(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
4 December 2014 (04.12.2014)

WIPO | PCT

(10) International Publication Number
WO 2014/191771 A2

(51) International Patent Classification:
G06F 17/14 (2006.01)

(21) International Application Number:
PCT/GB2014/051671

(22) International Filing Date:
30 May 2014 (30.05.2014)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
1309689.6 30 May 2013 (30.05.2013) GB

(71) Applicants *(for all designated States except US)*: **IMPERIAL INNOVATIONS LIMITED** [GB/GB]; 52 Princes Gate, Exhibition Road, London SW7 2PG (GB). **THE CHINESE UNIVERSITY OF HONG KONG** [CN/CN]; Shatin, NT, Hong Kong (CN).

(72) Inventors; and
(71) Applicants *(for US only)*: **ORIGUEN, Jose Antonio** [ES/ES]; Coldn De Larreategni 33, E-48009 Bilbao (ES). **DRAGOTTI, Pier Luigi** [IT/GB]; 69 Haydon Park Road, London SW19 8JH (GB). **BLU, Thierry Albert Jean-Louis** [FR/CN]; 1/2F 28C Tai Yeung Che, Lam Tsuen, Tai Po N.T., Hong Kong (CN).

(74) Agent: **LEACH, Sean**; Mathys & Squire LLP, The Shard, 32 London Bridge Street, London SE1 9SG (GB).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

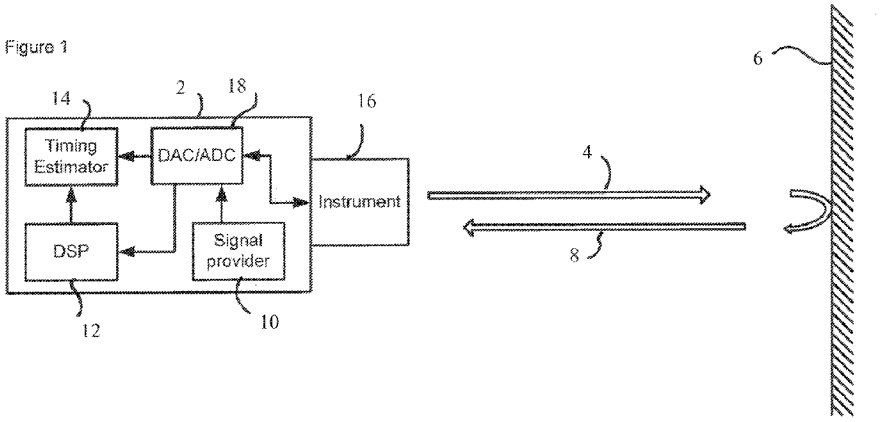
(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

(54) Title: METHOD AND APPARATUS

Figure 1



(57) Abstract: A signal processing method for estimating a frequency domain representation of signal from a series of samples distorted by an instrument function, the method comprising obtaining the series of samples; obtaining a set of coefficients that fit a set of basis functions to a complex exponential function wherein the set of basis functions comprises a plurality of basis functions each defined by a shifted version of the instrument function in a signal domain; estimating the frequency domain representation of the signal based on the series of samples and the coefficients. This is wherein the estimate of the instrument function is based on a characterisation of the instrument function in the frequency domain at frequencies associated with the complex exponential function.

WO 2014/191771 A2

Method and Apparatus

The present invention relates to sampling and reconstruction of signals, and more particularly to the sampling and reconstruction of non-bandlimited signals, such as signals with a finite rate of innovation, FRI.

The fundamental problem of all sampling is to estimate an original signal from a discrete set of samples of that signal. Typically, the signal is distorted by the instrument which collects the measurement, for example by convolving the true signal with the sampling kernel of the instrument. In most signal processing, this distorted signal is then converted into discrete samples. It is well known that the signal domain spacing of the samples imposes a limit on the frequency domain bandwidth of the measurement.

Any particular signal collecting instrument has physical characteristics which effect the measurement performed by that instrument. These characteristics can distort the original signal before it is sampled. The nature of this distortion can be characterised by an instrument function. One example of such an instrument function is the point spread function of an optical system. This instrument function describes the distorting effect that the instrument has on the original signal before it is sampled.

There are typically two approaches to improving the accuracy of estimation of a signal based on samples of a measurement that signal. The first is to increase the sampling bandwidth (increase the sampling frequency). The second is to reduce the distorting effect of the instrument function, to take the example of an optical system, this could be done by reducing the spatial extent of the point spread function.

For certain classes of signal, such as signals with a finite rate of innovation, FRI, one theoretical approach to the problem of signal recovery assumes that the instrument function belongs to a class of functions known as exponential reproducing kernels. This class of functions is defined by the ability of each such kernel to reconstruct an arbitrary complex

exponential function from a linear combination of shifted representations of that kernel.

In real practical instruments, the instrument function is determined by practical physical limitations of the instrument, for example in an optical system the aperture and other components in the optical train define the instrument function. Therefore, the instrument functions of real instruments generally are not exponential reproducing kernels. In addition, the theoretical approaches developed from FRI reconstruction using exponential reproducing kernels suffer from very poor stability in the presence of noise, e.g. deviations of the original signal from model behaviour. Of course, noise is inherent in any practical physical measurement. Therefore, although they have been known in the art for many years, these theoretical methods are not applicable to real world practical systems.

Aspects of the present disclosure are directed to the practical sampling, frequency space estimation, and recovery of, signals in practical real world systems, and in the presence of measurement noise. The signals may comprise FRI signals and other types of signal.

Embodiments of the present disclosure address this problem by modifying the discrete samples of a signal based on knowledge of the instrument function to provide a frequency domain estimate of that signal. In particular, the samples are scaled by a set of coefficients, Cmn. These coefficients are selected to fit a set of basis functions, hn, to a complex exponential function. Each of these basis functions is defined by a signal domain representation of the instrument function, shifted in the signal domain.

The versions of the instrument function are based on a characterisation of the instrument function in the frequency domain at selected frequencies, αm. These frequencies are also associated with the complex exponential function used to select the coefficients Cmn. The complex exponential function used to select the coefficients Cmn may comprise $\exp\{-\alpha mt\}$, where t indicates a signal domain displacement. For example, in the case of a time domain function, the parameter t indicates a time delay, whilst, in the case of a spatial function, t, would indicate a spatial displacement rather than a time delay.

The frequencies αm may be complex frequencies, in the sense of Laplace transforms, where the imaginary part of the complex frequency indicates a frequency of an oscillatory component and the real part of the complex frequency indicates a damping or driving of the signal.

Embodiments of the disclosure relate to methods of configuring a signal processing apparatus 12 by selecting these frequencies and generating the corresponding coefficients Cmn for the instrument based on an estimate of an instrument function.

To assist in understanding the invention, a selection of particular implementations of the invention is discussed below. The first of these implementations relates to range finding based on the time of flight of a series of pulses. This example is chosen because it relates to the estimation of a signal comprising a sequence of pulses. As will be appreciated by the skilled addressee in the context of the present disclosure, most practical signals and a train of pulses in particular, may be described by the timing and amplitude of a series of single points (e.g. a train of Dirac-delta functions). Therefore, although this example is provided as one way of understanding the present disclosure, it will be understood that the apparatus and methods of the disclosure can be applied to signal processing of any signal.

The present disclosure finds particular application in FRI signals, but may also be applied to other types of signal. FRI signals, as the name implies have a finite number of degrees of freedom per unit time, and they may be either periodic or aperiodic. Examples of FRI signals include streams of Dirac delta functions, non-uniform splines, sinusoids, and piecewise polynomials.

Each of these types of signal may be characterised by their "innovation parameters". For example, the innovation parameters of a train of Dirac delta functions are the location, tk, in the signal domain (e.g timing, or position) of the delta functions, and their amplitudes ak. This class of FRI signal is used as the basis of most of the discussion which follows, because

nearly all signals can be reduced to a representation as a series of delta functions. It will however be appreciated that the teaching of the present disclosure is capable of application to any type of signal.

In an aspect there is provided a signal processing method for estimating a frequency domain representation (X) of signal (x) from a series of samples (yn) distorted by an instrument function, (h) the method comprising: obtaining the series of samples (yn); obtaining a set of coefficients (Cmn) that fit a set of basis functions (hn) to a complex exponential function wherein the set of basis functions comprises a plurality of basis functions, each defined by a shifted version of the instrument function, in a signal domain; estimating the frequency domain representation (X) of the signal (x) based on the series of samples (y) and the coefficients (Cmn); wherein the estimate of the instrument function is based on a characterisation of the instrument function in the frequency domain at frequencies associated with the complex exponential function.

In an aspect there is also provided a digital signal processing apparatus comprising: a data obtainer configured to obtain a series of samples, yn, of a signal, y(t), that corresponds to a signal x distorted by an instrument function, h; and a data scaler, configured to scale the series of samples, yn, using coefficients Cmn wherein the coefficients cmn are selected to approximate a complex exponential function from a set of basis functions, hn, defined by the instrument function shifted by a number of samples, n; a data provider for providing a frequency domain representation, X, of the signal, x, based on the scaled samples; wherein the instrument function is based on a characterisation of the instrument function in the frequency domain at frequencies associated with the complex exponential function.

The samples of the signal may be separated in the signal domain by an inter-sample interval, or sample period. Each basis function may be shifted in the signal domain with respect to the others by an integer number, n, of inter-sample intervals, in these examples, the fit can comprise a projection into the space defined by the basis functions. In some examples, each basis function may be shifted in the signal domain by a non-integer number of inter-sample

intervals, and in these examples the fit can be performed by reducing a merit function.

The coefficients (cmn) may comprise a frequency domain representation, (Hn) of a respective corresponding one of the basis functions (hn). The complex exponential function may comprise signal domain oscillatory components, and the frequencies at which the instrument function (h) is characterised comprise the frequencies of the oscillatory components. The complex exponential function may consist solely of signal domain oscillatory components. This has been found to provide signal estimation that is particularly robust to noise.

In some possibilities the coefficients comprise the reciprocal of the frequency domain representation, Hn, of the basis functions hn at the selected frequencies. In some possibilities the coefficients comprise reversed versions of the frequency domain representation, Hn, of the basis functions hn. The coefficients may be scaled by the z-transform of the autocorrelation of the instrument function in the signal domain.

The selected frequencies may be evenly spaced in the frequency domain, and in some cases the frequencies are selected to reduce the variance in the magnitude of the instrument function at the frequencies whilst increasing the spacing between the frequencies. This has also been found to provide improved resilience to noise. In some possibilities, the signal has K degrees of freedom, where K is an integer, where K is an integer, n=0, 1,…,P, and $P \leq 2K-1$.

In some possibilities the signal to be estimated belongs to a class of finite rate innovation, FRI, signal, and the frequencies are selected based on the performance of a maximum likelihood estimator of innovation parameters of that class of signal. This may be achieved by reducing a Cramer-Rao bound.

In some possibilities the method comprises obtaining an error in the representation, provided by the coefficients (cmn) and the basis functions (hn), of the complex exponential function; and, based on this error, determining whether to re-estimate the frequency domain representation (X) of the signal. Determining whether to re-estimate the frequency domain

representation (X) of the signal may comprise obtaining an estimate of noise in the original signal, and performing a comparison based on the estimated noise and the error in the representation of the complex exponential function.

In some possibilities, re-estimating the frequency domain representation of the signal comprises determining a filter configured to minimise the energy of the frequency domain estimation representation (X) of the signal (x), and adjusting the frequency domain representation based on the filter. This may be provided by an approximation to an annihilating filter. In some possibilities adjusting based on the filter comprises obtaining an estimate of parameters describing the location, in the signal domain, of the signal using the roots of the filter.

The annihilating filter method need not be used, for example in some possibilities re-estimating the frequency domain representation of the signal comprises obtaining an estimate of parameters describing the location, in the signal domain, of the signal by deriving eigenvalues from a matrix pencil problem.

In some possibilities the signal comprises an image, and the method further comprising scaling the frequency domain representation of the image based on the frequencies associated with the complex exponential function, determining a set of filter coefficients, bkj, configured to minimise the energy of the frequency domain representation of the image, and determining a boundary in the image from the filter coefficients. As will be appreciated, image data may be treated by estimating a frequency domain representation of each row, and then estimating the frequency domain representation of each column.

In some possibilities the image processing methods may be used for coregistration of images, in these possibilities the method comprises: determining a transform for coregistering the image with a second image based on the boundary, and based on a boundary obtained for the second image. There is also provided a method of enhancing image resolution, comprising identifying at least one boundary for each of a plurality of images, and determining, for each

boundary, a transform for coregistering each boundary with a selected reference boundary, and determining a combined image based on the plurality of images and the transforms. Determining the combined image may comprise determining the value of pixels in the combined image based on samples from at least one of the images, and the transform associated with that image.

In an aspect there is provided a method of calibrating a signal acquisition device comprising a signal acquisition instrument, the method comprising: obtaining an estimate of the transfer function of the instrument; selecting a set of frequencies, that are evenly spaced in the frequency domain; and storing, in a memory of the device, coefficients (Cmn) based on samples of the transfer function at the selected frequencies. These frequencies may be selected to reduce the variance of the magnitude of the instrument function at the frequencies, whilst increasing the spacing between the frequencies. In some possibilities the signal acquisition device is adapted for acquiring finite rate innovation, FRI, signals, and the frequencies are selected based on the performance of a maximum likelihood estimator of innovation parameters of that class of signal.

The frequencies may be selected to reduce a Cramer-Rao bound determined based on those frequencies, as discussed in Annex 1.

