



## UNIVERSITÀ DEGLI STUDI DI PADOVA

---

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE, PADOVA.  
DEPARTMENT OF INFORMATION TECHNOLOGY AND ELECTRICAL  
ENGINEERING, ETH ZÜRICH.

*Corso di Laurea in Ingegneria delle Telecomunicazioni*

TESI DI LAUREA

### IDENTIFICATION OF WIRELESS SYSTEMS

*Laureando*

**IRENE PAPPALARDO**

*Relatore*

**PROF. MICHELE ZORZI**

*Correlatori*

**PROF. HELMUT BÖLCSKEI, ING. REINHARD HECKEL**

---

ANNO ACCADEMICO 2011/2012

---

*Ai miei fantastici genitori*

## Sommario

La presente tesi si propone di studiare il problema dell'identificazione di tre modelli di sistemi tempo varianti. Tale problema consiste nel ricostruire i parametri di sistema a partire da misurazioni ingresso-uscita. Inoltre sono oggetto di studio le condizioni sufficienti che permettono l'identificazione del sistema. Nella prima parte della tesi si studia l'identificazione di sistemi a banda larga (*wideband systems*). Viene presentata una caratterizzazione discreta del modello continuo del sistema. A questo scopo viene utilizzata la trasformata di Mellin che fornisce una rappresentazione del segnale in termini delle sue componenti di *scaling*. Nella seconda parte dell'elaborato, si considerano i sistemi la cui risposta consiste in una sovrapposizione pesata di versioni del segnale di ingresso, traslate nel tempo e in frequenza. Si assume inoltre che tali sistemi siano *sparsi*, ovvero che introducano solamente alcune traslazioni in tempo e frequenza, ignote. Si sono confrontati tramite simulazioni due diversi approcci di ricostruzione. Tali approcci differiscono sia in termini di segnale di ingresso che per l'algoritmo usato per la ricostruzione. Secondo il primo approccio, la formulazione dell'identificazione viene ricondotta a un problema di tipo *multiple measurement vector*, mentre nel secondo approccio rispecchia un problema di tipo *block-sparse*. Per ciascun approccio viene usata una particolare versione di un algoritmo greedy di ricostruzione, più precisamente l'orthogonal matching pursuit. Le simulazioni non solo dimostrano che i due approcci hanno prestazioni diverse, ma evidenziano anche, dal punto di vista quantitativo, come il secondo approccio abbia più successo rispetto al primo, in termini di probabilità di ricostruzione (*recovery probability*). Gli stessi risultati si ottengono nella condizione in cui il segnale ricevuto è corrotto dal rumore. Per ultimo, si sono analizzati dei particolari sistemi a banda stretta (*narrowband systems*), detti *parametric underspread linear systems*, che introducono sul segnale in ingresso un numero limitato di traslazioni in tempo e frequenza. Si sono estesi i risultati trovati in (Bajwa, 2011) al caso di sistemi a più ingressi e singola uscita. Più precisamente si sono trovate le condizioni sufficienti per l'identificazione e l'implementazione della ricostruzione di sistemi a ingressi multipli.



---

## Abstract

We consider the problem of identifying three different models of linear time-varying systems. The identification problem is to reconstruct the system parameters from an input-output measurement. Furthermore, we are interested in finding sufficient conditions that allow identification. Firstly, we consider identifiability of wideband systems. To this end, we present a discrete characterization of the continuous model. In order to derive that characterization, we use the Mellin transform that provides a representation of the signal in terms of its scale components. Secondly, we consider a system, where the output consists of a weighed superposition of time and frequency shifted versions of the input signal. We assume the system is sparse, in the sense that it induces only few, however unknown, time and frequency shifts. Two different recovery approaches are compared using simulations. The approaches differ in the probing signal and in the algorithm used for the recovery. For the first approach, the recovery problem is reduced to a multiple measurement vector problem, for the second to a block-sparse problem. For both problems, adaptations of a greedy algorithm, namely orthogonal matching pursuit, are used. The numerical experiments not only demonstrate that the approaches perform differently, but also provide the quantitative improvement of the second approach to the first one in terms of recovery probability. The same results are achieved for the noisy case when the received signal is corrupted by noise. Finally, we investigate on particular narrowband systems, called parametric underspread linear systems, that induce only finitely many time and frequency shifts on the input signal. We extend the results found in (Bajwa, 2011), i.e., sufficient conditions for identifiability and the implementation, to the multiple input single output case.



# Contents

<b>1. Introduction</b>	<b>1</b>
<b>2. Identification of Wideband Time-Varying Systems</b>	<b>7</b>
2.1. Mellin Transform and its Sampling Theorem . . . . .	7
2.2. Input-Output Relation of a Wideband Time-Varying System . . . . .	11
2.3. Discrete Time-Scale Characterization . . . . .	12
2.4. Finite Approximation . . . . .	18
2.5. Uncertainty Principle for Frequency and Scale Operators . . . . .	20
<b>3. Identification of Narrowband Time-Varying Sparse Systems</b>	<b>29</b>
3.1. Discrete Time-Frequency Characterization . . . . .	29
3.2. Matrix Representations . . . . .	32
3.3. Recovery Algorithms . . . . .	37
3.4. Simulation Results . . . . .	39
<b>4. Identification of Parametric Underspread Linear Systems</b>	<b>43</b>
4.1. System Characterization . . . . .	43
4.2. Assumptions for Identification . . . . .	44
4.3. Recovery Procedure: The Sampling Stage . . . . .	47
4.4. Matrix Formulation . . . . .	48
4.5. Recovery Procedure: The Recovery Stage . . . . .	51
4.5.1. First step: recovery of the measurement vectors. . . . .	51
4.5.2. Second step: ESPRIT algorithm and recovery of the delays . . . . .	52
4.5.3. Third step: recovery of the Doppler-shifts and attenuation factors . . . . .	56
4.6. Implementation of the Recovery Stage . . . . .	57
4.7. Sufficient Conditions for Identifiability . . . . .	59
4.8. Identification for Multiple Inputs . . . . .	62
<b>A. Proof of Theorem 5</b>	<b>67</b>
<b>B. Proof of Theorem 6</b>	<b>69</b>





# Introduction

*System identification* is an important problem in engineering and it has many practical applications. These include channel identification in wireless communication or underwater communication, control engineering and radar imaging. In this thesis we consider the identification of wireless systems.

A wireless channel is usually modeled as a linear time-varying system [1]. The simplified characterization of a wireless channel is given by the *multipath approximation*: the electromagnetic field from the transmitter to the receiver is described by individual waves that travel along specific propagation paths and undergo different effects, like diffraction, reflection and absorption. At every point in space and time, the individual waves interfere with each other and form the overall electromagnetic field. Furthermore, since the terminals of a wireless system are often mobile and the objects in the environment move as well, the resulting electromagnetic field is also time-varying [2], i.e., the received signal consists of the superposition of time- and frequency- shifted versions of the transmitted signal, weighted by a *spreading function*  $s_H$  that characterizes the environment in which the communication takes place. Under the approximation of infinite multipath components, the received signal can be represented as [3]

$$y(t) = \int_{\tau} \int_{\nu} s_H(\tau, \nu) x(t - \tau) e^{j2\pi\nu t} d\nu d\tau. \quad (1.1)$$

This characterization implicitly assumes that signals involved are narrowband. That is, during the transmission of the signals, the motions in the system are slow compared to the signal propagation speed and the objects in the environment do not change position relative to the positional resolution of the signal (slowly fluctuating objects) [4]. Suppose the transmitted signal has duration  $T$  and that the object velocity is  $v$ , then the *narrowband condition* on the signal bandwidth  $B$  is given by

$$\frac{2v}{c} \ll \frac{1}{TB} \quad (1.2)$$

where  $c$  is the medium propagation speed. In an environment with fast moving objects or if the signals have large time-bandwidth product, the requirement (1.2) is violated. In these systems, the delay-Doppler shift approximation (1.1) is not

precise and a wideband model is needed to account for the effects caused by high speed motion [6]. In the wideband signal model, the received signal is given by a superposition of the time shifted and time scaled versions of the input signal, weighted by a *wideband spreading function*  $\chi_H$ . Letting  $a$  the scale parameter, the input-output relation of a wideband system can be modeled as [6]

$$y(t) = \int_a \int_{\tau} \chi_H(a, \tau) \sqrt{a} x(a(t - \tau)) d\tau da. \quad (1.3)$$

The *identification problem*, which is depicted in Fig. 1.1, can be stated as determining the parameters of the operator  $\mathbb{H}$ , which models the channel/system, by input-output operations on the system response  $y(t)$  to a known probing signal  $x(t)$ . The crucial point is to select a suitable probing signal and make appropriate computations on the output in order to recover the parameters that completely characterize the system  $\mathbb{H}$ . In particular, in order to recover an operator of the form (1.1), we need to recover  $s_H$ ; and to recover an operator of the form (1.3), we need to recover  $\chi_H$ .

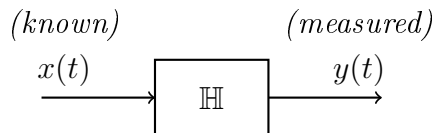


Figure 1.1.: Identification problem.

If we consider the particular case of a linear time-invariant system, we have  $s_H(\tau, \nu) = h(\tau)\delta(\nu)$  for narrowband systems and  $\chi_H(a, \tau) = g(\tau)\delta(a)$  for wideband systems, where  $h(t)$  and  $g(t)$  are the impulse responses of the respective systems. Since the output of the system to the input probing signal  $x(t) = \delta(t)$  is exactly its impulse response, which fully characterizes  $\mathbb{H}$ , linear time-invariant systems are always identifiable. The situation is fundamentally different for linear time-variant systems, where the response to the input  $x(t) = \delta(t)$  is not equal to the impulse response and does not allow a complete characterization of the system. We require then to investigate on some conditions of the operator  $\mathbb{H}$ , for which identification is possible. These requirements are often expressed in terms of the spreading functions support, which in both narrowband and wideband representations is supposed to be finite according to some physical limitations. In fact, due to physical restrictions on the system, time delays, Doppler shifts and scale shifts cannot take any real value but they are bounded within some ranges. For narrowband systems (1.1) Kailath [7] found that linear time-varying systems with spreading function  $s_H$  compactly supported on a rectangle are identifiable if and only if the spreading function's support area is not larger than 1. This result is generalized by allowing any shape of the support region, provided that the total area is not larger than 1 [8].

Analogous results are not yet provided in the case of wideband systems (1.3).

In the first part of the thesis, we consider the problem of identifying a wideband system. Our initial idea was to gain intuition by first considering the discrete case of (1.3). To this end, we try to adopt the proof Kailath [7] did for systems of the form

---

(1.1) to wideband systems. In order to discretize (1.3), we assume the input signal is band- and scale-limited and apply the sampling theorem both in the Fourier and in the scale domain. In this context the Mellin transform becomes relevant for dealing with the scale parameter.

Difficulties arose because to apply a similar argument as Kailath did, we would need a result that specifies the approximate number of degrees of freedom of a scale and band limited signal, paralleling the 2WT theorem [9]. Therefore, the problem remains open.

In the second part of this thesis, we analyze the identification of *sparse* systems. Sparsity is an important property which has wide applications in many communication channels. Practical examples are given by underwater environment, radar imaging and mobile communication. We consider a discretized version of (1.1) and assume the spreading function is sparse in a sense made precisely later. As for the case of wideband systems, some constraints on the input and output signals are needed in order to use the sampling theorem to derive a discrete representation of the system. In particular, we consider band limited input signals and assume the output is observed for a finite time duration. We focus on the possible procedures for recovery the samples of the spreading function, starting from the discrete representation of (1.1) developed in [15]. This allows to state the identification problem as solving a linear system of equations, i.e.,

$$\mathbf{Y} = \mathbf{X}\mathbf{S} \tag{1.4}$$

where  $\mathbf{X}$  is the matrix of the known input samples, while  $\mathbf{Y}$  and  $\mathbf{S}$  contain, respectively, the samples of the output, measured from the system response, and the samples of the spreading function to be recovered.

The necessary conditions for identifiability are connected with the notion of *compressive sensing*. The central idea is to identify the system, i.e., reconstruct the spreading function samples in  $\mathbf{S}$ , assuming  $\mathbf{S}$  has only a few nonzero entries. In this case the spreading function is referred to *sparse* and, as long as the system of equations (1.4) is underdetermined, it can admit a unique solution.

We analyze and compare two approaches that differ in the choices of the input signal and the algorithms used for recovery. The first approach allows to formulate the problem as a multiple measured vector (MMV) problem [16] where both  $\mathbf{Y}$  and  $\mathbf{S}$  in (1.4) are matrices whose columns constitute single measurements. The main property of the model is that the collection of sparse vectors share a common sparse support set, i.e., the positions of the nonzero samples in the columns of  $\mathbf{S}$  are unchanged. The second approach is to formulate the recovery as a block-sparse problem. In this case, the sparse vector of the spreading function has nonzero samples occurring in clusters, i.e., blocks. The algorithms we use for recovery are adaptations of the orthogonal matching pursuit (OMP) algorithm which is the canonical greedy algorithm for sparse approximation [18]. The simultaneous OMP (S-OMP) algorithm [19] is used for the first approach, while the block-sparse OMP (B-OMP) algorithm [20] is used for the second approach.

We compare the performances of the approaches in terms of the empirical recovery probability, for the noiseless case, and of the root mean square error, for the noisy

case. The main result is that the first approach is superior to the second one, especially when the number of the samples of the discrete spreading function is large, i.e., the approximation of the continuous spreading function to be recovered is more precise.

As the third part of this thesis, we study the identification of a particular multiple input narrowband system, supposing always the communication channel is sparse. An interesting question on system identification is whether the spreading function can be identified when dealing with multiple input multiple output (MIMO) systems. In this situation, each output can be represented as a superposition of the responses of different subsystems applied to each input. Given that the spreading functions of the subchannels are independent, to characterize the identifiability of MIMO systems we can without loss of generality consider multiple input single output (MISO) systems. We consider the extension to multiple inputs of a particular narrowband system, called *parametric underspread linear system (ULS)* [23], whose spreading function is described by a finite set of delays and Doppler-shifts, as expressed in the following. The input-output relation is given by

$$y(t) = \int_{\tau} \int_{\nu} \sum_i \tilde{s}_{H_i}(\tau, \nu) x_i(t - \tau) e^{j2\pi\nu t} d\nu d\tau \quad (1.5)$$

where the spreading function  $\tilde{s}_{H_i}$  of the  $i$ -th subchannel is given by

$$\tilde{s}_{H_i}(\tau, \nu) = \sum_{k=1}^{K_i} \alpha_{i,k} \delta(\tau - \tau_{i,k}) \delta(\nu - \nu_{i,k}). \quad (1.6)$$

In (1.6) the parameters  $\alpha_{i,k}$ ,  $\tau_{i,k}$  and  $\nu_{i,k}$  represent the attenuation factor, the time delay and the Doppler-shift, respectively, associated with the  $i$ -th subchannel, while  $K_i$  is the number of delay-Doppler pairs involved in  $\tilde{s}_{H_i}$ . The crucial difference to the model considered in the second part of this thesis is that, in order to identify the system, we need to specify all the parameters above, for every subsystem  $\tilde{s}_{H_i}$ . The condition for identifiability of a MIMO channel is provided in [21], given that the supports of the subsystems are known and in [15] given the support is unknown. The spreading functions  $\tilde{s}_{H_i}$  in (1.6) are identifiable if and only if the sum of the areas of the support regions of  $\tilde{s}_{H_i}$  is not larger than 1. We generalize the results of [22] and [23] for SISO systems to MISO systems. In particular, with some adjustments we adopt the multiple inputs problem to a single input problem, applying then the same recovery procedure.

We derive sufficient conditions on the bandwidth and temporal support of the equivalent input signal that ensure identification of the parametric ULS in (1.5) and (1.6). It can be shown that the system is identifiable as long as the time-bandwidth product of the input signal is proportional to the total number of the delay-Doppler pairs in the entire system. The recovery procedure is based on the ESPRIT algorithm ([22] and [24]) for the time delay estimation. The recovery of the frequencies and of the attenuation factors follows directly from the delay recovery and concludes the identification problem.

---

The organization of the thesis is as follows. In Chapter 1 we analyze the identification of the wideband system, using the Mellin transform and the corresponding sampling theorem, and we derive the equivalent discrete characterization of the continuous channel. In Chapter 2 we study the identification of sparse systems and compare empirically the performances of two different recovery approaches. In Chapter 3 we extend the identification of the parametric underspread linear system to a multiple input single output system and derive sufficient conditions for identification.

*Notations.*

Lowercase boldface letters stand for column vectors and uppercase letters designate matrices. For the matrix  $\mathbf{A}$  we write its transpose, complex conjugate, Hermitian (complex conjugate of the transpose) and Penrose pseudo-inverse by  $\mathbf{A}^T$ ,  $\mathbf{A}^*$ ,  $\mathbf{A}^H$  and  $\mathbf{A}^\dagger$ , respectively. The entry in the  $k$ th row and  $l$ th column of  $\mathbf{A}$  is denoted by  $[\mathbf{A}]_{k,l}$ . For the vector  $\mathbf{a}$ , the  $k$ th element is written  $[\mathbf{a}]_k$  and its Euclidean norm is denoted by  $\|\mathbf{a}\|_2$ . Finally, for two functions  $x(t)$  and  $y(t)$  defined for  $t \in \mathbb{R}$ , we write  $\langle x(t), y(t) \rangle \triangleq \int_{-\infty}^{+\infty} x(t)y^*(t)dt$  for the inner product between  $x(t)$  and  $y(t)$ .



# Identification of Wideband Time-Varying Systems

In the first part of this chapter (Section 2.1) we introduce the Mellin transform, which will be useful to obtain a discrete characterization of wideband time-varying systems. The Mellin transform can be seen as an equivalent representation of the signal in the scale domain as well as the Fourier transform provides the spectral representation in the frequency domain. We also discuss a sampling theorem which enables the exact reconstruction of a scale limited signal from its samples in the time domain.

In the second part of the chapter (Sections 2.2 - 2.4) we derive the discrete characterization of a wideband channel, using the sampling theorems both in frequency and scale domains.

We conclude the chapter (Section 2.5) making some remarks about the scale operator and deriving the uncertainty principle for frequency and scale domains.

## 2.1. Mellin Transform and its Sampling Theorem

All the material in this section can be found in [12].

The Mellin transform  $\mathcal{M}_f(\beta)$  of a function  $f(t)$ , whose support in time domain is  $(0, +\infty)$ , is defined [12] as

$$\mathcal{M}_f(\beta) = \int_0^{+\infty} \frac{1}{\sqrt{t}} f(t) e^{+j2\pi\beta \ln t} dt \quad (2.1)$$

If  $f(t)$  is defined also on  $(-\infty, 0)$ , the positive and negative parts have to be treated separately.

The inverse Mellin transform is defined as

$$f(t) = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{t}} \mathcal{M}_f(\beta) e^{-j2\pi\beta \ln t} d\beta \quad (2.2)$$

The properties of the Mellin transform, together with its derivation, are extensively studied by Cohen in [11]. We report here two important properties that are useful for the future analysis.

1. *Scale invariance.* Indicating with  $\mathcal{M}_f(\beta)$  the Mellin transform of  $f(t)$ , then the Mellin transform of  $f_0(t) = \sqrt{c_0}f(c_0t)$  is given by

$$\mathcal{M}_{f_0}(\beta) = e^{-j2\pi\beta \ln c_0} \mathcal{M}_f(\beta) \quad (2.3)$$

*Proof.* Applying the definition in (2.1) on  $f_0(t)$ ,

$$\begin{aligned} \mathcal{M}_{f_0}(\beta) &= \int_0^{+\infty} \frac{1}{\sqrt{t}} \sqrt{c_0} f(c_0t) e^{+j2\pi\beta \ln t} dt \\ \left( u = c_0t \Rightarrow dt = \frac{du}{c_0} \right) &= \int_0^{+\infty} \frac{c_0}{\sqrt{u}} f(u) e^{+j2\pi\beta \ln \frac{u}{c_0}} \frac{du}{c_0} \\ &= e^{-j2\pi\beta \ln c_0} \int_0^{+\infty} \frac{1}{\sqrt{u}} f(u) e^{+j2\pi\beta \ln u} du \\ &= e^{-j2\pi\beta \ln c_0} \mathcal{M}_f(\beta) \end{aligned}$$

□

Note that  $\mathcal{M}_f(\beta)$  and  $\mathcal{M}_{f_0}(\beta)$  have same supports.

2. *Multiplicative convolution.* Indicating with  $\mathcal{M}_{f_1}(\beta)$  and  $\mathcal{M}_{f_2}(\beta)$  the Mellin transforms of  $f_1(t)$  and  $f_2(t)$  respectively, then the Mellin transform of the multiplicative convolution of  $f_1(t)$  and  $f_2(t)$ , given by

$$f_1 \otimes f_2(t) = \int_0^{+\infty} f_1(tt') f_2^*(t') dt' \quad (2.4)$$

is simply  $\mathcal{M}_{f_1}(\beta) \mathcal{M}_{f_2}^*(\beta)$ .

*Proof.* Applying the definition in (2.1) on  $f(t) = f_1 \otimes f_2(t)$ ,

$$\begin{aligned} \mathcal{M}_f(\beta) &= \int_0^{+\infty} \frac{1}{\sqrt{t}} \int_0^{+\infty} f_1(tt') f_2^*(t') dt' e^{+j2\pi\beta \ln t} dt \\ \left( u = tt' \Rightarrow dt = \frac{du}{t'} \right) &= \int_0^{+\infty} \int_0^{+\infty} \sqrt{\frac{t'}{u}} f_1(u) f_2^*(t') e^{+j2\pi\beta \ln \frac{u}{t'}} dt' \frac{du}{t'} \\ &= \int_0^{+\infty} \frac{1}{\sqrt{u}} f_1(u) e^{+j2\pi\beta \ln u} du \int_0^{+\infty} \frac{1}{\sqrt{t'}} f_2^*(t') e^{-j2\pi\beta \ln t'} dt' \\ &= \mathcal{M}_{f_1}(\beta) \mathcal{M}_{f_2}^*(\beta) \end{aligned}$$

□

In the same way as the Fourier transformation is used to reconstruct a band limited function from its uniformly spaced samples, the Mellin transform can be used to recover a scale limited function from its samples, exponentially spaced in the time domain. A function  $f(t)$  is said to be *scale limited* to scale  $\beta_0$  if the support of its Mellin transform  $\mathcal{M}_f(\beta)$  is  $[-\beta_0, \beta_0]$ . The scale support width is indicated with  $B$ ,  $B = 2\beta_0$ .

We follow the same procedure reported in [12] to state and prove the following theorem. See [13] for a more exhaustive analysis.



**Theorem 1. Sampling Theorem.** A function  $f(t) \in L^2(\mathcal{R})$ , scale limited to  $\beta_0$ , can be exactly reconstructed from its samples in time domain if the samples are exponentially spaced along the time axis, i.e. from  $\{f(\tau^n)\}_{n=-\infty}^{+\infty}$ , where  $\tau = e^{1/B}$  and  $B = 2\beta_0$ .

*Proof.* Consider a signal  $f(t)$ , whose Mellin transform  $\mathcal{M}_f(\beta)$  is scale limited to  $\beta_0$ , and indicate with  $B = 2\beta_0$  the scale support width of  $\mathcal{M}_f(\beta)$ . Due to the Fourier series representation, for  $\beta \in [-\beta_0, \beta_0]$   $\mathcal{M}_f(\beta)$  can be expressed as

$$\mathcal{M}_f(\beta) = \sum_{m=-\infty}^{+\infty} a_m e^{j2\pi \frac{m}{B} \beta} \quad (2.5)$$

where  $a_m$  are the coefficients of the series expansion

$$a_m = \frac{1}{B} \int_{-\beta_0}^{\beta_0} \mathcal{M}_f(\beta) e^{-j2\pi \frac{m}{B} \beta} d\beta, \quad m \in \mathbb{Z} \quad (2.6)$$

Indicating with  $\tau = e^{1/B} = e^{1/(2\beta_0)}$ , the coefficients  $a_m$  in (2.6) are given by

$$a_m = \ln \tau \int_{-\beta_0}^{\beta_0} \mathcal{M}_f(\beta) e^{-j2\pi \beta m \ln \tau} d\beta = \ln \tau \int_{-\beta_0}^{\beta_0} \mathcal{M}_f(\beta) e^{-j2\pi \beta \ln \tau^m} d\beta, \quad m \in \mathbb{Z} \quad (2.7)$$

The function  $f(t)$  can be expressed as the inverse Mellin transform of  $\mathcal{M}_f(\beta)$  and, using that  $f(t)$  is scale limited, (2.2) becomes

$$f(t) = \int_{-\beta_0}^{\beta_0} \frac{1}{\sqrt{t}} \mathcal{M}_f(\beta) e^{-j2\pi \beta \ln t} d\beta \quad (2.8)$$

For  $t = \tau^m$ , (2.8) becomes

$$f(\tau^m) = \int_{-\beta_0}^{\beta_0} \frac{1}{\sqrt{\tau^m}} \mathcal{M}_f(\beta) e^{-j2\pi \beta \ln \tau^m} d\beta \quad (2.9)$$

Comparing (2.7) and (2.9), the relation between the coefficients  $a_m$  and the function  $f(t)$  is

$$a_m = \ln \tau \sqrt{\tau^m} f(\tau^m), \quad m \in \mathbb{Z} \quad (2.10)$$

Now, starting from (2.8) and substituting  $\mathcal{M}_f(\beta)$  from the Fourier series represen-

tation in (2.5) with  $\frac{m}{B} = \ln \tau^m$  yields

$$\begin{aligned}
 f(t) &= \int_{-\beta_0}^{\beta_0} \frac{1}{\sqrt{t}} \mathcal{M}_f(\beta) e^{-j2\pi\beta \ln t} d\beta \\
 &= \int_{-\beta_0}^{\beta_0} \frac{1}{\sqrt{t}} \sum_{m=-\infty}^{+\infty} a_m e^{j2\pi\beta \ln \tau^m} e^{-j2\pi\beta \ln t} d\beta \\
 &= \frac{1}{\sqrt{t}} \sum_{m=-\infty}^{+\infty} a_m \int_{-\beta_0}^{\beta_0} e^{-j2\pi\beta \ln(\tau^{-m}t)} d\beta \\
 &= \frac{1}{\sqrt{t}} \sum_{m=-\infty}^{+\infty} a_m \frac{1}{j2\pi \ln(\tau^{-m}t)} \left[ e^{j2\pi\beta_0 \ln(\tau^{-m}t)} - e^{-j2\pi\beta_0 \ln(\tau^{-m}t)} \right] \\
 &= \frac{1}{\sqrt{t}} \sum_{m=-\infty}^{+\infty} a_m \frac{\sin[\pi(2\beta_0) \ln(\tau^{-m}t)]}{\pi \ln(\tau^{-m}t)} \\
 &= \frac{1}{\sqrt{t}} \sum_{m=-\infty}^{+\infty} a_m \frac{\sin[\pi B \ln(\tau^{-m}t)]}{\pi \ln(\tau^{-m}t)}
 \end{aligned}$$

Finally, substituting  $a_m$  from (2.10),

$$f(t) = \frac{1}{\sqrt{t}} \sum_{m=-\infty}^{+\infty} \ln \tau \sqrt{\tau^m} f(\tau^m) \frac{\sin[\pi B \ln(\tau^{-m}t)]}{\pi \ln(\tau^{-m}t)} \quad (2.11)$$

$$= \ln \tau \sum_{m=-\infty}^{+\infty} f(\tau^m) \frac{\sin[\pi B \ln(\tau^{-m}t)]}{\pi \sqrt{\tau^{-m}t} \ln(\tau^{-m}t)} \quad (2.12)$$

$$= \ln \tau \sum_{m=-\infty}^{+\infty} f(\tau^m) \gamma(\tau^{-m}t) \quad (2.13)$$

where the function  $\gamma(t)$  is defined as

$$\gamma(t) = \frac{\sin[\pi B \ln t]}{\pi \sqrt{t} \ln t} = \frac{B}{\sqrt{t}} \text{sinc}(B \ln t) \quad (2.14)$$

with the *sinc* function defined as

$$\text{sinc } t = \frac{\sin(\pi t)}{\pi t}$$

□

For the sake of completeness, note that the Mellin transform of  $\gamma(t)$  is exactly the rectangular function with support  $[-\beta_0, \beta_0]$ , similarly to the Fourier case. Applying the definition in (2.1) on  $\gamma(t)$ ,

$$\begin{aligned}
 \mathcal{M}_\gamma(\beta) &= \int_0^{+\infty} \frac{1}{\sqrt{t}} \gamma(t) e^{+j2\pi\beta \ln t} dt \\
 &= \int_0^{+\infty} \frac{B}{t} \text{sinc}(B \ln t) e^{+j2\pi\beta \ln t} dt \\
 (u = \ln t \Rightarrow dt = t du) &= \int_{-\infty}^{+\infty} B \text{sinc}(Bu) e^{+j2\pi\beta u} du \\
 &= \mathcal{F}^{-1}[B \text{sinc}(Bu)] \\
 &= \text{rect}\left(\frac{\beta}{B}\right) = \\
 &= \text{rect}\left(\frac{\beta}{2\beta_0}\right) = \begin{cases} 1 & \text{if } |\beta| \leq \beta_0 \\ 0 & \text{elsewhere} \end{cases} \quad (2.15)
 \end{aligned}$$

where  $\mathcal{F}^{-1}$  denotes the inverse Fourier transform and the *rect* function is defined as

$$\text{rect } t = \begin{cases} 1 & \text{if } |t| \leq \frac{1}{2} \\ 0 & \text{elsewhere} \end{cases}$$

## 2.2. Input-Output Relation of a Wideband Time-Varying System

In a wideband time-varying system, the received signal  $y(t)$  is composed of a superposition of different versions of the transmitted signal  $x(t)$ , time shifted and Doppler scaled ([5] and [6]), whose input-output relation is given by

$$y(t) = \int_0^{+\infty} \int_{-\infty}^{+\infty} \chi(\tau, a) \sqrt{a} x(a(t - \tau)) d\tau da. \quad (2.16)$$

The propagation delay  $\tau$  is due to different scattering multipath lengths from transmitter to receiver and the time scale parameter  $a$  is due to relative motion of transmitter, scatters and receiver (Doppler effect). The scale parameter  $a$  can correspond either to a time expansion, if  $0 < a < 1$ , or to a time compression, if  $a > 1$ , depending on whether the scatter is moving away from the receiver or is approaching it. The *wideband spreading function*  $\chi(\tau, a)$  represents the strength of the scatterers.

