Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

C. R. Acad. Sci. Paris, Ser. I 340 (2005) 599–604

COMPTES RENDUS
MATHEMATIQUE

http://france.elsevier.com/direct/CRASS1/

Probability Theory/Statistics

# About the optimal density associated to the chiral index of a sample from a bivariate distribution

Don Coppersmith[a], Michel Petitjean[b]

[a] *IBM TJ Watson Research Center, Yorktown Heights, New York 10598, USA*
[b] *ITODYS (CNRS, UMR 7086, université Paris 7), 1, rue Guy de la Brosse, 75005 Paris, France*

Received 29 November 2004; accepted after revision 1 March 2005

Presented by Paul Deheuvels

**Abstract**

The complex quadratic form $z'Pz$, where $z$ is a fixed vector in $\mathbf{C}^n$ and $z'$ is its transpose, and $P$ is any permutation matrix, is shown to be a convex combination of the quadratic forms $z'P_\sigma z$, where $P_\sigma$ denotes the symmetric permutation matrices. We deduce that the optimal probability density associated to the chiral index of a sample from a bivariate distribution is symmetric. This result is used to locate the upper bound of the chiral index of any bivariate distribution in the interval $[1-1/\pi, 1-1/2\pi]$. ***To cite this article: D. Coppersmith, M. Petitjean, C. R. Acad. Sci. Paris, Ser. I 340 (2005).***

**Résumé**

**À propos de la densité optimale associée à l'indice chiral d'un échantillon d'une distribution bivariée.** Nous montrons que la forme quadratique complexe $z'Pz$, où $z$ est un vecteur donné dans $\mathbf{C}^n$ et $z'$ est son transposé, et $P$ est une matrice de permutation, est une combinaison convexe des formes quadratiques $z'P_\sigma z$, où les $P_\sigma$ sont des matrices de permutation symétriques. On en déduit que la densité de probabilité optimale associée à l'indice chiral d'un échantillon d'une distribution bivariée est symétrique. Ce résultat est utilisé pour localiser la borne supérieure de l'indice chiral d'une distribution bivariée quelconque dans l'intervalle $[1-1/\pi, 1-1/2\pi]$. ***Pour citer cet article : D. Coppersmith, M. Petitjean, C. R. Acad. Sci. Paris, Ser. I 340 (2005).***

*E-mail addresses:* dcopper@us.ibm.com (D. Coppersmith), petitjean@itodys.jussieu.fr (M. Petitjean).
*URLs:* http://www.research.ibm.com/people/c/coppersmith/, http://petitjeanmichel.free.fr/itoweb.petitjean.html.

1631-073X/$ – see front matter © 2005 Académie des sciences. Published by Elsevier SAS. All rights reserved.
doi:10.1016/j.crma.2005.03.011

## Version française abrégée

L'indice chiral $\chi$ d'une distribution multivariée est défini à partir de la distance de Wasserstein [5] entre cette distribution et son image par tranformation orthogonale de déterminant $-1$, cette distance étant minimisée pour toutes les rotations et translations de l'image, et normalisée à l'inertie [2].

L'indice chiral d'un échantillon de taille $n$ d'une distribution bivariée d'inertie $T$, s'exprime dans le plan complexe à l'aide de la forme quadratique complexe $z'Pz$, dans laquelle $z$ est un vecteur complexe à $n$ composantes, $z'$ est son transposé non conjugué, $T = \|z\|^2/n$, et $P$ est une matrice de permutation égale à $n$ fois la matrice des probabilités conjointes associée à la distance de Wasserstein :

$$\chi = 1 - \Big[\max_{\{P\}}(z'Pz)\Big]\Big/nT.$$

Dans une première partie, nous montrons qu'il existe toujours une permutation optimale symétrique.

