2 = \frac{w_2^{(n,\infty)}}{W} \geq \frac{1}{4n}.$$

We proceed with the lower bound on p_0. Using the explicit formula for $w_1^{(n,\infty)}$, we have

$$|w_1^{(n,\infty)}| = \frac{x_2 - x_0}{x_1 - x_0} \prod_{\ell=3}^{\infty} \frac{x_\ell - x_2}{x_\ell - x_1} = \frac{nB-1}{B-1} \prod_{\ell=3}^{\infty} \left(1 - \frac{x_2 - x_1}{x_\ell - x_1}\right).$$

Since $x_\ell \geq x_3 = (nB)^B \geq 2B$, we have $x_\ell - x_1 = x_\ell - B \geq x_\ell/2$, and therefore, $0 \leq \frac{x_2 - x_1}{x_\ell - x_1} = \frac{nB-B}{x_\ell - B} \leq \frac{2nB}{x_\ell}$. Moreover, $x_{\ell+1} \geq x_\ell^3 \geq 2x_\ell$ for all $\ell \geq 3$, so $\sum_{\ell=3}^{\infty} \frac{1}{x_\ell} \leq \frac{2}{x_3}$. Hence,

$$\sum_{\ell=3}^{\infty} \frac{x_2 - x_1}{x_\ell - x_1} \leq 2nB \sum_{\ell=3}^{\infty} \frac{1}{x_\ell} \leq \frac{4nB}{x_3} = \frac{4}{(nB)^{B-1}} \leq \frac{4}{(nB)^2}.$$

Now we first lower-bound $|w_1|$. Set

$$a_\ell := \frac{x_2 - x_1}{x_\ell - x_1} \qquad (\ell \geq 3).$$

Then

$$|w_1| = \frac{x_2 - x_0}{x_1 - x_0} \prod_{\ell=3}^{\infty} (1 - a_\ell) = \frac{nB-1}{B-1} \prod_{\ell=3}^{\infty} (1 - a_\ell).$$

Since $x_\ell \geq x_3 = (nB)^B$ and $n, B \geq 3$, we have $0 \leq a_\ell \leq \frac{2nB}{x_\ell} \leq \frac{2}{(nB)^{B-1}} < 1$. Hence, for every $L \geq 3$, $\prod_{\ell=3}^{L}(1 - a_\ell) \geq 1 - \sum_{\ell=3}^{L} a_\ell$. Letting $L \to \infty$ and using $\sum_{\ell=3}^{\infty} a_\ell \leq 4/(nB)^2$, we obtain $\prod_{\ell=3}^{\infty}(1 - a_\ell) \geq 1 - \frac{4}{(nB)^2}$. Therefore

$$|w_1| \geq \frac{nB-1}{B-1} \left(1 - \frac{4}{(nB)^2}\right).$$

Hence

$$|w_1| - n \geq \frac{nB-1}{B-1} \left(1 - \frac{4}{(nB)^2}\right) - n = \frac{n-1}{B-1} - \frac{4(nB-1)}{(B-1)n^2B^2}.$$

Since $n, B \geq 3$, we have $\frac{4(nB-1)}{n^2B^2} \leq \frac{4}{nB} \leq \frac{4}{9} < 1 \leq n - 1$, and therefore $\frac{4(nB-1)}{(B-1)n^2B^2} \leq \frac{n-1}{B-1}$. Thus $|w_1| \geq n$, and consequently $W = \sum_{j\in I_-} |w_j| \geq |w_1| \geq n$. We now use again the bound on the sum of the remaining weights for the even indices $j \geq 4$, $1 - p_0 = \frac{w_2^{(n,\infty)} + \sum_{\substack{j\geq 4\\ j\in I_+}} w_j}{W} \leq \frac{1+1}{n} = \frac{2}{n}$. We conclude

with the lower bound on q_1: since the negative tail beyond x_1 is supported on odd $j \geq 3$, Lemma 8.2 applied with $M = 2$ gives

$$\sum_{\substack{j\geq 3\\ j\in I_-}} |w_j| \leq 8(nB)^2 \sum_{j=3}^{\infty} \frac{1}{x_j^2} \leq \frac{16(nB)^2}{x_3^2} = \frac{16}{(nB)^{2B-2}}.$$

Since $W \geq n, B \geq 3$ and $nB^2 \geq 18$,

$$1 - q_1 = \sum_{\substack{j\geq 3\\ j\in I_-}} q_j = \frac{1}{W} \sum_{\substack{j\geq 3\\ j\in I_-}} |w_j| \leq \frac{16}{n(nB)^{2B-2}} \leq \frac{16}{n^3B^4} \leq \frac{1}{n^2B^2}.$$

This proves the lemma. □

8.1.4 Finishing the Proof

Proof of Theorem 3.1. Set $\nu := \nu^{(n,\infty)}$ and $\mu := \mu^{(n,\infty)}$. By Corollary 8.1, these are well-defined probability measures, and $W^+ = W^- = W$. By Lemma 8.5, μ and ν agree on all moments, and all of these moments are finite.

