# Identifiability Conditions for Blind and Semi–Blind Multichannel Estimation


Elisabeth de Carvalho and Dirk T.M. Slock[†]
Institut EURECOM,
B.P. 193, 06904 Sophia Antipolis Cedex, FRANCE
{carvalho, slock}@eurecom.fr



**ABSTRACT**

We investigate the identifiability conditions for blind and semi–blind FIR multichannel estimation in terms of channel characteristics, data length and input symbol excitation modes. Parameters are identifiable if they are determined uniquely by the probability distribution of the data. Two models are presented: in the deterministic model, both channel coefficients and input symbols are considered as deterministic quantities and in the Gaussian model, the input symbols as Gaussian random variables. The Gaussian model appears more robust than the deterministic one as it requires less demanding identifiability conditions. Furthermore, semi–blind methods appear superior to blind methods as they allow the estimation of any channel with only few known symbols.


## 1 Introduction

Blind multichannel identification has received considerable interest over the last decade. In particular, second-order methods have raised a lot of attention, due to their ability to perform channel identification with relatively short data bursts. These methods suffer from several drawbacks though. They leave an ambiguity in the channel determination (in a single-user context, they can only determine the channel up to a scale or phase factor) and cannot identify certain ill-conditioned channels. This motivates the development of various other methods to alleviate this problem. Semi-blind estimation techniques exploit the knowledge of certain input symbols and appear superior to purely blind and training sequence methods as much for their performance as for their ability to perform identification for any channel for few known symbols [1].

We present here the identifiability conditions for blind and semi–blind FIR multichannel estimation for two models. For the deterministic model, in which the input symbols are considered as deterministic quantities, we give necessary and sufficient conditions in the blind case and sufficient conditions in the semi-blind case. For the Gaussian model, in which the input symbols are considered as Gaussian random variables, we give sufficient conditions. The Gaussian model proves to be more robust and the deterministic one and semi–blind methods appear superior to blind methods as they allow the estimation of any channel with only few known symbols.


[†]The work of Elisabeth de Carvalho was supported by Laboratoires d'Electronique PHILIPS under contract Cifre 297/95. The work of Dirk Slock was supported by Laboratoires d'Electronique PHILIPS under contract LEP 95FAR008 and by the EURECOM Institute


## 2 Data Model and notations

Consider a sequence of symbols $a(k)$ received through $m$ channels of length N and coefficients $\boldsymbol{h}(i)$: $\boldsymbol{y}(k) = \sum_{i=0}^{N-1}\boldsymbol{h}(i)a(k-i) + \boldsymbol{v}(k)$, $\boldsymbol{v}(k)$ is an additive independent white Gaussian noise with $r_{\boldsymbol{vv}}(k-i) = \mathrm{E}\,\boldsymbol{v}(k)\boldsymbol{v}(i)^H = \sigma_v^2 I_m\,\delta_{ki}$. Assume we receive $M$ samples, concatenated in the vector $\boldsymbol{Y}_M(k)$:

$$\boldsymbol{Y}_M(k) = \mathcal{T}_M(h)\,A_{M+N-1}(k) + \boldsymbol{V}_M(k) \tag{1}$$

$\boldsymbol{Y}_M(k) = [\boldsymbol{y}^H(k-M+1)\cdots\boldsymbol{y}^H(k)]^H$, similarly for $\boldsymbol{V}_M(k)$, and $A_M(k) = \left[a^H(k-M-N+2)\cdots a^H(k)\right]^H$, where $(.)^H$ denotes hermitian transpose. $\mathcal{T}_M(h)$ is a block Toeplitz matrix filled out with the channel coefficients. The SIMO transfer function is: $\mathbf{H}(z) = \sum_{i=0}^{N-1}\boldsymbol{h}(i)z^{-i} = [\mathrm{H}_1^H(z)\ \cdots\ \mathrm{H}_m^H(z)]^H$.

$$\boldsymbol{H} = [\boldsymbol{h}(N-1)\cdots\boldsymbol{h}(0)] \text{ and } h = [\boldsymbol{h}^H(N-1)\cdots\boldsymbol{h}^H(0)]^H. \tag{2}$$

We shall simplify the notation in (3) with $k = M-1$ to:

$$\boldsymbol{Y} = \mathcal{T}(h)A + \boldsymbol{V} = \mathcal{T}_k(h)A_k + \mathcal{T}_u(h)A_u + \boldsymbol{V} \tag{3}$$

where $A_k$ are the $M_k$ known symbols in the burst and $A_u$ the $M_u$ unknown symbols.