Embodiments of the disclosure will now be described, by way of example only, with reference to the accompanying drawings, in which:

Figure 1 shows a schematic diagram of a range finding apparatus comprising a signal processing apparatus;

Figure 2 illustrates the process of sampling a signal;

Figure 3 shows a schematic diagram of a range finding apparatus which shows the signal processing apparatus 12 of Figure 1 in more detail;

Figure 4 illustrates one method of operation of a signal processing apparatus 12, such as that illustrated in Figure 3;

Figure 5 illustrates an image processing method and apparatus that can be implemented using

the signal processing apparatus 12 of Figure 3 and/or the method illustrated in Figure 4;
Figure 6 illustrates a method of enhancing the resolution of an image; and
Figure 7 further exemplifies the method illustrated in Figure 6.

Figure 1 shows a range finding apparatus 2 for determining a distance based on time of flight of a signal.

The range finding apparatus 2 is arranged to direct a signal 4 against an object 6, to receive the signal 8 reflected back from the object, and to determine the distance between the apparatus 2 and the object 6 based on the time delay between transmission of the outgoing signal 4 and receipt of the reflected signal 8 at the apparatus 2.

The range finding apparatus 2 comprises a signal provider 10 configured to transmit an outgoing signal 4 to the object 6. The signal provider 10 also provides the outgoing signal to a timing estimator.

The range finding apparatus 2 also comprises an instrument for detecting the reflected (incoming) signals. An ADC 18 is coupled to the detection instrument 16 for converting signals received by the instrument 16 into digital samples.

The ADC 18 is coupled to provide the samples of the received signals to a signal processing apparatus 12, DSP. The DSP 12 is configured to process the samples and to provide the processed samples to the timing estimator 14 configured to estimate the timing of the incoming reflected signal, and to determine the delay between the outgoing signal and the reflected signal. This enables the range finding apparatus to determine the distance to the object 6 based on the time of flight of the signal.

As illustrated in Figure 2, the reflected signal 8 received by the instrument, comprises a time domain signal x(t). The detection instrument 16 distorts this incoming signal, for example by introducing delay, and attenuation. This distortion may be described by an instrument

function, e.g. the impulse response of the instrument 16. The signal, x(t), distorted by the detection instrument, y(t), is then sampled by the ADC 18 to provide a discrete representation y(n) of the distorted signal.

In order to determine the distance between the range finding apparatus 2 and the object 8, the range finding apparatus 2 determines the delay between the transmission of the outgoing signal 4, and receipt of the reflected signal 6 based on this quantised representation of the signal, y(n). The resolution of this measurement is limited by the distortion introduced by the instrument function, and by the resolution of the ADC. Accordingly, the problem of range estimation in this context reduces to the fundamental problem of all sampling – mentioned above, e.g. how to recover an original signal from a discrete set of samples of that signal.

The example of range finding based on a series of pulses is useful for understanding the present disclosure, and it will be appreciated by the skilled addressee in the context of the present disclosure that many practical signals, including trains of pulses, may be expressed as a series of delta functions. Therefore, although this example is described in the context of range estimation, the principles can be applied to any signal, and in particular to signals which can be expressed as a series of delta functions.

Figure 3 illustrates a more detailed example of a range finding apparatus as described above with reference to Figure 1 and Figure 2. The apparatus shown in Figure 3 comprises an instrument for obtaining a measurement of a signal, an ADC 18 for providing a series of samples of the measured signal, a signal processing apparatus 12, and a timing estimator.

The signal processing apparatus 12 is configured to receive the samples from the ADC, and to provide an estimate of a frequency domain representation of the signal measured by the instrument to the timing estimator. A signal domain representation of the signal can be derived from this frequency domain representation. Accordingly, the timing estimator 14 can estimate a time domain delay, between a transmitted signal and the signal domain representation based on the output from the signal processing apparatus.

The signal processing apparatus 12 may be provided separately from the remainder of the range finding apparatus of Figure 3. The signal processing apparatus 12 and its methods of operation may be applied in other contexts. This signal processing apparatus 12 and its operation will now be described in greater detail.

The signal processing apparatus 12 comprises a data obtainer 22 for obtaining samples, y(n), of the signal from an ADC. The data obtainer 22 is coupled to a data scaler 30, and the data scaler 30 is coupled to a coefficient provider 20. The data scaler 30 is also coupled to a data provider 34 for providing an output from the signal processing apparatus. The signal processing apparatus 12 further comprises a data tuner 32 that is coupled to the data provider 34, to the coefficient provider 20, and to the data obtainer 22.

The data obtainer 22 comprises a memory 24 for storing signal samples, and an input interface 26 to enable data to be provided into the memory 24 from a source of digital samples such as the ADC 18. The data obtainer 22 also comprises a coupling 28 to enable the data scaler 30 and the data tuner 32 to read signal samples from the memory 24.

The coefficient provider 20 comprises a memory 21 configured to store coefficients, Cmn, and to provide these coefficients to the data scaler. The coefficients, Cmn, are selected to approximate a complex exponential function by defining a linear combination of a set of basis functions, e.g.

$$\sum_n C_{mn} h_n(t) \cong \exp\{\alpha_m t\} \qquad (1)$$

The basis functions hn(t) are defined by shifted versions of the instrument function in the signal domain, hn. Where n indicates a shift in the signal domain, e.g. by a number of sample periods, n, and the frequencies αm may be complex frequencies where the imaginary part of the complex frequency indicates a frequency of an oscillatory component and the real part of

the complex frequency indicates a damping or driving (decay or growth).

In the example illustrated in Figure 3, the coefficient provider 20 comprises coefficients, Cmn, that are selected according to the relation:

$$C_{mn} = \frac{\exp\{\alpha_m n\}}{H(\alpha_m)} \qquad (2)$$

Where H(αm) comprises the frequency domain version of the instrument function, hn, at the frequencies αm. The frequencies, αm, are selected to reduce the variance in the magnitude of the instrument function at those frequencies, $|H(\alpha_m)|$, whilst increasing the spacing between the frequencies. As noted above, the parameter, n, relates to a shift in the signal domain.

The data scaler 30 is configured to obtain the signal samples, y(n) via the data obtainer 22. The data scaler 30 is configured to scale these samples using the coefficients, Cmn, and to sum the scaled samples to estimate a frequency domain representation of the signal, X(αm), according to the relation:

$$X(\alpha_m) = \sum_n C_{mn} y_n \qquad (3)$$

The data scaler 30 is configured to provide this estimate, X(αm), of the frequency domain representation of the signal to the data provider. For completeness at this stage, we note that it can be shown (see Annex 1) that this estimate X(αm) can be written as:

$$X(\alpha_m) = \sum_{k=0}^{K-1} x_k u_k{}^m - \zeta_m \qquad (4)$$

Where:

$u_k = \exp\left\{\lambda \frac{t_k}{T}\right\}$, $x_k = a_k \exp\left\{\alpha_0 \frac{t_k}{T}\right\}$, in which ak and tk are the innovation parameters of the signal (e.g. the timing, tk, and amplitude, ak, of a train of delta functions representing the signal). ζ_m represents an error in the estimate as defined below. The parameters α0 and λ are complex parameters which define the frequencies αm in the complex plane according to the relation:

$\alpha m=\alpha 0 + m\lambda$,

where m is an integer $0 \leq m \leq 2K-1$, and K defines the number of degrees of freedom of the signal.

The data provider 34 comprises a memory 35 for storing data, and a coupling 36 to enable the data scaler 30 and the data tuner 32 to read and write data into the memory 35. The data provider 34 also comprises an output interface 38 to enable data to be provided as output, e.g. to the timing estimator.

The data tuner 32 is configured to obtain an estimate, E, of the measurement noise from the data obtainer, and to obtain the estimate $X(\alpha m)$ from the data provider. The data tuner 32 is also configured to estimate the error, ζ_m, in the estimate $X(\alpha m)$ of the frequency domain representation of the signal according to the relation:

$$\zeta_m = \sum_{k=0}^{K-1} a_k \varepsilon_m \left(\frac{t_k}{T} \right) \qquad (5)$$

Where the signal to be estimated may be considered as a series of K delta functions, tk represents the location of these delta functions in the signal domain (e.g their timing) and ak represents their amplitudes. The amplitudes ak can be estimated from the frequency domain representation, $X(\alpha m)$, according to the relation defined in Equation 4, above (e.g. by setting $\zeta_m = 0$ and solving Equation 4 above based on the current estimate of $X(\alpha m)$).

The parameter εm, in Equation 5 represents the mismatch between the exponential function $\exp\{\alpha m\}$ and a model assembled from shifted versions of the instrument function in the signal domain, hn(t), scaled by the coefficients, Cmn, viz:

$$\varepsilon_m(t) = \left| \exp\{\alpha_m t\} - \sum_n C_{mn} h_n(t) \right| \qquad (6)$$

The data tuner 32 is configured to compare the estimate of the error, ζ_m, with the estimate of

the measurement noise, E, and to determine whether to re-estimate the signal based on this comparison.

If the signal is to be re-esimated, the data tuner 32 is operable to re-estimate the signal by determining a filter, F, that reduces the energy of the frequency domain estimation representation, X(αm), e.g. by selecting coefficients of the filter F that adjust the convolution integral F*X towards zero. This may be performed using a Matrix Pencil method as defined in Y. Hua and T. K. Sakar, "Matrix Pencil Method for Estimating Parameters of Exponentially Damped Undamped Sinusoids in Noise," IEEE Transactions on Acoustics, Speech and Signal Processing. vol. 38, pp. 814-824, May 1990, or by use of an annihilating filter approach as defined in M. Vetterli, P. Marliliano. and T. Blu. "Sampling signals with finite rate of innovation:' IEEE Transactions on Signal Processing, vol. 50, pp. 1417-1428, June 2002.

The data tuner 32 is configured to obtain the roots, uk, of the z-transform of the filter, F, (e.g. the points in the complex plane where the z-transform of F is zero valued), and to determine a new estimate of the timings, tk of the series of K delta functions based on the relation:

$$u_k = \exp\left\{\lambda \frac{t_k}{T}\right\} \qquad (7)$$

Where T is the sampling period, and λ represents the spacing of the frequencies αm in the complex plane. The data tuner 32 is further configured to determine new estimates of the amplitudes, ak, based on the new estimate of the timings, tk.

The data tuner 32 is further configured to re-estimate the frequency domain representation of the signal by adjusting an existing estimate X(αm) based according to the relation:

$$X(\alpha_m)^{i+i} = X(\alpha_m)^i + \sum_{k=0}^{K-1} a_k^i \varepsilon_m\left(\frac{t_k}{T}\right) \qquad (8)$$

Where the superscripts represent iteration number, e.g. X(αm)i+1 represents the re-estimated frequency domain representation of the signal at the (i+1)th iteration, and X(αm)i represents the existing estimate at the ith iteration. The data tuner 32 is configured to repeat the process of determining the filter, F, and re-estimation of the frequency domain representation of the

signal, X, until the parameters ak, tk converge. Convergence criteria can be selected based on the absolute or relative variation in the estimated parameters ak, tk between consecutive iterations, and/or based on an estimate of the measurement noise in the original signal (e.g. a determination as to whether the adjustment between iterations is less than the estimated measurement noise).

Operation of the signal processing apparatus 12 will now be described with reference to Figure 4.

In operation, the data scaler 30 obtains 100 a series of samples yn from the data obtainer 22 and obtains 102 the coefficients Cmn from the coefficient provider 20. The data scaler 30 scales 104 the samples using the coefficients to provide an estimate of the frequency domain representation of the signal as in Equation 3, above.

The data tuner 32 obtains 106 an estimate of the measurement noise in the original signal from the data obtainer. The data tuner 32 then scales 110 the model mismatch using the estimated amplitudes, ak, obtained from X(αm) according to the relation defined in Equation 5 to determine an estimate of the error ζ_m in the estimate of X(αm).

The data tuner 32 compares 112 the estimate measurement noise with an estimate of the error ζ_m as defined in Equation 5, above. In the event that the measurement noise is greater than the error associated with the model mismatch, the data tuner 32 does not tune the estimate and the data provider 34 provides 114 the estimate X(αm).

The data tuner 32 obtains 108 an estimate of a "model mismatch" based on the error in the representation of the complex exponential function, provided by the coefficients Cmn and the basis functions hn, as defined in Equation 6, above.

The data tuner 32 obtains the amplitudes ak by determining 116 a filter, F, that reduces the energy of the frequency domain estimation representation, X, e.g. by selecting coefficients of

the filter F that adjust the convolution integral F*X towards zero. The data tuner 32 determines 118 the roots of the z-transform of this filter, uk (see Equation 7) which in turn yield the timings tk, and the amplitudes ak to re-estimate X. This may also be done using the Matrix Pencil approach as explained in Annex 1 and defined in Y. Hua and T. K. Sakar, "Matrix Pencil Method for Estimating Parameters of Exponentially Damped Undamped Sinusoids in Noise," IEEE Transactions on Acoustics, Speech and Signal Processing. vol. 38, pp. 814-824, May 1990.

The values of tk and ak are then tested 122 for convergence as described above with reference to Figure 3. In the event that the values have converged, the data provider provides the estimate of the frequency domain representation of the signal, X. In the event that the values have not converged steps 116, 118, 120 and 122 are repeated.

This frequency domain estimate, $X(\alpha m)$, of the signal, x(t), may be used to estimate the signal x(t) in the signal domain, thereby enabling the timing estimator 14 to determine the time delay between the transmitted and reflected signals. This estimation may be performed by determining the eigenvalues of a Matrix Pencil problem to identify the innovation parameters, or by using an "annihilating" filter method, or an approximation thereto, or by any other application of Prony's method as defined in Annex 1.

As will be appreciated, although the example described so far has related to time domain signals, spatial signals, such as images may be treated in the same way. Just as trains of pulses can be characterised by the timing and amplitude of Dirac-delta functions, so images may be defined by the location and intensity of pixels, or the location, amplitude and gradient of image boundaries, and accordingly they may be treated as FRI signals. The signal processing apparatus 12 and method described with reference to Figure 3 and Figure 4 may also be applied to determining the innovation parameters of an image.