Points (i)-(iii) are exactly the support-probability bounds proved in Lemma 8.6, together with the trivial support facts $q_0 = 0$, $p_1 = 0$, and $q_2 = 0$. Finally, point (iv) combines the bound $W \geq n$ from Lemma 8.6 with the case $M = 2$ of Lemma 8.2, which gives $|w_j| \leq \frac{8(nB)^2}{x_j^2}$ for all $j \geq 0$. This proves the proposition. □

8.2 Proof of the N-Way Separator

The goal of this section is to prove Theorem 3.2, of which we state an extended version below:

Proposition 8.2. *Fix integers $B, N \geq 2$ and define*

$$G_N(x) := \prod_{m=0}^{\infty} \sum_{u=0}^{N-1} (xB^{-m})^u = \sum_{j=0}^{\infty} w_j x^j.$$

For $h \in \{0, \ldots, N-1\}$ and $t \in \mathbb{N}_0$, let

$$W_h(t) := \sum_{\substack{j\geq 0\\ j\equiv h \pmod N}} w_j B^{tj}.$$

Then

$$W_h(0) \in (0, \infty) \qquad (h = 0, \ldots, N-1).$$

Define

$$\mu_h(\{B^j\}) := \begin{cases} \dfrac{w_j}{W_h(0)}, & j \equiv h \pmod N, \\ 0, & \text{otherwise.} \end{cases}$$

Then the measures $(\mu_h)_{h=0}^{N-1}$ agree on all moments, all of these moments are finite, and the normalizing constants satisfy $W_h(0) = \sum_{\substack{j\geq 0\\ j\equiv h \pmod N}} w_j \in (0, \infty)$ for $h \in \{0, \ldots, N-1\}$. Also, writing each j as $j = L(N-1) + b$ for $L \in \mathbb{N}_0, b \in \{0, \ldots, N-2\}$, and defining $e(j) := \frac{(N-1)L(L-1)}{2} + Lb$, it holds that

$$B^{-e(j)} \leq w_j \leq 2^j B^{-e(j)}.$$

Moreover, we have that

$$1 \le W_0(0) = \ldots = W_{N-1}(0) < 4,$$

as well as $j - e(j) \le N-1$ *with equality if and only if* $j \in \{N-1, N, \ldots, 2N-2\}$.

8.2.1 Structural Properties of the Construction

We next prove a sequence of structural properties of the construction that will establish the proposition:

Lemma 8.7. *Fix integers* $N \ge 2$ *and* $B \ge 2$*, and set* $q := 1/B$*. For each* $M \in \mathbb{N}_0$*, define*

$$G_{N,M}(x) := \prod_{m=0}^{M} \sum_{u=0}^{N-1} (xq^m)^u.$$

Then:

(i) The infinite product

$$G_N(x) := \prod_{m=0}^{\infty} \sum_{u=0}^{N-1} (xq^m)^u$$

converges locally uniformly on $\mathbb{C}$*, and hence defines an entire function.*

(ii) Writing

$$G_{N,M}(x) = \sum_{j \ge 0} w_j^{(M)} x^j, \qquad G_N(x) = \sum_{j \ge 0} w_j x^j,$$

one has, for every $j \ge 0$*,*

$$w_j^{(M)} = \sum_{\substack{(c_0,\ldots,c_M) \\ c_m \in \{0,\ldots,N-1\} \\ \sum_{m=0}^{M} c_m = j}} q^{\sum_{m=0}^{M} m c_m}.$$

(iii) For every fixed $j \ge 0$*, the sequence* $w_j^{(M)}$ *is nondecreasing in* M *and* $w_j^{(M)} \uparrow w_j$ *as* $M \to \infty$*. Consequently,*

$$w_j = \sum_{\substack{(c_m)_{m \ge 0} \\ c_m \in \{0,\ldots,N-1\} \\ \sum_{m \ge 0} c_m = j}} q^{\sum_{m \ge 0} m c_m} = \sum_{\substack{(c_m)_{m \ge 0} \\ c_m \in \{0,\ldots,N-1\} \\ \sum_{m \ge 0} c_m = j}} B^{-\sum_{m \ge 0} m c_m}.$$