Dans une seconde partie, nous utilisons ce résultat pour localiser la borne supérieure de $\chi$ dans le plan complexe. Nous exhibons une famille de distributions dont l'indice chiral est arbitrairement proche de $1 - 1/\pi$, puis nous obtenons un majorant égal à $1 - 1/2\pi$.

L'extension aux distributions bivariées quelconques (d'inertie finie et non nulle) est faite via un théorème de convergence de la littérature.

## 1. Introduction

The chiral index $\chi$ of a finite variance $d$-variate probability distribution $\mathcal{P}$ is the Wasserstein distance [5] between the distribution $\mathcal{P}$ and its inverted image $\overline{\mathcal{P}}$, minimized for all rotations and translations of $\overline{\mathcal{P}}$, and normalized to the inertia of $\mathcal{P}$ [2]. It takes values over $[0, 1]$. It is a skewness measure offering various applications in computer sciences [3]. In the case of a sample of size $n$, the optimal joint density between and $\mathcal{P}$ and $\overline{\mathcal{P}}$ is known to exist [5]. The matrix associated to this optimal density is shown to be $(1/n)$ times a permutation matrix [2,4]. In the univariate case, this permutation matrix is symmetric [1]. We extend the result in this Note to the bivariate case: the optimal joint density is symmetric. The upper bound of the chiral index of a $d$-variate distribution is unknown, except in the univariate case, for which it is 1/2 [2]. In the bivariate case, the symmetry of the optimal joint density of a sample is used to locate the upper bound in $[1 - 1/\pi, 1 - 1/2\pi]$.

## 2. Symmetry of the optimal permutation

We first need to establish two theorems in the complex plane. We fix a complex vector $z = (z_1, z_2, \ldots, z_n) \in\mathbf{C}^n$. Given a permutation $\sigma$ on $n$ indices $\{1, 2, \ldots, n\}$, define the quadratic form $z'P_\sigma z = \sum_{j=1}^n z_j z_{\sigma(j)}$, where $P_\sigma$ is the permutation matrix associated to $\sigma$.

**Theorem 2.1.** *For any permutation $\tau$, the complex number $z'P_\tau z$ is in the convex hull of the set $\{z'P_\sigma z\colon\ P_\sigma = P'_\sigma\}$.*

$\Re()$ denoting the real part of a complex number, this following lemma will be crucial:

**Lemma 2.2.** *If $\tau$ is an $n$-cycle, then there is a symmetric permutation $\sigma$ satisfying $\Re(z'P_\sigma z) \leqslant \Re(z'P_\tau z)$.*

**Proof of Lemma 2.2.** Let $\tau = (1, 2, 3, \ldots, n)$. When $n \in \{1, 2\}$ the result is immediate: $\tau$ itself is a symmetric permutation. When $n$ is even, define two permutations (as products of $\frac{n}{2}$ disjoint 2-cycles): $\alpha = (1, 2)(3, 4)\cdots(n - 1, n)$, and $\beta = (2, 3)(4, 5)\cdots(n, 1)$, and compute that $z'P_\tau z = [z'P_\alpha z + z'P_\beta z]/2$. So $z'P_\tau z$ is a convex combination of $z'P_\alpha z$ and $z'P_\beta z$, whence $\min\{\Re(z'P_\alpha z), \Re(z'P_\beta z)\} \leqslant \Re(z'P_\tau z)$.

We are left with the case where $n$ is odd, $n \geqslant 3$. Consider the following $2n$ permutations:

$$\left.\begin{aligned} \alpha_j &= (j)(j+1, j+2)(j+3, j+4)\cdots(j-2, j-1) \\ \beta_j &= (j-1, j+1)(j)(j+2, j+3)(j+4, j+5)\cdots(j-3, j-2) \end{aligned}\right\} \quad 1 \leqslant j \leqslant n.$$

We compute: $z'P_{\alpha_j}z + z'P_{\alpha_{j+1}}z - 2z'P_\tau z = z_j^2 + z_{j+1}^2 - 2z_jz_{j+1} = (z_{j+1} - z_j)^2$, where we are considering the indices modulo $n$, so that if $j = n$ then $z_{j+1} = z_1$.

