Model Order Selection: Estimating the Number of Signals Observed by Multiple Sensors

Marco Chiani
CNIT, University of Bologna, Italy

Based on joint works with Moe Z. Win

PhD Summer School on Signal Processing for ICT
Bressanone, 7-11 Luglio 2014

Research supported in part by the EC under the FP7 grant agreement n. 288502

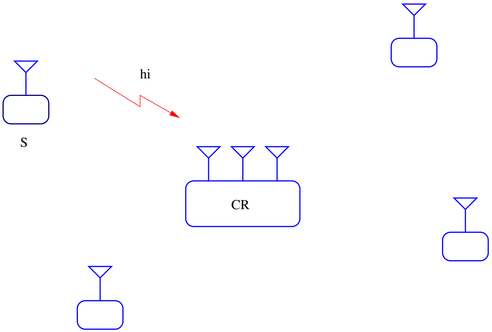
Analyzing the radio scene

- An energy detector can be used to distinguish between "noise only" and "signal(s) + noise".

- But how can we estimate the number of signals?

Scenario

Assuming the receiver is equipped with multiple antenna elements (sensors), spatial correlation can be exploited to discriminate signal sources from the spatially uncorrelated thermal noise.



The mathematical problem

Received n_{R}-dimensional vector

$$\mathbf{y}(t) = \sum_{i=1}^{n_{\mathrm{T}}} \mathbf{h}_i x_i(t) + N(t) \tag{1}$$

where

- n_{T} is the (unknown) number of signals (active transmitters)
- $N(t) \in \mathbb{C}^{n_{\mathrm{R}}}$ is the thermal noise, $N(t) \sim \mathcal{CN}_{n_{\mathrm{R}}}\left(\mathbf{0}, \sigma^2 \mathbf{I}_{n_{\mathrm{R}}}\right)$
- $x_i(t) \in \mathbb{C}$ is the symbol transmitted by the $i^{\underline{\text{th}}}$ source at time t
- $\mathbf{h}_i \in \mathbb{C}^{n_{\mathrm{R}}}$, $i = 1, \dots, n_{\mathrm{T}}$, are linearly independent vectors describing the gain of the radio channel between $i^{\underline{\text{th}}}$ transmitting source and the n_{R} receiving antennas.

Problem: determine the number of sources n_{T} by observing n_{S} samples $\mathbf{y}(t_1), \cdots, \mathbf{y}(t_{n_{\mathrm{S}}})$ over which the number of sources and the channel do not change, assuming **nothing** is known at the receiver.

Equivalent $n_R \times n_T$ virtual MIMO formulation

$$\mathbf{y}(t) = \mathbf{H}\,\mathbf{x}(t) + N(t) \tag{2}$$

where $\mathbf{x}(t) = (x_1(t), \ldots, x_{n_T}(t))^T$ and $\mathbf{H} = [\mathbf{h}_1|\cdots|\mathbf{h}_{n_T}]$.

Note: **the only information available is the structure of the model.**

($\mathbf{H}$, n_T, σ^2, and obviously $\mathbf{x}(t)$ and $N(t)$ are unknowns)

Assumptions:

$\mathbf{x}(t_i) \sim \mathcal{CN}_{n_T}(\mathbf{0}, \mathbf{R}_x)$, with $\mathbf{x}(t_1), \mathbf{x}(t_2), \ldots$ statistically independent.

$N(t_i) \sim \mathcal{CN}_{n_R}(\mathbf{0}, \sigma^2\mathbf{I})$, with $N(t_1), N(t_2), \ldots$ statistically independent.

Observed vectors in a $n_R \times n_S$ observation matrix

$$\mathbf{Y} = [\mathbf{y}(t_1)|\cdots|\mathbf{y}(t_{n_S})] \tag{3}$$

For a given $\mathbf{H}$ the observed vector has covariance matrix

$$\begin{aligned}\mathbf{R} &= \mathbb{E}\left\{\mathbf{y}(t)\mathbf{y}^\dagger(t)\right\} = \mathbf{H}\,\mathbb{E}\left\{\mathbf{x}(t)\mathbf{x}^\dagger(t)\right\}\mathbf{H}^\dagger + \sigma^2\mathbf{I} \\ &= \mathbf{H}\mathbf{R}_x\mathbf{H}^\dagger + \sigma^2\mathbf{I}.\end{aligned} \tag{4}$$

The observed matrix is $\mathbf{Y} \sim \mathcal{CN}_{n_R,n_S}(\mathbf{0}, \mathbf{R}, \mathbf{I}_{n_S})$, and the matrix $\mathbf{W} = \mathbf{Y}\mathbf{Y}^\dagger \in \mathbb{C}^{n_R \times n_R}$ is Wishart.

