Adaptive Algorithms
for
Blind Separation
of
Dependent Sources

George V. Moustakides
INRIA, Sigma 2

Outline

- Problem definition–Motivation
- Existing adaptive scheme–Independence
- General adaptive scheme–Dependence
  - Imposing the desired limits
  - Imposing symmetric behavior
  - Non-separable sources
- Performance measure
- Optimum algorithms
- Conclusion

Problem Definition-Motivation

$$\begin{bmatrix} x_1(n) \\ x_2(n) \\ \vdots \\ x_L(n) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1L} \\ a_{21} & a_{22} & \cdots & a_{2L} \\ \vdots & \vdots & \vdots & \vdots \\ a_{L1} & a_{L2} & \cdots & a_{LL} \end{bmatrix} \begin{bmatrix} s_1(n) \\ s_2(n) \\ \vdots \\ s_L(n) \end{bmatrix}$$

$$X(n) = \mathbf{A}S(n)$$

If the observation sequence $X(n)$ is available, estimate the source sequence $S(n)$. Equivalently estimate $\mathbf{B} = \mathbf{A}^{-1}$.

Interested in adaptive techniques. At every time n, available $X(n)$ and form an estimate $\mathbf{B}(n)$ of $\mathbf{A}^{-1}$.

Why Dependent Sources ?

- Because it is analytically challenging.
- Here is a "practical reason" to motivate such an analysis

$$f(s_1,\ldots,s_L) = 0.99\prod_{i=1}^{L} f_i(s_i) + 0.01\prod_{i=1}^{L} g_i(s_i)$$

ε-contamination model

$f(s_1,\ldots,s_L)$ no longer corresponds to independent sources!!

Question: Do BSS algorithms still work under such mild divergence from the independence assumption?

Existing adaptive technique

Consider $L = 2$ sources and let $\mathbf{B}(n)$ denote the estimate of $\mathbf{A}^{-1}$ at time n. At time n available: $\mathbf{B}(n\text{-}1)$, $X(n)$

$$\hat{S}(n) = \mathbf{B}(n-1)X(n)$$

$$\mathbf{B}(n) = \mathbf{B}(n-1) - \mu\mathbf{H}\left(\hat{S}(n)\right)\mathbf{B}(n-1), \mathbf{B}(0) = \mathbf{I}$$

μ is a constant (step size) with $0 < \mu << 1$.

$$\mathbf{H}(Z) = \begin{bmatrix} z_1^2 - 1 & z_1 z_2 \\ z_1 z_2 & z_2^2 - 1 \end{bmatrix} + \begin{bmatrix} 0 & g_1(z_1)z_2 - g_2(z_2)z_1 \\ g_1(z_2)z_1 - g_2(z_1)z_2 & 0 \end{bmatrix}$$

$g_i(z)$ are odd univariate nonlinear function (i.e. $g_i(-z) = -g_i(z)$).

Acceptable solution(s)

$$\hat{S}(n) = \mathbf{B}(n-1)X(n) = \mathbf{B}(n-1)\mathbf{A}S(n) = \mathbf{C}(n-1)S(n)$$

$$\mathbf{C}(n) \rightarrow \begin{bmatrix} \pm c_1 & 0 \\ 0 & \pm c_2 \end{bmatrix} \quad \hat{S}(n) \rightarrow \begin{bmatrix} \pm c_1 s_1(n) \\ \pm c_2 s_2(n) \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 0 & \pm c_2 \\ \pm c_1 & 0 \end{bmatrix} \qquad \rightarrow \begin{bmatrix} \pm c_2 s_2(n) \\ \pm c_1 s_1(n) \end{bmatrix},$$

Does $\mathbf{C}(n) = \mathbf{B}(n)\mathbf{A}$ converge to one of the eight matrices ?

$$\begin{bmatrix} \pm c_1 & 0 \\ 0 & \pm c_2 \end{bmatrix} \text{ or } \begin{bmatrix} 0 & \pm c_2 \\ \pm c_1 & 0 \end{bmatrix} ?$$

Theorem 1: If the sources $s_i(n)$ are independent random variables, with symmetric probability density functions; and at most one is Gaussian, then

- $$\lim_{n\to\infty} E\{\mathbf{C}(n)\} = \begin{bmatrix} \pm c_1 & 0 \\ 0 & \pm c_2 \end{bmatrix} \text{or} \begin{bmatrix} 0 & \pm c_2 \\ \pm c_1 & 0 \end{bmatrix}$$

- $$\lim_{n\to\infty} E\{(\mathbf{C}(n)\text{-}E\{\mathbf{C}(\infty)\})^2\} = O(\mu)$$

Observations

- No knowledge of the exact pdfs is required; only symmetry (and independence (?)).

$$\mathbf{H}(z_1,z_2)=\begin{bmatrix} z_1^2-1 & z_1z_2 \\ z_1z_2 & z_2^2-1 \end{bmatrix}$$

$$+\begin{bmatrix} 0 & g_1(z_1)z_2-g_2(z_2)z_1 \\ -g_1(z_1)z_2+g_2(z_2)z_1 & 0 \end{bmatrix}$$