Proof. For each $m \ge 0$, set $F_m(x) := \sum_{u=0}^{N-1} (xq^m)^u$. Fix $R > 0$. For $|x| \le R$,

$$|F_m(x) - 1| \le \sum_{u=1}^{N-1} |x|^u q^{um} \le \sum_{u=1}^{N-1} R^u q^{um}.$$

Hence

$$\sum_{m=0}^{\infty} \sup_{|x| \le R} |F_m(x) - 1| \le \sum_{u=1}^{N-1} R^u \sum_{m=0}^{\infty} q^{um} < \infty.$$

Therefore the infinite product $\prod_{m=0}^{\infty} F_m(x)$ converges locally uniformly on $\mathbb{C}$ to a holomorphic function, proving (i).

Now fix $M \in \mathbb{N}_0$. Since $G_{N,M}$ is a finite product of polynomials, expanding it gives

$$G_{N,M}(x) = \sum_{(c_0,\ldots,c_M)\in\{0,\ldots,N-1\}^{M+1}} \prod_{m=0}^{M} (xq^m)^{c_m}.$$

Collecting powers of x, we obtain

$$G_{N,M}(x) = \sum_{(c_0,\ldots,c_M)\in\{0,\ldots,N-1\}^{M+1}} x^{\sum_{m=0}^{M} c_m} q^{\sum_{m=0}^{M} m c_m},$$

so the coefficient of x^j is exactly

$$w_j^{(M)} = \sum_{\substack{(c_0,\ldots,c_M)\\ c_m\in\{0,\ldots,N-1\}\\ \sum_{m=0}^{M} c_m = j}} q^{\sum_{m=0}^{M} m c_m}.$$

This proves (ii). Next, because $G_{N,M+1}(x) = G_{N,M}(x)\, F_{M+1}(x)$ and all coefficients of $F_{M+1}(x)$ are nonnegative, every coefficient of $G_{N,M+1}$ is at least the corresponding coefficient of $G_{N,M}$. Thus, for each fixed j, $w_j^{(M+1)} \geq w_j^{(M)}$. So $(w_j^{(M)})_{M\geq 0}$ is nondecreasing.

Since $G_{N,M} \to G_N$ locally uniformly on $\mathbb{C}$, the Taylor coefficients converge coefficientwise. For example, fixing any $r > 0$, Cauchy's coefficient formula gives

$$w_j^{(M)} = \frac{1}{2\pi i} \int_{|z|=r} \frac{G_{N,M}(z)}{z^{j+1}}\, dz \longrightarrow \frac{1}{2\pi i} \int_{|z|=r} \frac{G_N(z)}{z^{j+1}}\, dz = w_j.$$

Hence $w_j^{(M)} \uparrow w_j$. Now, for fixed $j \geq 0$, let

$$\mathcal{E}_{j,M} := \left\{ (c_m)_{m\geq 0} : c_m \in \{0,\ldots,N-1\},\ \sum_{m\geq 0} c_m = j,\ c_m = 0\ \forall m > M \right\},$$

and let $\mathcal{E}_j := \bigcup_{M\geq 0} \mathcal{E}_{j,M}$. Then part (ii) gives $w_j^{(M)} = \sum_{(c_m)\in\mathcal{E}_{j,M}} q^{\sum_{m\geq 0} m c_m}$. Since $\mathcal{E}_{j,M} \uparrow \mathcal{E}_j$ and all summands are nonnegative, monotone convergence yields $\lim_{M\to\infty} w_j^{(M)} = \sum_{(c_m)\in\mathcal{E}_j} q^{\sum_{m\geq 0} m c_m}$. Because $w_j^{(M)} \to w_j$, we obtain

$$w_j = \sum_{\substack{(c_m)_{m\geq 0}\\ c_m\in\{0,\ldots,N-1\}\\ \sum_{m\geq 0} c_m = j}} q^{\sum_{m\geq 0} m c_m}.$$

This is well defined because every admissible sequence $(c_m)_{m\geq 0}$ has finite support once $\sum_{m\geq 0} c_m = j < \infty$. Recalling $q = 1/B$ gives the last identity. □

With these coefficients, we now show that the residue-class sums $W_h(t)$ are equal for every moment order $t \in \mathbb{N}_0$.

