Some Issues on Errors–in–Variables Identification

I. Determination of linear relations in the Frisch scheme context

Roberto Guidorzi, Roberto Diversi, Umberto Soverini

rguidorzi,rdiversi,usoverini@deis.unibo.it

Bologna University

Estimating linear relations from noisy data

One of the most common problems in science and technology concerns the estimation of linear relations from data affected by errors. Consider noiseless data given by N samples of n variables ($N > n$):

$$x_1 \; x_2 \; \ldots \; x_n. \tag{1}$$

Linear relations between the variables (if any) can be described in the form

$$a_1 \, x_1 + a_2 \, x_2 + \; \ldots \; + a_n \, x_n = 0 \tag{2a}$$

or

$$y = \alpha_1 \, \xi_1 + \alpha_2 \, \xi_2 + \ldots + \alpha_{n-1} \, \xi_{n-1} \tag{2b}$$

By denoting with X the $N \times n$ matrix with rows given by the observations

$$X = \begin{bmatrix} x_{11} & x_{21} & \dots & x_{n1} \\ x_{12} & x_{22} & \dots & x_{n2} \\ \vdots & \vdots & & \vdots \\ x_{1N} & x_{2N} & \dots & x_{nN} \end{bmatrix}, \tag{3}$$

sets of one or more linear relations can be expressed in the form

$$X\,A = 0 \tag{4}$$

where A is a $(n \times q)$ matrix with columns given by the q sets of coefficients describing the $q = n - \mathrm{rank}\, X$ (independent) linear relations linking the data. Relation (4) can be rewritten also by substituting X with

$$\Sigma = \frac{X^T X}{N} \tag{5}$$

i.e., under the assumption of null mean value of the variables, with the sample covariance matrix of the data.

In absence of noise and in presence of linear relations

$$\Sigma \geq 0 \tag{6}$$

and every solution, A, with maximal rank, of

$$\Sigma A = 0 \tag{7}$$

is a basis of $\ker \Sigma$.

When the data are corrupted by noise, $\operatorname{rank} X = n$, no linear relations are compatible with the observations and

$$\Sigma > 0. \tag{8}$$

In situations of this kind, linear relations can be extracted only by modifying X or Σ i.e. the data.

Definition 1 (Kalman, 1982a,b) – A *scheme* is a systematic procedure to extract linear relations from data affected by errors.

Assumptions behind estimation schemes

If no assumptions are introduced, *any* set of noisy data is compatible with *any* solution. The assumptions on the noise are, usually, the following:

1) The noise is additive; every observation is the sum of an unknown exact part $\hat{x}_i$, and of a noise term $\tilde{x}_i$:

$$x_i = \hat{x}_i + \tilde{x}_i \tag{9}$$

2) The mean value of $\hat{x}_i$ and $\tilde{x}_i$ is null:

$$\sum_{t=1}^{N} \hat{x}_{it} = 0, \qquad \sum_{t=1}^{N} \tilde{x}_{it} = 0. \tag{10}$$

3) The sequences of noise samples are orthogonal to the sequences of noiseless variables:

$$\sum_{t=1}^{N} \tilde{x}_{it}\, \hat{x}_{jt} = 0 \qquad \text{for every } i, j. \tag{11}$$

Under these assumptions:

$$X = \hat{X} + \tilde{X} \tag{12}$$

$$\hat{X}^T \tilde{X} = 0 \tag{13}$$

$$\Sigma = \hat{\Sigma} + \tilde{\Sigma} \tag{14}$$

$$\Sigma > 0 \tag{15}$$

$$\tilde{\Sigma} \geq 0 \;\text{ or }\; \tilde{\Sigma} > 0 \tag{16}$$

$$\hat{\Sigma} \geq 0. \tag{17}$$

The problem of determining linear relations compatible with noisy data can be formulated as follows:

Problem 1 (Kalman, 1982a,b) – Given a sample covariance matrix of noisy observations, Σ, determine positive definite or semidefinite noise covariance matrices $\tilde{\Sigma}$ such that

$$\hat{\Sigma} = \Sigma - \tilde{\Sigma} \geq 0. \tag{18}$$

All corresponding solutions are described by any basis of $\ker \hat{\Sigma}$.

The Frisch scheme

This scheme, proposed by Ragnar Frisch in 1934, assumes the additivity of noise terms

$$x_i = \hat{x}_i + \tilde{x}_i, \tag{19}$$

the independence between noise and data sequences and the mutual independence of the noise sequences. This corresponds, by introducing the suffix n in the sample covariance matrices, to

$$\tilde{\Sigma}_n = \mathrm{diag}\,[\,\tilde{\sigma}_1^2, \ldots, \tilde{\sigma}_n^2\,] \geq 0 \quad \text{or} \quad > 0 \tag{20}$$

where $\tilde{\sigma}_1^2, \ldots, \tilde{\sigma}_n^2$ are the sample variances of the noise terms $\tilde{x}_1, \ldots, \tilde{x}_n$.

Every positive definite or semidefinite diagonal matrix $\tilde{\Sigma}_n$ such that

$$\hat{\Sigma}_n = \Sigma_n - \tilde{\Sigma}_n \geq 0 \tag{21}$$

is a *solution* of the Frisch scheme. The corresponding point $P = (\tilde{\sigma}_1^2, \ldots, \tilde{\sigma}_n^2) \in \mathcal{R}^n$ is a solution in the noise space.

Properties of the solutions in the noise space

Theorem 1 (Beghelli, Guidorzi and Soverini, 1990) – All admissible solutions in the noise space lie on a convex (hyper)surface $\mathcal{S}(\Sigma_n)$ whose concavity faces the origin and whose intersections with the coordinate axes are the points $[0, \ldots, \tilde{\sigma}_i^2, \ldots, 0]$ corresponding to the n ordinary least squares solutions.

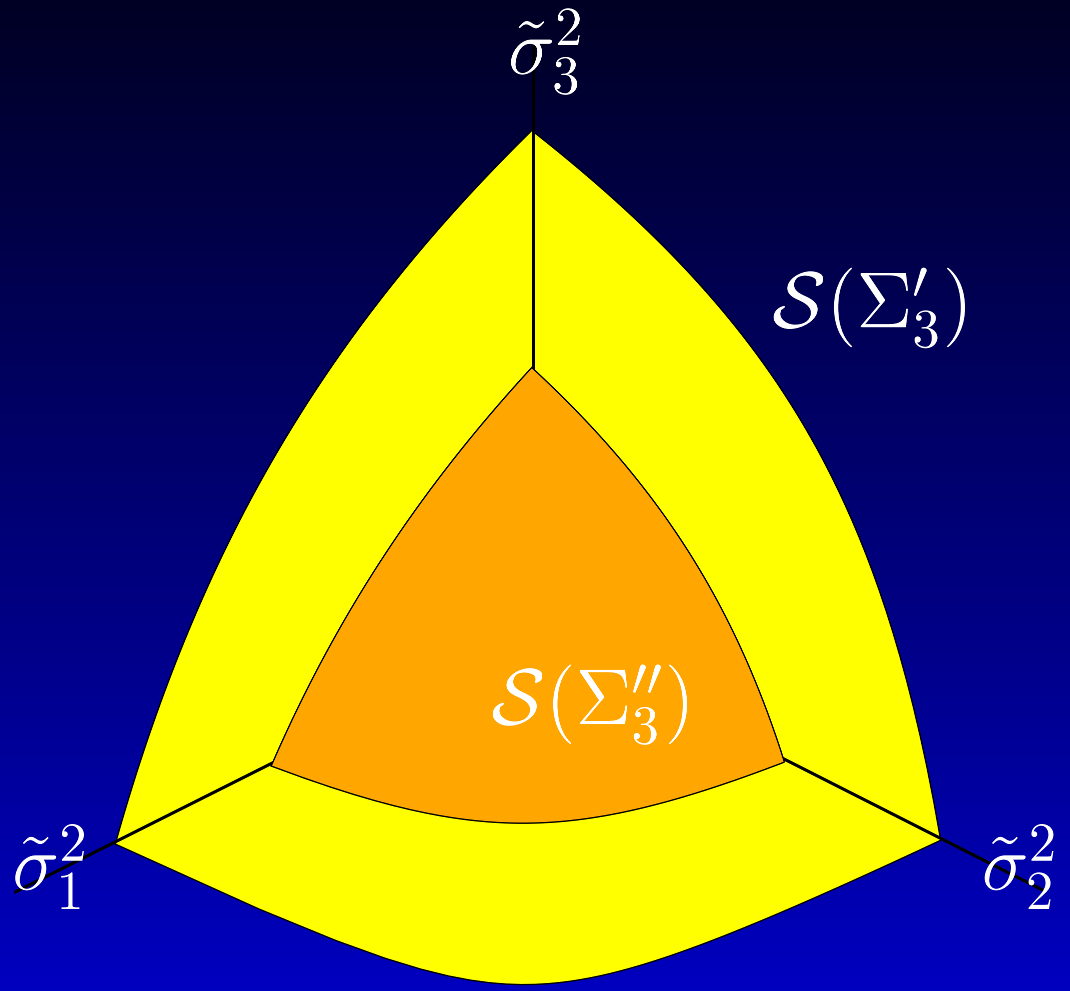


Figure 1: Loci $\mathcal{S}(\Sigma_3)$ of admissible noise points

Definition 2 (Guidorzi, 1995) – The (hyper)surface $\mathcal{S}(\Sigma_n)$ will be called *singularity (hyper)surface* of Σ_n because its points define noise covariance matrices $\tilde{\Sigma}_n$ associated with singular matrices $\hat{\Sigma}_n$.

A problem of great relevance concerns the conditions under which a covariance matrix is compatible with more linear relations i.e. the evaluation of the maximal dimension of $\ker \Sigma_n$ $(\mathrm{Maxcor}_F(\Sigma_n))$ in the contextt of the Frisch scheme.

Theorem 2 (Kalman, 1982a) – $\mathrm{Maxcor}_F(\Sigma_n) = 1$ if and only if Σ_n^{-1} is Frobenius–like or becomes Frobenius–like by changing the sign of some variables.

Theorem 3 – When $\mathrm{Maxcor}_F(\Sigma_n) = 1$ the coefficients $a_1, \ldots, a_n$ of all linear relations compatible with the Frisch scheme lie (by normalizing one of the coefficients to 1) inside the simplex whose vertices are defined by the n LS solutions.

Theorem 4 – When $\mathrm{Maxcor}_F(\Sigma_n) = 1$ the points of the simplex of solutions in the parameter space are isomorphic with the points of $\mathcal{S}(\Sigma_n)$.

Theorem 5 (Schachermayer and Deistler, 1998) – $\mathcal{S}(\Sigma_n)$ is nonuniformly convex.

Theorem 6 (Deistler and Scherrer, 1992) – All points of $\mathcal{S}(\Sigma_n)$ where $\mathrm{Cor}\,(\Sigma_n) = k$ $(k > 1)$ are accumulation points for those where $\mathrm{Cor}\,(\Sigma_n) = k - 1$.

