Chebyshev Rollout Schedules for Polynomial Extrapolation under Low-Order Interference

CausalSmith*

August 25, 2026

Abstract

This paper studies how to place measurement rounds in a finite-population rollout experiment when the target is the full-adoption contrast but the rollout budget stops at a treated fraction $q < 1$. Under static rollout consistency, a degree-β polynomial restriction on the rollout mean curve, and a common round-mean variance envelope—restrictions that low-order interference motivates but that we impose rather than derive from a microfounded interference model—unbiased linear estimation reduces to a polynomial extrapolation problem from $[0, q]$ to 1. For a fixed schedule, the worst-case variance over the positive-semidefinite covariance class meeting this diagonal envelope equals the variance scale times the squared ℓ^1 (total-variation) norm of the polynomial-exact weights. The resulting diagonal-envelope amplification criterion, defined over linear unbiased estimators, attains its minimax value up to multiplicative constants at shifted Chebyshev-Lobatto measurement fractions when $q \leq q_{\max} < 1$ and the number of rollout intervals satisfies $k \geq c\beta$ for some $c > 1$. The minimax amplification is of order

$$\left(\frac{(1+\sqrt{1-q})^2}{q}\right)^{2\beta},$$

up to constants depending only on $q_{\max}$ and the oversampling ratio. For the exact nested rollout covariance problem, the same Chebyshev schedule is rate-feasible through the envelope upper bound, and exact optimality under the true rollout covariance structure is posed as a separate covariance-specific design question.

1 Introduction

Rollout experiments often measure outcomes before full adoption. A platform may expose a small treated fraction before scaling a feature, a policy maker may expand access in stages, or a researcher may observe several rollout fractions strictly below the endpoint at which everyone is treated. Under interference, the treated-fraction mean response follows a rollout curve that can extend beyond the convex-combination structure of two no-interference potential outcomes. The full-adoption contrast must then be inferred from the response curve traced out by the rollout.

This paper studies the design problem of where to place the rollout measurement fractions. The estimand is the endpoint contrast $m_P(1) - m_P(0)$, while the experiment observes round

*Machine-generated by the CausalSmith research pipeline. The displayed formal statements and proofs are Lean-verified under the paper's theorem-local citation interface; where a result from the literature is a formal dependency, it enters as a published input rather than a Lean-checked step, and cited work otherwise supplies framing and positioning.

means on the truncated interval $[0, q]$, with $q \leq q_{\max} < 1$. The design variable is the ordered schedule of treated fractions between the untreated baseline and the budget endpoint. The analysis is finite-population and design-based. Assumption 1 gives a static rollout potential-outcome interpretation indexed by contemporaneous assignments; Assumption 2 restricts the rollout mean curve to be a polynomial of degree at most β; and Assumption 3 bounds each round mean variance by a common scale σ_0^2/n.

Under these conditions, the statistical problem becomes a polynomial extrapolation problem. Lemma 1 identifies the endpoint contrast with the sum of the nonconstant coefficients of the rollout mean curve. Definition 3 then characterizes linear weights that are unbiased for every degree-β rollout mean curve by exactness on the monomials $1, u, \ldots, u^{\beta}$. Theorem 1 shows that, when cross-round covariances are left unrestricted except for positive semidefiniteness and the diagonal variance envelope, the worst-case variance bound for a linear unbiased estimator, sharp over the positive-semidefinite covariance class meeting the diagonal envelope, is

$$\frac{\sigma_0^2}{n}\left(\sum_{j=0}^{k} |w_j|\right)^2 .$$

Thus the schedule affects the envelope risk only through the smallest squared ℓ^1 (total-variation, TV) norm of polynomial-exact weights, the amplification criterion $A_{\beta}(p)$ in Definition 5.

The main design result is Theorem 2. For any fixed low-budget cap $q_{\max} \in (0, 1)$, any oversampling ratio $c > 1$, and any $k \geq c\beta$, every admissible schedule has amplification at least a constant times

$$\left(\frac{(1+\sqrt{1-q})^2}{q}\right)^{2\beta},$$

uniformly over $0 < q \leq q_{\max}$. The shifted Chebyshev-Lobatto schedule in Definition 4 attains the same rate up to a constant depending only on c and $q_{\max}$. The result is therefore a rate-minimax statement for the envelope criterion: low-budget extrapolation carries an exponential cost in the polynomial order, and endpoint-clustered Chebyshev-Lobatto placement controls the exponent that is unavoidable under the stated envelope.

The full-budget boundary clarifies what drives this cost. Proposition 2 shows that when $q = 1$, $k = \beta$, and the equal-spacing schedule $p^{\mathrm{eq}}(\beta, 1)$ is used, the endpoint contrast can be estimated by the difference between the final and baseline round means with amplification at most 4, independent of β. The exponential factor in Theorem 2 is therefore specifically the price of estimating an endpoint outside the observed rollout interval.

The exact nested covariance problem is kept separate. Definition 6 leaves inside the minimax problem the covariance matrix that a fully specified rollout law would generate, and Remark 1 poses exact optimality for that narrower problem. The proved one-sided comparison in Lemma 2 bounds the exact finite-population risk by the envelope amplification criterion at every admissible schedule. Consequently, Proposition 3 establishes the displayed low-budget exponential upper bound for the exact rollout risk attained by the Chebyshev-Lobatto schedule.

The paper contributes to the design-based interference literature by making the treated-fraction schedule itself the object of minimax analysis. Classical potential-outcome formulations treat no-interference experiments as the baseline case [Rubin, 1974, Imbens and Rubin, 2015], while social-interaction and interference settings make spillovers part of the estimand and the dependence structure [Manski, 1993, 2013]. Modern work has developed exposure mappings, randomization-based estimators, graph cluster designs, and randomization tests under interference [Aronow and Samii, 2017, Sävje et al., 2017, Eckles et al., 2017, Ugander et al., 2013, Athey et al., 2017, Basse et al., 2019, Puelz et al., 2019]. The closest line of work is the low-order network-interference and rollout literature, where polynomial structure permits identification of global treatment effects from partial exposure or staged adoption [Cortez et al., 2022, Cortez-Rodriguez et al., 2022, 2024, Eichhorn et al., 2024]. The present analysis keeps that polynomial-extrapolation structure and answers a new design question: conditional on a rollout budget, where should measurement fractions be placed to minimize the worst-case diagonal-envelope variance amplification among polynomial-exact linear estimators? Low-order interference motivates the imposed degree-β polynomial mean restriction, while the contribution translates that restriction and the variance envelope into a sharp rollout-schedule design result.

The answer uses classical ideas from optimal design and approximation theory. Polynomial-regression design and equivalence theory provide the language for optimizing measurement locations [Smith, 1918, Kiefer and Wolfowitz, 1959, Kiefer, 1974, Pukelsheim, 1993], and Chebyshev systems explain why endpoint-spread nodes control polynomial extrapolation [Karlin and Studden, 1966]. Here those approximation facts enter through the dual representation in Lemma 3: minimizing the ℓ^1 norm of unbiased weights is equivalent to controlling the largest endpoint contrast among degree-β polynomials bounded at the schedule nodes. The paper translates the classical Chebyshev endpoint-extrapolation phenomenon and minimum-ℓ^1 optimal-recovery duality into a finite-population rollout measurement-schedule criterion under the round-mean variance envelope. To our knowledge, the resulting amplification criterion and Chebyshev-Lobatto rollout schedule are new design objects in that literature. The appendix records the formal verification scope for the paper's formal statements and proofs.

2 Setup and Assumptions

We work with a finite-population rollout design. The population consists of a finite unit set U_n with n units. The experimenter chooses a sequence of rollout measurement rounds indexed by $j = 0, \ldots, k$, and the contemporaneous assignment vector at round j is Z_j. Let Ω denote the finite assignment space generated by the rollout randomization, and let π be the probability mass function on Ω. For any design-measurable random variable X, $\mathbb{E}_\pi X$, $\mathrm{Var}_\pi(X)$, and $\mathrm{Cov}_\pi(X_1, X_2)$ denote the corresponding finite-sum expectation, variance, and covariance under π. The specific assignment law—how treatment is nested across rounds and how each target fraction is implemented—is not further specified here, because every result uses only the round-mean variance envelope of Assumption 3. The exact cross-round covariance a fully specified rollout law would induce is not part of the verified formal layer; it enters only through the symbolic benchmark R_{exact} of the discussion, which we treat as a placeholder for future covariance-restricted design problems rather than as a fully specified finite-population object. Throughout, the schedule entries p_j are target treated fractions and the formal results treat them as real design points in $[0, 1]$; in a finite population these targets may require rounding to feasible treatment counts. The formal results treat the realized measurement fractions as the schedule nodes, so unbiasedness holds exactly at those realized fractions; any gap between a target fraction and the realized fraction is a finite-sample perturbation of the moment conditions in Definition 3, outside the formal scope.

The potential-outcome notation is finite-population notation: $Y_i(Z_j)$ is unit i's outcome under the assignment vector realized at measurement j, and $\bar{Y}_j$ is the observed finite-population round mean. The static rollout condition fixes the interpretation of these objects.

Assumption 1. [ass:static-rollout-consistency] (Static Rollout Consistency) *At each rollout measurement j, the observed round mean is*

$$\bar{Y}_j = n^{-1} \sum_{i \in U_n} Y_i(Z_j).$$

The potential outcomes $Y_i(Z_j)$ depend on the contemporaneous assignment vector Z_j, but not on earlier rollout steps.

Assumption 1 is the static rollout potential-outcomes condition with no carryover. It permits arbitrary contemporaneous interference through the assignment vector while ruling out lagged treatment histories as separate arguments of the potential outcome, as in static rollout formulations for interference designs [Cortez-Rodriguez et al., 2024].

A schedule is an ordered vector p of treated fractions. The rollout budget q is the final treated fraction, so the design observes the response curve only on $[0, q]$, even though the target contrast compares $u = 0$ with $u = 1$.

Definition 1. [def:budgeted-schedule-class] (Budgeted schedule class $S_{k,q}$) *The budgeted schedule class is*

$$S_{k,q} = \left\{ p \in [0,1]^{k+1} : p_0 = 0 \text{ and } 0 < p_1 < \cdots < p_k = q \right\}.$$

Definition 1 makes the design variable explicit. The econometrician controls the placement of the rollout measurement fractions between the untreated baseline and the budget q; the strict inequalities guarantee distinct interior interpolation nodes for the polynomial problem below.

The mean response curve under a law P is $m_P(u)$. The low-order interference restriction used here is that the design mean along the rollout path is a polynomial of degree at most β.

Assumption 2. [ass:beta-order-polynomial] (Polynomial Rollout Mean) *For every law P, the rollout mean curve satisfies*

$$m_P(u) = \sum_{\ell=0}^{\beta} a_{P,\ell} u^{\ell}$$

for all $u \in [0, 1]$.

Assumption 2 is the beta-order rollout polynomial identity condition. It summarizes low-order interference along the treated-fraction path by requiring the average response to be exhausted by powers $u^0, \ldots, u^{\beta}$, a restriction used in polynomial exposure and rollout representations of interference [Cortez et al., 2022]. This form arises, for instance, when each unit's outcome depends on the assignment vector only through interactions of order at most β among units, and the rollout assigns treatment so that, at target fraction u, the probability that any fixed set of $\ell \leq \beta$ units is jointly treated is a polynomial in u of degree ℓ (as under independent Bernoulli assignment at rate u; fixed-count or rounded rollout schemes satisfy this only approximately, since their joint-treatment probabilities are not exactly u^{ℓ}); averaging the bounded-order response over that assignment then yields a degree-β polynomial mean curve, with β the maximum interaction order. Absent such bounded-order structure the result should be read as an optimal-extrapolation theorem for an imposed degree-β mean curve rather than a generic interference-design guarantee; because the target $u = 1$ lies outside the observed interval, misspecifying β biases the extrapolated endpoint contrast, and this bias is amplified precisely in the low-budget regime. The endpoint contrast studied below is $\tau_P = m_P(1) - m_P(0)$.

The next restriction is a sharp variance envelope that isolates the effect of treated-fraction placement from additional covariance structure.

Assumption 3. [`ass:round-mean-variance-envelope`] (Round-Mean Variance Envelope) *For every law P, every admissible schedule p, and every rollout measurement j,*

$$\mathrm{Var}_\pi(\bar{Y}_j) \le \frac{\sigma_0^2}{n}.$$

Assumption 3 is specific to this analysis. It bounds each round mean by the common scale σ_0^2/n while leaving the cross-round covariance signs and magnitudes otherwise unrestricted subject to positive semidefiniteness. This is the envelope under which the worst-case variance reduces to an amplification problem for the estimator weights. The scale σ_0^2 absorbs both the outcome magnitudes and the design's cross-unit dependence. Boundedness of outcomes alone does not control it: under common or strongly dependent randomization a finite-population round mean can have variance of order 1 rather than $1/n$. The envelope holds with σ_0^2 a fixed constant when outcomes are bounded *and* the design induces only weak (for example, bounded-neighborhood) cross-unit dependence, as under independent or locally dependent Bernoulli assignment; under stronger dependence the round-mean variance decays more slowly in n, and the results then apply with σ_0^2 read as the relevant (possibly n-dependent) envelope scale, with the amplification bounds interpreted at that scaling.

For the low-budget results, the budget remains uniformly below full rollout.

Assumption 4. [`ass:low-budget-cap`] (Low-Budget Cap) *The rollout budget and its cap satisfy*

$$0 < q_{\max} \quad \textit{and} \quad q \le q_{\max} < 1.$$

Assumption 4 is also specific to this analysis. The cap $q_{\max}$ keeps the problem in the extrapolation regime: the target endpoint $u = 1$ lies outside the observed interval $[0, q]$. The boundary case $q = 1$ is therefore separated from the low-budget rate calculations. Every result invoking this assumption also imposes $q > 0$; when $q = 0$ the budgeted schedule class $S_{k,q}$ is empty, so the operative low-budget regime is $0 < q \le q_{\max} < 1$.

The admissible law class collects the preceding restrictions and pins the design expectation of each round mean to the polynomial curve at the corresponding schedule point.