One particular example is illustrated in Figure 5, and relates to defining a boundary between regions in a two-dimensional image.

Figure 5 shows an example of an image processing apparatus. Also represented in Figure 5 is original image, and a second version of the image as modified by an instrument function and by being sampled into pixels. In the example of Figure 5, the pixel values of the second version of the image correspond to the samples, y(n), of the original signal.

The apparatus of Figure 5 comprises a signal processing apparatus 12 as described above with reference to Figure 3 and Figure 4. The signal processing apparatus 12 is coupled to a boundary determiner, and the boundary determiner 144 is coupled to an interpolator.

The signal processing apparatus 12 is configured to receive the pixel values, and to operate on the pixel values according to the method described with reference to Figure 4 to provide a frequency domain representation of the image. The frequency domain representation of the image, as with the example described in Figure 3 and Figure 4, is based on the frequencies αm at which the instrument function of the imaging has been characterised. Naturally, these may comprise horizontal and vertical components, so the frequency domain representation can be denoted $X(\alpha mx,\alpha my)$.

The boundary determiner 144 is configured to scale the values $X(\alpha mx,\alpha my)$ by the local frequency in the image to provide scaled frequency domain image data, I(x,y), as follows

$$I(\alpha_{mx},\alpha_{my}) = (\alpha_{mx} + j\alpha_{my})X(\alpha_{mx},\alpha_{my}), \qquad (9)$$

Where, $\alpha_{mx} = \frac{j2\pi xp}{M}$,

$\alpha_{my} = \frac{j2\pi yq}{N}$,

j denotes the square root of minus one, M and N represent the dimensions of the image in the x and y directions respectively, and p and q are indices.

The boundary determiner 144 is further configured to determine a set of coefficients, b, that reduce the quantity I toward zero by solving, or finding an approximate solution (e.g. in a best

fit sense) to the equations defined by:

$$b_{pq} * I_{pq} = 0 \qquad (10)$$

where * indicates a convolution. This is analagous to finding the filter which annihilates the quantity I, and may be approached using the Annihilating Filter or Matrix Pencil methods as described above and in Annex 1. Once bpq has been determined, the boundary determiner 144 is configured to solve the system of equations:

$$f(x,y) = \sum_{p=1}^{K}\sum_{q=1}^{L} b_{pq} \exp\{-\alpha_{mx}\}\exp\{-\alpha_{my}\} = 0 \qquad (11)$$

where, $\alpha_{mx} = \frac{j2\pi xp}{M}$,

$\alpha_{my} = \frac{j2\pi yq}{N}$,

to identify a boundary as the locations (x,y) in the image where f(x,y)=0.

In this system of equations M and N are the dimensions of the original image, and j denotes the square root of minus one. Verification of this result may be found in H. Pan, T. Blu, and P.L. Dragotti, `Sampling Curves with Finite Rate of Innovation' in Proc. of Sampling Theory and Application Conference, Singapore, May 2011. This provides one particular way of identifying a boundary in an image. However, as will be appreciated, where an image comprises FRI signals, the innovation parameters of the image can be used to define the location of structures (such as boundaries) in that image, so other methods of recovering boundaries using the estimate X(αmx,αmy) may also be used, for example based on determining innovation parameters of the image.

The interpolator 146 is configured to obtain the boundary data from the boundary determiner, and to interpolate between the boundaries based on the samples of the image y(n). The interpolated image then incorporates the information obtained from the boundaries, and the initial image data values to provide an enhanced version of the image.

In addition to interpolating between boundaries to provide image enhancement, embodiments of the disclosure comprise methods of combining images to provide improved resolution. One such method is described with reference to Figure 6 and Figure 7.

As illustrated in Figure 6 and Figure 7, in this method a first image 812 is obtained 702, and the frequency space representation of the image X1(αmx,αmy) estimated as described with reference to Figure 3, Figure 4, and Figure 5. A first boundary is identified 704 in this first image 812 based on Equations 9 to 11, above

A second image 810 is also obtained 706, the frequency space representation of this image X2(αmx,αmy) is estimated, and a second boundary is identified 708 in this second image. This can be done in the same way as the first image but as will be appreciated the first and second boundaries need not be obtained in the same way, nor at the same time.

A transform, such as an affine transform, is determined 710 for co-registering the first boundary with the second boundary. This may, for example, be performed by finding a transform that reduces (e.g. minimises) a merit function based on the difference between the two boundaries.

A set of image pixel locations 806 is then determined 714 from the transform, to provide a set of pixel locations 806 having a resolution higher than the first image 812 and the second image 810. The transform is then applied to the first image data 812 to provide transformed first image data 802.

The transformed first image data 802, and the second image data 810 are then combined 716 by interpolating from the transformed first image data 802, and the second image data 810 to determine combined image data 820 at the new set of pixel locations 806.

Naturally, although this method has been described with reference to the co-registration of two images, it may be applied to any number of images.

To assist in understanding the disclosure, the functionality of the signal processing apparatus 12 of Figure 2 has been notionally subdivided into functional elements. It will however be appreciated that these functional elements need not be provided as physically distinct structures, and some or all aspects of the functionality may be provided in a common physical structure, and some or all aspects of the functionality may be further subdivided. As one particular example, the signal processing apparatus 12 may be provided by a single processor, and the data provider 34 and the data obtainer 22 may comprise memory elements, such as registers or buffers that are readable and/or writable by the chip. For example the data provider 34 and the data obtainer 22 may be memory caches, such as on-chip caches of the processor. In addition, this functionality may be provided by any appropriate processor such as an application specific integrated circuit, ASIC, a field programmable gate array, FPGA, or an assembly of digital logic gates. In some embodiments the functionality is provided by a computer program. The computer program comprises program instructions operable to program a processor to provide the functionality of the signal processing apparatus. This computer program can be stored in a transitory or non-transitory form.

Figure 1 introduces one example of how an embodiment of the disclosure may be used. This particular example relates to range finding, and this example is chosen because it relates to the estimation of a signal comprising a sequence of pulses. As will be appreciated by the skilled addressee in the context of the present disclosure, any signal may be described by a series of single points (e.g. a train of Dirac delta functions). Therefore, although this example is provided as one way of understanding the present disclosure, it will be understood that the apparatus and methods of the disclosure can be applied to signal processing in many other circumstances. Estimating a train of pulses is just one example. Examples of instrument functions include the impulse response of a filter, and the point spread function of an optical system.

A variety of signals could be used in the range finding example of Figure 1, such as optical or acoustic (e.g. ultrasound) signals. However, as will be appreciated in the context of the

present disclosure, the temporal resolution of the ADC, and the speed of the outgoing and reflected signals (e.g. the speed of light for optical signals), limit the accuracy of the measurement.

The instrument function may be characterised by its response to a unit impulse, another way is to measure the response of the instrument to signals of known frequency and amplitude, for example using sinusoids of known frequencies. Other methods of characterising detection instruments will be apparent to the skilled addressee in the context of the present disclosure.

In the discussion above, the parameter n is used to index samples in the signal domain, and is used to denote shifts in the signal domain example, where the signal is a time domain signal, the parameter n relates to a time delay, and where the signal is a spatially varying signal, the parameter n relates to a spatial displacement.

The delays, n, may comprise integer numbers of samples, in which case the coefficients Cmn may be obtained from an orthogonal projection of the exponential function exp{αm} onto the shifted basis functions. However, in some embodiments the delays, n, may comprise integer and non-integer numbers of samples. In these possibilities the coefficients Cmn can be obtained from fitting procedure based on a reduction of a merit function that describes the model mismatch, such as:

$$\varepsilon_m(t) = \left| \exp\{\alpha_m t\} - \sum_n C_{mn} h_n(t) \right|$$

One way to select these coefficients, Cmn, is according to the relation defined in Equation 2 above. However, other numerical and analytical solutions may be used. For example:

$$C_{mn} = \frac{H(-\alpha_m)}{\hat{a}_H(\exp\{\alpha_m\})} \exp\{\alpha_m n\},$$

Where $\hat{a}_H(\exp\{\alpha_m\}) = \sum_{l \in Z} a_H[l] \exp\{-\alpha_m l\}$ is the z-transform of the auto-correlation of the

signal domain representation of the instrument function $a_H = \langle h(t-l), h(t) \rangle$ evaluated at z=exp{αm}.

The coefficients Cmn may be selected in other ways. For example, coefficients may be selected based on constraining the model mismatch to have particular values (e.g. to be zero to within a selected tolerance) at particular locations. One example of this approach is to ensure that the model ΣCmnhn interpolates the exponential function exp{αmt) exactly where t is an integer. Taking this approach, the coefficients can be defined:

$$C_{mn} = \frac{1}{\sum_{l \in Z} \exp\{\alpha_m l\} h(l)} \exp\{\alpha_m n\}$$

Other coefficients which reduce a merit function determined by the model mismatch may also be used. Examples of these coefficients typically comprise frequency domain representations of the instrument function.

The frequencies, αm, are described as being selected to reduce the variance in the magnitude of the instrument function at these frequencies, H(αm) whilst increasing the spacing between the frequencies. This may be achieved in a variety of ways. In some embodiments, the signal belongs to a class of finite rate innovation, FRI, signal, and the frequencies, αm are selected based on the performance of a maximum likelihood estimator of innovation parameters of that class of signal. One way to achieve this is by selecting the frequencies, αm, that reduce (for example minimise) the Cramer-Rao bound, which is defined by Equation 16 of Annex 1. Other approaches may be used to select these frequencies. One or more sets of coefficients Cmn can be stored by the coefficient provider 20. In addition, the model mismatch $\varepsilon_m(t)$ may also be stored in a memory, for example for access by the data tuner.

In the range finding example, the original signal is a time domain signal, and so references to frequency domain and signal domain should be understood in this context. Other examples relate to image processing, in which case the signal is a spatial domain signal, and so references to signal domain and frequency domain should be understood in that context. As

will be appreciated by the skilled reader in the context of the present disclosure the link between the frequency domain and the signal domain may be provided by any one of a selection of transforms such as Fourier series, Fourier transforms, Laplace transforms, Z-transforms or any other frequency transform.

In particular the signal processing methods and apparatus of the disclosure may be used in the estimation of neurological signals, or in seismology, or image processing. Examples of the disclosure provide multi-channel data acquisition apparatus, where each channel comprises a signal processing apparatus 12 such as that described with reference to Figure 3.

It will be appreciated that a signal processing method for estimating a frequency domain representation of a signal from a series of samples of the signal collected by an instrument and distorted by an instrument function associated with the collection of the signal by the instrument has been disclosed. It will also be appreciated that this method may be applied in any one of the various ways described herein. For example the method may comprise: obtaining the series of samples; obtaining a set of coefficients that fit a set of basis functions to a complex exponential function wherein the set of basis functions comprises a plurality of basis functions each defined by a shifted version of the instrument function in a signal domain; estimating the frequency domain representation of the signal based on the series of samples and the coefficients; wherein the instrument function is based on a characterisation of the instrument function in the frequency domain at frequencies associated with the complex exponential function. Each of these features of this method may be implemented, generalised, or further refined in any one or more of the ways described herein.

It will therefore be understood that the embodiments described are to be understood as illustrative examples. Further embodiments are envisaged. It is to be understood that any feature described in relation to any one embodiment may be used alone, or in combination with other features described, and may also be used in combination with one or more features of any other of the embodiments, or any combination of any other of the embodiments. Furthermore, equivalents and modifications not described above may also be employed

without departing from the scope of the invention, which is defined in the accompanying claims.

With reference to the drawings in general, it will be appreciated that schematic functional block diagrams are used to indicate functionality of systems and apparatus described herein. It will be appreciated however that the functionality need not be divided in this way, and should not be taken to imply any particular structure of hardware other than that described and claimed below. The function of one or more of the elements shown in the drawings may be further subdivided, and/or distributed throughout apparatus of the disclosure. In some embodiments the function of one or more elements shown in the drawings may be integrated into a single functional unit.

In some examples, one or more memory elements can store data and/or program instructions used to implement the operations described herein. Embodiments of the disclosure provide tangible, non-transitory storage media comprising program instructions operable to program a processor to perform any one or more of the methods described and/or claimed herein and/or to provide data processing apparatus as described and/or claimed herein.

The activities and apparatus outlined herein may be implemented using controllers and/or processors which may be provided by fixed logic such as assemblies of logic gates or programmable logic such as software and/or computer program instructions executed by a processor. Other kinds of programmable logic include programmable processors, programmable digital logic (e.g., a field programmable gate array (FPGA), an erasable programmable read only memory (EPROM), an electrically erasable programmable read only memory (EEPROM)), an application specific integrated circuit, ASIC, or any other kind of digital logic, software, code, electronic instructions, flash memory, optical disks, CD-ROMs, DVD ROMs, magnetic or optical cards, other types of machine-readable mediums suitable for storing electronic instructions, or any suitable combination thereof.

ANNEX 1

The following is a detailed description of the way in which the present invention may be considered. Whilst various features may be described as being essential or necessary, this may only be the case for the specific example discussed, for example due to other requirements imposed by the particular mathematical situation under consideration. These statements should not, therefore, be construed as limiting the present invention in any way.

Approximate Strang-Fix:

FRI Sampling with Arbitrary Kernels

Jose Antonio Urigüen*†, Thierry Blu+, and Pier Luigi Dragotti†,

Abstract

In the last few years, several new methods have been developed for the sampling and the exact reconstruction of specific classes of non-bandlimited signals known as signals with finite rate of innovation (FRI). This is achieved by using adequate sampling kernels and reconstruction schemes. The sampling kernels used in [1] are called exponential reproducing kernels and satisfy the generalized Strang-Fix conditions. These ensure that proper linear combinations of the kernel with its shifted versions reproduce polynomials or exponentials exactly.

