When the Out-of-Sample-Optimal Schur Portfolio Lies Between HRP and Minimum Variance

Peter Cotton*

Abstract

We give analytical conditions under which the Schur portfolio with the best out-of-sample variance lies strictly inside the bridge, the one-parameter family we introduced to unify hierarchical risk parity with minimum variance, and exactly provable examples on both sides. Under stated endpoint conditions, cross-block estimation error separates the optimum from both ends, and an exchangeable four-asset family is solved exactly, its optimizer decreasing from full coupling to the middle of the bridge. The interior mechanism persists under whole-matrix noise, with a sharp second-order sign criterion at full coupling, and in sample-covariance experiments the median-optimal coupling climbs with sample size.

Keywords: portfolio construction; hierarchical risk parity; Schur complement; estimation error; shrinkage; bias–variance trade-off.
JEL: G11, C13, C58. **MSC2020:** 91G10, 62H12.

1 Introduction

Hierarchical risk parity [1] allocates by recursive bisection without inverting a covariance matrix. Minimum variance inverts it in full. The two are ends of a single construction: augment each block with a γ-scaled Schur complement of its sibling and the recursion is exactly HRP at $\gamma = 0$, while recovering minimum variance as $\gamma \to 1$. This unification first appeared in a blog [2], then in a textbook [3, §12.3.4], with a fuller development in [4]. It quickly found its way into practice: `skfolio`, a popular Python portfolio library, ships it as `SchurComplementary`. The shipped variant is exactly HRP at $\gamma = 0$ and returns minimum variance at $\gamma = 1$ only under additional structure, a distinction Section 6 makes exact. On the hierarchical graph structures of [6] the $\gamma = 1$ recovery is exact. The endpoints have since been compared analytically: [5] derive the weight noise of HRP and of Markowitz under covariance estimation error and prove HRP is the more robust.

The question behind the bridge is a form of the classical estimation-risk problem. [9] quantified the sampling error of efficient portfolios; [10] named the enigma; [11] showed that imposing the wrong constraints acts as shrinkage and helps; [12] documented how hard the $1/N$ portfolio is to beat; and [13] made shrinkage of the input systematic. The Schur coupling gives that trade-off a single parameter, and the present paper gives the parameter provable structure.

This paper asks where on the bridge one should sit when the covariance is estimated with error. The two ends turn out to be special in different ways. The hierarchical end never reads the cross-block, so cross-block noise cannot hurt it, and the question at that end is whether coupling has

*Microprediction. Email: `peter.cotton@microprediction.com`. The machine-verifiable certificates for the exact claims are in `paper/neither_end_certificate.py` at `github.com/microprediction/schur`; implementations of the recursion are in the `allocation` and `precise` packages.

enough population value to pay its noise cost. The minimum-variance end is flat in the population, so any positive marginal noise sensitivity moves the optimum inside. Between the two observations sits an exchangeable four-asset family whose expected out-of-sample variance can be computed exactly, and there the optimum is unique, interior, and slides from full coupling toward the middle of the bridge as the noise grows.

The restriction to cross-block noise is then dropped. With noise in every entry of the covariance the same family still prefers the interior, under an explicit and sharp condition on the noise's second moments, and in Monte Carlo under the sample covariance of T Gaussian draws the median-optimal coupling climbs from 0.62 to 0.96 over the sample sizes examined: sample size selects the seat. We are explicit about what is not claimed. Covariance error does not always force an interior optimum, and either endpoint can be optimal on a nontrivial interval of noise levels. Exact examples mark those boundaries, and every exact claim about them is verified in rational arithmetic.

2 The recursion

The family. Following [4], fix an asset order and a balanced bisection tree; assume for notational simplicity that n is a power of two, so sibling blocks have equal size (a step-up matrix enters in the general case [4]). A hierarchical allocation assigns a node's capital to its children (L, R) in the ratio $1/\nu(\cdot_L) : 1/\nu(\cdot_R)$ for an inverse-fitness functional ν, and recurses. Writing a node's block of the covariance as $\left(\begin{smallmatrix} A & B \\ B^\top & D \end{smallmatrix}\right)$, hierarchical risk parity passes A and D to the children unchanged. The Schur-complementary family instead passes *augmented* blocks built from the γ-damped Schur complement and its companion vector,

$$A^c(\gamma) \;=\; A - \gamma\, BD^{-1}B^\top, \qquad b_A(\gamma) \;=\; \mathbf{1} - \gamma\, BD^{-1}\mathbf{1}, \tag{1}$$

and symmetrically for $D^c(\gamma), b_D(\gamma)$. The pair form is primary. For SPD Q with blocks A, B, D and any $b = (b_L; b_R)$,

$$Q^{-1}b \;=\; \begin{pmatrix} (A - BD^{-1}B^\top)^{-1}\,(b_L - BD^{-1}b_R) \\ (D - B^\top A^{-1}B)^{-1}\,(b_R - B^\top A^{-1}b_L) \end{pmatrix},$$

so recursive exactness follows by induction on the tree with the companion vectors carried alongside the blocks. The canonical presentation of [4] encodes the pair in a single matrix, dividing $A^c(\gamma)$ elementwise by $b_A b_A^\top$ and evaluating the inter-group fitness on $\left(A^c(\gamma)^{-1} \circ b_A b_A^\top\right)^{-1}$. That encoding requires every coordinate of b_A to be nonzero, which can fail on ordinary SPD input (a 4×4 example with $b_A = (0, 1)^\top$ has leading principal minors $2, 4, \frac{7}{2}, 3$), so we treat it as an optional representation and work with the pair. At $\gamma = 0$ every member is HRP. At $\gamma = 1$ the scheme replicates the minimum-variance portfolio; the one-step form of that equivalence is Proposition 1, following Appendix A of [4].

Proposition 1 ($\gamma = 1$ replicates minimum variance)**.** *Let $\Sigma \succ 0$ have blocks A, B, D as above and write $A^c = A - BD^{-1}B^\top$, $b_A = \mathbf{1} - BD^{-1}\mathbf{1}$, and symmetrically D^c, b_D. Then*

$$\Sigma^{-1}\mathbf{1} \;=\; \begin{pmatrix} (A^c)^{-1}b_A \\ (D^c)^{-1}b_D \end{pmatrix}. \tag{2}$$

For any SPD Q and $b \neq 0$, the constrained problem $w(Q, b) = \arg\min_{w:\, b^\top w = 1} w^\top Q w$ has solution and attained variance

$$w(Q, b) \;=\; \frac{Q^{-1}b}{b^\top Q^{-1}b}, \qquad \nu(Q, b) \;=\; w(Q, b)^\top Q\, w(Q, b) \;=\; \frac{1}{b^\top Q^{-1}b}, \tag{3}$$

so that $Q^{-1}b = w(Q,b)/\nu(Q,b)$. Consequently the two-block scheme that allocates between blocks in the ratio $1/\nu(A^c, b_A) : 1/\nu(D^c, b_D)$ and within blocks by $w(A^c, b_A)$, $w(D^c, b_D)$ reproduces $w \propto \Sigma^{-1}\mathbf{1}$, and recursive application to the sub-problems (each again of the form $Q^{-1}b$) extends this down the tree. If a block has $b_A = 0$ then by (2) that block of $\Sigma^{-1}\mathbf{1}$ is exactly zero and the block receives nothing; the normalized $w(Q,b)$ is simply not formed. (For instance $\Sigma = \left(\begin{smallmatrix} 2 & 1 \\ 1 & 1 \end{smallmatrix}\right)$ has $b_A = 0$ and $\Sigma^{-1}\mathbf{1} = (0,1)^\top$.)

Proof. For (2), multiply the block inversion identity

$$\begin{pmatrix} A & B \\ B^\top & D \end{pmatrix}^{-1} = \begin{pmatrix} (A^c)^{-1} & 0 \\ 0 & (D^c)^{-1} \end{pmatrix} \begin{pmatrix} I & -BD^{-1} \\ -B^\top A^{-1} & I \end{pmatrix}$$

on the right by $\mathbf{1}$; the top block is $(A^c)^{-1}(\mathbf{1} - BD^{-1}\mathbf{1}) = (A^c)^{-1}b_A$ and the bottom block is analogous. For (3), the Lagrangian condition is $Qw = \lambda b$, so $w = \lambda Q^{-1}b$; the constraint $b^\top w = 1$ forces $\lambda = 1/(b^\top Q^{-1} b)$, giving the stated $w(Q,b)$, and substituting back gives $\nu(Q,b) = \lambda$. Rearranging, $Q^{-1}b = w(Q,b)\,(b^\top Q^{-1}b) = w(Q,b)/\nu(Q,b)$. Apply this to each block of (2): the unnormalized block solutions are the normalized sub-portfolios scaled by exactly the inverse fitness $1/\nu$, which is the stated two-block scheme. □

Remark 1 (Minimum variance in disguise). The constrained problem (3) was stated for general b, and that generality is not idle. Several portfolio problems that look different are minimum variance with a prescribed unit exposure, minimizing $w^\top Q w$ subject to $b^\top w = 1$, and are therefore minimum variance in disguise: the tangency portfolio ($b = \mu$), maximum diversification ($b = \sigma$, the vector of volatilities), and the variance-optimal combination of forecasts, where Q is the covariance of forecast errors and $b = \mathbf{1}$ [7]. Tracking error against a benchmark becomes minimum variance in active weights, and hedging a single beta exposure takes b to be that exposure; their generic versions carry several affine constraints at once and sit outside (3). The block identity (2) and the damped bridge apply to each with $\mathbf{1}$ replaced by b throughout. The noise calculus of this paper is developed for $b = \mathbf{1}$; we expect, but do not prove here, that the endpoint results carry over with the companion vectors $b_A(\gamma)$ built from b.

The variant analyzed. Appendix A of [4] licenses an equivalent "collapse" of the pair (A^c, b_A) into a single matrix $\widetilde{A}$ with $\widetilde{A}^{-1}\mathbf{1} = A^c(\gamma)^{-1}b_A(\gamma)$. The reference implementations take exactly this route.[1] It is the recursion we analyze. Each child block is replaced by

$$A_\gamma \;=\; \mathrm{sym}\Big[\big(I - \gamma B D^{-1}\big)^{-1} A^c(\gamma)\Big], \qquad D_\gamma \;=\; (A \leftrightarrow D,\; B \to B^\top), \tag{4}$$

where $\mathrm{sym}(X) = \frac{1}{2}(X + X^\top)$; note $(I - \gamma BD^{-1})\mathbf{1} = b_A(\gamma)$, so (4) is the collapse of (1), up to the symmetrization. The fitness is the *naive portfolio variance*

$$\nu(X) \;=\; \tilde{w}^\top X\, \tilde{w}, \qquad \tilde{w}_i \;=\; \frac{1/X_{ii}}{\sum_j 1/X_{jj}}, \tag{5}$$

so the left child receives the fraction $\nu(D_\gamma)/\big(\nu(A_\gamma) + \nu(D_\gamma)\big)$ of the node's capital, and the recursion continues inside each augmented block down to singletons. Nodes with a singleton child perform no augmentation; the implementations skip the augmentation whenever *either* child has one asset, so a two-asset node splits by the inverse of its already augmented diagonal entries. The variant is

[1] `skfolio`'s `SchurComplementary` and the `allocation` package use identical augmentation and fitness routines; see `github.com/skfolio/skfolio` and `github.com/microprediction/allocation`.

best read as a principled top-down family, inspired by HRP, that allows the user to move "in the direction" of an explicit global optimization in a manner we make precise. Its endpoints deserve care at the outset. At $\gamma = 0$ the augmentation (4) is the identity and the recursion is exactly HRP over the fixed order. At $\gamma = 1$ it reproduces minimum variance only under additional structure: the exchangeable structure of Lemma 3 suffices, and Section 6.2 gives an integer counterexample. The symmetrization is exact for a single level of coupling, because one level reads only diagonals and quadratic forms of the augmented blocks, and both see only the symmetric part; on deeper trees it matters, and the operative gap at $\gamma = 1$ is the naive splits. Implementations optionally cap γ at an in-sample variance turning point (`keep_monotonic`); we analyze the raw recursion. One thing cannot be simplified away. The inter-group fitness must be a quadratic form evaluated on the *augmented* block. A diagonal-only ν, as used in some HRP variants, generally does not preserve the Schur block-inversion identity (2) that exact minimum-variance recovery requires.

Two structural facts hold across the entire family and not just the variant (4): by (1) the cross-block B enters every member only through the products $\gamma B D^{-1}(\cdot)$, and at $\gamma = 0$ the augmentation is the identity with $b_A = \mathbf{1}$.

