# Measure Aspects of Cut Polyhedra: $\ell_1$-embeddability and Probability

Michel DEZA
Monique LAURENT

Laboratoire d'Informatique, URA 1327 du CNRS
Département de Mathématiques et d'Informatique
Ecole Normale Supérieure

LIENS - 93 - 17

September 1993

# Measure aspects of cut polyhedra: $\ell_1$-embeddability and probability

Michel Deza and Monique Laurent

LIENS - Ecole Normale Supérieure
45 rue d'Ulm
75230 Paris cedex 05, France

# 1 Preliminaries

We recall in this Section all the definitions that we need for this Chapter and, in particular, the definitions about distance spaces, isometric embeddings, measure spaces, and our main host spaces, namely, the Banach $\ell_p$- and $L_p$-spaces for $1 \leq p \leq \infty$.

### 1.0.1 Distance spaces and $\ell_p$-spaces

Let $X$ be a set. A function $d : X \times X \to \mathbb{R}_+$ is called a **distance** on $X$ if $d$ is **symmetric**, i.e., satisfies $d(i,j) = d(j,i)$ for all $i,j \in X$, and if $d(i,i) = 0$ holds for all $i \in X$. Then, $(X,d)$ is called a **distance space**. If $d$ satisfies, in addition, the following inequalities

$$d(i,j) \leq d(i,k) + d(j,k) \quad \text{for all } i,j,k \in X, \tag{1.1}$$

called **triangle inequalities**, then $d$ is called a **semimetric** on $X$. Moreover, if $d(i,j) = 0$ holds only for $i = j$, then $d$ is a **metric** on $X$.

Suppose $d$ is a distance on the set $V_n = \{1,\ldots,n\}$. Set $E_n = \{ij : i,j \in V_n, i \neq j\}$, where the symbol $ij$ denotes the unordered pair of the integers $i,j$, i.e., $ij$ and $ji$ are considered identical. Because of symmetry and since $d(i,i) = 0$ for $i \in V_n$, we can view the distance $d$ as a vector $d = (d_{ij})_{1 \leq i < j \leq n} \in \mathbb{R}^{E_n}$ and, vice versa, each vector $d \in \mathbb{R}^{E_n}$ yields a symmetric function that is zero on the main diagonal. We will use both representations as a function on $V_n \times V_n$ or as a vector of $\mathbb{R}^{E_n}$ for a distance on $V_n$.

Given a normed space $(E, \|\ .\ \|)$, a metric $d_{\|.\|}$ is defined on $E$, called **norm** or **Minkowski metric**, by setting

$$d_{\|.\|}(x,y) = \| x - y \|$$

for all $x,y \in E$.

We will consider, in particular, the norm metric, denoted by $d_{\ell_p}$ and called the $\ell_p$-**metric**, of the Banach $\ell_p$-space $(\mathbb{R}^m, \|\ .\ \|_p)$ for $p \geq 1$. Recall that

$$\| x \|_p = \Big( \sum_{1 \leq k \leq m} |x_k|^p \Big)^{\frac{1}{p}}$$

for all $x \in \mathbb{R}^m$. The metric space $(\mathbb{R}^m, d_{\ell_p})$ is denoted by $\ell_p^m$. Similarly, $\ell_\infty^m$ denotes the metric space $(\mathbb{R}^m, d_{\ell_\infty})$, where $d_{\ell_\infty}$ denotes the norm metric associated with the norm $\|\ .\ \|_\infty$ which is defined by

$$\| x \|_\infty = \max(|x_k| : 1 \leq k \leq m),$$

for all $x \in \mathbb{R}^m$.

For $1 \leq p < \infty$, the metric space $\ell_p^\infty$ consists of the set of infinite sequences $x = (x_i)_{i\geq 0} \in \mathbb{R}^{\mathbb{N}}$ for which the sum $\sum_{i\geq 0} |x_i|^p$ is finite, endowed with the distance $d(x,y) = \left(\sum_{i\geq 0} |x_i - y_i|^p\right)^{\frac{1}{p}}$. In the same way $\ell_\infty^\infty$ is the set of bounded infinite sequences $x \in \mathbb{R}^{\mathbb{N}}$, endowed with the distance $d(x,y) = \max(|x_i - y_i| : i \geq 0)$.

If $(X,d)$ and $(X',d')$ are two distance spaces, $(X,d)$ is said to be **isometrically embeddable** into $(X',d')$ if there exists a map $\phi$ (the **embedding**) from $X$ to $X'$ such that $d(x,y) = d'(\phi(x),\phi(y))$ for all $x,y \in X$. One says also that $(X,d)$ is an **isometric subspace** of $(X',d')$. All the embeddings considered here are isometric, so we sometimes omit the adjective "isometric".

A distance space $(X,d)$ is said to be **$\ell_p$-embeddable** if $(X,d)$ is isometrically embeddable into the space $\ell_p^m$ for some integer $m \geq 1$. The smallest such integer $m$ is called the **$\ell_p$-dimension** of $(X,d)$ and is denoted by $m_p(X,d)$. Then, we denote by

$$m_p(n) = \max(m_p(X,d) : |X| = n \text{ and } (X,d) \text{ is } \ell_p\text{-embeddable}) \tag{1.2}$$

the minimum dimension $m$ such that each $\ell_p$-embeddable distance on $n$ points can be embedded in $\ell_p^m$. It is known that $m_p(n)$ is finite; in fact, $m_p(n) \leq \binom{n}{2}$ for all $n$ and $p$ (see Section 7.1).

$(X,d)$ is said to be **$\ell_p^\infty$-embeddable** if it is an isometric subspace of $\ell_p^\infty$.

We are interested here in the study of the distances spaces which can be isometrically embedded in one of the following host spaces: $\ell_p^m$, $\ell_p^\infty$, or $L_p(\Omega,\mathcal{A},\mu)$ (see the definition below) for $p \geq 1$ and we are mainly concerned with the cases $p = 1, 2$.

The case $p = 1$ is directly relevant to the central topic of this book. Indeed, the distances on $n$ points that are $\ell_1$-embeddable are precisely the members of the cut cone $\mathrm{CUT}_n$; see Theorem 2.11.

The case $p = 2$ is also closely related to our topic; see Section 4.3.

On the other hand, it is well known that the distances on $n$ points that are $\ell_\infty$-embeddable are precisely the semimetrics on $n$ points, i.e., the members of the semimetric cone $\mathrm{MET}_n$. To see it, note that if $d$ is a semimetric on the set $V_n = \{1,\ldots,n\}$, then the mapping $i \in V_n \mapsto (d(1,i), d(2,i), \ldots, d(n-1,i)) \in \mathbb{R}^{n-1}$ is an isometric embedding of $(V_n,d)$ into $\ell_\infty^{n-1}$. This shows that

$$m_\infty(n) \leq n-1. \tag{1.3}$$

We also consider isometric embeddings into the **hypercube metric space** $(\{0,1\}^m, d_{\ell_1})$, which is a subspace of $\ell_1^m$. Note that the hypercube metric space can also be defined as the graphic metric space $(V(H_m), d_{H_m})$, where $H_m$ denotes the 1-skeleton of the $m$-dimensional hypercube and $d_{H_m}$ is its path metric defined on the nodes of $H_m$. A distance space $(X, d)$ is said to be **hypercube embeddable** if it can be isometrically embedded in some hypercube metric space. Hence, each hypercube embeddable distance space is $\ell_1$-embeddable and, in fact, if $d$ is rational valued, then the space $(X, d)$ is hypercube embeddable if and only if $(X, \lambda d)$ is $\ell_1$-embeddable for some scalar $\lambda$; see Proposition 2.8.

### 1.0.2 Measure spaces and $L_p$-spaces

For defining the distance space $L_p(\Omega, \mathcal{A}, \mu)$, we need to recall some definitions on measure spaces. Let $\Omega$ be a set and let $\mathcal{A}$ be a **$\sigma$-algebra** of subsets of $\Omega$, i.e., $\mathcal{A}$ satisfies the following properties

$$\begin{cases} \Omega \in \mathcal{A}, \\ \text{if } A \in \mathcal{A} \text{ then } \Omega \setminus A \in \mathcal{A}, \\ \text{if } A = \bigcup_{1 \leq k \leq \infty} \text{ with } A_k \in \mathcal{A} \text{ for all } k, \text{ then } A \in \mathcal{A}. \end{cases}$$

A function $\mu : \mathcal{A} \longrightarrow \mathbb{R}_+$ is a **measure** on $\mathcal{A}$ if it is additive, i.e., $\mu(\bigcup_{k \geq 1} A_k) = \sum_{k \geq 1} \mu(A_k)$ for all pairwise disjoint sets $A_k \in \mathcal{A}$, and satisfies $\mu(\emptyset) = 0$. Note that measures are always assumed to be nonnegative. A **measure space** is a triple $(\Omega, \mathcal{A}, \mu)$ consisting of a set $\Omega$, a $\sigma$-algebra $\mathcal{A}$ of subsets of $\Omega$, and a measure $\mu$ on $\mathcal{A}$. A **probability space** is a measure space with total measure $\mu(\Omega) = 1$.

Given a function $f : \Omega \longrightarrow \mathbb{R}$, its **$L_p$-norm** is defined by

$$\| f \|_p = \left( \int_\Omega |f(\omega)|^p \mu(d\omega) \right)^{\frac{1}{p}}.$$

Then, $L_p(\Omega, \mathcal{A}, \mu)$ denotes the set of functions $f : \Omega \longrightarrow \mathbb{R}$ which satisfy $\| f \|_p < \infty$. The $L_p$-norm defines a metric structure on $L_p(\Omega, \mathcal{A}, \mu)$, namely, by taking $\| f - g \|_p$ as distance between two functions $f, g \in L_p(\Omega, \mathcal{A}, \mu)$.

A distance space $(X, d)$ is said to be **$L_p$-embeddable** if it is a subspace of $L_p(\Omega, \mathcal{A}, \mu)$ for some measure space $(\Omega, \mathcal{A}, \mu)$.

The most classical example of an $L_p$-space is the space $L_p(\Omega, \mathcal{A}, \mu)$, where $\Omega$ is the open interval $(0, 1)$, $\mathcal{A}$ is the family of Borel subsets of $(0, 1)$, and $\mu$ is the Lebesgue measure; it is simply denoted by $L_p(0, 1)$.

We now make precise the connections existing between $L_p$-spaces and $\ell_p$-spaces.

If $\Omega = \mathbb{N}$, $\mathcal{A} = 2^\Omega$ is the collection of all subsets of $\Omega$, and $\mu$ is the cardinality measure, i.e., $\mu(A) = |A|$ if $A$ is a finite subset of $\Omega$ and $\mu(A) = \infty$ otherwise, then $L_p(\mathbb{N}, 2^\mathbb{N}, |.|)$ coincides with the space $\ell_p^\infty$.

If $\Omega = V_m$ is a set of cardinality $m$, $\mathcal{A} = 2^\Omega$, and $\mu$ is the cardinality measure, then $L_p(V_m, 2^{V_m}, |.|)$ coincides with $\ell_p^m$.

In other words, $\ell_p^m$ is an isometric subspace of $\ell_p^\infty$ which, in turn, is $L_p$-embeddable.

Finally, we introduce one more semimetric space. Let $(\Omega, \mathcal{A}, \mu)$ be a measure space. Set $\mathcal{A}_\mu = \{A \in \mathcal{A} : \mu(A) < \infty\}$. One can define a distance $d_\mu$ on $\mathcal{A}_\mu$ by setting

$$d_\mu(A, B) = \mu(A \triangle B)$$

for all $A, B \in \mathcal{A}_\mu$. Then, $d_\mu$ is a semimetric on $\mathcal{A}_\mu$. We call it a **measure semimetric**, and the space $(\mathcal{A}_\mu, d_\mu)$ is called a **measure semimetric space**. The semimetric $d_\mu$ is also called a Fréchet-Nikodym-Aronszajn distance in the literature. We will consider in Section 6.3 the related Steinhaus distance, which is defined by

$$\frac{\mu(A \triangle B)}{\mu(A \cap B)}$$

for $A, B \in \mathcal{A}_\mu$. Note that the measure semimetric space $(\mathcal{A}_\mu, d_\mu)$ is the subspace of $L_1(\Omega, \mathcal{A}, \mu)$ consisting of its 0-1 valued functions. Moreover, if $\Omega = V_m$ is a finite set of cardinality $m$, $\mathcal{A} = 2^\Omega$, and $\mu$ is the cardinality measure, then the space $(\mathcal{A}_\mu, d_\mu)$ coincides with the hypercube metric space $(\{0,1\}^m, d_{\ell_1})$.

### 1.0.3 Finitude result for $L_p$-embeddability

Though we are mainly concerned with finite distance spaces, i.e., distance spaces $(X, d)$ with $X$ finite, we also present some results involving infinite distance spaces. For instance, we study in Section 5.1 the normed spaces whose norm metric is $L_1$-embeddable. However, the following fundamental result shows that the study of $L_p$-embeddable spaces can be reduced to the finite case.

THEOREM 1.4 [BCK66] *Let $p \geq 1$ and let $(X, d)$ be a distance space. Then, $(X, d)$ is $L_p$-embeddable if and only if every finite subspace of $(X, d)$ is $L_p$-embeddable.*

We consider in detail $L_1$-embeddable distance spaces in Section 2. In particular, we show that, for a finite distance space, the properties of being $\ell_1$-, $\ell_1^\infty$, or $L_1$-embeddable are all equivalent to the property of belonging to the cut cone (see Theorem 2.11).

Similar results are known for the case $p = 2$ (see, e.g., [WW75]). Namely, for a finite distance space $(X, d)$, the properties of being $\ell_2$-, $\ell_2^\infty$-, or $L_2$-embeddable are all equivalent (and, then, $(X, d)$ embeds in $\ell_2^{|X|-1}$; see relation (4.18)). Moreover, if $X$ is countable, then $(X, d)$ is $\ell_2^\infty$-embeddable (or, equivalently, $L_2$-embeddable) if and only if each subspace of $(X,d)$ on $n+1$ points embeds in $\ell_2^n$. A crucial result in the case $p=2$ is the well known result by Schoenberg [Sch38b], which shows that $L_2$-embeddable spaces can be characterized by the negative type inequalities; see Theorem 4.16 in Section 4. This contrasts with the case $p=1$ where no complete characterization by inequalities is known for $L_1$-embeddable spaces.

**Isometric embeddings among the $L_p$-spaces.**
There is a vast literature on the topic of isometric embeddings among the various $L_p$-spaces; see, e.g., [WW75, Dor76, Bal87, LV93]. We summarize here some of the main results.

Theorem 1.5 [Dor76] *Let $1 \leq p < \infty$, $1 \leq r \leq \infty$ and $m \in \mathbb{N}, m \geq 2$. Then, $\ell_r^m$ is an isometric subspace of $L_p(0,1)$ if and only if one of the following assertions holds.*
*(i) $p \leq r < 2$.*
*(ii) $r = 2$.*
*(iii) $m = 2$ and $p = 1$.*

Hence, for instance, $\ell_r^3$ does not embed isometrically in $L_p(0,1)$ if $r > 2$. It was already shown in [BCK66] that $L_p(0,1)$ embeds isometrically in $L_1(0,1)$ for all $1 \leq p \leq 2$.

As a reformulation of the above Theorem, we have the following implications for a distance space $(X,d)$.
- If $(X,d)$ is $\ell_p^2$-embeddable for some $1 \leq p \leq \infty$, then $(X,d)$ is $L_1$-embeddable.
- If $(X,d)$ is $\ell_p$-embeddable for some $1 \leq p \leq 2$, then $(X,d)$ is $L_1$-embeddable.
- If $(X,d)$ is $\ell_2$-embeddable, then $(X,d)$ is $L_p$-embeddable for all $1 \leq p \leq \infty$.

Let $r \neq p$ such that $1 \leq r, p < \infty$ and let $m \geq 1$ be an integer. Then, $\ell_r^m$ embeds isometrically in $\ell_p^n$ for some integer $n \geq 1$ if and only if $r = 2$ and $p$ is an even integer ([LV93]). Given an even integer $p$, we can define $N(m,p)$ as the smallest integer $n \geq 1$ for which $\ell_2^m$ embeds isometrically into $\ell_p^n$. It is shown in [LV93] that $N(2,p) = \frac{p}{2} + 1$ and that, for any $p \geq 2$ and $m \geq 1$, $\max(N(m-1,p), N(m,p-2)) \leq N(m,p) \leq \binom{m+p-1}{m-1}$. An exact evaluation of $N(m,p)$ is known for small values of $p, m$; for instance, $N(3,4) = 6, N(3,6) = 11, N(3,8) = 16, N(7,4) = 28, N(8,6) = 120, N(23,4) = 276, N(23,6) = 2300, N(24,10) = 98280$.
Therefore, given $r$ and $m \in \mathbb{N}$ such that $1 < r \leq 2 < m$, we have that $\ell_r^m$ does not embed isometrically into $\ell_1^n$ ($n$ positive integer), but $\ell_r^m$ embeds into $L_1(0,1)$ and, moreover, every finite subspace of $\ell_r^m$ on $s$ points embeds into $\ell_1^{\binom{s}{2}}$.

## 2 The cut cone and $\ell_1$-metrics

In this Section, we show how the members of the cut cone can be interpreted in terms of metrics and measure spaces. We essentially follow [Ass80b] and [AD82].

In order to make the Chapter self-contained, we recall the definition of the cut cone.

Given a subset $S$ of $V_n = \{1, \ldots, n\}$, $\delta(S)$ denotes the **cut semimetric** defined by

$$\delta(S)_{ij} = 1 \text{ if } |S \cap \{i,j\}| = 1, \text{ and } \delta(S)_{ij} = 0 \text{ otherwise,} \tag{2.1}$$

for all $1 \leq i < j \leq n$. Then, $\mathrm{CUT}_n$ denotes the cone in $\mathbb{R}^{E_n}$ generated by the cut semimetrics $\delta(S)$ for all subsets $S \subseteq V_n$, $\mathrm{CUT}_n$ is called the **cut cone** on $n$ points, and the **cut polytope** $\mathrm{CUT}_n^\square$ denotes the polytope in $\mathbb{R}^{E_n}$ whose vertices are the cut semimetrics $\delta(S)$ for all subsets $S$ of $V_n$. So,

$$\mathrm{CUT}_n = \{\sum_{S\subseteq V_n} \lambda_S \delta(S) : \lambda_S \geq 0 \text{ for all } S \subseteq V_n\}, \tag{2.2}$$

$$\mathrm{CUT}_n^\square = \{\sum_{S\subseteq V_n} \lambda_S \delta(S) : \sum_{S\subseteq V_n} \lambda_S = 1 \text{ and } \lambda_S \geq 0 \text{ for all } S \subseteq V_n\}. \tag{2.3}$$

If one considers an arbitrary finite set $X$ instead of $V_n$, then one defines similarly the cut cone $\mathrm{CUT}(X)$ and the cut polytope $\mathrm{CUT}^\square(X)$ on $X$. So, $\mathrm{CUT}(V_n) = \mathrm{CUT}_n$ and $\mathrm{CUT}^\square(V_n) = \mathrm{CUT}_n^\square$.

## 2.1 $\ell_1$-spaces (finite case)

Clearly, every member $d$ of the cut cone $\mathrm{CUT}_n$ defines a semimetric on $n$ points. Hence arises the question of characterizing the class of semimetrics that belong to the cut cone. Several equivalent characterizations are stated in Theorem 2.11. We now present several intermediate results.

PROPOSITION 2.4 *Let $d = (d_{ij})_{1\leq i<j\leq n} \in \mathbb{R}^{E_n}$. The following assertions are equivalent.*
*(i) $d \in CUT_n$ (resp. $d \in CUT_n^\square$).*
*(ii) There exist a measure space (resp. a probability space) $(\Omega, \mathcal{A}, \mu)$ and $A_1, \ldots, A_n \in \mathcal{A}$ such that $d_{ij} = \mu(A_i \triangle A_i)$ for all $1 \leq i < j \leq n$.*

PROOF. Assume $d \in \mathrm{CUT}_n$. Then, $d = \sum_{S\subseteq\{1,\ldots,n\}} \lambda_S \delta(S)$ for some $\lambda_S \geq 0$. We define a measure space $(\Omega, \mathcal{A}, \mu)$ as follows. Let $\Omega$ denote the family of subsets of $\{1, \ldots, n\}$, let $\mathcal{A}$ denote the family of subsets of $\Omega$ and let $\mu$ denote the measure on $\mathcal{A}$ defined by $\mu(A) = \sum_{S\in A} \lambda_S$ for each $A \in \mathcal{A}$ (i.e., $A$ is a collection of subsets of $\{1, \ldots, n\}$). Define $A_i = \{S \in \Omega : i \in S\}$. Then, $\mu(A_i \triangle A_j) = \mu(\{S \in \Omega : |S \cap \{i,j\}| = 1\})$ $= \sum_{S\in\Omega:|S\cap\{i,j\}|=1} \lambda_S = d_{ij}$ holds, for all $1 \leq i < j \leq n$. Moreover, if $d \in \mathrm{CUT}_n^\square$, then we have $\sum_S \lambda_S = 1$, i.e. $\mu(\Omega) = 1$, that is $(\Omega, \mathcal{A}, \mu)$ is a probability space.

Conversely, assume $d_{ij} = \mu(A_i \triangle A_j)$ for $1 \leq i < j \leq n$, where $(\Omega, \mathcal{A}, \mu)$ is a measure space and $A_1, \ldots, A_n \in \mathcal{A}$. Set $A^S = \bigcap_{i\in S} A_i \cap \bigcap_{i\notin S} (\Omega \setminus A_i)$ for each $S \subseteq \{1, \ldots, n\}$. Then, $A_i = \bigcup_{S:i\in S} A^S$, $A_i \triangle A_j = \bigcup_{S:|S\cap\{i,j\}|=1} A_S$ and $\Omega = \bigcup_S A^S$. Therefore, $d = \sum_{S\subseteq\{1,\ldots,n\}} \mu(A^S)\delta(S)$, showing that $d$ belongs to the cut cone $\mathrm{CUT}_n$. Moreover, if

$(\Omega, \mathcal{A}, \mu)$ is a probability space, i.e., $\mu(\Omega) = 1$, then $\sum_S \mu(A^S) = 1$, implying that $d$ belongs to the cut polytope $\mathrm{CUT}_n^\square$. ∎

Proposition 2.5 *Let $d \in \mathbb{R}^{E_n}$ and $(V_n, d)$ the associated distance space. The following assertions are equivalent.*
*(i) $d \in CUT_n$.*
*(ii) $(V_n, d)$ is $\ell_1$-embeddable, i.e., there exist $n$ vectors $u_1, \ldots, u_n \in \mathbb{R}^m$ for some $m$ such that $d_{ij} = \| u_i - u_j \|_1$ for all $1 \le i < j \le n$.*

Proof. $(i) \Rightarrow (ii)$. Suppose that $d \in \mathrm{CUT}_n$. Then, $d = \sum_{1 \le k \le m} \lambda_k \delta(S_k)$ with $\lambda_1, \ldots, \lambda_m \ge 0$. For $1 \le i \le n$, define the vector $u_i \in \mathbb{R}^m$ with components $(u_i)_k = \lambda_k$ if $i \in S_k$ and $(u_i)_k = 0$ otherwise, for $1 \le k \le m$. Then $d_{ij} = \| u_i - u_j \|_1$ holds, showing that $(V_n, d)$ is $\ell_1$-embeddable.
$(ii) \Rightarrow (i)$. Assume that $(V_n, d)$ is $\ell_1$-embeddable, i.e., there exist $n$ vectors $u_1, \ldots, u_n \in \mathbb{R}^m$ for some $m \ge 1$ such that $d_{ij} = \| u_i - u_j \|_1$, for $1 \le i < j \le n$. We show that $d \in \mathrm{CUT}_n$. It suffices to show the result for the case $m = 1$ by additivity of the $\ell_1$-norm. Hence, $d_{ij} = |u_i - u_j|$ where $u_1, \ldots, u_n \in \mathbb{R}$. Without loss of generality, we can suppose that $0 = u_1 \le u_2 \le \ldots \le u_n$. Then, $d = \sum_{1 \le k \le n-1} (u_{k+1} - u_k) \delta(\{1, 2, \ldots, k-1, k\})$ holds, showing that $d \in \mathrm{CUT}_n$. ∎

Remark 2.6 The proof of Proposition 2.5 shows that, if $d$ is a distance on $V_n$, then $d$ is $\ell_1^m$-embeddable whenever $d$ can be decomposed as a nonnegative combination of $m$ cut semimetrics.