A channel is said irreducible if its subchannels $\mathrm{H}_i(z)$ have no zeros in common, and reducible otherwise. A reducible channel can be decomposed as: $\mathbf{H}(z) = \mathbf{H}_o(z)\mathrm{H}_c(z)$, where $\mathbf{H}_o(z)$ of length $N_o$ is irreducible and $\mathrm{H}_c(z)$ of length $N_c$ is a monochannel, $\boldsymbol{H}_c(0) = 1$. A channel is said minimum–phase if all its zeros are inside the unit circle.

We introduce the notion of effective number of subchannels denoted $m_e$: it is the rank of $\boldsymbol{H}$ in (2). Indeed, certain identifiability conditions will be based on the fact that $\mathcal{T}(h)$ is full column rank. When $\mathbf{H}(z)$ is irreducible, $\mathcal{T}(h)$ is full column rank if the rank of its lines is greater or equal to the number of columns, *i.e.* $m_e M \geq M + N - 1$, or $M \geq \left\lceil \frac{N-1}{m_e-1} \right\rceil$ (where $\lceil x \rceil$ gives the closest integer greater than $x$). Throughout the paper, we denote:

$$\underline{L} = \left\lceil \frac{N-1}{m_e-1} \right\rceil \qquad (4)$$

when $m_e = 1$, $\underline{L} = 0$.

## 3 Identifiability Definition

Let $\theta$ be the parameter to be estimated and $\boldsymbol{Y}$ the observations. $\theta$ is said identifiable if:

$$\forall\, \boldsymbol{Y}, \quad f(\boldsymbol{Y}|\theta) = f(\boldsymbol{Y}|\theta') \quad \Rightarrow \quad \theta = \theta' \qquad (5)$$

This definition has to be adapted in the blind identification case because blind techniques can at best identify the channel up to a multiplicative factor $\alpha$: $\alpha \in \mathbb{C}$ in the deterministic model and $|\alpha| = 1$ in the Gaussian model. The identifiability condition (5) will be for $\theta$ to equal $\theta'$ up to the blind indeterminacy.

For both deterministic and Gaussian models, $f(\boldsymbol{Y}|\theta)$ describes a Gaussian distribution: identifiability will be identifiability from the mean and the covariance of $\boldsymbol{Y}$.

## 4 Deterministic Model

In the deterministic model, both (unknown) input symbols and channel coefficients are assumed to be unknown deterministic quantities. Lots of blind algorithms fall into this category, among which we find: the least squares approach in [2], the (unweighed) subspace fitting approaches, the blocking equalizers method [3], the deterministic ML approaches in their blind version and in their semi–blind version [4].

In the deterministic model, $\boldsymbol{Y} \sim \mathcal{N}(\mathcal{T}(h)A, \sigma_v^2 I)$ and $\theta = [A_u^H \; h^H]^H$ (the estimation of $\theta$ is decoupled from the estimation of $\sigma_v^2$). Identifiability is based on the mean only. $A_u$ and $h$ are identifiable if:

$$\mathcal{T}(h)A = \mathcal{T}(h')A' \Rightarrow A_u = A_u' \text{ and } h = h' \qquad (6)$$

($A = \frac{1}{\alpha}A'$ and $h = \alpha h'$ in the blind case). Identifiability is then defined from the noise–free data that we will denote: $\boldsymbol{X} = \mathcal{T}(h)A$.

### 4.1 Training Sequence Based Identifiability

Training sequence based estimation is a particular case of the deterministic model for which all the input symbols are known. $\mathcal{T}(h)A = \mathcal{A}h$ (where $\mathcal{A}$ is some structured matrix containing the $a(k)$'s). $h$ is determined uniquely if and only if $\mathcal{A}$ is full column rank, which corresponds to conditions $(i - ii - iii)$ below.

