HAL open science

Statistical properties of line spectrum pairs.

Jean-Yves Tourneret

To cite this version:

Jean-Yves Tourneret. Statistical properties of line spectrum pairs.. Signal Processing, 1998, 6 (2), pp.239-255. 10.1016/S0165-1684(97)00221-1 . hal-03485289

HAL Id: hal-03485289

https://hal.science/hal-03485289v1

Submitted on 17 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Statistical properties of line spectrum pairs

Jean-Yves Tourneret*

ENSEEIHT/GAPSE, 2 rue Camichel, 31071 Toulouse, France

Abstract

An original method for computing the Probability Density Function (PDF) of the Line Spectrum Pair (LSP) coefficient vector as a function of the PDF of the AutoRegressive (AR) parameter vector is derived. Being linked by non-linear relations, AR parameters and LSP coefficients cannot both be Gaussian. This strict theoretical result could hide an approximately Gaussian distribution for the LSP parameters. Some distances between the LSP parameter PDF and its asymptotic Gaussian PDF are then studied. In a pattern recognition application, the LSP parameter statistics allow us to study the convexity of LSP point clusters, conditioned to each class: the LSP parameters seem to be attractive for pattern recognition when used with the centroid distance rule. © 1998 Elsevier Science B.V. All rights reserved.

Zusammenfassung

In diesem Artikel wird eine neue Methode zur Berechnung der Wahrscheinlichkeitsdichtefunktion (engl. *probability density function*, PDF) des Koeffizientenvektors von Linienspektrumpaaren (LSP) abgeleitet. Die neue Dichte ergibt sich dabei als Funktion der Dichte des Parametervektors eines autoregressiven (AR) Prozesses. Da AR Parameter und LSP Koeffizienten durch nichtlineare Beziehungen miteinander verknüpft sind, können nicht beide Größen Gauss-verteilt sein. Dieses strikte theoretische Resultat könnte eine approximative Gaussverteilung der LSP Parameter verbergen. Daraufhin werden einige Distanzmaße zwischen den LSP Parameterdichten und ihre asymptotischen Gaußschen Dichten untersucht. Bei Anwendung auf ein Mustererkennungsproblem erlauben die Verteilungseigenschaften der LSP Parameter die Untersuchung der Konvexität der LSP Punkthaufen, die die vorliegenden Klassen bestimmen. Daher bieten sich LSP Parameter als geeignete Werkzeuge bei der Mustererkennung an, wenn sie im Zusammenhang mit der *centroid distance rule* eingesetzt werden. © 1998 Elsevier Science B.V. All rights reserved.

Résumé

Nous présentons une méthode originale permettant de calculer la densité de probabilité des coefficients LSP en fonction de celle des paramètres AR. La relation liant ces deux ensembles de paramètres étant non-linéaire, ils ne peuvent être simultanément gaussiens. Ce résultat purement théorique peut cacher une loi très proche de la loi normale pour les coefficients LSP. Plusieurs distances entre la loi des coefficients LSP et la loi normale asymptotique associée sont alors étudiées. La statistique des coefficients LSP nous permet d'analyser la convexité des nuages de points obtenus lors d'une

* Tel.: (33) 5 61 58 83 14; fax: (33) 5 61 58 82 37; e-mail: tourneret@len7.enseeiht.fr.

classification avec ces coefficients: les coefficients LSP semblent bien adaptés à la reconnaissance des formes avec la règle de la distance au barycentre.

Keywords: AR parameters; Line spectrum pairs; Non-Gaussian; Non-linear; Distances; Pattern recognition; Estimation

1. Introduction

Many sets of parameters equivalent to AutoRegressive (AR) parameters have been used for quantization and coding. Reflection coefficients or PARCOR coefficients have shown interesting properties which make them attractive for quantization. They ensure stability of the AR filter after quantization and there exists a natural order in these parameters [10]. Log-Area Ratios (LARs) and arcsine reflection coefficients have then been shown to be optimal in the sense of minimizing a maximum spectral deviation [7]. Recently, the Line Spectrum Pair coefficients (LSP) have been shown to outperform these AR parameter representations in many speech processing applications [12,15,16]. This result is mainly due to an uncorrelated property of these coefficients [5,6].

This paper derives some statistical properties of the LSP parameters in the case of a Gaussian AR parameter vector denoted as a with mean m_a and covariance matrix C_a. The Gaussian assumption for AR parameters can be justified in three applications [19–21]:

- In Estimation theory, AR parameters are deterministic and have to be estimated. According to the Mann and Wald theorem, most commonly used AR parameter estimators can be assumed to be Gaussian when the parameters are estimated for sufficiently large data records [8]. Moreover, most of these estimators are asymptotically efficient. It means that, for large number of samples, the covariance matrix of a can be approximated by the Cramer–Rao bound

$$C_a \simeq \frac{\sigma_e^2}{n} R^{-1},$$

  σ_e^2 is the driving noise variance, n is the number of samples and R is the AR process covariance matrix.
- In Pattern Recognition, AR parameters are random. To derive the optimal Bayesian Classifier, the Probability Density Function (PDF) of the AR parameter vector 'a' has to be known, conditioned on each class. According to [4] (p. 22), *the multivariate normal density is an appropriate model for an important situation, viz., the case where the features vectors for a given class are continuous valued, mildly corrupted versions of a single typical or prototype mean vector.* When the features are statistically independent and have the same variance σ^2, the covariance matrix of each class is of the form $\sigma^2 I$ (I being the identity matrix). When the features are not independent, the covariance matrix of each class is no longer diagonal. It can then be expressed as $C_a = \sigma^2 C$. C is a unit norm matrix and σ^2 is the variance of the class (sometimes called within-class scatter) [4] (p. 26).
- In the theory of random coefficient AR models, many methods have been developed using the Gaussian assumption [11].

The paper focused on the pattern recognition case. Consequently, in what follows, the parameter vector of an AR process will be considered as a Gaussian variable denoted by a with mean $m_a = E(a)$ and covariance matrix $C_a = \sigma^2 C$, C being a unit norm matrix independent of σ. Note that all results can be extended to the estimation case with

$$\sigma^2 = \frac{\sigma_e^2}{n} \|R^{-1}\| \quad \text{and} \quad C = \frac{R^{-1}}{\|R^{-1}\|}.$$

In the first part of the paper, an original method for determining the LSP parameter PDF, as a function of the PDF of the AR parameter vector, is proposed.

A statistical study can show that the LSP parameter PDF tends to the Gaussian PDF when $\sigma^2 \to 0$ (pattern recognition) or $n \to +\infty$ (estimation) (the proof is given in [8] for reflection coefficients). However, for non-zero values of σ^2, due to their

non-linear relationship, AR parameters and LSP coefficients cannot both be Gaussian. This strict theoretical result could hide an approximately Gaussian distribution for the LSP parameters. This is very important, since a fine knowledge of the LSP parameter PDF is not necessary for most applications (optimal vector quantization, Bayesian classification, etc.). Thus, determining under which conditions the Gaussian approximation is valid or not, is of major importance. Usual distances, such as the Kullback divergence or the Bhattacharyya distance, are effective tools for measuring the closeness between two PDFs. Thus, they can be used to determine the distance between the LSP parameter PDF (determined in the first part of the paper) and its asymptotic Gaussian PDF. However, due to the very complicated expression of the LSP parameter PDF, these distances do not lead to closed-form expressions. A new distance (denoted as M-distance) has revealed to be an effective tool for measuring the closeness between two random variables [21].

In the second part of the paper, the M-distance is used to determine a simple closed-form expression of the closeness between the LSP parameter PDF and its asymptotic Gaussian PDF.

Non-Gaussian parameters can lead to surprising results when they are used for pattern recognition. For instance, it is well-known that the centroid distance rule has to be used with parameters whose PDF is convex [4]. Recent works have shown that the PDF of reflection and cepstrum coefficients could be non-convex [18,19], which makes them unattractive for pattern recognition (when used with the centroid distance rule).

In the third part of the paper, the shape of LSP parameter clusters is studied: LSP parameters seem to be well suited to classification with the centroid distance rule.

The paper is organized as follows. Section 2 reminds the reader of some LSP parameter definitions and properties. Section 3 presents a recursive way of determining the LSP parameter PDF as a function of the PDF of the AR parameter vector. An original method for determining the 'closeness' between the LSP parameter statistics and its asymptotic Gaussian distribution is studied in Section 4. Section 5 studies some LSP parameter properties for classification. Conclusions are reported in Section 6.

