What Should We Say About the Kurtosis?

A. Mansour, *Member, IEEE,* and C. Jutten, *Member, IEEE*

***Abstract*—In this work, we point out some important properties of the normalized fourth-order cumulant (i.e., the kurtosis). In addition, we emphasize the relation between the signal distribution and the sign of the kurtosis. One should mention that in many situations, authors claim that the sign of the kurtosis depends on the nature of the signal (i.e., over- or sub-Gaussian). For unimodal probability density function, that claim is true and is clearly proved in the letter. But for more complex distributions, it has been shown that the kurtosis sign may change with parameters and does not depend only on the asymptotic behavior of the distributions.**

Finally, these results give theoretical explanation to techniques, like nonpermanent adaptation, used in nonstationary situations.

***Index Terms*—Blind identification, blind separation, high order statistics, kurtosis, over- and sub-Gaussian, probability density function.**

I. Introduction

In various works [1]–[6] concerning the problem of blind separation of sources, authors propose algorithms whose efficacy demands conditions on the source kurtosis, and sometimes that all the sources have the same sign of kurtosis. In fact, this assumption seems very strong and in this work we studied the relationship between the signal distribution and the sign of its kurtosis.

II. Definition and Properties

Definition 1: Let us denote $p(x)$ the probability density function (pdf) of a random process $x(t)$ and $E()$ the average. By definition [7], [8] the kurtosis $K[p(x)]$ is as in (1), shown at the bottom of the next page. Clearly, the kurtosis sign $ks(x)$ is equal to the fourth-order cumulant sign. Some properties can be easily derived.

1) $Cum_4(ax+b) = a^4 Cum_4(x)$, so $ks(x)$ is invariant by any linear transformation $ks(ax+b) = ks(x)$.
2) Let $p(x) = p_e(x) + p_o(x)$, where $p_e(x)$ is even and $p_o(x)$ is odd. It is easy to prove that $ks(x)$ only depends on $p_e(x)$ and that $p_e(x)$ can be considered as a pdf.

Therefore, in the following, the study may be restricted to a zero-mean process $x(t)$ whose the pdf $p(x)$ is even and has a variance $\sigma_x^2 = 1$.

It is well known that the kurtosis of a Gaussian distribution is equal to zero. Intuitively, the sign of the kurtosis seems related to the comparison between $p(x)$ and Gaussian distribution, by considering the asymptotic properties of the distributions and the following definition:

Definition 2: A pdf $p(x)$ is said over-Gaussian (respectively sub-Gaussian) [9]–[11], if

$\exists\, x_0 \in \mathbb{R}^+ \,|\, \forall\, x \geq x_0,\ p(x) > g(x)$ (respectively, $p(x) < g(x)$), where $g(x)$ is the normalized Gaussian pdf. In many examples, it seems that $ks(x)$ is positive[1] for over-Gaussian signals and negative for sub-Gaussian signals.

III. Theoretical Results

Let us consider an (even) pdf $p(x)$ and a zero-mean normalized Gaussian pdf $g(x)$.

Theorem 1: If $p(x) = g(x)$ have only two solutions, then

$$Ks(x) > 0 \iff p(x) \text{ is over-Gaussian}$$
$$Ks(x) < 0 \iff p(x) \text{ is sub-Gaussian.}$$

A demonstration is given in the Appendix. This theorem shows that the intuitive claim, given in the previous section, is true under the specific condition of Theorem 1.

When $p(x) = g(x)$ has more than two solutions, then there is no longer simple rule to predict $ks(x)$. More precisely, over-Gaussian as well as sub-Gaussian pdf's can lead to positive as well as negative sign of kurtosis. As an example, let us consider the following over-Gaussian pdf:

$$p(x) = \frac{b}{4}(\exp(-b|x-a|) + \exp(-b|x+a|)). \qquad (2)$$

It is easy to compute the kurtosis of (2):

$$K[p(x)] = 2\frac{6-(ab)^4}{4+4a^2b^2+a^4b^4}.$$

Clearly, $ks(x)$ is not always negative, but may change according to the values of the parameters a and b: $K[p(x)] \geq 0$ if $0 < ab \leq \sqrt[4]{6}$ and $K[p(x)] < 0$ if $ab > \sqrt[4]{6}$.

