Nonparametric Independent Process Analysis

Zoltán Szabó
Faculty of Informatics, Eötvös Loránd University,
Pázmány Péter sétány 1/C, Budapest, H-1117, Hungary
email: `szzoli@cs.elte.hu`
web: `http://nipg.inf.elte.hu`

Barnabás Póczos
School of Computer Science, Carnegie Mellon University,
5000 Forbes Ave, 15213, Pittsburgh, PA, USA
email: `bapoczos@cs.cmu.edu`
web: `http://www.autonlab.org`

1. Introduction

- Linear dynamical systems (LDS): $\mathbf{s}_t = \mathbf{F}\mathbf{s}_{t-1} + \boldsymbol{\mu}_t$, $\mathbf{x}_t = \mathbf{A}\mathbf{s}_t + \boldsymbol{\nu}_t$
  - Limitations: linear dynamics, Gaussian driving noise.
- Non-Gaussian driving noises:
  - ICA = separation of mixed non-Gaussian, one-dimensional sources: ✓
  - ISA = ICA with multidimensional sources. $\mathbf{x}_t = \mathbf{A}\mathbf{s}_t$.
    - Limitations: Unknown, nonparametric dynamics is hardly touched:
      - stationary + ergodic sources, constrained mixing.
      - block-decorrelatedness for all time-shifts.
      - Additionally both assume: known and equal component dimensions.
- Our **contributions**:
  - ISA with nonparametric, asymptotically stationary dynamics.
  - unknown and possibly different dimensional components.
  - simple separation based solution: kernel regression + ISA.

2. Problem

- Task: estimate linearly mixed ($\mathbf{A}$), multidimensional sources ($\mathbf{s}$) of unknown functional autoregressive (fAR) dynamics ($\mathbf{f}$) with independent driving noises ($\mathbf{e}$)

$$\mathbf{s}_t = \mathbf{f}(\mathbf{s}_{t-1}, \ldots, \mathbf{s}_{t-L_s}) + \mathbf{e}_t, \quad (1)$$
$$\mathbf{x}_t = \mathbf{A}\mathbf{s}_t. \quad (2)$$

- Assumptions: $\mathbf{A}$: full column rank; $\mathbf{e} = [\mathbf{e}^1; \ldots; \mathbf{e}^M]$ ($\mathbf{e}^m \in \mathbb{R}^{d_m}$): traditional ISA.
- Goal (fAR-IPA): estimate $\mathbf{A}$ and $\mathbf{s}_t$ by using observations $\mathbf{x}_t$ only.
- Special cases:
  - if $\mathbf{f}$ were known, linear: autoregressive IPA (AR-IPA).
  - if order $L_s = 0$: traditional ISA.
  - ISA with one-dimensional independent subspaces ($d_m = 1, \forall m$): ICA.

3. Method

- In the ISA special case: ISA separation principle
  - ISA = ICA up to permutation – conjecture of Cardoso ('98).
  - recently has been proved – sufficient conditions (Szabó et al., JMLR 2007).
- We derive a similar reduction scheme for the fAR-IPA problem:

fAR-IPA = fAR identification + ISA.

- According to (1)-(2) $\mathbf{x}_t$ is fAR with innovation $\mathbf{n}_t = \mathbf{A}\mathbf{e}_t$:

$$\mathbf{x}_t = \mathbf{A}\mathbf{f}(\mathbf{A}^{-1}\mathbf{x}_{t-1}, \ldots, \mathbf{A}^{-1}\mathbf{x}_{t-L_s}) + \mathbf{A}\mathbf{e}_t = \mathbf{g}(\mathbf{x}_{t-1}, \ldots, \mathbf{x}_{t-L_s}) + \mathbf{n}_t. \quad (3)$$

Idea: (3) = nonparametric regression problem $\xrightarrow{\text{e.g., Nadaraya-Watson}} \hat{\mathbf{g}}, \hat{\mathbf{n}}_t \xrightarrow{\text{ISA}} \mathbf{A}, \mathbf{e}_t$.

4. Illustration

- Dataset: *d-geom* ($d_1 = 2$, $d_2 = d_3 = 3$, $d_4 = 4$), *ikeda* ($M = 2$, $d_m = 2$); see Fig. 3.
- Performance (Amari-index): ISA ambiguities ⇒ measure the block-permutation property of

$$\mathbf{G} = \mathbf{W}_{\text{ISA}}\mathbf{A}. \quad (4)$$

- Experiences:
  - *d-geom* (different dimensional sources; 12D):
    - amenable for sample size $T \geq 100,000$, see Fig. 4.
  - *ikeda* (Fig.5):
    - AR-IPA: can not find the proper subspaces.
    - LDS: EM + Kalman smoother ⇒ Amari-index = 0.48 – poor.
    - fAR-IPA: precise estimation for sample number $T \geq 10,000$, Amari-index = 0.0055.