Lemma 8.8. *For every $t \in \mathbb{N}_0$ and every $h \in \{0,\ldots,N-1\}$, let the residue-class sums be defined as $W_h(t) := \sum_{\substack{j\geq 0\\ j\equiv h \pmod N}} w_j B^{tj}$. Then, for all $h \in \{0,\ldots,N-1\}$, it holds that: $W_h(t) = \frac{G_N(B^t)}{N}$. Moreover, letting $W := W_0(0) = \ldots = W_{N-1}(0)$, we have that $1 \leq W < 4$.*

Proof. Let $\omega := e^{2\pi \mathrm{i}/N}$ be a primitive N-th root of unity. Fix $t \in \mathbb{N}_0$ and $\ell \in \{1, \dots, N-1\}$. In the defining product for $G_N(\omega^\ell B^t)$, the factor with index $m = t$ is $\sum_{u=0}^{N-1} (\omega^\ell)^u = 0$, hence $G_N(\omega^\ell B^t) = 0$. Since $\sum_{j=0}^{\infty} \left| w_j (\omega^\ell B^t)^j \right| = \sum_{j=0}^{\infty} w_j B^{tj} = G_N(B^t) < \infty$, the series is absolutely convergent, so we may regroup it by residue classes modulo N. Therefore

$$0 = \sum_{j=0}^{\infty} w_j (\omega^\ell B^t)^j = \sum_{h=0}^{N-1} \sum_{\substack{j \geq 0 \\ j \equiv h \pmod N}} w_j (\omega^\ell B^t)^j = \sum_{h=0}^{N-1} \omega^{\ell h} \sum_{\substack{j \geq 0 \\ j \equiv h \pmod N}} w_j B^{tj} = \sum_{h=0}^{N-1} \omega^{\ell h} W_h(t).$$

Note also that, by definition of $W_h(t)$'s, we have that

$$\sum_{h=0}^{N-1} W_h(t) = \sum_{h=0}^{N-1} \sum_{\substack{j \geq 0 \\ j \equiv h \pmod N}} w_j B^{tj} = \sum_{j=0}^{\infty} w_j B^{tj} = G_N(B^t).$$

We can summarize, for every fixed $t \in \mathbb{N}_0$, all of the above in a system of N linear equations (indexed by $\ell \in \{0, \dots, N-1\}$) in N unknowns (indexed by $h \in \{0, \dots, N-1\}$). Namely, consider the matrix $\mathbf{\Omega} \in \mathbb{C}^{N \times N}$ that consists of $\mathbf{\Omega}_{\ell h} = \omega^{\ell h}$'s as entries and the vector $\mathbf{W}(t) \in \mathbb{R}_+^N$ consists of $W_h(t)$'s as entries. Then,

$$\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \cdots & \omega^{N-1} \\ 1 & \omega^2 & \omega^4 & \cdots & \omega^{2(N-1)} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \omega^{N-1} & \omega^{2(N-1)} & \cdots & \omega^{(N-1)^2} \end{pmatrix} \begin{pmatrix} W_0(t) \\ W_1(t) \\ W_2(t) \\ \vdots \\ W_{N-1}(t) \end{pmatrix} = \begin{pmatrix} G_N(B^t) \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

We stress that $\mathbf{\Omega}$ is the Vandermonde matrix of the N-th roots of unity and we know that this is invertible because $1 \neq \omega \neq \cdots \neq \omega^{N-1}$, since ω is a primitive N-th root of unity. Thus, the solution to the above must be unique. Suppose it takes the form $\mathbf{W}(t) = \gamma \mathbf{1}$ for some $\gamma \in \mathbb{R}$, then it must satisfy $\gamma \mathbf{1}^\top \mathbf{1} = G_N(B^t)$, i.e., $\gamma = \frac{G_N(B^t)}{N}$, which, for all $\ell \in \{1, \dots, N-1\}$, also satisfies $\gamma \sum_{h=0}^{N-1} \omega^{\ell h} = 0$, since $\sum_{h=0}^{N-1} \omega^{\ell h} = \frac{1-\omega^{\ell N}}{1-\omega^\ell} = 0$. Hence, $W_0(t) = W_1(t) = \cdots = W_{N-1}(t) = \frac{G_N(B^t)}{N}$.

