

Average Convergence Behavior of the FastICA Algorithm for Blind Source Separation

Scott C. Douglas¹, Zhijian Yuan², and Erkki Oja²

¹ Department of Electrical Engineering
Southern Methodist University
Dallas, Texas 75275 USA

² Neural Networks Research Centre
Helsinki University of Technology
Espoo 02015 Finland

Abstract. The FastICA algorithm is a popular procedure for independent component analysis and blind source separation. In this paper, we analyze the average convergence behavior of the single-unit FastICA algorithm with kurtosis contrast for general m -source noiseless mixtures. We prove that this algorithm causes the average inter-channel interference (ICI) to converge exponentially with a rate of $(1/3)$ or -4.77dB at each iteration, independent of the source mixture kurtoses. Explicit expressions for the average ICI for the three- and four-source mixture cases are also derived, along with an exact expression for the average ICI in a particular situation. Simulations verify the accuracy of the analysis.

1 Introduction

The FastICA algorithm is a popular procedure for independent component analysis and blind source separation. The technique is simple to implement and converges quickly when applied to mixtures of independent non-Gaussian sources. The algorithm's convergence speed is locally-quadratic, and it is cubic when a kurtosis-based cost is employed [1, 2]. This cubic convergence behavior can be described using the analytical expressions for the evolution of the combined system coefficient vector $\mathbf{c}_t = [c_{1,t} \cdots c_{m,t}]^T$ for infinite data measurements, as given by

$$c_{i,t+1} = \frac{\kappa_i c_{i,t}^3}{\sqrt{\sum_{j=1}^m \kappa_j^2 c_{j,t}^6}}, \quad (1)$$

where κ_i is the i th source kurtosis. The vector \mathbf{c}_t corresponds to the weight vector \mathbf{w}_t of the single-unit FastICA algorithm in a transformed coordinate system where the independent components are explicitly included.

The FastICA algorithm's convergence behavior depends on the initial point of the algorithm, represented by \mathbf{w}_0 or \mathbf{c}_0 . As this point is usually chosen fully at random in lack of any prior knowledge of the mixtures, an interesting question arises: What is the *average* convergence behavior of the algorithm across a distribution of initial points? Consider the performance metric known as the *inter-channel interference* (ICI) defined for the m -source case as

$$ICI_t^{(m)} = \frac{\sum_{i=1}^m c_{i,t}^2 - \max_{1 \leq i \leq m} c_{i,t}^2}{\max_{1 \leq i \leq m} c_{i,t}^2}. \quad (2)$$

Recently, an interesting observation about the FastICA algorithm with kurtosis contrast was made [3]: For a random initial \mathbf{w}_0 or \mathbf{c}_0 , convergence of the average ICI appears to follow the “(1/3)rd Rule” given by

$$E\{ICI_t^{(m)}\} = \left(\frac{1}{3}\right) E\{ICI_{t-1}^{(m)}\}, \quad (3)$$

over almost the entire convergence period. Additional work has shown that this convergence behavior can be proven for the single-unit FastICA algorithm applied to simple two-source mixtures [4, 5], but it is not clear if such behavior extends to the general m -source mixture case.

In this paper, we prove that the FastICA algorithm with kurtosis contrast indeed obeys the “(1/3)rd Rule” for general m -source mixtures over a large portion of the convergence period. Our analysis employs a norm-constrained Gaussian prior for the initial separation system vector. Moreover, explicit expressions for the average ICI in the three- and four-source mixture cases are provided, and simulations are used to verify the analytical performance predictions.

2 Average Behavior of the FastICA Algorithm for a Three-Source Mixture

Before presenting the general m -source performance analysis, we introduce the analytical tools used in our derivations for a three-source separation task. The unconstrained (*e.g.* non-normalized) combined system vector at iteration t is

$$\mathbf{c}_t = [\kappa_1^{\frac{q}{2}} x^{\frac{p}{2}} \quad \kappa_2^{\frac{q}{2}} y^{\frac{p}{2}} \quad \kappa_3^{\frac{q}{2}} z^{\frac{p}{2}}]^T, \quad (4)$$

where $p = 2(3^t)$ and $q = 3^t - 1$. Employing this choice within (2) results in the general ICI expression derived in [1]. At time $t = 0$, we have $\mathbf{c}_0 = [x \quad y \quad z]^T$, where x , y , and z are random variables with some assumed probability density function (p.d.f.). A reasonable joint p.d.f. choice for $\{x, y, z\}$ would give a uniform prior for the direction of \mathbf{c}_0 . We can induce such a p.d.f by letting x , y , and z be zero mean, uncorrelated, and jointly Gaussian. We can then express $ICI_t^{(3)}$ using ratios of powers of x , y , and z *without normalization*, and the resulting expectations can be evaluated without trigonometric functions.