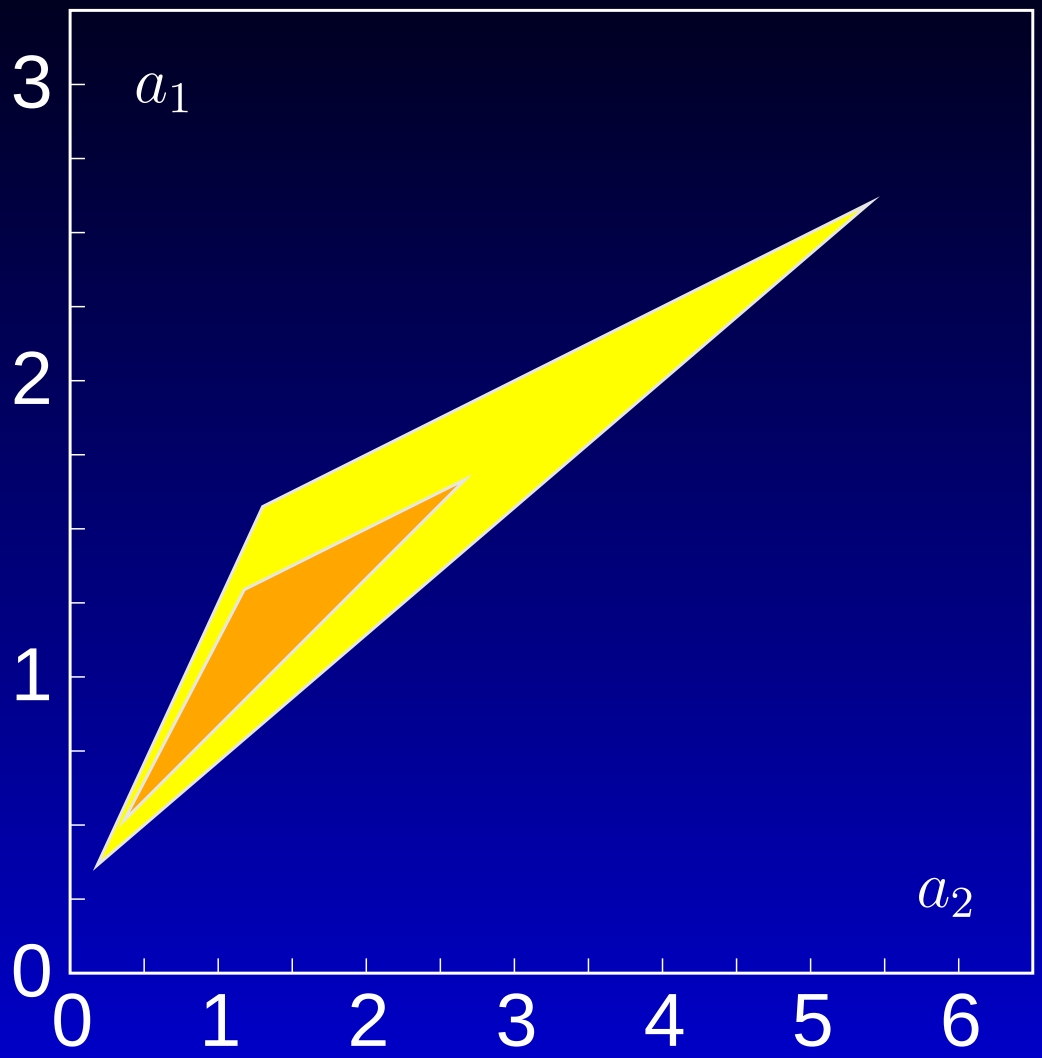


Figure 2: Loci of admissible parameters

Computation of $\mathrm{Maxcor_F}\,(\Sigma_n)$

Define the singularity (hyper)surface $\mathcal{S}(\Sigma_{n/r})$ as the locus of the points $[\tilde{\sigma}_1^2, \ldots, \tilde{\sigma}_r^2]^T \in \mathcal{R}^r$ such that

$$\Sigma_n - \mathrm{diag}\,[\,\tilde{\sigma}_1^2, \ldots, \tilde{\sigma}_r^2, 0, \ldots, 0\,] \geq 0 \qquad (22)$$

and Σ_r as the sample covariance matrix of the first r variables. Then the following geometric relations hold:

Theorem 7 (Guidorzi and Stoian, 1994) – $\mathcal{S}(\Sigma_{n/r})$ always lies under or on $\mathcal{S}(\Sigma_r)$.

Theorem 8 (Guidorzi, 1995) – $\mathrm{Maxcor_F}\,(\Sigma_n) \geq q$ if and only if $\mathcal{S}(\Sigma_{n-q+1}) \cap \mathcal{S}(\Sigma_{n/n-q+1}) \neq 0$ for every subset of $n-q+1$ variables, i.e. for every permutation of the data leading to different subgroups in the first $n-q+1$ positions.

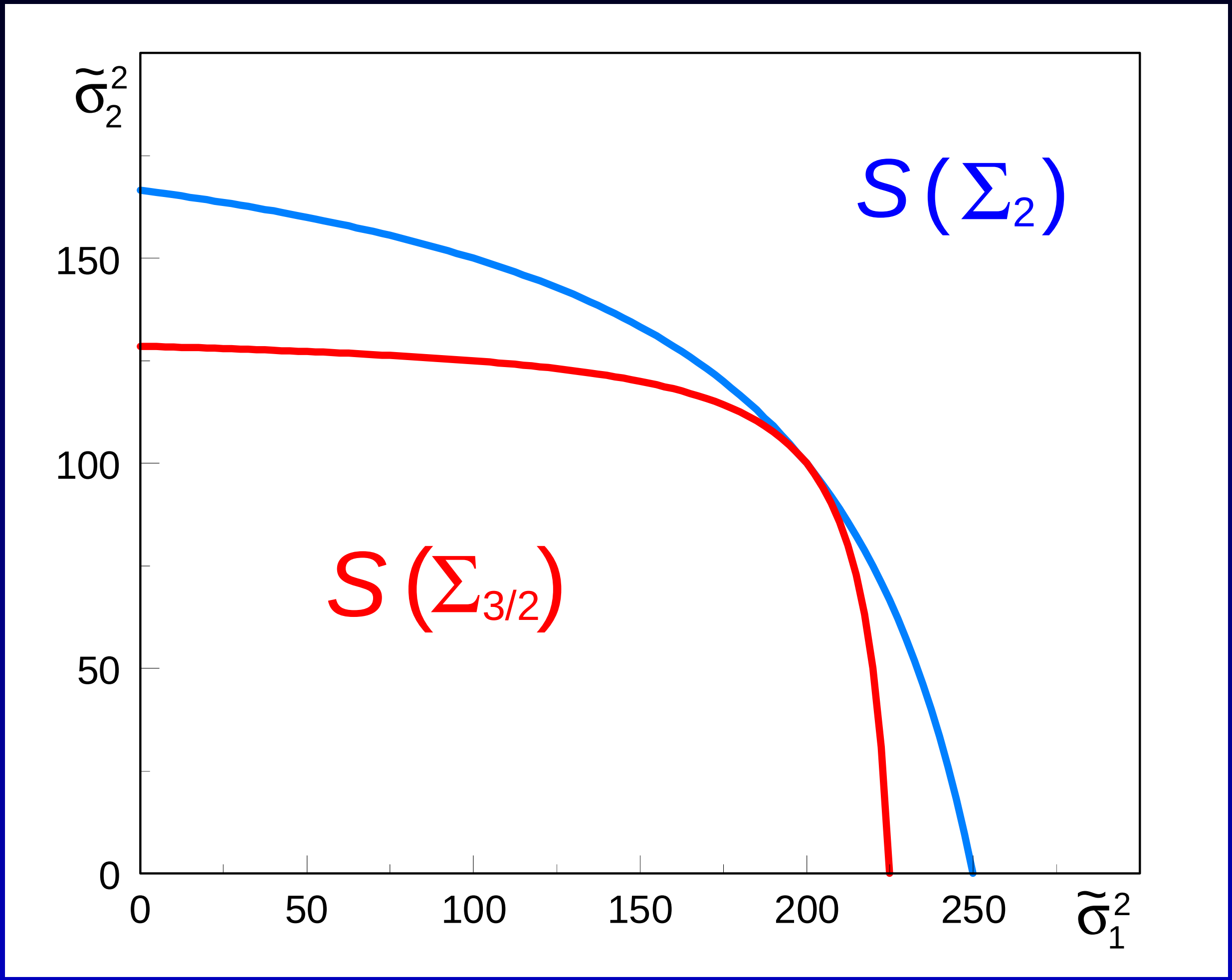


Figure 3: Common points between $\mathcal{S}(\Sigma_2)$ and $\mathcal{S}(\Sigma_{3/2})$ in a (3×3) covariance matrix with Maxcor$_F$ = 2

Complete sets of data for the Frisch scheme

The Frisch scheme is considered as less affected by prejudices than other schemes leading to unique solutions or to a limited number of solutions, because it treats all variables in a symmetric way and no *a priori* decomposition of Σ is preferred to any other.

Remark 1 – The linear relation linking the noiseless data belongs to the set of Frisch solutions; the knowledge of Σ allows computing the whole set of solutions but not discriminating one solution against others.

The definitions and properties that follow concern the asymptotic case (infinite sequence of data).

Definition 3 (Guidorzi, 1991) – Two noise–free data covariance matrices of the same linear algebraic process, $\hat{\Sigma}_1$ and $\hat{\Sigma}_2$ are defined as *independent* if

$$\dim \ker \hat{\Sigma}_1 = \dim \ker \hat{\Sigma}_2 = \dim \ker (\hat{\Sigma}_1 - \hat{\Sigma}_2) = 1. \tag{23}$$

Property 1 – If $\hat{\Sigma}_1$ and $\hat{\Sigma}_2$ are independent there exists a unique (modulo scaling) vector a satisfying the conditions

$$\hat{\Sigma}_1\, a = \hat{\Sigma}_2\, a = (\hat{\Sigma}_1 - \hat{\Sigma}_2)\, a = 0. \tag{24}$$

Definition 4 (Guidorzi, 1991) – Two noisy data covariance matrices of the same linear algebraic process, $\Sigma_1 > 0$ and $\Sigma_2 > 0$ are defined as *independent* if

$$\dim \ker (\Sigma_1 - \Sigma_2) = 1. \tag{25}$$

Theorem 9 (Guidorzi, 1991) – Two independent noisy covariance matrices, Σ_1 and Σ_2, satisfy the following conditions under the Frisch scheme

$$\dim \ker (\Sigma_1 - \tilde{\Sigma}\,) = \dim \ker (\Sigma_2 - \tilde{\Sigma}\,) = 1 \tag{26}$$

$$(\Sigma_1 - \Sigma_2)\, a = (\Sigma_1 - \tilde{\Sigma})\, a = (\Sigma_2 - \tilde{\Sigma})\, a = 0, \tag{27}$$

where $a = [a_1\; a_2\; \ldots\; a_n]^T$ defines the process model (1) and $\tilde{\Sigma} \geq 0$ is a diagonal matrix satisfying the conditions

$$\Sigma_1 - \tilde{\Sigma} \geq 0, \qquad \Sigma_2 - \tilde{\Sigma} \geq 0. \tag{28}$$

Theorem 10 (Guidorzi, 1991) – Among all points common to the hypersurfaces of admissible noise points associated with the independent noisy covariance matrices Σ_1 and Σ_2, one and only one point is mapped, according to Σ_1 and Σ_2, into the same point of the parameter space.

Corollary 1 – The Frisch scheme leads to a unique solution determined by every pair of independent noisy data covariance matrices of the process.

Corollary 2 – Two independent noisy data covariance matrices of a process constitute a *complete* set of data under the Frisch scheme.