Definition 2. [`def:rollout-law-class`] (Rollout law class $\mathcal{P}_\beta$) *The rollout law class is*

$$\mathcal{P}_\beta = \left\{ P : \begin{array}{l} P \textit{ satisfies Assumption 1, Assumption 2, and Assumption 3,} \\ \textit{and } \mathbb{E}_\pi[\bar{Y}_j] = m_P(p_j) \textit{ for every } j = 0, \ldots, k \end{array} \right\}.$$

Thus $\mathcal{P}_\beta$ is a finite-population law class indexed by the common polynomial order and variance scale, with the schedule entering through the evaluation points p_j. The condition $\mathbb{E}_\pi[\bar{Y}_j] = m_P(p_j)$ is the bridge between the randomization distribution and the deterministic polynomial approximation problem.

We estimate the endpoint contrast with a linear rollout estimator

$$\hat{\tau}_{w,p} = \sum_{j=0}^{k} w_j \bar{Y}_j,$$

where w is a vector of weights applied to the observed round means. Unbiasedness over all degree-β rollout mean curves is equivalent to exactness on the monomial basis.

Definition 3. [`def:unbiased-weight-set`] (Unbiased weight set $W_\beta(p)$) *For a schedule $p = (p_0, \ldots, p_k)$, the set of unbiased weights of order β is*

$$W_\beta(p) = \left\{ w \in \mathbb{R}^{k+1} : \sum_{j=0}^{k} w_j p_j^0 = 0 \text{ and } \sum_{j=0}^{k} w_j p_j^\ell = 1 \text{ for every } \ell = 1, \ldots, \beta \right\}.$$

Here p_j^0 denotes the constant monomial 1 for every j, including $p_0 = 0$. The first equation in Definition 3 removes the intercept $m_P(0)$. The remaining equations reproduce the contribution of each nonconstant monomial to $m_P(1) - m_P(0)$, so any $w \in W_\beta(p)$ gives a design-unbiased linear estimator for every law in Definition 2, whenever the set is nonempty.

The schedules compared in the main theorem are shifted Chebyshev-Lobatto schedules. They place more measurements near the endpoints of the observed interval, the usual geometry for controlling polynomial extrapolation.

Definition 4. [`def:chebyshev-schedule`] (Chebyshev schedule $p^{\mathrm{Ch}}(k,q)$) *For $j = 0, \ldots, k$, the shifted Chebyshev-Lobatto node on $[0,q]$ is*

$$p_j^{\mathrm{Ch}}(k,q) = \frac{q}{2}\left\{1 - \cos\left(\frac{\pi j}{k}\right)\right\}.$$

The shifted Chebyshev-Lobatto schedule is

$$p^{\mathrm{Ch}}(k,q) = \left(p_0^{\mathrm{Ch}}(k,q), \ldots, p_k^{\mathrm{Ch}}(k,q)\right).$$

We write p^{Ch} for this schedule when k and q are clear. For comparison with a more elementary design, the equally spaced $\beta + 1$-node schedule is $p^{\mathrm{eq}}(\beta, q)$, with entries $p_j^{\mathrm{eq}}(\beta, q) = qj/\beta$ for $j = 0, \ldots, \beta$.

The variance envelope makes the relevant design criterion the smallest possible ℓ^1 amplification among unbiased weights.

Definition 5. [`def:amplification-criterion`] (Amplification criterion $A_\beta(p)$) *For a schedule p, the amplification criterion is*

$$A_\beta(p) = \inf_{w \in W_\beta(p)} \left(\sum_{j=0}^{k} |w_j|\right)^2.$$

The minimax amplification over budgeted schedules is

$$M_{\beta,k,q} = \inf_{p \in S_{k,q}} A_\beta(p).$$

The quantity $A_\beta(p)$ records the best worst-case variance multiplier available at a fixed schedule under the total-variation variance envelope. The minimax quantity $M_{\beta,k,q}$ then optimizes that multiplier over the admissible treated-fraction placements in Definition 1.

Finally, it is useful to distinguish the envelope criterion from the exact finite-population covariance problem. For a law P and schedule p, let $\Gamma_P(p)$ be the design covariance matrix of the vector $(\bar{Y}_0, \ldots, \bar{Y}_k)$.

Definition 6. [`def:exact-nested-risk`] (Exact nested risk $R_{\mathrm{exact}}(\beta, k, q)$) *The exact nested risk is*

$$R_{\mathrm{exact}}(\beta, k, q) = \inf_{p \in S_{k,q}} \inf_{w \in W_\beta(p)} \sup_{P \in \mathcal{P}_\beta} w' \Gamma_P(p) w,$$

where $\Gamma_P(p)$ is the design covariance matrix of the rollout round means under law P and schedule p.

Definition 6 leaves the true nested-rollout covariance matrix inside the minimax problem. The exact risk R_{exact} is therefore the finite-population benchmark, while Definition 5 is the sharp envelope problem solved in the main results.

3 Main Results

The setup in Definition 2 reduces the identification problem to a polynomial endpoint problem. The first result records the corresponding identity for the target contrast: under the static rollout interpretation and the degree-β rollout mean restriction, the estimand is the sum of the nonconstant coefficients of the rollout mean curve.

Lemma 1. [`prop:rollout-polynomial-identity`] (Rollout Polynomial Identity) *For any finite assignment space Ω, population size n, number of rollout intervals k, polynomial order β, contemporaneous assignment vectors Z_j, potential outcomes $Y_i(Z_j)$, round means $\bar{Y}_j$, rollout mean curve m_P, and polynomial coefficients $(a_{P,\ell})_{\ell \geq 0}$, suppose Assumption 1 and Assumption 2 hold. Then*

$$m_P(1) - m_P(0) = \sum_{\ell=1}^{\beta} a_{P,\ell}.$$

Thus the endpoint treatment contrast satisfies

$$\tau_P = m_P(1) - m_P(0) = \sum_{\ell=1}^{\beta} a_{P,\ell}.$$

The proof is deferred to Section B.

Lemma 1 identifies the polynomial target: any design-unbiased estimator of the nonconstant polynomial coefficients is also unbiased for the endpoint contrast. The subsequent results add sampling variability and optimize the placement of rollout measurements.

The next theorem turns this identity into a design criterion under the variance envelope in Assumption 3. The point is that, once cross-round covariances are left unrestricted except for positive semidefiniteness and the diagonal envelope, the worst-case variance of a linear unbiased estimator is governed by the squared ℓ^1 norm of its weights.

Theorem 1. [`thm:tv-envelope-design`] (TV Envelope Design Bound) *Let n, k, β be integers, let $q \in \mathbb{R}$, and let $\sigma_0^2 \in \mathbb{R}$. Fix a finite assignment space Ω, a finite design π on Ω, rollout assignment vectors $Z_0, \dots, Z_k$, round means $\bar{Y}_0, \dots, \bar{Y}_k$, a rollout mean curve m_P, polynomial coefficients $a_{P,\ell}$, and a schedule $p = (p_0, \dots, p_k)$. Suppose that*

- ***(Order.)*** $1 \leq \beta \leq k$.
- ***(Envelope scale.)*** $\sigma_0^2 \geq 0$.

- ***(Schedule.)*** *$p \in S_{k,q}$, in the sense of Definition 1.*
- ***(Law class.)*** *The design, rollout variables, round means, mean curve, coefficients, envelope scale, and schedule satisfy the rollout law-class restrictions $\mathcal{P}_\beta$ in Definition 2.*

Then $W_\beta(p)$, as defined in Definition 3, is nonempty. Moreover, for every $w \in W_\beta(p)$,

$$\mathbb{E}_\pi\left[\sum_{j=0}^{k} w_j \bar{Y}_j\right] = m_P(1) - m_P(0)$$

and

$$\mathrm{Var}_\pi\left(\sum_{j=0}^{k} w_j \bar{Y}_j\right) \le \frac{\sigma_0^2}{n}\left(\sum_{j=0}^{k} |w_j|\right)^2.$$

For every such w, there also exists a positive semidefinite $(k+1)\times(k+1)$ matrix Γ such that

$$\Gamma_{jj} \le \frac{\sigma_0^2}{n} \quad \text{for every } j = 0, \ldots, k,$$

and

$$\sum_{i=0}^{k}\sum_{j=0}^{k} w_i \Gamma_{ij} w_j = \frac{\sigma_0^2}{n}\left(\sum_{j=0}^{k} |w_j|\right)^2.$$

Consequently, with $A_\beta(\cdot)$ and $M_{\beta,k,q}$ as in Definition 5,

$$\inf\left\{ v : \exists w \in W_\beta(p) \text{ with } v = \frac{\sigma_0^2}{n}\left(\sum_{j=0}^{k} |w_j|\right)^2 \right\} = \frac{\sigma_0^2}{n} A_\beta(p),$$

and

$$\inf\left\{ v : \exists p' \in S_{k,q} \text{ with } v = \frac{\sigma_0^2}{n} A_\beta(p') \right\} = \frac{\sigma_0^2}{n} M_{\beta,k,q}.$$

Finally, if $p^\star \in S_{k,q}$ satisfies

$$A_\beta(p^\star) = M_{\beta,k,q},$$

then for every $p' \in S_{k,q}$,

$$\frac{\sigma_0^2}{n} A_\beta(p^\star) \le \frac{\sigma_0^2}{n} A_\beta(p').$$

Theorem 1 separates three roles that are otherwise easy to conflate. The law class determines which linear combinations are unbiased; the envelope scale contributes the common factor σ_0^2/n; and the schedule affects risk only through the feasible set of polynomial-exact weights. Thus the schedule problem is the minimization of $A_\beta(p)$ over $S_{k,q}$, a finite-dimensional optimal-design problem in the tradition of polynomial extrapolation and minimax design [Smith, 1918, Kiefer and Wolfowitz, 1959, Kiefer, 1974, Pukelsheim, 1993, Karlin and Studden, 1966].

A simple benchmark is obtained by using exactly $\beta + 1$ equally spaced measurements on $[0, q]$. This benchmark is useful because it shows that unbiased extrapolation is feasible without oversampling. It also gives a (possibly large) universal upper bound on the amplification of naive spacing when q is small; because Proposition 1 is only an upper bound, it does not by itself certify that equal spacing is unstable or suboptimal in a matching lower-bound sense.

Proposition 1. [`prop:equal-spacing-benchmark`] (Equal-Spacing Amplification Bound) *For every integer $\beta \geq 1$ and every $q \in (0, 1]$, let*

$$p^{\mathrm{eq}}(\beta, q)_j = \frac{qj}{\beta}, \qquad j = 0, \dots, \beta.$$

Then $p^{\mathrm{eq}}(\beta, q)_j \in [0, 1]$ for every $j = 0, \dots, \beta$, and, with A_β as in Definition 5,

$$A_\beta(p^{\mathrm{eq}}(\beta, q)) \leq 9 \left(\frac{\beta}{q}\right)^{2\beta}.$$

Proposition 1 is not a minimax statement. It provides a scale against which the Chebyshev result below should be read: equal spacing gives a universal upper bound, but the base of this upper bound grows with β/q, consistent with the difficulty of extrapolating from $[0, q]$ to 1 using evenly spaced nodes (the bound is one-sided and does not by itself establish a matching instability).

The full-budget boundary is qualitatively different. When $q = 1$, the target endpoint is observed, so no extrapolation is needed.

Proposition 2. [`prop:no-extrapolation-boundary`] (Full-Budget Endpoint Rule) *For every polynomial order $\beta \geq 1$, take $k = \beta$ and use the equal-spacing schedule $p^{\mathrm{eq}}(\beta, 1)$. Define the weight vector $w = (w_0, \dots, w_\beta)$ by*

$$w_0 = -1, \qquad w_\beta = 1, \qquad w_j = 0 \quad \text{for } 1 \leq j \leq \beta - 1.$$

Then $w \in W_\beta(p^{\mathrm{eq}}(\beta, 1))$, where $W_\beta(\cdot)$ is the unbiased weight set in Definition 3. Consequently, for the amplification criterion in Definition 5,

$$A_\beta(p^{\mathrm{eq}}(\beta, 1)) \leq 4.$$

Proposition 2 explains why the low-budget assumption is imposed separately in Assumption 4. At full rollout, the endpoint contrast can be estimated by the difference between the final and baseline round means, with amplification bounded independently of β. The exponential behavior in the next theorem is therefore specific to endpoint extrapolation.

Chebyshev-Lobatto placement is the schedule that controls this extrapolation cost. The theorem below compares every admissible schedule to the shifted Chebyshev-Lobatto schedule from Definition 4, uniformly over budgets bounded away from one. The oversampling ratio c permits more than the minimum $\beta + 1$ measurements, which is the regime in which the discrete Lobatto grid norms the relevant continuous polynomial problem.

Theorem 2. [`thm:chebyshev-minimax`] (Chebyshev Minimax Amplification) *Fix a budget cap $q_{\max} \in (0, 1)$. There exists a constant $C_-(q_{\max}) > 0$ such that, for every oversampling ratio $c > 1$, there exists a constant $C_+(c, q_{\max}) > 0$ with the following property.*

For every pair of integers β, k, every rollout budget q, and every schedule p, suppose that

- ***(Order and oversampling.)*** *$\beta \geq 1$ and $k \geq c\beta$.*
- ***(Low-budget regime.)*** *$q > 0$ and the low-budget cap condition of Assumption 4 holds for $(q, q_{\max})$.*
- ***(Admissible schedule.)*** *When p is used below, $p \in S_{k,q}$ as in Definition 1.*

Then, with $A_\beta(\cdot)$ and $M_{\beta,k,q}$ as in Definition 5 and $p^{\mathrm{Ch}}(k,q)$ as in Definition 4,

$$A_\beta(p) \geq C_-(q_{\max}) \left(\frac{(1+\sqrt{1-q})^2}{q} \right)^{2\beta},$$

and

$$A_\beta\Big(p^{\mathrm{Ch}}(k,q)\Big) \leq C_+(c, q_{\max}) \left(\frac{(1+\sqrt{1-q})^2}{q} \right)^{2\beta}.$$

Moreover, the minimax amplification satisfies the matching two-sided bound

$$C_-(q_{\max}) \left(\frac{(1+\sqrt{1-q})^2}{q} \right)^{2\beta} \leq M_{\beta,k,q} \leq C_+(c, q_{\max}) \left(\frac{(1+\sqrt{1-q})^2}{q} \right)^{2\beta}.$$

Theorem 2 gives the main schedule result. For fixed $q \leq q_{\max} < 1$ and $k \geq c\beta$, no admissible schedule can improve the amplification exponent beyond the displayed Chebyshev base, while the shifted Chebyshev-Lobatto schedule attains the same rate up to constants depending only on $q_{\max}$ and c. Under the envelope criterion, the rate-optimal schedule clusters measurements near 0 and q, as in Chebyshev-Lobatto extrapolation; heuristically the experimenter concentrates measurements near the interval endpoints, where controlling the polynomial extrapolation to $u = 1$ is most demanding. The low-budget cost remains exponential in the polynomial order, but Chebyshev-Lobatto placement controls the exponent that the total-variation variance envelope makes unavoidable.

