ON THE SIGNATURE OF SQUARED DISTANCE MATRICES OF METRIC MEASURE SPACES

ALEXEY KROSHNIN, TIANYU MA, AND EUGENE STEPANOV

Dedicated to the memory of Anatoly Moiseevich Vershik

Abstract. We consider the numbers of positive and negative eigenvalues of matrices of squared distances between randomly sampled i.i.d. points in a given metric measure space. These numbers and their limits, as the number of points grows, contain some important information about the whole space. We show that these limits exist almost surely, are nonrandom and independent on the measure once the latter has full support. We also relate them to the signature of the pseudo-Euclidean space in which the given metric space can be isometrically embedded. In addition, we provide several examples of explicit calculations or just estimates of those limits for sample metric spaces, in particular, for the universal countable Rado–Erdős–Rényi graph. For a large class of countable spaces (for instance, containing all graphs with bounded intrinsic metrics), we get that the number of negative eigenvalues increases to infinity as the size of samples grows. However, we are able to provide examples when the number of samples grows to infinity and the numbers of both negative and positive eigenvalues increases to infinity, or the number of positive eigenvalues is bounded (but as large as desired), and the number of positive ones is fixed.

1. Introduction

Let (X, d) be a metric space (i.e., X a nonempty set equipped with a distance d) and μ be a Borel probability measure on X. We further assume, without loss of generality, that (X, d) is separable. The so-called *learning* problem for the metric measure space (X, d, μ) is to recover the information on the triple (X, d, μ) from the information on distances between points of an appropriately chosen subset of X. In view of the famous result of M. Gromov (so-called "*mm-reconstruction theorem*" [8, section $3.\frac{1}{2}.5$]) further reproved and generalized by A. Vershik in [14] and [15], this problem is completely solved. A Polish (i.e., separable and homeomorphic to complete) space (X, d, μ) can be reconstructed from distances between points obtained from the following randomized procedure: for every $m \in \mathbb{N}$, one chooses i.i.d. random points $\xi_i \in X$ drawn from the distribution μ, $i = 1, \dots, m$, thus forming a finite random sequence of points $X_m := (\xi_i)_{i=1}^m$, and calculates the $m \times m$ random distance matrix $(d(\xi_i, \xi_j))_{i,j=1}^m$. The joint distribution law of all these matrices (i.e., for all $m \in \mathbb{N}$) then determines (X, d, μ) uniquely up to a measure-preserving isometry. To reconstruct the space (X, d, μ), this approach requires the knowledge of the statistics of random distance matrices of all sizes, which is quite difficult to accomplish in practice. The natural question is to what extent

E.S. acknowledges the MIUR Excellence Department Project awarded to the Department of Mathematics, University of Pisa, CUP I57G22000700001. The work is also partially within the framework of HSE University Basic Research Program.

one may limit himself to the statistics of something less than the whole set of distance matrices, e.g., to the spectra of the latter, or even to some function of the spectra (e.g., the numbers of positive and negative eigenvalues). In other words, what kind of information on (X, d, μ) is encoded by the spectra of random distance matrices, or even by some functions of the spectra. Several attempts to answer these questions have been made, even numerically in [3], and for some functions of Euclidean distances in [4, 17, 18, 10]. However, they still remain very far from even being partially understood.

In this paper, we mainly concentrate on information on signatures. More precisely, we focus on the numbers of positive and negative eigenvalues of random *squared* distance matrices $(d^2(\xi_i, \xi_j))_{i,j=1}^m$, as $m \to \infty$. The methods in the above-cited Gromov–Vershik result on distance matrices can be applied to our research on squared distances at little cost: the joint distribution law of squared distance (or any monotone function of a distance) matrices still determines (X, d, μ) uniquely up to a measure-preserving isometry. However, working with d^2 instead of d looks much more natural precisely in terms of the relationship between the spectra of the respective matrices and the properties of these metric measure spaces. For instance, it follows from the Schoenberg theorem [1, theorem 3.1] that all matrices $(d^2(x_i, x_j))_{i,j=1}^m$ with $x_i \in X$, $m \in \mathbb{N}$ are conditionally negative semidefinite (i.e., negative semidefinite on the hyperplane $x_1 + \ldots + x_m = 0$) if and only if (X, d) is isometrically embeddable in a Hilbert space. In this case, when X is finite, an embedding can be explicitly reconstructed using the *multidimensional scaling* (MDS) algorithm well known in data science [16]. Thus, we prefer to state our results in terms of the signatures of the matrices

$$S(X_m) := -\frac{1}{2}\left(d^2(\xi_i, \xi_j)\right)_{i,j=1}^m.$$

The purely "aesthetic" coefficient $-1/2$ here is only to reflect the fact that exactly these matrices are involved directly in the MDS algorithm.

We will show that

(i) the random numbers $s_+(S(X_m))$ of positive and $s_-(S(X_m))$ of negative random eigenvalues of matrices $S(X_m)$, almost surely converge to some fixed numbers s_+ and s_- respectively, and provide several ways to compute them,

(ii) in particular, under a rather mild condition on (X, d, μ), the numbers s_+ and s_- count the numbers of positive and negative eigenvalues of the operator $K\colon L^2(X, \mu) \to L^2(X, \mu)$ defined by the formula

$$(Ku)(x) := -\frac{1}{2}\int_X d^2(x, y)u(u)\, d\mu(y).$$

Therefore, we write in this case $s_\pm = s_\pm(K)$. This is true when μ has finite 4-th order moment, i.e.,

$$\int_X d^4(x_0, y)\, d\mu(y) < +\infty$$

for some (and hence for all) $x_0 \in X$, in particular, when (X, d) is bounded. Under this condition, the spectra of $S(X_m)$ converge almost surely as $m \to \infty$ in many reasonable senses to the spectrum of K [11]. The operator K is involved in the construction of MDS for general metric measure spaces

(X, d, μ) as the limit of MDS for finite samples from X when the number of samples grows.

We will further show that

(iii) s_+ and s_- do not depend on μ if the latter has full support, i.e., $\operatorname{supp}\mu = X$. This seems to be quite counterintuitive in view of (ii), since in most cases $s_\pm = s_\pm(K)$, while K clearly depends on μ.

(iv) The number s_+ may be viewed as a natural lower bound for the dimension of the MDS embedding of (X, d, μ). More generally, the couple of numbers (s_-, s_+) give the lower bounds for the signature of a pseudo-Euclidean space in which (X, d) can be isometrically embedded (the respective embedding can be constructed explicitly). These bounds are in fact almost tight in the sense that (X, d, μ) cannot be isometrically embedded in a pseudo-Euclidean space with signature (σ_-, σ_+) with $\sigma_\pm < s_\pm - 1$. As an easy corollary, we get that every finite metric space can be isometrically embedded in a pseudo-Euclidean one;

(v) one has that $s_+ = +\infty$ when (X, d) is bounded, X is countable and d is separated from zero, e.g., for countable graphs equipped with bounded intrinsic distance. As a corollary, the embedding of (X, d, μ) produced by MDS is not finite-dimensional.

We also

(vi) provide several examples of calculations of $s_\pm$ for particular metric spaces. In many interesting cases of infinite metric spaces, we get $s_+ = s_- = +\infty$.

(vii) However, we provide examples of metric spaces with arbitrarily large finite s_-. It seems easier to provide examples when both s_+ and s_- are large. This corresponds well to the observation of [6] about so-called hollow symmetric non-negative (HSN) matrices (symmetric matrices with nonnegative entries and with zeros on the diagonal) that they "normally" have bigger number of negative rather than positive eigenvalues (recall that squared distance matrices are a particular case of HSN matrices).

(viii) Nevertheless, we are able to show a construction of a finite metric space with s_- as large as desired and s_+ fixed (up to an error of at most 1). In particular, this provides a construction of HSN matrices (even stronger, squared distance matrices) with arbitrarily large number of positive eigenvalues which is alternative to a seemingly much heavier construction of [6].

(ix) We also provide some natural estimates on $s_\pm$.

Finally,

(viii) in the quite natural example of the Rado–Erdős–Rényi graph (which is a universal graph containing all the finite graphs as its proper subgraphs), one has $s_+ = s_- = \infty$;

(ix) however, the limits of the ratios $s_+(S(X_m))/s_-(S(X_m))$ as $m \to \infty$ may depend on μ as illustrated by the example of the Rado–Erdős–Rényi graph.

2. Notation and preliminaries

2.1. **Notation.** The metric measure space (X, d, μ) is always assumed to be separable with μ a Borel probability measure.

For a set D, we denote by $\mathbf{1}_D$ its characteristic function. If D is a subset of a metric space, we let $\bar{D}$ stand for its closure. For a compact linear operator K over

some normed space, we denote by $s_\pm(K)$ the numbers of its positive and negative eigenvalues, respectively. In this paper, unless otherwise mentioned explicitly, the numbers of eigenvalues are always counted with multiplicity. We write $\operatorname{supp}\mu$ for the support of a measure μ.

For a Hilbert space H we denote by $(\cdot,\cdot)$ and $\|\cdot\|$ the respective scalar product and the Hilbert norm. The same notation will stand for the (indefinite) scalar product and the norm (sometimes called "interval" in physics) in a pseudo-Euclidean space. For a linear operator A between two Banach spaces, we denote by $\operatorname{Im} A$ its range, by $\ker A$ its kernel (i.e. the zero space) and by A^* its adjoint. If L is a subset of a linear space, then $\operatorname{span} L$ stands for its linear span.

By $L^p(X,\mu)$ we denote the usual Lebesgue space of integrable (with respect to μ) with power $p \geq 1$ (or μ-essentially bounded when $p = +\infty$) functions on X, the canonical norms in these spaces being denoted by $\|\cdot\|_p$. Denote by $\mathbf{1}^\perp$ the orthogonal complement of constant function $\mathbf{1}$ in $L^2(X,\mu)$, and by P_μ the orthogonal projector from $L^2(X,\mu)$ to $\mathbf{1}^\perp$. Let

$$(K_\mu)(x) := -\frac{1}{2}\int_X d^2(x,y)u(y)d\mu(y) \tag{2.1}$$

$$T_\mu := (P_\mu K_\mu P_\mu)(u), \tag{2.2}$$

be linear operators defined on $L^2(X,\mu)$. We usually avoid the subscript μ and write the above introduced operators as P, K, T instead of P_μ, K_μ, T_μ respectively, unless we would like to emphasize their dependence on the measure μ.

If the 4-th moment of μ is infinite, then the domains of K and T are not all of $L^2(X,\mu)$. In the opposite case, both K and T are self-adjoint Hilbert-Schmidt operators over the whole $L^2(X,\mu)$. Throughout this paper, we call the operator K and T the MDS defining operators associated with the metric measure space (X,d,μ). The infinite MDS maps of (X,d,μ) are constructed from the positive eigenvalues and their corresponding $L^2(X,\mu)$-normalized eigenfunctions of T. Note that

$$\begin{aligned}(Tf)(x) &= \int_X k_T(x,y)f(y)\,d\mu(y), \quad \text{where}\\ k_T(x,y) &:= k(x,y) - \int_X k(x,y')\,d\mu(y') - \int_X k(x',y)\,d\mu(x')\\ &\quad + \int_X\int_X k(x',y')\,d\mu(x')\,d\mu(y').\end{aligned} \tag{2.3}$$

where for brevity we denoted $k(x,y) := -d^2(x,y)/2$. In particular, (2.3) implies

$$k_T(x,x) + k_T(y,y) - 2k_T(x,y) = d^2(x,y). \tag{2.4}$$

For the general metric measure space (X,d,μ) and $X_N := \{x_i\}_{i=1}^N \subset X$ a finite *subset*, we denote $S(X_N) := \left(-\frac{1}{2}d^2(x_i,x_j)\right)_{i,j=1}^N$. Note that for a finite *sequence* of i.i.d. random elements ξ_i of X, we write $X_m := (\xi_i)_{i=1}^m \subset X$ in lower-case index for the finite random sequence and $S(X_m) := \left(-\frac{1}{2}d^2(\xi_i,\xi_j)\right)_{i,j=1}^m$ for the respective random $m\times m$ matrix. It is important to note that X_m does not necessarily always have exactly m elements because some of ξ_i may occasionally coincide. Therefore,

$S(X_m)$ can be different from $S(\hat{X}_N)$, where $\hat{X}_N$ is the *set* of distinct elements of the *sequence* X_m.[1]

We denote by $\mathbf{1}_n$ the vector $\mathbf{1}_n := (1, \dots, 1)^T \subset \mathbb{R}^n$ and by Id_n the identity $n \times n$ matrix (the superscript T in this context stands for the transpose of matrices), and $\Pi_n := \mathrm{Id}_n - \frac{\mathbf{1}_n \mathbf{1}_n^T}{n} \in \mathbb{R}^{n \times n}$.

2.2. **Spectral theory for matrices.** In the sequel, we frequently use the Cauchy interlacing theorem [13, theorem 17.17]. We state it here for convenience in the form we need in the sequel.

Proposition 2.1 (Cauchy interlacing theorem)**.** *Let B be a real symmetric matrix of order N. Let $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_N$ be the eigenvalues of B counting multiplicity. For any principal submatrix B' of B of order $N-1$, the eigenvalues $\lambda'_1 \leq \lambda'_2 \leq \cdots \leq \lambda'_{N-1}$ of B' are interlacing such that*

$$\lambda_i \leq \lambda'_i \leq \lambda_{i+1} \quad \text{for all } 1 \leq i \leq N-1.$$

In particular, for any real δ, if the symmetric matrix B has a principal block admitting at least m eigenvalues greater (resp., less) than or equal to δ, so does the matrix B.

3. Gromov–Vershik scheme

For a general metric measure space (X, d, μ), we refer to the experiment consisting of choosing i.i.d. random points $\xi_i \sim \mu$ for $i = 1, \dots, m$, and calculating the random matrix $S(X_m)$ where X_m is the finite random sequence of points $X_m := (\xi_i)_{i=1}^m \subset X$, as the *Gromov–Vershik scheme*. Namely, let $(\Omega, \Sigma, \mathbb{P})$ be a probability space and $\xi_i \colon \Omega \to X$, $i \in \mathbb{N}$, be i.i.d. random elements of X such that $\mathrm{law}(\xi_i) = \mu$ for all $i \in \mathbb{N}$. We set then $X_m(\omega) := (\xi_i(\omega))_{i=1}^m$ for all $\omega \in \Omega$. Note that X_m is not necessarily a random *set*, but rather a random finite sequence of elements of X, some of which may be repeating. cancelling the repeating elements of the latter, we obtain a random set $\hat{X}_N(\omega) \subset X$ with $N = N(m, \omega) := \#X_m(\omega)$. We denote its elements by $\hat{\xi}_1(\omega), \dots, \hat{\xi}_N(\omega)$, i.e., $\hat{X}_N(\omega) := \{\hat{\xi}_i(\omega)\}_{i=1}^N$. We will then say that the random set $\hat{X}_N$ and the respective random matrix $S(\hat{X}_N)$ are obtained by *Gromov–Vershik scheme without repetitions*.

The random matrix $S(X_m)$ may be viewed as the $m \times m$ principal submatrix of the infinite random matrix $\left(-\frac{1}{2} d^2(\xi_i, \xi_j)\right)_{i,j=1}^\infty$. The distribution law of the latter is clearly the measure $\nu := f_\# \mu^\infty$ (which may be called *matrix distribution* following [14, 15]) over the space of infinite matrices $\mathbb{R}^{\mathbb{N} \times \mathbb{N}}$, where μ^∞ is the countable tensor product of μ, $f \colon X^\mathbb{N} \to \mathbb{R}^{\mathbb{N} \times \mathbb{N}}$ is defined by

$$f((x_1, x_2, \dots, x_k, \dots)) := \left(-\frac{1}{2} d^2(x_i, x_j)\right)_{i,j=1}^\infty,$$

and $f_\#$ stands for the usual pushforward of a measure by the map f.