Finally, we may consider that artificial signals (for instance telecommunication signals) are bounded, and consequently their pdf are sub-Gaussian. It is often claimed that the kurtosis of such signals is negative. The following example shows that this claim is wrong. Let us consider the distribution of Fig. 1. It is easy to evaluate the kurtosis of this signal:

$$K(p(x)) = \frac{\alpha}{5\sigma_x^4}(b^5-a^5) + \frac{\beta}{5\sigma_x^4}(d^5-c^5) - \frac{\alpha^2}{3\sigma_x^4}(b^3-a^3)^2 - \frac{\beta^2}{3\sigma_x^4}(d^3-c^3)^2 - 2\frac{\alpha\beta}{3\sigma_x^4}(b^3-a^3)(d^3-c^3) \qquad (3)$$

with the normalization condition $\alpha(b-a) + \beta(d-c) = 1$. For scale reasons, we do not draw directly $K(p(x))$ but $K^\star(p(x)) = (1/2)[K(p(x)) + |K(p(x))|]$. Thus, if $K(p(x)) >$

[1] When the signal has a positive kurtosis sign, it is said that it is a positive kurtotic signal.

Manuscript received March 1, 1999. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. H. Messer-Yaron.

A. Mansour is with the BMC Research Center (RIKEN), Nagoya 463, Japan (e-mail: mansour@nagoya.riken.go.jp).

C. Jutten is with INPG-LIS, 38031 Grenoble, France.

Publisher Item Identifier S 1070-9908(99)09290-1.

1070–9908/99$10.00 © 1999 IEEE

Authorized licensed use limited to: The University of Toronto. Downloaded on April 11, 2009 at 11:24 from IEEE Xplore. Restrictions apply.

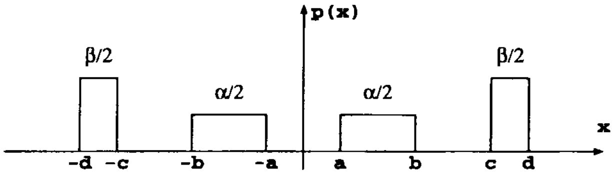

Fig. 1. Example of x-limited pdf.

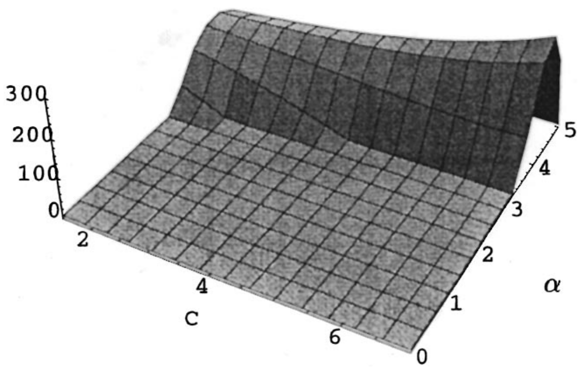

Fig. 2. Representation of $K^\star(p(x))$ according to parameters c and α, with $a = 0.9$, $b = 1.1$, and $d = 9$.

0, $K^\star(p(x)) = K(p(x))$, otherwise, if $K(p(x)) \leq 0$, $K^\star(p(x)) = 0$. Now, we remark that the sign of the kurtosis may be easily controlled with adequate values of the pdf parameters (see Fig. 2).

IV. Conclusion

In this letter, we point out some relations between pdf and kurtosis sign. First, we show that the kurtosis sign is not modified by any scale or translation factors, and it only depends on the even part of the pdf.

Usually, one associates the kurtosis sign of a distribution $p(x)$ to its over-Gaussian or sub-Gaussian nature. Here, we prove that this claim is only relevant for unimodal pdf $p(x)$. Generally, even for bounded pdf, we show by a few examples that the kurtosis sign can change according to the pdf parameters.

From a practical point of view, kurtosis sign of non stationary signals, which must be estimated on short moving windows, can change. A previous experimental study proves that the kurtosis sign of speech signal can be affected by the silent period [12]. Additionally, this work gives a theoretical explanation to the necessity and the efficacy of intermittent adaptation which is used for separation of nonstationary sources [13].

Appendix
Proof of Theorem 1

Let us consider that for $x > 0$, the equation $p(x) = g(x)$ only has one solution $\rho > 0$. It is known that the fourth-order cumulant of a Gaussian distribution is zero. As a consequence, we can write: $\int_{\mathbb{R}} x^4 g(x)\,dx = 3\int_{\mathbb{R}} x^2 g(x)\,dx = 3$. In addition, we just may study the sign of $\Upsilon = (1/2)K[p(x)]$. Using (1) and the unit variance signal, one can prove that

$$\Upsilon = \int_0^\rho x^4(p(x)-g(x))\,dx + \int_\rho^\infty x^4(p(x)-g(x))\,dx. \quad (4)$$

