# Aliasing is Good for You: Joint Registration and Reconstruction for Super-Resolution

Patrick Vandewalle, *Member, IEEE,* Luciano Sbaiz, *Member, IEEE,*
Joos Vandewalle, *Fellow, IEEE,* and Martin Vetterli, *Fellow, IEEE*

**Abstract**

In many applications, the sampling frequency is limited by the physical characteristics of the components: the pixel pitch, the rate of the A/D converter, etc. A low-pass filter is then often applied before the sampling operation to avoid aliasing. However, when multiple copies are available, it is possible to use the information that is inherently present in the aliasing to reconstruct a higher resolution signal. If the different copies have unknown relative offsets, this is a non-linear problem in the offsets and the signal coefficients. They are not easily separable in the set of equations describing the super-resolution problem. Thus, we perform joint registration and reconstruction from multiple unregistered sets of samples. We give a mathematical formulation for the problem when there are $M$ sets of $N$ samples of a signal that is described by $L$ expansion coefficients. We prove that the solution of the registration and reconstruction problem is generically unique if $MN \geq L + M - 1$. We describe two subspace-based methods to compute this solution. Their complexity is analyzed, and some heuristic methods are proposed. Finally, some numerical simulation results on one and two-dimensional signals are given to show the performance of these methods.

**Index Terms**

Aliasing, sampling, offset estimation, shift estimation, image registration, super-resolution.

Patrick Vandewalle, Luciano Sbaiz and Martin Vetterli are with the School of Computer and Communication Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland. Martin Vetterli is also with the Dept. of EECS, University of California, Berkeley. E-mail: {Patrick.Vandewalle, Luciano.Sbaiz, Martin.Vetterli}@epfl.ch

Joos Vandewalle is with the Dept. of Electrical Engineering (ESAT), Katholieke Universiteit Leuven, Belgium. E-mail: Joos.Vandewalle@esat.kuleuven.be

The work presented in this paper was supported by the National Competence Center in Research on Mobile Information and Communication Systems (NCCR-MICS), a center supported by the Swiss National Science Foundation under grant number 5005-67322. The work of Joos Vandewalle was supported by the Research Council KULeuven: GOA-Ambiorics and Center of Excellence on Optimization in Engineering; the Belgian Federal Science Policy Office : IUAP V-22. Parts of this work appeared earlier in [1] and [2].

**EDICS Category: DSP-RECO**

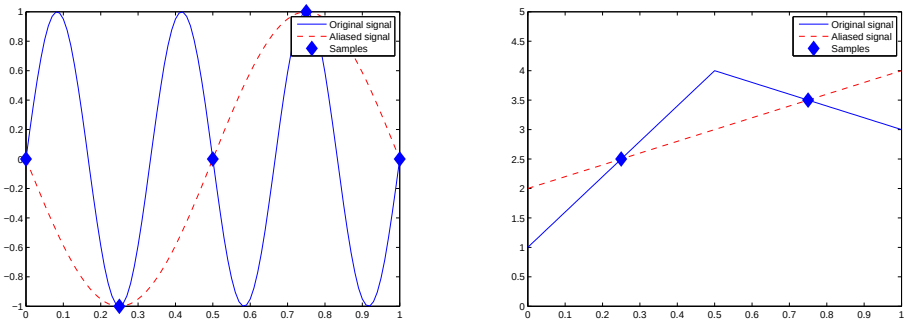
## I. Introduction

Aliasing, caused by undersampling of a signal, is generally considered as a nuisance and has to be avoided. To this effect, an anti-aliasing low-pass filter is often placed in front of the actual sampling operation, such that the sampled signal is not aliased. However, in super-resolution signal reconstruction, aliased components contain valuable high frequency components that can be used to recover a higher resolution reconstruction. In this paper, we study methods to reconstruct a signal, including its high frequency information, from multiple aliased sampled signals with relative offsets. The offsets are unknown, and need to be computed first. For the estimation of those offsets, we explicitly use the information available in the aliased part of the spectrum.

Applications can be found in various domains, such as super-resolution imaging. A camera is held manually while taking multiple images of the same scene. In satellite imaging, a satellite acquires images of approximately the same region at different moments. For one-dimensional signals, we find a similar setup in high rate A/D converters. A single high rate A/D converter is replaced by multiple lower rate converters that operate with small offsets. These low rate converters are hard to synchronize precisely, and can therefore be considered as converters with unknown offsets.

Aliasing is most often considered in a Fourier domain setup. It is seen as the replication of frequencies above half the sampling frequency in the base spectrum of the sampled signal. In this way, a frequency above half the sampling frequency is mapped onto a frequency below this limit after sampling and the two cannot be distinguished from their samples. In this paper, we will consider not only bandlimited functions, but more generally functions that belong to a finite-dimensional Hilbert space. In this case, the interpretation of frequency replication is not valid anymore, and the idea of aliasing needs to be broadened to this new signal space. We can then consider the sampling process as a projection of a continuous-time signal in a finite-dimensional Hilbert space onto the space of discrete-time signals. Aliasing occurs when the signal is undersampled, meaning that different signals from the Hilbert space are projected onto the same set of samples. The mapping between the continuous-time input signals and the discrete-time sampled signals is not a one-to-one mapping anymore. This interpretation is valid for signals in arbitrary finite-dimensional Hilbert spaces, including the bandlimited functions from the more restricted analysis. Some examples for different signal types are shown in Figure 1.

Sampling methods can be classified into different categories, according to the type of sampling that is performed (see Figure 2). A first important distinction is between uniform and non-uniform sampling methods. With uniform sampling, a signal is sampled periodically at constant time intervals.




(a) (b)

Fig. 1. Examples of aliasing for different types of signals. (a) For a bandlimited signal, input signals of different frequencies map into the same low frequency after sampling. (b) For a piecewise linear continuous signal, different input signals map into the same set of samples (and reconstruction) after sampling.

This is the standard setup, which is described by the Shannon-Nyquist sampling theorem [3]. An overview of some more recent results is given by Unser [4]. Among the non-uniform sampling methods, a distinction needs to be made between methods using known sampling instants [5], [6], and other methods where the sampling locations are unknown. If the sampling locations are unknown and completely arbitrary, the problem cannot be solved. This can be shown using a simple counting argument. Assume that the signal to be reconstructed has $L$ unknown parameters. For every additional sample, there is also an additional unknown (its location). Therefore the number of unknowns is always larger than the number of measurements, and this problem is unsolvable. However, for discrete signals, where the sampling locations can only take a finite number of values, a combinatorial solution can be found, as described by Marziliano and Vetterli [7].

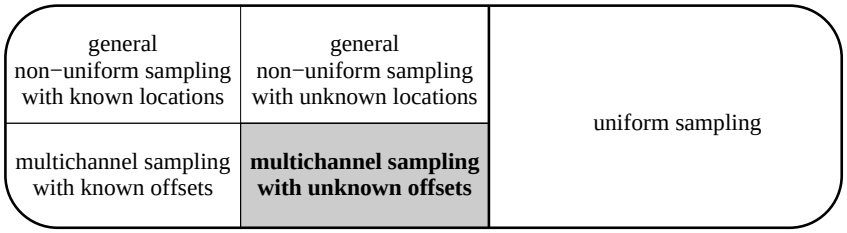



Fig. 2. Classification of sampling methods. Sampling methods can be divided into uniform and non-uniform methods. The non-uniform sampling methods can be subdivided depending on whether the locations are known and whether the samples are grouped in uniform sets with only unknown offsets. In this paper, we discuss multichannel sampling methods with unknown offsets.

An important subset of the non-uniform sampling methods is formed by multichannel sampling methods. In these methods, multiple sets of uniformly spaced samples are measured. Each of the sets of

samples is uniform, but the different sets have a non-uniform offset. Papoulis [8] described a solution for multichannel sampling with known sampling locations. He showed that a bandlimited signal can be perfectly reconstructed from $M$ sets of samples that are sampled at $1/M$ the Nyquist sampling rate. This result was extended by Unser and Zerubia in their generalized sampling approach [9], [10]. The problem with multiple sets and unknown sampling locations was solved for discrete-time signals by Marziliano and Vetterli [7]. In this paper, we will study the continuous-time case: multichannel sampling with unknown, real-valued offsets.

A typical application of such a setup is super-resolution imaging. The goal in super-resolution imaging is to reconstruct a high resolution image from a set of multiple images that are taken from almost the same point of view. Most of the times, the images are assumed to differ by small, sub-pixel shifts. This problem was first described by Tsai and Huang [11] in 1984. A good overview of existing super-resolution methods is given by Borman and Stevenson [12], Farsiu et al. [13], and in the special issues on super-resolution imaging in the IEEE Signal Processing Magazine [14] and in the EURASIP Journal on Applied Signal Processing [15].

Different methods exist to compute a high resolution image from a set of low resolution images. They mainly differ by the amount of prior information that is assumed to be available about the images. The most generally applicable methods only require that the images belong to a broad signal class, such as bandlimited signals [11], or signals with finite rate of innovation [16]. Other methods, such as the hallucination method by Baker and Kanade [17], use more specific knowledge about the image type (faces, printed text, etc.) to reconstruct a high resolution image. In general, there is a trade-off between the amount of prior information and the number of images required to produce good results. The methods described in this paper belong to the first category, and only use basic assumptions about the signal class. As opposed to other methods, which consider the aliasing as part of the noise, we will include the aliasing into our signal model, and use its information for both registration and reconstruction.

This paper is structured as follows. First, a mathematical description of the problem is given in Section II. Next, the uniqueness of the solution to this problem is discussed in Section III. Two subspace-based solution methods are described in Sections IV and V, respectively. We discuss the advantages and disadvantages of the different methods in Section VI, and we also present heuristic solution methods. In Section VII, we derive the computational complexity of the different algorithms. Results are presented in Section VIII, and finally some conclusions are drawn in Section IX. All the results and figures presented in this paper are reproducible [18]. The Matlab code and data to reproduce the results are available online at

http://lcavwww.epfl.ch/reproducible_research/VandewalleSVV06a.

## II. MATHEMATICAL DESCRIPTION

We aim to reconstruct a continuous-time signal from multiple uniform sets of samples with unknown offsets. The sampling frequency of each set is too low to allow direct reconstruction, but reconstruction is possible from the joint set of all samples together. In this section, we formulate the mathematical framework for this problem and fix notations.

Assume we have a finite dimensional[1] Hilbert space $\mathcal{H}$, for which we have a basis $\mathcal{B} = \{\varphi_l(t)\}_{l=0..L-1}$. That is, $\mathcal{H} = \text{span}\{\varphi_l(t)\}_{l=0..L-1}$. For simplicity, let us consider the functions $\varphi_l(t)$ defined on the interval $[0, 1]$. For periodic functions, we will assume the period to be 1, such that we consider one period. An arbitrary signal from this space can be expressed as

$$f(t) = \sum_{l=0}^{L-1} \alpha_l \varphi_l(t), \tag{1}$$

where $\alpha_l$ is the $l$-th expansion coefficient of $f(t)$ in the basis $\mathcal{B}$. Denote $\boldsymbol{\alpha}$ as the vector of unknown signal coefficients. Possible examples of spaces with associated bases include truncated Fourier series, wavelets, splines, etc. Assume now that we sample $f(t)$ uniformly at a rate $N$ over the signal period $[0, 1]$. This results in a first set of $N$ samples taken at times

$$\mathcal{T}_0 = \begin{pmatrix} 0 & \frac{1}{N} & \frac{2}{N} & \cdots & \frac{N-1}{N} \end{pmatrix}. \tag{2}$$

We take another set of samples at times

$$\mathcal{T}_m = \begin{pmatrix} t_m & \frac{1}{N} + t_m & \frac{2}{N} + t_m & \cdots & \frac{N-1}{N} + t_m \end{pmatrix}. \tag{3}$$

These are the sampling times of the first set, shifted by an unknown offset $t_m$ ($t_m \in \mathcal{R}$). As the offsets are relative to the first set of samples, we have $t_0 = 0$. We take $M$ such sets of samples $y_m$. The samples can be written as

$$\begin{aligned} y_m(n) &= f\left(\frac{n}{N} + t_m\right) \\ &= \sum_{l=0}^{L-1} \alpha_l \varphi_l\left(\frac{n}{N} + t_m\right), \text{ for } 0 \leq n < N. \end{aligned} \tag{4}$$

We can write this more compactly using vectors and matrices as

$$\mathbf{y}_m = \boldsymbol{\Phi}_{t_m} \boldsymbol{\alpha}. \tag{5}$$

In this equation, $\mathbf{y}_m$ is the $N \times 1$ vector with the $m$-th set of samples, and $\boldsymbol{\alpha}$ is the $L \times 1$ vector of expansion coefficients. The $N \times L$ matrix $\boldsymbol{\Phi}_{t_m}$ contains the undersampled basis functions that

[1]The extension to an infinite dimensional Hilbert space is theoretically possible, but of little practical interest since all data and algorithms are finite.

are uniformly sampled with an offset $t_m$. In our setup, $N < L$, and the set of equations (5) is underdetermined.

