MODELLING SIGNALS OF ARBITRARY KURTOSIS FOR TESTING BSS METHODS

Vicente Zarzoso and Asoke K. Nandi

Department of Electronic and Electrical Engineering, University of Strathclyde,
204 George Street, Glasgow G1 1XW, U.K.
`e-mail: vicente@spd.eee.strath.ac.uk`

Indexing terms: Blind source separation, kurtosis, Bernoulli distribution.

Abstract

It is shown how a very simple distribution, the Bernoulli random variable, exhibits all possible normalized kurtosis values when varying the probability of its two events. This result means that signals of any normalized kurtosis may be modelled by this distribution, which is of relevance within the framework of blind source separation based on fourth-order cumulants.

1 Introduction

A problem of growing interest in signal processing is the recovery of a set of signals — the *sources* — from another set of linear mixtures. In many applications neither the source signals nor the coefficients of the linear transformation are known, and then one is facing a *blind source separation (BSS)* problem. Most successful methods for BSS rely on higher-order statistics (HOS) of the available data, and in particular on their fourth-order cumulants [1 − 5]. For this reason, the performance of some of these methods is strongly dependent on the sign or value of source (normalized) kurtosis [1 − 4]. In order to carry out a thorough testing of these procedures, it is thus desirable to have available distributions with a range of kurtosis values.

`IEE Electronics Letters, Vol. 34, No. 1, 8th January 1998, pp. 29-30.`
`(Errata: Vol. 34, No. 7, 2nd April 1998, p. 703.)`

Consider the Bernoulli random variable (r.v.), with events $x = a$ and $x = b$ of probability

$$P(x) = \begin{cases} 1-p, & x = a, \\ p, & x = b. \end{cases} \tag{1}$$

In this communication it is proved that such a distribution presents a normalized kurtosis lying in the range $[-2, +\infty[$ as the parameter p varies along the probability interval $[0, 1]$; that is, this distribution is able to produce all allowed values for the statistic of interest. This outcome suggests that straightforward binary sequences of suitable probabilities may be used instead of other more complicated source distributions to assess the performance of fourth-order cumulant based BSS methods.

2 Normalized kurtosis of the Bernoulli distribution

The moments and cumulants up to the fourth order of the Bernoulli r.v. (1) are summarized in table 1. The normalized kurtosis is defined as

$$\gamma = \frac{\text{kurtosis}}{\text{variance}^2} = \frac{\kappa_4}{\kappa_2^2}. \tag{2}$$

This is equivalent to the evaluation of the kurtosis when the variance of the distribution is unity, i.e., when the r.v. has been normalized to unit variance. Table 2 shows the value of this parameter for several well-known random distributions. With the help of table 1, the statistic γ can be easily worked out for the Bernoulli distribution:

$$\gamma = \frac{1 - 6p + 6p^2}{p\,(1-p)} = \frac{6(p-p_1)(p-p_2)}{p\,(1-p)}, \tag{3}$$

p_1 and p_2 being the roots of the polynomial in the numerator of γ, given by

$$p_1,\, p_2 = \frac{1}{2}\left(1 \pm \frac{1}{\sqrt{3}}\right). \tag{4}$$

Figure 1 plots γ against p. In that plot it is seen how the normalized kurtosis takes all values from -2 to $+\infty$ as p varies across its allowed range $[0, 1]$. Furthermore, it is easy to show [6, page 107] that $\gamma \geq -2$ for any distribution. The equality holds true for the only case of a Bernoulli pdf with $p = 0.5$.

As a conclusion, with an appropriate choice of the parameter p, the Bernoulli r.v. is able to provide any normalized kurtosis value. Specifically, from equations (3) and (4), the value of probability p necessary to obtain a desired γ is given by:

$$p = \frac{1}{2}\left(1 \pm \sqrt{\frac{\gamma+2}{\gamma+6}}\right). \tag{5}$$

A Bernoulli distribution naturally defines a binary stochastic process. From the BSS standpoint, this observation, along with the previous result, indicates that binary sequences of proper probability values may be employed in the place of other source random processes to evaluate BSS methods. This is empirically checked below.

