THE STABILISATION CONJECTURE FOR THE CONSTRAINT POLYTOPES OF m-DEPENDENCE TAIL DEPENDENCE MATRICES IS FALSE FOR EVERY $m \geq 3$

KOREA SUPERINTELLIGENCE LABS

Abstract. A tail dependence matrix is the matrix of pairwise tail dependence coefficients of a random vector; the set $\mathcal{T}_d$ of $d \times d$ tail dependence matrices is a polytope whose facets are known only for $d \leq 6$. Inside it, Shyamalkumar and Tao single out the *m-dependence* family — symmetric Toeplitz matrices with unit diagonal, α_k on the k-th off-diagonal for $1 \leq k \leq m$ and 0 beyond — and its *constraint polytope* $P_d^{(m)}$ of admissible parameter vectors. They determine $P_d^{(2)}$, find it to be the same set for every $d \geq 6$, and conjecture that such stabilisation occurs for every m. We refute that conjecture at every $m \geq 3$. For every $m \geq 3$ and every d we exhibit a single closed-form rational point of $P_d^{(m)}$, realised by a linear ramp between two truncated mirror windows, at which $\alpha_{m-1} + 2\alpha_m = 1 + (m-1)/(3d + 2m - 1) > 1$, and we prove, by an explicit nonnegative combination of $4K + 5$ constraint rows, that $\alpha_{m-1} + 2\alpha_m \leq 1 + 1/(K + 1)$ on all of $P_d^{(m)}$ whenever $2m + K(m-1) < d$. Together these show that the point built at dimension d_0 is no longer admissible at dimension $3d_0 + 4m$, so no dimension is a stabilisation dimension, at any $m \geq 3$; with the two-dependence theorem of Shyamalkumar and Tao the conjecture holds exactly for $m \leq 2$. The same two bounds sandwich the gap: the maximum of $\alpha_{m-1} + 2\alpha_m$ over $P_d^{(m)}$ exceeds 1 by $\Theta_m(1/d)$, so the constraint polytopes converge to a limit in this direction and never reach it. The three theorems are formally verified in Lean 4; they are uniform in m and d, use no per-dimension certificate, and each of the eighteen declarations that carries them returns exactly the axiom set {`propext`, `Classical.choice`, `Quot.sound`}, so no step of the refutation is discharged by compiled evaluation. We also record the finite, certificate-based route by which the phenomenon was found, and the constraint polytope $P_8^{(3)}$.

1. Introduction

Tail dependence matrices. Shyamalkumar and Tao [7] work with the *lower* tail dependence coefficient, which they define, verbatim, as follows: "For a random vector (X_1, X_2) with $X_1 \sim F_1$ and $X_2 \sim F_2$, it is defined as

$$\lim_{u \downarrow 0} \frac{\Pr\left(F_1(X_1) \leq u, F_2(X_2) \leq u\right)}{u},$$

given the limit exist; we denote the above limit by $\chi(X_1, X_2)$." Two features of that definition matter below and neither is incidental: the limit is at the *lower* tail, and its existence is assumed rather than proved — the source adds at once that the limit "need not always exist". Nothing is lost by the choice of tail: a footnote there records that for a copula C the lower tail dependence coefficient with respect to C equals the upper tail dependence coefficient with respect to $C^*(u_1, u_2) := C(1 - u_1, 1 - u_2)$, again a copula. The matrix is then defined, again verbatim,

> "The matrix of tail dependence coefficients $(\chi(X_i, X_j))_{1 \leq i,j \leq d}$ corresponding to a d dimensional random vector $\mathbf{X}_{d\times 1}$ is called a Tail Dependence Matrix (TDM). The set of all such matrices is denoted by $\mathcal{T}_d$."

The *realization problem* asks which matrices belong to $\mathcal{T}_d$. The set $\mathcal{T}_d$ is a polytope in $\mathbb{R}^{d(d-1)/2}$; Fiebig, Strokorb and Schlather [2] give its complete facet description for $d \leq 6$ and none is known beyond, and deciding membership is NP-complete — conjectured in [7] and proved by Janßen, Neblung and Stoev [3]. The tractable route into $\mathcal{T}_d$ is Bernoulli compatibility, a notion introduced in [1] and defined there, verbatim, as

> "A $d \times d$ matrix B is a *Bernoulli-compatible matrix*, if $B = \mathbf{E}[\mathbf{X}\mathbf{X}^\top]$ for some $\mathbf{X} \in \mathcal{V}_d$. The set of all $d \times d$ Bernoulli-compatible matrices is denoted by $\mathcal{B}_d$."

(Definition 2.2 of [1]; $\mathcal{V}_d$ is defined in its Definition 2.1 as the random vectors supported by $\{0,1\}^d$.) The source restates the same class as the matrices "which can be expressed as $\mathbf{E}[\mathbf{X}\mathbf{X}^\mathsf{T}]$ where $\mathbf{X}_{d\times 1}$ is a random vector taking values in $\{0,1\}^d$", and writes $\mathcal{B}_d$ for it as well. Theorem 3.3 of [1] then reads, verbatim, "A square matrix with diagonal entries being 1 is a tail-dependence matrix if and only if it is a Bernoulli-compatible matrix multiplied by a constant", identifying $\mathcal{T}_d$ with the Bernoulli compatible matrices of unit diagonal up to a scalar; and the sharpened form used throughout [7], its Theorem 2.1 — stated there without proof and attributed to [2, Remark 17] and to Theorem 1 of Krause, Scherer, Schwinn and Werner [4] — reads, verbatim,

> "Let $\mathbf{T}$ be any $d \times d$ matrix with unit diagonal elements. Then $\mathbf{T} \in \mathcal{T}_d$ if and only if $(1/d)\mathbf{T} \in \mathcal{B}_d$."

Composed with the linear-programming description of $\mathcal{B}_d$ in Theorem 3.1 of [7] — B is Bernoulli compatible if and only if a certain $\big(d(d+1)/2+1\big) \times 2^d$ incidence system $C_d x = p_d(B)$ has a nonnegative solution — this turns membership in $\mathcal{T}_d$ into an explicit linear program over the 2^d subsets of the index set; we write it out as (1) below.

The m-dependence family and the conjecture. The family this paper is about is the m-dependence class of [7], defined in its Proposition 4.1, verbatim, as

> "the class, $\mathcal{T}_{d:m}$, of d-dimensional matrices which are symmetric Toeplitz with elements equal to zero that are not on any of the k-diagonals for $k = 0, \ldots, m$",

that is, the tail dependence matrices of d consecutive observations of a stationary m-dependent process. Writing $\Lambda_d(\alpha)$ for the $d \times d$ symmetric matrix with unit diagonal, α_k on the k-th off-diagonal for $1 \le k \le m$ and 0 beyond, the object of study is the *constraint polytope*

$$P_d^{(m)} \;=\; \big\{\alpha = (\alpha_1, \ldots, \alpha_m) \;:\; \Lambda_d(\alpha) \in \mathcal{T}_d\big\}.$$

The zero entries force the linear program to have support only on the index sets of diameter at most m, of which there are $(d-m+1)2^m - 1$, so the program is of size linear in d; that reduction is the content of Proposition 4.1 of [7]. For $m \le 2$ the source solves the problem completely. Writing (α, β) for (α_1, α_2), it proves, verbatim,

> "For any dimension $d \ge 6$, the two-dependence matrix of the form (14) is in $\mathcal{T}_d$ if and only if α and β satisfy the following constraints: $\alpha, \beta \ge 0$; $\alpha + 4\beta \le 2$; $2\alpha - \beta \le 1$."

This is Proposition 4.2 of the preprint v2 — Proposition 21 of the journal version — and its equation (14) is the matrix $\Lambda_d(\alpha, \beta)$ displayed in banded form. In Remark 4.7 the source tabulates three *different* regions at $d = 3, 4, 5$:

d	$P_d^{(2)}$
3	$\alpha \ge 0$, $0 \le \beta \le 1$, $2\alpha - \beta \le 1$
4	$\alpha, \beta \ge 0$, $\alpha + \beta \le 1$, $2\alpha - \beta \le 1$
5	$\alpha \ge 0$, $0 \le \beta \le \frac{1}{2}$, $\alpha + \beta \le 1$, $2\alpha - \beta \le 1$
≥ 6	$\alpha, \beta \ge 0$, $\alpha + 4\beta \le 2$, $2\alpha - \beta \le 1$

So at $m = 2$ the constraint polytopes shrink three times and then stop. The paper closes with the conjecture that this is a general phenomenon. We quote it verbatim from the electronic preprint, adding only the emphasis:

> "In this connection, it is interesting to note that in the two-dependence class of Proposition 4.2, the polytopes coincide for $d \ge 6$. **We conjecture that the latter stabilization of the constraint polytopes holds for all m-dependence TDMs of Proposition 4.1.**"

All numbered references to [7] in this paper are to the numbering of its arXiv version v2 (2019-08-01), recovered by compiling that e-print; the only journal number we have been able to check is Proposition 21, which is Proposition 4.2 there.