To make the exponential base concrete, Table 1 reports the low-budget Chebyshev base $(1+\sqrt{1-q})^2/q$ alongside the equal-spacing *upper-bound* base β/q implied by Proposition 1, for several budgets q. The Chebyshev base is at most $4/q$ and is independent of β, whereas the equal-spacing upper-bound base grows linearly in β (Proposition 1 bounds equal-spacing amplification from above only, so this column is an upper-bound base, not a certified rate); the amplification is the corresponding base raised to the power 2β, and the constants $C_-(q_{\max}), C_+(c, q_{\max})$ of Theorem 2 depend only on $q_{\max}$ and c.

q	Chebyshev base $(1+\sqrt{1-q})^2/q$	bound $4/q$	equal-spacing upper-bound base β/q
0.05	78.0	80.0	20β
0.1	38.0	40.0	10β
0.2	17.9	20.0	5β
0.5	5.83	8.0	2β

Table 1: Low-budget amplification bases (the amplification is the base raised to the power 2β). The Chebyshev base is β-independent and at most $4/q$; the equal-spacing base scales with β/q.

4 Discussion and Extensions

The envelope result in Theorem 1 is deliberately conservative about the covariance structure of a nested rollout. It permits any positive semidefinite covariance matrix whose diagonal entries satisfy the round-mean variance envelope, and it therefore turns the schedule problem into a polynomial amplification problem. The exact finite-population problem in Definition 6 is narrower: the covariance matrix is the true design covariance matrix $\Gamma_P(p)$ induced by the rollout law that generates the observed round means. The specific covariance restrictions this law imposes beyond the diagonal variance envelope are not encoded in the verified formal layer; accordingly R_{exact} is treated here as a symbolic benchmark, and the only verified comparison between the exact and envelope problems is the one-sided variance-envelope inequality of Lemma 2. That extra structure may matter for finite-sample optimality, even though the envelope criterion already gives an interpretable design rule that is sharp over the positive-semidefinite diagonal-envelope covariance class.

4.1 Exact-covariance scope and limitations

The distinction is important for the status of Chebyshev-Lobatto placement. Theorem 2 proves rate-minimax optimality up to constants for the total-variation envelope criterion $A_\beta(p)$ and $M_{\beta,k,q}$. Exact optimality with $\Gamma_P(p)$ left inside the minimax risk is the stronger covariance-specific question recorded separately below.

Remark 1. [`oeq:exact-nested-minimax`] (Exact nested optimality) *Fix a finite assignment space Ω with rollout randomization π, population size n, and variance-envelope scale σ_0^2. Fix a budget cap $q_{\max}$ with $0 < q_{\max} < 1$, let $c > 1$, let $\beta \geq 1$, set $k = \lceil c\beta \rceil$, and take q with $0 < q \leq q_{\max}$. Under these hypotheses the shifted Chebyshev-Lobatto schedule $p^{\text{Ch}}(k, q)$ is admissible, so that*

$$p^{\text{Ch}}(k, q) \in S_{k,q}.$$

For the exact finite-population nested-rollout minimax problem

$$R_{\text{exact}}(\beta, k, q) = \inf_{p \in S_{k,q}} \inf_{w \in W_\beta(p)} \sup_{P \in \mathcal{P}_\beta} w' \Gamma_P(p) w,$$

where $\Gamma_P(p)$ is the true design covariance matrix of the monotone Bernoulli rollout and $\mathcal{P}_\beta$ includes the static-rollout, β-polynomial, variance-envelope, and mean-curve-pinning restrictions, the open question is whether the fixed Chebyshev-Lobatto exact risk attains the nested infimum:

$$\inf_{w \in W_\beta(p^{\text{Ch}}(k,q))} \sup_{P \in \mathcal{P}_\beta} w' \Gamma_P(p^{\text{Ch}}(k, q)) w = R_{\text{exact}}(\beta, k, q).$$

This statement is only the exact-optimality question for $p^{\text{Ch}}(k, q)$; it does not assert the separate rate-feasibility bound.

Remark 1 keeps exact optimality separate from rate feasibility. The equality displayed there characterizes the optimizer of the true nested covariance problem and is stronger than achieving the same exponential order as the envelope benchmark. Establishing such an equality calls for the covariance restrictions implied by monotone Bernoulli rollout designs in addition to the diagonal variance envelope. Related finite-population interference designs suggest that these restrictions can be substantive, especially when exposure mappings, clustered assignments, or dependency neighborhoods impose additional structure on covariances [Cortez-Rodriguez et al., 2024, Eichhorn et al., 2024, Jiang and Wang, 2023, Cai et al., 2023].

The results below state the rate implication delivered by the one-sided envelope comparison. The exact nested risk is bounded above by the envelope amplification criterion at every admissible schedule. Therefore any schedule that is good for the envelope problem is automatically feasible for the exact finite-population risk.

Lemma 2. [`lem:exact-risk-envelope-upper`] (Exact Risk Envelope Bound) *Let Ω be a finite assignment space equipped with rollout randomization π. For any population size n, any integers k and β, any real numbers q and σ_0^2, and any schedule $p = (p_0, \ldots, p_k)$, suppose that*

- ***(Order.)*** $\beta \geq 1$.
- ***(Number of rollout intervals.)*** $k \geq \beta$.
- ***(Budget.)*** $0 < q \leq 1$.
- ***(Variance scale.)*** $\sigma_0^2 \geq 0$.
- ***(Budgeted schedule.)*** $p \in S_{k,q}$, *as in Definition 1.*

Then

$$\inf_{w \in W_\beta(p)} \sup_{P \in \mathcal{P}_\beta} w' \Gamma_P(p) w \leq \frac{\sigma_0^2}{n} A_\beta(p),$$

with $W_\beta(p)$ as in Definition 3, $\mathcal{P}_\beta$ as in Definition 2, and $A_\beta(p)$ as in Definition 5. Consequently,

$$R_{\text{exact}}(\beta, k, q) \leq \frac{\sigma_0^2}{n} M_{\beta,k,q},$$

where $R_{\text{exact}}(\beta, k, q)$ is as in Definition 6 and $M_{\beta,k,q}$ is as in Definition 5.

The proof is deferred to Section B.

Lemma 2 is a one-sided comparison. It places the exact covariance minimax risk below the envelope risk because the envelope allows a larger class of covariance matrices than those generated by the rollout law. A matching lower bound and characterization of the exact-risk minimizing schedule constitute the remaining exact-covariance problem.

Combining this comparison with the Chebyshev upper bound gives the finite-population feasibility statement relevant for low-budget rollout design. With $k = \lceil c\beta \rceil$, the shifted Chebyshev-Lobatto schedule attains the same exponential rate for the exact risk upper bound as it does for the envelope criterion.

Proposition 3. [lem:exact-chebyshev-rate-feasible] (Feasible Chebyshev Rate) *Fix an oversampling constant c and a budget cap $q_{\max}$ such that $c > 1$ and $0 < q_{\max} < 1$. Then there exists a constant $C_+(c, q_{\max}) > 0$ such that, for every population size n, polynomial order β, number of rollout intervals k, finite assignment space Ω, finite design π on Ω, rollout budget q, and variance-envelope scale σ_0^2, the following conditions imply the stated bound:*

- ***(Order.)*** $\beta \geq 1$.
- ***(Budget.)*** $q > 0$ *and q satisfies the low-budget cap with $q_{\max}$ in Assumption 4.*
- ***(Variance scale.)*** $\sigma_0^2 \geq 0$.
- ***(Chebyshev size.)*** $k = \lceil c\beta \rceil$.
- ***(Schedule.)*** $p = p^{\text{Ch}}(k, q)$ *is the shifted Chebyshev-Lobatto schedule of Definition 4.*

Under these conditions,

$$\inf_{w \in W_\beta(p^{\text{Ch}}(k,q))} \sup_{P \in \mathcal{P}_\beta} \operatorname{Var}_\pi \left(\sum_{j=0}^{k} w_j \bar{Y}_j \right) \leq \frac{\sigma_0^2}{n} C_+(c, q_{\max}) \left(\frac{(1 + \sqrt{1-q})^2}{q} \right)^{2\beta},$$

where $W_\beta(p^{\text{Ch}}(k,q))$ is the unbiased weight set of Definition 3, and the supremum ranges over rollout laws $P \in \mathcal{P}_\beta$ satisfying Definition 2 at the schedule $p^{\text{Ch}}(k,q)$ with variance-envelope scale σ_0^2.

Proposition 3 gives a practical interpretation of the main theorem for the original rollout experiment. Independently of exact nested optimality, the Chebyshev-Lobatto schedule has infimum worst-case exact variance over unbiased linear estimators bounded by the displayed low-budget exponential rate. Thus the envelope solution supplies a feasible exact-risk design with controlled finite-population risk.

4.2 Future work

Several extensions are natural. Clustered rollouts would replace unit-level monotone Bernoulli assignment with group-level assignment and would require tracking how cluster size and within-cluster exposure affect the round-mean covariance matrix. Mixed designs that combine rollout fractions with saturation, encouragement, or exposure-based randomization would add further design variables beyond the treated-fraction schedule, connecting the present extrapolation problem to recent work on clustered and network experiments [Chen and Li, 2025, Fan et al., 2025, Bhadra and Schweinberger, 2025, Schröder et al., 2026]. Adjacent staggered-rollout efficiency work outside network interference is also relevant for positioning the rollout-design question [Roth and Sant'Anna, 2021]. The Chebyshev argument suggests that endpoint placement should remain central whenever the estimand extrapolates beyond the observed rollout range, with a design-specific minimax object for each extension.

Appendices

A Proofs and Verification

This appendix collects the auxiliary statements that support the envelope design result and the low-budget Chebyshev rate. The first part translates the rollout-weight problem into a real-polynomial optimal-recovery problem; the second part records the Chebyshev approximation inequalities; the final part returns these inequalities to the admissible rollout schedules and to the variance envelope. The appendix states the mathematical ingredients under the modeling assumptions introduced in the main text.

The bridge from rollout weights to approximation is the duality between minimum ℓ^1 norm weights and bounded polynomial extrapolation. In the notation of Definition 3, the moment equations make each admissible weight vector a linear functional on degree-β polynomials. The dual norm of this functional is exactly the largest endpoint contrast among polynomials bounded at the chosen schedule nodes.

Lemma 3. [`lem:amplification-dual-norm`] (Dual Norm Amplification) *Let $\beta, k \in \mathbb{N}$ and let $p = (p_0, \dots, p_k)$ be a real-valued schedule whose nodes are distinct. Suppose that the unbiased weight set $W_\beta(p)$ in Definition 3 is nonempty. Then*

$$\inf_{w \in W_\beta(p)} \sum_{j=0}^{k} |w_j| = \sup \left\{ |r(1) - r(0)| : r \in \mathbb{R}[x],\ \deg(r) \le \beta,\ \max_{0 \le j \le k} |r(p_j)| \le 1 \right\}.$$

Consequently, for the amplification criterion $A_\beta(p)$ in Definition 5,

$$A_\beta(p) = \left(\sup \left\{ |r(1) - r(0)| : r \in \mathbb{R}[x],\ \deg(r) \le \beta,\ \max_{0 \le j \le k} |r(p_j)| \le 1 \right\} \right)^2.$$

Proof of Lemma 3. Throughout write

$$\mathcal{N} = \left\{ \sum_{j=0}^{k} |w_j| : w \in W_\beta(p) \right\}, \qquad \mathcal{A} = \left\{ \left(\sum_{j=0}^{k} |w_j| \right)^2 : w \in W_\beta(p) \right\},$$

and

$$D = \sup \left\{ |r(1) - r(0)| : r \in \mathbb{R}[x],\ \deg(r) \le \beta,\ \max_{0 \le j \le k} |r(p_j)| \le 1 \right\}.$$

Call a polynomial r *admissible* if $\deg(r) \le \beta$ and $\max_{0 \le j \le k} |r(p_j)| \le 1$, so that D is the supremum of $|r(1) - r(0)|$ over admissible r.

Step 1: the moment form of unbiasedness. Rewrite the unbiasedness equations in the single moment form

$$\sum_{j=0}^{k} w_j p_j^\ell = \begin{cases} 0, & \ell = 0, \\ 1, & 1 \le \ell \le \beta, \end{cases} \qquad 0 \le \ell \le \beta.$$

This is equivalent to $w \in W_\beta(p)$: the case $\ell = 0$ is the zeroth-moment equation in Definition 3, while each positive $\ell \le \beta$ is one of its remaining moment equations. Consequently $\mathcal{N}$ is exactly the set of attainable ℓ^1 norms for this moment system, and the assumed nonemptiness of $W_\beta(p)$ gives primal feasibility: $\mathcal{N} \ne \emptyset$ and $\mathcal{A} \ne \emptyset$.

Step 2: representation identity. For every $w \in W_\beta(p)$ and every real polynomial r with $\deg(r) \le \beta$,

$$r(1) - r(0) = \sum_{j=0}^{k} w_j\, r(p_j).$$

Indeed, writing $r(x) = \sum_{\ell=0}^{\beta} c_\ell x^\ell$ and exchanging the two finite sums,

$$\sum_{j=0}^{k} w_j\, r(p_j) = \sum_{\ell=0}^{\beta} c_\ell \sum_{j=0}^{k} w_j p_j^\ell = \sum_{\ell=1}^{\beta} c_\ell = r(1) - r(0),$$

where the middle equality is the moment form of Step 1 and the last uses $r(1) = \sum_{\ell=0}^{\beta} c_\ell$ and $r(0) = c_0$.