Due to physical limitations of the system, the support of the spreading function is assumed to be finite.  $\chi(\tau, a)$  can be considered effectively nonzero only for  $(\tau, a) \in [0, T_d] \times [A_l, A_u]$ , where  $T_d$  is the multipath delay spread and  $A_s = A_u - A_l$  is the Doppler scale spread. According with this assumption, the input-output relation in (2.16) becomes

$$y(t) = \int_{A_l}^{A_u} \int_0^{T_d} \chi(\tau, a) \sqrt{a} x(a(t - \tau)) d\tau da. \quad (2.17)$$

## 2.3. Discrete Time-Scale Characterization

We follow the analysis presented in [10], to express (2.16) according to a corresponding discrete representation. The assumptions for the discrete time-scale model are presented here.

1. The Fourier transform  $X(f)$  of the transmitted signal  $x(t)$  has bounded support within  $f \in [-W/2, W/2]$ .
2. The Mellin transform  $\mathcal{M}_x(\beta)$  of  $x(t)$  is band limited within  $\beta \in [-\beta_0/2, \beta_0/2]$ .

These conditions could violate the uncertainty principle, particularly if the product of the scale support width and the bandwidth is small (see Section 2.5, Theorem 3). The analysis of the time-scale model is related only to the classes of input and output signals, for which the uncertainty principle is (approximately) valid.

Due to the assumptions above, it is reasonable to use the sampling theorems in the Fourier and in the Mellin domain (see Section 2.1), and consider the spreading function  $\chi(\tau, a)$ , sampled both in time delay  $\tau$  and in Doppler scale  $a$ .

Starting from the input-output relation (2.16) and using assumptions 1 and 2 above, a discrete time-scale characterization is derived as follows.

Function  $\theta(\tau, t)$  is defined as

$$\theta(\tau, t) \triangleq \chi^*(\tau, t)\sqrt{t}. \quad (2.18)$$

Indicating with  $\mathcal{M}_\theta(\tau, \beta)$  the Mellin transform of  $\theta(\tau, t)$  with respect to  $t$ ,  $\mathcal{M}_\theta^*(\tau, \beta)$  is given by

$$\begin{aligned} \mathcal{M}_\theta^*(\tau, \beta) &= \int_0^{+\infty} \frac{1}{\sqrt{t}} \theta^*(\tau, t) e^{-j2\pi\beta \ln t} dt \\ &= \int_0^{+\infty} \frac{1}{\sqrt{t}} \chi(\tau, t) \sqrt{t} e^{-j2\pi\beta \ln t} dt \\ &= \int_0^{+\infty} \chi(\tau, t) e^{-j2\pi\beta \ln t} dt. \end{aligned} \quad (2.19)$$

Using the definition (2.18),  $\chi(\tau, t) = \theta^*(\tau, t)/\sqrt{t}$ . Substituting this expression into (2.16), the input-output relation can be expressed as

$$\begin{aligned} y(t) &= \int_0^{+\infty} \int_{-\infty}^{+\infty} \chi(\tau, a) \sqrt{a} x(a(t - \tau)) d\tau da \\ &= \int_0^{+\infty} \int_{-\infty}^{+\infty} \theta^*(\tau, a) \frac{1}{\sqrt{a}} \sqrt{a} x(a(t - \tau)) d\tau da \\ &= \int_{-\infty}^{+\infty} \int_0^{+\infty} x(a(t - \tau)) \theta^*(\tau, a) da d\tau \\ \text{(definition (2.4))} &= \int_{-\infty}^{+\infty} (x \circledast \theta(\tau, \cdot))(t - \tau) d\tau. \end{aligned} \quad (2.20)$$

Now, applying the inverse Mellin transform (2.2) on function  $(x \circledast \theta(\tau, \cdot))(t - \tau)$  and using the multiplicative convolution property, the input-output relation (2.20) becomes

$$y(t) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{t-\tau}} \mathcal{M}_x(\beta) \mathcal{M}_\theta^*(\tau, \beta) e^{-j2\pi\beta \ln(t-\tau)} d\beta d\tau. \quad (2.21)$$

Assuming that the support of the Mellin transform of  $x(t)$  is bounded to  $[-\beta_0/2, \beta_0/2]$ ,  $\mathcal{M}_x(\beta)$  can be replaced with  $\mathcal{M}_x(\beta)P_{\beta_0}(\beta)$ , where

$$P_{\beta_0}(\beta) = \text{rect} \left( \frac{\beta}{\beta_0} \right) = \begin{cases} 1 & \text{if } |\beta| \leq \frac{\beta_0}{2} \\ 0 & \text{elsewhere} \end{cases} \quad (2.22)$$

After this replacement,  $\mathcal{M}_f(\tau, \beta) = \mathcal{M}_\theta^*(\tau, \beta)P_{\beta_0}(\beta)$  is a scale limited Mellin transform within  $[-\beta_0/2, \beta_0/2]$ . Its inverse Mellin transform  $f(\tau, a)$  is, by definition (2.2),

$$f(\tau, a) = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{a}} \mathcal{M}_\theta^*(\tau, \beta) P_{\beta_0}(\beta) e^{-j2\pi\beta \ln a} d\beta \quad (2.23)$$

and, according to the sampling theorem in scale domain (section 2.1, theorem 1), can be exactly reconstructed from its exponentially spaced samples,  $\{f(\tau, e^{m/B})\}_{m=-\infty}^{+\infty}$ , with  $B = \beta_0$ . Substituting the definitions (2.19) and (2.22) into (2.23) yields

$$\begin{aligned} f(\tau, a) &= \int_{-\frac{\beta_0}{2}}^{\frac{\beta_0}{2}} \frac{1}{\sqrt{a}} \int_0^{+\infty} \chi(\tau, t) e^{-j2\pi\beta \ln t} dt e^{-j2\pi\beta \ln a} d\beta \\ &= \int_0^{+\infty} \frac{1}{\sqrt{a}} \chi(\tau, t) \int_{-\frac{\beta_0}{2}}^{\frac{\beta_0}{2}} e^{-j2\pi\beta \ln(ta)} d\beta dt \\ &= \int_0^{+\infty} \frac{1}{\sqrt{a}} \chi(\tau, t) \beta_0 \text{sinc}(\beta_0 \ln(ta)) dt. \end{aligned} \quad (2.24)$$

From (2.24), the exponentially spaced samples of  $f(\tau, a)$ , with  $a = e^{m/\beta_0}$ ,  $m \in \mathbb{Z}$ , are given by

$$\begin{aligned} f(\tau, e^{m/\beta_0}) &= \int_0^{+\infty} \frac{1}{\sqrt{e^{m/\beta_0}}} \chi(\tau, t) \beta_0 \text{sinc}(\beta_0 \ln(te^{m/\beta_0})) dt \\ &= \int_0^{+\infty} \frac{1}{\sqrt{e^{m/\beta_0}}} \chi(\tau, t) \beta_0 \text{sinc}(m + \beta_0 \ln t) dt, \quad m \in \mathbb{Z}. \end{aligned} \quad (2.25)$$

According to the final result (2.12) of the sampling theorem in scale domain and using the samples in (2.25), function  $f(\tau, a)$  in (2.24) can be expressed as

$$\begin{aligned} f(\tau, a) &= \ln(e^{1/\beta_0}) \sum_{m=-\infty}^{+\infty} f(\tau, e^{m/\beta_0}) \frac{\sin(\pi\beta_0 \ln(ae^{-m/\beta_0}))}{\pi\sqrt{ae^{-m/\beta_0}} \ln(ae^{-m/\beta_0})} \\ &= \frac{1}{\beta_0} \sum_{m=-\infty}^{+\infty} \int_0^{+\infty} \frac{1}{\sqrt{e^{m/\beta_0}}} \chi(\tau, t) \beta_0 \text{sinc}(m + \beta_0 \ln t) dt \beta_0 \frac{\text{sinc}(\beta_0 \ln(ae^{-m/\beta_0}))}{\sqrt{ae^{-m/\beta_0}}} \\ &= \sum_{m=-\infty}^{+\infty} \int_0^{+\infty} \chi(\tau, t) \text{sinc}(m + \beta_0 \ln t) dt \frac{1}{\sqrt{a}} \text{sinc}(\beta_0 \ln(ae^{-m/\beta_0})). \end{aligned} \quad (2.26)$$

Finally, using the definition of the Mellin transform (2.1) on  $f(\tau, a)$  in (2.26), with respect to  $a$ ,  $\mathcal{M}_f(\tau, \beta) = \mathcal{M}_\theta^*(\tau, \beta)P_{\beta_0}(\beta)$  becomes

$$\begin{aligned}\mathcal{M}_\theta^*(\tau, \beta)P_{\beta_0}(\beta) &= \int_0^{+\infty} \frac{1}{\sqrt{a}} f(\tau, a) e^{j2\pi\beta \ln a} da \\ &= \sum_{m=-\infty}^{+\infty} \int_0^{+\infty} \chi(\tau, t) \operatorname{sinc}(m + \beta_0 \ln t) dt \\ &\quad \times \int_0^{+\infty} \frac{1}{a} \beta_0 \operatorname{sinc}(\beta_0 \ln(ae^{-m/\beta_0})) e^{j2\pi\beta \ln a} da. \quad (2.27)\end{aligned}$$

Using the definition (2.14) of  $\gamma(t)$  (with  $B = \beta_0$  in this case) on the last integral in (2.27),

$$\begin{aligned}\mathcal{M}_\theta^*(\tau, \beta)P_{\beta_0}(\beta) &= \sum_{m=-\infty}^{+\infty} \int_0^{+\infty} \chi(\tau, t) \operatorname{sinc}(m + \beta_0 \ln t) dt \\ &\quad \times \int_0^{+\infty} \frac{1}{\sqrt{a}} \sqrt{e^{-m/\beta_0}} \frac{\beta_0}{\sqrt{ae^{-m/\beta_0}}} \operatorname{sinc}(\beta_0 \ln(ae^{-m/\beta_0})) e^{j2\pi\beta \ln a} da \\ &= \sum_{m=-\infty}^{+\infty} \int_0^{+\infty} \chi(\tau, t) \operatorname{sinc}(m + \beta_0 \ln t) dt \\ &\quad \times \int_0^{+\infty} \frac{1}{\sqrt{a}} \left( \sqrt{e^{-m/\beta_0}} \gamma(ae^{-m/\beta_0}) \right) e^{j2\pi\beta \ln a} da. \quad (2.28)\end{aligned}$$

Using the expression of the Mellin transform  $\mathcal{M}_\gamma(\beta)$  in (2.15) and the scale invariance property (2.3) on the last integral in (2.28),

$$\begin{aligned}\mathcal{M}_\theta^*(\tau, \beta)P_{\beta_0}(\beta) &= \sum_{m=-\infty}^{+\infty} \int_0^{+\infty} \chi(\tau, t) \operatorname{sinc}(m + \beta_0 \ln t) dt e^{-j2\pi\beta \ln e^{-m/\beta_0}} \operatorname{rect}\left(\frac{\beta}{\beta_0}\right) \\ &= \sum_{m=-\infty}^{+\infty} \int_0^{+\infty} \chi(\tau, t) \operatorname{sinc}(m + \beta_0 \ln t) dt e^{j2\pi m\beta/\beta_0} \operatorname{rect}\left(\frac{\beta}{\beta_0}\right) \\ \left(\frac{1}{\beta_0} = \ln a_0\right) &= \sum_{m=-\infty}^{+\infty} \int_0^{+\infty} \chi(\tau, t) \operatorname{sinc}\left(m + \frac{\ln t}{\ln a_0}\right) dt e^{j2\pi m\beta \ln a_0} \operatorname{rect}\left(\frac{\beta}{\beta_0}\right). \quad (2.29)\end{aligned}$$

From (2.29), with a sign change of variable  $m$ , the final expression of  $\mathcal{M}_\theta^*(\tau, \beta)P_{\beta_0}(\beta)$  is given by

$$\mathcal{M}_\theta^*(\tau, \beta)P_{\beta_0}(\beta) = \sum_{m=-\infty}^{+\infty} \int_0^{+\infty} \chi(\tau, a) \operatorname{sinc}\left(m - \frac{\ln a}{\ln a_0}\right) da e^{-j2\pi m\beta \ln a_0} \operatorname{rect}\left(\frac{\beta}{\beta_0}\right). \quad (2.30)$$

Substituting (2.30) into (2.21), the input-output relation becomes

$$\begin{aligned}
 y(t) &= \sum_{m=-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_0^{+\infty} \int_{-\infty}^{+\infty} \chi(\tau, a) \operatorname{sinc} \left( m - \frac{\ln a}{\ln a_0} \right) \\
 &\quad \times \frac{1}{\sqrt{t-\tau}} \mathcal{M}_x(\beta) e^{-j2\pi\beta \ln(a_0^m(t-\tau))} d\beta da d\tau \\
 &= \sum_{m=-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_0^{+\infty} \chi(\tau, a) \operatorname{sinc} \left( m - \frac{\ln a}{\ln a_0} \right) \\
 &\quad \times \sqrt{a_0^m} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{a_0^m(t-\tau)}} \mathcal{M}_x(\beta) e^{-j2\pi\beta \ln(a_0^m(t-\tau))} d\beta da d\tau.
 \end{aligned}$$

Applying the inverse Mellin transform (2.2) in the last integral,

$$\begin{aligned}
 y(t) &= \sum_{m=-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_0^{+\infty} \chi(\tau, a) \operatorname{sinc} \left( m - \frac{\ln a}{\ln a_0} \right) \sqrt{a_0^m} x(a_0^m(t-\tau)) da d\tau \\
 &= \sum_{m=-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_0^{+\infty} \chi(\tau, a) \operatorname{sinc} \left( \frac{\ln a_0^m - \ln a}{\ln a_0} \right) da a_0^{\frac{m}{2}} x(a_0^m(t-\tau)) d\tau \\
 &= \sum_{m=-\infty}^{+\infty} \int_{-\infty}^{+\infty} \tilde{\chi}(\tau, a_0^m) a_0^{\frac{m}{2}} x(a_0^m(t-\tau)) d\tau \tag{2.31}
 \end{aligned}$$

where function  $\tilde{\chi}(\tau, a)$  is a scale-smoothed version of the spreading function  $\chi(\tau, a)$

$$\tilde{\chi}(\tau, a) = \int_0^{+\infty} \chi(\tau, a') \operatorname{sinc} \left( \frac{\ln a - \ln a'}{\ln a_0} \right) da'. \tag{2.32}$$

The final discrete time-scale model results from (2.31), using a similar procedure as before but based on the Fourier transform, which is detailed as follows.

The integral in (2.31) is the convolution of the functions  $\tilde{\chi}(\tau, a_0^m)$  and  $a_0^{\frac{m}{2}} x(a_0^m \tau)$ . Applying the definition of the inverse Fourier transform and using the convolution property of the Fourier transform, the input-output relation in (2.31) becomes

$$\begin{aligned}
 y(t) &= \sum_{m=-\infty}^{+\infty} \left( \tilde{\chi}(\tau, \cdot) * a_0^{\frac{m}{2}} x(a_0^m \tau) \right) (t) \\
 &= \sum_{m=-\infty}^{+\infty} \int_{-\infty}^{+\infty} \tilde{U}(f, a_0^m) a_0^{-\frac{m}{2}} X(a_0^{-m} f) e^{j2\pi f t} df \tag{2.33}
 \end{aligned}$$

where  $\tilde{U}(f, a)$  is the Fourier transform of  $\tilde{\chi}(\tau, a)$ , with respect to  $\tau$ , and  $X(f)$  is the Fourier transform of  $x(\tau)$ . Note that  $a_0^{-\frac{m}{2}} X(a_0^{-m} f)$  is the Fourier transform of  $a_0^{\frac{m}{2}} x(a_0^m \tau)$ .

Assuming that the support of the Fourier transform  $X(f)$  is band limited to  $[-W/2, W/2]$ , then the Fourier transform  $a_0^{-\frac{m}{2}} X(a_0^{-m} f)$  is band limited to the range

$[-a_0^m W/2, a_0^m W/2]$  and can be replaced with  $a_0^{-\frac{m}{2}} X(a_0^{-m} f) P_{a_0^m W}(f)$ , where

$$P_{a_0^m W}(f) = \text{rect}\left(\frac{f}{a_0^m W}\right) = \begin{cases} 1 & \text{if } |f| \leq \frac{a_0^m W}{2} \\ 0 & \text{elsewhere} \end{cases} \quad (2.34)$$

After this replacement,  $G(f, a_0^m) = \tilde{U}(f, a) P_{a_0^m W}(f)$  is a band limited Fourier transform. Its inverse Fourier transform is

$$g(t, a_0^m) = \int_{-\infty}^{+\infty} \tilde{U}(f, a) P_{a_0^m W}(f) e^{j2\pi ft} df \quad (2.35)$$

and, according to the sampling theorem in Fourier domain, can be exactly reconstructed from its uniformly spaced samples,  $\{g(n/(a_0^m W), a_0^m)\}_{n=-\infty}^{+\infty}$ . Substituting the definitions of  $\tilde{U}(f, a)$  and  $P_{a_0^m W}(f)$  into (2.35) yields

$$\begin{aligned} g(t, a_0^m) &= \int_{-a_0^m \frac{W}{2}}^{a_0^m \frac{W}{2}} \int_{-\infty}^{+\infty} \tilde{\chi}(\tau', a_0^m) e^{-j2\pi\tau'f} d\tau' e^{j2\pi ft} df \\ &= \int_{-\infty}^{+\infty} \tilde{\chi}(\tau', a_0^m) \int_{-a_0^m \frac{W}{2}}^{a_0^m \frac{W}{2}} e^{-j2\pi(\tau'-t)f} df d\tau' \\ &= \int_{-\infty}^{+\infty} \tilde{\chi}(\tau', a_0^m) a_0^m W \text{sinc}(a_0^m W(\tau' - t)) d\tau'. \end{aligned} \quad (2.36)$$

From (2.36), the uniformly spaced samples of  $g(t, a_0^m)$  are given by

$$\begin{aligned} g\left(\frac{n}{a_0^m W}, a_0^m\right) &= \int_{-\infty}^{+\infty} \tilde{\chi}(\tau', a_0^m) a_0^m W \text{sinc}\left(a_0^m W\left(\tau' - \frac{n}{a_0^m W}\right)\right) d\tau' \\ &= \int_{-\infty}^{+\infty} \tilde{\chi}(\tau', a_0^m) a_0^m W \text{sinc}(n - a_0^m W\tau') d\tau', \quad n \in \mathbb{Z}. \end{aligned} \quad (2.37)$$

According to the sampling theorem in Fourier domain, function  $g(t, a_0^m)$  can be expressed as

$$\begin{aligned} g(t, a_0^m) &= \sum_{n=-\infty}^{+\infty} g\left(\frac{n}{a_0^m W}, a_0^m\right) \text{sinc}\left(a_0^m W\left(t - \frac{n}{a_0^m W}\right)\right) \\ &= \sum_{n=-\infty}^{+\infty} \int_{-\infty}^{+\infty} \tilde{\chi}(\tau', a_0^m) a_0^m W \text{sinc}(n - a_0^m W\tau') d\tau' \\ &\quad \times \text{sinc}\left(a_0^m W\left(t - \frac{n}{a_0^m W}\right)\right). \end{aligned} \quad (2.38)$$



Finally, applying the Fourier transform in (2.38), with respect to  $t$ ,  $G(f, a_0^m) = \tilde{U}(f, a)P_{a_0^m W}(f)$  becomes

$$\begin{aligned}
 \tilde{U}(f, a)P_{a_0^m W}(f) &= \int_{-\infty}^{+\infty} g(t, a_0^m) e^{-j2\pi ft} dt \\
 &= \sum_{n=-\infty}^{+\infty} \int_{-\infty}^{+\infty} \tilde{\chi}(\tau', a_0^m) \operatorname{sinc}(n - a_0^m W \tau') d\tau' \\
 &\quad \times \int_{-\infty}^{+\infty} a_0^m W \operatorname{sinc}\left(a_0^m W \left(t - \frac{n}{a_0^m W}\right)\right) e^{-j2\pi ft} dt \\
 &= \sum_{n=-\infty}^{+\infty} \int_{-\infty}^{+\infty} \tilde{\chi}(\tau', a_0^m) \operatorname{sinc}(n - a_0^m W \tau') d\tau' \\
 &\quad \times e^{-j2\pi \frac{nf}{a_0^m W}} \operatorname{rect}\left(\frac{f}{a_0^m W}\right). \tag{2.39}
 \end{aligned}$$

Substituting (2.39) into (2.33), the input-output relation becomes

$$\begin{aligned}
 y(t) &= \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \tilde{\chi}(\tau', a_0^m) \operatorname{sinc}(n - a_0^m W \tau') d\tau' \\
 &\quad \times e^{-j2\pi \frac{nf}{a_0^m W}} a_0^{-\frac{m}{2}} X(a_0^{-m} f) e^{j2\pi ft} df \\
 &= \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} \int_{-\infty}^{+\infty} \tilde{\chi}(\tau', a_0^m) \operatorname{sinc}(n - a_0^m W \tau') d\tau' \\
 &\quad \times \int_{-\infty}^{+\infty} a_0^{-\frac{m}{2}} X(a_0^{-m} f) e^{j2\pi f \left(t - \frac{n}{a_0^m W}\right)} df \\
 &= \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} \int_{-\infty}^{+\infty} \tilde{\chi}(\tau', a_0^m) \operatorname{sinc}(n - a_0^m W \tau') d\tau' a_0^{\frac{m}{2}} x\left(a_0^m t - \frac{n}{W}\right) \\
 &\quad \text{(substituting the expression of } \tilde{\chi}(\tau', a_0^m) \text{ from (2.32))} \\
 &= \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} \int_0^{+\infty} \int_{-\infty}^{+\infty} \chi(\tau, a') \operatorname{sinc}\left(\frac{\ln a - \ln a'}{\ln a_0}\right) \operatorname{sinc}(n - a_0^m W \tau') d\tau' da' \\
 &\quad \times a_0^{\frac{m}{2}} x\left(a_0^m t - \frac{n}{W}\right) \\
 &= \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} \hat{\chi}\left(\frac{n}{a_0^m W}, a_0^m\right) a_0^{\frac{m}{2}} x\left(a_0^m t - \frac{n}{W}\right) \tag{2.40}
 \end{aligned}$$

where  $\hat{\chi}(\tau, a_0^m)$  is a time-smoothed version of  $\tilde{\chi}(\tau, a_0^m)$

$$\hat{\chi}(\tau, a_0^m) = \int_{-\infty}^{+\infty} \tilde{\chi}(\tau', a_0^m) \operatorname{sinc}(a_0^m W(\tau - \tau')) d\tau'. \tag{2.41}$$

This result is summarized in the following theorem.

**Theorem 2.** *In a wideband time-varying system, if the transmitted signal  $x(t)$  is band limited and scale limited, i.e. its Fourier transform  $X(f) \equiv 0$  if  $f \notin [-W/2, W/2]$  and its Mellin transform  $\mathcal{M}_x(\beta) \equiv 0$  if  $\beta \notin [-\beta_0/2, \beta_0/2]$ , the received signal  $y(t)$  is decomposed into discrete time shifts and Doppler scalings on the input signal  $x(t)$ , weighted by a smoothed and sampled version of the wideband spreading function.*

$$y(t) = \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} \hat{\chi} \left( \frac{n}{a_0^m W}, a_0^m \right) a_0^{\frac{m}{2}} x \left( a_0^m t - \frac{n}{W} \right) \quad (2.42)$$

where

$$\hat{\chi}(\tau, a_0^m) = \int_0^{+\infty} \int_{-\infty}^{+\infty} \chi(\tau', a') \operatorname{sinc} \left( \frac{\ln a - \ln a'}{\ln a_0} \right) \operatorname{sinc}(a_0^m W(\tau - \tau')) d\tau' da' \quad (2.43)$$

The relation (2.43) is obtained combining (2.32) and (2.41).

## 2.4. Finite Approximation

Both summations in (2.42) involve infinitely many terms. However, due to the physical system restrictions described in section 2.2, the support of spreading function  $\chi(\tau, a)$  is bounded in both time and scale domains, as expressed in (2.17). Specifically, if  $\chi(\tau, a)$  is nonzero only when  $A_l \leq a \leq A_u$ , then  $\tilde{\chi}(\tau, a_0^m)$  defined in (2.32) can be expressed as

$$\begin{aligned} \tilde{\chi}(\tau, a_0^m) &= \int_{A_l}^{A_u} \chi(\tau, a') \operatorname{sinc} \left( \frac{\ln a_0^m - \ln a'}{\ln a_0} \right) da' \\ (\gamma = \ln a' \Rightarrow da' = e^\gamma d\gamma) &= \int_{\ln A_l}^{\ln A_u} \chi(\tau, e^\gamma) \operatorname{sinc} \left( \frac{\ln a_0^m - \gamma}{\ln a_0} \right) e^\gamma d\gamma \\ &= \int_{\ln A_l}^{\ln A_u} \chi(\tau, e^\gamma) \operatorname{sinc} \left( m - \frac{\gamma}{\ln a_0} \right) e^\gamma d\gamma. \end{aligned} \quad (2.44)$$

Considering the approximation that the sinc function in (2.44) is nonzero only in the mainlobe, values of  $m$  corresponding to nonzero coefficients of  $\tilde{\chi}(\tau, a_0^m)$  are

$$\begin{aligned} -1 &\leq m - \frac{\gamma}{\ln a_0} \leq 1 \\ \frac{\gamma}{\ln a_0} - 1 &\leq m \leq \frac{\gamma}{\ln a_0} + 1 \end{aligned} \quad (2.45)$$

From the integral in (2.44), substituting maximum and minimum values of  $\gamma$ , the condition (2.45) becomes

$$\frac{\ln A_l}{\ln a_0} - 1 \leq m \leq \frac{\ln A_u}{\ln a_0} + 1$$

or, since  $m$  assumes only integer values,

$$\left\lfloor \frac{\ln A_l}{\ln a_0} \right\rfloor \leq m \leq \left\lceil \frac{\ln A_u}{\ln a_0} \right\rceil \quad (2.46)$$

Similarly, if  $\chi(\tau, a)$  is nonzero only when  $0 \leq \tau \leq T_d$ , than  $\hat{\chi}(n/(a_0^m W), a_0^m)$  defined in (2.41) can be expressed as

$$\hat{\chi}\left(\frac{n}{a_0^m W}, a_0^m\right) = \int_0^{T_d} \tilde{\chi}(\tau', a_0^m) \operatorname{sinc}(n - a_0^m W \tau') d\tau'. \quad (2.47)$$

Applying in (2.47) the sinc approximation as before, values of  $n$  corresponding to nonzero coefficients of  $\hat{\chi}(n/(a_0^m W), a_0^m)$  are

$$\begin{aligned} -1 &\leq n - a_0^m W \tau' \leq 1 \\ a_0^m W \tau' - 1 &\leq n \leq a_0^m W \tau' + 1 \end{aligned} \quad (2.48)$$

From the integral in (2.47), substituting maximum and minimum values of  $\tau'$  in (2.48), it results that

$$-1 \leq n \leq a_0^m W T_d + 1$$

or, approximately,

$$0 \leq n \leq \lceil a_0^m W T_d \rceil \quad (2.49)$$

Combining the approximations (2.46) and (2.49) due to the limited system support, the input-output relation in (2.16) admits the following finite-dimensional representation

$$y(t) \approx \sum_{m=M_0}^{M_1} \sum_{n=0}^{N(m)} \chi_{n,m} x_{n,m}(t) \quad (2.50)$$

where  $M_0 = \lfloor \ln A_l / \ln a_0 \rfloor$ ,  $M_1 = \lceil \ln A_u / \ln a_0 \rceil$ ,  $N(m) = \lceil a_0^m W T_d \rceil$ ,

$$\chi_{n,m} = \hat{\chi}\left(\frac{n}{a_0^m W}, a_0^m\right) \quad (2.51)$$

and  $x_{n,m}(t)$ , a time shifted and scaled version of  $x(t)$ ,

$$x_{n,m}(t) = a_0^{\frac{m}{2}} x\left(a_0^m t - \frac{n}{W}\right). \quad (2.52)$$

The number of nonzero coefficients in (2.50) are given by

$$\begin{aligned} M &= \sum_{m=M_0}^{M_1} (N(m) + 1) \\ &= (M_1 - M_0 + 1) + \sum_{m=M_0}^{M_1} N(m) \\ &= (M_1 - M_0 + 1) + \sum_{m=M_0}^{M_1} \lceil a_0^m W T_d \rceil \end{aligned} \quad (2.53)$$

or, without considering the ceil function in the right hand side of (2.53),

$$\begin{aligned}
 M &\approx (M_1 - M_0 + 1) + \sum_{m=M_0}^{M_1} a_0^m W T_d \\
 &= (M_1 - M_0 + 1) + W T_d \sum_{m=M_0}^{M_1} a_0^m \\
 &= (M_1 - M_0 + 1) + W T_d \frac{a_0^{M_1+1} - a_0^{M_0}}{a_0 - 1} \tag{2.54}
 \end{aligned}$$

From (2.54), approximating  $M_0 = \ln A_l / \ln a_0$  and  $M_1 = \ln A_u / \ln a_0$

$$\begin{aligned}
 M &\approx \left( \frac{\ln A_u}{\ln a_0} - \frac{\ln A_l}{\ln a_0} + 1 \right) + W T_d \frac{a_0^{\frac{\ln A_u}{\ln a_0} + 1} - a_0^{\frac{\ln A_l}{\ln a_0}}}{a_0 - 1} \\
 \text{(change of base formula)} &= \left( \frac{\ln \frac{A_u}{A_l}}{\ln a_0} + 1 \right) + W T_d \frac{a_0 A_u - A_l}{a_0 - 1} \\
 &= \left( \frac{\ln \frac{A_u}{A_l}}{\ln a_0} + 1 \right) + W T_d A_l \frac{a_0 \frac{A_u}{A_l} - 1}{a_0 - 1} \tag{2.55}
 \end{aligned}$$

## 2.5. Uncertainty Principle for Frequency and Scale Operators

In this section we express the constraints intrinsically related to a signal with bounded Fourier and Mellin transforms, since our discrete model in section 2.3 is based on these assumptions. At first we state the uncertainty principle for two arbitrary operators, as given in [11] and [14], and then derive the results for frequency and scale operators.