Now, all the sets of samples $\mathbf{y}_m$ are combined into a single vector $\mathbf{y}$ and similarly the basis matrices $\mathbf{\Phi}_{t_m}$ are combined into $\mathbf{\Phi_t}$, with $\mathbf{t} = \begin{pmatrix} t_0 & t_1 & \cdots & t_{M-1} \end{pmatrix}$ denoting the offset vector. This can be written as

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_{M-1} \end{pmatrix} = \begin{pmatrix} \mathbf{\Phi}_{t_0} \\ \mathbf{\Phi}_{t_1} \\ \vdots \\ \mathbf{\Phi}_{t_{M-1}} \end{pmatrix} \boldsymbol{\alpha} = \mathbf{\Phi_t} \boldsymbol{\alpha}. \tag{6}$$

The matrix $\mathbf{\Phi_t}$ has size $MN \times L$. Assuming that the total number of samples is larger than or equal to the number of expansion coefficients, or $MN \geq L$, this set of equations is in general well- or over-determined if $\mathbf{t}$ is known. If, additionally,

$$MN \geq L + M - 1, \tag{7}$$

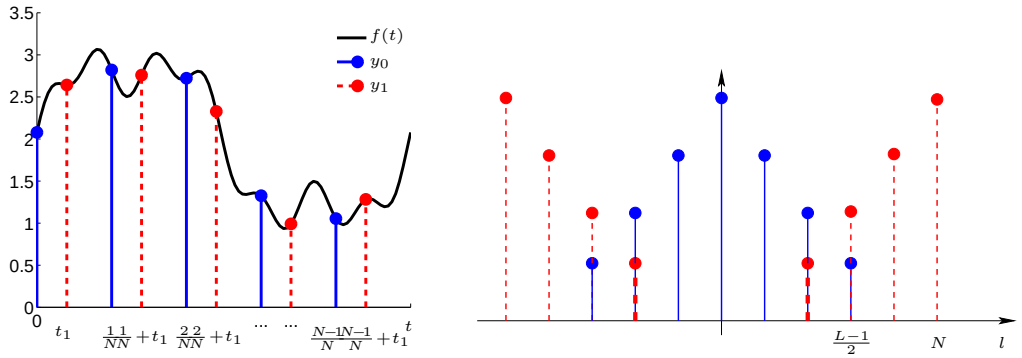
the number of equations is also larger than or equal to the number of unknowns ($L$ expansion coefficients and $M-1$ offsets), and it should be possible to remove the uncertainty of the unknown offsets. As we will show in the next sections, these additional equations allow us in general to compute the unknown offsets. Note that the challenging part of the problem is that it is non-linear in the offsets, and that $\boldsymbol{\alpha}$ and $\mathbf{t}$ have to be found simultaneously. In summary, the most important variables in this reconstruction problem are listed here (see also Figure 3):

- $N$ : the number of samples in each set $\mathbf{y}_m$,
- $\mathbf{y}_m$ : the length $N$ vector of the $m$-th set of samples,
- $L$ : the number of unknown expansion coefficients,
- $\boldsymbol{\alpha}$ : the length $L$ vector of the expansion coefficients $\alpha_l$ to be reconstructed,
- $M$ : the number of sets of samples,
- $\mathbf{t}$ : the length $M$ vector of the offsets $t_m$ between the different sets of samples.

The unknown variables are the expansion coefficients $\boldsymbol{\alpha}$ and the offsets $\mathbf{t}$. We assume that all the other variables are known. This is evident for the sets of samples $\mathbf{y}_m$, the number of samples per set $N$ and the number of sets $M$, as they form the input of the problem. We will also require that the number of signal coefficients $L$, or at least an estimate for $L$, is available.

As we will often use bandlimited functions throughout this paper, we specifically write out the equations (1)-(6) for this case. A bandlimited function can be expressed in the Fourier basis as

$$f(t) = \sum_{l=-K}^{K} \alpha_l e^{j2\pi l t}, \tag{8}$$




(a) (b)

Fig. 3. Illustration of the different variables with $M = 2$ and a Fourier basis. (a) Time domain representation of the signal $f(t)$ and its sets of samples $y_0$ (—) and $y_1$ (– –). (b) Frequency domain representation of the absolute values of the signal spectrum (—) and its aliased copies after sampling (– –).

where $\alpha_l$ are the Fourier coefficients. Note that we index the coefficients from $-K$ to $K$, instead of 0 to $L-1$, because this is the usual way of indexing Fourier coefficients. We will assume that $L$ is odd, such that $K = (L-1)/2$.

The first set of samples can be written as

$$y_0(n) = f\left(\frac{n}{N}\right) = \sum_{l=-K}^{K} \alpha_l W^{ln}, \tag{9}$$

with $W = e^{j2\pi/N}$. We can also write this in matrix notation as

$$\mathbf{y}_0 = \begin{pmatrix} 1 & \cdots & 1 & 1 & 1 & \cdots & 1 \\ W^{-K} & \cdots & W^{-1} & 1 & W & \cdots & W^{K} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ W^{-(N-1)K} & \cdots & W^{-(N-1)} & 1 & W^{N-1} & \cdots & W^{(N-1)K} \end{pmatrix} \begin{pmatrix} \alpha_{-K} \\ \vdots \\ \alpha_{-1} \\ \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_K \end{pmatrix} = \mathbf{F}\boldsymbol{\alpha}, \tag{10}$$

where $\mathbf{F}$ is an $N \times L$ Inverse Discrete Fourier Transform (IDFT) matrix. Due to the undersampling ($L > N$), some of the columns in $\mathbf{F}$ are repeated.

Similarly, for the $m$-th set of samples $y_m(n)$, we obtain

$$y_m(n) = f\left(\frac{n}{N} + t_m\right) = \sum_{l=-K}^{K} \alpha_l e^{j2\pi l(n/N+t_m)} = \sum_{l=-K}^{K} \alpha_l W^{ln} W_{t_m}^{l}, \tag{11}$$

with $W_{t_m} = e^{j2\pi t_m}$. In matrix notation, this can be expressed as

$$\mathbf{y}_m = \mathbf{F}\mathbf{D}_{t_m}\boldsymbol{\alpha}, \tag{12}$$

with $\mathbf{F}$ the $N \times L$ IDFT matrix defined above, and $\mathbf{D}_{t_m}$ an $L \times L$ diagonal matrix with elements $\mathbf{D}_{t_m}(l,l) = W_{t_m}^l$ $(-K \leq l \leq K)$. If we put all the $M$ sets of $N$ samples together, we obtain the Fourier basis equivalent of (6):

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_{M-1} \end{pmatrix} = \begin{pmatrix} \mathbf{F} \\ \mathbf{F}\mathbf{D}_{t_1} \\ \vdots \\ \mathbf{F}\mathbf{D}_{t_{M-1}} \end{pmatrix} \boldsymbol{\alpha}. \tag{13}$$

The DFT of a set of samples $\mathbf{y}_m$ can be written as

$$\mathbf{Y}_m = \frac{1}{N}\mathbf{F}_N^* \mathbf{y}_m = \frac{1}{N}\mathbf{F}_N^* \mathbf{F}\mathbf{D}_{t_m}\boldsymbol{\alpha}, \tag{14}$$

where $\mathbf{F}_N^*$ is a square $N \times N$ DFT matrix, and $\mathbf{F}$ is the $N \times L$ matrix defined in (10). The notation $\mathbf{F}_N^*$ is used for the Hermitian transpose of $\mathbf{F}_N$. The resulting vector $\mathbf{Y}_m$ has length $N$ and is an aliased and phase shifted version of $\boldsymbol{\alpha}$. If we take for example $L = 3N$, we get

$$\begin{aligned} \mathbf{Y}_m &= \frac{1}{N}\mathbf{F}_N^*\mathbf{F}\mathbf{D}_{t_m}\boldsymbol{\alpha} = \frac{1}{N}\mathbf{F}_N^* \begin{pmatrix} \mathbf{F}_N & \mathbf{F}_N & \mathbf{F}_N \end{pmatrix} \mathbf{D}_{t_m}\boldsymbol{\alpha} \\ &= \begin{pmatrix} \mathbf{I} & \mathbf{I} & \mathbf{I} \end{pmatrix} \mathbf{D}_{t_m}\boldsymbol{\alpha} = \sum_{i=-1}^{1} W_{t_m}^{iN} \mathbf{D}'_{t_m}\boldsymbol{\alpha}_i, \end{aligned} \tag{15}$$

where $\mathbf{D}'_{t_m}$ is the $N \times N$ central part of the $L \times L$ matrix $\mathbf{D}_{t_m}$, and $\boldsymbol{\alpha}_i$ is the $i$-th block of length $N$ from the vector $\boldsymbol{\alpha}$. In general, if $L$ is not a multiple of $N$, we can still perform the same decomposition by adding zeros to $\boldsymbol{\alpha}$ up to the next multiple of $N$.

Before tackling the problem in general, let us first analyze some very simple and intuitive examples.

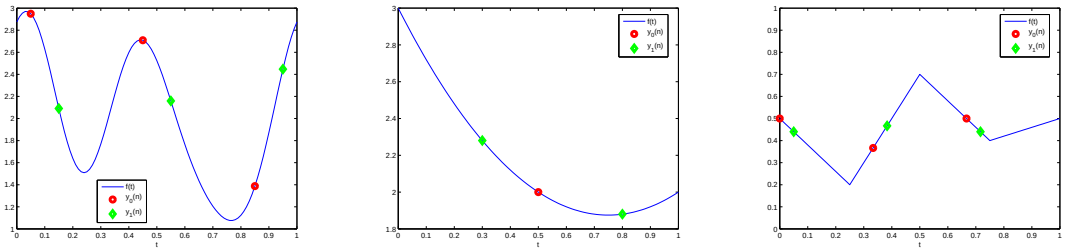
*Example 2.1 (Bandlimited functions):* Take $f(t)$ a bandlimited function (see Figure 4(a)), and $M = 2$, $N = 2$, and $L = 3$. Using (6), we can write

$$\mathbf{y} = \begin{pmatrix} y_0(0) \\ \hline y_0(1) \\ y_1(0) \\ y_1(1) \end{pmatrix} = \left(\begin{array}{ccc} 1 & 1 & 1 \\ -1 & 1 & -1 \\ \hline W_{t_1}^{-1} & 1 & W_{t_1} \\ -W_{t_1}^{-1} & 1 & -W_{t_1} \end{array}\right) \begin{pmatrix} \alpha_{-1} \\ \alpha_0 \\ \alpha_1 \end{pmatrix} = \boldsymbol{\Phi}_{\mathbf{t}}\boldsymbol{\alpha}. \tag{16}$$

The unknown offset in the matrix $\boldsymbol{\Phi}_{\mathbf{t}}$ is multiplied with the expansion coefficients in $\boldsymbol{\alpha}$ and makes the problem nonlinear. As described in (12), we can write $\boldsymbol{\Phi}_{t_1}$ (the lower part of $\boldsymbol{\Phi}_{\mathbf{t}}$) as an IDFT matrix multiplied with a diagonal matrix:

$$\boldsymbol{\Phi}_{t_1} = \begin{pmatrix} W_{t_1}^{-1} & 1 & W_{t_1} \\ -W_{t_1}^{-1} & 1 & -W_{t_1} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & -1 \end{pmatrix} \begin{pmatrix} W_{t_1}^{-1} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & W_{t_1} \end{pmatrix} = \mathbf{F}\mathbf{D}_{t_1}. \tag{17}$$

□



(a) (b) (c)

Fig. 4. Examples of the general setup. (a) Bandlimited function with two sets of three samples and unknown offset between the sets. (b) Second-degree polynomial with two sets of two samples and unknown offset between the sets. (c) One period of a periodic, piecewise linear and continuous function with two sets of three samples and unknown offset between the sets.