**Necessary and sufficient conditions** *The $m$–channel $\boldsymbol{H}$ is identifiable by training sequence estimation if and only if*

*(i) Burst Length $\geq N$.*

*(ii) Number of known symbols $\geq 2N - 1$.*

*(iii) Number of independent input symbol modes $\geq N$.*

### 4.2 Blind Channel Identifiability

We give here necessary and sufficient conditions for deterministic blind identifiability in terms of channel characteristics, burst length and input symbol modes. Only sufficient conditions were derived in [5], necessary and sufficient conditions were given in [2], but one of their conditions is useless.

**Necessary and sufficient conditions** *In the deterministic model, the $m$–channel $\boldsymbol{H}$ and the input symbols $A$ are blindly identifiable up to a scale factor if and only if*

*(i) $\mathbf{H}(z)$ is irreducible.*

*(ii) Burst length $\geq N + 2\underline{L}$.*

*(iii) Number of independent input symbol modes$\geq N$+$\underline{L}$.*

*Proof:* **Sufficiency** It is sufficient to prove that $h$ and $A$ can be uniquely identified by a blind deterministic method. It has been shown in [6] that, under condition $(i)$, a minimum parameterization of the noise subspace of the data is given by $\overline{\boldsymbol{P}}_{\underline{L}}$ of size $(m_e - 1) \times m_e(\underline{L} + 1)$. The notation $\overline{\boldsymbol{P}}_{\underline{L}}$ indicates that it can be obtained by linear prediction. $\overline{\boldsymbol{P}}_{\underline{L}}$ verifies $\overline{\boldsymbol{P}}_{\underline{L}} \mathcal{T}_{\underline{L}+1}(h) = 0$: under condition $(i)$, this relation determines uniquely the channel from $\overline{\boldsymbol{P}}_{\underline{L}}$ up to a scale factor.

$\overline{\boldsymbol{P}}_{\underline{L}}$ can be obtained from the mean of $\boldsymbol{Y}$, *i.e.* the noise-free data $\boldsymbol{X}$, if the matrix $\mathcal{A}_M$ is full row rank. Indeed:

$$\overline{\boldsymbol{P}}_{\underline{L}} \mathcal{X}_M = \overline{\boldsymbol{P}}_{\underline{L}} \mathcal{T}_{\underline{L}+1}(h) \mathcal{A}_M = 0 \qquad (7)$$

where $\mathcal{X}_M$ is of size $m_e(\underline{L} + 1) \times M - \underline{L}$ and

$$\mathcal{X}_M = \begin{bmatrix} \boldsymbol{x}(M-1) & \cdots & \boldsymbol{x}(\underline{L}) \\ \vdots & \ddots & \vdots \\ \boldsymbol{x}(M+\underline{L}-1) & \cdots & \boldsymbol{x}(0) \end{bmatrix}, \qquad (8)$$

$$\mathcal{A}_M = \begin{bmatrix} a(M-1) & \cdots & a(\underline{L}) \\ \vdots & \ddots & \vdots \\ a(M-\underline{L}-N) & \cdots & a(-N+1) \end{bmatrix}. \qquad (9)$$

where $\boldsymbol{x}(k) = \boldsymbol{y}(k)$ in the noiseless case. If $\mathcal{A}_M$ is full row rank, (7) $\Rightarrow \overline{\boldsymbol{P}}_{\underline{L}} \mathcal{T}_{\underline{L}+1}(h) = 0$. $(ii)$ and $(iii)$ are necessary and sufficient conditions for $\mathcal{A}_M$ to be of full row rank.

Under conditions $(i)$, we can determine uniquely $h' = \alpha h$ from $\overline{\boldsymbol{P}}_{\underline{L}}$. Under conditions $(i - ii)$, $\mathcal{T}(h')$ is full column rank and $A$ can be estimated up to a scale factor: $A' = \mathcal{T}^+(h')\boldsymbol{X} = A/\alpha$, where we denote $B^+ = (B^H B)^{-1} B^H$.