$g_i(z)=\alpha_i z^3$ common choice

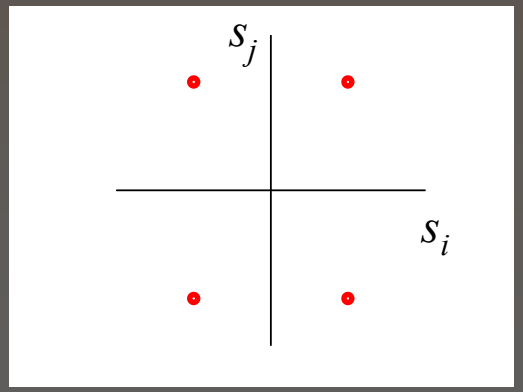
- Two or more Gaussian sources are NOT separable, in the sense that for ANY adaptive algorithm there is a continuum of limits instead of the 8 desirable ones.

Dependent Source Statistics

- If sources are dependent and the joint pdf is known, in principle we can construct adaptive algorithms that solve the source separation problem.
- We are interested in dependent sources for which it is not required knowledge of the pdfs.

All bivariate densities must satisfy the following quadratal symmetry

$$f_{ij}(s_i, s_j) = f_{ij}(-s_i, s_j) = f_{ij}(s_i, -s_j)$$



$$f(s_1, \ldots, s_L) = 0.99 \prod_{i=1}^{L} f_i(s_i) + 0.01 \prod_{i=1}^{L} g_i(s_i)$$

General adaptive scheme

$$\hat{S}(n) = \mathbf{B}(n-1)X(n), \qquad X(n) = \mathbf{A}S(n)$$

$$\mathbf{B}(n) = \mathbf{B}(n-1) - \mu\mathbf{H}\left(\hat{S}(n)\right)\mathbf{B}(n-1), \mathbf{B}(0) = \mathbf{I}$$

$$\mathbf{H}(z_1, z_2) = \begin{bmatrix} h(z_1) & q(z_2, z_1) \\ q(z_1, z_2) & h(z_2) \end{bmatrix}$$

$$\hat{S}(n) = \mathbf{C}(n-1)S(n), \qquad \mathbf{C}(n) = \mathbf{B}(n)\mathbf{A}$$

$$\mathbf{C}(n) = \mathbf{C}(n-1) - \mu\mathbf{H}\left(\hat{S}(n)\right)\mathbf{C}(n-1), \mathbf{C}(0) = \mathbf{A}$$

$$\mathbf{C}(n) = \mathbf{C}(n-1) - \mu\mathbf{H}\left(\mathbf{C}(n-1)S(n)\right)\mathbf{C}(n-1)$$

Imposing the desired limits

From the theory of adaptive algorithms we know that an algorithm of the form

$$\mathbf{C}(n) = \mathbf{C}(n-1) - \mu \mathbf{H}\left(\mathbf{C}(n-1)S(n)\right)\mathbf{C}(n-1)$$

converges in the mean $\lim_{n\to\infty} E\{\mathbf{C}(n)\} = \mathbf{C}$, where $\mathbf{C}$ is an equilibrium matrix satisfying the equation

$$\mathbf{C} = \mathbf{C} - \mu E\left\{\mathbf{H}\left(\mathbf{C}S(n)\right)\right\}\mathbf{C}$$

or equivalently $E\left\{\mathbf{H}\left(\mathbf{C}S(n)\right)\right\} = 0$

Since

$$\mathbf{H}(z_1, z_2) = \begin{bmatrix} h(z_1) & q(z_2, z_1) \\ q(z_1, z_2) & h(z_2) \end{bmatrix}$$

relation $E\left\{\mathbf{H}\left(\mathbf{C}S(n)\right)\right\} = 0$

is equivalent to a system of four (in general nonlinear) equations

$$E\left\{h\left(\mathbf{C}S(n)\right)\right\} = E\left\{q\left(\mathbf{C}S(n)\right)\right\} = 0$$

in four unknowns (the four elements of matrix $\mathbf{C}$) that can be solved to identify the equilibrium matrices.

If we want a specific matrix $\mathbf{C}_0$ to become an equilibrium, then the elements of $\mathbf{H}(Z)$ must satisfy

$$E\left\{h\left(\mathbf{C}_0 S(n)\right)\right\} = E\left\{q\left(\mathbf{C}_0 S(n)\right)\right\} = 0$$

If in particular we like the following eight matrices

$$\mathbf{C}_0 = \begin{bmatrix} \pm c_1 & 0 \\ 0 & \pm c_2 \end{bmatrix}, \begin{bmatrix} 0 & \pm c_2 \\ \pm c_1 & 0 \end{bmatrix}$$

to become equilibriums, then

$$E\left\{h\left(\pm c_i s_i\right)\right\} = E\left\{q\left(\pm c_i s_i, \pm c_j s_j\right)\right\} = 0$$

Imposing symmetric behavior

For every trajectory converging to one limit we require existence of symmetric trajectories converging to the other seven limits with the same probability.