Some important notions are provided as a preliminary to the uncertainty principle.

For a physical quantity  $a$ , there is always an operator  $\mathcal{A}$  associated to. For two physical quantities  $a$  and  $b$ , with operators  $\mathcal{A}$  and  $\mathcal{B}$ , the operator  $\mathcal{A}\mathcal{B}$  means to operate first with  $\mathcal{B}$  and then with  $\mathcal{A}$ .

The *commutator* of two operators  $\mathcal{A}$  and  $\mathcal{B}$  is an operator defined as

$$[\mathcal{A}, \mathcal{B}] \triangleq \mathcal{A}\mathcal{B} - \mathcal{B}\mathcal{A}. \tag{2.56}$$

If  $[\mathcal{A}, \mathcal{B}] = 0$  or, equivalently, if  $\mathcal{A}\mathcal{B} = \mathcal{B}\mathcal{A}$ , the two operators commute.

The *anticommutator* is defined as

$$[\mathcal{A}, \mathcal{B}]_+ \triangleq \mathcal{A}\mathcal{B} + \mathcal{B}\mathcal{A}. \tag{2.57}$$

There is an uncertainty principle for any two quantities which are represented by operators which do not commute.

Given a signal  $f(a)$  with unitary energy, i.e.  $\int |f(a)|^2 da = 1$ , the *average* or *mean* of  $a$  is

$$\langle a \rangle = \int a |f(a)|^2 da. \quad (2.58)$$

In the above definition, function  $|f(a)|^2$  can be considered as the density function of  $a$ . The *variance*  $\sigma_a^2$  of  $a$  is defined as the average of  $(a - \langle a \rangle)^2$

$$\begin{aligned} \sigma_a^2 &= \int (a - \langle a \rangle)^2 |f(a)|^2 da \\ &= \int a^2 |f(a)|^2 da + \int \langle a \rangle^2 |f(a)|^2 da - 2 \int a \langle a \rangle |f(a)|^2 da \\ &= \langle a^2 \rangle + \langle a \rangle^2 - 2 \langle a \rangle^2 \\ &= \langle a^2 \rangle - \langle a \rangle^2. \end{aligned} \quad (2.59)$$

The *standard deviation*  $\sigma_a$  is defined as the square root of the variance.

Suppose to represent signal  $f(a)$  in a different domain, through an unitary transformation that associates to  $f(a)$  a new signal  $s(t)$ . Then, the *operator*  $\mathcal{A}$  is associated to the variable  $a$  if the following relation is satisfied

$$\langle \mathcal{A} \rangle \triangleq \int s^*(t) \mathcal{A} s(t) dt = \langle a \rangle. \quad (2.60)$$

Depending on the signal  $s(t)$  and on the domain where it is defined, the same operator  $\mathcal{A}$  can be expressed in different forms, as it will be clarified later for time and frequency operators.

The operator  $\mathcal{A}$  associated to  $a$  is used in (2.60) to express the mean  $\langle a \rangle$  in (2.58), but it can also be used to express the variance  $\sigma_a^2$  in (2.59) as the average of  $(\mathcal{A} - \langle \mathcal{A} \rangle)^2$ , according to

$$\begin{aligned} \sigma_a^2 &= \int s^*(t) (\mathcal{A} - \langle \mathcal{A} \rangle)^2 s(t) dt \\ &= \int s^*(t) \mathcal{A}^2 s(t) dt + \int s^*(t) \langle \mathcal{A} \rangle^2 s(t) dt - 2 \int s^*(t) \mathcal{A} \langle \mathcal{A} \rangle s(t) dt \\ &= \langle \mathcal{A}^2 \rangle + \langle \mathcal{A} \rangle^2 - 2 \langle \mathcal{A} \rangle^2 \\ &= \langle \mathcal{A}^2 \rangle - \langle \mathcal{A} \rangle^2. \end{aligned} \quad (2.61)$$

Given two operators  $\mathcal{A}$  and  $\mathcal{B}$ , associated to the variables  $a$  and  $b$  respectively, then the *covariance*  $\text{Cov}_{ab}$  between  $a$  and  $b$  is defined as

$$\begin{aligned} \text{Cov}_{ab} &= \frac{1}{2} \langle \mathcal{A} \mathcal{B} + \mathcal{B} \mathcal{A} \rangle - \langle \mathcal{A} \rangle \langle \mathcal{B} \rangle \\ &= \frac{1}{2} \langle [\mathcal{A}, \mathcal{B}]_+ \rangle - \langle \mathcal{A} \rangle \langle \mathcal{B} \rangle. \end{aligned} \quad (2.62)$$

An operator  $\mathcal{A}$  is *Hermitian* if for any pair of functions  $f(t)$  and  $g(t)$

$$\int f^*(t) \mathcal{A} g(t) dt = \int g(t) \{ \mathcal{A} f(t) \}^* dt. \quad (2.63)$$

If the operator  $\mathcal{A}$  is Hermitian, then the following properties hold.

1. The mean  $\langle \mathcal{A} \rangle$  is real, as

$$\begin{aligned}
 \langle \mathcal{A} \rangle &= \int s^*(t) \mathcal{A} s(t) dt \\
 \text{(Hermitian def. (2.63))} &= \int s(t) \{ \mathcal{A} s(t) \}^* dt \\
 &= \left\{ \int s^*(t) \mathcal{A} s(t) dt \right\}^* \\
 &= \langle \mathcal{A} \rangle^* \tag{2.64}
 \end{aligned}$$

2. The operator  $\mathcal{A} - \langle \mathcal{A} \rangle$  is also Hermitian, as

$$\begin{aligned}
 \int f^*(t) (\mathcal{A} - \langle \mathcal{A} \rangle) g(t) dt &= \int [f^*(t) \mathcal{A} g(t) - f^*(t) \langle \mathcal{A} \rangle g(t)] dt \\
 \text{((2.64) and Hermitian def. (2.63))} &= \int [g(t) \{ \mathcal{A} f(t) \}^* - f^*(t) \langle \mathcal{A} \rangle^* g(t)] dt \\
 &= \int [g(t) \{ \mathcal{A} f(t) \}^* - g(t) \{ \langle \mathcal{A} \rangle f(t) \}^*] dt \\
 &= \int g(t) \{ (\mathcal{A} - \langle \mathcal{A} \rangle) f(t) \}^* dt \tag{2.65}
 \end{aligned}$$

3. The mean of the operator  $\mathcal{A}^2$  is

$$\langle \mathcal{A}^2 \rangle = \int s^*(t) \mathcal{A}^2 s(t) dt = \int |\mathcal{A} s(t)|^2 dt \tag{2.66}$$

$$\begin{aligned}
 \langle \mathcal{A}^2 \rangle &= \int s^*(t) \mathcal{A} (\mathcal{A} s(t)) dt \\
 \text{(Hermitian def. (2.63) on } s(t) \text{ and } \mathcal{A} s(t)) &= \int \mathcal{A} s(t) \{ \mathcal{A} s(t) \}^* dt \\
 &= \int |\mathcal{A} s(t)|^2 dt
 \end{aligned}$$

The *adjoint*  $\mathcal{A}^\dagger$  of an operator  $\mathcal{A}$  is another operator for which the following equality holds

$$\int f^*(t) \mathcal{A} g(t) dt = \int g(t) \{ \mathcal{A}^\dagger f(t) \}^* dt. \tag{2.67}$$

If  $\mathcal{A}^\dagger = \mathcal{A}$ , then condition (2.67) becomes the definition of a Hermitian operator, and  $\mathcal{A}$  is called *self adjoint* operator. The adjoint of a product of operators is given by

$$(\mathcal{A}\mathcal{B})^\dagger = \mathcal{B}^\dagger \mathcal{A}^\dagger. \tag{2.68}$$

For an arbitrary operator  $\mathcal{A}$ , the following decomposition in Hermitian operators holds

$$\mathcal{A} = \frac{1}{2}(\mathcal{A} + \mathcal{A}^\dagger) + \frac{1}{2}j(\mathcal{A} - \mathcal{A}^\dagger)/j. \tag{2.69}$$

In fact, operators  $\mathcal{A} + \mathcal{A}^\dagger$  and  $(\mathcal{A} - \mathcal{A}^\dagger)/j$  are Hermitian whether  $\mathcal{A}$  is Hermitian or not.

The anticommutator of two Hermitian operators  $\mathcal{A}$  and  $\mathcal{B}$  is also Hermitian, as

$$\begin{aligned}
 [\mathcal{A}, \mathcal{B}]_+^\dagger &= (\mathcal{A}\mathcal{B} + \mathcal{B}\mathcal{A})^\dagger \\
 &= (\mathcal{A}\mathcal{B})^\dagger + (\mathcal{B}\mathcal{A})^\dagger \\
 \text{(property (2.68))} &= \mathcal{B}^\dagger\mathcal{A}^\dagger + \mathcal{A}^\dagger\mathcal{B}^\dagger \\
 (\mathcal{A}^\dagger = \mathcal{A} \text{ and } \mathcal{B}^\dagger = \mathcal{B}) &= \mathcal{B}\mathcal{A} + \mathcal{A}\mathcal{B} \\
 &= [\mathcal{A}, \mathcal{B}]_+
 \end{aligned} \tag{2.70}$$

Having  $[\mathcal{A}, \mathcal{B}]_+^\dagger = [\mathcal{A}, \mathcal{B}]_+$ , the anticommutator is Hermitian. The product of two Hermitian operators  $\mathcal{A}\mathcal{B}$  is not necessarily Hermitian. However, using the decomposition (2.69), the operator  $\mathcal{A}\mathcal{B}$  can be also expressed as

$$\begin{aligned}
 \mathcal{A}\mathcal{B} &= \frac{1}{2}(\mathcal{A}\mathcal{B} + (\mathcal{A}\mathcal{B})^\dagger) + \frac{1}{2}(\mathcal{A}\mathcal{B} - (\mathcal{A}\mathcal{B})^\dagger) \\
 \text{(property (2.68))} &= \frac{1}{2}(\mathcal{A}\mathcal{B} + \mathcal{B}^\dagger\mathcal{A}^\dagger) + \frac{1}{2}(\mathcal{A}\mathcal{B} - \mathcal{B}^\dagger\mathcal{A}^\dagger) \\
 (\mathcal{A}^\dagger = \mathcal{A} \text{ and } \mathcal{B}^\dagger = \mathcal{B}) &= \frac{1}{2}(\mathcal{A}\mathcal{B} + \mathcal{B}\mathcal{A}) + \frac{1}{2}(\mathcal{A}\mathcal{B} - \mathcal{B}\mathcal{A}) \\
 \text{(def. (2.56) and (2.57))} &= \frac{1}{2}[\mathcal{A}, \mathcal{B}]_+ + \frac{j}{2}[\mathcal{A}, \mathcal{B}]/j
 \end{aligned} \tag{2.71}$$

**Theorem 3. Uncertainty Principle.** For any two quantities  $a$  and  $b$  represented by the respective Hermitian operators  $\mathcal{A}$  and  $\mathcal{B}$ , which do not commute, the uncertainty principle is

$$\sigma_a\sigma_b \geq \frac{1}{2}|\langle[\mathcal{A}, \mathcal{B}]\rangle| \tag{2.72}$$

where  $\sigma_a$  and  $\sigma_b$  are the standard deviations of  $a$  and  $b$ , respectively, and  $\langle[\mathcal{A}, \mathcal{B}]\rangle$  is the mean of the commutator of  $\mathcal{A}$  and  $\mathcal{B}$ .

*Proof.* Define operators  $\mathcal{A}_0 = \mathcal{A} - \langle\mathcal{A}\rangle$  and  $\mathcal{B}_0 = \mathcal{B} - \langle\mathcal{B}\rangle$ . From property (2.65), they are Hermitian and their mean is zero. The commutator of  $\mathcal{A}_0$  and  $\mathcal{B}_0$  is equal to that of  $\mathcal{A}$  and  $\mathcal{B}$ , as

$$\begin{aligned}
 [\mathcal{A}_0, \mathcal{B}_0] &= \mathcal{A}_0\mathcal{B}_0 - \mathcal{B}_0\mathcal{A}_0 \\
 &= (\mathcal{A} - \langle\mathcal{A}\rangle)(\mathcal{B} - \langle\mathcal{B}\rangle) - (\mathcal{B} - \langle\mathcal{B}\rangle)(\mathcal{A} - \langle\mathcal{A}\rangle) \\
 &= \mathcal{A}\mathcal{B} - \mathcal{B}\mathcal{A} = [\mathcal{A}, \mathcal{B}].
 \end{aligned} \tag{2.73}$$

The anticommutator of  $\mathcal{A}_0$  and  $\mathcal{B}_0$  is

$$\begin{aligned}
 [\mathcal{A}_0, \mathcal{B}_0]_+ &= \mathcal{A}_0\mathcal{B}_0 + \mathcal{B}_0\mathcal{A}_0 \\
 &= (\mathcal{A} - \langle\mathcal{A}\rangle)(\mathcal{B} - \langle\mathcal{B}\rangle) + (\mathcal{B} - \langle\mathcal{B}\rangle)(\mathcal{A} - \langle\mathcal{A}\rangle) \\
 &= \mathcal{A}\mathcal{B} + \mathcal{B}\mathcal{A} - 2\langle\mathcal{B}\rangle\mathcal{A} - 2\langle\mathcal{A}\rangle\mathcal{B} + 2\langle\mathcal{A}\rangle\langle\mathcal{B}\rangle \\
 &= [\mathcal{A}, \mathcal{B}]_+ - 2\langle\mathcal{B}\rangle\mathcal{A} - 2\langle\mathcal{A}\rangle\mathcal{B} + 2\langle\mathcal{A}\rangle\langle\mathcal{B}\rangle.
 \end{aligned} \tag{2.74}$$

Taking the expectation of both sides of (2.74) and using the definition (2.62) yields

$$\langle [\mathcal{A}_0, \mathcal{B}_0]_+ \rangle = 2 \left\{ \frac{1}{2} \langle [\mathcal{A}, \mathcal{B}]_+ \rangle - \langle \mathcal{A} \rangle \langle \mathcal{B} \rangle \right\} = 2 \text{Cov}_{ab}. \quad (2.75)$$

Using the definition of variance in (2.61),

$$\begin{aligned} \sigma_a^2 \sigma_b^2 &= \int s^*(t) (\mathcal{A} - \langle \mathcal{A} \rangle)^2 s(t) dt \times \int s^*(t) (\mathcal{B} - \langle \mathcal{B} \rangle)^2 s(t) dt \\ (\text{def. of } \mathcal{A}_0 \text{ and } \mathcal{B}_0) &= \int s^*(t) \mathcal{A}_0^2 s(t) dt \times \int s^*(t) \mathcal{B}_0^2 s(t) dt \\ (\text{property (2.66)}) &= \int |\mathcal{A}_0 s(t)|^2 dt \times \int |\mathcal{B}_0 s(t)|^2 dt \\ (\text{Schwarz inequality}) &\geq \left| \int \{ \mathcal{A}_0 s(t) \}^* \{ \mathcal{B}_0 s(t) \} dt \right|^2 \\ (\text{Hermitian def. (2.63) on } \mathcal{B}_0) &= \left| \int s^*(t) \mathcal{A}_0 \mathcal{B}_0 s(t) dt \right|^2 \\ (\text{def. (2.60)}) &= |\langle \mathcal{A}_0 \mathcal{B}_0 \rangle|^2. \end{aligned} \quad (2.76)$$

Expressing the product  $\mathcal{A}_0 \mathcal{B}_0$  as in (2.71)

$$\mathcal{A}_0 \mathcal{B}_0 = \frac{1}{2} [\mathcal{A}_0, \mathcal{B}_0]_+ + \frac{j}{2} [\mathcal{A}_0, \mathcal{B}_0]/j \quad (2.77)$$

and taking the expectation of both sides of (2.77) yields

$$\begin{aligned} \langle \mathcal{A}_0 \mathcal{B}_0 \rangle &= \frac{1}{2} \langle [\mathcal{A}_0, \mathcal{B}_0]_+ \rangle + \frac{j}{2} \langle [\mathcal{A}_0, \mathcal{B}_0]/j \rangle \\ (\text{prop. (2.75)}) &= \text{Cov}_{ab} + \frac{j}{2} \langle [\mathcal{A}_0, \mathcal{B}_0]/j \rangle. \end{aligned} \quad (2.78)$$

Note that  $\text{Cov}_{ab}$  and  $\frac{1}{2} \langle [\mathcal{A}_0, \mathcal{B}_0]/j \rangle$  are the real and imaginary parts of  $\langle \mathcal{A}_0 \mathcal{B}_0 \rangle$ . Substituting (2.78) into (2.76) yields

$$\begin{aligned} \sigma_a^2 \sigma_b^2 &\geq \left| \text{Cov}_{ab} + \frac{j}{2} \langle [\mathcal{A}_0, \mathcal{B}_0]/j \rangle \right|^2 \\ &= \text{Cov}_{ab}^2 + \frac{1}{4} |\langle [\mathcal{A}_0, \mathcal{B}_0] \rangle|^2 \\ (\text{prop. (2.73)}) &= \text{Cov}_{ab}^2 + \frac{1}{4} |\langle [\mathcal{A}, \mathcal{B}] \rangle|^2. \end{aligned} \quad (2.79)$$

Equation (2.79) is equivalent to

$$\sigma_a \sigma_b \geq \frac{1}{2} \sqrt{4 \text{Cov}_{ab}^2 + |\langle [\mathcal{A}, \mathcal{B}] \rangle|^2}.$$

which is a more general result of the uncertainty principle. Since  $\text{Cov}_{ab}^2$  is non negative, it can be dropped to obtain the more standard uncertainty principle in (2.72).  $\square$



The uncertainty principle holds in particular for frequency and scale operators. Before applying the theorem in this particular case, the time  $\mathcal{T}$ , frequency  $\mathcal{F}$  and scale  $\mathcal{C}$  operators are here defined.

Assume that  $s(t)$  is an arbitrary function defined in time domain and  $S(f)$  is its Fourier transform. According to definition (2.58), the mean time  $\langle t \rangle$  is

$$\langle t \rangle = \int t |s(t)|^2 dt = \int s^*(t) t s(t) dt. \quad (2.80)$$

Comparing (2.80) with (2.60), the time operator  $\mathcal{T}$ , expressed in time domain, is simply

$$\mathcal{T} = t. \quad (2.81)$$

Starting from (2.80) and expressing  $s^*(t)$  as inverse Fourier transform yields

$$\begin{aligned} \langle t \rangle &= \int \int S^*(f) e^{-j2\pi ft} df t s(t) dt \\ &= \int S^*(f) \int t s(t) e^{-j2\pi ft} dt df \\ \text{(derivative property)} &= \int S^*(f) \frac{1}{-j2\pi} \frac{dS(f)}{df} df \\ &= \int S^*(f) \frac{j}{2\pi} \frac{d}{df} S(f) df \end{aligned} \quad (2.82)$$

$$= \int S^*(f) \mathcal{T} S(f) df. \quad (2.83)$$

Comparing (2.82) and (2.83) with (2.60), the time operator  $\mathcal{T}$  in frequency domain, is given by

$$\mathcal{T} = \frac{j}{2\pi} \frac{d}{df}. \quad (2.84)$$

In the same way, the frequency operator  $\mathcal{F}$  in frequency domain is

$$\mathcal{F} = f \quad (2.85)$$

as, according to the frequency mean  $\langle f \rangle$  in Fourier domain

$$\langle f \rangle = \int f |S(f)|^2 df = \int S^*(f) f S(f) df. \quad (2.86)$$

Starting from (2.86) and expressing  $S^*(f)$  as Fourier transform yields

$$\begin{aligned} \langle f \rangle &= \int \int s^*(t) e^{+j2\pi ft} dt f S(f) df \\ &= \int s^*(t) \int f S(f) e^{+j2\pi ft} df dt \\ &= \int s^*(t) \frac{1}{j2\pi} \int (j2\pi f) S(f) e^{+j2\pi ft} df dt \\ \text{(derivative property)} &= \int s^*(t) \frac{1}{j2\pi} \frac{d}{dt} s(t) dt \end{aligned} \quad (2.87)$$

$$= \int s^*(t) \mathcal{F} s(t) dt. \quad (2.88)$$

From (2.87) and (2.88), the frequency operator  $\mathcal{F}$  in time domain is

$$\mathcal{F} = \frac{1}{j2\pi} \frac{d}{dt} = -\frac{j}{2\pi} \frac{d}{dt}. \quad (2.89)$$

Summarizing, the time  $\mathcal{T}$  and frequency  $\mathcal{F}$  operators in both time and frequency domains are

$$\mathcal{T} = t ; \mathcal{F} = -\frac{j}{2\pi} \frac{d}{dt} \quad (\text{time domain}) \quad (2.90)$$

$$\mathcal{T} = \frac{j}{2\pi} \frac{d}{df} ; \mathcal{F} = f \quad (\text{frequency domain}) \quad (2.91)$$

Note that the time and frequency operators are Hermitian. To prove this, consider frequency operator in time domain  $\mathcal{F} = \frac{1}{j2\pi} \frac{d}{dt}$  and substitute it into (2.63). For any two functions  $f(t)$  and  $g(t)$ , with finite energy, left hand side of (2.63) becomes

$$\begin{aligned} \int f^*(t) \mathcal{F} g(t) dt &= \int f^*(t) \frac{1}{j2\pi} \frac{d}{dt} g(t) dt \\ (\text{integration by parts}) &= \frac{1}{j2\pi} f^*(t) g(t) \Big|_{+\infty}^{-\infty} - \frac{1}{j2\pi} \int g(t) \frac{d}{dt} f^*(t) dt \\ (\text{finite energy assumption}) &= \int g(t) \left( \frac{1}{j2\pi} \frac{d}{dt} f(t) \right)^* dt \\ &= \int g(t) (\mathcal{F} f(t))^* dt \end{aligned} \quad (2.92)$$

which is the condition (2.63) for operator  $\mathcal{F}$ . An equal proof can be done for time operator  $\mathcal{T}$  in frequency domain.

The scale operator  $\mathcal{C}$  is defined as

$$\mathcal{C} \triangleq \frac{1}{2} [\mathcal{T}, \mathcal{F}]_+ = \frac{1}{2} (\mathcal{T}\mathcal{F} + \mathcal{F}\mathcal{T}) \quad (2.93)$$

and its representations in both time and frequency domains are

$$\mathcal{C} = \frac{1}{4\pi j} \left( t \frac{d}{dt} + \frac{d}{dt} t \right) \quad (\text{time domain}) \quad (2.94)$$

$$\mathcal{C} = \frac{j}{4\pi} \left( f \frac{d}{df} + \frac{d}{df} f \right) \quad (\text{frequency domain}) \quad (2.95)$$

The operator  $\mathcal{C}$  is also Hermitian because the anticommutator of two Hermitian operators is Hermitian (property (2.70)).

The uncertainty principle can be applied for frequency and scale operators, as they do not commute. In fact, the commutator of  $\mathcal{F}$  and  $\mathcal{C}$  is given by

$$[\mathcal{F}, \mathcal{C}] = \mathcal{F}\mathcal{C} - \mathcal{C}\mathcal{F} = \frac{1}{2} (\mathcal{F}\mathcal{T}\mathcal{F} + \mathcal{F}\mathcal{F}\mathcal{T} - \mathcal{T}\mathcal{F}\mathcal{F} - \mathcal{F}\mathcal{T}\mathcal{F}) = \frac{1}{2} (\mathcal{F}\mathcal{F}\mathcal{T} - \mathcal{T}\mathcal{F}\mathcal{F}). \quad (2.96)$$

To evaluate  $[\mathcal{F}, \mathcal{C}]$ , the operator in the right hand side of (2.96) is applied on an arbitrary function  $s(t)$ , using the definitions of  $\mathcal{T}$  and  $\mathcal{F}$  in time domain (2.90).

$$\begin{aligned}
 [\mathcal{F}, \mathcal{C}]s(t) &= \frac{1}{2}(\mathcal{F}\mathcal{F}\mathcal{T} - \mathcal{T}\mathcal{F}\mathcal{F})s(t) \\
 &= \frac{1}{2} \left[ -\frac{j}{2\pi} \frac{d}{dt} \left( -\frac{j}{2\pi} \frac{d}{dt} (s(t)t) \right) - t \left( -\frac{j}{2\pi} \frac{d}{dt} \left( -\frac{j}{2\pi} \frac{ds(t)}{dt} \right) \right) \right] \\
 &= \frac{1}{2} \left[ \left( -\frac{j}{2\pi} \right)^2 \frac{d}{dt} \left( t \frac{ds(t)}{dt} + s(t) \right) - t \left( -\frac{j}{2\pi} \right)^2 \frac{d^2s(t)}{dt^2} \right] \\
 &= \frac{1}{2} \left( -\frac{j}{2\pi} \right)^2 \left[ t \frac{d^2s(t)}{dt^2} + \frac{ds(t)}{dt} + \frac{ds(t)}{dt} - t \frac{d^2s(t)}{dt^2} \right] \\
 &= \left( -\frac{j}{2\pi} \right)^2 \frac{ds(t)}{dt} \\
 &= -\frac{j}{2\pi} \mathcal{F}s(t) \tag{2.97}
 \end{aligned}$$

Substituting the operator  $[\mathcal{F}, \mathcal{C}] = -\frac{j}{2\pi} \mathcal{F}$  into (2.72), the uncertainty principle for frequency and scale operators is

$$\sigma_f \sigma_c \geq \frac{1}{2} |\langle [\mathcal{F}, \mathcal{C}] \rangle| = \frac{1}{4\pi} |\langle \mathcal{F} \rangle|. \tag{2.98}$$

Condition (2.98) implies some restrictions on the signals involved in the discrete time-scale model of sections 2.3 and 2.4. In particular, for an input signal  $x(t)$ , with given variances  $\sigma_f^2$  and  $\sigma_c^2$ , in frequency and scale domains respectively, the mean frequency of  $x(t)$  is upper limited by  $4\pi\sigma_f\sigma_c$ .



# Identification of Narrowband Time-Varying Sparse Systems

In this chapter we study and compare two approaches for the identification of a sparse narrowband time-varying system.

In Section 3.1 we derive a discrete characterization of the system and describe the corresponding model for the discrete spreading function to be identified.

In Section 3.2 we formulate the recovery according to two different approaches, reducing the identification problem both to a MMV problem and to a block-sparse problem, providing also the matrix representation in both cases.

The algorithms used for recovery are described in detail in Section 3.3, while in Section 3.4 we provide numerical results to compare the performances of the two approaches.

In the following, we refer to the two approaches simply as *approach 1* and *approach 2*, respectively.

## 3.1. Discrete Time-Frequency Characterization

The output  $y(t)$  of a narrowband time-varying system can be represented as a weighted superposition of time-frequency shifted versions of the input signal  $x(t)$

$$y(t) = \int_{\tau} \int_{\nu} s_H(\tau, \nu) x(t - \tau) e^{j2\pi\nu t} d\nu d\tau \quad (3.1)$$

where  $s_H(\tau, \nu)$  is the spreading function of the operator  $H$ , which describes the system. Indicating with  $T_{\tau}$  and  $M_{\nu}$  the time- and frequency-shift operators on the signal  $x(t)$ ,

$$(T_{\tau}x)(t) \triangleq x(t - \tau) \quad (3.2)$$

$$(M_{\nu}x)(t) \triangleq e^{j2\pi\nu t} x(t) \quad (3.3)$$

the operator  $H$  is the continuous weighted superposition of time-frequency shift operators, i.e., the input-output relation can be expressed as

$$y(t) = (Hx)(t) = \int_{\tau} \int_{\nu} s_H(\tau, \nu) (M_{\nu}T_{\tau}x)(t) d\nu d\tau. \quad (3.4)$$

Due to physical limitations of the system (see e.g. [17]), the support of the spreading function is assumed to be finite, i.e.,  $s_H(\tau, \nu) \equiv 0$  for  $(\tau, \nu) \notin [0, \tau_{max}) \times [0, \nu_{max})$ . In order to study the discrete characterization of (3.4), we base on the model presented in [15], where the region  $[0, \tau_{max}) \times [0, \nu_{max})$  is divided in rectangular cells, that are either *active*, if  $s_H$  is nonzero on these cells, or null. More precisely, suppose to choose  $L \in \mathbb{R}$  and  $T \in \mathbb{R}$  such that the area  $[0, \tau_{max}) \times [0, \nu_{max})$  of the  $(\tau, \nu)$ -plane can be divided in  $L^2$  rectangular cells of the same area  $1/L$ , partitioning the  $\tau$ -axis in  $L$  parts of length  $T$  and the  $\nu$ -axis in  $L$  parts of length  $1/(TL)$ , as shown in Fig. 3.1.

