Supplementary Materials for Submission "Second-order Approximation of Minimum Discrimination Information in Independent Component Analysis"

1 Details of experiments with simulated signals

1.1 Details of 18 distributions

Page 4, Column 1, Section IV, Subsection A, Fig1
The details of the sources' distributions used in the first experiment are shown in Figure 1.

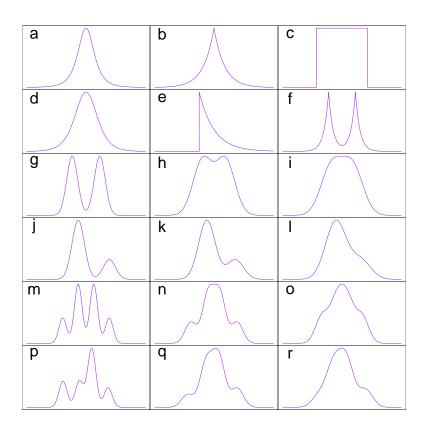


Figure 1: Probability density functions of sources: (a) Student with 3 degrees of freedom; (b) double exponential; (c) uniform; (d) Student with 5 degrees of freedom; (e) exponential; (f) mixture of two double exponentials; (g) - (h) - (i) symmetric mixtures of two Gaussians: multimodal, transitional and unimodal; (j) - (k) - (l) nonsymmetric mixtures of two Gaussians, multimodal, transitional and unimodal; (m) - (n) - (o) symmetric mixtures of four Gaussians: multimodal, transitional and unimodal; (p) - (q) - (r) nonsymmetric mixtures of four Gaussians: multimodal, transitional and unimodal.

1.2 Details of 18 two dimensional experiments

Page 4, Column 1, Section IV, Subsection B

The Amaric metrics and CPU elapsed time are recorded in Table 1, Table 2. Since the public EFICA package was implemented on Matlab development platform and other ICA algorithms we used were implemented on R development platform, it might not be appropriate to compare the CPU elapsed time between EFICA and other ICA algorithms (we denoted the EFICA's CPU elapsed time with superscript *).

FastICA works well only when its single nonlinear function is close to the unknown sources' distribution, otherwise its separation performance or negentropy estimation might degenerate due to the unwanted density mismatching. As can be seen in Figure 1, F-G0 and F-G1 perform well in symmetric distributions (f, g, h, i) thanks to their symmetric nonlinear functions (F-G0: $G0 = \frac{y_i^4}{4}$, F-G1: $G1 = \log \cosh(y_i)$) used in negentropy estimation (density matching), whereas their separation performance deteriorates in the cases of nonsymmetric or nontrivial distributions (j, k, l, p, q, r, n). Compared with the single nonlinear function used in F-G0 and F-G1, the nonparametric estimation used in PICA or several nonlinear functions used in MICA are more flexible in these nontrivial cases. Although PICA performed best in general (thanks to its flexible nonparametric estimation concerning negentropy), its CPU elapsed time is demanding. The CPU elapsed time required by MICA (88.89 ms) is less than PICA (556.11 ms), as a result of transforming a sequence of IRLS (in PICA) into a single weighted least squares (in MICA).

Table 1: Amari metrics for two-component ICA with 1000 samples .For each pdf (from a to r), averages over 100 replicates are presented.