The portion of the average ICI at iteration t in which the first kurtosis component is being extracted, such that the first element of (4) is the largest, is

$$E\{ICI_{1,t}^{(3)}\} = \frac{8}{(2\pi)^{3/2}} \int_0^\infty \frac{e^{-\frac{x^2}{2}}}{x^p} dx \left[\int_0^{ax} \left(\left(\frac{y}{a}\right)^p + \left(\frac{z}{b}\right)^p \right) e^{-\frac{y^2}{2}} dy \right] \int_0^{bx} e^{-\frac{z^2}{2}} dz \quad (5)$$

where $a = \left(\frac{\kappa_1}{\kappa_2}\right)^{\frac{q}{p}}$ and $b = \left(\frac{\kappa_1}{\kappa_3}\right)^{\frac{q}{p}}$. (6)

The integral in brackets on the right-hand side of (5) can be approximated as

$$\int_0^{ax} \left(\left(\frac{y}{a}\right)^p + \left(\frac{z}{b}\right)^p \right) e^{-\frac{y^2}{2}} dy \approx \left[e^{-\frac{a^2 x^2}{2}} \int_0^{ax} \left(\frac{y}{a}\right)^p dy \right] + \left(\frac{z}{b}\right)^p \int_0^{ax} e^{-\frac{y^2}{2}} dy \quad (7)$$

$$= \left[a e^{-\frac{a^2 x^2}{2}} \frac{x^{p+1}}{p+1} \right] + \left(\frac{z}{b}\right)^p \int_0^{ax} e^{-\frac{y^2}{2}} dy. \quad (8)$$

Substituting (8) into (5), we obtain

$$E\{ICI_{1,t}^{(3)}\} = \frac{8}{(2\pi)^{3/2}(p+1)} \left[\int_0^\infty ax e^{-\frac{x^2(1+a^2)}{2}} \int_0^{bx} e^{-\frac{z^2}{2}} dz dx + \int_0^\infty bx e^{-\frac{x^2(1+b^2)}{2}} \int_0^{ax} e^{-\frac{y^2}{2}} dy dx \right] \quad (9)$$

We can evaluate the integrals within brackets on the right-hand-side of (9) as

$$\int_0^\infty x e^{-\frac{x^2(1+a^2)}{2}} \left[\int_0^{bx} e^{-\frac{z^2}{2}} dz \right] dx = \frac{b}{1+a^2} \sqrt{\frac{\pi}{2(1+a^2+b^2)}} \quad (10)$$

$$\text{Therefore, } E\{ICI_{1,t}^{(3)}\} = \frac{2}{\pi(p+1)} \frac{1}{\sqrt{1+a^2+b^2}} \left[\frac{b}{a^{-1}+a} + \frac{a}{b^{-1}+b} \right]. \quad (11)$$

Now, as t increases, we have

$$\lim_{t \rightarrow \infty} \frac{q}{p} = \frac{1}{2}, \quad \lim_{t \rightarrow \infty} a = \sqrt{\frac{\kappa_1}{\kappa_2}}, \quad \lim_{t \rightarrow \infty} b = \sqrt{\frac{\kappa_1}{\kappa_3}}, \quad \text{and } 2(3^t) \gg 1. \quad (12)$$

Substituting these results into (11), we obtain

$$E\{ICI_{1,t}^{(3)}\} = \frac{1}{\pi} \left(\frac{1}{3} \right)^t \frac{1}{\sqrt{\kappa_1\kappa_2 + \kappa_1\kappa_3 + \kappa_2\kappa_3}} \left[\frac{\kappa_1\kappa_2}{\kappa_1 + \kappa_2} + \frac{\kappa_1\kappa_3}{\kappa_1 + \kappa_3} \right]. \quad (13)$$