2. LSP coefficient properties

For a given mth-order minimum phase AR polynomial $A_m(z) = 1 + \sum_{i=1}^{m} a_i z^{-i}$, two LSP polynomials denoted by $P_m(z)$ and $Q_m(z)$ can be constructed by setting the $(m+1)$th reflection coefficient (PARCOR coefficient) k_{m+1} to $+1$ or -1:

$$P_m(z) = A_m(z) - z^{-(m+1)} A_m(z^{-1}), \tag{1}$$

$$Q_m(z) = A_m(z) + z^{-(m+1)} A_m(z^{-1}). \tag{2}$$

For speech signals, the conditions $k_{m+1} = +1$ and $k_{m+1} = -1$ correspond to a complete closure and a complete opening of the glottis in the acoustic tube model, respectively. Let

$$P_m(z) = \sum_{i=0}^{m+1} p_i z^{-i}, \tag{3}$$

$$Q_m(z) = \sum_{i=0}^{m+1} q_i z^{-i} \tag{4}$$

with $p_0 = 1, p_{m+1} = -1$ and $q_0 = 1, q_{m+1} = +1$. With these conditions, the coefficients of the LSP polynomials $P_m(z)$ and $Q_m(z)$ are linked to the AR parameters by the following relations:

$$p_i = a_i - a_{m+1-i}, \quad i \in \{1, \ldots, m\}, \tag{5}$$

$$q_i = a_i + a_{m+1-i}, \quad i \in \{1, \ldots, m\}. \tag{6}$$

For m even (the cases of m odd and m even only differ in some details), the LSP polynomials can be expressed as

$$P_m(z) = (1 - z^{-1}) \prod_{i=2,4,\ldots,m} (1 - 2z^{-1}\cos\omega_i + z^{-2}), \tag{7}$$

$$Q_m(z) = (1 + z^{-1}) \prod_{i=1,3,\ldots,m-1} (1 - 2z^{-1}\cos\omega_i + z^{-2}). \tag{8}$$

The parameters $\{\omega_i\}_{i=1,\ldots,m}$ are the LSP parameters. For a stable AR polynomial $A_m(z)$, the LSP polynomials have very interesting properties for quantization and coding [15,16]:

- all roots of $P_m(z)$ and $Q_m(z)$ lie on the unit circle.
- all roots of $P_m(z)$ and $Q_m(z)$ alternate with each other on the unit circle: the LSP parameters $\{\omega_i\}$

satisfy the 'ordering property'

$$0 < \omega_1 < \omega_2 < \cdots < \omega_{m-1} < \omega_m < \pi.$$

In what follows, $\omega = T(a)$ denotes the LSP parameter vector and T the non-linear relation between AR parameters and LSP coefficients.

3. LSP coefficient PDF

This part presents a recursive method for determining the PDF of the LSP coefficient vector as a function of the PDF of the AR parameter vector $a = [a_1, \ldots, a_m]^T$. Define two polynomials $H_{m-2}(z) = \sum_{i=0}^{m-1} h_i z^{-i}$ and $K_{m-2}(z) = \sum_{i=0}^{m-1} k_i z^{-i}$ such that

$$P_m(z) = H_{m-2}(z)(1 - 2z^{-1}\cos\omega_m + z^{-2}), \tag{9}$$

$$Q_m(z) = K_{m-2}(z)(1 - 2z^{-1}\cos\omega_{m-1} + z^{-2}). \tag{10}$$

According to Eqs. (7) and (8), $H_{m-2}(z)$ and $K_{m-2}(z)$ satisfy the following equations:

$$H_{m-2}(z) = (1 - z^{-1}) \prod_{i=2,4,\ldots,m-2} (1 - 2z^{-1}\cos\omega_i + z^{-2}), \tag{11}$$

$$K_{m-2}(z) = (1 + z^{-1}) \prod_{i=1,3,\ldots,m-3} (1 - 2z^{-1}\cos\omega_i + z^{-2}). \tag{12}$$

These two equations show that $H_{m-2}(z)$ and $K_{m-2}(z)$ are LSP polynomials of order $m-2$. The AR parameters corresponding to these two LSP polynomials will be denoted $b_1, \ldots, b_{m-2}$. Eqs. (9) and (10) show that the coefficients of $H_{m-2}(z)$ and $K_{m-2}(z)$, for $i = 1, \ldots, m/2$, satisfy

$$p_i = h_i - 2h_{i-1}\cos\omega_m + h_{i-2}, \tag{13}$$

$$q_i = k_i - 2k_{i-1}\cos\omega_{m-1} + k_{i-2} \tag{14}$$

with $p_i = a_i - a_{m+1-i}$, $q_i = a_i + a_{m+1-i}$, $i \in \{1, \ldots, m\}$, and

$$h_0 = k_0 = 1, \quad h_{-1} = k_{-1} = 0, \tag{15}$$

$$h_j = b_j - b_{m-1-j}, \; k_j = b_j + b_{m-1-j}, \quad j \in \{1, \ldots, m-2\}. \tag{16}$$

These relations allow the determination of the PDF of the vector $V_{m-2} = [b_1, \ldots, b_{m-2}, \omega_{m-1}, \omega_m]^T$ as a function of the PDF of the AR parameter vector $a = [a_1, \ldots, a_m]^T = V_m$ (see Appendix A). In a similar way, the PDF of the vector $V_{m-4} = [c_1, \ldots, c_{m-4}, \omega_{m-3}, \omega_{m-2}, \omega_{m-1}, \omega_m]^T$ can be determined as a function of the PDF of the vector V_{m-2} ($c_1, \ldots, c_{m-4}$ are the AR parameters of a $(m-4)$th order LSP polynomial). With $m/2$ iterations (assuming the AR parameter vector PDF is known), the LSP parameter vector PDF can be computed. As an example, the second- and fourth-order LSP parameter vector PDFs are derived in Appendix B.

For simplicity, simulations are performed for a second-order Gaussian AR parameter vector, with two conjugate poles $p_1 = \rho e^{j\varphi}$ and $p_2 = \rho e^{-j\varphi}$. The mean and covariance matrix of this random vector are

$$m_a = \begin{bmatrix} -2\rho\cos\varphi \\ \rho^2 \end{bmatrix}, \tag{17}$$

$$C_a = \sigma^2 C. \tag{18}$$

C is a unit norm matrix and σ^2 characterizes the intra-class scattering. Simulations are performed with $\rho = 0.8$, $\varphi = \pi/4$ and

$$C_a = 10^{-2}\begin{bmatrix} 1 & 0.9 \\ 0.9 & 1 \end{bmatrix}.$$

Fig. 1 shows a comparison between the theoretical and estimated LSP parameter PDF with 95% confidence intervals. The theoretical LSP parameter PDF is determined using Eq. (B.9) (see Appendix B). The computation of 95% confidence intervals is made following [13] (p. 251). The LSP parameter histograms are computed from 1000 Gaussian AR parameter vectors. Fig. 1 shows that the LSP parameter PDFs and histograms are in good agreement.

4. Distance between the LSP parameter PDF and its asymptotic Gaussian PDF

In what follows, the variables $\{X_{\sigma^2}\}$ converge in probability to zero when σ^2 tends to zero, written

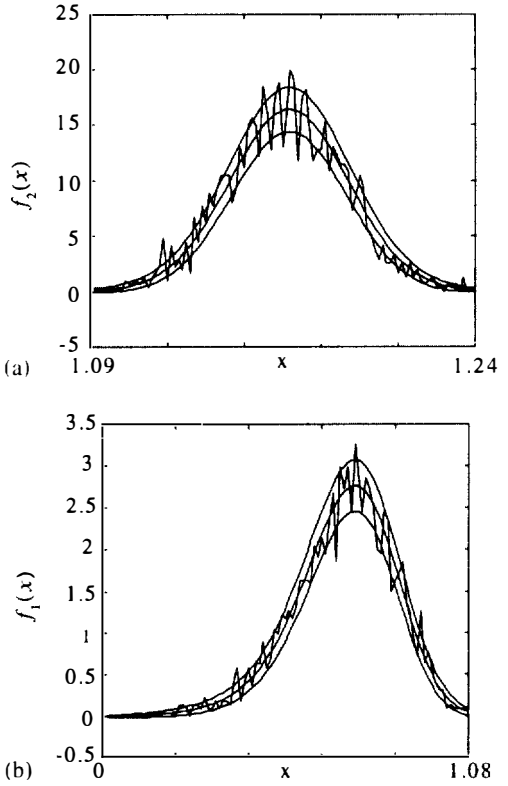

Fig. 1. Theoretical and estimated LSP parameter PDFs with 95% confidence intervals. (a) Coefficient ω_2. (b) Coefficient ω_1.

