hexagon:2611.00659v1

THE EXACT FDR–FNR FRONTIER AND A SEPARABLE GAP UNDER ONE-FACTOR GAUSSIAN DEPENDENCE

SAMUIL PETKOV

Abstract. We determine the exact asymptotic trade-off between the false discovery rate and the false non-discovery rate in a two-group Gaussian model generated by one unobserved common factor of fixed strength. For fixed mixture weight, signal strength, equicorrelation, and false discovery target, the optimum over all randomized full-vector procedures converges to the lower convex envelope of the independent Gaussian threshold curve with effective signal $\mu/\sqrt{1-\rho}$. Thus not observing the factor causes no additional asymptotic loss relative to the factor-revealed experiment. The converse reveals the factor and reduces exactly to an independent experiment. For achievability, factor estimation, a pathwise threshold-crossing bound, and random-denominator control yield an observable optimal rule with finite-sample FDR calibration. If the three model parameters are replaced by consistent estimates, a plug-in rule attains the same frontier with asymptotic FDR control. We also analyze the broad coordinate-local class $\delta_{n,i} = f_{n,i}(X_i, A_n)$, where the functions may be nonidentical and dimension-dependent and the shared external seed is independent of the model. Every oracle supporting line lying strictly below the two trivial endpoint values has a positive asymptotic gap from this separable class; an analytic parameter choice certifies that such a line exists. For the constrained problem, the gap vanishes at every level exactly when the oracle frontier is the chord joining the reject-none and reject-all rules; otherwise it is strict at every level strictly between the two trivial endpoint levels. For this fixed rank-one Gaussian family, the first result exactly evaluates the compound frontier and the second proves existence of a strict gap from the stated coordinate-local class. Broad dependent model classes and the complete constrained separable frontier remain open.

1. Introduction

Benjamini and Hochberg introduced the false discovery rate as the expected proportion of false rejections among all rejections [1, p. 289]. Genovese and Wasserman introduced the dual false non-discovery proportion—the proportion of false non-rejections among all non-rejections—and studied its expectation [8, pp. 499–500 and eqs. (1)–(2)]. In the notation used here, FDR and FNR are these two expectations. Their joint optimization is materially different from optimization of the corresponding marginal ratios.

In the i.i.d. two-group random-mixture model, Nie and Wu show that the all-rule asymptotic optimum is the greatest convex minorant of the optimal marginal-FNR curve. Their Gaussian example is generated by one-sided Neyman–Pearson thresholds, and its oracle may use one common random choice between two such rules [15, Theorem 1, eq. (8), and Appendix A]. In their terminology that common choice makes the oracle compound and can improve strictly on fixed coordinatewise rules. The coordinate-local class studied here is broader because it permits arbitrary external randomization shared across coordinates. Nie and Wu also prove a Gaussian extension under a spectral-norm weak-dependence condition and leave stronger dependence as an open direction [15, Section 6.3 and Theorem 24].

The broader problem is to characterize the dependent-model value $\Psi_{\mathrm{dep}}(\alpha; P)$ for substantial classes of high-dimensional laws and to compare compound and separable restrictions. This paper does not resolve that broad problem. It gives a complete pointwise answer for the compound value in one tractable but genuinely strongly dependent family and proves a qualified strict compound–separable gap within that family. Conditional on the latent labels, the test statistics have fixed equicorrelation $\rho \in (0, 1)$. The largest covariance eigenvalue grows linearly with the number of hypotheses, and $\|\Sigma_n - I_n\|_{\mathrm{op}} = \rho(n-1)$, so the weak-dependence theorem just cited does not apply. At the same time, the dependence comes from one common Gaussian factor that can be estimated from the full observation vector at the parametric rate.

Date: 18 July 2026.

2020 *Mathematics Subject Classification.* Primary 62J15; Secondary 62H15, 62C20.

Key words and phrases. false discovery rate, false non-discovery rate, multiple testing, compound decision rule, separable decision rule, Gaussian factor model, equicorrelation.

Contributions and novelty. For the fixed known one-factor family in (2.1), the new results are threefold.

(i) **Exact compound frontier.** We evaluate the pointwise asymptotic FDR–FNR frontier over all randomized full-vector Markov kernels as $B_{\mu/\sqrt{1-\rho}}$. This fixed-strength model lies outside the spectral-norm hypothesis of [15, Theorem 24].

(ii) **Constructive observable achievability.** We prove that revealing the latent factor has no asymptotic value for this frontier and give an observable factor-adjusted procedure with finite-sample FDR feasibility and asymptotically optimal FNR. Factor and empirical-mean residualization have precedents; the contribution here is all-procedure decision-theoretic optimality for the expectation-of-ratios objective. When π_1, μ, and ρ are unknown, consistent plug-in estimates give an observable rule with asymptotic FDR control and the same limiting FNR (theorem 7.7).

(iii) **A precise separable gap.** For nonidentical and dimension-dependent raw-coordinate rules with arbitrary independent shared randomization, we identify the exact limiting scalarized risk J_κ, prove the endpoint-strict criterion for $J_\kappa > H_\kappa$, and derive a supported-face gap with an analytic Gaussian witness. We also show that the constrained gap vanishes at every level when the oracle frontier is the endpoint chord and is strict at every interior level otherwise. The permitted shared seed is broad enough to implement the common threshold-selection coin in the independent oracle; the gap therefore isolates the value of data-derived cross-coordinate information rather than the value of a common coin.

The compound proof reveals the factor for a Blackwell lower bound, thereby transforming the model into an independent Gaussian mixture with effective signal

$$a = \frac{\mu}{\sqrt{1-\rho}}.$$

For the reverse inequality, subtracting a full-sample factor estimate shifts every oracle residual by the same random quantity $\Delta_n = O_{\mathbb{P}}(n^{-1/2})$. Coordinatewise convergence alone is insufficient because a common shift may change order $\sqrt{n}$ decisions. We instead bound the fraction of changed decisions pathwise by the empirical mass in a shrinking interval around the threshold, then control the random FDP and FNP denominators explicitly. A single external coin that replaces the procedure by the reject-none rule enforces the FDR constraint nonasymptotically.

The separable proof uses a different route. After conditioning on the shared seed, weak-star compactness reduces each triangular array to an averaged response. Conditional Neyman–Pearson rearrangement lower-bounds its scalarized FDR–FNR risk by the oracle supporting value. Equality would require one fixed raw threshold to remain an optimal residual threshold under almost every Gaussian factor translation, which is impossible when the supporting line lies strictly below both trivial endpoints. Compactness upgrades this obstruction to a uniform positive gap.

Relation to prior work. Several decision-theoretic results are close in objective but different in model or scope. Sun and Cai construct oracle and asymptotically optimal local-index-of-significance procedures that minimize FNR subject to FDR in a two-state hidden Markov model [18]. Xie, Cai, Maris and Li derive a full-data oracle for mFNR subject to mFDR under dependent two-group models and study tractable marginal approximations; their corrected asymptotic result uses a different strong-signal regime [21, 20]. Wu and Peña derive finite-dimensional compound Bayes actions for scalarized FDP–FNP losses, including a specific copula dependence construction [19]. Heller and Rosset derive fixed-dimensional policies that maximize expected true discoveries under FDR or pFDR constraints in a general two-group model [11, 12]. None of these results evaluates the present fixed-equicorrelation all-rule frontier or the shared-seed coordinate-local gap.

The same or related dependence structures have also been studied for specified procedures and other risks. Finner, Dickhaus and Roters obtain limiting error rates for linear step-up rules under exchangeability, including equicorrelated Gaussian extreme configurations [6]. Roquain and Villers derive exact FDP distributions and moments for step-up rules, together with FDR and power formulas for step-down rules, in independent and equicorrelated Gaussian mixtures [16]. Delattre and Roquain study BH FDP convergence when the equicorrelation tends to zero [2]. Roy and Bhandari estimate the common component by the empirical mean in a fixed-equicorrelation Gaussian model with a sparse two-variance prior and prove asymptotic Bayes optimality for residualized fixed-cutoff rules [17]; their sparse additive-risk regime differs from the fixed-mixture expectation-of-ratios objective here.

Principal-factor methods estimate realized FDP under known or estimated covariance [4, 3], Kluger and Owen establish a BH FDP central limit theorem under factor models [13], and other factor-adjusted methods pursue FDR

control or power rather than an all-rule frontier [7, 14, 5]. Recent preprints by Ghosh and Chakrabarti propose covariance-adaptive residual step-down testing, with admissibility and simulation evidence, and study a Bayes oracle for additive misclassification loss under sparse one-factor dependence [10, 9]. Neither preprint derives the constrained expectation-of-ratios frontier or compares the compound and coordinate-local classes analyzed here.

Among the works reviewed above, we are not aware of one that evaluates the fixed-ρ, fixed-π_1, fixed-signal asymptotic constrained expectation-of-FDP versus expectation-of-FNP frontier over all randomized full-vector procedures. The closest all-rule frontier result reviewed here is the weak-dependence theorem of Nie and Wu, whose spectral-norm hypothesis is violated in the present model.

2. Model, loss functions, and the oracle curve

Fix known constants, not depending on n,

$$\pi_1 \in (0,1), \qquad \pi_0 = 1-\pi_1, \qquad \mu > 0, \qquad \rho \in (0,1).$$

Let $\Theta_1, \Theta_2, \dots$ be independent Bernoulli(π_1) variables. Let $Z_0, Z_1, Z_2, \dots$ be independent standard normal variables, independent of all labels. For every n and $i \le n$, observe

$$X_i = \mu\Theta_i + \sqrt{\rho}\, Z_0 + \sqrt{1-\rho}\, Z_i. \tag{2.1}$$

Procedures may depend on π_1, μ, ρ. All expectations below are joint over the labels, all Gaussian variables, and any auxiliary randomization. Write ϕ and Φ for the standard normal density and distribution function, and $\overline{\Phi} = 1 - \Phi$. Equivalently, conditional on $\Theta = (\Theta_1, \dots, \Theta_n)$,

$$X \mid \Theta \sim N(\mu\Theta, \Sigma_n), \qquad \Sigma_n = (1-\rho)I_n + \rho \mathbf{1}_n \mathbf{1}_n^{\mathsf{T}}.$$

A randomized multiple-testing procedure is a Markov kernel from $\mathbb{R}^n$ to $\{0,1\}^n$. The decision $\delta_i = 1$ means that coordinate i is rejected. Write

$$R = \sum_{i=1}^n \delta_i, \qquad V = \sum_{i=1}^n (1-\Theta_i)\delta_i,$$

$$S = n - R, \qquad U = \sum_{i=1}^n \Theta_i(1-\delta_i).$$

Use the zero-denominator conventions

$$\mathrm{FDP} = \frac{V}{R \vee 1}, \qquad \mathrm{FNP} = \frac{U}{S \vee 1}, \tag{2.2}$$

and define

$$\mathrm{FDR}(\delta) = \mathbb{E}[\mathrm{FDP}], \qquad \mathrm{FNR}(\delta) = \mathbb{E}[\mathrm{FNP}].$$

For $\alpha \in [0,1]$, let

$$\mathrm{OPT}_n(\alpha; \pi_1, \mu, \rho) = \inf\big\{\mathrm{FNR}(\delta) : \mathrm{FDR}(\delta) \le \alpha\big\}, \tag{2.3}$$

where the infimum ranges over all randomized full-vector procedures.

Let $P^{\mathrm{1f}}_{\pi_1,\mu,\rho} = (P^{\mathrm{1f}}_n)_{n\ge 1}$ denote this sequence of one-factor experiments. In the notation of the broader dependent-model problem,

$$\Psi_{\mathrm{dep}}(\alpha; P^{\mathrm{1f}}_{\pi_1,\mu,\rho}) = \liminf_{n\to\infty} \mathrm{OPT}_n(\alpha; \pi_1, \mu, \rho). \tag{2.4}$$

Set

$$a = \frac{\mu}{\sqrt{1-\rho}}. \tag{2.5}$$

Write $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ with its usual compact topology. For $t \in \mathbb{R}$, define

$$p_t = \pi_0\overline{\Phi}(t) + \pi_1\overline{\Phi}(t-a), \tag{2.6}$$

$$s_t = 1 - p_t = \pi_0\Phi(t) + \pi_1\Phi(t-a), \tag{2.7}$$

$$q_a(t) = \frac{\pi_0\overline{\Phi}(t)}{p_t}, \tag{2.8}$$

$$r_a(t) = \frac{\pi_1\Phi(t-a)}{s_t}. \tag{2.9}$$

At the extended thresholds, set

$$(q_a(+\infty), r_a(+\infty)) = (0, \pi_1), \qquad (q_a(-\infty), r_a(-\infty)) = (\pi_0, 0). \tag{2.10}$$

Thus p_t and s_t are the marginal rejection and nonrejection probabilities of the oracle threshold rule, while $q_a(t)$ is its conditional null probability among rejections and $r_a(t)$ is its conditional signal probability among nonrejections.

Definition 2.1 (Convexified oracle frontier)**.** Let

$$\Gamma_a = \big\{(q_a(t), r_a(t)) : t \in \overline{\mathbb{R}}\big\}, \qquad K_a = \mathrm{conv}(\Gamma_a).$$

For $\alpha \in [0, 1]$, define the constrained lower-envelope value

$$B_a(\alpha) = \min\big\{r : (q, r) \in K_a,\ q \le \alpha\big\}. \tag{2.11}$$

The minimum exists because Γ_a, and hence K_a, is compact. For $\alpha \in [0, \pi_0]$, $(\alpha, B_a(\alpha))$ is the lower boundary of K_a; the activity assertion is proved in proposition 8.2. For $\alpha \ge \pi_0$, $B_a(\alpha) = 0$. We prove below that a minimizing point always uses at most two points of Γ_a.

Theorem 2.2 (Exact one-factor frontier)**.** For every fixed

$$\alpha \in [0, 1], \qquad \pi_1 \in (0, 1), \qquad \mu > 0, \qquad \rho \in (0, 1),$$

the limit exists and

$$\lim_{n\to\infty} \mathrm{OPT}_n(\alpha; \pi_1, \mu, \rho) = B_{\mu/\sqrt{1-\rho}}(\alpha). \tag{2.12}$$

The converse holds over all deterministic or randomized full-vector Markov kernels. There is an explicit sequence of unobserved-factor procedures satisfying

$$\mathrm{FDR}(\delta^{(n)}) \le \alpha$$

for every n, and

$$\limsup_{n\to\infty} \mathrm{FNR}(\delta^{(n)}) \le B_{\mu/\sqrt{1-\rho}}(\alpha).$$

Moreover,

$$\mathrm{OPT}_n(0; \pi_1, \mu, \rho) = \pi_1 \tag{2.13}$$

for every n, while

$$\mathrm{OPT}_n(\alpha; \pi_1, \mu, \rho) = 0 \tag{2.14}$$

for every n and every $\alpha \ge \pi_0$.

Corollary 2.3 (The one-factor value of the dependent frontier)**.** For the experiment sequence in (2.4),

$$\Psi_{\mathrm{dep}}(\alpha; P^{\mathrm{1f}}_{\pi_1,\mu,\rho}) = B_{\mu/\sqrt{1-\rho}}(\alpha), \qquad \alpha \in [0, 1].$$

Thus the theorem characterizes Ψ_{dep} exactly for this homogeneous rank-one Gaussian family.

Remark 2.4 (Scope of the open problem)*.* The corollary is a special-case solution, not a characterization for broad dependent laws. The compound optimum is evaluated for the homogeneous rank-one Gaussian family only. In Section 9 we separately prove a lower bound and a strict gap for raw-coordinate rules with model-independent external randomization. That result does not cover data-dependent shared couplings disguised as marginally coordinatewise kernels, and it does not imply a gap for arbitrary dependent model classes. Within the present family, theorem 9.8 shows that the constrained gap is either absent at every level or strict at every level $\alpha \in (0, \pi_0)$.

Remark 2.5*.* The theorem is pointwise in the fixed parameter quadruple. No uniformity is claimed as π_1 approaches an endpoint, $\mu \downarrow 0$, $\rho \downarrow 0$, $\rho \uparrow 1$, $\alpha \downarrow 0$, or as an optimizer threshold diverges. The boundary models $\mu = 0$, $\rho = 1$, and $\pi_1 \in \{0, 1\}$ are not covered by the theorem and can have different geometry or degeneracies.