Consequently,

$$\tau_{max} = TL, \nu_{max} = \frac{1}{T} \quad (3.5)$$

and

$$\tau_{max}\nu_{max} = L. \quad (3.6)$$

Then the arbitrary, possibly fragmented, support of the spreading function of operator  $H$ , indicated as  $M_\Gamma$ , can be expressed as the union of a particular subset of these cells

$$M_\Gamma \triangleq \bigcup_{(k,m) \in \Gamma} \left( U + \left( kT, \frac{m}{TL} \right) \right) \quad (3.7)$$

where  $U \triangleq [0, T) \times [0, 1/(TL))$  is the fundamental cell in the  $(\tau, \nu)$ -plane and, consequently,  $U + (kT, m/(TL))$  is the cell with the below-left corner at  $(kT, m/(TL))$ .  $\Gamma$  specifies the *active cells* that identify the support of the spreading function,

$$\Gamma \subseteq \Sigma \triangleq \{(0, 0), (0, 1), \dots, (L-1, L-1)\} \quad (3.8)$$

and  $|\Gamma|$  is the number of active cells.

Starting from the continuous input-output relation (3.1), a discrete characterization of the system can be derived, using the following assumptions.

1. The probing signal  $x(t)$  is bandlimited to  $[-B, B]$ .
2. The signal  $y(t)$  is observed for the finite time interval  $[-V, V]$ . The truncated version of the system response  $y(t)$  on  $[-V, V]$  is indicated as  $\bar{y}(t)$ .

We report here the basic steps in [15] in order to have an intuition of the discrete model we use in the following analysis, and refer to [15] for more details.

We define  $\bar{s}_H(\tau, \nu)$  as the effective spreading function of the system, obtained from  $s_H(\tau, \nu)$  after considering the assumptions 1 and 2 above. It can be shown that  $\bar{s}_H(\tau, \nu)$  is a smoothed version of  $s_H(\tau, \nu)$ , both in time and frequency domain and is given as

$$\bar{s}_H(\tau, \nu) = 4BV s_H(\tau, \nu) * \text{sinc}(2B\tau) * \text{sinc}(2V\nu). \quad (3.9)$$

From (3.9),  $\bar{s}_H(\tau, \nu)$  is not supported on  $[0, \tau_{max}) \times [0, \nu_{max})$  but, considering the approximation that the sinc functions are nonzero only in their mainlobes, most of the volume of  $\bar{s}_H(\tau, \nu)$  is supported on  $[-1/(2B), \tau_{max} + 1/(2B)) \times [-1/(2V), \nu_{max} + 1/(2V))$ . It can be seen that the continuous output signal  $\bar{y}(t)$  is approximately

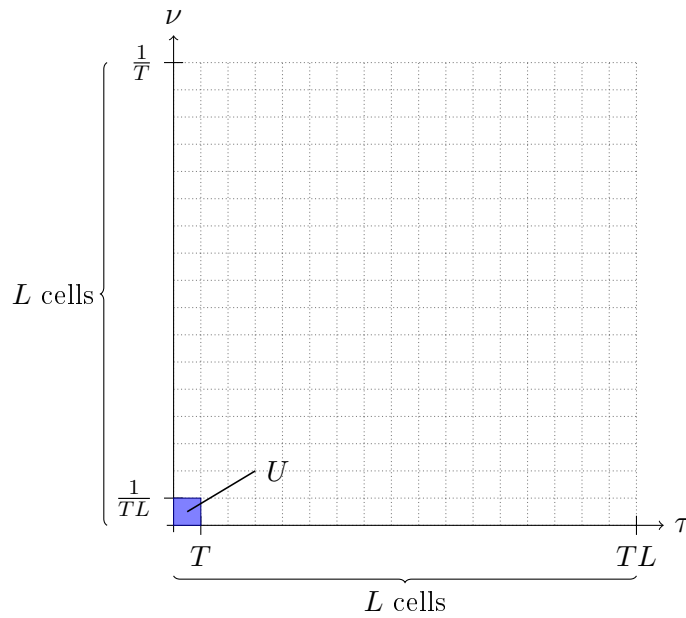


Figure 3.1.: Model for the support of the spreading function.

bandlimited within  $[-B, B + \nu_{max}]$ , hence it can be sampled at rate  $f_s = 2B + \nu_{max}$ , yielding the discrete output  $\bar{y}[n]$ . We define then function  $\bar{s}_H[m, l]$  as a sampled version of  $\bar{s}_H(\tau, \nu)$ , given by

$$\bar{s}_H[m, l] \triangleq \bar{s}_H\left(\frac{m}{f_s}, \frac{l}{2V}\right) e^{-j\pi l}. \quad (3.10)$$

Relation (3.10) suggests the model for the support of the discrete spreading function of the system, where the  $(\tau, \nu)$ -plane is discretized both in  $\tau$ -direction, with resolution of  $1/f_s$ , and in  $\nu$ -direction, with resolution  $1/(2V)$ , i.e., the spreading function is uniformly sampled in the  $(\tau, \nu)$ -plane.

Finally, the integer parameters  $E$  and  $D$  are set such that

$$\tau_{max} = \frac{EL}{f_s}, \quad \nu_{max} = \frac{DL}{2V}. \quad (3.11)$$

According to definition (3.11),  $E$  and  $D$  are the number of samples taken on an active cell of  $\bar{s}_H(\tau, \nu)$  in  $\tau$  and  $\nu$  direction, respectively, as indicated in Fig. 3.2.

The continuous input-output relation in (3.1) is then equivalent to the following discrete characterization of the system

$$\bar{y}[n] = \sum_{m \in \mathbb{Z}} \sum_{l \in \mathbb{Z}} \bar{s}_H[m, l] x[n - m] e^{j2\pi \frac{ln}{DEL}}, \quad n = 0, \dots, DEL - 1 \quad (3.12)$$

where the discrete signal  $x[n]$  is a sampled version of  $x(t)$ .

From the approximation  $\bar{s}_H(\tau, \nu) \equiv 0$  for  $(\tau, \nu) \notin [-1/(2B), \tau_{max} + 1/(2B)] \times [-1/(2V), \nu_{max} + 1/(2V)]$  and from the definitions of  $\tau_{max}$  and  $\nu_{max}$  in (3.11), the discrete spreading function  $\bar{s}_H[m, l]$  is nonzero only for  $(m, l) \in \{0, 1, \dots, EL - 1\} \times$

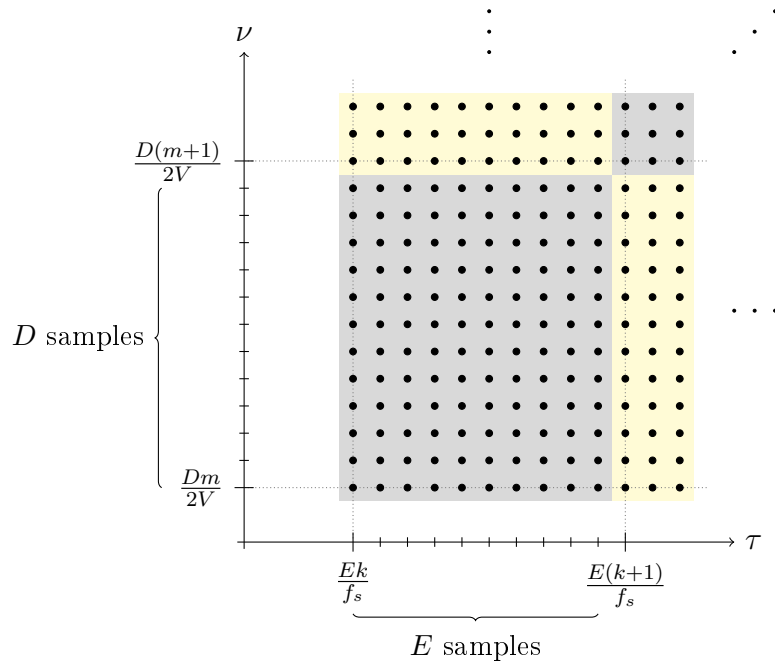


Figure 3.2.: Model for the support of the discrete spreading function.

$\{0, 1, \dots, DL - 1\}$  and, consequently, the two summations in (3.12) involve only finitely many terms.

$$\bar{y}[n] = \sum_{m=0}^{EL-1} \sum_{l=0}^{DL-1} \bar{s}_H[m, l] x[n - m] e^{j2\pi \frac{ln}{DEL}}, \quad n = 0, \dots, DEL - 1. \quad (3.13)$$

Furthermore, if most of the volume of  $\bar{s}_H(\tau, \nu)$  is approximately supported on  $M_\Gamma$ , given in (3.7), then the samples of the spreading function in (3.13) satisfy

$$\bar{s}_H[m, l] \equiv 0 \text{ for } \left( \frac{m}{f_s}, \frac{l}{2V} \right) \notin M_\Gamma. \quad (3.14)$$

## 3.2. Matrix Representations

In the following we compare two approaches to identify a system with input-output relation (3.13). The *approach 1* allows to formulate the problem as a multiple measured vectors (MMV) problem, in which equation (3.13) admits the following matrix representation

$$\mathbf{Z} = \mathbf{A}_c \mathbf{S}. \quad (3.15)$$

The matrices  $\mathbf{Z} \in \mathbb{C}^{L \times DE}$ ,  $\mathbf{A}_c \in \mathbb{C}^{L \times L^2}$  and  $\mathbf{S} \in \mathbb{C}^{L^2 \times DE}$  depend on  $\bar{y}[n]$ ,  $x[n]$  and  $\bar{s}_H[m, l]$ , respectively. In particular,  $\mathbf{Z}$  contains the discrete Zak transform of  $y[n]$ ,  $\mathbf{A}_c$  translates in time and frequency the input samples  $x[n]$  and each row of  $\mathbf{S}$  contains the samples  $\bar{s}_H[m, l]$  inside each active cell. The precise definitions are presented later. A graphical representation of (3.15) is presented in Fig. 3.3. This



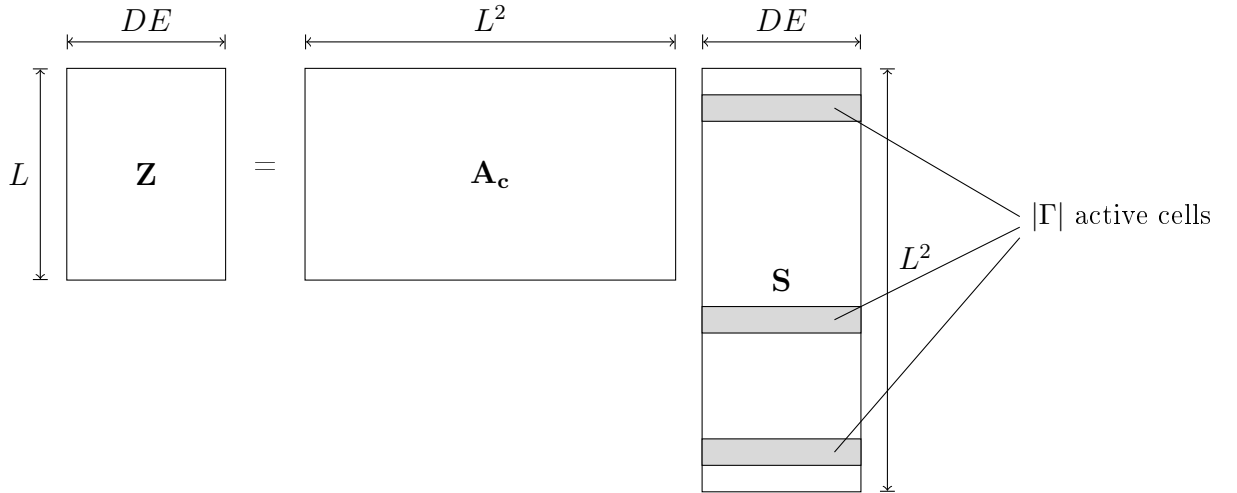


Figure 3.3.: Model for the MMV problem.

problem involves a particular probing signal  $x[m]$ , given as a train of Dirac impulses

$$x[m] = \begin{cases} c_{-k} & \text{if } m = Ek \\ 0 & \text{otherwise} \end{cases} \quad (3.16)$$

with the additional property that

$$c_k = c_{k+L}. \quad (3.17)$$

The coefficients  $c_k$ ,  $k = 0, \dots, L - 1$  are independent and uniformly at random, chosen on the complex unit disc.

The *approach 2* leads to a matrix representation as a Block-Sparsity model, given by

$$\mathbf{y} = \mathbf{X}\mathbf{s}. \quad (3.18)$$

The vectors  $\mathbf{y} \in \mathbb{C}^{DEL \times 1}$  and  $\mathbf{s} \in \mathbb{C}^{DEL^2 \times 1}$  contain samples of the output  $\bar{y}[n]$  and of the spreading function  $\bar{s}_H[m, l]$ , respectively, while matrix  $\mathbf{X} \in \mathbb{C}^{DEL \times DEL^2}$  depends on the probing signal  $x[m]$ , which is a sequence of independent symbols with distribution  $\mathcal{CN}(0, 1)$ . A graphical representation of (3.18) is presented in Fig. 3.4.

The two models are detailed below.

### 1. MMV model.

To derive the matrix representation (3.15) from the discrete input-output relation (3.13), we calculate first the discrete Zak transform of  $y[n]$ , defined as

$$\mathcal{Z}_y^{EL, D}[n, r] \triangleq \frac{1}{D} \sum_{q=0}^{D-1} y[n + ELq] e^{-j2\pi \frac{qr}{D}}, \quad n = 0, \dots, EL - 1, \quad r = 0, \dots, D - 1. \quad (3.19)$$

Substituting  $n$  in (3.19) with  $n = n' + Ep$ , where  $n' = 0, \dots, E - 1$  and  $p = 0, \dots, L - 1$  yields

$$\begin{aligned}
 \mathcal{Z}_y^{EL,D}[n' + Ep, r] &= \frac{1}{D} \sum_{q=0}^{D-1} y[n' + Ep + ELq] e^{-j2\pi \frac{qr}{D}} \\
 \text{(from (3.13))} &= \frac{1}{D} \sum_{q=0}^{D-1} \sum_{m=0}^{EL-1} \sum_{l=0}^{DL-1} \overline{s}_H[m, l] x[n' + Ep + ELq - m] \\
 &\quad \times e^{j2\pi \frac{l(n'+Ep+ELq)}{DEL}} e^{-j2\pi \frac{qr}{D}} \\
 (m = m' + Ek) &= \frac{1}{D} \sum_{q=0}^{D-1} \sum_{m'=0}^{E-1} \sum_{k=0}^{L-1} \sum_{l=0}^{DL-1} \overline{s}_H[m' + Ek, l] \\
 &\quad \times x[n' - m' + E(p + Lq - k)] e^{j2\pi \frac{l(n'+Ep+ELq)}{DEL}} e^{-j2\pi \frac{qr}{D}}
 \end{aligned}$$

From the definition (3.16) of the probing signal and using (3.17)

$$x[n' - m' + E(p + Lq - k)] = \begin{cases} c_{k-p} & \text{if } m' = n' \\ 0 & \text{if } m' \neq n' \end{cases} \quad (3.20)$$

Substituting (3.20) into the equation above yields

$$\begin{aligned}
 \mathcal{Z}_y^{EL,D}[n' + Ep, r] &= \frac{1}{D} \sum_{q=0}^{D-1} \sum_{k=0}^{L-1} \sum_{l=0}^{DL-1} c_{k-p} \overline{s}_H[n' + Ek, l] e^{j2\pi \frac{l(n'+Ep+ELq)}{DEL}} e^{-j2\pi \frac{qr}{D}} \\
 (l = l' + Dh) &= \frac{1}{D} \sum_{q=0}^{D-1} \sum_{k=0}^{L-1} \sum_{l'=0}^{D-1} \sum_{h=0}^{L-1} c_{k-p} \overline{s}_H[n' + Ek, l' + Dh] \\
 &\quad \times e^{j2\pi \frac{(l'+Dh)(n'+Ep+ELq)}{DEL}} e^{-j2\pi \frac{qr}{D}} \\
 &= \frac{1}{D} \sum_{q=0}^{D-1} \sum_{k=0}^{L-1} \sum_{l'=0}^{D-1} \sum_{h=0}^{L-1} c_{k-p} \overline{s}_H[n' + Ek, l' + Dh] \\
 &\quad \times e^{j2\pi \frac{(l'+Dh)(n'+Ep)}{DEL}} e^{j2\pi \frac{ql'}{D}} e^{-j2\pi \frac{qr}{D}} \\
 &= \frac{1}{D} \sum_{q=0}^{D-1} \sum_{k=0}^{L-1} \sum_{l'=0}^{D-1} \sum_{h=0}^{L-1} c_{k-p} \overline{s}_H[n' + Ek, l' + Dh] \\
 &\quad \times e^{j2\pi \frac{(l'+Dh)(n'+Ep)}{DEL}} e^{j2\pi \frac{q(l'-r)}{D}} \\
 &= \sum_{k=0}^{L-1} \sum_{h=0}^{L-1} c_{k-p} \overline{s}_H[n' + Ek, r + Dh] e^{j2\pi \frac{(r+Dh)(n'+Ep)}{DEL}} \quad (3.21)
 \end{aligned}$$

where the last equality is due to the DFT. Define  $z_p[n, r]$  as

$$z_p[n, r] \triangleq \mathcal{Z}_y^{EL,D}[n + Ep, r] = \sum_{k=0}^{L-1} \sum_{h=0}^{L-1} c_{k-p} \overline{s}_H[n + Ek, r + Dh] e^{j2\pi \frac{(r+Dh)(n+Ep)}{DEL}} \quad (3.22)$$

where  $n = 0, \dots, E-1$ ,  $r = 0, \dots, D-1$  and  $p = 0, \dots, L-1$ .

The  $p$ -th entry of the  $(n, r)$ th column  $\mathbf{z}[n, r]$  of  $\mathbf{Z}$  in (3.15) is defined as

$$[\mathbf{z}[n, r]]_p \triangleq z_p[n, r] e^{-j2\pi \frac{rp}{DL}}, \quad p = 0, \dots, L-1. \quad (3.23)$$

Parameters  $n$  and  $r$  identify the columns of  $\mathbf{Z}$ , which are ordered as follows

$$(n, r) \in \{(0, 0), (0, 1), \dots, (E - 1, D - 1)\}. \quad (3.24)$$

The column of matrix  $\mathbf{S}$  in (3.15), identified by parameters  $n$  and  $r$  as in (3.24), is defined as

$$\mathbf{s}[n, r] \triangleq [s_{0,0}[n, r], s_{0,1}[n, r], \dots, s_{L-1,L-1}[n, r]]^T \quad (3.25)$$

where

$$s_{k,m}[n, r] \triangleq \overline{s_H}[n + Ek, r + Dm] e^{j2\pi \frac{n(r+Dm)}{DEL}}, \quad k, m = 0, \dots, L - 1. \quad (3.26)$$

From definition (3.26), each row of  $\mathbf{S}$  contains all the samples of  $\overline{s_H}[m, l]$  inside the same active cell (see Fig. 3.2), with the exception of a phase shift. As we assume that there are only  $|\Gamma|$  active cells, then  $\mathbf{S}$  has  $|\Gamma|$  nonzero rows.

Finally, matrix  $\mathbf{A}_c$  in (3.15) is given by

$$\mathbf{A}_c \triangleq [\mathbf{A}_{c,0} | \mathbf{A}_{c,1} | \dots | \mathbf{A}_{c,L-1}] \quad (3.27)$$

with the square  $L \times L$  submatrices defined as

$$\mathbf{A}_{c,k} \triangleq \mathbf{C}_{c,k} \mathbf{F}^H \quad (3.28)$$

where  $\mathbf{C}_{c,k} \triangleq \text{diag}\{c_k, c_{k-1}, \dots, c_{k-(L-1)}\}$  and  $[\mathbf{F}]_{p,m} \triangleq e^{-j2\pi \frac{pm}{L}}$ , with  $p, m = 0, \dots, L - 1$ .

With the definitions above, (3.15) describes exactly the discrete input-output relation (3.13). Denote the matrix obtained from  $\mathbf{S}$  by selecting only the rows corresponding to the active cells of  $\overline{s_H}[m, l]$  with  $\mathbf{S}_\Gamma$  and let  $\mathbf{A}_\Gamma$  be the matrix containing the columns of  $\mathbf{A}_c$  corresponding to the same cells. Then (3.15) is equivalent to

$$\mathbf{Z} = \mathbf{A}_\Gamma \mathbf{S}_\Gamma. \quad (3.29)$$

## 2. Block-Sparsity model.

The matrix representation of this problem follows directly from the input-output relation (3.13). With a change of variables  $m = m' + Ek$  and  $l = l' + Dr$ , (3.13) becomes

$$\begin{aligned} \bar{y}[n] = \sum_{m=0}^{E-1} \sum_{k=0}^{L-1} \sum_{l=0}^{D-1} \sum_{r=0}^{L-1} \overline{s_H}[m + Ek, l + Dr] x[n - (m + Ek)] e^{j2\pi \frac{(l+Dr)n}{DEL}}, \\ n = 0, \dots, DEL - 1. \end{aligned} \quad (3.30)$$

The column vector  $\mathbf{y}$  in (3.18) contains the  $DEL$  samples of the output  $\bar{y}[n]$ ,  $n = 0, \dots, DEL - 1$ . Define the column vector  $\mathbf{s}$  in (3.18) as

$$\mathbf{s} \triangleq \left[ \mathbf{s}_H^{0,0T} \mid \mathbf{s}_H^{0,1T} \mid \dots \mid \mathbf{s}_H^{L-1,L-1T} \right]^T \quad (3.31)$$

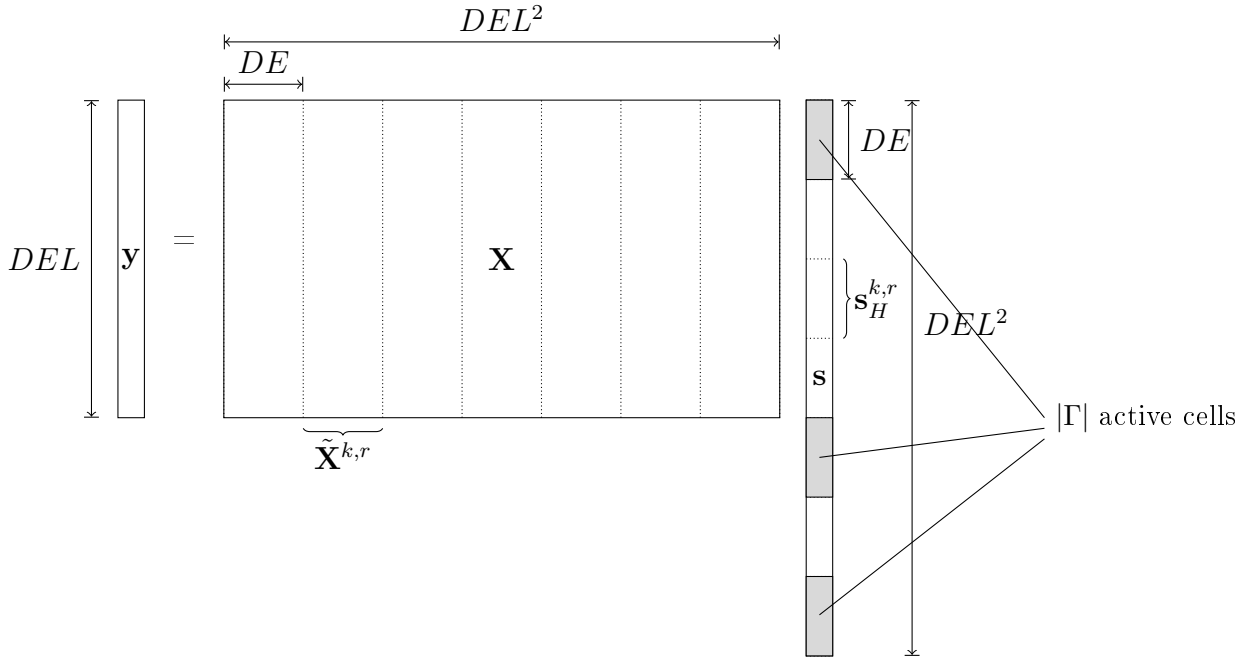


Figure 3.4.: Model for the B-sparsity problem.

where each subvector  $\mathbf{s}_H^{k,r} \in \mathbb{C}^{DE \times 1}$ ,  $k, r = 0, \dots, L-1$ , contains all the samples of the spreading function  $\overline{s}_H[m, l]$  on the same cell of Fig. 3.2,

$$\mathbf{s}_H^{k,r} \triangleq [s_H^{k,r}(0, 0), s_H^{k,r}(0, 1), \dots, s_H^{k,r}(E-1, D-1)]^T \quad (3.32)$$

with  $s_H^{k,r}(m, l) \triangleq \overline{s}_H[m + Ek, l + Dr]$ . Assuming that the number of active cells is  $|\Gamma|$ , only  $|\Gamma|$  blocks are nonzero, so that  $\mathbf{s}$  is block-sparse.

Finally, matrix  $\mathbf{X}$  is given by

$$\mathbf{X} = [\tilde{\mathbf{X}}^{0,0} | \tilde{\mathbf{X}}^{0,1} | \dots | \tilde{\mathbf{X}}^{L-1,L-1}] \quad (3.33)$$

where the submatrices  $\tilde{\mathbf{X}}^{k,r} \in \mathbb{C}^{DEL \times DE}$ ,  $k, r = 0, \dots, L-1$ , are defined as

$$\tilde{\mathbf{X}}^{k,r} \triangleq [\tilde{\mathbf{x}}^{k,r}(0, 0) | \tilde{\mathbf{x}}^{k,r}(0, 1) | \dots | \tilde{\mathbf{x}}^{k,r}(E-1, D-1)] \quad (3.34)$$

with the  $n$ -th entry of vector  $\tilde{\mathbf{x}}^{k,r}(m, l) \in \mathbb{C}^{DEL \times 1}$ ,  $m = 0, \dots, E-1$ ,  $l = 0, \dots, D-1$  given by

$$[\tilde{\mathbf{x}}^{k,r}(m, l)]_n \triangleq x[n - (m + Ek)] e^{j2\pi \frac{(l+Dr)n}{DEL}}, \quad n = 0, \dots, DEL-1. \quad (3.35)$$

Denote the vector obtained from  $\mathbf{s}$  by selecting only the blocks corresponding to the active cells of  $\overline{s}_H[m, l]$  by  $\mathbf{s}_\Gamma$  and let  $\mathbf{X}_\Gamma$  be the matrix containing the columns of  $\mathbf{X}$  corresponding to those cells. Then (3.18) is equivalent to

$$\mathbf{y} = \mathbf{X}_\Gamma \mathbf{s}_\Gamma. \quad (3.36)$$

### 3.3. Recovery Algorithms

The goal of the two algorithms described in this section is to recover both the support  $\Gamma$  and the samples  $\overline{s}_H[m, l]$  of the spreading function in (3.13), given the input signal  $x[n]$  in both cases. The algorithms are adaptations of the orthogonal matching pursuit (OMP) algorithm [18]. Before starting the detailed description of the two algorithms, some practical observations are presented. It can be shown that the identification of a system of the form (3.13) fails if the value of sparsity satisfies  $|\Gamma| \geq L$ , since in this case there are more unknowns than knowns. The simulations performed in section 3.4 are applied under the necessary condition  $|\Gamma| \leq L$ . The recovery depends on  $\mathbf{A}_c$  and  $\mathbf{X}$ . In particular, the more the samples of the spreading function are well approximated by the same set of elements in  $\mathbf{A}_c$  and  $\mathbf{X}$ , the more successful the recovery is. In the case of high recovery probability, matrices  $\mathbf{A}_c$  and  $\mathbf{X}$  are well-conditioned.

#### 1. Simultaneous Orthogonal Matching Pursuit (S-OMP).

Recovery of  $\mathbf{S}$  from  $\mathbf{Z}$  obtained as  $\mathbf{Z} = \mathbf{A}_c \mathbf{S}$ , when we assume that the columns of  $\mathbf{S}$  share the same sparsity pattern, is a MMV problem. The S-OMP algorithm is one approach to solve it, i.e., to reconstruct  $\mathbf{S}$  from  $\mathbf{Z}$ , obtained as  $\mathbf{Z} = \mathbf{A}_c \mathbf{S}$ . The S-OMP algorithm deals with the approximation of several signals at once (column vectors  $\mathbf{s}[n, r]$  of matrix  $\mathbf{S}$ ) using different linear combinations of the same elementary signals (columns of matrix  $\mathbf{A}_c$ ) [19].

#### Algorithm 1. S-OMP

*Input:*

- Matrix  $\mathbf{Z}$  of the observed samples.
- Matrix  $\mathbf{A}_c$  (dictionary).
- Value of sparsity  $|\Gamma|$  of matrix  $\mathbf{S}$  (number of nonzero rows of  $\mathbf{S}$ ).