There is a characterization for hypercube embeddable semimetrics analogue to that of Proposition 2.5.

Proposition 2.7 *Let $d \in \mathbb{R}^{E_n}$ and $(V_n, d)$ be the associated distance space. The following assertions are equivalent.*
*(i) $d = \sum_S \lambda_S \delta(S)$ for some nonnegative integer scalars $\lambda_S$.*
*(ii) $(V_n, d)$ is hypercube embeddable, i.e., there exist $n$ vectors $u_1, \ldots, u_n \in \{0, 1\}^m$ for some $m$ such that $d_{ij} = \| u_i - u_j \|_1$ for all $1 \le i < j \le n$.*
*(iii) There exist a finite set $\Omega$ and $n$ subsets $A_1, \ldots, A_n$ of $\Omega$ such that $d_{ij} = |A_i \triangle A_j|$ for all $1 \le i < j \le n$.*
*(iv) $(V_n, d)$ is an isometric subspace of $(\mathbb{Z}^m, d_{\ell_1})$ for some integer $m \ge 1$.*

PROOF. The proof of $(i) \iff (ii)$ is analogous to that of Proposition 2.5. Namely, for $(i) \Longrightarrow (ii)$, assume $d = \sum_{1 \leq k \leq m} \delta(S_k)$ (allowing repetitions). Consider the binary $n \times m$ matrix $M$ whose columns are the incidence vectors of the sets $S_1, \ldots, S_m$. If $u_1, \ldots, u_n$ denote the rows of $M$, then $d_{ij} = \| u_i - u_j \|_1$ holds, providing an embedding of $(V_n, d)$ in the hypercube of dimension $m$. Conversely, for $(ii) \Longrightarrow (i)$, consider the matrix $M$ whose rows are the $n$ given vectors $u_1, \ldots, u_n$. Let $S_1, \ldots, S_m$ be the subsets of $\{1, \ldots, n\}$ whose incidence vectors are the columns of $M$. Then, $d = \sum_{1 \leq k \leq m} \delta(S_k)$ holds, giving a decomposition of $d$ as a nonnegative integer combination of cuts.
$(iii)$ is a reformulation of $(ii)$, $(iii) \Longrightarrow (iv)$ is obvious, and $(iv) \Longrightarrow (i)$ follows from the proof of the implication $(ii) \Longrightarrow (i)$ of Proposition 2.5. ∎

The next result follows immediately from Propositions 2.5 and 2.7.

PROPOSITION 2.8 *Let $(V_n, d)$ be a distance space where $d$ is rational valued. Then, $(V_n, d)$ is $\ell_1$-embeddable if and only if $(V_n, \lambda d)$ is hypercube embeddable for some scalar $\lambda$.*

The equivalence $(i) \iff (ii)$ from Proposition 2.7 can be generalized in the context of Hamming spaces and multicuts.
Recall that, given $x, y \in \mathbb{R}^k$, their **Hamming distance** $d_H(x, y)$ is defined as the number of positions where the coordinates of $x$ and $y$ differ. Hence, when considered on binary vectors, the Hamming distance coincides with the $\ell_1$-distance.
Let $q \geq 2$ be an integer and let $S_1, \ldots, S_q$ be $q$ subsets of $V_n$ that partition $V_n$. Then, the **multicut vector** $\delta(S_1, \ldots, S_q)$ is the vector of $\mathbb{R}^{E_n}$ defined by
$\delta(S_1, \ldots, S_q)_{ij} = 0$ if $i, j \in S_h$ for some $h$, $1 \leq h \leq p$,
$\delta(S_1, \ldots, S_q)_{ij} = 1$ otherwise
for $1 \leq i < j \leq n$.

PROPOSITION 2.9 *Let $d \in \mathbb{R}^{E_n}$ and $(V_n, d)$ be the associated distance space. The following assertions are equivalent.*
*(i) $d = \sum_{(S_1, \ldots, S_q) \text{ partition of } V_n} \lambda_{S_1 \ldots S_q} \delta(S_1, \ldots, S_q)$ for some nonnegative integers $\lambda_{S_1 \ldots S_q}$.*
*(ii) $(V_n, d)$ is an isometric subspace of the Hamming space $(\{0, 1, \ldots, q-1\}^m, d_H)$ for some integer $m \geq 1$.*

The following result will permit us to link $\ell_1$- and $L_1$-embeddability.

LEMMA 2.10 *Let $(X, d)$ be a distance space. The following assertions are equivalent.*
*(i) $(X, d)$ is $L_1$-embeddable.*
*(ii) $(X, d)$ is a subspace of a measure semimetric space $(\mathcal{A}_\mu, d_\mu)$ for some measure space $(\Omega, \mathcal{A}, \mu)$.*

PROOF. The implication $(ii) \Rightarrow (i)$ is clear, since $(\mathcal{A}_\mu, d_\mu)$ is a subspace of $L_1(\Omega, \mathcal{A}, \mu)$. We check $(i) \Rightarrow (ii)$. It suffices to show that each space $L_1(\Omega, \mathcal{A}, \mu)$ is a subspace of $(\mathcal{B}_\nu, d_\nu)$ for some measure space $(T, \mathcal{B}, \nu)$. Set $T = \Omega \times \mathbb{R}$, $\mathcal{B} = \mathcal{A} \times \mathcal{R}$ where $\mathcal{R}$ is the family of Borel subsets of $\mathbb{R}$, and $\nu = \mu \otimes \lambda$ where $\lambda$ is the Lebesgue measure on $\mathbb{R}$. For $f \in L_1(\Omega, \mathcal{A}, \mu)$, let $E(f) = \{(\omega, s) \in \Omega \times \mathbb{R} : s > f(\omega)\}$ denote its epigraph. Then, the map $f \longmapsto E(f) \triangle E(0)$ provides an isometric embedding from $L_1(\Omega, \mathcal{A}, \mu)$ to $(\mathcal{B}_\nu, d_\nu)$, since $\| f - g \|_1 = \nu(E(f) \triangle E(g))$ holds. ∎

We summarize in the next Theorem the equivalent characterizations that we have obtained for members of the cut cone $\mathrm{CUT}_n$.

THEOREM 2.11 *Let $d \in \mathbb{R}^{E_n}$ and $(V_n, d)$ be the associated distance space. The following assertions are equivalent.*
*(i) $d \in CUT_n$.*
*(ii) $(V_n, d)$ is $\ell_1$-embeddable.*
*(iii) $(V_n, d)$ is $L_1$-embeddable.*
*(iv) There exist a measure space $(\Omega, \mathcal{A}, \mu)$ and $A_1, \ldots, A_n \in \mathcal{A}$ such that $d_{ij} = \mu(A_i \triangle A_i)$ for all $1 \leq i < j \leq n$.*
*(v) $(V_n, d)$ is an isometric subspace of $\ell_1^\infty$.*

## 2.2 $L_1$-spaces (infinite case)

Theorem 2.11 remains partially valid for the case of a distance space $(X, d)$ where the set $X$ is infinite. Indeed, the equivalence $(iii) \Longleftrightarrow (iv)$ holds by Lemma 2.10 and the implication $(ii) \Longrightarrow (iii)$ holds trivially. In fact, there is an infinite analogue of the cut cone, as we now see.

For each subset $Y$ of $X$, let $\delta_Y$ denote the cut function induced by $Y$ defined by $\delta_Y(x, y) = 1$ if $|Y \cap \{x, y\}| = 1$, $\delta_Y(x, y) = 0$ otherwise, for $x, y \in X$; so $\delta_Y$ is just the symmetric function corresponding to the cut semimetric $\delta(Y)$. Let $\mathcal{D}(X)$ denote the set of all cut functions $\delta_Y$ for $Y \subseteq X$.

Let $C_1(X)$ denote the set of all semimetrics $d$ on $X$ for which $(X, d)$ is $L_1$-embeddable.

THEOREM 2.12 *Let $(X, d)$ be a distance space. The following assertions are equivalent.*
*(i) $(X, d)$ is $L_1$-embeddable.*
*(ii) There exists a measure $\nu$ on $\mathcal{D}(X)$ such that $d(x, y) = \int_{\mathcal{D}(X)} \delta(x, y) \nu(d\delta)$ for $x, y \in X$.*

PROOF. $(i) \Rightarrow (ii)$. Assume $(X, d)$ is $L_1$-embeddable. Then, by Lemma 2.10, there exist a measure space $(\Omega, \mathcal{A}, \mu)$ and a map $x \longmapsto A_x$ from $X$ to $\mathcal{A}_\mu$ such that $d(x, y) = \mu(A_x \triangle A_y)$

for $x, y \in X$. For $\omega \in \Omega$, set $A^\omega = \{x \in X : \omega \in A_x\}$. We define a measure $\nu$ on $\mathcal{D}(X)$ additively by setting $\nu(\{\delta_Y\}) = \mu(\{\omega \in \Omega : A^\omega = Y\})$ for each $Y \subseteq X$.
Note that $\omega \in A_x$ if and only if $x \in A^\omega$ and $\omega \in A_x \triangle A_y$ if and only if $|A^\omega \cap \{x, y\}| = 1$. Therefore,

$$\begin{aligned} d(x,y) &= \mu(A_x \triangle A_y) = \mu(\{\omega \in \Omega : |A^\omega \cap \{x,y\}| = 1\}) \\ &= \mu(\{\omega \in \Omega : \delta_{A^\omega}(x,y) = 1\}) \\ &= \mu(\textstyle\bigcup_{Y \subseteq X : \delta_Y(x,y)=1} \{\omega \in \Omega : A^\omega = Y\}) \\ &= \textstyle\int_{\mathcal{D}(X)} \delta(x,y) \nu(d\delta). \end{aligned}$$

$(ii) \Rightarrow (i)$. Conversely, assume that $d = \int_{\mathcal{D}(X)} \delta \nu(d\delta)$ for some non negative measure on $\mathcal{D}(X)$. Fix $s \in X$ and set $A_x = \{\delta \in \mathcal{D}(X) : \delta(s,x) = 1\}$ for each $x \in X$. Then, $d(x,y) = \nu(A_x \triangle A_y)$ holds, since $\delta(x,y) = 0$ if $\delta \notin A_x \triangle A_y$ and $\delta(x,y) = 1$ if $\delta \in A_x \triangle A_y$. This shows, using Lemma 2.10, that $(X,d)$ is $L_1$-embeddable. ∎

THEOREM 2.13 *(i)* $C_1(X)$ *is a convex cone.*
*(ii) The extremal rays of* $C_1(X)$ *are the rays generated by the nonzero cut functions* $\delta_Y$ *for* $Y \subseteq X$, $\emptyset \neq Y \neq X$.

PROOF. $(i)$ follows from Corollary 4.25 $(i)$.
We check $(ii)$. It is easy to see that each cut function lies on an extreme ray of $C_1(X)$ (it lies, in fact, on an extreme ray of the semimetric cone). Consider now $d \in C_1(X)$ which is not a cut function. We can suppose that $d(x_1,x_2) = 1$, $d(x_1,x_3) = \alpha > 0$ and $d(x_2,x_3) = \beta > 0$ for some $x_1, x_2, x_3 \in X$ with $\alpha \geq \beta$. Set $d_1 = \int_{\mathcal{D}(X)} \delta(x_1,x_2)\delta(x_1,x_3)\delta\nu(d\delta)$ and $d_2 = d - d_1$. Then, $d_1, d_2 \in C_1(X)$ by Theorem 2.12. But $d_1(x_1,x_2) = \frac{1+\alpha-\beta}{2} > 0$, since $2\delta(x_1,x_2)\delta(x_1,x_3) = \delta(x_1,x_2) + \delta(x_1,x_3) - \delta(x_2,x_3)$ for each cut function $\delta$. Also, $d_1(x_2,x_3) = 0$ and $d_2(x_2,x_3) = \beta$. Therefore $d$ does not lie on an extreme ray of $C_1(X)$ since $d = d_1 + d_2$ where $d_1$ and $d_2$ are not proportional to $d$. ∎

By Theorem 1.4 applied in the case $p = 1$, we have that a distance space is $L_1$-embeddable if and only if every finite subspace of it is $L_1$-embeddable. Therefore, $L_1$-embeddability can be characterized by a family of linear inequalities, each involving only a finite number of variables. In other words, characterizing $L_1$-embeddability amounts to finding the facet definig inequalities of the cut cone $\mathrm{CUT}_n$ for all $n \geq 2$.

# 3 The correlation cone and $\{0,1\}$-covariances

As before, we set $V_n = \{1, \ldots, n\}$ and $E_n = \{ij : i,j \in V_n, i \neq j\}$ denotes the set of unordered pairs of elements of $V_n$. Given a subset $S$ of $V_n$, let $\pi(S) = (\pi(S)_{ij})_{1 \leq i \leq j \leq n} \in$

$\mathbb{R}^{V_n \cup E_n}$ (identifying the diagonal pair $ii$ with the element $i \in V_n$) be defined by

$$\pi(S)_{ij} = 1 \text{ if } i,j \in S \text{ and } \pi(S)_{ij} = 0 \text{ otherwise} \tag{3.1}$$

for all $i,j \in V_n$; $\pi(S)$ is called a **correlation vector**. The **correlation cone** $\mathrm{COR}_n$ is the cone generated by all correlation vectors $\pi(S)$ for $S \subseteq V_n$ and the **correlation polytope** $\mathrm{COR}_n^\square$ is the convex hull of the correlation vectors $\pi(S)$ for $S \subseteq V_n$. So,

$$\mathrm{COR}_n = \{\sum_{S \subseteq V_n} \lambda_S \pi(S) : \lambda_S \geq 0 \text{ for all } S \subseteq V_n\}, \tag{3.2}$$

$$\mathrm{COR}_n^\square = \{\sum_{S \subseteq V_n} \lambda_S \pi(S) : \sum_{S \subseteq V_n} \lambda_S = 1 \text{ and } \lambda_S \geq 0 \text{ for all } S \subseteq V_n\}. \tag{3.3}$$

The correlation cone and/or polytope were considered in many papers, among them, [MD72, Dez73, Erd87, Isa89, Pad89, Sim90, BH91, Pit86, Pit91]. The polytope $\mathrm{COR}_n^\square$ is also known under the name of **boolean quadric polytope** ([Pad89]). The terminology "correlation polytope" was introduced by Pitowsky ([Pit91]). It is motivated by the fact that $\mathrm{COR}_n^\square$ arises naturally in the context of probability; see Proposition 3.13. Actually, this interpretation was already used in [MD72] in the context of quantum mechanics for describing the pair distributions of particles in lattice sites.

It is sometimes convenient to consider an arbitrary finite subset $X$ instead of $V_n$. Then, the correlation cone is denoted by $\mathrm{COR}(X)$ and the correlation polytope by $\mathrm{COR}^\square(X)$.

## 3.1 The covariance mapping

A simple but fundamental property is that the cut cone $\mathrm{CUT}_{n+1}$ and the correlation cone $\mathrm{COR}_n$ (resp. the cut polytope $\mathrm{CUT}_{n+1}^\square$ and the correlation polytope $\mathrm{COR}_n^\square$) are in one-to-one correspondance via the following covariance mapping.

The **covariance mapping** $\xi$ is the mapping from the space $\mathbb{R}^{E_{n+1}}$ (indexed by the $\binom{n+1}{2}$ pairs of elements of $V_{n+1}$) to the space $\mathbb{R}^{V_n \cup E_n}$ (indexed by the $n$ elements of $V_n$ and the $\binom{n}{2}$ pairs of elements of $V_n$) defined by $p = \xi(d)$ for $d = (d_{ij})_{1 \leq i < j \leq n+1}$ and $p = (p_{ij})_{1 \leq i \leq j \leq n}$ with

$$p_{ij} = \frac{1}{2}(d_{i,n+1} + d_{j,n+1} - d_{ij}) \text{ for all } 1 \leq i \leq j \leq n \tag{3.4}$$

or, equivalently,

$$(3.5) \qquad \begin{cases} d_{ij} = p_{ii} + p_{jj} - 2p_{ij} & \text{for } 1 \leq i < j \leq n \\ d_{i,n+1} = p_{ii} & \text{for } 1 \leq i \leq n. \end{cases}$$

The covariance mapping $\xi$ is a linear bijection from $\mathbb{R}^{E_{n+1}}$ to $\mathbb{R}^{V_n \cup E_n}$. One can easily check that, for any subset $S$ of $V_n$, $\xi(\delta(S)) = \pi(S)$ holds. Therefore,

$$(3.6) \qquad \xi(\mathrm{CUT}_{n+1}) = \mathrm{COR}_n \text{ and } \xi(\mathrm{CUT}^{\square}_{n+1}) = \mathrm{COR}^{\square}_{n+1}.$$

In the same way, given a finite subset $X$ and an element $x_0 \in X$, the cut cone $\mathrm{CUT}(X)$ and the correlation cone $\mathrm{COR}(X \setminus \{x_0\})$ (resp. the cut polytope $\mathrm{CUT}^{\square}(X)$ and the correlation polytope $\mathrm{COR}^{\square}(X \setminus \{x_0\})$) are in one-to-one linear correspondance, via the covariance mapping $\xi$, also denoted as $\xi_{x_0}$ if one wants to stress the choice of the point $x_0$. For the sake of clarity, we rewrite the definition.

Let $X$ be a set (not necessarly finite), $x_0 \in X$, let $d$ be a distance on $X$ and let $p$ be a symmetric function on $X \setminus \{x_0\}$. Then, $p = \xi(d) = \xi_{x_0}(d)$ if

$$(3.7) \qquad p(x,y) = \frac{1}{2}(d(x,x_0) + d(y,x_0) - d(x,y)) \text{ for all } x,y \in X \setminus \{x_0\}$$

or, equivalently,

$$(3.8) \qquad \begin{cases} d(x,x_0) = p(x,x) & \text{for all } x \in X \setminus \{x_0\}, \\ d(x,y) = p(x,x) + p(y,y) - 2p(x,y) & \text{for all } x,y \in X \setminus \{x_0\}. \end{cases}$$

Therefore, for $X$ finite,

$$\xi_{x_0}(\mathrm{CUT}(X)) = \mathrm{COR}(X \setminus \{x_0\}) \text{ and } \xi_{x_0}(\mathrm{CUT}^{\square}(X)) = \mathrm{COR}^{\square}(X \setminus \{x_0\}).$$

Note that, if one uses relation (3.7) for computing $p(x,x_0)$, then one obtains that $p(x,x_0) = 0$ for all $x \in X$. This explains why we consider $p$ as being defined only on the pairs of elements from $X \setminus \{x_0\}$.

The covariance mapping appeared in many different areas of mathematics. See, for instance, [Cri88], [CP93] (where, for a metric space $(X,d)$ and its image $p = \xi(d)$, the quantity $p(x,y)$ is known as the **Gromov product** of $x,y \in X \setminus \{x_0\}$), [Fic87] (where it is called a **linear generalized similarity function**).

The connection between cut and correlation polyhedra, which is formulated in (3.6), was rediscovered independently by several authors (e.g., in [Ham65, Dez73, Sim90]).

## 3.2 Covariances

We now introduce the notion of $M$-covariance. This notion is studied in [Ass79, Ass80b] for $M$ being a subset of a Hilbert space. We consider here only the case when $M = \mathbb{R}$ or $M = \{0, 1\}$.

DEFINITION 3.9 *Let $M$ be a subset of $\mathbb{R}$. A symmetric function $p : X \times X \longrightarrow \mathbb{R}$ is called an $M$-**covariance** if there exist a measure space $(\Omega, \mathcal{A}, \mu)$ and functions $f_x \in L_2(\Omega, \mathcal{A}, \mu)$ taking values in $M$, for all $x \in X$, such that*

$$p(x, y) = \int_\Omega f_x(\omega) f_y(\omega) \mu(d\omega) \text{ for all } x, y \in X.$$

*In particular, $p$ is a $\{0, 1\}$-**covariance** if and only if there exist a measure space $(\Omega, \mathcal{A}, \mu)$ and sets $A_x \in \mathcal{A}_\mu$, for all $x \in X$, such that*

$$p(x, y) = \mu(A_x \cap A_y) \text{ for all } x, y \in X.$$

The next two Lemmas show how $\mathbb{R}$-covariances and $\{0, 1\}$-covariances are related to $L_2$- and $L_1$-embeddable distance spaces, respectively, via the covariance mapping. These facts will be extensively used in the sequel.

LEMMA 3.10 *Let $X$ be a set and $x_0 \in X$. Let $d$ be a distance on $X$ and let $p = \xi_{x_0}(d)$ be the corresponding symmetric function on $X \setminus \{x_0\}$. Then, $(X, \sqrt{d})$ is $L_2$-embeddable if and only if $p$ is an $\mathbb{R}$-covariance on $X \setminus \{x_0\}$.*

PROOF. It is an immediate verification. ∎

LEMMA 3.11 [Dez73] *Let $X$ be a set and $x_0 \in X$. Let $d$ be a distance on $X$ and let $p = \xi_{x_0}(d)$ be the corresponding symmetric function on $X \setminus \{x_0\}$. Then, $(X, d)$ is $L_1$-embeddable if and only if $p$ is a $\{0, 1\}$-covariance on $X \setminus \{x_0\}$.*

PROOF. By Lemma 2.10, $(X, d)$ is $L_1$-embeddable if and only if there exist a measure space $(\Omega, \mathcal{A}, \mu)$ and sets $A_x \in \mathcal{A}_\mu$ for $x \in X$ such that $d(x, y) = \mu(A_x \triangle A_y)$ for all $x, y \in X$. Without loss of generality, we can suppose that $A_{x_0} = \emptyset$. Then, it follows from relation (3.7 ) that $p(x, y) = \frac{1}{2}(\mu(A_x) + \mu(A_y) - \mu(A_x \triangle A_y)) = \mu(A_x \cap A_y)$ for all $x, y \in X \setminus \{x_0\}$. ∎

The following finitude result is a consequence of Lemma 3.11 and Theorem 1.4.

PROPOSITION 3.12 *Let $p$ be a symmetric function on $X$. Then, $p$ is a $\{0,1\}$-covariance on $X$ if and only if, for each finite subset $Y$ of $X$, the restriction of $p$ to $Y$ is a $\{0,1\}$-covariance on $Y$.*

We now give an interpretation of the members of the correlation cone and polytope which is an analogue of Proposition 2.4 (via the covariance mapping). It was rediscovered in [Pit86].

PROPOSITION 3.13 *Let $p = (p_{ij})_{1 \leq i \leq j \leq n} \in \mathbb{R}^{V_n \cup E_n}$. The following assertions are equivalent.*
*(i) $p \in COR_n$ (resp. $p \in COR_n^{\square}$).*
*(ii) There exist a measure space (resp. a probability space) $(\Omega, \mathcal{A}, \mu)$ and $A_1, \ldots, A_n \in \mathcal{A}$ such that $p_{ij} = \mu(A_i \cap A_j)$ for all $1 \leq i \leq j \leq n$.*

Therefore, for $X$ finite, $p$ is a $\{0,1\}$-covariance on $X$ if and only if $p$ belongs to the correlation cone $\mathrm{COR}(X)$.

For the members of the correlation cone which can be written as a nonnegative *integer* combination of correlation vectors, we can assume that the measure space in Proposition 3.13 $(ii)$ is endowed with the cardinality measure. Namely, we have the following result, which is an analogue of Proposition 2.7 $(i) \Longleftrightarrow (iii)$ (via the covariance mapping).

PROPOSITION 3.14 *Let $p = (p_{ij})_{1 \leq i \leq j \leq n} \in \mathbb{R}^{V_n \cup E_n}$. The following assertions are equivalent.*
*(i) $p = \sum_{S \subseteq V_n} \lambda_S \pi(S)$ for some nonnegative integers $\lambda_S$.*
*(ii) There exist a finite set $\Omega$ and $n$ subsets $A_1, \ldots, A_n$ of $\Omega$ such that $p_{ij} = |A_i \cap A_j|$ for all $1 \leq i \leq j \leq n$.*

A vector $p$ satisfying the conditions of Proposition 3.14 is sometimes called an **intersection pattern** in the literature (see, e.g., [DR84]). Testing whether a given vector $p$ is an intersection pattern is an NP-complete problem ([Chv80]). However, this problem is polynomial when restricted to some classes of vectors; for instance, it is polynomial when retricted to the class of the vectors $p$ such that $p_{ii} = 2$ for all $i \in V_n$. We refer to Chapter ??? (on hypercube embedding) for results related to these questions.

## 3.3 The Boole problem

We now describe an application of the interpretation of the correlation cone and polytope given in Proposition 3.13 for the following **Boole problem**.