3. The observed-factor experiment

3.1. **Strong dependence.**

Proposition 3.1 (Spectrum of the equicorrelation matrix)**.** The matrix Σ_n has eigenvalue $1-\rho$ with multiplicity $n-1$, and eigenvalue $1+(n-1)\rho$ with multiplicity one. Consequently,

$$\|\Sigma_n\|_{\mathrm{op}} = 1+(n-1)\rho = \Theta(n), \qquad \|\Sigma_n - I_n\|_{\mathrm{op}} = \rho(n-1).$$

Proof. If $v \perp \mathbf{1}_n$, then $\mathbf{1}_n\mathbf{1}_n^\top v = 0$, so $\Sigma_n v = (1-\rho)v$. On the other hand,

$$\Sigma_n\mathbf{1}_n = (1-\rho)\mathbf{1}_n + \rho n\mathbf{1}_n = [1+(n-1)\rho]\mathbf{1}_n.$$

These eigenspaces span $\mathbb{R}^n$, which proves the claim. □

In particular, the weak-dependence condition $\|\Sigma_n - \sigma^2 I_n\|_{\mathrm{op}} \to 0$ used in the Gaussian extension of [15] fails for every fixed $\rho > 0$.

3.2. **Exact reduction when the factor is observed.**

Suppose temporarily that Z_0 is revealed and define

$$Y_i = \frac{X_i - \sqrt{\rho}\, Z_0}{\sqrt{1-\rho}}. \tag{3.1}$$

Then

$$Y_i = a\Theta_i + Z_i, \tag{3.2}$$

so the pairs (Θ_i, Y_i) are independent and identically distributed, with

$$Y_i \mid \Theta_i = 0 \sim N(0,1), \qquad Y_i \mid \Theta_i = 1 \sim N(a,1).$$

Proposition 3.2 (Ancillary-factor reduction)**.** Let $\mathrm{OPT}_n^{\mathrm{obs}}(\alpha)$ denote the optimum when Z_0 is observed, and let $\mathrm{OPT}_n^{\mathrm{ind}}(\alpha; a)$ denote the optimum in the independent Gaussian mixture (3.2). Then

$$\mathrm{OPT}_n^{\mathrm{obs}}(\alpha) = \mathrm{OPT}_n^{\mathrm{ind}}(\alpha; a) \tag{3.3}$$

for every integer $n \geq 1$ and $\alpha \in [0,1]$. Both optima allow randomized Markov kernels.

Proof. The transformation $(X, Z_0) \leftrightarrow (Y, Z_0)$ is a measurable bijection. Pull any kernel in the (X, Z_0)-experiment back through this bijection and denote it by $K(\,\mathrm{d}\delta \mid y, z)$. Moreover,

$$Z_0 \perp (\Theta, Y).$$

Let $K(\,\mathrm{d}\delta \mid y, z)$ be any randomized rule in the revealed-factor experiment. Define a Y-only kernel by

$$K'(\,\mathrm{d}\delta \mid y) = \int_{\mathbb{R}} K(\,\mathrm{d}\delta \mid y, z)\phi(z)\,\mathrm{d}z. \tag{3.4}$$

The integral in (3.4) defines a measurable Markov kernel. Because the joint distribution factorizes as

$$P_{\Theta,Y,Z_0} = P_{\Theta,Y} \otimes N(0,1),$$

the marginal joint law of (Θ, Y, δ) induced by K' is exactly the one induced by K. Indeed, for every measurable set A in the (Θ, Y)-space and every decision vector d, Tonelli's theorem gives

$$\begin{aligned}\int \mathbf{1}\{(\theta,y) \in A\} K(\{d\} \mid y, z)\, P_{\Theta,Y}(\,\mathrm{d}\theta,\,\mathrm{d}y)\,\phi(z)\,\mathrm{d}z \\ = \int \mathbf{1}\{(\theta,y) \in A\} K'(\{d\} \mid y)\, P_{\Theta,Y}(\,\mathrm{d}\theta,\,\mathrm{d}y).\end{aligned}$$

Thus the two kernels induce the same joint law of (Θ, Y, δ), and hence the same FDP and FNP laws. Conversely, the observed experiment can compute Y from (X, Z_0) and apply any Y-only kernel. □

4. The independent Gaussian frontier

This section gives a self-contained reconstruction of the independent-model limit in the conventions of (2.2). It also records the exact finite-n risks of threshold rules used later.

4.1. **Local false-discovery rates and conditional sorting.** For the independent model (3.2), define

$$W_i = \mathbb{P}(\Theta_i = 0 \mid Y_i). \tag{4.1}$$

The Gaussian likelihood ratio gives

$$W_i = \left[1 + \frac{\pi_1}{\pi_0} \exp\left(aY_i - \frac{a^2}{2}\right)\right]^{-1}. \tag{4.2}$$

Because $a > 0$, the map $y \mapsto W(y)$ is a continuous strictly decreasing bijection from $\mathbb{R}$ onto $(0, 1)$. Since Y_i has a positive density everywhere, W_i has an atomless distribution with full support $(0, 1)$.

Lemma 4.1 (Conditional sorting)**.** Condition on $Y = (Y_1, \ldots, Y_n)$ and on the external randomization of a procedure. Among all rejection sets of a fixed size K, rejecting the K smallest values of W_i simultaneously minimizes the conditional expected FDP and the conditional expected FNP.

Proof. Because the action space is finite, every Markov kernel can be represented by a Borel function of Y and an independent uniform random variable. We may therefore condition on both Y and that auxiliary variable. Let $D \subseteq \{1, \ldots, n\}$ be the realized rejection set and $|D| = K$. If $K > 0$, posterior independence gives

$$\mathbb{E}[\mathrm{FDP} \mid Y, D] = \frac{1}{K} \sum_{i \in D} W_i; \tag{4.3}$$

for $K = 0$, the conditional expectation is zero. If $K < n$,

$$\mathbb{E}[\mathrm{FNP} \mid Y, D] = \frac{1}{n - K} \sum_{i \notin D} (1 - W_i); \tag{4.4}$$

for $K = n$, it is zero. For fixed K, replacing a rejected coordinate with larger W_i by a nonrejected coordinate with smaller W_i weakly decreases both displayed quantities. Repeating the exchange yields the claimed sorted set. To define the replacement kernel on every input, including the null set of ties, order the pairs (W_i, i) lexicographically and reject the first K; this is a Borel rule and preserves the original conditional distribution of K. □

Write the order statistics as

$$W_{(1)} \le \cdots \le W_{(n)}.$$

Define, for $k = 0, \ldots, n$,

$$A_n(k) = \begin{cases} 0, & k = 0, \\ \dfrac{1}{k} \displaystyle\sum_{j=1}^{k} W_{(j)}, & k > 0, \end{cases} \tag{4.5}$$

Similarly, define

$$C_n(k) = \begin{cases} \dfrac{1}{n - k} \displaystyle\sum_{j=k+1}^{n} (1 - W_{(j)}), & k < n, \\ 0, & k = n. \end{cases} \tag{4.6}$$

After the sorting reduction, a procedure is determined by a possibly data-dependent and randomized rejection count K, and

$$\mathrm{FDR} = \mathbb{E}[A_n(K)], \qquad \mathrm{FNR} = \mathbb{E}[C_n(K)]. \tag{4.7}$$

4.2. **Quantile limit and convexification.** Let F_W be the distribution function of W_1 and use the generalized-inverse convention

$$Q(u) = \inf\{w \in [0, 1] : F_W(w) \ge u\}, \qquad 0 < u < 1.$$

Full support and continuity give $Q(u) \downarrow 0$ as $u \downarrow 0$ and $Q(u) \uparrow 1$ as $u \uparrow 1$. Define

$$A(u) = \begin{cases} 0, & u = 0, \\ \dfrac{1}{u} \displaystyle\int_0^u Q(s)\, \mathrm{d}s, & u > 0, \end{cases} \tag{4.8}$$

and

$$C(u) = \begin{cases} \dfrac{1}{1-u}\displaystyle\int_u^1 [1 - Q(s)]\,\mathrm{d}s, & u < 1, \\ 0, & u = 1. \end{cases} \tag{4.9}$$

Both functions are continuous on $[0, 1]$.

Lemma 4.2 (Uniform order-statistic risk convergence)**.** Almost surely,

$$\max_{0\le k\le n} |A_n(k) - A(k/n)| \longrightarrow 0, \tag{4.10}$$

and

$$\max_{0\le k\le n} |C_n(k) - C(k/n)| \longrightarrow 0. \tag{4.11}$$

Proof. Let $Q_n(s) = W_{(\lceil ns\rceil)}$ for $s \in (0, 1]$. For $0 < k < n$,

$$A_n(k) = \frac{1}{k/n}\int_0^{k/n} Q_n(s)\,\mathrm{d}s, \qquad C_n(k) = \frac{1}{1-k/n}\int_{k/n}^1 [1 - Q_n(s)]\,\mathrm{d}s.$$

Since the variables are bounded, empirical weak convergence implies convergence in the 1-Wasserstein metric, hence

$$\int_0^1 |Q_n(s) - Q(s)|\,\mathrm{d}s \longrightarrow 0 \tag{4.12}$$

almost surely. For $k/n \ge \varepsilon$,

$$|A_n(k) - A(k/n)| \le \frac{1}{\varepsilon}\int_0^1 |Q_n(s) - Q(s)|\,\mathrm{d}s.$$

For $k \le \varepsilon n$, monotonicity of the quantiles gives

$$A_n(k) \le W_{(\lceil \varepsilon n\rceil)}, \qquad A(k/n) \le Q(\varepsilon).$$

Because F_W is continuous and strictly increasing, the order statistic converges to $Q(\varepsilon)$, so

$$\limsup_{n\to\infty} \max_{k\le \varepsilon n} |A_n(k) - A(k/n)| \le Q(\varepsilon)$$

almost surely. Letting rational $\varepsilon \downarrow 0$ proves (4.10). The proof of (4.11) is symmetric: away from one, use (4.12); for $k \ge (1-\varepsilon)n$, use

$$C_n(k) \le 1 - W_{(\lfloor (1-\varepsilon)n\rfloor)}, \qquad C(k/n) \le 1 - Q(1-\varepsilon),$$

for all sufficiently large n, to obtain a limsup bounded by $1 - Q(1-\varepsilon)$, and let rational $\varepsilon \downarrow 0$. □

Lemma 4.3 (Two-point representation of a constrained boundary)**.** Let $\Gamma \subset \mathbb{R}^2$ be compact, let $H = \mathrm{conv}(\Gamma)$, and suppose

$$b(\alpha) = \min\{y : (x, y) \in H,\ x \le \alpha\}$$

has a nonempty feasible set. Then one minimizer is a convex combination of at most two points of Γ.

Proof. Compactness of H gives attainment. If the affine dimension of H is at most one, the one-dimensional form of Carathéodory's theorem gives the claim. Otherwise, a minimizer $z = (x, y)$ lies on ∂H: if it were an interior point, then $z - \varepsilon(1, 1)$ would remain in H for small $\varepsilon > 0$, remain feasible, and have a smaller second coordinate.

Choose a supporting line of H through z. By Carathéodory's theorem, z is a convex combination of at most three points of Γ. Equality in the supporting-line inequality forces every point with positive weight to lie on that same line. These points are collinear, so their convex combination reduces to a combination of at most two of them. □

Applying lemma 4.3 to Γ_a gives the equivalent two-threshold formula

$$\begin{aligned} B_a(\alpha) = \min\big\{&\lambda r_a(t_1) + (1-\lambda) r_a(t_2) : \lambda \in [0, 1],\ t_1, t_2 \in \overline{\mathbb{R}}, \\ &\lambda q_a(t_1) + (1-\lambda) q_a(t_2) \le \alpha\big\}. \end{aligned} \tag{4.13}$$

Proposition 4.4 (Independent limiting frontier)**.** For every $\alpha \in [0, 1]$,

$$\lim_{n\to\infty} \mathrm{OPT}_n^{\mathrm{ind}}(\alpha; a) = B_a(\alpha). \tag{4.14}$$

Proof. First identify the deterministic curve. For $u \in (0,1)$, the map $t \mapsto \mathbb{P}(Y_1 \geq t)$ is continuous and strictly decreasing from one to zero, so there is a unique t such that

$$\mathbb{P}(Y_1 \geq t) = u.$$

Because $W(Y)$ is strictly decreasing,

$$\{W \leq Q(u)\} = \{Y \geq t\}$$

up to a null set. Therefore

$$\begin{aligned}\int_0^u Q(s)\,\mathrm{d}s &= \mathbb{E}[W\mathbf{1}\{Y \geq t\}] \\ &= \mathbb{P}(\Theta = 0, Y \geq t) = \pi_0\overline{\Phi}(t),\end{aligned}$$

while $u = p_t$. Hence $A(u) = q_a(t)$. Similarly,

$$\begin{aligned}\int_u^1 [1 - Q(s)]\,\mathrm{d}s &= \mathbb{E}[(1-W)\mathbf{1}\{Y < t\}] \\ &= \mathbb{P}(\Theta = 1, Y < t) = \pi_1\Phi(t-a),\end{aligned}$$

so $C(u) = r_a(t)$. The endpoint values are

$$(A(0), C(0)) = (0, \pi_1), \qquad (A(1), C(1)) = (\pi_0, 0).$$

Thus $\{(A(u), C(u)) : 0 \leq u \leq 1\} = \Gamma_a$.

For the lower bound, write $v_n = \mathrm{OPT}_n^{\mathrm{ind}}(\alpha; a)$, choose a subsequence n_m along which $v_{n_m} \to \liminf_{n\to\infty} v_n$, and choose a feasible procedure at dimension n_m whose FNR is at most $v_{n_m} + m^{-1}$. By lemma 4.1, replace it by a sorted rule without increasing either risk. Write $\mathrm{FDR}_m, \mathrm{FNR}_m$ for its risks and let K_m be its rejection count. Compactness gives a further subsequence such that

$$K_m/n_m \Rightarrow T$$

for some $T \in [0,1]$. Moreover,

$$\begin{aligned}|\mathrm{FDR}_m - \mathbb{E}[A(K_m/n_m)]| &\leq \mathbb{E}\max_{0\leq k\leq n_m} |A_{n_m}(k) - A(k/n_m)| \longrightarrow 0, \\ |\mathrm{FNR}_m - \mathbb{E}[C(K_m/n_m)]| &\leq \mathbb{E}\max_{0\leq k\leq n_m} |C_{n_m}(k) - C(k/n_m)| \longrightarrow 0.\end{aligned}$$

Here dominated convergence applies because both maxima are bounded by one. Since A and C are bounded and continuous, bounded-continuous convergence under $K_m/n_m \Rightarrow T$ yields

$$(\mathrm{FDR}_m, \mathrm{FNR}_m) \longrightarrow \big(\mathbb{E}[A(T)], \mathbb{E}[C(T)]\big). \tag{4.15}$$

Feasibility gives $\mathbb{E}[A(T)] \leq \alpha$. The pair in (4.15) is a barycenter of Γ_a, hence belongs to the compact convex set K_a, and therefore $\mathbb{E}[C(T)] \geq B_a(\alpha)$. Since the sorted near-minimizer remains feasible,

$$v_{n_m} \leq \mathrm{FNR}_m \leq v_{n_m} + m^{-1}.$$

Near-optimality now gives $\liminf_{n\to\infty} v_n \geq B_a(\alpha)$.