Step 3: weak duality. Let $w \in W_\beta(p)$ and let r be admissible. By Step 2 and the triangle inequality,

$$|r(1) - r(0)| = \left| \sum_{j=0}^{k} w_j\, r(p_j) \right| \le \sum_{j=0}^{k} |w_j|\, |r(p_j)| \le \sum_{j=0}^{k} |w_j|,$$

the last step because $|r(p_j)| \le 1$ for every node. Thus every value $|r(1) - r(0)|$ is bounded above by every element of $\mathcal{N}$. Since $\mathcal{N} \ne \emptyset$, the set of dual values is bounded above, so D is a finite real number, and

$$D \le \inf \mathcal{N}.$$

Moreover the zero polynomial is admissible with $|r(1) - r(0)| = 0$, so 0 belongs to the dual value set and

$$D \ge 0.$$

Step 4: strong duality. There exists $w^\star \in W_\beta(p)$ with

$$\sum_{j=0}^{k} |w_j^\star| \le D.$$

Indeed, by Step 2 the endpoint value $r(1) - r(0)$ depends on r only through the node-value vector $(r(p_j))_{j=0}^k$, so it defines a linear functional on the subspace of $\mathbb{R}^{k+1}$ spanned by such vectors; by the definition of D that functional is dominated there by the sublinear majorant $x \mapsto D \max_{0\le j\le k} |x_j|$. Extending it to all of $\mathbb{R}^{k+1}$ under the same majorant, its coordinate values $w_j^\star$ again satisfy the moment equations of Step 1, hence $w^\star \in W_\beta(p)$; and evaluating the extension at the sign vector $\varepsilon_j = \operatorname{sign}(w_j^\star)$ gives $\sum_j |w_j^\star| \le D \max_j |\varepsilon_j| = D$. In particular $\inf \mathcal{N} \le D$, which combined with Step 3 gives the first assertion,

$$\inf_{w\in W_\beta(p)} \sum_{j=0}^{k} |w_j| = \sup\left\{ |r(1) - r(0)| : \deg(r) \le \beta,\ \max_{0\le j\le k} |r(p_j)| \le 1 \right\}.$$

Step 5: the squared identity. The set $\mathcal{A}$ is nonempty by Step 1 and bounded below by 0, so $\inf \mathcal{A}$ is a well-defined real number. For every $w \in W_\beta(p)$ we have $0 \le D \le \sum_j |w_j|$ by Steps 3 and 4, hence $D^2 \le \left(\sum_j |w_j|\right)^2$; taking the infimum over $\mathcal{A}$ gives

$$D^2 \le \inf \mathcal{A}.$$

Conversely, the weight $w^\star$ of Step 4 satisfies $0 \le \sum_j |w_j^\star| \le D$, and therefore

$$\inf \mathcal{A} \le \left(\sum_{j=0}^{k} |w_j^\star|\right)^2 \le D^2.$$

Thus $\inf \mathcal{A} = D^2$. Since Definition 5 defines $A_\beta(p)$ as precisely this infimum, the claimed squared formula follows. □

Lemma 3 is the point at which the econometric design problem becomes the real polynomial optimal-recovery problem. The schedule affects the right-hand side only through the locations at which the polynomial is constrained. Hence minimax placement of rollout fractions is equivalent, under the envelope criterion, to choosing nodes that make bounded interpolation on $[0, q]$ as stable as possible for the endpoint contrast $r(1) - r(0)$.

The approximation bounds used for that comparison are standard Chebyshev facts adapted to the extrapolation geometry. The first is the exterior extremal property: among degree-β polynomials bounded by one on $[-1, 1]$, the Chebyshev polynomial controls growth outside the interval [Karlin and Studden, 1966, Pukelsheim, 1993].

Lemma 4. [`lem:chebyshev-exterior-extremal`] (Exterior Chebyshev Extremal Bound) *Let β be a positive integer. For every real polynomial P and every exterior point x_0, suppose that*

- ***(Degree bound.)*** $\deg(P) \le \beta$.
- ***(Unit interval bound.)*** *For every* $x \in [-1, 1]$, $|P(x)| \le 1$.
- ***(Exterior point.)*** $x_0 > 1$.

Then

$$|P(x_0)| \le T_\beta(x_0),$$

where T_β is the Chebyshev polynomial of the first kind of degree β.

Proof of Lemma 4. Let $s = \{0, \dots, \beta\}$, and let v_i, $i \in s$, be the Chebyshev extremal nodes

$$v_i = \cos \frac{\pi i}{\beta}, \qquad i = 0, \dots, \beta.$$

The node map $i \mapsto v_i$ is injective on s, so the interpolation nodes are distinct. For $i = 0, \dots, \beta$, let L_i be the Lagrange basis polynomial for these nodes,

$$L_i(x) = \prod_{\substack{0 \le j \le \beta \\ j \ne i}} \frac{x - v_j}{v_i - v_j}.$$

Since $|s| = \beta + 1$, any polynomial Q with $\deg(Q) \le \beta$ has degree strictly less than the number of nodes, and the same is true for T_β. Hence both polynomials are recovered from their node values by Lagrange interpolation.

First prove the one-sided claim: if $\deg(Q) \le \beta$ and $|Q(x)| \le 1$ for every $x \in [-1, 1]$, then

$$Q(x_0) \le T_\beta(x_0).$$

Interpolating Q at the nodes and evaluating at x_0 gives

$$Q(x_0) = \sum_{i=0}^{\beta} Q(v_i) L_i(x_0).$$

The same interpolation argument for T_β gives

$$T_\beta(x_0) = \sum_{i=0}^{\beta} T_\beta(v_i) L_i(x_0).$$

For $x_0 > 1$, the signed exterior coefficients satisfy

$$(-1)^i L_i(x_0) \ge 0, \qquad i = 0, \dots, \beta.$$

Indeed, every numerator factor obeys $x_0 - v_j > 0$, because each Chebyshev node lies in $[-1, 1]$ and $x_0 > 1$. For the denominator, the Chebyshev-node monotonicity used in this sign calculation gives $v_i < v_j$ when $j < i$ and $v_j < v_i$ when $i < j$. Thus the denominator has exactly i negative factors, and multiplying by $(-1)^i$ makes the denominator sign positive. Combining these signs in the product formula for L_i yields the displayed inequality.

Since $v_i \in [-1, 1]$, the interval bound on Q gives $|Q(v_i)| \le 1$, and multiplying by the sign $(-1)^i$ preserves this absolute-value bound. Therefore

$$(-1)^i Q(v_i) \le 1.$$

The Chebyshev extremal-node identity gives

$$T_\beta(v_i) = (-1)^i, \qquad i = 0, \dots, \beta.$$

Inserting the factor $(-1)^i(-1)^i = 1$ and multiplying the last inequality by the nonnegative coefficient $(-1)^i L_i(x_0)$ yields, for each i,

$$Q(v_i)L_i(x_0) = \big((-1)^i Q(v_i)\big)\big((-1)^i L_i(x_0)\big) \le (-1)^i L_i(x_0) = T_\beta(v_i)L_i(x_0).$$

Summing over $i = 0, \ldots, \beta$ and using the two interpolation displays proves the one-sided claim.

Apply this claim first to $Q = P$, obtaining $P(x_0) \le T_\beta(x_0)$. Apply it again to $Q = -P$; the degree bound and interval bound are unchanged, so $-P(x_0) \le T_\beta(x_0)$. The two inequalities give

$$-T_\beta(x_0) \le P(x_0) \le T_\beta(x_0),$$

and hence

$$|P(x_0)| \le T_\beta(x_0).$$

□

For the rollout application, the polynomial is controlled at the finite set of measurement nodes. Oversampled Chebyshev-Lobatto grids pass from this nodewise control to uniform control on the interval. The next statement records the required discrete norming step, conditional on the Ehlich–Zeller mesh inequality.

Lemma 5. [`lem:oversampled-chebyshev-lobatto-norming`] (Oversampled Lobatto Norming) *Suppose the Ehlich–Zeller Chebyshev–Lobatto mesh inequality holds: for every pair of nonnegative integers β, k and every real polynomial R with $\deg(R) \le \beta$ and $\beta < k$,*

$$|R(x)| \le \frac{1}{\cos(\pi\beta/(2k))} \max_{0\le j\le k} \left| R\left(-\cos\left(\frac{\pi j}{k}\right)\right)\right|$$

for every $x \in [-1, 1]$. Then, for every real oversampling ratio c with $c > 1$, there exists a real constant $K(c) > 0$ such that the following holds. For every pair of nonnegative integers β, k and every real polynomial R satisfying

- ***(Positive order.)*** $\beta \ge 1$,
- ***(Oversampling.)*** $k \ge c\beta$,
- ***(Degree bound.)*** $\deg(R) \le \beta$,

one has, for every $x \in [-1, 1]$,

$$|R(x)| \le K(c) \max_{0\le j\le k} \left| R\left(-\cos\left(\frac{\pi j}{k}\right)\right)\right|.$$

Proof of Lemma 5. Fix $c > 1$, and set

$$K(c) = \frac{1}{\cos(\pi/(2c))}.$$

Since $c > 1$, the angle $\pi/(2c)$ lies strictly between 0 and $\pi/2$, so $\cos(\pi/(2c)) > 0$ and $K(c) > 0$. Now let $\beta \ge 1$, $k \ge c\beta$, and $\deg(R) \le \beta$, and abbreviate

$$N = \max_{0\le j\le k} \left| R\left(-\cos\left(\frac{\pi j}{k}\right)\right)\right|.$$

Since $c > 1$ and $\beta \geq 1 > 0$, we have

$$0 < \beta < c\beta \leq k,$$

so in particular $\beta < k$ and $k > 0$. The assumed Ehlich–Zeller mesh inequality therefore applies and gives, for every $x \in [-1, 1]$,

$$|R(x)| \leq \frac{1}{\cos(\pi\beta/(2k))} N.$$

It remains only to bound the secant factor uniformly. From $k \geq c\beta$, $k > 0$ and $c > 0$,

$$\frac{\beta}{k} \leq \frac{1}{c}, \qquad \text{hence} \qquad 0 \leq \frac{\pi\beta}{2k} = \frac{\pi}{2} \cdot \frac{\beta}{k} \leq \frac{\pi}{2c}.$$

Also $\pi/(2c) < \pi/2 \leq \pi$, so both angles lie in $[0, \pi]$. Since cosine is decreasing on $[0, \pi]$, this implies

$$\cos(\pi/(2c)) \leq \cos(\pi\beta/(2k)),$$

and positivity of $\cos(\pi/(2c))$ gives

$$\frac{1}{\cos(\pi\beta/(2k))} \leq \frac{1}{\cos(\pi/(2c))} = K(c).$$

The quantity N is nonnegative, being a maximum of absolute values, so multiplying the last inequality by N and chaining with the Ehlich–Zeller bound gives, for every $x \in [-1, 1]$,

$$|R(x)| \leq \frac{1}{\cos(\pi\beta/(2k))} N \leq K(c)\, N = K(c) \max_{0 \leq j \leq k} \left| R\left(-\cos\left(\frac{\pi j}{k} \right) \right) \right|,$$

which is the claim. □

Lemma 5 explains why the main theorem imposes $k \geq c\beta$ with $c > 1$. Oversampling prevents a degree-β polynomial from taking large values between adjacent Lobatto nodes while remaining small at the nodes. The constant $K(c)$ is allowed to depend on the oversampling ratio, but not on β, k, or the rollout budget.

The affine map from $[0, q]$ to $[-1, 1]$ sends the target endpoint $u = 1$ to the exterior point $x_q = 2/q - 1$. The corresponding Chebyshev multiplier $\lambda(q) = x_q + \sqrt{x_q^2 - 1}$ gives the low-budget exponential base. Uniformity over $q \leq q_{\max} < 1$ is recorded next.

Lemma 6. [`lem:continuous-chebyshev-endpoint-bound`] (Continuous Chebyshev Endpoint Bound) *Fix $q_{\max}$ with $0 < q_{\max} < 1$. Then there exist constants $C(q_{\max}) > 0$ and $c(q_{\max}) > 0$ such that, for every $\beta \in \mathbb{N}$ and every $q \in \mathbb{R}$, if*

- ***(Degree.)*** $1 \leq \beta$,
- ***(Budget.)*** $0 < q \leq q_{\max}$,
- ***(Exterior notation.)*** $x_q = 2/q - 1$ *and* $\lambda(q) = x_q + \sqrt{x_q^2 - 1}$,

then both of the following hold:

for every real polynomial R with $\deg(R) \le \beta$ *and* $\sup_{x\in[-1,1]} |R(x)| \le 1, \qquad |R(x_q) - R(-1)| \le C(q_{\max})\,\lambda(q)^\beta,$

and

$$T_\beta(x_q) - 1 \ge c(q_{\max})\,\lambda(q)^\beta.$$

Proof of Lemma 6. Put

$$\bar{x} = \frac{2}{q_{\max}} - 1, \qquad \bar{\lambda} = \bar{x} + \sqrt{\bar{x}^2 - 1}.$$

Since $0 < q_{\max} < 1$ we have $2/q_{\max} > 2$, hence $\bar{x} > 1$; and since $\sqrt{\bar{x}^2 - 1} \ge 0$ this gives $\bar{\lambda} \ge \bar{x} > 1$. Choose

$$C(q_{\max}) = 2, \qquad c(q_{\max}) = \frac{1}{2}(1 - \bar{\lambda}^{-1})^2.$$

Both are positive: $\bar{\lambda} > 1$ gives $0 < \bar{\lambda}^{-1} < 1$, so $1 - \bar{\lambda}^{-1} > 0$.