3 Experimental results

An adaptive source separation algorithm is chosen: Cardoso's and Laheld's EASI method [1]. Consider the noiseless BSS problem

$$\mathbf{y}_t = M\mathbf{x}_t, \qquad t = 1, 2, \ldots, \tag{6}$$

where vector $\mathbf{y}_t$ contains the samples of the m sensors, vector $\mathbf{x}_t$ the n source signals, and M represents the $(m \times n)$ mixing matrix. The aim is to search for a separating matrix W such that the global mixing-unmixing system $G = WM$ may be factorized as the product PD, where P is a permutation matrix and D an invertible diagonal matrix. Basically, EASI employs the following adaptation rule for the separating matrix:

$$W_{t+1} = (I - \lambda_t H_{\mathrm{g}}(\mathbf{s}_t))\, W_t. \tag{7}$$

This rule boils down to the specification of a set of non-linear functions $\mathbf{g}(\cdot)$ of the estimated sources $\mathbf{s}_t = W\mathbf{y}_t$. There exists a very close relationship among these functions, the sources statistics and the stability of the algorithm. In fact, for the case of cubic nonlinearities, the convergence condition becomes that the sum of normalized kurtosis of any two sources must be less than zero:

$$\gamma^{(i)} + \gamma^{(j)} < 0, \qquad 1 \le i < j \le n, \tag{8}$$

$\gamma^{(i)}$ being the ith-source normalized kurtosis.

For the following simulations, a two-source two-sensor scenario is chosen, with a fixed non-singular mixing matrix:

$$M = \begin{bmatrix} 0.2 & -0.3 \\ 0.1 & 0.4 \end{bmatrix}. \tag{9}$$

A fixed value of the adaptation step is also selected as $\lambda_t = 0.01$. The choice of cubic nonlinearities and uniformly distributed independent sources meets the stability condition (8). The identity matrix is taken as the initial separating matrix. The trajectories of the square global matrix elements as a function of the iteration number (equal to the sample number, in this case) are recorded for each separation. Separations are obtained in 100 Monte Carlo runs with these conditions, but with different source realizations each time. Finally, the trajectories are averaged over the 100 trials. Figure 2 displays the results obtained for the first 500 iterations. The global matrix diagonal elements converge to 1, while the two off-diagonal elements converge to 0, as expected for a successful source extraction.

Next, the uniformly distributed sources are substituted by two independent binary sequences with the same normalized kurtosis value as the uniform distribution, $\gamma = -1.2$. From (5), the probability of one of their two levels is obtained as $p = 0.7041$ (note that the choice of the plus or the minus sign in that expression is irrelevant to the results). The above experiment is carried out once again, in the same conditions but with the new set of sources and over a new independent set of 100 Monte Carlo trials. The trajectories of the mean square value of the global system elements are represented in figure 3, again for the first 500 iterations. Note that although the actual waveforms and pdf shapes inputted to the algorithm in both cases are completely different, figures 2 and 3 are essentially identical. Only the value of the normalized kurtosis affects the method performance.

Analogous results should be obtained for batch-processing BSS algorithms based on fourth-order cumulants, providing their asymptotic performance is solely dependent on the kurtosis of the source signals.

4 Conclusion

A Bernoulli random variable characterized by a probability p exhibits a normalized kurtosis γ which is entirely dependent on p and takes *all* possible values as p runs across [0, 1]. Hence it is possible to model *any* signal distribution, as far as γ is concerned, by means of a simple binary sequence and an appropriate choice of the probability of its two only values. In the context of BSS this result allows one to employ binary sequences for testing BSS methods based on fourth-order cumulants, instead of its continuous-distribution counterparts of the same kurtosis values, as seen in the simulations. Moreover, a continuous pdf may be difficult to find for a desired value of γ, in contrast to the binary sequence, for which there always exists a value of p that provides the required value of γ, according to expression (5).

One has to keep in mind, however, that this procedure for simulating distributions via binary sequences to assess BSS methods is only valid as long as the BSS method being used does not take advantage of the discrete nature of the signals at its input.

References

1. J.-F. Cardoso and B. H. Laheld, "Equivariant Adaptive Source Separation", *IEEE Transactions on Signal Processing*, Vol. 44, No. 12, December 1996, pp. 3017–3030.

2. F. Harroy and J.-L. Lacoume, "Maximum likelihood estimators and Cramer-Rao bounds in source separation", *Signal Processing*, Vol. 55, December 1996, pp. 167–177.

3. A. Mansour and C. Jutten, "Fourth-Order Criteria for Blind Sources Separation", *IEEE Transactions on Signal Processing*. Vol. 43, No. 8, August 1995, pp. 2022–2025.

4. A. K. Nandi and V. Zarzoso, "Fourth-Order Cumulant Based Blind Source Separation", *IEEE Signal Processing Letters*. Vol. 3, No. 12, December 1996, pp. 312–314.

5. P. Comon, "Independent component analysis, A new concept?", *Signal Processing*, Vol. 36, No. 3, April 1994, pp. 287–314.

6. M. G. Kendall and A. Stuart, *The Advanced Theory of Statistics*, Vol. I, Fifth Edition, London: Charles Griffin & Company Ltd., 1987.