Let $(\Omega, \mathcal{A}, \mu)$ be a probability space and let $A_1, \ldots, A_n$ be $n$ events of $\mathcal{A}$. A classical question, which goes back to Boole [Boo54], is the following:

*Suppose we are given the values $p_i = \mu(A_i)$ for $1 \leq i \leq n$, what is the best estimation of $\mu(A_1 \cup \ldots A_n)$ ?*

It is easy to see that the answer is $\max(p_1, \ldots, p_n) \leq \mu(A_1 \cup \ldots A_n) \leq \min(1, \sum_{1 \leq i \leq n} p_i)$. The next natural question is the following:

*Suppose we are given the values $p_i = \mu(A_i)$ for $1 \leq i \leq n$ and the values of the joint probabilities $p_{ij} = \mu(A_i \cap A_j)$ for $1 \leq i < j \leq n$. What is the best estimation of $\mu(A_1 \cup \ldots \cup A_n)$ in terms of the $p_i$'s and the $p_{ij}$'s ?*

In fact, an answer to this problem is given by the facet defining inequalities for the correlation polytope $\mathrm{COR}_n^{\square}$. Namely,

$$\mu(A_1 \cup \ldots \cup A_n) \geq \max(w^T p : w^T z \leq 1 \text{ is facet defining for } \mathrm{COR}_n^{\square})$$

(see Proposition 3.17 and relation (3.20)).

The approach described below for obtaining estimations of $\mu(A_1 \cup \ldots \cup A_n)$ uses linear programming; it was considered in [KM76], [Pit91]. Given $p \in \mathrm{COR}_n$, consider the following two linear programming problems.

$$(3.15) \qquad \begin{array}{lll} z_{min} := & \min & \sum_{\emptyset \neq S \subseteq V_n} \lambda_S \\ & \text{subject to} & \sum_{\emptyset \neq S \subseteq V_n} \lambda_S \pi(S) = p \\ & & \lambda_S \geq 0 \text{ for } \emptyset \neq S \subseteq V_n \end{array}$$

$$(3.16) \qquad \begin{array}{lll} z_{max} := & \max & \sum_{\emptyset \neq S \subseteq V_n} \lambda_S \\ & \text{subject to} & \sum_{\emptyset \neq S \subseteq V_n} \lambda_S \pi(S) = p \\ & & \lambda_S \geq 0 \text{ for } \emptyset \neq S \subseteq V_n \end{array}$$

PROPOSITION 3.17 *Let $z_{min}$ and $z_{max}$ be defined by the relations* (3.15) *and* (3.16)*. Then, $z_{min} \leq \mu(A_1 \cup \ldots \cup A_n) \leq z_{max}$.*

PROOF. Let $p \in \mathrm{COR}_n$ be defined by $p_{ij} = \mu(A_i \cap A_j)$ for all $1 \leq i \leq j \leq n$ (setting $p_{ii} = p_i$). For $S \subseteq V_n$, set $A^S = \bigcap_{i \in S} A_i \cap \bigcap_{i \notin S} (\Omega \setminus A_i)$. Then, $A_i \cap A_j = \bigcup_{S \subseteq V_n : i,j \in S} A^S$, $\Omega = \bigcup_{S \subseteq V_n} A^S$ and $A_1 \cup \ldots \cup A_n = \bigcup_{S \subseteq V_n : S \neq \emptyset} A^S$. Therefore, $p = \sum_{S \subseteq V_n : S \neq \emptyset} \mu(A^S) \pi(S)$ holds, with $\mu(A^S) \geq 0$ for all $S$. Hence $(\mu(A^S) : \emptyset \neq S \subseteq V_n)$ is a feasible solution to the programs (3.15) and (3.16), with objective value $\mu(A_1 \cup \ldots \cup A_n)$. This shows the result. ∎

The dual programs to (3.15) and (3.16) are the following programs (3.18) and (3.19), respectively.

$$\text{(3.18)} \qquad \begin{array}{lll} \max & w^T p & \\ \text{subject to} & w^T \pi(S) \leq 1 & \text{for } \emptyset \neq S \subseteq V_n \end{array}$$

$$\text{(3.19)} \qquad \begin{array}{lll} \min & w^T p & \\ \text{subject to} & w^T \pi(S) \geq 1 & \text{for } \emptyset \neq S \subseteq V_n \end{array}$$

By linear programming duality, we obtain that

$$\text{(3.20)} \qquad z_{min} = \max(w^T p : w^T z \leq 1 \text{ is a valid inequality for } \mathrm{COR}_n^\square)$$

and one can easily verify that, in relation (3.20), it is sufficient to consider facet defining inequalities. Similarly,

$$z_{max} = \min(w^T p : w^T z \geq 1 \quad \text{is facet defining for the polytope } \mathrm{Conv}(\{\pi(S) : \emptyset \neq S \subseteq V_n\}).$$

(The latter polytope is distinct from $\mathrm{COR}_n^\square$ since it does not contain the origin.)

Therefore, by (3.20), every valid inequality for $\mathrm{COR}_n^\square$ yields a lower bound for $\mu(A_1 \cup \ldots \cup A_n)$ in terms of the joint probabilities $p_{ij} = \mu(A_i \cap A_j)$ for $1 \leq i \leq j \leq n$. We now give some examples of such lower bounds.

Suppose $p = \sum_S \lambda_S \pi(S)$ with $\lambda_S \geq 0$ for all $S$. Let $u \in \mathbb{R}^{V_n \cup E_n}$ be defined by $u_i = n$ for $i \in V_n$ and $u_{ij} = -2$ for $ij \in E_n$. By taking the scalar product of both sides of $p = \sum_S \lambda_S \pi(S)$ with $u$, we obtain that $n \sum_{1 \leq i \leq n} p_i - 2 \sum_{1 \leq i < j \leq n} p_{ij} = \sum_S \lambda_S |S|(n+1-|S|)$, where $n \leq |S|(n+1-|S|) \leq \lfloor \frac{n+1}{2} \rfloor \lceil \frac{n+1}{2} \rceil$ if $S \neq \emptyset$. Hence, we deduce that

$$\frac{n \sum_{1 \leq i \leq n} p_i - 2 \sum_{1 \leq i < j \leq n} p_{ij}}{\lfloor \frac{n+1}{2} \rfloor \lceil \frac{n+1}{2} \rceil} \leq \sum_{\emptyset \neq S \subseteq V_n} \lambda_S \leq \frac{n \sum_{1 \leq i \leq n} p_i - 2 \sum_{1 \leq i < j \leq n} p_{ij}}{n}$$

and, therefore, from the definition of $z_{min}$, $z_{max}$ and from Proposition 3.17,

$$\text{(3.21)} \qquad \frac{n \sum\limits_{1 \leq i \leq n} p_i - 2 \sum\limits_{1 \leq i < j \leq n} p_{ij}}{\lfloor \frac{n+1}{2} \rfloor \lceil \frac{n+1}{2} \rceil} \leq \mu(A_1 \cup \ldots \cup A_n) \leq \frac{n \sum\limits_{1 \leq i \leq n} p_i - 2 \sum\limits_{1 \leq i < j \leq n} p_{ij}}{n}.$$

The inequality

$$\text{(3.22)} \qquad 2k \sum_{1 \leq i \leq n} p_i - 2 \sum_{1 \leq i < j \leq n} p_{ij} \leq k(k+1)$$

is valid for the correlation polytope $\mathrm{COR}_n^\square$, for all $1 \le k \le n-1$; it is facet defining if $1 \le k \le n-2$ and $n \ge 4$. Setting $b_{n+1} = 2k+1-n$ and $b_1 = \ldots = b_n = 1$, the inequality (3.22) corresponds (via the covariance map) to the inequality

$$\sum_{1 \le i < j \le n+1} b_i b_j x_{ij} \le k(k+1) \tag{3.23}$$

which is valid for the cut polytope $\mathrm{CUT}_{n+1}^\square$. Note that the inequality (3.23) is a switching of the hypermetric inequality $Hyp_{n+1}(1,\ldots,1,-1,\ldots,-1,2k+1-n)$ (with $n-k$ coefficients $+1$ and $k$ coefficients -1) (see Section 4.1 for definitions). Therefore, we have the following lower bound for $\mu(A_1 \cup \ldots \cup A_n)$:

$$\frac{2}{k+1} \sum_{1 \le i \le n} p_i - \frac{2}{k(k+1)} \sum_{1 \le i < j \le n} p_{ij} \le \mu(A_1 \cup \ldots \cup A_n) \tag{3.24}$$

for each $k$, $1 \le k \le n-1$. The bound (3.24) was found independently by several authors, including [Chu41, DS67, Gal77]. Note that (3.24) coincides with the lower bound of (3.21) in the case $n = 2k$. The case $k = 1$ of (3.24) gives the bound

$$\sum_{1 \le i \le n} p_i - \sum_{1 \le i < j \le n} p_{ij} \le \mu(A_1 \cup \ldots \cup A_n)$$

which is a special case of the Bonferroni bound (3.29) mentioned below.

More generally, given integers $b_1, \ldots, b_n$ and $k \ge 0$, the inequality

$$\sum_{1 \le i \le n} b_i(2k+1-b_i)p_i - 2 \sum_{1 \le i < j \le n} b_i b_j p_{ij} \le k(k+1) \tag{3.25}$$

is valid for $\mathrm{COR}_n^\square$. This yields the bound

$$\frac{1}{k(k+1)}\Big( \sum_{1 \le i \le n} p_i b_i (2k+1-b_i) - 2 \sum_{1 \le i < j \le n} b_i b_j p_{ij} \Big) \le \mu(A_1 \cup \ldots \cup A_n).$$

The inequality (3.25) can alternatively be written as

$$\Big( \sum_{1 \le i \le n} b_i p_i - k \Big)\Big( \sum_{1 \le i \le n} b_i p_i - k - 1 \Big) \ge 0 \tag{3.26}$$

with the convention that, when developing the product, the expression $p_i p_j$ is replaced by the variable $p_{ij}$ (setting $p_{ii} = p_i$). The inequality (3.25) (or (3.26)) (or special cases of it) was considered by many authors (e.g., [Yos70, MD72, KM76, Erd87, Mes87, Pit91]).

(Note that, if we set $b_{n+1} = 2k+1-\sum_{1\le i\le n} b_i$, the inequality (3.25) corresponds (via the covariance mapping) to the inequality

$$\sum_{1\le i<j\le n+1} b_i b_j x_{ij} \le k(k+1),$$

which is valid for the cut polytope $\mathrm{CUT}_{n+1}^{\square}$.)

## 3.4 Generalization to higher order correlations

Clearly, much of the treatment of Section 3.3 can be generalized to higher order correlations. Namely, let $\mathcal{I}$ be a family of subsets of $V_n$. Given a subset $S$ of $V_n$, its **$\mathcal{I}$-correlation vector** $\pi^{\mathcal{I}}(S) \in \mathbb{R}^{\mathcal{I}}$ is defined by

$$\begin{cases} \pi^{\mathcal{I}}(S)_I = 1 & \text{if } I \subseteq S, \\ \pi^{\mathcal{I}}(S)_I = 0 & \text{otherwise,} \end{cases}$$

for all $I \in \mathcal{I}$. Then, the cone $\mathrm{COR}_n(\mathcal{I})$ (resp. the polytope $\mathrm{COR}_n^{\square}(\mathcal{I})$) is defined as the conic hull (resp. the convex hull) of all $\mathcal{I}$-correlation vectors $\pi^{\mathcal{I}}(S)$ for $S \subseteq V_n$.

Given an integer $1 \le m \le n$, let $\mathcal{I}_{\le m}$ denote the collection of all subsets of $V_n$ of cardinality less or equal to $m$. Hence, $\mathcal{I}_{\le 2}$ consists of all singletons and pairs of elements of $V_n$ and $\mathrm{COR}_n(\mathcal{I}_{\le 2})$, $\mathrm{COR}_n^{\square}(\mathcal{I}_{\le 2})$ coincide with $\mathrm{COR}_n$, $\mathrm{COR}_n^{\square}$, respectively.

For $\mathcal{I} = 2^{V_n}$, which consists of all subsets of $V_n$, $\mathrm{COR}_n^{\square}(2^{V_n})$ is a simplex of dimension $2^n-1$ and $\mathrm{COR}_n(2^{V_n})$ is a simplicial cone of dimension $2^n-1$. This implies, in particular, that every correlation polytope $\mathrm{COR}_n^{\square}(\mathcal{I})$ arises as a projection of the simplex $\mathrm{COR}_n^{\square}(2^{V_n})$ (namely, on the subspace $\mathbb{R}^{2^{V_n}\setminus\mathcal{I}}$).

Proposition 3.13 remains valid for the case of arbitrary $\mathcal{I}$-correlations.

PROPOSITION 3.27 *Let $\mathcal{I}$ be a nonempty collection of subsets of $\{1,\ldots,n\}$ and let $p = (p_I)_{I\in\mathcal{I}} \in \mathbb{R}^{\mathcal{I}}$. The following assertions are equivalent.*
*(i) $p \in COR_n(\mathcal{I})$ (resp. $p \in COR_n^{\square}(\mathcal{I})$).*
*(ii) There exist a measure space (resp. a probability space) $(\Omega, \mathcal{A}, \mu)$ and $A_1,\ldots,A_n \in \mathcal{A}$ such that $p_I = \mu(\cap_{i\in I} A_i)$ for all $I \in \mathcal{I}$.*

We have the following general formulation of the Boole problem.

*Suppose we are given the values of the joint probabilities $p_I = \mu(\cap_{i\in I} A_i)$, for all $I \in \mathcal{I}$. What is the best estimation of $\mu(A_1 \cup \ldots \cup A_n)$ in terms of the $p_I$'s ?*

The same reasoning as in the case of the usual pairwise correlations permits to show the following generalization of Proposition 3.17.

PROPOSITION 3.28 *Let* $p_I = \mu(\cap_{I\in\mathcal{I}} A_i)$ *for* $I \in \mathcal{I}$. *Then,*
$z_{min} \leq \mu(A_1 \cup \ldots \cup A_n) \leq z_{max}$, *where*

$$\begin{aligned} z_{min} &= \min(\textstyle\sum_{\emptyset\neq S\subseteq V_n} \lambda_S : p = \sum_{\emptyset\neq S\subseteq V_n} \lambda_S \pi^{\mathcal{I}}(S) \textit{ with } \lambda_S \geq 0 \textit{ for all } S) \\ &= \max(w^T p : w^T z \leq 1 \textit{ is facet defining for } COR_n^{\square}(\mathcal{I})) \end{aligned}$$

*and*

$$\begin{aligned} z_{max} &= \max(\textstyle\sum_{\emptyset\neq S\subseteq V_n} \lambda_S : p = \sum_{\emptyset\neq S\subseteq V_n} \lambda_S \pi^{\mathcal{I}}(S) \textit{ with } \lambda_S \geq 0 \textit{ for all } S) \\ &= \min(w^T p : w^T z \leq 1 \textit{ is facet defining for } Conv(\{\pi^{\mathcal{I}}(S) : \emptyset \neq S \subseteq V_n\})). \end{aligned}$$

Boros and Prekopa [BP89] consider in detail the case $\mathcal{I} = \mathcal{I}_{\leq m}$; they describe a method for finding bounds on $\mu(A_1\cup\ldots\cup A_n)$ in terms of the quantities $S_k = \sum_{1\leq i_1<\ldots<i_k\leq n} \mu(A_{i_1} \cap \ldots \cap A_{i_k})$ for $1 \leq k \leq n$. For instance, the following bounds hold

$$\left\{ \begin{array}{ll} \mu(A_1 \cup \ldots \cup A_n) \geq \sum_{1\leq i\leq m} (-1)^{i-1} S_i & \text{for } m \text{ even ,} \\ \mu(A_1 \cup \ldots \cup A_n) \leq \sum_{1\leq i\leq m} (-1)^{i-1} S_i & \text{for } m \text{ odd .} \end{array} \right. \tag{3.29}$$

They were first discovered by Bonferroni [Bon36]. Several improvments of them have been proposed, most recently in [Gra93].

The inequality (3.26) can be easily generalized to the case of the polytope $\mathrm{COR}_n^{\square}(\mathcal{I}_{\leq 2m})$ for any $m \geq 1$. Given integers $b_1, \ldots, b_n$ and $k_1, \ldots, k_m \geq 0$, the inequality

$$\prod_{1\leq l\leq m} (\sum_{1\leq i\leq n} b_i p_i - k_l)(\sum_{1\leq i\leq n} b_i p_i - k_l - 1) \geq 0$$

is clearly valid for the polytope $\mathrm{COR}_n^{\square}(\mathcal{I}_{\leq 2m})$. Thus arises the question of determining the parameters $b_1, \ldots, b_n, k_1, \ldots, k_m$ for which it is facet defining. This problem is, however, already difficult for the case $m = 1$ of the correlation polytope $\mathrm{COR}_n^{\square}$.

# 4 Conditions for $L_1$-embeddability

We present in Section 4.1 some of the most important known necessary conditions for $L_1$-embeddability, namely, the hypermetric and the negative type conditions. There are many other known necessary conditions for $L_1$-embeddability, arising from known valid inequalities for the cut cone; they are described in Chapter ????( on facets). We focus here on the hypermetric and negative type conditions since they will be used repeatedly in this Chapter and throughout the book. In Section 4.3, we show the implications exisiting between the properties of being $L_1$-, $L_2$-embeddable, of negative type, or hypermetric, for a distance space. In Section 4.4, we present two operations, the direct sum and the tensor product, which preserve, respectively, $L_1$-embeddable distance spaces and $\{0,1\}$-covariances.

## 4.1 Hypermetric and negative type conditions

### 4.1.1 Hypermetric and negative type inequalities

Let $n \geq 2$, $b_1, \ldots, b_n$ be integers such that $\epsilon := \sum_{1 \leq i \leq n} b_i \in \{0, 1\}$. We consider the inequality

$$\sum_{1 \leq i < j \leq n} b_i b_j d_{ij} \leq 0. \tag{4.1}$$

If $\sum_{1 \leq i \leq n} b_i = 1$, then the inequality (4.1) is called a **hypermetric inequality** and is denoted by $Hyp_n(b_1, \ldots, b_n)$. If $\sum_{1 \leq i \leq n} b_i = 0$, then the inequality (4.1) is called an **inequality of negative type** and is denoted by $Neg_n(b_1, \ldots, b_n)$. The inequality (4.1) is said to be **pure** if $|b_i| = 0, 1$ for all $i \in V_n$. We can suppose that at least two of the $b_i$'s are nonzero (else, the inequality (4.1) is void). Hence, $\sum_{1 \leq i \leq n} |b_i| = 2k + \epsilon$ for some integer $k \geq 1$. The inequality (4.1) is then said to be $(2k + \epsilon)$**-gonal**.

In particular, the 2-gonal inequality is the inequality of negative type $Neg_n(b_1, \ldots, b_n)$, where $b_i = 1$, $b_j = -1$ and $b_h = 0$ for $h \in V_n \setminus \{i, j\}$, for some distinct $i, j \in V_n$; it is nothing but the nonnegativity constraint

$$d_{ij} \geq 0.$$

The pure 3-gonal inequality is the hypermetric inequality $Hyp_n(b_1, \ldots, b_n)$, where $b_i = b_j = 1$, $b_k = -1$ and $b_h = 0$ for $h \in V_n \setminus \{i, j, k\}$, for some distinct $i, j, k \in V_n$; it is nothing but the triangle inequality (1.1).

For $\epsilon = 0, 1$, the pure $(2k + \epsilon)$-gonal inequality reads

$$\sum_{1 \leq r < s \leq k+\epsilon} d_{i_r i_s} + \sum_{1 \leq r < s \leq k} d_{j_r j_s} - \sum_{\substack{1 \leq r \leq k+\epsilon \\ 1 \leq s \leq k}} d_{i_r j_s} \leq 0,$$

where $i_1, \ldots, i_k, i_{k+\epsilon}, j_1, \ldots, j_k$ are distinct indices of $V_n$.

Figure 1 shows the pure 4-gonal and 5-gonal inequalities or, rather, their left hand sides. It should be understood as follows: a plain edge between two nodes $i$ and $j$ indicates a coefficient $+1$ for the variable $d_{ij}$ and a dotted edge indicates a coefficient -1.

Figure 1

The negative type inequalities are classical inequalities in analysis; they were used, in particular, by Schoenberg [Sch37, Sch38a, Sch38b]. The hypermetric inequalities were introduced by Deza [Dez61, Dez62] and later, independently, by Kelly [Kel70].

### 4.1.2 Hypermetric and negative type distance spaces

We now turn to the definition of a hypermetric distance space, or of a negative type distance space. Basically, a distance space $(X, d)$ is said to be hypermetric (resp. of negative type) if $d$ satisfies all hypermetric inequalities (resp. all inequalities of negative type). More precisely, we have the following definitions.

Let $(X, d)$ be a distance space. Then, $(X, d)$ is said to be of **negative type** if, for all $n \geq 2$, $x_1, \ldots, x_n \in X$, $b_1, \ldots, b_n \in \mathbb{Z}$ with $\sum_{1 \leq i \leq n} b_i = 0$, the following inequality holds

$$\sum_{i,j=1}^{n} b_i b_j d(x_i, x_j) \leq 0. \tag{4.2}$$

$(X, d)$ is said to be **hypermetric** if, for all $n \geq 2$, $x_1, \ldots, x_n \in X$, $b_1, \ldots, b_n \in \mathbb{Z}$ with $\sum_{1 \leq i \leq n} b_i = 1$, the inequality (4.2) holds. For $\epsilon = 0, 1$ and $k \in \mathbb{Z}, k \geq 1$, $(X, d)$ is said to be **$(2k + \epsilon)$-gonal** if, for all $n \geq 2$, $x_1, \ldots, x_n \in X$, $b_1, \ldots, b_n \in \mathbb{Z}$ with $\sum_{1 \leq i \leq n} b_i = \epsilon$ and $\sum_{1 \leq i \leq n} |b_i| = 2k + \epsilon$, the inequality (4.2) holds.

Note that, in the above definitions, we do not require that the points $x_1, \ldots, x_n$ be distinct. For instance, suppose that $x_1 = x_2$. Then, $d(x_1, x_2) = 0$ and $d(x_1, x_i) = d(x_2, x_i)$ for all $i$ and, therefore, the inequality (4.2) reads

$$\sum_{2 \leq i < j \leq n} b'_i b'_j d(x_i, x_j) \leq 0$$

after setting $b'_2 = b_1 + b_2, b'_3 = b_3, \ldots, b'_n = b_n$. In other words, we could have assumed in the above definitions that the inequality (4.2) is pure, i.e., that $|b_i| = 0, 1$ for all $i$. For instance, the distance space $(X,d)$ is $(2k+\epsilon)$-gonal if and only if, for all (not necessarly distinct) $x_1,\ldots,x_k,x_{k+\epsilon},y_1,\ldots,y_k \in X$, the following inequality holds

$$\sum_{1\leq i<j\leq k+\epsilon} d(x_i,x_j)+\sum_{1\leq i<j\leq k} d(y_i,y_j)-\sum_{\substack{1\leq i\leq k+\epsilon\\ 1\leq j\leq k}} d(x_i,y_j)\leq 0. \tag{4.3}$$

In particular, $(X,d)$ is 5-**gonal** if and only if, for all $x_1,x_2,x_3,y_1,y_2 \in X$, the following inequality holds

$$\sum_{1\leq i<j\leq 3} d(x_i,x_j)+d(y_1,y_2)-\sum_{\substack{1\leq i\leq 3\\ j=1,2}} d(x_i,y_j)\leq 0. \tag{4.4}$$

Clearly, if $(X,d)$ is of negative type, then the inequality (4.2) holds for all $b_1,\ldots,b_n\in\mathbb{R}$ with $\sum_{1\leq i\leq n} b_i=0$. Some more implications among the $k$-gonal conditions are summmarized in the next result.