*Example 2.2 (Second order polynomials):* Let $f(t)$ be a second degree polynomial that is sampled using two sets of two samples (see Figure 4(b)). We can write $f(t)$ as $f(t) = \alpha_0 t^2 + \alpha_1 t + \alpha_2$. In this case, we have $M = 2$, $N = 2$, and $L = 3$. Note that here, we do not consider a periodic signal, but the samples are still taken in the interval $[0, 1]$. The rest of the setup is exactly as described above, and (6) becomes

$$\mathbf{y} = \left(\begin{array}{c} y_0(0) \\ \hline y_0(1) \\ y_1(0) \\ y_1(1) \end{array}\right) = \left(\begin{array}{ccc} 0 & 0 & 1 \\ \hline 0.25 & 0.5 & 1 \\ t_1^2 & t_1 & 1 \\ (0.5+t_1)^2 & 0.5+t_1 & 1 \end{array}\right) \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{pmatrix} = \mathbf{\Phi_t}\boldsymbol{\alpha}. \tag{18}$$

We can again see that the unknown offset $t_1$ appears in the matrix $\mathbf{\Phi_t}$, and is multiplied with the unknown signal coeffi cients in $\boldsymbol{\alpha}$, making the equations nonlinear. □

*Example 2.3 (Periodic piecewise linear and continuous functions):* Consider a continuous, periodic function with period 1 that is piecewise linear on intervals $[l/L, (l+1)/L]$ (with $0 \le l < L$). This is the space $\mathcal{S} = \text{span}\{\varphi_{l,\text{circ}}(t)\}$ for $0 \le l < L$, with

$$\varphi(t) = \begin{cases} -L + 1 + Lt & 1 - 1/L < t < 1 \\ 1 - Lt & 0 \le t < 1/L \\ 0 & \text{otherwise,} \end{cases} \qquad \text{with } t \in [0, 1) \tag{19}$$

and $\varphi_{l,\text{circ}}(t)$ the periodic extension of $\varphi(t - l/L)$ with period 1. An example is given in Figure 4(c) for $L = 4$. In this case, we can write the signal $f(t)$ as

$$f(t) = \alpha_0\varphi_0(t) + \alpha_1\varphi_1(t) + \alpha_2\varphi_2(t) + \alpha_3\varphi_3(t). \tag{20}$$

If we take two sets of three samples, we obtain the following equations:

$$\left(\begin{array}{c} y_0(0) \\ y_0(1) \\ y_0(2) \\ \hline y_1(0) \\ y_1(1) \\ y_1(2) \end{array}\right) = \left(\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 2/3 & 1/3 & 0 \\ 0 & 0 & 1/3 & 2/3 \\ \hline 1-t_1 & t_1 & 0 & 0 \\ 0 & 2/3-t_1 & 1/3+t_1 & 0 \\ 0 & 0 & 1/3-t_1 & 2/3+t_1 \end{array}\right) \left(\begin{array}{c} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{array}\right). \tag{21}$$

We assume that the second set of samples is taken in the same $[l/L, (l+1)/L]$ interval as the first set (in this example $0 < t_1 \leq 1/12$). Other values of $t_1$ do not cause any fundamental changes, they simply change the above equations because the basis function $\varphi(t)$ is described by different equations on the intervals $[1-1/L, 1]$, $[0, 1/L]$ and outside of these intervals, as can be seen in (19). Because of the finite support of the basis functions, the matrices $\mathbf{\Phi}_{t_m}$ have a banded structure. □

From the above examples, we see that the unknown offsets and signal coefficients are mixed up, which makes the problem hard to solve. Once the offsets are known, (6) is reduced to a set of linear equations in the unknown signal coefficients. This can be solved straightforwardly using a least squares method. From this description, we clearly see that the hardest part of the reconstruction problem is the estimation of the offsets. In the next sections, we will concentrate on this problem, and present methods to compute these offsets based on subspaces.

## III. Uniqueness of the solution

In this section, we discuss the existence and uniqueness of a solution to the super-resolution problem with unregistered sets of samples. It follows straightforwardly from the description above that a solution exists (at least in the ideal, noiseless case). The uniqueness of such a solution is less trivial.

### *A. General case*

Let us start with an intuitive statement for the general problem. A unique solution exists if there is a unique mapping from each set of samples to a single space generated by the columns of $\mathbf{\Phi_t}$. In other words, the intersection of the spaces generated for two different offset vectors $\mathbf{t}$ only contains the trivial zero vector, except for possibly some degenerate cases. This argument follows the one made by Marziliano for discrete sampling with unknown locations [19]. More formally, we have the following lemma.

*Lemma 3.1 (Uniqueness of the general problem):* If the $L$-dimensional subspaces (in the $MN$-dimensional measurement space) generated by the columns of $\mathbf{\Phi_t}$ for any pair of different offset

vectors $\mathbf{t}$ and $\mathbf{t}'$ have a low-dimensional intersection, then a unique signal $f(t)$ belonging to the finite-dimensional Hilbert space $\mathcal{H}$ can be reconstructed from the sample vector $\mathbf{y}$, except if the sample vector $\mathbf{y}$ is an element of the intersection.

**Proof:** We prove this by contradiction. Assume that there are two different offset vectors $\mathbf{t}$ and $\mathbf{t}'$, and that the sample vector $\mathbf{y}$ is not in the intersection between the spaces generated by the columns of $\mathbf{\Phi_t}$ and $\mathbf{\Phi_{t'}}$. We can then write the sample vector $\mathbf{y}$ in two ways, namely:

$$\mathbf{y} = \mathbf{\Phi_t}\boldsymbol{\alpha} = \mathbf{\Phi_{t'}}\boldsymbol{\alpha}'. \tag{22}$$

Because the vector $\mathbf{y}$ is not in the intersection of the spaces spanned by $\mathbf{\Phi_t}$ and $\mathbf{\Phi_{t'}}$, this cannot be true. Our assumption was wrong, and $\mathbf{t} = \mathbf{t}'$. ■

In general, the space of sample vectors that belong to the intersection is low-dimensional and thus, the set of sample vectors for which a unique solution exists, is dense in the Hilbert space $\mathcal{H}$. We shall therefore ignore the degenerate cases. Nevertheless, we will specifically calculate such an example of a degenerate case for polynomial signals in Section III-C.

Unfortunately, this theorem does not give any specific conditions for the uniqueness of the solution. We will therefore analyze a few specific signal types to gain understanding.

### *B. Bandlimited signals*

First of all, we consider the space of bandlimited functions with the Fourier basis. As we discussed in Section II, the matrix $\mathbf{\Phi}_{t_m}$ can then be written as the product of an $N \times L$ IDFT matrix $\mathbf{F}$ with an $L \times L$ diagonal matrix $\mathbf{D}_{t_m}$. For aliased signals, some of the columns in $\mathbf{F}$ are repeated, and the corresponding columns of $\mathbf{\Phi}_{t_m}$ only differ by their multiplication factor $W_{t_m}^l$. These columns correspond to overlapping frequencies in the sampled spectrum.

If all the sets of samples are considered together, the $l$-th column $\boldsymbol{\phi}_{\mathbf{t}}^l$ of the matrix $\mathbf{\Phi_t}$ is a repetition of $M$ times the same basis vector $\mathbf{F}^l$, the $l$-th column of $\mathbf{F}$, multiplied by the different factors $W_{t_m}^l$:

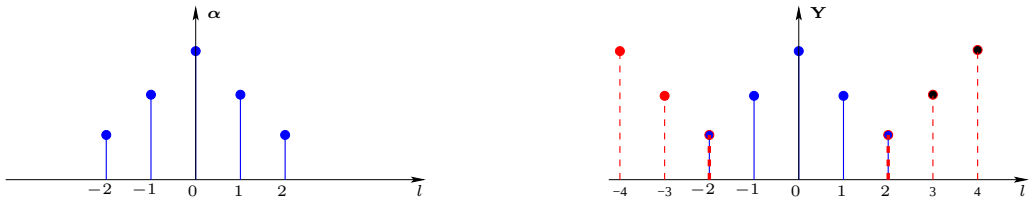
$$\left(\boldsymbol{\phi}_{\mathbf{t}}^l\right)^T = \begin{pmatrix} \mathbf{F}^{l^T} & W_{t_1}^l \mathbf{F}^{l^T} & \cdots & W_{t_{M-1}}^l \mathbf{F}^{l^T} \end{pmatrix}. \tag{23}$$

*Example 3.1:* For example, if we take $M = 2$, $L = 5$, and $N = 4$, there are two overlapping

frequency components (see Figure 5). We have

$$\mathbf{\Phi_t} = \begin{pmatrix} \mathbf{\Phi}_0 \\ \mathbf{\Phi}_{t_1} \end{pmatrix} = \begin{pmatrix} \mathbf{F} \\ \mathbf{F}\mathbf{D}_{t_1} \end{pmatrix} = \left(\begin{array}{ccccc} 1 & 1 & 1 & 1 & 1 \\ W^2 & W^3 & 1 & W & W^2 \\ 1 & W^2 & 1 & W^2 & 1 \\ W^2 & W & 1 & W^3 & W^2 \\ \hline W_{t_1}^{-2} & W_{t_1}^{-1} & 1 & W_{t_1} & W_{t_1}^2 \\ W_{t_1}^{-2}W^2 & W_{t_1}^{-1}W^3 & 1 & W_{t_1}W & W_{t_1}^2W^2 \\ W_{t_1}^{-2} & W_{t_1}^{-1}W^2 & 1 & W_{t_1}W^2 & W_{t_1}^2 \\ W_{t_1}^{-2}W^2 & W_{t_1}^{-1}W & 1 & W_{t_1}W^3 & W_{t_1}^2W^2 \end{array}\right), \qquad (24)$$

with $W = e^{j2\pi/4}$. The first and the fifth column of $\mathbf{F}$ are equal, such that the corresponding columns of $\mathbf{\Phi_t}$ only differ by their factors $W_{t_1}^{-2}$ and $W_{t_1}^2$ from $\mathbf{D}_{t_1}$. □




(a) (b)

Fig. 5. Signal spectrum before (a) and after sampling (b), for $L = 5$ and $N = 4$. The base spectrum (—) and aliased spectrum (– –) overlap for the first and last spectral component, corresponding to the first and fifth column in (24).