MOMENTS	CUMULANTS
$\mu_1 = a\,(1-p) + b\,p$	$\kappa_1 = \mu_1$
$\mu_2 = a^2(1-p) + b^2 p$	$\kappa_2 = (b-a)^2 p\,(1-p)$
$\mu_3 = a^3(1-p) + b^3 p$	$\kappa_3 = (b-a)^3 p\,(1-p)(1-2p)$
$\mu_4 = a^4(1-p) + b^4 p$	$\kappa_4 = (b-a)^4 p\,(1-p)(1-6p+6p^2)$

Table 1: Statistics up to the 4th-order of the Bernoulli distribution.

DISTRIBUTION	γ
Uniform-phase sinusoid	-1.5
Uniform	-1.2
Gaussian	0
Laplacian	3
Exponential	6

Table 2: Normalized kurtosis values of some random distributions.

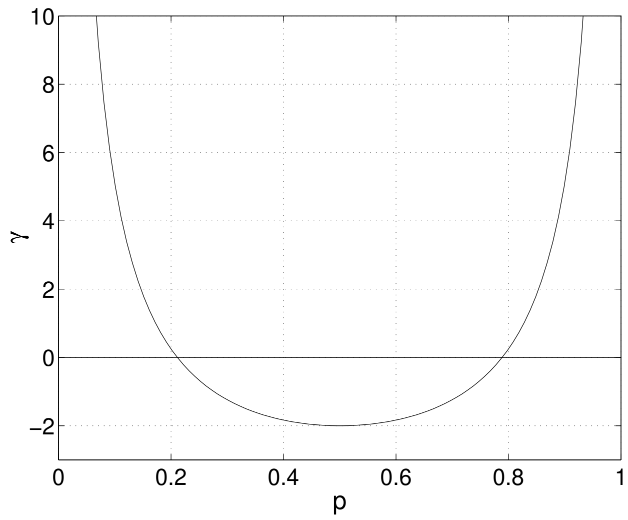

Figure 1: Normalized kurtosis versus probability of an event in the Bernoulli random variable.

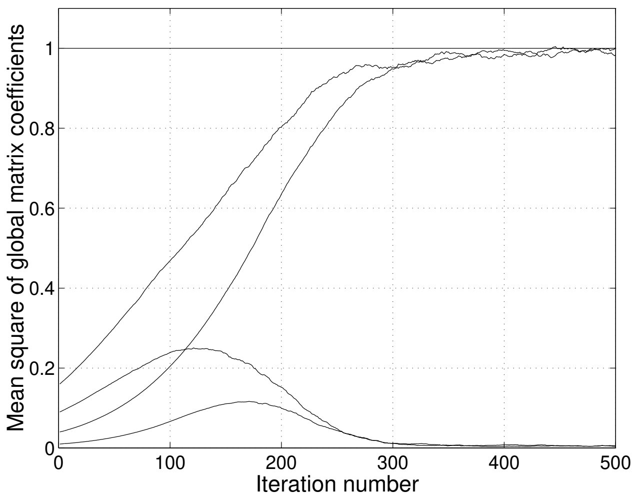

Figure 2: Trajectories averaged over 100 Monte Carlo runs of the square global matrix elements obtained by the EASI algorithm for two uniformly distributed sources. Cubic nonlinearities, $\lambda_t = 0.01$.

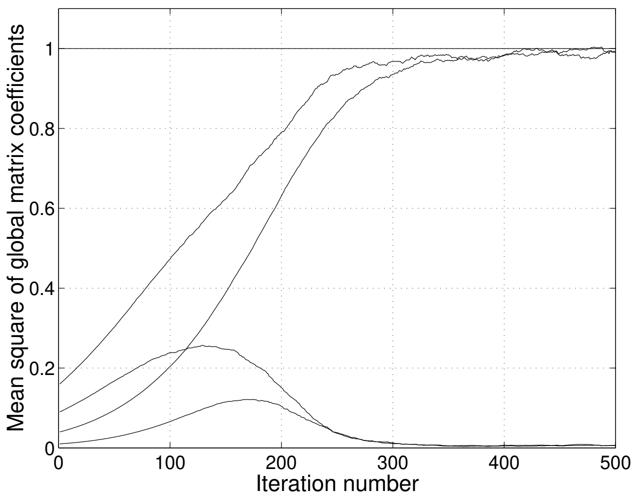

Figure 3: Trajectories averaged over 100 Monte Carlo trials of the square global matrix coefficients obtained by the EASI method for equivalent binary sources ($p = 0.7041$). Cubic nonlinearities, $\lambda_t = 0.01$.