*Output:*

- Matrix  $\mathbf{S}$  of the samples of the spreading function.

*Procedure:*

- a) Initialize the residual matrix  $\mathbf{R}_0 = \mathbf{Z}$  and the index set  $\Lambda_0 = \emptyset$ . The iteration counter is  $t = 1$ .
- b) Find an index  $\lambda_t$  that solves the optimization problem

$$\lambda_t = \arg \max_{i=1, \dots, L^2} \|\mathbf{R}_{t-1}^H(\mathbf{A}_c \mathbf{e}_i)\|_2 \quad (3.37)$$

where  $\mathbf{e}_i$  is the  $i$ -th canonical basis column vector in  $\mathbb{C}^{L^2}$  and, consequently,  $\mathbf{A}_c \mathbf{e}_i$  is the  $i$ -th column of  $\mathbf{A}_c$ .

- c) Set  $\Lambda_t = \Lambda_{t-1} \cup \lambda_t$ .

d) Determine matrix  $\tilde{\mathbf{A}}$ , selecting from  $\mathbf{A}_c$  the columns with indexes in  $\Lambda_t$ .

e) Calculate the new approximation of  $\mathbf{S}$  and the new residual matrix

$$\tilde{\mathbf{S}}_t = \tilde{\mathbf{A}}^\dagger \mathbf{Z} \quad (3.38)$$

$$\mathbf{R}_t = \mathbf{Z} - \tilde{\mathbf{S}}_t \quad (3.39)$$

f) Increment  $t$ . If  $t = |\Gamma|$ , stop; otherwise return to step (b).

g) Set  $\mathbf{S}_\Gamma = \tilde{\mathbf{S}}_{|\Gamma|}$ . Matrix  $\mathbf{S}$  is obtained from  $\mathbf{S}_\Gamma$  adding zero rows, corresponding to indexes in  $\{1, \dots, L^2\}$  not included in  $\Lambda_{|\Gamma|}$ .

Since the matrix  $\mathbf{S}$  has a sparsity of  $|\Gamma|$ , the algorithm is performed for  $|\Gamma|$  times, so that  $\mathbf{A}_\Gamma$  in (3.29) contains the  $|\Gamma|$  columns of  $\mathbf{A}_c$  that best approximate the spreading function.

Step (b) is referred to as the *greedy selection* of the algorithm. Maximizing the  $\ell_2$ -norm in (3.37) means finding the element of the dictionary that can contribute a lot of energy to every column of the matrix  $\mathbf{S}$ .

## 2. Block-sparse Orthogonal Matching Pursuit (B-OMP).

Because of the structure of vector  $\mathbf{s}$  in (3.18), where nonzero entries appear in blocks, one approach for identification of the B-sparsity model is the B-OMP algorithm [20]. Since a MMV model is a special case of a block-sparse model, S-OMP algorithm is equivalent to B-OMP algorithm if MMV is formulated as a block-sparse model. That makes the two approaches comparable.

### Algorithm 2. B-OMP

*Input:*

- Vector  $\mathbf{y}$  of the observed samples.
- Matrix  $\mathbf{X}$  (dictionary).
- Value of sparsity  $|\Gamma|$  of vector  $\mathbf{s}$  (number of nonzero blocks of  $\mathbf{s}$ ).

*Output:*

- Vector  $\mathbf{s}$  of the samples of the spreading function.

*Procedure:*

- a) Initialize the residual vector  $\mathbf{r}_0 = \mathbf{y}$  and the block index set  $E_0 = \emptyset$ . The iteration counter is  $t = 1$ .
- b) Find an index  $\lambda_t$  that solves the optimization problem

$$\lambda_t = \arg \max_{i=1, \dots, L^2} \left\| \tilde{\mathbf{X}}^H [i] \mathbf{r}_{t-1} \right\|_2 \quad (3.40)$$

where  $\tilde{\mathbf{X}}[i] \in \mathbb{C}^{DEL \times DE}$  is the  $i$ -th block of  $\mathbf{X}$ .  $\tilde{\mathbf{X}}[i] = \tilde{\mathbf{X}}^{k,r}$  in (3.33) for some proper values of  $k$  and  $r$ .

- c) Set  $E_t = E_{t-1} \cup \lambda_t$ .
- d) Determine matrix  $\tilde{\mathbf{X}}$ , selecting from  $\mathbf{X}$  the blocks with indexes in  $E_t$ .
- e) Calculate the new approximation of  $\mathbf{s}$  and the new residual vector

$$\tilde{\mathbf{s}}_t = \tilde{\mathbf{X}}^\dagger \mathbf{y} \quad (3.41)$$

$$\mathbf{r}_t = \mathbf{y} - \tilde{\mathbf{s}}_t \quad (3.42)$$

- f) Increment  $t$ . If  $t = |\Gamma|$ , stop; otherwise return to step (b).
- g) Set  $\mathbf{s}_\Gamma = \tilde{\mathbf{s}}_{|\Gamma|}$ . Vector  $\mathbf{s}$  is obtained from  $\mathbf{s}_\Gamma$  adding zero blocks, corresponding to indexes in  $\{1, \dots, L^2\}$  not included in  $E_{|\Gamma|}$ .

### 3.4. Simulation Results

The two approaches presented above are compared in the noiseless and noisy cases.

1. *Noiseless case.* The identification of the spreading function from the observed samples is performed without introducing noise. The two models are compared in terms of *recovery probability*. Recovery is regarded as successful if the relative error between the recovered spreading function  $\hat{s}_H$  and the original one  $\overline{s}_H$  is less than a fixed tolerance, admitted because of limited precision in the simulations, i.e.,

$$\frac{\|\overline{s}_H - \hat{s}_H\|_2}{\|\hat{s}_H\|_2} \leq 10^{-5} \quad (3.43)$$

The recovery probability is the average of the successful recoveries upon 1000 trials.

2. *Noisy case.* In this case the observed samples are corrupted by complex additive white Gaussian noise, with a SNR of 20 dB. The two models are compared in terms of the *root mean square error*, *sqrt-MSE*, of the recovered spreading function  $\hat{s}_H$ . The average of the relative error defined in (3.43) is evaluated upon 1000 trials.

Some parameters are set before the simulation is performed:

- Parameter  $L$  is fixed and set to  $L = 19$  and, consequently, the number of cells on the  $(\tau, \nu)$ -plane is  $L^2 = 361$ .
- The cardinality  $|\Gamma|$  of the support set of the spreading function is varied from 1 to 19.
- The product  $DE$  assumes four different values:  $DE = \{1, 7, 13, 19\}$ . In fact, the identification depends on  $DE$ , rather than on  $D$  and  $E$  individually.

- For each pair  $(DE, |\Gamma|)$ , 1000 trials are performed to obtain the recovery probability and the root MSE.

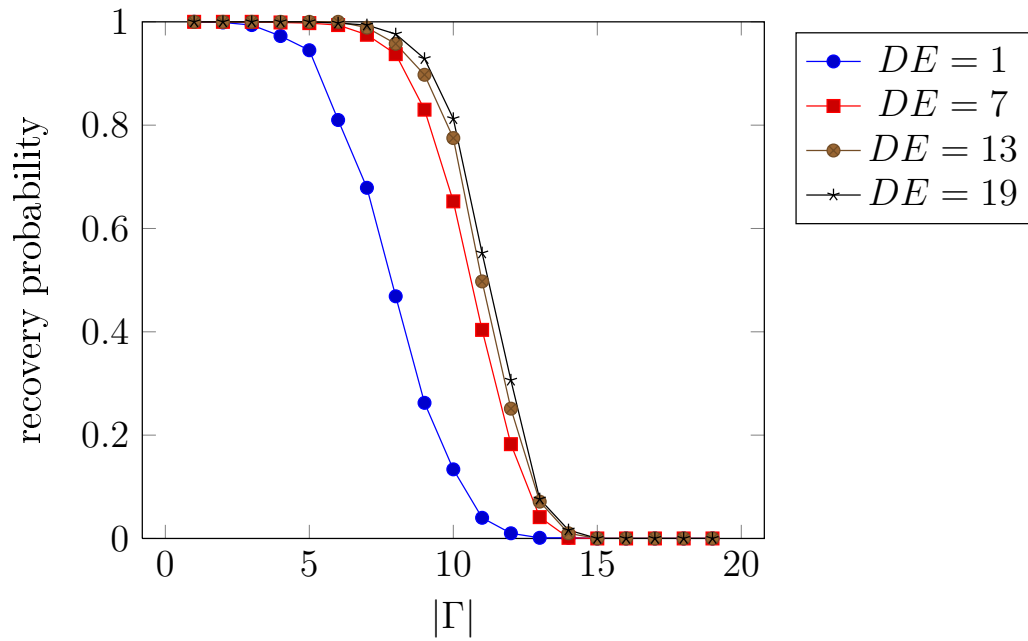
Spreading functions are generated by choosing uniformly at random a support set  $\Gamma$  with cardinality  $|\Gamma|$ , which corresponds to an area of  $\Delta = |\Gamma|/L$  in the continuous setting. Samples  $\overline{s_H}[m, l]$  are then chosen independently with distribution  $\mathcal{CN}(0, 1)$ .

In Fig. 3.5 and 3.6 we report the results, for the noiseless and the noisy case, respectively.

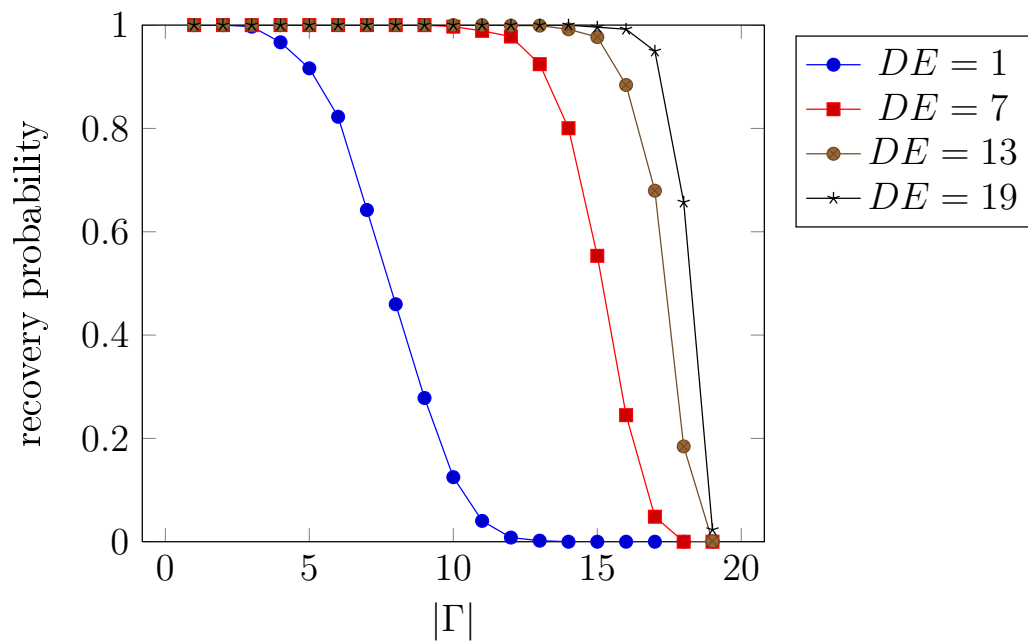
From Fig. 3.5, it can be seen that the *approach 2* performs significantly better than the *approach 1*, especially for large values of  $DE$ . The choice of taking iid symbols as the probing signal for the *approach 2* guarantees that submatrices of  $\mathbf{X}$  in (3.18) are much more well-conditioned than submatrices of  $\mathbf{A}_c$  in (3.15). For  $DE = 1$  the approaches perform almost equal.

The same results are obtained for the noisy case of Fig. 3.6, where the *approach 2* is superior especially for large values of  $|\Gamma|$ . It can be seen that for small values of  $|\Gamma|$ , the plots lie on a linear slope, indicating that the support set has been correctly identified, but the recovery of the samples of the spreading function is corrupted by noise. When  $|\Gamma|$  increases, the support set is not identified any more and the plots deviate from the linear slope.



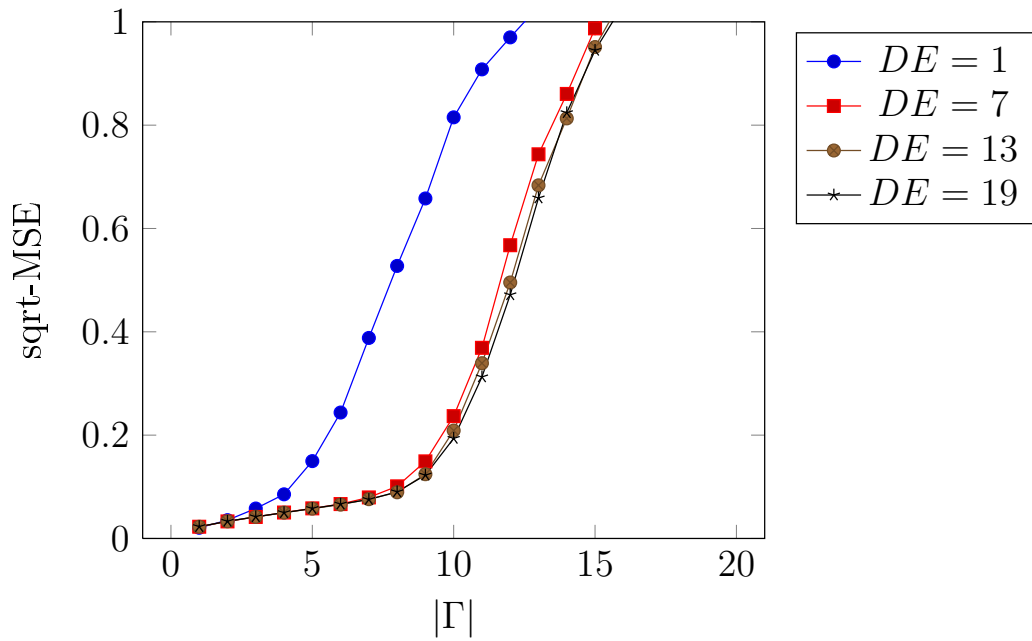


(a) MMV model, S-OMP

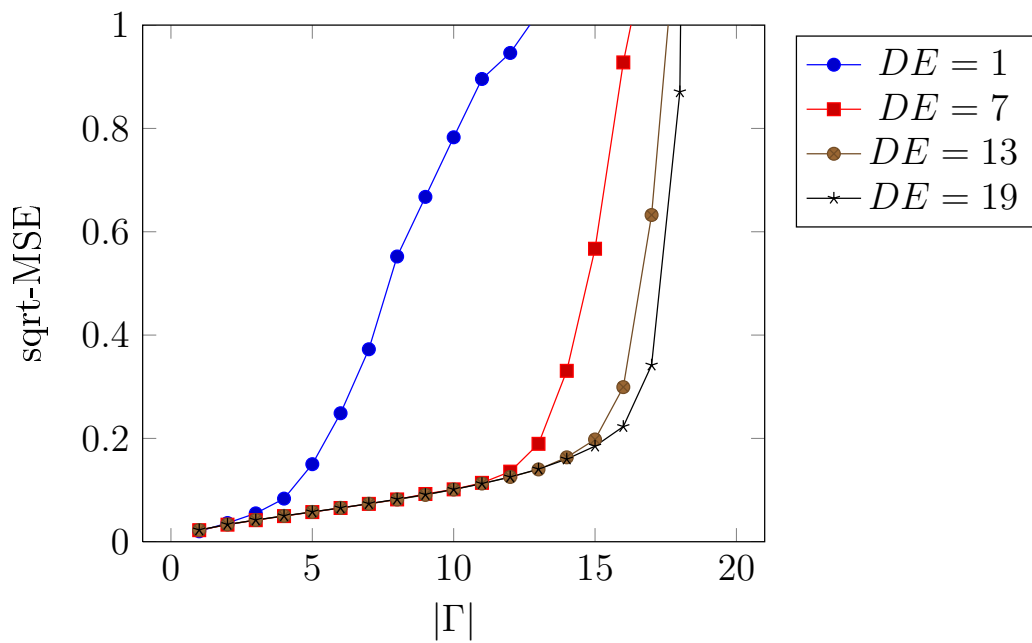


(b) B-sparsity model, B-OMP

Figure 3.5.: Recovery probability in the noiseless case.



(a) MMV model, S-OMP



(b) B-sparsity model, B-OMP

Figure 3.6.: Sqrt-MSE in the noisy case, with a SNR of 20 dB.

# Identification of Parametric Underspread Linear Systems

In this chapter we study the identification of a particular class of narrowband systems, called parametric underspread linear systems. The identification conditions seen for a narrowband system in general hold also in this case. We extend the analysis presented for a single input system to a multiple input system, using some ideas as in [23].

In Section 4.1 we provide the characterization of a parametric ULS. In Section 4.2 we formalize the problem of identification and give the system assumptions.

A matrix formulation of the problem is given in Section 4.4. In Sections 4.3 and 4.5 we propose a recovery procedure of the parameters that describe the system, while in section 4.6 we discuss on the implementation of the recovery procedure.

In Section 4.7 we specify the sufficient conditions on the input signal needed to guarantee the unique identification using the proposed procedure.

Finally, in Section 4.8 we extend the results [23] found previously for the system identification with a single input, to multiple inputs.

## 4.1. System Characterization

The general input-output relation of a linear time-varying system, as seen in chapter 3, is given by

$$y(t) = \int_{\tau} \int_{\nu} s_H(\tau, \nu) x(t - \tau) e^{j2\pi\nu t} d\nu d\tau \quad (4.1)$$

where the received signal  $y(t)$  consists of the continuous superposition of time- and frequency-shifted versions of the transmitted signal  $x(t)$ , weighted according to the *spreading function*  $s_H(\tau, \nu)$ .

We consider systems with spreading function characterized by a finite set of delays  $\tau_k$  and Doppler-shifts  $\nu_k$  (Fig. 4.1), i.e.,

$$s_H(\tau, \nu) = \sum_{k=1}^K \alpha_k \delta(\tau - \tau_k) \delta(\nu - \nu_k) \quad (4.2)$$

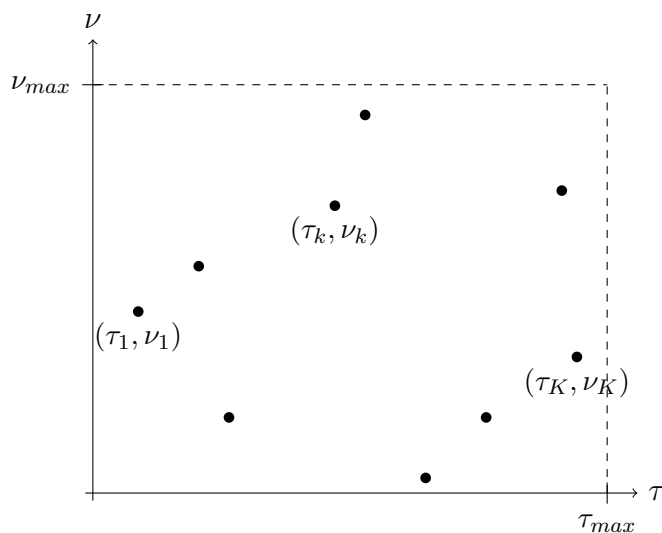


Figure 4.1.: Spreading function of a parametric ULS.

where parameters  $\alpha_k \in \mathbb{C}$  are the attenuation factors associated with the delay-Doppler pairs  $(\tau_k, \nu_k)$  and the delays  $\tau_k$  are assumed to be *distinct*. Relation (4.1) can be expressed, using (4.2), as

$$y(t) = \sum_{k=1}^K \alpha_k x(t - \tau_k) e^{j2\pi\nu_k t}. \quad (4.3)$$

Systems described by (4.3) are referred to as *parametric underspread linear systems* in [23]. The term *underspread* is referred to as systems, whose spreading function is supported within a region in the delay-Doppler plane of area smaller than 1. Such systems are identifiable as noticed in Kailath's work [7].

The identification of (4.3) involves finding a probing signal  $x(t)$  that guarantees the system parameters to be recovered from the observed signal  $y(t)$ . The parameters that completely characterize the system are the triplets  $(\tau_k, \nu_k, \alpha_k)$ , for  $k = 1, \dots, K$ . In the following, we give conditions on the probing signal that ensure the identification of (4.3) and derive a recovery procedure that estimates the parameters above.

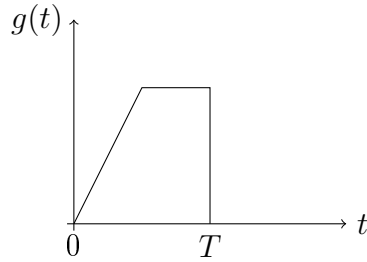
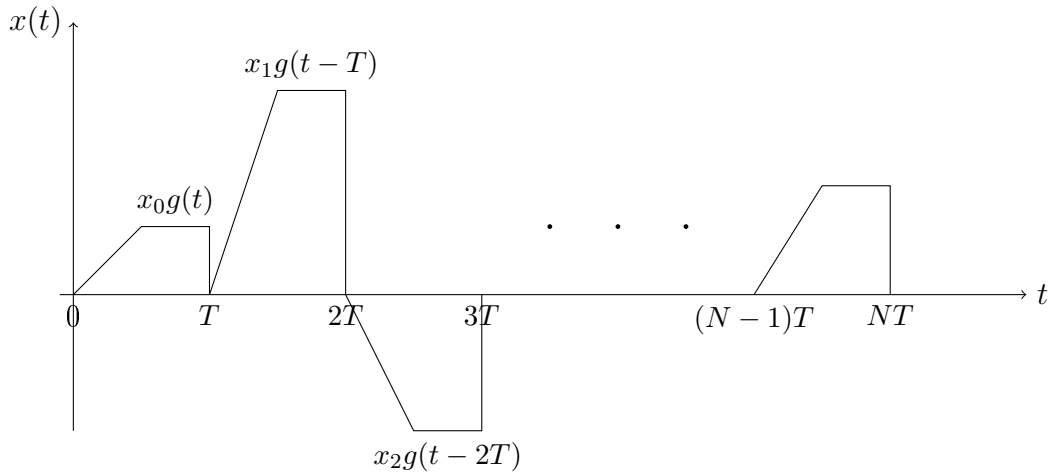
## 4.2. Assumptions for Identification

The first assumption, already mentioned above, is that the delays of the spreading function are distinct. This could seem a restrictive property but it is largely verified in practical situations.

The probing signal  $x(t)$  is chosen as a finite train of pulses, i.e.,

$$x(t) = \sum_{n=0}^{N-1} x_n g(t - nT) \quad (4.4)$$

where


 Figure 4.2.: Example of the prototype pulse  $g(t)$ .

 Figure 4.3.: Example of the probing signal  $x(t)$ .

- $g(t)$  is a prototype pulse, supported on  $[0, T]$  and with unit energy,  $\int |g(t)|^2 dt = 1$  (an example is depicted in Fig. 4.2);
- $\{x_n \in \mathbb{C}\}$  is an  $N$ -length probing sequence.

An example of the probing signal  $x(t)$  is shown in Fig. 4.3. The temporal support of  $x(t)$ , indicated as  $\mathcal{T}$ , is defined as  $\mathcal{T} = NT$ , while its two-sided bandwidth, indicated as  $\mathcal{W}$ , is the same as that of  $g(t)$ .

Since for an arbitrary pulse  $g(t)$  the bandwidth and the temporal support are related to each other as  $\mathcal{W} \propto 1/T$ , the parameter  $N$  is proportional to the time-bandwidth product of  $x(t)$ ,

$$N = \frac{\mathcal{T}}{T} \propto \mathcal{T}\mathcal{W} \quad (4.5)$$

which is according to the  $2WT$ -Theorem [9] the number of temporal degrees of freedom available for estimating the system.

The following assumptions are made in order to have some restrictions on the support of the spreading function  $s_H(\tau, \nu)$ .

- A) The support of the spreading function  $s_H(\tau, \nu)$  in (4.2) lies within a rectangular region of the delay-Doppler plane, i.e.,  $(\tau_k, \nu_k) \in [0, \tau_{max}] \times [0, \nu_{max}]$ , for  $k = 1, \dots, K$ . The parameters  $\tau_{max}$  and  $\nu_{max}$  are referred to as *delay spread* and *Doppler spread* of the system, respectively.

- B) The delay spread is strictly smaller than the temporal support of  $g(t)$ , or  $\tau_{max} < T$ .
- C) The Doppler spread is much smaller than the bandwidth of  $g(t)$ ,  $\nu_{max} \ll \mathcal{W}$ . Noting that  $\mathcal{W} \propto 1/T$ , this assumption is equivalent to  $\nu_{max}T \ll 1$ .

The main result for the identification of a ULS with a single input is summarized in the following theorem.

**Theorem 4. Identification of Parametric Underspread Linear Systems.** Suppose that a parametric ULS is completely described by  $K$  triplets  $(\tau_k, \nu_k, \alpha_k)$ , where delays  $\tau_k$  are distinct. Then the system can be identified as long as it satisfies assumptions A), B), C), the probing sequence  $\{x_n\}$  is nonzero for all  $n = 0, \dots, N-1$  and the time-bandwidth product of the known input signal  $x(t)$  satisfies the condition

$$\mathcal{T}\mathcal{W} \geq 4K \quad (4.6)$$

The probing signal that guarantees identification is not arbitrary but needs to be of the form of (4.4), as it will be clear in the following.

Before describing the recovery procedure in detail, a proper expression for the response signal  $y(t)$  is derived. According to the probing signal in (4.4), the input-output relation in (4.3) can be expressed as

$$y(t) = \sum_{k=1}^K \sum_{n=0}^{N-1} \alpha_k x_n e^{j2\pi\nu_k t} g(t - \tau_k - nT) \quad (4.7)$$

$$\approx \sum_{k=1}^K \sum_{n=0}^{N-1} \alpha_k x_n e^{j2\pi\nu_k nT} g(t - \tau_k - nT) \quad (4.8)$$

$$= \sum_{k=1}^K \sum_{n=0}^{N-1} a_k[n] g(t - \tau_k - nT) \quad (4.9)$$

where the sequences  $a_k[n]$ ,  $k = 1, \dots, K$  in (4.9) are defined as

$$a_k[n] = \alpha_k x_n e^{j2\pi\nu_k nT}, \quad n = 0, \dots, N-1 \quad (4.10)$$

The approximation in (4.8) follows from the assumptions B) and C), which is seen as follows. Since  $g(t)$  is compactly supported on  $[0, T]$ , then  $g(t - \tau_k - nT) = 0$  for  $t \notin [nT + \tau_k, (n+1)T + \tau_k]$ . First of all, due to the assumption B),  $\tau_k < T \forall k$  and  $g(t - \tau_k - nT) = 0$  for  $t \notin [nT, (n+2)T]$ . Therefore, for each value of  $k$  and  $n$  in (4.7),  $e^{j2\pi\nu_k t} g(t - \tau_k - nT) = e^{j2\pi\nu_k [(n+1)T + \tilde{t}]} g(t - \tau_k - nT)$ , where  $\tilde{t} \in [-T, T]$ . Secondly, due to the assumption C),  $\nu_k T \ll 1 \forall k$  and then  $e^{j2\pi\nu_k T} \approx 1$  and  $e^{j2\pi\nu_k \tilde{t}} \approx 1$ . As a consequence, the largest error committed by approximating  $e^{j2\pi\nu_k nT} \approx e^{j2\pi\nu_k t}$  in (4.8) is very small, given by  $e^{j2\pi\nu_{max} 2T} \approx 1$ .

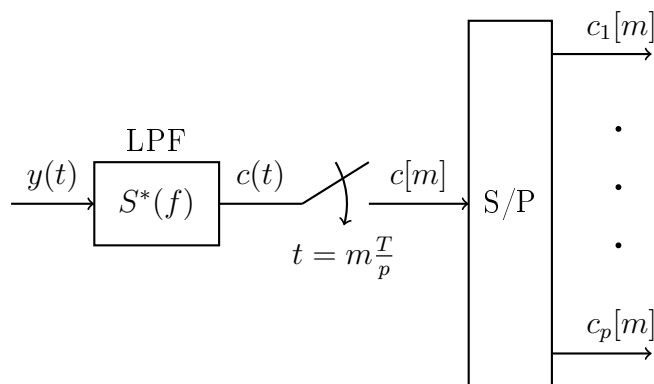


Figure 4.4.: Sampling stage for the identification procedure.

The procedure used for identification of the system (4.9) can be divided into two stages, the *sampling stage* and the *recovery stage*. In the first one the observed signal  $y(t)$  is sampled and processed in a suitable form, while in the second one the estimation of the triplets  $(\tau_k, \nu_k, \alpha_k)$  is performed.

### 4.3. Recovery Procedure: The Sampling Stage

The operations performed in the sampling stage are reported in Fig. 4.4.

The output of the system  $y(t)$  is first processed through a lowpass filter (LPF) and then sampled by an uniform sampler. The LPF is required before the sampler as an anti-aliasing filter and it also rejects the noise outside the working spectral band. The impulse response of the LPF is given by  $s^*(-t)$  and its frequency response  $S^*(f)$  has the following frequency support,

$$\mathcal{F} = \left[ -\frac{p}{2T}, \frac{p}{2T} \right]. \quad (4.11)$$

Parameter  $p$  is assumed to be even and to satisfy the condition  $p \geq 2K$ .