Now let $\beta \ge 1$ and $0 < q \le q_{\max}$, and write

$$x_q = \frac{2}{q} - 1, \qquad \lambda(q) = x_q + \sqrt{x_q^2 - 1}.$$

Then $q \le q_{\max} < 1$, so exactly as above $x_q > 1$ and $\lambda(q) > 1$. Moreover $q \le q_{\max}$ gives $2/q_{\max} \le 2/q$ and therefore

$$1 \le \bar{x} \le x_q,$$

and on $[1, \infty)$ the map $t \mapsto t + \sqrt{t^2 - 1}$ is nondecreasing, since it equals $e^{\operatorname{arcosh} t}$ and both arcosh and the exponential are nondecreasing. Hence

$$\bar{\lambda} \le \lambda(q).$$

Chebyshev evaluation at the exterior point. Write $x_q = \cosh a$ with $a = \operatorname{arcosh}(x_q) \ge 0$, which is legitimate because $x_q \ge 1$. Then $T_\beta(\cosh a) = \cosh(\beta a)$, while $e^a = x_q + \sqrt{x_q^2 - 1} = \lambda(q)$, so

$$T_\beta(x_q) = \cosh(\beta a) = \frac{e^{\beta a} + e^{-\beta a}}{2} = \frac{\lambda(q)^\beta + \lambda(q)^{-\beta}}{2}.$$

Upper bound. Let R be a real polynomial with $\deg(R) \le \beta$ and $|R(x)| \le 1$ for $x \in [-1, 1]$. By Lemma 4, applied at the exterior point $x_q > 1$,

$$|R(x_q)| \le T_\beta(x_q).$$

Also $-1 \in [-1, 1]$, so $|R(-1)| \le 1$, and therefore

$$|R(x_q) - R(-1)| \le |R(x_q)| + |R(-1)| \le T_\beta(x_q) + 1.$$

Since $\lambda(q) \ge 1$ and $\beta \ge 1$ we have $\lambda(q)^{-\beta} \le 1 \le \lambda(q)^\beta$, whence

$$\frac{\lambda(q)^\beta + \lambda(q)^{-\beta}}{2} + 1 \le \frac{\lambda(q)^\beta + \lambda(q)^\beta}{2} + \lambda(q)^\beta = 2\lambda(q)^\beta.$$

Combining the last three displays proves

$$|R(x_q) - R(-1)| \le 2\lambda(q)^\beta = C(q_{\max})\lambda(q)^\beta.$$

Lower bound. The same evaluation formula and the algebraic identity

$$\frac{\lambda(q)^{\beta}+\lambda(q)^{-\beta}}{2}-1=\frac{1}{2}\left(1-\lambda(q)^{-\beta}\right)^{2}\lambda(q)^{\beta}$$

(valid because $\lambda(q)^{\beta}>0$; expand the right-hand side) reduce the claim to a monotonicity step. Since $1<\bar{\lambda}\le\lambda(q)$ and $\beta\ge 1$,

$$\bar{\lambda}\le\lambda(q)\le\lambda(q)^{\beta},\qquad\text{hence}\qquad\lambda(q)^{-\beta}\le\bar{\lambda}^{-1},$$

so that $0\le 1-\bar{\lambda}^{-1}\le 1-\lambda(q)^{-\beta}$ and therefore $(1-\bar{\lambda}^{-1})^{2}\le\left(1-\lambda(q)^{-\beta}\right)^{2}$. Multiplying by $\frac{1}{2}\lambda(q)^{\beta}>0$ gives

$$T_{\beta}(x_q)-1=\frac{\lambda(q)^{\beta}+\lambda(q)^{-\beta}}{2}-1\ge\frac{1}{2}(1-\bar{\lambda}^{-1})^{2}\lambda(q)^{\beta}=c(q_{\max})\lambda(q)^{\beta}.$$

This proves both assertions. □

Lemma 6 supplies the continuous upper and lower endpoint calculations behind Theorem 2. The first inequality controls the extrapolation contrast for every bounded polynomial; the second shows that the Chebyshev polynomial itself has the corresponding order of growth. After rescaling from $[-1,1]$ back to $[0,q]$, this is the source of the factor $((1+\sqrt{1-q})^{2}/q)^{2\beta}$ in the amplification criterion.

The remaining deterministic ingredients connect the approximation statements to the rollout schedule class. First, the shifted Chebyshev-Lobatto nodes used in Definition 4 are admissible budgeted schedules.

Lemma 7. [`lem:chebyshev-schedule-admissible`] (Chebyshev Schedule Admissibility) *For every integer $k\ge 1$ and every $q\in(0,1]$, the shifted Chebyshev–Lobatto schedule $p^{\mathrm{Ch}}(k,q)$ of Definition 4 is an element of the budgeted schedule class $S_{k,q}$ of Definition 1. Equivalently, its coordinates all lie in $[0,1]$, it starts at 0, is strictly increasing across the rollout measurement indices, and ends at q.*

Proof of Lemma 7. Let $k\ge 1$ and $0<q\le 1$, so that $k>0$ and the nodes

$$p_j^{\mathrm{Ch}}(k,q)=\frac{q}{2}\left(1-\cos\frac{\pi j}{k}\right),\qquad j=0,\dots,k,$$

are well defined. By Definition 1 it suffices to check four things: that every coordinate lies in $[0,1]$, that the schedule starts at 0, that it is strictly increasing in j, and that it ends at q.

Range. The inequalities $-1\le\cos(\pi j/k)\le 1$ give $0\le 1-\cos(\pi j/k)\le 2$, and multiplying by $q/2>0$ yields

$$0\le p_j^{\mathrm{Ch}}(k,q)=\frac{q}{2}\left(1-\cos\frac{\pi j}{k}\right)\le\frac{q}{2}\cdot 2=q\le 1.$$

Hence $p_j^{\mathrm{Ch}}(k,q)\in[0,1]$ for every $j=0,\dots,k$.

Endpoints. At $j=0$, $\cos 0=1$, so

$$p_0^{\mathrm{Ch}}(k,q)=\frac{q}{2}(1-1)=0;$$

at $j = k$, $\pi k/k = \pi$ and $\cos \pi = -1$, so

$$p_k^{\mathrm{Ch}}(k, q) = \frac{q}{2}\big(1 - (-1)\big) = q.$$

Strict increase. If $0 \le a < b \le k$, then, since $k > 0$,

$$0 \le \frac{\pi a}{k} < \frac{\pi b}{k} \le \pi.$$

Cosine is strictly decreasing on $[0, \pi]$, hence

$$\cos \frac{\pi b}{k} < \cos \frac{\pi a}{k},$$

so that

$$1 - \cos \frac{\pi a}{k} < 1 - \cos \frac{\pi b}{k},$$

and multiplication by the positive number $q/2$ gives

$$p_a^{\mathrm{Ch}}(k, q) < p_b^{\mathrm{Ch}}(k, q).$$

The coordinates are therefore in $[0, 1]$, start at 0, are strictly increasing, and end at q, so $p^{\mathrm{Ch}}(k, q) \in S_{k,q}$. □

The admissibility statement matters because Theorem 2 compares the Chebyshev schedule to all elements of $S_{k,q}$. It verifies that endpoint clustering at 0 and q is compatible with the strict monotonicity imposed by Definition 1.

A separate feasibility issue is the existence of polynomially unbiased weights. Distinct budgeted nodes provide enough interpolation information once the number of rollout intervals is at least the polynomial order.

Lemma 8. [`lem:unbiased-weight-set-nonempty`] (Nonempty Unbiased Weights) *Let $q \in \mathbb{R}$, let β and k be integers, and let $p = (p_0, \ldots, p_k)$ be a rollout schedule. Suppose that*

- ***(Order.)*** *$\beta \ge 1$.*
- ***(Enough nodes.)*** *$k \ge \beta$.*
- ***(Budgeted schedule.)*** *$p \in S_{k,q}$, with $S_{k,q}$ as in Definition 1.*

Then the unbiased weight set $W_\beta(p)$ of Definition 3 is nonempty. Equivalently, there exists a weight vector $w = (w_0, \ldots, w_k)$ such that $w \in W_\beta(p)$.

Proof of Lemma 8. Because $p \in S_{k,q}$, the entries $p_0 < p_1 < \cdots < p_k$ are strictly increasing in Definition 1, hence the $k+1$ nodes are pairwise distinct. Let L_j be the Lagrange basis polynomial for the full collection of nodes,

$$L_j(t) = \prod_{\substack{0 \le i \le k \\ i \ne j}} \frac{t - p_i}{p_j - p_i}, \qquad j = 0, \ldots, k,$$

and define the endpoint weights

$$w_j = L_j(1) - L_j(0), \qquad j = 0, \ldots, k.$$

Fix $0 \le \ell \le \beta$ and let $e_\ell(t) = t^\ell$, a polynomial of degree $\ell \le \beta \le k$, hence of degree strictly less than the number $k+1$ of nodes. Lagrange interpolation on the $k+1$ distinct nodes therefore reproduces it exactly,

$$e_\ell(t) = \sum_{j=0}^{k} e_\ell(p_j)\, L_j(t) = \sum_{j=0}^{k} p_j^\ell\, L_j(t).$$

Evaluating at $t = 1$ and at $t = 0$ gives

$$1^\ell = \sum_{j=0}^{k} p_j^\ell\, L_j(1), \qquad 0^\ell = \sum_{j=0}^{k} p_j^\ell\, L_j(0),$$

and subtracting the second from the first yields

$$\sum_{j=0}^{k} w_j p_j^\ell = 1^\ell - 0^\ell.$$

For $\ell = 0$ this is $1 - 1 = 0$, while for every $1 \le \ell \le \beta$ it is $1 - 0 = 1$. These are exactly the moment equations of Definition 3 defining $W_\beta(p)$. Thus the vector $w = (w_0, \ldots, w_k)$ belongs to $W_\beta(p)$, and the set is nonempty. □

Lemma 8 justifies the infima over $W_\beta(p)$ in Definition 5 for the schedules used in the main results. Econometrically, $k \ge \beta$ gives at least $\beta + 1$ distinct measurements, which is the finite-dimensional requirement for exactness on the monomial basis $1, u, \ldots, u^\beta$.

The last auxiliary statement is the variance-envelope calculation. It establishes the squared ℓ^1 norm as the sharp worst-case envelope over positive semidefinite covariance matrices with the stated diagonal bound.

Lemma 9. [`lem:variance-envelope-sharpness`] (Sharp Variance Envelope) *For any population size n, final rollout index k, finite assignment space Ω with rollout randomization π, rollout-round statistics $X_0, \ldots, X_k$, variance scale σ_0^2, and weight vector $w = (w_0, \ldots, w_k)$, suppose that*

- ***(Nonnegative scale.)*** $\sigma_0^2 \ge 0$.
- ***(Round variance envelope.)*** *Assumption 3 holds for $X_0, \ldots, X_k$ with scale σ_0^2/n.*

Then

$$\mathrm{Var}_\pi\left(\sum_{j=0}^{k} w_j X_j\right) \le \frac{\sigma_0^2}{n}\left(\sum_{j=0}^{k} |w_j|\right)^2.$$

Moreover, there exists a positive semidefinite matrix $\Gamma \in \mathbb{R}^{(k+1)\times(k+1)}$ such that

$$\Gamma_{jj} \le \frac{\sigma_0^2}{n} \quad \textit{for every } j = 0, \ldots, k,$$

and

$$\sum_{i=0}^{k}\sum_{j=0}^{k} w_i\, \Gamma_{ij}\, w_j = \frac{\sigma_0^2}{n}\left(\sum_{j=0}^{k} |w_j|\right)^2.$$

Proof of Lemma 9. Write

$$s = \frac{\sigma_0^2}{(n:\mathbb{R})},$$

interpreting the quotient as the total real division used in the statement; in particular the quotient is defined also when $n = 0$. Since $\sigma_0^2 \geq 0$ and the real cast of a natural number is nonnegative, this gives $s \geq 0$.

Covariance bound. For real random variables U, V on Ω, the design covariance is the finite sum

$$\operatorname{Cov}_\pi(U,V) = \sum_{z\in\Omega} \pi(z)\big(U(z) - \mathbb{E}_\pi U\big)\big(V(z) - \mathbb{E}_\pi V\big).$$

Applying Cauchy–Schwarz to the vectors with entries $\sqrt{\pi(z)}(U(z) - \mathbb{E}_\pi U)$ and $\sqrt{\pi(z)}(V(z) - \mathbb{E}_\pi V)$ gives

$$\operatorname{Cov}_\pi(U,V) \leq \sqrt{\operatorname{Var}_\pi(U)}\sqrt{\operatorname{Var}_\pi(V)}.$$

Applying the same bound to $-U$ and using $\operatorname{Var}_\pi(-U) = \operatorname{Var}_\pi(U)$ gives

$$|\operatorname{Cov}_\pi(U,V)| \leq \sqrt{\operatorname{Var}_\pi(U)}\sqrt{\operatorname{Var}_\pi(V)}.$$

For each i, j, Assumption 3 gives $\operatorname{Var}_\pi(X_i) \leq s$ and $\operatorname{Var}_\pi(X_j) \leq s$. Hence

$$|\operatorname{Cov}_\pi(X_i,X_j)| \leq \sqrt{\operatorname{Var}_\pi(X_i)}\sqrt{\operatorname{Var}_\pi(X_j)} \leq \sqrt{s}\sqrt{s} = s,$$

where the last equality uses $s \geq 0$.

Variance bound. By bilinearity of the design covariance, the variance of the weighted sum expands as

$$\operatorname{Var}_\pi\left(\sum_{j=0}^{k} w_j X_j\right) = \sum_{i=0}^{k}\sum_{j=0}^{k} w_i w_j \operatorname{Cov}_\pi(X_i,X_j).$$

Termwise,

$$w_i w_j \operatorname{Cov}_\pi(X_i,X_j) \leq |w_i w_j \operatorname{Cov}_\pi(X_i,X_j)| = |w_i|\,|w_j|\,|\operatorname{Cov}_\pi(X_i,X_j)| \leq |w_i|\,|w_j|\,s.$$

Summing and factoring the double sum therefore gives

$$\operatorname{Var}_\pi\left(\sum_{j=0}^{k} w_j X_j\right) \leq \sum_{i=0}^{k}\sum_{j=0}^{k} |w_i|\,|w_j|\,s = s\left(\sum_{j=0}^{k} |w_j|\right)^2.$$

Substituting $s = \sigma_0^2/(n:\mathbb{R})$ gives the stated variance bound.