[1] In what follows we maintain the following notation: if X_m is a finite sequence with possibly repeating elements, then $\hat{X}_N$ will stand for the set of its elements, or, equivalently, the sequence obtained from X_m by cancelling repetitions. However, if the sequence (random or deterministic) X_N does not contain repetitions, i.e., can be viewed as a set, then we do not write $\hat{X}_N$ in this case.

We now reformulate the Gromov–Vershik theorem ("*mm-reconstruction theorem*", see [14] or alternatively [15]) which in its original version uses the random distance matrix $(d(\xi_i,\xi_j))_{i,j=1}^{\infty}$ instead of $\left(-\frac{1}{2}d^2(\xi_i,\xi_j)\right)_{i,j=1}^{\infty}$.

Theorem 3.1. *If (X,d) is Polish, then the measure ν determines (X,d,μ) uniquely up to a measure-preserving isometry.*

4. Limit distance signature of separable metric spaces

We introduce the following notion crucial for this paper.

Definition 4.1. *Let (X,d) be a metric space. We call the limit distance signature of (X,d) the couple $(s_-(X,d),s_+(X,d))$ with $s_\pm(X,d))\in\mathbb{N}\cup\{+\infty\}$ defined by*

$$s_\pm(X,d)=\sup\{s_\pm(S(X_N)) : X_N\subset X \text{ finite}\}. \tag{4.1}$$

In this section we prove the following result.

Theorem 4.2. *Let (X,d) be a separable metric space. Then the following assertions hold.*

(i) *If X is at most countable, then*
$$s_\pm(X,d)=\lim_N s_\pm(S(X_N))=\sup_N s_\pm(S(X_N))$$
for an arbitrary sequence of finite subsets $X_N\nearrow X$ as $N\to\infty$. In particular, if X is finite, then $s_\pm(X,d)=s_\pm(S(X))$.

(ii) *If X is infinite, then $s_\pm(X,d)=s_\pm(\tilde{X},d)$, where $\tilde{X}\subset X$ is an arbitrary countable dense subset of X endowed with the same distance d.*

(iii) *If μ is a Borel probability measure over X with full support, i.e., with $\operatorname{supp}\mu=X$, then*
$$s_\pm(X,d)=\lim_m s_\pm(S(X_m(\omega)))=\lim_m s_\pm(S(\hat{X}_N(\omega)))\quad a.s.,$$
where X_m are random finite sequences of i.i.d. points in X chosen according to the Gromov–Vershik scheme and $\hat{X}_N$ are the respective random finite sets of points in X obtained by cancelling repeated elements in the sequence X_m, i.e., according to the Gromov–Vershik scheme without repetitions.

(iv) *If μ has the finite 4-th moment and full support in X, then*
$$s_\pm(K_\mu)=s_\pm(X,d),$$
where K_μ is the operator defined by (2.1), although the numbers $s_\pm(X,d)$ do not depend on μ. Furthermore,
$$s_\pm(K_\mu)-1\le s_\pm(T_\mu)\le s_\pm(K_\mu),$$
and if $\mathbf{1}\not\perp\ker K_\mu$, then $s_\pm(T_\mu)=s_\pm(K_\mu)$, where T_μ is the operator defined by (2.2).

(v) *$s_\pm(T_\mu)$ also does not depend on μ once the latter has finite 4-th moment and full support in X.*

Remark 4.3. Note that $s_\pm(X,d)\ge 1$ unless X is a singleton (in which case clearly $s_\pm(X,d)=0$ by (i)). In fact, for every finite $X_N\subset X$ one has $s_-(S(X_N))\ge 1$ since the Perron–Frobenius theorem guarantees the existence of a strictly negative eigenvalue of maximum absolute value among the other eigenvalues. Moreover, the trace of the above matrices is zero, so we have $s_+(S(X_N))\ge 1$.

Proof. Claim (i) is Proposition 4.4. Claim (ii) is in fact the definition of limit distance signature, the correctness of which is Proposition 4.6. Claim (iii) is Proposition 4.7. Claim (iv) follows from Proposition 4.11 and Proposition 4.14. Finally, Claim (v) is part of Proposition4.15. □

The subsections below are dedicated to the detailed proof of the above Theorem 4.2.

4.1. **Deterministic characterization of limit distance signatures.** The following purely deterministic statement is valid.

Proposition 4.4. *Let (X, d) be an at most countable metric space. Then for all sequences of finite subsets $X_N \subset X$ such that $X_N \nearrow X$ as $N \to \infty$, the limits always satisfy*

$$s_\pm(X, d) = \lim_N s_\pm(S(X_N)).$$

In particular, the above limits are independent of the sequence $(X_N)_N$.

Proof. By the Cauchy interlacing theorem, the sequences $\big(s_\pm(S(X_N))\big)_N$ are non-decreasing and hence admit a finite or infinite limit. Moreover, by the definition of $s_\pm(X, d)$, we have $s_\pm(X, d) \geq s_\pm(S(X_N))$ for every N, and hence

$$s_\pm(X, d) \geq \lim_N s_\pm(S(X_N)).$$

To show the other side of the inequality, note for any finite subset $X' \subset X$, there is an $M \in \mathbb{N}$ such that $X' \subset X_M$. Thus, by the Cauchy interlacing theorem, $s_\pm(S(X')) \leq s_\pm(S(X_M))$. Since X' is arbitrary, we also have the reverse inequality

$$s_\pm(X, d) \leq s_\pm(S(X_M)) \leq \sup_N s_\pm(S(X_N)) = \lim_N s_\pm(S(X_N)),$$

concluding the proof. □

We now consider the limit distance signature of a general separable metric spaces.

Proposition 4.5. *Let (X, d_X) be isometrically embeddable in (Y, d_Y), both metric spaces being separable. Then*

$$s_\pm(X, d_X) \leq s_\pm(Y, d_Y).$$

Proof. By embedding X isometrically into Y we may assume without loss of generality that X is a subset of Y. The assertion is then straightforward from the definition of $s_\pm$. □

Proposition 4.6. *Let (X, d) be a separable metric space. Then*

$$s_\pm(X, d) = s_\pm(\tilde{X}, d),$$

where $\tilde{X} \subset X$ is an arbitrary countable dense subset of X.

Proof. Consider an arbitrary finite set $X_N \subset X$. Since $\tilde{X}$ is dense in X, we can find a finite $\tilde{X}_N \subset \tilde{X}$ close enough to X_N so that $s_\pm(S(X_N)) \leq s_\pm(S(\tilde{X}_N))$. This implies $s_\pm(X, d) \leq s_\pm(\tilde{X}, d)$ and the reverse inequality follows immediately from the definition of $s_\pm(X, d)$ (alternatively, from Proposition 4.5) concluding the proof. □

4.2. **Random characterization of limit distance signature.** Consider now

- the random finite sequences X_m and the respective random matrices $S(X_m)$ obtained by the Gromov–Vershik scheme,
- as well as the random finite sets $\hat{X}_N \subset X$ with $N \leq m$ random which is obtained from X_m by cancelling repeating elements, and respective random matrices $S(\hat{X}_N)$ (in other words, the random set $\hat{X}_N$ and the matrix $S(\hat{X}_N)$ obtained by Gromov–Vershik scheme without repetitions).

The following assertion is valid.

Proposition 4.7. *Let (X, d) be a separable metric space. If μ has full support, one has*

$$s_\pm(X, d) = \lim_m s_\pm(S(\hat{X}_{N(m,\omega)}(\omega))) = \lim_m s_\pm(S(X_m(\omega)))$$

for $\mathbb{P}$-a.e. $\omega \in \Omega$.

Proof. Since by Lemma 4.8, one has

$$\overline{\bigcup_m \hat{X}_{N(m,\omega)}(\omega)} = \operatorname{supp}\mu = X$$

for $\mathbb{P}$-a.e. $\omega \in \Omega$, i.e., $\bigcup_m \hat{X}_N$ is a countable dense set in X almost surely. Therefore, we have

$$s_\pm(X, d) = \lim_m s_\pm(S(\hat{X}_{N(m,\omega)}(\omega)))$$

$\mathbb{P}$-a.s. in view of Proposition 4.6. By Lemma 4.9 below, we also get

$$s_\pm(S(X_m(\omega))) = s_\pm(S(\hat{X}_{N(m,\omega)}(\omega)))$$

for all $\omega \in \Omega$, concluding the proof. □

The following lemma has been used in the proof of the above Proposition 4.7.

Lemma 4.8. *Let (X, d) be a separable metric space, and μ be a Borel probability measure over X. Then*

$$\overline{\bigcup_m \hat{X}_{N(m,\omega)}(\omega)} = \operatorname{supp}\mu$$

for $\mathbb{P}$-a.e. $\omega \in \Omega$, i.e., $\bigcup_m \hat{X}_N$ is a countable dense set in $\operatorname{supp}\mu$. In particular, if $\operatorname{supp}\mu$ is infinite, then $\lim_m N(m, \omega) = +\infty$ for $\mathbb{P}$-a.e. $\omega \in \Omega$ and if additionally $\operatorname{supp}\mu$ is countable, then also $\hat{X}_N \nearrow \operatorname{supp}\mu$ a.s. as $m \to \infty$.

Proof. To prove the first claim, consider an arbitrary open set $U \subset X$ satisfying

$$U \bigcap \overline{\bigcup_m \hat{X}_{N(m,\omega)}(\omega)} = \emptyset,$$

hence $\xi_i(\omega) \in X \setminus U$ for all $i \in \mathbb{N}$. The probability of this event does not exceed $\mu(X \setminus U)^n$ for all $n \in \mathbb{N}$ and hence is zero unless $\mu(U) = 0$, i.e. $U \cap \operatorname{supp}\mu = \emptyset$, showing that

$$\operatorname{supp}\mu \subset \overline{\bigcup_m \hat{X}_{N(m,\omega)}(\omega)} = \emptyset.$$

On the other hand, $\mathbb{P}(\{\xi_i \in \operatorname{supp}\mu\}) = \mu(\operatorname{supp}\mu) = 1$, and hence

$$\operatorname{supp}\mu \supset \overline{\bigcup_m \hat{X}_{N(m,\omega)}(\omega)}$$

a.s., showing the claim.

To prove the second claim, denote for every $k \in \mathbb{N}$ the set

$$\Omega_k := \{\omega \in \Omega : \lim_m N(m,\omega) \leq k\}$$

Observe that for all $\omega \in \Omega_k$, one has

$$\#\bigcup_m \hat{X}_{N(m,\omega)}(\omega) \leq k.$$

Hence, $\overline{\bigcup_m \hat{X}_{N(m,\omega)}(\omega)} \neq \operatorname{supp}\mu$ when $\operatorname{supp}\mu$ is infinite. The first claim just proven implies $\mathbb{P}(\Omega_k) = 0$ in this case. Thus,

$$\mathbb{P}\left(\left\{\lim_m N(m,\omega) < +\infty\right\}\right) = \mathbb{P}\Big(\Omega \setminus \bigcup_k \Omega_k\Big) = 1$$

as claimed. □

Lemma 4.9. *Let the symmetric matrix $\hat{S} \in \mathbb{R}^{p\times p}$ be obtained from the symmetric matrix $S \in \mathbb{R}^{q\times q}$ cancelling all the repeating rows and all the repeating columns. Then $s_\pm(S) = s_\pm(\hat{S})$.*

Proof. For each k and j, interchanging the k-th with j-th row and the k-th with j-th column simultaneously does not change the signature of a matrix, so the proof reduces just to an inductive application of Lemma 4.10 below. □

Lemma 4.10. *Let the symmetric matrix $S \in \mathbb{R}^{(p+1)\times(p+1)}$ be of the block form*

$$S = \begin{bmatrix} \hat{S} & v \\ v^t & a \end{bmatrix},$$

where $\hat{S} \in \mathbb{R}^{p\times p}$ is a symmetric matrix, $v \in \mathbb{R}^p$ and $a \in \mathbb{R}$ be such that the p-th and $(p+1)$-th row as well as the p-th and $(p+1)$-th column of S are identical. Then $s_\pm(S) = s_\pm(\hat{S})$.

Proof. Clearly, we have (e.g., by the Cauchy interlace theorem) that $s_\pm(S) \geq s_\pm(\hat{S})$. On the other hand, since the p-th and $(p+1)$-th row as well as the p-th and $(p+1)$-th column of S are identical, we have that $s_0(S) = s_0(\hat{S}) + 1$, where $s_0(S)$ stands for the number of zero eigenvalues of S (i.e. the dimension of the kernel of the latter). Since also

$$s_-(S) + s_0(S) + s_+(S) = p+1, \qquad s_-(\hat{S}) + s_0(\hat{S}) + s_+(\hat{S}) = p,$$

we get

$$\begin{aligned} s_-(S) + s_0(S) + s_+(S) &= (s_-(\hat{S}) + s_0(\hat{S}) + s_+(\hat{S})) + 1 = \\ &= s_-(\hat{S}) + s_0(S) + s_+(\hat{S}), \end{aligned}$$

which implies $s_-(S) + s_+(S) = s_-(\hat{S}) + s_+(\hat{S})$, hence the claim. □

4.3. **Limit distance signature of the space and signatures of MDS defining operators.** The following result relates the limit distance signature of (X,d,μ) to the numbers of negative and positive eigenvalues of any MDS defining operator.

Proposition 4.11. *Let (X,d) be a separable metric space. If μ has finite 4-th moment and full support in X, one has*

$$s_\pm(K_\mu) = s_\pm(X,d),$$

where K_μ is defined by (2.1).

Remark 4.12. When μ has finite 4-th moment but does not have full support in X, we only can assert that

$$s_\pm(K_\mu) \le s_\pm(X,d). \tag{4.2}$$

In fact, by the above Proposition 4.11 one has

$$s_\pm(K_\mu) = s_\pm(\operatorname{supp}\mu, d),$$

and $s_\pm(\operatorname{supp}\mu, d) \le s_\pm(X,d)$ by Proposition 4.5. The inequality in (4.2) can be sharp as easily seen from the example, say, of X just a Euclidean plane (i.e., $X=\mathbb{R}^2$ equipped with the Euclidean distance), so that $s_+(X,d)=2$. and μ having a line as a support, so that $s_+(K_\mu) = s_+(\operatorname{supp}\mu,d) = s_+(\mathbb{R},d) = 1$, see Example 5.5.