The signal  $c(t)$ , returned by the LPF, is sampled at times  $\{t = mT/p\}$ ,  $m \in \mathbb{Z}$ , which yields the discrete signal  $c[m]$ . A serial-to-parallel conversion (S/P) is then applied to obtain the subsequences  $\{c_l[m]\}$ ,  $l = 1, \dots, p$ . The  $m$ th sample of the  $l$ th sequence is given by

$$\begin{aligned} c_l[m] &= c(mT + (l-1)T/p) \\ &= (y(t) * s^*(-t))(mT + (l-1)T/p) \\ &= \int_{-\infty}^{+\infty} y(t) s^* \left( mT + (l-1)\frac{T}{p} + t \right) dt \\ &= \left\langle y(t), s \left( t + mT + (l-1)\frac{T}{p} \right) \right\rangle. \end{aligned} \quad (4.12)$$

The sequences  $\{c_l[m]\}$ ,  $l = 1, \dots, p$  correspond to periodically splitting the samples at the output of the sampler, which are generated at a rate of  $p/T$ , into  $p$  sequences at a rate of  $1/T$  each.

In the next section, we express the relationship between the sequences  $\{c_l[m]\}$ ,  $l = 1, \dots, p$ , obtained from (4.12) and the unknown sequences  $\{a_k[m]\}$ ,  $k = 1, \dots, K$ , given in (4.10). We show that it is possible to formulate the following system of equations, given in the matrix form as

$$\mathbf{d}[m] = \mathbf{N}(\tau)\mathbf{b}[m], \quad m \in \mathbb{Z} \quad (4.13)$$

where the  $p$ -length column vectors  $\mathbf{d}[m]$  depend on the sequences  $\{c_l[m]\}$ , the  $K$ -length column vectors  $\mathbf{b}[m]$  depend on the sequences  $\{a_k[m]\}$  and the  $p \times K$  matrix  $\mathbf{N}(\tau)$  depends only on the unknown delays  $\tau_k$ ,  $k = 1, \dots, K$ . In fact,  $\mathbf{N}(\tau)$  is a generalized Vandermonde matrix, defined as

$$\mathbf{N}(\tau) = \begin{bmatrix} \sigma_1^{\tau_1} & \sigma_1^{\tau_2} & \cdots & \sigma_1^{\tau_K} \\ \sigma_2^{\tau_1} & \sigma_2^{\tau_2} & \cdots & \sigma_2^{\tau_K} \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_p^{\tau_1} & \sigma_p^{\tau_2} & \cdots & \sigma_p^{\tau_K} \end{bmatrix} \quad (4.14)$$

where

$$\begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \vdots \\ \sigma_p \end{bmatrix} = \begin{bmatrix} e^{-j\frac{2\pi}{T}(-\frac{p}{2})} \\ e^{-j\frac{2\pi}{T}(-\frac{p}{2}+1)} \\ \vdots \\ e^{-j\frac{2\pi}{T}(\frac{p}{2}-1)} \end{bmatrix} \quad (4.15)$$

## 4.4. Matrix Formulation

Before describing the recovery step, we derive the matrix formulation (4.13) of the identification problem.

Define the  $p$ -length column vector  $\mathbf{c}(f)$ , whose  $l$ th element is the discrete-time Fourier transform (DTFT)  $C_l(e^{j2\pi fT})$  of  $\{c_l[m]\}$ . Similarly, define the  $K$ -length column vector  $\mathbf{a}(f)$ , whose  $k$ th element is the DTFT  $A_k(e^{j2\pi fT})$  of  $\{a_k[m]\}$ , defined in (4.10). The relation between these two vectors is expressed as follows.

$$\begin{aligned} C_l(e^{j2\pi fT}) &\triangleq \sum_{m \in \mathbb{Z}} c_l[m] e^{-j2\pi f m T} \\ \text{(from (4.12))} &= \sum_{m \in \mathbb{Z}} \left\langle y(t), s \left( t + mT + (l-1)\frac{T}{p} \right) \right\rangle e^{-j2\pi f m T} \\ \text{(from (4.9))} &= \sum_{k=1}^K \sum_{n=0}^{N-1} a_k[n] \sum_{m \in \mathbb{Z}} \left\langle g(t - nT - \tau_k), s \left( t + mT + (l-1)\frac{T}{p} \right) \right\rangle e^{-j2\pi f m T} \\ &= \sum_{k=1}^K \sum_{n=0}^{N-1} a_k[n] \sum_{m \in \mathbb{Z}} (\hat{g} * \hat{s})(mT) e^{-j2\pi f m T} \end{aligned} \quad (4.16)$$

where the functions  $\hat{g}(t)$  and  $\hat{s}(t)$  are defined as

$$\hat{g}(t) \triangleq g(t - nT - \tau_k) \quad (4.17)$$

$$\hat{s}(t) \triangleq s^* \left( -t - (l-1)\frac{T}{p} \right) \quad (4.18)$$



According to the Poisson summation formula, equation (4.16) becomes

$$C_l(e^{j2\pi fT}) = \sum_{k=1}^K \sum_{n=0}^{N-1} a_k[n] \frac{1}{T} \sum_{i \in \mathbb{Z}} \hat{G}\left(f + \frac{i}{T}\right) \hat{S}\left(f + \frac{i}{T}\right) \quad (4.19)$$

where  $\hat{G}(f)$  and  $\hat{S}(f)$  are the continuous Fourier transforms of  $\hat{g}(t)$  and  $\hat{s}(t)$ , respectively. According to (4.17) and (4.18), they are given by

$$\hat{G}(f) = G(f)e^{-j2\pi f(nT+\tau_k)} \quad (4.20)$$

$$\hat{S}(f) = S^*(f)e^{+j2\pi f(l-1)\frac{T}{p}} \quad (4.21)$$

where  $G(f)$  is the continuous Fourier transform of  $g(t)$ . Substituting (4.20) and (4.21) into (4.19) yields

$$\begin{aligned} C_l(e^{j2\pi fT}) &= \sum_{k=1}^K \sum_{n=0}^{N-1} a_k[n] \frac{1}{T} \sum_{i \in \mathbb{Z}} G\left(f + \frac{i}{T}\right) S^*\left(f + \frac{i}{T}\right) e^{j2\pi(f+\frac{i}{T})((l-1)\frac{T}{p}-nT-\tau_k)} \\ &= \sum_{k=1}^K e^{-j2\pi f\tau_k} \left[ \sum_{n=0}^{N-1} a_k[n] e^{-j2\pi fnT} \right] \sum_{i \in \mathbb{Z}} G_{l,i} N_{i,k} \\ &= \sum_{k=1}^K e^{-j2\pi f\tau_k} A_k(e^{j2\pi fT}) \sum_{i \in \mathbb{Z}} G_{l,i} N_{i,k} \end{aligned} \quad (4.22)$$

where  $G_{l,i}$  and  $N_{i,k}$  are defined as

$$G_{l,i} \triangleq \frac{1}{T} G\left(f + \frac{i}{T}\right) S^*\left(f + \frac{i}{T}\right) e^{j2\pi(f+\frac{i}{T})(l-1)\frac{T}{p}} \quad (4.23)$$

$$N_{i,k} \triangleq e^{-j2\pi \frac{i}{T} \tau_k} \quad (4.24)$$

Now, since  $C_l(e^{j2\pi fT})$  and  $A_k(e^{j2\pi fT})$  are  $1/T$  periodic, assume that  $f \in [0, 1/T]$ . As a consequence,  $f + i/T \in [\frac{i}{T}, \frac{i+1}{T}]$ ,  $i \in \mathbb{Z}$ . Comparing this interval with the support of  $S^*(f)$  in (4.11), the values of  $i$  that make  $[\frac{i}{T}, \frac{i+1}{T}]$  fit into  $\mathcal{F}$  are  $i = \{-p/2, \dots, p/2 - 1\}$ , since  $S^*(f + \frac{i}{T}) = 0$  for  $i < p/2$  and for  $i > p/2 - 1$ . Therefore the last summation in (4.22) involves only  $p$  nonzero terms, i.e.,

$$C_l(e^{j2\pi fT}) = \sum_{k=1}^K e^{-j2\pi f\tau_k} A_k(e^{j2\pi fT}) \sum_{i=-p/2}^{p/2-1} G_{l,i} N_{i,k}. \quad (4.25)$$

With the change of variable  $i' = i + p/2 + 1$ , if  $i = \{-p/2, \dots, p/2 - 1\}$  then  $i' = \{1, \dots, p\}$ , which can be used as a row/column index. The relation (4.25) can be expressed in the following matrix form

$$\mathbf{c}(f) = \mathbf{W}(f) \mathbf{N}(\tau) \mathbf{D}(f, \tau) \mathbf{a}(f) \quad (4.26)$$

where

- The  $(l, i')$ th element of the  $p \times p$  matrix  $\mathbf{W}(f)$  is given by  $G_{l,i}$  in (4.23), with  $i = i' - p/2 - 1$ .
- The matrix  $\mathbf{N}(\tau)$  is a  $p \times K$  generalized Vandermonde matrix with  $(i', k)$ th element given by  $N_{i',k}$  in (4.24), with  $i = i' - p/2 - 1$ ; note that  $\mathbf{N}(\tau)$  is the same matrix involved in the final formulation, already defined in (4.14).
- The matrix  $\mathbf{D}(f, \tau)$  is a  $K \times K$  diagonal matrix whose  $k$ th diagonal element is given by  $e^{-j2\pi f \tau_k}$ .

Assuming that matrix  $\mathbf{W}(f)$  is invertible, (4.26) can be written as

$$\mathbf{W}^{-1}(f)\mathbf{c}(f) = \mathbf{N}(\tau)\mathbf{D}(f, \tau)\mathbf{a}(f). \quad (4.27)$$

The conditions under which matrix  $\mathbf{W}(f)$  is invertible are given in section 4.7. Now, denoting the  $p$ -length column vector  $\mathbf{d}(f)$  and the  $K$ -length column vector  $\mathbf{b}(f)$  as follows

$$\mathbf{d}(f) \triangleq \mathbf{W}^{-1}(f)\mathbf{c}(f) \quad (4.28)$$

$$\mathbf{b}(f) \triangleq \mathbf{D}(f, \tau)\mathbf{a}(f) \quad (4.29)$$

(4.27) becomes simply

$$\mathbf{d}(f) = \mathbf{N}(\tau)\mathbf{b}(f). \quad (4.30)$$

Since  $\mathbf{N}(\tau)$  is not a function of  $f$ , the inverse DTFT can be applied to both the left and right hand sides of (4.30) yielding

$$\mathbf{d}[m] = \mathbf{N}(\tau)\mathbf{b}[m], \quad m \in \mathbb{Z} \quad (4.31)$$

where the set of the  $l$ th elements of each vector  $\mathbf{d}[m]$ ,  $m \in \mathbb{Z}$ , indicated as  $\{d_l[m]\}$ , forms the inverse DTFT of the  $l$ th element of  $\mathbf{d}(f)$ , indicated as  $D_l(e^{j2\pi f T})$ ,  $l = 1, \dots, p$ ,

$$d_l[m] = T \int_0^{1/T} D_l(e^{j2\pi f T}) e^{j2\pi f m T} df, \quad m \in \mathbb{Z}. \quad (4.32)$$

In the same way,

$$b_k[m] = T \int_0^{1/T} B_k(e^{j2\pi f T}) e^{j2\pi f m T} df, \quad m \in \mathbb{Z} \quad (4.33)$$

where  $B_k(e^{j2\pi f T})$  is the  $k$ th element of  $\mathbf{b}(f)$ ,  $k = 1, \dots, K$ , and the sequence  $\{b_k[m]\}$  is its inverse DTFT, given by the set of the  $k$ th elements of each vector  $\mathbf{b}[m]$ ,  $m \in \mathbb{Z}$ .

Note that the vectors  $\mathbf{d}[m]$  can be reconstructed from the system response  $y(t)$  as, from (4.28), they depend only on the samples  $\{c_l[m]\}$  and on the structure of Fig. 4.4. On the other hand,  $\mathbf{N}(\tau)\mathbf{b}[m]$  in (4.31) depends on the unknown parameters  $(\tau_k, \nu_k, \alpha_k)$ .

In the next section, we show how to recover the triplets  $(\tau_k, \nu_k, \alpha_k)$  from the infinite set of measurement vectors  $\mathbf{d}[m]$  in (4.31).

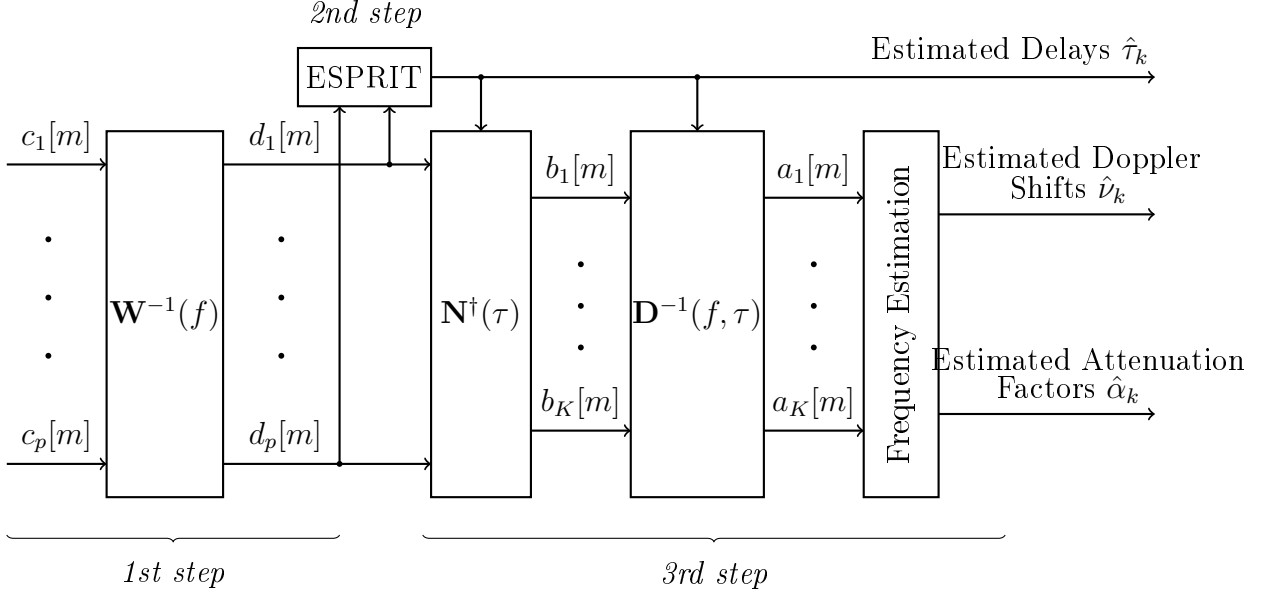


Figure 4.5.: Recovery stage for the identification procedure.

## 4.5. Recovery Procedure: The Recovery Stage

The recovery stage consists of three steps.

In the first step, the measurement vectors  $\mathbf{d}[m]$ , or equivalently the sequences  $\{d_l[m]\}$ , are obtained from the sequences  $\{c_l[m]\}$ , according to (4.28). In the second step, the ESPRIT algorithm is used to recover the unknown delays  $\tau_k$  from  $\mathbf{d}[m]$ . Finally, in the third step, the recovered delays are used to obtain the Doppler-shifts  $\nu_k$  and the attenuation factors  $\alpha_k$  associated with each delay. The architecture of the recovery stage is shown in Fig. 4.5, while the three steps are discussed in detail below.

### 4.5.1. First step: recovery of the measurement vectors.

From the relation in (4.28) and according to the definition in (4.32) the measurement vectors  $\mathbf{d}[m]$  can be expressed as

$$\mathbf{d}[m] = \text{IDTFT} \{ \mathbf{d}(f) \} [m] = \text{IDTFT} \{ \mathbf{W}^{-1}(f) \mathbf{c}(f) \} [m]. \quad (4.34)$$

The direct relationship between the sequences  $\{d_l[m]\}$  and  $\{c_l[m]\}$  is here derived, while the implementation of  $\text{IDTFT} \{ \mathbf{W}^{-1}(f) \mathbf{c}(f) \}$  is reported in the next section. Denoting with  $W_{lj}(f)$  the  $(l, j)$ th element of the matrix  $\mathbf{W}^{-1}(f)$ , the  $l$ th entry of  $\mathbf{d}[m]$  in (4.34) becomes

$$\begin{aligned} d_l[m] &= \text{IDTFT} \{ D_l(e^{j2\pi fT}) \} [m] \\ &= \text{IDTFT} \{ W_{l1}(f)C_1(e^{j2\pi fT}) + \dots + W_{lp}(f)C_p(e^{j2\pi fT}) \} [m] \\ &= \text{IDTFT} \{ W_{l1}(f)C_1(e^{j2\pi fT}) \} [m] + \dots + \text{IDTFT} \{ W_{lp}(f)C_p(e^{j2\pi fT}) \} [m] \\ &= \text{IDTFT} \{ W_{l1}(f) \} [m] * c_1[m] + \dots + \text{IDTFT} \{ W_{lp}(f) \} [m] * c_p[m]. \end{aligned} \quad (4.35)$$

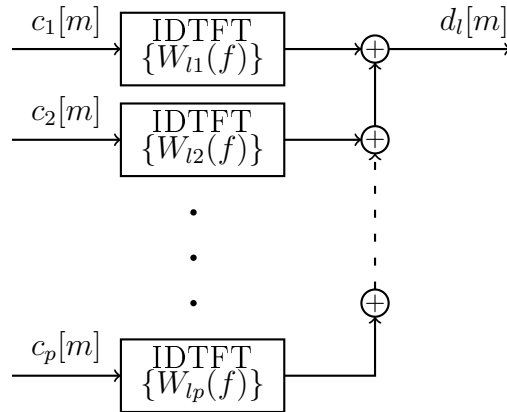


Figure 4.6.: Recovery of the sequences  $\{d_l[m]\}$  from  $\{c_l[m]\}$ .

Each sample  $c_j[m]$ ,  $j = 1, \dots, p$ , is filtered with a filter whose frequency response is given by  $W_{l_j}(f)$ . The outputs are then summed to obtain the sample  $d_l[m]$ . A graphical representation of (4.35) is reported in Fig. 4.6.

#### 4.5.2. Second step: ESPRIT algorithm and recovery of the delays

Once the measurement vectors  $\mathbf{d}[m]$  are reconstructed, relation (4.31) is used to obtain the unknown delays  $\tau_k$ . In fact, (4.31) describes an infinite set of measurement vectors  $\mathbf{d}[m]$ ,  $m \in \mathbb{Z}$ , each obtained by the same measurement matrix  $\mathbf{N}(\tau)$ , which has a fixed structure and is completely determined by the unknown delays  $\tau_k$ , according to (4.14) and (4.15). The same situation has been studied in the field of direction-of-arrival estimation (DOA) and a method used for this type of problem is the ESPRIT algorithm [24]. This algorithm belongs to a class of techniques, known as *subspace methods*, which are based on separating the space generated by the measurement vectors into two subspaces, the signal and the noise subspaces. Although the ESPRIT algorithm needs the system to have the *rotational invariance property*, which is satisfied in our case, it is one of the most efficient subspace method for this type of problem. Before describing in detail the ESPRIT algorithm, we give an idea of the rotational invariance property. It requires that the subsystems that recover the sequences  $\{d_l[m]\}$  are related by a proper phase shift, as it will be clear in the following. In our case, the structure of the Vandermonde matrix  $\mathbf{N}(\tau)$  guarantees the rotational invariance property.

We refer to the procedure described in [22], summarized also in [27].

Consider the matrix  $\mathbf{N}(\tau)$ , given in (4.14). Create now two subsystems where  $p - 1$  sub channels upon the  $p$  sequences  $\{d_l[m]\}$  are considered. In particular, the first subsystem has  $\{d_1[m]\}, \{d_2[m]\}, \dots, \{d_{p-1}[m]\}$  and the second subsystem has  $\{d_2[m]\}, \{d_3[m]\}, \dots, \{d_p[m]\}$ . We denote with  $\mathbf{N}_\downarrow(\tau)$  and  $\mathbf{N}_\uparrow(\tau)$  the matrices that relate the  $p - 1$  outputs of the first and the second subsystem, respectively, with the input vectors  $\mathbf{b}[m]$  in (4.31). In particular,  $\mathbf{N}_\downarrow(\tau)$  is defined as the submatrix extracted from  $\mathbf{N}(\tau)$  by deleting the last row (referred to  $\{d_p[m]\}$ ), while  $\mathbf{N}_\uparrow(\tau)$  is

the submatrix extracted from  $\mathbf{N}(\tau)$  by deleting the first row (referred to  $\{d_1[m]\}$ ). Since  $\mathbf{N}(\tau)$  is a Vandermonde matrix, it satisfies the rotational invariance property, as

$$\mathbf{N}_\uparrow(\tau) = \mathbf{N}_\downarrow(\tau)\mathbf{R}(\tau) \quad (4.36)$$

where  $\mathbf{R}(\tau)$  is a diagonal  $K \times K$  matrix, given by

$$\mathbf{R}(\tau) = \text{diag} \left\{ e^{-j\frac{2\pi}{T}\tau_1}, \dots, e^{-j\frac{2\pi}{T}\tau_K} \right\}. \quad (4.37)$$

In order to recover the unknown delays  $\tau_k$ , the goal of the ESPRIT algorithm is to recover matrix  $\mathbf{R}(\tau)$  from vectors  $\mathbf{d}[m]$ . Define the  $(p-1) \times p$  selection matrices  $\mathbf{J}_\downarrow$  and  $\mathbf{J}_\uparrow$  such that

$$\mathbf{N}_\downarrow(\tau) = \mathbf{J}_\downarrow\mathbf{N}(\tau) \quad (4.38)$$

$$\mathbf{N}_\uparrow(\tau) = \mathbf{J}_\uparrow\mathbf{N}(\tau) \quad (4.39)$$

It is easy to see that

$$\mathbf{J}_\downarrow = [\mathbf{I}_{p-1} | \mathbf{0}_{p-1}] \quad (4.40)$$

$$\mathbf{J}_\uparrow = [\mathbf{0}_{p-1} | \mathbf{I}_{p-1}] \quad (4.41)$$

where  $\mathbf{I}_{p-1}$  is the identity matrix of size  $p-1$  and  $\mathbf{0}_{p-1}$  is a  $(p-1)$ -length column vector of zero elements. Substituting (4.38) and (4.39) into (4.36) yields

$$\mathbf{J}_\uparrow\mathbf{N}(\tau) = \mathbf{J}_\downarrow\mathbf{N}(\tau)\mathbf{R}(\tau). \quad (4.42)$$

We compute now the *signal space*, i.e., the vector space spanned the measurement vectors  $\mathbf{d}[m]$ , reconstructed in the previous step. To this end, calculate the  $p \times p$  correlation matrix of  $\mathbf{d}[m]$ , as

$$\mathbf{R}_{dd} \triangleq \sum_{m \in \mathbb{Z}} \mathbf{d}[m]\mathbf{d}^H[m]. \quad (4.43)$$

Substituting (4.31) into (4.43) yields

$$\begin{aligned} \mathbf{R}_{dd} &= \sum_{m \in \mathbb{Z}} (\mathbf{N}(\tau)\mathbf{b}[m])(\mathbf{N}(\tau)\mathbf{b}[m])^H \\ &= \mathbf{N}(\tau) \left( \sum_{m \in \mathbb{Z}} \mathbf{b}[m]\mathbf{b}^H[m] \right) \mathbf{N}^H(\tau) \\ &= \mathbf{N}(\tau)\mathbf{R}_{bb}\mathbf{N}^H(\tau) \end{aligned} \quad (4.44)$$

where  $\mathbf{R}_{bb}$  is the  $K \times K$  correlation matrix of  $\mathbf{b}[m]$ .

Because of its Vandermonde structure, the columns of  $\mathbf{N}(\tau)$  are all linearly independent. Since  $\mathbf{N}(\tau)$  and  $\mathbf{N}^H(\tau)$  have full column- and row-rank, respectively, from relation (4.44) matrices  $\mathbf{R}_{dd}$  and  $\mathbf{R}_{bb}$  have the same rank. We assume for simplicity that matrix  $\mathbf{R}_{bb}$  is nonsingular, i.e. it has full rank. This assumption is satisfied if all the sequences  $\{b_k[m]\}$  defined in (4.33) are uncorrelated or partially correlated. We analyze later the case of singular matrix  $\mathbf{R}_{bb}$ . From the observations above

$$\text{rank}(\mathbf{R}_{dd}) = \text{rank}(\mathbf{R}_{bb}) = K. \quad (4.45)$$

According to the relation in (4.45), matrix  $\mathbf{R}_{dd}$  is singular and has  $p - K$  zero eigenvalues. Indicating with  $\mathbf{\Lambda}_s$  the  $K \times K$  diagonal matrix containing the nonzero eigenvalues of  $\mathbf{R}_{dd}$ , the eigenvalue decomposition of  $\mathbf{R}_{dd}$  becomes

$$\mathbf{R}_{dd} = \mathbf{E}_s \mathbf{\Lambda}_s \mathbf{E}_s^H \quad (4.46)$$

where the columns of the  $p \times K$  matrix  $\mathbf{E}_s$  are the  $K$  eigenvectors corresponding to the nonzero eigenvalues of  $\mathbf{R}_{dd}$ . The columns of  $\mathbf{E}_s$  constitute the signal space, the space spanned by vectors  $\mathbf{d}[m]$ . We show now that  $\mathbf{E}_s$  and  $\mathbf{N}(\tau)$  have the same column span. Indicating with  $\mathbf{E}_n$  the  $p \times (p - K)$  matrix whose columns are the eigenvectors corresponding to the zero eigenvalues of  $\mathbf{R}_{dd}$ , the column span of  $\mathbf{E}_s$  is the orthogonal complement of the column span of  $\mathbf{E}_n$ . Since the eigenvectors in  $\mathbf{E}_n$  are orthogonal to the columns of  $\mathbf{N}(\tau)$ , as we prove in the following, the eigenvectors in  $\mathbf{E}_s$  span the same space spanned by the columns of  $\mathbf{N}(\tau)$ . In fact, from the definition of  $\mathbf{E}_n$  and from (4.44)

$$\mathbf{R}_{dd} \mathbf{E}_n = \mathbf{0}_{p \times (p-K)} \quad (4.47)$$

$$\mathbf{N}(\tau) \mathbf{R}_{bb} \mathbf{N}^H(\tau) \mathbf{E}_n = \mathbf{0}_{p \times (p-K)} \quad (4.48)$$

$$\mathbf{N}^H(\tau) \mathbf{E}_n = \mathbf{0}_{K \times (p-K)} \quad (4.49)$$

where the step in (4.49) is due to the fact that the  $p \times K$  matrix  $\mathbf{N}(\tau) \mathbf{R}_{bb}$  has full column rank and therefore  $(\mathbf{N}(\tau) \mathbf{R}_{bb})^\dagger (\mathbf{N}(\tau) \mathbf{R}_{bb}) = \mathbf{I}_K$ .

Since  $\mathbf{E}_s$  and  $\mathbf{N}(\tau)$  have the same column span, there exists an invertible  $K \times K$  matrix  $\mathbf{T}$  such that

$$\mathbf{N}(\tau) = \mathbf{E}_s \mathbf{T}. \quad (4.50)$$

Substituting (4.50) into (4.42) yields

$$\mathbf{J}_\uparrow \mathbf{E}_s \mathbf{T} = \mathbf{J}_\downarrow \mathbf{E}_s \mathbf{T} \mathbf{R}(\tau) \quad (4.51)$$

$$\mathbf{J}_\uparrow \mathbf{E}_s = \mathbf{J}_\downarrow \mathbf{E}_s \mathbf{T} \mathbf{R}(\tau) \mathbf{T}^{-1} \quad (4.52)$$

$$\mathbf{J}_\uparrow \mathbf{E}_s = \mathbf{J}_\downarrow \mathbf{E}_s \mathbf{\Phi} \quad (4.53)$$

where the  $K \times K$  matrix  $\mathbf{\Phi}$  is defined as

$$\mathbf{\Phi} \triangleq \mathbf{T} \mathbf{R}(\tau) \mathbf{T}^{-1}. \quad (4.54)$$

From the definition of selection matrices in (4.40) and (4.41), equation (4.53) becomes

$$\mathbf{E}_{s\uparrow} = \mathbf{E}_{s\downarrow} \mathbf{\Phi} \quad (4.55)$$

where the  $(p - 1) \times K$  matrices  $\mathbf{E}_{s\uparrow}$  and  $\mathbf{E}_{s\downarrow}$  are the submatrices extracted from  $\mathbf{E}_s$  by deleting the last and the first row, respectively. Since  $\mathbf{E}_s$  has full column rank, also  $\mathbf{E}_{s\downarrow}$  has full column rank and matrix  $\mathbf{\Phi}$  can be recovered from (4.55) as

$$\mathbf{\Phi} = \mathbf{E}_{s\downarrow}^\dagger \mathbf{E}_{s\uparrow}. \quad (4.56)$$

The recovery of the unknown delays  $\tau_k$  is completed since equation (4.54) represents the eigenvalue decomposition of matrix  $\mathbf{\Phi}$  and thus the delays  $\tau_k$  in the diagonal of matrix  $\mathbf{R}(\tau)$  can be recovered from the eigenvalues of matrix  $\mathbf{\Phi}$ , which are equal to the eigenvalues of  $\mathbf{R}(\tau)$ .

The ESPRIT algorithm steps are reported below.

**Algorithm 3. Delay Recovery Algorithm: ESPRIT algorithm**

*Input:*

- Measurement vectors  $\mathbf{d}[m]$ ,  $m \in \mathbb{Z}$ .

*Output:*

- The cardinality  $K$  of the set of unknown delays.
- The delays  $\tau_k$ , for  $k = 1, \dots, K$ .