LEMMA 4.5 *Let $(X,d)$ be a distance space.*
*(i) If $(X,d)$ is $(2k+1)$-gonal, then $(X,d)$ is $(2k+2)$-gonal, for any integer $k\geq 1$.*
*(ii) If $(X,d)$ is $(k+2)$-gonal, then $(X,d)$ is $k$-gonal, for any integer $k\geq 2$.*

PROOF. $(i)$ Let $x_1,\ldots,x_{k+1},y_1,\ldots,y_{k+1}$ be $2k+2$ points of $X$. By assumption, $(X,d)$ satisfies each of the $k+1$ inequalities (4.3) obtained by considering all $y_i$'s except one. Similarly, $(X,d)$ satisfies each of the $k+1$ inequalities (4.3) obtained by considering all $x_i$'s except one (and exchanging the role of the $x_i$'s and $y_i$'s). If we sum up these $2k+2$ inequalities, we deduce that $(X,d)$ satisfies the $(2k+2)$-gonal inequality (4.3) relative to the points $x_1,\ldots,x_{k+1},y_1,\ldots,y_{k+1}$.
$(ii)$ We check that $(X,d)$ is $(2k-1)$-gonal whenever it is $(2k+1)$-gonal; the other case is similar. Let $x_1,\ldots,x_k,y_1,\ldots,y_{k-1}$ be $2k-1$ points of $X$. Let $x\in X$ and set $x_{k+1}=y_k=x$. By assumption, $(X,d)$ satisfies the $(2k+1)$-gonal inequality (4.3) relative to the points $x_1,\ldots,x_{k+1},y_1,\ldots,y_k$. But, the latter inequality, after some cancellations, is nothing but the $(2k-1)$-gonal inequality (4.3) relative to the points $x_1,\ldots,x_k,y_1,\ldots,y_{k-1}$. ∎

REMARK 4.6 The proof of Lemma 4.5 $(i)$ shows, in fact, that the pure $(2k+2)$-gonal inequality follows from the pure $(2k+1)$-gonal inequalities. However, for $k\geq 2$ integer, the $k$-gonal inequalities do **not** follow from the $(k+2)$-gonal inequalities (the proof of Lemma 4.5 $(ii)$ works indeed at the level of distance spaces since we make the assumption that the two points $x_{k+1}$ and $y_k$ of $X$ coincide).

**Equality case in the hypermetric and negative type inequalities.** The following question is considered in [Kel70, Ass84, Bal90]. What are the distance spaces, within a given class, that satisfy a given hypermetric or negative type inequality at equality ?
For instance, Kelly [Kel70] characterizes the finite subspaces of $(\mathbb{R}, d_{\ell_1})$ that satisfy the $(2k+1)$-gonal inequality at equality. Namely, given $x_1, \ldots, x_{k+1}, y_1, \ldots, y_k \in \mathbb{R}$, the equality

$$\sum_{1\leq i<j\leq k+1} |x_i - x_j| + \sum_{1\leq i<j\leq k} |y_i - y_j| - \sum_{\substack{1\leq i\leq k+1\\ 1\leq j\leq k}} |x_i - y_j| = 0$$

holds if and only if $y_1, \ldots, y_k$ separate $x_1, \ldots, x_{k+1}$, i.e., there exist a permutation $\alpha$ of $\{1, \ldots, k+1\}$ and a permutation $\beta$ of $\{1, \ldots, k\}$ such that

$$x_{\alpha(1)} \leq y_{\beta(1)} \leq x_{\alpha(2)} \leq y_{\beta(2)} \leq \cdots \leq y_{\beta(k)} \leq x_{\alpha(k+1)}.$$

Generalizations and related results can be found in [Kel70, Ass84].
Ball (Lemma 4,[Bal90]) characterizes the reals $x_1, \ldots, x_n$ for which the distance space $(\{x_1, \ldots, x_n\}, d_{\ell_1})$ satisfies the negative type inequality $Neg_n(-(n-4), 1, \ldots, 1, -2)$ at equality. This result is used for Proposition 7.4 $(i)$ in Section 7.1, for deriving a lower bound on the minimum $\ell_1$-dimension of a distance space.

The **hypermetric cone** HYP$(X)$ (resp. the **negative type cone** NEG$(X)$) is defined as the set of all the distances $d$ on $X$ that are hypermetric (resp. of negative type). In the finite case, $X = V_n$, the cone HYP$(V_n)$ is simply denoted by HYP$_n$ and the cone NEG$(V_n)$ by NEG$_n$; both are assumed to be cones in $\mathbb{R}^{E_n}$. Hence, HYP$_n$ (resp. NEG$_n$) is the cone in $\mathbb{R}^{E_n}$ defined by the inequalities (4.1) for all integers $b_1, \ldots, b_n$ with $\sum_{1\leq i\leq n} b_i = 1$ (resp. $\sum_{1\leq i\leq n} b_i = 0$).

### 4.1.3 Analogues of the hypermetric and negative type conditions for covariances

We now introduce the notion of a function of positive type, which will turn out to be closely related to that of a distance of negative type.

DEFINITION 4.7 *A symmetric function $p : X \times X \longrightarrow \mathbb{R}$ is said to be of* **positive type** *on $X$ if, for all $n \geq 2$, $x_1, \ldots, x_n \in X$, the matrix $(p(x_i, x_j))_{1\leq i,j\leq n}$ is positive semidefinite, i.e., the inequality $\sum_{1\leq i,j\leq n} b_i b_j p(x_i, x_j) \geq 0$ holds for all $b_1, \ldots, b_n \in \mathbb{R}$ (or, equivalently, for all $b_1, \ldots, b_n \in \mathbb{Z}$).*

The next Lemma 4.8 shows that the notions of functions of positive type and distances of negative type are, in fact, equivalent (via the covariance mapping). Then, Lemma 4.9 is an analogue of Lemma 4.8 for hypermetric inequalities. Both results will be used very often in the book and, in particular, in this Chapter and in Chapter ??? (on hypermetrics).

LEMMA 4.8 *Let $X$ be a set and $x_0 \in X$. Let $d$ be a distance on $X$ and $p = \xi_{x_0}(d)$ be the corresponding symmetric function on $X \setminus \{x_0\}$. The following assertions are equivalent.*
*(i) $(X, d)$ is of negative type, i.e., for all $n \geq 2$, $x_1, \ldots, x_n \in X$, $b_1, \ldots, b_n \in \mathbb{Z}$ with $\sum_{1 \leq i \leq n} b_i = 0$, $\sum_{1 \leq i,j \leq n} b_i b_j d(x_i, x_j) \leq 0$.*
*(ii) $p$ is a function of positive type on $X \setminus \{x_0\}$, i.e., for all $n \geq 2$, $x_1, \ldots, x_n \in X \setminus \{x_0\}$, $b_1, \ldots, b_n \in \mathbb{Z}$, $\sum_{1 \leq i,j \leq n} b_i b_j p(x_i, x_j) \geq 0$.*

PROOF. The proof is based on the following observation. Given $x_1, \ldots, x_n \in X \setminus \{x_0\}$, $b_1, \ldots, b_n \in \mathbb{Z}$, we have that

$$\begin{aligned} \sum_{1 \leq i,j \leq n} b_i b_j p(x_i, x_j) &= \sum_{1 \leq i,j \leq n} \tfrac{1}{2}(d(x_i, x_0) + d(x_j, x_0) - d(x_i, x_j)) \\ &= \sum_{1 \leq i \leq n} b_i (\sum_{1 \leq j \leq n} b_j) d(x_0, x_i) - \sum_{1 \leq i,j \leq n} b_i b_j d(x_i, x_j) \\ &= - \left( \sum_{0 \leq i,j \leq n} b_i b_j d(x_i, x_j) \right), \end{aligned}$$

after setting $b_0 = - \sum_{1 \leq j \leq n} b_j$. ∎

LEMMA 4.9 *Let $X$ be a set and $x_0 \in X$. Let $d$ be a distance on $X$ and $p = \xi_{x_0}(d)$ be the corresponding symmetric function on $X \setminus \{x_0\}$. The following assertions are equivalent.*
*(i) $(X, d)$ is hypermetric, i.e., for all $n \geq 2$, $x_1, \ldots, x_n \in X$, $b_1, \ldots, b_n \in \mathbb{Z}$ with $\sum_{1 \leq i \leq n} b_i = 1$, $\sum_{1 \leq i,j \leq n} b_i b_j d(x_i, x_j) \leq 0$ holds.*
*(ii) For all $n \geq 2$, $x_1, \ldots, x_n \in X \setminus \{x_0\}$, $b_1, \ldots, b_n \in \mathbb{Z}$, $p$ satisfies the inequality $\sum_{1 \leq i,j \leq n} b_i b_j p(x_i, x_j) - \sum_{1 \leq i \leq n} b_i p(x_i, x_i) \geq 0$.*

PROOF. The proof is based on the following observation. Given $x_1, \ldots, x_n \in X \setminus \{x_0\}$, $b_1, \ldots, b_n \in \mathbb{Z}$, then

$$\begin{aligned} &\sum_{1 \leq i,j \leq n} b_i b_j p(x_i, x_j) - \sum_{1 \leq i \leq n} b_i p(x_i, x_i) \\ &= \sum_{1 \leq i \leq n} b_i (\sum_{1 \leq j \leq n} b_j - 1) d(x_0, x_i) - \sum_{1 \leq i,j \leq n} b_i b_j d(x_i, x_j) \\ &= -(\sum_{0 \leq i,j \leq n} b_i b_j d(x_i, x_j)), \end{aligned}$$

after setting $b_0 = 1 - \sum_{1 \leq j \leq n} b_j$. ∎

## 4.2 Characterization of $L_2$-embeddability

We present in this Section several equivalent characterizations for $L_2$-embeddable distance spaces.

We start with some preliminary results. Given a set of vectors $v_1, \ldots, v_n \in \mathbb{R}^k$ $(k \geq 1)$, their **Gram matrix** is the $n \times n$ matrix whose $(i, j)$-th entry is $v_i^T v_j$; its rank is equal to the rank of the system $(v_1, \ldots, v_n)$. The next Lemma 4.10, which states the connection existing between Gram matrices and positive semidefinite matrices, is well known; we give the proof for completeness.

LEMMA 4.10 *Let $A = (a_{ij})_{1\leq i,j\leq n}$ be a symmetric matrix which is positive semidefinite and let $k \leq n$ be its rank. Then, $A$ is a Gram matrix, i.e., there exist vectors $v_1,\ldots,v_n \in \mathbb{R}^k$ such that $a_{ij} = v_i^T v_j$ for $1 \leq i,j \leq n$. Moreover, if $v'_1,\ldots,v'_n$ are other vectors of $\mathbb{R}^k$ such that $a_{ij} = v_i'^T v'_j$ for $1 \leq i,j \leq n$, then $v'_i = T(v_i)$, $1 \leq i \leq n$, for some orthogonal transformation $T$ of $\mathbb{R}^k$. The system $(v_1,\ldots,v_n)$ has rank $k$.*

PROOF. By assumption, $A$ has $k$ non zero eigenvalues which are positive. Hence, there exists an $n\times n$ matrix $Q_0$ such that $A = Q_0 D Q_0^T$, where $D$ is an $n \times n$ matrix whose entries are all zero except $k$ diagonal entries, say with indices $(1,1),\ldots,(k,k)$, equal to 1. Denote by $Q$ the $n\times k$ submatrix of $Q_0$ consisting of its first $k$ columns. Then, $A = QQ^T$ holds, i.e., $a_{ij} = v_i^T v_j$ for $1 \leq i,j \leq n$, where $v_1,\ldots,v_n$ denote the rows of $Q$. It is easy to see that $(v_1,\ldots,v_n)$ has the same rank $k$ as $A$.

Let $Q'$ be another $n \times k$ matrix such that $A = Q'Q'^T$. Both matrices $Q, Q'$ have rank $k$, hence there exists a $k\times k$ non singular matrix $B$ such that $Q' = QB$. Let $Q_1$ be a non singular $k \times k$ submatrix of $Q$ formed, say, by its first $k$ rows, and let $Q'_1$ denote the $k\times k$ submatrix of $Q'$ formed by its first $k$ rows. Then, $Q'_1 = Q_1 B$. From the equality $Q_1Q_1^T = Q'_1(Q'_1)^T$, we obtain that $BB^T$ is the identity matrix, i.e., $B$ is an orthogonal transformation of $\mathbb{R}^k$. ∎

Let $M$ be a symmetric $n \times n$ matrix. The **inertia** $\mathrm{In}(M)$ of $M$ is defined as the triple $(p,q,s)$, where $p$ (resp. $q$,$s$) denotes the number of positive (resp. negative, zero) eigenvalues of $M$; hence, $n = p+q+s$. If $P$ is a nonsingular matrix, then it is well known that the two matrices $M$ and $PMP^T$ have the same inertia (this result is known as **Sylvester's law of inertia**).

LEMMA 4.11 *Let $M$ be a symmetric matrix with the following block decomposition $M = \left(\begin{array}{c|c} A & B \\ \hline B^T & C \end{array}\right)$, where $C$ is nonsingular. Then, $In(M) = In(C) + In(A - BC^{-1}B^T)$.*
(The matrix $A - BC^{-1}B^T$ is also known as the Schur complement of $C$ in $M$.)

PROOF. One verifies easily the following identity:

$$\left(\begin{array}{c|c} I & BC^{-1} \\ \hline 0^T & I \end{array}\right)\left(\begin{array}{c|c} A - BC^{-1}B^T & 0 \\ \hline 0^T & C \end{array}\right)\left(\begin{array}{c|c} I & 0 \\ \hline C^{-1}B^T & I \end{array}\right) = \left(\begin{array}{c|c} A & B \\ \hline B^T & C \end{array}\right) = M.$$

By Sylvester's law of inertia, we obtain that the matrices $M$ and $\left(\begin{array}{c|c} A - BC^{-1}B^T & 0 \\ \hline 0^T & C \end{array}\right)$

have the same inertia. Therefore, $\mathrm{In}(M) = \mathrm{In}(C) + \mathrm{In}(A - BC^{-1}B^T)$. ∎

LEMMA 4.12 *Let $M$ be a symmetric $n \times n$ matrix and let $U$ be a subspace of $\mathbb{R}^n$ such that $x^T M x \leq 0$ holds for all $x \in U$. If $U$ has dimension $n-1$, then $M$ has at most one positive eigenvalue.*

PROOF. Suppose, for contradiction, that $M$ has two positive eigenvalues $\lambda_1$ and $\lambda_2$. Let $u_1$ and $u_2$ be eigenvectors for $\lambda_1$ and $\lambda_2$, respectively, with $u_1^T u_2 = 0$ and $\| u_1 \|_2 = \| u_2 \|_2 = 1$. Let $V$ denote the subspace of $\mathbb{R}^n$ spanned by $u_1$ and $u_2$. Then, $x^T M x > 0$ holds for all $x \in V$, $x \neq 0$; indeed, if $x = a_1 u_1 + a_2 u_2$, then $x^T M x = a_1^2 \lambda_1 + a_2^2 \lambda_2 > 0$ if $(a_1, a_2) \neq (0,0)$. As $U$ and $V$ have respective dimensions $n-1$ and 2, there exists $x \in U \cap V$ with $x \neq 0$. Then, $x^T M x \leq 0$ since $x \in U$ and $x^T M x > 0$ since $x \in V$, yielding a contradiction. ∎

Let $(X, d)$ be a distance space with $X = \{1, \ldots, n\}$. Let $D$ denote the corresponding distance matrix; it is the $n \times n$ symmetric matrix whose $(i,j)$-th entry is $d(i,j)$ (with diagonal terms all equal to 0). We also consider the symmetric $(n+1) \times (n+1)$ matrix $M(X, d)$ defined by

$$M(X,d) = \left( \begin{array}{c|c} D & -e \\ \hline -e^T & 0 \end{array} \right) \tag{4.13}$$

where $e = (1, \ldots, 1) \in \mathbb{R}^n$. (The bordered matrix $M(X,d)$ is closely related to the Cayley-Menger matrix of the distance space $(X,d)$; indeed, the latter is defined in the same way but it has the borders $e, e^T$ instead of $-e, -e^T$.)

Let $x_0 \in X$ and let $p = \xi_{x_0}(d)$ denote the image of $d$ under the covariance mapping $\xi_{x_0}$ (recall relation (3.8)). Let $P(X,d)$ denote the $(n-1) \times (n-1)$ matrix whose $(i,j)$-th entry is $p(i,j)$ for $i, j \in X \setminus \{x_0\}$. The next Lemma 4.14 establishes a relation between the ranks of the matrices $M(X,d)$, $P(X,d)$, and $(I - \frac{1}{n}J)D(I - \frac{1}{n}J)$, where $I$ denotes the identity matrix and $J$ the all ones matrix.

LEMMA 4.14 *$\mathrm{rank}(P(X,d)) = \mathrm{rank}(M(X,d)) - 2 = \mathrm{rank}((I - \frac{1}{n}J)D(I - \frac{1}{n}J))$.*

PROOF. The equalities $\mathrm{rank}(M(X,d)) = \mathrm{rank}(P(X,d)) + 2$ and $\mathrm{rank}((I - \frac{1}{n}J)D(I - \frac{1}{n}J)) = \mathrm{rank}(P(X,d))$ can be checked by doing some row/column manipulations on $M(X,d)$ and $(I - \frac{1}{n}J)D(I - \frac{1}{n}J)$, and applying relation (3.8). ∎

We now present a classical result, due to Schoenberg [Sch35, Sch38b], on the characterization of $L_2$-embeddable distance spaces. Theorem 4.16 gives a characterization of the distance spaces that are isometrically $L_2$-embeddable in terms of the negative type inequalities and Theorem 4.15 is an equivalent formulation in the context of covariances.

Theorem 4.15 *Let $p$ be a symmetric function on $X$. Then, $p$ is of positive type on $X$ if and only if $p$ is an $\mathbb{R}$-covariance on $X$.*

Proof. Suppose first that $p$ is an $\mathbb{R}$-covariance on $X$. Then, $p(x,y) = \int_\Omega f_x(\omega) f_y(\omega)\mu(d\omega)$ for all $x, y \in X$, where $f_x$ are real valued functions of $L_2(\Omega, \mathcal{A}, \mu)$. Let $b \in \mathbb{Z}^X$. Then, $\sum_{x,y\in X} b_x b_y p(x,y) = \int_\Omega |f_x(\omega)|^2 \mu(d\omega) \geq 0$. This shows that $p$ is of positive type on $X$. Conversely, suppose that $p$ is of positive type on $X$. We show that $p$ is an $\mathbb{R}$-covariance on $X$. From the finitude result from Theorem 1.4 (for $p = 2$) and Lemma 3.10, we can suppose that $X$ is finite. By assumption, the matrix $(p(i,j))_{i,j\in X}$ is positive semidefinite and, thus, is a Gram matrix, by Lemma 4.10. This shows that $p$ is an $\mathbb{R}$-covariance on $X$. ∎

Theorem 4.16 *Let $(X,d)$ be a distance space. Then, $(X,d)$ is of negative type if and only if $(X,\sqrt{d})$ is $L_2$-embeddable.*

Proof. It follows from Theorem 4.15, after applying Lemmas 3.10 and 4.8. ∎

When the distance space $(X,\sqrt{d})$ is $\ell_2$-embeddable with $X$ finite, the associated distance matrix $D$ is also known as a **Euclidian distance matrix** in the literature (see, e.g., [Gow85, HWLT91]).

Proposition 4.17 *Let $(X,d)$ be a finite distance space of negative type. Then, $(X,\sqrt{d})$ is $\ell_2$-embeddable and its minimum $\ell_2$-dimension $m_2(X,\sqrt{d})$ satisfies:*
$m_2(X,\sqrt{d}) = rank(P(X,d)) = rank(M(X,d)) - 2 = rank((I - \frac{1}{n}J)D(I - \frac{1}{n}J))$.

Proof. In view of Lemma 4.14, we have only to check that $m_2(X,\sqrt{d}) = \mathrm{rank}(P(X,d))$. Set $k = m_2(X,\sqrt{d})$ and $r = \mathrm{rank}(P(X,d))$. By Lemma 4.10, $P(X,d)$ is the Gram matrix of a system of vectors $v_i \in \mathbb{R}^r$ for $i \in X \setminus \{x_0\}$. Then, $v_i,\ i \in X$, provide an $\ell_2$-embedding of $(X,\sqrt{d})$, if we set $v_{x_0} = 0$. This implies that $r \geq k$. On the other hand, there exist vectors $u_i \in \mathbb{R}^k,\ i \in X$, such that $d(i,j) = \| u_i - u_j \|_2^2$ for all $i,j \in X$. We can suppose without loss of generality that $u_{x_0} = 0$. Then, $P(X,d)$ coincides with the Gram matrix of $u_i,\ i \in X \setminus \{x_0\}$, which implies that $r \leq k$. Hence, $r = k$ holds. ∎

Therefore, the parameter $m_2(n)$, which we recall is defined as the minimum $\ell_2$-dimension of an $\ell_2$-embeddable distance space on $n$ points, satisfies

$$m_2(n) = n - 1. \tag{4.18}$$

We now present two additional equivalent characterizations for distance spaces of negative type.

THEOREM 4.19 *Let $(X,d)$ be a finite distance space with $X = \{1,\dots,n\}$. Let $D$ be the associated $n \times n$ distance matrix and let $M(X,d)$ be the bordered matrix defined by* (4.13)*. Consider the following assertions.*
*(i) $(X,d)$ is of negative type.*
*(ii) The matrix $(I - \frac{1}{n}J)(-D)(I - \frac{1}{n}J)$ is positive semidefinite.*
*(iii) The matrix $M(X,d)$ has exactly one positive eigenvalue.*
*(iv) The matrix $D$ has exactly one positive eigenvalue.*
*Then, $(i) \Longleftrightarrow (ii)$* [Gow85]*, $(i) \Longleftrightarrow (iii)$* [HW88]*, and $(i) \Longrightarrow (iv)$. Moreover, if $D$ has a constant row sum, then $(i) \Longleftrightarrow (iv)$* [KS93]*.*

PROOF. $(i) \Longleftrightarrow (ii)$ Set $K = I - \frac{1}{n}J$ and $A = K(-D)K$. Then, for $x \in \mathbb{R}^n$, we have that $x^T A x = y^T(-D)y$, setting $y = Kx$. One checks easily that the range of $K$ consists of the vectors $y \in \mathbb{R}^n$ such that $\sum_{1 \le i \le n} y_i = 0$. Therefore, we obtain that $A$ is positive semidefinite if and only if $y^T(-D)y \geq 0$ for all $y \in \mathbb{R}^n$ such that $\sum_{1 \le i \le n} y_i = 0$, i.e., $(X,d)$ is of negative type.
$(i) \Longleftrightarrow (iii)$ Let $Q$ be an orthogonal $n \times n$ matrix such that $Qe$ is equal to the vector $e' = (0,\dots,0,1)$. Set $D' = QDQ^T$, $Q' = \left(\begin{array}{c|c} Q & 0 \\ \hline 0^T & 1 \end{array}\right)$, and $M' = Q'M(X,d)Q'^T$. Hence,

$$M' = \left(\begin{array}{c|c} D' & -e' \\ \hline -e'^T & 0 \end{array}\right) = \left(\begin{array}{c|c|c} D'_0 & b & 0 \\ \hline b^T & \beta & -1 \\ \hline 0 & -1 & 0 \end{array}\right),$$ if we let $D'$ be pictured as $\left(\begin{array}{c|c} D'_0 & b \\ \hline b^T & \beta \end{array}\right)$,

where $D'_0$ is an $(n-1) \times (n-1)$ matrix and $b \in \mathbb{R}^{n-1}$. As the matrix $\left(\begin{array}{c|c} \beta & -1 \\ \hline -1 & 0 \end{array}\right)$ is nonsingular, we can apply Lemma 4.11 for computing the inertia of $M'$. We obtain that $\mathrm{In}(M') = \mathrm{In}(D'_0) + \mathrm{In}\left(\begin{array}{c|c} \beta & -1 \\ \hline -1 & 0 \end{array}\right)$. By Sylvester's law of inertia, $M(X,d)$ and $M'$

have the same inertia and, in particular, both $M(X,d)$ and $M'$ have the same number of positive eigenvalues. One checks easily that the matrix $\left(\begin{array}{c|c} \beta & -1 \\ \hline -1 & 0 \end{array}\right)$ has exactly one positive eigenvalue. Therefore, $M(X,d)$ has one positive eigenvalue if and only if $D'_0$ has no positive eigenvalue, i.e., $x^T D'_0 x \leq 0$ for all $x \in \mathbb{R}^{n-1}$. But, $x^T D'_0 x \leq 0$ holds for all $x \in \mathbb{R}^{n-1}$ if and only if $y^T D' y \leq 0$ holds for all $y \in \mathbb{R}^n$ such that $e'^T y = 0$ (because $y^T D' y = x^T D'_0 x + x_n(2b^T x + \beta x_n)$, if $y = (x, x_n)$) or, equivalently, $z^T D z \leq 0$ holds for all $z \in \mathbb{R}^n$ such that $e^T z = 0$, i.e., $(X,d)$ is of negative type.
$(i) \Longrightarrow (iv)$ The matrix $D$ has at least one positive eigenvalue since $D$ has its diagonal terms equal to 0. If $(X,d)$ is of negative type then, by Lemma 4.12, $D$ has at most one positive eigenvalue since $x^T D x \leq 0$ holds for all $x$ in an $(n-1)$-dimensional subspace of $\mathbb{R}^n$. Therefore, $D$ has one positive eigenvalue.
Finally, suppose that $D$ has a constant row sum, equal to $s$, and that $(iv)$ holds. Then, $s$ is the positive eigenvalue of $D$ since $De = se$ holds. This implies that the matrix $\frac{s}{n}J - D$ is positive semidefinite. Hence, by Lemma 4.10, there exists an $n \times k$ matrix $X$ such that $\frac{s}{n}J - D = XX^T$. Let $v_1, \ldots, v_n$ denote the row vectors of $X$. Then, we have that $\frac{s}{n} - d(i,j) = v_i^T v_j$ for all $i,j$ or, equivalently, $d(i,j) = \| v_i - v_j \|_2^2$ for all $i,j$. This shows that $(X, \sqrt{d})$ is $\ell_2$-embeddable, i.e., by Theorem 4.16, that $(X,d)$ is of negative type. ∎

## 4.3 A chain of implications

We summarize in this Section the implications existing between the properties of being $L_1$-, $L_2$-embeddable, of negative type, and hypermetric.