$$z'P_{\alpha_j}z + z'P_{\beta_j}z - 2z'P_\tau z = 2z_j^2 + 2z_{j-1}z_{j+1} - 2z_{j-1}z_j - 2z_jz_{j+1} = -2(z_{j+1} - z_j)(z_j - z_{j-1}).$$

Now suppose the lemma is false, so that for all $j$, $\Re(z'P_{\alpha_j}z) > \Re(z'P_\tau z)$ and $\Re(z'P_{\beta_j}z) > \Re(z'P_\tau z)$. Then for all $j$, $\Re[(z_{j+1} - z_j)^2] > 0$ and $\Re[-2(z_{j+1} - z_j)(z_j - z_{j-1})] > 0$.

Fix $j$. Define $b = z_{j+1} - z_j$, $c = z_j - z_{j-1}$. We have just seen that $\Re(b^2) > 0$ and $\Re(c^2) > 0$ and $\Re(bc) < 0$ (so that $\Re(b^2) \neq 0$ and $\Re(c^2) \neq 0$ and $\Re(bc) \neq 0$). Observe also that $(\Re(b)c - \Re(c)b)$ is pure imaginary, so its square is real and nonpositive: $\Re(b)^2c^2 + \Re(c)^2b^2 - 2\Re(b)\Re(c)bc \leqslant 0$. Taking the real parts of all terms and rearranging, $2\Re(b)\Re(c)\Re(bc) \geqslant \Re(b)^2\Re(c^2) + \Re(c)^2\Re(b^2) > 0$, and from $\Re(bc) < 0$ we conclude $\Re(b)\Re(c) < 0$. So the sign of $\Re(b) = \Re(z_{j+1} - z_j)$ and the sign of $\Re(c) = \Re(z_j - z_{j-1})$ are opposite. As $j$ cycles around $1, 2, \ldots, n, 1$, the signs of $\Re(z_{j+1} - z_j)$ alternate. But $n$ is odd, so this alternation is impossible. The contradiction proves Lemma 2.2. □

**Lemma 2.3.** *If $\tau$ is an $n$-cycle, then $z'P_\tau z$ is in the convex hull of the set $\{z'P_\sigma z\colon\ P_\sigma = P'_\sigma\}$.*

**Proof of Lemma 2.3.** Suppose the conclusion is false. Then there is a line $\ell$ through $z'P_\tau z$ in the complex plane, with all $\{z'P_\sigma z\colon\ \sigma = \sigma^{-1}\}$ lying on one side of the line. If $\ell$ has direction $\theta$, we have $\Re[z'P_\sigma z\,\mathrm{e}^{\mathrm{i}(\pi/2-\theta)}] > \Re[z'P_\tau z\,\mathrm{e}^{\mathrm{i}(\pi/2-\theta)}]$ for all $\sigma$ with $\sigma = \sigma^{-1}$. Now set $w = z\,\mathrm{e}^{\mathrm{i}(\pi/2-\theta)/2}$ so that $w'P_\tau w = \mathrm{e}^{\mathrm{i}(\pi/2-\theta)}z'P_\tau z$, and apply Lemma 2.2. □

**Proof of Theorem 2.1.** Express an arbitrary permutation $\tau$ as a product of disjoint cycles $\tau_j$; apply the proof of Lemma 2.3 to each cycle.

Then Theorem 2.4 is deduced immediately from Theorem 2.1:

**Theorem 2.4.** *The modulus of $z'Pz$ is maximized by a symmetric permutation matrix $P_\sigma$.*

It is pointed out that non symmetric permutations may be optimal (e.g. when $z$ has several identical elements).

