Kurtosis: Definition and Properties.

Ali MANSOUR
BMC Research Center (RIKEN),
2271-130, Anagahora, Shimoshidami, Moriyama-ku, Nagoya 463 (JAPAN)
email: mansour@nagoya.riken.go.jp
http://www.bmc.riken.go.jp/sensor/Mansour/mansour.html

Christian JUTTEN
INPG - LIS
46, avenue Félix Viallet, 38031 Grenoble Cedex
email: chris@tirf.inpg.fr

Noboru OHNISHI
BMC Research Center (RIKEN),
2271-130, Anagahora, Shimoshidami, Moriyama-ku, Nagoya 463 (JAPAN)
email: ohnishi@sheraton.ohnishi.nuie.nagoya-u.ac.jp

Abstract *In various studies on blind separation of sources, one assumes that sources have the same sign of kurtosis. In fact, this assumption seems very strong and in this paper we studied relation between signal distribution and the sign of the kurtosis. A theoretical result has been found in a simple case. However, for more complex distributions, the kurtosis sign cannot be predicted and may change with parameters. The results give theoretical explanation to tricks, like non-permanent adaptation, used in non stationary situations.*

Keywords: kurtosis, high order statistics, blind identification and separation, probability density function.

1 Introduction

In various works [8, 5, 4, 3, 9, 2] 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 paper we studied relation between signal distribution and the sign of its kurtosis.

2 Definition and Properties

Let us denote by $x(t)$ a zero-mean real signal and by $p(x)$ its probability density function (pdf). By definition, the kurtosis $K[p(x)]$ is the normalized fourth-order cumulant of the signal [1, 7]:

$$\begin{aligned} K[p(x)] &= \frac{Cum_4(x)}{E(x^2)^2} \\ &= \frac{E(x^4) - 3E(x^2)^2}{E(x^2)^2}, \end{aligned} \tag{1}$$

where $E()$ is the average. If the signal is not a zero-mean signal, the equation (1) becomes [7]:

$$\begin{aligned} K[p(x)] &= \frac{Cum_4(x)}{E(x^2)^2} \\ &= \frac{E(x^4) - 3E(x^2)^2}{E(x^2)^2} + \frac{12E(x)^2E(x^2)}{E(x^2)^2} \\ &\quad - \frac{4E(x)E(x^3) + 6E(x)^4}{E(x^2)^2}. \end{aligned} \tag{2}$$

Clearly, the kurtosis has the same sign than the fourth-order cumulant, then we will only study the sign of the fourth-order cumulant.

Let us denote the kurtosis sign $ks(x)$. Some properties can be easily derived :

1. The kurtosis sign $ks(x)$ is invariant by any linear transformation. From (2), we deduce:

$$Cum_4(ax+b) = a^4 Cum_4(x), \quad (3)$$

   then $ks(ax+b) = ks(x)$. So in the following, we consider zero-mean signal with a variance $\sigma_x^2 = 1$ (where σ_x is the standard deviation of $x(t)$).

2. If we express the pdf $p(x)$ as a sum of two functions: $p(x) = p_e(x) + p_o(x)$, where $p_e(x)$ is even and $p_o(x)$ is odd, then:

   - $ks(x)$ only depends on the even function $p_e(x)$,
     because the fourth-order cumulant (1) depends only on the fourth and second-order moments (so it depends just on the even moments).
   - The even function $p_e(x)$ has the properties of a pdf:

$$p_e(x) \geq 0, \quad \forall x$$
$$\text{and} \quad \int_{\mathbb{R}} p(x)dx = \int_{\mathbb{R}} p_e(x)dx = 1.$$

Therefore, in the following, the study may be restricted to a zero-mean signal $x(t)$ with a variance $\sigma_x^2 = 1$, whose pdf $p(x)$ is even. As examples, in the table 1, we computed $ks(x)$ for four well known distributions.

Table 1: known distributions.

Signal	$Cum_4(x)$	$ks(x)$	fig.
Uniform	$\frac{a^4+b^4+6a^2b^2+3.5(a^3b+b^3a)}{-30}$	$-$	1.a
Discrete	$\frac{-N(N+1)(2N^2+2N+1)}{15}$	$-$	1.b
Gamma	$\frac{26}{3}$, if $\sigma_x = 1$	$+$	1.c
Cosine	$\frac{192}{\pi^4} - 2 - 2\alpha^4$	$-$	1.d