Theorem 4.20 *Let $(X,d)$ be a distance space. Consider the following assertions.*
*(i) $(X,d)$ is $L_2$-embeddable.*
*(ii) $(X,d)$ is $L_1$-embeddable.*
*(iii) $(X,d)$ is hypermetric.*
*(iv) $(X,d)$ is of negative type.*
*(v) $(X,\sqrt{d})$ is $L_2$-embeddable.*
*We have the chain of implications $(i) \Longrightarrow (ii) \Longrightarrow (iii) \Longrightarrow (iv) \Longleftrightarrow (v)$.*

Proof. $(i) \Longrightarrow (ii)$ is a classical result in analysis. For $(ii) \Longrightarrow (iii)$, it suffices to check that every finite subspace of $(X,d)$ is hypermetric, i.e., that every member of the cut cone satisfies the hypermetric inequalities or, equivalently, that each cut semimetric satisfies the hypermetric inequalities. Indeed, given a subset $S$ of $V_n$ and $b_1, \ldots, b_n \in \mathbb{Z}$ with $\sum_{1 \leq i \leq n} b_i = 1$, we have that
$\sum_{1 \leq i < j \leq n} b_i b_j \delta(S)_{ij} = (\sum_{i \in S} b_i)(\sum_{i \in V_n \setminus S} b_i) = \sum_{i \in S} b_i (1 - \sum_{i \in S} b_i) \leq 0$

since all $b_i$'s are integers. For $(iii) \Longrightarrow (iv)$, we use the equivalent formulations of the negative type and hypermetric conditions given in Lemmas 4.8 and 4.9. Finally, $(iv) \Longleftrightarrow (v)$ holds by Theorem 4.16. ∎

Example 4.21 Let $d(K_{2,3})$ denote the path metric of the complete bipartite graph $K_{2,3}$ with node set $\{x_1, x_2, x_3\} \cup \{y_1, y_2\}$. Then, $d(K_{2,3})(x_i, x_j) = d(K_{2,3})(y_1, y_2) = 1$ and $d(K_{2,3})(x_i, y_j) = 2$. Hence, $d(K_{2,3})$ violates the pentagonal inequality (4.4). Therefore, $d(K_{2,3})$ is not hypermetric and, thus, not $L_1$-embeddable.

From the implication $(ii) \Longrightarrow (iii)$ from Theorem 4.20, we have the inclusion $C_1(X) \subseteq \mathrm{HYP}(X)$ (recall that $C_1(X)$ is the cone of all $L_1$-embeddable distances on $X$). This inclusion is, in general, strict. It is strict, in particular, if $7 \leq |X| < \infty$ or $X = \mathbb{N}$.

For showing the strict inclusion $\mathrm{CUT}_n \subset \mathrm{HYP}_n$ for $n \geq 7$, it suffices to exhibit an inequality which defines a facet for $\mathrm{CUT}_n$ and is not hypermetric. Many such inequalities are described in Chapter ??? (on facets).

For showing the strict inclusion $C_1(\mathbb{N}) \subset \mathrm{HYP}(\mathbb{N})$, consider the distance $d$ on $\mathbb{N}$ obtained by taking the spherical $t$-extension of the path metric of the Schläfli graph $G_{27}$, i.e., $d_{ij}$ is the shortest length of a path joining $i$ and $j$ in $G_{27}$ if $i$ and $j$ are both nodes of $G_{27}$ and $d_{ij} = t$ otherwise. For $t \geq \frac{4}{3}$, $d$ is hypermetric but $d$ is not $L_1$-embeddable (since the path metric of $G_{27}$ lies on an extreme ray of the hypermetric cone on 27 points) ([Gri92]).

However, there are many examples of classes of distance spaces $(X, d)$ for which the properties of being hypermetric and $L_1$-embeddable are equivalent. We present such examples with $X$ infinite in Sections 5.1 and 5.2.

We summarize in Remark 4.22 below a list of distance spaces $(X, d)$ for which $L_1$-embeddability can be characterized by a set $\mathcal{I}$ of inequalities that are all hypermetric or of negative type.

Remark 4.22 $(i)$ $(V_n, d)$ with $n \leq 6$; $\mathcal{I}$ consists of the hypermetric inequalities, i.e., $\mathrm{CUT}_n = \mathrm{HYP}_n$ for $n \leq 6$ ([Dez61] for $n \leq 5$ and [AM89] for $n = 6$).

$(ii)$ A normed space $(\mathbb{R}^m, d_{||.||})$; $\mathcal{I}$ consists of the negative type inequalities,
a normed space $(\mathbb{R}^m, d_{||.||})$ whose unit ball is a polytope; $\mathcal{I}$ consists of the 7-gonal inequalities (see Theorems 5.1 and 5.2).

$(iii)$ $(L, d_v)$ where $(L, \preceq)$ is a lattice with distance $d_v(x, y) = v(v \vee y) - v(x \wedge y)$ for $x, y \in L$; $\mathcal{I}$ consists of the 5-gonal inequalities or, equivalently, $\mathcal{I}$ consists of the negative type inequalities (see Theorem 5.6 and Example 5.9).

$(iv)$ $(\mathcal{A}, d)$ where $\mathcal{A}$ is a family of subsets of a set $\Omega$ which is stable under the symmetrical difference and $d(A,B) = v(A \triangle B)$ for $A, B \in \mathcal{A}$ with $v$ nonnegative and $v(\emptyset) = 0$; $\mathcal{I}$ consists of the inequalities of negative type (see Example 5.10).

$(v)$ The graphic space $(V, d(G))$ where $G$ is a connected bipartite graph with node set $V$; $\mathcal{I}$ consists of the 5-gonal inequalities (see Chapter ????(on graphs)).

$(vi)$ The graphic space $(V, d(G))$ where $G$ is a connected graph on at least 38 nodes and having a node adjacent to all other nodes; $\mathcal{I}$ consists of the negative type inequalities and the 5-gonal inequalities (see Chapter ????(on graphs)).

## 4.4 The direct sum and tensor product operations

We present two operations, the direct sum and the tensor product, which preserve, respectively, $L_1$-embeddability and $\{0,1\}$-covariances.

DEFINITION 4.23 *(i) Let $(X_1, d_1)$ and $(X_2, d_2)$ be two distance spaces. Their* **direct sum** *is the distance space $(X_1 \times X_2, d_1 \oplus d_2)$, where*

$$d_1 \oplus d_2((x_1,x_2),(y_1,y_2)) = d_1(x_1,y_1) + d_2(x_2,y_2) \text{ for all } x_1, y_1 \in X_1, x_2, y_2 \in X_2.$$

*(ii) Let $p_1 : X_1 \times X_1 \longrightarrow \mathbb{R}$ and $p_2 : X_2 \times X_2 \longrightarrow \mathbb{R}$ be two symmetric functions. Their* **tensor product** *is the symmetric function $p_1 \otimes p_2 : (X_1 \times X_2) \times (X_1 \times X_2) \longrightarrow \mathbb{R}$ defined by*

$$p_1 \otimes p_2((x_1,x_2),(y_1,y_2)) = p_1(x_1,y_1)p_2(x_2,y_2) \text{ for all } x_1, y_1 \in X_1, x_2, y_2 \in X_2.$$

PROPOSITION 4.24 [Ass79, Ass80b] *(i) If $(X_1, d_1)$ and $(X_2, d_2)$ are $L_1$-embeddable, then their direct sum $(X_1 \times X_2, d_1 \oplus d_2)$ is $L_1$-embeddable.*
*(i) If $p_1$ is a $\{0,1\}$-covariance on $X_1$ and $p_2$ is a $\{0,1\}$-covariance on $X_2$, then their tensor product $p_1 \otimes p_2$ is a $\{0,1\}$-covariance on $X_1 \times X_2$.*

PROOF. $(i)$ By assumption, there exist a measure space $(\Omega_i, \mathcal{A}_i, \mu_i)$ and an isometric embedding $\phi_i$ of $(X_i, d_i)$ into $L_1(\Omega_i, \mathcal{A}_i, \mu_i)$, for $i = 1, 2$. Let $(\Omega = \Omega_1 \cup \Omega_2, \mathcal{A}, \mu)$ denote the measure space obtained by extending $\mathcal{A}_i$ and $\mu_i$ to $\Omega_1 \cup \Omega_2$. We obtain an isometric embedding of $(X_1 \times X_2, d_1 \oplus d_2)$ into $(\Omega, \mathcal{A}, \mu)$ by setting $\phi(x_1,x_2)(\omega) = \phi_i(x_i)(\omega)$ if $\omega \in \Omega_i$, for $i = 1, 2$. Indeed,

$$\begin{aligned} d_1 \oplus d_2((x_1,x_2),(y_1,y_2)) &= d_1(x_1,y_1) + d_2(x_2,y_2) \\ &= \| \phi_1(x_1) - \phi_1(y_1) \| + \| \phi_2(x_2) - \phi_2(y_2) \| \\ &= \| \phi(x_1,x_2) - \phi(y_1,y_2) \| . \end{aligned}$$

$(ii)$ By assumption, there exist a measure space $(\Omega_i, \mathcal{A}_i, \mu_i)$ and a mapping $x \in X_i \mapsto A_x^{(i)} \in (\mathcal{A}_i)_{\mu_i}$ such that $p_i(x,y) = \mu_i(A_x^{(i)} \cap A_y^{(i)})$ for all $x, y \in X_i$, for $i = 1, 2$. Set

$\Omega = \Omega_1 \times \Omega_2$, $\mathcal{A} = \{A_1 \times A_2 : A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2\}$, and take $\mu = \mu_1 \otimes \mu_2$. Then, we have that $p_1 \otimes p_2((x_1, x_2), (y_1, y_2)) = p_1(x_1, y_1) p_2(x_2, y_2) = \mu_1(A^{(1)}_{x_1} \cap A^{(1)}_{y_1}) \mu_2(A^{(2)}_{x_2} \cap A^{(2)}_{y_2})$ $= \mu_1 \otimes \mu_2((A^{(1)}_{x_1} \times A^{(2)}_{x_2}) \cap (A^{(1)}_{y_1} \times A^{(2)}_{y_2}))$. This shows that $p_1 \otimes p_2$ is a $\{0,1\}$-covariance on $X_1 \times X_2$. ∎

COROLLARY 4.25 *(i) If $(X, d_1)$ and $(X, d_2)$ are $L_1$-embeddable, then $(X, d_1 + d_2)$ is $L_1$-embeddable.*
*(ii) If $p_1$ and $p_2$ are $\{0,1\}$-covariances on $X$, then $p_1 p_2$ is a $\{0,1\}$-covariance on $X$.*

PROOF. $(i)$ follows from Theorems 1.4 and 2.11. In a more elementary way, $(i)$ follows from Proposition 4.24 $(i)$, since $(X, d_1 + d_2)$ is a subspace from $(X_1 \times X_2, d_1 \oplus d_2)$ (via the embedding $x \mapsto (x, x)$). $(ii)$ follows from Proposition 4.24 $(ii)$, since $p_1 p_2$ identifies with the restriction of $p_1 \otimes p_2$ to the diagonal subset $\{(x,x) : x \in X\}$ of $X \times X$. ∎

# 5 Two cases of complete characterization of $L_1$-embeddability

We present in this Section two classes of distance spaces for which $L_1$-embeddability can be fully characterized using only hypermetric or negative type inequalities. The first class consists of metric spaces arising from normed spaces and the second one consists of metric spaces arising from metric lattices equipped with a valuation.

## 5.1 $L_1$-metrics from normed spaces

Let $(E, \|\,.\,\|)$ be a normed space. We consider the associated metric space $(E, d_{\|.\|})$, where $d_{\|.\|}$ is the norm metric defined by

$$d_{\|.\|}(x, y) = \| x - y \|$$

for all $x, y \in E$.

In this Section, we give a characterization of the norms on $E = \mathbb{R}^m$ for which the metric space $(\mathbb{R}^m, d_{\|.\|})$ is $L_1$-embeddable.

We first recall some definitions.

Let $K$ be a **convex body** in $\mathbb{R}^m$, i.e., $K$ is a nonempty convex compact subset of $\mathbb{R}^m$. We suppose that the origin 0 belongs to the interior of $K$. $K$ is said to be **centrally symmetric** if $-x \in K$ for all $x \in K$. The **polar** $K^*$ of a convex body $K$ is the convex body defined by

$$K^* = \{x \in \mathbb{R}^m : x^T y \leq 1 \text{ for all } y \in K\}.$$

The notions of convex bodies and of norms are, in fact, equivalent, in the following sense. First, if $\| \, . \, \|$ is a norm on $\mathbb{R}^m$ then its **unit ball**

$$B = \{x \in \mathbb{R}^m : \| x \| \leq 1\}$$

is clearly a centrally symmetric convex body. Conversely, let $K$ be a centrally symmetric convex body (containing the origin in its interior). The **support function** $h(K, .) : \mathbb{R}^m \rightarrow \mathbb{R}$ of $K$ is defined by

$$h(K, x) = \max(x^T y : y \in K)$$

for $x \in \mathbb{R}^m$. One can easily check that $h(K, .)$ defines a norm on $\mathbb{R}^m$, whose unit ball is the polar $K^*$ of $K$. This norm can be alternatively defined by

$$h(K, x) = \min(\lambda > 0 : \frac{x}{\lambda} \in K^*)$$

for all $x \in \mathbb{R}^m$.

A convex polytope is called a **zonotope** if it is the vector sum of some line segments. A convex body which can be approximated by zonotopes with respect to the Blaschke-Hausdorff metric is called a **zonoid**. Zonotopes and zonoids are central objects in convex geometry and they are also relevant to many other fields (see, e.g., [Bol69, SW83] for a survey). They are, in particular, relevant to the topic of $L_1$-metrics as we explain below.

We now present several equivalent characterizations for $L_1$-embeddability of a normed metric space $(\mathbb{R}^m, d_{\|.\|})$.

THEOREM 5.1 (see [Bol69, SW83]) *Let $\| \, . \, \|$ be a norm on $\mathbb{R}^m$ and let $B$ be its unit ball. The following assertions are equivalent.*
*(i) $d_{\|.\|}$ is of negative type.*
*(ii) $d_{\|.\|}$ is hypermetric.*
*(iii) $(\mathbb{R}^m, d_{\|.\|})$ is $L_1$-embeddable.*
*(iv) The polar of $B$ is a zonoid.*

Precise reference for the equivalence $(i) \Longleftrightarrow (ii) \Longleftrightarrow (iv)$ can be found in [SW83] and $(iii) \Longleftrightarrow (iv)$ is proved in [Bol69]. $L_1$-embeddability of norm metrics can be characterized by much simpler inequalities when the unit ball of the normed space is a polytope.

THEOREM 5.2 ([Ass80a, Ass84, Wit78], see [SW83]). *Let* $\| \,.\, \|$ *be a norm on* $\mathbb{R}^m$ *for which the unit ball* $B$ *is a polytope. The following assertions are equivalent.*
*(i)* $\| \,.\, \|$ *satisfies Hlawka's inequality*

$$\| x \| + \| y \| + \| z \| + \| x+y+z \| \geq \| x+y \| + \| x+z \| + \| y+z \|$$

*for all* $x, y, z \in \mathbb{R}^m$.
*(ii)* $\| \,.\, \|$ *satisfies the 7-gonal inequality*

$$\sum_{1 \leq i < j \leq 4} \| x_i - x_j \| + \sum_{1 \leq h < k \leq 3} \| y_h - y_k \| \leq \sum_{\substack{1 \leq i \leq 4 \\ 1 \leq k \leq 3}} \| x_i - y_k \|$$

*for all* $x_1, x_2, x_3, x_4, y_1, y_2, y_3 \in \mathbb{R}^m$.
*(iii) The polar of* $B$ *is a zonotope.*
*(iv)* $(\mathbb{R}^m, d_{\|.\|})$ *is* $L_1$*-embeddable.*

Actually, the implication $(ii) \Longrightarrow (i)$ of Theorem 5.2 remains valid for general norms. Namely, if an arbitrary norm on $\mathbb{R}^m$ satisfies the 7-gonal inequality, then it also satisfies Hlawka's inequality ([Ass84]).

The above results can be partially extended to the more general concept of projective metrics. A continuous metric $d$ on $\mathbb{R}^m$ is called a **projective metric** if it satisfies $d(x,z) = d(x,y)+d(y,z)$ for any collinear points $x, y, z$ lying in that order on a common line. Clearly, every norm metric is projective. The cone of projective metrics is the object considered by the fourth Hilbert problem in $\mathbb{R}^m$ (see [Ale88], [Amb82]).

We have the following characterization of $L_1$-embeddability for projective metrics.

THEOREM 5.3 [Ale88] *Let* $d$ *be a projective metric on* $\mathbb{R}^m$*. The following assertions are equivalent.*
*(i)* $(\mathbb{R}^m, d)$ *is* $L_1$*-embeddable.*
*(ii)* $d$ *is hypermetric.*
*(iii) There exists a positive Borel measure* $\mu$ *on the hyperplanesets of* $\mathbb{R}^m$ *satisfying*

$$\begin{cases} \mu([[x]]) = 0 & \text{for all } x \in \mathbb{R}^m \\ 0 < \mu([[x,y]]) < \infty & \text{for all } x \neq y \in \mathbb{R}^m \end{cases}$$

*and such that* $d$ *is defined by the following formula (called* **Crofton formula***):*

$$d(x,y) = \mu([[x,y]]) \text{ for } x, y \in \mathbb{R}^m,$$

*where* $[[x,y]]$ *denotes the set of hyperplanes meeting the segment* $[x,y]$.

In dimension $m = 2$, Theorem 5.3 (ii) always holds (see [Ale88]), i.e., every projective metric on $\mathbb{R}^2$ is $L_1$-embeddable. On the other hand, the norm metric $d_{\ell_\infty}$ arising from the norm $\| x \|_\infty = \max(|x_1|, |x_2|, |x_3|)$ in $\mathbb{R}^3$ is not $L_1$-embeddable since it is not hypermetric. Indeed, the points $x_1 = (1,1,0)$, $x_2 = (1,-1,0)$, $x_3 = (-1,1,0)$, $y_1 = (0,0,0)$ and $y_2 = (0,0,1)$ violate the 5-gonal inequality (4.4) ([Kel70]).

## 5.2 $L_1$-metrics from lattices

In this Section, we consider a class of metric spaces arising from lattices. A good reference on lattices is [Bir67].

Let $(L, \preceq)$ be a **lattice** (possibly infinite), i.e., a partially ordered set in which any two elements $x, y \in L$ have a join $x \vee y$ and a meet $x \wedge y$. A function $v : L \longrightarrow \mathbb{R}_+$ satisfying

(5.4) $$v(x \vee y) + v(x \wedge y) = v(x) + v(y) \text{ for all } x, y \in L.$$

is called a **valuation** on $L$. The valuation $v$ is said to be **isotone** if $v(x) \leq v(y)$ whenever $x \preceq y$ and **positive** if $v(x) < v(y)$ whenever $x \preceq y, x \neq y$. Set

(5.5) $$d_v(x, y) = v(x \vee y) - v(x \wedge y) \text{ for all } x, y \in L.$$

One can easily check that $(L, d_v)$ is a semimetric space if $v$ is an isotone valuation on $L$ and $(L, d_v)$ is a metric space if $v$ is a positive valuation on $L$; in the latter case, $L$ is called a **metric lattice** (see [Bir67]). Clearly, every metric lattice is **modular**, i.e., satisfies $x \wedge (y \vee z) = (x \wedge y) \vee z$ for all $x, y, z$ with $z \preceq x$. A lattice is called **distributive** if $x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$ for all $x, y, z$. The following result gives a characterization of the $L_1$-embeddable metric lattices.

THEOREM 5.6 [Kel70] *Let $L$ be a metric lattice with positive valuation $v$. The following assertions are equivalent.*
*(i) $L$ is a distributive lattice.*
*(ii) $(L, d_v)$ is 5-gonal.*
*(iii) $(L, d_v)$ is hypermetric.*
*(iv) $(L, d_v)$ is $L_1$-embeddable.*

PROOF. It suffices to show the implications $(ii) \Rightarrow (i)$ and $(i) \Rightarrow (iv)$.
$(ii) \Rightarrow (i)$. Using the definition of the valuation $v$ and applying the 5-gonal inequality (4.4) to the points $x_1 = x \vee y$, $x_2 = x \wedge y$, $x_3 = z$, $y_1 = x$, $y_2 = y$, we obtain the inequality: $2(v(x \vee y \vee z) - v(x \wedge y \wedge z)) \leq v(x \vee y) + v(x \vee z) + v(y \vee z) - v(x \wedge y) - v(x \wedge z) - v(y \wedge z)$. By applying again the 5-gonal inequality to the points $x_1 = x$, $x_2 = y$, $x_3 = z$, $y_1 = x \vee y$,

$y_2 = x \wedge y$, we obtain the reverse inequality. Therefore, the equality holds in the above inequality. In fact, this condition of equality is equivalent to $L$ being distributive (see [Bir67]).
$(i) \Rightarrow (iv)$. Take a finite subset $L_0$ of $L$. We show that $(L_0, d_v)$ is $L_1$-embeddable. Let $K$ be the sublattice of $L$ generated by $L_0$. Suppose $K$ has length $n$. Then, $K$ is isomorphic to a ring $\mathcal{N}$ of subsets of a set $X$, $|X| = n$ ("ring" means closed under $\cup$ and $\cap$)(see [Bir67] p.58). Via this isomorphism, we have a valuation, again denoted by $v$, defined on $\mathcal{N}$. We can assume without loss of generality that $v(\emptyset) = 0$. Then, $v$ can be extended to a valuation $v^*$ on $2^X$ satisfying $v^*(S) = \sum_{x \in S} v^*(\{x\})$ for $S \subseteq X$. Now, if $x \longmapsto S_x$ is the isomorphism from $K$ to $\mathcal{N}$, then we have the embedding $x \longmapsto S_x$ from $(L_0, d_v)$ to $(2^X, v^*)$ which is isometric. Indeed, $d_v(x, y) = v(x \vee y) - v(x \wedge y) = v(S_x \cup S_y) - v(S_x \cap S_y)$ $= v^*(S_x \cup S_y) - v^*(S_x \cap S_y) = v^*(S_x \triangle S_y)$. This shows that every finite subset of $(L, d_v)$ is $L_1$-embeddable, and thus, from Theorem 1.4, $(L, d_v)$ is $L_1$-embeddable. ∎

EXAMPLE 5.7 [Ass79] Let $(\mathbb{N}^*, \preceq)$ denote the lattice consisting of the set $\mathbb{N}^*$ of positive integers with order relation $x \preceq y$ if $x$ divides $y$. Then, for $x, y \in \mathbb{N}^*$, $x \wedge y$ is the g.c.d. of $x$ and $y$ and $x \vee y$ is their l.c.m.. One checks easily that $(\mathbb{N}^*, \preceq)$ is a distributive lattice. Therefore, $(\mathbb{N}^*, d_v)$ is $L_1$-embeddable for every positive valuation $v$ on $\mathbb{N}^*$. For instance, $x \in \mathbb{N}^* \longmapsto v(x) := \log x$ is a positive valuation on $\mathbb{N}^*$ ; hence, the metric $d_v$, defined by $d_v(x, y) = \log(\frac{l.c.m.(x,y)}{g.c.d.(x,y)})$ for all integers $x, y \geq 1$, is $L_1$-embeddable.