**Necessity** $(i)$ If the channel is not irreducible, then $\mathcal{T}(h)$ is not full column rank. If $A$ is in the null space of $\mathcal{T}(h)$, $\boldsymbol{X} = \mathcal{T}(h)A = 0$ and identifiability is not possible: either $A = 0$ and $h$ cannot be identified, either $A \neq 0$ and $A' = 0$ verifies $\mathcal{T}(h)A' = 0$. If $A$ is not in the null space of $\mathcal{T}(h)$, we can find $A'$ verifying $\mathcal{T}(h)A' = 0$ and $A'' = A + A'$ linearly independent from $A$ verifies $\mathcal{T}(h)A'' = \boldsymbol{X}$.

$(ii-iii)$ If either $(ii)$ or $(iii)$ are not satisfied, $\mathcal{A}_M$ is not full rank: we can find $\overline{\boldsymbol{P}}' \neq \overline{\boldsymbol{P}}$ such that $\overline{\boldsymbol{P}}'\mathcal{X} = 0$ but $\overline{\boldsymbol{P}}'\mathcal{T}(h) \neq 0$ and hence another $h'$ linearly independent from $h$ such that $\overline{\boldsymbol{P}}'\mathcal{T}(h') = 0$ exists for which $\boldsymbol{X} = \mathcal{T}(h')\boldsymbol{A}$, which shows that $(ii-iii)$ are necessary conditions. □

### 4.3 Semi–Blind Channel Identifiability

Consider the general case of a reducible channel: $\mathbf{H}(z) = \mathbf{H}_o(z)H_c(z)$. Sufficient conditions for semi–blind identifiability are given in the case of grouped known symbols. We denote $\underline{L}_o = \lceil \frac{N_o-1}{m_e-1} \rceil$.

**Sufficient conditions** *In the deterministic model, the m–channel* $\mathbf{H}$ *and the unknown input symbols* $A_u$ *are semi–blindly identifiable if*

*(i) Burst length* $\geq \min(N_o + 2\underline{L}_o, N_c)$.

*(ii) Number of independent input symbol modes: at least* $\geq N_o + \underline{L}_o$ *which are not zeros of* $\mathbf{H}(z)$.

*(iii) Known symbols: number* $\geq 2N_c - 1$, *grouped, number of independent modes* $\geq N_c$.

*Proof:* The semi–blind problem is decomposed as a purely blind and a purely training sequence problem. Conditions for identifying the part of $\boldsymbol{H}$ that can be blindly identifiable, *i.e.* $\boldsymbol{H}_o(z)$ up to a scale factor, and then conditions for identifying by training sequence the rest *i.e.* the parameters in $\mathrm{H}_c(z)$ and the scale factor are derived.

$\overline{\boldsymbol{P}}_{\underline{L}}\mathcal{X}_M = \overline{\boldsymbol{P}}_{\underline{L}}\mathcal{T}_{\underline{L}+1}(h_o)\mathcal{T}_{\underline{L}+N_o-1}(h_c)\mathcal{A}_M$ implies that $\overline{\boldsymbol{P}}_{\underline{L}}\mathcal{T}_{\underline{L}+1}(h_o) = 0$ if and only $\mathcal{T}_{\underline{L}+N_o-1}(h_c)\mathcal{A}_M$ is full row rank, which conditions $(i-ii)$ guarantee.

Under conditions $(i-ii)$, we can uniquely identify $h'_o = \alpha h_o$ linear combination of the true channel: $\mathcal{T}_M^{+}(h'_o)\boldsymbol{X} = \mathcal{T}(h_c)A/\alpha$. Under conditions $(i-iii)$, $h_c$ and this scale factor is identified by training sequence estimation. □

For an irreducible channel, 1 known symbol is sufficient. For a monochannel, $2N-1$ grouped known symbol are sufficient. If you have $2N-1$ grouped known symbols containing $N$ independent modes, conditions $(i)$ and $(ii)$ are useless.

We do not prove identifiability in the case where the known symbols are not grouped. We however think that identifiability is guaranteed even in that case.

## 5 Gaussian Model

### 5.1 Gaussian Model

In the Gaussian model, the unknown input symbols are considered as i.i.d. Gaussian random variables of mean 0 and variance $\sigma_a^2$, and the known symbols as deterministic (of mean $A_k$ and variance 0) [4]. The prediction method in [7] or the covariance matching method belong to this category [8].