The noise model. Let $\Sigma \succ 0$ be the population covariance and let the estimate be

$$\widehat{\Sigma}_\tau \;=\; \Sigma + \tau N, \tag{6}$$

where N is a random symmetric matrix. The objects of study are the expected out-of-sample variance and its population limit,

$$F(\gamma;\tau) \;=\; \mathbb{E}\big[\, w(\gamma;\widehat{\Sigma}_\tau)^\top \Sigma\, w(\gamma;\widehat{\Sigma}_\tau)\,\big], \qquad V_0(\gamma) \;=\; F(\gamma;0), \tag{7}$$

where $w(\gamma; M)$ denotes the allocation the recursion produces from an input matrix M. The allocation w and the realized variance $w^\top \Sigma w$ are rational functions of (γ, M), hence C^∞ wherever their denominators do not vanish. When N has finite support (as in all examples below) F is itself a rational function of γ; for general N it is merely smooth, under the assumptions we now state.

Assumption 1. *There are $\bar{\gamma} \in (0, 1]$, $\bar{\tau} > 0$, and a compact set $K = [0, \bar{\gamma}] \times \mathcal{M}$, with $\mathcal{M}$ a compact neighborhood of Σ in the symmetric matrices, such that every denominator appearing in the recursion (the inverses in* (4)*, the diagonal entries in* (5)*, and the fitness sums) is bounded away from zero on K; and N is bounded with $\Sigma + \tau N \in \mathcal{M}$ for all $|\tau| \leq \bar{\tau}$.*

Assumption 2. *N is supported on the top-level cross-block ($N_{LL} = 0$, $N_{RR} = 0$) and its law is symmetric: $N \overset{d}{=} -N$.*

Under Assumption 1, $(\gamma, \tau) \mapsto w(\gamma; \Sigma + \tau N)$ is C^∞ on $[0, \bar{\gamma}] \times [-\bar{\tau}, \bar{\tau}]$ for every realization, with all derivatives bounded uniformly over realizations; F is then C^∞ by differentiation under the expectation (the derivatives are dominated by their uniform bounds). Assumption 2 makes F even in τ.

Remark 2 (Model scope). The model (6)–(7) with Assumption 2 is stylized on purpose: it isolates the cross-block channel, which is the only channel the coupling γ touches. A natural generative alternative is partial information about Σ itself, for instance $\widehat{\Sigma}$ equal to the sample covariance of T independent draws from $\mathcal{N}(0, \Sigma)$, which corresponds to $\tau \asymp T^{-1/2}$ with noise supported on *every* block. Under whole-matrix noise the exact invariance of Lemma 1 applies only to the cross-block component of the noise: $F(0;\tau)$ is then itself noisy, and the endpoint comparison becomes a comparison of noise costs rather than a comparison against a noise-free anchor. Sections 3–5 are exact for (6) under Assumptions 1–2. Section 6 then drops Assumption 2 for the solved family and answers the whole-matrix question directly.

3 The hierarchical end

Lemma 1 (The HRP end is blind to the cross-block)**.** *For a fixed tree, $w(0;M)$ is a function of the diagonal blocks (A,D) alone. Consequently, under Assumption 2, $w(0;\widehat{\Sigma}_\tau) = w(0;\Sigma)$ almost surely and $F(0;\tau) = V_0(0)$ exactly, for every τ.*

Proof. At $\gamma = 0$ the augmentation (4) returns its first argument. The top-level split therefore uses $\nu(A)$ and $\nu(D)$, and (5) reads only entries of A and of D. The recursion then continues inside the unmodified blocks A and D; every quantity computed below the top level is a function of entries of A or entries of D. The top-level cross-block B appears nowhere. (Cross-blocks interior to A or D *are* read, through the naive variances of their parents; only the top-level B is never read.) The second sentence follows because $\widehat{\Sigma}_\tau$ and Σ share their diagonal blocks under Assumption 2. □

Remark 3 (Scope). The lemma concerns a fixed tree and this particular variant: the recursion reads the naive variance of full child sub-blocks, as in [1]. A split on diagonal variances alone would read nothing off-diagonal. When the tree itself is data-driven (a dendrogram, or Fiedler seriation [4]), the ordering step does read the cross-block, but only through the discrete tree choice, which for a generic estimate is locally constant; the differential statements below are unaffected for noise small enough to leave the tree unchanged. (In Example 2 the seriation returns the same top partition on both noise branches, so fixing the order there is without loss.)

Theorem 1 (Uniform separation from HRP under small noise)**.** *Suppose Assumption 1 holds, N is bounded (no support or symmetry restriction needed), and $V_0'(0) < 0$, i.e. coupling has first-order population value. Then there exist $\varepsilon > 0$, $\tau_0 > 0$, and $c > 0$ such that*

$$\partial_\gamma F(\gamma;\tau) \;\le\; -c \qquad \text{for all } 0 \le \gamma \le \varepsilon,\; 0 \le \tau \le \tau_0.$$

Consequently every minimizer of $F(\cdot\,;\tau)$ over $[0,\bar{\gamma}]$ lies in $[\varepsilon,\bar{\gamma}]$, for every $\tau \le \tau_0$.

Proof. Write $V(\gamma;M) = w(\gamma;M)^\top \Sigma\, w(\gamma;M)$. Under Assumption 1, $\partial_\gamma V$ exists, is continuous, and is bounded on K, so $\partial_\gamma F(\gamma;\tau) = \mathbb{E}[\partial_\gamma V(\gamma;\Sigma+\tau N)]$ by differentiation under the expectation, and $(\gamma,\tau) \mapsto \partial_\gamma F(\gamma;\tau)$ is jointly continuous on $[0,\bar{\gamma}]\times[0,\bar{\tau}]$ by dominated convergence. Its value at $(0,0)$ is $V_0'(0) < 0$. By continuity there is a relatively open neighborhood of $(0,0)$ on which $\partial_\gamma F \le \frac{1}{2}V_0'(0) =: -c < 0$; choose ε, τ_0 so that $[0,\varepsilon]\times[0,\tau_0]$ sits inside it. For $\tau \le \tau_0$ and any $\gamma \in [0,\varepsilon)$, integrating the derivative gives $F(\varepsilon;\tau) \le F(\gamma;\tau) - c(\varepsilon-\gamma) < F(\gamma;\tau)$, so no point of $[0,\varepsilon)$ can be a minimizer. □

3.1 The size of the noise cost

How large is $F - V_0$ near the HRP end? The uniform bound is of order $\gamma\tau^2$ and the exponent is generically sharp, and it is worth displaying the mechanism, because a plausible shortcut gives the wrong exponent. Write the estimated cross-block as $\widehat{B} = B + \tau E$ for a realization E of N's cross-block. Expanding (4) at small γ,

$$A_\gamma(\widehat{B}) \;=\; A + \gamma\,\mathrm{sym}\big(\widehat{B}D^{-1}A - \widehat{B}D^{-1}\widehat{B}^\top\big) + O(\gamma^2). \tag{8}$$

Averaging over the symmetric branches $\widehat{B} = B \pm \tau E$, the terms linear in τ cancel, but the term quadratic in $\widehat{B}$ does not:

$$\tfrac{1}{2}\sum_{\pm} A_\gamma(B \pm \tau E) \;-\; A_\gamma(B) \;=\; -\,\gamma\,\tau^2\, ED^{-1}E^\top \;+\; O(\gamma^2\tau^2). \tag{9}$$

The expected augmented block is displaced at order $\gamma\tau^2$, and the allocation inherits the displacement unless the weight functional happens to annihilate it. In terms of derivatives of the realized variance,

$$\frac{\partial^2}{\partial\tau^2}(w^\top \Sigma w) \;=\; 2\, w_\tau^\top \Sigma\, w_\tau \;+\; 2\, w_{\tau\tau}^\top \Sigma\, w, \tag{10}$$

and while $w_\tau = O(\gamma)$ (each appearance of $\widehat{B}$ in (4) carries a factor γ), the second derivative $w_{\tau\tau}$ is only $O(\gamma)$ as well, through the quadratic term visible in (9). The first term of (10) is $O(\gamma^2)$; the second is generically $O(\gamma)$. The shortcut "$w_\tau = O(\gamma)$, therefore $F - V_0 = O(\gamma^2\tau^2)$" bounds only the first term and is false; Example 1 makes the failure exact.

Theorem 2 (Noise cost near the HRP end)**.** *Under Assumptions 1 and 2 there is a continuous bounded function H on $[0,\bar{\gamma}] \times [0,\bar{\tau}]$ with*

$$F(\gamma;\tau) - V_0(\gamma) \;=\; \gamma\,\tau^2\, H(\gamma,\tau), \qquad \textit{hence} \qquad \big|F(\gamma;\tau) - V_0(\gamma)\big| \;\le\; C\,\gamma\,\tau^2 \tag{11}$$

with $C = \sup|H|$. The exponent of γ cannot be improved: in Example 1, $H(0,0) = \frac{8}{189} \neq 0$.

Proof. Let $\Phi(\gamma,\tau) = F(\gamma;\tau) - V_0(\gamma)$, which is C^∞ on the compact rectangle $R = [0,\bar{\gamma}] \times [-\bar{\tau},\bar{\tau}]$ by Assumption 1. Three exact identities hold on R: $\Phi(0,\tau) = 0$ for all τ (Lemma 1); $\Phi(\gamma,0) = 0$ for all γ (no noise); and $\Phi(\gamma,-\tau) = \Phi(\gamma,\tau)$ (Assumption 2). Fix γ and Taylor-expand in τ with integral remainder: since $\Phi(\gamma,0) = 0$ and, by evenness, $\partial_\tau\Phi(\gamma,0) = 0$,

$$\Phi(\gamma,\tau) \;=\; \tau^2 \int_0^1 (1-s)\,\partial^2_{\tau\tau}\Phi(\gamma,s\tau)\,ds \;=:\; \tau^2\,\Psi(\gamma,\tau).$$

Ψ is C^1 on R (differentiation under the integral, with the third derivatives of Φ bounded on R). Differentiating the identity $\Phi(0,\tau) = 0$ twice in τ gives $\partial^2_{\tau\tau}\Phi(0,\tau) = 0$ for all τ, hence $\Psi(0,\tau) = 0$. Taylor in γ with integral remainder then gives

$$\Psi(\gamma,\tau) \;=\; \gamma \int_0^1 \partial_\gamma\Psi(u\gamma,\tau)\,du \;=:\; \gamma\, H(\gamma,\tau),$$

with H continuous and bounded on R by the same uniform derivative bounds. Combining the two displays yields (11). The sharpness claim is Example 1, where $\partial_\gamma(F - V_0)(0;\tau) = \frac{8}{189}\tau^2$ exactly. $\square$

With C^4 regularity in τ (available under Assumption 1), the same identities give the uniform expansion

$$F(\gamma;\tau) \;=\; V_0(\gamma) + \tau^2 G(\gamma) + O(\tau^4), \qquad G(\gamma) \;=\; \tfrac{1}{2}\,\partial^2_{\tau\tau}F(\gamma;0) \;=\; \gamma\, h(\gamma), \tag{12}$$

with h continuous, so $G(0) = 0$, but in general $G'(0) = h(0) \neq 0$: G vanishes only linearly at the HRP end. The exact boundary derivative is

$$\partial_\gamma F(0;\tau) \;=\; V_0'(0) + \tau^2\, H(0,\tau). \tag{13}$$

Remark 4 (When is the stronger $\gamma^2\tau^2$ bound valid?). The bound $|F - V_0| \le C\gamma^2\tau^2$ is recoverable under one additional, substantive hypothesis: that the first coupling variation is unbiased under the noise, i.e. $\partial_\gamma F(0;\tau) = V_0'(0)$ for all small τ. Then both the value and the first γ-derivative of $F - V_0$ vanish at $\gamma = 0$ and a third application of the Taylor argument yields $F - V_0 = \gamma^2\tau^2\widetilde{H}$. The recursion does not satisfy this generically, because

$$\mathbb{E}\big[(B + \tau N_B)\, D^{-1}(B + \tau N_B)^\top\big] \;=\; BD^{-1}B^\top + \tau^2\,\mathbb{E}\big[N_B D^{-1} N_B^\top\big],$$

and that quadratic bias enters (4) multiplied by a single power of γ. Both noise models on the solved family's covariance (16) happen to possess a further downstream cancellation (Remark 10): the common cross-correlation noise of Section 5 and the one-entry noise of Example 2. Numerical evidence from them alone can therefore suggest, wrongly, that the quadratic bound is general. Example 1 has no such cancellation.

Example 1 (HRP locally optimal at a finite noise level)**.** *Let $n = 4$ with*

$$A = D = \begin{pmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 2 \end{pmatrix}, \qquad B = \tfrac{1}{10}\,\mathbf{1}\mathbf{1}^\top, \qquad E = e_1 e_1^\top, \qquad \widehat{B} = B \pm \tau E \ \ \textit{w.p. } \tfrac{1}{2} \textit{ each,}$$

and $\Sigma = \left(\begin{smallmatrix} A & B \\ B & A \end{smallmatrix}\right)$, which is SPD (both $A + B$ and $A - B$ are positive definite). Then, exactly and for every τ,

$$\partial_\gamma F(0;\tau) \;=\; -\frac{1}{700} + \frac{8}{189}\,\tau^2. \tag{14}$$

In particular $V_0'(0) = -\frac{1}{700} < 0$, so the premise of Theorem 1 holds, yet $\partial_\gamma F(0;\tau) > 0$ as soon as $\tau^2 > \frac{27}{800}$, i.e. $\tau > \sqrt{27/800} \approx 0.1837$: the HRP endpoint is then a one-sided local minimum. In addition $F(\gamma;\tau) - V_0(\gamma) = \frac{8}{189}\gamma\tau^2 + O(\gamma^2)$, so no bound of the form $C\gamma^2\tau^2$ can hold on any neighborhood of $\gamma = 0$.