We now present an analogue of Theorem 5.6 in the context of semigroups. We recall that a commutative semigroup $(S, +)$ consists of a set $S$ equipped with a composition rule $+$ which is commutative and associative. We assume the existence of a neutral element denoted by 0.

THEOREM 5.8 ([BC76], see [Ass79, Ass80b]) *Let $(S, +)$ be a commutative semigroup with neutral element 0 and let $v : S \longmapsto \mathbb{R}_+$ be a mapping such that $v(0) = 0$. Set*

$$D_v(x, y) = 2v(x + y) - v(2x) - v(2y) \text{ for } x, y \in S.$$

*Assume that one of the following assertions (i) or (ii) holds.*
*(i) $(S, +)$ is a group.*
*(ii) For each $x \in S$, there exists an integer $n \geq 1$ such that $2nx = x$.*
*Then, $(S, D_v)$ is $L_1$-embeddable if and only if $(S, D_v)$ is of negative type.*

PROOF. We only give a sketch of the proof of the implication: if $(S, D_v)$ is of negative type, then $(S, D_v)$ is $L_1$-embeddable.

Let $p_v : S^2 \longrightarrow \mathbb{R}$ denote the symmetric function obtained by applying the covariance transformation $\xi$ to $D_v$, i.e., $p_v(x,y) = \frac{1}{2}(D_v(x,0) + D_v(y,0) - D_v(x,y)) = v(x) + v(y) - v(x+y)$ for $x, y \in S$. By Lemma 3.11, showing that $(S, D_v)$ is $L_1$-embeddable amounts to showing that $p_v$ is a $\{0,1\}$-covariance on $S$. By assumption, $(S, D_v)$ is of negative type or, equivalently, by Lemma 4.8, $p_v$ is of positive type on $S$. Berg and Christensen [BC76] show that, under this assumption, the function $v$ is of the form

$$v(x) = h(x) + \int_{\hat{S}-\{\hat{1}\}} (1-\rho(x))\mu(d\rho) \text{ for all } x \in S,$$

where
- the function $h : S \longrightarrow \mathbb{R}_+$ satisfies $h(x+y) = h(x) + h(y)$ for all $x, y \in S$,
- $\hat{S}$ denotes the set of characters on $S$, i.e., of the functions $\rho : S \longrightarrow [-1,1]$ satisfying $\rho(x+y) = \rho(x)\rho(y)$ for all $x, y \in S$ and $\rho(0) = 1$, and $\hat{1}$ is the unit character defined by $\hat{1}(x) = 1$ for all $x \in S$,
- $\mu$ is a nonnegative Radon measure on $\hat{S} - \{\hat{1}\}$ such that $\int_{\hat{S}-\{\hat{1}\}}(1-\rho(x))\mu(d\rho) < \infty$ for all $x \in S$.

Therefore, we have that $p_v(x,y) = \int_{\hat{S}-\{\hat{1}\}}(1-\rho(x))(1-\rho(y))\mu(d\rho)$ for all $x, y \in S$. In case $(i)$, every character on $S$ takes only values $\pm 1$. Setting $A_x = \{\rho \in \hat{S} : \rho(x) = -1\}$ for $x \in S$, we obtain that $p_v(x,y) = 4\mu(A_x \cap A_y)$ for all $x, y \in S$. In case $(ii)$, every character on $S$ takes only values $0, 1$. Setting $A_x = \{\rho \in \hat{S} : \rho(x) = 0\}$ for $x \in S$, we obtain that $p_v(x,y) = \mu(A_x \cap A_y)$ for all $x, y \in S$. Therefore, $p_v$ is a $\{0,1\}$-covariance, i.e., $(S, D_v)$ is $L_1$-embeddable. ∎

EXAMPLE 5.9 Let $(L, \preceq)$ be a lattice and let $S$ be a subset of $L$ which is stable under the join operation $\vee$ of $L$ and contains the least element $0$ of $L$. Then, $(S, \vee)$ is a commutative semigroup satisfying Theorem 5.8 $(ii)$. Therefore, given a mapping $v : S \longrightarrow \mathbb{R}_+$ such that $v(0) = 0$, $(S, D_v)$ is $L_1$-embeddable if and only if $(S, D_v)$ is of negative type, with $D_v(x,y) = 2v(x \vee y) - v(x) - v(y)$ for $x, y \in S$. In particular, if $v$ is a valuation on $L$, i.e., satisfies (5.4), then $D_v$ coincides with $d_v$ (which is defined in (5.5)) and, therefore, we have the following variation of Theorem 5.6: $(L, d_v)$ is $L_1$-embeddable if and only if $(L, d_v)$ is of negative type.

EXAMPLE 5.10 Let $\mathcal{A}$ be a family of subsets of a set $\Omega$ and suppose that $\mathcal{A}$ is stable under the symmetrical difference. Then, $(\mathcal{A}, \triangle)$ is a commutative group. Let $v : \mathcal{A} \mapsto \mathbb{R}_+$ be a mapping such that $v(\emptyset) = 0$ and set $d(A,B) = v(A \triangle B)(= \frac{D_v(A,B)}{2})$ for $A, B \in \mathcal{A}$. Then, by Theorem 5.8, $(\mathcal{A}, d)$ is $L_1$-embeddable if and only if $(\mathcal{A}, d)$ is of negative type.

# 6 Metric transforms preserving $L_1$-embeddability

Let $(X, d)$ be a distance space and let $F : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ be a function. We define the distance space $(X, F(d))$, called **metric transform** of $(X, d)$, by setting $F(d)_{ij} = F(d_{ij})$ for all $i, j \in X$.

A general question is to find nontrivial functions $F$ which preserve certain properties, e.g., metricity, $L_1$- or $L_2$-embeddability, of the original distance space. We present in this Section several results of this type about metric transforms. We refer to [WW75] for a good exposition of some of these results.

LEMMA 6.1 [Blu70] *Suppose that $F$ is nondecreasing, concave and $F(0) = 0$. If $(X, d)$ is a semimetric space, then $(X, F(d))$ is also a semimetric space.*

PROOF. Let $i, j, k \in X$. We show that $F(d_{ij}) \leq F(d_{ik}) + F(d_{jk})$ holds. But, $F(d_{ij}) \leq F(d_{ik} + d_{jk})$ holds since $F$ is nondecreasing and $F(d_{ik} + d_{jk}) \leq F(d_{ik}) + F(d_{jk})$ holds since $F$ is concave (to see it, verify that the function $t \longmapsto F(t + d_{jk}) - F(t) - F(d_{jk})$ is nondecreasing). ∎

A function satisfying the conditions of Lemma 6.1 is called a **scale**. Examples of scales include:
• $F(t) = \frac{t}{1+t}$, $t \geq 0$.
• $F(t) = 1 - \exp(-\lambda t)$ for $t \geq 0$, where $\lambda$ is a positive scalar; it is called the **Schoenberg scale**.
• $F(t) = t^\alpha$ for $t \geq 0$, where $0 < \alpha \leq 1$; it is called the **power scale**.

## 6.1 Metric transforms of $\ell_2$-spaces

We present two classical results about the functions $F$ for which the metric transform $F(\ell_2^m)$ of the $m$-dimensional Euclidian space $\ell_2^m = (\mathbb{R}^m, d_{\ell_2})$ is $L_2$-embeddable or embeddable into some Euclidian space $\ell_2^n$.

THEOREM 6.2 *(i)* [Sch38a] *Let $2 \leq m \leq n$ be integers. The functions $t \in \mathbb{R}_+ \longmapsto F(t)$ that are nonnegative, continuous, satisfy $F(0) = 0$, and for which $F(\ell_2^m)$ is isometrically embeddable in $\ell_2^n$ are of the form $F(t) = ct$ ($t \geq 0$), where $c \geq 0$.*
*(ii)* [vNS41] *Let $m \geq 1$ be an integer. The functions $t \in \mathbb{R}_+ \longmapsto F(t)$ that are nonnegative, continuous, satisfy $F(0) = 0$, and for which $F(\ell_2^m)$ is isometrically $L_2$-embeddable are of the form*

$$F(t) = \left( \int_0^\infty \frac{1 - \Omega_m(tu)}{u^2} \sigma'(u) du \right)^{1/2} \quad (t \geq 0)$$

*where* $u \in \mathbb{R}_+ \longmapsto \sigma(u)$ *is nondecreasing,* $\sigma(0) = 0$*, and* $\int_1^\infty \frac{1}{u^2}\sigma'(u)du < \infty$ *(with* $\sigma'$ *denoting the first derivative of* $\sigma$*). The function* $\Omega_m$ *is defined by*

$$\Omega_m(t) = 1 - \frac{t^2}{2\cdot m} + \frac{t^4}{2\cdot 4\cdot m\cdot(m+2)} - \frac{t^6}{2\cdot 4\cdot 6\cdot m\cdot(m+2)\cdot(m+4)} + \ldots .$$

*For* $m = 1$*, we have* $\Omega_1(t) = \cos(t)$ *and, thus, the functions* $F$ *are of the form*

$$F(t) = \left(\int_0^\infty \frac{\sin^2(tu)}{u^2}\sigma'(u)du\right)^{1/2} \quad (t \in \mathbb{R}_+).$$

PROOF. As an illustration, let us give the proof of the easy implication in $(ii)$ for the case $m = 1$. Let $F$ be defined as in the case $m = 1$ of Theorem 6.2 $(ii)$. By Theorem 4.16, in order to show that $F(\ell_2^1)$ is $L_2$-embeddable, it suffices to check that $F^2(\ell_2^1)$ is of negative type. By Lemma 4.8, this is equivalent to checking that its image under the covariance mapping is of positive type. Let $b_1, \ldots, b_k \in \mathbb{R}$ and $x_1, \ldots, x_k \in \mathbb{R}$; we show that the inequality $\sum_{1\le i,j\le k} b_ib_j(F^2(x_i) + F^2(x_j) - F^2(x_i - x_j)) \ge 0$ holds. For this, we use the identity
$\sin^2(x_iu) + \sin^2(x_ju) - \sin^2((x_i - x_j)u) = 2\sin^2(x_iu)\sin^2(x_ju) + \frac{\sin(2x_iu)\sin(2x_ju)}{2}$.
Indeed, we deduce from it that
$\sum_{1\le i,j\le k} b_ib_j(F^2(x_i) + F^2(x_j) - F^2(x_i - x_j))$
$= \int_0^\infty \left(2\left(\sum_{i=1}^k b_i\sin^2(x_iu)\right)^2 + \frac{1}{2}\left(\sum_{i=1}^k b_i\sin(2x_iu)\right)^2\right)\frac{d\sigma(u)}{u^2} \ge 0.$ ∎

EXAMPLE 6.3 [Sch37] Consider the function $F(t) = t^\alpha$ $(t \ge 0)$ where $0 < \alpha < 1$. Then, $F(\ell_2^1)$ is $L_2$-embeddable. Indeed, $F$ satisfies the conditions of Theorem 6.2 $(ii)$; this fact relies on the following integral formula

$$t^{2\alpha} = c_\alpha^{-1}\int_0^\infty u^{-1-2\alpha}\sin^2(tu)du\ (t \ge 0) \ \text{ where } c_\alpha = \int_0^\infty u^{-1-2\alpha}\sin^2(u)du.$$

## 6.2 The Schoenberg scale

We now consider the **Schoenberg scale**, $F(t) = 1 - \exp(-\lambda t)$ $(t \in \mathbb{R}_+)$, where $\lambda$ is a positive scalar. We show below that this scale preserves $L_1$-embeddability and the negative type property.

THEOREM 6.4 [Sch38b] *Let* $(X, d)$ *be a distance space. The following assertions are equivalent.*

*(i)* $(X, d)$ *is of negative type.*
*(ii) The symmetric function* $p : X \times X \longrightarrow \mathbb{R}$, *defined by* $p(x, y) = \exp(-\lambda d(x, y))$ *for* $x, y \in X$, *is of positive type for all* $\lambda > 0$.
*(iii)* $(X, 1 - \exp(-\lambda d))$ *is of negative type for all* $\lambda > 0$.

Proof. Note that the properties involved in Theorem 6.4 are all of finite type, i.e., they hold if and only if they hold for any finite subset of $X$. Hence, we can assume that $X$ is finite, say $X = \{1, \ldots, n\}$.
$(i) \Longrightarrow (ii)$ Since $(X, d)$ is of negative type then, by Theorem 4.16, $(X, \sqrt{d})$ is $\ell_2$-embeddable, i.e., there exist $x^{(1)}, \ldots, x^{(n)} \in \mathbb{R}^m$ $(m \geq 1)$ such that $d_{jk} = (\| x^{(j)} - x^{(k)} \|_2)^2$ for all $j, k \in X$. Let $b_1, \ldots, b_n \in \mathbb{R}$. We show that $\sum_{1 \leq j,k \leq n} b_j b_k \exp\left(-\lambda(\| x^{(j)} - x^{(k)} \|_2)^2\right) \geq 0$. For this, we use the following classical identity

$$\exp(-x^2) = 2^{-1}\pi^{-\frac{1}{2}} \int_{-\infty}^{\infty} \exp(ixu)\exp(-\frac{u^2}{4})du.$$

(Here, $i$ denotes the complex square root of unity.) Indeed, we have that
$\sum_{j,k \in X} b_j b_k \exp\left(-\lambda(\| x^{(j)} - x^{(k)} \|_2)^2\right)$
$= \sum_{j,k \in X} b_j b_k \prod_{1 \leq h \leq m} \exp(-\lambda(x_h^{(j)} - x_h^{(k)})^2)$
$= \sum_{i,j \in X} b_j b_k 2^{-m}\pi^{-\frac{m}{2}} \prod_{1 \leq h \leq m} \int_{-\infty}^{\infty} \exp(i\sqrt{\lambda}(x_h^{(j)} - x_h^{(k)})u_h)\exp(-\frac{u_h^2}{4})du_h$
$= \sum_{j,k \in X} b_j b_k 2^{-m}\pi^{-\frac{m}{2}} \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} \exp(i\sqrt{\lambda}(x^{(j)} - x^{(k)})^T u)\exp(-\frac{1}{4}\sum_{1 \leq h \leq m} u_h^2)du_1 \ldots du_m$
$= 2^{-m}\pi^{-\frac{m}{2}} \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} \left|\sum_{j \in X} b_j \exp(i\sqrt{\lambda}{x^{(j)}}^T u)\right|^2 \exp(-\frac{1}{4}(\sum_{1 \leq h \leq m} u_h^2))du_1 \ldots du_m \geq 0.$

$(ii) \Longrightarrow (iii)$ Set $d'_{ij} = \exp(-\lambda d_{ii}) + \exp(-\lambda d_{jj}) - 2\exp(-\lambda d_{ij})) = 2(1 - \exp(-\lambda d_{ij}))$ for $i, j \in X$, i.e., $d'$ arises from $p = \exp(-\lambda d)$ by applying the inverse of the covariance mapping (defined in (3.8)). Applying Lemma 4.8, we obtain that $(X, d')$ is of negative type, i.e., $(X, 1 - \exp(-\lambda d))$ is of negative type.

$(iii) \Longrightarrow (i)$ Let $b_1, \ldots, b_n \in \mathbb{R}$ with $\sum_{1 \leq i \leq n} b_i = 0$. We show that the inequality $\sum_{1 \leq i < j \leq n} b_i b_j d_{ij} \leq 0$ holds. By expanding in series the exponential function, we obtain that
$\sum_{1 \leq i < j \leq n} b_i b_j (1 - \exp(-\lambda d_{ij}))$
$= \lambda\left(\sum_{1 \leq i < j \leq n} b_i b_j d_{ij} - \frac{\lambda}{2}\sum_{1 \leq i < j \leq n} b_i b_j d_{ij}^2 + \frac{\lambda^2}{3!}\sum_{1 \leq i < j \leq n} b_i b_j d_{ij}^3 - \ldots\right) \leq 0$
for all $\lambda > 0$, since $1 - \exp(-\lambda d)$ is of negative type. By taking the limit when $\lambda \rightarrow 0$, we obtain that $\sum_{1 \leq i < j \leq n} b_i b_j d_{ij} \leq 0$. ∎

Remark that Theorem 6.4 remains valid if we assume only that $(ii)$ and $(iii)$ hold for a set of positive $\lambda$'s admitting 0 as accumulation point. The same remark also applies to the next Theorem 6.5 $(i)$.

Theorem 6.5 [Ass79, Ass80b] *Let $(X,d)$ be a distance space. Then,*
*$(i)$ $(X,d)$ is $L_1$-embeddable if and only if $(X, 1-\exp(-\lambda d))$ is $L_1$-embeddable for all $\lambda > 0$.*
*$(ii)$ Let $\nu$ be a positive measure on $\mathbb{R}$ and set $f(t) = \int_0^\infty (1-\exp(-\lambda t))\nu(d\lambda)$ for $t \geq 0$. If $(X,d)$ is $L_1$-embeddable, then $(X, f(d))$ is $L_1$-embeddable.*

Proof. $(i)$ Assume that $(X,d)$ is $L_1$-embeddable. Hence, $(X,d)$ is an isometric subspace of a measure semimetric space $(\mathcal{A}_\mu, d_\mu)$ for some measure space $(\Omega, \mathcal{A}, \mu)$. Set $v(A) = 1-\exp(-\lambda\mu(A))$ for $A \in \mathcal{A}_\mu$. Then, the distance space $(\mathcal{A}_\mu, 1-\exp(\lambda d_\mu))$ coincides with the space $(\mathcal{A}_\mu, d_v)$, which is defined as in Example 5.10. Therefore, in order to show that $(\mathcal{A}_\mu, 1-\exp(-\lambda d_\mu))$ is $L_1$-embeddable, it suffices to show that $(\mathcal{A}_\mu, 1-\exp(-\lambda d_\mu))$ is of negative type. But, we know from Theorem 6.4 that $(\mathcal{A}_\mu, 1-\exp(-\lambda d_\mu))$ is of negative type, since $(\mathcal{A}_\mu, d_\mu)$ is $L_1$-embeddable and, thus, of negative type.
The proof of the converse implication is analogue to that of the implication $(iii) \Longrightarrow (i)$ of Theorem 6.4 (replacing the negative type inequality by an arbitrary inequality valid for the cut cone $\mathrm{CUT}(Y)$ where $Y$ is a finite subset of $X$).
$(ii)$ Again we may suppose that $(X,d)$ is an isometric subspace of $(\mathcal{A}_\mu, d_\mu)$ for some measure space $(\Omega, \mathcal{A}, \mu)$. Set $v(A) = f(\mu(A))$ for $A \in \mathcal{A}_\mu$. Then, the space $(\mathcal{A}_\mu, f(d_\mu))$ coincides with the space $(\mathcal{A}_\mu, d_v)$, which is constructed as in Example 5.10. We check that $(\mathcal{A}_\mu, d_v)$ is of negative type, which will imply that $(\mathcal{A}_\mu, f(d_\mu))$ is $L_1$-embeddable. Indeed, let $A_i \in \mathcal{A}_\mu$, $b_i \in \mathbb{R}$ for $i \in X$ such that $\sum_{i\in X} b_i = 0$. Then,
$\sum_{i,j\in X} b_i b_j d_v(A_i \triangle A_j) = \int_0^\infty \sum_{i,j\in X} b_i b_j (1-\exp(-\lambda\mu(A_i\triangle A_j))\nu(d\lambda) \geq 0$
because $1-\exp(-\lambda d)$ is of negative type and $\nu$ is a positive measure. ∎

Example 6.6 [Ass79, Ass80b] If $(X,d)$ is $L_1$-embeddable, then $(X, d^\alpha)$ is $L_1$-embeddable for all $0 \leq \alpha \leq 1$.
This is a consequence of Theorem 6.5 $(ii)$ and of the following integral formula

$$t^\alpha = e_\alpha^{-1} \int_0^\infty (1-\exp(-\lambda^2 t))\lambda^{-1-2\alpha} d\lambda \ (t \geq 0) \ \text{ where } e_\alpha = \int_0^\infty (1-\exp(-\lambda^2))\lambda^{-1-2\alpha} d\lambda.$$

### 6.3 The biotope transform

We mention another transformation which preserves $L_1$-embeddability. Let $d$ be a distance on a set $X$ and let $s$ be a point of $X$. We define a new distance $d^{(s)}$ on $X$ by setting

$$d^{(s)}(i,j) = \frac{d(i,j)}{d(i,s)+d(j,s)+d(i,j)}$$

for all $i, j \in X$. In particular, if $(\Omega, \mathcal{A}, \mu)$ is a measure space and $(X, d)$ is the measure semimetric space $(\mathcal{A}_\mu, d_\mu)$, then its transform $d_\mu^{(\emptyset)}$ takes the form

$$d_\mu^{(\emptyset)}(A, B) = \frac{\mu(A \triangle B)}{\mu(A) + \mu(B) + \mu(A \triangle B)} = \frac{\mu(A \triangle B)}{2\mu(A \cup B)}$$

for $A, B \in \mathcal{A}_\mu$. The distance $(A, B) \in \mathcal{A}_\mu \times \mathcal{A}_\mu \mapsto \frac{\mu(A \triangle B)}{\mu(A \cup B)}$ is called the **Steinhaus distance**. The distance $(A, B) \mapsto \frac{|A \triangle B|}{|A \cup B|}$, which is obtained in the special case when $\mu$ is the cardinality measure, is also called the **biotope distance**. This terminology comes from the fact that this distance is used in some biological problems for the study of biotopes (see [MS58]). As a consequence of the next Proposition 6.7, the Steinhaus and biotope distances are $L_1$-embeddable.

PROPOSITION 6.7 *(i)* [MS58] *If $d$ is a semimetric on $X$, then $d^{(s)}$ is a semimetric on $X$.*
*(ii)* [Ass80b] *If $(X, d)$ is $L_1$-embeddable, then $(X, d^{(s)})$ is also $L_1$-embeddable.*

PROOF. $(i)$ follows from $(ii)$ and the fact that a distance space on at most 4 points is $L_1$-embeddable if and only if it is a semimetric space (see Remark 4.22 $(i)$).
$(ii)$ We can suppose that $(X, d)$ is an isometric subspace of some measure semimetric space $(\mathcal{A}_\mu, d_\mu)$, i.e., $d(i, j) = \mu(A_i \triangle A_j)$ where $A_i \in \mathcal{A}_\mu$ for all $i, j \in X$, and we can suppose without loss of generality that $A_s = \emptyset$. Hence, as was already observed, $d^{(s)}(i, j) = \frac{\mu(A_i \triangle A_j)}{2\mu(A_i \cup A_j)}$ for all $i, j \in X$. By Lemma 3.11, showing that $(X, d^{(s)})$ is $L_1$-embeddable amounts to showing that $p = \xi_s(d^{(s)})$ is a $\{0,1\}$-covariance. From (3.7), $p$ is defined by $p(i, j) = \frac{1}{2}(d^{(s)}(i, s) + d^{(s)}(j, s) - d^{(s)}(i, j))$ for $i, j \in X \setminus \{s\}$. Hence, $p(i, j) = \frac{1}{4} + \frac{1}{4} \frac{\mu(A_i \cap A_j)}{\mu(A_i \cup A_j)}$ for $i, j \in X \setminus \{s\}$. Therefore, it suffices to show that the symmetric function $(i, j) \in (X \setminus \{s\})^2 \longmapsto \frac{\mu(A_i \cap A_j)}{\mu(A_i \cup A_j)}$ is a $\{0,1\}$-covariance. For this, we use the identity $\frac{\mu(A \cap B)}{\mu(A \cup B)} = \frac{\mu(A \cap B)}{\mu(\Omega)} \left( \sum_{i \geq 0} \left( \frac{\mu(\bar{A} \cap \bar{B})}{\mu(\Omega)} \right)^i \right)$ (which follows from the identity $\sum_{i \geq 0} (1 - u)^i = \frac{1}{u}$ for all $0 < u \leq 1$) and the fact that $\{0,1\}$-covariances are preserved under taking sum, product and limit ([Ass80b], see Corollary 4.25 $(ii)$). ∎

## 6.4 The power scale

We finally consider the **power scale**, $F(t) = t^\alpha$ for $t \geq 0$, where $0 < \alpha < 1$. The question is to determine the largest exponent $\alpha$ for which the power scale preserves some metric properties as hypermetricity, $\ell_1$-, or $\ell_2$-embeddability.

We consider the parameters $g(n)$, $h(n)$, $t(n)$, $c_1(n)$, $c_2(n)$ which are defined as follows: $g(n)$ (resp. $h(n)$, $t(n)$, $c_1(n)$, $c_2(n)$) is the maximum exponent $\alpha$, $0 \leq \alpha \leq 1$, such that the distance space $(V_n, d^\alpha)$ is $i$-gonal for all $i \leq 2\lfloor \frac{n}{2} \rfloor + 1$ (resp. hypermetric, of negative type, $\ell_1$-embeddable, $\ell_2$-embeddable).