For the upper bound, choose the two-threshold minimizer in (4.13) and use one global coin to select its threshold. Lemma 4.5 gives finite-n FDR at most the corresponding convex combination of the q_a-coordinates, hence at most α, while its FNR converges to the corresponding convex combination of the r_a-coordinates, namely $B_a(\alpha)$. Thus $\limsup_{n\to\infty} v_n \leq B_a(\alpha)$, completing the proof. □

Lemma 4.5 (Exact finite-sample threshold risks)**.** For every integer $n \geq 1$ and $t \in \overline{\mathbb{R}}$, let $\delta_i^{t,n} = \mathbf{1}\{Y_i \geq t\}$, interpreting the endpoints as reject none and reject all, and write $\mathrm{FDR}_n(t)$ and $\mathrm{FNR}_n(t)$ for its risks. With $p_{+\infty} = 0$ and $p_{-\infty} = 1$,

$$\mathrm{FDR}_n(t) = q_a(t)\,[1 - (1-p_t)^n] \leq q_a(t), \tag{4.16}$$

and

$$\mathrm{FNR}_n(t) = r_a(t)\,[1 - p_t^n] \leq r_a(t). \tag{4.17}$$

Proof. For finite t, the four events obtained by crossing the label with rejection or nonrejection form an i.i.d. multinomial classification. In particular,

$$V \mid R = r \sim \mathrm{Bin}(r, q_a(t)), \qquad r > 0,$$

so

$$\mathbb{E}[V/R \mid R = r] = q_a(t).$$

Multiplying by $\mathbb{P}(R > 0) = 1 - (1 - p_t)^n$ proves (4.16). Similarly,

$$U \mid S = s \sim \mathrm{Bin}(s, r_a(t)), \qquad s > 0,$$

and $\mathbb{P}(S > 0) = 1 - p_t^n$, which proves (4.17). At $t = +\infty$, the risks are $(0, \pi_1)$; at $t = -\infty$, they are $(\pi_0, 0)$. These agree with the displayed formulas. □

5. Universal converse

For brevity, write

$$\mathrm{OPT}_n^{\mathrm{unobs}}(\alpha) = \mathrm{OPT}_n(\alpha; \pi_1, \mu, \rho).$$

Proposition 5.1 (Oracle lower bound)**.** For every integer $n \geq 1$ and $\alpha \in [0, 1]$,

$$\mathrm{OPT}_n^{\mathrm{unobs}}(\alpha) \geq \mathrm{OPT}_n^{\mathrm{obs}}(\alpha) = \mathrm{OPT}_n^{\mathrm{ind}}(\alpha; a). \tag{5.1}$$

Consequently,

$$\liminf_{n\to\infty} \mathrm{OPT}_n^{\mathrm{unobs}}(\alpha) \geq B_a(\alpha). \tag{5.2}$$

Proof. Every X-only Markov kernel can be used in the augmented experiment (X, Z_0) by ignoring Z_0. The feasible set in the observed-factor experiment is therefore at least as large, which gives the first inequality in (5.1). The equality is proposition 3.2. Taking lower limits in the pointwise inequality and applying proposition 4.4 gives

$$\liminf_{n\to\infty} \mathrm{OPT}_n^{\mathrm{unobs}}(\alpha) \geq \liminf_{n\to\infty} \mathrm{OPT}_n^{\mathrm{ind}}(\alpha; a) = B_a(\alpha).$$

No separability, symmetry, threshold, or deterministic restriction is imposed on the original procedure. □

6. Factor adjustment and quantitative stability

6.1. **Full-sample factor estimation.** Because $\mathbb{E}[\overline{X}_n \mid Z_0] = \mu\pi_1 + \sqrt{\rho}\, Z_0$, define the sample-mean factor estimate

$$\widehat{Z}_0 = \frac{\overline{X}_n - \mu\pi_1}{\sqrt{\rho}}, \qquad \overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i, \tag{6.1}$$

and the adjusted observations

$$\widehat{Y}_i = \frac{X_i - \sqrt{\rho}\,\widehat{Z}_0}{\sqrt{1-\rho}} = \frac{X_i - \overline{X}_n + \mu\pi_1}{\sqrt{1-\rho}}. \tag{6.2}$$

The last expression makes implementation explicit: the adjusted vector is a function of X and the known parameters only. This is a full-sample, rather than sample-split, estimate, so its error is not independent of any tested coordinate. The proof uses instead the exact structural fact that data reuse creates one common scalar displacement across all coordinates.

Write

$$\overline{\Theta}_n = \frac{1}{n}\sum_{i=1}^n \Theta_i, \qquad \overline{Z}_n = \frac{1}{n}\sum_{i=1}^n Z_i, \qquad \overline{Y}_n = \frac{1}{n}\sum_{i=1}^n Y_i.$$

Lemma 6.1 (Exact common displacement and concentration)**.** Let

$$\Delta_n = \sqrt{\frac{\rho}{1-\rho}}(\widehat{Z}_0 - Z_0). \tag{6.3}$$

Then, for every coordinate,

$$\widehat{Y}_i = Y_i - \Delta_n, \tag{6.4}$$

and

$$\Delta_n = a(\overline{\Theta}_n - \pi_1) + \overline{Z}_n = \overline{Y}_n - a\pi_1. \tag{6.5}$$

Moreover,

$$\mathbb{E}[(\widehat{Z}_0 - Z_0)^2] = \frac{\mu^2\pi_0\pi_1 + 1 - \rho}{n\rho}, \tag{6.6}$$

$$\mathbb{E}[\Delta_n^2] = \frac{1 + a^2\pi_0\pi_1}{n}, \tag{6.7}$$

and, with

$$h_a = 1 + \frac{a^2}{4}, \tag{6.8}$$

one has for every $\eta > 0$,

$$\mathbb{P}(|\Delta_n| \geq \eta) \leq 2\exp\left(-\frac{n\eta^2}{2h_a}\right). \tag{6.9}$$

Equivalently, for every $x > 0$,

$$\mathbb{P}(|\widehat{Z}_0 - Z_0| \geq x) \leq 2\exp\left[-\frac{n\rho x^2}{2(1 - \rho + \mu^2/4)}\right]. \tag{6.10}$$

Proof. Averaging (2.1) gives

$$\overline{X}_n = \mu\overline{\Theta}_n + \sqrt{\rho}\, Z_0 + \sqrt{1-\rho}\,\overline{Z}_n.$$

Substitution into (6.1) yields

$$\widehat{Z}_0 - Z_0 = \frac{\mu(\overline{\Theta}_n - \pi_1) + \sqrt{1-\rho}\,\overline{Z}_n}{\sqrt{\rho}}. \tag{6.11}$$

The two terms are independent and centered, which proves (6.6). Equations (6.4), (6.5), and (6.7) follow directly.

For the tail bound, Hoeffding's lemma gives

$$\mathbb{E}\exp\{sa(\Theta_i - \pi_1)\} \leq \exp\left(\frac{s^2a^2}{8}\right).$$

Combining this with the standard normal moment-generating function, and using independence across i, yields

$$\begin{aligned}\mathbb{E}e^{s\Delta_n} &= \prod_{i=1}^{n} \mathbb{E}\exp\left\{\frac{s}{n}[a(\Theta_i - \pi_1) + Z_i]\right\} \\ &\leq \exp\left\{\frac{s^2}{2n}\left(1 + \frac{a^2}{4}\right)\right\} = \exp\left(\frac{s^2h_a}{2n}\right).\end{aligned}$$

The Chernoff bound gives (6.9); applying the scaling in (6.3) gives (6.10). □

6.2. **Threshold-crossing counts.** Fix a finite threshold t. Couple the observed-factor oracle rule and the factor-adjusted rule by

$$d_i = \mathbf{1}\{Y_i \geq t\}, \qquad \widehat{d}_i = \mathbf{1}\{\widehat{Y}_i \geq t\} = \mathbf{1}\{Y_i \geq t + \Delta_n\}. \tag{6.12}$$

Let

$$D = \sum_{i=1}^{n} |\widehat{d}_i - d_i|, \tag{6.13}$$

and let

$$D_0 = \sum_{i=1}^{n} (1 - \Theta_i)|\widehat{d}_i - d_i|, \qquad D_1 = \sum_{i=1}^{n} \Theta_i|\widehat{d}_i - d_i|.$$

Lemma 6.2 (Pathwise crossing bound)**.** Let $M = (2\pi)^{-1/2}$. For every deterministic $\eta > 0$,

$$\frac{D}{n} \leq \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}\{|Y_i - t| \leq \eta\} + \mathbf{1}\{|\Delta_n| > \eta\} \tag{6.14}$$

pathwise. Consequently,

$$\mathbb{E}[D/n] \leq 2M\eta + 2\exp\left(-\frac{n\eta^2}{2h_a}\right), \tag{6.15}$$

$$\mathbb{E}[D_0/n] \leq 2\pi_0 M\eta + 2\exp\left(-\frac{n\eta^2}{2h_a}\right), \tag{6.16}$$

and

$$\mathbb{E}[D_1/n] \le 2\pi_1 M\eta + 2\exp\left(-\frac{n\eta^2}{2h_a}\right). \tag{6.17}$$

For every $x > 0$,

$$\mathbb{P}\left(\frac{D}{n} > 2M\eta + x\right) \le e^{-2nx^2} + 2\exp\left(-\frac{n\eta^2}{2h_a}\right). \tag{6.18}$$

Proof. On the event $|\Delta_n| \le \eta$, a decision can change only if Y_i lies between t and $t + \Delta_n$, hence only if $|Y_i - t| \le \eta$. This proves (6.14) without conditioning and without any independence between Δ_n and the tested coordinates.

The same pathwise inclusion, restricted by the label, gives

$$\frac{D_0}{n} \le \frac{1}{n}\sum_{i=1}^{n}(1-\Theta_i)\mathbf{1}\{|Y_i - t| \le \eta\} + \mathbf{1}\{|\Delta_n| > \eta\},$$

and the analogous inequality with D_1 and Θ_i.

The marginal density of Y_i is

$$f(y) = \pi_0\phi(y) + \pi_1\phi(y - a),$$

so $\sup_y f(y) \le M$. Moreover,

$$\mathbb{E}[(1-\Theta_i)\mathbf{1}\{|Y_i - t| \le \eta\}] = \pi_0\int_{t-\eta}^{t+\eta}\phi(y)\,\mathrm{d}y \le 2\pi_0 M\eta,$$

and the analogous alternative-label expectation is at most $2\pi_1 M\eta$. Taking expectations in the three pathwise bounds and applying (6.9) proves (6.15)–(6.17). For (6.18), apply Hoeffding's inequality to the i.i.d. indicators $\mathbf{1}\{|Y_i - t| \le \eta\}$, whose mean is at most $2M\eta$, and then use a union bound with (6.9). □

For $n \ge 2$, choose

$$\eta_n = \sqrt{\frac{4h_a \log n}{n}} \tag{6.19}$$

and define

$$d_n = 4M\sqrt{\frac{h_a \log n}{n}} + 2n^{-2}. \tag{6.20}$$

Then $\mathbb{E}[D/n] \le d_n$ and $d_n \to 0$.

6.3. **Random-ratio stability.** The estimate $D/n \to 0$ does not by itself imply convergence of FDP or FNP because either ratio is unstable when its denominator is small. At a fixed finite threshold, however, $p_t, s_t \in (0, 1)$, and binomial lower-tail bounds keep the oracle rejection and acceptance counts of order n with exponentially high probability.

Lemma 6.3 (Deterministic ratio perturbation)**.** Let (V, R) and $(\widehat{V}, \widehat{R})$ be the false-discovery and rejection counts of two decision vectors that differ in at most D coordinates. If $R > 0$, then

$$\left|\frac{\widehat{V}}{\widehat{R} \vee 1} - \frac{V}{R}\right| \le \frac{2D}{R}. \tag{6.21}$$

If $S > 0$, the analogous inequality holds for FNP after replacing (V, R) by (U, S).

Proof. If $D \ge R/2$, the right side is at least one and the conclusion follows because both ratios lie in $[0, 1]$. If $D < R/2$, then $\widehat{R} \ge R - D > R/2$, and

$$\begin{aligned}\left|\frac{V}{R} - \frac{\widehat{V}}{\widehat{R}}\right| &\le \frac{|V - \widehat{V}|}{R} + \frac{\widehat{V}|R - \widehat{R}|}{R\widehat{R}} \\ &\le \frac{D}{R} + \frac{D}{R}.\end{aligned}$$

The FNP statement is identical. □

Lemma 6.4 (Expected FDP and FNP stability)**.** For $n \geq 2$ and a finite threshold t, let the oracle and plug-in decisions be coupled as in (6.12). Define

$$e_n^R(t) = \frac{4d_n}{p_t} + e^{-np_t/8} \tag{6.22}$$

and

$$e_n^S(t) = \frac{4d_n}{s_t} + e^{-ns_t/8}. \tag{6.23}$$

Then

$$\mathbb{E}\left|\mathrm{FDP}_{\mathrm{plug}} - \mathrm{FDP}_{\mathrm{oracle}}\right| \leq e_n^R(t), \tag{6.24}$$

and

$$\mathbb{E}\left|\mathrm{FNP}_{\mathrm{plug}} - \mathrm{FNP}_{\mathrm{oracle}}\right| \leq e_n^S(t). \tag{6.25}$$

At $t = \pm\infty$, define both error terms to be zero because the endpoint rules are implemented directly.

Proof. For the oracle threshold, $R \sim \mathrm{Bin}(n, p_t)$, so the Chernoff bound gives

$$\mathbb{P}(R < np_t/2) \leq e^{-np_t/8}.$$

On $\{R \geq np_t/2\}$, lemma 6.3 gives

$$|\,\mathrm{FDP}_{\mathrm{plug}} - \mathrm{FDP}_{\mathrm{oracle}}\,| \leq \frac{4D}{np_t}.$$

The ratio difference is at most one on the complementary event. Taking expectations and using $\mathbb{E}[D/n] \leq d_n$ proves (6.24). The proof of (6.25) uses $S \sim \mathrm{Bin}(n, s_t)$ and is identical. The events $R = 0$ and $S = 0$ are contained in the respective Chernoff remainders. □

7. The calibrated factor-adjusted procedure

By (4.13), the minimum is attained by two extended thresholds. Fix

$$\lambda \in [0, 1], \qquad t_1, t_2 \in \overline{\mathbb{R}} \tag{7.1}$$

such that, with

$$\beta = \lambda q_a(t_1) + (1 - \lambda)q_a(t_2), \tag{7.2}$$

one has

$$\beta \leq \alpha, \qquad B_a(\alpha) = \lambda r_a(t_1) + (1 - \lambda)r_a(t_2). \tag{7.3}$$

For $0 < \alpha < \pi_0$, the optimizer may be chosen with $\beta = \alpha$; this is proved in proposition 8.2.

For the selected two-threshold mixture, define

$$e_n^R = \lambda e_n^R(t_1) + (1 - \lambda)e_n^R(t_2), \tag{7.4}$$

$$e_n^S = \lambda e_n^S(t_1) + (1 - \lambda)e_n^S(t_2), \tag{7.5}$$

where the endpoint convention of lemma 6.4 is used. For every fixed optimizer,

$$e_n^R \longrightarrow 0, \qquad e_n^S \longrightarrow 0. \tag{7.6}$$

Definition 7.1 (Calibrated factor-adjusted rule)**.** For $0 < \alpha < \pi_0$ and $n \geq 2$, perform the following steps.

(i) Compute $\widehat{Z}_0$ from (6.1) and the adjusted observations $\widehat{Y}_i$ from (6.2).

(ii) Draw one external Bernoulli variable J, independent of all model variables, with $\mathbb{P}(J = 1) = \lambda$. Let

$$T = \begin{cases} t_1, & J = 1, \\ t_2, & J = 0 \end{cases}.$$

If $T \in \mathbb{R}$, set

$$\delta_i^{\text{base}} = \mathbf{1}\{\widehat{Y}_i \geq T\}.$$

Ties are rejected. If $T = +\infty$, reject none; if $T = -\infty$, reject all.

(iii) Set

$$\gamma_n = \min\left\{1, \frac{\alpha}{\beta + e_n^R}\right\}, \tag{7.7}$$

with $\gamma_n = 1$ if the denominator is zero. Draw a second global coin $\xi_n \sim \text{Bernoulli}(\gamma_n)$, independent of J and all model variables, and return

$$\delta_i^{(n)} = \xi_n \delta_i^{\text{base}}. \tag{7.8}$$

Thus $\xi_n = 0$ invokes the reject-none rule.

For $n = 1$ and $0 < \alpha < \pi_0$, reject none. For $\alpha = 0$, reject none for every n. For $\alpha \geq \pi_0$, reject all for every n.

Remark 7.2 (Global versus coordinatewise randomization). Both randomizations in definition 7.1 are global. One coin selects the threshold for the entire vector; the other replaces the entire base decision vector by the reject-none rule. Independent coordinatewise thinning would not, in general, scale FDR linearly and is not used. The second coin is a nonasymptotic proof device based on a conservative analytic bound; $\gamma_n \to 1$, but the bound can be numerically conservative at moderate n.

Proposition 7.3 (Nonasymptotic FDR control)**.** The procedure in definition 7.1 satisfies

$$\text{FDR}(\delta^{(n)}) \leq \alpha \tag{7.9}$$

for every n.