Proof. The computation is elementary and worth walking through. *Step 1 (symmetry).* Since $A = D$ and each realized $\widehat{B}$ is symmetric, the two augmented blocks in (4) coincide, so the top-level split is $\frac{1}{2} : \frac{1}{2}$ for every γ, τ, and branch, and $w = \frac{1}{2}(u;u)$ where u are the inverse-diagonal weights (5) of the common augmented block $A_\gamma(\widehat{B})$. *Step 2 (objective).* With the true Σ,

$$V \;=\; w^\top \Sigma w \;=\; \tfrac{1}{4}(u^\top A u + 2\,u^\top B u + u^\top A u) \;=\; \tfrac{1}{2}\,u^\top (A+B)\,u.$$

Step 3 (first-order augmentation). By (8) with $D = A$ and $\widehat{B}$ symmetric, $A_\gamma(\widehat{B}) = A + \gamma H(\widehat{B}) + O(\gamma^2)$ with $H(\widehat{B}) = \widehat{B} - \widehat{B}A^{-1}\widehat{B}$; the symmetrization does not alter this since $H(\widehat{B})$ is symmetric. *Step 4 (weight derivative).* The weights (5) depend only on the diagonal. With $\operatorname{diag}(A) = (1, 2)$, $u(0) = (\frac{2}{3}, \frac{1}{3})$, and differentiating $u(\gamma) = (A_{22} + \gamma H_{22},\, A_{11} + \gamma H_{11})/(3 + \gamma(H_{11} + H_{22}))$ at $\gamma = 0$ gives

$$u'(0) \;=\; \frac{H_{22} - 2H_{11}}{9}\begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Step 5 (variance derivative). With $M = A + B$ one computes $M(1,-1)^\top = (\frac{3}{2}, -\frac{5}{2})^\top$ and $u(0)^\top M(1,-1)^\top = \frac{2}{3}\cdot\frac{3}{2} - \frac{1}{3}\cdot\frac{5}{2} = \frac{1}{6}$, so

$$\partial_\gamma V\big|_{\gamma=0} \;=\; u(0)^\top M\,u'(0) \;=\; \frac{H_{22} - 2H_{11}}{54}.$$

Step 6 (average over branches). H is quadratic in $\widehat{B}$, so averaging is exact:

$$\tfrac{1}{2}\sum_{\pm} H(B \pm \tau E) \;=\; B - BA^{-1}B - \tau^2\, EA^{-1}E.$$

With $A^{-1} = \frac{1}{7}\left(\begin{smallmatrix} 8 & 2 \\ 2 & 4 \end{smallmatrix}\right)$ one gets $BA^{-1}B = \frac{4}{175}\,\mathbf{1}\mathbf{1}^\top$ and $EA^{-1}E = \frac{8}{7}\,e_1 e_1^\top$, so the averaged H has $\bar{H}_{11} = \frac{1}{10} - \frac{4}{175} - \frac{8}{7}\tau^2$ and $\bar{H}_{22} = \frac{1}{10} - \frac{4}{175}$. Substituting into Step 5,

$$\partial_\gamma F(0;\tau) \;=\; \frac{\bar{H}_{22} - 2\bar{H}_{11}}{54} \;=\; \frac{-\frac{27}{350} + \frac{16}{7}\tau^2}{54} \;=\; -\frac{1}{700} + \frac{8}{189}\,\tau^2.$$

Since F is smooth in γ near 0 (denominators reduce to those of A at $\gamma = 0$), the exchange of expectation and derivative is justified by the finite support of the noise. The certificate script confirms (14) in exact rational arithmetic, by dual-number (first-order jet) evaluation of the full recursion, at $\tau \in \{0, \frac{1}{10}, \frac{3}{20}, \frac{1}{5}, \frac{3}{10}\}$: for instance $\partial_\gamma F(0; \frac{3}{20}) = -\frac{1}{2100}$ and $\partial_\gamma F(0; \frac{1}{5}) = +\frac{1}{3780}$. $\square$

4 The minimum-variance end

Suppose the bridge under study reaches minimum variance exactly at $\gamma = 1$: writing $w_* = \Sigma^{-1}\mathbf{1}/(\mathbf{1}^\top\Sigma^{-1}\mathbf{1})$, suppose $w(1;\Sigma) = w_*$ while $\mathbf{1}^\top w(\gamma;\Sigma) = 1$ for all γ. This holds for the canonical member of the family (Proposition 1) and, in the solved family of Section 5, for the implemented variant as well; it is not automatic for the implemented variant in general.

Lemma 2 (Flatness at an exact minimum-variance endpoint)**.** *Under the above, assume also that $\gamma \mapsto w(\gamma;\Sigma)$ is twice differentiable in a neighborhood of $\gamma = 1$; the recursion is rational in γ, so this holds wherever its denominators are nonzero. Then*

$$V_0'(1) \;=\; 0, \qquad V_0''(1) \;=\; 2\, w_\gamma(1)^\top \Sigma\, w_\gamma(1) \;\geq\; 0,$$

where $w_\gamma(1) = \partial_\gamma w(\gamma;\Sigma)|_{\gamma=1}$; the second derivative is strictly positive whenever the bridge arrives with a nonzero tangent, $w_\gamma(1) \neq 0$.

Proof. Since $\Sigma w_* = \lambda\mathbf{1}$ with $\lambda = 1/(\mathbf{1}^\top\Sigma^{-1}\mathbf{1})$, differentiating $V_0(\gamma) = w(\gamma)^\top\Sigma\, w(\gamma)$ at $\gamma = 1$ gives $V_0'(1) = 2\, w_\gamma(1)^\top\Sigma w_* = 2\lambda\, w_\gamma(1)^\top\mathbf{1} = 0$, because $\mathbf{1}^\top w(\gamma) \equiv 1$ implies $\mathbf{1}^\top w_\gamma \equiv 0$. Differentiating once more, $V_0''(1) = 2w_\gamma(1)^\top\Sigma w_\gamma(1) + 2w_{\gamma\gamma}(1)^\top\Sigma w_*$, and the second term vanishes for the same reason ($\mathbf{1}^\top w_{\gamma\gamma} \equiv 0$). □

Theorem 3 (Small noise moves the optimum inside)**.** *Assume the setting of Lemma 2 with $V_0''(1) > 0$; assume $F(\gamma;\tau)$ is smooth in τ and even in τ, uniformly for γ near 1, as holds under Assumption 1 together with any sign-symmetric noise law, so that (12) holds in C^2 with $G = \frac{1}{2}\partial_\tau^2 F|_{\tau=0}$; and assume $G'(1) > 0$. Then there are $\tau_1 > 0$ and a C^1 curve $\gamma^*(\tau)$, defined for $0 \leq \tau < \tau_1$, of local minimizers of $F(\cdot\,;\tau)$ with*

$$\gamma^*(\tau) \;=\; 1 - \frac{G'(1)}{V_0''(1)}\,\tau^2 + O(\tau^4) \;<\; 1 \quad (\tau > 0). \tag{15}$$

If moreover $F(\cdot\,;\tau) \to V_0$ uniformly on $[0,\bar{\gamma}]$ as $\tau \to 0$, as follows from (11) under Assumptions 1 and 2, and $\gamma = 1$ is the unique global minimizer of V_0 on $[0,\bar{\gamma}]$, then for all sufficiently small τ the curve gives the unique global minimizer of $F(\cdot\,;\tau)$. The sign condition is essential: if $G'(1) < 0$, full coupling remains locally optimal for small noise.

Proof. Let $g(\gamma, s) = \partial_\gamma F(\gamma;\sqrt{s})$ for $s = \tau^2 \geq 0$. This is C^1 through $s = 0$: F is smooth and even in τ by hypothesis, and a smooth even function of τ is a smooth function of τ^2 (Whitney), so the regularity does not rest on the $O(s^2)$ notation alone. By (12) in C^2, $g(\gamma, s) = V_0'(\gamma) + s\, G'(\gamma) + O(s^2)$ with the error uniform in γ near 1. By Lemma 2, $g(1, 0) = V_0'(1) = 0$ and $\partial_\gamma g(1, 0) = V_0''(1) > 0$. The implicit function theorem gives a C^1 curve $\gamma^*(s)$ with $g(\gamma^*(s), s) = 0$, $\gamma^*(0) = 1$, and

$$\frac{d\gamma^*}{ds}\Big|_{s=0} \;=\; -\,\frac{\partial_s g(1,0)}{\partial_\gamma g(1,0)} \;=\; -\,\frac{G'(1)}{V_0''(1)} \;<\; 0,$$

which is (15). Since $\partial_\gamma g > 0$ near $(1, 0)$, each $\gamma^*(s)$ is a strict local minimizer. For global optimality: $F(\cdot\,;\tau) \to V_0$ uniformly on $[0,\bar{\gamma}]$ as $\tau \to 0$ (by (11) or directly by uniform continuity), so if $\gamma = 1$ is the unique global minimizer of V_0, minimizers of F must enter any prescribed neighborhood of 1 for small τ (outside such a neighborhood V_0 exceeds $V_0(1)$ by a positive margin, by compactness), where strict convexity of F in γ (from $\partial_\gamma g > 0$) makes $\gamma^*(s)$ the only candidate. If $G'(1) < 0$ the same computation gives $d\gamma^*/ds > 0$: the unconstrained stationary point exits the interval at $\gamma = 1$ and full coupling is locally optimal. □

5 An exactly solved family

Fix the four-asset covariance

$$\Sigma \;=\; \begin{pmatrix} 1 & \frac{1}{2} & \frac{3}{5} & \frac{3}{5} \\ \frac{1}{2} & 1 & \frac{3}{5} & \frac{3}{5} \\ \frac{3}{5} & \frac{3}{5} & 4 & 2 \\ \frac{3}{5} & \frac{3}{5} & 2 & 4 \end{pmatrix}, \tag{16}$$

i.e. blocks $\{1,2\}$ and $\{3,4\}$ with unit and double volatilities, within-block correlation $\frac{1}{2}$, and cross-block correlation $\rho_c = \frac{3}{10}$. Let the noise perturb the cross-block correlation *in common*: all four cross entries move together, $\rho_c \to \rho_c \pm \tau$ with probability $\frac{1}{2}$ each. At the level of the two blocks' equal-weight portfolios, the population variances and cross covariance are

$$a = \tfrac{3}{4}, \qquad d = 3, \qquad c = \tfrac{3}{5}, \qquad S := a + d = \tfrac{15}{4}, \qquad K := S - 2c = \tfrac{51}{20},$$

and the estimated group-level cross covariance is $Z = c \pm e$ with $e = 2\tau$ (the blocks have volatilities 1 and 2).