The multiplicity of the smallest eigenvalue of $\mathbf{R}$

Ordered eigenvalues of $\mathbf{R}$: $\lambda_1 \geq \ldots \geq \lambda_{n_\mathrm{R}}$.
If $n_\mathrm{T} < n_\mathrm{R}$, the smallest $n_\mathrm{R} - n_\mathrm{T}$ eigenvalues of $\mathbf{R}$ are all equal to σ^2.
Example: $n_\mathrm{R} = 5$ antennas, $n_\mathrm{T} = 2$ sources

$$\mathbf{R} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{U}^\dagger = \mathbf{U}\begin{bmatrix} \lambda_1 & 0 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 & 0 \\ 0 & 0 & \lambda_3 = \sigma^2 & 0 & 0 \\ 0 & 0 & 0 & \lambda_4 = \sigma^2 & 0 \\ 0 & 0 & 0 & 0 & \lambda_5 = \sigma^2 \end{bmatrix}\mathbf{U}^\dagger$$

$\Rightarrow$ determining the number of source $\equiv$ determining of the multiplicity of the smallest eigenvalue of $\mathbf{R}$.

Unfortunately, $\mathbf{R}$ is not available at the receiver $\Rightarrow$ use its estimate based on the available n_S observations.

The problem seen as estimate of the multiplicity of the smallest eigenvalue of **R** based on the sample covariance matrix

The sample covariance matrix (SCM) is

$$\hat{\mathbf{R}} = \frac{1}{n_S} \sum_{i=1}^{n_S} \mathbf{y}(t_i)\mathbf{y}^\dagger(t_i) = \frac{1}{n_S}\mathbf{Y}\mathbf{Y}^\dagger = \frac{1}{n_S}\mathbf{W}. \tag{5}$$

This is an estimate of the covariance matrix **R**.

The problem of determining the number of sources is reformulated as:

find an estimate of the multiplicity
of the smallest eigenvalue of **R** based on $\hat{\mathbf{R}}$.

Model selection based on Information Criteria: AIC, MDL, BIC

Consider a set of observed data $\mathbf{Y}$ generated according to a probability distribution within a set of possible distributions $\left\{f(\mathbf{Y};\boldsymbol{\theta}^{(k)})\right\}$, where the model k is characterized by the unknown parameter vector $\boldsymbol{\theta}^{(k)} = \left\{\theta_1^{(k)}, \ldots, \theta_{\nu(k)}^{(k)}\right\}$ of size $\nu(k)$. The modeling problem consists of selecting k.

Known methods: Akaike information criterion (AIC), minimum description length (MDL), Bayesian information criterion (BIC).

Given the observation $\mathbf{Y}$, the selected model is

$$\hat{k} = \arg\min_{k} \left\{- \log f(\mathbf{Y};\hat{\boldsymbol{\theta}}^{(k)}) + L(\nu(k), n_{\mathrm{S}})\right\} \quad (6)$$

where $\hat{\boldsymbol{\theta}}^{(k)}$ for a fixed k is the maximum likelihood (ML) estimate of $\boldsymbol{\theta}^{(k)}$and $L(\nu(k), n_{\mathrm{S}})$ is a penalty function which depends on the number $\nu(k)$ of free-adjusted parameters for the model k.