To conclude the proof, observe that

$$G_N(1) = N \prod_{m=1}^{\infty} \sum_{u=0}^{N-1} q^{um} = N \prod_{m=1}^{\infty} \frac{1 - q^{Nm}}{1 - q^m},$$

and so

$$W = \frac{G_N(1)}{N} = \prod_{m=1}^{\infty} \frac{1 - q^{Nm}}{1 - q^m} \leq \prod_{m=1}^{\infty} \frac{1}{1 - 2^{-m}} \leq \frac{32}{9} < 4,$$

For the last bound, $\prod_{m=3}^{\infty}(1 - 2^{-m}) \geq 1 - \sum_{m=3}^{\infty} 2^{-m} = 3/4$; including the first two factors gives $\prod_{m=1}^{\infty}(1 - 2^{-m}) \geq 9/32$. Finally, $w_0 = 1$ and $W = W_0(0) \geq w_0$, so $W \geq 1$. □

Based on the above result, using the definition $W := W_0(0) = \cdots = W_{N-1}(0)$ and, for each $h \in \{0, \dots, N-1\}$, the distributions descriptions simplify to

$$\mu_h(\{B^j\}) := \begin{cases} \dfrac{w_j}{W_h(0)} & j \equiv h \pmod N \\ 0 & \text{otherwise} \end{cases} = \begin{cases} \dfrac{w_j}{W}, & j \equiv h \pmod N \\ 0, & \text{otherwise} \end{cases}.$$

We can now observe that these distributions share all moments, which are additionally finite:

Lemma 8.9. *For every $h, h' \in \{0, \dots, N-1\}$ and every $t \in \mathbb{N}_0$, $\int x^t \, d\mu_h(x) = \int x^t \, d\mu_{h'}(x) < \infty$.*

Proof. By Lemma 8.8,

$$\int x^t d\mu_h(x) = \frac{1}{W} \sum_{\substack{j \geq 0 \\ j \equiv h \pmod N}} w_j B^{tj} = \frac{W_h(t)}{W} = \frac{W_{h'}(t)}{W} = \int x^t d\mu_{h'}(x).$$

Since G_N is entire and the coefficients are nonnegative, the series $\sum_{j=0}^{\infty} w_j B^{tj} = G_N(B^t)$ converges for every $t \in \mathbb{N}_0$, so all moments are finite. □

8.2.2 Estimating Decay Rates of Limiting Weights

We can next provide bounds on the coefficients w_j that define the distributions above. To this end, it is useful to write for the remainder of the proof each $j \in \mathbb{N}_0$ as $j = L(N-1) + b$ for $L \in \mathbb{N}_0, b \in \{0, \dots, N-2\}$, and to define $e(j) := \frac{(N-1)L(L-1)}{2} + Lb$.

Lemma 8.10. *Assume $B \geq 2$. Then for every $j \in \mathbb{N}_0$, $B^{-e(j)} \leq w_j \leq 2^j B^{-e(j)}$. Moreover, $j - e(j) \leq N-1$ with equality if and only if $j \in \{N-1, N, \dots, 2N-2\}$.*

Proof. By Lemma 8.7, for every $j \in \mathbb{N}_0$,

$$w_j = \sum_{\substack{(c_m)_{m \geq 0} \\ c_m \in \{0, \dots, N-1\} \\ \sum_{m \geq 0} c_m = j}} B^{-\sum_{m \geq 0} m c_m}.$$

We use this formula to prove both bounds. Write $j = L(N-1) + b$ for $L \in \mathbb{N}_0, b \in \{0, \dots, N-2\}$, and recall that $e(j) := \frac{(N-1)L(L-1)}{2} + Lb$.

We begin with the lower bound. Consider the admissible sequence $c_0 = \cdots = c_{L-1} = N-1, \quad c_L = b, \quad c_m = 0$ for all $m \geq L+1$. Its total mass is $\sum_{m \geq 0} c_m = L(N-1) + b = j$, so it contributes to the above sum for w_j. The corresponding exponent is

$$\sum_{m \geq 0} m c_m = \sum_{m=0}^{L-1} m(N-1) + Lb = (N-1)\frac{L(L-1)}{2} + Lb = e(j).$$

Since all terms in the coefficient formula are nonnegative, this single contribution yields $w_j \geq B^{-e(j)}$.

We next prove the upper bound. For each admissible sequence $(c_m)_{m \geq 0}$, let $0 \leq t_1 \leq t_2 \leq \cdots \leq t_j$ be the nondecreasing sequence obtained by listing each index m exactly c_m times. Then $\sum_{m \geq 0} m c_m = \sum_{i=1}^{j} t_i$. Moreover, because each $c_m \leq N-1$, no integer can occur more than $N-1$ times in the sequence $(t_i)_{i=1}^{j}$. It follows that $t_i \geq \left\lfloor \frac{i-1}{N-1} \right\rfloor$ for every $i = 1, \dots, j$. Therefore we may write $t_i = \left\lfloor \frac{i-1}{N-1} \right\rfloor + \delta_i$, for $\delta_i \in \mathbb{N}_0$. Summing over i gives

$$\sum_{i=1}^{j} t_i = \sum_{i=1}^{j} \left\lfloor \frac{i-1}{N-1} \right\rfloor + \sum_{i=1}^{j} \delta_i.$$