Invoking symmetry for the terms $E\{ICI_{2,t}^{(3)}\}$ and $E\{ICI_{3,t}^{(3)}\}$, we find an approximate expression for the average ICI to be

$$E\{ICI_t^{(3)}\} = \sum_{n=1}^3 E\{ICI_{n,t}^{(3)}\} = g_3(\kappa_1, \kappa_2, \kappa_3) \left(\frac{1}{3} \right)^t, \quad (14)$$

$$g_3(\kappa_1, \kappa_2, \kappa_3) = \frac{2}{\pi} \frac{1}{\sqrt{\kappa_1\kappa_2 + \kappa_1\kappa_3 + \kappa_2\kappa_3}} \left[\frac{\kappa_1\kappa_2}{\kappa_1 + \kappa_2} + \frac{\kappa_1\kappa_3}{\kappa_1 + \kappa_3} + \frac{\kappa_2\kappa_3}{\kappa_2 + \kappa_3} \right] \quad (15)$$

Eqn. (14) states that the average ICI for arbitrary three-source mixtures asymptotically obeys the “(1/3)rd Rule” in (3). Numerical evaluations of this expression show that it is extremely accurate in predicting the average ICI during the algorithm’s convergence period. Moreover, across all source kurtosis combinations, the maximum value of $E\{ICI_t^{(3)}\}$ occurs when $\kappa_1 = \kappa_2 = \kappa_3$, for which

$$E\{ICI_t^{(3)}\} = \frac{\sqrt{3}}{\pi} \left(\frac{1}{3} \right)^t. \quad (16)$$

3 Average Behavior of the FastICA Algorithm for General m -Source Mixtures

Eqn. (14) provides evidence that the kurtosis-based FastICA algorithm has exponential convergence with a rate that is independent of the source distributions.

Can this result be extended to m -source mixtures? And how accurate is the approximation used in (8)? The following theorem addresses these issues, the proof of which is outlined in the Appendix.

Theorem 1: *Assume that a single-unit FastICA algorithm with kurtosis contrast is applied to an m -source noiseless mixture with infinite data, and that the initial combined system coefficient vector \mathbf{c}_0 is uniformly-distributed on the m -dimensional unit hypersphere. Then, the average ICI at iteration t is*

$$E\{ICI_t^{(m)}\} = g_m(\kappa_1, \dots, \kappa_m) \left(\frac{1}{3}\right)^t + R(t, \kappa_1, \dots, \kappa_m), \quad (17)$$

where the m -dimensional function $g_m(\cdot)$ does not depend on t and $R(t, \cdot)$ decreases to zero faster than $(1/3)^t$ as $t \rightarrow \infty$.

The above theorem indicates that the “(1/3)rd Rule” holds in general for the single-unit FastICA procedure with kurtosis contrast. The convergence rate of the algorithm does not depend on the source kurtoses, which only affect the overall magnitude of the average ICI during the convergence period. This result explains why the FastICA algorithm can be called “fast” – the average convergence speed for the ICI is linear with a constant rate in all source scenarios.

The methodology used to derive the above theorem can in theory be used to find an asymptotic expression for the average ICI in (17) by determining an explicit expression for $g_m(\kappa_1, \dots, \kappa_m)$ for any m . For $m = 2$, see [4], or set $\kappa_3 = 0$ in (14). For $m = 4$, one can show that

$$g_4(\kappa_1, \kappa_2, \kappa_3, \kappa_4) = \frac{4}{\pi^2} (h_{1234} + h_{1243} + h_{1324} + h_{1342} + h_{1423} + h_{1432} + h_{2314} + h_{2341} + h_{2413} + h_{2431} + h_{3412} + h_{3421}) \quad (18)$$

$$h_{ijkl} = \frac{\sqrt{\kappa_i \kappa_j}}{\sqrt{\kappa_i} + \sqrt{\kappa_j}} \left[\frac{\sqrt{\kappa_k^{-1}}}{\sqrt{\kappa_i^{-1} + \kappa_j^{-1} + \kappa_k^{-1}}} \arctan \left(\frac{\sqrt{\kappa_l^{-1}}}{\sqrt{\kappa_i^{-1} + \kappa_j^{-1} + \kappa_k^{-1}}} \right) \right], \quad (19)$$

which reduces to (14) when $\kappa_4 = 0$. When $\kappa_1 = \kappa_2 = \kappa_3 = \kappa_4$, we have

$$E\{ICI_t^{(4)}\} = \frac{4}{\pi\sqrt{3}} \left(\frac{1}{3}\right)^t, \quad (20)$$

which is $4/3$ times larger than the maximum ICI in the three-source case with $\kappa_1 = \kappa_2 = \kappa_3$ and $4/\sqrt{3} = 2.31$ times larger than the maximum ICI in the two source case with $\kappa_1 = \kappa_2$. For $m > 4$, the integrals become difficult to evaluate.