- $L(\nu, n_{\mathrm{S}}) = L_{AIC}(\nu, n_{\mathrm{S}}) = \nu$ for the AIC
- $L(\nu, n_{\mathrm{S}}) = L_{BIC}(\nu, n_{\mathrm{S}}) = \frac{\nu}{2} \log n_{\mathrm{S}}$ for the BIC

Methods for the detection of the number of sources

[Wax, Kailath 1985]: the parameter vector is

$$\boldsymbol{\theta}^{(k)} = (\lambda_1, \cdots, \lambda_k, \mathbf{v}_1, \cdots, \mathbf{v}_k, \sigma^2) \tag{7}$$

where $\lambda_i, \mathbf{v}_i$ are the eigenvalues and eigenvectors of $\mathbf{R}$.
The number of free-adjusted parameters is $k(2n_{\mathrm{R}} - k) + 1$.
[Wax, Kailath 1985]: The estimate of the number of signals is

$$\begin{aligned} \hat{n}_{\mathrm{T}} &= \arg \min_{k \in \{0,\ldots,k_{\max}\}} \left\{ \log \left(\frac{\frac{1}{n_{\mathrm{R}}-k} \sum_{i=k+1}^{n_{\mathrm{R}}} l_i}{\prod_{i=k+1}^{n_{\mathrm{R}}} l_i^{1/(n_{\mathrm{R}}-k)}} \right)^{(n_{\mathrm{R}}-k)n_{\mathrm{S}}} \right. \\ &+ \left. L(k(2n_{\mathrm{R}}-k), n_{\mathrm{S}}) \right\} \end{aligned} \tag{8}$$

where l_i are the ordered eigenvalues of the SCM.
Good performance for a sufficiently large number of observations n_{S}.
Performance degradation for small sample sizes.

Main contributions

- Exact distribution of the eigenvalues of the SCM for the Gaussian multivariate case for population covariance matrix (PCM) with eigenvalues of arbitrary multiplicities. This problem is of large interest in multivariate statistical analysis and only approximate solutions, valid for large number of samples or large dimensions of the observed vectors, were previously known [Anderson, Muirhead].
- The exact ML estimate of the PCM eigenvalues.
- New method to estimate the number of sources embedded in Gaussian noise by means of information-theoretic criteria model order selection.
- Numerical results.

Exact distribution of the eigenvalues of the SCM

Lemma

Let $\tilde{\mathbf{Y}} \sim \mathcal{CN}_{n_R,n_S}(\mathbf{0}, \mathbf{I}_{n_R}, \mathbf{I}_{n_S})$ *and let* $\mathbf{R}$ *be an* $n_R \times n_R$ *positive definite matrix. The joint p.d.f. of the (real) non-zero ordered eigenvalues* $\mathbf{l} = (l_1, \ldots, l_{n_{min}})$ *of the* $n_S \times n_S$ *quadratic form* $\tilde{\mathbf{Y}}^\dagger \mathbf{R} \tilde{\mathbf{Y}}$ *or of the* $n_R \times n_R$ *Wishart matrix* $\mathbf{R}^{1/2}\tilde{\mathbf{Y}}\tilde{\mathbf{Y}}^\dagger\mathbf{R}^{1/2}$ *is given by*

$$f(\mathbf{l}; \boldsymbol{\mu}, \mathbf{m}) = K(\boldsymbol{\mu}, \mathbf{m}) \det \mathbf{V}(\mathbf{l}) \det \mathbf{G}(\mathbf{l}; \boldsymbol{\mu}, \mathbf{m}) \prod_{i=1}^{n_{min}} l_i^{n_S - n_{min}} \qquad (9)$$

where $n_{min} = \min(n_R, n_S)$, $\mathbf{V}(\mathbf{l})$ *is the Vandermonde matrix,* $K(\boldsymbol{\mu}, \mathbf{m}) = \frac{(-1)^{n_S(n_R - n_{min})}}{\Gamma_{(n_{min})}(n_S) \prod_{i=1}^{L} \Gamma_{(m_i)}(m_i)} \frac{\prod_{i=1}^{L} \mu_{(i)}^{m_i n_S}}{\prod_{i<j}\left(\mu_{(i)} - \mu_{(j)}\right)^{m_i m_j}}$, $\Gamma_{(m)}(a) \triangleq \prod_{i=1}^{m}(a-i)!$ *and the vector* $\boldsymbol{\mu} = (\mu_{(1)}, \ldots, \mu_{(L)})$ *contains the* L *distinct ordered eigenvalues of* $\mathbf{R}^{-1}$*, with corresponding multiplicities described by the vector* $\mathbf{m} = (m_1, \ldots, m_L)$.
The $n_R \times n_R$ *matrix* $\mathbf{G}(\mathbf{l}; \boldsymbol{\mu}, \mathbf{m})$ *has elements*

$$g_{i,j} = \begin{cases} \left(-l_j\right)^{d_i} e^{-\mu_{(e_i)} l_j} & j = 1, \ldots, n_{min} \\ [n_R - j]_{d_i} \, \mu_{(e_i)}^{n_R - j - d_i} & j = n_{min} + 1, \ldots, n_R \end{cases} \qquad (10)$$