From Theorem 4.20 and Theorem 4.16, we have the relations
$c_2(n) \leq c_1(n) \leq h(n) \leq \min(t(n), g(n))$ and $t(n) = 2c_2(n)$, respectively.

Set $\gamma(s) = \log_2(1 + \frac{1}{s})$ for $s > 0$. The following results are given in [DM90].
• $g(2n) = \gamma(n-1)$ for $n \geq 2$ and $g(2n+1) = \gamma(n)$ for $n \geq 1$.
Note that, if $d(K_{n,n})$ denotes the path metric of the complete bipartite graph $K_{n,n}$, then $(d(K_{n,n}))^\alpha$ violates the $2n$-gonal inequality if $\alpha > \gamma(n-1)$. This shows that $g(2n) \leq \gamma(n-1)$. Similarly, $(d(K_{n,n+1}))^\alpha$ violates the $2n+1$-gonal inequality if $\alpha > \gamma(n)$, showing that $g(2n+1) \leq \gamma(n)$.

• $h(n) \geq \gamma(n-1)$ for all $n \geq 2$. This implies that $c_2(n) \geq \frac{\gamma(n-1)}{2}$ for all $n \geq 2$.

• Let $d(K_{m,n})$ denote the path metric of the complete bipartite graph $K_{m,n}$. Then, $d(K_{m,n})^\alpha$ is $\ell_2$-embeddable if and only if $c \leq \frac{1}{2}\gamma(\frac{2mn}{m+n} - 1)$. This implies that

$$\begin{cases} c_2(2n) \leq \frac{1}{2}\gamma(n-1) & \text{for all } n \geq 2, \\ c_2(2n+1) \leq \frac{1}{2}\gamma(\frac{2n(n+1)}{2n+1} - 1) & \text{for all } n \geq 1. \end{cases}$$

Deza and Maehara [DM90] conjecture that the above inequalities for $c_2(2n)$ and $c_2(2n+1)$ hold at equality. It is known that $c_2(3) = 1 = \frac{1}{2}\gamma(\frac{1}{3})$ (easy), $c_2(4) = \frac{1}{2} = \frac{1}{2}\gamma(1)$ ([Blu70]) and $c_2(6) = \frac{1}{2}\gamma(2)$ ([DM90]), i.e., the conjecture holds for $n = 3, 4, 6$.

We summarize the known information for $n = 3, 4, 5, 6$:
• $g(3) = h(3) = c_1(3) = c_2(3) = 1$, and $t(3) = 2$,
• $t(4) = g(4) = h(4) = c_1(4) = 1$ and $c_2(4) = \frac{1}{2}$,
• $g(5) = h(5) = c_1(5) = \gamma(2) = \log_2(\frac{3}{2})$ and $\frac{1}{2}\gamma(2) \leq c_2(5) \leq \frac{1}{2}\gamma(\frac{7}{5}) = \frac{1}{2}\log_2(\frac{12}{7})$.
• $t(6) = g(6) = h(6) = c_1(6) = \gamma(2) = \log_2(\frac{3}{2})$ and $c_2(6) = \frac{1}{2}\gamma(2) = \frac{1}{2}\log_2(\frac{3}{2})$.

# 7 Additional questions on $\ell_1$-embeddings

In this Section, we address the following two questions.
- Evaluate the minimum $\ell_p$-dimension $m_p(n)$ of an $\ell_p$-embeddable distance space on $n$ points.
- Determine the smallest integer $c(m)$ such that, for every distance space $(X, d)$, $(X, d)$ is $\ell_1^m$-embeddable if and only if every subspace of $(X, d)$ on $c(m)$ points embeds in $\ell_1^m$.

## 7.1 On the minimum $\ell_p$-dimension

We consider here the problem of evaluating the minimum $\ell_p$-dimension $m_p(n)$ of an $\ell_p$-embeddable space on $n$ points. We recall the definition of $m_p(n)$ from relation (1.2), i.e., $m_p(n)$ is the smallest integer $m$ such that any $\ell_p$-embeddable space on $n$ points can embedded in $\ell_p^m$. The results we present come essentially from [Bal90] and can be stated as follows.

As was already observed in relations (1.3) and (4.18), $m_\infty(n) \leq n-1$ and $m_2(n) = n-1$ but, for general $p$, it is not immediate that $m_p(n)$ is finite. Wolfe [Wol67] showed that $m_\infty(n) \leq n-2$. Witsenhausen proved that $m_1(n) \leq \binom{n}{2}$ and Ball extended the result for any $p \geq 1$. In other words, every $\ell_p$-embeddable distance on $n$ points can embedded in $\ell_p^m$, where $m = \binom{n}{2}$. Moreover, for $1 \leq p < 2$, this result is essentially best possible, since $m_p(n) \geq \binom{n-1}{2}$ for $1 < p < 2$, $n \geq 3$, and $m_1(n) \geq \binom{n-2}{2}$ for $n \geq 4$. One can also show that $m_1(4) = m_\infty(4) = 2$, $m_1(5) = 3$ and $m_1(6) = 6$. It is conjectured in [Bal90] that $m_1(n) = \binom{n-2}{2}$ for all $n \geq 5$.

Ball's proof for the upper bound $m_p(n) \leq \binom{n}{2}$ is based on an application of Caratheodory's Theorem to the cut cone (if $p = 1$) or the cone $\mathrm{NOR}_n(p)$ (for $p \geq 1$; see the definition below). We first present the result in the case $p = 1$.

PROPOSITION 7.1 $m_1(n) \leq \binom{n}{2}$.

PROOF. Let $d$ be a distance on $n$ points that is $\ell_1$-embeddable, i.e., $d$ belongs to the cut cone $\mathrm{CUT}_n$. We show that $d$ can be embedded in $\ell_1^m$, where $m = \binom{n}{2}$. For this, it suffices to show that $d$ can be decomposed as a nonnegative linear combination of at most $\binom{n}{2}$ distinct cut semimetrics (recall Remark 2.6). Let $H$ denote the hyperplane in $\mathbb{R}^{E_n}$ defined by the equation $\sum_{1\leq i<j\leq n} x_{ij} = 1$. Then, the section $\mathrm{CUT}_n \cap H$ of the cut cone $\mathrm{CUT}_n$ by $H$ is a polytope of dimension $\binom{n}{2}-1$ whose vertices are the vectors $\frac{\delta(S)}{|\delta(S)|}$ for all subsets $S$ of $V_n$. Set $a = \sum_{1\leq i<j\leq n} d_{ij}$. Then, $\frac{d}{a} \in \mathrm{CUT}_n \cap H$ and, thus, by Caratheodory's Theorem, $\frac{d}{a}$ can be written as the convex hull of at most $\binom{n}{2}$ members of $\{\frac{\delta(S)}{|\delta(S)|} : S \subseteq V_n\}$. This shows that $d$ can written as the conic hull of at most $\binom{n}{2}$ cut semimetrics. ∎

The result from Proposition 7.1 can extended for any $p \geq 1$, using the following cone $\mathrm{NOR}_n(p)$ instead of the cut cone $\mathrm{CUT}_n$.

Given an integer $p \geq 1$, let $\mathrm{NOR}_n(p)$ denote the set of all distances $d$ on $V_n$ for which $d^{\frac{1}{p}}$ is $\ell_p$-embeddable, i.e., there exist $n$ vectors $v_1, \ldots, v_n \in \mathbb{R}^m$ ($m \geq 1$) such that $d_{ij} = \| v_i - v_j \|_p^p$ for all $1 \leq i < j \leq n$.

Note that, if $p = 1$, then $\mathrm{NOR}_n(1)$ coincides with the cut cone $\mathrm{CUT}_n$ (by Proposition 2.5). An element $d \in \mathrm{NOR}_n(p)$ is said to be **linear** if $d^{\frac{1}{p}}$ is $\ell_p^1$-embeddable, i.e., if there exist $x_1, \ldots, x_n \in \mathbb{R}$ such that $d_{ij} = |x_i - x_j|^p$ for all $1 \leq i < j \leq n$. For example, each cut semimetric belongs to $\mathrm{NOR}_n(p)$ and is linear, i.e.,

$$\mathrm{CUT}_n \subseteq \mathrm{NOR}_n(p) \text{ for all } 1 \leq i < j \leq n.$$

We collect in the next result some properties of the set $\mathrm{NOR}_n(p)$.

LEMMA 7.2 *(i) $NOR_n(p)$ is a cone.*
*(ii) Let $d \in NOR_n(p)$. Then, $d^{\frac{1}{p}}$ is $\ell_p^m$-embeddable if and only if $d$ is the sum of $m$ linear members of $NOR_n(p)$. In particular, if $d$ lies on an extreme ray of $NOR_n(p)$, then $d$ is linear.*

PROOF. $(i)$ Let $d, d' \in \mathrm{NOR}_n(p)$. We show that $d + d' \in \mathrm{NOR}_n(p)$. By assumption, there exist some vectors $u_1, \ldots, u_n, v_1, \ldots, v_n \in \mathbb{R}^m$ $(m \geq 1)$ such that $d_{ij} = \| u_i - u_j \|_p^p$ and $d'_{ij} = \| v_i - v_j \|_p^p$ for all $1 \leq i < j \leq n$. Set $w_i = (u_i, v_i) \in \mathbb{R}^{2m}$ for all $1 \leq i \leq n$. Then, $d_{ij} + d'_{ij} = \| w_i - w_j \|_p^p$ for all $1 \leq i < j \leq n$. This shows that $d + d' \in \mathrm{NOR}_n(p)$. Hence, $\mathrm{NOR}_n(p)$ is a cone.
$(ii)$ Let $d \in \mathrm{NOR}_n(p)$. If $d^{\frac{1}{p}}$ is $\ell_p^m$-embeddable, then there exist $u_1, \ldots, u_n \in \mathbb{R}^m$ such that $d_{ij} = \| u_i - u_j \|_p^p = \sum_{1 \leq h \leq m} |(u_i)_h - (u_j)_h|^p$ for $1 \leq i < j \leq n$. Hence, $d = d^1 + \ldots + d^m$, where $d^h$ denotes the distance on $V_n$ defined by $(d^h)_{ij} = |(u_i)_h - (u_j)_h|$ for $1 \leq i < j \leq n$. This shows that $d$ is the sum of $m$ linear members of $\mathrm{NOR}_n(p)$, since $d^1, \ldots, d^m$ belong to $\mathrm{NOR}_n(p)$ and are linear, by construction. The converse implication holds similarly. ∎

PROPOSITION 7.3 *$m_p(n) \leq \binom{n}{2}$.*

PROOF. We sketch the proof. Let $H$ denote again the hyperplane in $\mathbb{R}^{E_n}$ defined by the equation $\sum_{1 \leq i < j \leq n} x_{ij} = 1$. Set $L = \{d \in \mathrm{NOR}_n(p) : d \in H \text{ and } d \text{ is linear}\}$. One can show that $L$ is a compact set and that $\mathrm{NOR}_n(p) \cap H$ is a $\binom{n}{2} - 1$-dimensional convex set which coincides with the convex hull of $L$. As in the proof of Proposition 7.1, Caratheodory's Theorem implies that every member of $\mathrm{NOR}_n(p)$ can be written as the sum of $\binom{n}{2}$ linear members of $\mathrm{NOR}_n(p)$. Now, suppose $d$ is an $\ell_p$-embeddable distance on $n$ points. Then, $d^p \in \mathrm{NOR}_n(p)$ and, thus, $d^p$ is the sum of $\binom{n}{2}$ members of $\mathrm{NOR}_n(p)$, i.e., $d$ embeds in $\ell_p^m$, where $m = \binom{n}{2}$. ∎

PROPOSITION 7.4 *(i) $m_1(n) \geq \binom{n-2}{2}$ for $n \geq 4$.*
*(ii) $m_p(n) \geq \binom{n-1}{2}$ for $1 < p < 2$ and $n \geq 3$.*

PROOF. $(i)$ Set $m = \binom{n-2}{2}$. We exhibit a semimetric $d$ on $V_n$ which embeds in $\ell_1^m$ but not in $\ell_1^k$ if $k < m$. Set $d = \sum_{2\leq r<s\leq n-1} \delta(\{r,s,n\})$, i.e.,

$$\begin{cases} d_{1n} &= \binom{n-2}{2} \\ d_{1i} &= n-3 \qquad \text{for } 2 \leq i \leq n-1 \\ d_{ij} &= 2(n-4) \quad \text{for } 2 \leq i < j \leq n-1 \\ d_{in} &= \binom{n-3}{2} \qquad \text{for } 2 \leq i \leq n-1. \end{cases}$$

By construction, $d$ embeds isometrically in $\ell_1^m$. We show that $d$ cannot be embedded in $\ell_1^k$ if $k < m$. For this, we consider the following inequality of negative type $Neg_n(-(n-4), 1, \ldots, 1, -2)$, i.e., the inequality

$$2(n-4)x_{1n} - (n-4) \sum_{2\leq i\leq n-1} x_{1i} - 2 \sum_{2\leq i\leq n-1} x_{in} + \sum_{2\leq i<j\leq n-1} x_{ij} \leq 0. \tag{7.5}$$

Let $F$ denote the face of the cone $\mathrm{NOR}_n(1)$ ($=\mathrm{CUT}_n$) defined by the inequality (7.5). One can easily check that $d$ satisfies the inequality (7.5) at equality, i.e., $d$ lies on the face $F$, and that the only cut semimetrics lying on $F$ are the cut semimetrics $\delta(\{r,s,n\})$ for $2 \leq r < s \leq n-1$. Moreover, these cut semimetrics are linearly independent, i.e., $F$ is a simplicial face of $\mathrm{NOR}_n(1)$. One can also show (Lemma 4, [Bal90]) that
$(*)$ the only linear members of $\mathrm{NOR}_n(1)$ lying on $F$ are of the form $\lambda\delta(\{r,s,n\})$ for $2 \leq r < s \leq n-1$ and $\lambda > 0$.

Let us suppose that $d$ embeds in $\ell_1^k$. Then, $d$ is the sum of $k$ linear members of $\mathrm{NOR}_n(1)$. From $(*)$ above, we deduce that $d$ can be written as a nonnegative linear combination of $k$ of the cuts $\delta(\{r,s,n\})$ for $2 \leq r < s \leq n-1$. Since $d$ lies on a simplicial face, the latter decomposition of $d$ must coincide with the initial decomposition $d = \sum_{2\leq r<s\leq n-1} \delta(\{r,s,n\})$. Therefore, $k = \binom{n-2}{2}$.

$(ii)$ We only sketch the proof, which is along the same lines as for $(i)$. Set $m = \binom{n-1}{2}$. Consider the vectors $v_1, \ldots, v_n \in \mathbb{R}^m$ defined by

$$(v_i)_{rs} = \begin{cases} 1 & \text{if } r = i \\ -1 & \text{if } s = i \\ 0 & \text{otherwise} \end{cases}$$

for $1 \leq r < s \leq n$. Define the distance $d$ on $V_n$ by setting $d_{ij} = \| v_i - v_j \|_p$ for $1 \leq i < j \leq n$. So $d$ embeds in $\ell_p^m$ by construction. One can show that $d$ does not embed in $\ell_p^k$ if $k < m$ by using, as in case $(i)$, a special inequality which is valid for the cone $\mathrm{NOR}_n(p)$ and is satisfied at equality by $d^p$. Namely, one uses the inequality

$$\sum_{1\leq i<j\leq n} \| u_i - u_j \|_p^p - (n + 2^{p-1} - 2) \| u_i \|_p^p \leq 0$$

which holds for any set of $n$ vectors $u_1, \ldots, u_n \in \mathbb{R}^h$ $(h \geq 1)$ if $1 \leq p \leq 2$ ([Bal87]). ∎

Remark 7.6 The proof of the lower bound $\binom{n-2}{2}$ from Proposition 7.4 $(i)$ uses essentially the fact that the inequality (7.5) defines a simplicial face $F$ of $\mathrm{NOR}_n(1)$ that contains only $\binom{n-2}{2}$ linear members (up to multiple). Observe that the hypermetric inequality $\mathrm{Hyp}_n(-(n-4), 1, \ldots, 1, -1)$, i.e.,

$$(n-4)x_{1n} - (n-4) \sum_{2 \leq i \leq n-1} x_{1i} - \sum_{2 \leq i \leq n-1} x_{in} + \sum_{2 \leq i < j \leq n-1} x_{ij} \leq 0$$

defines a simplicial facet $G$ of $\mathrm{NOR}_n(1)$ that contains the face $F$. Hence, $G$ contains $\binom{n}{2} - 1$ cut semimetrics but $G$ may contain additional linear points. For this reason, the lower bound from Proposition 7.4 $(i)$ cannot be improved using the facet $G$ instead of the face $F$.

Linial, London and Rabinovitch [LLR93] define the **metric dimension** of a graph $G$ as the smallest integer $m$ for which there exists a norm $\| \,.\, \|$ on $\mathbb{R}^m$ such that the graphic space $(V, d_G)$ of the graph $G$ can be isometrically embedded into the space $(\mathbb{R}^m, d_{\|.\|})$. The definition extends clearly to an arbitrary semimetric space. Hence, rather than looking only at embeddings in a fixed Banach $\ell_p$-space, [LLR93] considers embeddings in an arbitrary normed space.
Actually, the notion of metric dimension is linked with $\ell_\infty$-embeddings in the following way. Let $(V_n, d)$ be a semimetric space. Then, its metric dimension is equal to the minimum rank of a sytem of vectors $v_1, \ldots, v_n \in \mathbb{R}^k$ $(k \geq 1)$ providing an $\ell_\infty$-embedding of $(V_n, d)$, i.e., such that $d_{ij} = \| v_i - v_j \|_\infty$ for all $1 \leq i < j \leq n$.
The metric dimension of several graphs is computed in [LLR93]. In particular, $dim(K_n) = \lceil \log_2(n) \rceil$, $dim(T) = O(\log_2(n))$ for a treee on $n$ nodes (both being realized by an $\ell_\infty$-embedding), $dim(C_{2n}) = n$ for a cycle on $2n$ nodes (realized by an $\ell_1$-embedding), $dim(K_{n \times 2}) \geq n-1$ for the cocktail party graph (i.e., $K_{2n}$ minus a perfect matching). It is also shown in [LLR93] that, if $G$ is a graph on $n$ nodes with metric dimension $d$, then each vertex has degree $\leq 3^d - 1$, $G$ has diameter $\geq \frac{1}{2}(n^{\frac{1}{d}} - 1)$, and there exists a subset $S$ of $O(dn^{1-\frac{1}{d}})$ nodes whose deletion disconnects $G$ and so that each connected component of $G \backslash S$ has no more than $(1 - \frac{1}{d} + o(1))n$ nodes.

## 7.2 Compactness results for $\ell_1$-embeddability in the plane

Let $m \geq 1$ be an integer and let $p \geq 1$. Define $c_p(m)$ as the smallest integer such that an arbitrary distance space $(X, d)$ is $\ell_p^m$-embeddable if and only if every subspace of $(X, d)$ on $c_p(m)$ points is $\ell_p^m$-embeddable. By convention, we set $c_p(m) = \infty$ if $c_p(m)$ does not exist.

The study of the parameter $c_p(m)$ is motivated by the following result of Menger for the case $p = 2$. Menger [Men28] showed that a distance space $(X, d)$ embeds isometrically in the Euclidian space $(\mathbb{R}^m, d_{\ell_2})$ if and only if each subspace of $(X, d)$ on $m+3$ points embeds isometrically in $(\mathbb{R}^m, d_{\ell_2})$. In other words, $c_2(m) = m+3$ for each $m \geq 1$.

Thus arises naturally the question of looking for analogues of Menger's Theorem for the case of arbitrary $\ell_p$-metrics and, in particular, in the case $p = 1$.

Since the spaces $(\mathbb{R}, d_{\ell_1})$ and $(\mathbb{R}, d_{\ell_2})$ are identical, we deduce from Menger's Theorem that $c_1(1) = 4$. Malitz and Malitz [MM92] show that $6 \leq c_1(2) \leq 11$ and $c_1(m) \geq 2m+1$ for all $m \geq 1$. The latter result is improved by J. Schmerl who proves that $c_1(m) \geq 2m+2$ for all $m \geq 1$.

It is conjectured in [MM92] that $c_1(m)$ exists for all $m$ and that $c_1(m) = 2m+2$ for all $m$.

Note that if we know that $c_1(m)$ exists, then this implies the existence of a polynomial time algorithm for checking embeddability of a finite distance space in the space $(\mathbb{R}^m, d_{\ell_1})$, for any given $m$. We recall that, on the other hand, checking $\ell_1$-embeddability of a finite distance space (i.e., embeddability into some $(\mathbb{R}^m, d_{\ell_1})$ for unrestricted $m$) is NP-complete ([Kar85]).

# 8 Examples of the use of the $L_1$-metric

## 8.1 The $L_1$-metric in probability theory

Let $(\Omega, \mathcal{A}, \mu)$ be a probability space and let $X : \Omega \longrightarrow \mathbb{R}$ be a random variable with finite expected value $E(X) = \int_\Omega |X(\omega)|\mu(d\omega) < \infty$, i.e., $X \in L_1(\Omega, \mathcal{A}, \mu)$. Let $F_X$ denote the distribution function of $X$, i.e., $F_X(x) = \mu(\{\omega \in \Omega : X(\omega) = x\})$ for $x \in \mathbb{R}$; when it exists, its derivative $F'_X$ is called the density of $X$. A great variety of metrics on random variables are studied in the monography [Rac91]; among them, the following are based on the $L_1$-metric.

• The usual $L_1$-metric between the random variables:
$L_1(X,Y) = E(|X-Y|) = \int_\Omega |X(\omega) - Y(\omega)|\mu(d\omega)$.

• The Monge-Kantorovich-Wasserstein metric (i.e., the $L_1$-metric between the distribution functions): $k(X,Y) = \int_\mathbb{R} |F_X(x) - F_Y(x)|dx$.

• The total valuation metric (i.e., the $L_1$-metric between the densities when they exist): $\sigma(X,Y) = \frac{1}{2}\int_\mathbb{R} |F'_X(x) - F'_Y(x)|dx$.

• The engineer metric (i.e., the $L_1$-metric between the expected values): $EN(X,Y) = |E(X) - E(Y)|$.

• The indicator metric: $i(X,Y) = E(1_{X \neq Y}) = \mu(\{\omega \in \Omega : X(\omega) \neq Y(\omega)\})$.

In fact, the $L_p$-analogues ($1 \leq p \leq \infty$) of the above metrics, especially of the first two, are also used in probability theory.

Several results are known, establishing links among the above metrics. One of the main such results is the Monge-Kantorovich mass-transportation theorem which shows that the second metric $k(X,Y)$ can be viewed as a minimum of the first metric $L_1(X,Y)$ over all joint distributions of $X$ and $Y$ with fixed marginal. A relationship between the $L_1(X,Y)$ and the engineer metric $EN(X,Y)$ is given by [Rac91] as solution of a moment problem. Similarly, a connection between the total valuation metric $\sigma(X,Y)$ and the indicator metric $i(X,Y)$ is given in Dobrushin's theorem on the existence and uniqueness of Gibbs fields in statistical physics. See [Rac91] for a detailed account of the above topics.

We mention another example of use of the $L_1$-metric in probability theory, namely for Gaussian random fields. We refer to [Nod87, Nod89] for a detailed account. Let $B = (B(x); x \in M)$ be a centered Gaussian system with parameter space $M$, $0 \in M$. The variance of the increment is denoted by

$$d(x,y) := E((B(x) - B(y))^2) \text{ for } x,y \in M.$$

When $(M,d)$ is a metric space which is $L_1$-embeddable, the Gaussian system is called a Lévy's Brownian motion with parameter space $(M,d)$. The case $M = \mathbb{R}^n$ and $d(x,y) = \parallel x - y \parallel_2$ gives the usual Brownian motion with $n$-dimensional parameter. By Lemma 2.10, $(M,d)$ is $L_1$-embeddable if and only if there exist a non negative measure space $(H,\nu)$ and a map $x \mapsto A_x \subseteq H$ with $\nu(A_x) < \infty$ for $x \in M$, such that $d(x,y) = \nu(A_x \triangle A_y)$ for $x,y \in M$. Hence, a Gaussian system admits a representation called of Chentsov type

$$B(x) = \int_{A_x} W(dh) \text{ for } x \in M$$

in terms of a Gaussian random measure based on the measure space $(H,\nu)$ with $d(x,y) = \nu(A_x \triangle A_y)$ if and only if $d$ is $L_1$-embeddable.