Proof. Without loss of generality, we may assume (X,d) to be complete (if not, we may extend d and μ to the completion of (X,d) and work in the latter completion). Then by the estimate (A.4) from Lemma A.3(ii) (applied with $A := K_\mu$, $a := -d^2/2$), one has for some sequence of finite subsets $X_N \subset X$, the following chain of inequalities

$$s_+(K_\mu) \le \liminf_N s_+(S(X_N)) \le s_+(X,d). \tag{4.3}$$

On the other hand,

$$s_+(K_\mu) \ge s_+(S(X_N))$$

for every finite $X_N \subset X$ by estimate (A.2) in Lemma A.3(i), applied with $A := K_\mu$, $a := -d^2/2$. Hence

$$s_+(K_\mu) \ge s_+(X,d).$$

This inequality together with (4.3) gives the claim

$$s_+(K_\mu) = s_+(X,d).$$

The analogous claim for s_- is completely symmetric. □

Remark 4.13. An alternative proof of (4.3) may be obtained by observing that for $\mathbb{P}$-a.e. $\omega\in\Omega$, one has

$$\lim_m \inf\left\{\left|\pi\left(\sigma\left(\frac{1}{m}S(X_m(\omega))\right)\right) - \sigma(K_\mu)\right| : \pi\in S_\infty\right\} = 0 \tag{4.4}$$

using theorem 3.1 from [11]. Here $\sigma(K_\mu)$ stands for the spectrum of K_μ, $\sigma(\frac{1}{m}S(X_m))$ stands for the spectrum of the random matrix $\frac{1}{m}S(X_m)$, both being considered as elements of the space ℓ^2 of square summable sequences; and S_∞ stands for the set of all possibly infinite permutations of a countable set. This provides

$$\lim_m s_\pm\left(\frac{1}{m}S(X_m(\omega))\right) \ge s_\pm(K_\mu)$$

for $\mathbb{P}$-a.e. $\omega\in\Omega$. Hence (4.3) holds because

$$s_\pm\left(S(X_m(\omega))\right) = s_\pm\left(\frac{1}{m}S(X_m(\omega))\right)$$

for all $\omega\in\Omega$ and

$$\lim_m s_\pm\left(S(X_m(\omega))\right) = s_\pm(X,d)$$

for $\mathbb{P}$-a.e. $\omega\in\Omega$ by Proposition 4.7. The nice feature of this argument is that it requires neither the completion of the space nor the use of continuity of the operator kernel.

The following result on the relationship between the signatures of the operators K_μ and T_μ is also worth being mentioned.

Proposition 4.14. *Let (X, d, μ) be a metric measure space with a Borel probability measure μ having finite 4-th moment and full support in X, and the operators $K = K_\mu$ and $T = T_\mu$ are defined by (2.1) and (2.2), respectively. Then*

$$s_\pm(K) - 1 \le s_\pm(T) \le s_\pm(K). \tag{4.5}$$

Moreover, if $\mathbf{1} \not\perp \ker K$, then $s_\pm(T) = s_\pm(K)$.

Proof. The estimates (4.5) are just Lemma A.4(i) and (ii) applied with $A := K$, $Q := P$, and $B := T$, $H_1 = H_2 := L^2(X, \mu)$, $k = 1$. Lemma A.4(iii) with the same notations implies also the last claim since $(\operatorname{Im} P)^\perp = \mathbf{1}^\perp$ and $\mathbf{1} \notin \ker K$ (otherwise μ is a Dirac delta measure, $T = K = 0$, and $\ker K = \{0\}$, a contradiction), and hence $\mathbf{1} \not\perp \ker K$ implies $(\operatorname{Im} P)^\perp \cap (\ker K)^\perp = \{0\}$. □

Note that if $\mathbf{1}$ is an eigenfunction of K with a nonzero eigenvalue, then $\mathbf{1} \perp \ker K$.

Finally, we provide the following characterization of $s_\pm(T_\mu)$, similar to Proposition 4.11 for $s_\pm(K_\mu)$.

Proposition 4.15. *Let (X, d) be a separable metric space. If μ has finite 4-th moment and full support in X, then one has*

$$s_\pm(T_\mu) = \sup\{s_\pm(\Pi_N S(X_N)\Pi_N) : X_N \subset X \text{ finite}, \#X_N = N\}, \tag{4.6}$$

where T_μ is defined by (2.2). In particular, $s_\pm(T_\mu)$ are the same for all μ with finite 4-th moment and full support in X.

Proof. It suffices to replicate word-to-word the proof of Proposition 4.11 replacing the use of the estimate (A.2) by (A.3) and (A.4) by (A.5). □

5. The role of s_+ and s_-

5.1. **Embedding in a Hilbert space.** We feel almost obliged to mention first the following theorem on existence of an isometric embedding into a Hilbert space, which is essentially just a reformulation of the classical Schoenberg theorem [1, theorem 3.1].

Proposition 5.1. *The following statements for a separable metric space (X, d) are equivalent.*

(i) *The space (X, d) is isometrically embeddable in a Hilbert space.*
(ii) *For some Borel probability measure μ on X with full support and finite 4-th moment, one has $s_-(T_\mu) = 0$, i.e., T_μ is positive semidefinite.*
(iii) *For every Borel probability measure μ on X, one has $s_-(T_\mu) = 0$.*

In case (ii), for an isometric embedding of X in ℓ^2, we may take the map $f\colon X \to \ell^2$ defined by

$$f(x) := \left(\sqrt{\lambda_i} u_i(x)\right)_i, \tag{5.1}$$

where $\lambda_i \in \mathbb{R}$ be strictly positive eigenvalues (counting multiplicities) and $u_i \in L^2(X, \mu)$ the respective eigenfunctions of the operator $T_\mu\colon L^2(X, \mu) \to L^2(X, \mu)$ (defined by (2.2)) normalized so that $\|u_i\|_2 = 1$ for all i.

In all these cases, one necessarily has $s_-(X, d) = 1$.

Proof. If (i) holds, we identify X with the image of the respective isometric embedding and μ with its push-forward through this embedding. We may assume thus without loss of generality that $X \subset H$, for some Hilbert space H, and μ be a measure in H. Then every finite set $X_N \subset X$ is isometrically embedded in a finite-dimensional Euclidean space (the subspace $\operatorname{span} X_N \subset H$). By Theorem 5.6[2], for the operator $T_N := T_{\mu_N}$ with μ_N an arbitrary Borel probability measure supported on X_N, we have $s_-(T_N) = 0$. Thus, by Lemma A.3(ii) with $A := K_\mu$ and H instead of X, we have

$$s_-(T_\mu) \le \liminf s_-(T_N) = 0,$$

thus proving (i)⇒(iii).

The implication (iii)⇒(ii) is trivial. We prove (ii)⇒(i). To this aim, let μ be as in (ii) and let $T := T_\mu$. Observe that by Lemma A.2 (with $A := T$, $a := k_T$), one has

$$\sum_i \lambda_i u_i(x) u_i(y) = k_T(x,y), \tag{5.2}$$

with the equality formally valid only in the sense of $L^2(X \times X, \mu \otimes \mu)$. But since T is positive definite, by Mercer's theorem, the convergence in (5.2) is uniform so that this equation holds for all $(x,y) \in X \times X$. We act now exactly as in the proof of the first part of Theorem 5.6. Namely, from (5.2), we get

$$\sum_i \lambda_i u_i^2(x) = k_T(x,x) \tag{5.3}$$

for all $x \in X$. Now, (5.3) and (5.2) imply

$$\begin{aligned}\sum_i \lambda_i \left(u_i(x) - u_i(y)\right)^2 &= \sum_i \lambda_i u_i^2(x) + \sum_i \lambda_i u_i^2(y) - 2\sum_i \lambda_i u_i(x) u_i(y)\\ &= k_T(x,x) + k_T(y,y) - 2k_T(x,y),\end{aligned}$$

which in view of (2.4) gives

$$\sum_i \lambda_i \left(u_i(x) - u_i(y)\right)^2 = d^2(x,y) \tag{5.4}$$

for all $(x,y) \in X \times X$. The latter implies for the map f defined by (5.1) the relationship holds

$$\|f(x) - f(y)\|_{\ell_2}^2 = \sum_i \lambda_i \left(u_i(x) - u_i(y)\right)^2 = d^2(x,y),$$

i.e., f is an isometry onto the image, proving (i).

The final claim of the statement follows from the estimate $s_-(X,d) \ge 1$ valid for every metric space, and from Theorem 4.2(iv) which implies the estimates

$$s_-(X,d) = s_-(K_\mu) \le s_-(T_\mu) + 1 = 1$$

for every μ Borel probability measure with full support and finite 4-th moment in X. This completes the proof. □

[2]We may hare referred here to the celebrated Schoenberg's theorem. We use a more general Theorem 5.6 instead to be self-consistent

Remark 5.2. When the metric spaces (X, d) is isometrically embeddable in a Hilbert space, then $s_-(X, d) = 1$, but the converse is false. There exist even finite metric spaces (X, d) not embeddable isometrically in a Hilbert space, which still satisfy $s_-(X, d) = 1$, as the following example shows.

Example 5.3. Let $X := \{1, 2, 3, 4\}$ with $d(i, j) := 2$ unless either $i = j$, in which case $d(i, i) = 0$ or either $i = 1, 2, 3$ and $j = 4$, or, symmetrically, $j = 1, 2, 3$ and $i = 4$. In the latter cases $d(1, 4) = d(2, 4) = d(3, 4) = 1$. Note that the points $1, 2, 3$ embed isometrically in the Euclidean plane $\mathbb{R}^2$ as vertices of some equilateral triangle of sidelength $a = 2$. However, the points $1, 2, 3, 4$ cannot embed isometrically in $\mathbb{R}^3$ (hence in any Hilbert space) since 4 should be otherwise the midpoint of each of the segments $[ij]$, $i, j = 1, \dots, 3$. Nevertheless, $s_-(X, d) = 1$ and $s_+(X, d) = 3$.

The following easy example shows the metric space isometrically embeddable in an infinite-dimensional Hilbert space.

Example 5.4. Let $X := \mathbb{N}$, with $d(i, j) := 1$ when $i \neq j$ (and, of course, $d(i, i) = 0$). Then $S(X_N)$ for every N-element subset $X_N \subset X$ has an eigenvalue $1/2$ with multiplicity $N - 1$ and one eigenvalue $-(N - 1)/2$. Then $s_+(S(X_N)) = N - 1$ and $s_-(S(X_N)) = 1$ implies $s_+(X) = +\infty$ and $s_-(X) = 1$. Note that this space can be easily isometrically embedded in a Hilbert space (e.g., in the space ℓ^2 of square summable sequences, via the map $k \mapsto e_k/\sqrt{2}$, e_k standing for the k-th coordinate vector).

Example 5.5. Let $\mathbb{R}^n$ stand for the usual Euclidean n-dimensional space. Then $s_-(\mathbb{R}^n) = 1$, and $s_+(\mathbb{R}^n) = n$, and hence for a metric space (X, d) to be isometrically embeddable in $\mathbb{R}^n$, it is necessary and sufficient that $s_-(T) = 0$ and $s_+(X, d) \leq n$ (clearly, in this case $s_+(X, d) = n$ unless (X, d) can be isometrically embedded in some $\mathbb{R}^m$ with $m < n$).

5.2. **Embedding in pseudo-Euclidean spaces.** The pseudo-Euclidean space $\mathbb{R}^{n,p}$ with signature (n, p) is the linear space of vectors $\mathbb{R}^{n+p}$ equipped with the bilinear form

$$(u, v) := -\sum_{i=1}^{n} u_i v_i + \sum_{j=n+1}^{n+p} u_j v_j.$$

The latter defines a pseudo-distance $d_{n,p}(u, v) := \sqrt{(u - v, u - v)}$ on every set $\Sigma \subset \mathbb{R}^{n,p}$ such that the Minkowski difference $\Sigma - \Sigma := \{u - v : u \in \Sigma, v \in \Sigma\}$ belongs to the positive cone

$$C_{n,p} := \left\{ v \in \mathbb{R}^{n,p} : (v, v) := -\sum_{i=1}^{n} v_i^2 + \sum_{j=n+1}^{n+p} v_j^2 \geq 0 \right\}.$$

The following statement is valid.

Theorem 5.6. *Let μ be a Borel probability measure with full support and finite 4-th moment in the metric space (X, d). Suppose that $s_-(T_\mu) = n \in \mathbb{N}$ and $s_+(T_\mu) = p \in \mathbb{N}$. Let $\lambda_i \in \mathbb{R}$ be nonzero eigenvalues (counting multiplicities) and $u_i \in L^2(X, \mu)$ be the respective eigenfunctions of the operator $T_\mu \colon L^2(X, \mu) \to L^2(X, \mu)$ (defined by (2.2)), normalized so that $\|u_i\|_2 = 1$, $i = 1, \dots, n+p$. Without loss of generality, assume $\lambda_i < 0$ for $i = 1, \dots n$ and $\lambda_i > 0$ for $i = n + 1, \dots n + p$. Then the map*

$f\colon X \to \mathbb{R}^{n,p}$ defined by

$$f(x) := \sum_{i=1}^{n+p} \sqrt{|\lambda_i|}u_i(x),$$

is an isometry onto the image such that its image $\Sigma := f(X)$ satisfies $\Sigma - \Sigma \subset C_{n,p}$. In particular, every finite metric space can be isometrically embedded into some pseudo-Euclidean space with the image Σ such that $\Sigma - \Sigma$ belongs to the positive cone.

Vice versa, if μ is a Borel probability measure with finite 4-th moment (not necessarily of full support in X) and there is an $f\colon X \to \mathbb{R}^{n,p}$ isometry onto the image such that $\Sigma := f(X)$ satisfies $\Sigma - \Sigma \subset C_{n,p}$, then $s_-(T_\mu) \le n$ and $s_+(T_\mu) \le p$.

Proof. By Lemma A.2 (with $A := T$, $a := k_T$), one has

$$\sum_{i=1}^{\nu} \lambda_i u_i(x)u_i(y) = k_T(x,y), \tag{5.5}$$

where $\nu := n + p$, with the equality in (5.5) formally valid only in the sense of $L^2(X \times X, \mu \otimes \mu)$. Since the functions on both sides of this equality are continuous, then (5.5) is also valid pointwise, i.e., for all $(x, y) \in X \times X$. In particular, we also get

$$\sum_{i=1}^{\nu} \lambda_i u_i^2(x) = k_T(x,x) \tag{5.6}$$

for all $x \in X$.

From (5.6) and (5.5), we get

$$\begin{aligned}\sum_{i=1}^{\nu} \lambda_i \left(u_i(x) - u_i(y)\right)^2 &= \sum_{i=1}^{\nu} \lambda_i u_i^2(x) + \sum_{i=1}^{\nu} \lambda_i u_i^2(y) - 2\sum_{i=1}^{\nu} \lambda_i u_i(x)u_i(y)\\ &= k_T(x,x) + k_T(y,y) - 2k_T(x,y)\end{aligned}$$

for all (x, y). The identity (2.4) gives then

$$\sum_{i=1}^{\nu} \lambda_i \left(u_i(x) - u_i(y)\right)^2 = d^2(x,y) \tag{5.7}$$

for all (x, y). We observe now that one can view f as a map between X and $\mathbb{R}^{n,p}$ with (5.7) reading as

$$\begin{aligned}d_{n,p}(f(x), f(y))^2 &= -\sum_{i=1}^{n} |\lambda_i| \left(u_i(x) - u_i(y)\right)^2 + \sum_{j=n+1}^{n+p} |\lambda_j| \left(u_j(x) - u_j(y)\right)^2\\ &= \sum_{i=1}^{\nu} \lambda_i \left(u_i(x) - u_i(y)\right)^2 = d^2(x,y).\end{aligned}$$

Hence, f is an isometry onto the image $\Sigma := f(X)$ with $\Sigma - \Sigma \subset C_{n,p}$, concluding the proof of the first part.