Because the Fourier basis vectors $\mathbf{F}^l$ of the $N \times L$ matrix $\mathbf{F}$ are orthogonal (note that we use $W = e^{j2\pi/N}$ in this matrix, and not $e^{j2\pi/L}$), the vectors $\boldsymbol{\phi}_{\mathbf{t}}^l$ are also orthogonal to each other for any set of offset values $\mathbf{t}$. Only the vectors corresponding to overlapping spectrum coefficients (like the first and fifth column in Example 3.1) are not orthogonal, because they are composed by the same Fourier vector $\mathbf{F}^l$ but have different coefficients. Each vector describes a trajectory in $\mathcal{C}^{MN}$ for varying values of $\mathbf{t}$. This trajectory is defined more precisely in the following theorem:

*Theorem 3.1:* For varying $\mathbf{t}$, any vector $\boldsymbol{\phi}_{\mathbf{t}}^l$ describes a trajectory in the $M$-dimensional subspace $\mathcal{V}_l$ of $\mathcal{C}^{MN}$:

$$\mathcal{V}_l = \text{span}\left\{\mathbf{A} \otimes \mathbf{F}^l\right\} \quad \text{with } \mathbf{A} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & -1 & & 1 \\ \vdots & & \ddots & \vdots \\ 1 & 1 & \cdots & -1 \end{pmatrix}, \qquad (25)$$

where $\mathbf{A}$ is an $M \times M$ matrix, and $\otimes$ represents the Kronecker product. The vectors $\boldsymbol{\phi}_{\mathbf{t}}^{l+iN}$ corresponding to overlapping spectrum coefficients belong to the same space $\mathcal{V}_l$, while other vectors $\boldsymbol{\phi}_{\mathbf{t}}^k$

belong to orthogonal subspaces $\mathcal{V}_k \perp \mathcal{V}_l$.

**Proof:** The trajectory of $\boldsymbol{\phi}_{\mathbf{t}}^l$ as a function of $\mathbf{t}$ is in $\mathcal{V}_l$ iff we can write any arbitrary $\boldsymbol{\phi}_{\mathbf{t}}^l$ as a linear combination of the columns of $\mathbf{A} \otimes \mathbf{F}^l$. From (23) and (25), we can write the vector $\boldsymbol{\phi}_{\mathbf{t}}^l$ as a linear combination of the columns of $\mathbf{A} \otimes \mathbf{F}^l$ by solving the set of linear equations

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & -1 & & 1 \\ 1 & & \ddots & \vdots \\ 1 & 1 & \cdots & -1 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{M-1} \end{pmatrix} = \begin{pmatrix} 1 \\ W_{t_1}^l \\ \vdots \\ W_{t_{M-1}}^l \end{pmatrix}. \tag{26}$$

As this set of $M$ equations is of full rank, it always has a unique solution, and our vector is therefore part of $\mathcal{V}_l$. Vectors corresponding to overlapping Fourier coefficients are composed from the same Fourier basis vectors $\mathbf{F}^l$, and therefore belong to the same space $\mathcal{V}_l$.

The orthogonality between two subspaces $\mathcal{V}_k$ and $\mathcal{V}_l$ for $k \neq l$ can easily be seen. As each of the subspaces is generated by the columns of the matrices defined in (25), it is sufficient that we prove that any arbitrary column of $\mathbf{A} \otimes \mathbf{F}^k$ is orthogonal to any column of $\mathbf{A} \otimes \mathbf{F}^l$. Denoting the $i$-th column of $\mathbf{A}$ as $\mathbf{A}^i$, we can write the inner product between two such vectors as

$$\left\langle \mathbf{A}^i \otimes \mathbf{F}^k, \mathbf{A}^j \otimes \mathbf{F}^l \right\rangle = \sum_{n=0}^{M-1} (\pm 1) \left\langle \mathbf{F}^k, \mathbf{F}^l \right\rangle = 0 \text{ if } k \neq l. \tag{27}$$

This is valid because $\left\langle \mathbf{F}^k, \mathbf{F}^l \right\rangle = 0$ from the orthogonality of the Fourier basis. ■

From this analysis, we can see that the problem can be considered in each of the $N$ subspaces separately by projecting $\mathbf{y}$ onto the different $M$-dimensional subspaces. The projection of $\mathbf{y}$ onto the subspace $\mathcal{V}_l$ is called

$$\mathbf{y}^{(l)} = \mathbf{P}_{\mathcal{V}_l} \mathbf{y}. \tag{28}$$

An illustration for two overlapping vectors in a three-dimensional space is sketched in Figure 6.

Assuming that $MN \geq L + M - 1$, at most $M$ columns of $\boldsymbol{\Phi}_{\mathbf{t}}$ belong to the same subspace $\mathcal{V}_l$. If there are $M$ vectors $\boldsymbol{\phi}_{\mathbf{t}}^{l+iN}$ from $\boldsymbol{\Phi}_{\mathbf{t}}$ in a particular subspace $\mathcal{V}_l$, these vectors form a basis for $\mathcal{V}_l$. As the vector $\mathbf{y}^{(l)}$ is also in $\mathcal{V}_l$, it belongs to the subspace spanned by these vectors $\boldsymbol{\phi}_{\mathbf{t}}^{l+iN}$ for any value of $\mathbf{t}$. However, for the subspaces containing less than $M$ vectors $\boldsymbol{\phi}_{\mathbf{t}}^{l+iN}$, the span of these vectors does not cover the entire space $\mathcal{V}_l$. Hence, the projection of $\mathbf{y}$ onto $\mathcal{V}_l$ generally only belongs to $\text{span}(\boldsymbol{\phi}_{\mathbf{t}}^{l+iN})_{0 \leq i < M-1}$ for a single offset vector $\mathbf{t}$. The offset vector $\mathbf{t}$ can then be uniquely determined from these subspaces, and the signal coefficients $\boldsymbol{\alpha}$ are thus also uniquely determined.

Using the Fourier basis, there are certain cases for which it is not sufficient to require $MN \geq L + M - 1$. The vectors in the space $\mathcal{V}_0$ (the columns $l = iN$) do not depend on $\mathbf{t}$, and therefore, no information about the offsets can be derived from these vectors. So if only the space $\mathcal{V}_0$ contains less

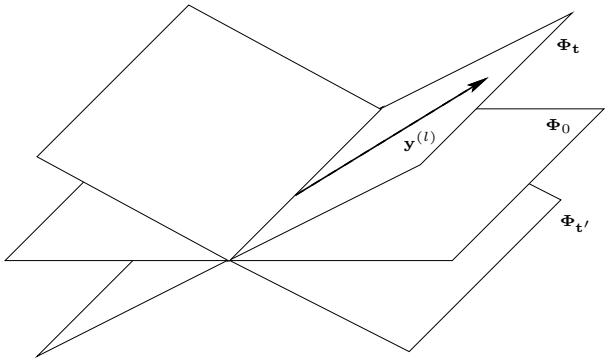



Fig. 6. Illustration of the trajectory of the span of two columns $\phi^l_{\mathbf{t}}$ (the planes in this drawing) in a three-dimensional space $\mathcal{V}_l$ for different offset values (0, $\mathbf{t}$ and $\mathbf{t}'$). Only for the correct offsets $\mathbf{t}$, the vector $\mathbf{y}^{(l)}$ belongs to the space spanned by the two columns.

than $M$ vectors, it is not possible to derive the offsets from the sets of samples. This happens when the number of overlapping parts of the frequency spectrum is even (and therefore also the number of sets of samples $M$ is even). In that case, we need $MN \geq L + M$.

*Example 3.2:* Looking back at Example 3.1, we can see that the first and fifth column of $\mathbf{\Phi_t}$ span the subspace $\mathcal{V}_2$, except in the degenerate case when $t_1 = 1$. The projection of $\mathbf{y}$ onto these two columns will therefore always be the same as its projection onto $\mathcal{V}_2$. No information about $t_1$ can be derived from this subspace. However, for $\mathcal{V}_{-1}$, $\mathcal{V}_0$ and $\mathcal{V}_1$, there is only one column of $\mathbf{\Phi_t}$ in each of these subspaces, and the projections onto that column or onto the space $\mathcal{V}_l$ only coincide for the correct value of $t_1$. We need at least one such subspace containing only 1 column from $\mathbf{\Phi_t}$ in order to be able to compute $t_1$. As we do not want this to be the zero frequency, we require that $2N \geq L+2$, or $N = 4$. □

This leads to the following theorem. Its proof follows from the above argument.

*Theorem 3.2 (Bandlimited signals):* If $MN \geq L + M - 1$ for $M$ odd, or $MN \geq L + M$ for $M$ even, a bandlimited signal with $L$ Fourier coefficients can be uniquely reconstructed from $M$ uniform sets of $N$ samples with unknown offsets.

### C. Polynomials

Let us now also analyze the uniqueness for second degree polynomials, using the same setup as in Example 2.2. In this case, we can prove the following theorem.

*Theorem 3.3 (Second order polynomials):* When an arbitrary second order polynomial is sampled at four non-coinciding points $\{0, 0.5, t, t + 0.5\}$, then this polynomial $ax^2 + bx + c$ and the shift $t$ can be uniquely retrieved, except in the particular case $at + 0.5a + b = 0$, when an infinite number

of solutions exist.

The proof for this theorem is given in appendix. Such an example of multiple solutions is given in Figure 7.

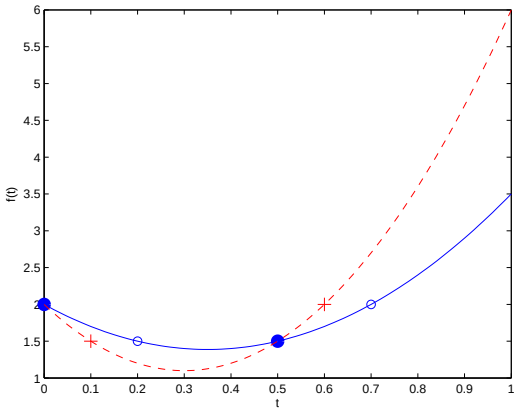



Fig. 7. Example of parasitic solutions for second order polynomials. The polynomial $f(x) = 5x^2 - 3.5x + 2$ (—) is sampled at times $\{0, 0.5\}$ and $\{0.2, 0.7\}$ ($t_1 = 0.2$, plotted with 'o'). From the two sets of samples, other second degree polynomials can also be reconstructed using other offsets $t_1$. For example, setting $t_1 = 0.1$, and thus shifting the second set of samples to the left (plotted with '+'), we would obtain the polynomial $g(x) = 10x^2 - 6x + 2$ (– –).

A similar theorem can be shown for polynomials of third order. With increasing order, the proofs by contradiction like the one in the appendix become very complex. However, it seems reasonable that also for higher-order polynomials, the solution will be unique except for a lower-dimensional set of degenerate cases.

## IV. Solution using matrix rank

In this section, we describe a first solution method that uses specific properties of the Fourier transform. It is therefore only applicable to the space of bandlimited functions.

### A. Method

As shown in (14), the discrete Fourier transform $\mathbf{Y}_m$ of the $m$-th set of samples can be written as

$$\mathbf{Y}_m = \frac{1}{N}\mathbf{F}_N^* \mathbf{F} \mathbf{D}_{t_m} \boldsymbol{\alpha}, \tag{29}$$

with $\mathbf{F}_N^*$ a square $N \times N$ DFT matrix, and $\mathbf{F}$ an $N \times L$ IDFT matrix. If we extend $\boldsymbol{\alpha}$ to a length that is a multiple of $N$, we can split the Fourier coefficient vector $\boldsymbol{\alpha}$ in blocks $\boldsymbol{\alpha}_i$ of length $N$, and obtain (see also (15)):

$$\mathbf{Y}_m = \mathbf{D}'_{t_m} \sum_{i=\lceil -(S-1)/2 \rceil}^{\lceil (S-1)/2 \rceil} W_{t_m}^{iN} \boldsymbol{\alpha}_i. \tag{30}$$

The vectors $\boldsymbol{\alpha}_i$ represent the overlapping parts of the Fourier spectrum due to undersampling, and there are $S = \lceil L/N \rceil$ such parts. This means that for any set of samples, $\mathbf{D}'^{-1}_{t_m}\mathbf{Y}_m$ is a linear combination of the $S$ parts of the Fourier spectrum $\boldsymbol{\alpha}_i$:

$$\mathbf{D}'^{-1}_{t_m}\mathbf{Y}_m = \sum_{i=\lceil -(S-1)/2 \rceil}^{\lceil (S-1)/2 \rceil} W_{t_m}^{iN}\boldsymbol{\alpha}_i. \tag{31}$$

In other words, each vector $\mathbf{D}'^{-1}_{t_m}\mathbf{Y}_m$ belongs to the $S$-dimensional subspace span$(\{\boldsymbol{\alpha}_i\}_{i=\lceil -(S-1)/2 \rceil .. \lceil (S-1)/2 \rceil})$. If we have more than $S$ sets of samples ($M > S$), and $N \geq S$, the rank of the matrix containing all the sets of samples should therefore be $S$:

$$\text{rank}\left(\mathbf{Y}^D_{\mathbf{t}}\right) = S \qquad \text{with } \mathbf{Y}^D_{\mathbf{t}} = \begin{pmatrix} \mathbf{Y}_0 & \mathbf{D}'^{-1}_{t_1}\mathbf{Y}_1 & \cdots & \mathbf{D}'^{-1}_{t_{M-1}}\mathbf{Y}_{M-1} \end{pmatrix}. \tag{32}$$

In general, if the estimated offsets $\hat{\mathbf{t}}$ do not have the correct values, rank$(\mathbf{Y}^D_{\hat{\mathbf{t}}}) > S$. For most offset values $\hat{\mathbf{t}}$, the matrix $\mathbf{Y}^D_{\hat{\mathbf{t}}}$ will be of full rank $M$ (and $M > S$). The correct values of the offsets $\mathbf{t}$ can then be found as the values for which the matrix rank becomes $S$. A schematic description of the complete reconstruction algorithm built on this method is given in Figure 8.