In the Gaussian model, the parameters to estimate are the channel coefficients and the noise covariance: $\theta = [h^H \ \sigma_v^2]^H$. Identifiability means identifiability from the mean and covariance matrix. Note that identifiability from the Gaussian model implies identifiability from any stochastic model, since such a model can be described in terms of the mean and the covariance plus higher–order moments.

### 5.2 Blind Channel Identifiability

In the blind case, $m_Y(\theta) = 0$, so identifiability is based on the covariance matrix only. In the Gaussian model, the channel and the noise variance are said identifiable if:

$$C_{YY}(h, \sigma_v^2) = C_{YY}(h', \sigma_v^{2'}) \Rightarrow h' = e^{j\varphi}h, \text{ and } \sigma_v^{2'} = \sigma_v^2 \quad (10)$$

When the input symbols are real, the phase factor is a sign, when they are complex, it is a complex unitary value.

We show here that it is possible to identify blindly the channel based on the second-order moments even for a reducible channel provided that its zeros are minimum-phase. We give conditions on the channel and the correlation sequence length.

#### 5.2.1 *Irreducible Channel*

We give here sufficient conditions in the case of an irreducible channel.

**Sufficient conditions** *In the Gaussian model, the m–channel* $\boldsymbol{H}$ *is identifiable blindly up to a phase factor if*

*(i) The channel is irreducible.*

*(ii)* $M \geq \underline{L} + 1$.

*Proof:* When condition $(ii)$ is verified, $\mathcal{T}_u(h)$ is (strictly) tall and $\sigma_v^2$ can then be uniquely identified as the minimal eigenvalue of $C_{YY}(\theta)$. $\boldsymbol{H}(z)$ can then be identified up to a phase factor from the denoised covariance matrix $C_{YY}(\theta) - \sigma_v^2 I$ by linear prediction [6]: under conditions $(i-ii)$, you can find $P(z)$, the multivariate prediction filter of length $\underline{L}+1$ obtained from the denoised covariance matrix, which verifies P(z) $\mathbf{H}$(z) = $\boldsymbol{h}(0)$. This relationship allows to recover uniquely $\mathbf{H}(z)$ from $P(z)$ up to a phase factor.□

Note that you do not need all the non zero correlations (time 0 to $N$) for identification but only the $\underline{L}+1$ first.

#### 5.2.2 Reducible channel

Let $\mathbf{H}(z)$ be a reducible channel: $\mathbf{H}(z) = \mathbf{H}_o(z)H_c(z)$. We prove that a reducible channel is identifiable in the Gaussian model if its zeros are minimum-phase.

**Sufficient conditions** *In the Gaussian model, the m–channel* $\boldsymbol{H}$ *is identifiable blindly up to a phase factor if*

*(i)* $H_c(z)$ *is minimum phase.*

*(ii)* $M \geq \max(L_o{+}1, N_c{-}N_o{+}1)$.

*Proof:* Under condition *(ii)*, $\mathcal{T}(h)$ is strictly tall and $\sigma_v^2$ can be identified as the minimal eigenvalue of $C_{YY}(\theta)$. The irreducible part $\boldsymbol{H}_o$ can be identified up to a scale factor thanks to the deterministic method described in section 4.2 [6] provided that $M \geq \underline{L}_o$: let $h_o' = \alpha h_o$ be this estimate of $h_o$. $\mathcal{T}(h_o')^+ \left(C_{YY}(\theta) - \sigma_v^2 I\right) \mathcal{T}(h_o')^{H+} = \sigma_a^2 \mathcal{T}(\alpha h_c)\mathcal{T}^H(\alpha^* h_c)$. $\alpha H_c(z)$ can be now identified by spectral factorization provided that $\alpha H_c(z)$ or $H_c(z)$ is minimum phase and $\mathcal{T}(h_c)\mathcal{T}^H(h_c)$ contains the N non zero correlations, *i.e.* $M + N_o - 1 \geq N_c$ or $M \geq N_c{-}N_o{+}1$. □

### 5.3 Semi–Blind Channel Identifiability

In the semi–blind case, the channel is identifiable from the mean and the covariance matrix.