Lemma 3 (Exact reduction to a scalar bridge)**.** *Assume $a < d$ and $4c^2 < ad$, as in* (16)*, and fix $\gamma \in [0,1]$ and a realized $z = c \pm e$ with $0 \le z \le 2c$. The augmentation* (4) *of the D block requires the factor $I - \gamma B^\top A^{-1}$ to be invertible, and that holds precisely when $\gamma z \ne a$; the A-side factor $I - \gamma B D^{-1}$ is invertible for every such (γ, z), since $\gamma z \le 2c < d$. Wherever the recursion is defined, its within-block weights are exactly $(\frac{1}{2}, \frac{1}{2})$ in each block, the augmented fitnesses are*

$$\nu_{A_\gamma}(z) \;=\; \frac{ad - \gamma z^2}{d - \gamma z}, \qquad \nu_{D_\gamma}(z) \;=\; \frac{ad - \gamma z^2}{a - \gamma z}, \tag{17}$$

and the left-block weight is

$$x_\gamma(z) \;=\; \frac{d - \gamma z}{a + d - 2\gamma z}, \tag{18}$$

with realized out-of-sample variance $Q(x_\gamma(z))$, where $Q(x) = ax^2 + d(1-x)^2 + 2c\,x(1-x)$. The formula (18) *is a rational function of (γ, z) whose denominator obeys $S - 2\gamma z \ge S - 4c > 0$. It therefore extends the recursion's group weight continuously through the point $\gamma z = a$, where the extension gives the D block the weight $1 - x_\gamma(z) = (a - \gamma z)/(S - 2\gamma z) = 0$. This is the analog of the $b = 0$ convention in Proposition 1: a block whose augmented fitness diverges receives nothing. We call the allocation defined by* (18) *on all of $[0,1]$ the* extended *allocation and write F_{ext} for its expected out-of-sample variance. For the numerical family* (16) *the ordering $c = \frac{3}{5} < \frac{3}{4} = a$ holds, so the lower branch has $z = c - e < a$ throughout and never meets $\gamma z = a$; the sole possible singularity is on the upper branch, at $\gamma = a/(c+e)$, and it lies in $[0,1]$ precisely when $e \ge a - c = \frac{3}{20}$. In addition, for* (16) *every other denominator the recursion divides by (the augmented diagonal entries and the fitness sum) is nonzero on the admissible range, so $\gamma z = a$ is the only point at which the recursion is undefined. Without $c < a$ both branches can be singular; $a = 1$, $d = 6$, $c = \frac{11}{10}$ satisfies $a < d$, $4c^2 < ad$, and $S > 6c$, yet at $e = \frac{1}{20}$ both $\gamma = \frac{20}{21}$ (lower branch) and $\gamma = \frac{20}{23}$ (upper branch) are singular.*

Proof. Write the asset-level blocks as $A_2 = \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix}$, $D_2 = \left(\begin{smallmatrix} 4 & 2 \\ 2 & 4 \end{smallmatrix}\right)$, and $\widehat{B} = z\,\mathbf{1}\mathbf{1}^\top$ (each cross entry equals $1 \cdot 2 \cdot (\rho_c \pm \tau) = z$). All three matrices are exchangeable (invariant under swapping the two assets of a block), and the augmentation (4) preserves exchangeability, so each augmented block has equal diagonal entries and (5) gives $(\frac{1}{2}, \frac{1}{2})$ within blocks. Invertibility first. Since $A_2^{-1}\mathbf{1} = \mathbf{1}/(2a)$, the D-side factor is $I - \gamma B^\top A_2^{-1} = I - \gamma \frac{z}{2a}\mathbf{1}\mathbf{1}^\top$, whose eigenvalues are 1 (on the difference vector) and $1-\gamma z/a$ (on $\mathbf{1}$); it is singular precisely when $\gamma z = a$. The A-side factor $I-\gamma\frac{z}{2d}\mathbf{1}\mathbf{1}^\top$ has eigenvalue $1-\gamma z/d > 0$, because $\gamma z \le 2c < d$; here $2c < d$ follows from $4c^2 < ad \le d^2$. The common numerator in (17) also never vanishes: $ad-\gamma z^2 \ge ad-4c^2 > 0$. The matrix factors are not the only denominators. The recursion also divides by the augmented diagonal entries and by the fitness sum, and exchangeability settles both. The difference vector $v=(1,-1)^\top$ satisfies $\mathbf{1}\mathbf{1}^\top v = 0$, so $\widehat{B}v = \widehat{B}^\top v = 0$: the Schur correction in (4) annihilates v and the collapse factor fixes v, hence each augmented block keeps its population difference eigenvalue, $\frac{1}{2}$ for the A block $(1-\frac{1}{2})$ and 2 for the D block $(4-2)$. An exchangeable 2×2 matrix has both diagonal entries equal to the average of its two eigenvalues, and its $\mathbf{1}$ eigenvalue is twice its fitness, so

$$(A_\gamma)_{ii} \;=\; \nu_{A_\gamma}(z)+\tfrac{1}{4}, \qquad (D_\gamma)_{ii} \;=\; \nu_{D_\gamma}(z)+1 \;=\; \frac{a(d+1)-\gamma z\,(z+1)}{a-\gamma z}.$$

The first is positive outright, since $\nu_{A_\gamma}(z) > 0$ by (17). The second has numerator at least $a(d+1)-2c\,(2c+1) = 3-\frac{66}{25} = \frac{9}{25} > 0$ for (16), so it can vanish nowhere and blows up only through the factor $a-\gamma z$ already identified. Finally

$$\nu_{A_\gamma}(z)+\nu_{D_\gamma}(z) \;=\; \frac{(ad-\gamma z^2)\,(S-2\gamma z)}{(d-\gamma z)(a-\gamma z)},$$

whose numerator is positive on the admissible range; so for (16) the point $\gamma z = a$ is the recursion's only hole. For (17): since D_2 is exchangeable with row sums $2d$, $D_2^{-1}\mathbf{1} = \mathbf{1}/(2d)$, so $\widehat{B}D_2^{-1} = \frac{z}{2d}\mathbf{1}\mathbf{1}^\top$ and $\widehat{B}D_2^{-1}\widehat{B}^\top = \frac{z^2}{d}\mathbf{1}\mathbf{1}^\top$ (using $\mathbf{1}\mathbf{1}^\top\mathbf{1}\mathbf{1}^\top = 2\,\mathbf{1}\mathbf{1}^\top$). The matrix $r = I-\gamma\frac{z}{2d}\mathbf{1}\mathbf{1}^\top$ satisfies $\mathbf{1}^\top r^{-1} = (1-\gamma z/d)^{-1}\mathbf{1}^\top$ (Sherman–Morrison), and sym does not change $\mathbf{1}^\top(\cdot)\mathbf{1}$. With within weights $(\frac{1}{2},\frac{1}{2})$, the fitness is $\nu_{A_\gamma} = \frac{1}{4}\mathbf{1}^\top A_\gamma\mathbf{1} = \frac{1}{4}\,\dfrac{\mathbf{1}^\top A_2\mathbf{1}-\gamma\frac{z^2}{d}\cdot 4}{1-\gamma z/d} = \dfrac{a-\gamma z^2/d}{1-\gamma z/d} = \dfrac{ad-\gamma z^2}{d-\gamma z}$, using $\mathbf{1}^\top A_2\mathbf{1} = 4a$; the D-side is symmetric. For (18): the inter-group ratio is $1/\nu_{A_\gamma} : 1/\nu_{D_\gamma}$, the common factor $ad-\gamma z^2$ cancels, and $x_\gamma(z) = (d-\gamma z)/((d-\gamma z)+(a-\gamma z))$. The final sentence is the definition of Q at the group level. Each identity is also verified exactly by the certificate script at a grid of rational (γ, z) away from the singular point; at the singular point the script verifies that the recursion is undefined (a division by an exact zero) and that the recursion's weight approaches (18) on both sides. □

Proposition 2 (Exact objective). *Let $x_* = x_1(c) = (d-c)/K$ be the population minimum-variance group weight. Then $Q(x) = Q(x_*) + K\,(x-x_*)^2$ (exact square completion),*

$$x_\gamma(z)-x_* \;=\; \frac{(d-a)\,(\gamma z-c)}{(S-2\gamma z)\,K}, \tag{19}$$

and therefore, with $r(u) = \big((u-c)/(S-2u)\big)^2$,

$$F_{\text{ext}}(\gamma;e) \;=\; Q(x_*) \;+\; \frac{(d-a)^2}{K}\cdot\frac{1}{2}\sum_{\sigma=\pm 1} r\big(\gamma(c+\sigma e)\big). \tag{20}$$

This is the exact expected out-of-sample variance of the extended allocation; no Taylor surrogate is involved. It agrees with the objective of the raw recursion at every (γ, e) where the recursion is defined; by Lemma 3, for the numerical family (16) *that is everywhere except the single point $\gamma = a/(c+e)$ when $e \ge a-c$. In particular*

$$V_0(\gamma) = \frac{9\,(44\gamma^2-200\gamma+275)}{5\,(8\gamma-25)^2}, \qquad G(\gamma) = \frac{1500\,\gamma^2\,(16\gamma+1)}{(8\gamma-25)^4}, \tag{21}$$

where G is the coefficient of e^2 in the small-noise expansion of $F_{\text{ext}}(\gamma; e)$; in the τ parametrization of (12)*, with $e = 2\tau$, the coefficient of τ^2 is $4G$. For $e < a - c$ there is no singular point and F_{ext} is the raw objective outright.*

Proof. Q is a quadratic in x with leading coefficient $a + d - 2c = K$ and vertex at x_* (set $Q'(x) = 2ax - 2d(1-x) + 2c(1-2x) = 2Kx - 2(d-c)$), giving the square completion. For (19), put both terms of $x_\gamma(z) - x_*$ over the common denominator $(S - 2\gamma z)K$: the numerator is $(d - \gamma z)K - (d-c)(S - 2\gamma z) = (\gamma z - c)(2d - S) = (d-a)(\gamma z - c)$, expanding and using $K = S - 2c$. Then $Q(x_\gamma(z)) - Q(x_*) = K(x_\gamma(z) - x_*)^2 = \frac{(d-a)^2}{K} r(\gamma z)$, and averaging over $z = c \pm e$ gives (20). The closed forms (21) follow by substituting the family's numbers into (20). The certificate script verifies V_0 exactly at a grid of rational γ; G follows from (20) by the expansion in e^2. □

Theorem 4 (Exact interior optimum, monotone in the noise)**.** *Assume $a < d$, $c > 0$, $4c^2 < ad$, and $S > 6c$ (all satisfied by* (16)*: $4c^2 = \frac{36}{25} < \frac{9}{4} = ad$ and $S = \frac{15}{4} > \frac{18}{5} = 6c$). Then for every noise amplitude $0 < e \le c$, with F_{ext} the extended objective of Lemma 3:*

(i) *$F_{\text{ext}}(\cdot\,; e)$ is strictly convex on $[0, 1]$;*

(ii) *it has a unique minimizer $\gamma^*(e) \in (\frac{1}{2}, 1)$ for $0 < e < c$, with the boundary values $\gamma^*(0) = 1$ and $\gamma^*(c) = \frac{1}{2}$;*

(iii) *γ^* is strictly decreasing in e;*

(iv) *the minimizer solves the exact first-order condition*

$$\sum_{\sigma=\pm 1} \frac{(c + \sigma e)\left[\gamma(c + \sigma e) - c\right]}{\left[S - 2\gamma(c + \sigma e)\right]^3} \;=\; 0, \tag{22}$$

and for small e,

$$\gamma^*(e) \;=\; 1 - \frac{S + 4c}{c^2\,(S - 2c)}\, e^2 + O(e^4) \;=\; 1 - \frac{1025}{153}\, e^2 + O(e^4) \tag{23}$$

for the numbers of (16)*;*

(v) *for the numbers of* (16) *the minimizer lies strictly inside the regular region: whenever the singular point $\gamma_e = a/(c + e)$ of Lemma 3 lies in $[0, 1]$, that is for $e \ge a - c = \frac{3}{20}$,*

$$\gamma^*(e) \;<\; \frac{a}{c + e}.$$

Consequently the raw recursion is defined on a neighborhood of $\gamma^*(e)$ and coincides there with F_{ext}, so (i)–(iv) describe the optimizer of the implemented recursion and not only of the extension.

Consequently every coupling in $[\frac{1}{2}, 1]$ is optimal at exactly one noise amplitude $e \in [0, c]$. In the noise parametrization of the paper, $e = 2\tau$, so the theorem covers $0 < \tau \le \frac{3}{10}$.

Proof. Throughout, drop the positive constant $(d-a)^2/K$ and study $\bar{F}(\gamma; e) = \frac{1}{2}[r(\gamma(c+e)) + r(\gamma(c-e))]$ with $r(u) = (u - c)^2/(S - 2u)^2$, defined for $u < S/2$; note $\gamma z \le 2c < S/2$ on the admissible range, since $z \le c + e \le 2c$ and $S > 6c > 4c$. By (20), $\bar{F}$ is F_{ext} up to a positive affine map, and it is smooth on all of $[0, 1] \times [0, c]$; the point $\gamma z = a$, where the raw recursion is undefined, is unremarkable for $\bar{F}$.

Derivatives of r. Direct differentiation gives

$$r'(u) \;=\; \frac{2K\,(u-c)}{(S-2u)^3}, \qquad r''(u) \;=\; \frac{2K\,(S-6c+4u)}{(S-2u)^4}.$$

(For r': the quotient rule gives numerator $2(u-c)(S-2u)+4(u-c)^2 = 2(u-c)(S-2c)$; for r'', differentiate r' and simplify the numerator to $2K[(S-2u)+6(u-c)]$.)

(i) Convexity. $\partial^2_{\gamma\gamma}\bar F(\gamma;e) = \frac12\sum_\sigma z_\sigma^2\, r''(\gamma z_\sigma)$ with $z_\sigma = c+\sigma e \ge 0$. Since $S > 6c$ and $u = \gamma z_\sigma \ge 0$, the numerator $S-6c+4u$ is strictly positive, and $(S-2u)^4 > 0$, so $r'' > 0$ wherever $z_\sigma > 0$; at $e = c$ the lower branch has $z_- = 0$ and contributes nothing, while the upper branch still gives strict positivity. Hence $\bar F$ is strictly convex in γ on $[0,1]$.