4 An Exact Expression for the Average ICI for m -Source Mixtures in a Particular Situation

Given our reliance on a uniform distribution for the direction of \mathbf{c}_0 , one might wonder whether the “(1/3)rd Rule” for FastICA requires this assumption. The following analysis suggests that this behavior likely holds in other contexts.

Suppose the elements of $\mathbf{c}_0 = [c_{1,0} \cdots c_{m,0}]^T$ are uniformly-distributed on the interval $[0, 1]$. Of course, \mathbf{c}_0 is normally of unit length, but as scaling doesn't matter, we choose a scaled version of \mathbf{c}_0 instead. Assuming $c_{i,t} \geq 0$ does not change the value of $ICI_t^{(m)}$, either. When projected onto the unit hypersphere, this distribution tends to concentrate probability in the $[\pm 1 \pm 1 \cdots \pm 1]^T$ directions of m -dimensional space, making convergence somewhat more challenging for the algorithm. Moreover, we shall assume that $\kappa_i = \kappa_j$ for all i and j . Under this situation, the value of $E\{ICI_t^{(m)}\}$ is easy to compute.

Theorem 2: *For the situation above, the average ICI at iteration t is exactly*

$$E\{ICI_t^{(m)}\} = \frac{m-1}{2(3^t)+1}. \quad (21)$$

Proof: The proof relies on the facts that (a) ordering of the coefficients within the update relations does not matter in the convergence analysis, and (b) the order statistics of i.i.d. Unif-[0, 1]-distributed random variables are jointly uniform over the integration volume [6]. Hence, the average ICI is given by

$$E\{ICI_t^{(m)}\} = m! \int_0^1 \int_0^{c_1} \int_0^{c_2} \cdots \int_0^{c_{m-1}} \frac{\sum_{j=2}^m c_j^{2(3^t)}}{c_1^{2(3^t)}} dc_m dc_{m-1} \cdots dc_1 \quad (22)$$

which easily integrates to (21). Simulations corroborate this exact result.

Theorem 2 is not meant as a replacement for the more-general result in Theorem 1. Rather, it shows that the average exponential convergence behavior of the kurtosis-based FastICA algorithm holds for at least one other distribution of \mathbf{c}_0 than a uniform angular distribution. It has been our experience that (3) predicts the average behavior of the original single-unit FastICA algorithm with kurtosis contrast quite well, and to date, all theoretical results concerning the convergence performance of this algorithm reflect (3) in one form or another.

5 Simulations

To verify our theoretical results, simulations in MATLAB were carried out. Three and four-source mixtures have been generated, in which the sources are zero-mean unit-variance binary ($|\kappa| = 2$), uniform ($|\kappa| = 6/5$), and/or Laplacian ($|\kappa| = 3$) distributed. MATLAB's `randn` function was used to generate $M = 20000$ different initial coefficient vectors \mathbf{c}_0 for the numerical simulations. $N = 5000$ snapshots were used to evaluate the FastICA algorithm on sampled data.

Figure 1 shows plots of the averaged value of $ICI_t^{(3)}$ as predicted by simulations of the analytical convergence expressions of the FastICA algorithm, as determined by the FastICA algorithm on sampled data, and calculated from (14) for a three-source separation task with binary, uniform, and Laplacian sources. All of the curves agree quite well up to iteration $k = 5$. For $k > 5$, our simulation method for estimating $E\{ICI_t^{(3)}\}$ using non-uniform sampling of the unit sphere