$$\mathbf{H}(z_1, z_2) = \begin{bmatrix} h(z_1) & q(z_2, z_1) \\ q(z_1, z_2) & h(z_2) \end{bmatrix}$$

$$h(-z) = h(z)$$

$$q(-z_1, z_2) = q(z_1, -z_2) = -q(z_1, z_2)$$

$$E\{h(c_1 s_1)\} = E\{h(c_2 s_2)\} = 0$$

$$E\{q(c_1 s_1, c_2 s_2)\} = E\{q(c_2 s_2, c_1 s_1)\} = 0 \text{ (for free)}$$

Necessary conditions

1. $\mathbf{H}(z_1, z_2) = \begin{bmatrix} h(z_1) & q(z_2, z_1) \\ q(z_1, z_2) & h(z_2) \end{bmatrix}$
2. $h(-z) = h(z)$
3. $q(-z_1, z_2) = q(z_1, -z_2) = -q(z_1, z_2)$
4. $E\{h(c_1 s_1)\} = E\{h(c_2 s_2)\} = 0$ (existence of c_1, c_2)

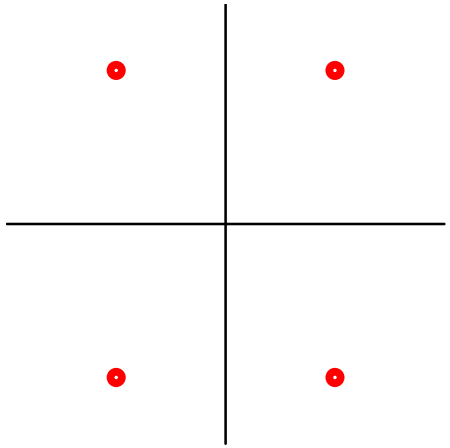
$$\mathbf{H}(z_1, z_2) = \begin{bmatrix} z_1^2 - 1 & z_1 z_2 + g(z_1) z_2 - g(z_2) z_1 \\ z_1 z_2 - g(z_1) z_2 + g(z_2) z_1 & z_2^2 - 1 \end{bmatrix}$$

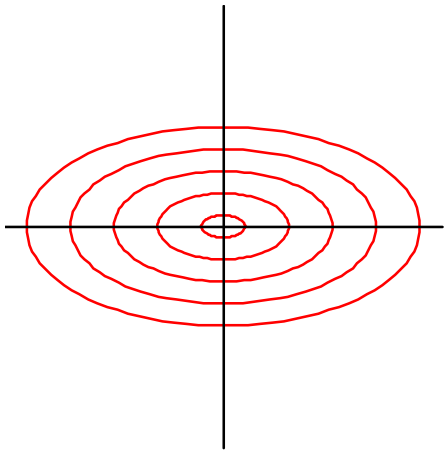
Theorem 2: If the sources $s_i(n)$ are random variables, with symmetric bivariate densities and no bivariate density is of the form

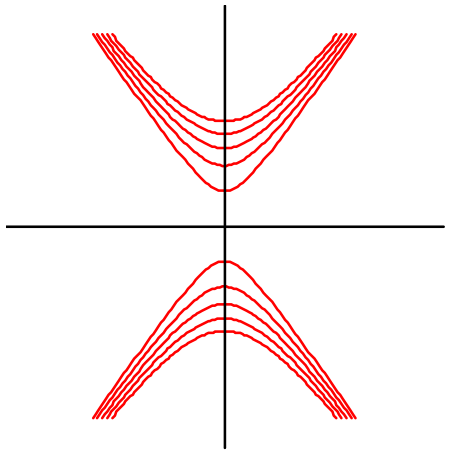
$$f_{ij}(s_i, s_j) = \beta(K_i s_i^2 + K_j s_j^2)$$

that is, no bivariate density exhibits either elliptical or hyperbolic symmetry (with respect to the two axis), then the generalized algorithm has exactly the same convergence properties for the dependent class, as the existing algorithm for the independent class.

Quadrantally symmetric bivariate pdfs







Example

Existing scheme

$$\mathbf{H}(z_1, z_2) = \begin{bmatrix} z_1^2 - 1 & z_1 z_2 + g(z_1) z_2 - g(z_2) z_1 \\ z_1 z_2 - g(z_1) z_2 + g(z_2) z_1 & z_2^2 - 1 \end{bmatrix}$$

Select

$$\mathbf{H}(z_1, z_2) = \begin{bmatrix} |z_1| - 1 & z_1 z_2^2 \operatorname{sgn}(z_2) \\ z_2 z_1^2 \operatorname{sgn}(z_1) & |z_2| - 1 \end{bmatrix}$$