Now

$$\sum_{i=1}^{j} \left\lfloor \frac{i-1}{N-1} \right\rfloor = \sum_{r=0}^{L-1} r(N-1) + Lb = \frac{(N-1)L(L-1)}{2} + Lb = e(j),$$

because among the integers $0, 1, \ldots, j-1 = L(N-1)+b-1$, each value $r \in \{0, \ldots, L-1\}$ occurs exactly $N-1$ times in the quotient $\lfloor (i-1)/(N-1) \rfloor$, and the value L occurs exactly b times. Hence $\sum_{m\geq 0} m c_m = \sum_{i=1}^{j} t_i = e(j) + \sum_{i=1}^{j} \delta_i$.

Moreover, the admissible sequence $(c_m)_{m\geq 0}$ uniquely determines the nondecreasing list $0 \leq t_1 \leq \cdots \leq t_j$, and hence uniquely determines the tuple $\delta_i := t_i - \left\lfloor \frac{i-1}{N-1} \right\rfloor \in \mathbb{N}_0 \quad (i = 1, \ldots, j)$. Thus, the map from admissible sequences to δ-tuples is injective. Therefore,

$$w_j = \sum_{\text{admissible } (c_m)} B^{-e(j)-\sum_{i=1}^{j}\delta_i} \leq B^{-e(j)} \sum_{\delta_1,\ldots,\delta_j \geq 0} B^{-\sum_{i=1}^{j}\delta_i}.$$

The variables now separate, so

$$\sum_{\delta_1,\ldots,\delta_j \geq 0} B^{-\sum_{i=1}^{j}\delta_i} = \left(\sum_{h=0}^{\infty} B^{-h}\right)^j = \left(\frac{1}{1-B^{-1}}\right)^j \leq 2^j,$$

since $B \geq 2$. Therefore $w_j \leq 2^j B^{-e(j)}$. This proves $B^{-e(j)} \leq w_j \leq 2^j B^{-e(j)}$. It remains to estimate $j - e(j)$. Using $j = L(N-1)+b$, we compute $j - e(j) = L(N-1) + b - \frac{(N-1)L(L-1)}{2} - Lb$ and distinguish cases:

1. If $L = 0$, then $j = b$ and $j - e(j) = b \leq N-2$.
2. If $L = 1$, then $j - e(j) = (N-1) + b - b = N-1$. Since in this case $j = (N-1)+b$ with $b \in \{0, \ldots, N-2\}$, this is exactly the range $j \in \{N-1, N, \ldots, 2N-3\}$.
3. If $L = 2$, then $j - e(j) = 2(N-1) + b - (N-1) - 2b = N-1-b$. Thus, $j - e(j) = N-1$ if and only if $b = 0$, i.e., if and only if $j = 2N-2$, and otherwise $j - e(j) \leq N-2$.
4. Finally, if $L \geq 3$, then $j - e(j) = \frac{(3L-L^2)(N-1)}{2} + (1-L)b$. Here $3L - L^2 \leq 0$ and $1 - L \leq 0$, so $j - e(j) \leq 0 \leq N-2$.

Combining all cases, we conclude that $j - e(j) \leq N-1$, with equality if and only if $j \in \{N-1, N, \ldots, 2N-2\}$. This completes the proof. □

8.2.3 Finishing the Proof

We now conclude the proof of Theorem 3.2:

Proof of Theorem 3.2. By Lemma 8.7, all coefficients w_j are nonnegative. Moreover, by Lemma 8.8,

$$W_0(0) = \cdots = W_{N-1}(0) =: W, \qquad 1 \leq W < 4.$$

Hence, for each $h \in \{0, \ldots, N-1\}$, the formula

$$\mu_h(\{B^j\}) := \begin{cases} \frac{w_j}{W}, & j \equiv h \pmod{N}, \\ 0, & \text{otherwise}, \end{cases}$$

defines a probability measure supported on $\{B^j : j \equiv h \pmod{N}\}$.