For the converse statement, we consider $X_N = \{x_i\}_{i=1}^N \subset \operatorname{supp}\mu$. Identifying X_N with $f(X_N)$, we may assume that $X_N \subset \mathbb{R}^{n,p}$ with $X_N - X_N \subset C_{n,p}$. Then

$T_N := \Pi_N S(X_N)\Pi_N$ is the matrix

$$T_{ij} = (\bar{x}_i, \bar{x}_j)_{n,p}, \quad \text{where } \bar{x}_i := x_i - \bar{x}, \quad \bar{x} := \frac{1}{N}\sum_{i=1}^{N} x_i.$$

Thus, one can view T_N as the difference between two matrices, $T_N = T_+ - T_-$, where

$$(T_-)_{ij} = (P_n^- \bar{x}_i) \cdot (P_n^- \bar{x}_j), \qquad (T_+)_{ij} = (P_p^+ \bar{x}_i) \cdot (P_p^+ \bar{x}_j),$$

P_n^- (resp. P_p^+) standing for the projections from $\mathbb{R}^{n+p}$ to $\mathbb{R}^n$ identified with the subspace of $\mathbb{R}^{n+p}$ with all coordinates zero except the first n ones (resp. to $\mathbb{R}^p$ identified with the subspace of $\mathbb{R}^{n+p}$ with all coordinates zero except the last p ones). Also, $u \cdot v$ stands for the usual dot product between two vectors in $\mathbb{R}^{n+p}$. Both T_- and T_+ are clearly positive semidefinite with $s_+(T_-) \le n$, $s_+(T_+) \le p$, and $s_-(T_-) = s_-(T_+) = 0$. From Lemma A.1(i), we get that

$$s_+(T_N) \le s_+(T_+) + s_+(-T_-) = s_+(T_+) + s_-(T_-) \le p.$$

A completely symmetric reasoning yields $s_-(T_N) \le n$. Then Proposition 4.15 applied to the space $(\operatorname{supp}\mu, d)$ yields the desired claim

$$s_+(T_\mu) \le p, \quad s_-(T_\mu) \le n,$$

since $L^2(X,\mu)$ is isomorphic to $L^2(\operatorname{supp}\mu, \mu)$. □

Note that not every subset Σ of a pseudo-Euclidean space $\mathbb{R}^{n,p}$ is an image of an isometric embedding of a metric space. In particular, the condition $\Sigma - \Sigma \in C_{n,p}$ is a strong restriction, see the example below.

Example 5.7. If $n = 1$, and $\Sigma \subset \mathbb{R}^{n,p}$ satisfies $\Sigma - \Sigma \subset C_{n,p}$, then the intersection of Σ with every line parallel to x_1 axis must be either empty or a singleton. In other words, Σ belongs to a hypersurface which, if seen in $\mathbb{R}^{n+p} = \mathbb{R}^{p+1}$, is a graph G of some function: $x_1 = g(x_2, \dots, x_{p+1})$. Clearly, $g\colon \mathbb{R}^p \to \mathbb{R}$ is a 1-Lipschitz function, since $G - x \subset C_{n,p}$ for every $x \in G$.

5.3. **On feasible signatures.** The following result shows that for a generic finite metric space there exists an arbitrarily small perturbation of the distance which preserves s_+ and makes s_- maximum possible.

Proposition 5.8. *Let (X,d) be a finite metric space with $X = \{x_1, \dots, x_N\}$ such that the triangle inequality is strict for any distinct triplet from X. Consider the matrix $T := \Pi_N S(X)\Pi_N$. Then for any sufficiently small $\varepsilon > 0$, there is a distance d_ε over X such that*

$$\max_{i,j} |d(x_i, x_j) - d_\varepsilon(x_i, x_j)| \le \varepsilon$$

and the corresponding matrix T_ε has signature

$$s_+(T_\varepsilon) = s_+(T), \quad s_-(T_\varepsilon) = N - 1 - s_+(T).$$

Proof. Fix linearly independent vectors $v_1, \dots, v_N \in \mathbb{R}^N$. For any $\varepsilon > 0$ define

$$g_\varepsilon(x_i, x_j) = d^2(x_i, x_j) - \varepsilon \|v_i - v_j\|^2.$$

Note $g_\varepsilon \to d^2$ uniformly as $\varepsilon \to 0^+$. For sufficiently $\varepsilon > 0$, we have $d_\varepsilon := \sqrt{g_\varepsilon} \ge 0$. Moreover, since for any distinct i, j, k, the strict triangle inequality

$$d(x_i, x_j) < d(x_i, x_k) + d(x_k, x_j)$$

holds, then for all sufficiently small ε, one has the triangle inequality

$$d_\varepsilon(x_i, x_j) \le d_\varepsilon(x_i, x_k) + d_\varepsilon(x_k, x_j).$$

Thus, d_ε is a distance over X for small $\varepsilon > 0$. We set then $X_\varepsilon := (X, d_\varepsilon)$.

Now, it is easy to see that the matrix corresponding to (X, d_ε) has the form

$$T_\varepsilon := \Pi_N S(X_\varepsilon)\Pi_N = T - \varepsilon G,$$

where $G := \big((v_i - \bar{v}, v_j - \bar{v})\big)_{i,j=1}^N$ with $\bar{v} := \frac{1}{N}\sum_{i=1}^N v_i$. In particular, G is a Gram matrix with $s_+(G) = N - 1$, $s_-(G) = 0$, thus due to Lemma A.1(i) one has

$$\begin{aligned} s_+(T_\varepsilon) &\le s_+(T) + s_+(-\varepsilon G) = s_+(T), \\ s_-(T_\varepsilon) &\ge s_-(-\varepsilon G) - s_-(-T) = N - 1 - s_+(T). \end{aligned}$$

Moreover, for small enough ε, one has $s_+(T_\varepsilon) = s_+(T)$ by Lemma A.1(ii). Since $\ker T$ is non-trivial, we have $s_-(T_\varepsilon) \le N - 1 - s_+(T_\varepsilon) = N - 1 - s_+(T)$. Thus, we get the equality for $s_-(T_\varepsilon)$. The claim follows. □

Corollary 5.9. *For any $p \ge 2$ and $n \ge 1$, there is an $(n+p+1)$-point metric space (X, d) such that the corresponding matrix T has signature $s_+(T) = p$, $s_-(T) = n$.*

Proof. It is enough to apply the above Proposition 5.8 to $N := n + p + 1$ distinct points in a general position (i.e. spanning $\mathbb{R}^p$) on the unit sphere in $\mathbb{R}^p$, endowed with the Euclidean distance. □

Remark 5.10. Under the conditions of the above Corollary 5.9, one clearly has $s_+(X, d) \in \{p, p+1\}$, $s_-(X, d) \in \{n, n+1\}$ in view of Theorem 4.2(iv). Thus, in particular, this corollary gives a construction of hollow symmetric nonnegative (HSN) matrices (in this particular case even stronger, squared distance matrices) with arbitrarily large number of positive eigenvalues. This provides an alternative proof to the main theorem of [6] (the latter seems to be much more complicated and gives just HSN matrices, not necessarily related to any distances).

We are also able to provide the following example of an infinite metric space (X, d) with $s_-(X, d) = +\infty$ and $s_+(X, d)$ finite (arbitrary up to the error of 1).

Example 5.11. Let $p \in \mathbb{N}$ be fixed. In the following construction we assume the pseudo-Euclidean spaces $\mathbb{R}^{n,p}$ with different $n \in \mathbb{N}$ to be hierarchically embedded one into another, in such a way that a vector $x \in \mathbb{R}^{n,p}$ is identified with the vector $y \in \mathbb{R}^{m,p}$, $m > n$, defined by

$$\begin{aligned} y_i &:= x_i, \quad i = 1, \dots, n, \\ y_i &:= 0, \quad i = n+1, \dots, m, \\ y_{m+j} &:= x_{n+j}, \quad j = 1, \dots, p. \end{aligned}$$

This gives a natural embedding of $\mathbb{R}^{n,p}$ into $\mathbb{R}^{m,p}$. In the same way, we identify $\mathbb{R}^n$ (resp. $\mathbb{R}^p$) with the subspace of $\mathbb{R}^{n,p}$ with last p coordinates (resp. first n coordinates) zero. As in the proof of Theorem 5.6, we denote by P_n^- (resp. P_p^+) the projections from $\mathbb{R}^{n+p}$ to $\mathbb{R}^n$ (resp. $\mathbb{R}^p$), both identified with subspaces of $\mathbb{R}^{n+p}$, defined by $(P_n^- z)_i := z_i$, and $(P_n^- z)_j := 0$ (resp. $(P_p^+ z)_i := 0$, $(P_p^+ z)_j := z_j$), where $i = 1, \dots, n$, $j = n+1, \dots, n+p$).

We are going to construct inductively a sequence of increasing sets

$$X_N = \{z^1, \dots, z^N\} \subset \mathbb{R}^{N-1,p}, \quad N \ge p+1,$$

with special properties. For brevity, we denote

$$S_N := S(X_N), \quad T_N := \Pi_N S_N \Pi_N.$$

As in the proof of Theorem 5.6, one has

$$\begin{aligned} T_N &= T_N^+ - T_N^-, \quad \text{where} \\ (T_N^-)_{ij} &:= (P_{N-1}^- \bar{z}^i) \cdot (P_{N-1}^- \bar{z}^j), \\ (T_N^+)_{ij} &:= (P_p^+ \bar{z}^i) \cdot (P_p^+ \bar{z}^j), \quad i,j = 1, \dots N, \\ \bar{z}^i &:= z^i - \frac{1}{N} \sum_{j=1}^N z^j, \end{aligned}$$

where the dot product is the Euclidean one. Clearly, $T_N^{\pm}$ are both positive semidefinite. We will construct $(X_N)_{N=p+1}^{\infty}$ in such a way that

(i) both $P_{N-1}^- X_N$ and $P_p^+ X_N$ are sets of points in a general position in $\mathbb{R}^{N-1}$ and $\mathbb{R}^p$ respectively (i.e. not contained in any hyperplane in these spaces),
(ii) $d_{N-1,p}$ is a distance over X_N and for every triple of distinct points in X_N, the triangle inequality for $d_{N-1,p}$ is strict,
(iii) $P_p^+ X_N \subset S$, where S is the unit sphere in $\mathbb{R}^p$,
(iv) $u \cdot P_p^+(z^k - z^1) < d_{N-1,p}(z^k, z^1)|u|$ for all $k = 2, \dots, N$ and non-zero $u \in T_{y^1}(S)$, where $T_{y^1}(S)$ is the tangent space to S at point $y^1 := P_p^+ z^1$, and $|\cdot|$ stands for the Euclidean norm in $\mathbb{R}^p$.

In view of (i) we will have then

$$s_+(T_N^-) = N - 1, \quad s_+(T_N^+) = p, \quad s_-(T_N^-) = s_-(T_N^+) = 0.$$

Thus from $T_N = T_N^+ - T_N^-$ with the help of Lemma A.1(i) we get

$$\begin{aligned} s_-(T_N) &\geq s_-(-T_N^-) - s_-(-T_N^+) \\ &= s_+(T_N^-) - s_+(T_N^+) = N - p - 1, \end{aligned} \tag{5.8}$$

as well as $s_+(T_N) \leq s_+(T_N^+) + s_-(T_N^-) = p$.

We start with $N := p + 1$: take $p + 1$ points $\{y^1, \dots, y^{p+1}\}$ in a general position on the unit sphere $S \subset \mathbb{R}^p$. Now, let $\{x^1, \dots, x^{p+1}\} \subset \mathbb{R}^p$ be in a general position and with the norms $|x^k|$, $k = 1, \dots, p + 1$, so small that for $d_{p,p}$ satisfies strict triangle inequalities over the set

$$X_{p+1} := \{z^1, \dots, z^{p+1}\}, \quad z^k := (x^k, y^k) \in \mathbb{R}^{p,p},$$

the property (iv) holds for $N := p + 1$ and $s_+(T_{p+1}) = s_+(T_{p+1}^+) = p$ (where the last equality is due to Example 5.5). Then, in particular,

$$s_+(S(X_{p+1})) = p \tag{5.9}$$

in view of Theorem 4.2(iv) and the fact that $s_-(S(X_{p+1})) \geq 1$. The possibility to satisfy (iv) follows from the strict convexity of the unit ball in $\mathbb{R}^p$ which implies

$$u \cdot (y^k - y^1) < |y^k - y^1| \cdot |u|$$

for all $k \neq 1$ and nonzero $u \in T_{y^1}(S)$, and

$$d_{p,p}(z^k, z^1) \to |y^k - y^1| \quad \text{as } x_k \to 0,\ x_1 \to 0.$$

Once $X_N = \{z^1, \dots, z^N\} \subset \mathbb{R}^{N-1,p}$ is constructed, we denote

$$y^k := P_p^+ z^k \in S, \quad x^k := P_{N-1}^- z^k.$$

We will choose $z^{N+1} = (x^{N+1}, y^{N+1})$ such that $y^{N+1} \in S$ and $x^{N+1} \in \mathbb{R}^N$ has the form

$$(x^{N+1})_i := (x^1)_i, \quad i = 1, \dots, N-1,$$
$$(x^{N+1})_N := \varepsilon_N,$$

where $\varepsilon_N > 0$. First, note that

$$\operatorname{dist}\left(\frac{y - y^1}{|y - y^1|}, T_{y^1}(S)\right) \to 0 \quad \text{as } y \to y^1,$$

where dist stands for the distance in $\mathbb{R}^p$ between a point and a set. Then, due to (iv), one can choose y^{N+1} close enough to y^1, so that for an auxiliary point $\tilde{z}^{N+1} := (x^1, y^{N+1})$ one has

$$\begin{aligned}(\tilde{z}^{N+1} - z^1, z^1 - z^k)_{N,p} &= (y^{N+1} - y^1) \cdot (y^1 - y^k)\\ &< d_{N,p}(z^k, z^1)|y^{N+1} - y^1|,\\ (z^1 - \tilde{z}^{N+1}, \tilde{z}^{N+1} - z^k)_{N,p} &= (y^1 - y^{N+1}) \cdot (y^{N+1} - y^k)\\ &< d_{N,p}(z^k, \tilde{z}^{N+1})|y^{N+1} - y^1|,\end{aligned}$$

and

$$|y^{N+1} - y^1| < d_{N,p}(z^1, z^k),$$

hence

$$d_{N,p}(\tilde{z}^{N+1}, z^1) = |y^{N+1} - y^1| < d_{N,p}(z^1, z^k) + d_{N,p}(z^k, \tilde{z}^{N+1}),$$

for all $k = 2, \dots, N$, with $a \cdot b$ standing for the Euclidean dot product in $\mathbb{R}^p$. Thus, by Lemma A.5,

$$d_{N,p}(\tilde{z}^{N+1}, z^k) < d_{N,p}(\tilde{z}^{N+1}, z^1) + d_{N,p}(z^1, z^k),$$
$$d_{N,p}(\tilde{z}^1, z^k) < d_{N,p}(z^1, \tilde{z}^{N+1}) + d_{N,p}(\tilde{z}^{N+1}, z^k).$$

for all $k = 2, \dots, N$, and hence $d_{N,p}$ satisfies the strict triangle inequality on the set $X_N \cup \{\tilde{z}^{N+1}\}$. Now, we can choose ε_N small enough, such that $d_{N,p}$ is a distance over $X_{N+1} := X_N \cup \{z^{N+1}\}$ satisfying the strict triangle inequality, and

$$u \cdot (y^{N+1} - y^1) < d_{N,p}(z^{N+1}, z^1)|u|$$

for all nonzero $u \in T_{y_1} S$. The latter follows from the fact that

$$u \cdot (y^{N+1} - y^1) < |y^{N+1} - y^1| \cdot |u|$$

due to the strict convexity of the unit ball in $\mathbb{R}^p$, and

$$d_{N,p}(z^{N+1}, z^1) = \sqrt{|y^{N+1} - y^1|^2 - \varepsilon_N^2} \to |y^{N+1} - y^1| \quad \text{as } \varepsilon_N \to 0.$$

Therefore, X_{N+1} satisfies (i)-(iv).