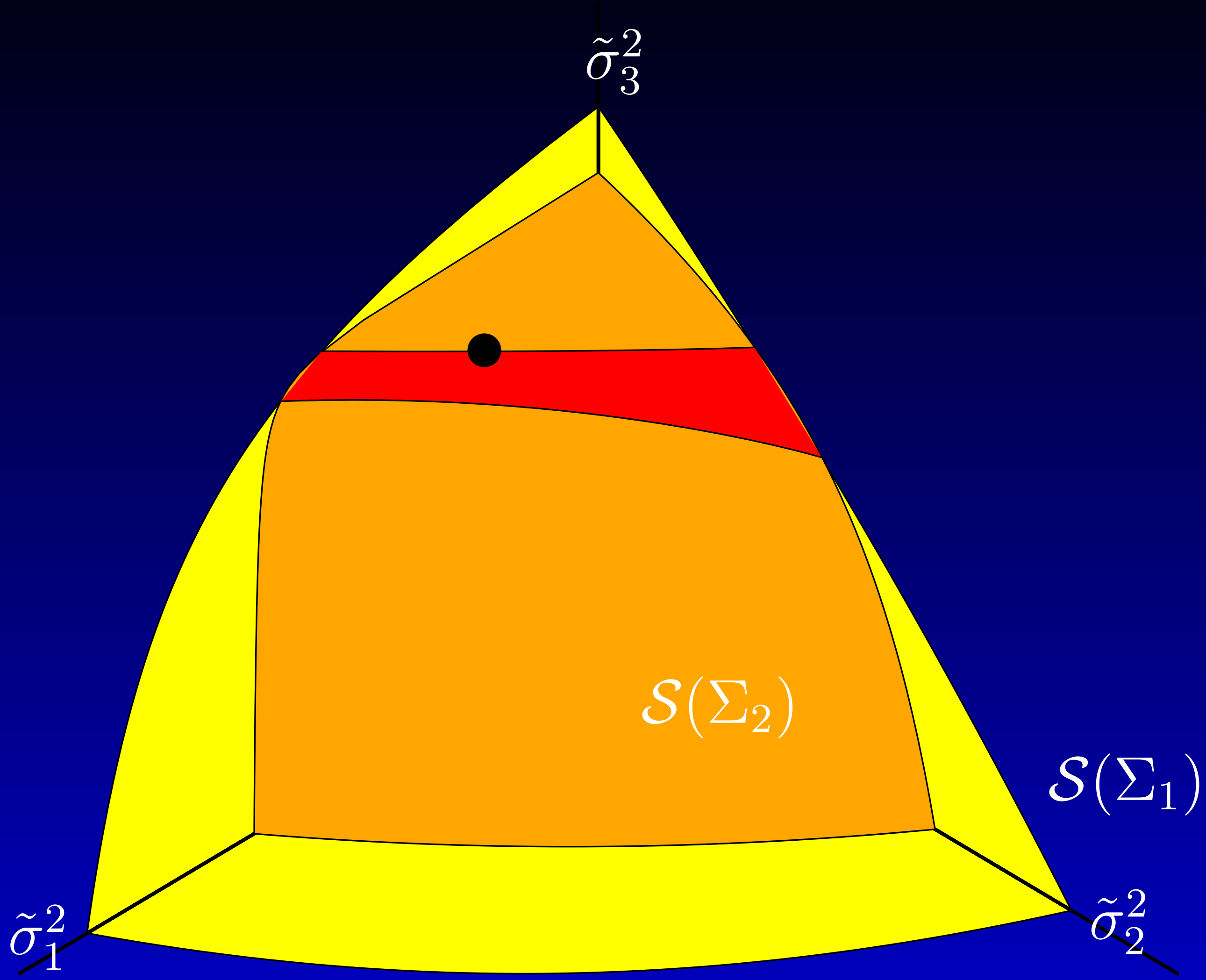


Figure 4: Admissible noise points

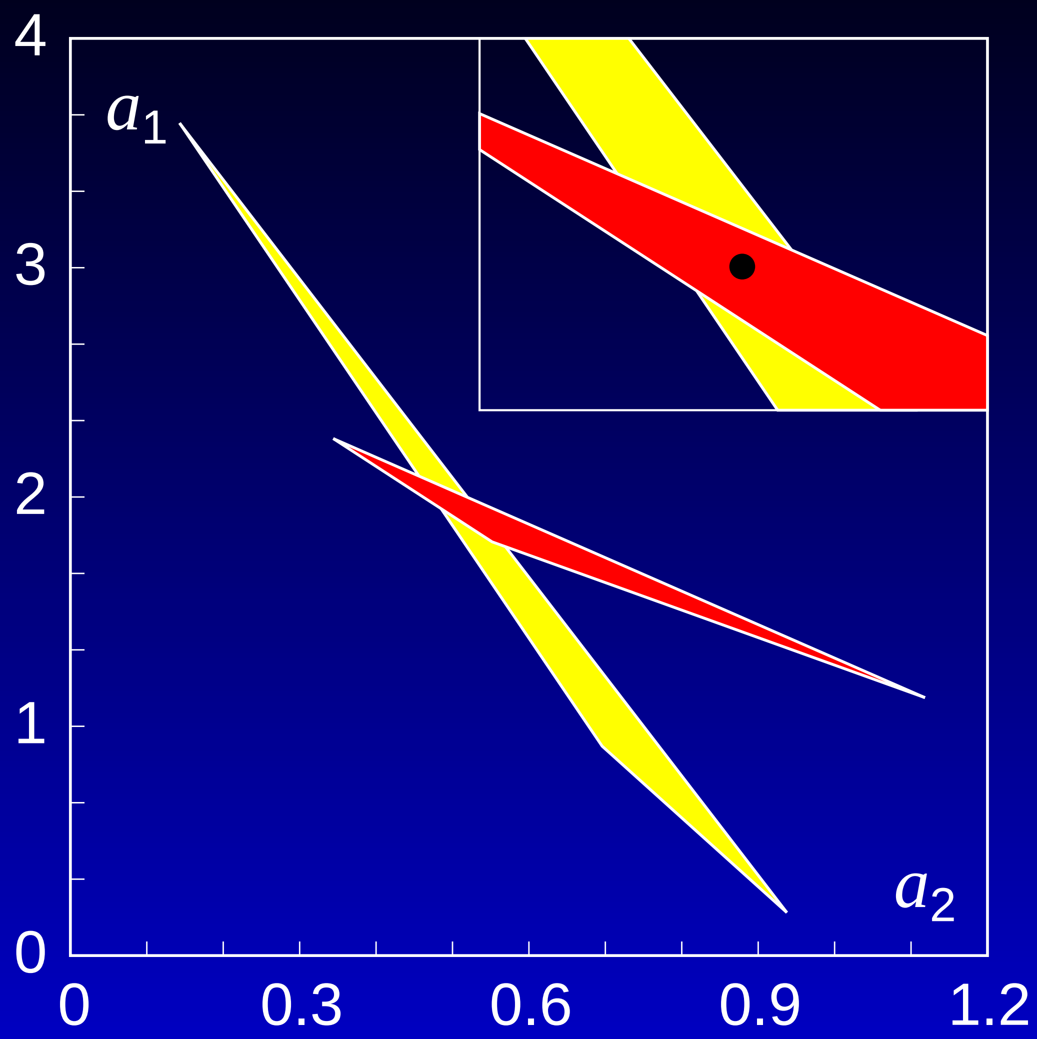


Figure 5: Admissible model parameters

Determination of the Frisch solution from real data

Theorem 10 allows defining a consistent criterion to search for solutions even when the intersection between $\mathcal{S}(\Sigma_1)$ and $\mathcal{S}(\Sigma_2)$ does not contain any point mapped, by Σ_1 and Σ_2 into the same point of the parameter space.

Criterion 1 (Guidorzi and Diversi, 2006) – Consider a pair of covariance matrices Σ_1 and Σ_2 and their loci of solutions, $\mathcal{S}(\Sigma_1)$, $\mathcal{S}(\Sigma_2)$ in the noise space. The best approximation of the actual noise variances will be given by the point $P \in \mathcal{S}(\Sigma_1) \cap \mathcal{S}(\Sigma_2)$ that minimizes the euclidean norm of the distance between the parameter vectors a' and a'' associated to P by Σ_1 and Σ_2.

Remark 2 – Criterion 1 is consistent since the cost function $f(P) = \|a' - a''\|_2$ annihilates when Σ_1 and Σ_2 are independent.

Remark 3 – Once that the minimum of $f(P)$ has been found, two solutions, a' and a'' will be available and their distance is a measure of the reliability of the procedure. Their mean value can be assumed as problem solution.

Remark 4 – It can be observed that the outlined procedure can be applied even in the case of simplexes without common points.

A numerical example

Two independent sets of noise–free data, concerning $N = 100$ observations of $n = 3$ variables are characterized by the sample covariance matrices

$$
\begin{aligned}
\hat{\Sigma}_1 &= \frac{\hat{X}_1^T \hat{X}_1}{N} = \begin{bmatrix} 3 & 12 & -12 \\ 12 & 56 & -52 \\ -12 & -52 & 50 \end{bmatrix} \\
\hat{\Sigma}_2 &= \frac{\hat{X}_2^T \hat{X}_2}{N} = \begin{bmatrix} 14 & 18 & -37 \\ 18 & 36 & -54 \\ -37 & -54 & 101 \end{bmatrix}.
\end{aligned}
$$

$\hat{\Sigma}_1$ and $\hat{\Sigma}_2$ have rank 2 and are associated with the same linear relation described by $a_1 = 2$, $a_2 = 0.5$ and $a_3 = 1$.

A Monte Carlo simulation of 100 runs has been performed by generating, in every run, two independent sets of three gaussian white sequences and by adding these sequences to the noise–free data in order to obtain the noisy ones.

	a_1	a_2
true	2	0.5
estim. 1	2.0320 ± 0.1437	0.4945 ± 0.0610
estim. 2	2.0305 ± 0.1424	0.4802 ± 0.0724

Table 1: True and estimated values of the coefficients a_1, a_2

References

S. Beghelli, R. Guidorzi and U. Soverini (1990). The Frisch scheme in dynamic system identification, Automatica, Special Issue on Identification and System Parameter Estimation, vol. 26, n. 1, pp. 171–176.

M. Deistler and W. Scherrer (1992). Identification of linear systems from noisy data, in New Directions in Time Series Analysis, Part II, D. Brillinger et al., eds., IMA Vol. Math. Appl., 46, Springer-Verlag, pp. 21–42.

R. Frisch (1934). Statistical confluence analysis by means of complete regression systems, pub. no. 5, Economic Institute, Oslo University.

R. Guidorzi (1991). Certain models from uncertain data: the algebraic case, Systems and Control Letters, vol. 17, n. 6, pp. 415–424.

R. Guidorzi (1995). Identification of the maximal number of linear relations from noisy data, Systems and Control Letters, vol. 24, n. 3, pp. 159-166.

R. Guidorzi and R. Diversi (2006). Determination of linear relations from real data in the Frisch scheme context, Proceedings of the 17th International Symposium on Mathematical Theory of Networks and Systems, pp. 530–535, Kyoto.

R. Guidorzi and A. Stoian (1994). On the computation of the maximal corank of a covariance matrix under the Frisch scheme, Preprints of the 10th IFAC Symposium on System Identification, vol. 3, pp. 171–173, Copenhagen.

R.E. Kalman (1982a). System identification from noisy data, in Dynamical Systems II, A.R. Bednarek and L. Cesari, eds., Academic Press, pp. 135–164.

R.E. Kalman (1982b). Identification from real data, in Current Developments in the Interface: Economics, Econometrics, Mathematics, M. Hazewinkel and H.G. Rinnooy Kan, eds., Reidel, Doordrecht, pp. 161–196.

W. Schachermayer and M. Deistler (1998). The set of observationally equivalent errors–in–variables models, Systems & Control Letters, vol. 34, pp. 101–104.

Some Issues on Errors–in–Variables Identification

II. The Frisch scheme in the dynamic case

Roberto Guidorzi, Roberto Diversi, Umberto Soverini

`rguidorzi,rdiversi,usoverini@deis.unibo.it`

Bologna University

The Frisch scheme in the dynamic case

Consider a dynamic SISO system of order n

$$\hat{y}(t+n) = \sum_{k=1}^{n} \alpha_k \, \hat{y}(t+k-1) + \sum_{k=1}^{n+1} \beta_k \, \hat{u}(t+k-1) \qquad (1)$$

and noisy input/output measures

$$u(t) = \hat{u}(t) + \tilde{u}(t) \qquad (2a)$$

$$y(t) = \hat{y}(t) + \tilde{y}(t) \qquad (2b)$$

where $\tilde{u}(t)$ and $\tilde{y}(t)$ are white processes with zero mean, mutually uncorrelated and uncorrelated with $\hat{u}(t)$.