Proof. Let δ^{oracle} be the observed-factor global mixture that uses the same threshold-selection coin J. By lemma 4.5,

$$\begin{aligned} \text{FDR}(\delta^{\text{oracle}}) &\leq \lambda q_a(t_1) + (1 - \lambda) q_a(t_2) \\ &= \beta. \end{aligned} \tag{7.10}$$

Conditioning on J and applying lemma 6.4 gives

$$\text{FDR}(\delta^{\text{base}}) \leq \beta + e_n^R. \tag{7.11}$$

Pathwise, $\text{FDP}(\xi_n \delta^{\text{base}}) = \xi_n \,\text{FDP}(\delta^{\text{base}})$. Since ξ_n is independent of the base procedure,

$$\begin{aligned} \text{FDR}(\delta^{(n)}) &= \gamma_n \,\text{FDR}(\delta^{\text{base}}) \\ &\leq \gamma_n(\beta + e_n^R) \leq \alpha. \end{aligned}$$

For $n = 1$ and for $\alpha = 0$, the separately defined reject-none rule has FDR zero. For $\alpha \geq \pi_0$, reject-all has FDR $\pi_0 \leq \alpha$. Thus every separately defined case is feasible. □

Proposition 7.4 (Asymptotic FNR optimality)**.** For the procedure in definition 7.1,

$$\limsup_{n \to \infty} \text{FNR}(\delta^{(n)}) \leq B_a(\alpha). \tag{7.12}$$

Proof. The endpoint cases are immediate from the separately defined rules and proposition 8.1, so assume $0 < \alpha < \pi_0$. Again using the observed-factor mixture and lemma 4.5,

$$\text{FNR}(\delta^{\text{oracle}}) \leq \lambda r_a(t_1) + (1 - \lambda) r_a(t_2) = B_a(\alpha). \tag{7.13}$$

Conditioning on J and applying the FNP part of lemma 6.4 gives

$$\text{FNR}(\delta^{\text{base}}) \leq B_a(\alpha) + e_n^S. \tag{7.14}$$

The zero-denominator convention gives the pathwise identity

$$\text{FNP}(\xi_n \delta^{\text{base}}) = \xi_n \,\text{FNP}(\delta^{\text{base}}) + (1 - \xi_n) \frac{1}{n} \sum_{i=1}^{n} \Theta_i.$$

Taking expectations and using independence therefore gives

$$\begin{aligned}\mathrm{FNR}(\delta^{(n)}) &= \gamma_n \,\mathrm{FNR}(\delta^{\mathrm{base}}) + (1-\gamma_n)\pi_1 \\ &\le \gamma_n[B_a(\alpha) + e_n^S] + (1-\gamma_n)\pi_1. \end{aligned} \tag{7.15}$$

It remains to note that $\gamma_n \to 1$. If $\beta < \alpha$, then $\beta + e_n^R < \alpha$ for all sufficiently large n, so $\gamma_n = 1$. If $\beta = \alpha > 0$, then

$$\gamma_n = \frac{\alpha}{\alpha + e_n^R} \longrightarrow 1.$$

Together with (7.6), this proves (7.12). □

Proof of theorem 2.2. For $0 < \alpha < \pi_0$, proposition 5.1 gives

$$\liminf_{n\to\infty} \mathrm{OPT}_n(\alpha;\pi_1,\mu,\rho) \ge B_a(\alpha),$$

while feasibility of $\delta^{(n)}$ gives

$$\mathrm{OPT}_n(\alpha;\pi_1,\mu,\rho) \le \mathrm{FNR}(\delta^{(n)}),$$

and propositions 7.3 and 7.4 give

$$\limsup_{n\to\infty} \mathrm{OPT}_n(\alpha;\pi_1,\mu,\rho) \le B_a(\alpha).$$

Thus the limit exists and equals $B_a(\alpha)$. The endpoint identities are proved in proposition 8.1 below. □

7.1. **Unknown parameters.** The procedure of definition 7.1 uses the true (π_1,μ,ρ). We now show that consistent estimates may be substituted, at the price of replacing finite-sample FDR control by asymptotic FDR control. Fix $\alpha \in (0,1)$. Let $\widehat{\pi}_1, \widehat{\mu}, \widehat{\rho}$ be measurable functions of X such that

$$(\widehat{\pi}_1, \widehat{\mu}, \widehat{\rho}) \longrightarrow (\pi_1,\mu,\rho) \quad \text{in probability.} \tag{7.16}$$

We do not construct such estimators here. Conditional on Z_0, the coordinates are i.i.d. from $\pi_0 N(c, 1-\rho) + \pi_1 N(c+\mu, 1-\rho)$ with $c = \sqrt{\rho}\, Z_0$, so the three targets are parameters of the within-vector distribution and do not involve the unobservable c.

For $\vartheta = (\varpi, b) \in \Xi = (0,1)\times(0,\infty)$, let $p_\vartheta, s_\vartheta, q_\vartheta, r_\vartheta$ be defined by (2.6)–(2.10) with (π_1,π_0,a) replaced by $(\varpi, 1-\varpi, b)$, and let $B_\vartheta(\alpha)$ be defined by (2.11) with the corresponding hull K_ϑ. Thus $\vartheta_0 = (\pi_1, a)$ gives the original objects. On $\overline{\mathbb{R}}$ use the metric $\mathsf{d}(s,t) = |\arctan s - \arctan t|$, with $\arctan(\pm\infty) = \pm\pi/2$; it induces the compact topology of $\overline{\mathbb{R}}$.

Lemma 7.5 (Parameter continuity and measurable near-optimizers)**.** (i) The map $(\vartheta, t) \mapsto (q_\vartheta(t), r_\vartheta(t))$ is continuous on $\Xi \times \overline{\mathbb{R}}$. (ii) For every $\vartheta_1 \in \Xi$, $\limsup_{\vartheta\to\vartheta_1} B_\vartheta(\alpha) \le B_{\vartheta_1}(\alpha)$. (iii) For every $\varepsilon > 0$ there is a Borel map $\vartheta \mapsto (\lambda_\vartheta, t_{1,\vartheta}, t_{2,\vartheta})$ from Ξ to $[0,1]\times\overline{\mathbb{R}}^2$ such that, for every $\vartheta \in \Xi$,

$$\begin{aligned} &\lambda_\vartheta q_\vartheta(t_{1,\vartheta}) + (1-\lambda_\vartheta) q_\vartheta(t_{2,\vartheta}) \le \alpha, \\ &\lambda_\vartheta r_\vartheta(t_{1,\vartheta}) + (1-\lambda_\vartheta) r_\vartheta(t_{2,\vartheta}) \le B_\vartheta(\alpha) + \varepsilon. \end{aligned} \tag{7.17}$$

Proof. (i) Continuity on $\Xi\times\mathbb{R}$ is clear. Let $N = [\varpi_-, \varpi_+]\times[b_-, b_+] \subset \Xi$ and $\vartheta = (\varpi, b) \in N$. The elementary bounds used in the proof of lemma 9.7 give, for $t \ge 0$,

$$0 \le q_\vartheta(t) \le \frac{\overline{\Phi}(t)}{\varpi_- \overline{\Phi}(t-b_-)}, \qquad 0 \le \varpi - r_\vartheta(t) \le \frac{\overline{\Phi}(t-b_+)}{\Phi(t)}.$$

For $t \le 0$, the identity $1 - \varpi - q_\vartheta(t) = (1-\varpi)\varpi[\Phi(t) - \Phi(t-b)]/p_\vartheta(t)$ and $p_\vartheta(t) \ge (1-\varpi_+)/2$ give

$$0 \le 1 - \varpi - q_\vartheta(t) \le \frac{2\Phi(t)}{1-\varpi_+}, \qquad 0 \le r_\vartheta(t) \le \frac{\Phi(t-b_-)}{(1-\varpi_+)\Phi(t)}.$$

All four bounds tend to zero as $t \to \pm\infty$, uniformly on N. Since the endpoint values $(0,\varpi)$ and $(1-\varpi, 0)$ are continuous in ϑ, this proves (i).

(ii) Write $\vartheta_1 = (\varpi_1, b_1)$. Let $(q^*, r^*) \in K_{\vartheta_1}$ attain $B_{\vartheta_1}(\alpha)$, and write it as a finite convex combination $\sum_k w_k (q_{\vartheta_1}(t_k), r_{\vartheta_1}(t_k))$ with $t_k \in \overline{\mathbb{R}}$. Fix $\eta \in (0,1)$, and for $\vartheta = (\varpi, b)$ consider

$$P_\vartheta = (1-\eta)\sum_k w_k \big(q_\vartheta(t_k), r_\vartheta(t_k)\big) + \eta(0,\varpi) \in K_\vartheta.$$

By (i), as $\vartheta \to \vartheta_1$ its first coordinate tends to $(1-\eta)q^* \le (1-\eta)\alpha < \alpha$, and its second tends to $(1-\eta)r^* + \eta\varpi_1 \le B_{\vartheta_1}(\alpha) + \eta$. Hence $\limsup_{\vartheta\to\vartheta_1} B_\vartheta(\alpha) \le B_{\vartheta_1}(\alpha) + \eta$. Let $\eta \downarrow 0$.

(iii) For $x = (\lambda, t_1, t_2) \in [0,1] \times \overline{\mathbb{R}}^2$, write $Q_\vartheta(x)$ and $\mathcal{R}_\vartheta(x)$ for the two mixtures on the left of (7.17). By (i), both are continuous in ϑ for fixed x and in x for fixed ϑ. Let $\mathcal{X}$ be the countable set of such x with $\lambda \in \mathbb{Q}$ and $t_1, t_2 \in \mathbb{Q} \cup \{\pm\infty\}$, enumerated as $x_1, x_2, \ldots$. Fix ϑ, $\varepsilon' > 0$, and a two-point minimizer x^* in (4.13) for $B_\vartheta(\alpha)$. We claim that some $x \in \mathcal{X}$ has $Q_\vartheta(x) \le \alpha$ and $\mathcal{R}_\vartheta(x) < B_\vartheta(\alpha) + \varepsilon'$. If $Q_\vartheta(x^*) = 0$, every threshold with positive weight equals $+\infty$, because $q_\vartheta > 0$ on $[-\infty, \infty)$; then $x = (1, +\infty, +\infty) \in \mathcal{X}$ has the same two coordinates. Otherwise, replace each positive-weight threshold $t_j < +\infty$ by $t_j + \eta$ if it is finite and by $-1/\eta$ if it equals $-\infty$. At least one threshold changes, and strict monotonicity of q_ϑ on $\overline{\mathbb{R}}$ (proposition 8.2) makes the new mixture strictly feasible; continuity on $\overline{\mathbb{R}}$ keeps its second coordinate below $B_\vartheta(\alpha) + \varepsilon'/2$ for small $\eta > 0$. Density of $\mathcal{X}$ and continuity in x give the claim. Every $x \in \mathcal{X}$ with $Q_\vartheta(x) \le \alpha$ yields a feasible point of K_ϑ, so the claim for all $\varepsilon' > 0$ gives

$$B_\vartheta(\alpha) = \inf_{k\ge 1}\big\{\mathcal{R}_\vartheta(x_k) + \infty \cdot \mathbf{1}\{Q_\vartheta(x_k) > \alpha\}\big\},$$

a countable infimum of Borel functions of ϑ. Hence $k(\vartheta) = \min\{k : Q_\vartheta(x_k) \le \alpha,\ \mathcal{R}_\vartheta(x_k) \le B_\vartheta(\alpha) + \varepsilon\}$ is finite, by the claim with $\varepsilon' = \varepsilon$, and Borel. The map $\vartheta \mapsto x_{k(\vartheta)}$ is the required one. □

Definition 7.6 (Plug-in factor-adjusted rule)**.** Fix positive sequences $\varepsilon_n \to 0$ and $c_n \in (0, 1/4]$ with

$$c_n \to 0, \qquad c_n\sqrt{n/\log n} \to \infty, \tag{7.18}$$

for example $c_n = \min\{1/4, n^{-1/3}\}$. Let $\mathcal{E}_n = \{\widehat{\pi}_1 \in (0,1),\ \widehat{\mu} > 0,\ \widehat{\rho} \in (0,1)\}$. On $\mathcal{E}_n^c$, reject none. On $\mathcal{E}_n$, perform the following steps.

(i) Set $\widehat{\vartheta} = (\widehat{\pi}_1, \widehat{\mu}/\sqrt{1-\widehat{\rho}})$ and $\widehat{Y}_i = (X_i - \overline{X}_n + \widehat{\mu}\widehat{\pi}_1)/\sqrt{1-\widehat{\rho}}$.

(ii) Let $(\widehat{\lambda}, \widehat{t}_1, \widehat{t}_2)$ be the value at $\widehat{\vartheta}$ of the Borel map of lemma 7.5(iii) with $\varepsilon = \varepsilon_n$.

(iii) Using an external uniform variable independent of all model variables, draw $J \in \{1,2\}$ with $\mathbb{P}(J = 1 \mid X) = \widehat{\lambda}$, and put $T = \widehat{t}_J$.

(iv) If $T = +\infty$, reject none; if $T = -\infty$, reject all. Otherwise let $\widehat{R} = \#\{i : \widehat{Y}_i \ge T\}$. If $\widehat{R} < c_n n$, reject none; else if $n - \widehat{R} < c_n n$, reject all; else reject exactly the coordinates with $\widehat{Y}_i \ge T$.

Every step uses only X and external randomization. The count truncation in step (iv) is a proof device: it keeps both random denominators at least $c_n n$ when the selected threshold is data dependent.

Theorem 7.7 (Plug-in parameters)**.** Fix the true parameters $\pi_1 \in (0,1)$, $\mu > 0$, $\rho \in (0,1)$ and a level $\alpha \in (0,1)$. Let $\widehat{\pi}_1, \widehat{\mu}, \widehat{\rho}$ be measurable functions of X satisfying (7.16) under these parameters, and let $\delta^{(n)}$ be the rule of definition 7.6, with sequences ε_n and c_n as there. Then

$$\limsup_{n\to\infty} \mathrm{FDR}(\delta^{(n)}) \le \alpha, \qquad \lim_{n\to\infty} \mathrm{FNR}(\delta^{(n)}) = B_a(\alpha). \tag{7.19}$$

Moreover, every sequence of procedures with $\limsup_n \mathrm{FDR}(\delta_n) \le \alpha$ satisfies $\liminf_n \mathrm{FNR}(\delta_n) \ge B_a(\alpha)$.

Proof. Step 1: oracle scale. Let $Y_i = a\Theta_i + Z_i$ be the oracle residuals of (3.1), put $\sigma = \sqrt{1-\rho}$ and, on $\mathcal{E}_n$, $\widehat{\sigma} = \sqrt{1-\widehat{\rho}}$. Since $X_i - \overline{X}_n = \sigma(Y_i - \overline{Y}_n)$,

$$\widehat{Y}_i \ge t \iff Y_i \ge \tau(t) := \widehat{\kappa}t + \widehat{m}, \qquad \widehat{\kappa} = \frac{\widehat{\sigma}}{\sigma}, \quad \widehat{m} = \overline{Y}_n - \frac{\widehat{\mu}\widehat{\pi}_1}{\sigma},$$

with $\tau(\pm\infty) = \pm\infty$. (Here $\widehat{\kappa}$ and $\widehat{m}$ are a scale and a shift local to this proof; $\widehat{m}$ is not an estimate of the second coordinate of ϑ.) By (7.16) and the law of large numbers $\overline{Y}_n \to a\pi_1 = \mu\pi_1/\sigma$, we have $\mathbb{P}(\mathcal{E}_n) \to 1$, $\widehat{\vartheta} \to \vartheta_0$, $\widehat{\kappa} \to 1$, and $\widehat{m} \to 0$ in probability. For $\kappa \in [1/2, 2]$, $m \in \mathbb{R}$, and $t \in \overline{\mathbb{R}}$, every s between t and κt has $1 + s^2 \ge 1 + t^2/4$, so, since arctan is 1-Lipschitz and $\max_t |t|/(1 + t^2/4) = 1$,

$$\mathsf{d}(\kappa t + m, t) \le |m| + \Big|\int_t^{\kappa t} \frac{\mathrm{d}s}{1+s^2}\Big| \le |m| + \frac{|\kappa - 1||t|}{1 + t^2/4} \le |m| + |\kappa - 1|.$$

Step 2: uniform empirical control. For $t \in \overline{\mathbb{R}}$, let $R_n(t) = \sum_i \mathbf{1}\{Y_i \geq t\}$, $V_n(t) = \sum_i (1-\Theta_i)\mathbf{1}\{Y_i \geq t\}$, and $U_n(t) = \sum_i \Theta_i \mathbf{1}\{Y_i < t\}$, and let ζ_n be the supremum over $t \in \overline{\mathbb{R}}$ of

$$\max\Big\{\Big|\frac{R_n(t)}{n} - p_t\Big|,\ \Big|\frac{V_n(t)}{n} - \pi_0\overline{\Phi}(t)\Big|,\ \Big|\frac{U_n(t)}{n} - \pi_1\Phi(t-a)\Big|\Big\},$$

with the limiting values at $t = \pm\infty$. (This uniform deviation is unrelated to the common displacement Δ_n of lemma 6.1. Each empirical mean is left-continuous in t and each mean is continuous, so the supremum may be taken over $\mathbb{Q} \cup \{\pm\infty\}$ and ζ_n is measurable.) Each of the three empirical means $G_n(t)$ is an average of i.i.d. indicators that are monotone in t, and its mean $g(t)$ is continuous and strictly monotone on $\overline{\mathbb{R}}$ with range of length at most one. For $k \geq 1$, choose $-\infty = t_0 < \cdots < t_k = +\infty$ splitting the range of g into k equal parts. On each $[t_{j-1}, t_j]$, monotonicity bounds $|G_n - g|$ by the larger of the two endpoint deviations plus $1/k$. Hoeffding's inequality at the $k+1$ grid points therefore gives $\mathbb{P}(\sup_t |G_n(t) - g(t)| > x + 1/k) \leq 2(k+1)e^{-2nx^2}$. With $k = \lceil\sqrt{n}\rceil$, $x = \sqrt{\log n/n}$, and a union bound over the three families, for $n \geq 3$,

$$\mathbb{P}\Big(\zeta_n > 2\sqrt{\log n/n}\Big) \leq \frac{6(\sqrt{n}+2)}{n^2}. \tag{7.20}$$

By (7.18), $\zeta_n/c_n \to 0$ in probability.