Endpoint signs. At $\gamma = 0$: $\partial_\gamma \bar F(0;e) = \frac12\sum_\sigma z_\sigma r'(0) = r'(0)\,c = -\,2Kc^2/S^3 < 0$, independently of e. At $\gamma = 1$: define $\psi(z) = z\,r'(z) = 2Kz(z-c)/(S-2z)^3$, so $\partial_\gamma\bar F(1;e) = \frac12[\psi(c+e)+\psi(c-e)]$. One computes

$$\psi''(z) \;=\; \frac{4K\,\big[\,S(S-6c)+4z\,(2S-3c)+4z^2\,\big]}{(S-2z)^5} \;>\; 0 \qquad (0\le z\le 2c),$$

since every bracketed term is positive under $S > 6c$ and $S-2z \ge S-4c > 0$. Thus ψ is strictly convex on $[0,2c]$, and since $\psi(c) = 0$, strict midpoint convexity gives $\frac12[\psi(c+e)+\psi(c-e)] > \psi(c) = 0$ for $0 < e \le c$. So the γ-derivative is strictly negative at $\gamma = 0$ and strictly positive at $\gamma = 1$.

(ii) Unique interior minimizer. Strict convexity plus the endpoint signs give a unique root $\gamma^*(e) \in (0,1)$ of $\partial_\gamma\bar F = 0$, which is the unique minimizer, and (22) is $\partial_\gamma \bar F = 0$ written out. At $e = c$ the lower branch has $z_- = 0$, so (22) reduces to $r'(2c\gamma) = 0$, whose unique solution is $2c\gamma = c$, i.e. $\gamma^*(c) = \frac12$; and $\gamma^*(e) > \frac12$ for $e < c$ follows from (iii). At $e = 0$ both branches coincide, $\partial_\gamma\bar F \propto r'(\gamma c)\,c$, whose root is $\gamma = 1$; this is the noise-free optimum, consistent with $V_0'(1) = 0$ (Lemma 2 applies: the reduction reaches the group-level minimum variance at $\gamma = 1$ exactly).

(iii) Monotonicity. For fixed $\gamma > 0$ write $\psi_\gamma(z) = z\,r'(\gamma z)$, so that $\partial_\gamma\bar F(\gamma;e) = \frac12\big[\psi_\gamma(c+e) + \psi_\gamma(c-e)\big]$ and

$$\partial_e\partial_\gamma\bar F(\gamma;e) \;=\; \tfrac12\big[\psi_\gamma'(c+e) - \psi_\gamma'(c-e)\big].$$

The identity $\psi_\gamma(z) = \psi(\gamma z)/\gamma$ gives $\psi_\gamma''(z) = \gamma\,\psi''(\gamma z) > 0$ on the admissible range, so ψ_γ' is strictly increasing, $\psi_\gamma'(c+e) > \psi_\gamma'(c-e)$, and the mixed partial is strictly positive. By the implicit function theorem on (22), $d\gamma^*/de = -\,\partial_e\partial_\gamma\bar F\,/\,\partial^2_{\gamma\gamma}\bar F < 0$.

(iv) Expansion. Insert $\gamma = 1-\delta$ into (22) and expand to first order in δ and e^2. The e-expansion of the sum at $\gamma = 1$ is $\frac{2e^2(S+4c)}{K^4} + O(e^4)$ (expand each branch of (22) in e; odd terms cancel), while the δ-derivative at $e = 0$ is $-\,2c^2/K^3 + O(\delta, e^2)$. Setting the sum of the two contributions to zero gives $\delta = e^2(S+4c)/(c^2K) + O(e^4)$, which is (23); the coefficient evaluates to $\frac{1025}{153}$ for (16).

(v) The singular point lies uphill. Let $\gamma_e = a/(c+e) \le 1$ and write $z_\pm = c\pm e$, so that $\gamma_e z_+ = a$ and $\gamma_e z_- = az_-/z_+$. By strict convexity it suffices to show $\partial_\gamma\bar F(\gamma_e;e) > 0$. Substituting into $\partial_\gamma\bar F = \frac12\sum_\sigma z_\sigma\, r'(\gamma z_\sigma)$ and clearing the inner fractions,

$$\partial_\gamma\bar F(\gamma_e;e) \;=\; K\left[\frac{z_+\,(a-c)}{(S-2a)^3} \;+\; \frac{z_-\,(az_- - cz_+)\,z_+^2}{(Sz_+ - 2az_-)^3}\right].$$

Both denominators are positive: $S - 2a = \frac94$ for (16), and $Sz_+ - 2az_- \ge z_+(S-2a) > 0$ since $z_- \le z_+$. Multiplying by the positive quantity $(S-2a)^3(Sz_+ - 2az_-)^3/K$ preserves the sign and produces the polynomial

$$P(e) \;=\; z_+(a-c)\,(Sz_+ - 2az_-)^3 \;+\; z_-(az_- - cz_+)\,z_+^2\,(S-2a)^3,$$

which for the numbers of (16) expands to

$$P(e) \;=\; \tfrac{177147}{400000} \;+\; \tfrac{6561}{800}\, e^2 \;+\; \tfrac{1215}{32}\, e^3 \;+\; \tfrac{23733}{640}\, e^4.$$

Every coefficient is nonnegative and the constant term is positive, so $P(e) > 0$ for every $e \geq 0$, and in particular on $[\frac{3}{20}, \frac{3}{5}]$. Hence $\partial_\gamma \bar{F}(\gamma_e; e) > 0$, and by strict convexity $\gamma^*(e) < \gamma_e$. The expansion of P and the sign agreement with the first-order condition are verified in exact rational arithmetic by the certificate script.

The final sentence of the theorem follows from (ii)–(iii): $\gamma^* : [0, c] \to [\frac{1}{2}, 1]$ is a continuous strictly decreasing bijection. □

Corollary 1 (The surrogate frontier spans the whole bridge in this family)**.** *For the family* (16)*, the closed forms* (21) *satisfy* $V_0' < 0$*,* $V_0'' > 0$*,* $G' > 0$*,* $G'' > 0$ *on* $(0,1)$*, and*

$$t(\gamma) \;=\; -\frac{V_0'(\gamma)}{G'(\gamma)} \;=\; \frac{9\,(1-\gamma)\,(25-8\gamma)^2}{25\,\gamma\,(64\gamma^2+608\gamma+25)}$$

decreases strictly from $+\infty$ *(at* $\gamma \downarrow 0$*) to* 0 *(at* $\gamma \uparrow 1$*). Theorem 6 below is applied here with* e *as the noise coordinate,* $s = e^2$*; in the* τ *parametrization the coefficient is* $4G$ *and* t *is divided by four, which changes neither endpoint limit. In its language,* $t_0 = \infty$ *and* $t_1 = 0$*: for this family, and only because of these endpoint limits, the surrogate optimizer sweeps the entire open bridge and neither endpoint is surrogate-optimal at any positive finite noise level. The surrogate optimizer is not the exact optimizer at finite noise. At* $e = c$ *the exact minimizer is* $\frac{1}{2}$ *(Theorem 4), while* $t(\frac{1}{2}) = \frac{1323}{2875} \neq e^2 = \frac{9}{25}$ *and the surrogate optimizer sits near* 0.537*. The two have the same small-noise expansion through order* e^2*: each is* $1 - \frac{1025}{153}\, e^2 + O(e^4)$*, by* (23) *for the exact optimizer, and for the surrogate because* $t(\gamma) = \frac{V_0''(1)}{G'(1)}(1-\gamma) + O\big((1-\gamma)^2\big)$ *with* $G'(1)/V_0''(1) = \frac{1025}{153}$*, verified exactly by the certificate script. Their proven ranges differ: the surrogate sweeps the entire open bridge* $(0,1)$*, while the exact optimizer is proved to sweep* $[\frac{1}{2}, 1]$ *over the admissible range* $0 \leq e \leq c$*, and nothing is claimed about it beyond* $e = c$*.*

Remark 5. Theorem 4 reproduces, exactly, the numerical optimizer sweep previously obtained by brute force on this family: with $e = 2\tau$, the exact minimizers at $\tau = 0.10, 0.15, 0.20, 0.25, 0.30$ are $\gamma^* \approx 0.8308,\ 0.7237,\ 0.6339,\ 0.5603,\ 0.5000$, the last being exact. The value $\tau = 0.35$ falls outside the range $e \leq c$ covered by the theorem (one noise branch then has negative estimated cross covariance); it could be handled by direct polynomial certification but we do not pursue it. The range $e \leq c$ is also not the boundary of positive definiteness, which at the group level extends to $e < \sqrt{ad} - c = \frac{9}{10}$ for (16). On the extended objective the minimizer keeps falling past the theorem's range, to approximately 0.450, 0.408, and 0.376 at $e = 0.7$, 0.8, and 0.89. We claim nothing there.

6 Whole-matrix noise

The model of Section 2 puts noise on the cross-block alone, and a fair objection runs: if the only noisy channel is the one the coupling controls, an interior optimum is unsurprising. Theorem 1 already resists the objection at one end, since it needs no support restriction and it excludes the endpoint that cross-block noise favors. The interiority results, however, were proved under Assumption 2. In this section we remove the restriction for the family of Section 5 and place noise in every entry of the covariance. We do this in three models of increasing generality. The first is an exchangeable model solved by the same reduction, with a sharp second-order criterion for interiority at the full-coupling end. The second noises all ten distinct entries independently and is settled by exact enumeration.

The third is the sample covariance of T Gaussian draws, the estimation model in which [5] prove hierarchical risk parity more robust than Markowitz.

The exchangeable model. Perturb every block by a rank-one exchangeable displacement,

$$\widehat{A} = A_2 + \tau\alpha\,\mathbf{1}\mathbf{1}^\top, \qquad \widehat{D} = D_2 + \tau\delta\,\mathbf{1}\mathbf{1}^\top, \qquad \widehat{B} = (c+\tau\zeta)\,\mathbf{1}\mathbf{1}^\top, \tag{24}$$

where (α, δ, ζ) is a bounded random vector, jointly sign-symmetric: $(\alpha, \delta, \zeta) \overset{d}{=} -(\alpha, \delta, \zeta)$. Every entry of $\widehat{\Sigma}$ is now noisy. The group-level quantities become $\hat{a} = a + \tau\alpha$, $\hat{d} = d + \tau\delta$, $\hat{z} = c + \tau\zeta$, and the anchor of Lemma 1 is lost: $F(0;\tau)$ is itself noisy, and the endpoint comparison becomes a comparison of noise costs.

Lemma 4 (Exact objective under whole-matrix noise)**.** *The perturbations* (24) *preserve the exchangeability of each block, so Lemma 3 applies verbatim with $(\hat{a}, \hat{d}, \hat{z})$ in place of (a, d, z): within-block weights are exactly $(\frac{1}{2}, \frac{1}{2})$, and the group weight is $x_\gamma = (\hat{d} - \gamma\hat{z})/(\hat{S} - 2\gamma\hat{z})$ with $\hat{S} = \hat{a} + \hat{d}$. Wherever the recursion is defined, the expected out-of-sample variance of the extended allocation is, exactly,*

$$F_{\text{ext}}(\gamma;\tau) \;=\; Q(x_*) \;+\; \frac{1}{K}\,\mathbb{E}\left[\frac{\big[(d-a)(\gamma\hat{z}-c) + \tau h\big]^2}{(\hat{S}-2\gamma\hat{z})^2}\right], \qquad h \;:=\; (a-c)\,\delta - (d-c)\,\alpha. \tag{25}$$

For $0 < c < a$ there is $\bar{\tau} > 0$ such that for $|\tau| \le \bar{\tau}$ every denominator the recursion divides by is bounded away from zero uniformly on $[0,1]$ and on every realization, so the recursion is defined on all of $[0,1]$ and coincides with F_{ext}.