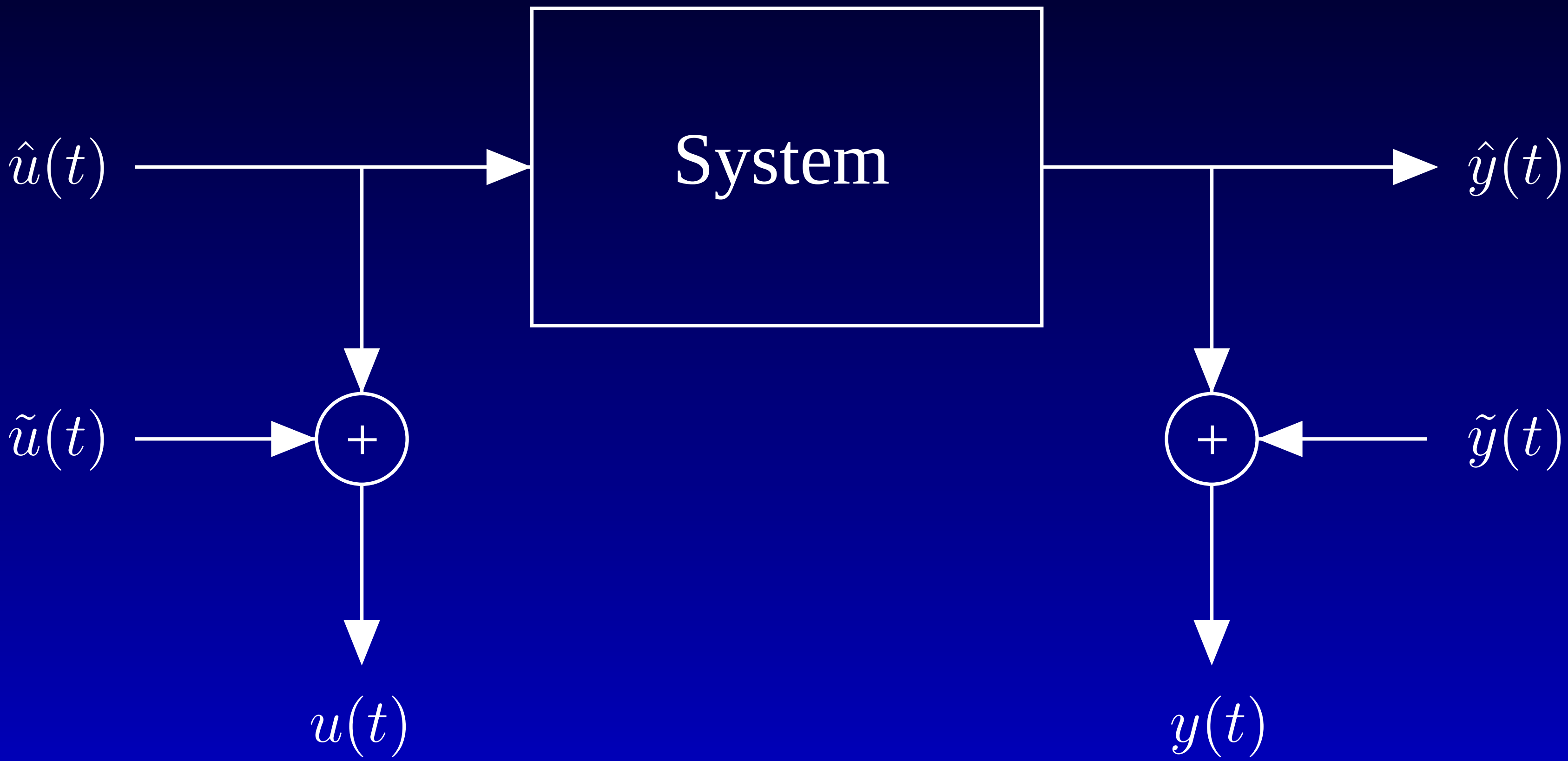


Figure 1: The dynamic Frisch scheme context

Define now the Hankel matrices

$$X_k(y) = \begin{bmatrix} y(1) & \dots & y(k) \\ y(2) & \dots & y(k+1) \\ \vdots & \ddots & \vdots \\ y(N) & \dots & y(k+N-1) \end{bmatrix}, \tag{3a}$$

$$X_k(u) = \begin{bmatrix} u(1) & \dots & u(k) \\ u(2) & \dots & u(k+1) \\ \vdots & \ddots & \vdots \\ u(N) & \dots & u(k+N-1) \end{bmatrix}, \tag{3b}$$

the matrix of input/output samples

$$X_k = [\, X_{k+1}(y) \; X_{k+1}(u) \,] \tag{4}$$

and the sample covariance matrices Σ_k given by

$$\Sigma_k = \frac{X_k^T X_k}{N} = \begin{bmatrix} \Sigma(yy) & \Sigma(yu) \\ \Sigma(uy) & \Sigma(uu) \end{bmatrix}. \tag{5}$$

Denoting with $\tilde{\sigma}_u^{*2}$ and $\tilde{\sigma}_y^{*2}$ the variances of $\tilde{u}(t)$ and $\tilde{y}(t)$ and with P^* the point

$$P^* = (\,\tilde{\sigma}_y^{*2}, \tilde{\sigma}_u^{*2}\,), \tag{6}$$

the previous assumptions establish, when $N \to \infty$, that

$$\Sigma_k = \hat{\Sigma}_k + \tilde{\Sigma}_k^* \tag{7}$$

where

$$\tilde{\Sigma}_k^* = \mathrm{diag}\,[\tilde{\sigma}_y^{*2} I_{k+1},\ \tilde{\sigma}_u^{*2} I_{k+1}]. \tag{8}$$

The identification problem

The identification problem, in the context of the Frisch scheme, consists in determining the order and the parameters of model (1), or of any equivalent state–space model, and the additive noise variances $\tilde{\sigma}_y^{*2}$, $\tilde{\sigma}_u^{*2}$ on the basis of the knowledge of the noisy sequences $u(\cdot)$, $y(\cdot)$ or, equivalently, of the sequence of increasing–dimension matrices Σ_k for $k = 1, 2, \ldots$.

Model (1) implies, for every input sequence persistently exciting of order n, the nonsingularity of $\hat{\Sigma}_1, \ldots, \hat{\Sigma}_{n-1}$ and the singularity of $\hat{\Sigma}_k$ for $k \geq n$.

For any value of k (lower, equal or larger than n), a point $P = (\tilde{\sigma}_y^2, \tilde{\sigma}_u^2)$ belonging to the first orthant of the noise space, defines an admissible solution if and only if

$$\dim \ker(\Sigma_k - \tilde{\Sigma}_k) = 1, \tag{9}$$

$$\Sigma_k - \tilde{\Sigma}_k \geq 0 \tag{10}$$

where $\tilde{\Sigma}_k$ is the noise covariance matrix defined by P

$$\tilde{\Sigma}_k = \tilde{\Sigma}_k(P) = \operatorname{diag}\left[\tilde{\sigma}_y^2\, I_{k+1},\; \tilde{\sigma}_u^2\, I_{k+1}\right]. \tag{11}$$

The corresponding solution in the parameter space, $\theta(P)$, is univocally defined by $\ker\left(\Sigma_k - \tilde{\Sigma}_k\right)$, i.e. by the relation

$$\hat{\Sigma}_k\, \theta(P) = (\Sigma_k - \tilde{\Sigma}_k)\, \theta(P) = 0. \tag{12}$$

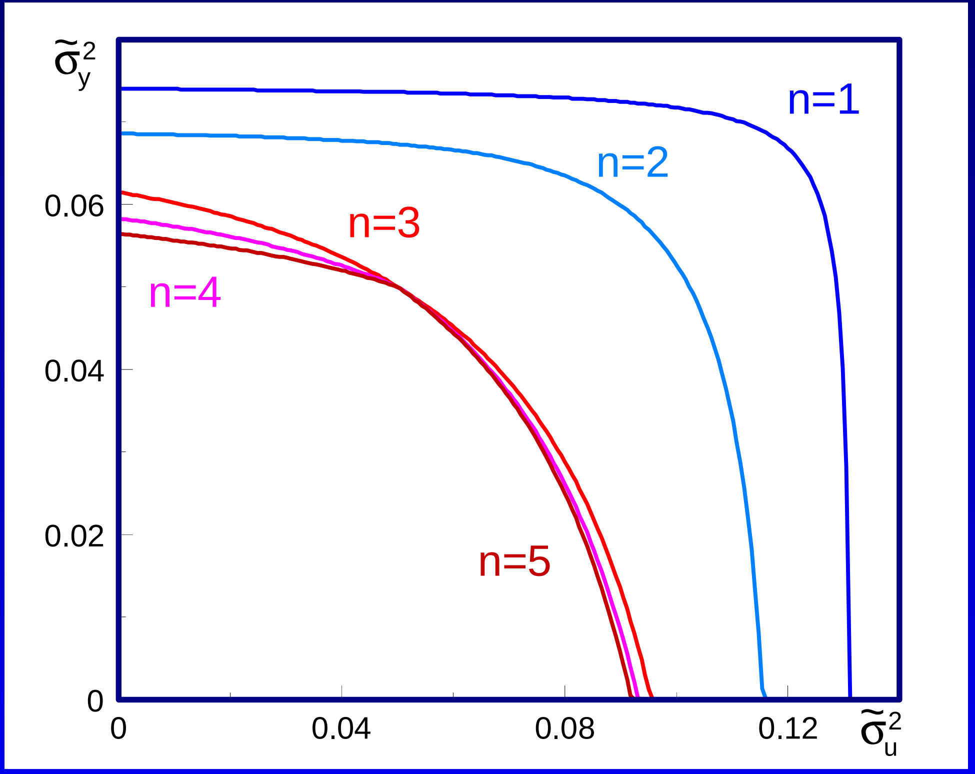
Theorem 1 (Beghelli, Guidorzi and Soverini, 1990) – For every $k > 0$ all admissible points define a convex curve $\mathcal{S}(\Sigma_k)$ in the first quadrant of the noise plane $\mathcal{R}^2$ with a concavity facing the origin. The point $P^* = (\tilde{\sigma}_y^{2*},\, \tilde{\sigma}_u^{2*})$ associated with the actual noise variances belongs to all curves $\mathcal{S}(\Sigma_k)$ when $k \geq n$ and $\theta(P^*)$ is the true parameter vector, θ^*.

Theorem 2 (Beghelli, Guidorzi and Soverini, 1990) – If i and j are integers with $j > i$, then $\mathcal{S}(\Sigma_j)$ lies under or on $\mathcal{S}(\Sigma_i)$.

Example 1 – The figure that follows shows the curves $\mathcal{S}(\Sigma_1), \ldots$
$\ldots, \mathcal{S}(\Sigma_5)$ for data generated by the third order system

$$\begin{aligned}\hat{y}(t+3) &= 0.4\,\hat{y}(t+2) - 0.3\,\hat{y}(t+1) - 0.1\,\hat{y}(t) \\ &\quad +0.2\,\hat{u}(t+2) - 0.38\,\hat{u}(t+1) + 0.58\,\hat{u}(t)\end{aligned}$$

for $\sigma_u^{*2} = 0.05$ and $\sigma_y^{*2} = 0.05$.