$X_{\sigma^2} = o_p(1)$ [2] (p. 198), if

$$\forall \varepsilon > 0 \quad \lim_{\sigma^2 \to 0} P[|X_{\sigma^2}| > \varepsilon] = 0. \tag{19}$$

The variables $\{X_{\sigma^2}\}$ are bounded in probability, written $X_{\sigma^2} = O_p(1)$, if

$$\forall \varepsilon > 0 \quad \exists \delta(\varepsilon) > 0 \quad \text{such that}$$

$$P[|X_{\sigma^2}| > \delta(\varepsilon)] < \varepsilon \text{ for all } \sigma^2. \tag{20}$$

In the pattern recognition case, the covariance matrix of the AR process a is of the form $C_a = \sigma^2 C$, C being a unit norm matrix, such that

$$a - m_a = O_p(\sigma) \Leftrightarrow \frac{1}{\sigma}(a - m_a) = O_p(1), \tag{21}$$

which means that $(1/\sigma)(a - m_a)$ is bounded in probability. Denote $T'(a)$ and $T''(a)$ the first- and second-order derivatives of $T(a)$. The computation of the second-order derivative $T''(a)$ can be made from the Hessian $H(a)$ of the transformation T (see [3] (p. 143), or Appendix C for an example). For instance, for a transformation from $\mathbb{R}^m$ into $\mathbb{R}$, $T''(a)$ is defined by

$$T''(a)(u,v) = u^{\mathrm{T}} H(a) v, \tag{22}$$

with $[H(a)]_{ij} = \partial^2 T/\partial a_i \partial a_j$. According to [2] (p. 202), $\omega = T(a)$ admits the following Taylor expansions:

$$\omega = T(a) = T(m_a) + T'(m_a)(a - m_a) + o_p(\sigma), \tag{23}$$

$$\omega = T(a) = T(m_a) + T'(m_a)(a - m_a) + \tfrac{1}{2} T''(m_a)(a - m_a, a - m_a) + o_p(\sigma^2), \tag{24}$$

which means that the random vectors

$$\frac{1}{\sigma}[\omega - T(m_a) - T'(m_a)(a - m_a)], \tag{25}$$

$$\frac{1}{\sigma^2}\left[\omega - T(m_a) - T'(m_a)(a - m_a) - \frac{1}{2} T''(m_a)(a - m_a, a - m_a)\right] \tag{26}$$

converge in probability to zero. Denote

$$G(\sigma) = T'(m_a)(a - m_a), \tag{27}$$

$$\mathrm{NG}(\sigma) = \tfrac{1}{2} T''(m_a)(a - m_a, a - m_a). \tag{28}$$

$G(\sigma)$ is a zero-mean Gaussian vector with covariance matrix $\sigma^2 DCD^{\mathrm{T}}$, where D is a matrix defined in [2] (p. 211). It is interesting to note that $(1/\sigma)G(\sigma)$ is a zero-mean Gaussian vector with covariance matrix DCD^{T} whose PDF is independent of σ. $\mathrm{NG}(\sigma)$ is quadratic in a and thus is non-Gaussian. In what follows, to make shorter, we will write G and NG, instead of $G(\sigma)$ and $\mathrm{NG}(\sigma)$. Eq. (23) shows that the vector $(1/\sigma)[\omega - T(m_a)]$ converges in Probability to the Gaussian vector $(1/\sigma)G(\sigma)$ which is $\mathcal{N}(0, DCD^{\mathrm{T}})$. This section addresses the problem of determining the distance between the distribution of $(1/\sigma)[\omega - T(m_a)]$ and its asymptotic Gaussian distribution $\mathcal{N}(0, DCD^{\mathrm{T}})$.

The Kullback divergence [1] has been shown to be an efficient tool to measure the 'closeness' between the PDFs of two random vectors. The Kullback divergence between two random vectors with

PDFs p_1 and p_2 is defined by

$$d_K(p_1, p_2) = \int_{\mathbb{R}^m} [p_2(x) - p_1(x)] \ln \frac{p_2(x)}{p_1(x)} \, dx. \tag{29}$$

This distance can be used to determine the closeness between the LSP parameter vector PDF (computed from the previous section) and the Gaussian PDF $\mathcal{N}(0, DCD^T)$. However, this distance depends on m_a and C_a through non-linear relations which are difficult to study. The next part of the paper presents a new way of measuring the closeness between the distribution of $(1/\sigma)[\omega - T(m_a)]$ and its asymptotic Gaussian distribution $\mathcal{N}(0, DCD^T)$.

For small values of σ, the development (24) reduces to $\omega \simeq T(m_a) + G$: the distribution of the LSP parameter vector $(1/\sigma)[\omega - T(m_a)]$ is close to the distribution of $(1/\sigma)G$, which is $\mathcal{N}(0, DCD^T)$. On the other hand, when the second-order term NG is not negligible, the distribution of the LSP parameter vector $(1/\sigma)[\omega - T(m_a)] \simeq (1/\sigma)(G + \mathrm{NG})$ is close to the distribution of $(1/\sigma)(G + \mathrm{NG})$, which is not $\mathcal{N}(0, DCD^T)$. The validity of the approximations $\omega - T(m_a) \simeq G$ and $\omega - T(m_a) \simeq G + \mathrm{NG}$ depends on the closeness between the random vectors G and $G + \mathrm{NG}$. The next part of the paper defines a new distance between random variables, denoted as M-distance. This distance is then used to give a measure of the closeness between G and $G + \mathrm{NG}$.

For simplicity, the study is restricted to random variables: for a mth order AR model, instead of studying the transformation

$$\omega = T(a) = [T_1(a), \ldots, T_m(a)]^T \tag{30}$$

from $\mathbb{R}^m$ into $\mathbb{R}^m$, the m transformations $\omega_i = T_i(a)$, $i \in \{1, \ldots, m\}$, from $\mathbb{R}^m$ into $\mathbb{R}$ are considered separately. Let

$$G = [G_1, \ldots, G_m]^T \tag{31}$$

and

$$\mathrm{NG} = [\mathrm{NG}_1, \ldots, \mathrm{NG}_m]^T. \tag{32}$$

Define the set S of variables X satisfying the two following conditions:

$$\forall k \in \mathbb{N} \quad M_X^k = E(X^k) < +\infty, \tag{33}$$

$$\frac{1}{R} = \limsup_{k \to +\infty} \left(\frac{|M_X^k|}{k!} \right)^{1/k} < +\infty. \tag{34}$$

Variables satisfying conditions (33) and (34) are characterized by their moments [14] (p. 290).

The M-distance between variables X and Y belonging to the set S is defined by [21]

$$d_M(X, Y) = \sum_{k=1}^{+\infty} \frac{|M_X^k - M_Y^k|}{k!}, \tag{35}$$

where M_X^k and M_Y^k are the kth-order moments of X and Y, respectively. This distance comes from the l_1 norm applied to infinite sequences of the form

$$1, \frac{iM_X^1}{1!}, \ldots, \frac{(i)^n M_X^n}{n!}, \ldots, \tag{36}$$

which appears naturally in the development of the characteristic function in terms of its moments. In general, d_M is not a distance because two different random variables can have the same moments [9] (p. 12). However, variables satisfying conditions (33) and (34) are characterized by their moments such that d_M is a distance on S.

Since a is Gaussian, variables G_i and $G_i + \mathrm{NG}_i$ belong to the set S (and are characterized by their moments). The M-distance can then be used to measure the closeness between G_i and $G_i + \mathrm{NG}_i$. It can be shown that

$$\lim_{\sigma^2 \to 0} d_M\left[\frac{1}{\sigma} G_i, \frac{1}{\sigma}(G_i + \mathrm{NG}_i)\right] = 0, \quad \forall i \in \{1, \ldots, m\}. \tag{37}$$

This means that the M-distance between $(1/\sigma)(G_i + \mathrm{NG}_i)$ and the Gaussian variable $(1/\sigma)G_i$ tends to zero when $\sigma^2 \to 0$. For small values of σ^2, the distribution of $(1/\sigma)[\omega_i - T_i(m_a)] \simeq (1/\sigma)(G_i + \mathrm{NG}_i)$ is very close to the distribution of the Gaussian variable $(1/\sigma)G_i$.

Note that this result can be easily extended to the estimation case (for which $C_a = (\sigma_e^2/n)R^{-1}$) by

$$\lim_{n \to +\infty} d_M[nG_i, n(G_i + \mathrm{NG}_i)] = 0, \quad \forall i \in \{1, \ldots, m\}. \tag{38}$$

This case illustrates the asymptotic normality of any parameter obtained from a 'regular' one-to-one transformation of the estimated AR parameter vector [8].