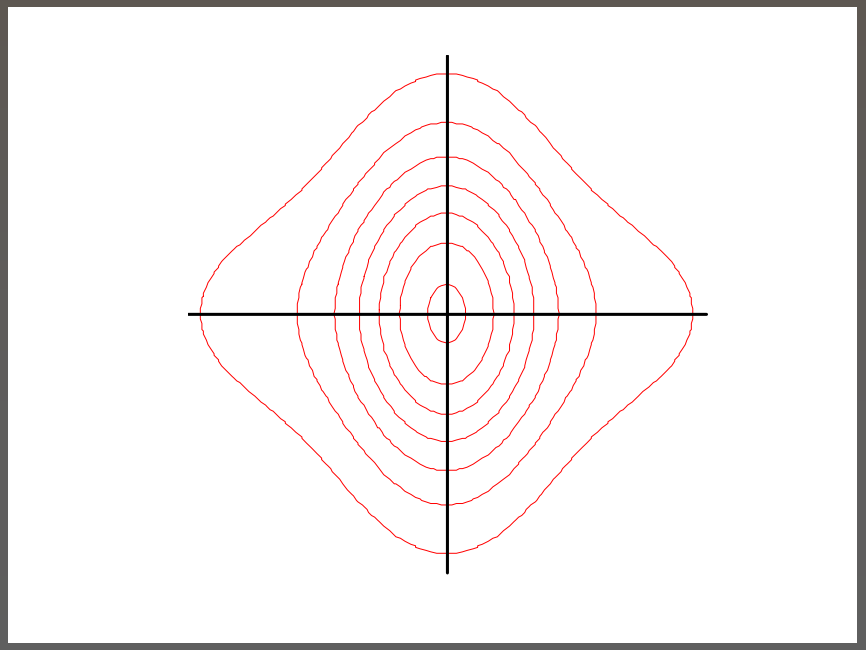
There is no whitening involved!!!

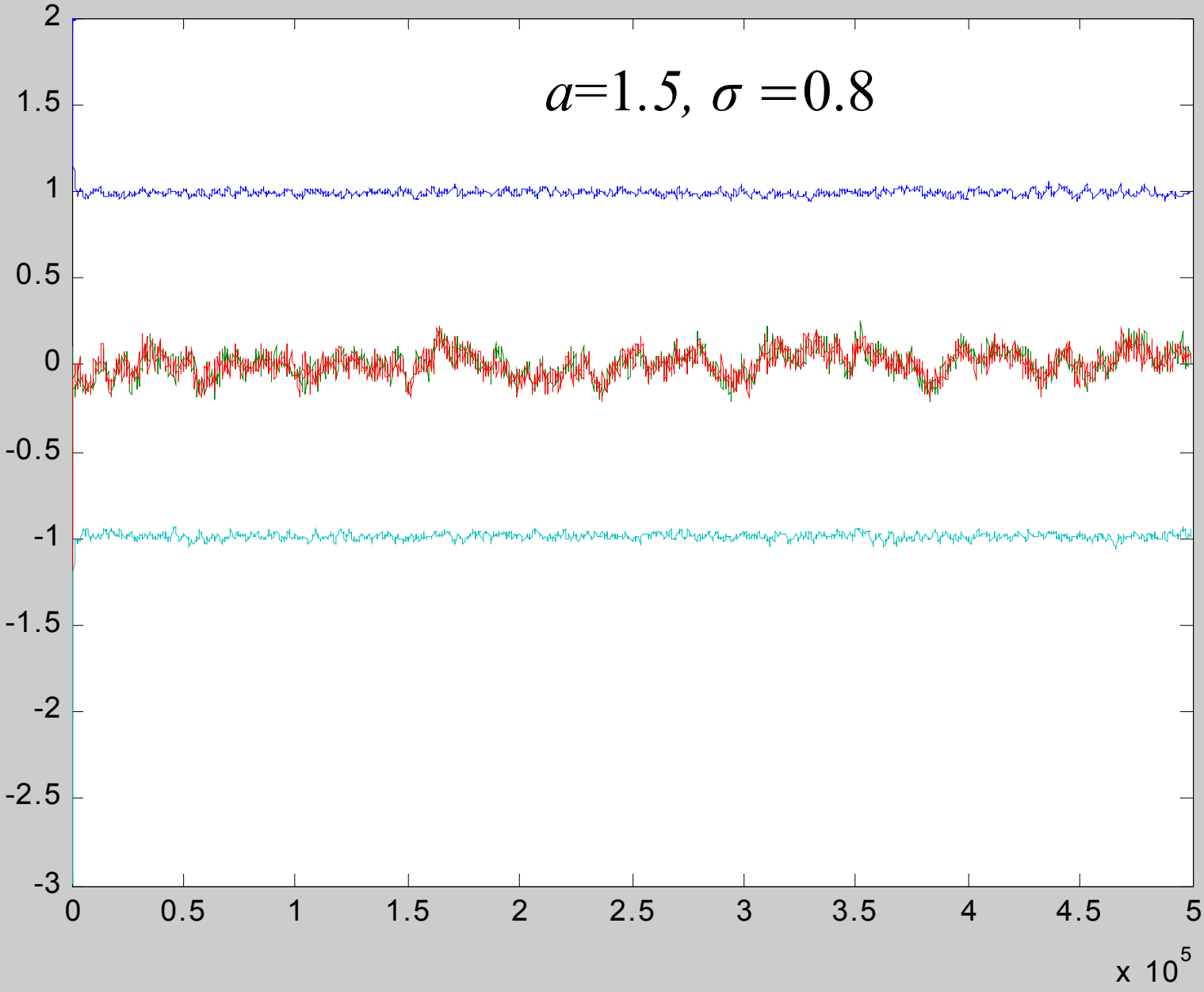
$$f(s_1,s_2)=0.5\frac{e^{-s_1^2/2}}{\sqrt{2\pi}}\frac{e^{-s_2^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}}+0.5\frac{e^{-s_1^2/2a^2\sigma^2}}{\sqrt{2\pi a^2\sigma^2}}\frac{e^{-s_2^2/2a^2}}{\sqrt{2\pi a^2}}$$