#### 5.3.1 Identifiability for any channel

In the semi–blind case, the Gaussian model presents the advantage to allow identification from the mean only. $m_Y(\theta) = \mathcal{T}_k(h)A_k = \mathcal{A}_k h$: if $\mathcal{A}_k$ is full column rank, $h$ can be identified. The difference with the training sequence case is that in the identification of $\boldsymbol{H}$ from $m_Y(\theta) = \mathcal{T}_k(h)A_k$, the zeros before and after the different blocks of known symbols also serve as training sequence symbols, which lowers the requirements of the classical training sequence case. For one non-zero known symbol $a(k)$, with $0 \leq k \leq M{-}N$, $\mathcal{A}_k$ contains only a non-zero submatrix of dimension $Nm \times Nm$: $a(k)I_{Nm}$. The Gaussian model presents the great advantage to allow identification of any channel, reducible or not, multi or monochannel, for only one non-zero known symbol not located at the edges of the input burst.

**Sufficient conditions** *In the Gaussian model, the m–channel* $\boldsymbol{H}$ *is identifiable blindly up to a phase factor if*

*(i) Burst length* $\geq N$.

*(ii) At least one non-zero known symbol* $a(k)$ *not located at the edges* $(0 \leq k \leq M{-}N)$.

**Sufficient conditions** *In the Gaussian model, the m-channel* $\boldsymbol{H}$ *is identifiable blindly up to a phase factor if*

*(i) Channel irreducible*

*(ii) At least 1 non-zero known symbol (located anywhere)*

*Proof:* Let's assume that $\boldsymbol{Y}$ contains a block of at least $\underline{L}+1$ samples $\boldsymbol{y}(k)$ that contain only unknown symbols (this gives a condition on the burst length which we do not give above because it depends on the number of known symbols and their position). Then $h$ can be identified blindly up to a unitary constant from the corresponding covariance matrix as indicated in section 4.2: $h' = e^{j\varphi}h$. This unitary scale can then be identified thanks to the mean $\mathcal{T}_k^+(h)m_Y = e^{-j\varphi}A_k$: one non-zero element of this quantity sufficies to identify $\varphi$.

### References


[1] E. de Carvalho and D. T. M. Slock. "Asymptotic Performance of ML Methods for Semi-Blind Channel Estimation". Pacific Grove, CA, Nov. 1997. Proc. Asilomar Conference on Signals, Systems & Computers.

[2] G. Xu, H. Liu, L. Tong, and T. Kailath. "A Least Squares Approach to Blind Channel Identification". *IEEE Transactions on Signal Processing*, 43(12):2982–2993, Dec. 1995.

[3] D.T.M. Slock. "Blind Fractionally-Spaced Equalization, Perfect-Reconstruction Filter Banks and Multichannel Linear Prediction". In *Proc. ICASSP 94 Conf.*, Adelaide, Australia, April 1994.

[4] E. de Carvalho and D. T. M. Slock. "Maximum-Likelihood FIR Multi-Channel Estimation with Gaussian Prior for the Symbols". In *Proc. ICASSP 97 Conf.*, Munich, Germany, April 1997.

[5] Y. Hua. "Fast Maximum Likelihood for Blind Identification of Multiple FIR Channels". *IEEE Transactions on Signal Processing*, 44(3):661–672, March 1996.

[6] D.T.M. Slock. "Blind Joint Equalization of Multiple Synchronous Mobile Users Using Oversampling and/or Multiple Antennas". In *Proc. 28th Asilomar Conference on Signal, Systems & Computers*, Pacific Grove, CA, Nov. 1994.

[7] K. Abed Meraim, E. Moulines, and P. Loubaton. "Prediction Error Method for Second-Order Blind Identification". *IEEE Transactions on Signal Processing*, 45(3):694–705, March 1997.

[8] G.B. Giannakis and S.D. Halford. "Asymptotically Optimal Blind Fractionally-Spaced Channel Estimation and Performance Analysis". *IEEE Transactions on Signal Processing*, 45(3):694–705, March 1997.