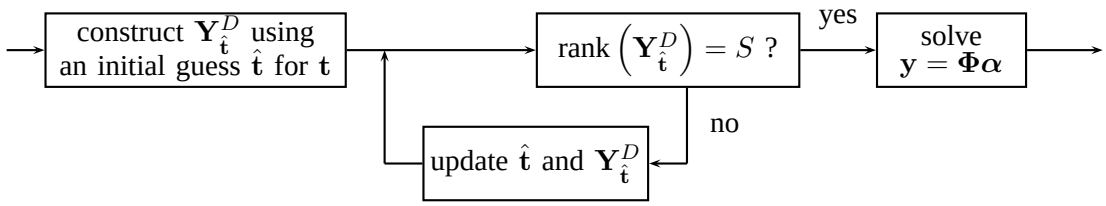



Fig. 8. Signal reconstruction algorithm using the matrix rank method from Section IV. The estimate for the offsets $\hat{\mathbf{t}}$ and the corresponding sample matrix $\mathbf{Y}^D_{\hat{\mathbf{t}}}$ are updated iteratively. Once the estimate is good enough, the signal parameters $\boldsymbol{\alpha}$ are reconstructed.

### *B. Discussion*

As there are $S$ overlapping parts in the sampled Fourier spectrum, we need at least $M = S + 1 = \lceil L/N \rceil + 1$ sets of samples to find the offsets using this method. If more sets of samples are available, we still have rank $\left(\mathbf{Y}^D_{\hat{\mathbf{t}}}\right) = S$, and this adds some robustness to both the offset estimation and the signal reconstruction. However, each new set of samples also adds a new unknown offset, thereby increasing the complexity of the estimation.

Remark that the operation performed by multiplying a sample vector $\mathbf{Y}_m$ with a matrix $\mathbf{D}'^{-1}_{t_m}$ does not change its norm $\|\mathbf{Y}_m\|_2^2$. It merely performs a rotation to align the sample vector into the $S$-dimensional subspace.

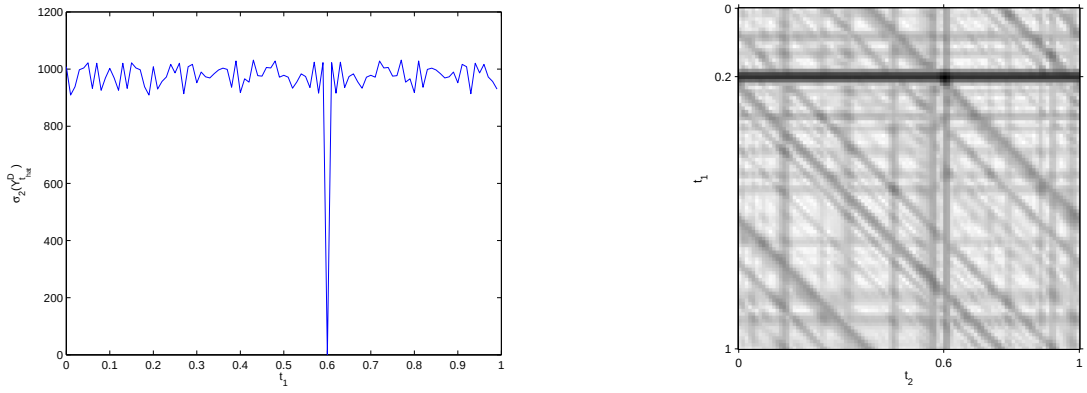
The computation of the rank of a matrix has a rather 'binary' outcome: either it is $S$, or it has an integer value larger than $S$. Hence, if the measurements are noisy, this test is very likely to fail even for the correct values of $\mathbf{t}$. It is therefore much better to evaluate the $S+1$-th singular value of the matrix $\mathbf{Y}_{\hat{\mathbf{t}}}^{D}$, or the determinant of the square matrix $\mathbf{Y}_{\hat{\mathbf{t}}}^{D*}\mathbf{Y}_{\hat{\mathbf{t}}}^{D}$. This can also give an indication about the quality of the current approximation. While the determinant requires less computations, it is also numerically less stable than a singular value decomposition. The $S+1$-th singular value could be computed using the inverse power method described by Strang [20]. This is an iterative method to approximate the smallest singular value of a matrix. It requires the solution of a linear $M \times M$ system of equations in each iteration. The number of iterations needed will generally be very small, because the $S+1$-th singular value is typically much smaller than the first $S$ singular values.

The objective function based on (32) can thus be written as

$$\min_{\hat{\mathbf{t}}} \sigma_{S+1}\left(\mathbf{Y}_{\hat{\mathbf{t}}}^{D}\right), \tag{33}$$

where the operation $\sigma_{S+1}(\mathbf{A})$ stands for computing the $S+1$-th singular value of the matrix $\mathbf{A}$.

*Example 4.1 (Bandlimited functions):* Let us consider a bandlimited function with $L = 81$ unknown Fourier coefficients. It is sampled with two sets of 90 samples, with offsets $\mathbf{t} = \begin{pmatrix} 0 & 0.6 \end{pmatrix}$. The objective function from (33) is shown as a function of $t_1$ in Figure 9(a). Similarly, we sample the same function with $M = 3$ sets of $N = 41$ samples and offsets $\mathbf{t} = \begin{pmatrix} 0 & 0.2 & 0.6 \end{pmatrix}$. The objective function is shown in Figure 9(b) as a function of the offsets $t_1$ and $t_2$. □




(a) (b)

Fig. 9. Examples of the objective function in (33). Next to the global minimum, it also contains many local minima. (a) Two aliased sets of 90 samples, with 81 unknown coefficients. The exact offset is $t_1 = 0.6$. (b) Three aliased sets of 41 samples, with 81 unknown coefficients. The exact offsets are $t_1 = 0.2$ and $t_2 = 0.6$.

In the above examples, the objective function (33) is very 'flat', except for a small region around the optimal offset values. At an arbitrary value away from the correct solution, it is therefore not

possible to predict what would be a better approximation. Standard minimization algorithms such as gradient descent will generally not converge to the right solution for such functions. However, from the example with three sets of samples, we can see that the objective function is not completely arbitrary. Horizontal, vertical and diagonal lines appear. They correspond to relative alignments of two out of the three sets of samples. It is therefore interesting to use this information in the search for the minimum. Such a heuristic algorithm will be described in Section VI.

## V. Solution using projections

In this section, we describe a method that is applicable to any kind of signals in a finite-dimensional Hilbert space defined in Section II. It uses projections onto a subspace basis to find the relative offsets.

### A. Method

From equation (6), we can see that $\mathbf{y}$ belongs to span($\mathbf{\Phi_t}$). However, this is only true if the matrix $\mathbf{\Phi_t}$ is constructed using the correct values for the offset vector $\mathbf{t}$ (except for degenerate cases). For another (incorrect) set of offset values $\hat{\mathbf{t}}$, (6) no longer holds and the sample vector $\mathbf{y}$ is not in span($\mathbf{\Phi}_{\hat{\mathbf{t}}}$) anymore. Using Lemma 3.1, one can test the correctness of the offset vector by verifying if the projection $\hat{\mathbf{y}}$ of the sample vector $\mathbf{y}$ onto span($\mathbf{\Phi}_{\hat{\mathbf{t}}}$) gives the sample vector $\mathbf{y}$ again. Or mathematically:

$$\begin{cases} \hat{\mathbf{y}} &= \mathbf{P}_{\mathbf{\Phi}_{\hat{\mathbf{t}}}}\mathbf{y} = \mathbf{\Phi}_{\hat{\mathbf{t}}}(\mathbf{\Phi}_{\hat{\mathbf{t}}}^*\mathbf{\Phi}_{\hat{\mathbf{t}}})^{-1}\mathbf{\Phi}_{\hat{\mathbf{t}}}^*\mathbf{y} = \mathbf{y}, & \text{for } \hat{\mathbf{t}} = \mathbf{t} \\ \hat{\mathbf{y}} &= \mathbf{P}_{\mathbf{\Phi}_{\hat{\mathbf{t}}}}\mathbf{y} \neq \mathbf{y}, & \text{for } \hat{\mathbf{t}} \neq \mathbf{t}. \end{cases} \tag{34}$$

Therefore, this can be used to build an optimization problem for finding the correct offsets $\mathbf{t}$. We search for the value of $\hat{\mathbf{t}}$ such that the following function is minimized:

$$\min_{\hat{\mathbf{t}}} \|\mathbf{y} - \hat{\mathbf{y}}\|^2. \tag{35}$$

Using this method to find the offsets between the sets of samples $\mathbf{y}_m$, we can derive an algorithm to reconstruct the signal $f(t)$ from its combined sets of samples $\mathbf{y}$. A block diagram of such an algorithm is given in Figure 10.

### B. Discussion

Let us study some examples of the objective function $\|\mathbf{y} - \hat{\mathbf{y}}\|^2$ (see (35)) as a function of the offset values $\hat{\mathbf{t}}$.

*Example 5.1 (Bandlimited functions):* A bandlimited function with 81 unknown Fourier coefficients is sampled with two sets of 90 samples. The offset vector that we used is $\mathbf{t} = \begin{pmatrix} 0 & 0.6 \end{pmatrix}$. The objective function $\|\mathbf{y} - \hat{\mathbf{y}}\|^2$ from (35) is shown as a function of the offset $t_1$ in Figure 11(a).

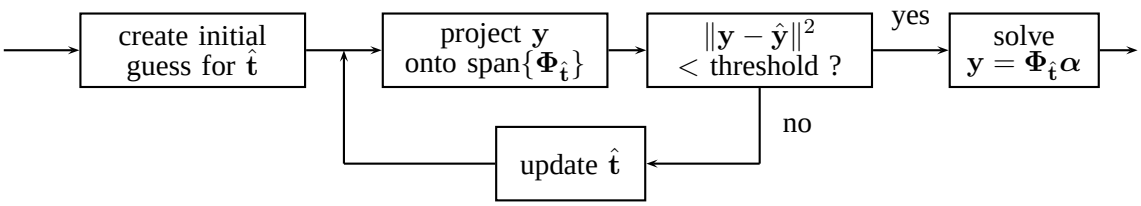
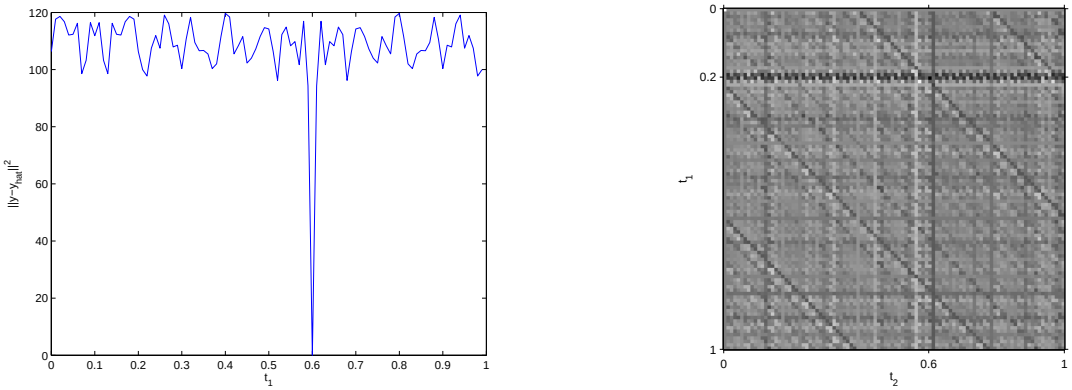



Fig. 10. Signal reconstruction algorithm using the projection method from Section V. The estimate for the offsets $\hat{\mathbf{t}}$ is updated iteratively. Once the estimate is considered good enough, the signal parameters $\boldsymbol{\alpha}$ are reconstructed.