## 3. Application to the chiral index

As mentioned in the introduction, the chiral index $\chi$ is a parameter measuring the degree of asymmetry of a multivariate distribution $\mathcal{P}$ having a finite and non null inertia $T$. It takes values in the interval $[0, 1]$. It is null if and only if the distribution is identical to any of its images $\overline{\mathcal{P}}$ generated by the composition of a translation and an orthogonal transformation with determinant $-1$.

Here we need to locate the upper bound of $\chi$ for bivariate distributions, and provide informations on the extreme chirality distributions. The results hereafter are obtained via complex analysis techniques applied to samples of bivariate distributions, rather than via probability calculations. The extension to parent distributions will be made with a published convergence theorem.

We set $z = x + \mathbf{i}y$ and $X = [x|y]$, $x$ and $y$ being fixed vectors in $\mathbf{R}^d$, so that $|z'Pz| = \lambda_1 - \lambda_2$, where $\lambda_1$ and $\lambda_2$ are the eigenvalues of $X'(\frac{P+P'}{2})X$.

The chiral index $\chi$ is computed at null expectation from Eqs. (6) and (7) in [1]. For a sample of a bivariate distribution $\mathcal{P}$ with inertia $T = \|z\|^2/n$ and $\|z\|^2 = x'x + y'y$, it is known [1] to be:

$$\chi = 1 - \Big[\max_{\{P\}}(\lambda_1 - \lambda_2)\Big]\Big/ nT. \tag{1}$$

The matrix associated to the joint density between $\mathcal{P}$ and $\overline{\mathcal{P}}$ is $[P/n]$, and the minimized Wasserstein distance between $\mathcal{P}$ and $\overline{\mathcal{P}}$ is $T\chi$ [2].

Thus Theorem 3.1 is deduced immediately from Theorem 2.4:

**Theorem 3.1.** *The optimal joint density matrix* $[P/n]$ *of the finite discrete bivariate distributions* $\mathcal{P}$ *and* $\overline{\mathcal{P}}$ *is symmetric.*

We consider now the more general situation where the $n$ points are partitioned into groups of colors [1,2,4]. Permutations involving cycles over two groups are no more considered, and the optimal permutation is taken over a subset of the $n!$ permutations. Obviously, Theorems 2.1 and 2.4 stand again, and the optimal joint density matrix is still symmetric. Colors are not further considered in this Note.

## 4. Localization of the upper bound of the chiral index: part 1

We exhibit here a family of centered sets for which the ratio $\max_{\{P\}}|z'Pz|/\|z\|^2$ is arbitrarily close to $1/\pi$. 'Centering' means working at null expectation. It means here that $\mathbf{1}'z = 0$, where $\mathbf{1}$ is a vector in $\mathbf{C}^n$ each of whose elements is 1. It is also recalled that the ratio is insensitive to an arbitrary planar rotation (phase).

**Lemma 4.1.** *The upper bound of the chiral index of a bivariate sample cannot be smaller than* $1 - 1/\pi$.

**Proof of Lemma 4.1.** Fix $\varepsilon > 0$. Choose an even integer $m > 1/\varepsilon$. Let $\omega = \mathrm{e}^{\mathbf{i}2\pi/(2m)}$ be a complex root of unity, so that $\omega^{2m} = 1$. Select an integer $r > m^4/\varepsilon^2$ and an even integer $k > r^{m-1}/\varepsilon$. The complex vector $z$ has $n = (1 + r + r^2 + \cdots + r^{m-1} + 2k)$ elements as follows. There are $m + 3$ blocks labelled $j = 0, \ldots, m + 2$, each consisting of identical elements. For $j < m$, block $j$ has $r^j$ identical elements with value $\omega^j/r^{j/2}$. Let $S$ denote the sum of these elements: $S = \sum_{j=0}^{j=m-1}\omega^j r^{j/2}$. Block $m$ contains $k$ identical elements with value $-S/k$; block $m+1$ contains $k/2$ elements with value $iS/k$; and block $m+2$ contains $k/2$ elements with value $-iS/k$. The sum of elements of $z$ is zero: block $m$ cancels the first $m$ blocks, and blocks $m + 1$, $m + 2$ cancel each other. Also, the sum of squares of elements of $z$ is zero: the squares of elements in the first $m$ blocks add to $\sum_{j=0}^{m-1}\omega^{2j} = 0$, while blocks $m + 1$ and $m + 2$ cancel block $m$. One can compute $x'x = y'y = m/2 + \mathrm{O}(\varepsilon)$ and $x'y = 0$.