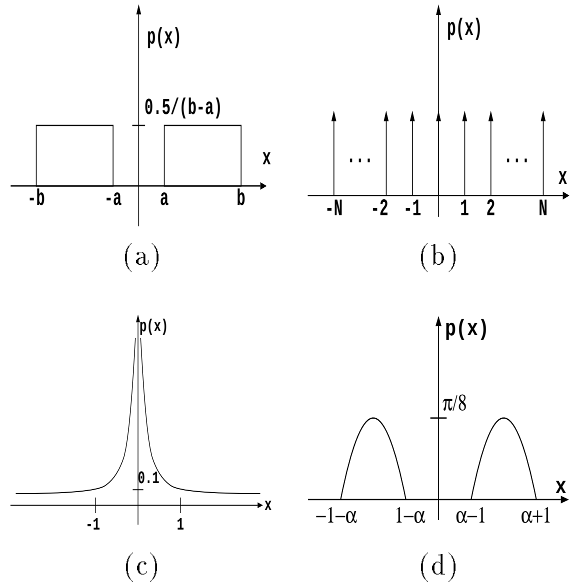


Figure 1: pdf signals

Clearly, $ks(x)$ is strongly related to the comparison between $p(x)$ and the normalized Gaussian distribution, whose kurtosis is equal to zero. Therefore, let us consider a pdf $p(x)$, we say $p(x)$ is an over-Gaussian pdf (respectively sub-Gaussian), if:

$$\exists \quad x_0 \in \mathbb{R}^+ \mid \forall x \geq x_0, \ p(x) > g(x) \quad (4)$$

(respectively $p(x) < g(x)$), where $g(x)$ is the normalized Gaussian pdf. From the previous examples, it seems that $ks(x)$ is positive for over-Gaussian signals and negative for sub-Gaussian signals.

3 Theoretical result

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

Theorem 1 *Assuming that $p(x)$ and $g(x)$ have only two intersections, $ks(x)$ is positive (respectively negative) if $p(x)$ is over-Gaussian (respectively sub-Gaussian).*

The demonstration is given in appendix A.

3.1 General cases

In the general case, if there are more then two intersection points, then there is no rule to predict $ks(x)$. More precisely, over-Gaussian as well as sub-Gaussian signals can lead to positive as well as negative sign of kurtosis. As an example, let us consider the signal $x(t)$ whose pdf is a sum of two exponential functions (sotef):

$$p(x) = \frac{b}{4}(\exp(-b|x-a|)+\exp(-b|x+a|)) \quad (5)$$

Figure 2 (a) shows the general form of $p(x)$ for $x \geq 0$. Figure 2 (b) and figure 2 (c) give examples (parameterized by a and b) where $ks(x) > 0$ and $ks(x) < 0$, respectively. On these figures, we show also the normalized Gaussian pdf $g(x)$: the previous theorem is not applicable, because there are at least 2 intersections points (in $\mathbb{R}^+$) between $p(x)$ and $g(x)$.

Using the equation (5), it is easy to prove that:

$$\begin{aligned} E(x^4) &= a^4 + \frac{12}{b^2}a^2 + \frac{24}{b^4} \quad (6) \\ E(x^2) &= a^2 + \frac{2}{b^2} \quad (7) \end{aligned}$$

From equations (6) and (7), we can derive the kurtosis of this signal:

$$k[p(x)] = 2\frac{6-(ab)^4}{4+4a^2b^2+a^4b^4}, \quad (8)$$

Then by choosing adequate values of the parameters a and b, it is possible to change $ks(x)$:

$$K[p(x)] \geq 0 \text{ if } 0 < ab \leq \sqrt[4]{6}. \quad (9)$$

With respect to the definition (4), $p(x)$ is an even over-Gaussian pdf, and nevertheless $ks(x)$ is not always negative, but may change according to the values of the parameters a and b.

3.2 Case of bounded pdf

In practical cases, we may consider that physical signals are bounded, and consequently their pdf are sub-Gaussian. Let us consider for instance quaternary sources $x(t)$ (see Fig 3), whose the fourth order cumulant is

$$\begin{aligned} Cum_4(x) &= a^4p(1-3p) - 6a^2b^2p(1-p) \\ &\quad -b^4(1-p)(2-3p). \quad (10) \end{aligned}$$

It is clear that the sign of $Cum_4(x)$ may change with the values of the parameters. For example let be $a = 0$, then $Cum_4(x) < 0$ if $p < 2/3$, and *vice versa*.