It is 0 at the correct offset value ($t_1 = 0.6$), and is larger at other values of $\hat{t}_1$. Next to this global minimum, the function also has many local minima.

Similarly, we also take three sets of 41 samples from a bandlimited function with again 81 Fourier coeffi cients. The offset vector for this example is $\mathbf{t} = \begin{pmatrix} 0 & 0.2 & 0.6 \end{pmatrix}$. Figure 11(b) shows the objective function as a function of the offsets $t_1$ and $t_2$. Smaller values are represented by darker graylevels. Again, the minimum can clearly be seen, and is at the intersection of the dark horizontal, vertical and diagonal lines corresponding to pairwise alignments between the sets of samples. However, the objective function also has many local minima. □

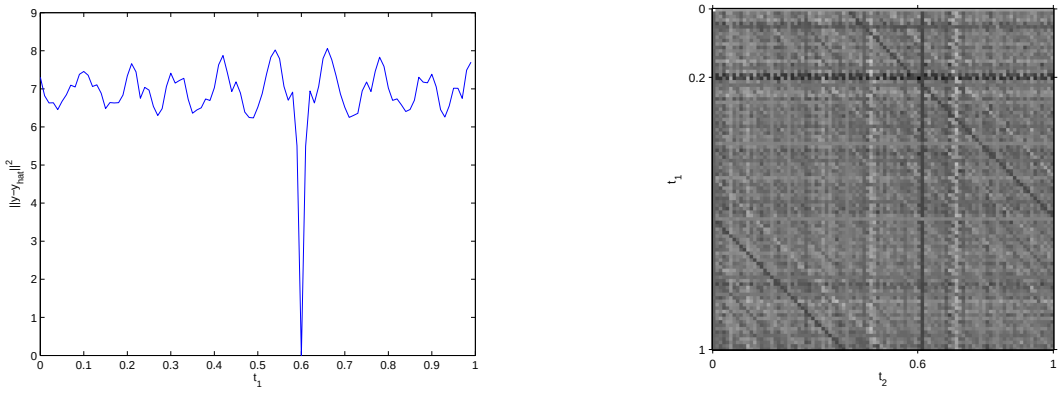



(a) (b)

Fig. 11. Examples of the objective function in (35) where $\mathcal{H}$ is the Fourier space. Next to the global minimum, it also contains many local minima. (a) Two aliased sets of 90 samples, with 81 unknown coeffi cients. The exact offset is $t_1 = 0.6$. (b) Three aliased sets of 41 samples, with 81 unknown coeffi cients. The exact offsets are $t_1 = 0.2$ and $t_2 = 0.6$. Small values are represented by dark pixels.

*Example 5.2 (Periodic piecewise linear and continuous functions):* A similar example is repeated for piecewise linear and continuous functions, as described in Example 2.3. The same setup as in Example 5.1 is used, with fi rst $M = 2$, $N = 90$ and $L = 81$ (Figure 12(a)), and next $M = 3$,

$N = 41$, and $L = 81$ (Figure 12(b)). The results are similar to those obtained in the bandlimited case. They also have many local minima, and are zero only for the correct offsets $\hat{\mathbf{t}} = \mathbf{t}$. □




(a) (b)

Fig. 12. Examples of the objective function in (35) on a piecewise linear and continuous function. Next to the global minimum, it also contains many local minima. (a) Two aliased sets of 90 samples, with 81 unknown coeffi cients. The exact offset is $t_1 = 0.6$. (b) Three aliased sets of 41 samples, with 81 unknown coeffi cients. The exact offsets are $t_1 = 0.2$ and $t_2 = 0.6$. Small values are represented by dark pixels.

From these examples, it is clearly visible that (35) is not easy to minimize (just like in Section IV). Next to the global minimum, the objective function has many local minima. We cannot use a standard algorithm like gradient descent for minimization. However, as can also be seen from Figure 11(b) and Figure 12(b), the objective function shows the same structure of horizontal, vertical and diagonal lines as in Figure 9(b).

Let us now consider specifi cally bandlimited functions, with the Fourier basis, to give some further interpretation of the above results. As it was discussed in Section III, the $MN$-dimensional sample space can be divided into $N$ orthogonal subspaces of dimension $M$. It is therefore possible to split the function from (35) into independent terms according to these $M$-dimensional subspaces:

$$\|\mathbf{y} - \hat{\mathbf{y}}\|^2 = \sum_{n=0}^{N-1} \|\mathbf{y}^{(n)} - \hat{\mathbf{y}}^{(n)}\|^2. \tag{36}$$

In this equation, $\mathbf{y}^{(n)}$ is the projection of $\mathbf{y}$ onto the subspace $\mathcal{V}_n$ spanned by the $n$-th vector for different values of $\hat{\mathbf{t}}$, and $\hat{\mathbf{y}}^{(n)}$ is the projection of $\mathbf{y}$ onto the vectors $\boldsymbol{\phi}_{\hat{\mathbf{t}}}^{n+iN}$ that belong to this space. An example of such a decomposition is given in Figure 13.

As we also described in Section III, for the frequencies $n$ that have $M$ overlapping spectrum coeffi cients, the spaces $\mathcal{V}_n$ and span($\boldsymbol{\phi}_{\hat{\mathbf{t}}}^{n+iN}$) are the same, regardless of the value of $\hat{\mathbf{t}}$. For these frequencies, $\|\mathbf{y}^{(n)} - \hat{\mathbf{y}}^{(n)}\|^2 = 0$, and they do not contribute to the objective function in (36). The

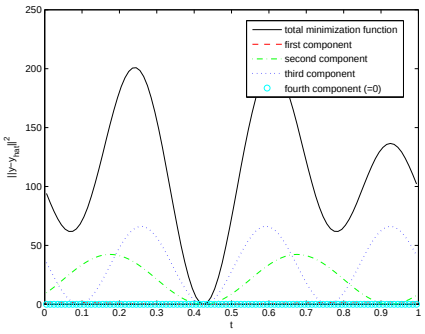



Fig. 13. Example of the decomposition of the objective function into its different components belonging to orthogonal $M$-dimensional subspaces.

other terms each contribute a periodic term to the global objective function. The minimum of (36) can therefore be found by minimizing the different components individually (see Figure 13). For each of the components, we can now minimize over an $M$-dimensional subspace instead of the original $MN$-dimensional space.

## VI. Practical Issues

### A. Sampling kernel

In the different methods described above, no sampling kernel was considered. The signals are sampled using Diracs. Although this is not very realistic, it is an approximation that is often made to simplify the analysis. For the frequency domain methods described above, it is in most cases not very difficult to take a different sampling kernel into account. In this case the sampling operation can be considered as a convolution with a sampling kernel, followed by the actual sampling. In frequency domain, this convolution can be seen as a multiplication with the Fourier transform of the sampling function. Therefore, as long as the sampling kernel does not remove frequencies, the reconstructed function can always be divided by the sampling kernel function again. This should cancel the effect of the sampling kernel. Such an operation can be performed after the reconstruction of an aliasing-free signal using the methods described in this paper. Of course, this supposes that the sampling kernel is known, space/time invariant, and not too ill-conditioned.

### B. Images and higher-dimensional signals

The above approaches were described for one-dimensional functions. However, all the methods can be extended to higher-dimensional functions. Nothing in the descriptions is limited to a one-dimensional signal, so the algorithms could be directly applied to images or higher-dimensional

signals. Of course, the complexity increases fast with the dimensionality. While a minimum of $\lceil L/N \rceil + 1$ sets of samples are required for a 1D signal using the rank-based method, and $\lceil (L-1)/(N-1) \rceil$ sets of samples for the projection-based method, we need at least $\lceil L/N \rceil^2 + 1$ 2D images for the rank-based method, and $\lceil (L^2-1)/(N^2-1) \rceil$ images for the projection-based method. As signals can be shifted along each of their dimensions, the number of offsets (which is also the dimensionality of the search space) also increases rapidly. Examples of the above methods on images are shown in Section VIII.

### *C. Minimization*

The main difficulty with the methods described in Sections IV and V is to find the global minimum of the objective functions. A good approximation of the offsets $\mathbf{t}$ is required for an approach like gradient descent to converge to the global minimum. Such an approximation could be obtained by evaluating the error functions from (33) or (35) on a uniform (dense) grid of possible values $\hat{\mathbf{t}}$. The global minimum can then be found with high probability close to the smallest value obtained on this uniform grid. However, this is a computationally very intensive method. It only makes sense if the number of sets of samples $M$ - and therefore also the number of offsets - is small, or when a first estimate of the offsets is available. Although this may sound like moving the problem to obtaining such a first estimate, it is a reasonable assumption. For example in super-resolution imaging, the shifts between the images are generally very small. Therefore, our search could be restricted to a small area around the origin (for example $10 \times 10$ pixels). Similarly, a first estimate could also be obtained from another registration method that does not use the aliasing. Such a method is described in the next subsection. Other approaches are heuristic methods that rely on the lines that can be seen in Figure 9(b), 11(b) and 12(b). They will also be described in the next subsection.

During the minimization, it would be interesting to know whether the global minimum has been found, or whether the current optimum is a local minimum. Such an indication can be obtained by evaluating (33) or (35) for the current value of $\hat{\mathbf{t}}$. Even for noisy measurements, there is a large difference between the average values of these functions and their value at the global minimum. This can therefore be used to check whether the algorithm has converged, and whether the result is reliable.

### *D. Heuristic approaches*

The high computational complexity of the algorithms from Sections IV and V is mainly due to the coupling between the offsets, i.e. the need to search the different offsets between the signals jointly. The joint minimum is not necessarily located at the intersection of the minima from individual optimizations.