The next part of the paper studies, for a fixed value of σ^2, the M-distance between variables $G_i + \mathrm{NG}_i$ (whose distribution approximates the LSP parameter distribution) and G_i as a function of the position of the AR model poles in the unit circle. Since a is Gaussian, $d_M(G_i, G_i + \mathrm{NG}_i)$ can be expressed as

$$d_M(G_i, G_i + \mathrm{NG}_i) = \sum_{k=1}^{M} d_i(k) + o(\sigma^{2M}), \tag{39}$$

with $d_i(k) = O(\sigma^{2k})$. For small values of σ^2, $d_M(G_i, G_i + \mathrm{NG}_i)$ can be approximated by the first few terms $d_i(k)$ in Eq. (39) such that

$$d_M(G_i, G_i + \mathrm{NG}_i) \simeq \sum_{k=1}^{M} d_i(k). \tag{40}$$

The lower $d_M(G_i, G_i + \mathrm{NG}_i)$, the lower the distance between the distribution of the ith LSP parameter ω_i and the distribution of the Gaussian variable G_i. For simplicity, the previous second-order Gaussian AR parameter vector, with two poles $p_1 = \rho e^{j\varphi}$ and $p_2 = \rho e^{-j\varphi}$, is considered. According to Appendix C, the following results are obtained:

$$d(1) = \begin{bmatrix} d_1(1) \\ d_2(1) \end{bmatrix} = \frac{\sigma^2}{2} \begin{bmatrix} \dfrac{(c_{11} + c_{22} + 2c_{12})|x_1|}{[4 - x_1^2]^{3/2}} \\ \dfrac{(c_{11} + c_{22} - 2c_{12})|x_2|}{[4 - x_2^2]^{3/2}} \end{bmatrix}, \tag{41}$$

with

$$C_a = \sigma^2 \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix},$$

$$x_1 = \rho^2 - 2\rho\cos\varphi - 1 \quad \text{and} \quad x_2 = 1 - 2\rho\cos\varphi - \rho^2.$$

Higher order terms $d(k)$, with $k > 1$, could be derived in a similar way. However, Figs. 2 and 3 show that the variations of $d(1)$ (continuous line) and $d(1) + d(2)$ (dashed line), for the two LSP coefficients, as a function of the modulus ρ and the phase φ, are very similar: for $k > 1$, the $d(k) = O(\sigma^{2k})$ are negligible with respect to $d(1) = O(\sigma^2)$. In that case, the difference between the higher-order moments of $G_i + \mathrm{NG}_i$ and G_i is small when compared to the difference between the means. Consequently, when $E(G_i + \mathrm{NG}_i) \simeq E(G_i)$, the second-order term NG_i can be neglected in the Taylor expansion (24): the ith LSP parameter distribution is close to the distribution of G_i (which is Gaussian).

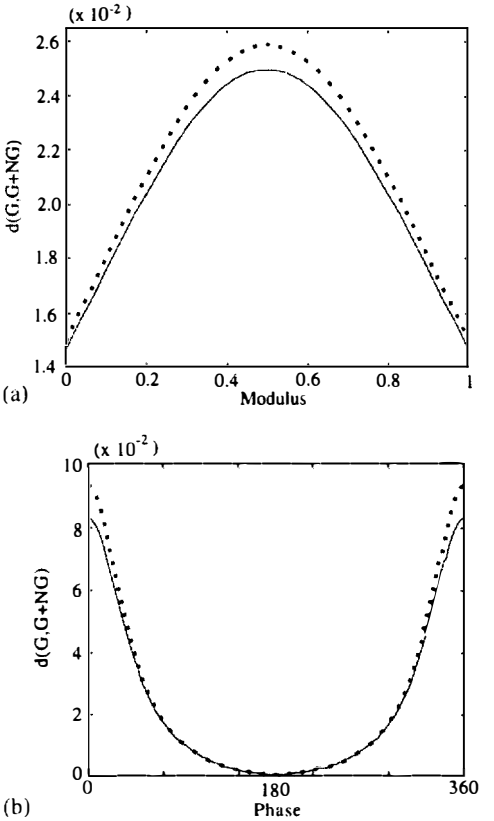

Fig. 2. Variations of $d_1(1)$ (continuous line) and $d_1(1) + d_1(2)$ (dashed line) for the first LSP coefficient (a) as a function of ρ ($\varphi = \pi/3$), (b) as a function of φ ($\rho = 0.4$).

The results obtained with the M-distance are then compared with the Kullback divergence between the LSP parameter PDF and the Gaussian PDF $\mathcal{N}(0, DCD^{\mathrm{T}})$. Figs. 2 and 4 (respectively Figs. 3 and 5) show that the qualitative behavior of the two approaches is very similar. However, the main advantage of the M-distance is a very simple closed-form expression (given in Eq. (41)) for the closeness between the LSP coefficient PDF and its asymptotic Gaussian PDF $\mathcal{N}(0, DCD^{\mathrm{T}})$. The convergence of the LSP coefficient PDF to its asymptotic Gaussian distribution depends on the AR model pole position inside the unit circle. The closed-form expression of d_M describes precisely the convergence dependence on the AR model poles. The smaller d_M, the smaller the distance between the LSP coefficient PDF and the Gaussian PDF

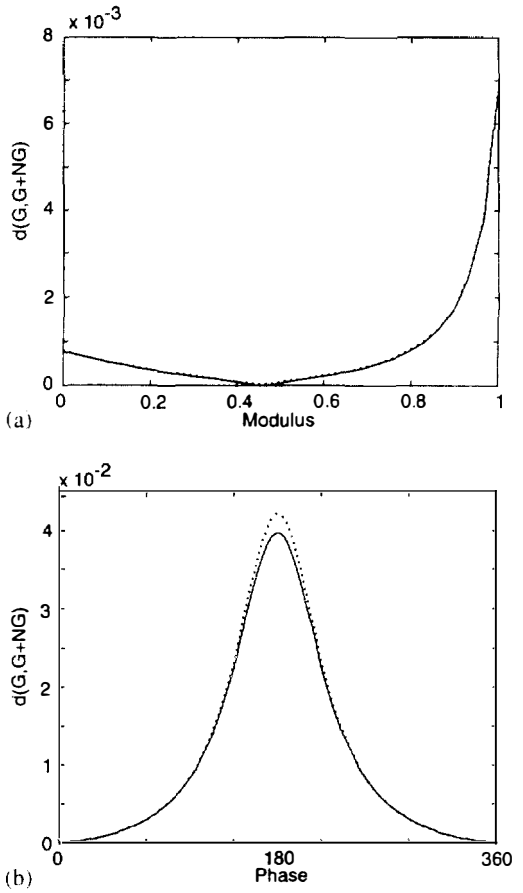

Fig. 3. Variations of $d_2(1)$ (continuous line) and $d_2(1) + d_2(2)$ (dashed line) for the second LSP coefficient (a) as a function of $\rho(\varphi = \pi/6)$, (b) as a function of φ ($\rho = 0.4$).

$\mathcal{N}(0, DCD^{\mathrm{T}})$. The smaller d_M, the faster the convergence of the LSP coefficient PDF to $\mathcal{N}(0, DCD^{\mathrm{T}})$. The variations of the M-distance, as a function of the modulus and the phase of the AR model (for a fixed variance $\sigma^2 = 0.01$ and for the two LSP coefficients), are plotted in Fig. 6. Fig. 6 shows that the M-distance is always very small when the AR model poles are close to the origin.

As a conclusion, the convergence of the LSP coefficient PDF to its asymptotic Gaussian distribution depends upon the AR model pole position inside the unit circle: for a fixed variance σ^2, the faster convergence occurs when the AR model poles are close to the origin.

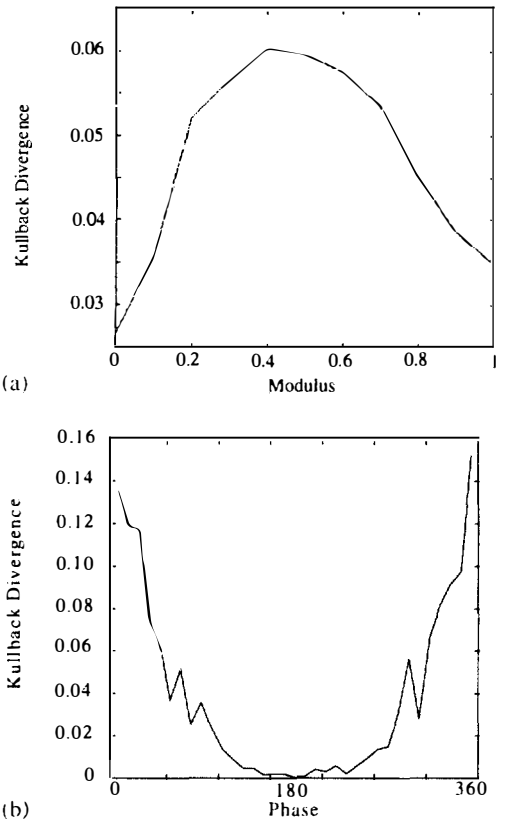

Fig. 4. Variations of the Kullback divergence between the first LSP coefficient PDF and the Gaussian PDF (a) as a function of $\rho(\varphi = \pi/3)$, (b) as a function of φ ($\rho = 0.4$).