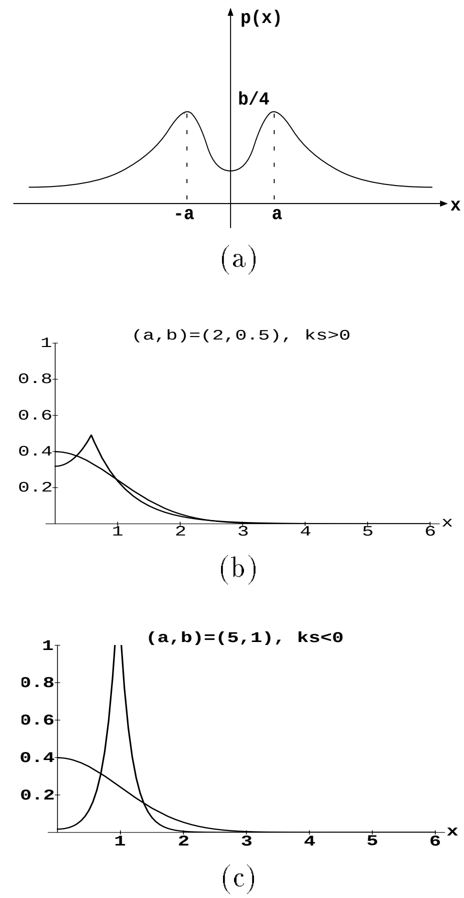


Figure 2: The normalized exponential pdf (sotef) and normalized Gaussian pdf $g(x)$.

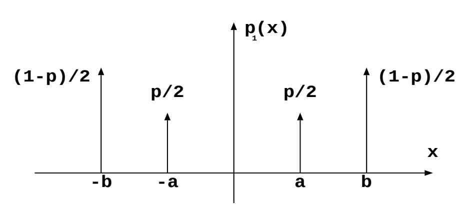
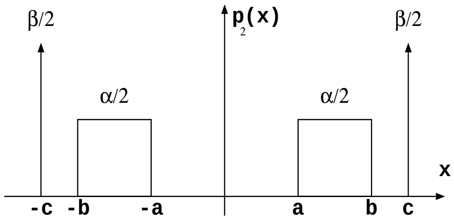
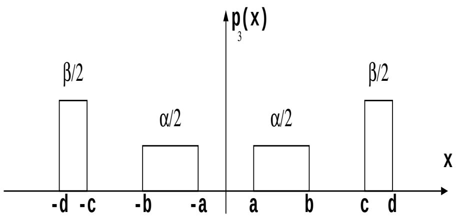


Figure 3: Pdf of quaternary sources.



(a)



(b)

Figure 4: Two exemples of x-limited pdf

Finally, let us consider the exemples in figure 4. It is easy to evaluate the kurtosis of these signals (Fig 4). The kurtosis of first signal (Fig 4 (a)) can be written as:

$$K(p_2(x)) = \frac{\alpha}{5\sigma_x^4}(b^5 - a^5) + \frac{\beta}{\sigma_x^4}c^5 - \frac{\alpha^2}{3\sigma_x^4}(b^3 - a^3)^2 - \frac{3\beta^2}{\sigma_x^4}c^6 - 2\frac{\alpha\beta}{\sigma_x^4}(b^3 - a^3)c^3, \quad (11)$$

with the condition:

$$\alpha(b - a) + \beta = 1. \quad (12)$$

The kurtosis of second signal (figure 4 (b)) is equal to:

$$K(p_3(x)) = \frac{\alpha}{5\sigma_x^4}(b^5 - a^5) + \frac{\beta}{5\sigma_x^4}(d^5 - c^5) - \frac{\alpha^2}{3\sigma_x^4}(b^3 - a^3)^2 - \frac{\beta^2}{3\sigma_x^4}(d^3 - c^3)^2 - 2\frac{\alpha\beta}{3\sigma_x^4}(b^3 - a^3)(d^3 - c^3), \quad (13)$$

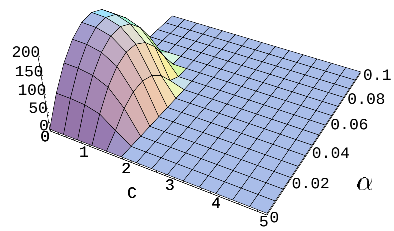
with the condition:

$$\alpha(b - a) + \beta(d - c) = 1. \quad (14)$$

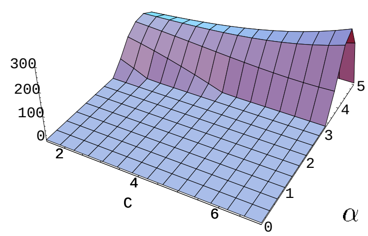
For scale reasons, we do not draw directly $K(p(x))$ but:

$$K^\star(p(x)) = \frac{1}{2}[K(p(x)) + \mid K(p(x)) \mid]. \quad (15)$$

Thus, if $K(p(x)) > 0$, $K^\star(p(x)) = K(p(x))$, otherwise, if $K(p(x)) < 0$, $K^\star(p(x)) = 0$. Then, we remark that the sign of the kurtosis may be easily controlled with adequate values of the pdf parameters (see Fig 5).