Definition 1.1. Call d_0 a *stabilisation dimension* for m when $P_d^{(m)} = P_{d_0}^{(m)}$ for every $d \ge d_0$, and write $d_0(m)$ for the least one if it exists.

The conjecture says that $d_0(m)$ exists for every m. The only values in print are $d_0(2) = 6$ and $d_0(1) = 3$, the latter implicit in the same paper's Remark 2.2: "as shown later in Proposition 4.2 and Remark 4.7, in the 1-dependent case while $\alpha \in [0, 1/2]$ is necessary and sufficient for $d \ge 3$, it is only sufficient when $d = 2$". Nothing in the literature bounds $d_0(m)$ from either side for any $m \ge 3$, and the only part of $P_d^{(m)}$ described anywhere for $m \ge 3$ is its face $\alpha_1 = 0$, determined by [8] (Remark 3.6). The conjecture is repeated, unresolved, in the 2020 doctoral thesis of the second author of [7] [9].

What is settled here. The conjecture is false, at every $m \geq 3$ and at every dimension. We prove three theorems, all uniform in m and d: a construction, a matching bound, and their combination. Throughout, $m \geq 3$ and $N = N(d,m) = 3d + 2m - 1$.

(A) *(Theorem 4.1.)* For every $m \geq 3$ and every $d \geq 1$ the parameter vector

$$\alpha_1^{(d,m)} = \frac{d}{N}, \qquad \alpha_{m-1}^{(d,m)} = \frac{d+m}{N}, \qquad \alpha_m^{(d,m)} = \frac{d+m-1}{N}, \qquad \alpha_k^{(d,m)} = 0 \ \text{ otherwise},$$

lies in $P_d^{(m)}$, and $\alpha_{m-1}^{(d,m)} + 2\alpha_m^{(d,m)} = 1 + \frac{m-1}{N} > 1$. The realisation is one closed formula: a linear ramp of slope $1/N$ between two truncated mirror windows.

(B) *(Theorem 5.1.)* For every $m \geq 3$, every d and every integer $K \geq 0$ with $2m + K(m-1) < d$, every $\alpha \in P_d^{(m)}$ satisfies $\alpha_{m-1} + 2\alpha_m \leq 1 + \frac{1}{K+1}$. The proof is an explicit nonnegative combination of $4K+5$ rows of the linear program, valid by a three-case argument and not by enumeration.

(C) *(Theorem 6.1.)* For every $m \geq 3$ and every d_0,

$$\alpha^{(d_0,m)} \in P_{d_0}^{(m)} \qquad \text{and} \qquad \alpha^{(d_0,m)} \notin P_{3d_0+4m}^{(m)}.$$

Hence no dimension is a stabilisation dimension: $d_0(m)$ does not exist for any $m \geq 3$.

Three consequences are worth stating at once. The conjecture is true *exactly* in the range where it was verified: at $m = 2$ by the source's own Proposition 4.2 quoted above, at $m = 1$ by its Remark 2.2, and at no $m \geq 3$. The bounds (A) and (B) sandwich the support value $h_d^{(m)} = \max\{\alpha_{m-1} + 2\alpha_m : \alpha \in P_d^{(m)}\}$ between $1 + \frac{m-1}{3d+2m-1}$ and $1 + \frac{m-1}{d-2m-1}$, so $h_d^{(m)} - 1 = \Theta_m(1/d)$: in this one direction the constraint polytopes converge to a limit at rate exactly $1/d$ and never reach it (Remark 6.3). And the mechanism is visible. For $m = 3, 4, 5, 6$ — the range in which we have computed it (Remark 2.7) — the limiting value 1 is exactly the maximum of $\alpha_{m-1} + 2\alpha_m$ over the *moving-maximum polytope* $C^{(m)}$ of Proposition 2.6, the region generated by the window patterns, so what (A) produces at those m is a banded tail dependence matrix lying outside $C^{(m)}$ at every finite dimension, by a margin that vanishes.

Two hypotheses are load-bearing and we check both by explicit counterexample (Proposition 6.4): $m \geq 3$ in (A), which fails at $m = 2$, where the two mirror windows coincide and the construction breaks — exactly why the conjecture is true there — and the fit hypothesis $2m + K(m-1) < d$ in (B), which fails at small d, where the bound is false.

Sections 7 and 8 record the finite, certificate-based route by which the phenomenon was found before (A)–(C) were available, because it carries information that (A)–(C) do not: thirteen explicit witnesses of $P_{d+1}^{(3)} \subsetneq P_d^{(3)}$ at consecutive dimensions $4 \leq d \leq 16$ (Theorem 7.1) — Theorem 6.1 gives a drop between d_0 and $3d_0 + 4m$, not at every step — and the exact facet description of $P_8^{(3)}$ (Theorem 8.1 and Remark 8.2), as far as we know the first constraint polytope of an m-dependence class with $m \geq 3$ to be determined.

What is not settled. Three things. The exact value of $h_d^{(3)}$, which exact linear programming gives as $1 + 1/\lceil 3(d-4)/2 \rceil$ for every d from 8 to 32 (Remark 7.3); Theorem 4.1 has the right leading constant and Theorem 5.1 is loose by a factor tending to 3. Whether $P_{d+1}^{(m)}$ is *strictly* smaller than $P_d^{(m)}$ at every single step: the computation says yes at $m = 3$ up to $d = 32$ and Theorem 7.1 certifies it to $d = 16$, but Theorem 6.1 gives a drop only over a gap of ratio 3. And whether $\bigcap_d P_d^{(m)}$ equals the moving-maximum polytope: at $m = 3, 4, 5, 6$ the two have the same support value in the direction $(0, \dots, 0, 1, 2)$, by Remarks 2.7 and 6.3, and every other direction is open.

Related work, and the search behind the novelty claim. To the best of our knowledge Theorems 4.1, 5.1 and 6.1, the finite chain of Theorem 7.1 and the facet description of Theorem 8.1 are new; the negative should be read as "not found", not as "does not exist". We record what was searched, on 2–3 September 2026.

Citation closure was taken in OpenAlex on six hubs — the source by DOI and by arXiv identifier, [2], [1], Strokorb and Schlather's paper on extremal coefficient and extremal correlation functions, and [3] — together with the complete works lists of Strokorb, Tao and Shyamalkumar (32, 14 and 47 works); no title or abstract concerns m-dependence, banding, Toeplitz tail dependence matrices or polytope stabilisation, and Semantic Scholar returns the same citation sets. The live arXiv API was queried on about twenty phrase keys with a positive control first: `all:"tail correlation function"` returns exactly two papers ever, while `"moving maxima"` with `"tail correlation"`, `"m-dependent"` with `"max-stable"`, `"extremal coefficient"` with `"Toeplitz"`, `"cut polytope"` with `"tail dependence"` and seven further pairs all return nothing. A local full-text arXiv corpus of 86 GB (all 369 shards) was swept three times: only eleven documents in it contain both `tail dependence` and `Toeplitz`, the source being the only on-topic one, and the phrase `stabilization of the constraint` occurs in exactly one document in the whole corpus, namely the source itself.

None of the works citing the source touches this conjecture. [3] settles the source's *other* conjecture, the NP-completeness of the realization problem, and its text contains no occurrence of `m-dependen`, `Toeplitz` or `stabiliz`. The closest published follow-up is [8], by the second author of [7]: it treats symmetric Toeplitz matrices with exactly *two* nonzero off-diagonal bands, never three, never mentions the stabilisation conjecture, and describes [7] as giving conditions "for 1- and 2-dependent matrices". It also explains why the construction of Theorem 4.1 was not found earlier: [8, Theorem 3.2] proves stabilisation in every two-band case, and the two-band shadow $\alpha_1 = 0$ of our point violates that theorem's bound (Remark 3.6), so the first off-diagonal entry $\alpha_1 > 0$ is essential and no work inside the two-band class could have produced the point.