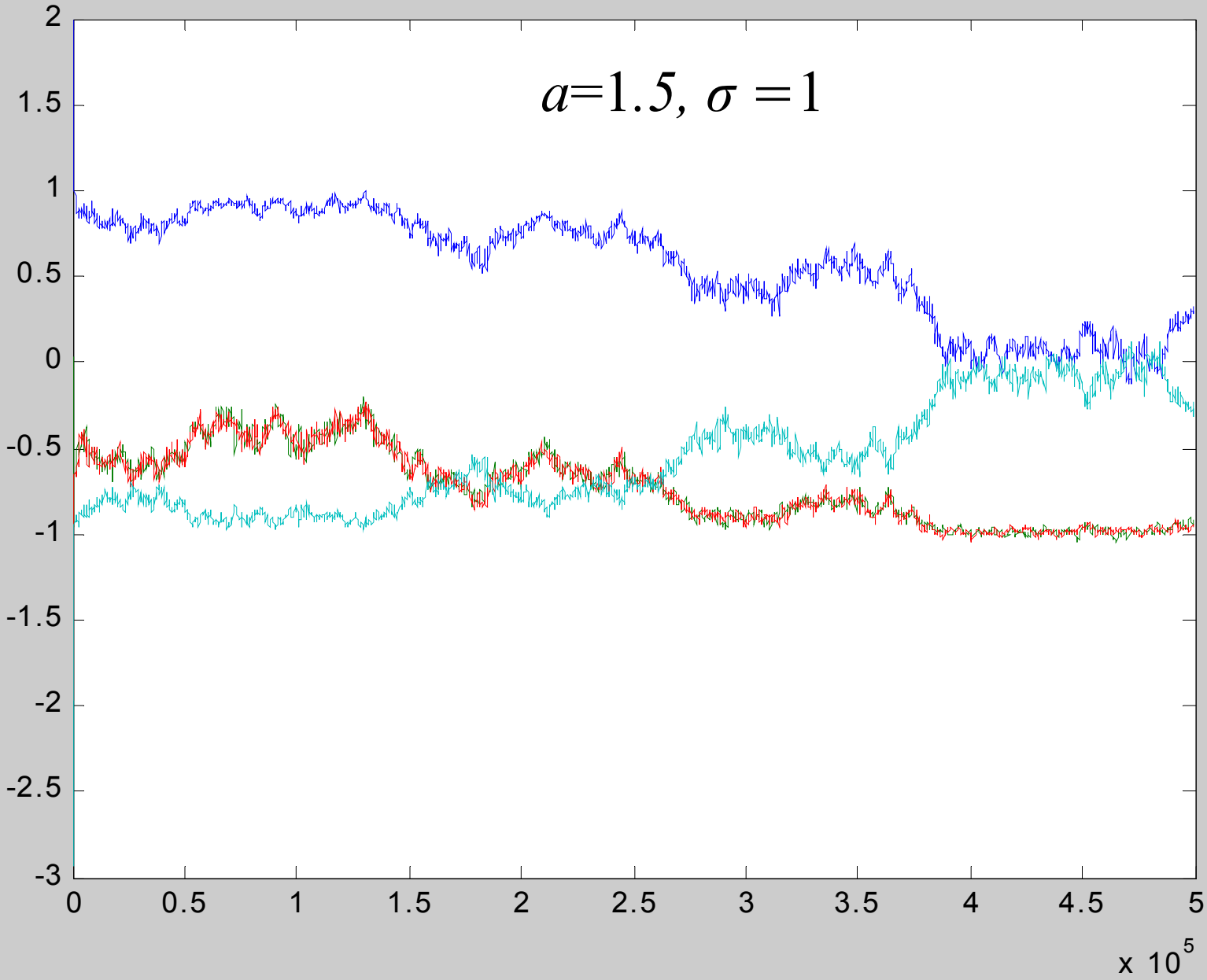
If $\sigma \neq 1$ the two sources are separable

If $\sigma = 1$ the two sources are not separable

a=1.5, σ =0.8




a=1.5, σ =0.8
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
-3
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
x 10^5


a=1.5, σ =1
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
-3
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
x 10^5

Performance measure

Define a performance measure for the algorithmic class

$$\mathbf{C}(n) = \mathbf{C}(n-1) - \mu\mathbf{H}\left(\mathbf{C}(n-1)S(n)\right)\mathbf{C}(n-1)$$

with $\mathbf{H}(z_1, z_2)$ satisfying the five necessary conditions.