(a) $K^\star(p_2(x))$, where $a = 2$ and $b = 9$.



(b) $K^\star(p_3(x))$, where $a = 0.9$, $b = 1.1$ and $d = 9$.

Figure 5: Representation of $K^\star(p(x))$ according to parameters c and α

4 Experimental results

In the case of real signals, the kurtosis estimation will be done on finite slipping windows [6]. For stationary signals, the window may be very long. But for non-stationnary signals (speech signals for exemple, see Fig 6), the length of the window must be short enough.

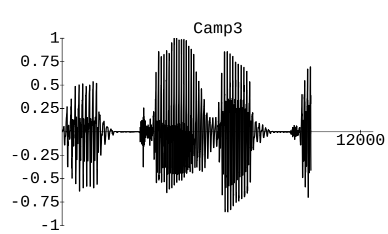


Figure 6: Speech signal: Camp3

Moreover, in case of non stationary signals, the pdf can vary a lot : for instance, silent periods in speech signals imply a peak around $x = 0$ for the pdf (see figure 7).

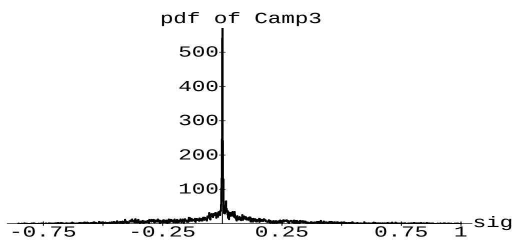


Figure 7: Estimated pdf of speech signal "CAMP3"

According to the size of the window, and its location, we observe changes in the kurtosis sign. Figure 8 shows the kurtosis time evolution of the speech signal of figure 6. The kurtosis is estimated on 500-sample windows, every 50 samples. We remark that the kurtosis is negative during a silent period, and it becomes positive during the speech transient.

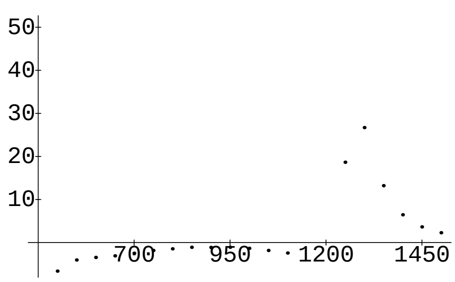
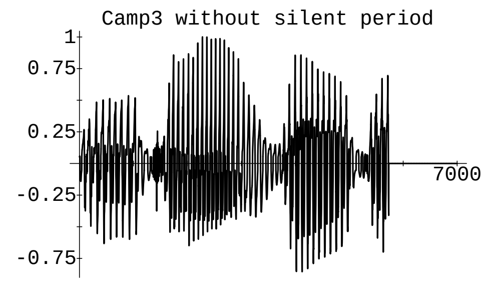
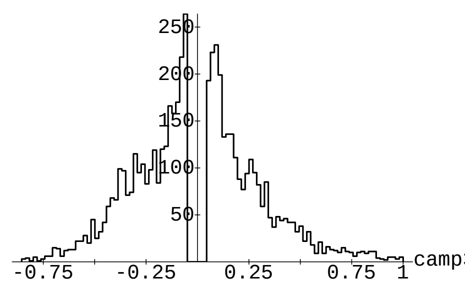


Figure 8: The estimated kurtosis of speech signal "Camp3"

Finally, by an experimental study we remark that: In the case of speech signal, the kurtosis sign fluctuations can be eliminated by estimate the kurtosis on the signal without the silent periods (see Fig 9).
This result can explain that the non-permanent learning (freezing the parameter estimation) in speech separation algorithms [9] enforces source pdf to have a negative kurtosis sign and then allows algorithm convergence.



(a) Speech signal "Camp3" without its silent periods.



(b) Estimated pdf of this signal.

Figure 9: Experimental results

5 Conclusion

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

Basically, people refer to comparison with Gaussian pdf. We prove the comparison is only relevant for unimodal pdf $p(x)$ having only two intersections (in $\mathbb{R}$) with the Gaussian pdf.

In the general case, even for bounded pdf, we show by a few examples that the kurtosis sign can be controlled by the pdf parameters.

From a practical point of view, kurtosis sign of non stationary signals, which must be estimated on short slipping windows, can change. It gives a theoretical explanation to the necessity and the efficacy of intermittent adaptation which is used for separation of non stationary sources [9].