Two residual risks are worth naming. *Non-arXiv venues:* [8] is itself the proof that this matters — it is in an open-access journal, is not on arXiv, and was invisible to every corpus and arXiv query, reached only through the citation graph; the mitigation is that the citation closure of all six hubs was taken completely, and a paper refuting this conjecture would almost certainly cite [7]. *Vocabulary:* someone could state the closely related max-stable fact — for every d there is an m-dependent stationary tail correlation function on d consecutive indices that is not the tail correlation function of any moving maximum with window $\{0, \dots, m\}$ — in language matching none of our phrase keys; we mitigated this by enumerating complete works lists rather than only citation lists, and did not eliminate it. Finally, the journal version of [7] is paywalled and its body was not read here; every quotation above is from the electronic preprint (v2, 2019-08-01), whose file was re-downloaded and checksummed during this work. What the journal version's open front matter does settle is that its abstract, its appendix figures and its reference list agree with that preprint item for item, and that the two-dependence result is its Proposition 21 — the caption of the last appendix figure there reads "Construction of Sets A_i, $i = 1, \dots, d$, for any d: 2-Dependent TDM of Proposition 21", where the preprint's caption names the same unnumbered result. That, the description of [7] in [8] and the recurrence of the conjecture in [9] are the indirect evidence that the journal version also stops at $m \le 2$; we have not verified it directly.

Organisation. Section 2 sets up the linear program, the certificate formats and the moving-maximum polytope, and Section 3 reproduces known results as controls on the instrument. Sections 4–6 prove Theorems 4.1–6.1; Sections 7 and 8 record the finite route and $P_8^{(3)}$; and Section 9 describes the formal verification and delimits it.

2. The realization problem as a linear program

Throughout, $I_d = \{0, 1, \dots, d-1\}$ indexes the coordinates, and subsets $S \subseteq I_d$ are identified with their indicator vectors $\mathbf{1}_S \in \{0,1\}^d$. For a weight function $v : 2^{I_d} \to \mathbb{Q}$ put

$$\mathrm{cov}_1(v, i) = \sum_{S \ni i} v_S, \qquad \mathrm{cov}_2(v, i, j) = \sum_{S \supseteq \{i,j\}} v_S.$$

We write $\operatorname{diam} S = \max S - \min S$.

Definition 2.1. Let $\Lambda \in \mathbb{Q}^{d \times d}$ be symmetric with unit diagonal. Composing Theorems 2.1 and 3.1 of [7], recorded in Section 1, at $B = \Lambda/d$ gives: $\Lambda \in \mathcal{T}_d$ if and only if there is $x : 2^{I_d} \to \mathbb{Q}$ with

$$x \ge 0, \qquad \sum_{S \subseteq I_d} x_S = 1, \qquad \mathrm{cov}_1(x, i) = \tfrac{1}{d} \ \ (i \in I_d), \qquad \mathrm{cov}_2(x, i, j) = \tfrac{\Lambda_{ij}}{d} \ \ (i \ne j). \tag{1}$$

Here x_S is the probability that the Bernoulli vector realising Λ/d equals $\mathbf{1}_S$. We call (1) the *normalised form*, and the rescaled system

$$v \ge 0, \qquad \mathrm{cov}_1(v, i) = 1 \ \ (i \in I_d), \qquad \mathrm{cov}_2(v, i, j) = \Lambda_{ij} \ \ (i \ne j) \tag{2}$$

the *scaled form*. Parameter vectors are rational throughout, as in the formal development; since $P_d^{(m)}$ is cut out by rational data, equality of two of these polytopes over $\mathbb{R}$ implies equality of their rational points, so nothing below is weakened by reading $P_d^{(m)}$ as its set of rational points.

Proposition 2.2. *For every $d \geq 1$ and every symmetric $\Lambda \in \mathbb{Q}^{d\times d}$, the systems* (1) *and* (2) *are simultaneously solvable.*

Proof. From x to v, take $v = d\,x$. Conversely, given v, put $A = \sum_{S\neq\emptyset} v_S$. Counting each nonempty S at least once, $A \leq \sum_S |S|\, v_S = \sum_{i\in I_d} \operatorname{cov}_1(v,i) = d$, so $x_\emptyset = 1 - A/d \geq 0$ and $x_S = v_S/d$ for $S \neq \emptyset$ solves (1): the empty set lies in no coverage sum, so only the total-mass equation sees it, and that equation is exactly what $x_\emptyset$ was chosen to satisfy. The one inequality used is the double count, which is where the total-mass row of (1) goes: it is slack in the scaled form. □

All arguments below use the scaled form. Both directions of membership carry finite certificates, and the soundness of both formats is linear-programming duality specialised to (2).

Proposition 2.3. *Let $\Lambda \in \mathbb{Q}^{d\times d}$ be symmetric with unit diagonal.*

(1) (Membership.) *Let $T_1, \dots, T_n \subseteq I_d$ and $c_1, \dots, c_n \in \mathbb{Q}_{\geq 0}$ satisfy $\sum_{k:\, i\in T_k} c_k = 1$ for every $i \in I_d$ and $\sum_{k:\, \{i,j\}\subseteq T_k} c_k = \Lambda_{ij}$ for all $i \neq j$. Then Λ solves* (2)*, hence $\Lambda \in \mathcal{T}_d$.*

(2) (Non-membership.) *Let $u \in \mathbb{Q}^d$ and $z \in \mathbb{Q}^{d\times d}$, and put*

$$g(S) = \sum_{i\in S} u_i + \sum_{\substack{i<j\\ i,j\in S}} z_{ij}, \qquad f(\Lambda) = \sum_{i\in I_d} u_i + \sum_{i<j} z_{ij}\Lambda_{ij}.$$

If $g(S) \leq 0$ for every *$S \subseteq I_d$ and $f(\Lambda) > 0$, then Λ does not solve* (2)*, hence $\Lambda \notin \mathcal{T}_d$.*

(3) (Objective identity.) *For the 3-dependence matrix, $f(\Lambda_d(\alpha)) = \sum_i u_i + \zeta_1\alpha_1 + \zeta_2\alpha_2 + \zeta_3\alpha_3$, where $\zeta_k = \sum_{j-i=k,\, i<j} z_{ij}$ is the sum of the coefficients of z along the k-th diagonal.*

Proof. (1) is the substitution $v_S = \sum_{k:T_k=S} c_k$. For (2), if v solved (2) then, exchanging the two summations,

$$0 \;\geq\; \sum_{S\subseteq I_d} v_S\, g(S) \;=\; \sum_{i\in I_d} u_i \operatorname{cov}_1(v,i) + \sum_{i<j} z_{ij} \operatorname{cov}_2(v,i,j) \;=\; f(\Lambda),$$

contradicting $f(\Lambda) > 0$. Here $g(S)$ is the value of the linear form $M \mapsto \sum_i u_i M_{ii} + \sum_{i<j} z_{ij} M_{ij}$ at the rank-one 0/1 matrix $\mathbf{1}_S\mathbf{1}_S^\intercal$ and $f(\Lambda)$ its value at Λ, so a functional nonpositive at every such rank-one matrix and positive at Λ separates Λ from the cone they generate. (3) is bookkeeping: $\Lambda_d(\alpha)_{ij}$ depends only on $j-i$ and vanishes for $j - i > 3$. □

Two consequences of the zero pattern are used constantly. The first is the support reduction already present in [7]; we record it because it is what makes Theorem 5.1 a statement about few subsets rather than about all 2^d of them.

Lemma 2.4. *Let $v \geq 0$ solve* (2) *for a matrix Λ with $\Lambda_{ij} = 0$, and let $i, j \in S$. Then $v_S = 0$. In particular, if $\Lambda = \Lambda_d(\alpha)$ is an m-dependence matrix then $v_S = 0$ for every S with* $\operatorname{diam} S > m$.

Proof. $\operatorname{cov}_2(v,i,j) = 0$ is a sum of nonnegative terms one of which is v_S. □

Lemma 2.5 (proved by hand; outside the formal development)**.** *$P_{d+1}^{(m)} \subseteq P_d^{(m)}$ for every $d \geq 1$ and every m.*

Proof. Given a solution v of (2) at dimension $d+1$ for $\Lambda_{d+1}(\alpha)$, push it forward along $S \mapsto S \cap I_d$. Coverage sums over indices in I_d are unchanged, so the image solves the system at dimension d for $\Lambda_d(\alpha)$, which is the leading principal $d\times d$ submatrix of $\Lambda_{d+1}(\alpha)$. (The same principal-submatrix argument is used in [7] in the proof of the two-dependence region.) □

Lemma 2.5 is what turns a point of $P_{d_0}^{(m)} \setminus P_{d_1}^{(m)}$ into a *strict inclusion* of sets. It is deliberately *not* part of the formal development, and Theorem 6.1 does not use it: Definition 1.1 asks for equality of the two sets, so exhibiting one parameter vector in one and not the other refutes stabilisation on its own. Every statement below says exactly which of the two it is.