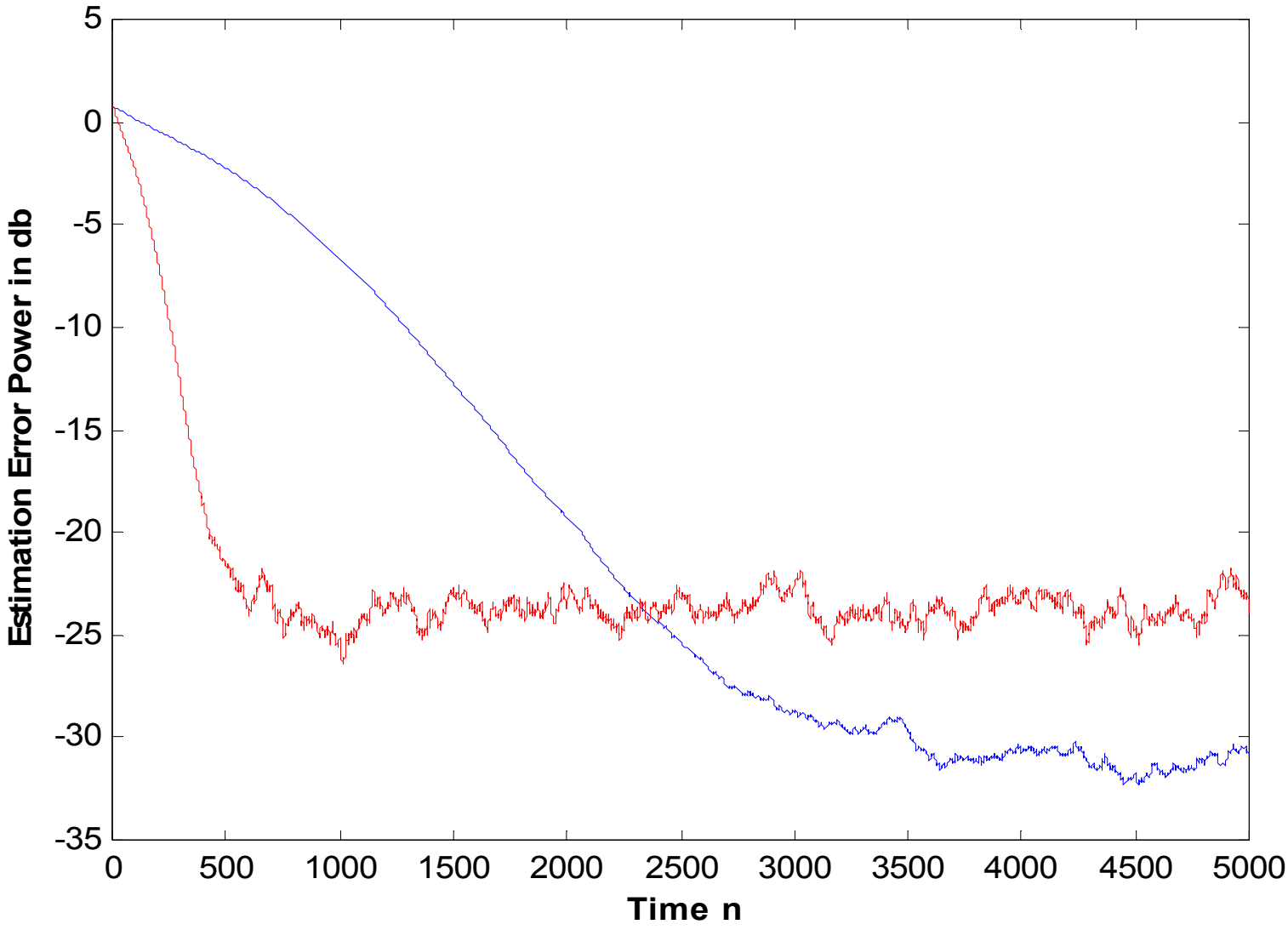
For performance there are two important quantities:

- Rate of convergence
- Steady state estimation error power

Both quantities depend on μ.