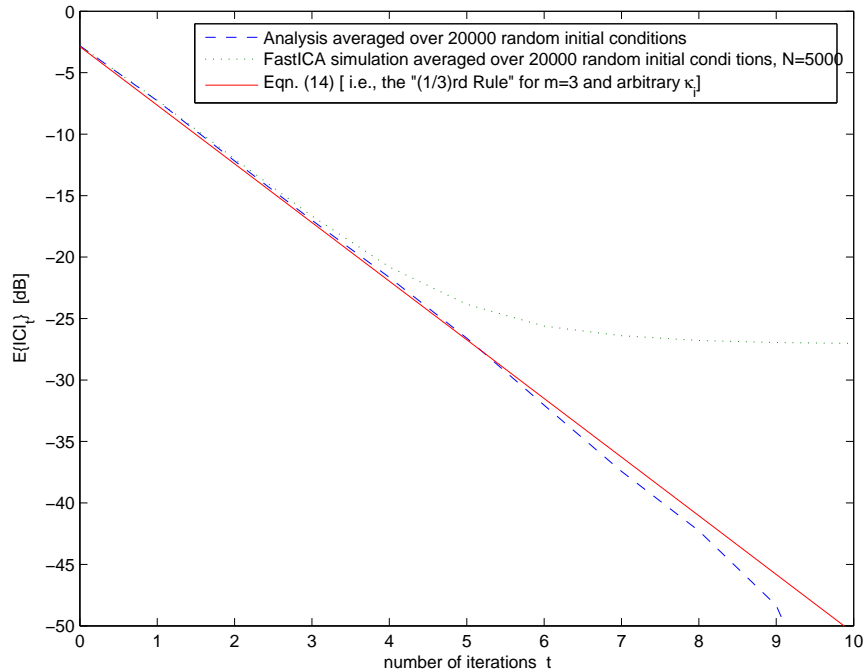


Fig. 1. Evolutions of the average ICI as determined by various methods, $m = 3$.

via the `randn` function is not accurate enough to verify (14). The scaling factor $g_3(2, 6/5, 3) = 131\sqrt{3}/(140\pi)$ is correct for the “(1/3)rd Rule” in this case.

Figure 2 shows plots of the averaged value of $ICI_t^{(4)}$ as predicted by analytical simulations, actual performance, and the prediction in (20) of FastICA convergence behavior for $m = 4$ binary sources. The scaling factor of $4/(\sqrt{3}\pi)$ is correct for the “(1/3)rd Rule” in the four-equal-kurtosis-source case.

6 Conclusions

In this paper, we analyze the average convergence behavior of the single-unit FastICA algorithm with kurtosis contrast on m -source mixtures, showing that its behavior is exponential with rate (1/3). Accurate expressions for $m = 3$ and $m = 4$ -source mixtures are provided, and simulations verify the analyses.

Appendix: Proof of Theorem 1

Consider a single-unit m -source FastICA algorithm with cubic nonlinearity. The combined system coefficient vector at iteration t is $\mathbf{c}_t = [\kappa_1^{\frac{a}{2}} x_1^{\frac{b}{2}} \cdots \kappa_m^{\frac{a}{2}} x_m^{\frac{b}{2}}]^T$, where $p = 2(3^t)$ and $q = 3^t - 1$. Consider the portion of the ICI at iteration t in which the first kurtosis component is being extracted, as given by

$$E\{ICI_{1,t}^{(m)}\} = \frac{2^m}{(2\pi)^{m/2}} \sum_{i=2}^m \int_0^\infty dx_1 \int_0^{a_2 x_1} dx_2 \cdots \int_0^{a_m x_1} dx_m \frac{(\frac{x_i}{a_i})^p}{x_1^p} \exp(-\frac{\sum_{i=1}^m x_i^2}{2}) \quad (23)$$

where

$$a_i = \left(\frac{\kappa_1}{\kappa_i}\right)^{\frac{a}{p}}, \quad i \in \{2, 3, \dots, m\}. \quad (24)$$