Step 3: realized risks at an arbitrary threshold. For $\tau \in \overline{\mathbb{R}}$, let $\mathrm{FDP}_n(\tau)$ and $\mathrm{FNP}_n(\tau)$ be the realized ratios of the rule that rejects $\{i : Y_i \geq \tau\}$, subject to the truncation of step (iv) with $\widehat{R} = R_n(\tau)$. Put

$$\omega_+(\eta) = \sup_{t \in \overline{\mathbb{R}}:\, p_t \leq \eta} \{q_a(t) + \pi_1 - r_a(t)\}, \qquad \omega_-(\eta) = \sup_{t \in \overline{\mathbb{R}}:\, s_t \leq \eta} \{\pi_0 - q_a(t) + r_a(t)\}.$$

Because p_t is continuous and strictly decreasing from one to zero on $\overline{\mathbb{R}}$, $\{p_t \leq \eta\} = [t_\eta, +\infty]$ with $t_\eta \to +\infty$ as $\eta \downarrow 0$, so $\omega_+(\eta) \to 0$ by (2.10) and continuity; symmetrically $\omega_-(\eta) \to 0$. We claim that, pathwise and simultaneously for every $\tau \in \overline{\mathbb{R}}$,

$$\big|\mathrm{FDP}_n(\tau) - q_a(\tau)\big| + \big|\mathrm{FNP}_n(\tau) - r_a(\tau)\big| \leq \varepsilon_n' := \frac{4\zeta_n}{c_n} + \omega_+(c_n+\zeta_n) + \omega_-(c_n+\zeta_n) + 2\zeta_n. \tag{7.21}$$

If the rule rejects none, then either $\tau = +\infty$ or $p_\tau \leq R_n(\tau)/n + \zeta_n < c_n + \zeta_n$; its ratios are 0 and $U_n(+\infty)/n$, and $|U_n(+\infty)/n - \pi_1| \leq \zeta_n$. The reject-all case is symmetric, with $\mathrm{FDP} = V_n(-\infty)/n$ and $s_\tau < c_n + \zeta_n$ unless $\tau = -\infty$. Otherwise $\tau \in \mathbb{R}$, $R = R_n(\tau) \geq c_n n$, and $S = n - R \geq c_n n$. With $V = V_n(\tau)$,

$$\Big|\frac{V}{R} - \frac{\pi_0\overline{\Phi}(\tau)}{p_\tau}\Big| \leq \frac{|V/n - \pi_0\overline{\Phi}(\tau)|}{R/n} + \frac{\pi_0\overline{\Phi}(\tau)}{p_\tau}\cdot\frac{|p_\tau - R/n|}{R/n} \leq \frac{2\zeta_n}{c_n},$$

and the same bound holds for $U_n(\tau)/S$ and $r_a(\tau)$, because $|S/n - s_\tau| = |R/n - p_\tau|$. This proves (7.21), and $\varepsilon_n' \to 0$ in probability by Step 2.

Step 4: transfer to the estimated curve. On $\mathcal{E}_n$, let

$$\varepsilon_n'' = \sup_{t \in \overline{\mathbb{R}}} \big\|(q_a, r_a)(\tau(t)) - (q_{\widehat{\vartheta}}, r_{\widehat{\vartheta}})(t)\big\|_1,$$

and set $\varepsilon_n'' = 0$ on $\mathcal{E}_n^c$. The supremand is continuous in t, so $\varepsilon_n'' \leq 2$ is measurable. Let $N \subset \Xi$ be a compact neighbourhood of ϑ_0. By lemma 7.5(i), $(\vartheta, t) \mapsto (q_\vartheta, r_\vartheta)(t)$ is uniformly continuous on the compact set $N \times (\overline{\mathbb{R}}, \mathsf{d})$. On the event $\{\widehat{\vartheta} \in N,\ \widehat{\kappa} \in [1/2, 2]\}$, whose probability tends to one, ε_n'' is therefore at most a modulus of continuity evaluated at $|\widehat{m}| + |\widehat{\kappa} - 1|$ plus a modulus evaluated at $\|\widehat{\vartheta} - \vartheta_0\|$. By Step 1, $\varepsilon_n'' \to 0$ in probability.

Step 5: risks. Given (X, Θ, Z_0), which determines Y, ζ_n, ε_n', ε_n'', $\widehat{\lambda}$, and $\widehat{t}_1, \widehat{t}_2$, the external coin selects $\widehat{t}_1$ with probability $\widehat{\lambda}$, and on $\mathcal{E}_n$ the rule with $T = \widehat{t}_j$ has the realized ratios of Step 3 at $\tau = \tau(\widehat{t}_j)$. Because (7.21) holds simultaneously for all τ, it applies at these data-dependent thresholds. Since $\mathrm{FDP} \leq 1$ and $\min\{1, x+y\} \leq x + \min\{1, y\}$ for $x, y \geq 0$, Steps 3–4 and (7.17) give

$$\begin{aligned}\mathrm{FDR}(\delta^{(n)}) &\leq \mathbb{E}\Big[\mathbf{1}_{\mathcal{E}_n}\big\{\widehat{\lambda}q_{\widehat{\vartheta}}(\widehat{t}_1) + (1-\widehat{\lambda})q_{\widehat{\vartheta}}(\widehat{t}_2)\big\}\Big] + \mathbb{E}\min\{1, \varepsilon_n' + \varepsilon_n''\} \\ &\leq \alpha + o(1),\end{aligned}$$

where $\mathbb{E}\min\{1, \varepsilon_n' + \varepsilon_n''\} \to 0$ by bounded convergence, because $\varepsilon_n' + \varepsilon_n'' \to 0$ in probability. Off $\mathcal{E}_n$, $\mathrm{FNP} \leq 1$. Therefore

$$\mathrm{FNR}(\delta^{(n)}) \leq \mathbb{E}\big[\mathbf{1}_{\mathcal{E}_n}\min\{1, B_{\widehat{\vartheta}}(\alpha) + \varepsilon_n\}\big] + \mathbb{P}(\mathcal{E}_n^c) + \mathbb{E}\min\{1, \varepsilon_n' + \varepsilon_n''\}.$$

By lemma 7.5(ii) and $\widehat{\vartheta} \to \vartheta_0$ in probability, $\mathbb{P}(\mathcal{E}_n \cap \{B_{\widehat{\vartheta}}(\alpha) > B_a(\alpha) + \eta\}) \to 0$ for every $\eta > 0$. Hence $\limsup_n \mathrm{FNR}(\delta^{(n)}) \le B_a(\alpha)$.

Step 6: lower bound. Let $\limsup_n \mathrm{FDR}(\delta_n) \le \alpha$ and $\alpha' \in (\alpha, 1]$. Then $\mathrm{FDR}(\delta_n) \le \alpha'$ for all large n, so $\mathrm{FNR}(\delta_n) \ge \mathrm{OPT}_n(\alpha'; \pi_1, \mu, \rho)$, and proposition 5.1 gives $\liminf_n \mathrm{FNR}(\delta_n) \ge B_a(\alpha')$. Continuity of B_a (proposition 8.2) and $\alpha' \downarrow \alpha$ give the claim. Applied to $\delta^{(n)}$, it completes (7.19). □

Remark 7.8 (Scope of the plug-in result). Theorem 7.7 gives asymptotic, not finite-sample, FDR control: the calibration in definition 7.1 uses error terms that depend on the unknown parameters. The theorem assumes consistent estimators without supplying them, and it gives no rate. It is pointwise in the true parameter: consistency is required only at that parameter, and no uniformity of the FDR or FNR convergence over a parameter set is claimed. The proof uses the estimates only through $\widehat{\vartheta}$, $\widehat{\kappa}$, and $\widehat{m}$. The truncation in step (iv) of definition 7.6 changes the rule only when fewer than $c_n n$ or more than $(1 - c_n)n$ coordinates would be rejected; it is used to make (7.21) hold uniformly over thresholds chosen from the data.

8. Boundary cases and convex-envelope geometry

Proposition 8.1 (Finite-dimensional endpoint optima)**.** For every integer $n \ge 1$,

$$\mathrm{OPT}_n(0; \pi_1, \mu, \rho) = \pi_1.$$

For every $\alpha \in [\pi_0, 1]$,

$$\mathrm{OPT}_n(\alpha; \pi_1, \mu, \rho) = 0.$$

Consequently, $B_a(0) = \pi_1$ and $B_a(\alpha) = 0$ for $\alpha \ge \pi_0$.

Proof. Fix i, and let $f_{i,j}(x)$ be the subdensity of $X = x$ jointly with $\Theta_i = j$, for $j \in \{0, 1\}$. Each subdensity is strictly positive at every $x \in \mathbb{R}^n$, because it is a finite positive mixture of nonsingular Gaussian densities. We use the everywhere-defined posterior version

$$\mathbb{P}(\Theta_i = 0 \mid X = x) = \frac{f_{i,0}(x)}{f_{i,0}(x) + f_{i,1}(x)} \in (0, 1) \tag{8.1}$$

for every x.

Because a randomized rule is a Markov kernel based only on X, one has $\Theta \perp \delta \mid X$. Condition on X and on the realized external randomization. Whenever $R > 0$,

$$\mathbb{E}[\mathrm{FDP} \mid X, \delta] = \frac{1}{R} \sum_{i:\, \delta_i = 1} \mathbb{P}(\Theta_i = 0 \mid X) > 0.$$

Since FDP is nonnegative, FDR zero forces $\mathbb{P}(R > 0) = 0$. Thus every feasible rule at $\alpha = 0$ rejects none almost surely, and its FNR equals

$$\mathbb{E}\left[\frac{1}{n} \sum_{i=1}^{n} \Theta_i\right] = \pi_1.$$

If $\alpha \ge \pi_0$, reject all. Then

$$\mathrm{FDR} = \mathbb{E}\left[\frac{1}{n} \sum_{i=1}^{n} (1 - \Theta_i)\right] = \pi_0 \le \alpha,$$

and $S = 0$, so FNP and FNR are zero under the convention (2.2). Nonnegativity gives optimality.

For $B_a(0)$, every point of Γ_a has nonnegative first coordinate, and the unique curve point with first coordinate zero is $(0, \pi_1)$. Hence every feasible convex combination at level zero is supported there and has second coordinate π_1. Symmetrically, $(\pi_0, 0)$ is the unique point of K_a with second coordinate zero. For $\alpha \ge \pi_0$, that reject-all point is feasible, and nonnegativity gives $B_a(\alpha) = 0$. □

Proposition 8.2 (Geometry of the optimizer)**.** The maps $t \mapsto q_a(t)$ and $t \mapsto r_a(t)$ are continuous on $\overline{\mathbb{R}}$, with q_a strictly decreasing from π_0 to 0 and r_a strictly increasing from 0 to π_1. The function B_a is convex and continuous on $[0, 1]$, strictly decreasing on $[0, \pi_0]$, and zero on $[\pi_0, 1]$. For $0 < \alpha < \pi_0$, an optimizer in (4.13) can be chosen with active constraint

$$\lambda q_a(t_1) + (1 - \lambda) q_a(t_2) = \alpha. \tag{8.2}$$

If the lower convex envelope touches the threshold curve at a single point, no threshold-selection randomization is needed. If α lies in a nontrivial affine segment joining points (q_1, r_1) and (q_2, r_2), with $q_1 < \alpha < q_2$, the mixing weight on the first point is

$$\lambda = \frac{q_2 - \alpha}{q_2 - q_1}. \tag{8.3}$$

Proof. Continuity follows from the Gaussian tail formulas. Gaussian tail-ratio asymptotics show that the alternative upper tail dominates the null upper tail as $t \to +\infty$, while the null lower tail dominates the alternative lower tail as $t \to -\infty$; these give the endpoint limits in (2.10). The posterior representation

$$q_a(t) = \mathbb{E}[W(Y) \mid Y \geq t]$$

shows strict decrease because $W(Y)$ is strictly decreasing and the mixture density assigns positive mass to every nonempty interval: raising the truncation point removes a positive-probability set on which W is strictly larger. Similarly,

$$r_a(t) = \mathbb{E}[1 - W(Y) \mid Y < t]$$

is strictly increasing. The image of the compactified line $\overline{\mathbb{R}}$ is compact, and its convex hull is compact. Attainment and the two-point representation are definition 2.1 and lemma 4.3.

Suppose an optimizer for $0 < \alpha < \pi_0$ had FDR coordinate $\beta < \alpha$. Its FNR coordinate cannot be zero, because the only point of K_a with FNR zero is reject-all, whose FDR coordinate is $\pi_0 > \alpha$. A sufficiently small convex combination with reject-all would therefore remain below α and strictly lower FNR, contradicting optimality. Hence the constraint is active.

Convexity of B_a follows from convexity of K_a. If $0 \leq \alpha_1 < \alpha_2 < \pi_0$, mix an optimizer at α_1 with reject-all until its first coordinate is α_2; its positive second coordinate decreases strictly, so $B_a(\alpha_2) < B_a(\alpha_1)$. The same conclusion at $\alpha_2 = \pi_0$ follows from $B_a(\pi_0) = 0$. A finite convex function is continuous on the interior of its domain.

For right-continuity at zero, take any sequence $\alpha_m \downarrow 0$ and corresponding minimizers $(x_m, B_a(\alpha_m))$, where $0 \leq x_m \leq \alpha_m$. Every convergent subsequence has a limit in K_a with first coordinate zero, hence equals $(0, \pi_1)$ by the uniqueness just proved. Therefore $B_a(\alpha_m) \to \pi_1$. For $0 \leq \alpha \leq \pi_0$, mixing the two endpoint rules gives

$$0 \leq B_a(\alpha) \leq \pi_1 \left(1 - \frac{\alpha}{\pi_0}\right),$$

which proves left-continuity at π_0; the function is zero thereafter. Finally, (8.3) is obtained by solving $\lambda q_1 + (1-\lambda) q_2 = \alpha$. □

9. A strict gap for raw-coordinate separable rules

The compound rule used for achievability estimates the common factor from the entire observation vector. We now show that this use of the other coordinates is essential at some points of the frontier. The conclusion is deliberately formulated for a precise separable class; it is not a characterization of the entire separable frontier.

Put

$$\widetilde{X}_i = \frac{X_i}{\sqrt{1-\rho}} = a\Theta_i + G + Z_i, \qquad G = \sqrt{\frac{\rho}{1-\rho}}\, Z_0. \tag{9.1}$$

Thus $G \sim N(0, \rho/(1-\rho))$ has a strictly positive density on all of $\mathbb{R}$.