Proof. The displacements $\mathbf{1}\mathbf{1}^\top$ commute with the swap of a block's two assets, so each perturbed block is exchangeable and $\widehat{B}$ retains the form $z\,\mathbf{1}\mathbf{1}^\top$. The proof of Lemma 3 therefore goes through with constants $(\hat{a}, \hat{d}, \hat{z})$. Within-block weights $(\frac{1}{2}, \frac{1}{2})$ are the population-optimal within weights, so the realized variance is $Q(x_\gamma)$ with the population Q, and $Q(x) - Q(x_*) = K(x - x_*)^2$ as in Proposition 2. For the displacement, the numerator identity behind (19) is linear in its inputs: with $\hat{d} = d + \tau\delta$, $\hat{S} = S + \tau(\alpha+\delta)$,

$$K(\hat{d} - \gamma\hat{z}) - (d-c)(\hat{S} - 2\gamma\hat{z}) \;=\; (d-a)(\gamma\hat{z} - c) \;+\; \tau\big[K\delta - (d-c)(\alpha+\delta)\big],$$

and $K - (d-c) = a - c$ gives the bracket as τh. Squaring and multiplying by K yields (25). For the final claim, note that at $\tau = 0$ every denominator is a continuous function of γ that is nonzero on $[0,1]$. In particular the factors $d - \gamma c \ge d - c$, $a - \gamma c \ge a - c$, $ad - \gamma c^2 \ge ad - c^2$, and $S - 2\gamma c \ge K$ are positive, and the augmented diagonal numerators are handled as in Lemma 3. By compactness of $[0,1]$ and boundedness of (α, δ, ζ), all of them stay uniformly positive for all sufficiently small $|\tau|$. □

Theorem 5 (Interior optimum under whole-matrix noise)**.** *Assume $a < d$, $0 < c < a$, $4c^2 < ad$, and $S > 6c$, as in* (16)*. Let $u := (d-a)\zeta + h$ and define*

$$\Xi \;:=\; (d-a)\,K\,\mathbb{E}[\zeta u] \;+\; 2c\,\mathbb{E}[u^2] \;-\; 2(d-a)\,c\,\mathbb{E}\big[u\,(\alpha+\delta-2\zeta)\big]. \tag{26}$$

Then $F_{\text{ext}}(\gamma;\tau) = V_0(\gamma) + \tau^2 G_W(\gamma) + O(\tau^4)$ in C^2 uniformly on $[0,1]$, with

$$G_W'(1) \;=\; \frac{2\,\Xi}{K^4}, \qquad V_0''(1) \;=\; \frac{2\,(d-a)^2c^2}{K^3} \;>\; 0,$$

and:

(i) *if $\Xi > 0$, then for all sufficiently small $\tau > 0$ the objective has a unique global minimizer $\gamma^*(\tau)$ on $[0,1]$, it is interior, and*

$$1 - \gamma^*(\tau) \;=\; \frac{\Xi}{(d-a)^2\, c^2\, K}\, \tau^2 \;+\; O(\tau^4); \tag{27}$$

(ii) *if $\Xi < 0$, full coupling is locally optimal for all sufficiently small τ;*

(iii) *cross-only noise ($\alpha = \delta = 0$) gives $\Xi = (d-a)^2(S+4c)\,\mathbb{E}[\zeta^2]$; under the two-point noise of Section 5, $\zeta = \pm 2$ so $\mathbb{E}[\zeta^2] = 4$, and (27) reduces exactly to (23);*

(iv) *for pairwise-uncorrelated channels,*

$$\Xi \;=\; (d-a)^2(S+4c)\,\mathbb{E}[\zeta^2] \;+\; 2c\,(d-c)(2d-a-c)\,\mathbb{E}[\alpha^2] \;+\; 2c\,(a-c)(2a-c-d)\,\mathbb{E}[\delta^2]. \tag{28}$$

The first two coefficients are positive under the standing hypotheses; the third is negative precisely when $2a < c+d$.

Proof. By Lemma 4, for small τ the objective is the expectation of a rational function of (γ,τ) whose denominators are uniformly bounded away from zero, so it is C^∞ on $[0,1]\times[-\bar\tau,\bar\tau]$, and sign symmetry of (α,δ,ζ) makes it even in τ. The expansion is then the Taylor expansion in τ^2, valid in C^2 uniformly. Write the integrand of (25) as $(u_0+\tau u_1)^2/(D_0+\tau D_1)^2$ with

$$u_0 = (d-a)\,c\,(\gamma-1), \quad u_1 = (d-a)\gamma\zeta + h, \quad D_0 = S - 2\gamma c, \quad D_1 = \alpha+\delta-2\gamma\zeta.$$

Expanding and discarding odd powers (which vanish in expectation),

$$G_W(\gamma) \;=\; \frac{1}{K}\left[\frac{\mathbb{E}[u_1^2]}{D_0^2} \;-\; \frac{4\,u_0\,\mathbb{E}[u_1 D_1]}{D_0^3} \;+\; \frac{3\,u_0^2\,\mathbb{E}[D_1^2]}{D_0^4}\right]. \tag{29}$$

At $\gamma = 1$: $u_0 = 0$, $u_0' = (d-a)c$, $u_1 = u$, $\partial_\gamma u_1 = (d-a)\zeta$, $D_0 = K$, $D_0' = -2c$, and $D_1 = \alpha+\delta-2\zeta$, so

$$G_W'(1) = \frac{1}{K}\left[\frac{2(d-a)\mathbb{E}[\zeta u]}{K^2} + \frac{4c\,\mathbb{E}[u^2]}{K^3} - \frac{4(d-a)c\,\mathbb{E}[uD_1]}{K^3}\right] = \frac{2\,\Xi}{K^4}.$$

The noise-free part is $V_0(\gamma) - Q(x_*) = (d-a)^2c^2(\gamma-1)^2/(KD_0^2)$, giving $V_0''(1) = 2(d-a)^2c^2/K^3$ and $V_0'(0) = 2c^2(d-a)^2(2c-S)/(KS^3) < 0$. Under $S > 6c$, V_0 is strictly convex on $[0,1]$ (Theorem 4(i) at $e=0$), so $\gamma = 1$ is its unique global minimizer. Theorem 3 applies with $G = G_W$. If $G_W'(1) > 0$ there is a C^1 curve of local minimizers with $1-\gamma^*(\tau) = (G_W'(1)/V_0''(1))\tau^2 + O(\tau^4)$, which is (27), and it is the unique global minimizer for small τ by the uniform convergence $F_{\text{ext}} \to V_0$. If $G_W'(1) < 0$ full coupling is locally optimal. Since $V_0'(0) < 0$ and the noise is bounded with no support restriction, Theorem 1 applies and keeps every minimizer uniformly separated from $\gamma = 0$, so the global minimizer in case (i) is interior. For (iii) and (iv), substitute $u = (d-a)\zeta + (a-c)\delta - (d-c)\alpha$ into (26) and drop the vanishing cross moments. The identity (25) is checked exactly at a grid of rational points by the certificate script, which also certifies $F(\hat\gamma) < \min(F(0), F(1))$ at the coupling $\hat\gamma$ that (27) predicts. For positivity in (iv), note $2d-a-c > d-c > 0$ and $a > c$. □

Remark 6 (The numerical family under whole-matrix noise). For the numbers of (16), (28) reads

$$\Xi \;=\; \tfrac{9963}{320}\,\mathbb{E}[\zeta^2] \;+\; \tfrac{1674}{125}\,\mathbb{E}[\alpha^2] \;-\; \tfrac{189}{500}\,\mathbb{E}[\delta^2].$$

Three consequences follow. First, with equal unit second moments on three pairwise-uncorrelated channels, the shift coefficient in (27) is $\Xi/((d-a)^2c^2K) = \frac{13081}{1377} \approx 9.50$, against $\frac{1025}{153} \approx 6.70$ for cross-only noise (23). In this equal-amplitude comparison the diagonal channels add a net inward push, the positive α-channel term dominating the negative δ-channel term. The exact recursion confirms the coefficient: at $\tau = \frac{1}{100}$ the exact minimizer over an exhaustive fine grid agrees with $1 - \frac{13081}{1377}\tau^2$ to five decimal places. For $\tau \in \{\frac{1}{50}, \frac{1}{100}\}$, the inequality $F(\hat{\gamma}) < \min\big(F(0), F(1)\big)$ holds as an exact rational inequality at $\hat{\gamma} = 1 - \frac{13081}{1377}\tau^2$. Second, the condition is sharp. Noise confined to the $\hat{d}$ channel has $\Xi = -\frac{189}{500}\,\mathbb{E}[\delta^2] < 0$, and the certificate script verifies exactly, at amplitudes $\tau|\delta| \in \{\frac{1}{10}, \frac{1}{4}\}$, that $F(1)$ is strictly below $F(\gamma)$ at every grid point $\gamma = k/50 < 1$. The support of the noise does not decide the question. The sign of (26) does, and at $\Xi = 0$ the second order is silent: a fourth-order criterion would be needed, and we do not pursue it. Third, at the hierarchical end (29) gives

$$G_W(0) \;=\; \frac{1}{K}\left[\frac{\mathbb{E}[h^2]}{S^2} + \frac{4c(d-a)\,\mathbb{E}[h(\alpha+\delta)]}{S^3} + \frac{3c^2(d-a)^2\,\mathbb{E}[(\alpha+\delta)^2]}{S^4}\right],$$

which involves no ζ moments. The cross channel is invisible at $\gamma = 0$, as in Lemma 1, but the diagonal channels now charge HRP a noise cost of its own: for the family it is $\frac{704\mathbb{E}[\alpha^2]+164\mathbb{E}[\delta^2]}{9375} > 0$ under uncorrelated channels. Interiority therefore wins a comparison between two noise costs, with no noise-free anchor left to lean on.

6.1 Entrywise noise, by exact enumeration

Independent noise entry by entry breaks exchangeability, and no scalar reduction is available. With two-point noise, however, the objective is a finite average of exact rationals, so interiority is still certifiable. We perturb each correlation and variance independently,

$$\widehat{\Sigma}_{ij} = \sigma_i\sigma_j\,(\rho_{ij} + \tau\,\varepsilon_{ij}) \;\; (i \neq j), \qquad \widehat{\Sigma}_{ii} = \sigma_i^2\,(1 + \tau\,\varepsilon_{ii}),$$

with independent signs $\varepsilon = \pm 1$, and run the recursion through every sign state ($2^6 = 64$ states for the six off-diagonal entries; $2^{10} = 1024$ when the four variances are noised as well). The certificate script verifies the following as exact rational inequalities, with every state defined at each coupling evaluated:

noised entries	τ	$\hat{\gamma}$	$F(0) - F(\hat{\gamma})$	$F(1) - F(\hat{\gamma})$
six off-diagonal	$\frac{3}{20}$	$\frac{7}{10}$	$\approx 3.57 \times 10^{-2}$	≈ 8.28
all ten	$\frac{1}{10}$	$\frac{11}{12}$	$\approx 4.62 \times 10^{-2}$	$\approx 8.97 \times 10^{-2}$

The noise levels are chosen so that every enumerated matrix is a genuine covariance: all 64 off-diagonal states at $\tau = \frac{3}{20}$ and all 1024 states at $\tau = \frac{1}{10}$ are positive definite (smallest eigenvalues ≈ 0.078 and ≈ 0.148), certified exactly through leading principal minors. At $\tau = \frac{1}{5}$ two of the 64 off-diagonal states leave the cone, so we do not certify there. The large second margin in the first row is itself informative: it comes from near-poles of the implemented recursion at full coupling even though every input is a valid covariance. One more step is needed before concluding, because F need not be continuous on all of $[0,1]$: an individual sign state can have an interior fitness-sum zero, where its loss has a second-order pole. The repair is a lower bound. Every weight vector the recursion produces sums to one, so its realized variance is at least $1/(\mathbf{1}^\top\Sigma^{-1}\mathbf{1}) > 0$. Each state's loss is a nonnegative rational function of γ with finitely many exceptional points. Extend it at each exceptional point by its finite limit when the singularity is removable and by $+\infty$ otherwise. The extension is lower semicontinuous, so the finite average is lower semicontinuous and attains its minimum on $[0,1]$, and the certified inequalities exclude both endpoints: every minimizer is strictly interior in both models. At larger noise the full-coupling end degrades by orders of magnitude. At $\tau = \frac{3}{20}$ with all ten entries noised, 56 of the 1024 branches are exactly singular at $\gamma = 1$ (the collapse factor vanishes there, with finite one-sided limits), and, even discarding those branches, the mean loss over the 968 nonsingular branches is ≈ 75, against $F(\frac{12}{25}) \approx 0.77$. The comparison against HRP, $F(0) - F(\frac{12}{25}) \approx 2.66 \times 10^{-2} > 0$, remains a clean exact certificate.

6.2 The sample covariance

Finally we take the model of Remark 2 literally. Let $\widehat{\Sigma}_T = \frac{1}{T}\sum_{t=1}^{T} x_t x_t^\top$ with $x_t \sim \mathcal{N}(0, \Sigma)$ i.i.d., and write $L_T(\gamma) = w(\gamma; \widehat{\Sigma}_T)^\top \Sigma\, w(\gamma; \widehat{\Sigma}_T)$ for the random realized loss, so that the objective of (7) is $F_T(\gamma) = \mathbb{E}[L_T(\gamma)]$. This puts Wishart noise in every entry, and it is the estimation model in which [5] derive the weight noise of HRP and of Markowitz and prove HRP the more robust.

Two objects must be kept apart here. A sample covariance is not exchangeable, and on non-exchangeable input the implemented recursion at $\gamma = 1$ is *not* the sample minimum-variance portfolio: the naive within-block splits break the exact recovery of Proposition 1 (the symmetrization is exact at a single level of coupling, and $n = 4$ has only one), which is why Section 4 assumed that recovery rather than derived it. The gap is exact. For the SPD matrix

$$M = \begin{pmatrix} 8 & -7 & 1 & -6 \\ -7 & 10 & -1 & 7 \\ 1 & -1 & 2 & 1 \\ -6 & 7 & 1 & 11 \end{pmatrix}$$

(leading principal minors $8, 31, 58, 230$), the implemented recursion gives $w(1; M) = \frac{1}{3683}(1330, 1197, 952, 204)$, while $M\,(1,1,1,0)^\top = 2 \cdot \mathbf{1}$ gives $w_{\mathrm{MV}}(M) = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0)$: the two disagree in every entry, exactly (certificate script). Both maps are rational in the entries wherever the recursion is regular. After clearing the recursion denominators, at least one coordinate of $w(1;\cdot) - w_{\mathrm{MV}}(\cdot)$ has a polynomial numerator that is nonzero at the displayed matrix, so that numerator is not the zero polynomial and its zero set has Lebesgue measure zero: exact agreement has measure zero under any absolutely continuous sampling model. On the exchangeable family the two coincide identically. In this subsection $\gamma = 1$ therefore means *full coupling of the implemented recursion*, and the sample minimum-variance portfolio $w_{\mathrm{MV}}(\widehat{\Sigma}_T)$ enters as a separate benchmark.