*Procedure:*

- (a) Construct the correlation matrix  $\mathbf{R}_{dd}$  as

$$\mathbf{R}_{dd} = \sum_{m \in \mathbb{Z}} \mathbf{d}[m] \mathbf{d}^H[m]. \quad (4.57)$$

- (b) Recover  $K$  as the rank of  $\mathbf{R}_{dd}$ .

- (c) Perform an eigenvalue decomposition of  $\mathbf{R}_{dd}$  and construct the matrix  $\mathbf{E}_s$  consisting of the  $K$  eigenvectors corresponding to the  $K$  nonzero eigenvalues of  $\mathbf{R}_{dd}$  as its columns.

- (d) Compute the matrix  $\mathbf{\Phi}$  as

$$\mathbf{\Phi} = \mathbf{E}_{s\downarrow}^\dagger \mathbf{E}_{s\uparrow} \quad (4.58)$$

where  $\mathbf{E}_{s\downarrow}$  and  $\mathbf{E}_{s\uparrow}$  denote the submatrices extracted from  $\mathbf{E}_s$  by removing its first row and its last row, respectively.

- (e) Compute the eigenvalues of  $\mathbf{\Phi}$ , denoted by  $\lambda_k$ ,  $k = 1, \dots, K$ .

- (f) Recover the unknown delays as

$$\tau_k = -\frac{T}{2\pi} \arg(\lambda_k), \quad k = 1, \dots, K. \quad (4.59)$$

Before describing the third step of the recovery procedure, we analyze the case of singular matrix  $\mathbf{R}_{bb}$ . In this case the ESPRIT algorithm cannot be applied directly because the rank of  $\mathbf{R}_{dd}$  is smaller than  $K$  and its column span is not equal to the entire signal subspace. An additional stage is then required.

The technique is proposed in [25] and is based on the *smoothed correlation* matrix defined as follows. Define  $\mathbf{d}_l[m]$ ,  $l = 1, \dots, p - K$ , the  $(K + 1)$ -length subvectors of  $\mathbf{d}[m]$  given by

$$\mathbf{d}_l[m] = [d_l[m] \ d_{l+1}[m] \ \dots \ d_{l+K}[m]]^T, \quad l = 1, \dots, p - K. \quad (4.60)$$

The smoothed correlation matrix  $\overline{\mathbf{R}}_{dd}$  is constructed as

$$\overline{\mathbf{R}}_{dd} = \frac{1}{p-K} \sum_{l=1}^{p-K} \sum_{m \in \mathbb{Z}} \mathbf{d}_l[m] \mathbf{d}_l^H[m]. \quad (4.61)$$

Since  $p \geq 2K$ , then  $p-K \geq K$  and according to the following theorem [25] the rank of  $\overline{\mathbf{R}}_{dd}$  is  $K$  regardless of the rank of  $\mathbf{R}_{bb}$ .

**Theorem 5.** *If the number of subvectors  $\mathbf{d}_l[m]$  is greater or equal to the number of the sequences in  $\mathbf{b}[m]$ , i.e., if  $p-K \geq K$ , then the smoothed correlation matrix  $\overline{\mathbf{R}}_{dd}$  has rank  $K$ .*

The proof of the theorem is given in the appendix. We apply then the ESPRIT algorithm on  $\overline{\mathbf{R}}_{dd}$  and refer its column rank as the signal space. The algorithm steps are the same as before, substituting  $\overline{\mathbf{R}}_{dd}$  with  $\mathbf{R}_{dd}$ .

Finally, we define the smoothed correlation matrix  $\overline{\mathbf{R}}_{dd}$  if  $K$  is not known at the recovery stage. In this case, since  $p \geq 2K$ , it suffices to take  $M = p/2$  subvectors  $\mathbf{d}_l[m]$  and (4.61) becomes

$$\overline{\mathbf{R}}_{dd} = \frac{1}{M} \sum_{l=1}^M \sum_{m \in \mathbb{Z}} \mathbf{d}_l[m] \mathbf{d}_l^H[m] \quad (4.62)$$

where  $\mathbf{d}_l[m] = [d_l[m] \ d_{l+1}[m] \ \dots \ d_{l+M}[m]]^T$ .

### 4.5.3. Third step: recovery of the Doppler-shifts and attenuation factors

Once  $\tau_k$  are known, the matrix  $\mathbf{N}(\tau)$  is recovered from (4.14) and  $\mathbf{b}(f)$  can be found from (4.30), as

$$\mathbf{b}(f) = \mathbf{N}^\dagger(\tau) \mathbf{d}(f) \quad (4.63)$$

since  $\mathbf{N}^\dagger(\tau) \mathbf{N}(\tau) = \mathbf{I}_K$  because of the assumption  $p \geq 2K$ . Finally, vector  $\mathbf{a}(f)$  can be recovered from (4.29) as

$$\mathbf{a}(f) = \mathbf{D}^{-1}(f, \tau) \mathbf{b}(f) = \mathbf{D}^{-1}(f, \tau) \mathbf{N}^\dagger(\tau) \mathbf{d}(f) \quad (4.64)$$

since the diagonal matrix  $\mathbf{D}(f, \tau)$  is always invertible.

The equivalent relations (4.63) and (4.64) in the time domain are given by

$$\mathbf{b}[m] = \text{IDTFT} \{ \mathbf{b}(f) \} = \mathbf{N}^\dagger(\tau) \mathbf{d}[m] \quad (4.65)$$

$$\mathbf{a}[m] = \text{IDTFT} \{ \mathbf{a}(f) \} = \text{IDTFT} \{ \mathbf{D}^{-1}(f, \tau) \mathbf{b}(f) \} \quad (4.66)$$

while the implementation of (4.65) and (4.66) is reported in the next section.



The sequences  $\{a_k[n]\}$  defined in (4.10), for  $n = 0, \dots, N - 1$ , are obtained from vectors  $\mathbf{a}[m]$ , considering only  $m = 0, \dots, N - 1$ . In particular, for each  $m$ , the  $k$ th component of the vector  $\mathbf{a}[m]$  corresponds to the sample  $a_k[m]$ . Define the  $N$ -length column vector  $\mathbf{a}_k$ , whose  $n$ th element is  $a_k[n]$ , for  $k = 1, \dots, K$ . Then the relation (4.10) can be expressed in the following matrix form

$$\mathbf{a}_k = \alpha_k \mathbf{X} \mathbf{r}(\nu_k), \quad k = 1, \dots, K \quad (4.67)$$

where

- $\mathbf{X}$  is an  $N \times N$  diagonal matrix whose  $n$ th diagonal element is given by  $x_n$ .
- $\mathbf{r}(\nu_k)$  is a  $N$ -length column vector whose  $n$ th element is given by  $e^{j2\pi\nu_k nT}$ .

The matrix  $\mathbf{X}$  in (4.67) can be inverted under the assumption that the sequence  $\{x_n\}$  is nonzero for  $n = 0, \dots, N - 1$ , i.e.,  $|x_n| > 0, \forall n$ . In this case, defining  $\tilde{\mathbf{a}}_k = \mathbf{X}^{-1} \mathbf{a}_k$ , (4.67) becomes

$$\tilde{\mathbf{a}}_k = \alpha_k \mathbf{r}(\nu_k), \quad k = 1, \dots, K. \quad (4.68)$$

For each value of  $k$ , the attenuation factor  $\alpha_k$  and the Doppler shift  $\nu_k$  are given by

$$\alpha_k = \tilde{\mathbf{a}}_k[0] \quad (4.69)$$

$$e^{j2\pi\nu_k T} = \frac{1}{\alpha_k} \tilde{\mathbf{a}}_k[1] \rightarrow \nu_k = \frac{1}{j2\pi T} \ln \left( \frac{1}{\alpha_k} \tilde{\mathbf{a}}_k[1] \right). \quad (4.70)$$

## 4.6. Implementation of the Recovery Stage

In this section we show how to implement the blocks indicated with  $\mathbf{W}^{-1}(f)$ ,  $\mathbf{N}^\dagger(\tau)$  and  $\mathbf{D}^{-1}(f, \tau)$  in Fig. 4.5.

We refer to the relations in (4.35) and in Fig. 4.6 that express the relationship in the time domain between the input and the output sequences, whose DTFT vectors are related by a multiplication of a frequency dependent matrix. According to (4.35), if the matrix  $\mathbf{W}^{-1}(f)$  is diagonal, each output sequence is simply obtained by filtering only the corresponding input sequence with a filter whose frequency response is given in the diagonal of  $\mathbf{W}^{-1}(f)$ .

We compute first the implementation of the block  $\mathbf{W}^{-1}(f)$ . To this end, we express the matrix  $\mathbf{W}(f)$  as the product of simple matrices, i.e., diagonal or not frequency dependent matrices, in order to have a simple and efficient computation of  $\text{IDTFT}\{\mathbf{W}^{-1}(f)\}$ .

From the definition in (4.23) and substituting the variable  $i$  with  $i' - p/2 - 1$ , the

$(l, i')$ th element of  $\mathbf{W}(f)$  is given by

$$\begin{aligned}
 G_{l,i} &= \frac{1}{T}G\left(f + \frac{i}{T}\right) S^*\left(f + \frac{i}{T}\right) e^{j2\pi\left(f + \frac{i}{T}\right)(l-1)\frac{T}{p}} \\
 &= \frac{1}{T}G\left(f + \frac{i}{T}\right) S^*\left(f + \frac{i}{T}\right) e^{j2\pi f(l-1)\frac{T}{p}} e^{j\frac{2\pi}{p}i(l-1)} \\
 (i \leftarrow i' - p/2 - 1) &= \frac{1}{T}G\left(f + \frac{i}{T}\right) S^*\left(f + \frac{i}{T}\right) e^{j2\pi f(l-1)\frac{T}{p}} e^{j\frac{2\pi}{p}(i'-1)(l-1)} e^{-j\pi(l-1)} \\
 &= \left[(-1)^{l-1} e^{j2\pi f(l-1)\frac{T}{p}}\right] \left[e^{j\frac{2\pi}{p}(i'-1)(l-1)}\right] \left[\frac{1}{T}G\left(f + \frac{i}{T}\right) S^*\left(f + \frac{i}{T}\right)\right] \\
 &= \Phi_l(f) F_{l,i'}^* \Psi_{i'}(f) \tag{4.71}
 \end{aligned}$$

where

$$\Phi_l(f) \triangleq (-1)^{l-1} e^{j2\pi f(l-1)\frac{T}{p}} \tag{4.72}$$

$$F_{l,i'} \triangleq e^{-j\frac{2\pi}{p}(i'-1)(l-1)} \tag{4.73}$$

$$\Psi_{i'}(f) \triangleq \frac{1}{T}G\left(f + \frac{1}{T}(i' - p/2 - 1)\right) S^*\left(f + \frac{1}{T}(i' - p/2 - 1)\right) \tag{4.74}$$

From (4.71), the following decomposition for matrix  $\mathbf{W}(f)$  holds

$$\mathbf{W}(f) = \mathbf{\Phi}(f) \mathbf{F}^H \mathbf{\Psi}(f) \tag{4.75}$$

where

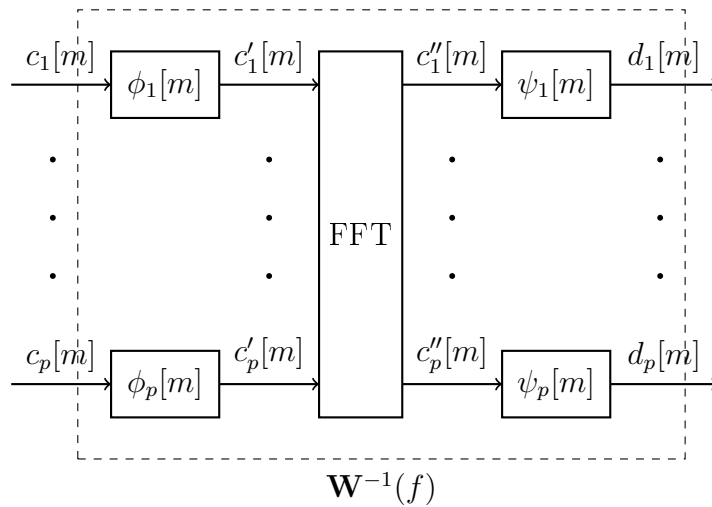
- The matrix  $\mathbf{\Phi}(f)$  is a  $p \times p$  diagonal matrix whose  $l$  diagonal element is given by  $\Phi_l(f)$ .
- The matrix  $\mathbf{F}$  is a  $p$ -point discrete Fourier transform matrix with  $(l, i')$ th element equal to  $F_{l,i'}$ .
- The matrix  $\mathbf{\Psi}(f)$  is a  $p \times p$  diagonal matrix whose  $i'$  diagonal element is given by  $\Psi_{i'}(f)$ .

Assuming that  $\mathbf{W}(f)$  is invertible and according to the decomposition in (4.75), the inverse matrix  $\mathbf{W}^{-1}(f)$  can be expressed as

$$\mathbf{W}^{-1}(f) = \mathbf{\Psi}^{-1}(f) \mathbf{F} \mathbf{\Phi}^{-1}(f). \tag{4.76}$$

As shown in Fig. 4.7, the implementation of  $\text{IDTFT}\{\mathbf{W}^{-1}(f)\}$  can be done in three stages, where each stage corresponds to one of the three matrices in (4.76). The first stage involves filtering the sequences  $\{c_l[m]\}$ ,  $l = 1, \dots, p$ , using the set of filters

$$\phi_l[m] \triangleq \text{IDTFT}\{\Phi_l^{-1}(f)\}[m] = \text{IDTFT}\left\{(-1)^{l-1} e^{-j2\pi f(l-1)\frac{T}{p}}\right\}[m], m \in \mathbb{Z} \tag{4.77}$$


 Figure 4.7.: Implementation of  $\mathbf{W}^{-1}(f)$ .

which return the sequences  $\{c'_l[m]\}$ ,  $m = 1, \dots, p$ . Next, multiplication with the DFT matrix  $\mathbf{F}$  can be efficiently implemented by applying the Fast Fourier Transform (FFT) to  $\{c'_l[m]\}$ . Note that the entries of  $\mathbf{F}$  are not frequency dependent and the filters of Fig. 4.6 reduce to simple multiplications by complex constants. The resulting sequences are  $\{c''_l[m]\}$ ,  $l = 1, \dots, p$ . Finally, the third step involves filtering  $\{c''_l[m]\}$  using the set of filters

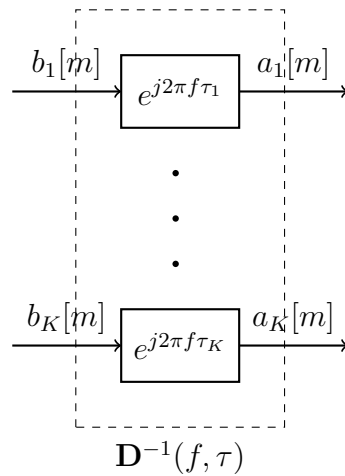
$$\begin{aligned} \psi_l[m] &\triangleq \text{IDTFT} \{ \Psi_l^{-1}(f) \} [m], m \in \mathbb{Z} \\ &= \text{IDTFT} \left\{ T \left[ G \left( f + \frac{1}{T}(l - p/2 - 1) \right) S^* \left( f + \frac{1}{T}(l - p/2 - 1) \right) \right]^{-1} \right\} [m] \end{aligned} \quad (4.78)$$

which get the desired sequences  $\{d_l[m]\}$ ,  $l = 1, \dots, p$  and compensate for the non-flatness of the frequency responses of the prototype pulse and the impulse response of the LPF. For each  $m$ , the vector  $\mathbf{d}[m]$  is given by  $[d_1[m] \ d_2[m] \ \dots \ d_p[m]]^T$ .

Implementations of  $\mathbf{N}^\dagger(\tau)$  and  $\mathbf{D}^{-1}(f, \tau)$  are simpler, since  $\mathbf{N}^\dagger(\tau)$  does not depend on frequency and  $\mathbf{D}^{-1}(f, \tau)$  is a diagonal matrix. In the latter case, each output sequence  $\{a_k[m]\}$  is obtained only from the corresponding sequences  $\{b_k[m]\}$ , filtered with a filter whose frequency response is given by  $e^{j2\pi f\tau_k}$ ,  $k = 1, \dots, K$ . Note that the frequency response of the  $k$ th filter is the  $k$ th element of the diagonal of  $\mathbf{D}^{-1}(f, \tau)$ . The implementation of the block  $\mathbf{D}^{-1}(f, \tau)$  is reported in Fig. 4.8.

## 4.7. Sufficient Conditions for Identifiability

In this section we derive the conditions under which the identification of the system (4.3) is possible and produces a unique solution. These conditions are expressed in terms of equivalent requirements on the time-bandwidth product  $\mathcal{TW}$  of the input signal  $x(t)$  in (4.4).


 Figure 4.8.: Implementation of  $\mathbf{D}^{-1}(f, \tau)$ .

First of all, we require that matrix  $\mathbf{W}(f)$  in (4.26) is invertible. We make use of the decomposition in (4.75) that expresses  $\mathbf{W}(f)$  as the product of matrices  $\mathbf{\Phi}(f)$ ,  $\mathbf{F}^H$  and  $\mathbf{\Psi}(f)$ . It is clear that  $\mathbf{W}(f)$  is invertible if the three matrices are separately invertible.

By construction, both the diagonal matrix  $\mathbf{\Phi}(f)$ , whose diagonal entries are defined in (4.72), and  $\mathbf{F}^H$ , whose entries are defined in (4.73), are always invertible. The invertibility of the diagonal matrix  $\mathbf{\Psi}(f)$  is ensured if its entries in (4.74) are nonzero, i.e., if the continuous-time Fourier transforms of the prototype pulse  $g(t)$  and of the impulse response  $s^*(-t)$  of the LPF are nonzero within the spectral band  $\mathcal{F}$  in (4.11). This requirement leads to the following conditions:

*Condition 1:*  $a \leq |G(f)| \leq b, \forall f \in \mathcal{F}$ , for some constants  $a > 0$  and  $0 < b < \infty$ .

*Condition 2:*  $c \leq |S(f)| \leq d, \forall f \in \mathcal{F}$ , for some constants  $c > 0$  and  $0 < d < \infty$ .

Condition 2 can be made always satisfied under a proper design of the LPF  $S^*(f)$ . Condition 1 requires that the bandwidth  $\mathcal{W}$  of the prototype pulse  $g(t)$  is larger than the bandwidth of  $s^*(-t)$ , supported in  $\mathcal{F}$ , i.e.,

$$\mathcal{W} \geq \frac{p}{T}. \quad (4.79)$$

If condition (4.79) is not satisfied, then  $|S(f)|$  would be zero for all the frequencies in  $\mathcal{F}$  not included in  $\mathcal{W}$ .

We provide now conditions for the unique recovery of the unknown delays  $\tau_k$ . To this end, equation (4.31) must admit a unique solution on the set of delays  $\tau_k$ , given by the ESPRIT procedure described before, and on the set of vectors  $\mathbf{b}[m]$ ,  $m \in \mathbb{Z}$ . We let

$$\mathbf{d}[\Lambda] = \{\mathbf{d}[m], m \in \mathbb{Z}\} \quad (4.80)$$

$$\mathbf{b}[\Lambda] = \{\mathbf{b}[m], m \in \mathbb{Z}\} \quad (4.81)$$

denote the set of vectors  $\mathbf{d}[m]$  and  $\mathbf{b}[m]$ , respectively. Using this notation, equation (4.31) becomes compactly

$$\mathbf{d}[\Lambda] = \mathbf{N}(\tau)\mathbf{b}[\Lambda]. \quad (4.82)$$

We follow the analysis carried out in [22] and make use of the following theorem, whose proof is given in the appendix, to find the conditions for the uniqueness of the solution in (4.82).

**Theorem 6.** *If  $(\bar{\tau}, \bar{\mathbf{b}}[\Lambda])$  is a solution of equation (4.82), if*

$$p > 2K - \dim(\text{span}(\bar{\mathbf{b}}[\Lambda])) \quad (4.83)$$

*and if*

$$\dim(\text{span}(\bar{\mathbf{b}}[\Lambda])) \geq 1 \quad (4.84)$$

*then  $(\bar{\tau}, \bar{\mathbf{b}}[\Lambda])$  is the unique solution of (4.82).*

The notation  $\text{span}(\bar{\mathbf{b}}[\Lambda])$  indicates the minimal dimension subspace containing the unknown vector set  $\bar{\mathbf{b}}[\Lambda]$ . The condition (4.84) is needed to avoid the case where  $\bar{\mathbf{b}}[\Lambda] = \mathbf{0}$ .

Theorem 6 suggests that a unique recovery of the set of delays  $\tau_k$  is guaranteed through a proper selection of the parameter  $p$ . In particular, since  $\dim(\text{span}(\bar{\mathbf{b}}[\Lambda]))$  is a positive number, the condition

$$p \geq 2K \quad (4.85)$$

is a sufficient condition for a unique recovery of the unknown delays.

Combining condition (4.85) with (4.79), we find the sufficient condition on the bandwidth of the input signal as

$$\mathcal{W} \geq \frac{2K}{T}. \quad (4.86)$$

Finally, the recovery of the Doppler-shifts  $\nu_k$  and of the attenuation factors  $\alpha_k$  is unique if simply the number  $N$  of pulses of the input signal  $x(t)$  is not less than 2, i.e.,

$$N \geq 2. \quad (4.87)$$

In fact, from equations (4.69) and (4.70) only 2 samples of vectors  $\mathbf{a}_k[m]$  suffices to recover the parameters  $\alpha_k$  and  $\nu_k$ , since 2 equations are needed to recover 2 unknowns in a linear system of equations.

Recalling from (4.5) that the time support  $\mathcal{T}$  of  $x(t)$  is given by  $N$  times the time support  $T$  of the prototype pulse  $g(t)$ , the sufficient condition on the time support  $\mathcal{T}$  is

$$\mathcal{T} = NT \geq 2T. \quad (4.88)$$

Combining (4.86) and (4.88), the sufficient conditions on the time-bandwidth product of the input signal that provide a unique identification of the system (4.3) are given by

$$\mathcal{T}\mathcal{W} \geq 4K \quad (4.89)$$

already presented in Theorem 4.

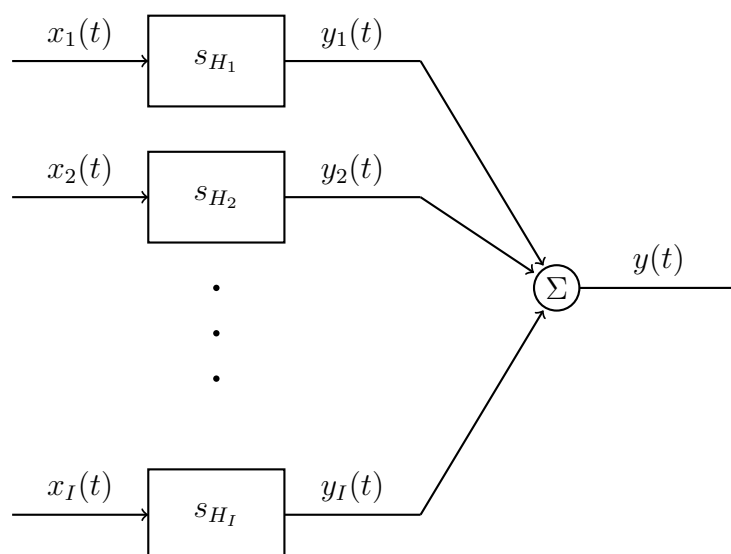


Figure 4.9.: System with multiple inputs.

## 4.8. Identification for Multiple Inputs

In the previous sections, the identification of the system (4.3) with one input signal was analyzed. The extension to a multiple inputs system is considered here. As shown in Fig. 4.9, the response of the system is given by the sum of the responses  $y_i(t)$  of the single subsystems, i.e.,

$$y(t) = \sum_{i=1}^I y_i(t) = \sum_{i=1}^I \int_{\tau} \int_{\nu} s_{H_i}(\tau, \nu) x_i(t - \tau) e^{j2\pi\nu t} d\nu d\tau \quad (4.90)$$

where  $I$  is the number of inputs and  $s_{H_i}(\tau, \nu)$  is the spreading function corresponding to the  $i$ th subsystem with input  $x_i(t)$  and output  $y_i(t)$ . Assuming that each spreading function  $s_{H_i}(\tau, \nu)$  has the same form as in (4.2), i.e.,

$$s_{H_i}(\tau, \nu) = \sum_{j=1}^{K_i} \alpha_{i,j} \delta(\tau - \tau_{i,j}) \delta(\nu - \nu_{i,j}) \quad (4.91)$$

the input-output relation in (4.90) becomes

$$y(t) = \sum_{i=1}^I \sum_{j=1}^{K_i} \alpha_{i,j} x_i(t - \tau_{i,j}) e^{j2\pi\nu_{i,j} t} \quad (4.92)$$

where  $K_i$  represents the cardinality of the set of delays  $\tau_{i,j}$  and Doppler-shifts  $\nu_{i,j}$ , associated to the  $i$ th input. As shown for the system with one input, the parameters that completely describe the system are the triplets  $(\alpha_{i,j}, \tau_{i,j}, \nu_{i,j})$ , for  $i = 1, \dots, I$  and  $j = 1, \dots, K_i$ . The number of triplets is indicated with  $K = \sum_{i=1}^I K_i$ .

Now, suppose that the assumption A) is still valid for each subsystem  $s_{H_i}(\tau, \nu)$  and indicate with  $\tau_{max_i}$  and  $\nu_{max_i}$  the delay spread and the Doppler spread of the

$i$ th system, respectively. Define parameters  $\tau_{max}$  and  $\nu_{max}$  as

$$\tau_{max} \triangleq \max_i \tau_{max_i} \quad (4.93)$$

$$\nu_{max} \triangleq \max_i \nu_{max_i} \quad (4.94)$$

and consider that each system  $s_{H_i}(\tau, \nu)$  has delay spread  $\tau_{max}$  and Doppler spread  $\nu_{max}$ . The probing signals  $x_i(t)$  are chosen as

$$x_i(t) = x(t - (i - 1)\tau_{max}), \quad i = 1, \dots, I \quad (4.95)$$

where the signal  $x(t)$  is defined in (4.4). In other words, the first probing signal is exactly the signal in (4.4),  $x_1(t) = x(t)$ , while each other probing signal  $x_i(t)$  is obtained from the previous one  $x_{i-1}(t)$  according to a time shift of  $\tau_{max}$ ,  $x_i(t) = x_{i-1}(t - \tau_{max})$ , for  $i = 2, \dots, I$ .

Substituting the inputs (4.95) into (4.92) yields

$$y(t) = \sum_{i=1}^I \sum_{j=1}^{K_i} \alpha_{i,j} x(t - (i - 1)\tau_{max} - \tau_{i,j}) e^{j2\pi\nu_{i,j}t} \quad (4.96)$$

$$= \sum_{i=1}^I \sum_{j=1}^{K_i} \alpha_{i,j} x(t - \tilde{\tau}_{i,j}) e^{j2\pi\nu_{i,j}t} \quad (4.97)$$

$$= \sum_{k=1}^K \alpha_k x(t - \tilde{\tau}_k) e^{j2\pi\nu_k t} \quad (4.98)$$

where in (4.97)  $\tilde{\tau}_{i,j} \triangleq (i - 1)\tau_{max} + \tau_{i,j}$ . Notice that the relation (4.98) in the case of multiple inputs has exactly the same form as the relation (4.3) for a single input. This suggests to use the same procedure described before to recover the triplets  $(\alpha_k, \tilde{\tau}_k, \nu_k)$ , considering  $y(t)$  in (4.98) as the response of a single system  $\tilde{s}_H$ , with the single input  $x(t)$ . In fact, relations (4.96)-(4.98) can describe the response of a system with spreading function given by

$$\tilde{s}_H(\tau, \nu) = \sum_{i=1}^I \sum_{j=1}^{K_i} \alpha_{i,j} \delta(\tau - (i - 1)\tau_{max} - \tau_{i,j}) \delta(\nu - \nu_{i,j}) \quad (4.99)$$

$$= \sum_{i=1}^I \sum_{j=1}^{K_i} \alpha_{i,j} \delta(\tau - \tilde{\tau}_{i,j}) \delta(\nu - \nu_{i,j}) \quad (4.100)$$

$$= \sum_{k=1}^K \alpha_k \delta(\tau - \tilde{\tau}_k) \delta(\nu - \nu_k). \quad (4.101)$$

Note that (4.101) is the same as (4.2).

An example of the spreading function  $\tilde{s}_H$  is depicted in Fig. 4.10, starting from a system with  $I = 3$  inputs. According to the choice (4.95), the equivalent system can be described by the single spreading function  $\tilde{s}_H$ , obtained from the subsystems  $s_{H_i}$  mapping their supports  $[0, \tau_{max}] \times [0, \nu_{max}]$  one after the other in the  $(\tau, \nu)$ -plane. The support of  $\tilde{s}_H$  is given by  $[0, I\tau_{max}] \times [0, \nu_{max}] = [0, \tilde{\tau}_{max}] \times [0, \nu_{max}]$ .