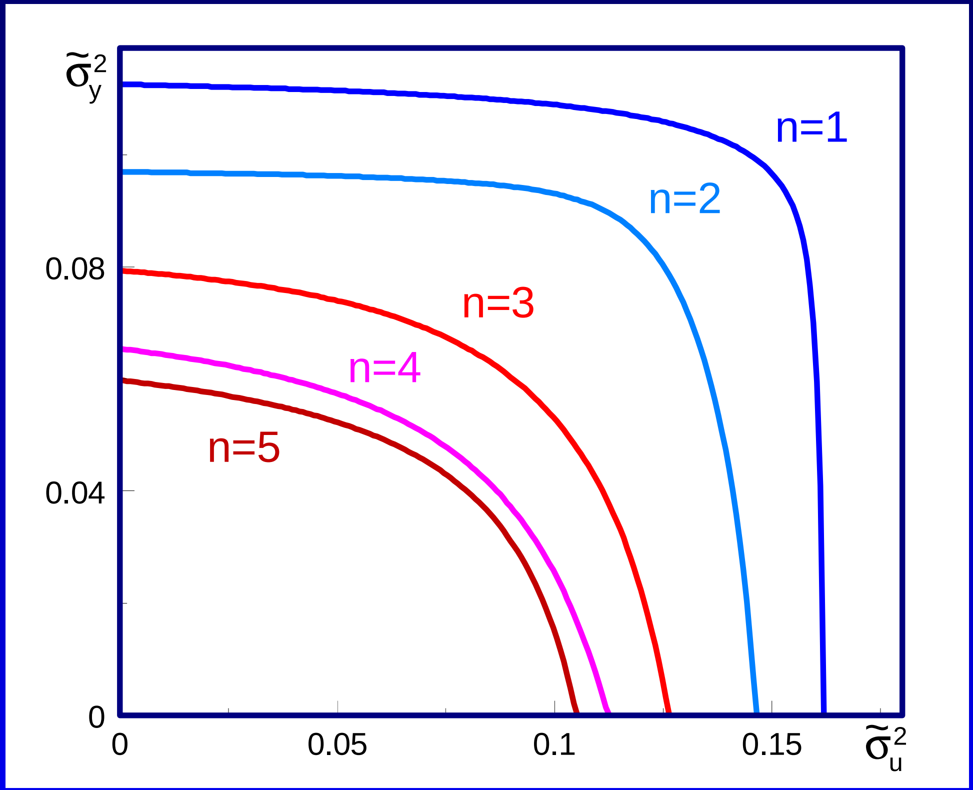
Problems in the Frisch identification of real processes

1) The key property described by Theorem 1 holds only when the (asymptotic) properties assumed for the additive noise sequences (mutual orthogonality and orthogonality with the input/output sequences) hold, i.e. when $\tilde{u}(\cdot)$ and $\tilde{y}(\cdot)$ are uncorrelated white sequences with infinite length.

2) Similar consequences follow from violations on the linearity and time–invariance assumptions.

3) The algorithms that can be developed to estimate a single solution from real data can exhibit robustness and reliability problems and require the development of suitable criteria.

Example 2 – The process considered is a natural gas reservoir converted to storage operations. The model orientation considers as input the total amount of injected/extracted gas and as output the mean reservoir pressure. The process exhibits a non stationary behavior because of the volume variations due to water encroaching.



Frisch identification criteria

The shifted relation criterion (Beghelli, Castaldi, Guidorzi and Soverini, 1993)

This method is based on the following rank deficiency property of the matrices $\hat{\Sigma}_k(P^*)$ for $k \geq n$:

- if $k \geq n$ the dimension of the null space of $\hat{\Sigma}_k(P^*)$ and, consequently, the multiplicity of its least eigenvalue, is equal to $(k - n + 1)$;
- for $k > n$ all linear dependence relations between the columns of the matrices $\hat{\Sigma}_k(P^*)$ are described by the same set of coefficients θ^*.

When $k = n$, $\ker \hat{\Sigma}_n(P^*) = \operatorname{im} \theta^*$ while when $k = n+1$

$$\ker \hat{\Sigma}_{n+1}(P^*) = \operatorname{im}\,[v' \; v''], \tag{13}$$

where

$$v' = [0 \;\; \alpha_1 \;\; \ldots \;\; \alpha_n \; -1 \;\; 0 \;\; \beta_1 \;\; \ldots \;\; \beta_{n+1}]^T \tag{14}$$

$$v'' = [\alpha_1 \;\; \ldots \;\; \alpha_n \; -1 \;\; 0 \;\; \beta_1 \;\; \ldots \;\; \beta_{n+1} \;\; 0]^T. \tag{15}$$

Consider now the intersections $P' = (\tilde{\sigma}_y^{2'}, \tilde{\sigma}_u^{2'})$, $P'' = (\tilde{\sigma}_y^{2''}, \tilde{\sigma}_u^{2''})$ of a line from the origin with $\mathcal{S}(\Sigma_n)$ and $\mathcal{S}(\Sigma_{n+1})$, so that

$$\frac{\tilde{\sigma}_y^{2'}}{\tilde{\sigma}_u^{2'}} = \frac{\tilde{\sigma}_y^{2''}}{\tilde{\sigma}_u^{2''}}, \tag{16}$$

and define the cost function (Diversi, Guidorzi and Soverini, 2004)

$$J(P', P'') = \mathrm{trace}\left([v'(P')\; v''(P')]^T \right. \tag{17}$$

$$\left. \times \hat{\Sigma}_{n+1}(P'')\,[v'(P')\; v''(P')]\right),$$

where $v'(P'), v''(P')$ have been constructed with the entries of $\theta(P')$. This function exhibits the following properties:

$$J(P', P'') \geq 0 \tag{18}$$

$$J(P', P'') = 0 \Leftrightarrow P' = P'' = P^*. \tag{19}$$

It is thus possible to perform the identification by searching for the solution that minimizes (17).

The covariance–matching criterion (Diversi, Guidorzi and Soverini, 2003)

Consider the residual $\gamma(t)$ of the EIV process

$$\begin{aligned}\gamma(t) &= \alpha_1\, y(t) + \cdots + \alpha_n\, y(t+n-1) - y(t+n) \\ &\quad + \beta_1\, u(t) + \cdots + \beta_{n+1}\, u(t+n); \qquad (20)\end{aligned}$$

that can also be written as

$$\begin{aligned}\gamma(t) &= \alpha_1\, \tilde{y}(t) + \cdots + \alpha_n\, \tilde{y}(t+n-1) - \tilde{y}(t+n) \\ &\quad + \beta_1\, \tilde{u}(t) + \cdots + \beta_{n+1}\, \tilde{u}(t+n), \qquad (21)\end{aligned}$$

i.e., as the sum of two MA processes driven by the white noises $\tilde{y}(t)$ and $\tilde{u}(t)$.

Because of the assumptions on $\tilde{y}(t)$ and $\tilde{u}(t)$ the autocorrelations of $\gamma(t)$, $r_\gamma(k) = E\left[\gamma(t)\,\gamma(t-k)\right]$, are given by

$$r_\gamma(0) = \tilde{\sigma}_y^{2*} \sum_{i=1}^{n+1} \alpha_i^2 + \tilde{\sigma}_u^{2*} \sum_{i=1}^{n+1} \beta_i^2 \tag{22}$$

$$r_\gamma(k) = \tilde{\sigma}_y^{2*} \sum_{i=1}^{n-k+1} \alpha_i\,\alpha_{i+k} + \tilde{\sigma}_u^{2*} \sum_{i=1}^{n-k+1} \beta_i\,\beta_{i+k} \tag{23}$$

$$\text{for } k = 1, \ldots, n$$

$$r_\gamma(k) = 0 \quad \text{for } k > n \tag{24}$$

where $\alpha_{n+1} = -1$.

Define now, for every point $P = (\tilde{\sigma}_y^2, \tilde{\sigma}_u^2)$ of $\mathcal{S}(\Sigma_n)$ the vector

$$r_k(P) = [r_\gamma(0,P)\, r_\gamma(1,P)\, \ldots\, r_\gamma(k,P)]^T, \qquad (25)$$

with entries computed by means of (22)–(24) using the variances $(\tilde{\sigma}_y^2, \tilde{\sigma}_u^2)$ and the parameters $\theta(P)$.

Compute also, by means of the available data and $\theta(P)$, the sample vector

$$\bar{r}_k(P) = [\bar{r}_\gamma(0,P)\, \bar{r}_\gamma(1,P)\, \ldots\, \bar{r}_\gamma(k,P)]^T, \qquad (26)$$

where

$$\bar{r}_\gamma(k,P) = \frac{1}{N}\sum_{t=1}^{n} \gamma(t)\,\gamma(t+k). \qquad (27)$$

Since, for $N \to \infty$

$$r_k(P^*) = \bar{r}_k(P^*) = [r_\gamma(0)\, r_\gamma(1)\ \ldots\ r_\gamma(k)]^T, \tag{28}$$

the following covariance–matching cost function can be considered

$$J(P) = \|r_k(P) - \bar{r}_k(P)\|_2\,, \tag{29}$$

that compares the theoretical statistical properties of $\gamma(t)$ with those computed from the data.

The identification problem can thus be solved by minimizing $J(P)$ along $\mathcal{S}(\Sigma_n)$.

A criterion based on Yule–Walker equations (Diversi, Guidorzi and Soverini, 2006)

Define the regressor vector

$$\varphi(t) = [y(t-n) \; \ldots \; y(t-1) \; -y(t) \, u(t-n) \; \ldots \; u(t)]^T \tag{30}$$

and the $q \times 1$ vector of delayed inputs

$$\varphi_u^h(t) = [u(t-n-q) \; \ldots \; u(t-n-1)]^T. \tag{31}$$

Consider then the $q \times (2n+2)$ matrix

$$\Sigma^h = E\,[\varphi_u^h(t)\, \varphi^T(t)]. \tag{32}$$

If $q \geq 2n+1$, it is easy to show that

$$\Sigma^h = E\left[\hat{\varphi}_u^h(t)\,\hat{\varphi}^T(t)\right], \tag{33}$$

where

$$\hat{\varphi}(t) = [\hat{y}(t-n)\ \ldots\ \hat{y}(t-1)\ -\hat{y}(t)\,\hat{u}(t-n)\ \ldots\ \hat{u}(t)]^T \tag{34}$$

$$\hat{\varphi}_u^h(t) = [\hat{u}(t-n-q)\ \ldots\ \hat{u}(t-n-1)]^T. \tag{35}$$

Since $\hat{\varphi}^T(t)\,\theta^* = 0$ it follows that

$$\Sigma^h\,\theta^* = 0. \tag{36}$$

- Relation (36) represents a set of high order Yule–Walker equations that could be directly used to estimate the parameter vector θ^*.
- This approach can also be seen as an instrumental variable method that uses delayed inputs as instruments.

• If $q \geq 2n+1$, θ^* can be consistently identified from Σ^h. The search for the point P^* on $\mathcal{S}(\Sigma_n)$ can be performed by means of the cost function

$$J(P) = \|\Sigma^h \, \theta(P)\|_2^2 = \theta^T(P)(\Sigma^h)^T \, \Sigma^h \, \theta(P) \qquad (37)$$

that exhibits the following properties

$$J(P) \geq 0$$

$$J(P) = 0 \Leftrightarrow P = P^*$$

Multivariable Frisch identification

The extension of Frisch identification techniques to the MISO case is straightforward but this is not the case for MIMO processes that face conceptual and practical congruence problems not present in the single–output case.

The solution of this problem (Guidorzi, 1996, Guidorzi, Soverini and Diversi, 2002) has required the introduction of new parametrizations of the Frisch singularity surfaces (Guidorzi and Pierantoni, 1995) that associate models to *all directions* in the noise space instead than to single (admissible) points.

Example 3 – A Monte Carlo simulation of 100 independent runs has been performed on the system

$$P(z)=\begin{bmatrix} z^2-0.4z+0.3 & 0.1975 \\ -0.2026z+0.1013 & z-0.4 \end{bmatrix},$$

$$Q(z)=\begin{bmatrix} 0.3426z+0.7194 \\ 0.7979 \end{bmatrix}$$

The input sequence $\hat{u}(\cdot)$ is a PRBS with length $N=300$. The variances of the noiseless output sequences $\hat{y}_1(\cdot)$, $\hat{y}_2(\cdot)$ are equal to 1. The noise variances are

$$\tilde{\sigma}_u^{2*}=0.04 \qquad \tilde{\sigma}_{y_1}^{2*}=0.16 \qquad \tilde{\sigma}_{y_2}^{2*}=0.36$$

corresponding to amounts of 20%, 40% and 60% in standard deviation.