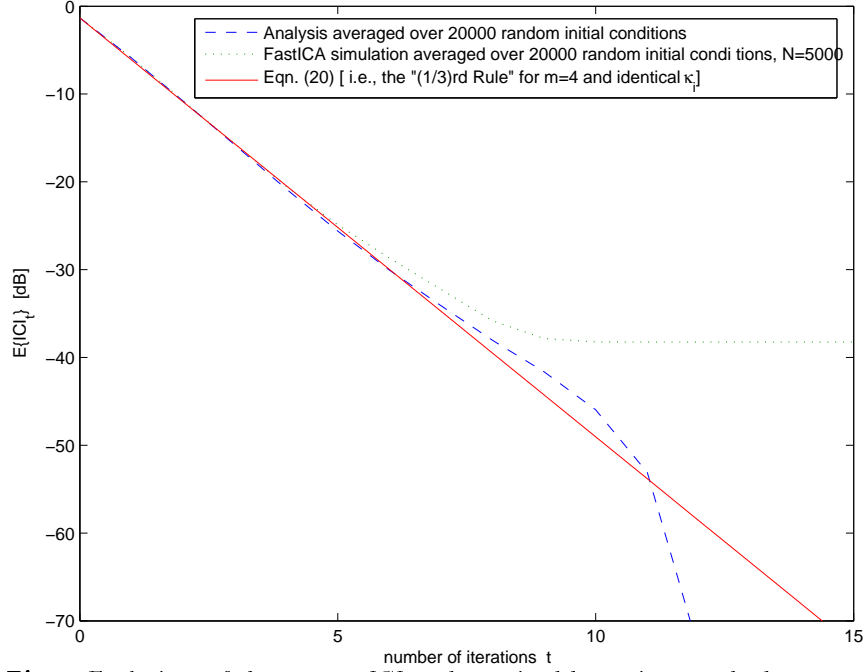


Fig. 2. Evolutions of the average ICI as determined by various methods, $m = 4$.

Using the transformation $\tilde{x}_i = x_i/x_1$ for $i \in \{2, \dots, m\}$, (23) becomes

$$E\{IC I_{1,t}^{(m)}\} = \frac{2^m}{(2\pi)^{m/2}} \sum_{i=2}^m \int_0^{a_2} d\tilde{x}_2 \cdots \int_0^{a_m} d\tilde{x}_m \int_0^\infty \left(\frac{\tilde{x}_i}{a_i}\right)^p x_1^{m-1} \exp\left(-\frac{x_1^2}{2}\left(1 + \sum_{i=2}^m \tilde{x}_i^2\right)\right) dx_1. \quad (25)$$

The most inside integral in (25) can be calculated as

$$\int_0^\infty x_1^{m-1} \exp\left(-\frac{x_1^2}{2}\left(1 + \sum_{i=2}^m x_i^2\right)\right) dx_1 = \begin{cases} \frac{(m-2)!!\sqrt{\pi}}{\sqrt{2}\left(1 + \sum_{i=2}^m x_i^2\right)^{m/2}} & m \text{ is odd} \\ \frac{[\frac{1}{2}(m-2)]!2^{(m-2)/2}}{\left(1 + \sum_{i=2}^m x_i^2\right)^{m/2}} & m \text{ is even} \end{cases}$$

where $(m-2)!! = 1 \cdot 3 \cdot 5 \cdot 7 \cdots (m-2)$. When $k \rightarrow \infty$, $\left(\frac{x_i}{a_i}\right)^p \rightarrow 0$ over the interval $0 \leq x_i < a_i$, and $a_i \rightarrow \sqrt{\frac{\kappa_1}{\kappa_i}} \equiv b_{1i}$. We can then approximate the integral

$$\int_0^{a_i} \left(\frac{x_i}{a_i}\right)^p \frac{1}{\left(1 + \sum_{j=2}^m x_j^2\right)^{m/2}} dx_i \approx \frac{1}{\left(1 + b_{1i}^2 + \sum_{j=2, j \neq i}^m x_j^2\right)^{m/2}} \frac{b_{1i}}{p}. \quad (26)$$

Noting that $p = 2(3^t)$, we have

$$E\{IC I_{1,t}^{(m)}\} \approx \begin{cases} \frac{1}{2(3)^t} \frac{2^m}{(2\pi)^{m/2}} \sum_{i=2}^m \int_0^{b_{12}} dx_2 \cdots \int_0^{b_{1m}} dx_m \frac{(m-2)!!\sqrt{\pi}}{\sqrt{2}\left(1 + b_{1i}^2 + \sum_{j=2, j \neq i}^m x_j^2\right)^{m/2}} & m \text{ is odd} \\ \frac{1}{2(3)^t} \frac{2^m}{(2\pi)^{m/2}} \sum_{i=2}^m \int_0^{b_{12}} dx_2 \cdots \int_0^{b_{1m}} dx_m \frac{[\frac{1}{2}(m-2)]!2^{(m-2)/2}}{\left(1 + b_{1i}^2 + \sum_{j=2, j \neq i}^m x_j^2\right)^{m/2}} & m \text{ is even} \end{cases}$$