Definition 9.1 (Shared-seed raw-coordinate separability)**.** At dimension n, a rule is *shared-seed raw-coordinate separable* if there are a standard Borel space $\mathcal{A}_n$, an auxiliary random variable A_n in $\mathcal{A}_n$, and measurable maps

$$h_{n,i} : \mathbb{R} \times \mathcal{A}_n \longrightarrow \{0, 1\}, \qquad i = 1, \ldots, n,$$

such that A_n is independent of all labels and Gaussian variables and

$$\delta_{n,i} = h_{n,i}(\widetilde{X}_i, A_n). \tag{9.2}$$

The maps may depend on n, i, π_1, μ, ρ, and the same seed may contain arbitrarily many common or private random bits. In particular, this class is larger than product randomization. It excludes any data-derived shared statistic, such as $\overline{X}_n$.

Write $\mathfrak{S}_n$ for this class. Let $\mathrm{OPT}_n^{\mathrm{sep}}(\alpha)$ be the infimum in (2.3) restricted to this class.

We next define the limiting factor-conditional risks of a possibly randomized one-coordinate response. Let

$$\mathcal{D} = \{d \in L^\infty(\mathbb{R}) : 0 \le d \le 1 \text{ a.e.}\}.$$

For $d \in \mathcal{D}$ and $z \in \mathbb{R}$, set

$$u_d(z) = \int_{\mathbb{R}} d(z+y)\phi(y)\,\mathrm{d}y, \qquad v_d(z) = \int_{\mathbb{R}} d(z+a+y)\phi(y)\,\mathrm{d}y, \tag{9.3}$$

$$p_d(z) = \pi_0 u_d(z) + \pi_1 v_d(z). \tag{9.4}$$

If d is neither zero nor one as an element of L^∞, positivity of the Gaussian density gives $0 < p_d(z) < 1$ for every z. In that case define

$$q_d(z) = \frac{\pi_0 u_d(z)}{p_d(z)}, \tag{9.5}$$

$$r_d(z) = \frac{\pi_1[1 - v_d(z)]}{1 - p_d(z)}. \tag{9.6}$$

At the two endpoints use

$$(q_0(z), r_0(z)) = (0, \pi_1), \qquad (q_1(z), r_1(z)) = (\pi_0, 0).$$

Writing $\nu_\rho = N(0, \rho/(1-\rho))$ for the law of G, let

$$Q(d) = \int q_d(z)\,\nu_\rho(\,\mathrm{d}z), \qquad \mathcal{R}(d) = \int r_d(z)\,\nu_\rho(\,\mathrm{d}z). \tag{9.7}$$

For $\kappa > 0$, introduce the separable and oracle weighted values

$$J_\kappa = \min_{d \in \mathcal{D}} \{\mathcal{R}(d) + \kappa Q(d)\}, \tag{9.8}$$

$$H_\kappa = \min_{t \in \overline{\mathbb{R}}} \{r_a(t) + \kappa q_a(t)\}. \tag{9.9}$$

The endpoint costs in (9.9) are

$$r_a(+\infty) + \kappa q_a(+\infty) = \pi_1, \qquad r_a(-\infty) + \kappa q_a(-\infty) = \kappa\pi_0. \tag{9.10}$$

Lemma 9.2 (Triangular-array reduction for separable rules)**.** For every sequence of shared-seed raw-coordinate separable rules and every $\kappa > 0$,

$$\liminf_{n\to\infty} \{\mathrm{FNR}(\delta_n) + \kappa\,\mathrm{FDR}(\delta_n)\} \ge J_\kappa. \tag{9.11}$$

Proof. First freeze the seed. For $b \in \mathcal{A}_n$, define $\mathrm{FDR}_n(b)$ and $\mathrm{FNR}_n(b)$ by integrating the two ratios directly over the model variables after replacing A_n by b in (9.2). These are measurable functions of b, and independence of the seed gives

$$\begin{aligned} \mathrm{FDR}(\delta_n) &= \int \mathrm{FDR}_n(b)\,P_{A_n}(\,\mathrm{d}b), \\ \mathrm{FNR}(\delta_n) &= \int \mathrm{FNR}_n(b)\,P_{A_n}(\,\mathrm{d}b). \end{aligned} \tag{9.12}$$

This direct definition avoids selecting a version of a conditional expectation at a seed value of probability zero.

Now take arbitrary deterministic seed values $b_n \in \mathcal{A}_n$, and put

$$L_n(b_n) = \mathrm{FNR}_n(b_n) + \kappa\,\mathrm{FDR}_n(b_n).$$

Select a subsequence along which $L_n(b_n)$ tends to its liminf and relabel that subsequence by n. Define the averaged response

$$d_n(x) = \frac{1}{n}\sum_{i=1}^{n} h_{n,i}(x, b_n) \in \mathcal{D}. \tag{9.13}$$

The set $\mathcal{D}$, equipped with the weak-star topology inherited from $L^\infty = (L^1)^*$, is compact and metrizable: Banach–Alaoglu gives compactness, $L^1(\mathbb{R})$ is separable, and the order interval $0 \le d \le 1$ is weak-star closed. After passage to a further subsequence,

$$d_n \xrightarrow{*} d \in \mathcal{D}. \tag{9.14}$$

Conditionally on $G = z$ and the frozen seed, the vectors $(\Theta_i, Z_i, \delta_{n,i})$, $i \le n$, are independent. The conditional means of the normalized rejection and false-rejection counts are

$$\mathbb{E}[R/n \mid G = z, A_n = b_n] = p_{d_n}(z), \qquad \mathbb{E}[V/n \mid G = z, A_n = b_n] = \pi_0 u_{d_n}(z). \tag{9.15}$$

The corresponding nonrejection means are $1 - p_{d_n}(z)$ and $\pi_1[1 - v_{d_n}(z)]$. For completeness, if $p_{d_n}(z) > 0$, the usual ratio concentration argument gives the explicit bound

$$\begin{aligned} &\left| \mathbb{E}[\mathrm{FDP} \mid G = z, A_n = b_n] - \frac{\pi_0 u_{d_n}(z)}{p_{d_n}(z)} \right| \\ &\qquad \le \frac{2}{\sqrt{n}\, p_{d_n}(z)} + \exp\{-n p_{d_n}(z)/8\}. \end{aligned} \tag{9.16}$$

Indeed, with

$$q = \frac{\mathbb{E}[V \mid G = z, A_n = b_n]}{\mathbb{E}[R \mid G = z, A_n = b_n]},$$

the centered sum $V - qR$ has conditional second moment at most n. On $\{R \ge n p_{d_n}(z)/2\}$, its absolute value divided by R is at most $2|V - qR|/[n p_{d_n}(z)]$; the complementary probability is at most $\exp\{-n p_{d_n}(z)/8\}$ by the multiplicative Chernoff bound. The same argument gives

$$\begin{aligned} &\left| \mathbb{E}[\mathrm{FNP} \mid G = z, A_n = b_n] - \frac{\pi_1[1 - v_{d_n}(z)]}{1 - p_{d_n}(z)} \right| \\ &\qquad \le \frac{2}{\sqrt{n}\,[1 - p_{d_n}(z)]} + \exp\{-n[1 - p_{d_n}(z)]/8\} \end{aligned} \tag{9.17}$$

whenever $p_{d_n}(z) < 1$.

Weak-star convergence in (9.14), tested against $x \mapsto \phi(x - z)$ and $x \mapsto \phi(x - z - a)$, implies

$$u_{d_n}(z) \longrightarrow u_d(z), \qquad v_{d_n}(z) \longrightarrow v_d(z) \tag{9.18}$$

for every fixed z. If $d \notin \{0, 1\}$, both limiting denominators are strictly positive. Equations (9.16)–(9.17), whose left sides are also bounded by one, give pointwise convergence of the conditional risks. Dominated convergence in z then yields

$$\mathrm{FDR}_n(b_n) \longrightarrow Q(d), \qquad \mathrm{FNR}_n(b_n) \longrightarrow \mathcal{R}(d). \tag{9.19}$$

If $d = 0$, then (9.18) and dominated convergence give $\mathbb{E}[R/n \mid A_n = b_n] \to 0$, with the conditional notation interpreted by the direct integration just defined. Put $T_n = \sum_{i=1}^n \Theta_i$. On $\{R \le n/2\}$, direct comparison with reject none gives

$$\left| \mathrm{FNP} - \frac{T_n}{n} \right| \le \frac{4R}{n};$$

on the complementary event the difference is at most one. Markov's inequality therefore gives

$$\mathbb{E}\left[\left| \mathrm{FNP} - \frac{T_n}{n} \right| \,\middle|\, A_n = b_n \right] \le 6\, \mathbb{E}[R/n \mid A_n = b_n] \longrightarrow 0. \tag{9.20}$$

Consequently the frozen-seed FNR tends to π_1, while its FDR is nonnegative. If $d = 1$, then $\mathbb{E}[(n - R)/n \mid A_n = b_n] \to 0$; the symmetric comparison with reject all gives

$$\mathbb{E}\left[\left| \mathrm{FDP} - \frac{1}{n} \sum_{i=1}^n (1 - \Theta_i) \right| \,\middle|\, A_n = b_n \right] \le 6\, \mathbb{E}[(n - R)/n \mid A_n = b_n] \longrightarrow 0. \tag{9.21}$$

Thus the frozen-seed FDR tends to π_0, while its FNR is nonnegative. In particular, the two endpoint costs are at least J_κ, since both endpoint responses are admissible in (9.8). Thus, in all three cases,

$$\liminf_{n \to \infty} L_n(b_n) \ge J_\kappa. \tag{9.22}$$

It remains to unfreeze the seed. If (9.11) failed, then along a subsequence (9.12) would give

$$\int L_n(b)\, P_{A_n}(\mathrm{d}b) < J_\kappa - \varepsilon$$

for some $\varepsilon > 0$. For every n on that subsequence one could choose b_n with $L_n(b_n) < J_\kappa - \varepsilon$, contradicting (9.22). □

Lemma 9.3 (Conditional Neyman–Pearson comparison)**.** Let $g : \mathbb{R} \to [0,1]$ be measurable and put

$$u = \int g(x)\phi(x)\,\mathrm{d}x, \qquad v = \int g(x)\phi(x-a)\,\mathrm{d}x, \qquad m = \pi_0 u + \pi_1 v.$$

Suppose $0 < m < 1$, and let $t \in \mathbb{R}$ be the unique threshold satisfying $p_t = m$. Define

$$q(g) = \frac{\pi_0 u}{m}, \qquad r(g) = \frac{\pi_1(1-v)}{1-m}.$$

Then

$$q(g) - q_a(t) = \frac{\pi_1(v_t - v)}{m}, \tag{9.23}$$

$$r(g) - r_a(t) = \frac{\pi_1(v_t - v)}{1-m}, \tag{9.24}$$

where $v_t = \overline{\Phi}(t-a)$, and $v_t \geq v$. Consequently,

$$r(g) + \kappa q(g) \geq r_a(t) + \kappa q_a(t) \geq H_\kappa. \tag{9.25}$$

Equality in the first inequality holds if and only if $g(x) = \mathbf{1}\{x \geq t\}$ for almost every x with respect to Lebesgue measure.

Proof. Among all responses with mixture rejection mass m, maximizing v is a Neyman–Pearson problem between the alternative density $\phi(x-a)$ and the mixture density $\pi_0\phi(x) + \pi_1\phi(x-a)$. Their likelihood ratio is a strictly increasing function of x. Since the mixture law is atomless, the unique maximizer up to null sets is the upper-threshold response $\mathbf{1}\{x \geq t\}$, where $p_t = m$. Hence $v_t \geq v$, with equality only for that response. Finally, $m = \pi_0 u + \pi_1 v$ gives

$$q(g) = 1 - \frac{\pi_1 v}{m},$$

which proves (9.23); (9.24) follows directly. Their nonnegative weighted sum proves (9.25) and the equality assertion. □

Theorem 9.4 (Exact scalarized separable value and strict gap)**.** For every $\kappa > 0$,

$$\lim_{n\to\infty} \inf_{\delta \in \mathfrak{S}_n} \{\mathrm{FNR}(\delta) + \kappa\,\mathrm{FDR}(\delta)\} = J_\kappa \geq H_\kappa. \tag{9.26}$$

Moreover,

$$J_\kappa > H_\kappa \quad \Longleftrightarrow \quad H_\kappa < \min\{\pi_1, \kappa\pi_0\}. \tag{9.27}$$

We call

$$H_\kappa < \min\{\pi_1, \kappa\pi_0\}, \tag{9.28}$$

the *endpoint-strict condition*. When it holds, define the positive separation

$$\varepsilon_\kappa := J_\kappa - H_\kappa > 0. \tag{9.29}$$

Proof. For fixed $d \notin \{0,1\}$ and $z \in \mathbb{R}$, apply lemma 9.3 to $g_z(y) = d(z+y)$. Its two conditional ratios are exactly $q_d(z)$ and $r_d(z)$. Integrating (9.25) over z proves $\mathcal{R}(d) + \kappa Q(d) \geq H_\kappa$. The endpoint cases follow from (9.10), proving $J_\kappa \geq H_\kappa$.

Give $\mathcal{D}$ the weak-star topology used in lemma 9.2. If $d_j \overset{*}{\rightharpoonup} d$ and $d \notin \{0,1\}$, Gaussian-convolution continuity and dominated convergence show

$$\mathcal{R}(d_j) + \kappa Q(d_j) \longrightarrow \mathcal{R}(d) + \kappa Q(d).$$

If $d = 0$, then $r_{d_j}(z) \to \pi_1$ for every z, so $\mathcal{R}(d_j) \to \pi_1$, whereas $Q(d_j) \geq 0$. If $d = 1$, then $q_{d_j}(z) \to \pi_0$, so $Q(d_j) \to \pi_0$, whereas $\mathcal{R}(d_j) \geq 0$. Hence the objective in (9.8) is lower semicontinuous on the compact set $\mathcal{D}$, and its infimum is attained.

We next identify J_κ with the scalarized asymptotic separable value. For the lower bound in (9.26), choose at each dimension a rule in $\mathfrak{S}_n$ within n^{-1} of the displayed finite-dimensional infimum and apply lemma 9.2. For the reverse bound, let d attain J_κ and choose a measurable representative. If $d = 0$ or $d = 1$, use the corresponding endpoint rule. Otherwise, use no external seed and set

$$h_{n,i}(x) = \mathbf{1}\left\{ d(x) \geq \frac{i - \frac{1}{2}}{n} \right\}, \qquad i = 1, \ldots, n. \tag{9.30}$$

Its averaged response obeys

$$\left|\frac{1}{n}\sum_{i=1}^{n} h_{n,i}(x) - d(x)\right| \le \frac{1}{2n}$$

for every x. The interior-limit part of the proof of lemma 9.2 therefore gives FDR and FNR limits $Q(d)$ and $\mathcal{R}(d)$, respectively. This proves the matching upper bound and hence the equality in (9.26).

Because the two extended thresholds are admissible in (9.9), one always has $H_\kappa \le \min\{\pi_1, \kappa\pi_0\}$. If equality holds, the endpoint response attaining the smaller cost is admissible in (9.8); combined with $J_\kappa \ge H_\kappa$, this gives $J_\kappa = H_\kappa$.