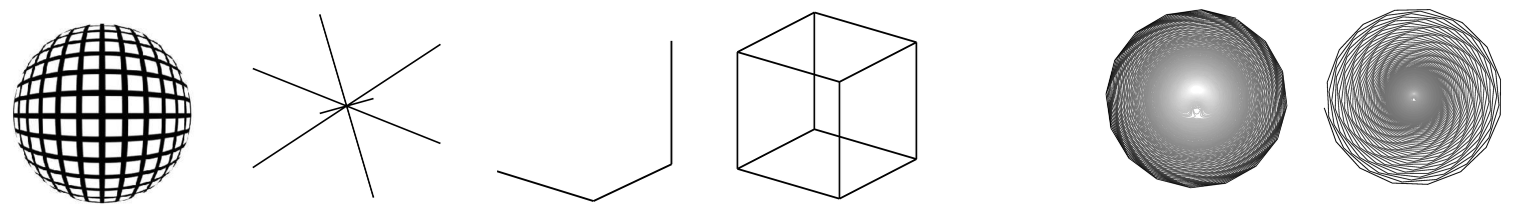


Figure 3: *Datasets. Left:* d-geom. *Right:* ikeda.

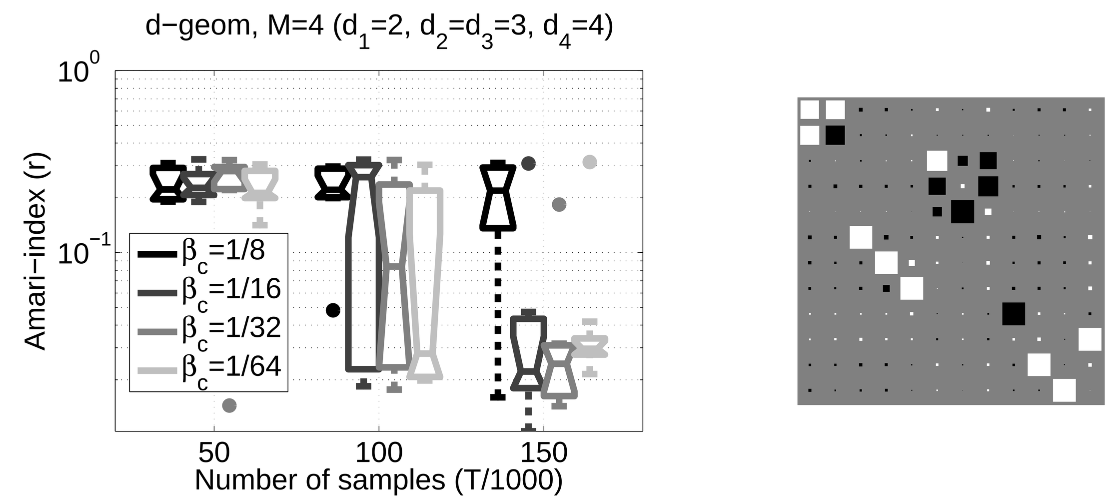
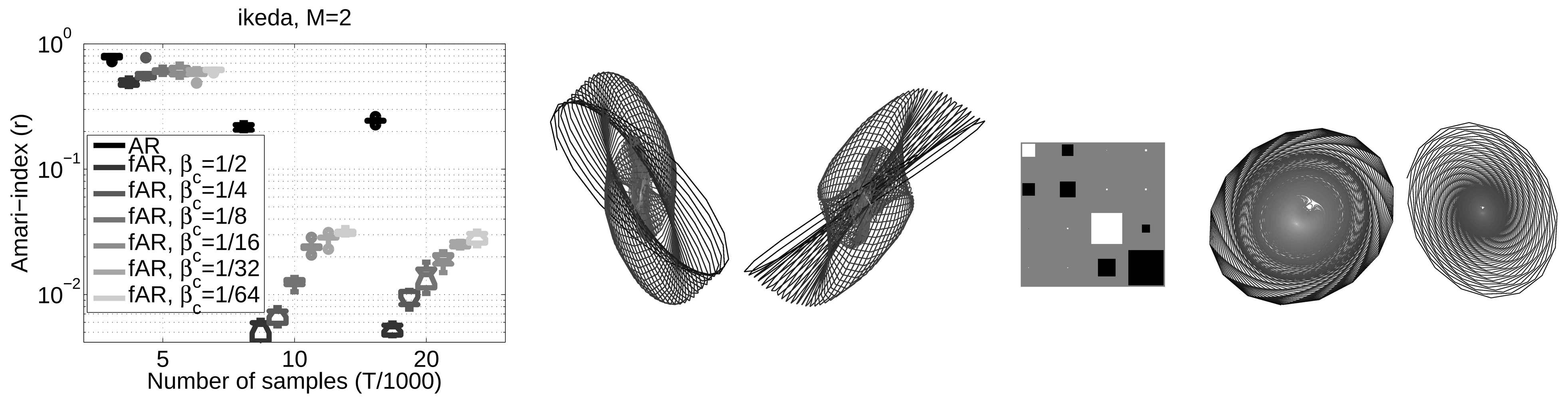


Figure 4: *Illustration on the* d-geom *dataset. Left: Amari-index. Right: Hinton-diagram of* $\mathbf{G}$.



(a) (b) (c) (d)

Figure 5: *Illustration on the* ikeda *dataset. (a): Amari-index. (b): Observation,* $\mathbf{x}_t$*. (c): Hinton-diagram of* $\mathbf{G}$*. (d): Estimated subspaces (*$\hat{\mathbf{s}}_t$*, fAR-IPA).*

Nemzeti Fejlesztési Ügynökség
www.ujszechenyiterv.gov.hu
06 40 638 638
MAGYARORSZÁG MEGÚJUL
The Project is supported by the European Union and co-financed by the European Social Fund (grant agreement no. TAMOP 4.2.1./B-09/1/KMR-2010-0003).

EUSIPCO 2011, 21st European Signal Processing Conference, August 29 – September 2, Barcelona, Spain