ML Estimate of the Eigenvalues of the Sample Cov. Matrix

Lemma

Let $\tilde{\mathbf{Y}} \sim \mathcal{CN}_{n_R,n_S}(\mathbf{0}, \mathbf{I}_{n_R}, \mathbf{I}_{n_S})$ *and let* $\mathbf{R} \in \mathbb{C}^{n_R \times n_R}$ *be the a positive definite matrix having* L *unknown distinct eigenvalues* $\lambda_{(1)} > \ldots > \lambda_{(L)}$ *with known multiplicities* $\mathbf{m} = (m_1, \ldots, m_L)$*. The ML estimates of the eigenvalues of* $\mathbf{R}$*, based on the observation of a given instantiation of the* $n_S \times n_S$ *quadratic form* $\tilde{\mathbf{Y}}^\dagger \mathbf{R} \tilde{\mathbf{Y}}$ *having non-zero ordered distinct eigenvalues* $\tilde{\mathbf{l}} = (\tilde{l}_1, \ldots, \tilde{l}_{n_{min}})$*, are given by*

$$\hat{\lambda}_{(i)} = \frac{1}{\hat{\mu}_{(L-i+1)}}, \quad i = 1 \ldots, L$$

where

$$(\hat{\mu}_{(1)}, \ldots, \hat{\mu}_{(L)}) = \arg\max_{\boldsymbol{\mu}} \left| K(\boldsymbol{\mu}, \mathbf{m}) \det \mathbf{G}(\tilde{\mathbf{l}}; \boldsymbol{\mu}, \mathbf{m}) \right|$$

the maximum is taken over all vectors $\boldsymbol{\mu} = (\mu_{(1)}, \ldots, \mu_{(L)})$ *with* $\mu_{(1)} > \ldots > \mu_{(L)}$*.*

New Methods for Estimating the Number of Sources

Lemma

The estimate of the number of sources based on information criteria and considering the exact eigenvalues distribution of the SCM in Lemma 1 is

$$\hat{n}_T = \arg\min_{k \in \{0,\ldots,k_{\max}\}} \left\{ -\log \left| K(\hat{\boldsymbol{\mu}}^{(k)}, \mathbf{m}^{(k)}) \ \det \mathbf{G}(\tilde{\mathbf{l}}; \hat{\boldsymbol{\mu}}^{(k)}, \mathbf{m}^{(k)}) \right| + L(k, n_S) \right\}$$

where $\mathbf{m}^{(k)} = (n_R - k, 1, 1, \ldots, 1)$, *and the elements of* $\hat{\boldsymbol{\mu}}^{(k)} = (\hat{\mu}_{(1)}^{(k)}, \ldots, \hat{\mu}_{(k+1)}^{(k)})$ *are the ML estimates of the* $k+1$ *distinct eigenvalues of* $\mathbf{R}^{-1}$ *in the hypothesis of* k *signal sources.*

Numerical results

(n. RX antennas, n. samples)	New (EXBIC)	BIC [WaxKai:85]	NE [NadEde:08]
(6,6)	0.0218	0.5672	0.0348
(6,8)	0.0198	0.1014	0.0352
(8,8)	0.0140	0.2962	0.0362
(16,16)	0.0038	0.0026	0.0478

Table : Probability of False Alarm, $k_{\max} = n_{\mathrm{R}} - 1$.