The sample mean of the raw loss is empirically unstable, and the following criterion explains a mechanism that would make F_T infinite.

Proposition 3 (Pole criterion)**.** *Let $L \geq 0$ and D be random variables on one probability space, and suppose there are constants $\varepsilon_0, c_0, n_0 > 0$ with $L \geq n_0/D^2$ on the event $\{|D| \leq \varepsilon_0\}$ and $\Pr[\,|D| \leq \varepsilon\,] \geq c_0\,\varepsilon$ for all $\varepsilon \leq \varepsilon_0$. Then $\mathbb{E}[L] = \infty$.*

Proof. By the layer-cake formula,

$$\mathbb{E}[L\,\mathbf{1}\{|D| \leq \varepsilon_0\}] \;\geq\; \int_0^\infty \Pr\left[\, n_0/D^2 > s,\; |D| \leq \varepsilon_0 \,\right] ds \;\geq\; \int_{n_0/\varepsilon_0^2}^\infty c_0 \sqrt{n_0/s}\, ds \;=\; \infty. \qquad \square$$

For the implemented recursion the relevant D is the fitness sum $\nu(A_\gamma) + \nu(D_\gamma)$, which the top-level split divides by. It can cross zero at small T because a sampled augmented block can be indefinite, and the split $\nu(D_\gamma)/(\nu(A_\gamma) + \nu(D_\gamma))$ then has a simple pole, so L_T has a second-order one. This is a different singularity from the collapse-factor zero $\gamma z = a$ of Lemma 3, which is removable for the loss: there the divergent-fitness block receives weight zero in the limit. Monte Carlo supports both hypotheses of the proposition at the full-coupling end: at $T = 15$, $\gamma = 1$

the frequency $\Pr[\,|D| \le \varepsilon\,]/\varepsilon$ is stable near 0.06 across $\varepsilon \in [3\times10^{-4}, 3\times10^{-2}]$, and on the blow-up draws the product $L_T D^2$ concentrates near 1. Consistently, the sample mean of $L_T(1)$ never settles while the sample median is stable to four decimals as the replication count grows (script `paper/wishart_experiment.py` prints both). We stop short of a theorem: the proposition needs a deterministic lower bound $L_T \ge n_0/D^2$ on the whole small-$|D|$ event, and what the numerics show is that the divergence is strongly indicated rather than established. A full proof would exhibit a regular zero of the fitness sum: an SPD matrix at which $D = 0$ with nonzero gradient, the portfolio residue does not vanish, and every other recursion denominator is nonzero. The proposition then applies on a neighborhood, where the Wishart density is positive. We have located such a zero numerically and leave its certification to future work. None of this is Markowitz estimation risk. The sample minimum-variance portfolio has finite mean, $\mathbb{E}[L_T^{\mathrm{MV}}] = \nu(\Sigma)\,(T-1)/(T-n)$ for $T > n$, the known-mean estimator $\widehat{\Sigma}_T$, and Gaussian draws,[2] and the canonical pair recursion of Proposition 1 has no pole of this kind at all, since its augmented block satisfies $A - \gamma B D^{-1} B^\top = (1-\gamma)\,A + \gamma\,(A - BD^{-1}B^\top)$, a convex combination of positive definite matrices whenever $\widehat{\Sigma}_T$ is positive definite. The pole belongs to the collapse and the naive splits. Implementations guard against it by forcing the augmented blocks positive definite, which truncates the tail.

Remark 7 (Toward an endpoint-exact bridge). These observations suggest a route toward a variant with exact endpoints and without the present fitness-sum pole, to be treated in a future paper. Damp the within-block off-diagonals with the same dial, replacing each augmented block Q by $(1-\gamma)\operatorname{diag}(Q) + \gamma Q$, and solve within blocks canonically: at $\gamma = 0$ the solve degenerates to $\operatorname{diag}(Q)^{-1}\mathbf{1}$, HRP's naive rule, and at $\gamma = 1$ the canonical pair returns exact sample minimum variance, with every inverted matrix a convex combination of positive definite matrices. The recursion cannot be called pole-free until the top-level split, which must interpolate HRP's inverse-fitness ratio and the canonical normalization, is specified and proved to stay away from zero. That is the future paper's problem.

We therefore summarize the raw recursion by medians $M_T(\gamma) = \operatorname{med} L_T(\gamma)$, which are stable. The table reports 2×10^5 replications per sample size, the same draws shared across the coupling grid (step 0.02), for the family (16); γ^* minimizes the median curve, and the last two columns are the sample minimum-variance benchmark and the frequency of catastrophic realizations at full coupling:

T	γ^*	$M_T(0)$	$M_T(\gamma^*)$	$M_T(1)$	$\operatorname{med} L_T^{\mathrm{MV}}$	$\Pr[L_T(1) > 10]$
15	0.62	0.804	0.777	0.800	0.883	0.018
25	0.72	0.798	0.764	0.773	0.820	0.008
50	0.84	0.795	0.753	0.755	0.778	0.002
100	0.92	0.794	0.747	0.747	0.759	0.000
250	0.96	0.793	0.744	0.744	0.748	0.000

The interior coupling of the implemented bridge has the best median at every sample size. It beats HRP on three quarters of draws at $T = 15$ and on essentially all draws at $T = 250$, it beats its own full-coupling end, and it beats actual sample Markowitz, whose median is the worst column at small T even though its mean is finite. The margins at the full-coupling end shrink below the table's precision at large T: $M_T(1) - M_T(\gamma^*) \approx 4\times10^{-4}$ at $T = 100$ and 5×10^{-5} at $T = 250$, consistent

[2] Write $W = T\,\Sigma^{-1/2}\widehat{\Sigma}_T\Sigma^{-1/2} \sim W_n(T, I)$ and $a = \Sigma^{-1/2}\mathbf{1}$, so $\|a\|^2 = 1/\nu(\Sigma)$. By rotational invariance take $a = \|a\|\,e_1$ and partition W with first diagonal entry q, first column tail b, and lower block C. Normalizing the first column of W^{-1} gives $L_T^{\mathrm{MV}}/\nu(\Sigma) = 1 + b^\top C^{-2} b$. Conditionally on C, $b \sim \mathcal{N}(0, C)$, so $\mathbb{E}[b^\top C^{-2} b \mid C] = \operatorname{tr}(C^{-1})$, and $\mathbb{E}\operatorname{tr}(C^{-1}) = (n-1)/(T-n)$ for $C \sim W_{n-1}(T, I)$, giving $(T-1)/(T-n)$.

with $\gamma^* \to 1$. The median comparison is qualitatively consistent with [5]: sample Markowitz overtakes HRP only as T grows. Full coupling of the implemented recursion carries the tail: 1.8% of draws at $T = 15$ land above 10, more than thirteen times the population minimum variance, and the frequency decays to zero by $T = 250$. Over the five sample sizes examined the median-optimal coupling increases with T, from 0.62 to 0.96, and the median margin at full coupling closes, as the small-noise theory says it should. Sample size selects the seat on the bridge.

Remark 8 (Ledoit–Wolf shrinkage of the input). A practitioner would shrink $\widehat{\Sigma}$ before allocating, so it is fair to ask whether the interior optimum is an artifact of feeding the recursion a raw estimate. For the family the question can be answered exactly. Shrinking toward μI with $\mu = \operatorname{tr}\widehat{\Sigma}/4$ preserves block exchangeability, so Lemma 4 applies with $\tilde{a} = (1-\rho)\hat{a} + \rho\mu/2$, $\tilde{d} = (1-\rho)\hat{d} + \rho\mu/2$, and $\tilde{z} = (1-\rho)\hat{z}$. Shrinkage moves the $\gamma = 1$ endpoint off the population minimum-variance portfolio, and the flatness of Lemma 2 is destroyed: at zero noise the displacement numerator at $\gamma = 1$ is $-\big[(d-a)c + \frac{7}{4}(a-c) + \frac{1}{2}(d-c)\big]\rho = -\frac{45}{16}\rho$ for (16), giving the strictly negative bias slope

$$\partial_\gamma F(1; 0, \rho) \;=\; -\frac{243}{32\,K^3}\,\rho \;+\; O(\rho^2).$$

The noise push and the shrinkage pull therefore compete at the same order when $\rho = \kappa\tau^2$, which is the Ledoit–Wolf-consistent scaling, and the stationarity condition near $\gamma = 1$ gives

$$1 - \gamma^*(\tau) \;=\; \frac{G'(1) - \frac{243}{32K^3}\,\kappa}{V_0''(1)}\,\tau^2 \;+\; O(\tau^4)\;:$$

the interior optimum survives for $\kappa < \kappa^* := \frac{32K^3}{243}\,G'(1)$ and is lost to second order for $\kappa > \kappa^*$; at $\kappa = \kappa^*$ the fourth-order term decides, as at $\Xi = 0$. Here G is the τ^2-coefficient of (12), four times the e-parametrized closed form (21). For the common cross-block noise of Section 5, $\kappa^* = \frac{656}{51} \approx 12.9$, while the small-noise Ledoit–Wolf intensity for the same noise is $\kappa_{\mathrm{LW}} = \|N\|_F^2/\|\Sigma - \mu I\|_F^2 = \frac{1600}{1019} \approx 1.57$ (the oracle small-noise intensity; the estimated intensity used in practice fluctuates around it). Optimal shrinkage sits an order of magnitude below the threshold, so interiority survives, with the shift coefficient scaled by $1 - \kappa_{\mathrm{LW}}/\kappa^* \approx 0.88$: Ledoit–Wolf absorbs about a tenth of the damping the bridge wants, and the coupling supplies the rest. The recursion confirms the arithmetic: at $\tau = 0.04$ the measured shifts $1-\gamma^*$ are 0.0387 unshrunk and 0.0338 at κ_{LW}, and the optimum is at the endpoint by $\kappa = \kappa^*$; the finite-τ transition sits slightly below κ^* (near 12.1 at $\tau = 0.04$) and approaches it as $\tau \to 0$. Over-shrinkage is the caveat: at fixed $\rho > 0$ not scaling with τ, the bias slope dominates and full coupling on the shrunk input is optimal for all small τ. This matches the walk-forward observation at `schur.microprediction.org` that the empirical minimum survives Ledoit–Wolf shrinkage of the inputs.

7 The surrogate in general

Away from solvable families one can still organize the problem through the second-order surrogate $F_2 = V_0 + \tau^2 G$ of (12).

Theorem 6 (Clipped monotone frontier, for the surrogate)**.** *Suppose V_0 and G are continuous on $[0,\bar{\gamma}]$ and C^2 on $(0,\bar{\gamma})$, with: $V_0' < 0$; $G' > 0$; $V_0'' \geq 0$ and $G'' \geq 0$; and at every γ at least one of $V_0''(\gamma) > 0$, $G''(\gamma) > 0$ holds. Define $t = -V_0'/G'$ on $(0,\bar{\gamma})$ and its endpoint limits $t_0 = \lim_{\gamma\downarrow 0} t(\gamma) \in (0,\infty]$ and $t_1 = \lim_{\gamma\uparrow\bar{\gamma}} t(\gamma) \in [0,\infty)$. Then t is continuous and strictly decreasing,* *and for every $s = \tau^2 > 0$ the surrogate $F_2(\cdot\,;\tau)$ has the unique minimizer*

$$\gamma_2^*(s) \;=\; \begin{cases} \bar{\gamma}, & s \le t_1, \\ t^{-1}(s), & t_1 < s < t_0, \\ 0, & s \ge t_0, \end{cases} \tag{30}$$

which is nonincreasing in s. In particular the endpoints are optimal for the surrogate on nontrivial intervals of noise levels precisely when $t_1 > 0$, respectively $t_0 < \infty$. If moreover $V_0, G \in C^1([0,\bar{\gamma}])$ with $V_0'(0) < 0$, as holds under Assumption 1 via (12), then $t_0 = -V_0'(0)/G'(0)$, finite exactly when $G'(0) > 0$. Endpoint regularity is needed for that formula: $V_0(\gamma) = V_0(0) - \sqrt{\gamma}$ near 0, with smooth G, satisfies every open-interval hypothesis and forces $t_0 = \infty$ regardless of $G'(0)$. In Example 1, $t_0 = \frac{27}{800}$; in the family of Section 5, $t_0 = \infty$ and $t_1 = 0$ (Corollary 1).