To conclude, let $X := \bigcup_{N=p+1}^{\infty} X_N$, and equip this set with the distance coinciding with $d_{N,p}$ over each X_N. Then

$$s_\pm(X, d) = \lim_n s_\pm(S_N).$$

by Theorem 4.2(i). Clearly, in view of Theorem 4.2(iv) and (5.8), one has

$$s_-(S_N) \geq N - p - 1,$$

and hence $s_-(X,d) = +\infty$. On the other hand, by Theorem 5.6 combined with Theorem 4.2(iv), we get $s_+(X,d) \le p+1$. Finally, $s_+(X,d) \ge s_+(S(X_{p+1})) = p$ by (5.9), thus

$$s_+(X,d) \in \{p, p+1\} \quad \text{and} \quad s_-(X,d) = +\infty,$$

concluding the example.

The following example shows another construction of an infinite metric space (X,d) with $s_+(X,d) = +\infty$ and $s_-(X,d)$ finite but different from 1.

Example 5.12. Let $X := \mathbb{N}$, and define the distance

$$d(i,j) := \begin{cases} 1, \ i<4, \ j=4 \quad \text{or } i=4, \ j<4, \\ 0, \ i=j, \\ 2, \ \text{otherwise.} \end{cases}$$

Note that the points $1,2,3,4$ form the same tripod from Example 5.3, which cannot be embedded in any Hilbert space. The subset $X_N := \{1,\dots,N\}$ has the $N \times N$ squared distance matrix

$$B_N := \begin{pmatrix} 0 & 4 & 4 & 1 & 4 & \dots & 4 \\ 4 & 0 & 4 & 1 & 4 & \dots & 4 \\ 4 & 4 & 0 & 1 & 4 & \dots & 4 \\ 1 & 1 & 1 & 0 & 4 & \dots & 4 \\ 4 & 4 & 4 & 4 & 0 & \dots & 4 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 4 & 4 & 4 & 4 & 4 & \dots & 0 \end{pmatrix}$$

and by the Haynsworth inertia additivity formula [9, theorem 1] its signature satisfies

$$(s_-, s_0, s_+)(B_N) = (s_-, s_0, s_+)(B_4) + (s_-, s_0, s_+)(B_N/B_4),$$

where B_N/B_4 stands for the Schur complement of B_4 in B_N. The explicit calculation shows $(s_-, s_0, s_+)(B_4) = (3,0,1)$ and

$$B_N/B_4 := \frac{4}{3} \begin{pmatrix} 8 & 11 & 11 & 11 & \dots & 11 \\ 11 & 8 & 11 & 11 & \dots & 11 \\ 11 & 11 & 8 & 11 & \dots & 11 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 11 & 11 & 11 & 11 & \dots & 8 \end{pmatrix}.$$

For $N := 4+k$ the latter $k \times k$ matrix has one eigenvalue -4 of multiplicity $k-1$ and one strictly positive simple eigenvalue (the Perron–Frobenius one), that is, $(s_-, s_0, s_+)(B_{4+k}/B_4) = (k-1, 0, 1)$. Thus, $(s_-, s_0, s_+)(B_{4+k}) = (k+2, 0, 2)$ for $k \ge 1$, and therefore, $s_-(X,d) = 2$ and $s_+(X,d) = +\infty$.

Now, we extend the previous Example 5.12 to show that s_- can be arbitrarily large.

Example 5.13. Let (X_i, d_i), $i = 1,\dots,m$, be a finite metric space with the squared distance matrix $D_i := (d_i^2(x_j, x_k))_{j,k}$. Assume $\operatorname{diam}(X_i) \le 2h$ for some $h > 0$.

Denote for the sake of brevity $N_i := \#X_i$, $S_i := S(X_i)$ and

$$R_i := \frac{h^2}{2} \mathbf{1}_{N_i} \mathbf{1}_{N_i}^T + S_i = \frac{1}{2}(h^2 \mathbf{1}_{N_i} \mathbf{1}_{N_i}^T - D_i).$$

Now we define a metric space (X, d) with $X := \bigsqcup_{i=1}^m X_i$ and

$$d(x,y) := \begin{cases} d_i(x,y), & x,y \in X_i, \\ h, & x \neq y \text{ otherwise.} \end{cases}$$

The respective squared distance matrix is given by

$$D := \begin{pmatrix} D_1 & h^2 & \dots & h^2 \\ h^2 & D_2 & \dots & h^2 \\ \vdots & \vdots & \ddots & \vdots \\ h^2 & h^2 & \dots & D_m \end{pmatrix}.$$

Let also

$$\begin{aligned} S &:= S(X) = -\tfrac{1}{2} D, \\ R &:= \tfrac{h^2}{2} \mathbf{1}_N \mathbf{1}_N^T + S = \tfrac{1}{2}(h^2 \mathbf{1}_N \mathbf{1}_N^T - D) = \operatorname{diag}(R_1, \dots, R_m), \end{aligned}$$

where $N := \sum_{i=1}^m N_i$. Note that $s_\pm(R) = \sum_{i=1}^m s_\pm(R_i)$. Further, since $\frac{h^2}{2} \mathbf{1}_N \mathbf{1}_N^T$ is positive semidefinite rank one matrix, by Lemma A.1, we get

$$s_+(S) \leq s_+(R) \leq s_+(S) + 1, \quad s_-(S) - 1 \leq s_-(R) \leq s_-(S).$$

Let us consider a specific example, where all X_i for $i = 1, \dots, m-1$, are isometric with

$$D_i = \begin{pmatrix} 0 & 4 & 4 & 1 \\ 4 & 0 & 4 & 1 \\ 4 & 4 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

as in the previous example, and X_m is such that $d_m(x,y) = 2$ if $x \neq y$. Since $\operatorname{diam}(X_i) = 2$ for all X_i, one can set $h := 1$. Thus, for $i = 1, \dots, m-1$

$$R_i = \frac{1}{2} \begin{pmatrix} 1 & -3 & -3 & 0 \\ -3 & 1 & -3 & 0 \\ -3 & -3 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

with $\sigma(R_i) = (-5/2, 1/2, 2, 2)$ and $s_+(R_i) = 3$, $s_-(R_i) = 1$. Furthermore, X_m can be isometrically embedded into a Euclidean space, and

$$D_m = 4\mathbf{1}_{N_m} \mathbf{1}_{N_m}^T - 4\mathrm{Id}_{N_m}, \quad S_m = 2\mathrm{Id}_{N_m} - 2, \quad R_m = 2\mathrm{Id}_{N_m} - \frac{3}{2} \mathbf{1}_{N_m} \mathbf{1}_{N_m}^T,$$

with $s_+(R_m) = N_m - 1$, $s_-(R_m) = 1$. Thus,

$$s_+(R) = 3(m-1) + N_m - 1, \quad s_-(R) = m,$$

and therefore,

$$s_+(X,d) = s_+(S) \in \{3m-5+N_m, 3m-4+N_m\}, \quad s_-(X,d) = s_-(S) \in \{m, m+1\}.$$

Note that the constructed space X can be represented as a graph endowed with its intrinsic distance.

6. Some further properties and examples of limit distance signatures

6.1. Further examples of infinite limit distance signature.

Example 6.1. Let S^n be the standard unit n-dimensional sphere equipped with its intrinsic distance d. Then $s_\pm(S^n, d) = +\infty$ (see [12, proof of proposition 6.1]). On the other hand, $s_+(S^n, d^{1/2}) = +\infty$ and $s_-(S^n, d^{1/2}) = 1$ since $(S^n, d^{1/2})$ can be isometrically embedded in the Hilbert space ℓ^2 by proposition 6.1 from [12].

Example 6.2. Let $\mathbb{T}^n$ be the standard n-dimensional flat torus equipped with its intrinsic distance d. Then $s_\pm(\mathbb{T}^n, d) = +\infty$ due to Example 6.1 above and Proposition 4.5.

Example 6.3. In view of the above Proposition 4.5, for the Urysohn universal space $\mathfrak{U}$ [15], one has $s_\pm(\mathfrak{U}) = +\infty$, because $\mathfrak{U}$ contains an isometric copy of every separable metric space, e.g., of the unit circle equipped with its intrinsic distance. For the same reason $s_\pm(c_0) = +\infty$, where c_0 stands for the Banach space of vanishing sequences, equipped with its usual supremum norm. Again by the same reason, we have that $s_\pm(X, d) = +\infty$ when (X, d) is either the cylinder $S^1 \times \mathbb{R}$ or the revolution torus in $\mathbb{R}^3$, equipped with their intrinsic distances.

Squared distance matrices are special cases of hollow symmetric non-negative (HSN) matrices, i.e., symmetric matrices with nonnegative entries and with zeros on the diagonal. Using the spectral theory of HSN matrices from [6], we obtain the following result for countable metric spaces.

Proposition 6.4. *Suppose (X, d) is a bounded infinite separable metric space, and there exists $\delta_0 > 0$ such that $\delta_0 \leq d(x, y)$ for all $x \neq y$. Then $s_+(X, d) = +\infty$, in particular, $s_+(T_\mu) = s_+(K_\mu) = +\infty$ for every Borel probability measure μ on X with full support.*

Proof. For every finite subset $X_N \subset X$, the squared distance matrix of X_N is a HSN matrix with off-diagonal entries from the interval $[\delta_0^2, \delta_1^2]$, where δ_1 is the diameter of (X, d). Since $\delta_0 > 0$ and $\delta_1 < +\infty$, by theorem 3.3 from [6], there exists some $n_0 \in \mathbb{N}$ such that any HSN matrix of order at least n_0 with off-diagonal entries from $[\delta_0^2, \delta_1^2]$ admits at least m strictly negative eigenvalues. Therefore, for any fixed m, there exists some n_0 such that for any finite $X_N \subset X$ with $\#X_N \geq n_0$, one has that $s_+(S(X_N)) \geq m$, proving the claim. □

The assumptions of the above corollary are satisfied, for instance, when the metric space (X, d) is induced by a connected countable graph with its intrinsic metric on vertices being bounded.

6.2. Limits and tangent cones.

Proposition 6.5. *Let (X_k, d_k) be a sequence of separable metric spaces with $p_k \in X_k$, and the sequence of pointed metric spaces (X_k, d_k, p_k) converge as $k \to \infty$ to a pointed separable metric space (X, d, p) in the pointed Gromov–Hausdorff sense. Assume that all (X_k, d_k) and (X, d) satisfy the Heine–Borel property (i.e., closed balls are compact). Then*

$$s_\pm(X, d) \leq \liminf_k s_\pm(X_k, d_k).$$

Proof. We prove the statement for s_+, and the proof for s_- is completely analogous. Fix a finite set $\Sigma_N \subset X$ and let $\bar{B}_r(p) \subset X$ be a closed ball centered at p such that $\Sigma_N \subset \bar{B}_r(p)$. It follows from the definition of the pointed Gromov–Hausdorff convergence [5, definition 8.1.1] that for an arbitrary $\varepsilon > 0$ there are $\bar{k} \in \mathbb{N}$ and maps $f_\varepsilon^k \colon \bar{B}_r(p) \subset X \to \bar{B}_{r+\varepsilon}(p_k) \subset X_k$ with $f_\varepsilon^k(p) = p_k$ such that

$$\left| d_k(f_\varepsilon^k(x), f_\varepsilon^k(y)) - d(x,y) \right| \le \varepsilon$$

for all $x, y \in \bar{B}_r(p)$ and $k \ge \bar{k}$. Choosing $\varepsilon > 0$ sufficiently small, by Lemma A.1(ii) one has

$$s_+(S(f_\varepsilon^k(\Sigma_N))) \ge s_+(S(\Sigma_N))$$

for all $k \in \mathbb{N}$ sufficiently large. We obtain the bound

$$\liminf_k s_+(X_k, d_k) \ge \liminf_k s_+(S(f_\varepsilon^k(\Sigma_N))) \ge s_+(S(\Sigma_N)).$$

Taking a supremum with respect to all finite subsets $\Sigma_N \subset X$, we get the claim. □

Corollary 6.6. *Suppose that the metric space (X, d) with Heine–Borel property has a tangent cone $(T_p, \tilde{d})$ at $p \in X$. Then $s_\pm(X, d) \ge s_\pm(T_p, \tilde{d})$. In particular, if (X, d) is a smooth Riemannian manifold equipped with its intrinsic distance, then $s_+(X, d) \ge \dim X$.*

Similarly, if (X, d) has a (Gromov–Hausdorff) asymptotic cone $(C_p, \tilde{d})$ at $p \in X$, then $s_\pm(X, d) \ge s_\pm(C_p, \tilde{d})$.

Proof. Recall that the pointed metric space $(T_p, \tilde{d}, p)$ (resp., $(C_p, \tilde{d}, p)$) is the pointed Gromov–Hausdorff limit of the sequence (X, kd, p) (resp., $(X, d/k, p)$) as $k \to +\infty$. Apply the above Proposition 6.5, having in mind that

$$s_\pm(X, kd) = s_\pm(X, d/k) = s_\pm(X, d).$$

If (X, d) is a smooth Riemannian manifold, it suffices to recall that $T_p = \mathbb{R}^n$ with $n = \dim X$, and $(T_p, \tilde{d})$ is isometric to $\mathbb{R}^n$ with the Euclidean distance so that $s_+(T_p, \tilde{d}) = s_+(\mathbb{R}^n) = n$ by Example 5.5. □

7. Limit signature of the Rado graph

7.1. **Preliminaries on the spectral theory of random matrices and the Wigner semi-circle law.** Random matrices play an important role in the study of the Rado graph, so we outline the relevant results on random matrices used in our paper.

Recall the following terminology in matrix spectral distribution. Let $W_N \in \mathbb{R}^{N \times N}$ be a real symmetric matrix. Denote by $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_N$ the eigenvalues of W_N counting multiplicity. Then the empirical spectral distribution (ESD) of the normalized matrix $\dfrac{1}{\sqrt{N}} W_N$ is the probability measure

$$\mu_N := \frac{1}{N} \sum_{j=1}^{N} \delta_{\lambda_j / \sqrt{N}}. \tag{7.1}$$

Now, we consider a sequence of random matrices $(W_N)_{N=1}^\infty$ constructed as follows. Let $(Y_k)_{k=1}^\infty$ be a sequence of i.i.d. random variables, and (Z_{ij}) with $1 \le i < j$

be a family of i.i.d. random variables. For each $N \geq 1$, W_N is a symmetric matrix of order N with entries given by

$$(W_N)_{k,k} = Y_k,\ (W_N)_{i,j} = (W_N)_{j,i} = Z_{ij}, \quad \text{for all } 1 \leq k \leq N,\ 1 \leq i < j \leq N.$$

Thus, W_N can be viewed as upper left submatrices of order N of the same infinite size random matrix. For each $N \geq 1$, by taking the ESD of the normalized random matrix $\frac{1}{\sqrt{N}} W_N$, we obtain a random probability measure of the form (7.1). The following version of the Wigner semi-circle law from [2, theorem 2.5] indicates the ESDs of these normalized random matrices converge to the semi-circle distribution.