We know from Theorem 2.4 that the optimal permutation $P$ pairs the elements of $z$, some being paired with themselves when $P$ contains 1-cycles. Let $B_j$ be the number of elements paired within the block $j$. We set $\beta_j = B_j/r^j$, so that $0 \leqslant \beta_j \leqslant 1$ for $j = 0, \ldots, m-1$. The contribution of these elements to $z'Pz$ is $\beta_j\omega^{2j}$.

One can see that the contribution to $z'Pz$ of the elements paired between two different blocks $j_1$ and $j_2$ is $\mathrm{O}(1/r^{(1/2)|j_1-j_2|}) = \mathrm{O}(\varepsilon/m^2)$ when $j_1 < m$ and $j_2 < m$, so that the $m^2 - m$ off-diagonal blocks contribute a total of $\mathrm{O}(\varepsilon)$ to $z'Pz$. The contribution of the elements paired between the blocks $j < m$ and blocks $m, m+1, m+2$ is $\mathrm{O}(r^{m-1}(\frac{1}{\sqrt{r^{m-1}}})\frac{|S|}{k}) = \mathrm{O}(\varepsilon)$. The contribution of the elements paired within the last three blocks is $\mathrm{O}(k(\frac{|S|}{k})^2) = \mathrm{O}(\varepsilon)$. All these contributions sum to at most $\mathrm{O}(\varepsilon)$, except for the diagonal terms $j_1 = j_2 < m$.

We look for $\lim_{m\to\infty}\max_{\{P\}}|z'Pz|/\|z\|^2$. When $m$ is arbitrarily large, we look for the $m$ values $\beta_j$ maximizing: $|\sum_{j=0}^{j=m-1}\beta_j\omega^{2j} + \mathrm{O}(\varepsilon)|/(m + \mathrm{O}(\varepsilon))$.

The complex number $\gamma = \sum_{j=0}^{j=m-1}\beta_j\omega^{2j}$ is the sum of $m$ terms having all modulus in the interval $[0, 1]$. Neglecting the term $\mathrm{O}(\varepsilon)$, we can see that only the terms offering a difference of phase $\phi$ such that $\cos(\phi) \geqslant$

$\beta_j/2|\gamma|$ will contribute to the modulus of $\gamma$. Since $|\gamma|$ tends to infinity when $m$ tends to infinity, only the terms having a difference of phase within $[-\pi/2, +\pi/2]$ with $\gamma$ will contribute to the modulus of $\gamma$.

For these latter we set $\beta_j = 1$, and we set $\beta_j = 0$ elsewhere. Working with a free arbitrary phase, we have:

$$\gamma = 1 + \omega^2 + \omega^4 + \cdots + \omega^{2(m-1)/2} = \frac{1-\omega^m}{1-\omega^2} = \frac{2}{(-\mathbf{i}\omega)(2\sin(2\pi/2m))}.$$

Its modulus is $|\gamma| = \frac{1}{\sin(\pi/m)} = \frac{m}{\pi} + \mathrm{O}(\frac{1}{m}) = \frac{m}{\pi} + \mathrm{O}(\varepsilon)$. Therefore:

$$\lim_{m\to\infty} \left\{ \max_{\{P\}} |z'Pz|/\|z\|^2 \right\} = \frac{1}{\pi}.$$

Our family of sets $z$ has a chiral index arbitrarily close to $1 - 1/\pi$, thus Lemma 4.1 is proved.