Let us consider that the pdf $p(x)$ is an over-Gaussian signal ($p(x) > g(x)$, when $x \to \infty$). Then, the sign of $p(x) - g(x)$ remains constant on each interval $[0,\ \rho]$, and $[\rho, \infty]$. Using the second mean value theorem, Υ can be rewritten as

$$\Upsilon = \lambda^4 \int_\rho^\infty (p(x)-g(x))\,dx - \xi^4 \int_0^\rho (g(x)-p(x))\,dx \quad (5)$$

where $0 < \xi < \rho < \lambda$. Using the fact that $p(x)$ and $g(x)$ are both pdf, we can deduce that $\int_0^\infty (p(x) - g(x))\,dx = \int_0^\rho (p(x) - g(x))\,dx + \int_\rho^\infty (p(x) - g(x))\,dx = 0$. Taking into account that $p(x)$ is over-Gaussian, we deduce

$$\int_\rho^\infty (p(x)-g(x))\,dx = \int_0^\rho (g(x)-p(x))\,dx > 0. \quad (6)$$

Using (5) and (6), we remark that

$$\Upsilon = (\lambda^4 - \xi^4)\int_\rho^\infty (p(x)-g(x))\,dx > 0. \quad (7)$$

Finally, if $p(x)$ is an over-Gaussian pdf (with the assumption of a unique solution to the equation $p(x) = g(x)$ for $x \in \mathbb{R}^+$, then its kurtosis is positive. Using the same reasoning and under the same condition, we can claim that a sub-Gaussian pdf has a negative kurtosis.

References

[1] B. Laheld and J. F. Cardoso, "Adaptive source separation without prewhitening," in *Signal Processing VII, Theories and Applications,* M. J. J. Holt, C. F. N. Cowen, P. M. Grant, and W. A. Sandham, Eds. Edinburgh, Scotland: Elsevier, Sept. 1994, pp. 183–186.

[2] N. Delfosse and P. Loubaton, "Adaptive blind separation of independent sources: A deflation approach," *Signal Process.,* vol. 45, pp. 59–83, July 1995.

[3] A. Mansour and C. Jutten, "Fourth order criteria for blind separation os sources," *IEEE Trans. Signal Processing,* vol. 43, pp. 2022–2025, Aug. 1995.

[4] A. Cichoki *et al.,* "Self-adaptive independent component analysis for sub-Gaussian and super-Gaussian mixtures with an unknown number of sources and additive noise," in *Proc. 1997 Int. Symp. Nonlinear Theory and its Applications,* Hawaii, pp. 731–734.

[5] O. Macchi and E. Moreau, "Self-adaptive source separation, Part I: Convergence analysis of a direct linear neural network controlled by the Hérault-Jutten algorithm," *IEEE Trans. Signal Processing,* vol. 45, pp. 918–925, Apr. 1997.

[6] A. Cichoki, I. Sabala, and S. I. Amari, "Intelligent neural networks for blind signal separation with unknown number of sources," in *Proc. Conf. Engineering of Intelligent Systems,* 1998.

[7] A. Benvéniste, M. Métivier, and P. Priouret, *Adaptive Algorithms and Stochastic Approximations.* Berlin, Germany: Springer-Verlag, 1990.

[8] P. McCullagh, *Tensor Methods in Statistics.* London, U.K.: Chapman & Hall, 1987.

[9] V. Zivojnovic, "Higher-order statistics and huber's robustness," in *Proc. HOS'95,* Girona, Spain, pp. 236–240.

[10] P. Comon, "Quelques developments recents en traitement du signal," Tech. Rep., Univ. Nice, Sophia-Antipolis, France, Sept. 18, 1995.

[11] J. L. Lacoume, P. O. Amblard, and P. Comon, *Statistiques d'Ordre Supérieur pour le Traitement du Signal.* Paris, France: Mason, 1997.

[12] A. Mansour, C. Jutten, and N. Ohnishi, "Definition and properties," in *Proc. Int. Conf. Multisource-Multisensor Information Fusion,* Las Vegas, NV, July 6–9, 1998, pp. 40–46.

[13] L. Nguyen Thi and C. Jutten, "Blind sources separation for convolutive mixtures," *Signal Process.,* vol. 45, pp. 209–229, 1995.

$$K[p(x)] = \frac{Cum_4(x)}{E(x^2)^2} = \frac{E(x^4) - 3E(x^2)^2 + 12E(x)^2E(x^2) - 4E(x)E(x^3) - 6E(x)^4}{E(x^2)^2} \quad (1)$$

Authorized licensed use limited to: The University of Toronto. Downloaded on April 11, 2009 at 11:24 from IEEE Xplore. Restrictions apply.