An analysis was given here for a second-order AR model but it could be extended to higher-order models. For instance, consider a fourth-order AR model with poles $p_1 = \rho e^{j\varphi}$, $p_2 = \rho e^{-j\varphi}$, $p_3 = \delta e^{j\psi}$ and $p_4 = \delta e^{-j\psi}$. Fig. 7(a–d) shows the M-distance for the four LSP coefficients as a function of ρ and φ, for a fixed $(\delta, \psi) = (0.2, \pi/3)$, $\sigma^2 = 0.01$ and $C = I$ (identity matrix): the convergence of the LSP coefficient PDF to its asymptotic Gaussian distribution is fast when the two poles p_1 and p_2 are close to the origin.

5. Application to pattern recognition

A major problem in pattern recognition is the determination of the 'optimal' classification rule for

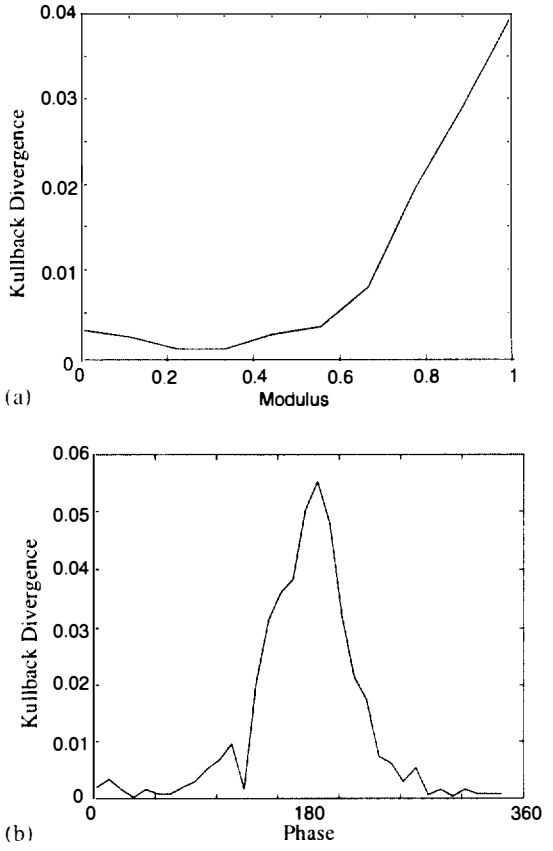

Fig. 5. Variations of the Kullback divergence between the second LSP coefficient PDF and the Gaussian PDF (a) as a function of ρ ($\varphi = \pi/3$), (b) as a function of φ ($\rho = 0.4$).

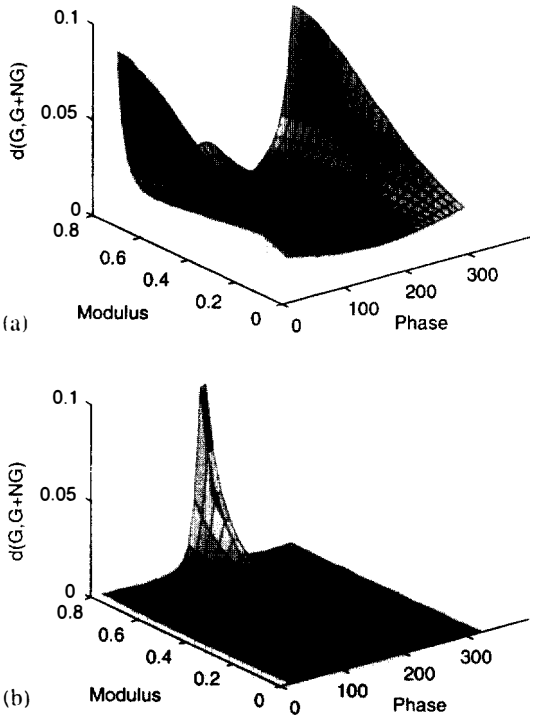

Fig. 6. Variations of the M-distance between the LSP coefficient PDF and the Gaussian PDF as a function of ρ and φ: (a) $d(G_1, G_1 + \mathrm{NG}_1)$ (first coefficient), (b) $d(G_2, G_2 + \mathrm{NG}_2)$ (second coefficient).

a given parameter vector. A solution to this problem is given by the Bayes classifier when the parameter vector statistics are known. However, implementation of the Bayes classifier is often difficult because of its complexity. The Bayes decision rule reduces to a centroid distance rule for Gaussian parameter vectors, leading to a simple classifier. For non-Gaussian parameter vectors with non-convex PDFs, surprising results can be obtained with the centroid distance rule. For instance, in Fig. 8 (non-convex classes), all points belonging to the first class highlighted with a star will be mis-classified with the centroid distance rule. A statistical analysis of the reflection and cepstrum coefficients has shown that their PDF can be non-convex. Thus, the centroid distance rule is not an effective classifier for these coefficients [18,19].

Unfortunately, it is not easy to theoretically prove the LSP point cluster convexity (or non-convexity) for any order. However, in supervised learning pattern recognition applications, parameter estimates m_a and C_a can be performed for each class. The LSP parameter distribution can then be determined (as a function of m_a and C_a), conditioned on each class. Thus, LSP point cluster convexity can be studied *qualitatively* by plotting the level lines for $f_0(\omega_1, \ldots, \omega_m)$ derived in Appendix A. For instance, Fig. 9(a, b) shows the 3D PDF and the level lines of a specific example: the second-order LSP parameter vector studied in the first section. The LSP parameter vector PDF is convex, yielding a case for which the centroid distance rule can be used.

In any case, simulations are not sufficient to prove the convexity property. However, the simulations provide (on every specific example)

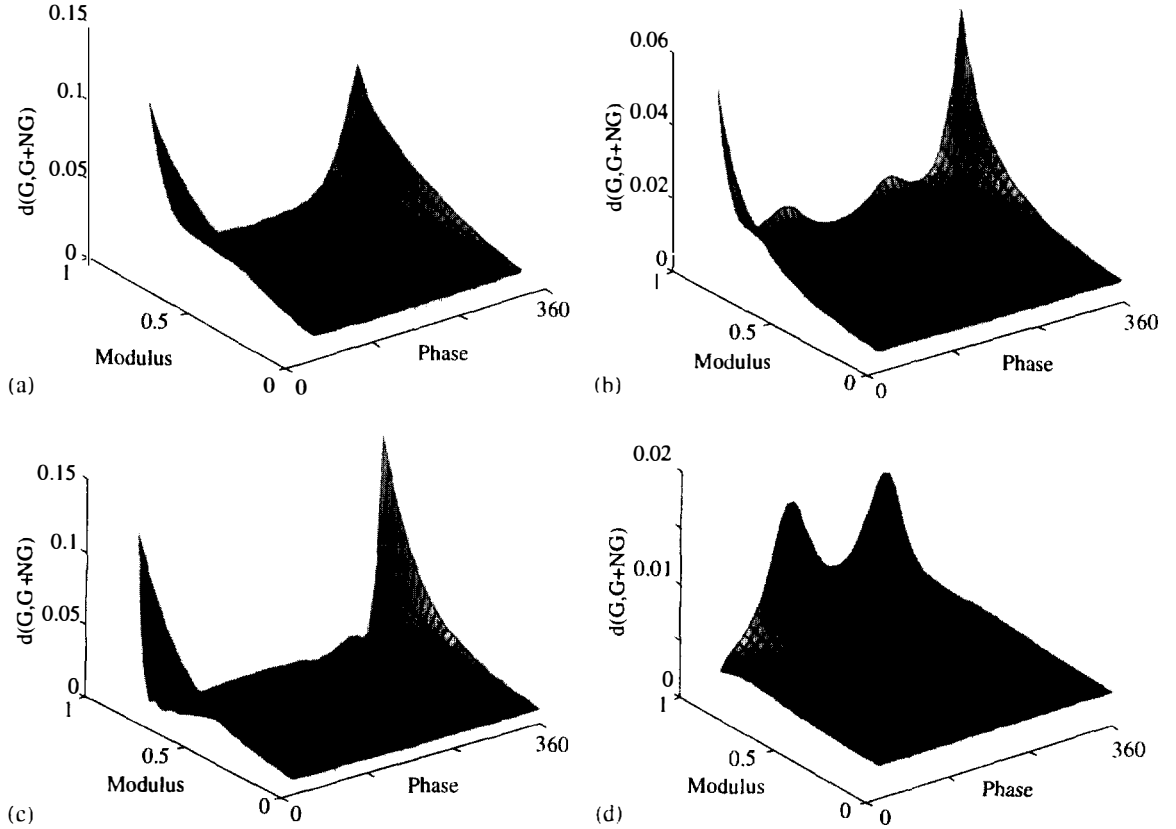

Fig. 7. Variations of the M-distance between the LSP coefficient PDF and the Gaussian PDF as a function of ρ and φ: (a) $d(G_1,G_1 + NG_1)$ (first coefficient), (b) $d(G_2,G_2 + NG_2)$ (second coefficient), (c) $d(G_3,G_3 + NG_3)$ (third coefficient), (d) $d(G_4,G_4 + NG_4)$ (second coefficient).