The last ingredient is the region the polytopes converge to.

Proposition 2.6 (proved by hand; outside the formal development)**.** *For $\emptyset \neq R \subseteq \{0,1,\ldots,m\}$ put $n_k(R) = \#\{r \in R : r+k \in R\}$, and let*

$$C^{(m)} = \operatorname{conv}\Big\{ \big(\tfrac{n_1(R)}{|R|}, \ldots, \tfrac{n_m(R)}{|R|}\big) : \emptyset \neq R \subseteq \{0,\ldots,m\} \Big\} \subseteq \mathbb{R}^m,$$

the moving-maximum polytope. *Then $C^{(m)} \subseteq P_d^{(m)}$ for every $d \geq 1$. Moreover $C^{(2)}$ is exactly the source's stabilised region $\{\alpha, \beta \geq 0,\ \alpha{+}4\beta \leq 2,\ 2\alpha{-}\beta \leq 1\}$, and $C^{(3)}$ has the seven vertices $(0,0,0)$, $(\frac12,0,0)$, $(0,\frac12,0)$, $(0,0,\frac12)$, $(\frac23,\frac13,0)$, $(\frac13,\frac13,\frac13)$, $(\frac34,\frac12,\frac14)$ and the nine facets $\alpha_1,\alpha_2,\alpha_3 \geq 0$; $2\alpha_1-\alpha_2 \leq 1$; $\alpha_1+4\alpha_2-3\alpha_3 \leq 2$; $-\alpha_1+4\alpha_2+3\alpha_3 \leq 2$; $-\alpha_1+2\alpha_2+2\alpha_3 \leq 1$; $2\alpha_1-2\alpha_2+2\alpha_3 \leq 1$; $\alpha_2+2\alpha_3 \leq 1$.*

Proof. Since $P_d^{(m)}$ is convex — a convex combination of two solutions of (2) is a solution — it suffices to realise each generator. Fix R and put weight $1/|R|$ on each of the sets $(t+R) \cap I_d$, $t \in \mathbb{Z}$ (finitely many are nonempty). For $i \in I_d$ the sets containing i are those with $t = i-r$, $r \in R$, so the coverage of i is $|R| \cdot (1/|R|) = 1$; the sets containing both i and $i+k$ are those with $t = i-r$ and $r+k \in R$, so that coverage is $n_k(R)/|R|$, independent of i; and it is 0 for $k > m$ because $R \subseteq \{0,\ldots,m\}$. Truncation at the two ends deletes only indices outside I_d and changes none of these counts. The vertex and facet lists of $C^{(2)}$ and $C^{(3)}$ are exact rational hull computations over the $2^{m+1}-1$ generators. ☐

The name is the source's own: at $m=2$ the region above is what [7] obtains analytically from the moving-maximum process $Y_i = \max\{cX_{i-2}, bX_{i-1}, aX_i\}$, with $\alpha = a \wedge b + b \wedge c$ and $\beta = a \wedge c$, and states that every (α,β) of its Proposition 4.2 is attained that way. The inequality $\alpha_{m-1} + 2\alpha_m \leq 1$, a facet of $C^{(3)}$ and valid on $C^{(m)}$ for $m = 3,4,5,6$ by Remark 2.7, is the direction this paper is about.

Remark 2.7 (computation outside the formal development). Over the generators of $C^{(m)}$ the maximum of $\alpha_{m-1}+2\alpha_m$ is exactly 1 for $m=3,4,5,6$ — attained at $R = \{0,m\}$, among others — while at $m=2$ it is $4/3$, attained only at $R=\{0,1,2\}$. So $\alpha_{m-1}+2\alpha_m \leq 1$ holds on all of $C^{(m)}$ for $m=3,4,5,6$ and fails on $C^{(2)}$: one of several places where $m=2$ is degenerate for the argument of this paper. This was computed at $m=3,4,5,6$ and is claimed for no larger m; no numbered statement below depends on it.

3. Controls: the published two-dependence theory

Before any $m \geq 3$ claim we reproduce what is already known. Write $\Lambda_d(\alpha,\beta)$ for the two-dependence matrix.

Remark 3.1 (computation outside the formal development). Each of the four published two-dependence regions displayed in Section 1 was recomputed here from Definition 2.1 alone, as a whole region and in both inclusions, by exact rational linear programming: starting from the box $0 \leq \alpha,\beta \leq 1$, each vertex of the current outer approximation is tested by an exact rational simplex, and a failing vertex is cut off by the Farkas vector the infeasible program returns, so the loop terminates with the exact H-description rather than with a grid sample. The output agrees with the source at $d=3,4,5$ and at $d=6,7,8,9$. The same computation gives $P_2^{(1)} = \{0 \leq \alpha_1 \leq 1\}$ and $P_d^{(1)} = \{0 \leq \alpha_1 \leq \frac12\}$ for every d with $3 \leq d \leq 9$. The one-dependence region is also the source's: its Remark 2.2 states that "in the 1-dependent case while $\alpha \in [0,1/2]$ is necessary and sufficient for $d \geq 3$, it is only sufficient when $d=2$", which is exactly what the computation returns in the range $3 \leq d \leq 9$ and which gives $d_0(1) = 3$ outright, not only in the computed range. All of this is a check on our instrument, not a result.

The next four propositions are the controls inside the formal development. Each is decided by the source's published two-dependence results, and each is proved by the certificate formats of Proposition 2.3, so that both formats are tested against published mathematics.

Proposition 3.2. *The three vertices $(\alpha,\beta) = (2/3,1/3)$, $(0,1/2)$ and $(1/2,0)$ of the source's $d \geq 6$ two-dependence region are realised at $d=6$: each carries an explicit nonnegative rational weight vector supported on at most eight subsets of I_6.*

Proposition 3.3. $(\alpha,\beta) = (0,3/5)$*, which violates $\alpha+4\beta \leq 2$, and $(\alpha,\beta) = (4/5,0)$, which violates $2\alpha-\beta \leq 1$, satisfy $\Lambda_6(\alpha,\beta) \notin \mathcal{T}_6$. Each is refuted by an explicit separating functional in the sense of Proposition 2.3 (2).*

Proposition 3.4. $\Lambda_3(1,1) \in \mathcal{T}_3$ *and* $\Lambda_4(1,1) \notin \mathcal{T}_4$.

Proposition 3.5. *$\Lambda_9(0,0,0)$, the 9×9 identity matrix, lies in $\mathcal{T}_9$.*

Proofs. All four are certificate checks. For Propositions 3.2, 3.4 and 3.5 the weight vectors are exhibited explicitly and Proposition 2.3 (1) applies; in Proposition 3.5 the weights are unit masses on the nine singletons. For Proposition 3.3 and the second half of Proposition 3.4, an explicit integer pair (u,z) is exhibited and the hypothesis $g(S)\le 0$ is checked over all $2^6=64$, respectively $2^4=16$, subsets; that check is exhaustive and is performed by machine. □

Proposition 3.3 is the negative control on which everything rests: a point a published theorem declares infeasible, refuted by our route. Proposition 3.4 shows that the source's small-d regions genuinely differ, so that the shrinking studied below is already visible at $m=2$; Proposition 3.5 rules out the degenerate reading in which the membership predicate is empty and every non-membership claim holds vacuously. Every control so far is at $m\le 2$. The next one is not.

Remark 3.6 (check against a published $m\ge 3$ theorem; outside the formal development). [8, Theorem 3.2] decides membership in $\mathcal{B}_d$ for a symmetric Toeplitz matrix with $1/d$ diagonal and exactly *two* nonzero off-diagonals, at positions m' and $m'+n$ (its own letters are m and n; we write m' for its m, which is not ours). Its Case 2, the case $m'\ne n$, reads, verbatim, "For $2m+2n+1\le d$, $\alpha,\beta\ge 0$; $2\alpha+2\beta\le 1/d$". At $(m',n)=(m-1,1)$ this is the face $\alpha_1=0$ of our polytope: for $d\ge 2m+1$, the m-dependence matrix with $\alpha_1=0$ is a tail dependence matrix if and only if $\alpha_{m-1},\alpha_m\ge 0$ and $2\alpha_{m-1}+2\alpha_m\le 1$. Our exact rational simplex reproduces that region at all sixteen tested points ($d=7,8,9,10$ against four parameter pairs, including the boundary point $(0,1/4,1/4)$ and the violating point $(0,1/4,13/50)$); this is the first check of the $m\ge 3$ instrument against a published theorem. It also locates Theorem 4.1: [8] proves stabilisation in every two-band case, and the two-band shadow of our point has $2(\alpha_{m-1}+\alpha_m)=2(2d+2m-1)/(3d+2m-1)\to 4/3>1$, so that shadow is refuted by [8] and the entry $\alpha_1>0$ is what puts the point outside the two-band world.