Proof. Strict monotonicity of t: on $(0,\bar{\gamma})$,

$$t' \;=\; \frac{-V_0''\,G' + V_0'\,G''}{(G')^2},$$

where the first numerator term is ≤ 0 (since $V_0'' \ge 0$, $G' > 0$) and the second is ≤ 0 (since $V_0' < 0$, $G'' \ge 0$); at each γ at least one is strictly negative by hypothesis, so $t' < 0$ pointwise, and t is a continuous strictly decreasing bijection from $(0,\bar{\gamma})$ onto (t_1, t_0). Sign analysis of $F_2' = V_0' + s\,G'$: since $G' > 0$, $F_2'(\gamma) < 0 \iff s < t(\gamma)$. If $t_1 < s < t_0$ then $F_2' < 0$ on $(0, t^{-1}(s))$ and $F_2' > 0$ on $(t^{-1}(s), \bar{\gamma})$, so $t^{-1}(s)$ is the unique minimizer. If $s \le t_1$ then $s < t(\gamma)$ for every interior γ, so $F_2' < 0$ throughout and the minimum sits at $\bar{\gamma}$; if $s \ge t_0$ then $F_2' > 0$ throughout and the minimum sits at 0. Monotonicity of (30) in s follows from the monotonicity of t^{-1} and the ordering of the three regimes. □

Proposition 4 (Local transfer to the exact objective)**.** *Fix τ and let $I \subseteq [0,\bar{\gamma}]$ be an interval containing $\gamma_2^*(\tau^2)$ on which $F_2'' \ge c > 0$ and $|F - F_2| \le K\tau^4$. If γ_F minimizes F over I, then*

$$|\gamma_F - \gamma_2^*| \;\le\; 2\sqrt{K/c}\,\tau^2.$$

No claim is made beyond I: uniqueness, global optimality, and the monotonicity of Theorem 6 are statements about the surrogate.

Proof. Strong convexity of F_2 on the interval I gives, for the minimizer γ_2^* of F_2, $\frac{c}{2}(\gamma_F - \gamma_2^*)^2 \le F_2(\gamma_F) - F_2(\gamma_2^*)$. Replace F_2 by F twice, paying $K\tau^4$ each time: $F_2(\gamma_F) - F_2(\gamma_2^*) \le F(\gamma_F) - F(\gamma_2^*) + 2K\tau^4$. Finally $F(\gamma_F) - F(\gamma_2^*) \le 0$ because γ_F minimizes F over I and $\gamma_2^* \in I$. Combine and solve for $|\gamma_F - \gamma_2^*|$. □

Remark 9 (The same dial in meteorology)*.* The trade-off studied here is not particular to finance. The damped Schur complement (1) coincides, term for term, with the reliability shrinkage that keeps Gaussian (Vecchia) pseudo-likelihoods of high-dimensional weather fields estimable when stations outnumber observations [8]. In that language V_0 is the loss of a spatial model that trusts its estimated conditional structure, and the clipped frontier of Theorem 6 says how much conditional structure to trust at a given sampling depth. The scale τ is estimation error, not depth itself; τ^2 is proportional to inverse effective sampling depth, as in the $\tau \asymp T^{-1/2}$ correspondence of Remark 2. The spatial literature also fits the damping intensity from data, which supplies a practical estimate of the noise level $s = \tau^2$ at which to read (30); the correspondence and a comparison of recipes are developed in [8].

8 A one-entry example

The solved family relies on the exchangeability of common cross-block noise. The next example removes that symmetry: the noise sits in a single covariance entry, and the endpoint comparisons are certified exactly.

Example 2 (An interior minimizer, exactly, without common-noise symmetry)**.** *Take the covariance* (16) *and let the estimate err in the* $(1,3)$ *entry only, by* $\pm\frac{3}{20}\cdot 2$ *with probability* $\frac{1}{2}$ *each. Then, in exact rational arithmetic,*

$$F(\tfrac{9}{10}) < F(0) \quad (\textit{margin } 4.879\times 10^{-2}), \qquad F(\tfrac{9}{10}) < F(1) \quad (\textit{margin } 5.702\times 10^{-4}),$$

so every minimizer of F *on* $[0,1]$ *is strictly interior. Both inequalities are strict, and the certified denominator bounds hold uniformly on* $[0,1]$ *and vary continuously with the covariance and noise parameters, so the same conclusion holds on an open neighborhood of these parameters: the phenomenon is robust, not a knife-edge rational coincidence. This is also the minimal non-trivial configuration: because nodes with a singleton child perform no augmentation,* $n \le 3$ *admits no* γ*-dependence at all, and* $n = 4$ *in two blocks of two is the smallest universe on which the bridge exists.*

Proof. $F(0)$, $F(\frac{9}{10})$, $F(1)$ are exact rational numbers, each a finite composition of field operations (4) on rational inputs, and the two inequalities are verified in exact fraction arithmetic by the certificate script. Three values settle the whole interval as follows. Every denominator arising in the recursion is a polynomial in γ with rational coefficients; the script certifies each strictly positive on $[0,1]$, on both noise branches, by exhibiting positive Bernstein coefficients (with exact subdivision where needed). Hence F is continuous on $[0,1]$ and attains its minimum, and $\min F \le F(\frac{9}{10}) < \min\{F(0), F(1)\}$, so no minimizer is an endpoint. The certified polynomials include the determinant numerators of the augmented blocks, and for a symmetric 2×2 block positive diagonal entries together with a positive determinant give positive definiteness, so the augmented blocks are SPD for every γ. Robustness: the certificate proves more than positivity at three points. Each denominator polynomial is strictly positive on all of $[0,1]$, so its minimum over $[0,1]$ is positive. The coefficients of these polynomials are rational functions of the covariance entries and the noise amplitude, with child-block determinants in the denominator, since (4) contains the child-block inverses. Those determinants are nonzero at the stated matrices, so the coefficients vary continuously near them. The minimum of a polynomial over a compact interval is a continuous function of its coefficients, so every denominator stays uniformly positive on an open neighborhood of the stated parameters. On that neighborhood F remains continuous on all of $[0,1]$ and attains its minimum, the three values $F(0)$, $F(\frac{9}{10})$, $F(1)$ vary continuously, and both strict inequalities persist: every minimizer stays interior. $\square$

Remark 10 (A structural cancellation, special to these families)*.* In both noise models on the covariance (16) (the common cross-correlation noise of Section 5 and the one-entry noise of Example 2), the boundary derivative does not move with the noise: exact jet evaluation gives

$$\partial_\gamma F(0;\tau) \;=\; V_0'(0) \;=\; -\tfrac{216}{3125} \qquad \text{exactly, at } \tau \in \{\tfrac{1}{10}, \tfrac{3}{20}, \tfrac{1}{5}\},$$

in both models. Finitely many evaluations suffice because the quantity is affine in τ^2: at $\gamma = 0$ the first γ-variation of each augmented block is quadratic in $\widehat{B} = B \pm \tau E$, by (8), the first variation of the weights is linear in the block variations, and symmetric averaging removes the term odd in τ. Equality with $V_0'(0)$ at $\tau = 0$ and at any one listed τ therefore forces the τ^2 coefficient to vanish.

In the notation of (12), $G'(0) = 0$ for these families; for the solved family this is also visible in (21), where $G(\gamma) \propto \gamma^2(16\gamma+1)$ vanishes quadratically. That is a cancellation particular to these families, not a general fact: Example 1 has $G'(0) = \frac{8}{189} > 0$. It explains why numerical experiments on these families alone can suggest, wrongly, that the noise cost is $O(\gamma^2\tau^2)$ in general.

Remark 11 (Sign, boundedness, and shorting). The sign of $F - V_0$ is not fixed: symmetric noise can reduce expected out-of-sample variance slightly at moderate γ. Nor does the noise cost blow up while Assumption 1 holds: the weights are a continuous function on the compact set K, hence bounded, say $\|w\| \leq L$, and $w^\top \Sigma w \leq \lambda_{\max}(\Sigma)L^2$. A tempting sharper bound fails. The weights sum to one but are not always long-only, so $w^\top \Sigma w \leq \lambda_{\max}(\Sigma)$ is not available. The solved family makes this exact: at $e = c$ on the upper noise branch, $\gamma = 1$ gives the group weight $x_1(2c) = \frac{4}{3}$ and asset weights $(\frac{2}{3}, \frac{2}{3}, -\frac{1}{6}, -\frac{1}{6})$. The estimated covariance there is still positive definite; shorting the volatile block is simply what minimum variance does once the estimated cross covariance exceeds the smaller block variance. What breaks at large noise is also worth stating precisely. The allocation is a rational function of its inputs and becomes undefined only where a denominator vanishes, as at the point $\gamma z = a$ of Lemma 3. Leaving the positive-definite cone does not by itself undefine the recursion; it removes the variance interpretation of the fitness, and implementations must repair or refuse.

9 Closing

Under estimation error the Schur coupling is a bias–variance dial, and the four layers say when to turn it. Whenever coupling has first-order population value, the optimum stays uniformly separated from hierarchical risk parity at all sufficiently small noise levels (Theorem 1), and the noise prices coupling at order $\gamma\tau^2$ (Theorem 2). At an exact minimum-variance endpoint, positive marginal noise sensitivity pushes the optimum inside by order τ^2 (Theorem 3). And in the solved family the exact optimizer is unique, interior, and slides monotonically from full coupling at zero noise to the middle of the bridge at the largest noise the theorem covers (Theorem 4). Dropping the cross-block restriction does not change the verdict: when nonzero, the sign of Ξ gives the second-order direction in which the optimizer leaves full coupling under whole-matrix noise, and in the sample-covariance experiments the median-optimal coupling increases with sample size (Section 6). The exact examples mark what cannot be claimed in general: hierarchical risk parity can become locally optimal at a finite noise level (Example 1), full coupling can stay locally optimal when the noise sits on a single diagonal channel (Remark 6), and the surrogate frontier is a clipped inverse, with each endpoint owning an interval of noise levels unless its limit condition holds (Theorem 6). A walk-forward bridge on Dow constituents is interactive at `schur.microprediction.org`; its minimum sits near $\gamma \approx 0.55$ and survives Ledoit–Wolf shrinkage of the inputs. Every theorem in this paper has a live check at `schur.microprediction.org/min-var-demos.html`, computed from the same recursion. Neither school dominates the other. Each may own an interval of noise levels, and in the second-order surrogate the border between them is drawn at $\tau^2 = -V_0'(\gamma)/G'(\gamma)$.

References

[1] M. López de Prado, *Building diversified portfolios that outperform out of sample*, Journal of Portfolio Management **42**(4) (2016), 59–69.

[2] P. Cotton, *Schur Complementary Portfolios: A Unification of Machine Learning and Optimization-Based Allocation*,

Geek Culture (Medium), 2022. https://medium.com/geekculture/schur-complementary-portfolios-fix-hierarchical-risk-parity-28b0efa1f35f

[3] D. P. Palomar, *Portfolio Optimization: Theory and Application*, Cambridge University Press, 2025. doi:10.1017/9781009428095.

[4] P. Cotton, *Schur complementary allocation: a unification of hierarchical and optimization-based portfolio construction*, 2024. arXiv:2411.05807. https://schur.microprediction.org

[5] A. Antonov, A. Lipton, and M. López de Prado, *Overcoming Markowitz's instability with the help of the hierarchical risk parity (HRP): theoretical evidence*, in Transactions of ADIA Lab, World Scientific, 2025, pp. 87–121. doi:10.1142/9789819813049_0003.

[6] G. Mograby, *Hierarchical minimum variance portfolios: a theoretical and algorithmic approach*, 2025. arXiv:2503.12328.

[7] J. M. Bates and C. W. J. Granger, *The combination of forecasts*, Journal of the Operational Research Society **20**(4) (1969), 451–468. doi:10.1057/jors.1969.103.

[8] P. Cotton, *Two sides of Schur damping: high-dimensional pseudo-likelihoods and portfolio allocation*, 2026. arXiv:2606.14798.

[9] J. D. Jobson and B. Korkie, *Estimation for Markowitz efficient portfolios*, Journal of the American Statistical Association **75**(371) (1980), 544–554. doi:10.1080/01621459.1980.10477507.

[10] R. O. Michaud, *The Markowitz optimization enigma: is 'optimized' optimal?*, Financial Analysts Journal **45**(1) (1989), 31–42. doi:10.2469/faj.v45.n1.31.

[11] R. Jagannathan and T. Ma, *Risk reduction in large portfolios: why imposing the wrong constraints helps*, The Journal of Finance **58**(4) (2003), 1651–1683. doi:10.1111/1540-6261.00580.

[12] V. DeMiguel, L. Garlappi, and R. Uppal, *Optimal versus naive diversification: how inefficient is the $1/N$ portfolio strategy?*, Review of Financial Studies **22**(5) (2009), 1915–1953. doi:10.1093/rfs/hhm075.

[13] O. Ledoit and M. Wolf, *A well-conditioned estimator for large-dimensional covariance matrices*, Journal of Multivariate Analysis **88**(2) (2004), 365–411.