The first contribution of the paper is to provide clear guidelines on how to choose the correct exponential reproducing kernel when estimating FRI signals from noisy samples. We then depart from the situation in which we can choose the sampling kernel and develop a new strategy that is universal in that it works with any kernel. We do so by noting that meeting the exact exponential reproduction condition is too stringent a constraint, we thus allow for a controlled error in the reproduction formula in order to use the exponential reproduction idea with any kernel and develop a reconstruction method which is more robust to noise.

Numerical results validate the various contributions of the paper and in particular show that the approximate exponential reproduction strategy leads to reconstruction results that are more stable than those obtained when using the traditional exact recovery method.

Index Terms

Sampling, Finite Rate of Innovation, Noise, MOMS, Matrix Pencil

EDICS Category: DSP-SAMP

I. Introduction

Most digital acquisition systems involve the conversion of signals from analog to digital, and sampling theorems provide the bridge between the continuous and the discrete-time worlds. Usually, the acquisition process is modelled as in Figure 1, where the smoothing function $\varphi(t)$ is called the sampling kernel, and normally models the distortion due to the acquisition device. The filtered continuous-time signal $y(t) = x(t) * \varphi(-\frac{t}{T})$ is then uniformly sampled at a rate $f_s = \frac{1}{T}$. Following this setup, the measurements are given by

$$y_n = \int_{-\infty}^{\infty} x(t)\varphi\left(\frac{t}{T} - n\right) \mathrm{d}t = \left\langle x(t), \varphi\left(\frac{t}{T} - n\right)\right\rangle.$$

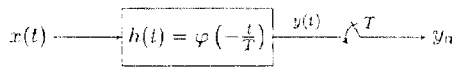


Figure 1. *Traditional sampling scheme.* The continuous-time input signal $x(t)$ is filtered with $h(t)$ and sampled every T seconds. The samples are then given by $y_n = (x * h)(t)|_{t=nT}$.

The fundamental problem of sampling is to recover the original continuous-time waveform $x(t)$ using the set of samples y_n. In the case where the signal is bandlimited, the answer due to Shannon is well known. Recently, it has been shown [2], [1], [3] that it is possible to sample and perfectly reconstruct specific classes of non-bandlimited signals. Such signals are called signals with finite rate of innovation (FRI) since they are completely described by a finite number of free parameters. Perfect reconstruction is achieved by using a variation of Prony's method, also known as annihilating filter method [4]. Signals that can be sampled within this framework include streams of pulses such as Diracs [2], [1], [3], [5], piecewise polynomial signals, piecewise sinusoidal signals [6] and classes of 2-D signals [7], [8], [9], [10]. In the presence of noise, FRI reconstruction techniques become unstable and methods to improve resiliency to noise have been presented in [11], [12], [13], [14], [15], [10].

We note that various sampling kernels can be used to perfectly reconstruct FRI signals such as the sinc and Gaussian functions first proposed in the original paper on FRI [2] and compact support kernels such as the family of Sum of Sincs (SoS) [3], and polynomial and exponential reproducing kernels [1], [16]. While they all allow perfect reconstruction in noiseless settings, their behavior changes in the presence of

noise. It is therefore natural to attempt to understand which factors cause a deterioration in performance and also provide precise guidelines on how to design the kernels that are the most resilient to noise.

In this paper, we focus on the family of exponential reproducing kernels [1], and provide guidelines on how to design the most stable ones. We focus on exponential reproducing kernels for two reasons: First, they can have compact support which is in itself a nice property when dealing with noisy measurements. Second and more important, any compact support kernel that has so far been used in FRI sampling using the setting of Figure 1 is a particular instance of the family of exponential reproducing kernels. We prove this fact in Appendix B.

Our contribution is twofold: We first explain how to design the most effective exponential reproducing kernels when sampling and reconstructing FRI signal in noisy environments. Since FRI recovery is equivalent to estimating a set of parameters in noise, we achieve this optimal design by finding the family of exponential reproducing kernels that minimise the Cramér-Rao bound of this FRI estimation problem. In the second contribution, we depart from the situation in which we are allowed to shape the kernel in the way that best meets our needs and consider the case where we have no control on the acquisition device. In this new scenario we develop a new strategy for FRI sampling that is universal in that it works with any possible kernel. In contrast to existing techniques that attempt at finding the exact parameters of the input, we describe how to perform an approximate recovery also capable of high accuracy when noise is present. The advantage of our new approach is that it can be applied to any sampling kernel. Moreover, it can replace the traditional exact recovery framework for kernels such as polynomial splines that can theoretically achieve perfect reconstruction, but are in practice very unstable. By replacing the exact recovery strategy with this new approximate framework we show that a reliable reconstruction of FRI signals in noisy settings is possible even when using 'unstable' kernels.

The outline of the paper is as follows. In Section II we introduce the noiseless scenario in which we sample and perfectly reconstruct the prototypical FRI signal: a train of Diracs. We also discuss exponential reproducing kernels [1] and the generalised Strang-Fix conditions [17]. In Section III we treat the more realistic setup when noise is present in the acquisition process. Here, we describe practical techniques to retrieve the train of Diracs and compute the Cramér-Rao bound for this problem. In Section IV we design a family of exponential reproducing kernels that is most resilient to noise. Then, in Section V we present the approximate FRI framework, and develop the basic ideas to sample FRI signals with any kernel. Interestingly, by properly defining the reconstruction stage, we also show that with this new framework we can improve reconstruction accuracy of existing kernels. In Section VI we present simulation results to validate the various contributions of the paper. Finally, we conclude in Section VII.

II. SAMPLING SIGNALS WITH FINITE RATE OF INNOVATION

In this section we provide a brief overview of the theory of sampling FRI signals. Specifically, we explain how to perfectly reconstruct a stream of Diracs from the samples. Note that the reconstruction of more sophisticated signals can always be reduced to this case.

A. *Perfect reconstruction of a stream of Diracs*

For the sake of clarity, we consider that the input $x(t)$ is a stream of K Diracs with amplitudes a_k located at different instants of time t_k where $t_k \in [0, \tau)$ and $k = 0, \ldots, K-1$. The signal has $2K$ degrees of freedom in total, and can be expressed as:

$$x(t) = \sum_{k=0}^{K-1} a_k \delta(t - t_k). \tag{1}$$

Now, based on the acquisition model of Figure 1, we filter the input with the kernel $\varphi(t)$ and obtain the measurements

$$y_n = \left\langle x(t), \varphi\left(\frac{t}{T} - n\right)\right\rangle = \sum_{k=0}^{K-1} a_k \varphi\left(\frac{t_k}{T} - n\right), \tag{2}$$

where $n = 0, 1, \ldots, N-1$. Here we assume that the sampling period T satisfies $\tau = NT$. Moreover, we also assume that $\varphi(t)$ is of compact support and satisfies

$$\sum_{n \in \mathbb{Z}} c_{m,n} \varphi(t - n) = e^{\alpha_m t}, \tag{3}$$

for proper coefficients $c_{m,n}$, with $m = 0, \ldots, P$ and $\alpha_m \in \mathbb{C}$. Kernels satisfying (3) are called exponential reproducing kernels of order $P+1$. We discuss them in detail in Section II-B. For the rest of the paper we restrict our analysis to parameters of the form $\alpha_m = \alpha_0 + m\lambda$, where $\alpha_0, \lambda \in \mathbb{C}$. The reason for this choice will become clear later on in the section.

Once we have sampled the input, we need to show that it can be unambiguously retrieved from the set of measurements y_n. In order to do so, first we linearly combine the samples y_n with the coefficients $c_{m,n}$ of (3), to obtain the new measurements:

$$s_m = \sum_{n=0}^{N-1} c_{m,n} y_n, \tag{4}$$

for $m = 0, \ldots, P$. Then, given that the signal $x(t)$ is a stream of Diracs (1) and combining (4) with (2)

we have [1]:

$$\begin{aligned} s_m &= \left\langle x(t), \sum_{n=0}^{N-1} c_{m,n}\varphi\left(\frac{t}{T}-n\right)\right\rangle \\ &= \int_{-\infty}^{\infty} x(t)\mathrm{e}^{\alpha_m \frac{t}{T}}\mathrm{d}t \\ &= \sum_{k=0}^{K-1} a_k \mathrm{e}^{\alpha_m \frac{t_k}{T}} = \sum_{k=0}^{K-1} x_k u_k^m, \end{aligned} \tag{5}$$

with $x_k = a_k \mathrm{e}^{\alpha_0 \frac{t_k}{T}}$ and $u_k = \mathrm{e}^{\lambda \frac{t_k}{T}}$. Note that the values s_m are the (exponential) moments of the signal $x(t)$ [6], and are equivalent to the projection of $x(t)$ onto the subspace spanned by the exponentials $\{\mathrm{e}^{\alpha_m \frac{t}{T}}\}_{m=0}^{P}$. The fact that we use parameters of the form $\alpha_m = \alpha_0 + m\lambda$, where $m = 0, \ldots, P$, is necessary for the values s_m to have a power sum series form, which is key to the recovery stage.

The new pairs of unknowns (u_k, x_k) for $k = 0, \ldots, K-1$ can then be retrieved from the measurements s_m using the annihilating filter method [2], [1], [12], also known as Prony's method in the spectral estimation community [4]. Let h_m with $m = 0, \ldots, K$ denote the filter with z-transform $\hat{h}(z) = \sum_{m=0}^{K} h_m z^{-m} = \prod_{k=0}^{K-1}\left(1 - u_k z^{-1}\right)$, that is, its roots correspond to the values u_k to be found. Then, it follows that h_m annihilates the observed series s_m:

$$h_m \star s_m = \sum_{i=0}^{K} h_i s_{m-i} = \sum_{k=0}^{K-1} x_k u_k^m \underbrace{\sum_{i=0}^{K} h_i u_k^{-i}}_{\hat{h}(u_k)} = 0. \tag{6}$$

Moreover, the zeros of this filter uniquely define the values u_k provided the locations t_k are distinct. The identity (6) can be written in matrix-vector form as:

$$\mathbf{S}\mathbf{h} = 0 \tag{7}$$

which reveals that the Toeplitz matrix $\mathbf{S}$ is rank deficient. By solving the above system, we find the filter coefficients h_m and then retrieve u_k by computing the roots of $\hat{h}(z)$. Given u_k we obtain the locations t_k since $u_k = \mathrm{e}^{\lambda \frac{t_k}{T}}$. We detail the main steps of the annihilating filter method in Algorithm 1.

Note that if λ is purely imaginary, u_k is periodic of period $2\pi \frac{T}{\mathrm{Im}\{\lambda\}}$ and it is necessary that the locations t_k satisfy $0 \leqslant t_k \leqslant 2\pi \frac{T}{\mathrm{Im}\{\lambda\}}$ for $k = 0, \ldots, K-1$ in order to retrieve them unambiguously. Finally, we determine the weights a_k by solving, for instance, the first K consecutive equations in (5). Notice that the problem can be solved only when there are at least as many equations as unknowns, implying that $P \geqslant 2K-1$. This indicates that the order $P+1$ of the exponential reproducing kernel has to be chosen according to the number of degrees of freedom of the input signal $x(t)$.

Algorithm 1 Annihilating filter: total least-squares method.

Retrieving the innovation parameters (t_k, a_k) of a train of K Diracs (1)

1: Calculate the sequence $s_m = \sum_{n=0}^{N-1} c_{m,n} y_n$ for $m = 0, \ldots, P$, from the N samples y_n of (2).

2: Build the system of equations (6) using the exponential moments s_m.

3: Retrieve the annihilating filter coefficients h_m, for $m = 0, \ldots, K$, by performing the SVD of the Toeplitz matrix $\mathbf{S}$ and choosing the eigenvector corresponding to the smallest eigenvalue.

4: Compute the roots $u_k = \mathrm{e}^{\lambda \frac{t_k}{T}}$ of the z-transform $\hat{h}(z) = \sum_{k=0}^{K} h_m z^{-m}$ and obtain $\{t_k\}_{k=0}^{K-1}$.

5: Calculate $\{a_k\}_{k=0}^{K-1}$ as the least mean square solution of the N equations $y_n - \sum_{k=0}^{K-1} a_k \varphi\left(\frac{t_k}{T} - n\right)$.

If the measurements y_n are noisy, then it is necessary to denoise them by using the methods in Section III.

All FRI reconstruction methods proposed so far ([2], [1], [12], [3]) use this same approach. Specifically, N samples of the form (2) are taken to form the vector $\mathbf{y} = (y_0, y_1, \ldots, y_{N-1})^T$. Then, they are combined linearly with coefficients $c_{m,n}$ as in (5), which is equivalent to multiplying vector $\mathbf{y}$ by a proper weighting matrix $\mathbf{C}$ leading to $\mathbf{s} = \mathbf{C}\mathbf{y} = (s_0, s_1, \ldots, s_P)^T$. Here $\mathbf{s}$ is the vector with the $P+1$ new measurements and $\mathbf{C}$ is a $(P+1) \times N$ matrix with coefficient $c_{m,n}$ at location (m, n). The values $\mathbf{s}$ can always be expressed in a power sum series form involving the innovation parameters (a_k, t_k) for $k = 0, \ldots, K-1$. These can then be retrieved from $\mathbf{s}$ using the annihilating filter method. The choice of $\mathbf{C}$ depends on the sampling kernel and, while in the noiseless setting perfect reconstruction is always achieved, the role of $\mathbf{C}$ becomes pivotal in noisy, non-ideal scenarios as discussed throughout the paper.