4. Theorem A: a ramp between two mirror windows

Fix $m\ge 3$ and $d\ge 1$ and put $N=3d+2m-1$. The construction uses two window shapes,

$$\{0,1,m\}\qquad\text{and}\qquad\{0,m-1,m\},$$

which are mirror images of one another inside $\{0,\dots,m\}$. Each realises, on its own, $\alpha_1=\alpha_{m-1}=\alpha_m=\frac{1}{3}$, so $\alpha_{m-1}+2\alpha_m=1$ exactly: they are among the generators of $C^{(m)}$ at which that quantity is largest. The point of Theorem 4.1 is obtained by sliding linearly from one to the other across the index set.

For $t=-m,-m+1,\dots,d-1$ let

$$S_t=(t+\{0,1,m\})\cap I_d,\qquad T_t=(t+\{0,m-1,m\})\cap I_d,$$

and give them the weights

$$c(S_t)=\frac{d-1-t}{N},\qquad c(T_t)=\frac{t+m}{N}.$$

Both weights are nonnegative exactly on the stated range of t, and both ends of the range sit on the boundary of the nonnegativity box: the slope $1/N$ is the steepest this two-window ansatz allows.

Theorem 4.1. *Let $m\ge 3$ and $d\ge 1$, and put $N=3d+2m-1$ and*

$$\alpha_1^{(d,m)}=\frac{d}{N},\qquad \alpha_{m-1}^{(d,m)}=\frac{d+m}{N},\qquad \alpha_m^{(d,m)}=\frac{d+m-1}{N},\qquad \alpha_k^{(d,m)}=0\ \ \text{for all other } k.$$

Then $\Lambda_d(\alpha^{(d,m)})\in\mathcal{T}_d$, that is $\alpha^{(d,m)}\in P_d^{(m)}$, realised by the weights above; and

$$\alpha_{m-1}^{(d,m)}+2\,\alpha_m^{(d,m)}\ =\ 1+\frac{m-1}{N}\ >\ 1.$$

Proof. Apply Proposition 2.3 (1) to the $2(d+m)$ listed sets with the listed weights; nonnegativity is the range of t. Truncation is invisible to the equations, because each equation asks only which t contribute at indices *inside* I_d. Since $m\ge 3$, the three differences within each window, namely 1, $m-1$ and m, are pairwise distinct; this is the only place the hypothesis is used, and it is where the argument fails at $m=2$ (Proposition 6.4).

Coverage of a point. For $i \in I_d$ the sets S_t containing i are those with $t \in \{i, i-1, i-m\}$ and the sets T_t containing i those with $t \in \{i, i-m+1, i-m\}$, and no other; a value of t outside $[-m, d-1]$ contributes nothing because the corresponding set is then empty at i. Summing the six weights,

$$\frac{(d-1-i)+(d-i)+(d-1-i+m)}{N} + \frac{(i+m)+(i+1)+i}{N} = \frac{3d+2m-1}{N} = 1.$$

Coverage of a pair. Let $i < j$ and $k = j - i$. A set S_t contains both i and j only if k is a difference of two elements of $\{0, 1, m\}$, that is $k \in \{1, m-1, m\}$, and then exactly one t works: $t = i$ for $k = 1$, $t = i-1$ for $k = m-1$, $t = i$ for $k = m$. The same holds for T_t with the differences of $\{0, m-1, m\}$, namely 1, $m-1$, m again: $t = i-m+1$ for $k = 1$, $t = i$ for $k = m-1$, $t = i$ for $k = m$. Adding the two contributions,

$$k=1: \ \frac{(d-1-i)+(i+1)}{N} = \frac{d}{N}, \qquad k=m-1: \ \frac{(d-i)+(i+m)}{N} = \frac{d+m}{N},$$
$$k=m: \ \frac{(d-1-i)+(i+m)}{N} = \frac{d+m-1}{N}, \qquad k \notin \{1, m-1, m\}: \ 0.$$

Each value is independent of i, which is the Toeplitz requirement, and they are exactly the entries of $\Lambda_d(\alpha^{(d,m)})$. Finally $\alpha_{m-1} + 2\alpha_m = \big((d+m) + 2(d+m-1)\big)/N = (3d+3m-2)/N = 1 + (m-1)/N$. ☐

No search is involved and no case distinction on d: the theorem is one algebraic identity in (m, d), and the formal proof is a proof of that identity, not an enumeration. At $(m, d) = (3, 4)$ the point is $(4, 7, 6)/17$, at $(3, 8)$ it is $(8, 11, 10)/29$, and at $(4, 10)$ it is $\alpha_1 = 10/37$, $\alpha_3 = 14/37$, $\alpha_4 = 13/37$, with excess $1 + 3/37$.

Remark 4.2 (computation outside the formal development). Every coverage equation of the weight vector above was also evaluated directly, in exact rational arithmetic, for $m = 3, \ldots, 8$ and $d = m+1, \ldots, 30$: no equation is violated, and the excess is $1 + (m-1)/N$ in every case. At $m = 2$ the same code breaks $d-1$ equations at each d. Combined with Remark 2.7, the point of Theorem 4.1 lies in $P_d^{(m)} \setminus C^{(m)}$ at every finite d, for $m = 3, 4, 5, 6$: a banded d-dimensional tail dependence matrix that the window patterns of Proposition 2.6 do not generate. Whether the moving maxima with window $\{0, \ldots, m\}$ realise exactly $C^{(m)}$ — they do at $m = 2$, by the source's own remark — is not decided here.

5. Theorem B: a dual bound on all of $P_d^{(m)}$

Theorem 4.1 is useless without a matching upper bound, and the bound has to be uniform in d with the right rate: what refutes stabilisation is not one point but the fact that the maximum of $\alpha_{m-1} + 2\alpha_m$ *strictly decreases* past any given dimension.

Theorem 5.1. *Let $m \geq 3$, let $K \geq 0$ be an integer and let d satisfy $2m + K(m-1) < d$. Then every $\alpha \in P_d^{(m)}$ satisfies*

$$\alpha_{m-1} + 2\alpha_m \ \leq \ 1 + \frac{1}{K+1}.$$

Proof. Put $a_j = m + j(m-1)$ for $0 \leq j \leq K$, an arithmetic progression of step $m-1$; the fit hypothesis says exactly that $a_K + m < d$, so all of $a_j - m = j(m-1)$, a_j, $a_j + m - 1$ and $a_j + m$ lie in I_d for every $j \leq K$. Let $v \geq 0$ solve (2) for $\Lambda_d(\alpha)$, and consider the nonnegative combination of $4K+5$ rows

$$Q = \sum_{j=0}^{K} \Big(\mathrm{cov}_1(v, a_j) - \mathrm{cov}_2(v, a_j - m, a_j) - \mathrm{cov}_2(v, a_j, a_j + m - 1) - \mathrm{cov}_2(v, a_j, a_j + m) \Big) \ + \ \mathrm{cov}_1(v, a_K).$$

The value of Q. Each cov_1 equals 1, and the three pair terms in the j-th bracket are the entries of $\Lambda_d(\alpha)$ at distances m, $m-1$ and m respectively, namely α_m, α_{m-1} and α_m. Writing $s = \alpha_{m-1} + 2\alpha_m$, each bracket equals $1 - s$ and

$$Q = (K+1)(1-s) + 1.$$

The coefficient of Q at a weight. Exchanging the order of summation, $Q = \sum_S v_S \, c(S)$ with

$$c(S) = \sum_{j=0}^{K} \tau_j(S) + \mathbf{1}[a_K \in S],$$

$$\tau_j(S) = \mathbf{1}[a_j \in S]\big(1 - \mathbf{1}[a_j - m \in S] - \mathbf{1}[a_j + m - 1 \in S] - \mathbf{1}[a_j + m \in S]\big).$$

The key inequality: $c(S) \geq 0$ *whenever* $\operatorname{diam} S \leq m$. By Lemma 2.4 only such S carry weight, so this suffices. Three observations.

(1) $\tau_j(S) \geq 0$ unless a_j, $a_j + m - 1$ and $a_j + m$ all lie in S, in which case $\tau_j(S) = -1$. Indeed, if $a_j - m \in S$ then neither $a_j + m - 1$ nor $a_j + m$ can lie in S, being at distance $2m - 1 > m$ and $2m > m$ from $a_j - m$; so the bracket is at least $1 - 1 = 0$.