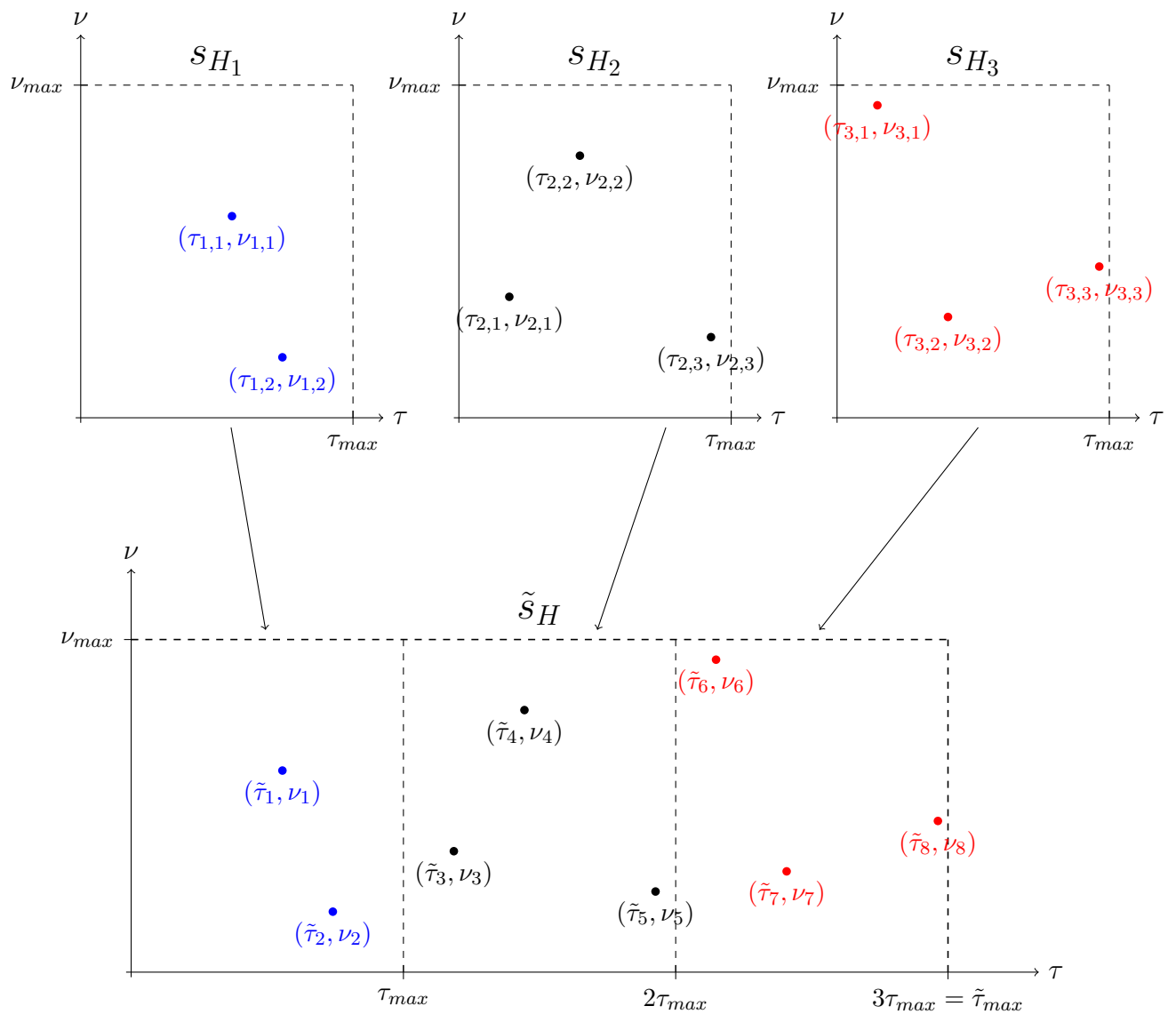


Figure 4.10.: Equivalent spreading function of a system with  $I = 3$  inputs.



Given that the assumptions B) and C) are satisfied for each single subsystem, then, if the number of inputs  $I$  is approximately small, they still remain valid also for the equivalent system. In this case, the procedure described in sections 4.3 and 4.5 is used for recovering the triplets  $(\alpha_k, \tilde{\tau}_k, \nu_k)$  in (4.98).

Finally, the original delays  $\tau_{i,j}$  are recovered from  $\tau_k$  or, equivalently, from  $\tilde{\tau}_{i,j}$  in (4.97), as

$$\begin{array}{ll} \text{if} & \tilde{\tau}_{i,j} \in [(i-1)\tau_{max}, i\tau_{max}] \\ \text{then} & \tau_{i,j} = \tilde{\tau}_{i,j} - (i-1)\tau_{max} \end{array} \quad (4.102)$$



## Proof of Theorem 5

To prove the theorem, we express the smoothed covariance matrix (4.61) in a proper form. To this end we follow the same steps already seen for the ESPRIT algorithm, in section 4.5, considering now  $p - K$  subsystems given by the sequences in the subvectors  $\mathbf{d}_l[m]$ ,  $l = 1, \dots, p - K$ . Indicating with  $\mathbf{N}_1(\tau)$  the  $K \times K$  matrix whose rows are the first  $K$  rows of  $\mathbf{N}(\tau)$ , the subvector  $\mathbf{d}_1[m]$  in (4.60) can be written as

$$\mathbf{d}_1[m] = \mathbf{N}_1(\tau)\mathbf{b}[m]. \quad (\text{A.1})$$

According to the rotational invariance property, the subvectors in (4.60) are given by

$$\mathbf{d}_l[m] = \mathbf{N}_1(\tau)\mathbf{R}^{l-1}(\tau)\mathbf{b}[m], \quad l = 1, \dots, p - K \quad (\text{A.2})$$

where the  $K \times K$  diagonal matrix  $\mathbf{R}(\tau)$  is defined in (4.37). For simplicity we omit the  $\tau$  dependence and indicate  $\mathbf{R}(\tau)$  simply with  $\mathbf{R}$ . Substituting (A.2) into (4.61),  $\bar{\mathbf{R}}_{dd}$  becomes

$$\begin{aligned} \bar{\mathbf{R}}_{dd} &= \frac{1}{p - K} \sum_{l=1}^{p-K} \sum_{m \in \mathbb{Z}} (\mathbf{N}_1(\tau)\mathbf{R}^{l-1}\mathbf{b}[m]) (\mathbf{N}_1(\tau)\mathbf{R}^{l-1}\mathbf{b}[m])^H \\ &= \frac{1}{p - K} \sum_{l=1}^{p-K} \sum_{m \in \mathbb{Z}} (\mathbf{N}_1(\tau)\mathbf{R}^{l-1}) (\mathbf{b}[m]\mathbf{b}^H[m]) (\mathbf{N}_1(\tau)\mathbf{R}^{l-1})^H \\ &= \frac{1}{p - K} \sum_{l=1}^{p-K} (\mathbf{N}_1(\tau)\mathbf{R}^{l-1}) \left( \sum_{m \in \mathbb{Z}} \mathbf{b}[m]\mathbf{b}^H[m] \right) (\mathbf{N}_1(\tau)\mathbf{R}^{l-1})^H \\ &= \mathbf{N}_1(\tau) \left( \frac{1}{p - K} \sum_{l=1}^{p-K} \mathbf{R}^{l-1} \mathbf{R}_{bb} \mathbf{R}^{*l-1} \right) \mathbf{N}_1^H(\tau) \\ &= \mathbf{N}_1(\tau) \bar{\mathbf{R}}_{bb} \mathbf{N}_1^H(\tau) \end{aligned} \quad (\text{A.3})$$

where  $\bar{\mathbf{R}}_{bb}$  is the *modified covariance matrix* of the sequences in  $\mathbf{b}[m]$ . Note that from definition in (4.37),  $\mathbf{R}^H = \mathbf{R}^*$ . Since  $\mathbf{N}_1(\tau)$  has full rank by definition of Vandermonde matrix, from (A.3)  $\bar{\mathbf{R}}_{dd}$  and  $\bar{\mathbf{R}}_{bb}$  have the same rank. Our task is then to prove that  $\bar{\mathbf{R}}_{bb}$  has rank  $K$ .

Matrix  $\bar{\mathbf{R}}_{bb}$  can be written as

$$\bar{\mathbf{R}}_{bb} = \begin{bmatrix} \mathbf{I}_K & \mathbf{R} & \cdots & \mathbf{R}^{p-K-1} \end{bmatrix} \text{diag} \left\{ \frac{1}{p-K} \mathbf{R}_{bb}, \dots, \frac{1}{p-K} \mathbf{R}_{bb} \right\} \begin{bmatrix} \mathbf{I}_K \\ \mathbf{R}^* \\ \vdots \\ \mathbf{R}^{*(p-K-1)} \end{bmatrix} \quad (\text{A.4})$$

or, equivalently, as

$$\bar{\mathbf{R}}_{bb} = \begin{bmatrix} \mathbf{C} & \mathbf{RC} & \cdots & \mathbf{R}^{p-K-1}\mathbf{C} \end{bmatrix} \begin{bmatrix} \mathbf{C}^H \\ \mathbf{C}^H \mathbf{R}^* \\ \vdots \\ \mathbf{C}^H \mathbf{R}^{*(p-K-1)} \end{bmatrix} \quad (\text{A.5})$$

where the  $K \times K$  matrix  $\mathbf{C}$  is defined such that

$$\mathbf{C}\mathbf{C}^H \triangleq \frac{1}{p-K} \mathbf{R}_{bb}. \quad (\text{A.6})$$

Indicating with  $\mathbf{G} \triangleq \begin{bmatrix} \mathbf{C} & \mathbf{RC} & \cdots & \mathbf{R}^{p-K-1}\mathbf{C} \end{bmatrix}$ , relation (A.5) becomes  $\bar{\mathbf{R}}_{bb} = \mathbf{G}\mathbf{G}^H$ . Now, if matrix  $\mathbf{G}$  has full rank  $K$ , also  $\bar{\mathbf{R}}_{bb}$  has rank  $K$ . We prove then that  $\mathbf{G}$  has rank  $K$ .

By a permutation of the columns of  $\mathbf{G}$ , we construct matrix  $\tilde{\mathbf{G}}$  that has the same rank as  $\mathbf{G}$ .

$$\tilde{\mathbf{G}} = \begin{bmatrix} c_{11}\mathbf{g}_1 & c_{12}\mathbf{g}_1 & \cdots & c_{1K}\mathbf{g}_1 \\ \vdots & \vdots & \cdots & \vdots \\ c_{K1}\mathbf{g}_K & c_{K2}\mathbf{g}_K & \cdots & c_{KK}\mathbf{g}_K \end{bmatrix} \quad (\text{A.7})$$

where  $c_{ij}$  is the  $(i,j)$ th element of the matrix  $\mathbf{C}$  and  $\mathbf{g}_i$  is the  $(p-K)$ -length row vector defined as

$$\mathbf{g}_i = \left[ 1 \quad e^{-j\frac{2\pi}{T}\tau_i} \quad \cdots \quad e^{-j\frac{2\pi}{T}(p-K-1)\tau_i} \right], \quad i = 1, \dots, K. \quad (\text{A.8})$$

From (A.7), it is easy to see that matrix  $\tilde{\mathbf{G}}$  has full rank if each row of matrix  $\mathbf{C}$  is nonzero and if the vectors  $\mathbf{g}_i$  are linearly independent. The former follows from the definition in (A.6). If the  $k$ th row of  $\mathbf{C}$  is zero, the  $k$ th element in the main diagonal of  $\mathbf{R}_{bb}$  is zero, which implies that the sequence  $\{b_k[m]\}$  in  $\mathbf{b}[m]$  has zero energy, which is not possible. The latter follows from the assumption that  $p-K \geq K$ . In this case, the Vandermonde matrix whose rows are the vectors  $\mathbf{g}_i$  has full rank ( $K$ ) by definition, i.e.,  $\mathbf{g}_i$  are linearly independent.

## Proof of Theorem 6

To prove the theorem, we make use of the following theorem, developed in [26].

**Theorem 7.** *Consider the following system of equations*

$$\mathbf{D} = \mathbf{N}(\tau)\mathbf{B} \tag{B.1}$$

where the  $p \times K$  matrix  $\mathbf{N}(\tau)$  is the Vandermonde matrix already defined in (4.14), while the  $K \times r$  matrix  $\mathbf{B}$  and the  $p \times r$  matrix  $\mathbf{D}$  collect  $r$  vectors  $\mathbf{b}[m]$  and  $\mathbf{d}[m]$ , respectively, in their columns. Parameter  $r$  is defined as the rank of  $\mathbf{B}$ .

$$r \triangleq \text{rank } \mathbf{B} \tag{B.2}$$

The recovery of matrix  $\mathbf{N}(\tau)$  from the measurement vectors in  $\mathbf{D}$  has a unique solution if the following condition holds

$$p > 2K - r. \tag{B.3}$$

*Proof.* In order to prove the theorem, we shall show that if condition (B.3) is true, then

$$\begin{cases} \mathbf{D} = \mathbf{N}(\tau)\mathbf{B} \\ \mathbf{D} \neq \mathbf{N}(\tau')\mathbf{B}' \end{cases} \tag{B.4}$$

for every set of  $K$  delays  $\tau'_k$ , different from the set of delays  $\tau_k$  at least for one value, and for every set signals in  $\mathbf{B}'$ .

We distinguish two different cases, according to the number of pairs of common delays between the sets  $\tau_k$  and  $\tau'_k$ . Let  $d$  indicate the number of pairs such that  $\tau_i = \tau'_j$  for some values of  $i, j = 1, \dots, K$ . Then,

$$\begin{cases} 0 \leq d < 2K - p & \Rightarrow \text{case 1) } \\ 2K - p \leq d < p & \Rightarrow \text{case 2) } \end{cases} \tag{B.5}$$

Note that the first case includes also the situation in which  $\tau_i \neq \tau'_j$  for all the values of  $i, j = 1, \dots, K$ , since this corresponds to  $d = 0$ .

*Case 1)* The system (B.4) can be expressed according to the following equation

$$\mathbf{N}(\tau)\mathbf{B} - \mathbf{N}(\tau')\mathbf{B}' \neq \mathbf{0} \tag{B.6}$$

or, in the matrix form,

$$\begin{bmatrix} \mathbf{N}(\tau) & \mathbf{N}(\tau') \end{bmatrix} \begin{bmatrix} \mathbf{B} \\ -\mathbf{B}' \end{bmatrix} \neq \mathbf{0}. \quad (\text{B.7})$$

By definition of the null space of a matrix, relation (B.7) holds if no columns of  $\begin{bmatrix} \mathbf{B} \\ -\mathbf{B}' \end{bmatrix}$  belong to the null space of  $[\mathbf{N}(\tau) \ \mathbf{N}(\tau')]$  or, in other words, if the nullity of the matrix  $[\mathbf{N}(\tau) \ \mathbf{N}(\tau')]$  is strictly smaller than the rank of the matrix  $\begin{bmatrix} \mathbf{B} \\ -\mathbf{B}' \end{bmatrix}$ .

Denote

$$\zeta = \text{null} \begin{bmatrix} \mathbf{N}(\tau) & \mathbf{N}(\tau') \end{bmatrix}. \quad (\text{B.8})$$

According to the Rank-Nullity Theorem,

$$\zeta = 2K - \text{rank} \begin{bmatrix} \mathbf{N}(\tau) & \mathbf{N}(\tau') \end{bmatrix}. \quad (\text{B.9})$$

Since  $\mathbf{N}(\tau)$  and  $\mathbf{N}(\tau')$  are Vandermonde matrices, the columns of matrix  $[\mathbf{N}(\tau) \ \mathbf{N}(\tau')]$  are linearly independent except for  $d$  repeated columns. Then,

$$\text{rank} \begin{bmatrix} \mathbf{N}(\tau) & \mathbf{N}(\tau') \end{bmatrix} = \min\{p, 2K - d\}. \quad (\text{B.10})$$

Now, the first case in (B.5) implies that  $p < 2K - d$  and, consequently, the rank in (B.10) is equal to  $p$ . Relation (B.9) becomes then

$$\zeta = 2K - p. \quad (\text{B.11})$$

Now, denoting

$$\nu = \text{rank} \begin{bmatrix} \mathbf{B} \\ -\mathbf{B}' \end{bmatrix} \quad (\text{B.12})$$

it is easy to see that  $\nu$  can not be smaller than the rank of  $\mathbf{B}$ , which implies that

$$\nu \geq r. \quad (\text{B.13})$$

Combining (B.11), (B.3) and (B.13) yields

$$\zeta = 2K - p < r \leq \nu \Rightarrow \zeta < \nu \quad (\text{B.14})$$

which concludes the proof of the first case.

*Case 2)* In this case, the system (B.4) is equivalent to the following expression

$$\mathbf{N}(\tau)\hat{\mathbf{B}} - \hat{\mathbf{N}}(\tau')\hat{\mathbf{B}}' \neq \mathbf{0} \quad (\text{B.15})$$

or, in matrix form, to

$$\begin{bmatrix} \mathbf{N}(\tau) & \hat{\mathbf{N}}(\tau') \end{bmatrix} \begin{bmatrix} \hat{\mathbf{B}} \\ -\hat{\mathbf{B}}' \end{bmatrix} \neq \mathbf{0} \quad (\text{B.16})$$

where the  $p \times (K - d)$  matrix  $\hat{\mathbf{N}}(\tau')$  is obtained from  $\mathbf{N}(\tau')$  by deleting the  $d$  columns equal to the columns of  $\mathbf{N}(\tau)$ , the  $(K - d) \times r$  matrix  $\hat{\mathbf{B}}'$  is obtained from  $\mathbf{B}'$  by deleting the  $d$  rows corresponding to the deleted columns in  $\mathbf{N}(\tau')$  and the  $K \times r$

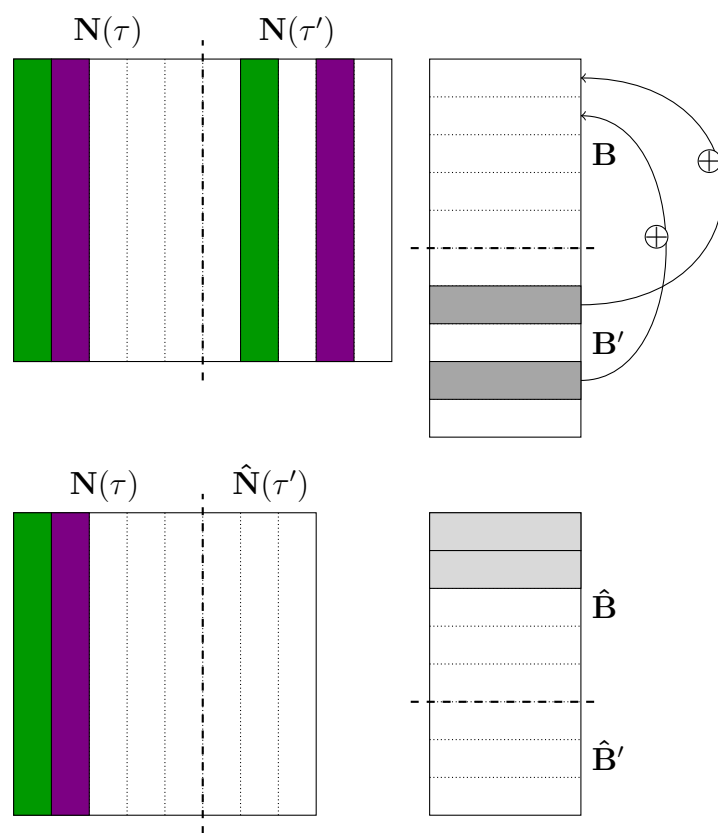


Figure B.1.: Graphical representation of the equivalence between (B.7) and (B.16).

matrix  $\hat{\mathbf{B}}$  is obtained from  $\mathbf{B}$  by adding the deleted rows of  $\hat{\mathbf{B}}'$  to the corresponding ones in  $\mathbf{B}$ . A graphical example that represents how to obtain (B.16) from (B.7) is depicted in Fig. B.1.

To prove the theorem, it suffices to show that the nullity of the matrix  $\begin{bmatrix} \mathbf{N}(\tau) & \hat{\mathbf{N}}(\tau') \end{bmatrix}$  is strictly smaller than the rank of the matrix  $\begin{bmatrix} \hat{\mathbf{B}} \\ -\hat{\mathbf{B}}' \end{bmatrix}$ . Following the same steps as before, denote

$$\hat{\zeta} = \text{null} \begin{bmatrix} \mathbf{N}(\tau) & \hat{\mathbf{N}}(\tau') \end{bmatrix}. \quad (\text{B.17})$$

According to the Rank-Nullity Theorem,

$$\hat{\zeta} = 2K - d - \text{rank} \begin{bmatrix} \mathbf{N}(\tau) & \hat{\mathbf{N}}(\tau') \end{bmatrix}. \quad (\text{B.18})$$

As before,

$$\text{rank} \begin{bmatrix} \mathbf{N}(\tau) & \hat{\mathbf{N}}(\tau') \end{bmatrix} = \min\{p, 2K - d\} = 2K - d \quad (\text{B.19})$$

since the second case in (B.5) implies that  $p \geq 2K - d$ . Relation (B.18) becomes then

$$\hat{\zeta} = 0. \quad (\text{B.20})$$

Now, denoting

$$\hat{\nu} = \text{rank} \begin{bmatrix} \hat{\mathbf{B}} \\ -\hat{\mathbf{B}}' \end{bmatrix} \quad (\text{B.21})$$

it is easy to see that  $\hat{\nu} \geq 1$ ; combining this result with (B.20) yields

$$\hat{\zeta} < \hat{\nu} \quad (\text{B.22})$$

which concludes the proof.  $\square$

Returning back to the proof of the theorem 6, to prove that the solution  $(\bar{\tau}, \bar{\mathbf{b}}[\Lambda])$  is unique for the infinite set of equations in (4.82) we need to express the system

$$\bar{\mathbf{d}}[\Lambda] = \mathbf{N}(\bar{\tau})\bar{\mathbf{b}}[\Lambda] \quad (\text{B.23})$$

according to the form (B.1) of the previous theorem. To this end, we denote  $r \triangleq \dim(\text{span}(\bar{\mathbf{b}}[\Lambda]))$ . From (4.84)  $r \geq 1$ . Then we can chose  $r$  times  $m \in \Lambda$  that form a finite subset  $\tilde{\Lambda} \subset \Lambda$  such that

$$\dim(\text{span}(\bar{\mathbf{b}}[\tilde{\Lambda}])) = r. \quad (\text{B.24})$$

Define now the matrices  $\mathbf{D}$  and  $\mathbf{B}$  whose columns are the vector sets  $\bar{\mathbf{d}}[\tilde{\Lambda}]$  and  $\bar{\mathbf{b}}[\tilde{\Lambda}]$ , respectively. Then, a finite representation of (B.23) is given by

$$\mathbf{D} = \mathbf{N}(\bar{\tau})\mathbf{B}. \quad (\text{B.25})$$

From its construction, the rank of matrix  $\mathbf{B}$  is  $r$  and (4.83) is equivalent to  $p > 2K - r$ . According to the theorem 7, the solution  $(\bar{\tau}, \mathbf{B})$  is the unique solution to (B.25).

Since the set of delays  $\bar{\tau}$  is the unique solution to (B.25), it is also the unique solution to the infinite set of equations in (B.23). Finally, since  $\mathbf{N}^\dagger(\bar{\tau})\mathbf{N}(\bar{\tau}) = \mathbf{I}_K$  because of the assumption  $p \geq 2K$ , for any set of vectors  $\bar{\mathbf{d}}[\Lambda]$ , if  $\bar{\mathbf{b}}[\Lambda]$  is a solution of (4.82), it is unique and given by

$$\bar{\mathbf{b}}[\Lambda] = \mathbf{N}^\dagger(\bar{\tau})\bar{\mathbf{d}}[\Lambda]. \quad (\text{B.26})$$



# Bibliography

- [1] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, June 2005.
- [2] H. Bölcskei, from the lecture notes of *Fundamentals of Wireless Communication*, Dec. 2009.
- [3] A. M. Sayeed and B. Aazhang, “Joint multipath-Doppler diversity in mobile wireless communications”, *IEEE Trans. Commun.*, Jan. 1999.
- [4] C. E. Cook and M. Bernfeld, *Radar Signals, An Introduction to Theory and Applications*, Academic Press, 1967.
- [5] L. G. Weiss, “Wavelets and wideband correlation processing”, *IEEE Signal Process. Mag.*, Jan. 1994.
- [6] L. S. Sibil, L. G. Weiss and T. L. Dixon, “Characterization of stochastic propagation and scattering via Gabor and wavelet transforms”, *J. Comput. Acoust.*, 1994.
- [7] T. Kailath, “Measurements on time-variant communication channels”, *IRE Trans. Inf. Theory*, Sept. 1962.
- [8] P. Bello, “Measurements of random time-variant linear channels”, *IEEE Trans. Inf. Theory*, July 1969.
- [9] D. Slepian, “On bandwidth”, *Proc. IEEE*, March 1976.
- [10] Y. Jiang and A. Papandreou-Suppappola, “Discrete time-scale characterization of wideband time-varying system”, *IEEE Trans. Signal Process.*, Apr. 2006.
- [11] L. Cohen, “The scale representation”, *IEEE Trans. Signal Process.*, Dec. 1993.
- [12] H. Sundaram, S. D. Joshi and R. K. P. Bhatt, “Scale periodicity and its sampling theorem”, *IEEE Trans. Signal Process.*, July 1997.
- [13] J. Bertrand, P. Bertrand and J. P. Ovarlez, “The mellin transform” in *The Transforms and Applications Handbook*, CRC Press, 1995.
- [14] L. Cohen, *Time-Frequency Analysis*, Prentice Hall, chapter 15, Dec. 1994.

- [15] R. Heckel and H. Bölcskei, “Compressive identification of linear operator”, preprint.
- [16] Y. Lu and M. Do, “A theory for sampling signals from a union of subspaces”, *IEEE Trans. Signal Process.*, June 2008.
- [17] M. Herman and T. Strohmer, “Pseudodifferential operators and Banach algebras in mobile communications”, *Appl. Comput. Harmon. Anal.*, March 2006.
- [18] M. A. Davenport and M. B. Wakin, “Analysis of orthogonal matching pursuit using the restricted isometry property”, *IEEE Trans. Inf. Theory*, Sept. 2010.
- [19] A. C. Gilbert, M. J. Strauss, and J. A. Tropp, “Algorithms for simultaneous sparse approximation. Part I: greedy pursuit”, special issue on sparse approximations in signal and image processing of *EURASIP J. Signal Processing*, April 2006.
- [20] A. Huang, G. Guan, Q. Wan and A. Mehdodniya, “A block orthogonal matching pursuit algorithm based on sensing dictionary”, *International Journal of the Physical Sciences*, March 2011.
- [21] G. E. Pfander, “Measurement of time-varying multiple-input multiple-output channels”, *Appl. Comp. Harmon. A.*, May 2008.
- [22] K. Gedalyahu and Y. C. Eldar, “Time-delay estimation from low-rate samples: a union of subspaces approach”, *IEEE Trans. Signal Process.*, June 2010.
- [23] W. U. Bajwa, K. Gedalyahu and Y. C. Eldar, “Identification of parametric underspread linear systems and super-resolution radar”, *IEEE Trans. Signal Process.*, June 2011.
- [24] R. Roy and T. Kailath, “ESPRIT-estimation of signal parameters via rotational invariance techniques”, *IEEE Trans. Acoust., Speech, Signal Process.*, July 1989.
- [25] T. J. Shan, M. Wax and T. Kailath, “On spatial smoothing for direction-of-arrival estimation for coherent signals”, *IEEE Trans. Acoust., Speech, Signal Process.*, Aug. 1985.
- [26] M. Wax and I. Ziskind, “On unique localization of multiple sources by passive sensor arrays”, *IEEE Trans. Acoust., Speech, Signal Process.*, July 1989.
- [27] A. F. Molisch, *Wireless Communications*, Appendix chapter 8, second edition, John Wiley & Sons, 2011.

# Acknowledgement

Prima di tutto voglio ringraziare i miei genitori, a cui dedico questa tesi, che mi hanno sempre aiutata in tutti i modi possibili, non facendomi davvero mancare nulla. Il GRAZIE più sentito va a Sere perchè mi conosce più di chiunque altro, sa come esultare e ridere con me nei momenti di gioia, sa come caricarmi nei momenti di tensione e sa come confortarmi nei momenti di sconforto.

Grazie al mio Albi che riesce a starmi sempre vicino nonostante la distanza *europea*; grazie per la pazienza, l'ironia, la telepatia e le incomprensioni che ci caratterizzano.

Un grande ringraziamento è rivolto al mio relatore che mi ha permesso di compiere l'esperienza di tesi al politecnico di Zurigo, la più bella esperienza mai vissuta. I thank Prof. Böleskei for making this project possible. Special thanks go to Reinhard for supervising this project, for his patience and assistance. Thanks also to Céline, Graeme, David and Michael for making my experience unique.

Grazie a tutte le persone che mi sono vicine e che hanno fatto il tifo per me. Grazie ai nonni, agli zii e ai miei bellissimi cugini. Grazie a Nadia e Franco per tutto l'affetto dimostrato da sempre. Grazie alla carissima Stefi che sento vicina nonostante la lontananza. Grazie a Claudio, Mariella e nonna Carla per l'infinita gentilezza. Grazie alla nonna Agnese per essere sempre presente con i moderni sms.

Grazie a tutti i miei amici, che mi sopportano nonostante il mio ritardo congenito. Un grazie particolare a Matteo, Laura, Lucia, Michela, Chiara, Pozzo, Poli, Don Daniele, per i bellissimi momenti passati insieme. Grazie a Ire-Marco e Cate-Checco per la particolare amicizia che ci lega. Thanks to Simon, Petra, Santhiago and Nikos for sharing the last 6 months; it was a pleasure sharing the same office with you. Grazie ai miei compagni di corso ma soprattutto grazie a Paolo che mi ha permesso di superare il mio blocco informatico (perchè non sono come lui?).