This Chentsov type representation can be compared with the Crofton formula for projective metrics from Theorem 5.3. Actually both come naturally together in [Amb82] (see parts A.8-A.9 of Appendix A there).

## 8.2 The $\ell_1$-metric in statistical data analysis

A **data structure** is a pair $(I,d)$, where $I$ is a finite set, called bf population, and $d : I \times I \longrightarrow \mathbb{R}_+$ is a symmetric map with $d_{ii} = 0$ for $i \in I$, called **dissimilarity index**. The typical problem in statistical data analysis is to choose a "good representation" of a data structure; usually, "good" means a representation allowing to represent the data structure visually by a graphic display. Each sort of visual display corresponds, in fact, to a special choice of the dissimilarity index as a distance and the problem turns out to be the classical isometric embedding problem in special classes of metrics.

For instance, in hierarchical classification, the case when $d$ is ultrametric corresponds to the possibility of a so-called indexed hierarchy (see [Joh67]). A natural extension is the case when $d$ is the path metric of a weighted tree, i.e., $d$ satisfies the four point condition (see Chapter ??? (on graphs)); then the data structure is called an **additive tree**. Also, data structures $(I, d)$ for which $d$ is $\ell_2$-embeddable are considered in factor analysis and multidimensional scaling. These two cases together with cluster analysis are the main three techniques for studying data structures. The case when $d$ is $\ell_1$-embeddable is a natural extension of the ultrametric and $\ell_2$ cases.

An $\ell_p$-approximation consists of minimizing the estimator $\| e \|_p$, where $e$ is a vector or a random variable (representing an error, deviation, etc. ). The following criteria are used in statistical data analysis:
- the $\ell_2$-norm, in the least square method; or its square,
- the $\ell_\infty$-norm, in the minimax or Chebychev method,
- the $\ell_1$-norm, in the least absolute values (LAV) method.

In fact, the $\ell_1$ criterion has been increasingly used. Its importance can be seen, for instance, from the volume [Dod87b] of proceedings of a conference entitled "Statistical data analysis based on the $L_1$ norm and related methods"; we refer, in particular, to [Dod87a], [Fic87], [Cal87], [Vaj87].

## 8.3 The $\ell_1$-metric in computer vision and pattern recognition

The $\ell_p$-metrics are also used in the new area called pattern recognition, or robot vision, or digital topology; see, e.g., [RK86], [Hor86].

A **computer picture** is a subset of $\mathbb{Z}^n$ (or of a scaling $\frac{1}{m}\mathbb{Z}^n$ of $\mathbb{Z}^n$) which is called a **digital** $n$-$D$-**space** (or an $n$-$D$ $m$-**quantized space**). Usually, pictures are represented in the digital plane $\mathbb{Z}^2$ or in the digital 3-$D$-space $\mathbb{Z}^3$. The points of $\mathbb{Z}^n$ are called the **pixels**.

Given a picture in $\mathbb{Z}^n$, i.e., a subset $A$ of $\mathbb{Z}^n$, one way to define its **volume** $vol(A)$ is by $vol(A) := |A|$, i.e., as the number of pixels contained in $A$. Then, the distance

$$d(A, B) = vol(A \triangle B)$$

is used in digital topology for evaluating the distance between pictures. It is a digital analogue of the symmetric difference metric used in convex geometry, where the distance between two convex bodies $A$ and $B$ in $\mathbb{R}^n$ is defined as the $n$-dimensional volume of their symmetric difference.

The above metric and other metrics on $\mathbb{Z}^n$ are used for studying analogues of clasical geometric notions as volume, perimeter, shape complexity, etc., for computer pictures.

The metrics on $\mathbb{Z}^n$ that are mainly used are the $\ell_1$-, $\ell_\infty$-metrics, as well as the $\ell_2$-metric after rounding to the nearest upper (or lower) integer.

When considered on $\mathbb{Z}^n$, the $\ell_1$-metric is also called the **grid metric** and the $\ell_\infty$-metric is called the **lattice metric** (or **Chebyshev metric**, or **uniform metric**). More specific names are used in the case $n = 2$. Then, the $\ell_1$-metric is also called the **city-block metric** (or **Manhattan metric**, or **taxi-cab metric**, or **rectilinear metric**), or the **4-metric** since each point of $\mathbb{Z}^2$ has exactly 4 closest neighbours in $\mathbb{Z}^2$ for the $\ell_1$-metric. Similarly, the $\ell_\infty$-metric on $\mathbb{Z}^2$ is called the **chessboard metric**, or the **8-metric** since each pixel has exactly 8 closest neighbours in $\mathbb{Z}^2$. Note indeed that the unit sphere $S^1_{\ell_1}$ (centered at the origin) for the $\ell_1$-norm in $\mathbb{R}^2$ contains exactly 4 integral points while the unit sphere $S^1_{\ell_\infty}$ for the $\ell_\infty$-norm contains 8 integral points.

Observe that the $\ell_1$-metric, when considered on $\mathbb{Z}^n$, can be seen as the path metric of an (infinite) graph on $\mathbb{Z}^n$. Namely, consider the graph on $\mathbb{Z}^n$ where two lattice points are adjacent if their $\ell_1$-distance is equal to 1; this graph is nothing but the usual grid. Then, the shortest path distance of two lattice points in the grid is equal to their $\ell_1$-distance. Similarly, the $\ell_\infty$-metric on $\mathbb{Z}^n$ is the path metric of the graph on $\mathbb{Z}^n$ where adjacency is defined by the pairs at $\ell_\infty$-distance one; actually, adjacency corresponds to the king move in chessboard terms.

There are some other useful metrics on $\mathbb{Z}^2$ which are obtained by combining the $\ell_1$- and $\ell_\infty$-metrics. The following two examples, the octogonal and the hexagonal distances, are path metrics; hence, in order to define them, it suffices to describe the pairs of lattice points at distance 1, i.e., to describe their unit balls.

The **octogonal distance** $d_{oct}$.
For each $(x,y) \in \mathbb{Z}^2$, its unit sphere $S^1_{oct}(x,y)$, centered at $(x,y)$, is defined by

$$S^1_{oct}(x,y) = S^3_{\ell_1}(x,y) \cap S^2_{\ell_\infty}(x,y),$$

where $S^3_{\ell_1}(x,y)$ denotes the $\ell_1$-sphere of radius 3 and $S^2_{\ell_\infty}(x,y)$ the $\ell_\infty$-sphere of radius 2, centered at $(x,y)$. Hence, $S^1_{oct}(x,y)$ contains exactly 8 integral points; note that moving from $(x,y)$ to its eight neighbours at distance 1 corresponds to the knight move in chessboard terms. Figure 2 shows the spheres $S^3_{\ell_1}$, $S^2_{\ell_\infty}$, and $S^1_{oct}$.

The **hexagonal distance** or **6-metric** $d_{hex}$.
Its unit sphere $S^1_{hex}(x,y)$, centered at $(x,y) \in \mathbb{Z}^2$, is defined by

$$S^1_{hex}(x,y) = S^1_{\ell_1}(x,y) \cup \{(x-1,y-1),(x-1,y+1)\} \text{ for } x \text{ even},$$

$$S^1_{hex}(x,y) = S^1_{\ell_1}(x,y) \cup \{(x+1,y-1),(x+1,y+1)\} \text{ for } x \text{ odd}.$$

The unit sphere $S^1_{hex}(x,y)$ contains exactly 6 integral points. Figure 3 shows the unit spheres $S^1_{hex}(0,0)$ and $S^1_{hex}(1,-3)$.

Several other modifications of the $\ell_1$-metric on the plane have been considered; see, e.g., [Ber91] and references there.

Figure 2

Figure 3

In practice, the subset $(\mathbb{Z}_k)^n := \{0, 1, \ldots, k-1\}^n$ is considered instead of the full space $\mathbb{Z}^n$. Note that $(\mathbb{Z}_2)^n$ is nothing but the vertex set of the $n$-dimensional hypercube and $((\mathbb{Z}_2)^n, d_{\ell_1})$ is the $n$-dimensional hypercube metric space. Note also that $(\mathbb{Z}_3)^2$ is the unit ball (centered at $(1,1)$) of the space $(\mathbb{Z}^n, d_{\ell_\infty})$. $(\mathbb{Z}_4)^n$ is known as the **tic-tac-toe board** (or **Rubik's $n$-cube**) and $(\mathbb{Z}_k)^2$, $(\mathbb{Z}_k)^3$ are called, respectively, the **$k$-grill** and the **$k$-framework**.

Other distances are used on $(\mathbb{Z}_k)^n$, in particular in coding theory, namely, the **Hamming distance** $d_H$ defined by

$$d_H(x, y) = |\{1 \leq i \leq n : x_i \neq y_i\}| \text{ for all } x, y \in (\mathbb{Z}_k)^n,$$

and the **Lee distance** $d_{Lee}$ defined by

$$d_{Lee}(x,y) = \sum_{1 \le i \le n} \min(|x_i - y_i|, k - |x_i - y_i|) \text{ for all } x, y \in (\mathbb{Z}_k)^n.$$

The metric space $(\mathbb{Z}_k, d_{Lee})$ can be seen as a discrete analogue of the elliptic metric space (which consists of the set of all the lines in $\mathbb{R}^2$ going through the origin and where the distance between two such lines is their angle).

The $\ell_1$-distance and the Hamming distance coincide when restricted to $(\mathbb{Z}_2)^n$, i.e., the spaces $((\mathbb{Z}_2)^n, d_{\ell_1})$ and $((\mathbb{Z}_2)^n, d_H)$ are identical. Also, $(\mathbb{Z}_k, d_{\ell_1})$ coincides with the graphic metric space of the path $P_k$ on $k$ nodes, $(\mathbb{Z}_k, d_H)$ coincides with the graphic space of the complete graph $K_k$ on $k$ nodes, and $(\mathbb{Z}_k, d_{Lee})$ coincides with the graphic space of the cycle $C_k$ on $k$ nodes. Therefore, the spaces $((\mathbb{Z}_k)^n, d_{\ell_1})$, $((\mathbb{Z}_k)^n, d_H)$ and $((\mathbb{Z}_k)^n, d_{Lee})$ coincide with the graphic space of the cartesian product $G^n$, where $G$ is $P_k$, $K_k$ and $C_k$, respectively.

One can easily check that
$(i)$ $P_k$ embeds isometrically in the $(k-1)$-dimensional hypercube, i.e., $(\mathbb{Z}_k, d_{\ell_1})$ is an isometric subspace of $((\mathbb{Z}_2)^{k-1}, d_{\ell_1})$ (simply, label each $x \in \mathbb{Z}_k$ by the binary string $1\ldots10\ldots0$ of lenght $k-1$ whose first $x$ letters are equal to 1). Hence, $((\mathbb{Z}_k)^n, \Phi d_{\ell_1})$ is an isometric subspace of $((\mathbb{Z}_2)^{n(k-1)}, d_{\ell_1})$.

$(ii)$ $((\mathbb{Z}_k)^n, d_H)$ is an isometric subspace of $((\mathbb{Z}_2)^{kn}, \frac{1}{2}d_{\ell_1})$ (label each $x \in \mathbb{Z}_k$ by the binary string of lenght $k$ whose letters are all equal to 0 except the $(x+1)$th one equal to 1).

$(iii)$ The even cycle $C_{2k}$ embeds isometrically into the $k$-dimensional hypercube. Therefore, $((\mathbb{Z}_{2k})^n, d_{Lee})$ is an isometric subspace of $((\mathbb{Z}_2)^{nk}, d_{\ell_1})$. Also, $(((\mathbb{Z}_{2k+1})^n, d_{Lee})$ is an isometric subspace of $((\mathbb{Z}_2)^{(2k+1)n}, \frac{1}{2}d_{\ell_1})$ (since the odd cycle $C_{2k+1}$ embeds isometrically into the $(2k+1)$-dimensional halfcube).

More details about the $\ell_1$-embeddings of the graphs $P_k$, $C_k$ and $K_k$ can be found in Chapter 3.

## References

[AD82] P. Assouad and M. Deza. Metric subspaces of $L^1$. *Publications mathématiques d'Orsay*, 3, 1982.

[Ale88] R. Alexander. Zonoid theory and Hilbert's fourth problem. *Geometriae Dedicata*, 28:199–211, 1988.

[AM89] D. Avis and Mutt. All facets of the six point Hamming cone. *European Journal of Combinatorics*, 10:309–312, 1989.

[Amb82] R.V. Ambarzumian. *Combinatorial integral geometry.* John Wiley and Sons, 1982.

[Ass79] P. Assouad. Produit tensoriel, distances extrémales et réalisation de covariances. *Comptes Rendus de l'Académie des Sciences de Paris*, 288(A):649–652 and 675–677, 1979.

[Ass80a] P. Assouad. Charactérisations de sous-espaces normés de $L_1$ de dimension finie. Number 19 in Séminaire d'Analyse Fonctionnelle. Ecole Polytechnique Palaiseau, 1979-1980.

[Ass80b] P. Assouad. Plongements isométriques dans $L^1$: aspect analytique. Number 14 in Séminaire Choquet (Initiation à l'Analyse). Paris VI, 1979-1980.

[Ass84] P. Assouad. Sur les inégalités valides dans $L^1$. *European Journal of Combinatorics*, 5:99–112, 1984.

[Bal87] K. Ball. Inequalities and sphere-packings in $L_p$. *Isreal Journal of Mathematics*, 58:243–256, 1987.

[Bal90] K. Ball. Isometric embedding in $\ell_p$-spaces. *European Journal of Combinatorics*, 11:305–311, 1990.

[BC76] C. Berg and P.R. Christensen. Positive definite functions on abelian semigroups. *Mathematische Annalen*, 223:253–272, 1976.

[BCK66] J. Bretagnolle, D. Dacuna Castelle, and J-L Krivine. Lois stables et espaces $L^p$. *Annales de l'Institut Henri Poincaré*, 2(3):231–259, 1966.

[Ber91] M. De Berg. On rectilinear link distances. *Computational Geometry: Theory and Applications*, 1:13–34, 1991.

[BH91] E. Boros and P.L. Hammer. The max-cut problem and quadratic 0-1 optimization; polyhedral aspects, relaxations and bounds. *Annals of Operations Research*, 33:151–180, 1991.

[Bir67] G. Birkoff. *Lattice theory.* American Mathematical Society, Providence, Rhode Island, 1967.

[Blu70] L.M. Blumenthal. *Theory and applications of discrete geometry.* Chelsea, New York, 1970.

[Bol69] E.D. Bolker. A class of convex bodies. *Transactions of the American Mathematical Society*, 145:323–345, 1969.

[Bon36] C.E. Bonferroni. Il calcolo delle assicurazioni su grouppi di teste. *Studi in Onore del Professor S.O. Carboni (Roma)*, 1936.

[Boo54] G. Boole. *An investigation of the laws of thought on which are founded the mathematical thories of logic and probabilities.* Dover Publications, original edition 1854.

[BP89] E. Boros and A. Prékopa. Closed form two-sided bounds for probabilities that at least $r$ and exactly $r$ out of $n$ events occur. *Mathematics of Operations Research*, 14(2):317–342, 1989.

[Cal87] G. Le Calve. $L_1$-embeddings of a data structure $(I, D)$. In Y. Dodge, editor, *Statistical data analysis based on the $L_1$-norm and related methods*, pages 195–202. Elsevier Science Publishers B.V. (North Holland), 1987.

[Chu41] K.L. Chung. On the probability of the occurrence of at least $m$ events among $n$ arbitrary events. *Annals of Mathematical Statistics*, 12:328–338, 1941.

[Chv80] V. Chvatal. Recognizing intersection patterns. In M. Deza and I.G. Rosenberg, editors, *Annals of Discrete Mathematics - Combinatorics 79, Part I*, volume 8, pages 249–251. North Holland, 1980.

[CP93] M. Coornaert and A. Papadopoulos. *Symbolic dynamics and hyperbolic groups*, volume 1539. Lecture Notes in Mathematics, 1993.

[Cri88] F. Critchley. On certain linear mappings between inner-product and squared-distance matrices. *Linear Algebra and its Applications*, 105:91–107, 1988.

[Dez61] M. Deza. On the Hamming geometry of unitary cubes. *Doklady Akademii Nauk SSR (in Russian) (resp. Soviet Physics Doklady (English translation))*, 134 (resp. 5):1037–1040 (resp. 940–943), 1960 (resp. 1961).

[Dez73] M. Deza. Matrices de formes quadratiques non négatives pour des arguments binaires. *Comptes rendus de l'Académie des Sciences de Paris*, 277(A):873–875, 1973.

[Dez62] M. Deza. Realizability of distance matrices in unit cubes (in Russian). *Problemy Kibernitiki*, 7:31–42, 19962.

[DM90] M. Deza and H. Maehara. Metric transforms and euclidian embeddings . *Transactions of the American Mathematical Society*, 317(2):661–671, 1990.

[Dod87a] Y. Dodge. An introduction to statistical data analysis $L_1$-norm based. In Y. Dodge, editor, *Statistical data analysis based on the $L_1$-norm and related methods*, pages 1–22. Elsevier Science Publishers B.V. (North Holland), 1987.

[Dod87b] Y. Dodge, editor. *Statistical data analysis based on the $L_1$-norm and related methods*. Elsevier Science Publishers B.V. (North Holland), 1987.

[Dor76] L.E. Dor. Potentials and isometric embeddings in $L_1$. *Israel Journal of Mathematics*, 24(3-4):260–268, 1976.

[DR84] M. Deza and I.R. Rosenberg. Intersection and distance patterns. *Utilitas Mathematica*, 25:191–212, 1984.

[DS67] D.A. Dawson and D. Sankoff. An inequality for probabilities. *Proceedings of the American Mathematical Society*, 18:504–507, 1967.

[Erd87] R.M. Erdahl. Representability conditions. In R.M. Erdahl and V.H. Smith, editors, *Density matrices and density functionals*, pages 51–75. D. Reidel, 1987.

[Fic87] B. Fichet. The role played by $L_1$ in data analysis. In Y. Dodge, editor, *Statistical data analysis based on the $L_1$-norm and related methods*, pages 185–193. Elsevier Science Publishers B.V. (North Holland), 1987.

[Gal77] J. Galambos. Bonferroni inequalities. *Annals of Probability*, 5:577–581, 1977.

[Gow85] J.C. Gower. Properties of euclidian and non-euclidian distance matrices. *Linear Algebra and its Applications*, 67:81–97, 1985.

[Gra93] D.A. Grable. Sharpened Bonferroni inequalities. *Journal of Combinatorial Theory B*, 57:131–137, 1993.

[Gri92] V.P. Grishukhin. On a $t$-extension of a distance space. Technical report SOCS-92.5, School of Computer Science, McGill University, Montreal, Canada, 1992.

[Ham65] P.L. Hammer. Some network flow problems solved with pseudo-boolean programming. *Operations Research*, 13:388–399, 1965.

[Hor86] B.K.P. Horn. *Robot vision*. MIT Press (Cambridge Mass.), 1986.

[HW88] T.L. Hayden and J. Wells. Approximations by matrices positive semidefinite on a subspace. *Linear Algebra and its Applications*, 109:115–130, 1988.

[HWLT91] T.L. Hayden, J. Wells, Wei-Min Liu, and P. Tarazaga. The cone of distance matrices. *Linear Algebra and its applications*, 144:153–169, 1991.

[Isa89] A.N. Isachenko. On the structure of the quadratic boolean problem polytope. In *Combinatorics and Graph Theory*, volume 25, pages 87–91. Banach Center Publications, 1989.

[Joh67] S.C. Johnson. Hierarchical clustering schemes. *Psychometrika*, 32(3):241–254, 1967.

[Kar85] A.V. Karzanov. Metrics and undirected cuts. *Mathematical Programming*, 32:183–198, 1985.

[Kel70] J.B. Kelly. Metric inequalities and symmetric differences. In O.Shisha, editor, *Inequalities II*, pages 193–212. Academic Press, New York, 1970.

[KM76] S. Kounias and J. Marin. Best linear Bonferroni bounds. *SIAM Journal of Applied Mathematics*, 30(2):307–323, 1976.

[KS93] J.H. Koolen and S.V. Shpectorov. Distance-regular graphs whose distance matrix has only one positive eigenvalue. 1993.

[LLR93] N. Linial, E. London, and Y. Rabinovitch. On the metric dimension of graphs. 1993.

[LV93] Y.I. Lyubich and L.N. Vaserstein. Isometric embeddings between classical banach spaces, cubature formulas, and spherical designs. *Geometriae Dedicata*, 47:327–362, 1993.

[MD72] W.B. McRae and E.R. Davidson. Linear inequalities for density matrices II . *Journal of Mathematical Physics*, 13(10):1527–1538, 1972.

[Men28] K. Menger. Untersuchungen über allgemeine metrik. *Mathematische Annalen*, 100:75–163, 1928.

[Mes87] M.M. Mestechkin. On the diagonal $N$-representability problem. In R.M. Erdahl and V.H. Smith, editors, *Density matrices and density functionals*, pages 77–87. D. Reidel, 1987.

[MM92] S.M. Malitz and J.I. Malitz. A bounded compactness theorem for $L^1$-embeddability of metric spaces in the plane. *Discrete Computational Geometry*, 8:373–385, 1992.

[MS58] F. Marczewski and H. Steinhaus. On a certain distance of sets and the corresponding distance of functions. *Colloquium Mathematicum*, 6:319–327, 1958.

[Nod87] A. Noda. Generalized Radon transform and Lévy's Brownian motion I and II. *Nagoya Mathematical Journal*, 105:71–87 and 89–107, 1987.

[Nod89] A. Noda. White noise representation of some Gaussian random fields. *The bulletin of Aichi University of Education*, 38:35–45, 1989.

[Pad89] M. Padberg. The boolean quadric polytope: some characteristics, facets and relatives. *Mathematical Programming*, 45:139–172, 1989.

[Pit86] I. Pitowsky. The range of quantum probability. *Journal of Mathematical Physics*, 27(6):1556–1565, 1986.

[Pit91] I. Pitowsky. Correlation polytopes: their geometry and complexity. *Mathematical Programming*, 50:395–414, 1991.

[Rac91] S.T. Rachev. *Probability metrics and the stability of stochastic models.* John Wiley and Sons, 1991.

[RK86] A. Rosenfeld and A.C. Kak. *Digital picture prcessing.* Academic Press, 1986.

[Sch35] I.J. Schoenberg. Remarks to M.Fréchet's article "Sur la définition axiomatique d'une classe d'espaces vectoriels distanciés applicables vectoriellement sur l'espace de Hilbert". *Annals of Mathematics*, 36:724–732, 1935.

[Sch37] I.J. Schoenberg. On certain metric spaces arising from Euclidian spaces by a change of metric and their imbedding in Hilbert space. *Annals of Mathematics*, 38(4):787–793, 1937.

[Sch38a] I.J. Schoenberg. Metric spaces and completely monotone functions. *Annals of Mathematics*, 39(4):811–841, 1938.

[Sch38b] I.J. Schoenberg. Metric spaces and positive definite functions. *Transactions of the American Mathematical Society*, 44:522–536, 1938.

[Sim90] C. De Simone. The cut polytope and the boolean quadric polytope. *Discrete Mathematics*, 79:71–75, 1989/90.

[SW83] R. Schneider and W. Weil. Zonoids and related topics. In P. Gruber and J. Wills, editors, *Convexity and its applications*, pages 296–317. Birkhäuser Verlag Basel, 1983.

[Vaj87] I. Vajda. $L_1$-distances in statistical inference: comparison of topological functional and statistical properties. In Y. Dodge, editor, *Statistical data analysis* *based on the $L_1$-norm and related methods*, pages 177–184. Elsevier Science Publishers B.V. (North Holland), 1987.

[vNS41] J. von Neumann and I.J. Schoenberg. Fourier integrals and metric geometry. *Transactions of the American Mathematical Society*, 50:226–251, 1941.

[Wit78] H.S. Witsenhausen. A support characterization of zonotopes. *Mathematika*, 25:13–16, 1978.

[Wol67] D. Wolfe. Imbedding of a finite metric set in an $n$-dimensional Minkowski space. *Nederl. Acad. Wetensch. Proc. Ser. A*, 70:136–140, 1967.

[WW75] J.H. Wells and L.R. Williams. *Embeddings and extensions in analysis.* Springer-Verlag, 1975.

[Yos70] M.L. Yoseloff. *A combinatorial approach to the diagonal $N$-representability problem.* PhD thesis, Princeton University, 1970.