Sharpness. Define

$$\varepsilon_j = \begin{cases} 1, & w_j \geq 0, \\ -1, & w_j < 0, \end{cases} \qquad j = 0,\dots,k,$$

and set

$$\Gamma_{ij} = s\,\varepsilon_i\varepsilon_j.$$

This matrix is positive semidefinite, since it is s times the rank-one outer product $\varepsilon\varepsilon'$ and $s \geq 0$. Equivalently, for every $v \in \mathbb{R}^{k+1}$,

$$\sum_{i=0}^{k}\sum_{j=0}^{k} v_i \Gamma_{ij} v_j = s\left(\sum_{j=0}^{k} v_j\varepsilon_j\right)^2 \geq 0.$$

Since $\varepsilon_j^2 = 1$, each diagonal entry satisfies

$$\Gamma_{jj} = s\varepsilon_j^2 = s = \frac{\sigma_0^2}{(n:\mathbb{R})},$$

and hence obeys the required diagonal bound. Finally, the definition of ε_j gives $w_j\varepsilon_j = |w_j|$ for every j, so

$$\sum_{i=0}^{k}\sum_{j=0}^{k} w_i\Gamma_{ij}w_j = s\left(\sum_{j=0}^{k} w_j\varepsilon_j\right)^2 = s\left(\sum_{j=0}^{k} |w_j|\right)^2.$$

Substituting $s = \sigma_0^2/(n:\mathbb{R})$ gives the desired positive semidefinite matrix and equality. $\square$

Lemma 9 completes the connection between approximation and variance. The upper bound follows from the diagonal variance envelope and positive semidefiniteness; the sharpness clause establishes $A_\beta(p)$ as the smallest uniform envelope criterion when cross-round covariances are otherwise unrestricted. Theorem 1 therefore optimizes this sharp criterion directly.

Verification note. The theorem statements and proofs for the formal results displayed in this paper are machine-checked in Lean 4. The checked development covers the finite-design variance algebra over a finite assignment space, the ℓ^1/ℓ^∞ dual amplification calculation, the positive-semidefinite variance-envelope sharpness argument, and the real-polynomial approximation objects used in the Chebyshev and Lobatto bounds. The potential-outcome consistency restriction in Assumption 1, the polynomial rollout restriction in Assumption 2, the law-class pinning in Definition 2, and the round-mean variance envelope in Assumption 3 enter as hypotheses. For modularity, Lemma 5 is displayed conditionally on the Ehlich–Zeller Chebyshev–Lobatto mesh inequality (Ehlich and Zeller, *Schwankung von Polynomen zwischen Gitterpunkten*, Math. Z. 86 (1964), 41–44), which appears as an explicit hypothesis of that lemma. The accompanying Lean development additionally proves the Ehlich–Zeller inequality through a Bernstein–Szegő differential-inequality argument, thereby discharging this classical approximation-theory input. A machine-readable crosswalk from each displayed statement to its checked Lean declaration accompanies the paper, together with the Lean and mathlib versions and the build instructions needed to reproduce the verification.

B Proofs of the main results

Proof of Lemma 1. The static consistency hypothesis of Assumption 1 is part of the rollout environment and fixes the interpretation of $\bar{Y}_j$, but the endpoint calculation uses only the polynomial mean assumption of Assumption 2.

Both endpoints $u = 1$ and $u = 0$ belong to $[0,1]$, so the degree-β expansion applies at each. Evaluating at $u = 1$ gives

$$m_P(1) = \sum_{\ell=0}^{\beta} a_{P,\ell}\,1^\ell = \sum_{\ell=0}^{\beta} a_{P,\ell},$$

because $1^\ell = 1$ for every ℓ. Evaluating the same expansion at $u = 0$ gives

$$m_P(0) = \sum_{\ell=0}^{\beta} a_{P,\ell}\,0^\ell = a_{P,0},$$

since $0^\ell = 0$ for every $\ell \geq 1$, while the $\ell = 0$ term is $a_{P,0} \cdot 0^0 = a_{P,0}$. Finally, the finite sum over $\{0, \dots, \beta\}$ splits into its constant term and the positive-degree terms:

$$\sum_{\ell=0}^{\beta} a_{P,\ell} = a_{P,0} + \sum_{\ell=1}^{\beta} a_{P,\ell}.$$

This finite-sum identity is obtained by induction on β, adjoining the top endpoint $\ell = \beta + 1$ to both sides at each step. Substituting the two endpoint evaluations into this split and cancelling the constant term $a_{P,0}$ yields

$$m_P(1) - m_P(0) = \left(a_{P,0} + \sum_{\ell=1}^{\beta} a_{P,\ell} \right) - a_{P,0} = \sum_{\ell=1}^{\beta} a_{P,\ell}.$$

□

Proof of Theorem 1. Let

$$\kappa := \frac{\sigma_0^2}{n},$$

with the convention that this real quotient is total, so the quotient is zero when the natural population size is zero. Since $\sigma_0^2 \geq 0$ and $n \geq 0$, this gives $\kappa \geq 0$.

We use this nonnegativity through the following elementary scaling rule for infima, applied twice below. Let $S \subseteq \mathbb{R}$ be nonempty and bounded below, and let $\kappa \geq 0$. Then $\kappa S = \{\kappa u : u \in S\}$ is again nonempty and bounded below, and

$$\inf\{\kappa\, u : u \in S\} = \kappa \inf S.$$

For $\kappa > 0$ this is the standard homogeneity of the infimum; for $\kappa = 0$ both sides equal 0. Each application below is preceded by a verification that the set being scaled is nonempty and bounded below.

1. First, $W_\beta(p)$ is nonempty. Indeed, the order hypothesis gives $\beta \geq 1$, the theorem assumes $k \geq \beta$, and $p \in S_{k,q}$. Hence Lemma 8 applies and gives a weight vector $w \in W_\beta(p)$.

2. Fix now any $w \in W_\beta(p)$. By linearity of the design expectation and by the defining mean-curve restriction $\mathbb{E}_\pi[\bar{Y}_j] = m_P(p_j)$ in $\mathcal{P}_\beta$ of Definition 2,

$$\mathbb{E}_\pi\left[\sum_{j=0}^{k} w_j \bar{Y}_j\right] = \sum_{j=0}^{k} w_j\, \mathbb{E}_\pi[\bar{Y}_j] = \sum_{j=0}^{k} w_j\, m_P(p_j).$$

   Since $p \in S_{k,q}$, every node p_j lies in $[0, 1]$, so the degree-β polynomial condition in $\mathcal{P}_\beta$ gives

$$m_P(p_j) = \sum_{\ell=0}^{\beta} a_{P,\ell} p_j^\ell.$$

   Substituting and interchanging the two finite sums,

$$\sum_{j=0}^{k} w_j m_P(p_j) = \sum_{\ell=0}^{\beta} a_{P,\ell} \sum_{j=0}^{k} w_j p_j^\ell.$$

Splitting off the constant term and inserting the moment equations in Definition 3,

$$\sum_{j=0}^{k} w_j p_j^0 = 0, \qquad \sum_{j=0}^{k} w_j p_j^\ell = 1 \quad (1 \le \ell \le \beta),$$

the last display equals

$$a_{P,0} \sum_{j=0}^{k} w_j p_j^0 + \sum_{\ell=1}^{\beta} a_{P,\ell} \sum_{j=0}^{k} w_j p_j^\ell = a_{P,0} \cdot 0 + \sum_{\ell=1}^{\beta} a_{P,\ell} \cdot 1 = \sum_{\ell=1}^{\beta} a_{P,\ell}.$$

By Lemma 1, whose hypotheses are the static-rollout and β-polynomial members of $\mathcal{P}_\beta$, this is $m_P(1) - m_P(0)$, proving the unbiasedness identity.

3. For the variance statement, apply Lemma 9 with $X_j = \bar{Y}_j$. Its hypotheses are the nonnegative-scale assumption $\sigma_0^2 \ge 0$ and the round-mean variance-envelope component of $\mathcal{P}_\beta$. Therefore

$$\mathrm{Var}_\pi\left(\sum_{j=0}^{k} w_j \bar{Y}_j\right) \le \frac{\sigma_0^2}{n}\left(\sum_{j=0}^{k} |w_j|\right)^2,$$

and the same lemma supplies a positive semidefinite matrix Γ with

$$\Gamma_{jj} \le \frac{\sigma_0^2}{n} \quad \text{for all } j, \qquad \sum_{i=0}^{k}\sum_{j=0}^{k} w_i \Gamma_{ij} w_j = \frac{\sigma_0^2}{n}\left(\sum_{j=0}^{k} |w_j|\right)^2.$$

4. For the fixed schedule p, the set whose infimum appears on the left-hand side of the first envelope identity is exactly the scalar multiple

$$\left\{ v : \ \exists w \in W_\beta(p), \ v = \kappa \left(\sum_{j=0}^{k} |w_j|\right)^2 \right\} = \kappa \cdot \left\{ u : \ \exists w \in W_\beta(p), \ u = \left(\sum_{j=0}^{k} |w_j|\right)^2 \right\}.$$

The unscaled set is nonempty by the first step and is bounded below by 0, since its elements are squares. Hence the usual nonnegative-scalar infimum rule applies to this set, and Definition 5 identifies its infimum as $A_\beta(p)$. With $\kappa \ge 0$,

$$\inf\left\{ v : \ \exists w \in W_\beta(p), \ v = \frac{\sigma_0^2}{n}\left(\sum_{j=0}^{k} |w_j|\right)^2 \right\} = \frac{\sigma_0^2}{n} A_\beta(p).$$

5. The same scalar-infimum step applies after also varying the schedule. Namely,

$$\left\{ v : \ \exists p' \in S_{k,q}, \ v = \kappa\, A_\beta(p') \right\} = \kappa \cdot \left\{ u : \ \exists p' \in S_{k,q}, \ u = A_\beta(p') \right\}.$$

This unscaled schedule set is nonempty because the fixed schedule p belongs to $S_{k,q}$. It is bounded below by 0: for every schedule r, $A_\beta(r) \ge 0$, because $A_\beta(r)$ is the real infimum of squared total-variation norms, with the same convention covering the empty feasible-weight case. Applying the nonnegative-scalar infimum rule and then Definition 5,

$$\inf\left\{ v : \ \exists p' \in S_{k,q}, \ v = \frac{\sigma_0^2}{n} A_\beta(p') \right\} = \frac{\sigma_0^2}{n} M_{\beta,k,q}.$$

6. Finally, suppose $p^\star \in S_{k,q}$ and

$$A_\beta(p^\star) = M_{\beta,k,q}.$$

For any $p' \in S_{k,q}$, the value $A_\beta(p')$ belongs to the set

$$\{u : \exists r \in S_{k,q},\ u = A_\beta(r)\}.$$

The same uniform nonnegativity $A_\beta(r) \geq 0$ for all schedules r makes this set bounded below by 0. Therefore the defining infimum satisfies

$$M_{\beta,k,q} \leq A_\beta(p').$$

Using $A_\beta(p^\star) = M_{\beta,k,q}$, we get $A_\beta(p^\star) \leq A_\beta(p')$. Multiplication by the nonnegative scalar $\kappa = \sigma_0^2/n$ preserves the inequality, so

$$\frac{\sigma_0^2}{n} A_\beta(p^\star) \leq \frac{\sigma_0^2}{n} A_\beta(p'),$$

as required.

□

Proof of Proposition 1. Let $p_j = qj/\beta$ for $j = 0, \ldots, \beta$. Since $\beta \geq 1$, $q > 0$, and $q \leq 1$, each coordinate is nonnegative, and $j \leq \beta$ gives

$$0 \leq p_j = \frac{qj}{\beta} \leq q \leq 1.$$

Thus $p_j \in [0, 1]$ for every $j = 0, \ldots, \beta$.

For the amplification bound, use the Lagrange endpoint weights on the equal grid. The grid is injective, since $p_j = p_m$ forces $qj/\beta = qm/\beta$ and hence $j = m$ because $q > 0$ and $\beta > 0$; so the Lagrange basis polynomials

$$L_j(t) = \prod_{\substack{0 \leq m \leq \beta \\ m \neq j}} \frac{t - p_m}{p_j - p_m}, \qquad j = 0, \ldots, \beta,$$

are well defined, and we set

$$w_j = L_j(1) - L_j(0), \qquad j = 0, \ldots, \beta.$$

By Lagrange interpolation on the $\beta + 1$ distinct nodes, these weights reproduce $r(1) - r(0)$ for every polynomial r of degree at most β; applied to the monomials $1, u, \ldots, u^\beta$, this is exactly $w \in W_\beta(p)$ as in Definition 3. Since $A_\beta(p)$ is the infimum in Definition 5 over the nonempty set $W_\beta(p)$, this feasible w gives

$$A_\beta(p) \leq \left(\sum_{j=0}^{\beta} |w_j| \right)^2.$$

Set $y = \beta/q$. Since $q \leq 1 \leq \beta$, we have $y \geq 1$, and hence $y^\beta \geq 1$.