A Proof of Theorem 1

Let us consider that for $x > 0$, there is one and only one intersection point ρ between $p(x)$ and $g(x)$.

It is known that the fourth-order cumulant of a Gaussian signal is zero. As a consequence, we can write:

$$\int_{\mathbb{R}} x^4 g(x)dx = 3\int_{\mathbb{R}} x^2 g(x)dx = 3. \qquad (16)$$

Using (1) and the unit variance signal, the kurtosis can be rewritten as:

$$K[p(x)] = \int_{\mathbb{R}} x^4(p(x) - g(x))dx. \qquad (17)$$

According to result of section 2, we may only consider the even pdf. In addition, we just may study the sign of Υ:

$$\begin{aligned}\Upsilon &= \frac{1}{2}K[p(x)] \\ &= \int_0^{\infty} x^4(p(x)-g(x))dx \\ &= \int_0^{\rho} x^4(p(x)-g(x))dx \\ &\quad + \int_{\rho}^{\infty} x^4(p(x)-g(x))dx. \qquad (18)\end{aligned}$$

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:

$$\begin{aligned}\Upsilon &= \xi^4\int_0^{\rho}(p(x)-g(x))dx \\ &\quad +\lambda^4\int_{\rho}^{\infty}(p(x)-g(x))dx \\ &= \lambda^4\int_{\rho}^{\infty}(p(x)-g(x))dx \\ &\quad -\xi^4\int_0^{\rho}(g(x)-p(x))dx \qquad (19)\end{aligned}$$

where:

$$0 < \xi < \rho < \lambda. \qquad (20)$$

In fact, $p(x)$ and $g(x)$ are both pdf, so we have:

$$\begin{aligned}\int_0^{\infty}(p(x)-g(x))dx &= \int_0^{\rho}(p(x)-g(x))dx + \\ &\quad \int_{\rho}^{\infty}(p(x)-g(x))dx \\ &= 0. \qquad (21)\end{aligned}$$

From (21), and taking into account that $p(x)$ is over-Gaussian, we deduce:

$$\int_{\rho}^{\infty}(p(x)-g(x))dx = \int_0^{\rho}(g(x)-p(x))dx > 0. \qquad (22)$$

Using (19), (20) and (22), we remark that:

$$\Upsilon = (\lambda^4 - \xi^4)\int_{\rho}^{\infty}(p(x)-g(x))dx > 0. \qquad (23)$$

Finally, if $p(x)$ is an over-Gaussian pdf (with the assumption of an unique intersection positive point between $p(x)$ and $g(x)$), then its kurtosis is positive. Using the same reasoning, we can claim that a sub-Gaussian pdf has a negative kurtosis.

References

[1] A. Benvéniste, M. Métivier, and P. Priouret. *Adaptive Algorithms and Stochastic Approximations.* Spring-Verlag, 1990.

[2] A. Cichocki, S. Douglas, S. I. Amari, and P. Mierzejewski. Independent component analysis for noisy data. In *Proc. of International Workshop on Independence & Artificial Neural Networks*, pages 52–58, Tenerife, SPAIN, February 9-10 1994.

[3] N. Delfosse and P. Loubaton. Adaptive blind separation of independent sources: A deflation approach. *Signal Processing*, 45(1):59–83, July 1995.

[4] B. Laheld and J. F. Cardoso. Adaptive source separation without pre-whitening. In M.J.J. Holt, C.F.N. Cowan, P.M. Grant, and W.A. Sandham, editors, *Signal Processing VII, Theories and Applications*, pages 183–186, Edinburgh, Scotland, September 1994. Elsevier.

[5] A. Mansour and C. Jutten. Fourth order criteria for blind separation of sources. *IEEE Trans. on Signal Processing*, 43(8):2022–2025, August 1995.

[6] A. Mansour, A. Kardec Barros, and N. Ohnishi. Comparison among three estimators for high order statistics. In *Fifth International Conference on Neural Information Processing (ICONIP'98)*, pages 899–902, Kitakyushu, Japan, 21-23 October 1998.

[7] P. McCullagh. *Tensor methods in statistics*. Chapman and Hall, 1987.

[8] E. Moreau and O. Macchi. New self-adaptive algorithms for source separation based on contrast functions. In *IEEE Signal Processing Workshop on Higher-Order Statistics*, pages 215–219, South Lac Tahoe, USA (CA), June 1993.

[9] L. Nguyen Thi and C. Jutten. Blind sources separation for convolutive mixtures. *Signal Processing*, 45(2):209–229, 1995.