(2) At most one j has $\tau_j(S) = -1$. If $j < i$ were two such indices then $a_j \in S$ and $a_i + m \in S$, at distance $m + (i - j)(m - 1) > m$.

(3) If $\tau_j(S) = -1$ and $j < K$ then $\tau_{j+1}(S) = +1$. For $S \subseteq [a_j, a_j + m]$ by (1) and $\operatorname{diam} S \leq m$, and $a_{j+1} = a_j + m - 1 \in S$, while $a_{j+1} - m = a_j - 1$, $a_{j+1} + m - 1 = a_j + 2m - 2$ and $a_{j+1} + m = a_j + 2m - 1$ all lie outside $[a_j, a_j + m]$ — the middle one because $m \geq 3$.

If no $\tau_j(S)$ is negative then $c(S) \geq 0$ because the extra indicator is nonnegative. If $\tau_j(S) = -1$ with $j < K$ then $\tau_{j+1}(S) = +1$ by (3) and all other terms are nonnegative by (1) and (2). If $\tau_K(S) = -1$ then $a_K \in S$ by (1), so the extra term $\mathbf{1}[a_K \in S] = 1$ pays for it. In all three cases $c(S) \geq 0$.

Conclusion. $Q = \sum_S v_S c(S) \geq 0$, so $(K + 1)(1 - s) + 1 \geq 0$, that is $s \leq 1 + 1/(K + 1)$. ☐

Note what the proof is not: it is not a check that a candidate functional is nonpositive at all 2^d rank-one 0/1 matrices. Lemma 2.4 restricts attention to the sets of diameter at most m, and there the three observations decide the sign; that is what makes Theorem 5.1 uniform in d.

Taking the largest admissible K, namely $K = \lfloor (d - 2m - 1)/(m - 1) \rfloor$, gives the bound in its sharpest form for a given d. At $m = 3$ this reads:

Corollary 5.2. *For every $d \geq 7$ and every $\alpha \in P_d^{(3)}$,*

$$\alpha_2 + 2\alpha_3 \ \leq\ 1 + \frac{1}{\lfloor (d - 5)/2 \rfloor}.$$

Proof. Theorem 5.1 with $m = 3$ and $K = \lfloor (d-7)/2 \rfloor$, for which $K+1 = \lfloor (d-5)/2 \rfloor$ and $6+2K \leq d-1 < d$. ☐

The bound is uniform in d but not tight: at $d = 16$ it gives $\alpha_2 + 2\alpha_3 \leq 1 + 1/5$ against the true maximum $1 + 1/18$ (Remark 7.3). Uniformity is all that is used.

Remark 5.3 (computation outside the formal development). The bound of Theorem 5.1 was also checked in the separating-functional format of Proposition 2.3 (2), by brute force in exact rational arithmetic over *all* 2^d subsets rather than only those of diameter at most m: for $m = 3$ at $d = 7, \ldots, 18$, for $m = 4$ at $d = 9, \ldots, 18$ and for $m = 5$ at $d = 11, \ldots, 18$, the functional assembled from the same rows (with a large negative z_{ij} on the pairs at distance more than m, harmless because $\Lambda_{ij} = 0$ there) satisfies $g(S) \leq 0$ for every S, with equality exactly at $S = \emptyset$, and its coefficient data $\left(\sum_i u_i, \zeta_1, \ldots, \zeta_m\right)$ is $(-(K + 2), 0, \ldots, 0, K + 1, 2K + 2)$ — the inequality $\alpha_{m-1} + 2\alpha_m \leq 1 + 1/(K + 1)$ cleared of denominators. This is a check on the certificate, not part of the proof, which is the three-case argument above.

6. Theorem C: the conjecture is false at every $m \geq 3$

Theorem 6.1. *Let $m \geq 3$ and let $d_0 \geq 1$ be arbitrary. Then, with $\alpha^{(d_0,m)}$ the parameter vector of Theorem 4.1,*

$$\alpha^{(d_0,m)} \in P_{d_0}^{(m)} \qquad \text{and} \qquad \alpha^{(d_0,m)} \notin P_{3d_0+4m}^{(m)}.$$

Consequently no d_0 is a stabilisation dimension in the sense of Definition 1.1: the conjecture of [7] *is false for every $m \geq 3$.*

Proof. Membership is Theorem 4.1. For the exclusion write $N_0 = 3d_0 + 2m - 1$ and take $K = \lfloor N_0/(m-1) \rfloor$. Then $K(m - 1) \leq N_0$, so

$$2m + K(m - 1) \ \leq\ 2m + 3d_0 + 2m - 1 \ =\ 3d_0 + 4m - 1 \ <\ 3d_0 + 4m,$$

which is the fit hypothesis of Theorem 5.1 at dimension $3d_0 + 4m$. On the other hand $N_0 < (K+1)(m-1)$, because $N_0 = (m - 1)K + (N_0 \bmod (m - 1))$ and $N_0 \bmod (m - 1) < m - 1$; hence

$$1 + \frac{1}{K + 1} \ <\ 1 + \frac{m - 1}{N_0} \ =\ \alpha_{m-1}^{(d_0,m)} + 2\,\alpha_m^{(d_0,m)}$$

by Theorem 4.1. If $\alpha^{(d_0,m)}$ were in $P^{(m)}_{3d_0+4m}$, Theorem 5.1 would bound the right-hand side by the left-hand side. For the last sentence, $P^{(m)}_{3d_0+4m} \neq P^{(m)}_{d_0}$ and $3d_0 + 4m \geq d_0$, so Definition 1.1 fails at d_0. □

The proof does not use Lemma 2.5: Definition 1.1 is an equality of sets, and one parameter vector in the symmetric difference refutes it. With Lemma 2.5 the statement upgrades to $P^{(m)}_{3d_0+4m} \subsetneq P^{(m)}_{d_0}$, a strict inclusion; that upgrade is the one step here that is not formally verified, and nothing else depends on it.

Corollary 6.2. *The stabilisation conjecture of* [7] *holds at* $m = 2$*, with* $d_0(2) = 6$*, and fails at every* $m \geq 3$*.*

Proof. The positive half is Proposition 4.2 of [7], quoted in Section 1; the negative half is Theorem 6.1. (At $m = 1$ the same paper's Remark 2.2 gives $P^{(1)}_d = \{0 \leq \alpha_1 \leq \frac{1}{2}\}$ for every $d \geq 3$, so $d_0(1) = 3$; we reproduce it for $3 \leq d \leq 9$ in Remark 3.1.) □

Remark 6.3 (proved by hand; outside the formal development). Let $h^{(m)}_d = \max\{\alpha_{m-1} + 2\alpha_m : \alpha \in P^{(m)}_d\}$. For $m \geq 3$ and $d > 2m + 1$, Theorems 4.1 and 5.1 give

$$\frac{m-1}{3d+2m-1} \;\leq\; h^{(m)}_d - 1 \;<\; \frac{m-1}{d-2m-1},$$

the upper bound by taking $K = \lfloor (d - 2m - 1)/(m - 1) \rfloor$ in Theorem 5.1. Hence $h^{(m)}_d - 1 = \Theta_m(1/d)$: the polytopes approach their limit at rate exactly $1/d$ in the direction $(0, \ldots, 0, 1, 2)$ and never reach it. Both bounds are formally verified; what is not is the existence of the maximum $h^{(m)}_d$ as an object, which is why this is a remark. The source itself notes an $O(d^{-1})$ phenomenon for its equi-correlation example in the same paragraph as the conjecture.

The next proposition is the set of controls on the pair of Theorems 4.1 and 5.1: each says that some hypothesis cannot be dropped, or that the bound is not vacuous.

Proposition 6.4.
(1) (The hypothesis $m \geq 3$ is load-bearing.) *At* $m = 2$ *and* $d = 6$ *the two windows of Section 4 coincide, and the first off-diagonal equation of the resulting certificate fails: the pair coverage at* $(0, 1)$ *is* $2/3$*, whereas the formula of Theorem 4.1 claims* $2/7$*.*
(2) (The fit hypothesis of Theorem 5.1 is load-bearing.) *At* $d = 4$ *the point* $(4, 7, 6)/17$ *lies in* $P^{(3)}_4$ *and has* $\alpha_2 + 2\alpha_3 = 19/17 > 10/9 = 1 + 1/9$*; so Theorem 5.1 with* $K = 8$ *would be false at* $d = 4$*, and it is the hypothesis* $2 \cdot 3 + 8 \cdot 2 < 4$ *that excludes it.*
(3) (The bound has teeth.) $\alpha = (0, 1, 1)$*, for which* $\alpha_2 + 2\alpha_3 = 3$*, is not in* $P^{(3)}_9$ *— by Theorem 5.1 alone, with* $K = 1$ *and bound* $3/2$*.*
(4) (The two theorems are consistent where they overlap.) *At* $d = 24$ *the point of Theorem 4.1 has* $\alpha_2 + 2\alpha_3 = 1 + 2/77$ *and Theorem 5.1 with* $K = 8$ *gives the bound* $1 + 1/9$*; the inequality* $1 + 2/77 \leq 1 + 1/9$ *is proved independently of both.*
(5) (The construction is the computed one.) $\alpha^{(4,3)} = (4, 7, 6)/17$ *and* $\alpha^{(8,3)} = (8, 11, 10)/29$*.*

Item (4) is a consistency control in the strict sense: a sign error in either theorem would show up as a contradiction between two separately proved statements; item (2) is tight by one part in 170, since $19 \cdot 9 = 171 > 170 = 10 \cdot 17$.