New: EXBIC, EXAIC. Literature: AIC, BIC [Wax, Kailath 1985], NE [Nadakuditi, Edelman 2008]

Numerical results

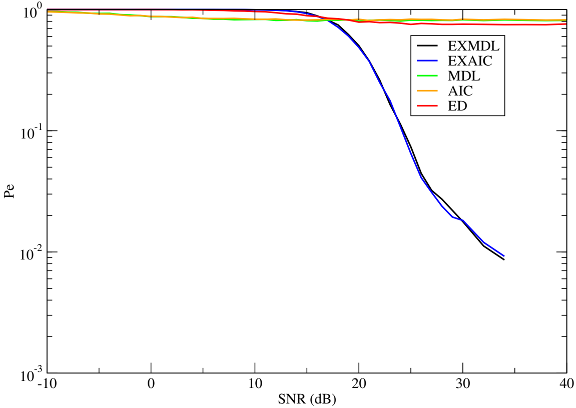


Figure : $n_R = 8$ RX antennas, $n_T = 6$ sources, $n_S = 8$ samples.
New: EXBIC, EXAIC. Literature: AIC, BIC [WaxKai:85], NE [NadEde:08]

Numerical results

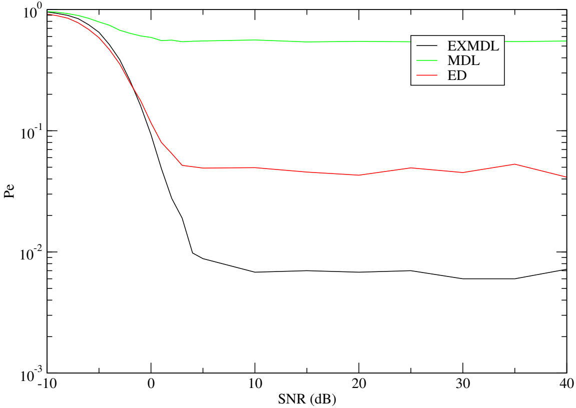


Figure : $n_R = 8$ RX antennas, $n_T = 1$ source, $n_S = 8$ samples.
New: EXMDL. Literature: AIC, MDL/BIC [WaxKai:85], ED [NadEde:08]

Example of observed samples (3 RX antennas, only real parts)

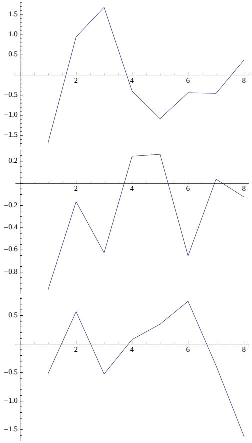


Figure : $n_T = 2$ sources, $n_S = 8$ samples, $SNR = 0$ dB

Numerical Results

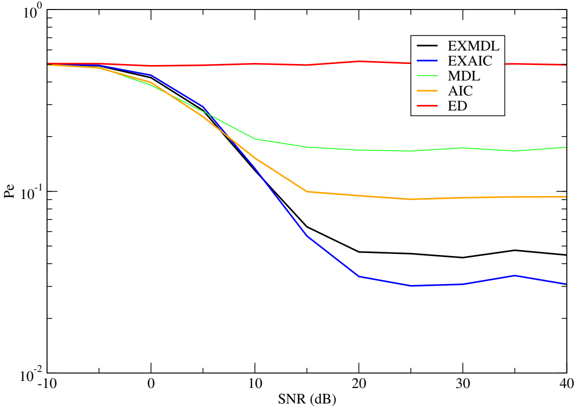


Figure : $n_R = 2$ RX antennas, $n_S = 4$ samples. The number of sources n_T is random over the set $\{0, 1, \ldots, \min(n_S, n_R) - 1\}$.

Numerical Results

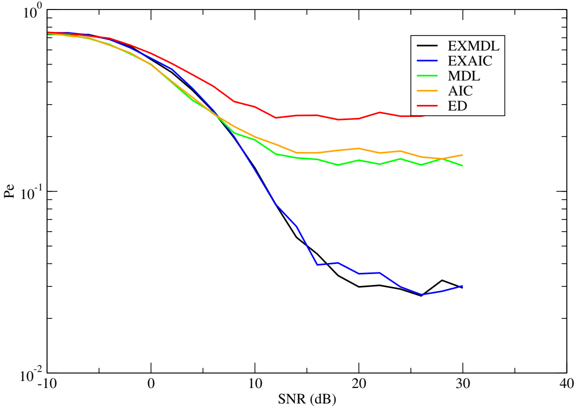


Figure : $n_R = 4, n_S = 8$. The number of sources n_T is random over the set $\{0, 1, \ldots, \min(n_S, n_R) - 1\}$.