pdfs	F-G0	F-G1	PICA	MICA	WICA	EICA
a b c	$4.06(\pm 4.06)$ $4.12(\pm 3.22)$ $1.56(\pm 0.93)$	$2.47(\pm 1.7)$ $2.92(\pm 1.47)$ $1.81(\pm 1.06)$	$\begin{array}{c} 2.47(\pm 1.83) \\ 2.62(\pm 1.29) \\ 3.44(\pm 12.86) \end{array}$	$3.29(\pm 8.24)$ $3.48(\pm 2.2)$ $5.59(\pm 15.84)$	$13.19(\pm 17.61) 33.31(\pm 25.76) 56.45(\pm 27.76)$	$3.54(\pm 2.82)$ $2.03(\pm 1.36)$ $1.42(\pm 1.22)$
d e f	$5.1(\pm 3.97)$ $3.94(\pm 3.76)$ $2.62(\pm 1.72)$	$3.78(\pm 2.45)$ $3.14(\pm 1.87)$ $1.7(\pm 0.97)$	$3.84(\pm 2.48)$ $1.6(\pm 0.91)$ $27.11(\pm 42.53)$	$4.8(\pm 3.75)$ $6.32(\pm 12.37)$ $1.94(\pm 1.42)$	$20.27(\pm 21.07)$ $30.01(\pm 24.74)$ $47.93(\pm 28.39)$	$\begin{array}{c} 6.42(\pm 5.42) \\ 3.27(\pm 2.11) \\ 6.16(\pm 7.54) \end{array}$
g h i	$1.49(\pm 0.84) 4.23(\pm 3.02) 6.75(\pm 4.76)$	$\begin{array}{c} 1.44(\pm 0.85) \\ 3.32(\pm 2.24) \\ 5.87(\pm 4.44) \end{array}$	$23.55(\pm 40.63) 4.94(\pm 12.45) 7.63(\pm 8.36)$	$4.5(\pm 6.2)$ $4.8(\pm 7.7)$ $8.87(\pm 9.44)$	$\begin{array}{c} 66.11(\pm 24.74) \\ 50.31(\pm 29.29) \\ 46.46(\pm 28.5) \end{array}$	$3.61(\pm 2.37)$ $4.18(\pm 3.14)$ $6.02(\pm 4.22)$
j k l	$7.52(\pm 8.11) 17.47(\pm 20.53) 28.44(\pm 27.52)$	$44.33(\pm 41.06) 34.22(\pm 34.17) 32.06(\pm 27.38)$	$5.68(\pm 17.99)$ $2.08(\pm 1.35)$ $4.03(\pm 2.46)$	$1.76(\pm 1.04)$ $2.2(\pm 1.39)$ $5.3(\pm 9.13)$	$37.97(\pm 28.9)$ $41.98(\pm 28.4)$ $42.89(\pm 26.42)$	$\begin{array}{c} 68.7(\pm 35.46) \\ 33.2(\pm 33.37) \\ 30.67(\pm 27.77) \end{array}$
m n o	$3.49(\pm 2.25)$ $14.67(\pm 15.28)$ $4.78(\pm 2.8)$	$5.06(\pm 3.38) 47.31(\pm 27.39) 6.68(\pm 4.28)$	$4.77(\pm 16.37) 10.13(\pm 23.51) 6.71(\pm 13.26)$	$12.17(\pm 14.97) \\ 38.1(\pm 26.83) \\ 11.01(\pm 14.14)$	$53.46(\pm 29.39)$ $50.46(\pm 25.77)$ $44.69(\pm 28.77)$	$3.48(\pm 2.23)$ $48.6(\pm 26.51)$ $5.51(\pm 3.56)$
p q r	$3.8(\pm 2.79)$ $41.18(\pm 27.54)$ $24.98(\pm 23.38)$	$15.67(\pm 15.49) 23.46(\pm 26.62) 44.55(\pm 28.95)$	$2.33(\pm 8.97)$ $5.26(\pm 13.14)$ $6.39(\pm 10.62)$	$4.09(\pm 3.18)$ $17.63(\pm 22.69)$ $12.38(\pm 13.41)$	$46.19(\pm 27.67) 43.2(\pm 27.64) 49.51(\pm 28.03)$	$8.37(\pm 15.7) 27.65(\pm 26.6) 43.72(\pm 28.27)$

1.3 Details of high dimensional experiments

We included the high dimensional experiments (originally shown in the first submission) in the supplementary materials due to the page limit of "IEEE Signal Processing Letters". As can be seen in Table 3 and Table 4, PICA performed best with the longest running time (thanks to its flexible nonparametric estimation, whereas a sequence of IRLS are required to solved), and MICA have the second lowest Amari metrics with the acceptable CPU elapsed time (thanks to the adding more nonlinear functions in the negentropy estimation without destroying the efficient fixed-point method compared with F-G0 and F-G1).