7. The finite route: thirteen strict inclusions

Theorem 6.1 gives a drop between d_0 and $3d_0+4m$. It does not say that the polytopes shrink at every single step, and the evidence for that is still the finite, certificate-based computation by which the phenomenon was first found; we record it because it is a strictly different statement. Write $P_d = P^{(3)}_d$. The witnesses are the primal optima of $\max\{\alpha_2 + 2\alpha_3\}$ at dimension d; for the single step $d = 5$ that objective does not separate and the direction $-\alpha_1 + 2\alpha_2 + 2\alpha_3$ is used instead, where the maximum drops from $9/4$ to 2.

Theorem 7.1. *For every* d *with* $4 \leq d \leq 16$*, the vector* $\alpha^{(d)}$ *of Table 1 satisfies*

$$\Lambda_d(\alpha^{(d)}) \in \mathcal{T}_d \qquad \textit{and} \qquad \Lambda_{d+1}(\alpha^{(d)}) \notin \mathcal{T}_{d+1},$$

that is, $\alpha^{(d)} \in P_d \setminus P_{d+1}$*. With Lemma 2.5,* $P^{(3)}_{d+1} \subsetneq P^{(3)}_d$ *at all thirteen consecutive steps.*

d	$\alpha^{(d)}$	#wts	d	$\alpha^{(d)}$	#wts
4	(1, 1, 1)	1	11	(7/22, 4/11, 4/11)	32
5	(1/4, 1/2, 3/4)	8	12	(7/24, 1/3, 3/8)	34
6	(0, 0, 1)	3	13	(9/28, 5/14, 5/14)	40
7	(1/3, 1/3, 1/2)	14	14	(3/10, 1/3, 11/30)	42
8	(1/4, 1/3, 5/12)	19	15	(11/34, 6/17, 6/17)	48
9	(5/16, 3/8, 3/8)	24	16	(11/36, 1/3, 13/36)	49
10	(5/18, 1/3, 7/18)	26			

Table 1. The thirteen witnesses of Theorem 7.1; "#wts" counts the subsets of positive weight in the membership certificate at dimension d.

Proof sketch. Each of the twenty-six statements is a finite certificate check in the sense of Proposition 2.3, and all twenty-six certificates are exhibited explicitly.

Membership at d. The certificate is the list of subsets carrying positive weight, given as bitmasks, with their rational weights; the count is in Table 1. Verifying it is d coverage equations and $\binom{d}{2}$ pair equations, each a sum of at most 49 rationals — *not* a search over the 2^d subsets.

Non-membership at $d+1$. The certificate is an integer pair (u, z) with $f(\Lambda_{d+1}(\alpha^{(d)})) > 0$, and the hypothesis $g(S) \le 0$ is verified **by exhaustive enumeration of all 2^{d+1} subsets** — $2^{17} = 131072$ cases at the last step. This is the one exhaustive computation in the theorem, and it is performed by machine inside the formal development; see Section 9. The value $f(\Lambda_{d+1}(\alpha^{(d)}))$ itself is not computed by enumeration but from Proposition 2.3 (3), from the three diagonal sums and $\sum_i u_i$.

How the certificates were found (outside the formal development, and not needed for the proof). An exact rational revised simplex over the subsets of diameter at most 3 maximises the chosen direction over P_{d+1} and returns both the primal optimum at dimension d and an optimal dual at dimension $d+1$. Two normalisations make the dual usable. The optimal dual satisfies $\zeta_k \ge c_k$ rather than $\zeta_k = c_k$; subtracting the excess from a single entry only decreases g, so one may assume equality, and then $f(\Lambda_{d+1}(\alpha)) = -h_{d+1} + c\cdot\alpha$ identically in α. And the dual constrains g only on the subsets of diameter at most 3; the remaining pairs are given a large negative z_{ij}, harmless because $\Lambda_{ij} = 0$ there, which lets the verification quantify over *all* subsets and dispense with a support lemma. (Theorem 5.1 does the opposite and proves the support lemma instead; that is why it is uniform in d and this is not.) Every certificate was re-verified by an independent exact-rational program before being entered into the formal development. □

Proposition 7.2. *No $d_0 \le 16$ is a stabilisation dimension for the 3-dependence constraint polytopes.*

Proof. Suppose $d_0 \le 16$ satisfies $P_d = P_{d_0}$ for all $d \ge d_0$ and put $e = \max(d_0, 4) \le 16$. Then $P_e = P_{d_0} = P_{e+1}$, contradicting the witness $\alpha^{(e)}$ of Theorem 7.1. Only the two memberships are used, not Lemma 2.5. □

Proposition 7.2 was the first statement of this kind we could prove, and by itself says only that $d_0(3) \ge 17$ if $d_0(3)$ exists; Theorem 6.1 supersedes it. What Theorem 7.1 still adds is the *consecutive* steps.

Remark 7.3 (computation outside the formal development). Exact rational linear programming, one solve per dimension, gives

$$h_d^{(3)} \;=\; 1 + \frac{1}{\lceil 3(d-4)/2 \rceil} \qquad \text{for every } d \text{ with } 8 \le d \le 32,$$

the closed form being compared against the computed value at each d, with no mismatch. Explicitly $h_4^{(3)} = 3$, $h_5^{(3)} = h_6^{(3)} = 2$, then $\frac{4}{3}, \frac{7}{6}, \frac{9}{8}, \frac{10}{9}, \frac{12}{11}, \frac{13}{12}, \frac{15}{14}$ at $d = 7, \dots, 13$, down to $\frac{43}{42}$ at $d = 32$. The value is strictly decreasing at every step from $d = 6$ to $d = 32$, and also from $d = 4$ to $d = 5$, so $P_{d+1}^{(3)} \subsetneq P_d^{(3)}$ at every one of those steps. Whether the exact value $1 + 1/\lceil 3(d-4)/2 \rceil$ persists for all d is open. Against it, the excess of Theorem 4.1 at $m = 3$ is $2/(3d+5)$ and the bound of Corollary 5.2 is $1/\lfloor (d-5)/2 \rfloor$, while the true value above is $1/\lceil 3(d-4)/2 \rceil$: as $d \to \infty$ these are $\sim \frac{2}{3} \cdot \frac{1}{d}$, $\sim 2 \cdot \frac{1}{d}$ and $\sim \frac{2}{3} \cdot \frac{1}{d}$, so the construction has the right leading constant $2/3$ and the bound is loose by a factor tending to 3.

8. The constraint polytope at $d = 8$

The source computes the m-dependence constraint polytope only for $m \le 2$. We record the $m = 3$, $d = 8$ region, as far as we know the first full such polytope determined with $m \ge 3$. Write $(\alpha_1, \alpha_2, \alpha_3) = (p, q, r)$.

Theorem 8.1. *Every $(p, q, r) \in P_8^{(3)}$ satisfies all fifteen inequalities*

$$\begin{array}{lll} p \ge 0, & q \ge 0, & r \ge 0, \\ q + 4r \le 2, & -2p + 2q + 2r \le 1, & p + 4q - 3r \le 2, \\ 2q \le 1, & p + 2q + 5r \le 3, & -p + 3q + 3r \le 2, \\ 3q + 2r \le 2, & p + 4r \le 2, & 6p - 5q + 4r \le 3, \\ 2p - 3q + 2r \le 1, & 3p - q + 5r \le 3, & 2p - q \le 1, \end{array}$$

and the twelve points

$$(0,0,0), \quad (0,0,\tfrac{1}{2}), \quad (0,\tfrac{1}{2},0), \quad (\tfrac{1}{4},\tfrac{1}{3},\tfrac{5}{12}), \quad (\tfrac{1}{4},\tfrac{1}{2},\tfrac{1}{4}), \quad (\tfrac{2}{7},\tfrac{2}{7},\tfrac{3}{7}),$$
$$(\tfrac{5}{14},\tfrac{3}{7},\tfrac{5}{14}), \quad (\tfrac{2}{5},\tfrac{1}{5},\tfrac{2}{5}), \quad (\tfrac{6}{13},\tfrac{4}{13},\tfrac{5}{13}), \quad (\tfrac{1}{2},0,0), \quad (\tfrac{2}{3},\tfrac{1}{3},0), \quad (\tfrac{3}{4},\tfrac{1}{2},\tfrac{1}{4})$$

all lie in $P_8^{(3)}$.