Suppose now that (9.28) holds but $J_\kappa = H_\kappa$, and let d attain J_κ. The endpoint costs are strictly larger than H_κ, so $d \notin \{0, 1\}$. Let

$$T_\kappa = \arg\min_{t\in\overline{\mathbb{R}}}\{r_a(t) + \kappa q_a(t)\}. \tag{9.31}$$

The objective is continuous on $\overline{\mathbb{R}}$. Condition (9.28) excludes both extended endpoints, so T_κ is a nonempty compact subset of $\mathbb{R}$. Now consider the function

$$c_d(z) := r_d(z) + \kappa q_d(z)$$

which is continuous, is everywhere at least H_κ, and has expectation H_κ under a law with a strictly positive density on $\mathbb{R}$. It must therefore equal H_κ for every $z \in \mathbb{R}$. For each z, equality must hold in both inequalities

$$c_d(z) \ge r_a(t_z) + \kappa q_a(t_z) \ge H_\kappa$$

from lemma 9.3. Hence the translated response is a threshold and its threshold t_z belongs to T_κ. In particular, at $z = 0$ the equality statement in lemma 9.3 says that $d(x) = \mathbf{1}\{x \ge T\}$ almost everywhere for some $T \in T_\kappa$. At a general z, the same response becomes

$$d(z + y) = \mathbf{1}\{y \ge T - z\}.$$

Equality in the lemma therefore forces $T - z \in T_\kappa$ for every $z \in \mathbb{R}$. This is impossible for a compact subset of $\mathbb{R}$. The contradiction proves $J_\kappa > H_\kappa$, so the positive separation in (9.29) is well defined. □

Corollary 9.5 (Strict compound–separable gap on supported faces)**.** Fix $\kappa > 0$ satisfying (9.28). If $\alpha \in (0, \pi_0)$ lies on the corresponding supported face of the oracle frontier, in the sense that

$$H_\kappa = B_a(\alpha) + \kappa\alpha, \tag{9.32}$$

then every sequence in the class of definition 9.1 satisfying $\mathrm{FDR}(\delta_n) \le \alpha$ for every n obeys

$$\liminf_{n\to\infty} \mathrm{FNR}(\delta_n) \ge B_a(\alpha) + \varepsilon_\kappa. \tag{9.33}$$

Consequently,

$$\liminf_{n\to\infty} \mathrm{OPT}_n^{\mathrm{sep}}(\alpha) \ge B_a(\alpha) + \varepsilon_\kappa > \lim_{n\to\infty} \mathrm{OPT}_n(\alpha; \pi_1, \mu, \rho). \tag{9.34}$$

In particular, if $t_* \in T_\kappa$ and $\alpha_* = q_a(t_*)$, then

$$B_a(\alpha_*) = r_a(t_*), \qquad H_\kappa = B_a(\alpha_*) + \kappa\alpha_*, \tag{9.35}$$

so (9.33) holds at α_*.

Proof. For a feasible separable sequence, lemma 9.2 and $\mathrm{FDR}(\delta_n) \le \alpha$ give

$$\begin{aligned}\liminf_{n\to\infty} \mathrm{FNR}(\delta_n) &\ge J_\kappa - \kappa\alpha \\ &= B_a(\alpha) + \varepsilon_\kappa,\end{aligned}$$

which proves (9.33). Taking nearly minimizing separable rules proves the first inequality in (9.34), and theorem 2.2 proves the last equality.

For the final assertion, the curve point $(q_a(t_*), r_a(t_*))$ shows that $B_a(\alpha_*) \le r_a(t_*)$. A strict inequality would produce a point $(q, r) \in K_a$ with $q \le \alpha_*$ and $r < r_a(t_*)$, and hence

$$r + \kappa q < r_a(t_*) + \kappa q_a(t_*) = H_\kappa.$$

This contradicts the definition of H_κ, because the continuous linear functional $(q, r) \mapsto r + \kappa q$ is bounded below by H_κ on all of K_a, the convex hull of the extended threshold curve. Thus (9.35) holds. □

Corollary 9.6 (A certified Gaussian gap witness)**.** Take

$$\pi_1 = \frac{1}{4}, \qquad \mu = 2, \qquad \rho = \frac{1}{2}, \qquad \kappa = \frac{2}{5}. \tag{9.36}$$

Then $a = 2\sqrt{2}$ and

$$H_{2/5} < \frac{13}{75} < \frac{1}{4} = \min\left\{\pi_1, \frac{2}{5}\pi_0\right\}. \tag{9.37}$$

Thus every $t_* \in T_{2/5}$ is finite, and at $\alpha_* = q_{2\sqrt{2}}(t_*) \in (0, 3/4)$ there is a strict asymptotic gap between all compound rules and the shared-seed raw-coordinate separable class:

$$\liminf_{n\to\infty} \mathrm{OPT}_n^{\mathrm{sep}}(\alpha_*) \geq B_{2\sqrt{2}}(\alpha_*) + \varepsilon_{2/5} > B_{2\sqrt{2}}(\alpha_*). \tag{9.38}$$

Moreover, the certified separation between $H_{2/5}$ and the smaller endpoint cost is greater than $23/300$.

Proof. At $t = a/2 = \sqrt{2}$, put $\eta = \overline{\Phi}(\sqrt{2})$. Symmetry gives

$$q_{2\sqrt{2}}(\sqrt{2}) = \frac{3\eta}{1+2\eta}, \qquad r_{2\sqrt{2}}(\sqrt{2}) = \frac{\eta}{3-2\eta}. \tag{9.39}$$

Mills' inequality yields

$$\eta < \frac{\phi(\sqrt{2})}{\sqrt{2}} = \frac{e^{-1}}{2\sqrt{\pi}} < \frac{1}{9}.$$

For example, the last strict inequality follows from the elementary lower bounds $e > 19/7$ and $\pi > 49/16$. Therefore

$$q_{2\sqrt{2}}(\sqrt{2}) < \frac{1}{3}, \qquad r_{2\sqrt{2}}(\sqrt{2}) < \frac{1}{25},$$

and hence

$$H_{2/5} \leq r_{2\sqrt{2}}(\sqrt{2}) + \frac{2}{5} q_{2\sqrt{2}}(\sqrt{2}) < \frac{1}{25} + \frac{2}{15} = \frac{13}{75}.$$

The two endpoint costs are $1/4$ and $(2/5)(3/4) = 3/10$, proving (9.37). Since

$$\frac{1}{4} - \frac{13}{75} = \frac{23}{300},$$

the endpoint margin statement follows. The remaining conclusions are theorem 9.4 and corollary 9.5. □

The supported-face corollary leaves open the levels α at which every supporting line of the oracle frontier passes through an endpoint. We now show that the constrained gap vanishes exactly when the oracle frontier is the chord joining the two endpoint rules, and is strict at every interior level otherwise. Put

$$\kappa_0 = \frac{\pi_1}{\pi_0}, \qquad \ell(\alpha) = \pi_1 \max\left\{0, 1 - \frac{\alpha}{\pi_0}\right\}, \qquad \alpha \in [0,1].$$

The extended endpoints $(0, \pi_1)$ and $(\pi_0, 0)$ of Γ_a lie on the line $r + \kappa_0 q = \pi_1$, and ℓ is the value of mixing the reject-none and reject-all rules.

Lemma 9.7 (The threshold curve is not the endpoint chord)**.** There is $t_0 \in \mathbb{R}$ such that $r_a(t) + \kappa_0 q_a(t) > \pi_1$ for every $t \geq t_0$. In particular, Γ_a is not contained in a line.

Proof. Since $s_t - \Phi(t-a) = \pi_0[\Phi(t) - \Phi(t-a)] \leq \pi_0 \overline{\Phi}(t-a)$,

$$\pi_1 - r_a(t) = \frac{\pi_1[s_t - \Phi(t-a)]}{s_t} \leq \frac{\pi_0 \pi_1 \overline{\Phi}(t-a)}{s_t}.$$

Because $a > 0$, $p_t \leq \overline{\Phi}(t-a)$, so $q_a(t) \geq \pi_0 \overline{\Phi}(t)/\overline{\Phi}(t-a) > 0$. For $t \geq a+1$, the elementary bounds $\overline{\Phi}(x) \leq \phi(x)/x \leq \phi(x)$ for $x \geq 1$, $\overline{\Phi}(t) \geq \int_t^{t+1} \phi(y)\,\mathrm{d}y \geq \phi(t+1)$, and $s_t \geq \pi_0 \Phi(t) \geq \pi_0/2$ give

$$0 \leq \frac{\pi_1 - r_a(t)}{q_a(t)} \leq \frac{\pi_1 \overline{\Phi}(t-a)^2}{s_t \overline{\Phi}(t)} \leq \frac{2\pi_1}{\pi_0} \frac{\phi(t-a)^2}{\phi(t+1)} = \frac{2\pi_1}{\pi_0\sqrt{2\pi}} \exp\left\{-\frac{t^2}{2} + (2a+1)t - a^2 + \frac{1}{2}\right\},$$

which tends to zero as $t \to \infty$. Hence, for all large t,

$$r_a(t) + \kappa_0 q_a(t) - \pi_1 = q_a(t)\left[\kappa_0 - \frac{\pi_1 - r_a(t)}{q_a(t)}\right] > 0.$$

A line containing Γ_a would contain both extended endpoints, hence would be the line $r + \kappa_0 q = \pi_1$; the display excludes this. □

Theorem 9.8 (Exactly when the separable gap vanishes)**.** Exactly one of the following two cases occurs.

(a) *Chord case:* $r_a(t) + \kappa_0 q_a(t) \geq \pi_1$ for every $t \in \mathbb{R}$. Then $B_a = \ell$ on $[0,1]$. For every $n \geq 1$ and $\alpha \in [0,1]$, the shared-seed rule that rejects all coordinates with probability $\min\{1, \alpha/\pi_0\}$ and otherwise rejects none belongs to $\mathfrak{S}_n$ and has FDR $= \min\{\alpha, \pi_0\} \leq \alpha$ and FNR $= B_a(\alpha)$. Consequently,

$$\lim_{n\to\infty} \mathrm{OPT}_n^{\mathrm{sep}}(\alpha) = \lim_{n\to\infty} \mathrm{OPT}_n(\alpha; \pi_1, \mu, \rho) = B_a(\alpha), \qquad \alpha \in [0,1].$$

(b) *Strict case:* $r_a(t) + \kappa_0 q_a(t) < \pi_1$ for some $t \in \mathbb{R}$. Then $B_a(\alpha) < \ell(\alpha)$ for every $\alpha \in (0, \pi_0)$. Moreover, for every $\alpha \in (0, \pi_0)$ there is a constant $V^{\mathrm{sep}}(\alpha) > B_a(\alpha)$, depending only on $(\alpha, \pi_1, \mu, \rho)$, such that every sequence $\delta_n \in \mathfrak{S}_n$ with $\limsup_{n\to\infty} \mathrm{FDR}(\delta_n) \leq \alpha$ satisfies

$$\liminf_{n\to\infty} \mathrm{FNR}(\delta_n) \geq V^{\mathrm{sep}}(\alpha).$$

In particular, $\liminf_{n\to\infty} \mathrm{OPT}_n^{\mathrm{sep}}(\alpha) > B_a(\alpha) = \lim_{n\to\infty} \mathrm{OPT}_n(\alpha; \pi_1, \mu, \rho)$ for every $\alpha \in (0, \pi_0)$.

In both cases there is no gap at $\alpha = 0$ or $\alpha \geq \pi_0$. With $w = \pi_0\pi_1$, $F_0 = \Phi(t)$, $F_1 = \Phi(t-a)$, $G_0 = \overline{\Phi}(t)$, and $G_1 = \overline{\Phi}(t-a)$, one has the identity

$$\frac{s_t p_t}{\pi_1}\big[r_a(t) + \kappa_0 q_a(t) - \pi_1\big] = (1-w)F_1G_0 + wF_1G_1 - wF_0(F_0 - F_1), \tag{9.40}$$

so the case distinction depends on π_1 only through $\pi_0\pi_1$.

Proof. Identity. Since $\kappa_0 q_a(t) = \pi_1 G_0/p_t$ and $r_a(t) = \pi_1 F_1/s_t$, the left side of (9.40) equals $F_1 p_t + G_0 s_t - s_t p_t$. Substituting $s_t = \pi_0 F_0 + \pi_1 F_1$, $p_t = \pi_0 G_0 + \pi_1 G_1$, and $G_0 - G_1 = F_1 - F_0$ and collecting terms gives the right side.

Endpoint levels. At $\alpha = 0$ the reject-none rule and at $\alpha \geq \pi_0$ the reject-all rule are separable and attain the compound optima of proposition 8.1 for every n.

Case (a)*.* Every finite curve point, and both extended endpoints, satisfy $r \geq \pi_1 - \kappa_0 q$, so K_a lies in this closed half-plane. If $(q,r) \in K_a$ and $q \leq \alpha \leq \pi_0$, then $r \geq \pi_1 - \kappa_0\alpha = \ell(\alpha)$; thus $B_a \geq \ell$ on $[0, \pi_0]$. Mixing the endpoint points of K_a gives $B_a \leq \ell$, and both functions vanish on $[\pi_0, 1]$. For the displayed rule, take $\mathcal{A}_n = \{0,1\}$, $A_n \sim \mathrm{Bernoulli}(\gamma)$ with $\gamma = \min\{1, \alpha/\pi_0\}$, and $h_{n,i}(x,b) = b$. Reject-all has FDP $= n^{-1}\sum_i (1 - \Theta_i)$ and FNP $= 0$; reject-none has FDP $= 0$ and FNP $= n^{-1}\sum_i \Theta_i$. Hence FDR $= \gamma\pi_0$ and FNR $= (1-\gamma)\pi_1 = \ell(\alpha) = B_a(\alpha)$. Therefore $\mathrm{OPT}_n(\alpha; \pi_1, \mu, \rho) \leq \mathrm{OPT}_n^{\mathrm{sep}}(\alpha) \leq B_a(\alpha)$ for every n, and theorem 2.2 gives the displayed limits.

Case (b)*, the oracle frontier.* Choose $\bar{t} \in \mathbb{R}$ with $r_a(\bar{t}) + \kappa_0 q_a(\bar{t}) < \pi_1$, and let $\bar{q} = q_a(\bar{t}) \in (0, \pi_0)$. Then $B_a(\bar{q}) \leq r_a(\bar{t}) < \ell(\bar{q})$. Suppose that $B_a(\alpha) = \ell(\alpha)$ for some $\alpha \in (0, \pi_0)$. For $x \in [0, \pi_0] \setminus \{\alpha\}$, choose the endpoint $e \in \{0, \pi_0\}$ on the other side of α and write $\alpha = \theta x + (1-\theta)e$ with $\theta \in (0,1)$. Convexity of B_a, $B_a(e) = \ell(e)$, and affinity of ℓ on $[0, \pi_0]$ give

$$\ell(\alpha) = B_a(\alpha) \leq \theta B_a(x) + (1-\theta)\ell(e),$$

hence $B_a(x) \geq \ell(x)$. Together with $B_a \leq \ell$, this forces $B_a = \ell$ on $[0, \pi_0]$, contradicting the strict inequality at $\bar{q}$.

Case (b)*, the separable risk set.* Let $\mathcal{G} = \{(Q(d), \mathcal{R}(d)) : d \in \mathcal{D}\} \subset [0,1]^2$, with the endpoint conventions for $d \in \{0,1\}$, and let $\mathcal{S} = \mathrm{conv}(\overline{\mathcal{G}})$, a compact convex set. For $\kappa \geq 0$, the continuous linear functional $L_\kappa(q,r) = r + \kappa q$ satisfies $\min_{\mathcal{S}} L_\kappa = \inf_{\mathcal{G}} L_\kappa = J_\kappa$, where $J_0 = 0$ is attained by $d = 1$.

We first classify the points $g \in \overline{\mathcal{G}}$ with $L_\kappa(g) = H_\kappa$ for some $\kappa > 0$. Write $g = \lim_j (Q(d_j), \mathcal{R}(d_j))$ and, by weak-star compactness of $\mathcal{D}$, assume $d_j \overset{*}{\rightharpoonup} d$. If $d \notin \{0,1\}$, the continuity argument in the proof of theorem 9.4, which gives convergence of Q and $\mathcal{R}$ separately, yields $g = (Q(d), \mathcal{R}(d))$, so $\mathcal{R}(d) + \kappa Q(d) = H_\kappa$. As in that proof, $c_d(z) = r_d(z) + \kappa q_d(z)$ is continuous, at least H_κ, and has ν_ρ-integral H_κ; hence $c_d \equiv H_\kappa$. The equality case of lemma 9.3 at $z = 0$ gives $d = \mathbf{1}\{x \geq T\}$ almost everywhere for some $T \in \mathbb{R}$; at general z the translated response $y \mapsto d(z+y)$ is the threshold response at $T - z$, and equality in (9.25) gives $r_a(T-z) + \kappa q_a(T-z) = H_\kappa$ for every $z \in \mathbb{R}$. Thus $t \mapsto r_a(t) + \kappa q_a(t)$ is constant on $\mathbb{R}$; letting $t \to \pm\infty$ gives $\pi_1 = H_\kappa = \kappa\pi_0$, so $\kappa = \kappa_0$ and Γ_a lies on the line $r + \kappa_0 q = \pi_1$, contradicting lemma 9.7. If $d = 0$, the proof of theorem 9.4 gives $\mathcal{R}(d_j) \to \pi_1$, so $g = (q, \pi_1)$ with $q \geq 0$; since $H_\kappa \leq \pi_1$, the equation $\pi_1 + \kappa q = H_\kappa$ forces $g = (0, \pi_1)$ and $H_\kappa = \pi_1$. If $d = 1$, then $Q(d_j) \to \pi_0$, so $g = (\pi_0, r)$ with $r \geq 0$; since $H_\kappa \leq \kappa\pi_0$, this forces $g = (\pi_0, 0)$ and $H_\kappa = \kappa\pi_0$. Both endpoints cannot occur for the same κ: that would give $\kappa = \kappa_0$ and $H_{\kappa_0} = \pi_1$, so $r_a(t) + \kappa_0 q_a(t) \geq \pi_1$ for every t, which is case (a) and is excluded here.