## 5. Localization of the upper bound of the chiral index: part 2

We first show that no set can have the ratio $\max_{\{P\}} |z'Pz|/\|z\|^2$ smaller than $1/\pi$ under the additional condition that at least half of the $n$ elements $z_j$ are null. The centering condition is not set here.

**Lemma 5.1.** *For any complex vector having at least half of its elements null, we have the following inequality*: $[\max_{\{P\}} |z'Pz|/\|z\|^2] \geqslant 1/\pi$.

**Proof of Lemma 5.1.** We consider an arbitrary phase $\theta$ and its associated permutation $P_\theta$ such that $z_j$ is paired with itself when $\Re(z_j^2\,\mathrm{e}^{\mathbf{i}\theta}) > 0$ and $z_j$ is paired with a null element when $\Re(z_j^2\,\mathrm{e}^{\mathbf{i}\theta}) \leqslant 0$. Setting $z_j^2 = r_j\,\mathrm{e}^{\mathbf{i}\phi_j}$, we have $\mathrm{e}^{\mathbf{i}\theta} z'P_\theta z = \sum r_j\,\mathrm{e}^{\mathbf{i}(\theta+\phi_j)}$, and since $|z'P_\theta z| \geqslant |\Re(\mathrm{e}^{\mathbf{i}\theta} z'P_\theta z)|$, we have:

$$\frac{|z'P_\theta z|}{\|z\|^2} \geqslant \frac{\sum r_j \max\{0, \cos(\theta+\phi_j)\}}{\sum r_j}.$$

The numerator of the right member of the inequality above is a continuous function of $\theta$ maximized for some unknown value of $\theta$, where $\theta/2$ is the phase of the free rotation. Although the maximum is difficult to locate, it cannot be smaller than the mean value of the function. This mean value is:

$$\frac{1}{2\pi} \int_{\theta=0}^{\theta=2\pi} \sum r_j \max\bigl\{0, \cos(\theta+\phi_j)\bigr\}\,\mathrm{d}\theta.$$

Permuting the two summation operators, we are left with a finite sum of integrals, each of them being equal to $2r_j$. The mean value of the function is $2\sum r_j/2\pi$, thus proving Lemma 5.1. □

The condition 'at least half of the elements are null' is asymptotically satisfied for the sets considered in the previous section. Now we remove this condition, and we set instead the centering condition $\mathbf{1}'z = 0$.

**Lemma 5.2.** *The chiral index of any bivariate sample cannot be greater than* $1 - 1/2\pi$.

**Proof of Lemma 5.2.** We know that there exists $\theta$ such that: $\sum r_j \max\{0, \cos(\theta+\phi_j)\} \geqslant (1/\pi)\sum r_j$.

Let $k$ be the number of elements $z_j$ such that $\Re(z_j^2) < 0$. We define $z_k$ as the $k$-dimensional vector such that $\Re(z_j^2) < 0$, and $z_{n-k}$ as the $(n-k)$-dimensional vector such that $\Re(z_j^2) \geqslant 0$, such that $\mathbf{1}'z_k + \mathbf{1}'z_{n-k} = 0$. We set the arbitrary phase such that $\theta = 0$, without loss of generality: $\sum \max\{0, \Re(z_j^2)\} \geqslant (1/\pi)\sum |z_j|^2$.