	α_{111}	α_{112}	α_{121}
true	0.3000	-0.4000	0.19750
ident.	0.2781 ± 0.08	-0.4246 ± 0.09	0.2363 ± 0.14
	α_{211}	α_{212}	α_{221}
true	0.1013	-0.2026	-0.4000
ident.	0.090 ± 0.10	-0.2165 ± 0.12	-0.3748 ± 0.16

	β_{111}	β_{112}	β_{121}
true	0.7194	0.3426	0.7979
ident.	0.7167 ± 0.04	0.3423 ± 0.03	0.7997 ± 0.06

	$\tilde{\sigma}_u^*$	$\tilde{\sigma}_{y_1}^*$	$\tilde{\sigma}_{y_2}^*$
true	0.0400	0.1600	0.3600
ident.	0.0400 ± 0.02	0.1537 ± 0.02	0.3746 ± 0.04

References

S. Beghelli, P. Castaldi, R. Guidorzi and U. Soverini (1993). A robust criterion for model selection in identification from noisy data, Proceedings of Ninth International Conference on Systems Engineering, Las Vegas, Nevada, pp. 480–484.

S. Beghelli, R. Guidorzi and U. Soverini (1990). The Frisch scheme in dynamic system identification, Automatica, Special Issue on Identification and System Parameter Estimation, vol. 26, n. 1, pp. 171–176.

R. Diversi, R. Guidorzi and U. Soverini (2003). A new criterion in EIV identification and filtering applications, Preprints of the 13th IFAC Symposium on System Identification, pp. 1993–1998, Rotterdam.

R. Diversi, R. Guidorzi and U. Soverini (2004). Frisch scheme-based algorithms for EIV identification, Proceedings of the 12th IEEE Mediterranean Conference on Control and Automation, Kusadasi, Turkey.

R. Diversi, R. Guidorzi and U. Soverini (2006). Yule-Walker equations in the Frisch scheme solution of errors-in-variables identification problems, Proceedings of the 17th International Symposium on Mathematical Theory of Networks and Systems, pp. 391–395, Kyoto.

R. Guidorzi (1996). Identification of multivariable processes in the Frisch scheme context, MTNS '96, St. Louis.

R. Guidorzi and M. Pierantoni (1995). A new parametrization of Frisch scheme solutions, Proc. of the XII Int. Conf. on Systems Science, pp. 114–120, Wroclaw.

R. Guidorzi, U. Soverini and R. Diversi (2002). Multivariable EIV identification, Proceedings of the 10th IEEE Mediterranean Conference on Control and Automation, Lisboa.

Some Issues on Errors–in–Variables Identification

III. Applications of Frisch identification, EIV interpolation and filtering

Roberto Guidorzi, Roberto Diversi, Umberto Soverini

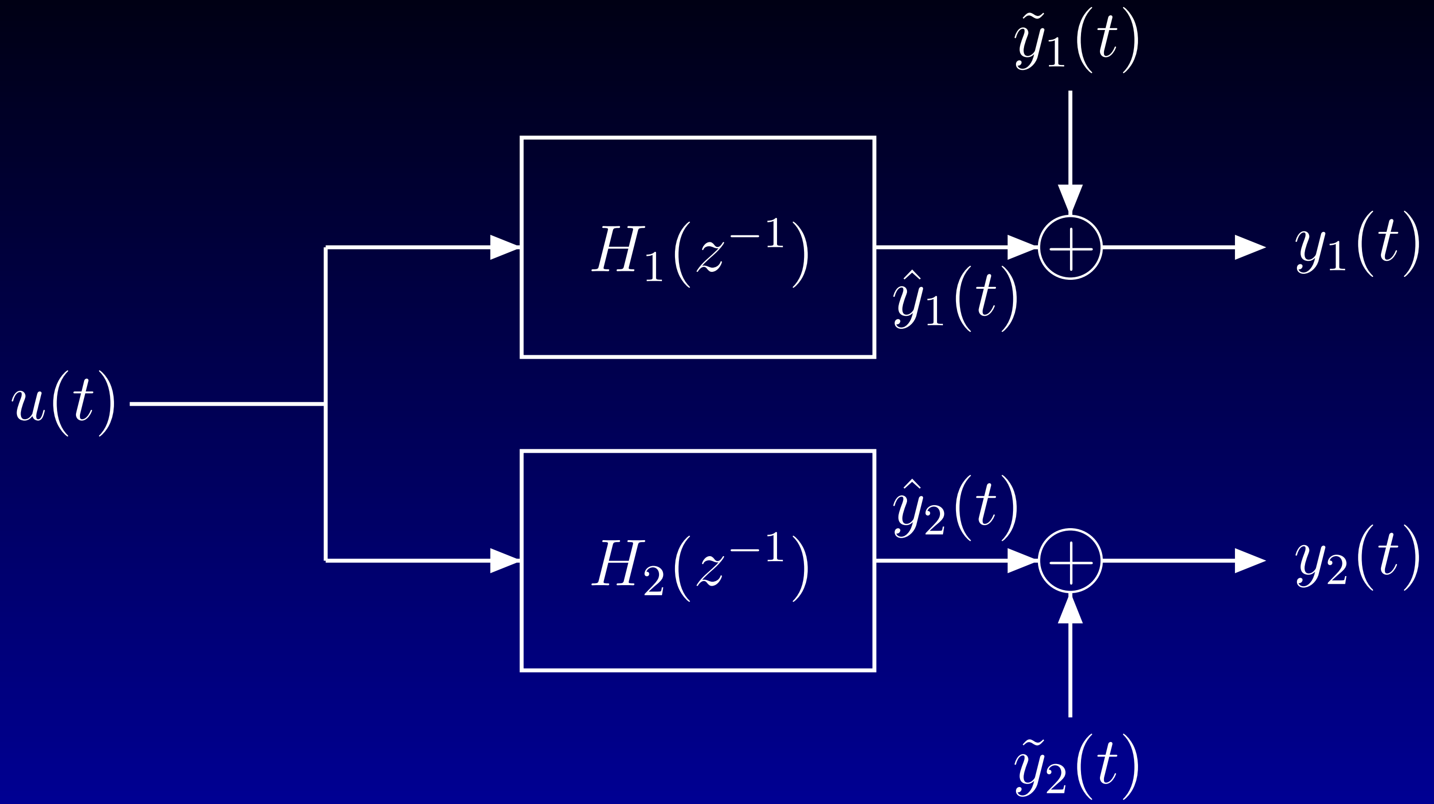
`rguidorzi,rdiversi,usoverini@deis.unibo.it`

Bologna University

Blind identification of SIMO FIR systems (Diversi, Guidorzi and Soverini, 2005a, Guidorzi, Diversi and Soverini, 2006)

- The blind identification of dynamic systems is of great relevance in many fields like telecommunications, sismology, radioastronomy, etc. The purpose is the reconstruction of the transfer function of a transmission channel starting from noisy measurements performed only on its output.

- Blind identification relies on linear models describing a set of parallel channels driven by an unknown sequence and characterized by a finite impulse response (FIR). These models can describe a single unknown source in presence of multiple spatially and/or temporally distributed sensors.



$$\hat{y}_i(t) = H_i(z^{-1})\,u(t) = \sum_{k=0}^{n} h_i(k)\,u(t-k), \qquad i = 1,2 \quad (1)$$

$$H_i(z^{-1}) = h_i(0) + h_i(1)\,z^{-1} + \cdots + h_i(n)\,z^{-n}, \quad i = 1,2 \quad (2)$$

$$y_i(t) = \hat{y}_i(t) + \tilde{y}_i(t), \quad i = 1,2 \qquad (3)$$

Relations (1) lead immediately to the well–known cross–relation property

$$H_2(z^{-1})\,\hat{y}_1(t) = H_1(z^{-1})\,\hat{y}_2(t). \tag{4}$$

It is thus possible to write

$$\Big[X_{n+1}(\hat{y}_1)\ \ X_{n+1}(\hat{y}_2)\Big]\, h = 0\ , \tag{5}$$

where

$$h = [h_2(n)\ \cdots\ h_2(0)\ \ -h_1(n)\ \cdots\ -h_1(0)]^T \tag{6}$$

and

$$X_{n+1}(\hat{y}_i) = \begin{bmatrix} \hat{y}_i(1) & \dots & \hat{y}_i(n+1) \\ \vdots & & \vdots \\ \hat{y}_i(N) & \dots & \hat{y}_i(N+n) \end{bmatrix}, \quad i = 1, 2. \tag{7}$$

Define now the covariance matrix

$$\hat{\Sigma}_n = \lim_{N\to\infty} \frac{1}{N} \Big[X_{n+1}(\hat{y}_1)\, X_{n+1}(\hat{y}_2)\Big]^T \Big[X_{n+1}(\hat{y}_1)\;\, X_{n+1}(\hat{y}_2)\Big]. \quad (8)$$

It follows that

$$\hat{\Sigma}_n\, h = 0 \quad (9)$$

$$\Sigma_n = \hat{\Sigma}_n + \tilde{\Sigma}_n^* \quad (10)$$

$$\tilde{\Sigma}_n^* = \mathrm{diag}\,[\tilde{\sigma}_{y1}^{2*}\, I_{n+1},\; \tilde{\sigma}_{y2}^{2*}\, I_{n+1}], \quad (11)$$

where Σ_n and $\tilde{\Sigma}_n^*$ can be obtained by inserting $X_{n+1}(y_i)$ and $X_{n+1}(\tilde{y}_i)$ in (8).

The blind identification problem has thus been mapped into an errors–in–variables identification problem.

Identification of noisy autoregressive models (Diversi, Soverini and Guidorzi, 2005, Diversi, Guidorzi and Soverini, 2005d)

- Autoregressive (AR) models are commonly used in a wide range of engineering applications, like spectral estimation, speech and image processing, noise cancellation etc.
- A considerable attention has been dedicated, in the literature, to the problem of estimating the AR parameters from signals corrupted by white noise.
- In this case the estimates obtained with classical AR identification methods (least–squares, Yule–Walker equations) are poor, particularly for low signal–to–noise ratio conditions.

Problem statement

Consider the noisy AR model:

$$x(t) = \alpha_1 \, x(t-1) + \cdots + \alpha_n \, x(t-n) + e(t), \qquad (12)$$

$$y(t) = x(t) + w(t), \qquad (13)$$

where $x(t)$ is the noise–free AR signal, $e(t)$ is the driving noise and $y(t)$ is the available observation affected by the additive noise $w(t)$.

Assumptions: $e(t)$ and $w(t)$ are zero–mean white processes, mutually uncorrelated, with unknown variances σ_e^{2*} and σ_w^{2*}.

Problem: Estimate $\alpha_1, \ldots, \alpha_n$ and σ_e^{2*}, σ_w^{2*} starting from the available measurements $y(1), y(2), \ldots, y(N)$.

Mapping the noisy AR problem into an EIV problem

By defining the vectors

$$\begin{aligned}\varphi_x(t) &= [x(t-n)\ \ldots\ x(t-1)\ x(t)]^T, \\ \varphi_y(t) &= [y(t-n)\ \ldots\ y(t-1)\ y(t)]^T, \\ \varphi_w(t) &= [w(t-n)\ \ldots\ w(t-1)\ w(t)]^T,\end{aligned}$$

and the parameter vector

$$\theta^* = [\,\alpha_n\ \cdots\ \alpha_1\ -1]^T,$$

it is possible to write model (12)–(13) in the form

$$\left(\varphi_x^T(t) - [0\ \ldots\ 0\ e(t)]\right)\ \theta^* = 0, \tag{14}$$

$$\varphi_y(t) = \varphi_x(t) + \varphi_w(t). \tag{15}$$

Define now the $(n+1) \times (n+1)$ covariance matrix

$$\begin{aligned}\hat{\Sigma}_n &= E\left[\varphi_x(t)\,\varphi_x^T(t)\right] - \operatorname{diag}\,[\underbrace{0\;\ldots\;0}_{n}\;\;\sigma_e^{2*}] \\ &= \begin{bmatrix} r_x(0) & r_x(1) & \cdots & r_x(n) \\ r_x(1) & r_x(0) & \cdots & r_x(n-1) \\ \vdots & \vdots & \ddots & \vdots \\ r_x(n) & r_x(n-1) & \cdots & r_x(0) - \sigma_e^{2*} \end{bmatrix},\end{aligned}$$

where $r_x(k) = r_x(-k) = E\left[x(t)\,x(t-k)\right]$.