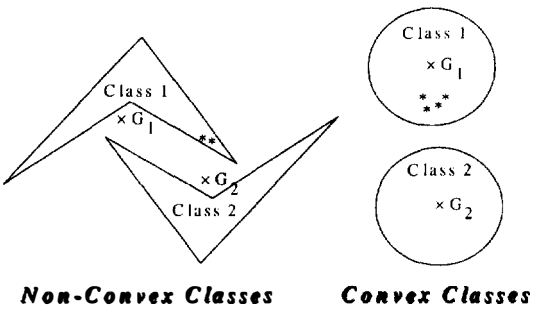

Fig. 8. Classification model.

qualitative information about the LSP point cluster convexity (sufficient in most pattern recognition applications). No counter example has been exhibited to date which shows the non-convexity of LSP point clusters. Consequently, LSP coefficients seem to be well suited to classification, contrary to reflection and cepstrum coefficients [17,19].

6. Conclusion

A recursive method, to compute the LSP coefficient PDF as a function of the AR parameter PDF, was derived.

Since they are non-linearly related, LSP coefficients and AR parameters cannot both be Gaussian. The M-distance was used to measure the closeness between the LSP coefficient distribution and its asymptotic Gaussian PDF. Other distances (such as the Kullback divergence) or other methods (such as expansions in Edgeworth or Gram Charlier series) could be used to measure these deviations from normality. However, the main advantage of

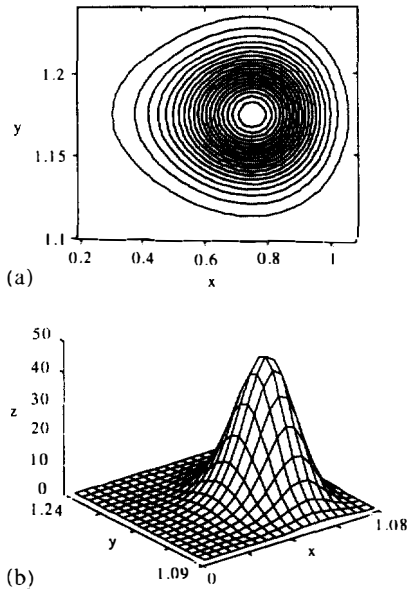

Fig. 9. (a) Level lines and (b) 3D plot of a 2nd-order LSP vector PDF.

the M-distance was a very simple closed-form expression of these deviations as a function of the AR parameter mean and covariance matrix. This closed-form expression showed that the distance between the LSP coefficient distribution and its asymptotic Gaussian distribution depends on the position of the AR model poles in the unit circle: the distance is very small when the AR model poles are close to the origin.

In a supervised learning pattern recognition application, the AR parameter PDF can be estimated conditioned on each class. The LSP coefficient PDF can then be determined conditioned on each class. The LSP coefficient PDF was used, on a specific example, to study qualitatively the LSP point cluster convexity. This kind of study can be performed in every pattern recognition application. No counter-example which shows the non-convexity of LSP coefficient point clusters has been exhibited to date. Thus, LSP seem to be well suited to pattern recognition with the centroid distance rule.

Appendix A. Recursive determination of the LSP coefficient PDF

This appendix describes the different steps necessary to determine the LSP coefficient PDF as a function of the PDF of the AR parameter vector. The concatenation of the following equations:

$$p_i = h_i - 2h_{i-1}\cos\omega_m + h_{i-2}, \tag{A.1}$$

$$q_i = k_i - 2k_{i-1}\cos\omega_{m-1} + k_{i-2}, \tag{A.2}$$

for $i = 1, \ldots, m$, with $p_i = a_i - a_{m+1-i}$, $q_i = a_i + a_{m+1-i}$ and

$$h_0 = k_0 = 1, \quad h_{-1} = k_{-1} = 0, \tag{A.3}$$

$$h_j = b_j - b_{m-1-j},\ k_j = b_j + b_{m-1-j}, \quad j \in \{1, \ldots, m-2\}, \tag{A.4}$$

leads to

$$\begin{aligned}
a_1 &= b_1 - \cos\omega_{m-1} - \cos\omega_m,\\
a_m &= b_{m-2} - \cos\omega_{m-1} + \cos\omega_m,\\
a_2 &= b_2 - (b_1 + b_{m-2})\cos\omega_{m-1} - (b_1 - b_{m-2})\cos\omega_m + 1,\\
a_{m-1} &= b_{m-3} - (b_1 + b_{m-2})\cos\omega_{m-1} + (b_1 - b_{m-2})\cos\omega_m,\\
&\vdots\\
a_i &= b_i - (b_{i-1} + b_{m-i})\cos\omega_{m-1} - (b_{i-1} - b_{m-i})\cos\omega_m + b_{i-2},\\
a_{m-i+1} &= b_{m-i-1} - (b_{i-1} + b_{m-i})\cos\omega_{m-1} + (b_{i-1} - b_{m-i})\cos\omega_m + b_{m-i+1},\\
&\vdots
\end{aligned} \tag{A.5}$$

The Jacobian matrix corresponding to the transformation from $V_m = a = (a_1, \ldots, a_m)^T$ to $V_{m-2} = [b_1, \ldots, b_{m-2}, \omega_{m-1}, \omega_m]^T$, denoted as $J(V_m, V_{m-2})$, can then be computed. If C_i denotes the ith column of this matrix, the following results are obtained:

$$C_1 = \begin{bmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ y_{m-1} \\ y_m \end{bmatrix}, \quad C_2 = \begin{bmatrix} -x_{m-1} - x_m \\ 1 \\ 0 \\ \cdot \\ 0 \\ -x_{m-1} + x_m \\ y_{m-1}(b_1 + b_{m-2}) \\ y_m(-b_1 + b_{m-2}) \end{bmatrix}, \quad C_{m-1} = \begin{bmatrix} -x_{m-1} + x_m \\ 0 \\ \cdot \\ 0 \\ 1 \\ -x_{m-1} - x_m \\ y_{m-1}(b_1 + b_{m-2}) \\ y_m(-b_1 + b_{m-2}) \end{bmatrix}, \quad C_m = \begin{bmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ 1 \\ y_{m-1} \\ -y_m \end{bmatrix}, \tag{A.6}$$

and, for $i \in \{1, \ldots, m/2\}$:

$$C_i = \begin{bmatrix} \cdot \\ C_i(i-2) \\ C_i(i-1) \\ C_i(i) \\ \cdot \\ C_i(m-i) \\ \cdot \\ C_i(m-1) \\ C_i(m) \end{bmatrix} = \begin{bmatrix} \cdot \\ 1 \\ -x_{m-1} - x_m \\ 1 \\ \cdot \\ -x_{m-1} + x_m \\ \cdot \\ y_{m-1}(b_{i-1} + b_{m-i}) \\ y_m(b_{i-1} - b_{m-i}) \end{bmatrix}, \tag{A.7}$$

$$C_{m-i+1} = \begin{bmatrix} \cdot \\ C_{m-i+1}(i-1) \\ \cdot \\ C_{m-i+1}(m-i-1) \\ C_{m-i+1}(m-i) \\ C_i(m-i+1) \\ \cdot \\ C_i(m-1) \\ C_i(m) \end{bmatrix} = \begin{bmatrix} \cdot \\ -x_{m-1} + x_m \\ \cdot \\ 1 \\ -x_{m-1} - x_m \\ 1 \\ \cdot \\ y_{m-1}(b_{i-1} + b_{m-i}) \\ -y_m(b_{i-1} - b_{m-i}) \end{bmatrix}. \tag{A.8}$$

The terms which are not specified in C_i and C_{m-i+1} are equal to zero and $x_{m-1}, x_m, y_{m-1}, y_m$ are defined by

$$x_{m-1} = \cos\omega_{m-1}, \quad x_m = \cos\omega_m, \quad y_{m-1} = \sin\omega_{m-1}, \quad y_m = \sin\omega_m. \tag{A.9}$$