Table 2: CPU elapsed time (ms) for two-component ICA with 1000 samples .For each pdf (from a to r), averages over 100 replicates are presented.

pdfs	F-G0	F-G1	PICA	MICA	WICA	EICA
a b c	$6.83(\pm 5.87)$ $8.05(\pm 4.03)$ $7.04(\pm 5.69)$	$9.34(\pm 6.15)$ $8.39(\pm 5.48)$ $5.77(\pm 2.94)$	$\begin{array}{c} 883.67(\pm 385.18) \\ 510.72(\pm 210.95) \\ 442.81(\pm 201.58) \end{array}$	$72.97(\pm 31.11)$ $54.28(\pm 16.98)$ $78.28(\pm 38.45)$	$5.23(\pm 4.75)$ $4.69(\pm 1.53)$ $4.67(\pm 1.1)$	$7.4^*(\pm 5.79^*)$ $5^*(\pm 5.41^*)$ $5.1^*(\pm 5.02^*)$
d e f	$8.55(\pm 6.81)$ $7.66(\pm 3.67)$ $5.85(\pm 2.55)$	$10.02(\pm 6.47)$ $10.33(\pm 7.4)$ $4.62(\pm 2.43)$	$740.89(\pm 267.51)$ $435.1(\pm 193.75)$ $357.74(\pm 173.23)$	$65.94(\pm 29.66)$ $77.59(\pm 30.53)$ $56.24(\pm 32.51)$	$4.71(\pm 0.94) 4.73(\pm 2.11) 4.79(\pm 1.57)$	$5.5^*(\pm 5.57^*)$ $5.2^*(\pm 5.41^*)$ $4.5^*(\pm 5^*)$
g h i	$5.64(\pm 3.7)$ $9.75(\pm 6.77)$ $11.93(\pm 7.34)$	$4.6(\pm 2.65) 7.69(\pm 4.87) 11.14(\pm 7.62)$	$372.12(\pm 202.82)$ $611.35(\pm 201.57)$ $668.38(\pm 149.08)$	$105.74(\pm 35.33) 104.21(\pm 29.99) 111.74(\pm 22)$	$4.76(\pm 1.75) 4.56(\pm 1.04) 4.61(\pm 1.44)$	$4.4^*(\pm 4.99^*)$ $3.3^*(\pm 4.93^*)$ $4.6^*(\pm 5.21^*)$
j k l	$10.51(\pm 6.21) 12.35(\pm 7.48) 13.03(\pm 8.06)$	$12.6(\pm 8.01) 14.87(\pm 8.34) 15.71(\pm 7.65)$	$373.46(\pm 214.14)$ $402.61(\pm 205.03)$ $593.5(\pm 198.13)$	$35.31(\pm 16.14)$ $45.31(\pm 25.61)$ $60.85(\pm 29.8)$	$4.55(\pm 0.67) 4.67(\pm 1.51) 4.76(\pm 1.78)$	$4.8^{*}(\pm 5.22^{*})$ $3.8^{*}(\pm 5.46^{*})$ $4.9^{*}(\pm 5.59^{*})$
m n o	$9.21(\pm 4.83)$ $14.26(\pm 7.52)$ $9.7(\pm 6.33)$	$9.78(\pm 5.99)$ $17.99(\pm 8.74)$ $11.56(\pm 8.27)$	$502.49(\pm 253.81)$ $644.78(\pm 206.62)$ $688.88(\pm 169.91)$	$128.87(\pm 22.69) 127.16(\pm 20.71) 127.19(\pm 30.57)$	$4.96(\pm 2.3) 5.08(\pm 1.76) 5.27(\pm 2.82)$	$5^*(\pm 5.95^*)$ $5.2^*(\pm 5.94^*)$ $3.8^*(\pm 4.88^*)$
p q r	$10.27(\pm 6.84) 11.96(\pm 10.44) 15.81(\pm 8.99)$	$14.06(\pm 7.76) 14.98(\pm 9.36) 17.66(\pm 9.11)$	$411.55(\pm 198.01) 661.84(\pm 222.94) 734.84(\pm 166.28)$	$66.94(\pm 35.25) 138.21(\pm 42.57) 116.72(\pm 41.27)$	$\begin{array}{c} 4.97(\pm 1.75) \\ 5.14(\pm 2.23) \\ 5.47(\pm 2.63) \end{array}$	$3.8^*(\pm 4.88^*)$ $5^*(\pm 5.6^*)$ $4.1^*(\pm 5.34^*)$