Then we build the matrix $[nW_+]$, such that $[W_+]$ is a joint density matrix and $[nW_+]$ is a doubly stochastic matrix, as follows:

$$(n+k)[nW_+] = \mathbf{1}\cdot\mathbf{1}' + \left(\begin{array}{c|c} -\mathbf{1}\cdot\mathbf{1}' + n\mathbf{I} & 0 \\ \hline 0 & \mathbf{1}\cdot\mathbf{1}' \end{array}\right)$$

in which $\mathbf{I}$ is the identity matrix of size $n-k$ and the vectors $\mathbf{1}$ have the appropriate size (either $k$, or $n-k$, or $n$). Building $z' = [z'_{n-k}|z'_k]$, we have: $z'[nW_+]z = \frac{n}{n+k}(z'_{n-k}z_{n-k})$, and since the real part cannot exceed the modulus, we get $|z'[nW_+]z| \geqslant \frac{n}{n+k}(\frac{1}{\pi})\sum |z_j|^2$.

The permutation matrices are the extreme points of the closed bounded convex set of bistochastic matrices. Then, $\max_{\{P\}} |z'Pz| \geqslant |z'[nW_+]z|$ and:

$$\max_{\{P\}} |z'Pz| \geqslant \frac{n}{n+k}\left(\frac{1}{\pi}\right)\|z\|^2, \tag{2}$$

and since $k \leqslant n$: $\max_{\{P\}} |z'Pz|/\|z\|^2 \geqslant 1/2\pi$, which proves Lemma 5.2.

A slight improvement is obtained when the condition $z'z = 0$ is added. We build the doubly stochastic matrix $[nW_-]$:

$$\bigl(n+(n-k)\bigr)[nW_-] = \mathbf{1}\cdot\mathbf{1}' + \left(\begin{array}{c|c} \mathbf{1}\cdot\mathbf{1}' & 0 \\ \hline 0 & -\mathbf{1}\cdot\mathbf{1}' + n\mathbf{I} \end{array}\right).$$

Then: $z'[nW_-]z = \frac{n}{n+(n-k)}(z'_k z_k)$. Since $z'_k z_k = -z'_{n-k}z_{n-k}$, we are led to the same inequalities as above, except that the factor $n/(n+k)$ is now replaced by $n/(n+(n-k))$:

$$\max_{\{P\}} |z'Pz| \geqslant \frac{n}{n+(n-k)}\left(\frac{1}{\pi}\right)\|z\|^2. \tag{3}$$

Depending which of $k$ or $(n-k)$ is the smaller, the largest of the ratios $n/(n+k)$ and $n/(n+(n-k))$ cannot be smaller than $2/3$, and thus: $\max_{\{P\}} |z'Pz|/\|z\|^2 \geqslant 2/3\pi$, corresponding to a chiral index upper bounded by $1-2/3\pi$.

The condition $z'z = 0$, i.e. $x'x = y'y$ and $x'y = 0$, means that the variance matrix of the centered set $[x|y]$ is proportional to the identity matrix. This condition is asymptotically satisfied by the sets described in the previous section.

From Lemmas 4.1 and 5.2, the upper bound of the chiral index of any bivariate sample is lying somewhere in the interval $[1-1/\pi, 1-1/2\pi]$. From the convergence theorem in section IV in [2], we deduce Theorem 5.3:

**Theorem 5.3.** *The upper bound of the chiral index of a bivariate distribution lies in the interval* $[1-1/\pi, 1-1/2\pi]$.

The family of bivariate distributions described in Section 4 are conjectured to be asymptotically of maximal chirality. In higher dimensions, finding the upper bound of the chiral index and exhibiting extreme chirality distributions are open problems.

## References

[1] M. Petitjean, About second kind continuous chirality measures. 1. Planar sets, J. Math. Chem. 22 (1997) 185–201.

[2] M. Petitjean, Chiral mixtures, J. Math. Phys. 43 (2002) 4147–4157.

[3] M. Petitjean, Chirality and symmetry measures: a transdisciplinary review, Entropy 5 (2003) 271–312, http://www.mdpi.net/entropy.

[4] M. Petitjean, From shape similarity to shape complementarity: toward a docking theory, J. Math. Chem. 35 (2004) 147–158.

[5] S.T. Rachev, Probability Metrics and the Stability of Stochastic Models, Wiley, New York, 1991.