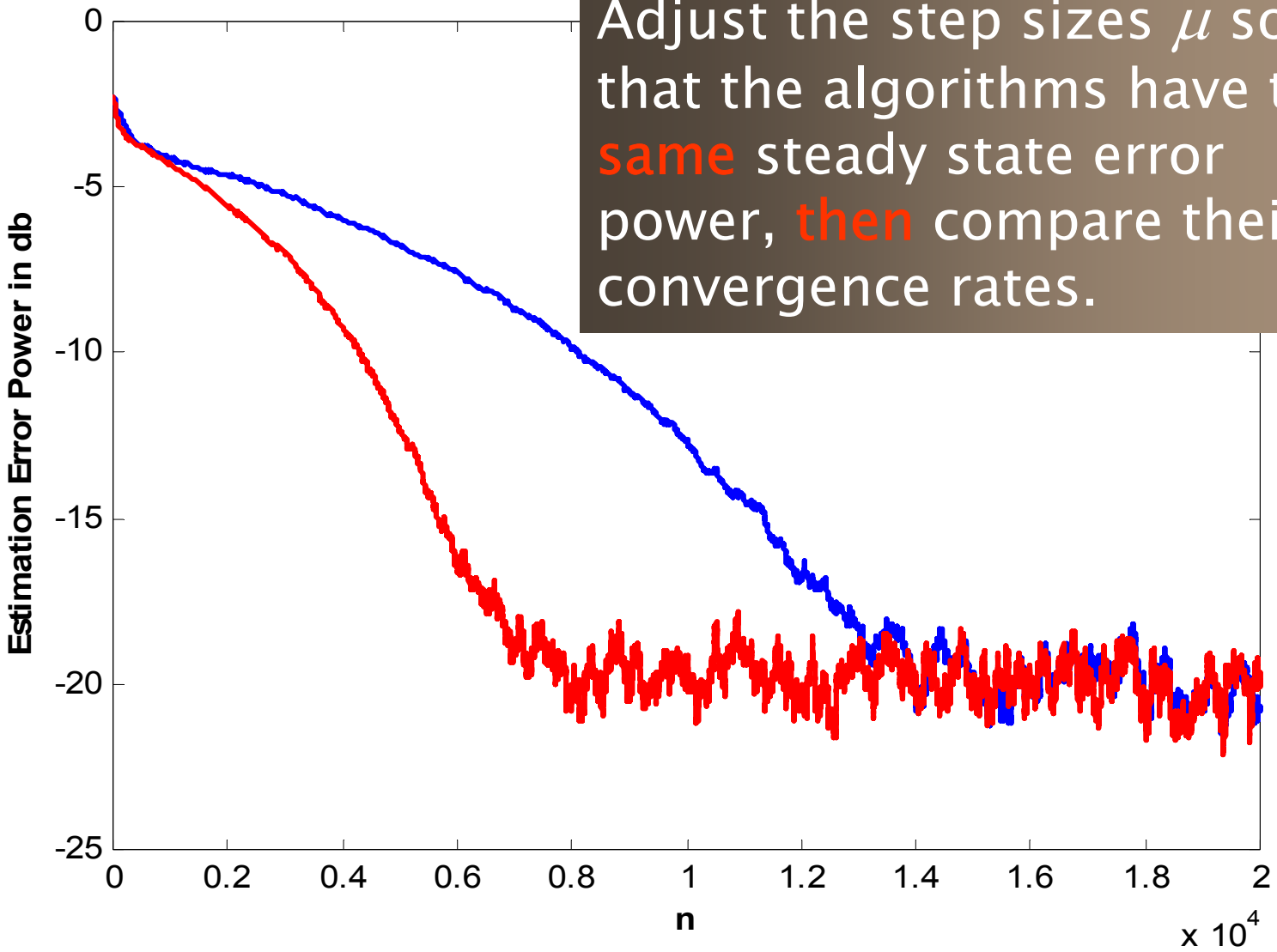
Combine the two quantities to define a fair performance measure.

$$\left\|\mathbf{C}(n)-E\left\{\mathbf{C}(\infty)\right\}\right\|^2=\sum_{i,j}\left(\mathbf{C}_{ij}(n)-E\left\{\mathbf{C}_{ij}(\infty)\right\}\right)^2$$



$$\left\|\mathbf{C}(n)-E\left\{\mathbf{C}(\infty)\right\}\right\|^2=\sum_{i,j}\left(\mathbf{C}_{ij}(n)-E\left\{\mathbf{C}_{ij}(\infty)\right\}\right)^2$$

Adjust the step sizes μ so that the algorithms have the same steady state error power, then compare their convergence rates.



It is possible to define an analytic performance measure that can characterize an adaptive algorithm.

$$\text{Efficacy} = \text{function}\left(h(z), q(z_1, z_2), f_{ij}(s_i, s_j)\right)$$

The important properties of the Efficacy are the following

- When comparing Efficacies of two algorithms, it is equivalent to comparing the two algorithms fairly!!
- The optimum algorithm is the one that maximizes the Efficacy over $h(z)$, $q(z_1,z_2)$.

Optimum adaptive algorithms

When the bivariate pdfs are different the optimization problem is still open.

When the sources are **independent** with the same pdf $f(z)$, then

$$h(z_1) = a\left(z_1 l(z_1) - 1\right)$$

$$q(z_1, z_2) = b z_1 l(z_2) + c z_2 l(z_1)$$

$$l(z) = -\frac{f'(z)}{f(z)},$$

where a, b, c, proper constants

Existing scheme

$$h(z_1) = z_1^2 - 1, \qquad q(z_1, z_2) = z_1 z_2 + \left(z_1 g(z_2) - z_2 g(z_1)\right)$$

Maximize Efficacy with respect to $g(z)$

$$q(z_1, z_2) = z_1 z_2 + d\left(z_1 l(z_2) - z_2 l(z_1)\right)$$

Conclusion

- We have presented a rich class of adaptive algorithms that can solve the blind source separation problem for dependent sources with symmetries.
- We have identified the sources that cannot be separated with the specific combination of source statistics and algorithmic model.
- We have proposed an analytic performance measure that can be optimized to yield the fastest converging algorithm in the equi-distributed source case.