The value at 1. Fix j. Each factor of $L_j(1)$ is controlled separately: from $0 \le p_m \le q \le 1$ we get $|1 - p_m| \le 1$, while

$$p_j - p_m = \frac{q}{\beta}(j - m), \qquad \text{so} \qquad |p_j - p_m| = \frac{q}{\beta}\,|j - m|.$$

There are exactly β indices $m \ne j$ in $\{0, \dots, \beta\}$, so multiplying the β factor bounds gives

$$|L_j(1)| = \prod_{m \ne j} \frac{|1 - p_m|}{|p_j - p_m|} \le \prod_{m \ne j} \frac{\beta/q}{|j - m|} = \frac{y^\beta}{\prod_{m \ne j} |j - m|}.$$

The remaining product is a factorial pair: splitting the indices below and above j,

$$\prod_{\substack{0 \le m \le \beta \\ m \ne j}} |j - m| = \left(\prod_{m=0}^{j-1} (j - m)\right)\left(\prod_{m=j+1}^{\beta} (m - j)\right) = j!\,(\beta - j)!.$$

Hence

$$|L_j(1)| \le \frac{y^\beta}{j!(\beta - j)!}.$$

Summing over j requires the binomial identity

$$\sum_{j=0}^{\beta} \frac{1}{j!(\beta - j)!} = \frac{1}{\beta!} \sum_{j=0}^{\beta} \frac{\beta!}{j!(\beta - j)!} = \frac{2^\beta}{\beta!} \le 2,$$

where the first equality rescales each term by $\beta!$, the second sums the binomial coefficients $\beta!/(j!(\beta - j)!)$ over $j = 0, \dots, \beta$, and the last inequality is $2^\beta \le 2\,\beta!$, valid for every $\beta \ge 1$ by induction (the base case $\beta = 1$ reads $2 \le 2$, and the step multiplies the left side by 2 and the right side by $\beta + 1 \ge 2$). Therefore

$$\sum_{j=0}^{\beta} |L_j(1)| \le y^\beta \sum_{j=0}^{\beta} \frac{1}{j!(\beta - j)!} \le 2y^\beta.$$

The value at 0. The node p_0 is exactly 0, so $L_j(0) = L_j(p_0)$ is the Lagrange basis polynomial evaluated at a node: it equals 1 for $j = 0$ and 0 for $j \ne 0$. Hence

$$\sum_{j=0}^{\beta} |L_j(0)| = 1.$$

Conclusion. By the triangle inequality applied termwise to $w_j = L_j(1) - L_j(0)$,

$$\sum_{j=0}^{\beta} |w_j| \le \sum_{j=0}^{\beta} |L_j(1)| + \sum_{j=0}^{\beta} |L_j(0)| \le 2y^\beta + 1 \le 3y^\beta,$$

where the last inequality uses $y^\beta \ge 1$. Both sides are nonnegative, so squaring and substituting $y = \beta/q$ yields

$$A_\beta(p^{\mathrm{eq}}(\beta, q)) \le \left(\sum_{j=0}^{\beta} |w_j|\right)^2 \le (3y^\beta)^2 = 9\left(\frac{\beta}{q}\right)^{2\beta}.$$

□

Proof of Proposition 2. At the full budget $q = 1$ and with $k = \beta$, the equal-spacing schedule has entries

$$p_j^{\mathrm{eq}}(\beta, 1) = \frac{j}{\beta}, \qquad j = 0, \dots, \beta, \qquad \text{so in particular} \qquad p_0^{\mathrm{eq}}(\beta, 1) = 0, \quad p_\beta^{\mathrm{eq}}(\beta, 1) = 1.$$

Write $p_j = p_j^{\mathrm{eq}}(\beta, 1)$ and

$$w_0 = -1, \qquad w_\beta = 1, \qquad w_j = 0 \quad (1 \le j \le \beta - 1).$$

Because $\beta \ge 1$, the first and last grid indices are distinct, i.e. $\beta \ne 0$, so the two prescriptions $w_0 = -1$ and $w_\beta = 1$ do not conflict and the vector w is well defined.

The moment equations. Only the indices $j = 0$ and $j = \beta$ carry nonzero weight. For the zeroth moment, $p_j^0 = 1$ for every j (including $p_0 = 0$), so

$$\sum_{j=0}^{\beta} w_j p_j^0 = w_0 + w_\beta = (-1) + 1 = 0.$$

Now fix $1 \le \ell \le \beta$. Since $p_0 = 0$ gives $p_0^\ell = 0^\ell = 0$ and $p_\beta = 1$ gives $p_\beta^\ell = 1$,

$$\sum_{j=0}^{\beta} w_j p_j^\ell = w_0\, p_0^\ell + w_\beta\, p_\beta^\ell = -0^\ell + 1^\ell = 1,$$

with all interior terms vanishing because their weights are zero. These are exactly the moment equations of Definition 3, so $w \in W_\beta(p^{\mathrm{eq}}(\beta, 1))$.

The amplification bound. By Definition 5, $A_\beta(p^{\mathrm{eq}}(\beta, 1))$ is the infimum of $\left(\sum_j |w'_j|\right)^2$ over $w' \in W_\beta(p^{\mathrm{eq}}(\beta, 1))$, a set of nonnegative numbers, so the feasible w just exhibited gives

$$A_\beta(p^{\mathrm{eq}}(\beta, 1)) \le \left(\sum_{j=0}^{\beta} |w_j| \right)^2 .$$

The absolute values contribute only at $j = 0$ and $j = \beta$, so

$$\sum_{j=0}^{\beta} |w_j| = |-1| + |1| = 1 + 1 = 2.$$

Therefore

$$A_\beta(p^{\mathrm{eq}}(\beta, 1)) \le 2^2 = 4.$$

□

Proof of Theorem 2. 1. *The constants.* By Lemma 6, applied with the cap $q_{\max}$, choose constants $C_{\mathrm{ep}}(q_{\max}) > 0$ and $c_{\mathrm{ep}}(q_{\max}) > 0$ such that, writing $x_q = 2/q - 1$ and

$$\lambda(q) = x_q + \sqrt{x_q^2 - 1},$$

the following two bounds hold for every $\beta \geq 1$ and every $0 < q \leq q_{\max}$: first, every real polynomial R with $\deg(R) \leq \beta$ and $\sup_{x \in [-1,1]} |R(x)| \leq 1$ satisfies

$$|R(x_q) - R(-1)| \leq C_{\mathrm{ep}}(q_{\max})\, \lambda(q)^{\beta};$$

and second,

$$T_\beta(x_q) - 1 \geq c_{\mathrm{ep}}(q_{\max})\, \lambda(q)^{\beta}.$$

These constants are chosen before the oversampling ratio is fixed, so set

$$C_-(q_{\max}) = c_{\mathrm{ep}}(q_{\max})^2 > 0.$$

Now fix $c > 1$. The Ehlich–Zeller Chebyshev–Lobatto mesh inequality used as the premise of Lemma 5 is available in the following form: for every pair of nonnegative integers β, k, every real polynomial R with $\deg(R) \leq \beta$ and $\beta < k$, and every $x \in [-1, 1]$,

$$|R(x)| \leq \frac{1}{\cos(\pi\beta/(2k))} \max_{0 \leq j \leq k} \left| R\left(-\cos\left(\frac{\pi j}{k} \right) \right) \right|.$$

Applying Lemma 5 with this mesh inequality, choose $K(c) > 0$ with the property that every real polynomial R of degree at most β obeys $\sup_{x \in [-1,1]} |R(x)| \leq K(c) \max_{0 \leq j \leq k} |R(-\cos(\pi j/k))|$ whenever $\beta \geq 1$ and $k \geq c\beta$, and set

$$C_+(c, q_{\max}) = \big(K(c) C_{\mathrm{ep}}(q_{\max}) \big)^2 > 0.$$

2. *Standing consequences of the hypotheses.* Fix β, k, q satisfying the hypotheses. The low-budget-regime hypothesis gives $q > 0$, and the low-budget cap of Assumption 4 gives $q \leq q_{\max} < 1$; hence $q \leq 1$, $q < 1$, and $0 < q \leq 1$. Since $k \geq c\beta$, $c > 1$, and $\beta \geq 1$, we have $\beta \leq c\beta \leq k$, and in particular $k \geq 1$. Therefore Lemma 7 gives

   $$p^{\mathrm{Ch}}(k, q) \in S_{k,q}.$$

   The exterior parameter is the exponential base in the statement:

   $$\lambda(q) = \frac{(1 + \sqrt{1 - q})^2}{q}.$$

   Indeed, $x_q^2 - 1 = (2/q - 1)^2 - 1 = 4/q^2 - 4/q = 4(1 - q)/q^2$, which is nonnegative because $q < 1$, so $\sqrt{x_q^2 - 1} = 2\sqrt{1 - q}/q$; adding $x_q = (2 - q)/q$ gives $\lambda(q) = \big(2 - q + 2\sqrt{1 - q}\big)/q$, and expanding $(1 + \sqrt{1 - q})^2 = 1 + 2\sqrt{1 - q} + (1 - q) = 2 - q + 2\sqrt{1 - q}$ identifies the two expressions. Finally, for any schedule $p' \in S_{k,q}$ we record the dual quantity of Lemma 3,

   $$D(p') = \sup \left\{ |r(1) - r(0)| : r \in \mathbb{R}[x],\ \deg(r) \leq \beta,\ \max_{0 \leq j \leq k} |r(p'_j)| \leq 1 \right\},$$

   and note that $D(p') \geq 0$ because the zero polynomial is feasible.

3. *Lower bound at an arbitrary $p \in S_{k,q}$.* Since $\beta \geq 1$, $\beta \leq k$, and $p \in S_{k,q}$, Lemma 8 makes $W_\beta(p)$ nonempty; and the strict ordering $p_0 < \cdots < p_k$ required by Definition 1 makes the nodes distinct. Thus Lemma 3 applies and gives

   $$A_\beta(p) = D(p)^2.$$

Consider the test polynomial

$$r(u) = T_\beta\left(\frac{2u}{q} - 1\right).$$

Its degree is at most β, being the composition of the degree-β polynomial T_β with an affine map. For every rollout node, $p \in S_{k,q}$ gives $p_j \ge p_0 = 0$ and, by monotonicity, $p_j \le p_k = q$; hence

$$0 \le p_j \le q, \qquad \text{so} \qquad \frac{2p_j}{q} - 1 \in [-1, 1],$$

and since $|T_\beta| \le 1$ on $[-1, 1]$ we get $|r(p_j)| \le 1$. Therefore r is feasible in the supremum defining $D(p)$. Its endpoint values are

$$r(1) = T_\beta\left(\frac{2}{q} - 1\right) = T_\beta(x_q), \qquad r(0) = T_\beta(-1).$$

The endpoint lower bound from Lemma 6 gives

$$T_\beta(x_q) - 1 \ge c_{\text{ep}}(q_{\max})\lambda(q)^\beta.$$

Since $x_q > 1$ we have $T_\beta(x_q) \ge 1 > 0$, so $|T_\beta(x_q)| = T_\beta(x_q)$; and $|T_\beta(-1)| \le 1$. The reverse triangle inequality therefore yields

$$T_\beta(x_q) - 1 \le |T_\beta(x_q)| - |T_\beta(-1)| \le |T_\beta(x_q) - T_\beta(-1)| = |r(1) - r(0)| \le D(p).$$

Hence $D(p) \ge c_{\text{ep}}(q_{\max})\lambda(q)^\beta \ge 0$, and squaring both nonnegative sides,

$$A_\beta(p) = D(p)^2 \ge c_{\text{ep}}(q_{\max})^2\lambda(q)^{2\beta} = C_-(q_{\max})\left(\frac{(1 + \sqrt{1-q})^2}{q}\right)^{2\beta}.$$

4. *Upper bound at the Chebyshev–Lobatto schedule.* The schedule $p^{\text{Ch}}(k, q)$ is admissible by Lemma 7. Together with $\beta \ge 1$ and $\beta \le k$, Lemma 8 makes $W_\beta(p^{\text{Ch}}(k, q))$ nonempty; the strict ordering in $p^{\text{Ch}}(k, q) \in S_{k,q}$ gives distinct nodes. Thus Lemma 3 gives $A_\beta(p^{\text{Ch}}(k, q)) = D(p^{\text{Ch}}(k, q))^2$, and it suffices to bound the dual supremum. Let r be any polynomial with $\deg(r) \le \beta$ and $\max_{0 \le j \le k} |r(p_j^{\text{Ch}}(k, q))| \le 1$, and put

$$R(x) = r\left(\frac{q}{2}(x + 1)\right).$$

Then $\deg(R) \le \beta$, again because the inner map is affine. At the Lobatto abscissae the inner map reproduces the schedule,

$$\frac{q}{2}\left(-\cos\frac{\pi j}{k} + 1\right) = \frac{q}{2}\left(1 - \cos\frac{\pi j}{k}\right) = p_j^{\text{Ch}}(k, q), \qquad j = 0, \ldots, k,$$

so $R(-\cos(\pi j/k)) = r(p_j^{\text{Ch}}(k, q))$ and therefore

$$\max_{0 \le j \le k}\left|R\left(-\cos\frac{\pi j}{k}\right)\right| \le 1.$$

The oversampled norming inequality supplied by Lemma 5 with the displayed Ehlich–Zeller mesh input applies because $\beta \ge 1$, $k \ge c\beta$, and $\deg(R) \le \beta$. It gives

$$|R(x)| \le K(c) \cdot 1 = K(c) \qquad (x \in [-1, 1]).$$

Thus $S = K(c)^{-1}R$ satisfies $\deg(S) \le \beta$ and $|S(x)| \le 1$ on $[-1,1]$. The inner affine map sends the two relevant points to the two endpoints,

$$\frac{q}{2}(x_q + 1) = \frac{q}{2} \cdot \frac{2}{q} = 1, \qquad \frac{q}{2}(-1+1) = 0,$$

so $R(x_q) = r(1)$ and $R(-1) = r(0)$, and hence $S(x_q) = K(c)^{-1}r(1)$ and $S(-1) = K(c)^{-1}r(0)$. Applying the endpoint upper bound from Lemma 6 to S gives

$$K(c)^{-1}|r(1) - r(0)| = |S(x_q) - S(-1)| \le C_{\mathrm{ep}}(q_{\max})\lambda(q)^{\beta},$$

that is, $|r(1) - r(0)| \le K(c)C_{\mathrm{ep}}(q_{\max})\lambda(q)^{\beta}$. Taking the supremum over all such r gives $0 \le D(p^{\mathrm{Ch}}(k,q)) \le K(c)C_{\mathrm{ep}}(q_{\max})\lambda(q)^{\beta}$, and squaring both nonnegative sides,

$$A_\beta\Big(p^{\mathrm{Ch}}(k,q)\Big) \le \big(K(c)C_{\mathrm{ep}}(q_{\max})\big)^2 \lambda(q)^{2\beta} = C_+(c, q_{\max}) \left(\frac{(1+\sqrt{1-q})^2}{q}\right)^{2\beta}.$$

5. *The two-sided minimax bound.* The first two displayed inequalities of the statement are exactly the bounds just proved. For the minimax lower bound, the universal lower estimate obtained from Lemma 8, Lemma 3, and Lemma 6 holds for every $p \in S_{k,q}$, and the schedule class is nonempty because $p^{\mathrm{Ch}}(k,q) \in S_{k,q}$; so the common bound is a lower bound for the nonempty set $\{A_\beta(p) : p \in S_{k,q}\}$, and passing to its infimum gives

$$M_{\beta,k,q} \ge C_-(q_{\max}) \left(\frac{(1+\sqrt{1-q})^2}{q}\right)^{2\beta}.$$

For the minimax upper bound, that set is bounded below by 0 and contains $A_\beta(p^{\mathrm{Ch}}(k,q))$, so its infimum is at most that value; chaining with the Chebyshev upper bound obtained from Lemma 7, Lemma 5, and Lemma 6,

$$M_{\beta,k,q} \le A_\beta\Big(p^{\mathrm{Ch}}(k,q)\Big) \le C_+(c, q_{\max}) \left(\frac{(1+\sqrt{1-q})^2}{q}\right)^{2\beta}.$$

This proves the claimed two-sided minimax estimate.