Now fix $\alpha \in (0, \pi_0)$ and define

$$V^{\mathrm{sep}}(\alpha) = \min\{r : (q, r) \in \mathcal{S},\ q \le \alpha\}.$$

The minimum exists because $\mathcal{S}$ is compact and contains $(0, \pi_1)$. We claim $V^{\mathrm{sep}}(\alpha) > B_a(\alpha)$. Since B_a is a finite convex function on $[0, 1]$, it has a subgradient $-\kappa_*$ at α: $B_a(x) \ge B_a(\alpha) - \kappa_*(x - \alpha)$ for $x \in [0, 1]$. Taking $x = \pi_0$ and using $B_a(\alpha) > 0 = B_a(\pi_0)$ from proposition 8.2 gives $\kappa_* > 0$. Every point of K_a has $q \in [0, \pi_0]$ and $r \ge B_a(q)$, so $r + \kappa_* q \ge B_a(\alpha) + \kappa_* \alpha$ on K_a, with equality at the minimizing point for $B_a(\alpha)$, whose first coordinate is α by proposition 8.2. Hence $H_{\kappa_*} = B_a(\alpha) + \kappa_* \alpha$. Suppose a point $(q', r') \in \mathcal{S}$ had $q' \le \alpha$ and $r' \le B_a(\alpha)$. Then

$$L_{\kappa_*}(q', r') \le B_a(\alpha) + \kappa_* \alpha = H_{\kappa_*} \le J_{\kappa_*} \le L_{\kappa_*}(q', r'),$$

where $H_{\kappa_*} \le J_{\kappa_*}$ is theorem 9.4. Write (q', r') as a convex combination of points of $\overline{\mathcal{G}}$ with positive weights. Each of them has L_{κ_*}-value at least J_{κ_*}, and their weighted average equals $H_{\kappa_*} \le J_{\kappa_*}$; so each of them satisfies $L_{\kappa_*} = H_{\kappa_*}$. By the classification they are all equal to one endpoint. If it is $(0, \pi_1)$, then $r' = \pi_1 > B_a(\alpha)$ because B_a is strictly decreasing on $[0, \pi_0]$ and $\alpha > 0$; if it is $(\pi_0, 0)$, then $q' = \pi_0 > \alpha$. Both contradict the choice of (q', r'), proving the claim.

Case (b)*, the lower bound.* The function V^{sep}, defined by the same formula on $[0, \infty)$, is finite, convex, and nonincreasing, because $\mathcal{S}$ is convex and contains $(0, \pi_1)$. It therefore has a subgradient $-\kappa \le 0$ at the interior point α. For $(q, r) \in \mathcal{S}$, $q \ge 0$ and

$$r \ge V^{\mathrm{sep}}(q) \ge V^{\mathrm{sep}}(\alpha) - \kappa(q - \alpha),$$

so $J_\kappa = \min_{\mathcal{S}} L_\kappa \ge V^{\mathrm{sep}}(\alpha) + \kappa\alpha$. If $\kappa = 0$, this would give $V^{\mathrm{sep}}(\alpha) \le J_0 = 0 \le B_a(\alpha)$, contrary to the claim; hence $\kappa > 0$. For a separable sequence with $\limsup_n \mathrm{FDR}(\delta_n) \le \alpha$, lemma 9.2 gives

$$\liminf_{n\to\infty} \mathrm{FNR}(\delta_n) \ge \liminf_{n\to\infty} \{\mathrm{FNR}(\delta_n) + \kappa\,\mathrm{FDR}(\delta_n)\} - \kappa\alpha \ge J_\kappa - \kappa\alpha \ge V^{\mathrm{sep}}(\alpha).$$

Applying this to rules in $\mathfrak{S}_n$ that are feasible at level α and within n^{-1} of $\mathrm{OPT}_n^{\mathrm{sep}}(\alpha)$, and using theorem 2.2, gives the final assertion. □

For the parameters of corollary 9.6, the bounds in its proof give $r_{2\sqrt{2}}(\sqrt{2}) + \frac{1}{3} q_{2\sqrt{2}}(\sqrt{2}) < \frac{1}{25} + \frac{1}{9} < \frac{1}{4} = \pi_1$, so case (b) holds and the separable gap is strict at every $\alpha \in (0, 3/4)$, not only at the supported point α_*.

Remark 9.9 (Numerical location of the two cases). The following observations are numerical and are not used in any proof. Determining the sign of the infimum over t of $r_a(t) + \kappa_0 q_a(t) - \pi_1$ on a fine grid in t, refined by a bounded scalar minimization, the chord case (a) occurs for small effective signal and the strict case (b) for large effective signal, with a single observed transition at $a \approx 1.349$ for $\pi_1 = 1/2$, $a \approx 1.464$ for $\pi_1 \in \{1/4, 3/4\}$, $a \approx 1.770$ for $\pi_1 \in \{0.1, 0.9\}$, and $a \approx 1.989$ for $\pi_1 \in \{0.05, 0.95\}$. Monotonicity of the case distinction in a is not proved here. The baseline point of Table 1, with $a = 2\sqrt{2}$, lies in case (b); this is also certified analytically by the preceding paragraph.

Remark 9.10 (What the gap does and does not establish). Corollary 9.6 proves existence of a strict gap at a certified, nonendpoint supporting point for the fixed one-factor family, and theorem 9.8 identifies exactly the levels at which the constrained gap vanishes. These results cover nonidentical, n-dependent coordinate rules with arbitrary independent shared randomization. They do not evaluate J_κ or V^{sep} in closed form, and they do not determine the whole constrained separable frontier; theorem 9.8 gives a strict lower bound, not the exact separable limit, in case (b). No positive gap is possible at the two endpoint regimes $\alpha = 0$ and $\alpha \ge \pi_0$, where the compound optima are attained by reject none and reject all. The translation obstruction in the proof also uses both $\rho > 0$, which gives G unbounded full support, and independence of the external seed from the data and the common factor.

10. Computational checks

The proof is analytic. The accompanying code is included only for falsification, implementation, and finite-sample diagnostics. It implements the observed-factor oracle, full-sample factor adjustment, leave-one-out adjustment, deterministic two-fold cross-fitting, the exact finite-threshold and two-threshold-mixture formulas, a numerical lower convex hull, the rigorous calibration in definition 7.1, and two independent data generators. Automated tests cover endpoint rules, extreme Gaussian tails, tiny FDR levels, exact finite-n formulas, input validation, and random-number-stream invariance.

For the parameter point

$$(\pi_1, \mu, \rho, \alpha) = (0.25, 2, 0.5, 0.1),$$

the numerical convex-envelope value is 0.0548568 to seven decimal places. Table 1 reports a fixed-seed run with $n = 400$ and 4,000 replications. The uncalibrated plug-in variants are close to the oracle, as predicted asymptotically. The final row is deliberately different: it implements the theorem's conservative analytic finite-sample calibration, whose run probability is only $\gamma_n = 0.01678$ at this moderate dimension. Its FDR is safely below the target, but its FNR is correspondingly close to the reject-none value $\pi_1 = 0.25$. Thus the calibration is a proof device, not a claim of practical sharpness.

TABLE 1. Baseline Monte Carlo diagnostic at $n = 400$. Parentheses contain Monte Carlo standard errors. The crossing fraction is measured relative to the coupled oracle; for the calibrated row the oracle uses the same calibration coin.

Method	FDR	FNR	Crossing fraction	γ_n
Oracle	0.09906 (0.00049)	0.05451 (0.00021)	—	—
Full-sample	0.09970 (0.00055)	0.05526 (0.00025)	0.00744	—
Leave-one-out	0.10018 (0.00055)	0.05504 (0.00025)	0.00741	—
Two-fold cross-fit	0.10188 (0.00056)	0.05530 (0.00025)	0.01050	—
Full-sample, calibrated	0.00186 (0.00023)	0.24701 (0.00054)	0.00012	0.01678

The baseline scaling run in Table 2 uses independent fixed seeds at each dimension. The crossing fraction decreases with n, and the full-sample FNR approaches the oracle target. Monte Carlo error and finite-sample bias prevent the rows from being monotone and are not used as evidence of a convergence rate.

TABLE 2. Baseline scaling diagnostic for the uncalibrated full-sample rule.

n	Reps	Oracle FDR	Oracle FNR	Full FDR	Full FNR	Crossing fraction
100	6000	0.09972	0.05580	0.10390	0.05767	0.01435
400	3000	0.10041	0.05503	0.10167	0.05535	0.00734
1600	1500	0.09925	0.05458	0.09955	0.05471	0.00371

A separate high-dependence case with $\rho = 0.9$ and $n = 400$ has target FNR 9.81×10^{-7}, oracle and full-sample FDR estimates 0.09978 and 0.10383, and crossing fraction 0.00717. Its analytic calibration probability is only 0.01061, again illustrating conservatism. A weak-signal case selects an endpoint mixture, for which factor adjustment changes no decisions and the analytic calibration probability is one.

The machine-readable result records every reported configuration and seed, software versions, exact-formula checks, a tiny-α endpoint regression, and a comparison between latent-factor and direct multivariate-normal generators. The summary was generated with Python 3.14.3, NumPy 2.3.5, and SciPy 1.17.0. These diagnostics are not part of the proof.

11. Discussion

The results isolate a setting in which strong dependence is asymptotically removable for the compound FDR–FNR frontier while producing a strict loss for raw-coordinate separable rules at every interior level whenever the oracle frontier is not the endpoint chord. The compound proof neither requires nor establishes total-variation or Le Cam equivalence of the full experiments; root-n factor recovery alone would not justify such a claim. Instead it uses a loss-specific stability mechanism: one common residual displacement creates a small empirical threshold-crossing set, and the random ratios are stable after explicit denominator control. The separable lower bound identifies the complementary obstruction: a single raw response cannot translate itself by the unobserved factor.

Several extensions require additional arguments. For unknown π_1, μ, and ρ, theorem 7.7 shows that consistent plug-in estimates attain the same frontier with asymptotic FDR control; finite-sample FDR control, rates, and the construction of the estimators are not addressed. Multiple factors would replace the common scalar displacement by a low-rank fitted component whose coordinatewise geometry depends on the loadings. Uniform results as

$\rho \uparrow 1$ would also require thresholds and constants that remain controlled as the effective signal a diverges. The separable-gap proof uses both the full support of the Gaussian factor and independence of the external seed; discrete factors or factor-correlated side information need separate treatment. None of these extensions follows automatically from the pointwise results here, and the present paper does not characterize $\Psi_{\mathrm{dep}}(\alpha; P)$ across broad dependent classes.

AI assistance disclosure

OpenAI GPT-5.6 Sol, accessed through ChatGPT and Codex, was used to develop and audit the proofs and numerical checks, and for drafting. Anthropic Claude Opus 5.5, used through Claude Code, was used for editing, an independent check of Sections 6–7, and to derive Lemma 9.7 and Theorem 9.8, as well as Lemma 7.5 and Theorem 7.7. The author reviewed and takes responsibility for every statement and proof. AI systems are not authors of this work, and the work has not been peer reviewed.

References

[1] Yoav Benjamini and Yosef Hochberg. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1):289–300, 1995. doi:10.1111/j.2517-6161.1995.tb02031.x.

[2] Sylvain Delattre and Etienne Roquain. On the false discovery proportion convergence under Gaussian equi-correlation. *Statistics & Probability Letters*, 81(1):111–115, 2011. doi:10.1016/j.spl.2010.09.025.

[3] Jianqing Fan and Xu Han. Estimation of the false discovery proportion with unknown dependence. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 79(4):1143–1164, 2017. doi:10.1111/rssb.12204.

[4] Jianqing Fan, Xu Han, and Weijie Gu. Estimating false discovery proportion under arbitrary covariance dependence. *Journal of the American Statistical Association*, 107(499):1019–1035, 2012. doi:10.1080/01621459.2012.720478.

[5] Jianqing Fan, Yuan Ke, Qiang Sun, and Wen-Xin Zhou. FarmTest: Factor-adjusted robust multiple testing with approximate false discovery control. *Journal of the American Statistical Association*, 114(528):1880–1893, 2019. doi:10.1080/01621459.2018.1527700.

[6] Helmut Finner, Thorsten Dickhaus, and Markus Roters. Dependency and false discovery rate: Asymptotics. *The Annals of Statistics*, 35(4):1432–1455, 2007. arXiv:0710.3171, doi:10.1214/009053607000000046.

[7] Chloé Friguet, Maela Kloareg, and David Causeur. A factor model approach to multiple testing under dependence. *Journal of the American Statistical Association*, 104(488):1406–1415, 2009. doi:10.1198/jasa.2009.tm08332.

[8] Christopher Genovese and Larry Wasserman. Operating characteristics and extensions of the false discovery rate procedure. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(3):499–517, 2002. doi:10.1111/1467-9868.00347.

[9] Prasenjit Ghosh and Arijit Chakrabarti. Bayesian model pursuit and near-oracle sparse signal discovery under dependence, 2026. arXiv:2606.22490.

[10] Prasenjit Ghosh and Arijit Chakrabarti. Covariance-adaptive residualization and stagewise calibration for dependent multiple testing, 2026. Version 2, July 2026. arXiv:2606.07466.

[11] Ruth Heller and Saharon Rosset. Optimal control of false discovery criteria in the two-group model. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 83(1):133–155, 2021. doi:10.1111/rssb.12403.

[12] Erratum: Optimal control of false discovery criteria in the two-group model. Journal of the Royal Statistical Society Series B: Statistical Methodology, 83(4), 882, 2021. doi:10.1111/rssb.12441.

[13] Dan M. Kluger and Art B. Owen. A central limit theorem for the Benjamini–Hochberg false discovery proportion under a factor model. *Bernoulli*, 30(1):743–769, 2024. doi:10.3150/23-BEJ1615.

[14] Jeffrey T. Leek and John D. Storey. A general framework for multiple testing dependence. *Proceedings of the National Academy of Sciences*, 105(48):18718–18723, 2008. doi:10.1073/pnas.0808709105.

[15] Yutong Nie and Yihong Wu. Large-scale multiple testing: Fundamental limits of false discovery rate control and compound oracle. *The Annals of Statistics*, 54(2):647–666, 2026. doi:10.1214/25-AOS2581.

[16] Etienne Roquain and Fanny Villers. Exact calculations for false discovery proportion with application to least favorable configurations. *The Annals of Statistics*, 39(1):584–612, 2011. doi:10.1214/10-AOS847.

[17] Rahul Roy and Subir Kumar Bhandari. Asymptotic Bayes' optimality under sparsity for exchangeable dependent multivariate normal test statistics. *Statistics & Probability Letters*, 207:110030, 2024. doi:10.1016/j.spl.2023.110030.

[18] Wenguang Sun and T. Tony Cai. Large-scale multiple testing under dependence. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 71(2):393–424, 2009. doi:10.1111/j.1467-9868.2008.00694.x.

[19] Wensong Wu and Edsel A. Peña. Bayes multiple decision functions. *Electronic Journal of Statistics*, 7:1272–1300, 2013. arXiv:1110.0043, doi:10.1214/13-EJS813.

[20] Jichun Xie, T. Tony Cai, and Hongzhe Li. Correction to the paper "Optimal False Discovery Rate Control for Dependent Data". *Statistics and Its Interface*, 9(1):33–35, 2016. doi:10.4310/SII.2016.v9.n1.e1.

[21] Jichun Xie, T. Tony Cai, John Maris, and Hongzhe Li. Optimal false discovery rate control for dependent data. *Statistics and Its Interface*, 4(4):417–430, 2011. doi:10.4310/SII.2011.v4.n4.a1.

Department of Physics, École normale supérieure, Université PSL, Paris, France
Email address: samuil.petkov@phys.ens.psl.eu