Table 3: Amari metrics for ICA algorithms in high dimensions. m is the dimension, reps is the number of replications, and N is the sample size

m	reps	N	F-G0	F-G1	PICA	WICA	MICA	EICA
4	100	1000	$22.82(\pm 16.69)$	$27.98(\pm 24.84)$	$8.22(\pm 7.34)$	$52.11(\pm 29.37)$	$16.28(\pm 9.07)$	25.94(±24)
4 8	100 50	$\frac{4000}{2000}$	$11.03(\pm 8.66)$ $39.09(\pm 25.19)$	$12.37(\pm 11.85)$ $42.44(\pm 28.6)$	$4.28(\pm 4.91)$ $13.66(\pm 11.98)$	$37.86(\pm 22.83)$ $120.41(\pm 36.36)$	$7.34(\pm 5.41)$ $26(\pm 9.8)$	$12.73(\pm 14.57)$ $39.35(\pm 27.45)$
8	50	4000	$25.2(\pm 11.41)^{'}$	$28.08(\pm 19.16)$	$7.79(\pm 1.76)$	$99.15(\pm 27.37)^{'}$	$16.07(\pm 4.74)$	$25.49(\pm 18.1)$
16 16	$\frac{25}{25}$	$\frac{2000}{4000}$	$118.16(\pm 29.06) \\ 61.35(\pm 10.91)$	$111.62(\pm 28.77) \\ 70.33(\pm 21.45)$	$53.19(\pm 39.23)$ $17.61(\pm 1.42)$	$309.51(\pm 29.57)$ $267.38(\pm 28.37)$	$60.14(\pm 15.57) \\ 37.3(\pm 4.96)$	$102.37(\pm 22.53) 58.96(\pm 14.32)$

2 Details of experiments with real data

Page 4, Column2, Section IV, Subsection C

The original images and recovered images are plotted in Figure 2. $\,$

Table 4: CPU elapsed time (ms) for ICA algorithms in high dimensions. m is the dimension, reps is the number of replications, and N is the sample size

m	reps	N	F-G0	F-G1	PICA	WICA	MICA	EICA
4 4 8	100 100 50	1000 4000 2000	$54.25(\pm 20.76)$ $214.62(\pm 85.41)$ $218.5(\pm 29.37)$	$47.12(\pm 15.14)$ $164.54(\pm 66.21)$ $172.78(\pm 22.36)$	$4112.41(\pm 1166.51)$ $6373.6(\pm 2616.8)$ $11681.34(\pm 1402.03)$	$13.95(\pm 7.91)$ $242.57(\pm 62.35)$ $62.44(\pm 51.3)$	$302.52(\pm 89.1)$ $359.83(\pm 145.28)$ $636.94(\pm 56.09)$	$ \begin{array}{c} 10^* (\pm 0^*) \\ 10^* (\pm 0^*) \\ 20^* (\pm 0^*) \end{array} $
8 16 16	50 25 25	4000 2000 4000	$444.8(\pm 110.54) 563.2(\pm 131.68) 1020.8(\pm 325.49)$	$\begin{array}{c} 293.68(\pm 79.85) \\ 414.24(\pm 75.91) \\ 742.52(\pm 250.03) \end{array}$	$\begin{array}{c} 12115(\pm 3489.61) \\ 27755.8(\pm 4026.31) \\ 30276.4(\pm 9758.36) \end{array}$	$\begin{array}{c} 229.88(\pm 88.4) \\ 76.72(\pm 18.64) \\ 236.28(\pm 91.03) \end{array}$	$750.5(\pm 165.15)$ $1154.6(\pm 26.73)$ $1412.04(\pm 78.95)$	$30^*(\pm 0^*)$ $80^*(\pm 0^*)$ $80^*(\pm 0^*)$

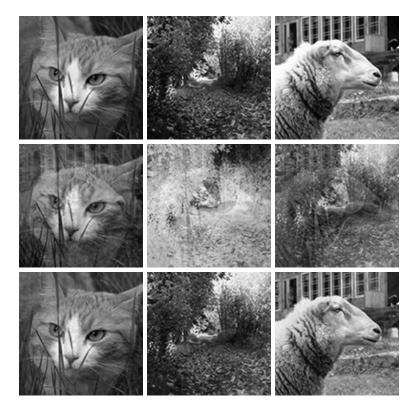


Figure 2: Recovery of image sources by MDIICA. From the top row to the bottom row, original sources, mixtures and recovered sources are plotted.