Similarly, by invoking symmetry, we get

$$E\{ICI_{n,t}^{(m)}\} \approx \begin{cases} \frac{1}{2(3)^t} \frac{2^m}{(2\pi)^{m/2}} \sum_{i=1, i \neq n}^m \int_0^{b_{n1}} dx_1 \cdots \int_0^{b_{nm}} dx_m \frac{(m-2)!! \sqrt{\pi}}{\sqrt{2(1+(b_{ni})^2 + \sum_{j=1, j \neq n, j \neq i}^m x_j^2)^{m/2}}} & m \text{ is odd} \\ \frac{1}{2(3)^t} \frac{2^m}{(2\pi)^{m/2}} \sum_{i=1, i \neq n}^m \int_0^{b_{n1}} dx_1 \cdots \int_0^{b_{nm}} dx_m \frac{[\frac{1}{2}(m-2)]! 2^{(m-2)/2}}{(1+(b_{ni})^2 + \sum_{j=1, j \neq n, j \neq i}^m x_j^2)^{m/2}} & m \text{ is even} \end{cases}$$

for $n \in \{2, \dots, m\}$ and

$$b_{ni} = \sqrt{\frac{\kappa_n}{\kappa_i}}. \quad (27)$$

Finally, the average ICI is $E\{ICI_t^{(m)}\} = \sum_{n=1}^m E\{ICI_{n,t}^{(m)}\}$, which results in (17) with

$$g(\kappa_1, \dots, \kappa_m) = \begin{cases} \frac{2^{m-1}}{(2\pi)^{m/2}} \sum_{n=1}^m \sum_{i=1, i \neq n}^m \int_0^{b_{n1}} dx_1 \cdots \int_0^{b_{nm}} dx_m \frac{(m-2)!! \sqrt{\pi}}{\sqrt{2(1+(b_{ni})^2 + \sum_{j=1, j \neq n, j \neq i}^m x_j^2)^{m/2}}} & m \text{ is odd} \\ \frac{2^{m-1}}{(2\pi)^{m/2}} \sum_{n=1}^m \sum_{i=1, i \neq n}^m \int_0^{b_{n1}} dx_1 \cdots \int_0^{b_{nm}} dx_m \frac{[\frac{1}{2}(m-2)]! 2^{(m-2)/2}}{(1+(b_{ni})^2 + \sum_{j=1, j \neq n, j \neq i}^m x_j^2)^{m/2}} & m \text{ is even} \end{cases}$$

Additional calculations show that the error introduced in (17) by (26) is

$$|R(t, \kappa_1, \dots, \kappa_m)| \leq c(m) \left(\frac{1}{3}\right)^t \sum_{n=1}^m \left[(m-1) \prod_{i=2}^m b_{ni} \left(1 - \prod_{i=2}^m b_{ni}^{-\frac{2}{p}}\right) + \sum_{i=2}^m \left(\left(1 - \frac{1}{\sqrt{p}}\right)^{p+1} b_{ni}^{1-\frac{2}{p}} + \frac{b_{ni}}{p} \left|1 - \frac{p}{p+1} b_{ni}^{-\frac{4}{p}}\right| \right) \prod_{i=2}^m b_{ni}^{1-\frac{2}{p}} \right] \quad (28)$$

Since $1 - \prod_{i=2}^m b_{ni}^{-\frac{2}{p}} \rightarrow 0$ and $\left(1 - \frac{1}{\sqrt{p}}\right)^{p+1} \rightarrow 0$ as $t \rightarrow \infty$, it is easy to see that

$$\lim_{t \rightarrow \infty} \frac{R(t, \kappa_1, \dots, \kappa_m)}{\left(\frac{1}{3}\right)^t} = 0. \quad (29)$$

References

1. A. Hyvärinen and E. Oja, "A fast fixed-point algorithm for independent component analysis," *Neural Computation*, vol. 9, no. 7, pp. 1483-1492, Oct. 1997.
2. A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis* (New York: Wiley, 2001).
3. S.C. Douglas, "On the convergence behavior of the FastICA algorithm," *Proc. Fourth Symp. Indep. Compon. Anal. Blind Signal Separation*, Kyoto, Japan, pp. 409-414, Apr. 2003.
4. S.C. Douglas, "A statistical convergence analysis of the FastICA algorithm for two-source mixtures," *Proc. 39th Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, Oct. 2005.
5. S.C. Douglas, "Relationships between the FastICA algorithm and the Rayleigh Quotient Iteration," to appear in *Proc. 6th Int. Conf. Indep. Compon. Anal. Blind Source Separation*, Charleston, SC, Mar. 2006.
6. H.A. David, *Order Statistics*, 2nd ed. (New York: Wiley, 1980).