Proposition 7.1 (Wigner semi-circle law)**.** *Let W_N be the order N random real symmetric matrix such that the diagonal terms are i.i.d. random variables, and the off-diagonal terms are i.i.d. random variables with variance $\sigma^2 > 0$, constructed as above. Then with probability 1, the ESDs of normalized random matrices $\frac{1}{\sqrt{N}} W_N$ weakly converge to the semi-circle law, i.e., to a measure with density*

$$\rho_\sigma(x) = \begin{cases} \frac{1}{2\pi\sigma^2}\sqrt{4\sigma^2 - x^2}, & -2\sigma \leq x \leq 2\sigma, \\ 0, & \text{otherwise.} \end{cases}$$

In particular, the above theorem implies that the signatures of the matrices $(W_N)_{N=1}^{\infty}$ almost surely satisfy

$$\lim_{N\to\infty} s_\pm(W_N) = \infty, \quad \lim_{N\to\infty} \frac{s_+(W_N)}{s_-(W_N)} = 1. \tag{7.2}$$

7.2. **Signatures of the MDS operators and limit signature of the Rado graph.** The Rado graph R is a homogeneous, countable, and universal graph in the sense that it contains all finite graphs as its induced subgraphs. The Rado graph also has the following finite extension property: given any two finite disjoint subsets U, V of R, there exists an $x \in R \setminus (U \cup V)$ such that x is adjacent to all points in U, but not adjacent to any point in V. The Rado graph is known to be unique up to graph isomorphism, see [7, chapter VII] for details. Let d be the intrinsic (graph) metric on R, and μ be any Borel probability measure on (R, d) having full support. The above properties immediately imply that the intrinsic metric d of R is bounded with $d \leq 2$.

We also recall that the Rado graph admits the following random construction (infinite Erdős–Rényi graph). Fix any real number $0 < p < 1$. Let (Z_{ij}) with $i < j \in \mathbb{N}$ be a family of i.i.d. Bernoulli random variables: $Z_{ij} \sim Be(p)$. Then we define a random graph G with the set of vertices $\mathbb{N}$ and the random adjacency matrix $\mathcal{A}$ defined by

$$\mathcal{A}_{ij} = \mathcal{A}_{ji} = Z_{ij},\ \mathcal{A}_{ii} = 0, \quad \text{for all } 1 \leq i < j. \tag{7.3}$$

With probability 1, this construction yields a graph isomorphic to the Rado graph—thus, it can be considered as a "random enumeration" of R. Furthermore, for each $N \geq 1$, let S_N be the random matrix of order N such that

$$(S_N)_{ij} = \frac{3}{2}\mathcal{A}_{ij} - 2, \quad (S_N)_{ii} = 0 \quad \text{for all } 1 \leq i \neq j \leq N. \tag{7.4}$$

Then almost surely one has $(S_N)_{ij} = -\frac{d_G^2(i,j)}{2}$ for $1 \leq i, j \leq N$, where d_G is the intrinsic metric on G. Therefore, provided that G is isomorphic to R, there

is a sequence of finite subsets $R_N \subset R$ such that $R_N \nearrow R$, $\#R_N = N$, and $S_N = S(R_N)$. Since the random matrices $(S_N)_{N\geq 1}$ also satisfy the assumptions of Proposition 7.1, we obtain the following theorem on the limit distance signature of (R, d, μ).

Theorem 7.2. *One has*

$$s_\pm(R, d) = +\infty$$

and $s_\pm(K) = s_\pm(T) = +\infty$ *for every probability measure* μ *over* R *having full support. Moreover, the only functions in* $L^2(R, \mu)$ *in the kernel of* T *are constants.*

Proof. By the Wigner semi-circle law (Proposition 7.1), with probability 1, we have

$$\lim_{N\to\infty} s_\pm(S_N) = +\infty.$$

On the other hand, by the random construction of the Erdős–Rényi–Rado graph, almost surely there is a sequence of finite subsets $R_N \subset R$ with $R_N \nearrow R$ as $N \to \infty$ such that

$$S_N = S(R_N).$$

Thus by Theorem 4.2(i), we have $s_\pm(R, d) = +\infty$, which by Theorem 4.2(v) implies $s_\pm(K) = +\infty$. Finally, Proposition 4.14 yields $s_\pm(T) = +\infty$.

It remains to compute the dimension of the kernel $\ker T$ of T. Since $T = PKP$, clearly $\mathbf{1} \in \ker T$. Since P is self-adjoint in $L^2(R, \mu)$, we have $0 \neq f \in \mathbf{1}^\perp \cap \ker T$ if and only if $K(f)$ is collinear to the constant function $\mathbf{1}$ in $L^2(R, \mu)$. Number the vertices of R by positive integers. The fact μ has full support implies $(Kf)_i = (Kf)_j$ for all $i, j \in \mathbb{N}$. The sequence $(Kf)_i$ has an infinite "matrix multiplication" representation. More specifically, set

$$S = -\frac{1}{2}\left(d^2(i, j)\right)_{i,j=1}^{\infty},\ v_j = f_j\mu(\{j\}).$$

We then have

$$(Kf)_i = \sum_{j=1}^{\infty} S_{ij}v_j.$$

For any $n \geq 1$, define

$$v_n^+ = \sum_{j\leq n,\ v_j\geq 0} v_j,\ v_n^- = \sum_{j\leq n,\ v_j<0} v_j$$

Because μ has full support and $f \neq 0$, we see $v_n^+ \nearrow v^+ > 0$ and $v_n^- \searrow v^- < 0$. Since $f \in L^2(R, \mu)$ and $\mu(R) = 1$, the sequence (v_j) is absolutely convergent and v^+, v^- are finite. Using the finite extension property of the Rado graph, for any fixed $n \geq 1$, we can find $i_1 > n$ such that i_1 is adjacent to all vertices in the summation of v_n^+ while non-adjacent to all vertices in the summation of v_n^-. Therefore, we obtain

$$(Kf)_{i_1} = -\frac{1}{2}\left(v_n^+ + 4v_n^- + \sum_{j>n} v_j d^2(i_1, j)\right)$$

Similarly we can find $i_2 > n$ such that

$$(Kf)_{i_2} = -\frac{1}{2}\left(4v_n^+ + v_n^- + \sum_{j>n} v_j d^2(i_2, j)\right)$$

Using $(Kf)_{i_1} = (Kf)_{i_2}$, we get

$$-3v_n^+ + 3v_n^- + \sum_{j>n}(d^2(i_1,j) - d^2(i_2,j))v_j = 0 \tag{7.5}$$

Since the diameter of (R,d) is 2, we find

$$\left|\sum_{j>n}\left(d^2(i_1,j) - d^2(i_2,j)\right)v_j\right| \leq 4\sum_{j>n}|v_j|$$

Since $\{v_j\}$ is absolutely convergent, we find a contradiction to (7.5) for sufficiently large n. It follows that $\ker T \cap \mathbf{1}^\perp = \{0\}$. Since $\mathbf{1}^\perp$ has codimension 1 in $L^2(R,\mu)$, the dimension of $\ker T$ cannot exceed 1. Because $\ker T$ contains constants, they are the only elements of the kernel of T as claimed. □

Remark 7.3. In [12], the authors showed for a homogeneous closed connected Riemannian manifold X equipped with the canonical distance d and Riemannian volume μ, the constants are in the kernel of the MDS defining operator T for (X,d,μ). The results on such smooth manifolds given by these authors appear to share similarities with the theorem above. However, in contrast to the Riemannian volume measure on homogeneous spaces, a probability measure with full support on a homogeneous countable metric space can never be homogeneous, i.e., invariant under automorphisms of the graph.

7.3. **Signature ratio limit of i.i.d. generated finite subsets of the Rado graph.** We examine the following question on the Rado graph. Let $(R_N)_{N\geq 1}$ be an increasing sequence of finite subsets of R such that $R_N \nearrow R$, endowed the restriction of the metric d. Denote by $\Delta(R_N)$ the ratio of the number of positive eigenvalues over the number of negative eigenvalues of the associated matrix $S(R_N)$, that is,

$$\Delta(R_N) := \frac{s_+(S(R_N))}{s_-(S(R_N))}.$$

We ask whether the limit $\lim_{N\to\infty}\Delta(R_N)$ exists on the extended real line, and whether this limit is independent of the choice of $(R_N)_{N\geq 1}$. Note first that the Wigner semicircle law and the random construction of R imply there exists a sequence $R_N \nearrow R$ such that $\lim_{N\to\infty}\Delta(R_N) = 1$, see Section 7.2. However, this is not true for all sequences, as the following example shows.

Example 7.4. Fix any point $R'_1 = \{r_1\}$. By the finite extension property, if $R'_N \subset R$ is finite such that all pairs of vertices are adjacent, there exists $r_{N+1} \in R \setminus R'_k$ such that $R'_{N+1} = R'_N \cup \{r_{N+1}\}$ has all pairs of vertices adjacent. Thus, we obtain a strictly increasing sequence $(R'_N)_{N\geq 1}$ of finite subsets of R such that the vertices of R'_N are all adjacent for each $N \geq 1$. Each R'_N with the metric induced by d can be isometrically embedded into $\mathbb{R}^{N-1}$, but not in $\mathbb{R}^{N-2}$. By the Proposition 5.1, we have $s_+(S(R'_N)) = N-1$ and $s_-(S(R'_N)) = 1$ for this sequence. Note that $R \setminus \cup_{N=1}^\infty R'_N$ is infinite, i.e.

$$R \setminus \cup_{N=1}^\infty R'_N = \{t_1, t_2, \ldots\}.$$

Define $R_N = R'_{N^2} \cup \{t_1, \ldots, t_N\}$, so we have $R_N \nearrow R$. using the Cauchy interlacing theorem, it is easy to show for the sequence $(R_N)_{N\geq 1}$ defined as above, we have $\lim_{N\to\infty}\Delta(R_N) = \infty$.

Thus we see the limit $\lim_{N\to\infty}\Delta(R_N)$ is not unique. Moreover, similarly one can show that it does not always exist on the extended real line. Although the limit of the ratios $\Delta(R_N)$ depends on the choice of $R_N \nearrow R$, they typically converge to 1 in the following sense.

Proposition 7.5. *Let G be a random Erdős–Rényi–Rado graph, μ be a probability measure on $\mathbb{N}$ with full support, and $\xi_i \in \mathbb{N}$, be i.i.d. random elements with $\operatorname{law}(\xi_i) = \mu$, $i \in \mathbb{N}$, independent of G. Then, for $X_m := (\xi_i)_{i=1}^m$ one has*

$$\lim_{m\to\infty}\Delta(X_m) = 1$$

almost surely.

Proof. Since μ has an infinite support, a random sequence $\xi = (\xi_i)_{i=1}^\infty$ a.s. has infinitely many distinct elements. For such a realization of ξ, define an auxiliary sequence $\hat\xi = (\hat\xi_i)_{i=1}^\infty \subset \mathbb{N}$ which is obtained from ξ by cancelling the repeating elements in the order of their appearance. Since G and ξ are independent, the conditional law of G given ξ (uniquely defined for a.e. ξ) does not depend on ξ and coincides with the marginal law of G, that is, the distributional law of an Erdős–Rényi–Rado random graph. The matrices $\hat S_N$ corresponding to S_N (defined by (7.4)) are given by the formulae

$$(\hat S_N)_{ij} := \frac{3}{2}\mathcal{A}_{\hat\xi_i,\hat\xi_j} - 2,\ (\hat S_N)_{ii} = 0, \quad 1 \le i \ne j \le N.$$

Since all $\hat\xi_i$ are by definition distinct, by the construction of an Erdős–Rényi–Rado graph and Proposition 7.1 we obtain that ESDs of $\hat S_N$ converge to the semicircle law a.s. (conditioned on ξ, i.e. with ξ and hence also $\hat\xi$ fixed). Therefore,

$$\lim_{N\to\infty}\frac{s_+(\hat S_N)}{s_-(\hat S_N)} = 1,$$

and by the Fubini theorem, this holds (unconditionally) a.s.

Finally, by the properties of the Rado graph, one a.s. has $(\hat S_N)_{ij} = -\frac12 d_G^2(\hat\xi_i,\hat\xi_j)$, i.e. $\hat S_N = S(\hat X_N)$, where $\hat X_N := \{\hat\xi_i\}_{i=1}^N$. Then by Lemma 4.9 we get

$$\lim_{m\to\infty}\Delta(X_m) = \lim_{N\to\infty}\Delta(\hat X_N) = \lim_{N\to\infty}\frac{s_+(S(\hat X_N))}{s_-(S(\hat X_N))} = \lim_{N\to\infty}\frac{s_+(\hat S_N)}{s_-(\hat S_N)} = 1 \text{ a.s.}$$

as claimed. □

The above Proposition 7.5 shows that the limit of the ratios $\lim_{m\to\infty}\Delta\left((\xi_i)_{i=1}^m\right)$ is equal to 1 almost surely for almost every realization of the random Rado graph G, once the random points ξ_i in the Gromov–Vershik scheme are independent of G. This is however not the case when we fix a specific realization of the Rado graph as the following examples show. In fact, we show that the limit of the ratios $\lim_{m\to\infty}\Delta\left((\xi_i)_{i=1}^m\right)$ may become as large as desired (Example 7.6) or even infinity (Example 7.7), all this of course, with some positive probability and depending on the choice of the measure μ.

Example 7.6. Let (R,d) be as before. Let $E := \{e_1, e_2, \dots\} \subset R$ be the infinite Hilbertian simplex, i.e., $d(e_i,e_j) = \delta_{ij}$ (we call it Hilbertian since (E,d) can be isometrically embedded in the Hilbert space ℓ^2). The set $E_0 := R \setminus E$ is clearly infinite, and we enumerate its elements so that

$$E_0 = \{e_1^0, e_2^0, \dots\}.$$

Fix any integer $j \geq 1$. For $k = 1, \ldots, j$, define

$$E_k := \{e_l^k := e_{lj+k} : l \in \mathbb{N}\},$$

so that $E = \bigcup_{k=1}^{j} E_k$, and the sets E_k, $k = 1, \ldots, j$, are disjoint. We construct a probability measure μ of full support on R such that $\mu(\{e_l^0\}) = \mu(\{e_l^k\})$ for all $l \geq 0$ and $k = 1, \ldots, j$. Note that in this way

$$\mu(E_0) = \mu(E_1) = \ldots = \mu(E_j) = \frac{1}{j+1} \tag{7.6}$$

Let $\xi = (\xi_i)_{i \geq 1}$ with $\xi_i \colon \Omega \to R$ be the i.i.d. random sequence from (R, d, μ). For each $k = 0, \ldots, j$, denote by $N_{E_k}(m)$ the number of *distinct* elements in E_k of the finite random sequence $(\xi_1, \xi_2, \ldots, \xi_m)$.

We show first that for $p_j := 1/(j+1) > 0$, one has

$$\mathbb{P}\left(\{N_{E_0}(m, \omega) \leq N_{E_k}(m, \omega), k = 1, \ldots, j\}\right) \geq p_j \tag{7.7}$$

for all $m \in \mathbb{N}$. In fact, setting

$$\Omega_k := \{\omega \in \Omega : N_{E_k}(m, \omega) \leq N_{E_i}(m, \omega), i = 0, \ldots, j\},$$

we clearly get

$$\mathbb{P}(\Omega_0) = \mathbb{P}(\Omega_1) = \ldots = \mathbb{P}(\Omega_j),$$

and $\bigcup_{k=0}^{j} \Omega_k = \Omega$. Therefore,

$$1 = \mathbb{P}(\Omega) \leq \sum_{k=0}^{j} \mathbb{P}(\Omega_k) = (j+1)\mathbb{P}(\Omega_0),$$

hence the claim (7.7).