Numerical I

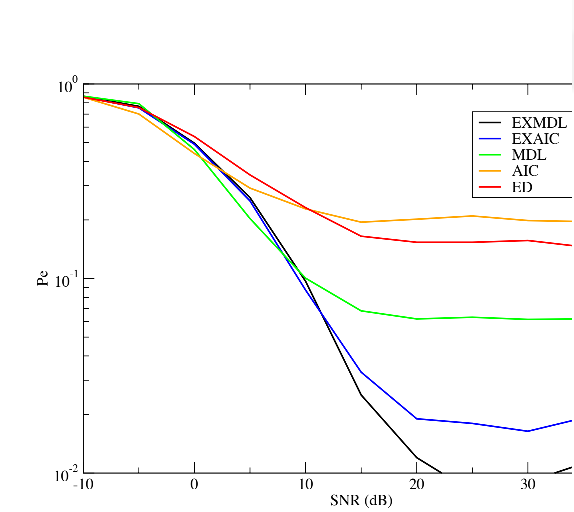


Figure : $n_R = 8, n_S = 16$. The number of sources n_T is random over the set $\{0, 1, \ldots, n_R - 1\}$.

Conclusions

- Found the exact distribution of the eigenvalues of the sample covariance matrix for the multivariate Gaussian case
- Proposed a new method for estimating the number of signals embedded in Gaussian noise
- The new method has good performance even with a small number of observations

References Model Selection

- M. Chiani, Moe Z. Win “Estimating the number of signals observed by multiple sensors, in IEEE 2010 IAPR Workshop on Cognitive Information Processing (CIP2010), 2010.
- M. Chiani, M. Z. Win, and H. Shin, “MIMO networks: the effects of interference, IEEE Trans. Inform. Theory, Jan. 2010.
- A. Giorgetti and M. Chiani, Time-of-arrival estimation based on information theoretic criteria, IEEE Trans. Signal Processing, 2013.
- A. Giorgetti, A. Mariani, M. Chiani, "Spectrum holes detection by information theoretic criteria," Proc. IEEE Int. Work. on Cognitive Radio Advanced Spectrum Management (CogART), 2011.

References about random matrices

- A.T. James, "Distributions of matrix variates and latent roots derived from normal samples," Ann. Math. Stat., 1964.
- R.J. Muirhead, "Aspects of Multivariate Statistical Theory," Wiley, New York, 1982.
- C. Tracy, H. Widom, "The distributions of random matrix theory and their applications," New Trends Math. Phys., 2009.
- M. Chiani, M. Z. Win, and A. Zanella, On the capacity of spatially correlated MIMO Rayleigh fading channels, IEEE Trans. Inform. Theory, 2003.
- M. Chiani, M. Z. Win, and H. Shin, MIMO networks: the effects of interference, IEEE Trans. Inform. Theory, 2010.
- H. Shin, M. Win, J. H. Lee, and M. Chiani, "On the capacity of doubly correlated MIMO channels," IEEE Trans. Wireless Commun., 2006.
- A. Zanella, M. Chiani, and M. Z. Win, On the marginal distribution of the eigenvalues of Wishart matrices," IEEE Trans. Commun., 2009.

References about random matrices

- M. Chiani and A. Zanella, Joint distribution of an arbitrary subset of the ordered eigenvalues of Wishart matrices," in Proc. IEEE PIMRC 2008.
- H. Shin, M. Win, J. H. Lee, and M. Chiani, Asymptotic statistics of mutual information for doubly correlated MIMO channels," IEEE Trans. Wireless Commun., 2008.
- A. Zanella and M. Chiani, Reduced complexity power allocation strategies for MIMO systems with singular value decomposition, IEEE Trans. Veh. Technol., 2012.
- A. Zanella, M. Chiani, and M. Z. Win, Statistical analysis of Steepest Descend and LMS detection algorithms for MIMO systems, IEEE Trans. Veh. Technol., Nov. 2011.
- M. Chiani, Distribution of the largest eigenvalue for real Wishart and Gaussian random matrices and a simple approximation for the Tracy-Widom distribution, Journal of Multivariate Analysis, 2014, see also arxiv 2012.