Proof sketch. Each of the fifteen inequalities is proved for *all* rational (p, q, r) by its own separating functional: an explicit integer pair (u, z) whose diagonal sums and $\sum_i u_i$ are the coefficients of that inequality up to sign, so that Proposition 2.3 (3) evaluates $f(\Lambda_8(p, q, r))$ as an affine function of (p, q, r), positive exactly when the inequality fails, and Proposition 2.3 (2) gives the contradiction. The one exhaustive step is the hypothesis $g(S) \le 0$, verified over all $2^8 = 256$ subsets, once per inequality. The twelve memberships are weight-vector certificates with between 8 and 21 subsets each. □

Remark 8.2 (computation outside the formal development). Exact rational linear programming, by the outer-approximation loop of Remark 3.1, shows that the fifteen inequalities of Theorem 8.1 are precisely the facets of $P_8^{(3)}$ and the twelve points precisely its vertices, so $P_8^{(3)}$ *is* the region cut out above. What the formal development contains is the sandwich $\mathrm{conv}(V) \subseteq P_8^{(3)} \subseteq \{\text{fifteen inequalities}\}$, whose two ends the computation shows to coincide; closing the loop internally needs the H-to-V direction, a triangulation argument in $\mathbb{R}^3$ that was not attempted.

Remark 8.3 (computation outside the formal development). The same computation gives the exact H- and V-descriptions of $P_d^{(3)}$ for $d = 4, \ldots, 9$, with

$$\#\text{facets} = 7,\ 9,\ 10,\ 12,\ 15,\ 16 \qquad \text{and} \qquad \#\text{vertices} = 7,\ 9,\ 8,\ 11,\ 12,\ 12.$$

Total cost 26 minutes of CPU, of which $d = 9$ alone is 857 seconds — this is where the whole-region computation stops being cheap. The facet count is still growing at $d = 9$, whereas at $m = 2$ it has settled at four by $d = 6$; by Theorem 6.1 the polytopes themselves never settle, whatever their facet counts do.

9. Verification

Every numbered statement above that is not labelled "proved by hand" or "computation outside the formal development" has been formally verified in Lean 4 against *Mathlib* [5, 6]; repository reference to be added at submission. The development is seven modules and about 2 300 lines; `#print axioms` was run on every declaration a statement above is pinned to, and `sorryAx` occurs nowhere.

The three theorems of this paper are in three of those modules, 803 lines in total: the ramp (Theorem 4.1), the dual combination (Theorem 5.1), and the refutation with its controls (Theorem 6.1, Corollary 5.2, Proposition 6.4). **Each of the eighteen named declarations that carry those statements returns exactly the axiom set** {`propext`, `Classical.choice`, `Quot.sound`}, and a sweep of *every* declaration those three modules add to the environment — fifty-nine of them, by `Lean.collectAxioms` over the compiled artefacts — finds no other axiom anywhere in them: no step of the refutation is discharged by compiled evaluation, none of it is a per-dimension certificate, and all of it is uniform in m and d. For their mathematics the three modules rest only on the definitions module, which is kernel-clean in the same sense; the third of them also imports the module of Section 7, but purely to borrow the statement of stabilisation at $m = 3$, and the sweep just described is what confirms that none of that module's compiled-evaluation certificates reaches any proof here.

The support reduction of Lemma 2.4 is what makes that possible. A certificate for Theorem 5.1 in the style of Section 7 would need a large negative entry on every pair at distance more than m, so that the hypothesis of Proposition 2.3 (2) could be checked over all 2^d subsets with no support hypothesis; proving the support reduction instead removes both those entries and the case analysis they force.

The older, certificate-based results are the exception, and they are exactly the ones stated with a finite range. Each of the thirteen shrinking steps of Theorem 7.1 carries, besides the three standard axioms, five axioms generated by compiled rational arithmetic (`native_decide`): three for the membership check at dimension d and two for the separating-functional check at dimension $d+1$. Proposition 7.2 carries the 65 axioms of the thirteen steps; in Theorem 8.1 the conjunction of the fifteen inequalities carries 75, five per inequality, and each of the twelve memberships three; and the controls of Section 3 carry three or two each. These counts were read off the compiled artefacts here with `Lean.collectAxioms`, not taken from the development's own notes. A reader who does not accept compiled evaluation should read each such statement as "the certificate has been checked by an independently written exact-rational program as well", which is also true; nothing in Sections 4–6 is affected either way.

Which claims rest on exhaustive computation. Exactly one kind, and it occurs only in Sections 3, 7 and 8: the hypothesis $g(S) \leq 0$ of Proposition 2.3 (2), checked over *all* 2^n subsets at the relevant dimension n — $2^{17} = 131072$ subsets at the last shrinking step, 2^8 subsets fifteen times in Theorem 8.1, 2^6 and 2^4 in the controls. No membership claim is exhaustive: a membership certificate is checked by $n + \binom{n}{2}$ rational identities over a list of at most 49 subsets, and the certificate of Theorem 4.1 is not a list at all but a closed formula. Every statement labelled as a computation — Remarks 2.7, 3.1, 3.6, 4.2, 5.3, 7.3, 8.2 and 8.3 — is exact rational linear programming or exact rational arithmetic, not a formal proof.

Cost. All computation outside Lean is exact rational arithmetic with no floating point. The whole development is about 1.5 CPU-hours: 26 minutes for the 3-dependence regions at $d = 4, \ldots, 9$, about 35 for the support-value sweeps to $d = 32$, about 12 to generate and independently verify the finite certificates, and about 10 for the checks of the remarks above. In Lean, the three modules carrying Theorems 4.1, 5.1 and 6.1 compile in 41 seconds of wall time together, against 142 seconds of CPU for the thirteen shrinking steps, whose cost doubles with every further step: the next two were priced at about 150 and 350 seconds, which is why that route stops near $d \approx 20$ and Theorem 5.1 does not.

References

[1] P. Embrechts, M. Hofert and R. Wang, *Bernoulli and tail-dependence compatibility*, Ann. Appl. Probab. **26**(3) (2016) 1636–1658, DOI 10.1214/15-AAP1128; arXiv:1606.08212.

[2] U.-R. Fiebig, K. Strokorb and M. Schlather, *The realization problem for tail correlation functions*, Extremes **20**(1) (2017) 121–168, DOI 10.1007/s10687-016-0250-8; arXiv:1405.6876.

[3] A. Janßen, S. Neblung and S. Stoev, *Tail-dependence, exceedance sets, and metric embeddings*, Extremes **26**(4) (2023) 747–785, DOI 10.1007/s10687-023-00471-z; arXiv:2212.01044.

[4] D. Krause, M. Scherer, J. Schwinn and R. Werner, *Membership testing for Bernoulli and tail-dependence matrices*, J. Multivariate Anal. **168** (2018) 240–260, DOI 10.1016/j.jmva.2018.07.014.

[5] L. de Moura and S. Ullrich, *The Lean 4 theorem prover and programming language*, in: Automated Deduction — CADE 28, Lecture Notes in Computer Science 12699, Springer, 2021, pp. 625–635, DOI 10.1007/978-3-030-79876-5_37.

[6] The mathlib Community, *The Lean mathematical library*, in: Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs (CPP 2020), ACM, 2020, pp. 367–381, DOI 10.1145/3372885.3373824.

[7] N. D. Shyamalkumar and S. Tao, *On tail dependence matrices: the realization problem for parametric families*, Extremes **23**(2) (2020) 245–285, DOI 10.1007/s10687-019-00366-y. Cited here by the numbering of the arXiv version, arXiv:1901.02157v2 (1 August 2019), whose title reads *On Tail Dependence Matrices — The Realization Problem for Parametric Families*.

[8] S. Tao, *On complete positivity of sparse symmetric Toeplitz matrices*, Journal of Comprehensive Pure and Applied Mathematics **3**(1) (2025) 1–9, DOI 10.33790/cpam1100114.

[9] S. Tao, *A contribution to modeling tail dependence*, PhD thesis, Department of Statistics and Actuarial Science, University of Iowa, 2020, DOI 10.17077/etd.005530.