Next, by Lemma 8.9, the distributions μ_h agree on all moments, and all these moments are finite. Lemma 8.10 also shows that, for every $j \in \mathbb{N}_0$, $B^{-e(j)} \leq w_j \leq 2^j B^{-e(j)}$, and also that $j - e(j) \leq N-1$ with equality if and only if $j \in \{N-1, N, \ldots, 2N-2\}$. This proves the proposition. □

9 Concluding Remarks

Our results all stem from the same gap between moment information and the local features of a distribution that matter for optimization. Distributions with the same full moment sequence can have quantiles separated by arbitrarily large factors, with consequences for Value-at-Risk, lower-tail average quantiles, chance-constrained capacity, and newsvendor decisions. They can also have disjoint ranges of nearly optimal posted prices. For secretary success probability, the N-way construction concentrates distributions near different first atoms; uniform mixtures of these laws and the Ramsey argument yield the classical finite-horizon value s_m, even for i.i.d. observations. For prophet inequalities, the binary construction creates the order-statistic gap behind the fixed-r barriers. Moments determine expectations of polynomials, but they need not reveal the local tail or threshold behavior on which these decisions depend.

Acknowledgements

A substantial portion of this work was completed by the authors in early 2026. In preparing the first version, we used generative AI tools for limited purposes, primarily language editing and presentation. ChatGPT 5.6 Pro/Ultra also assisted with reviewing the exposition and identifying connections to the classical literature. For the second version, GPT-6 Astra Ultra assisted with mathematical proof development for the i.i.d. extension of the secretary success-probability result.

All results of this paper have been formalized in Lean 4.32.2 using mathlib. ChatGPT 5.6 Pro/Ultra assisted with reviewing the formalization for the first version. For the second version, GPT-6 Astra Ultra helped update it to cover the i.i.d. extension and improve its overall structure. The development is available in the accompanying GitHub repository, which includes a guide linking the paper's results to the corresponding Lean declarations. The authors take full responsibility for the paper's content.

References

[AGK19] Richard Arratia, Larry Goldstein, and Fred Kochman. Size bias for one and all. *Probability Surveys*, 16:1–61, 2019.

[AHM51] Kenneth J Arrow, Theodore Harris, and Jacob Marschak. Optimal inventory policy. *Econometrica: Journal of the Econometric Society*, pages 250–272, 1951.

[BBDS17] Moshe Babaioff, Liad Blumrosen, Shaddin Dughmi, and Yaron Singer. Posting prices with unknown distributions. *ACM Trans. Economics and Comput.*, 5(2):13:1–13:20, 2017.

[CCL+26] José Correa, Andrés Cristi, Vasilis Livanos, Víctor Verdugo, and Jiechen Zhang. On the informativeness of moments in optimal stopping. In *Proceedings of the ACM Symposium on Theory of Computing (STOC)*, pages 777–787, 2026.

[CDFS22] José R. Correa, Paul Dütting, Felix A. Fischer, and Kevin Schewior. Prophet inequalities for independent and identically distributed random variables from an unknown distribution. *Mathematics of Operations Research*, 47(2):1287–1309, 2022.

[Che22] Ethan Che. Robustly optimal auction design under mean constraints. In *Proceedings of the 23rd ACM Conference on Economics and Computation*, pages 153–181, 2022.

[Chi70] T. S. Chihara. A characterization and a class of distribution functions for the Stieltjes–Wigert polynomials. *Canadian Mathematical Bulletin*, 13(4):529–532, 1970.

[CLK+18] Vinicius Carrasco, Vitor Farinha Luz, Nenad Kos, Matthias Messner, Paulo Monteiro, and Humberto Moreira. Optimal selling mechanisms under moment conditions. *Journal of Economic Theory*, 177:245–279, 2018.

[DDN21] Bikramjit Das, Anulekha Dhara, and Karthik Natarajan. On the heavy-tail behavior of the distributionally robust newsvendor. *Operations Research*, 69(4):1077–1099, 2021.

[DFKL20] Paul Dütting, Michal Feldman, Thomas Kesselheim, and Brendan Lucier. Prophet inequalities made easy: Stochastic optimization by pricing non-stochastic inputs. *SIAM Journal on Computing*, 49(3):540–582, 2020.

[DK19] Paul Dütting and Thomas Kesselheim. Posted pricing and prophet inequalities with inaccurate priors. In *Proceedings of the ACM Conference on Economics and Computation*, pages 111–129, 2019.

[DY10] Erick Delage and Yinyu Ye. Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations research*, 58(3):595–612, 2010.

[Edg88] Francis Y Edgeworth. The mathematical theory of banking. *Journal of the Royal Statistical Society*, 51(1):113–127, 1888.

[EHLM20] Hossein Esfandiari, MohammadTaghi Hajiaghayi, Brendan Lucier, and Michael Mitzenmacher. Prophets, secretaries, and maximizing the probability of choosing the best. In *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*, volume 108 of *Proceedings of Machine Learning Research*, pages 3717–3727. PMLR, 2020.