Denote by $h_i^m(v_{m-2})$ the closed-form expression of a_i as a function of the parameter vector $v_{m-2} = (b_1, \ldots, b_{m-2}, \omega_{m-1}, \omega_m)^T$ defined in Eq. (A.5). Denote $f_m(v_m)$ and $f_{m-2}(v_{m-2})$ the PDFs of vectors $V_m = (a_1, \ldots, a_m)^T$ and V_{m-2} respectively. The PDF of V_{m-2} satisfies the following equation:

$$f_{m-2}(v_{m-2}) = f_m(h_1^m(v_{m-2}), \ldots, h_m^m(v_{m-2}))|J(V_m, V_{m-2})|. \tag{A.10}$$

Using similar developments, the Jacobian matrix corresponding to the transformation from V_i to V_{i-2} is shown to have the following form:

$$J(V_i,V_{i-2}) = \begin{bmatrix} M_i & 0 \\ 0 & I_{m-i-1} \end{bmatrix}, \quad i \in \{m,\ldots,2\}. \tag{A.11}$$

In Eq. (A.11), I_{m-i-1} is the $(m-i)\times(m-i)$ identity matrix and M_i is a $i \times i$ matrix similar to $J(V_m,V_{m-2})$. The PDF of V_{i-2}, denoted $f_{i-2}(v_{i-2})$, can then be determined as a function of the PDF of V_i, denoted $f_i(v_i)$:

$$f_{i-2}(v_{i-2}) = f_i(h_1^i(v_{i-2}),\ldots,h_i^i(v_{i-2}),\omega_{i-1},\ldots,\omega_m)|J(V_i,V_{i-2})|. \tag{A.12}$$

Appendix B. LSP coefficient PDFs for a second- and fourth-order Gaussian AR parameter vector

This appendix derives the second- and fourth-order LSP coefficient PDFs for a Gaussian AR parameter vector.

B.1. Order 2

The two LSP coefficients are linked to AR parameters by the following relations:

$$\cos\omega_2 = \tfrac{1}{2}(a_2 - a_1 - 1), \tag{B.1}$$

$$\cos\omega_1 = \tfrac{1}{2}(1 - a_1 - a_2), \tag{B.2}$$

which lead to

$$a_1 = -\cos\omega_1 - \cos\omega_2 = g_1(\omega), \tag{B.3}$$

$$a_2 = 1 - \cos\omega_1 + \cos\omega_2 = g_2(\omega). \tag{B.4}$$

The Jacobian matrix corresponding to the transformation from $a = (a_1,a_2)^{\mathrm{T}}$ to $\omega = (\omega_1,\omega_2)^{\mathrm{T}}$ is

$$J = \begin{bmatrix} \sin\omega_1 & \sin\omega_2 \\ \sin\omega_1 & -\sin\omega_2 \end{bmatrix}, \tag{B.5}$$

with determinant

$$\det(J) = -2\sin\omega_1\sin\omega_2. \tag{B.6}$$

The PDF of a Gaussian vector a with mean $m_a = (m_1, m_2)^{\mathrm{T}}$ and covariance matrix C_a is

$$f(x) = \frac{1}{2\pi\sqrt{\det C_a}}\exp Q(x), \quad x \in \mathbb{R}^2, \tag{B.7}$$

where $Q(x)$ is the quadratic form

$$Q(x) = -\tfrac{1}{2}(x - m_a)^{\mathrm{T}}C_a^{-1}(x - m_a). \tag{B.8}$$

The PDF of the LSP coefficient vector $\omega = (\omega_1,\omega_2)^{\mathrm{T}}$ can then be determined

$$f(\omega) = \frac{\sin\omega_1\sin\omega_2}{\pi\sqrt{\det\Sigma}}\exp Q(g(\omega))I_A(\omega), \quad \omega \in \mathbb{R}^2, \tag{B.9}$$

with $g(\omega) = [g_1(\omega),g_2(\omega)]^{\mathrm{T}}$ and

$I_A(\omega) = 1$ if $0 \leqslant \omega_1 \leqslant \omega_2 \leqslant \pi$,

$I_A(\omega) = 0$ else.

B.2. Order 4

The fourth-order LSP polynomials are defined by

$$P_4(z) = 1 + \sum_{i=1}^{4} (a_i - a_{5-i})z^{-i} - z^{-5}, \qquad Q_4(z) = 1 + \sum_{i=1}^{4} (a_i + a_{5-i})z^{-i} + z^{-5}. \tag{B.10}$$

B.2.1. First iteration

The LSP polynomials can be expressed as

$$\begin{aligned} P_4(z) &= \left[1 + \sum_{i=1}^{2} (b_i - b_{3-i})z^{-i} - z^{-3}\right][1 - 2z^{-1}\cos\omega_4 + z^{-2}], \\ Q_4(z) &= \left[1 + \sum_{i=1}^{2} (b_i + b_{3-i})z^{-i} + z^{-3}\right][1 - 2z^{-1}\cos\omega_3 + z^{-2}]. \end{aligned} \tag{B.11}$$

Thus

$$\begin{aligned} a_1 - a_4 &= -2\cos\omega_4 + b_1 - b_2, \\ a_1 + a_4 &= -2\cos\omega_3 + b_1 + b_2, \\ a_2 - a_3 &= 1 - 2\cos\omega_4(b_1 - b_2) + (b_2 - b_1), \\ a_2 + a_3 &= 1 - 2\cos\omega_3(b_1 + b_2) + (b_2 + b_1). \end{aligned} \tag{B.12}$$

The closed-form expressions of a_i as a function of $v_2 = (b_1, b_2, \omega_3, \omega_4)^{\mathrm{T}}$ are

$$\begin{aligned} a_1 &= b_1 - \cos\omega_3 - \cos\omega_4 = h_1^4(v_2), \\ a_2 &= 1 + b_2 - \cos\omega_3(b_1 + b_2) - \cos\omega_4(b_1 - b_2) = h_2^4(v_2), \\ a_3 &= b_1 - \cos\omega_3(b_1 + b_2) + \cos\omega_4(b_1 - b_2) = h_3^4(v_2), \\ a_4 &= b_2 - \cos\omega_3 + \cos\omega_4 = h_4^4(v_2). \end{aligned} \tag{B.13}$$

The Jacobian matrix corresponding to the transformation from $v_4 = (a_1, \ldots, a_4)^{\mathrm{T}}$ to $v_2 = (b_1, b_2, \omega_3, \omega_4)^{\mathrm{T}}$ is

$$J(V_4, V_2) = \begin{bmatrix} 1 & -\cos\omega_3 - \cos\omega_4 & 1 - \cos\omega_3 + \cos\omega_4 & 0 \\ 0 & 1 - \cos\omega_3 + \cos\omega_4 & -\cos\omega_3 - \cos\omega_4 & 1 \\ \sin\omega_3 & (b_1 + b_2)\sin\omega_3 & (b_1 + b_2)\sin\omega_3 & \sin\omega_3 \\ \sin\omega_4 & (b_1 - b_2)\sin\omega_4 & (b_2 - b_1)\sin\omega_4 & -\sin\omega_4 \end{bmatrix}, \tag{B.14}$$

with determinant

$$\det J(V_4, V_2) = 2\sin\omega_3 \sin\omega_4 (1 + \cos\omega_4 + b_1 - b_2)(1 - 2\cos\omega_3 - b_1 - b_2). \tag{B.15}$$

The PDF of V_2 is then given by

$$f_2(v_2) = f_4(h_1^4(v_2), h_2^4(v_2), h_3^4(v_2), h_4^4(v_2))|\det J(V_4, V_2)|, \tag{B.16}$$

$f_4(a)$ being the PDF of a Gaussian vector a with mean $m_a = (m_1, m_2)^{\mathrm{T}}$ and covariance matrix C_a.