B. Exponential reproducing kernels

An exponential reproducing kernel is any function $\varphi(t)$ that, together with a linear combination of its shifted versions, can reproduce functions of the form $\mathrm{e}^{\alpha_m t}$, with complex parameters α_m. This can be expressed mathematically as follows:

$$\sum_{n \in \mathbb{Z}} c_{m,n} \varphi(t - n) = \mathrm{e}^{\alpha_m t}, \tag{8}$$

for properly chosen coefficients $c_{m,n} \in \mathbb{C}$ and where $m = 0, \ldots, P$ and $\alpha_m \in \mathbb{C}$. Exponential reproducing kernels for which (8) is true satisfy the so-called generalised Strang-Fix conditions [17] (see Appendix A). Specifically, Equation (8) holds if and only if

$$\hat{\varphi}(\alpha_m) \neq 0 \text{ and } \hat{\varphi}(\alpha_m + 2j\pi l) = 0. \tag{9}$$

for $m = 0, \ldots, P$ and $l \in \mathbb{Z} \backslash \{0\}$, where $\hat{\varphi}(\alpha_m)$ represents the bilateral Laplace transform of $\varphi(t)$ at $s = \alpha_m$.

Moreover, the coefficients $c_{m,n}$ in (8) are given by

$$c_{m,n} = \left\langle e^{\alpha_m t}, \tilde{\varphi}(t-n) \right\rangle = \int_{-\infty}^{\infty} e^{\alpha_m t} \tilde{\varphi}(t-n)\, dt = c_{m,0} e^{\alpha_m n}, \tag{10}$$

where the function $\tilde{\varphi}(t)$ forms a biorthonormal set with $\varphi(t)$, i.e. when $\tilde{\varphi}(t)$ is such that $\langle \tilde{\varphi}(t-n), \varphi(t-m) \rangle = \delta_{m-n}$, and where $c_{m,0} = \int_{-\infty}^{\infty} e^{\alpha_m x} \tilde{\varphi}(x) dx$.

Any exponential reproducing kernel can be written as $\varphi(t) = \gamma(t) * \beta_{\vec{\alpha}}(t)$ [1], [18], [19], where $\gamma(t)$ is an arbitrary function, even a distribution, and $\beta_{\vec{\alpha}}(t)$ is an E-Spline. A function $\beta_\alpha(t)$ with Fourier transform $\hat{\beta}_\alpha(\omega) = \frac{1-e^{\alpha-j\omega}}{j\omega-\alpha}$ is called cardinal exponential spline of first order, with $\alpha \in \mathbb{C}$ [18]. The time domain representation of such function is $\beta_\alpha(t) = e^{\alpha t}$ for $t \in [0,1)$ and zero elsewhere. The function $\beta_\alpha(t)$ is, therefore, of compact support and a linear combination of its shifted versions reproduces the exponential $e^{\alpha t}$. Higher order cardinal exponential splines (E-Splines) can be obtained through convolution of first order ones, so that for instance $\beta_{\vec{\alpha}}(t) = (\beta_{\alpha_0} * \beta_{\alpha_1} * \ldots * \beta_{\alpha_P})(t)$, where $\vec{\alpha} = (\alpha_0, \alpha_1, \ldots, \alpha_P)$, is an E-Spline of order $P+1$. This can also be written in the Fourier domain as follows:

$$\hat{\beta}_{\vec{\alpha}}(\omega) = \prod_{m=0}^{P} \frac{1 - e^{\alpha_m - j\omega}}{j\omega - \alpha_m}. \tag{11}$$

The higher order E-Splines are of length $P+1$ and their regularity increases with P (i.e. they have $P-1$ continuous derivatives). These functions can reproduce any exponential in $\mathrm{span}\{e^{\alpha_0 t}, e^{\alpha_1 t}, \ldots, e^{\alpha_P t}\}$ [1], [18]. We also note that whenever $\alpha_m = 0$ for $m = 0, \ldots, P$, the function $\beta_{\vec{\alpha}}(t)$ reduces to a B-Spline and no longer reproduces exponentials but polynomials up to degree P.

In this paper we work with real valued sampling kernels. Therefore, we require that $\gamma(t)$ and $\beta_{\vec{\alpha}}(t)$ be real valued. We note that this second condition (i.e. $\beta_{\vec{\alpha}}(t)$ real valued) is satisfied when the exponentials α_m are real or in complex conjugate pairs. We also restrict the exponential parameters to be of the form $\alpha_m = \alpha_0 + m\lambda$ with $m = 0, \ldots, P$. These conditions imply that, either λ and α_0 are real-valued, or otherwise $\lambda = j\lambda_0$ is a pure imaginary number, in which case $\mathrm{Im}\{\alpha_0\} = -P\lambda_0/2$.

III. Sampling FRI signals in the presence of noise

"Noise", or more generally model mismatch, is present in data acquisition, making the solution presented in Section II-A only ideal. We therefore assume that we have access to the measurements

$$\tilde{y}_n = y_n + \epsilon_n = \sum_{k=0}^{K-1} a_k \varphi\left(\frac{t_k}{T} - n\right) + \epsilon_n, \tag{12}$$

with $n = 0, \ldots, N-1$ and where ϵ_n are i.i.d. Gaussian random variables, of zero mean and standard deviation σ. When the samples are corrupted by noise, the new set of measurements s_m of Equation (4)

change, and perfect reconstruction is no longer possible. We now have the noisy moments:

$$\tilde{s}_m = \sum_{n=0}^{N-1} c_{m,n}\tilde{y}_n = \underbrace{\sum_{n=0}^{N-1} c_{m,n}y_n}_{s_m} + \underbrace{\sum_{n=0}^{N-1} c_{m,n}\epsilon_n}_{b_m} = \sum_{k=0}^{K-1} x_k u_k^m + b_m, \tag{13}$$

for $m = 0\ldots, P$ and where $x_k = a_k e^{\alpha_k \frac{t_k}{T}}$ and $u_k = e^{\lambda \frac{t_k}{T}}$ with $k = 0, \ldots, K-1$.

Next, we review some of the algorithms that have been used in the literature to recover FRI signals from noisy samples. We also introduce the Cramér–Rao lower bound (CRB) to analyse the accuracy with which FRI signals can be retrieved in the presence of noise. This is a lower bound on the mean square error (MSE) that applies to unbiased estimator [20].

A. Practical methods for noisy FRI reconstruction

As we have explained in Section II-A, the retrieval procedure for FRI signals is based on calculating the new set of measurements (4). When the samples (2) are corrupted by noise, the measurements s_m of Equation (4) become (13). Consequently, (7) is not satisfied any more because now $\tilde{\mathbf{S}} = \mathbf{S} + \mathbf{B}$ with $\mathbf{B}$ being the Toeplitz matrix with entries b_m from (13). The idea behind the procedure adopted in [12] is to look for a solution that minimises $\|\tilde{\mathbf{S}}\mathbf{h}\|^2$ under the constrain that $\|\mathbf{h}\|^2 = 1$. This is a classical total least-square (TLS) problem that can be solved using singular value decomposition (SVD). The algorithm may be further improved by denoising $\tilde{\mathbf{S}}$ before applying TLS, by using the Cadzow iterative algorithm [12], [21]. There exist other methods that attain a similar accuracy. One such approach, based on solving a matrix pencil problem [22], [23], was introduced for FRI in [11]. It has been used in other FRI publications such as [3], [16] and is used in the simulations of this paper as well.

All these methods operate effectively when the perturbation is white, that is when the covariance matrix of the noise $\mathbf{B}$ satisfies $\mathbf{R}_\mathbf{B} = \mathrm{E}\{\mathbf{B}^H\mathbf{B}\} = \alpha\mathbf{I}$, where α is a constant factor and $\mathbf{I}$ is the identity matrix. In such a case, noise has a similar effect on the principal values of the SVD decomposition of $\tilde{\mathbf{S}}$ [24], [11], which explains why the retrieval procedures reviewed so far are the most adequate. However, for many FRI kernels the white Gaussian noise assumption does not hold. As a consequence, the entries of the Toeplitz noise matrix $\mathbf{B}$ may have non-uniform variance or even become correlated. In order for SVD to operate properly it is necessary to "pre-whiten" the noise. This is a well known approach proposed by various authors in the spectral estimation community (for instance by De Moor in [24]). In our simulations we employ a standard approach consisting of using a weighting matrix $\mathbf{W} = \mathbf{R}_\mathbf{B}^{-1/2}$ [25] such that $\mathbf{R}_\mathbf{A} = \mathrm{E}\{\mathbf{A}^H\mathbf{A}\} = \mathbf{I}$ with $\mathbf{A} = \mathbf{B}\mathbf{W}$. Therefore, we perform the SVD decomposition of $\tilde{\mathbf{S}}\mathbf{W}$, now characterised by white noise.

B. Measuring the performance

FRI signals are completely characterised by their innovation parameters. For instance, a stream of K Diracs can be completely determined from the locations t_k and amplitudes a_k. We therefore introduce the vector $\Theta = (t_0, \ldots, t_{K-1}, a_0, \ldots, a_{K-1})^T$ and the goal is to estimate Θ from the vector of N noisy samples $\tilde{\mathbf{y}} = (\tilde{y}_0, \ldots, \tilde{y}_{N-1})^T$ given by (12). For simplicity we assume the sampling period is $T = 1$. A way to determine the CRB of this estimation problem was given in [12] assuming ϵ_n is a zero-mean Gaussian noise with covariance matrix $\mathbf{R} = \mathrm{E}\{\mathbf{e}\mathbf{e}^H\}$, where $\mathbf{e}$ is the vector of length N with values ϵ_n. In this set-up any unbiased estimate of the unknown parameters $\hat{\Theta}(\tilde{\mathbf{y}}) = (\hat{t}_0, \ldots, \hat{t}_{K-1}, \hat{a}_0, \ldots, \hat{a}_{K-1})^T$ has a covariance matrix that is lower bounded by

$$\mathrm{cov}(\hat{\Theta}(\tilde{\mathbf{y}})) \geqslant (\Phi_{\mathbf{y}}^T \mathbf{R}^{-1} \Phi_{\mathbf{y}})^{-1}, \tag{14}$$

where the matrix $\Phi_{\mathbf{y}}$ where the matrix $\Phi_{\mathbf{y}}$ is given by (15).[1]

$$\Phi_{\mathbf{y}} = \left(\begin{array}{ccc|ccc} a_0\varphi'(t_0) & \cdots & a_{K-1}\varphi'(t_{K-1}) & \varphi(t_0) & \cdots & \varphi(t_{K-1}) \\ a_0\varphi'(t_0-1) & \cdots & a_{K-1}\varphi'(t_{K-1}-1) & \varphi(t_0-1) & \cdots & \varphi(t_{K-1}-1) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_0\varphi'(t_0-(N-1)) & \cdots & a_{K-1}\varphi'(t_{K-1}-(N-1)) & \varphi(t_0-(N-1)) & \cdots & \varphi(t_{K-1}-(N-1)) \end{array}\right). \tag{15}$$

While this is the correct way to measure the performance of various FRI recovery techniques [12], [11] given the noisy samples $\tilde{\mathbf{y}}$, we also note that in practice FRI reconstruction methods operate on the sequence of moments $\tilde{\mathbf{s}} = \mathbf{C}\tilde{\mathbf{y}}$. It is therefore of interest to find the CRB associated to the measurements $\tilde{\mathbf{s}}$ (which may be fewer than the number of samples), since this will indicate the best performance that can be achieved in practice. In this new case, the bound is given by [26]:

$$\mathrm{cov}(\hat{\Theta}(\tilde{\mathbf{s}})) \geqslant (\Phi_{\mathbf{s}}^H \mathbf{R}_{\mathrm{b}}^{-1} \Phi_{\mathbf{s}})^{-1}. \tag{16}$$

where $\Phi_{\mathbf{s}}$ takes the form:

$$\Phi_{\mathbf{s}} = \left(\begin{array}{ccc|ccc} a_0\alpha_0 \mathrm{e}^{\alpha_0 t_0} & \cdots & a_{K-1}\alpha_0 \mathrm{e}^{\alpha_0 t_{K-1}} & \mathrm{e}^{\alpha_0 t_0} & \cdots & \mathrm{e}^{\alpha_0 t_{K-1}} \\ a_0\alpha_1 \mathrm{e}^{\alpha_1 t_0} & \cdots & a_{K-1}\alpha_1 \mathrm{e}^{\alpha_1 t_{K-1}} & \mathrm{e}^{\alpha_1 t_0} & \cdots & \mathrm{e}^{\alpha_1 t_{K-1}} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_0\alpha_P \mathrm{e}^{\alpha_P t_0} & \cdots & a_{K-1}\alpha_P \mathrm{e}^{\alpha_P t_{K-1}} & \mathrm{e}^{\alpha_P t_0} & \cdots & \mathrm{e}^{\alpha_P t_{K-1}} \end{array}\right) \tag{17}$$

[1] The matrix can be obtained calculating the derivative of $\tilde{y}_n$ with respect to each parameter in Θ. That is, the columns of $\Phi_{\mathbf{y}}$ to the left of | are $\frac{\partial \tilde{y}_n}{\partial t_k}$ and the columns of $\Phi_{\mathbf{y}}$ to the right of | are $\frac{\partial \tilde{y}_n}{\partial a_k}$.

and $\mathbf{R}_\mathbf{b} = \mathrm{E}\{\mathbf{b}\mathbf{b}^H\}$. Here, $(\cdot)^H$ is Hermitian transpose and $\mathbf{b}$ is the vector of $P+1$ noisy values b_m.