□

Proof of Lemma 2. Write $s = \sigma_0^2/n$, where division by n is real division by the natural cast of n used in the real-valued risk definitions; when $n = 0$, this quotient is the total real quotient and equals 0. Since $\sigma_0^2 \ge 0$ and $(n : \mathbb{R}) \ge 0$, we have $s \ge 0$. By Definition 6, $\Gamma_P(p)$ is the design covariance matrix of $(\bar{Y}_0, \ldots, \bar{Y}_k)$, so that for every weight vector w

$$w'\Gamma_P(p)w = \mathrm{Var}_\pi\left(\sum_{j=0}^{k} w_j \bar{Y}_j\right).$$

All infima and suprema below are the real-valued operations appearing in Definition 6 and Definition 5. Thus an empty law-value supremum has value 0, and if every element of a law-value set is at most a nonnegative number B, then its supremum is at most B.

An order-theoretic comparison of infima. Both assertions use the following elementary fact for these real-valued infima. Let $F, G \subseteq \mathbb{R}$, let $c \geq 0$, suppose F is nonempty and bounded below and G is nonempty, and suppose that

$$\text{for every } g \in G \text{ there is } f \in F \text{ with } f \leq c\,g.$$

Then

$$\inf F \leq c \inf G.$$

Indeed, scalar multiplication by $c \geq 0$ commutes with the real-valued infimum in the form $c \inf G = \inf\{c\,g : g \in G\}$. Since F is nonempty and bounded below, $\inf F$ is a lower bound for F. For each $g \in G$, the witness $f \in F$ gives $\inf F \leq f \leq c\,g$, so $\inf F$ is a lower bound for $\{c\,g : g \in G\}$. Hence $\inf F \leq \inf\{c\,g : g \in G\} = c \inf G$.

Step 1: the fixed-schedule bound for every admissible schedule. Fix any $p' \in S_{k,q}$ and apply the fact with $c = s$ and

$$F = \left\{ \sup_{P \in \mathcal{P}_\beta} \operatorname{Var}_\pi\left(\sum_{j=0}^{k} w_j \bar{Y}_j\right) : w \in W_\beta(p') \right\}, \qquad G = \left\{ \left(\sum_{j=0}^{k} |w_j|\right)^2 : w \in W_\beta(p') \right\}.$$

Since $\beta \geq 1$, $k \geq \beta$, and $p' \in S_{k,q}$, Lemma 8 supplies some $w \in W_\beta(p')$, so both F and G are nonempty. Moreover F is bounded below by 0: every element of each law-value set is a design variance, hence is nonnegative, and the real-valued supremum of such a set is nonnegative, including the empty case. For the link condition, take $g = \left(\sum_j |w_j|\right)^2 \in G$ and let $f \in F$ be the element indexed by the same w. For every law value appearing in this supremum, Definition 2 gives the round-mean variance envelope for $\bar{Y}_0, \ldots, \bar{Y}_k$ with scale σ_0^2/n. Hence Lemma 9 gives

$$\operatorname{Var}_\pi\left(\sum_{j=0}^{k} w_j \bar{Y}_j\right) \leq s \left(\sum_{j=0}^{k} |w_j|\right)^2.$$

The right-hand side is independent of the law value and is nonnegative, so the real-valued supremum over the law-value set is at most this same quantity. Thus $f \leq s\,g$. The fact therefore applies, and since $\inf G = A_\beta(p')$ by Definition 5,

$$\inf_{w \in W_\beta(p')} \sup_{P \in \mathcal{P}_\beta} w' \Gamma_P(p') w \leq s \inf_{w \in W_\beta(p')} \left(\sum_{j=0}^{k} |w_j|\right)^2 = \frac{\sigma_0^2}{n} A_\beta(p').$$

Applying this to the given schedule p proves the first assertion,

$$\inf_{w \in W_\beta(p)} \sup_{P \in \mathcal{P}_\beta} w' \Gamma_P(p) w \leq \frac{\sigma_0^2}{n} A_\beta(p).$$

Step 2: passing to the nested risk. Apply the same fact a second time, now with $c = s$ and

$$F = \left\{ \inf_{w \in W_\beta(p')} \sup_{P \in \mathcal{P}_\beta} w' \Gamma_P(p') w : p' \in S_{k,q} \right\}, \qquad G = \left\{ A_\beta(p') : p' \in S_{k,q} \right\}.$$

By Definition 6 and Definition 5, $\inf F = R_{\text{exact}}(\beta, k, q)$ and $\inf G = M_{\beta,k,q}$. Both sets are nonempty because the given schedule satisfies $p \in S_{k,q}$. The set F is bounded below by 0: for each schedule p', every fixed-weight law supremum is nonnegative by the preceding nonnegativity argument, and the corresponding real-valued infimum over weights is therefore nonnegative. The link condition is exactly the fixed-schedule bound from Step 1: for $g = A_\beta(p')$, take the element of F indexed by the same p'. Hence

$$R_{\text{exact}}(\beta, k, q) \leq \frac{\sigma_0^2}{n} \inf_{p' \in S_{k,q}} A_\beta(p') = \frac{\sigma_0^2}{n} M_{\beta,k,q}.$$

□

Proof of Proposition 3. By Theorem 2, applied with the fixed cap $q_{\max}$ and the oversampling constant $c > 1$, there is a constant $C_+(c, q_{\max}) > 0$ such that, whenever $\beta \geq 1$, $k \geq c\beta$, $q > 0$, and the low-budget cap holds,

$$A_\beta\left(p^{\text{Ch}}(k, q)\right) \leq C_+(c, q_{\max}) \left(\frac{(1 + \sqrt{1-q})^2}{q}\right)^{2\beta}.$$

This is the constant used in the statement; it depends only on c and $q_{\max}$.

Now fix $n, \beta, k, \Omega, \pi, q, \sigma_0^2$ satisfying the hypotheses of the lemma, and write $p^{\text{Ch}} = p^{\text{Ch}}(k, q)$.

The hypotheses of the two inputs hold. Since $k = \lceil c\beta \rceil$ is the least integer at or above $c\beta$, and since $c > 1$ and $\beta \geq 1$ give $\beta \leq c\beta$, we have

$$\beta \leq c\beta \leq k.$$

In particular $k \geq \beta \geq 1$, so $k \geq 1$. The low-budget cap of Assumption 4 gives $q \leq q_{\max} < 1$, which together with $q > 0$ yields

$$0 < q \leq 1.$$

Therefore Lemma 7 applies and gives

$$p^{\text{Ch}} \in S_{k,q}.$$

Finally, write $s = \sigma_0^2/n$, interpreting the quotient as the total real division used in the risk definitions; in particular the quotient is defined also when $n = 0$, where it equals 0. Since $\sigma_0^2 \geq 0$ and the real cast of a natural number is nonnegative, this gives

$$s = \frac{\sigma_0^2}{n} \geq 0$$

uniformly over the whole range of population sizes, with no case distinction.

Envelope comparison at the Chebyshev schedule. For any weight vector w, the exact quadratic form is by definition the design variance of the corresponding linear rollout estimator,

$$w' \Gamma_P\left(p^{\text{Ch}}\right) w = \text{Var}_\pi\left(\sum_{j=0}^{k} w_j \bar{Y}_j\right),$$

because $\Gamma_P(p)$ is the design covariance matrix of $(\bar{Y}_0, \ldots, \bar{Y}_k)$. Applying the fixed-schedule part of Lemma 2 to the admissible schedule p^{Ch}, whose hypotheses $\beta \geq 1$, $k \geq \beta$, $0 < q \leq 1$, $\sigma_0^2 \geq 0$ and $p^{\text{Ch}} \in S_{k,q}$ were just verified, gives

$$\inf_{w \in W_\beta(p^{\text{Ch}})} \sup_{P \in \mathcal{P}_\beta} \text{Var}_\pi\left(\sum_{j=0}^{k} w_j \bar{Y}_j\right) \leq \frac{\sigma_0^2}{n} A_\beta\left(p^{\text{Ch}}\right).$$

Conclusion. Since $s = \sigma_0^2/n \geq 0$, multiplying the Chebyshev upper bound for $A_\beta(p^{\mathrm{Ch}})$ by σ_0^2/n preserves the inequality, so chaining the last two displays gives

$$\inf_{w \in W_\beta(p^{\mathrm{Ch}})} \sup_{P \in \mathcal{P}_\beta} \mathrm{Var}_\pi\left(\sum_{j=0}^{k} w_j \bar{Y}_j\right) \leq \frac{\sigma_0^2}{n} A_\beta\left(p^{\mathrm{Ch}}\right) \leq \frac{\sigma_0^2}{n} C_+(c, q_{\max}) \left(\frac{(1+\sqrt{1-q})^2}{q}\right)^{2\beta},$$

which is the asserted bound. □

References

Peter M. Aronow and Cyrus Samii. Estimating average causal effects under general interference, with application to a social network experiment. *The Annals of Applied Statistics*, 11(4):1912–1947, 2017. doi: 10.1214/16-AOAS1005.

Susan Athey, Dean Eckles, and Guido W. Imbens. Exact¡i¿p¡/i¿-values for network interference. *Journal of the American Statistical Association*, 113(521):230–240, 2017. doi: 10.1080/01621459.2016.1241178.

Guillaume Basse, Avi Feller, and Panos Toulis. Randomization tests of causal effects under interference. *Biometrika*, 106(2):487–494, 2019. doi: 10.1093/biomet/asy072.

Subhankar Bhadra and Michael Schweinberger. Causal inference under network interference, 2025. URL https://arxiv.org/abs/2508.06808.

Chencheng Cai, Xu Zhang, and Edoardo M. Airoldi. Independent-set design of experiments for estimating treatment and spillover effects under network interference, 2023. URL https://arxiv.org/abs/2312.04026.

Xinyuan Chen and Fan Li. Model-assisted inference for dynamic causal effects in staggered rollout cluster randomized experiments, 2025. URL https://arxiv.org/abs/2502.10939.

Mayleen Cortez, Matthew Eichhorn, and Christina Lee Yu. Staggered rollout designs enable causal inference under interference without network knowledge, 2022. URL https://arxiv.org/abs/2205.14552.

Mayleen Cortez-Rodriguez, Matthew Eichhorn, and Christina Lee Yu. Exploiting neighborhood interference with low order interactions under unit randomized design, 2022. URL https://arxiv.org/abs/2208.05553.

Mayleen Cortez-Rodriguez, Matthew Eichhorn, and Christina Lee Yu. Analysis of two-stage rollout designs with clustering for causal inference under network interference, 2024. URL https://arxiv.org/abs/2405.05119.

Dean Eckles, Brian Karrer, and Johan Ugander. Design and analysis of experiments in networks: Reducing bias from interference. *Journal of Causal Inference*, 5(1), 2017. doi: 10.1515/jci-2015-0021.

Matthew Eichhorn, Samir Khan, Johan Ugander, and Christina Lee Yu. Low-order outcomes and clustered designs: Combining design and analysis for causal inference under network interference, 2024. URL https://arxiv.org/abs/2405.07979.

Xinyuan Fan, Chenlei Leng, and Weichi Wu. Causal inference under interference: Regression adjustment and optimality, 2025. URL https://arxiv.org/abs/2502.06008.

Guido W. Imbens and Donald B. Rubin. *Causal Inference for Statistics, Social, and Biomedical Sciences*. Cambridge University Press, 2015. doi: 10.1017/CBO9781139025751.

Yiming Jiang and He Wang. Causal inference under network interference using a mixture of randomized experiments, 2023. URL https://arxiv.org/abs/2309.00141.

Samuel Karlin and William J. Studden. *Tchebycheff Systems: With Applications in Analysis and Statistics*. Wiley-Interscience, New York, 1966.

J. Kiefer. General equivalence theory for optimum designs (approximate theory). *The Annals of Statistics*, 2(5):849–879, 1974. doi: 10.1214/aos/1176342810.

Jack Kiefer and Jacob Wolfowitz. Optimum designs in regression problems. *The Annals of Mathematical Statistics*, 30(2):271–294, 1959. doi: 10.1214/aoms/1177706252.

Charles F. Manski. Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies*, 60(3):531–542, 1993. doi: 10.2307/2298123.

Charles F. Manski. Identification of treatment response with social interactions. *The Econometrics Journal*, 16(1):S1–S23, 2013. doi: 10.1111/j.1368-423X.2012.00368.x.

David Puelz, Guillaume Basse, Avi Feller, and Panos Toulis. A graph-theoretic approach to randomization tests of causal effects under general interference, 2019. URL https://arxiv.org/abs/1910.10862.

Friedrich Pukelsheim. *Optimal Design of Experiments*. Wiley, New York, 1993.

Jonathan Roth and Pedro H. C. Sant'Anna. Efficient estimation for staggered rollout designs, 2021. URL https://arxiv.org/abs/2102.01291.

Donald B. Rubin. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5):688–701, 1974. doi: 10.1037/h0037350.

Maresa Schröder, Miruna Oprescu, Stefan Feuerriegel, and Nathan Kallus. Causal inference on networks under misspecified exposure mappings: A partial identification framework, 2026. URL https://arxiv.org/abs/2602.03459.

Kirstine Smith. On the standard deviations of adjusted and interpolated values of an observed polynomial function and its constants and the guidance they give towards a proper choice of the distribution of observations. *Biometrika*, 12(1/2):1–85, 1918. doi: 10.1093/biomet/12.1-2.1.

Fredrik Sävje, Peter M. Aronow, and Michael G. Hudgens. Average treatment effects in the presence of unknown interference, 2017. URL https://arxiv.org/abs/1711.06399.

Johan Ugander, Brian Karrer, Lars Backstrom, and Jon Kleinberg. Graph cluster randomization. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 329–337, 2013. doi: 10.1145/2487575.2487695.