Once (7.7) is proven, we consider the set $\hat{R}_{N(m)}$ of distinct elements of the finite random sequence $(\xi_i)_{i=1}^m$, and the respective square matrix $S(\hat{R}_{N(m)})$ as well as the square matrices

$$\hat{S}^0_{N_{(E_0,m)}} := -\frac{1}{2}(d^2(\xi_i, \xi_j))_{i,j=1,\ldots,m, \xi_i, \xi_j \in E_0 \cap \hat{R}_{N(m)}},$$
$$\hat{E}_{N_E(m)} := -\frac{1}{2}(d^2(\xi_i, \xi_j))_{i,j=1,\ldots,m, \xi_i, \xi_j \in E \cap \hat{R}_{N(m)}},$$

where the lower index stands for the size of the respective matrix. One has

$$s_+\left(E_{N_E(m)}\right) = N_E(m) - 1,$$

since $E_{N_E(m)}$ is $N_E(m) \times N_E(m)$ matrix and (E, d) is isometrically embedded in a Hilbert space. Hence

$$s_+\left(\hat{E}_{N_E(m)}\right) = N_E(m) - 1. \tag{7.8}$$

We observe now that the matrix $S(\hat{R}_{N(m)})$ has the block form

$$S(\hat{R}_{N(m)}) = \begin{pmatrix} \hat{S}^0_{N_{E_0}(m)} & * \\ * & \hat{E}_{N_E(m)} \end{pmatrix}$$

(where $*$ stands for some blocks). Thus in view of the Cauchy interlacing theorem, we get

$$s_+(S(\hat{R}_{N(m)})) \geq s_+(\hat{E}_{N_E(m)}) = N_E(m) - 1, \tag{7.9}$$

the latter equality due to (7.8). On the other hand, $S(\hat{R}_{N(m)})$ is an $N(m) \times N(m)$ matrix, and hence

$$\begin{aligned} s_-(S(\hat{R}_{N(m)})) &\le N(m) - s_+(S(\hat{R}_{N(m)})) \\ &\le N(m) - (N_E(m) - 1) \qquad \text{by (7.9)} \\ &= N_{E_0}(m) + 1. \end{aligned} \tag{7.10}$$

From (7.8) by (7.7), of probability at least $p_j > 0$, one has

$$s_+\left(\hat{E}_{N_E(m)}\right) = N_E(m) - 1 = \sum_{k=1}^{j} N_{E_k}(m) - 1 \ge jN_{E_0}(m) - 1 \tag{7.11}$$

for sufficiently large m. Hence, combining (7.10), (7.11), and (7.9), we arrive at

$$\Delta(\hat{R}_{N(m)}) \ge \frac{jN_{E_0}(m) - 1}{N_{E_0}(m) + 1} \ge \frac{j}{3}. \tag{7.12}$$

Hence, for any $j \ge 1$, we have

$$\liminf_{m\to\infty} \mathbb{P}\left(\left\{\Delta(\hat{R}_{N(m)}) \ge \frac{j}{3}\right\}\right) \ge p_j > 0.$$

Setting j sufficiently large, we may have $\lim_{m\to\infty} \Delta\left((\xi_i)_{i=1}^m\right)$ as large as desired with a positive probability.

Example 7.7. Let (R, d) be as before. According to the Example 7.4, there exists an enumeration $R = \{r_i\}_{i=1}^\infty$ such that $\Delta(R_N) \to \infty$ with $R_N := \{r_1, \dots, r_N\}$. For the sake of brevity, let us identify R with $\mathbb{N}$ according to this enumeration. Define a probability measure μ on $\mathbb{N}$ by $\mu(\{k\}) := C2^{-k^2}$ with the normalizing constant $C := \left(\sum_{k=1}^\infty 2^{-k^2}\right)^{-1}$. As usual, let $(\xi_i)_{i=1}^\infty$ be a sequence of i.i.d. random elements of R with law μ and $(\hat{\xi}_j)_{j=1}^\infty$ be the corresponding sequence obtained from ξ by cancelling the repeating elements in the order of their appearance. It is easy to see that $\hat{\xi}_i = i$ for all $i \in \mathbb{N}$ if and only if $\xi_j \le \max_{1\le i<j} \xi_i + 1$ for all $j \in \mathbb{N}$, with a convention $\max \emptyset = 0$. Furthermore,

$$\begin{aligned} \mathbb{P}\left(\left\{\exists j \in \mathbb{N} : \xi_j > \max_{1\le i<j} \xi_i + 1\right\}\right) &\le \sum_{j=1}^\infty \mathbb{P}\left(\left\{\xi_j > \max_{1\le i<j} \xi_i + 1\right\}\right) \\ &= \sum_{j=1}^\infty \sum_{k=0}^\infty \mathbb{P}\left(\left\{\xi_j = k+2, \max_{1\le i<j} \xi_i \le k\right\}\right). \end{aligned} \tag{7.13}$$

Since all ξ_i are i.i.d., we have (with convention $0^0 = 1$ in the case $k = 0$, $j = 1$)

$$\mathbb{P}\left(\left\{\xi_j = k+2, \max_{1\le i<j} \xi_i \le k\right\}\right) = \mathbb{P}(\{\xi_1 = k+2\})\big(\mathbb{P}(\{\xi_1 \le k\})\big)^{j-1}.$$

Note that

$$\mathbb{P}(\{\xi_1 \le k\}) = 1 - \mathbb{P}(\{\xi_1 \ge k+1\}) \le 1 - \mathbb{P}(\{\xi_1 = k+1\}) = 1 - \mu(\{k+1\}).$$

Thus, we continue the chain of estimates (7.13) by

$$\begin{aligned}\mathbb{P}\left(\left\{\exists j\in\mathbb{N}:\xi_j>\max_{1\le i<j}\xi_i+1\right\}\right)&\le\sum_{j=1}^{\infty}\sum_{k=0}^{\infty}\mu(\{k+2\})\big(1-\mu(\{k+1\})\big)^{j-1}\\&=\sum_{k=0}^{\infty}\mu(\{k+2\})\sum_{j=1}^{\infty}\big(1-\mu(\{k+1\})\big)^{j-1}\\&=\sum_{k=0}^{\infty}\frac{\mu(\{k+2\})}{\mu(\{k+1\})}=\sum_{k=0}^{\infty}2^{(k+1)^2-(k+2)^2}\\&=\sum_{k=0}^{\infty}2^{-2k-3}=\frac{1}{6}.\end{aligned}$$

Therefore, with probability at least $\frac{5}{6}$, one has $\xi_j\le\max_{1\le i<j}\xi_i+1$ for all $j\in\mathbb{N}$, hence $\hat{\xi}_i\equiv i$, and thus

$$\lim_{m\to\infty}\Delta\big((\xi_i)_{i=1}^m\big)=\lim_{N\to\infty}\Delta\big((\hat{\xi}_i)_{i=1}^N\big)=\lim_{N\to\infty}\Delta(R_N)=\infty.$$

Appendix A. Auxiliary lemmas on operators and their spectra

We collect here some technical results on the spectra of operators on Hilbert spaces used throughout the paper.

Lemma A.1. *Let H be a Hilbert space.*

(i) *If A, B are compact self-adjoint operators on H, then*

$$s_\pm(A+B)\le s_\pm(A)+s_\pm(B).$$

(ii) *If A and $\{A_n\}_{n\ge1}$, are compact self-adjoint operators on H such that $A_n\to A$ in operator norm, then*

$$s_\pm(A)\le\liminf s_\pm(A_n).$$

Proof. We prove the claims for s_+. The proof for s_- is identical.

We start with the proof of (i). Assume that $s_+(A)<\infty$ and $s_+(B)<\infty$, otherwise the claim is trivial. Suppose that $s_+(A+B)>s_+(A)+s_+(B)$. Then there are at least $m+1=s_+(A)+s_+(B)+1$ linearly independent eigenvectors $u_1,\dots,u_m$ of $A+B$ corresponding to positive eigenvalues. Let L_A (resp., L_B) be the linear span of all eigenvectors of A (resp., B) corresponding to positive eigenvalues. Clearly, we have that $\dim L_A=s_+(A)$ and $\dim L_B=s_+(B)$. Then $\dim(L_A+L_B)\le s_A(A)+s_+(B)<m$ and there is a non-zero vector $u\in\operatorname{span}\{u_i\}_{i=1}^m\cap(L_A+L_B)^\perp$. Since A and B are self-adjoint, this yields $(u,Au)\le0$, $(u,Bu)\le0$, hence $(u,(A+B)u)\le0$ what leads to a contradiction. Therefore, $s_+(A+B)\le s_+(A)+s_+(B)$ as claimed.

Now, let us prove (ii). Let $u_1,\dots,u_m$ be unit norm eigenvectors of A corresponding to positive eigenvalues $\lambda_1\ge\lambda_2\ge\dots\ge\lambda_m>0$. Define symmetric matrices $\bar{A}:=((u_i,Au_j))_{i,j=1}^m=\operatorname{diag}(\lambda_1,\dots,\lambda_m)$ and $\bar{A}_n:=((u_i,A_nu_j))_{i,j=1}^m$. Then for $\|A_n-A\|<\lambda_m$, we have

$$\lambda_{\min}(\bar{A}_n)\ge\lambda_{\min}(\bar{A})+\lambda_{\min}(\bar{A}_n-\bar{A})\ge\lambda_m-\|\bar{A}_n-\bar{A}\|\ge\lambda_m-\|A_n-A\|>0.$$

In this case, $s_+(A_n)\ge s_+(\bar{A}_n)\ge m$, therefore,

$$m\le\liminf_n s_+(A_n).$$

Taking supremum over all $m\le s_+(A)$, we get the result. □

Lemma A.2. *Let (X, d) be a metric space, μ be a Borel probability measure, and $A\colon L^2(X,\mu) \to L^2(X,\mu)$ be an integral linear compact operator defined by the formula*

$$(Au)(x) := \int_X a(x,y)u(y)\,d\mu(y)$$

with $a \in L^2(X \times X, \mu \otimes \mu)$ satisfying $a(x,y) = a(y,x)$ for $\mu\otimes\mu$-a.e. $(x,y) \in X \times X$, so that A is self-adjoint. Let $\lambda_i \in \mathbb{R}$ and $u_i \in L^2(X,\mu)$ be an enumeration of its eigenvalues (counting multiplicities) and the respective eigenfunctions of unit norm in $L^2(X,\mu)$. Then

$$a(x,y) = \sum_i \lambda_i u_i(x) u_i(y), \tag{A.1}$$

the equality in the sense of $L^2(X \times X, \mu \otimes \mu)$.

Proof. By the Hilbert–Schmidt theorem, $\{u_i\}$ form a topological basis in $L^2(X,\mu)$. Thus representing a in terms of the Fourier series in $u_i \otimes u_j$ in $L^2(X \times X, \mu \otimes \mu)$, we get

$$a(x,y) = \sum_{i,j} c_{ij} u_i(x) u_i(y).$$

For the Fourier coefficients c_{ij}, we have

$$\begin{aligned} c_{ij} &= \int_X \left(\int_X a(x,y) u_i(u)\,d\mu(y) \right) u_j(x)\,d\mu(x) = \int_X \lambda_i u_i(x) u_j(x)\,d\mu(x) \\ &= \lambda_i \delta_{ij}, \end{aligned}$$

where δ_{ij} stands for the Kronecker delta, which implies (A.1). □

The following lemmas have been used in the proof of Proposition 4.11.

Lemma A.3. *Let (X, d) be a metric space, μ be a Borel probability measure, and $A\colon L^2(X,\mu) \to L^2(X,\mu)$ be an integral linear operator defined by the formula*

$$(Au)(x) := \int_X a(x,y)u(y)\,d\mu(y)$$

with a continuous function $a \in L^2(X \times X, \mu \otimes \mu)$ satisfying $a(x,y) = a(y,x)$. The following assertions hold.

(i) *For every finite subset $X_N = \{x_1, \dots, x_N\} \subset \operatorname{supp}\mu$, one has*

$$s_\pm(A) \geq s_\pm(A(X_N)), \tag{A.2}$$

$$s_\pm(P_\mu A P_\mu) \geq s_\pm(\Pi_N A(X_N) \Pi_N), \tag{A.3}$$

where $A(X_N) := (a(x_i, x_j))_{i,j=1}^N \in \mathbb{R}^{N\times N}$.

(ii) *If μ is a Radon measure (which is automatically guaranteed when (X, d) is Polish), then there is a sequence of finite subsets $X_N \subset X$ such that*

$$s_\pm(A) \leq \liminf_N s_\pm(A(X_N)), \tag{A.4}$$

$$s_\pm(P_\mu A P_\mu) \leq \liminf_N s_\pm(\Pi_N A(X_N) \Pi_N). \tag{A.5}$$

Proof. To prove (i), let $\varepsilon > 0$ be such that all the positive eigenvalues of $A(X_N)$ belong to the interval $[\epsilon, +\infty)$. Choose an $r > 0$ sufficiently small such that all open

balls $B_i := B_r(x_i)$, $i = 1, \dots, N$, are pairwise disjoint, and for all $i, j = 1, \dots, N$ the estimate

$$\big|a(x_i', x_j') - a(x_i, x_j)\big| \le \frac{\varepsilon}{2}, \qquad \text{for all } x_i' \in B_i,\ x_j' \in B_j. \tag{A.6}$$

holds. Define

$$v_i := \frac{1}{\mu(B_i)} \mathbf{1}_{B_i} \in L^2(X, \mu), \qquad i = 1, \dots, N,$$

(recall that $\mu(B_i) > 0$ since $x_i \in \operatorname{supp} \mu$). Let $Q_N \colon L^2(X, \mu) \to L^2(X, \mu)$ be the orthogonal projector onto $L_N := \operatorname{span}\{v_i\}_{i=1}^N$. Define the linear operator $\tilde{A}_N \colon L^2(X, \mu) \to L^2(X, \mu)$ by

$$(\tilde{A}_N u)(x) := \sum_{i,j=1}^N \int_X a(x_i, x_j) \mathbf{1}_{B_i}(x) \mathbf{1}_{B_j}(y) u(y)\, d\mu(y).$$

For every $u \in L_N$, setting $U := \bigcup_{i=1}^N B_i$, we have

$$\|\mathbf{1}_U A u - \tilde{A}_N u\|_2 \le \|\mathbf{1}_U A u - \tilde{A}_N u\|_\infty \le \frac{\varepsilon}{2} \|u\|_1 \le \frac{\varepsilon}{2} \|u\|_2, \tag{A.7}$$

where the first and the last estimates are just Höder inequality minding that μ is a probability measure, and the second one is due to (A.6). Since $Q_N v = Q_N(\mathbf{1}_U v)$ for all $v \in L^2(X, \mu)$, then from (A.7) we get that the operator norm of $Q_N A Q_N - \tilde{A}_N = Q_N(A - \tilde{A}_N) Q_N$ is bounded by $\frac{\varepsilon}{2}$. Now, let us define the embedding operator $E_N \colon \mathbb{R}^N \to L_N$ as $E_N x := \sum_{i=1}^N x_i v_i$, so that $\operatorname{Im} E_N = L_N$ and

$$(E_N^* v_i)_k = \frac{1}{\mu(B_i)} \delta_{ik}$$

with δ_{ik} standing for the Kronecker delta. Since

$$\begin{aligned} \tilde{A}_N(v_k)(x) &= \sum_{i=1}^N \mathbf{1}_{B_i}(x) \sum_{j=1}^N \int_X a(x_i, x_j) 1_{B_j}(y) \frac{1}{\mu(B_k)} \mathbf{1}_{B_k}(y) d\mu(y), \\ &= \sum_{i=1}^N \mathbf{1}_{B_i}(x) a(x_i, x_k) = \sum_{i=1}^N a(x_i, x_k) \mu(B_i) v_i, \end{aligned}$$

then $E_N^* \tilde{A}_N E_N = A(X_N)$. We have therefore

$$\begin{aligned} s_+(A) &\ge s_+(Q_N A Q_N) \\ &\qquad \text{by Lemma A.4(i) with } Q := Q_N,\ B := Q_N A Q_N,\ H_1 = H_2 = L^2(X, \mu) \\ &\ge s_+(\tilde{A}_N) \quad \text{by the Weyl inequality} \\ &= s_+(E_N^* \tilde{A}_N E_N) \quad \text{by Lemma A.4(iii) since } (\ker \tilde{A}_N)^\perp \subset L_N = \operatorname{Im} E_N \\ &= s_+(A(X_N)). \end{aligned}$$

Now, let $\bar{Q}_N \colon L^2(X, \mu) \to L^2(X, \mu)$ be the orthogonal projector onto $L_N \cap \mathbf{1}^\perp$. In the same way as above, $\bar{Q}_N(A - \tilde{A}_N)\bar{Q}_N$ is bounded by $\frac{\varepsilon}{2}$. Furthermore,

$$(\mathbf{1}, E_N x) = \sum_{i=1}^N (\mathbf{1}, v_i) x_i = (\mathbf{1}_N, x),$$

hence $\mathrm{Im}(E_N\Pi_N) = L_N \cap \mathbf{1}^\perp$ and $\bar{Q}_N E_N \Pi_N = E_N \Pi_N$. Therefore,

$$\begin{aligned} s_+(P_\mu A P_\mu) &\geq s_+(\bar{Q}_N A \bar{Q}_N) \quad \text{by Lemma A.4 since } P_\mu \bar{Q}_N = \bar{Q}_N \\ &\geq s_+(\bar{Q}_N \tilde{A}_N \bar{Q}_N) \quad \text{by the Weyl inequality} \\ &\geq s_+(\Pi_N E_N^* \tilde{A}_N E_N \Pi_N) \quad \text{by Lemma A.4 since } \bar{Q}_N E_N \Pi_N = E_N \Pi_N \\ &= s_+(\Pi_N A(X_N) \Pi_N). \end{aligned}$$

The proofs of $s_-(A) \geq s_-(A(X_N))$ and $s_-(P_\mu A P_\mu) \geq s_-(\Pi_N A(X_N)\Pi_N)$ are completely symmetric.