From relation (14) it follows that

$$\hat{\Sigma}_n\,\theta^* = 0. \tag{16}$$

Since $\varphi_y(t) = \varphi_x(t) + \varphi_w(t)$, it follows that the covariance matrix of the noisy observation is given by

$$\Sigma_n = E\left[\varphi_y(t)\, \varphi_y^T(t)\right] = \hat{\Sigma}_n + \tilde{\Sigma}_n^* \tag{17}$$

where

$$\tilde{\Sigma}_n^* = \begin{bmatrix} \sigma_w^{2*} & 0 & \cdots & \cdots & 0 \\ 0 & \sigma_w^{2*} & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \sigma_w^{2*} & 0 \\ 0 & \cdots & \cdots & 0 & (\sigma_w^{2*} + \sigma_e^{2*}) \end{bmatrix} = \text{diag}\,[\sigma_w^{2*}\, I_n, \sigma_s^{2*}],$$

with $\sigma_s^{2*} = \sigma_w^{2*} + \sigma_e^{2*}$.

EIV Interpolation (Guidorzi, Diversi and Soverini, 2002, 2003)

Consider a SISO EIV process of order n

$$\hat{y}(t+n) = \sum_{k=1}^{n} \alpha_k \, \hat{y}(t+k-1) + \sum_{k=1}^{n+1} \beta_k \, \hat{u}(t+k-1) \tag{18}$$

$$u(t) = \hat{u}(t) + \tilde{u}(t) \tag{19a}$$

$$y(t) = \hat{y}(t) + \tilde{y}(t) \tag{19b}$$

When L input–output samples are available, equation (18) can be evaluated at times $t = n+1, \ldots, L$. It is thus possible to write $N = L - n$ relations in the form

$$G \, \hat{v} = 0 \tag{20}$$

where G is the $N \times 2(N+n)$ matrix

$$
G = \left[\begin{array}{ccccccccc}
\alpha_1 & \beta_1 & \alpha_2 & \beta_2 & \dots & \alpha_n & \beta_n & -1 & \beta_{n+1} \\
0 & 0 & \alpha_1 & \beta_1 & \alpha_2 & \beta_2 & \dots & \alpha_n & \beta_n \\
\vdots & & & & & & & & \vdots \\
0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots
\end{array} \right.
$$

$$
\left. \begin{array}{cccccccc}
0 & \dots & \dots & \dots & \dots & \dots & \dots & 0 \\
-1 & \beta_{n+1} & 0 & \dots & \dots & \dots & \dots & 0 \\
\vdots & & & & & & & \vdots \\
0 & \alpha_1 & \beta_1 & \dots & \alpha_n & \beta_n & -1 & \beta_{n+1}
\end{array} \right] \qquad (21)
$$

and $\hat{v}$ is the $2(N+n)$-dimensional vector

$$
\hat{v} = [\ \hat{y}(1)\ \ \hat{u}(1)\ |\ \hat{y}(2)\ \ \hat{u}(2)\ |\ldots|\ \hat{y}(N+n)\ \ \hat{u}(N+n)\ ]^T. \qquad (22)
$$

Relations (19) lead immediately to the condition

$$G\,v = G\,(\,\hat{v} + \tilde{v}\,) = G\,\tilde{v} = \Gamma, \tag{23}$$

where v and $\tilde{v}$ are vectors containing the observations and the noise samples, with the same structure as $\hat{v}$.

Problem 1 (Interpolation) – Given the model of the process and the noisy observations v, determine an unbiased and minimal variance estimate of $\hat{v}$.

Problem 2 (Filtering) – Given the model of the process and an increasing sequence of observations, estimate, as soon as a new observation $u(t)$, $y(t)$, becomes available, an unbiased and minimal variance estimate of $\hat{u}(t), \hat{y}(t)$.

When $\tilde{u}(t)$ and $\tilde{y}(t)$ are gaussian, the optimal interpolation is the ML estimation of $\hat{v}$ under the constraint $G\,\hat{v} = 0$:

$$\max_{\hat{v}^*} \, p\,(v|\hat{v}^*) = \max_{\hat{v}^*} \, \exp\left\{-\frac{1}{2}\,(v - \hat{v}^*)^T\,\tilde{\Sigma}^{-1}\,(v - \hat{v}^*)\right\} \tag{24}$$

with $G\,\hat{v}^* = 0$;

$$\tilde{\Sigma} = \mathrm{E}\,[\,\tilde{v}\,\tilde{v}^T] \tag{25}$$

$$= \begin{bmatrix} \tilde{\sigma}_y^2 & \tilde{\sigma}_{yu} & \dots & 0 & 0 \\ \tilde{\sigma}_{yu} & \tilde{\sigma}_u^2 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \tilde{\sigma}_y^2 & \tilde{\sigma}_{yu} \\ 0 & 0 & \dots & \tilde{\sigma}_{yu} & \tilde{\sigma}_u^2 \end{bmatrix} = \begin{bmatrix} \tilde{\Sigma}_{yu} & 0 & \dots & 0 \\ 0 & \tilde{\Sigma}_{yu} & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \tilde{\Sigma}_{yu} \end{bmatrix}$$

is the covariance matrix of $\tilde{v}$.

The ML estimation of $\hat{v}$ is

$$\hat{v}^*_{ML} = v - \tilde{\Sigma}\, G^T (G\, \tilde{\Sigma}\, G^T)^{-1} G\, v = \left[I - \tilde{\Sigma}\, G^T (G\, \tilde{\Sigma}\, G^T)^{-1} G \right] v.$$

The covariance matrix of the estimation error $e_{MV} = \hat{v} - \hat{v}^*_{MV}$ is

$$\begin{aligned} \operatorname{cov} e_{MV} = \\ &= \mathrm{E}\left[(\tilde{\Sigma} G^T (G \tilde{\Sigma} G^T)^{-1} G v - \tilde{v})(\tilde{\Sigma} G^T (G \tilde{\Sigma} G^T)^{-1} G v)^T - \tilde{v}^T \right] \\ &= \tilde{\Sigma}(I - G^T (G \tilde{\Sigma} G^T)^{-1} G \tilde{\Sigma})\,. \end{aligned} \tag{26}$$

EIV Filtering (Guidorzi, Diversi and Soverini, 2002, 2003, Diversi, Guidorzi and Soverini, 2003, Diversi, Guidorzi and Soverini, 2005bc, Markovsky and De Moor, 2005)

Passing from t to $t+1$, the update of $\hat{v}^*(t)$ and $G(t)$ is

$$\hat{v}^*(t+1) = [\, \hat{v}^*(t)^T \;\; \hat{y}^*(t+1) \;\; \hat{u}^*(t+1) \,]^T \tag{27}$$

$$G(t+1) = \tag{28}$$

$$\left[\begin{array}{cccccccccc} & & & & & & & & 0 & 0 \\ & & & G(t) & & & & & \vdots & \vdots \\ & & & & & & & & 0 & 0 \\ 0 & \dots & 0 & \alpha_1 & \beta_1 & \dots & \alpha_n & \beta_n & -1 & \beta_{n+1} \end{array}\right] = \left[\begin{array}{ccc} & 0 & 0 \\ G(t) & \vdots & \vdots \\ & 0 & 0 \\ m_1(t) & \multicolumn{2}{c}{m_2} \end{array}\right]$$

Denoting with $R(t)$ the $N \times N$ matrix

$$R(t) = \tilde{\Sigma}(t)\, G^T(t)(G(t)\tilde{\Sigma}(t)G^T(t))^{-1}G(t) \tag{28}$$

then

$$\hat{v}^*(t) = [I - R(t)]\, v(t) \tag{29}$$

and an iterative, finite–memory, solution of the optimal filtering problem is given by

$$\hat{v}^*(t+1)_{2n+2} = \begin{bmatrix} \hat{v}^*(t)_{2n} \\ y(t+1) \\ u(t+1) \end{bmatrix} + k(t)\, \tilde{\Sigma}(n+1)\, s^T(t)\, \varepsilon(t+1) \tag{30}$$

where $\hat{v}^*(t+1)_{2n+2}$ denotes the last $2n+2$ entries of $\hat{v}^*(t+1)$,

$$s(t) = \overline{m} \begin{bmatrix} (I - R(N))_{2n} & 0 \\ 0 & I_2 \end{bmatrix}, \tag{31}$$

$$\overline{m} = [\, \alpha_1,\, \beta_1, \ldots, -1, \beta_{n+1}], \tag{32}$$

$$k(t) = 1/(s(t)\,\tilde{\Sigma}(n+1)\,\overline{m}^T), \tag{33}$$

$$\varepsilon(t+1) = y(t+1) - \sum_{k=1}^{n} \alpha_k\,\hat{y}^*(t-n+k) - \beta_{n+1}\,u(t+1) - \sum_{k=1}^{n} \beta_k\,\hat{u}^*(t-n+k) \tag{34}$$

and the suffix $2n$ in a matrix denotes its lower right $2n \times 2n$ submatrix. $s(t)$ and $k(t)$ can, finally, be updated by means of the recursion

$$R(t+1)_{2n} = \begin{bmatrix} R(t)_{2n} & 0 \\ 0 & 0 \end{bmatrix}_{2n} + k(t)\,\tilde{\Sigma}(n)\,\left(s^T(t)\,s(t)\right)_{2n}. \tag{35}$$

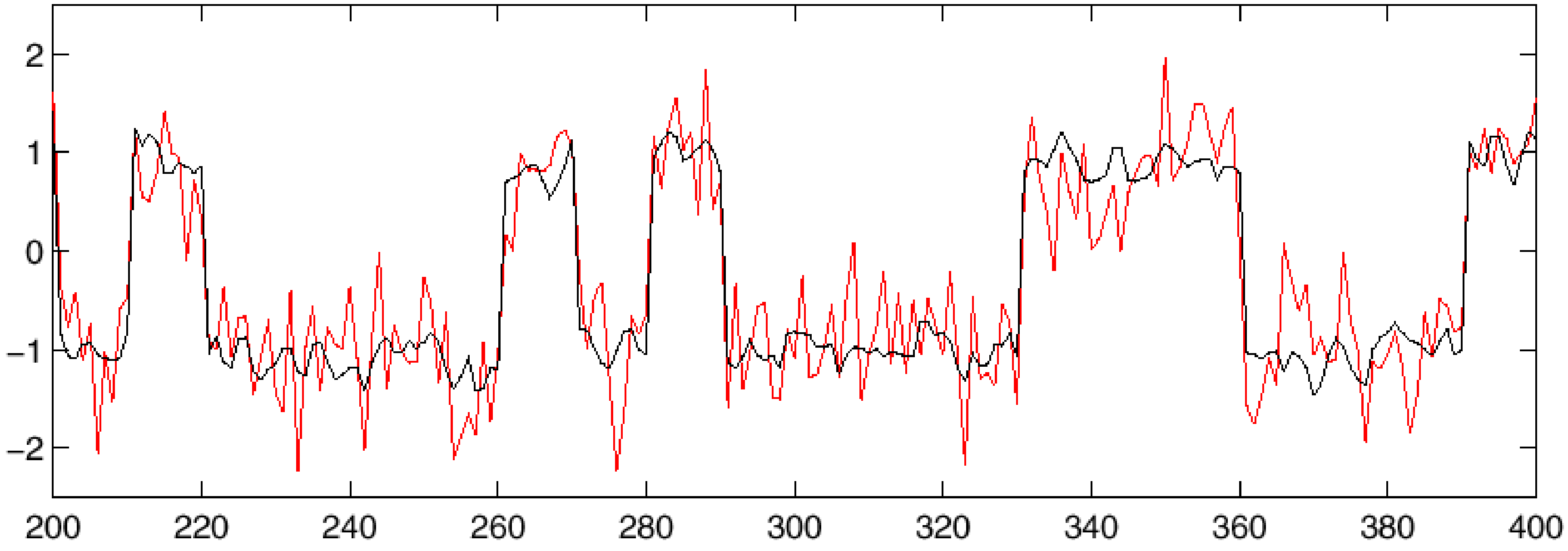
Example 1 – This example refers to a 100 runs Monte Carlo simulation concerning sequences of 400 samples generated by the model