B.2.2. Second iteration

The parameters b_1 and b_2 are linked to the two first LSP coefficients ω_1 and ω_2 by the following relations:

$$\begin{cases} b_1 - b_2 = -1 - 2\cos\omega_2 \\ b_1 + b_2 = 1 - 2\cos\omega_1 \end{cases} \Leftrightarrow \begin{cases} b_1 = -\cos\omega_1 - \cos\omega_2 = h_1^2(\omega), \\ b_2 = 1 - \cos\omega_1 + \cos\omega_2 = h_2^2(\omega). \end{cases} \tag{B.17}$$

The Jacobian matrix corresponding to the transformation from $v_2 = (b_1, b_2, \omega_3, \omega_4)^{\mathrm{T}}$ to $v_0 = \omega = (\omega_1, \omega_2, \omega_3, \omega_4)^{\mathrm{T}}$ is

$$J(V_2, V_0) = \begin{bmatrix} \sin\omega_1 & \sin\omega_1 & 0 & 0 \\ -\sin\omega_2 & \sin\omega_2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{B.18}$$

with determinant

$$\det J(V_2, V_0) = 2\sin\omega_1 \sin\omega_2. \tag{B.19}$$

As a conclusion, the PDF of the fourth-order LSP parameter vector is given by

$$f_0(\omega_1, \omega_2, \omega_3, \omega_4) = f_2(h_1^2(\omega), h_2^2(\omega), \omega_3, \omega_4)|2\sin\omega_1 \sin\omega_2| I_A(\omega), \quad \omega \in \mathbb{R}^2, \tag{B.20}$$

with

$$\begin{aligned} I_A(\omega) &= 1 \quad \text{if } 0 \leqslant \omega_1 \leqslant \omega_2 \leqslant \omega_3 \leqslant \omega_4 \leqslant \pi, \\ I_A(\omega) &= 0 \quad \text{else.} \end{aligned} \tag{B.22}$$

Appendix C. Computation of d_k for LSP coefficients

In this appendix, we determine the first moments of variables G and $G + \mathrm{NG}$ for LSP coefficients in the case of a second-order AR model. Similar results can be obtained for higher-order AR models. AR parameters and LSP coefficients are linked by the following relations:

$$\begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} = \begin{bmatrix} T_1(a) \\ T_2(a) \end{bmatrix} \begin{bmatrix} \arccos\left(\dfrac{1 - a_1 - a_2}{2}\right) \\ \arccos\left(\dfrac{-1 - a_1 + a_2}{2}\right) \end{bmatrix}. \tag{C.1}$$

The first- and second-order derivatives of T_1 and T_2 can then be computed:

$$T_1'(a) = \frac{1}{\sqrt{4 - (1 - a_1 - a_2)^2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \tag{C.2}$$

$$T_2'(a) = \frac{1}{\sqrt{4 - (1 + a_1 - a_2)^2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \tag{C.3}$$

$$T_1''(a) = \frac{-1 + a_1 + a_2}{[4 - (1 - a_1 - a_2)^2]^{3/2}} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \tag{C.4}$$

$$T_2''(a) = \frac{1 + a_1 - a_2}{[4 - (1 - a_1 - a_2)^2]^{3/2}} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}. \tag{C.5}$$

Let $x_1 = -1 + m_1 + m_2$ and $x_2 = 1 + m_1 - m_2$. Using Eqs. (27) and (28), we obtain

$$G_1 = T'_1(m_a)(a - m_a) = \frac{1}{\sqrt{4 - x_1^2}}(a_1 + a_2 - m_1 - m_2), \tag{C.6}$$

$$G_2 = T'_2(m_a)(a - m_a) = \frac{1}{\sqrt{4 - x_2^2}}(a_1 - a_2 - m_1 + m_2), \tag{C.7}$$

$$\mathrm{NG}_1 = \tfrac{1}{2}T''_1(m_a)(a - m_a, a - m_a) = \tfrac{1}{2}(a - m_a)^{\mathrm{T}}_1 T''_1(m_a)(a - m_a) \tag{C.8}$$

$$= \frac{x_1}{2[4 - x_1^2]^{3/2}}(a_1 + a_2 - m_1 - m_2)^2, \tag{C.9}$$

$$NG_2 = \tfrac{1}{2}T''_2(m_a)(a - m_a, a - m_a) = \frac{1}{2}(a - m_a)^{\mathrm{T}} T''_2(m_a)(a - m_a) \tag{C.10}$$

$$= \frac{x_2}{2[4 - x_2^2]^{3/2}}(a_1 - a_2 - m_1 + m_2)^2. \tag{C.11}$$

The moments of G_i and NG_i can be determined as a function of the AR parameter covariance matrix

$$C_a = \sigma^2 \begin{bmatrix} c_{11} & c_{12} \\ c_{12} & c_{22} \end{bmatrix}.$$

For instance

$$E(\mathrm{NG}_1) = \frac{\sigma^2 x_1}{2[4 - x_1^2]^{3/2}}(c_{11} + c_{22} + 2c_{12}), \tag{C.12}$$

$$E(\mathrm{NG}_2) = \frac{\sigma^2 x_2}{2[4 - x_2^2]^{3/2}}(c_{11} + c_{22} - 2c_{12}). \tag{C.13}$$

Hence

$$d(1) = \begin{bmatrix} d_1(1) \\ d_2(1) \end{bmatrix} = \frac{\sigma^2}{2}\begin{bmatrix} \dfrac{|x_1|(c_{11} + c_{22} + 2c_{12})}{[4 - x_1^2]^{3/2}} \\ \dfrac{|x_2|(c_{11} + c_{22} - 2c_{12})}{[4 - x_1^2]^{3/2}} \end{bmatrix}. \tag{C.14}$$

It is well known that the higher-order moments of the Gaussian distribution can be determined as a function of its mean and covariance matrix. In particular, for $i \in \{1,2\}$ and $j \in \mathbb{N}$:

$$E[(a_i - m_i)^{2j+1}] = 0, \qquad E[(a_i - m_i)^{2j-2}] = \frac{(2j+1)!}{2^j j!}\sigma^{2j+2} \tag{C.15}$$

This allows us to compute the higher-order moments of G_i and NG_i as well as the $d(k)$ for $k > 1$.

References

[1] M. Basseville, Distance measures for signal processing and pattern recognition, Signal Processing 18 (December 1989) 349–369.

[2] P.J. Brockwell, R.A. Davis, Time Series: Theory and Methods, 2nd ed., Springer, Berlin, 1990.

[3] P.G. Ciarlet, Introduction à l'analyse numérique matricielle et à l'optimisation, Masson, 1982.

[4] R.O. Duda, Pattern Classification and Scene Analysis, Wiley, New York, 1973.

[5] J.S. Erkelens, P.M.T. Broersen, On the statistical properties of line spectrum pairs, Proc. ICASSP'95, Detroit, 8–12 May 1995, pp. 768–772.

[6] W.R. Gardner, B.D. Rao, Optimal distortion measures for the high rate vector quantization of LPC parameters, Proc. ICASSP'95, Detroit, 8–12 May 1995, pp. 752–755.

[7] A.H. Gray, Jr., J.D. Markel, Quantization and bit allocation in speech processing, IEEE Trans. Acoust. Speech Signal Process. ASSP-24 (December 1976) 459–473.

[8] S. Kay, J. Makhoul, On the satistics of the estimated reflection coefficients of an autoregressive process, IEEE Trans. Acoust. Speech Signal Process. 31 (December 1983) 1447–1455.

[9] E. Lukacs, Characteristic Functions, Griffin, London, 1970.

[10] J. Makhoul, Linear prediction: a tutorial review, Proc. IEEE 63 (April 1975) 124–143.

[11] D.F. Nicholls, B.G. Quinn, Random coefficient autoregressive Models: an Introduction, Lecture Notes in Statistics, Vol. 11, Springer, New York, 1982.

[12] K.K. Paliwal, B.S. Atal, Efficient vector quantization of LPC parameters at 24 bits/frame, IEEE Trans. Speech Audio Process. 1 (January 1993) 3–14.

[13] A. Papoulis, Probability, Random Variables and Stochastic Processes, 3rd ed., McGraw-Hill, New York, 1991.

[14] A. Rényi, Calcul des Probabilités, Dunod, France, 1996.

[15] S. Saoudi, J.M. Boucher, A. Le Guyader, A new efficient algorithm to compute the LSP parameters for speech coding, Signal Processing 28 (2) (August 1992) 201–212.

[16] N. Sugamura, N. Farvardin, Quantizer design in LSP speech analysis and synthesis, IEEE Journal on Selected Areas in Communications 6 (2) (February 1988).

[17] J.Y. Tourneret, M. Ghogho, Line spectrum pairs in pattern recognition, Proc. EUSIPCO'96, Italy, 10–13 September 1996.

[18] J.Y. Tourneret, B. Lacaze, Study of the couple (Cepstrum Coefficient, k-NN Rule), Proc. 7th IEEE Workshop on Stat. and Array Proc. (SSAP'94), Quebec, 26–29 June 1994, pp. 357–360.

[19] J.-Y. Tourneret, B. Lacaze, On the statistics of estimated reflection and cepstrum coefficients of an autoregressive process, Signal Processing 43 (3) (May 1995) 253–267.

[20] J.Y. Tourneret, B. Lacaze, Statistical study of reflection coefficients derived from a random AR process, Proc. ICASSP'95, Detroit, 8–12 May 1995, pp. 1709–1712.

[21] J.Y. Tourneret, B. Lacaze, Convergence rate of the cepstrum coefficient distribution to the Gaussian, Proc. IEEE Work. HOS, Girona, Spain, 12–14 June 1995, pp. 200–203.