Furthermore, since we have assumed that the noise ϵ_n added to the samples is additive, white and Gaussian, with variance σ^2, the covariance matrix of the filtered noise is $\mathbf{R}_\mathbf{b} = \mathrm{E}\{\mathbf{b}\mathbf{b}^H\} = \sigma^2\mathbf{C}\mathbf{C}^H$. Consequently, there is a direct relation between formulations (14) and (16), that can be expressed through the matrix $\mathbf{C}$. In order to see this, note that $\hat{\mathbf{s}} = \mathbf{C}(\mathbf{y}+\mathbf{e}) = \mathbf{s}+\mathbf{b}$ and also that $\Phi_\mathbf{s} = \mathbf{C}\Phi_\mathbf{y}$. Moreover, we know that

$$\mathrm{cov}(\hat{\Theta}(\tilde{\mathbf{y}})) \geqslant (\Phi_\mathbf{y}^H\mathbf{R}^{-1}\Phi_\mathbf{y})^{-1} = \sigma^2(\Phi_\mathbf{y}^H\Phi_\mathbf{y})^{-1},$$

since the noise ϵ_n is AWGN. Consequently, it is true that

$$\mathrm{cov}(\hat{\Theta}(\hat{\mathbf{s}})) \geqslant (\Phi_\mathbf{s}^H\mathbf{R}_\mathbf{b}^{-1}\Phi_\mathbf{s})^{-1} = \sigma^2(\Phi_\mathbf{y}^H\mathbf{C}^H(\mathbf{C}\mathbf{C}^H)^{-1}\mathbf{C}\Phi_\mathbf{y})^{-1} = \sigma^2(\Phi_\mathbf{y}^H\mathbf{C}^\dagger\mathbf{C}\Phi_\mathbf{y})^{-1},$$

where $\mathbf{C}^\dagger = \mathbf{C}^H(\mathbf{C}\mathbf{C}^H)^{-1}$ is the left pseudo-inverse of $\mathbf{C}$. When the number of moments $P+1$ equals the number of samples N, then both formulations are equivalent. This is easily seen since $\mathbf{C}^\dagger\mathbf{C} = \mathbf{C}^{-1}\mathbf{C} = \mathbf{I}_N$. The predicted CRBs using either alternative coincide given that the matrix of coefficients $\mathbf{C}$ is square and invertible. Intuitively, this is the optimal configuration, because no linear combination of the samples should improve the estimation based on the samples themselves. Retrieving the innovation parameters through the moments when $P+1 < N$ is instead suboptimal.

Experimentally, we have noted that FRI algorithms reach the bound (16) when $\mathbf{C}$ is sufficiently well conditioned. Therefore our goal now is to design kernels that lead to properly conditioned $\mathbf{C}$ that minimise (16) for any choice of P.[2]

IV. Optimal exponential reproducing kernels

As mentioned before, an exponential reproducing kernel is any function $\varphi(t) = \gamma(t) * \beta_{\vec{\alpha}}(t)$ [1], [18], where $\gamma(t)$ is arbitrary and $\beta_{\vec{\alpha}}(t)$ is an E-Spline. In this section we want to find rules on how to choose the exponential parameters $\alpha_m = \alpha + j(m - P/2)\lambda_0$ for $m = 0, \ldots, P$ and the function $\gamma(t)$ in order to make FRI recovery techniques with these kernels as stable as possible. Finding the best parameters translates into the optimisation of the matrix of coefficients $\mathbf{C}$. Therefore we first determine the properties that $\mathbf{C}$ has to satisfy and then design the kernels that lead to the correct $\mathbf{C}$.

[2]Without entering in too many technical details, which are beyond the scope of this section, we observe that the condition $P+1 = N$ can be imposed only for blockwise sampling, e.g. when sampling periodic signals using N samples. This condition cannot be imposed on infinite length signals since the number of samples is in this case infinite, and sequential reconstruction algorithms will operate on blocks with possibly varying number of samples.

A. Correct choice of $\mathbf{C}$

The first step in the FRI reconstruction stage is that the vector of samples $\mathbf{y}$ is transformed into the vector of moments $\mathbf{s} = \mathbf{C}\mathbf{y}$. Therefore, our first aim is to get a well conditioned $\mathbf{C}$. Note that the relevance of this step was already highlighted in [3]. Note also that matrix $\mathbf{C}$ is composed of elements $c_{m,n} = c_{m,0}e^{\alpha_m n}$ at position (m, n), where $n = 0, \ldots, N-1$ and $m = 0, \ldots, P$:

$$\mathbf{C} = \underbrace{\begin{pmatrix} c_{0,0} & 0 & \cdots & 0 \\ 0 & c_{1,0} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & c_{P,0} \end{pmatrix}}_{\mathbf{D}} \underbrace{\begin{pmatrix} 1 & e^{\alpha_0} & \cdots & e^{\alpha_0(N-1)} \\ 1 & e^{\alpha_1} & \cdots & e^{\alpha_1(N-1)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & e^{\alpha_P} & \cdots & e^{\alpha_P(N-1)} \end{pmatrix}}_{\mathbf{V}},$$

where $\mathbf{D}$ is diagonal and $\mathbf{V}$ Vandermonde. Hence, to have a stable $\mathbf{C}$ we want the absolute values of the diagonal elements of $\mathbf{D}$ to be the same, i.e. $|c_{m,0}| = 1$. Moreover, we want the elements in $\mathbf{V}$ to be purely imaginary:

$$e^{\alpha_m} = e^{j\frac{\pi}{L}(2m-P)} \qquad \text{for } m = 0, \ldots, P, \text{ i.e. } \alpha = 0 \text{ and } \lambda_0 = 2\frac{\pi}{L}. \tag{18}$$

Clearly, purely imaginary α_m make the Vandermonde matrix $\mathbf{V}$ better conditioned [27]. We are therefore only left with the problem of finding the best L in (18). Since we have experimentally seen that FRI algorithms are able to reach the CRB (16) if $\mathbf{C}$ is well conditioned, one way to determine L is to chose the value that minimises (16) for the location of a single Dirac. It turns out the minimum is always achieved when $L = P + 1$, as shown in Figure 2 for various choices of P and L.

To some extent, this is not surprising since this choice ensures that the exponentials span the entire unit circle, which is well known to be the best configuration when recovering the parameters of a power sum series [28]. Finally, when we impose $P + 1 = N$ with $L = P + 1$, besides minimising (16), we also ensure that the moment-based CRB in (16) matches the sample-based bound in (14), leading to the best possible performance. In this situation, the matrix $\mathbf{C}$ ends up being square and unitary. This is the most stable numerical transformation since its condition number is one.

In summary, the best exponential reproducing kernels should reproduce exponentials with exponents of the form $\alpha_m = j\frac{\pi}{P+1}(2m - P)$ and have $|c_{m,0}| = 1$ for $m = 0, \ldots, P$. Finally, whenever possible, the order of the kernel should be $P + 1 = N$. In the next section we show how to obtain such kernels.

B. Exponential MOMS

Equipped with the analysis of the previous section, we now design optimal exponential reproducing kernels of maximum-order and minimum-support (eMOMS). We require that $|c_{m,0}| = 1$ for $m = 0, \ldots, P$

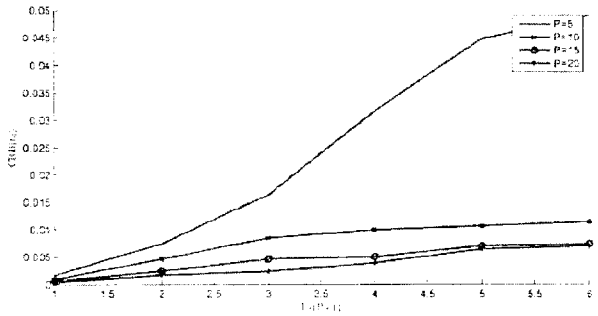


Figure 2. *CRB vs. L.* Here we plot various CRB values (16) ($\sigma = 1$) for coefficients satisfying $|c_{m,0}| = 1$, $m = 0, \ldots, P$ when we vary L in equation (18). For any value of P the CRB is minimised when $L = P + 1$ (note that all the lines are monotonically increasing).

and that the exponential parameters be of the form:

$$\alpha_m = j\omega_m = j\frac{\pi}{P+1}(2m - P) \qquad m = 0, \ldots, P, \tag{19}$$

however, we do not make any assumption on P since the following methodology can be used with any value of P.

By taking into account that any exponential reproducing kernel $\varphi(t)$ can be written as $\varphi(t) = \gamma(t) * \beta_0(t)$, we design $\gamma(t)$ so that $|c_{m,0}| = 1$ is satisfied. We note that, by using (3), we have that

$$\mathrm{e}^{\alpha_m t} = c_{m,0} \sum_{n \in \mathbb{Z}} \mathrm{e}^{\alpha_m n} \varphi(t - n).$$

Consequently

$$\begin{aligned} 1 &= c_{m,0} \sum_{n \in \mathbb{Z}} \mathrm{e}^{\alpha_m (n-t)} \varphi(t - n) \\ &\overset{(a)}{=} c_{m,0} \sum_{k \in \mathbb{Z}} \hat{\varphi}(\alpha_m + j2\pi k) \mathrm{e}^{j2\pi k t} \\ &\overset{(b)}{=} c_{m,0} \hat{\varphi}(\alpha_m), \end{aligned}$$

where (a) follows from Poisson summation formula[3] and (b) from the application of the generalised Strang-Fix conditions (9). Therefore, we have that for any exponential reproducing kernel $c_{m,0} =$

[3]Poisson summation formula: $\sum_{n \in \mathbb{Z}} f(t + nT) = \frac{1}{T} \sum_{k \in \mathbb{Z}} \hat{f}\left(\frac{2\pi k}{T}\right) \mathrm{e}^{j2\pi k \frac{t}{T}}$.

$\hat{\varphi}(\alpha_m)^{-1}$. We then realise that imposing $|c_{m,0}| = 1$ is equivalent to requiring $|\hat{\varphi}(\alpha_m)| = 1$. Finally, by using the fact that $\hat{\varphi}(\alpha_m) = \hat{\gamma}(\alpha_m)\hat{\beta}_{\vec{\alpha}}(\alpha_m)$ and evaluating the Laplace transforms at $\alpha_m = j\omega_m$, we arrive at the following condition on $\hat{\gamma}(\omega_m)$:

$$|\hat{\varphi}(\omega_m)| = |\hat{\gamma}(\omega_m)\hat{\beta}_{\vec{\alpha}}(\omega_m)| = 1 \tag{20}$$
$$\Leftrightarrow \quad |\hat{\gamma}(\omega_m)| = |\hat{\beta}_{\vec{\alpha}}(\omega_m)|^{-1},$$

where we now work with the Fourier transform of each function.

Among all the admissible kernels satisfying (20), we are interested in the one with the shortest support $P+1$. We thus consider the kernels given by a linear combination of various derivatives of the original E-Spline $\beta_{\vec{\alpha}}(t)$, i.e.:

$$\varphi(t) = \sum_{\ell=0}^{P} d_\ell \beta_{\vec{\alpha}}^{(\ell)}(t), \tag{21}$$

where $\beta_{\vec{\alpha}}^{(\ell)}(t)$ is the ℓth derivative of $\beta_{\vec{\alpha}}(t)$, also $\beta_{\vec{\alpha}}^{(0)}(t) = \beta_{\vec{\alpha}}(t)$, and d_ℓ is a set of coefficients. This is like saying that $\gamma(t)$ is a linear combination of the Dirac delta and its derivatives, up to order P [19]. These kernels are still able to reproduce the exponentials $e^{\alpha_m t}$ and are a variation of the maximal-order minimal-support (MOMS) kernels introduced in [29]. This is why we call them exponential MOMS (or eMOMS). They are also a specific case of the broader family of generalised E-Splines presented in [30]. The advantage of this formulation is twofold: first the modified kernel $\varphi(t)$ is of minimum support $P+1$, the same as that of $\beta_{\vec{\alpha}}(t)$; second we only need to find the coefficients d_ℓ that meet the constraint (20), in order to achieve $|c_{m,0}| = 1$. Using the Fourier transform of (21), which is given by:

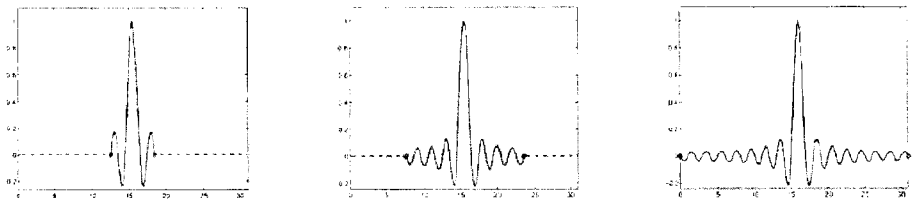
$$\hat{\varphi}(\omega) = \hat{\beta}_{\vec{\alpha}}(\omega) \sum_{\ell=0}^{P} d_\ell (j\omega)^\ell,$$

we realise that we can satisfy (20) by choosing the coefficients d_ℓ so that the resulting polynomial $\hat{\gamma}(\omega) = \sum_\ell d_\ell (j\omega)^\ell$ interpolates the set of points $(\omega_m, |\hat{\beta}_{\vec{\alpha}}(\omega_m)|^{-1})$ for $m = 0, 1, \ldots, P$.

Once we have designed the kernels satisfying that $c_{m,0}$ has modulus one for all m, we are left with a phase ambiguity. Determining the phase is equivalent to using a time shift Δ for the E-Spline in (21), introducing an additional degree of freedom. It is possible to show that, in order for exponential MOMS with $|c_{m,0}| = 1$ and parameters (19) to be continuous-time functions, then $c_{m,0} = |c_{m,0}|e^{j\omega_m\Delta}$ for $m = 0, \ldots, P$, where Δ is an integer bigger than or equal to 1 and smaller than or equal to P.

In Figure 3 we present some of the kernels obtained by implementing the procedure explained above. Interestingly, as shown in Appendix B, these specific functions always equal one period of the Dirichlet kernel. We also point out that when $P+1 = N$ the scenario derived using this family of exponential

reproducing kernels converges to the original FRI formulation of [2] when we periodise the input or, equivalently, the sampling kernel.



(a) $P = 5$ (b) $P = 15$ (c) $P = 30$

Figure 3. *Examples of exponential MOMS.* These are 3 of the 30 possible kernels with support $P + 1 \leq N = 31$ samples. They coincide with one period of the Dirichlet kernel of period $P + 1$ for P even or $2(P + 1)$ for P odd (see Appendix B). All of them are built selecting the phase of $c_{m,0}$ such that they are continuous-time functions centred around $\Delta = \lfloor \frac{P+1}{2} \rfloor$. They are illustrated placed in the middle of the sampling interval for $T = 1$.