$$y(t+3) = 1.2\,y(t+2) - 0.81\,y(t+1) + 0.27\,y(t) + 0.0885\,u(t+3)$$

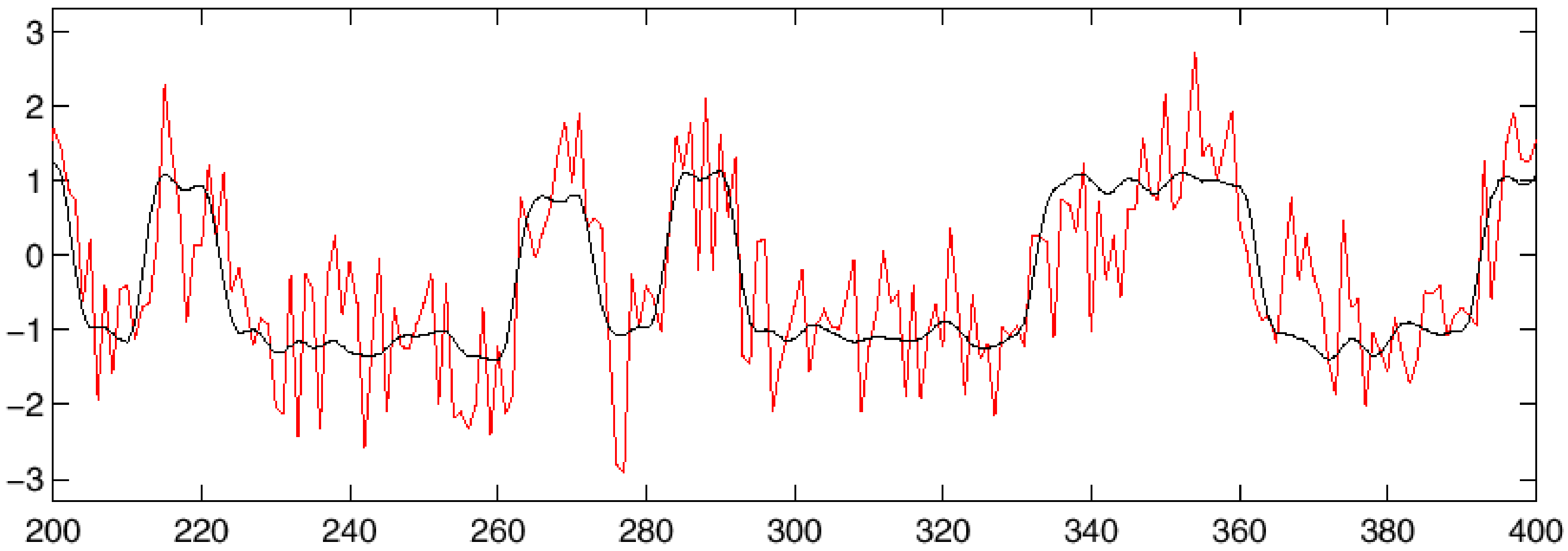
$$+ 0.168\,u(t+2) + 0.0788\,u(t+1) + 0.0385\,u(t)$$

The variances of $\hat{u}(\cdot)$ and $\hat{y}(\cdot)$ are equal to 1 while the variances of $\tilde{u}(\cdot)$ and $\tilde{y}(\cdot)$ are $\tilde{\sigma}_u^2 = 0.25$ and $\tilde{\sigma}_y^2 = 0.64$; their covariance is $\tilde{\sigma}_{yu} = 0.36$.

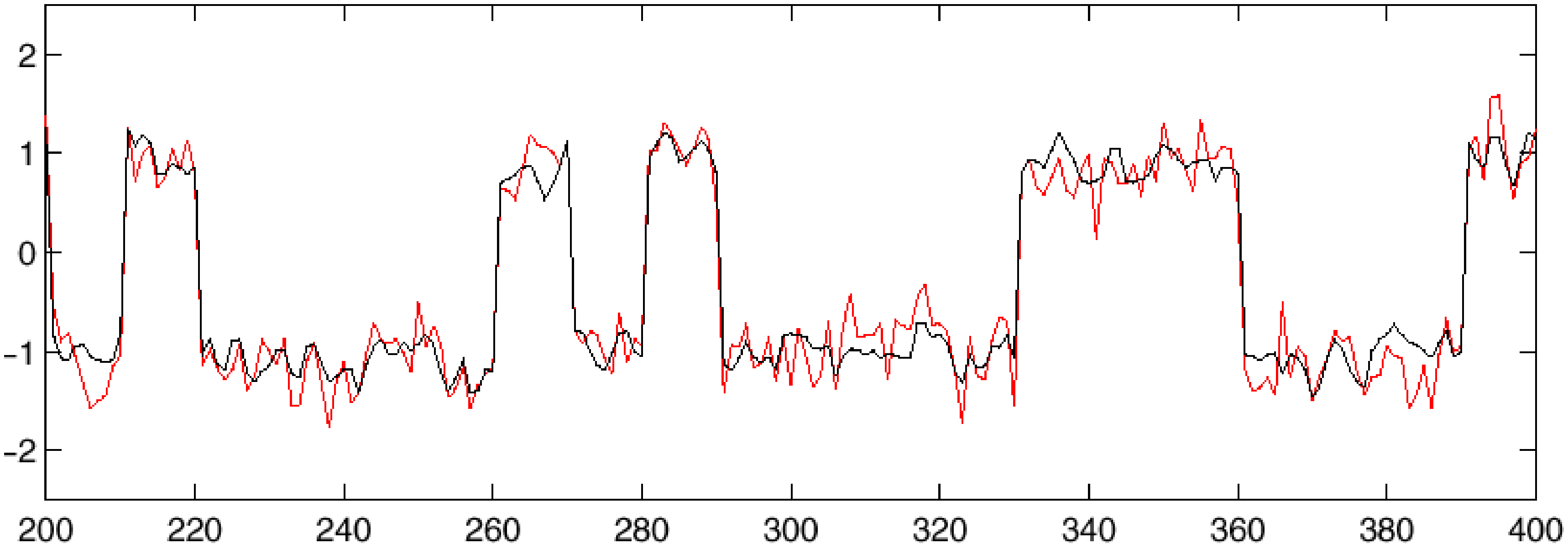
This corresponds to percent amounts of noise of 50% on the input and 80% on the output (on standard deviations).



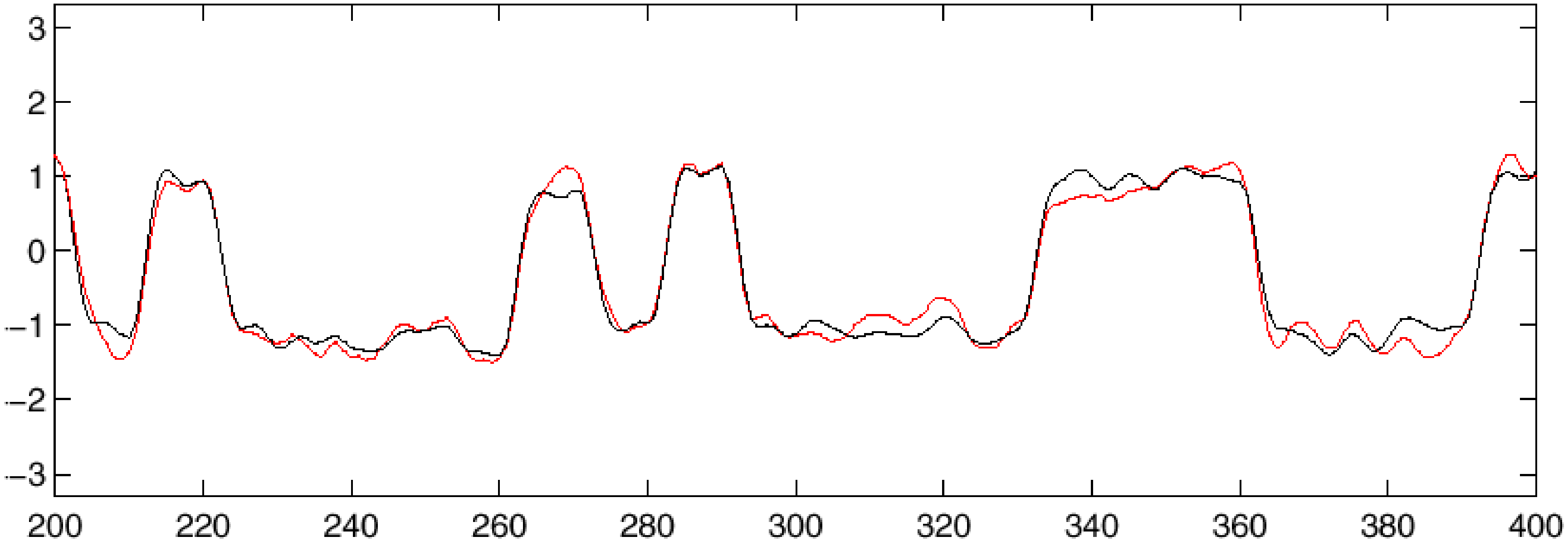
Noiseless input (black) and its observation (red)



Noiseless output (black) and its observation (red)



Noiseless (black) and filtered (red) input



Noiseless (black) and filtered (red) output

The expected performance of filtering is given by

$$\sigma_{eu}^2 = 0.2470$$

$$\sigma_{ey}^2 = 0.1667$$

The mean of the actual values obtained in 100 runs are

$$\sigma_{eu}^2 = 0.2473 \pm 0.0066$$

$$\sigma_{ey}^2 = 0.1698 \pm 0.0136$$

References

R. Diversi, R. Guidorzi and U. Soverini (2003). Algorithms for optimal errors-in-variables filtering, Systems & Control Letters, Vol. 48, pp. 1–13.

R. Diversi, R. Guidorzi and U. Soverini (2005a). Blind identification and equalization of two-channel FIR systems in unbalanced noise environments, Signal Processing, Vol. 85, pp. 215–225.

R. Diversi, R. Guidorzi and U. Soverini (2005b). Optimal errors-in-variables filtering in the MIMO case, Preprints of the 16th IFAC World Congress, Prague.

R. Diversi, R. Guidorzi and U. Soverini (2005c). Kalman filtering in extended noise environments, IEEE Transactions on Automatic Control, Vol. 50, pp. 1396–1402.

R. Diversi, R. Guidorzi and U. Soverini (2005d). A noise-compensated estimation scheme for AR processes, Proceedings of the 44th IEEE Conference on Decision and Control and 8th European Control Conference, pp. 4146–4151, Seville.

R. Diversi, U. Soverini and R. Guidorzi (2005). A new estimation approach for AR models in presence of noise, Preprints of the 16th IFAC World Congress, Prague.

R. Guidorzi, R. Diversi and U. Soverini (2002).
Errors-in-variables filtering in behavioural and state–space contexts, in Total Least Squares and Errors-in-Variables Modelling – Analysis, Algorithms and Applications, S. Van Huffel and P. Lemmerling (Eds.), pp. 281–291, Kluwer Academic Publishers, Dordrecht, The Netherlands.

R. Guidorzi, R. Diversi and U. Soverini (2003). Optimal errors-in-variables filtering, Automatica, Vol. 39, pp. 281–289.

R. Guidorzi, R. Diversi and U. Soverini (2006). Blind identification and equalization of multichannel FIR systems in unbalanced noise environments, Signal Processing (to appear).

I. Markovsky and B. De Moor (2005). Linear dynamic filtering with noisy input and output, Automatica, vol. 41, no. 1, pp. 167–171.