[Fer89] Thomas S. Ferguson. Who solved the secretary problem? *Statistical Science*, 4(3):282–289, 1989.

[FSZ21] Moran Feldman, Ola Svensson, and Rico Zenklusen. Online contention resolution schemes with applications to bayesian selection problems. *SIAM J. Comput.*, 50(2):255–300, 2021.

[GM66] John P. Gilbert and Frederick Mosteller. Recognizing the maximum of a sequence. *Journal of the American Statistical Association*, 61(313):35–73, 1966.

[GPTD22] Yiannis Giannakopoulos, Diogo Poças, and Alexandros Tsigonias-Dimitriadis. Robust revenue maximization under minimal statistical information. *ACM Transactions on Economics and Computation*, 10(4):1–33, 2022.

[GS10] Joel Goh and Melvyn Sim. Distributionally robust optimization and its tractable approximations. *Operations Research*, 58(4-part-1):902–917, 2010.

[Hau21a] Felix Hausdorff. Summationsmethoden und momentfolgen i. *Mathematische Zeitschrift*, 9:74–109, 1921.

[Hau21b] Felix Hausdorff. Summationsmethoden und momentfolgen ii. *Mathematische Zeitschrift*, 9:280–299, 1921.

[HJ24] Jason D. Hartline and Aleck C. Johnsen. Equivocal blends: Prior independent lower bounds. In *ITCS*, volume 287 of *LIPIcs*, pages 59:1–59:21. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2024.

[HKS07] MohammadTaghi Hajiaghayi, Robert Kleinberg, and Tuomas Sandholm. Automated online mechanism design and prophet inequalities. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 58–65, 2007.

[HR14] Jason D. Hartline and Tim Roughgarden. Optimal platform design. *CoRR*, abs/1412.8518, 2014.

[KS78] Ulrich Krengel and Louis Sucheston. On semiamarts, amarts, and processes with finite value. In *Probability on Banach Spaces*, pages 197–266. Dekker, 1978.

[KW12] Robert Kleinberg and Seth Matthew Weinberg. Matroid prophet inequalities. In *Proceedings of the forty-fourth annual ACM Symposium on Theory of Computing*, pages 123–136, 2012.

[Lei82] R. Leipnik. The lognormal distribution and strong nonuniqueness of the moment problem. *Theory of Probability and Its Applications*, 26(4):850–852, 1982.

[LPU15] Retsef Levi, Georgia Perakis, and Joline Uichanco. The data-driven newsvendor problem: New bounds and insights. *Operations Research*, 63(6):1294–1306, 2015.

[LS18] Euiwoong Lee and Sahil Singla. Optimal online contention resolution schemes via ex-ante prophet inequalities. In *26th Annual European Symposium on Algorithms (ESA 2018)*, 2018.

[Mye81] Roger B. Myerson. Optimal auction design. *Math. Oper. Res.*, 6(1):58–73, 1981.

[Nut22] Pranav Nuti. The secretary problem with distributions. In *Integer Programming and Combinatorial Optimization*, volume 13265 of *Lecture Notes in Computer Science*, pages 429–439. Springer, 2022.

[OT18] Sofiya Ostrovska and Mehmet Turan. Discrete stieltjes classes for log-heine type distributions. *arXiv preprint arXiv:1808.03086*, 2018.

[RM19] Hamed Rahimian and Sanjay Mehrotra. Distributionally robust optimization: A review. *arXiv preprint arXiv:1908.05659*, 2019.

[RS17] Aviad Rubinstein and Sahil Singla. Combinatorial prophet inequalities. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1671–1687. SIAM, 2017.

[SAK57] Herbert E Scarf, KJ Arrow, and S Karlin. A min-max solution of an inventory problem. Technical report, Rand Corporation Santa Monica, 1957.

[SC84] Ester Samuel-Cahn. Comparison of threshold stop rules and maximum for independent nonnegative random variables. *The Annals of Probability*, 12(4):1213–1216, 1984.

[Sod19] Sasha Sodin. The classical moment problem, 2019.

[ST43] James Alexander Shohat and Jacob David Tamarkin. *The problem of moments*. American Mathematical Soc., 1943.

[Sti94] T-J Stieltjes. Recherches sur les fractions continues. In *Annales de la Faculté des sciences de Toulouse: Mathématiques*, volume 8, pages J1–J122, 1894.

[WKS14] Wolfram Wiesemann, Daniel Kuhn, and Melvyn Sim. Distributionally robust convex optimization. *Operations research*, 62(6):1358–1376, 2014.