V. Universal sampling of signals with finite rate of innovation

In the previous section we have shown how to design optimal exponential reproducing kernels for noisy FRI sampling. In many practical circumstances, however, the freedom to choose the sampling kernel $\varphi(t)$ is a luxury we may not have.

Essential in the FRI setting is the ability of $\varphi(t)$ to reproduce exponential functions, because this allows us to map the signal reconstruction problem to Prony's method in spectral-line estimation theory. In this section we relax this condition and consider any function $\varphi(t)$ for which the exponential reproduction property (3) does not necessarily hold. For these functions it is still possible to find coefficients $c_{m,n}$ such that the reproduction of exponentials is approximate rather than exact. We propose to use this approximate reproduction and the corresponding coefficients $c_{m,n}$ to retrieve FRI signals from the samples obtained using these kernels.

This new approach has several advantages: First, it is universal in that it can be used with any kernel $\varphi(t)$. In fact, as we shall show in the following sections, this new formulation does not even require an exact knowledge of the kernel. Second, while reconstruction of FRI signals with this new method is not going to be exact, we will show that in many cases a proper iterative algorithm can make the reconstruction error arbitrarily small. Finally, it can be used to increase the resiliency to noise of some unstable kernels proposed in the FRI literature. For example, kernels like polynomial splines or the Gaussian function

lead to very ill-conditioned reconstruction procedures. We show that by replacing the original C with the one formed from properly chosen coefficients $c_{m,n}$, based on approximate reproduction of exponentials, we achieve a much more stable reconstruction with the same kernels.

A. Approximate reproduction of exponentials

Assume we want to use a function $\varphi(t)$ and its integer shifts to approximate the exponential $e^{\alpha t}$. Specifically, we want to find the coefficients c_n such that:

$$\sum_{n\in\mathbb{Z}} c_n \varphi(t-n) \approxeq e^{\alpha t}. \tag{22}$$

This approximation is exact only when $\varphi(t)$ satisfies the generalized Strang-Fix conditions (9). For any other function it is of particular interest to find the coefficients c_n that best fit (22). In order to do so, we directly use[4] $c_n = c_0 e^{\alpha n}$ and introduce the 1-periodic function

$$g_\alpha(t) = c_0 \sum_{n\in\mathbb{Z}} e^{-\alpha(t-n)} \varphi(t-n). \tag{23}$$

We then find that approximating the exponential $e^{\alpha t}$ with integer shifts of $\varphi(t)$ can be transformed into approximating $g_\alpha(t)$ by the constant value 1. The reason is that we can rewrite (22) in the form of the right-hand side of (23) by substituting $c_n = c_0 e^{\alpha n}$ and moving $e^{\alpha t}$ to the left-hand side.

As a consequence of Poisson summation formula, we have that the Fourier series expansion of $g_\alpha(t)$ is given by

$$g_\alpha(t) = \sum_{l\in\mathbb{Z}} g_l e^{j2\pi lt} = \sum_{l\in\mathbb{Z}} c_0 \hat{\varphi}(\alpha + j2\pi l) e^{j2\pi lt}$$

and that our approximation problem reduces to:

$$g_\alpha(t) = \sum_{l\in\mathbb{Z}} c_0 \hat{\varphi}(\alpha + j2\pi l) e^{j2\pi lt} \approxeq 1. \tag{24}$$

This shows more deeply the relation between the generalised Strang-Fix conditions (9) and the approximation of exponentials. If $\varphi(t)$ satisfies the generalised Strang-Fix conditions (9) then $\hat{\varphi}(\alpha + j2\pi l) = 0$ for $l \in \mathbb{Z}\backslash\{0\}$ and (24) holds exactly when $c_0\hat{\varphi}(\alpha) = 1$. Otherwise, the terms $\hat{\varphi}(\alpha - j2\pi l)$ for $l \in \mathbb{Z}\backslash\{0\}$ do not vanish, and we can only find the coefficient c_0 so that $g_\alpha(t) \approxeq 1$. However, the closer the values $\hat{\varphi}(\alpha + j2\pi l)$ are to zero, the better the approximation in (22) is.

In general $\varphi(t)$ can be any function and we can find different sets of coefficients c_n in order for (22) to hold. Regardless of the coefficients we use, we can determine the accuracy of our approximation by

[4]The exact exponential reproducing coefficients always satisfy $c_n = c_0 e^{\alpha n}$. We now anticipate that different sets of approximation coefficients we derive throughout the section also have the same form.

using the Fourier series expansion of $g_\alpha(t)$. In fact, the error of approximating $f(t) = e^{\alpha t}$ by the function $s(t) = \sum_{n\in\mathbb{Z}} c_n \varphi(t-n)$ with coefficients $c_n = c_0 e^{\alpha n}$ is equal to:

$$\begin{aligned} \varepsilon(t) = f(t) - s(t) &= e^{\alpha t}[1 - g(t)] \\ &= e^{\alpha t}\left[1 - c_0 \sum_{l\in\mathbb{Z}} \hat{\varphi}(\alpha + j2\pi l) e^{j2\pi l t}\right]. \end{aligned} \tag{25}$$

Note that, if the Laplace transform of $\varphi(t)$ decays sufficiently quickly, very few terms of the Fourier series expansion are needed to have an accurate bound for the error.

A natural choice of the coefficients $c_n = c_0 e^{\alpha n}$ is the one given by the *least-squares approximation*. Despite the fact that $f(t)$ is not square-integrable, we can still obtain the coefficients by computing the orthogonal projection of $f(t)$ onto the subspace spanned by $\varphi(t-n)$ [31]. They take the form

$$c_n = \frac{\hat{\varphi}(-\alpha)}{\hat{a}_\varphi(e^\alpha)} e^{\alpha n},$$

where $\hat{a}_\varphi(e^\alpha) = \sum_{l\in\mathbb{Z}} a_\varphi[l] e^{-\alpha l}$ is the z-transform of $a_\varphi[l] = \langle \varphi(t-l), \varphi(t)\rangle$, evaluated at $z = e^\alpha$.

The least-squares approximation has the disadvantage that it requires the exact knowledge of $\varphi(t)$. However, as we stated before, if the Laplace transform of $\varphi(t)$ decays sufficiently quickly, we can assume the terms $\hat{\varphi}(\alpha + j2\pi l)$ are close to zero for $l \in \mathbb{Z}\backslash\{0\}$. In this case we have that the error in (25) is easily minimised by choosing $c_0 = \hat{\varphi}(\alpha)^{-1}$. We denote this second type of approximation *constant least-squares*. Besides its simplicity, a second advantage of choosing $c_n = \hat{\varphi}(\alpha)^{-1} e^{\alpha n}$ is that it requires only the knowledge of the Laplace transform of $\varphi(t)$ at α. If we put ourselves in the FRI setting where we require the approximate reproduction of the exponentials $e^{\alpha_m t}$ with $m = 0, \ldots, P$, then this simplified formulation needs only the knowledge of the Laplace transform of $\varphi(t)$ at α_m, $m = 0, \ldots, P$.

Finally, a third interesting choice of coefficients is the one that ensures that $s(t)$ *interpolates* $f(t)$ exactly at $t = l \in \mathbb{Z}$ [32], [18]. These coefficients are as follows:

$$c_n = \frac{1}{\sum_{l\in\mathbb{Z}} e^{-\alpha l}\varphi(l)} e^{\alpha n}.$$

Note that in order to use the interpolation coefficients we only need information on $\varphi(t)$ at integer instants of time. We summarise the previous results in Table I.

According to our experience, the least-square coefficients provide a smaller error (25) when the exponential $e^{\alpha t}$ to approximate has an exponent α that is not purely imaginary. Otherwise, the constant least-squares coefficients are just as good. Interpolation coefficients are very easy to compute given the values of the kernel at integer points in time. However, they always provide a worse approximation quality. Therefore, for the rest of the paper, we use constant least-squares coefficients.

Table I

COEFFICIENTS FOR THE APPROXIMATE REPRODUCTION (22)

Type	Coefficients
Least-squares approximation	$c_n = \frac{\hat{\varphi}(-\alpha)}{\hat{a}_\varphi(\mathrm{e}^\alpha)}\mathrm{e}^{\alpha n}$
Constant least-squares	$c_n = \hat{\varphi}(\alpha)^{-1}\mathrm{e}^{\alpha n}$
Interpolation	$c_n = \frac{1}{\sum_{k\in\mathbb{Z}}\mathrm{e}^{-\alpha k}\varphi(k)}\mathrm{e}^{\alpha n}$

We show an example of the above analysis in Figure 4. Here we want to approximate exponentials using linear combinations of linear splines. First, note that this spline reproduces polynomials of orders 0 and 1 exactly. This is shown in Figure 4 (a-b). With the same function, we approximately reproduce 4 complex exponentials $\mathrm{e}^{j\frac{\pi}{16}(2m-7)t}$ for $m = 3, \ldots, 0$, with the constant least-squares coefficients $c_{m,n} = \hat{\varphi}(\frac{\pi}{16}(2m-7))^{-1}\mathrm{e}^{\frac{\pi}{16}(2m-7)n}$. We present the approximation of their real part in Figure 4 (c-f). We notice that some exponentials are better approximated than others, in this example the ones with lower frequency. If we used a higher order spline, the approximation quality would improve. However, we have chosen a linear spline for illustration purposes. Also note that the number of exponentials that can be approximated is now independent of the order of the spline.

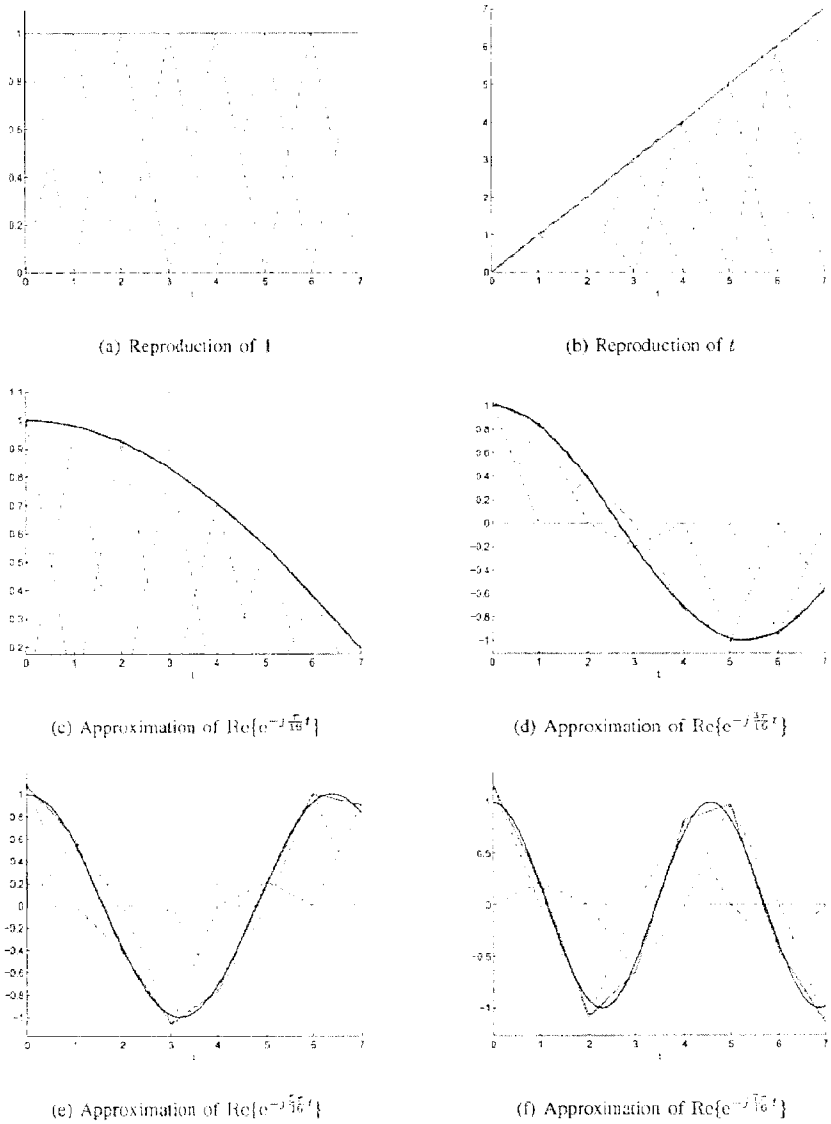
B. Approximate FRI recovery

Consider again the stream of Diracs $x(t) = \sum_{k=0}^{K-1} a_k\delta(t-t_k)$ and the samples

$$y_n = \left\langle x(t), \varphi\left(\frac{t}{T} - n\right)\right\rangle = \sum_{k=0}^{K-1} a_k\varphi\left(\frac{t_k}{T} - n\right). \tag{26}$$

We want to retrieve the locations and amplitudes of the Diracs from the samples (26), but now we make no assumption on the sampling kernel. We find proper coefficients for $\varphi(t)$ to approximate the exponentials $\mathrm{e}^{\alpha_m t}$, where $m = 0, \ldots, P$, $\alpha_m = \alpha_0 + m\lambda$, and $\alpha_0, \lambda \in \mathbb{C}$. From the previous section we know that a good approximation is achieved if we choose $c_{m,n} = c_{m,0}\mathrm{e}^{\alpha_m n}$ with $c_{m,0} = \hat{\varphi}(\alpha_m)^{-1}$. We thus only need to know the Laplace transform of $\varphi(t)$ at α_m, $m = 0, \ldots, P$. Also, note that P no longer needs to be related to the support of $\varphi(t)$, but we can use any value subject to $P \geqslant 2K-1$.