We now prove (ii). To this aim, for an arbitrary $\varepsilon > 0$, using that μ is the Radon measure, we can find a compact $K \subset X$ with $\mu(X \setminus K)$ be so small that

$$\int_{(X\times X)\setminus(K\times K)} a^2(x,y)\, d\mu(x) d\mu(y) < \varepsilon^2.$$

This can be done in view of by absolute continuity of the integral and of the fact that

$$\begin{aligned} \mu((X\times X)\setminus(K\times K)) &= \mu(((X\setminus K)\times X)\cup(K\times(X\setminus K))) \\ &\leq 2\mu(X\setminus K). \end{aligned}$$

We find a $\delta > 0$, a δ-net $X_N := \{x_1, \dots, x_N\} \subset K$, and a partition of K into N disjoint Borel subsets $B_1, \dots, B_N$ such that

$$\big|a(x_i', x_j') - a(x_i, x_j)\big| < \varepsilon, \qquad \text{for all } x_i' \in B_i,\ x_j' \in B_j \tag{A.8}$$

(e.g., B_i being a "Voronoi cell" for x_i).

Much similar to the prof of (i), let $Q_N \colon L^2(X,\mu) \to L^2(X,\mu)$ be the orthogonal projection onto $\mathrm{span}\{\mathbf{1}_{B_i}\}_{i=1}^N$ and denote by $\tilde{A}(X_N) \colon L^2(X,\mu) \to L^2(X,\mu)$ the linear operator defined by the formula

$$(\tilde{A}_N u)(x) := \sum_{i,j=1}^N \int_X a(x_i,x_j)\mathbf{1}_{B_i}(x)\mathbf{1}_{B_j}(y)u(y)\, d\mu(y).$$

For every $u \in L^2(X,\mu)$, we have

$$\|Q_N A Q_N u - \tilde{A}_N u\|_2 \leq \|Q_N A Q_N u - \tilde{A}_N u\|_\infty \leq \varepsilon \|u\|_1 \leq \varepsilon \|u\|_2,$$

and

$$\begin{aligned} \|Au - Q_N A Q_N u\|_2^2 &\leq \int_X d\mu(x) \left(\int_X |a(x,y) - \mathbf{1}_K(x) a(x,y) \mathbf{1}_K(y)| \cdot |u(y)|\, d\mu(y) \right)^2 \\ &\leq \|u\|_2^2 \int_X d\mu(x) \left(\int_X |a(x,y) - \mathbf{1}_K(x) a(x,y) \mathbf{1}_K(y)|^2\, d\mu(y) \right) \\ &\leq \varepsilon^2 \|u\|_2^2. \end{aligned}$$

This implies

$$\|Au - \tilde{A}_N u\|_2 \leq \|Au - Q_N A Q_N u\|_2 + \|Q_N A Q_N u - \tilde{A}_N u\|_2 \leq 2\varepsilon \|u\|_2.$$

In other words, we constructed a sequence of $X_N \subset X$ such that $\tilde{A}_N \to A$ as linear bounded operators over $L^2(X,\mu)$ as $N \to \infty$. This implies

$$s_\pm(A) \leq \liminf_N s_\pm(\tilde{A}_N)$$

by Lemma A.1(ii). In the same way one can show that

$$s_\pm(P_\mu \tilde{A}_N P_\mu) \le \liminf_N s_\pm(\Pi_N A(X_N)\Pi_N),$$

hence concluding the proof. □

Lemma A.4. *Let H_1, H_2 be Hilbert spaces, A and B be linear compact self-adjoint operators over H_2 and H_1, respectively, such that $B = Q^*AQ$, where $Q\colon H_1 \to H_2$ is a linear bounded operator. Then*

(i) $s_\pm(B) \le s_\pm(A)$.

(ii) *If, moreover, the codimension of the image of Q is $k \in \mathbb{N}$, then*

$$s_\pm(B) \ge s_\pm(A) - k.$$

(iii) *If $(\operatorname{Im} Q)^\perp \cap (\ker A)^\perp = \{0\}$, then $s_\pm(A) = s_\pm(B)$.*

Proof. Let $\{u_j\}_{j=1}^n \subset H_1$ be a set of linearly independent eigenvectors of B corresponding to strictly positive eigenvalues. The vectors Qu_j, $j = 1, \dots, n$ are also linearly independent since otherwise $\ker B \cap \operatorname{span}\{u_j\}_{j=1}^n \ne \{0\}$ which is impossible. Hence, $(Av, v) > 0$ for every $v \in \operatorname{span}\{Qu_j\}_{j=1}^n$, $v \ne 0$. This implies $s_+(A) \ge n$, and therefore, taking the supremum over all $n \in \mathbb{N}$ such that $n \le s_+(B)$, we get $s_\pm(B) \le s_\pm(A)$, proving Claim (i).

We now prove (ii). Let $\{w_j\}_{j=1}^m \subset H_2$ be a set of linearly independent eigenvectors of A corresponding to strictly positive eigenvalues. Then

$$\dim\left(\operatorname{span}\{w_j\}_{j=1}^m \cap \operatorname{Im} Q\right) \ge m - k.$$

Thus we may choose at least $m - k$ linearly independent elements

$$\{v_j\}_{j=1}^{m-k} \subset \operatorname{span}\{w_j\}_{j=1}^m \cap \operatorname{Im} Q$$

Pick then linearly independent $u_j \in H_1$ with $Qu_j = v_j$, $j = 1, \dots, m-k$. One has

$$(u, Bu) = (Qu, AQu) > 0$$

for all $u \in \operatorname{span}\{u_j\}_{j=1}^{m-k}$ unless $u = 0$. Thus implies $s_+(B) \ge m - k$ and taking the supremum over all $m \in \mathbb{N}$ satisfying $m \le s_+(A)$, we have $s_\pm(B) \ge s_\pm(A) - k$, i.e., Claim (ii).

Finally, to prove Claim (iii), we first note that for the orthogonal projector P_A from H_2 onto $(\ker A)^\perp$ one has $P_A A = A$. Hence, $A = A^* = A^* P_A^* = AP_A$, and this again implies

$$A = P_A A = P_A A P_A.$$

Therefore,

$$B = Q^*AQ = Q^*P_A A P_A Q = Q^* P_A^* A P_A Q = (P_A Q)^* A (P_A Q). \tag{A.9}$$

Furthermore,

$$(\operatorname{Im} Q)^\perp \cap (\ker A)^\perp = (\operatorname{Im} Q + \ker A)^\perp = \{0\},$$

hence $\operatorname{Im} Q + \ker A$ is dense in H_2, thus $\operatorname{Im}(P_A Q)$ is dense in $\operatorname{Im} P_A = (\ker A)^\perp$. Let again $\{w_j\}_{j=1}^m \subset H_2$ be a set of linearly independent eigenvectors of A corresponding to strictly positive eigenvalues, thus the matrix $(Aw_i, w_j)_{i,j=1}^m$ is positive definite. Since all w_j clearly belong to $(\ker A)^\perp$, then one can find $\{\tilde{w}_j\}_{j=1}^m \subset \operatorname{Im}(P_A Q)$ with each $\tilde{w}_j$ so close to w_j that they are all still linearly independent and the matrix $(A\tilde{w}_i, \tilde{w}_j)_{i,j=1}^m$ is positive definite (which gives in particular $(A\tilde{w}, \tilde{w}) > 0$ for all $\tilde{w} \in \operatorname{span}\{\tilde{w}_j\}_{j=1}^m$, $\tilde{w} \ne 0$).

Choosing $u_j \in H$ linearly independent such that $P_A Q u_j = \tilde{w}_j$, in view of (A.9) we get

$$(u, Bu) = ((P_A Q)u, A(P_A Q)u) > 0$$

for all $u \in \operatorname{span}\{u_j\}_{j=1}^m$, $u \neq 0$. Thus, $s_+(B) \geq m$ and taking the supremum over all $m \in \mathbb{N}$ satisfying $m \leq s_+(A)$, we have $s_\pm(B) \geq s_\pm(A)$. This together with Claim (i) gives Claim (iii). □

Lemma A.5. *Let $\Sigma = \{z_i\}_{i=1}^3 \subset \mathbb{R}^{n,p}$ be such that $\Sigma - \Sigma \subset C_{n,p}$ (thus, $d_{n,p}(z_i, z_j)$ is defined). Then the strict triangle inequality*

$$d_{n,p}(z_1, z_3) < d_{n,p}(z_1, z_2) + d_{n,p}(z_2, z_3)$$

is equivalent to the strict one-sided Cauchy-Schwarz inequality

$$(z_1 - z_2, z_2 - z_3)_{n,p} < d_{n,p}(z_1, z_2) d_{n,p}(z_2, z_3).$$

Proof. The strict triangle inequality is clearly equivalent to

$$d_{n,p}^2(z_1, z_3) < d_{n,p}^2(z_1, z_2) + 2 d_{n,p}(z_1, z_2) d_{n,p}(z_2, z_3) + d_{n,p}^2(z_2, z_3).$$

Since

$$\begin{aligned} d_{n,p}^2(z_1, z_3) &:= (z_1 - z_3, z_1 - z_3) \\ &= (z_1 - z_2, z_1 - z_2) + 2(z_1 - z_2, z_2 - z_3) + (z_2 - z_3, z_2 - z_3) \\ &= d_{n,p}^2(z_1, z_2) + 2(z_1 - z_2, z_2 - z_3) + d_{n,p}^2(z_2, z_3), \end{aligned}$$

we get the result. □

References

[1] A. Y. Alfakih et al. *Euclidean distance matrices and their applications in rigidity theory*. Springer, 2018.

[2] Z. Bai and J. W. Silverstein. *Spectral analysis of large dimensional random matrices*, volume 20. Springer, 2010.

[3] E. Bogomolny, O. Bohigas, and C. Schmit. Distance matrices and isometric embeddings. *J. Math. Phys. Anal. Geom.*, 4(1):7–23, 202, 2008.

[4] C. Bordenave. Eigenvalues of Euclidean random matrices. *Random Structures Algorithms*, 33(4):515–532, 2008.

[5] D. Burago, Y. Burago, S. Ivanov, et al. *A course in metric geometry*, volume 33. American Mathematical Society Providence, 2001.

[6] Z. B. Charles, M. Farber, C. R. Johnson, and L. Kennedy-Shaffer. Nonpositive eigenvalues of hollow, symmetric, nonnegative matrices. *SIAM Journal on Matrix Analysis and Applications*, 34(3):1384–1400, 2013.

[7] R. L. Graham, J. Nešetřil, and S. Butler. *The Mathematics of Paul Erdős II*. Springer, 2013.

[8] M. Gromov. *Metric structures for Riemannian and non-Riemannian spaces*, volume 152 of *Progress in Mathematics*. Birkhäuser Boston, Inc., Boston, MA, 1999. Based on the 1981 French original [MR0682063 (85e:53051)], With appendices by M. Katz, P. Pansu and S. Semmes, Translated from the French by Sean Michael Bates.

[9] E. V. Haynsworth. Determination of the inertia of a partitioned Hermitian matrix. *Linear algebra and its applications*, 1(1):73–81, 1968.

[10] T. Jiang. Distributions of eigenvalues of large Euclidean matrices generated from l_p balls and spheres. *Linear Algebra Appl.*, 473:14–36, 2015.

[11] V. Koltchinskii and E. Giné. Random matrix approximation of spectra of integral operators. *Bernoulli*, 6(1):113–167, 2000.

[12] A. Kroshnin, E. Stepanov, and D. Trevisan. Infinite multidimensional scaling for metric measure spaces. *ESAIM: Control, Optimisation and Calculus of Variations*, 28:58, 2022.

[13] J. McKee, C. Smyth, J. McKee, and C. Smyth. Symmetrizable matrices i: Introduction. *Around the Unit Circle: Mahler Measure, Integer Matrices and Roots of Unity*, pages 301–315, 2021.

[14] A. M. Vershik. The universal Uryson space, Gromov's metric triples, and random metrics on the series of natural numbers. *Uspekhi Mat. Nauk*, 53(5(323)):57–64, 1998.
[15] A. M. Vershik. Random and universal metric spaces. In *Dynamics and randomness II*, volume 10 of *Nonlinear Phenom. Complex Systems*, pages 199–228. Kluwer Acad. Publ., Dordrecht, 2004.
[16] J. Wang. *Geometric structure of high-dimensional data and dimensionality reduction*. Springer, 2012.
[17] X. Zeng. Distributions for the eigenvalues of large random matrices generated from four manifolds. *J. Phys. A*, 47(2):025206, 24, 2014.
[18] X. Zeng. Spectrum of large euclidean random matrices generated from l_p ellipsoids. *Linear Multilinear Algebra*, 70(8):1403–1422, 2022.

The Weierstrass Institute for Applied Analysis and Stochastics, Mohrenstrasse 39, 10117 Berlin, Germany
Email address: alex.kroshnin@gmail.com

Faculty of Mathematics, HSE University, Usacheva ul. 6, 119048 Moscow, Russian Federation
Email address: tma@hse.ru

Dipartimento di Matematica, Università di Pisa, Largo Bruno Pontecorvo 5, 56127 Pisa, Italy and St.Petersburg Branch of the Steklov Mathematical Institute of the Russian Academy of Sciences, St.Petersburg, Russian Federation and HSE University, Moscow, Russian Federation
Email address: stepanov.eugene@gmail.com