*Algorithm 6.1 (Hierarchical approach):* If the sets of samples are images, we know that their coefficients are not arbitrary Gaussian random variables. In general, the amplitude of the Fourier transform of a natural image decays like $1/f$ [21]. We can therefore assume that a good estimate for the offsets can be obtained from the low frequencies of the sampled sets. For these frequencies, the aliased coefficients are much smaller than the base spectrum coefficients, and can be neglected in a first estimate. We used the method from Vandewalle et al. [22] for this. Once we have such an initial estimate, we can use the methods from Section IV or V using a gradient descent algorithm to obtain a more precise estimate for the motion parameters, taking the aliasing into account. □

*Algorithm 6.2 (Search for lines in the objective function):* From the horizontal, vertical and diagonal lines that are often visible in figures like Figure 9(b), 11(b) and 12(b), it can be seen that an independent pairwise alignment often works. Such lines correspond to pairwise registrations of the first and second, first and third, and second and third signals, respectively. A heuristic algorithm has therefore been developed to search for these pairwise alignments without evaluating the complete $M$-dimensional grid. A horizontal line for example is searched by evaluating the objective function on five different vertical lines, and searching the minimum of the average values. The estimated offset can then be found at the intersection of the different lines. Even if, in a setup like the one shown in Figure 11(b), only one line can easily be found, this still results in a considerable reduction of the complexity. As it can also be seen from the results in Section VIII, this method performs very well in a setup with low noise values, but degrades rapidly when more noise is added. □

*Algorithm 6.3 (Keeping the $P$ best pairwise alignments):* Another heuristic approach based on this independence, is to search a fixed number $P$ of local minima for the pairwise registration between the first set of samples and each of the other sets. When performing such a pairwise registration between the first and the $m$-th set of samples, all the other offset values are kept constant. For each pair, a vector of $P$ possible offset values $\begin{pmatrix} \hat{t}_{m_1} & \hat{t}_{m_2} & \cdots & \hat{t}_{m_P} \end{pmatrix}$ is obtained. Next, the best combination $\hat{\mathbf{t}}$ of these pairwise local minima is searched among all possible combinations. The global minimum can then be searched in the neighbourhood of this value of $\hat{\mathbf{t}}$. Keeping only a single minimum for each of the pairwise alignments ($P = 1$) is generally not sufficient, due to the different approximations made by pairwise registration. However, by trying all the combinations of the best $P = 5$ pairwise alignments, the algorithm typically converges to the correct result. □

## VII. Complexity

In this section, we discuss the computational complexity of the different methods. For each of the methods described above, the complexity of computing the offsets $\mathbf{t}$ can be written as the number of times the error function has to be evaluated multiplied with the number of operations required to

evaluate the error function. In this analysis, we will assume that the variables $N$ and $L$ grow at the same rate, and $M$ grows at a lower rate. The number of operations required for the reconstruction is the same for all the algorithms.

Let us start by analyzing the complexity of the error function evaluation both for the matrix rank method (Section IV) and the projection method (Section V). In the matrix rank method, the smallest singular value of the matrix $\mathbf{Y}_{\hat{\mathbf{t}}}^D$ has to be computed for every error evaluation. The construction of $\mathbf{Y}_{\hat{\mathbf{t}}}^D$ for a specific set of offsets $\hat{\mathbf{t}}$ requires $O((M-1)N)$ operations. To compute the smallest singular value, we use the inverse power method on $\mathbf{Y}_{\hat{\mathbf{t}}}^{D*}\mathbf{Y}_{\hat{\mathbf{t}}}^D$ [20]. This matrix multiplication involves $O(M^2N)$ operations, and the inverse power method itself requires $O(M^3 + aM^2)$ operations, where $a$ is the number of iterations required for the method to converge. As we assume that the difference with the second smallest singular value is large, convergence will be fast. So the total number of operations becomes

$$C_{\text{rank}} = O((M-1)N + M^2N + M^3 + aM^2) = O(M^3 + M^2N). \tag{37}$$

For the projection method, we need to compute the projection of the sample vector $\mathbf{y}$ onto the space spanned by the columns of $\mathbf{\Phi}_{\hat{\mathbf{t}}}$. This can be done using (34). In the general case, for an arbitrary basis, such a projection has complexity

$$C_{\text{proj}} = O(MNL^2 + L^3). \tag{38}$$

However, if enough storage space is available, these projection matrices can be precomputed, because they do not depend on the actual signal. In that case, the complexity is reduced to

$$C'_{\text{proj}} = O(M^2N^2). \tag{39}$$

For specific types of bases, the complexity can be strongly reduced.

*Example 7.1 (Fourier basis):* For the Fourier basis, the blocks $\mathbf{\Phi}_{\hat{t}_m}$ of the matrix $\mathbf{\Phi}_{\hat{\mathbf{t}}}$ can be divided into an $N \times L$ IDFT inverse Fourier transform block $\mathbf{F}$, multiplied by an $L \times L$ diagonal offset matrix $\mathbf{D}_{t_m}$: $\mathbf{\Phi}_{\hat{t}_m} = \mathbf{F}\mathbf{D}_{t_m}$ (see also (13)). Using this decomposition, we can simplify the projection formula

$$\hat{\mathbf{y}} = \mathbf{P}_{\mathbf{\Phi}_{\hat{\mathbf{t}}}}\mathbf{y} = \mathbf{\Phi}_{\hat{\mathbf{t}}}(\mathbf{\Phi}_{\hat{\mathbf{t}}}^*\mathbf{\Phi}_{\hat{\mathbf{t}}})^{-1}\mathbf{\Phi}_{\hat{\mathbf{t}}}^*\mathbf{y}. \tag{40}$$

The multiplication $\mathbf{\Phi}_{\hat{\mathbf{t}}}^*\mathbf{y}$ can be decomposed into $M$ multiplications $\mathbf{D}_{\hat{t}_m}^*\mathbf{F}^*\mathbf{y}$, where $\mathbf{F}^*\mathbf{y}$ is the Fourier transform of $\mathbf{y}$. This Fourier transform can be computed in advance, because it is independent of the offsets $\mathbf{t}$. Only the multiplications with the diagonal matrices $\mathbf{D}_{\hat{t}_m}$ remain, which require $ML$ operations. The matrix $\mathbf{\Phi}_{\hat{\mathbf{t}}}^*\mathbf{\Phi}_{\hat{\mathbf{t}}}$ has nonzero elements only on the main diagonal and on the $iN$-th diagonals, due to the orthogonality of the Fourier vectors. Its inverse, $(\mathbf{\Phi}_{\hat{\mathbf{t}}}^*\mathbf{\Phi}_{\hat{\mathbf{t}}})^{-1}$, can therefore be computed efficiently in $O(L^3/N^2) = O(S^2L)$ operations. Finally, the multiplication with the first

matrix $\mathbf{\Phi}_{\hat{\mathbf{t}}}$ can be decomposed into $M$ multiplications with diagonal matrices, and $M$ inverse DFTs. As the error can as well be computed in the Fourier domain (using Parseval's theorem), we only need to multiply with the diagonal matrices, which requires again $ML$ operations. The overall complexity can then be approximated as

$$C_{\text{proj},F} = O(ML + S^2L + ML) = O(S^2L), \tag{41}$$

where $S$ is the maximum number of overlapping spectral components, which is typically of the same order as $M$. □

For the standard algorithm described in Section VI-C, the error function needs to be evaluated first on a regular grid. Assuming that no estimate of the offsets is available, this requires $N^{(M-1)}$ error evaluations. Next, a standard minimization algorithm is applied near the minimum value that was obtained from the first part. Let us call the number of error evaluations in this part $E$ (typical values are around $E = 30$). This is negligible compared to the number of evaluations on the regular grid, and the total number of evaluations with this standard algorithm can therefore be approximated as

$$O(N^{M-1} + E) = O(N^{M-1}). \tag{42}$$

For the hierarchical method from Algorithm 6.1, we need only $E$ evaluations. Of course, the complexity for computing the initial estimate needs to be added to this.

With the heuristic line search of Algorithm 6.2, the number of function evaluations is difficult to predict, as multiple functions are executed consecutively to find the minimum along horizontal, vertical, or diagonal lines. When the value of the error function is below a certain threshold, the algorithm is stopped. The maximum number of evaluations is

$$O(21N + (27 + 7^{M-1} + N)E) = O((7^{M-1} + N)E). \tag{43}$$

Algorithm 6.3, using pairwise alignments, requires

$$O((M-1)N + 5^{M-1} + E) = O(MN + 5^{M-1}) \tag{44}$$

error function evaluations. The complexity of the different algorithms is summarized in Table I.

For the reconstruction, the set of linear equations from (6) has to be solved. The complexity of this operation is $O(MNL^2 + L^3)$.

## VIII. Results

The above algorithms are tested and compared in a number of simulations. The simulations are performed on random bandlimited signals. A Fourier basis is used, such that both methods can be compared directly. The Fourier coefficients of the signals are generated as a white Gaussian random

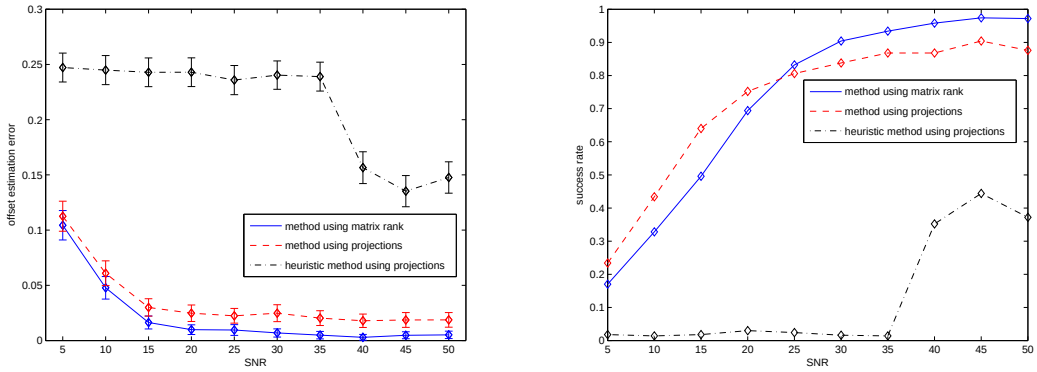
TABLE I

COMPUTATIONAL COMPLEXITY FOR THE DIFFERENT METHODS ON 1D SIGNALS. THE TOTAL COMPLEXITY IS OBTAINED BY MULTIPLYING THE NUMBER OF FUNCTION EVALUATIONS WITH THE COMPLEXITY OF A SINGLE FUNCTION EVALUATION. A WORST CASE SCENARIO WAS USED FOR ALG. 6.2. THE COST TO OBTAIN THE INITIAL ESTIMATE IN ALG. 6.1 IS DENOTED AS $C_{\text{INIT}}$.

| | matrix rank method<br>$M^3+M^2N$ | projection method<br>$MNL^2+L^3$ | projection method (Fourier)<br>$S^2L$ |
|---|---|---|---|
| regular grid<br>$N^{M-1}$ eval. | $N^{M-1}(M^3+M^2N)$ | $N^{M-1}(MNL^2+L^3)$ | $N^M S^2 L$ |
| hier. method (Alg. 6.1)<br>$E$ eval. + $C_{\text{init}}$ | $(M^3+M^2N)E+C_{\text{init}}$ | $(MNL^2+L^3)E+C_{\text{init}}$ | $S^2LE+C_{\text{init}}$ |
| line search (Alg. 6.2)<br>$(7^{M-1}+N)E$ eval. | $(7^{M-1}+N)(M^3+M^2N)E$ | $(7^{M-1}+N)(MNL^2+L^3)E$ | $(7^{M-1}+N)S^2LE$ |
| pairw. al. (Alg. 6.3)<br>$MN+5^{M-1}$ eval. | $(MN+5^{M-1})(M^3+M^2N)$ | $(MN+5^{M-1})(MNL^2+L^3)$ | $(MN+5^{M-1})S^2L$ |

process. Hence, the resulting (time domain) signals also form white Gaussian random processes. A number of simulations are performed with different random offset values, and different amounts of additive white Gaussian noise. The performance of the different algorithms on 1D signals is compared in Figures 14 and 15. Figures 14(a) and 14(b) show the mean absolute error in the shift estimates and the success rate as a function of the SNR, respectively. The success rate of the methods is defined as the relative number of simulations in which the error on the registration is smaller than $10^{-3}$. All the results were averaged over 250 simulations. Parameter values of $M = 3$, $L = 81$, and $N = 41$ were used in all the simulations. An offset value of 1 corresponds to a shift over the whole signal period. We can see that the method using matrix rank from Section IV performs slightly better than the method using projections described in Section V. The heuristic method performs clearly worse than the two methods that sample a uniform grid first and then perform a minimization around the minimum value. It works well for high SNR values, but breaks down for SNRs around 35 dB.

The absolute error in the shift estimates and the success rate are plotted as a function of the number of samples per set $N$ in Figure 15. In these simulations, parameters $M = 3$ and $L = 81$ were used, and for $N$, values from 25 to 45 were taken. No noise was added for these simulations. Again, the results were averaged over 250 simulations. The performance of the different algorithms increases with increasing number of samples. This is what we expected, as an increasing number of samples per set gives increasing amounts of information for the same number of unknowns. For the matrix rank




(a) (b)

Fig. 14. Results of the different algorithms as a function of the signal-to-noise ratio (SNR). The matrix rank algorithm from Section IV performs slightly better than the projection algorithm from Section V. The heuristic Algorithm 6.2 only performs good for high SNRs. Parameter values of $M = 3$, $L = 81$, and $N = 41$ were used. (a) Offset estimation error as a function of the SNR of the sampled signals. An offset error of 1 corresponds to a shift over the entire signal. (b) Success rate of the registration and reconstruction as a function of the SNR.