LABORATOIRE
Annuaire
Contacts & Accés
Intranet

RECHERCHE
Géomérie & Images
Modèles et Algorithmes Déterministes
Statistique
Séminaires & Colloquiums

RESSOURCES
Publications
Logiciels
Galerie
Plateforme **MIRAGE**
Offres d'Emplois

LIENS
Ecole Doctorale
UFR-**IMA**
ENSIMAG
INRIA Rhone-Alpes

UMR 5224

SÉMINAIRES DU LJK-STATISTIQUE

Ce séminaire est répertorié dans ACM

Les séminaires ont lieu le Jeudi à la tour IRMA en salle 1 à 14h00

Années 2002 2003 2004 2005 2006 2007 2008

(Téléchargement au format tex)

Séminaires 2008

Jeudi 31 Janvier — **Céline LACAUX (Institut Elie Cartan, Nancy)**

Champs stables à autosimilarité matricielle

Jeudi 24 Janvier — **Michæl BLUM (TIMB / TIMC)**

Méthodes bayésiennes sans vraisemblance

Jeudi 17 Janvier — **Clémentine PRIEUR (INSA Toulouse)**

Mesures de dépendance pour certains systèmes dynamiques

Jeudi 10 Janvier — **Mr Emmanuel TROUVÉ (LISTIC, Université de Savoie)**

Analyse d'images RSO interférométriques et polarimétriques. Application à l'observation des glaciers tempérés

Séminaires 2007

Jeudi 13 Décembre — **Marc HOFFMAN (Université Marne-la-Vallée)**

Estimation multi-échelle de la volatilié et bruit de microstructure

Jeudi 6 Décembre — **Alexander NAZIN (Institute of Control Sciences RAS, Moscow)**

Application of mirror descent method to a stochastic multi-armed bandit problem

Jeudi 29 Novembre — **Olivier RENAUD (Université de Genève)**

Tests simultanés dans le plan temps-fréquence: le cas de signaux d'électro-encéphalogrammes dans la recherche en psychologie

Jeudi 27 Mars
Amphithéâtre - Maison Jean Kuntzmann à 8h45

JOURNEE ONDELETTES

Ondelettes et Images

Jeudi 20 Mars

YANNICK LEFEBVRE (Université de Marne-la-Vallée, EDF Rpt. MRI)

Etude d'un système vieillissant: de la modélisation au calcul de disponiblité

Jeudi 13 Mars

Oleg LEPSKI (Université d'Aix-Marseille)

Problèmes statistiques pour des fonctions multidimensionnelles

Jeudi 6 Mars

Jean-Marc VESIN (EPFL)

Les défis en traitement des signaux biomédicaux

Jeudi 20 Février

P. LAFAYE DE MICHEAUX (Université de Montréal)

A multivariate empirical characteristic function test of independence with normal marginals

Jeudi 13 Février

C. VIAL (CREST/ENSAI)

Test d'une direction révélatrice :une approche géométrique

Jeudi 6 Février

L. GARDES (Université Montpellier 2)

Utilisation de seuils aléatoires pour l'estimation de quantiles extrêmes

Jeudi 30 Janvier

A. TSYBAKOV (Université Paris 6)

Classification non-paramétrique adaptative

Jeudi 23 Janvier

M. LAVIELLE (Universités Descartes & Paris Sud)

Estimation non paramétrique de processus autorégressifs non linéaires sous des contraintes dynamiques

Jeudi 16 Janvier

M. GARRIDO (LabSAD, Université Grenoble 2)

Une nouvelle approche bayésienne pour l'estimation des paramètres d'une loi de Pareto généralisée, application à l'estimation de quantiles extremes

Jeudi 9 Janvier

N. CHOPIN (CREST, Paris)

Applications des méthodes de Monte Carlo séquentielles à la statistique bayésienne

Séminaires 2002

Jeudi 19 Décembre

A. PHAM (SMS/LMC/IMAG)

	Jeudi 21 Février	**Frédérique LETUE (LABSAD-UPMF)** *Estimation de la fonction de régression de Cox par sélection de modèle*
	Jeudi 7 Février	**Bruno JEDYNAK (Université Lille 1)** *Grandes déviations et performances maximales en reconnaissance de formes*
	Jeudi 31 Janvier	**Yann GUERMEUR (LORIA Vandoeuvre-les-Nancy)** *Machines à vecteurs de support multiclasses*
	Jeudi 24 Janvier	**Antoine CHAMBAZ (Université Paris-Sud, Orsay)** *Détection de ruptures multidimensionnelles : théorie et application*
	Jeudi 17 Janvier	**Georges MOUSTAKIDES (IRISA, Rennes)** *Adaptive Algorithms for Blind Separation of Dependent Sources*
	Jeudi 10 Janvier	**Jean-François CARDOSO (Dépt. TSI, ENST, Paris)** *Séparation de sources en imagerie astronomique*
	Samedi 5 Janvier	**V. RIVOIRARD (Université Paris 6)** *Approche maxiset et estimation bayésienne sur les espaces de Besov faibles.*

contact: Webmaster