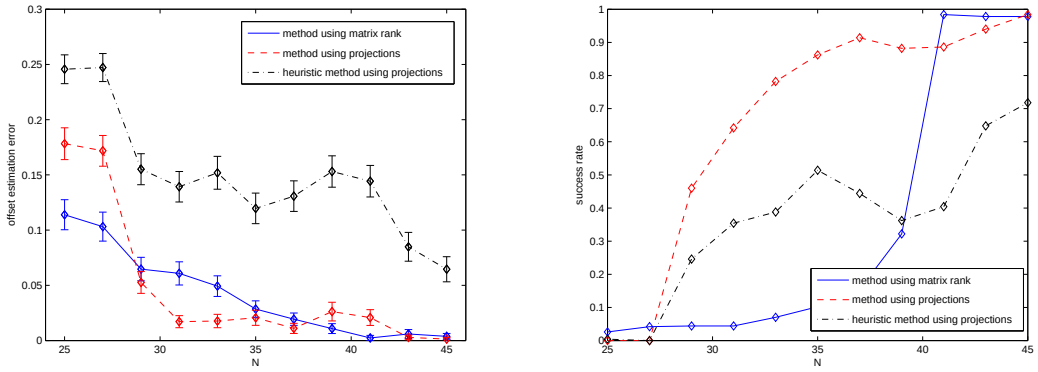
algorithm from Section IV, the minimum number of samples per set that are required in this setup is $N = 41$, while for the projection algorithm from Section V, $N = 28$ samples per set are needed. This explains the better performance of the projection algorithm for $28 \leq N < 41$. With more than 41 samples per set, both algorithms perform very well. When $N < 28$ samples are available per set, all algorithms have low accuracy.

Our algorithms are also applied to some 2D signals (images). In this case, the signals were undersampled by only two, because of the high complexity (see the discussion in Section VI). The results can be seen in Figure 16. A double resolution image is perfectly reconstructed from a set of low resolution images, both using the rank-based algorithm from Section IV (Figures 16(a) and 16(b)) and the projection-based algorithm from Section V (Figures 16(c) and 16(d)).

In all the simulations, we use periodic signals. This is not the case in most real applications, but we can generally assume that the shift between the sets of samples is small. The differences between the signals due to their aperiodicity are therefore small, and can be neglected.

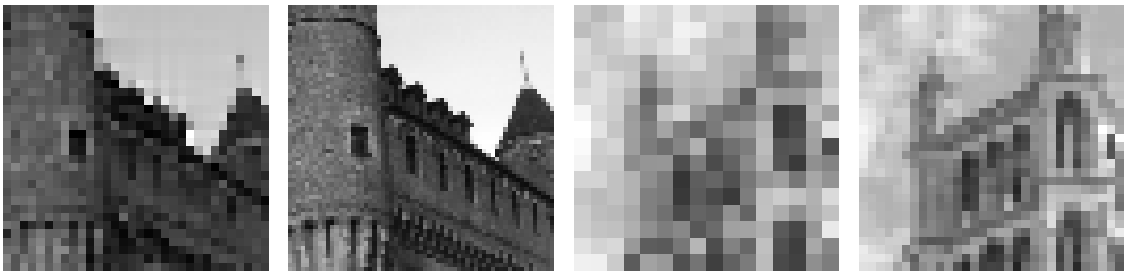
## IX. Conclusions

In this paper, we presented a theory of super-resolution from unregistered aliased sets of samples. We formulated the problem mathematically, and proved that the solution is unique if enough sets of samples are available. We described two methods to compute this unique solution. In the first method, the rank of a modified sample matrix is checked. The second method uses projections onto subspaces




(a) (b)

Fig. 15. Results of the different algorithms as a function of the number of samples $N$. Parameter values $M = 3$ and $L = 81$ were used in all simulations, and for $N$, values from 25 to 45 were used. No noise was added in this setup. The algorithm from Section V performs better than the algorithm from Section IV if the number of samples per set $N$ is not sufficiently high. The heuristic Algorithm 6.2 performs worse than the other algorithms. (a) Offset estimation error as a function of the number of samples per set $N$. An offset error of 1 corresponds to a shift over the entire signal. (b) Success rate of the registration and reconstruction as a function of the number of samples per set $N$.



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

Fig. 16. Results of the different algorithms on images. (a) One of the five $32 \times 32$ images used as input for Algorithm 6.3, with the rank-based method from Section IV. (b) Reconstructed $63 \times 63$ image from the images from (a). (c) One of the five $8 \times 8$ images used as input for Algorithm 6.1, with the projection-based method from Section V. (d) Reconstructed $15 \times 15$ image from the images from (c).

to compute the offsets between the different sets of samples. The first method is only applicable to bandlimited signals, while the second is applicable to any type of finite Hilbert space signals. Both methods can be used for the reconstruction of one or two-dimensional signals from multiple sets of aliased samples. The main limitation of these methods is their computational complexity. They are therefore mainly applicable in domains that do not require real-time reconstruction. For example in satellite imaging, a very high resolution is desired, even if this requires some computational effort. Future work is mainly oriented towards reducing the computational complexity of the methods.

## APPENDIX

**Proof of Theorem 3.3:** Let us call the two sets of samples $\mathbf{y}_0$ and $\mathbf{y}_1$. The samples satisfy

$$\left(\begin{array}{c}\mathbf{y}_0\\ \hline \mathbf{y}_1\end{array}\right) = \left(\begin{array}{c} y_0(0)\\ y_0(1)\\ \hline y_1(0)\\ y_1(1)\end{array}\right) = \left(\begin{array}{ccc} 0 & 0 & 1\\ 1 & 1 & 1\\ \hline t^2 & t & 1\\ (1+t)^2 & 1+t & 1\end{array}\right)\left(\begin{array}{c} a\\ b\\ \\ c\end{array}\right). \tag{45}$$

Let us assume that, next to the desired solution $\{a, b, c, t\}$, there exists another solution $\{a', b', c', t'\}$ resulting in the same samples. We can therefore also write (45) for this solution, and obtain the following set of equations:

$$\begin{cases} c = c' \\ 0.25a + 0.5b + c = 0.25a' + 0.5b' + c' \\ t^2a + tb + c = t'^2a' + t'b' + c' \\ (0.5+t)^2a + (0.5+t)b + c = (0.5+t')^2a' + (0.5+t')b' + c'. \end{cases} \tag{46}$$

Eliminating $c$ and $c'$ from the equations, and rearranging the terms, we get

$$\begin{cases} c' = c \\ b' = 0.5a + b - 0.5a' \\ t^2a - t'^2a' + tb - t'b' = 0 \\ t^2a - t'^2a' + ta - t'a' + 0.25a - 0.25a' + bt - b't' + 0.5b - 0.5b' = 0. \end{cases} \tag{47}$$

From the second equation, we can compute $b'$ as a function of $a'$ and the terms of the first solution $\{a, b, c, t\}$. Similarly, after simplification, $d$ can be computed from the fourth equation:

$$a' = at/t'. \tag{48}$$

We assume that $t' \neq 0$, which is a reasonable assumption, because this would mean that the two sets of samples $\mathbf{y}_0$ and $\mathbf{y}_1$ coincide, and there are only two distinct samples. After filling in $d$ and $b'$, the third equation can be factorized as follows:

$$(t - t')(at + 0.5a + b) = 0. \tag{49}$$

From this equation, we can see that either $at + 0.5a + b = 0$, or $t = t'$. Because we assumed that the two solutions $\{a, b, c, t\}$ and $\{a', b', c', t'\}$ are different, the second solution can be discarded (it involves immediately $a' = a$, $b' = b$, and $c' = c$).

It is therefore only when $at + 0.5a + b = 0$ that other solutions exist, which are given by

$$\begin{cases} a' = at/t' \\ b' = 0.5a + b - 0.5at/t' \\ c' = c, \end{cases} \tag{50}$$

with $t'$ a free parameter. In all the other cases, the solution $\{a, b, c, t\}$ is unique. ■

## REFERENCES


[1] P. Vandewalle, L. Sbaiz, J. Vandewalle, and M. Vetterli, "How to take advantage of aliasing in bandlimited signals," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 3, May 2004, pp. 948–951.

[2] P. Vandewalle, L. Sbaiz, M. Vetterli, and S. Süsstrunk, "Super-resolution from highly undersampled images," in *IEEE International Conference on Image Processing*, September 2005.

[3] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423, July 1948.

[4] M. Unser, "Sampling—50 years after Shannon," *Proceedings of the IEEE*, vol. 88, no. 4, pp. 569–587, April 2000.

[5] F. Marvasti, Ed., *Nonuniform sampling: theory and practice*. Springer, 2001.

[6] T. Strohmer, "Computationally attractive reconstruction of bandlimited images from irregular samples," *IEEE Transactions on Image Processing*, vol. 6, no. 4, pp. 540–548, April 1997.

[7] P. Marziliano and M. Vetterli, "Reconstruction of irregularly sampled discrete-time bandlimited signals with unknown sampling locations," *IEEE Transactions on Signal Processing*, vol. 48, no. 12, pp. 3462–3471, December 2000.

[8] A. Papoulis, "Generalized sampling expansion," *IEEE Transactions on Circuits and Systems*, vol. 24, no. 11, pp. 652–654, November 1977.

[9] M. Unser and J. Zerubia, "Generalized sampling without bandlimiting constraints," in *Proceedings IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 3, April 1997, pp. 2113–2116.

[10] ——, "Generalized sampling: stability and performance analysis," *IEEE Transactions on Signal Processing*, vol. 45, no. 12, pp. 2941–2950, December 1997.

[11] R. Y. Tsai and T. S. Huang, "Multiframe image restoration and registration," in *Advances in Computer Vision and Image Processing*, T. S. Huang, Ed. JAI Press, 1984, vol. 1, pp. 317–339.

[12] S. Borman and R. Stevenson, "Spatial resolution enhancement of low-resolution image sequences - a comprehensive review with directions for future research," University of Notre Dame, Tech. Rep., 1998.

[13] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, "Advances and challenges in super-resolution," *International Journal of Imaging Systems and Technology*, vol. 14, no. 2, pp. 47–57, Aug. 2004.

[14] *IEEE Signal Processing Magazine, Special Issue on Super-Resolution*, vol. 20, no. 3, May 2003.

[15] *EURASIP Journal on Applied Signal Processing, Special Issue on Super-Resolution*, 2006.

[16] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *IEEE Transactions on Signal Processing*, vol. 50, no. 6, pp. 1417–1428, June 2002.

[17] S. Baker and T. Kanade, "Limits on super-resolution and how to break them," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 9, pp. 1167–1183, September 2002.

[18] M. Schwab, M. Karrenbach, and J. Claerbout, "Making scientific computations reproducible," *Computing in Science & Engineering*, vol. 2, no. 6, pp. 61–67, November-December 2000.

[19] P. Marziliano, "Sampling innovations," Ph.D. dissertation, Ecole Polytechnique Fédérale de Lausanne (EPFL), 2001, no. 2369.

[20] G. Strang, *Linear algebra and its applications*, 3rd ed. Saunders College Publishing, 1988.

[21] A. Torralba and A. Oliva, "Statistics of natural image categories," *Network: Computation in Neural Systems*, vol. 14, no. 3, pp. 391–412, Aug. 2003.

[22] P. Vandewalle, S. Süsstrunk, and M. Vetterli, "A frequency domain approach to registration of aliased images with application to super-resolution," *accepted to EURASIP Journal on Applied Signal Processing, Special Issue on Super-Resolution Imaging*, 2005.