In order to retrieve the innovation parameters (t_k, a_k), we proceed in the same way as in the case of



(a) Reproduction of 1

(b) Reproduction of t

(c) Approximation of $\mathrm{Re}\{e^{-j\frac{\pi}{19}t}\}$

(d) Approximation of $\mathrm{Re}\{e^{-j\frac{3\pi}{16}t}\}$

(e) Approximation of $\mathrm{Re}\{e^{-j\frac{5\pi}{16}t}\}$

(f) Approximation of $\mathrm{Re}\{e^{-j\frac{7\pi}{16}t}\}$

Figure 4. *B-Spline kernel reproduction and approximation capabilities.* Figures (a-b) show the exact reproduction of polynomials of orders 0 and 1 with a linear spline. Figures (c-f) show the approximation of the real parts of 4 complex exponentials: $e^{j\frac{\pi}{16}(2m-7)t}$ for $m = 3, \ldots, 0$ with constant least-squares coefficients (Table I), using a linear spline. We plot the weighted and shifted versions of the splines with dashed blue lines, the reconstructed polynomials and exponentials with red solid lines, and the original functions to be reproduced with solid black lines.

exact reproduction of exponentials, but now we have that the moments are

$$s_m = \sum_{n=0}^{N-1} c_{m,n} y_n = \left\langle x(t), \sum_{n=0}^{N-1} c_{m,n} \varphi\left(\frac{t}{T} - n\right) \right\rangle \tag{27}$$
$$= \sum_{k=0}^{K-1} x_k u_k^m - \underbrace{\sum_{k=0}^{K-1} a_k \varepsilon_m\left(\frac{t_k}{T}\right)}_{\zeta_m}$$

where $x_k = a_k e^{\alpha_0 \frac{t_k}{T}}$ and $u_k = e^{\lambda \frac{t_k}{T}}$. There is a model mismatch due to the approximation error $\varepsilon_m(t)$ of (25), equal to ζ_m. We treat it as noise, and retrieve the parameters of the signal using the methods of Section III. The model mismatch depends on the quality of the approximation, dictated by the coefficients $c_{m,n}$, the values α_m and P, and the kernel $\varphi(t)$. If ζ_m is negligible when compared to other forms of noise then the procedure is sufficiently good. In close-to-noiseless settings, however, the estimation of the Diracs can be refined using the iterative algorithm shown in the box Algorithm 2. The basic idea of the algorithm is that, given an estimate of the locations of the Diracs, we can compute an approximation of ζ_m and use it to refine the computation of the moments s_m.

Algorithm 2 Recovery of a train of K Diracs using approximation of exponentials

1: Compute the moments $s_m^0 = \sum_n c_{m,n} y_n$, from the original data and set $s_m^1 = s_m^0$.

2: Build the system of equations (6) using s_m^i and retrieve the annihilating filter coefficients h_m, for $m = 0, \ldots, M$, where $M \geq K$.

3: Calculate the values u_k^i from the roots of h_m, and determine the locations t_k^i, for the ith iteration.

4: Find the amplitudes a_k^i from x_k^i, obtained by solving the first K consecutive equations in (5).

5: Recalculate the moments for the next iteration $i+1$ by removing the model mismatch from the moments calculated based on the original data by using

$$s_m^{i+1} = s_m^0 + \sum_{k=0}^{K-1} a_k^i \varepsilon_m(t_k^i),$$

for $m = 0, \ldots, P$ and where $\varepsilon_m(t)$ is the error of the approximation (25).

6: Repeat steps 2 to 5 until convergence of the values (a_k^i, t_k^i).

C. How to select the parameters α_m

In Section IV we have determined that, if we have full control on the design of the sampling kernel, we should use as many moments as samples: $P + 1 = N$, the exponential parameters should be purely

imaginary and of the form $\alpha_m = j\frac{\pi}{P+1}(2m-P)$ and the coefficients $c_{m,n}$ should be such that $|c_{m,0}| = 1$ for $m = 0, \ldots, P$. However, in the approximated FRI scenario, the sampling kernel is fixed and we can only choose the number of moments $P+1$ and the values α_m but we cannot impose $|c_{m,0}| = 1$.

This fact leads to a trade-off in the choice of α_m. On the one hand, we want them to be purely imaginary: $\alpha_m = j\omega_m$ and to span the entire unit circle. On the other hand, we want the values $c_{m,0} = \hat{\varphi}(\omega_m)^{-1}$ to have modulus as close as possible to 1. Since $|\hat{\varphi}(\omega)|$ is normally a low-pass filter, the condition $|c_{m,0}| \approxeq 1$ for $m = 0, \ldots, P$ is satisfied when all ω_m are very close to zero, which is contrary to spanning the unit circle. We therefore choose the exponential parameters to be of the form:

$$\alpha_m = j\omega_m = j\frac{\pi}{L}(2m-P) \quad m = 0, \ldots, P. \tag{28}$$

and then determine P and L that optimise the above trade-off.

Again the criterion we follow is to choose the values of P and L that minimise the CRB (16) when retrieving the location of a single Dirac. In most cases analysed, we found that the best P is normally greater or equal than the support of the sampling kernel $\varphi(t)$ and that L is in the range $P+1 \leqslant L \leqslant 4(P+1)$. We show an example of the choice of L in Figure 5 for the case where $\varphi(t)$ is a B-Spline of order 6. Here, we first determine a value for L and compute parameters (28). Then, we calculate the coefficients $c_{m,n} = \hat{\varphi}(\omega_m)^{-1}e^{j\omega_m n}$ for $m = 0, \ldots, P$, where $\hat{\varphi}(\omega)$ is the Fourier transform of $\varphi(t)$. We finally build matrix C and compute (16) for a single Dirac. The minima are always around $L = 1.5(P+1)$.

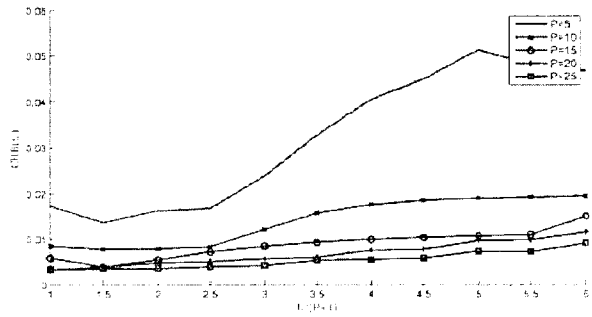


Figure 5. *CRB vs. L.* Here we plot different CRB values (16) ($\sigma = 1$) for exponential parameters (28) when we vary L. We use $c_{m,n} = \hat{\varphi}(\omega_m)^{-1}e^{j\omega_m n}$ for $m = 0, \ldots, P$, where $\hat{\varphi}(\omega)$ is the Fourier transform of a B-Spline of order 6. Note that the minima are always attained around $L = 1.5(P+1)$.

VI. Simulations

We now present simulation results to validate the main contributions of the paper. Specifically, we show the performance of the eMOMS kernels introduced in Section IV-B and of the approximate FRI recovery method introduced in Section V for the case where the sampling kernel is a polynomial spline.

A. The experimental setup

We take N samples following the scheme of Figure 1 by directly calculating $y_n = \sum_{k=0}^{K-1} a_k \varphi\left(\frac{t_k}{T} - n\right)$ for $n = 0, \ldots, N-1$, since we have a train of K Diracs as the input. We corrupt the samples with additive white Gaussian noise of variance σ^2, chosen according to the target signal-to-noise ratio defined as $\mathrm{SNR(dB)} = 10 \log \frac{\|\mathbf{y}\|^2}{N\sigma^2}$. We finally compute the $P+1$ noisy moments and then retrieve the innovation parameters $\{a_k, t_k\}_{k=0}^{K-1}$ of the input using the matrix pencil method as in [11].

We are mainly interested in the error in the estimation of the time locations, since these are the most challenging parameters to retrieve. For each Dirac, we show the standard deviation of this error:

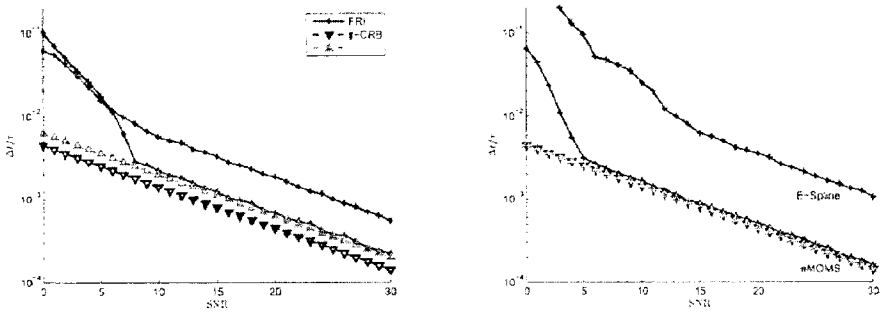
$$\Delta t_k = \sqrt{\frac{\sum_{i=0}^{I-1} (\hat{t}_k^{(i)} - t_k)^2}{I}} \qquad k = 0, \ldots, K-1, \tag{29}$$

where $\hat{t}_k^{(i)}$ are the estimated time locations at iteration i and I is the total number of iterations. We calculate (29) for a range of fixed signal-to-noise ratios and average the effects using $I = 1000$ noise realisations for each SNR. We compare the performance (29) with the square root of the variance predicted by the two different Cramér–Rao bounds (CRB) of Section III: the sample-based CRB (14) and the moment-based CRB (16).

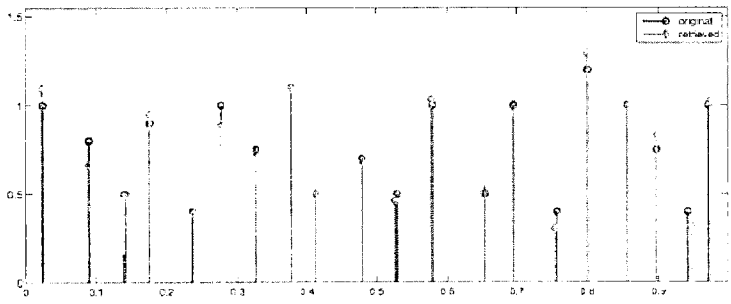
B. Exponential MOMS

In Figure 6 (a-b) we present simulation results when we retrieve $K = 2$ Diracs from $N = 31$ samples using a standard E-Spline and the exponential MOMS kernels of Section IV-B. The former are characterised by purely imaginary exponents $\alpha_m = j\frac{\pi}{2(P+1)}(2m - P)$ for $m = 0, \ldots, P$.

We see that for any order $P+1$ we consider, eMOMS outperform E-splines. Moreover, eMOMS always achieve the moment-based CRB (in red and denoted s-CRB in the legend). This bound gets closer to the sample-based CRB (in black and denoted y-CRB in the legend) as the value of $P+1$ increases and as expected matches it when $P+1 = N$. To further illustrate the stability of eMOMS, in Figure 6 (c) we show the retrieval of $K = 20$ Diracs randomly spaced over $\tau = NT = 1$ and with arbitrary amplitudes. The signal-to-noise ratio is 15dB, and we use $N = 61$ samples with $P+1 = N$.



(a) eMOMS $P = 15$ (b) eMOMS $P = 30$



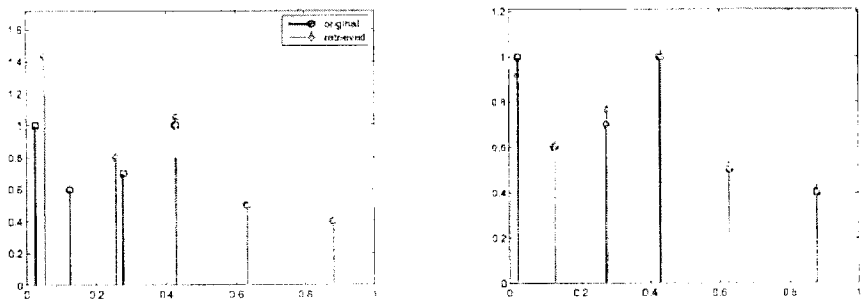
(c) Retrieval of $K = 20$ Diracs

Figure 6. *Performance of exponential MOMS kernels.* (a-b) compare the performance of E-Splines vs. exponential MOMS kernels of different orders $P + 1$ when noise is added to $N = 31$ samples. We show the recovery of the first of $K = 2$ Diracs. We note that eMOMS always outperform E-splines and achieve the moment-based CRB (s-CRB). This bound gets closer to the sample-based CRB (y-CRB) as the value of $P + 1$ increases and as expected matches it when $P + 1 = N$. Finally, (c) shows the retrieval of $K = 20$ Diracs randomly spaced over $\tau = NT = 1$. The signal-to-noise ratio is 15dB, and we use $N = 61$ samples and $P + 1 = N$ moments.

C. Approximate FRI

We now show results of the approximate retrieval procedure of SectionV-B when we take samples using a B-Spline kernel of order $P+1$ and approximate exponentials $e^{\alpha_m t}$, with $\alpha_m = j\frac{\pi}{2(P_2+1)}(2m - P_2)$ and $m = 0, \ldots, P_2$. We obtain the coefficients using the constant least squares formula $c_{m,n} = \hat{\varphi}(\alpha_m)^{-1} e^{\alpha_m n}$, where $\hat{\varphi}(s)$ represents the Laplace transform of the B-Spline. We compare the performance with the traditional recovery strategy based in this case on the exact reproduction of polynomials.

In Figure 7 (a-b) we compare the estimation of $K = 6$ Diracs located at random over $\tau = NT$, from $N = 31$ noisy samples, and with SNR = 20dB. For (a) we sample and reconstruct with a B-Spline of order $P + 1 = 26$. We note that not all the Diracs can be found. For (b) we sample with a B-Spline of order $P + 1 = 6$ but we build $P_2 + 1 = 26$ exponential moments in order to increase the accuracy of the approximate reconstruction. The average error of the estimation of the location of the six Diracs is around 10^{-2} for the Diracs properly recovered in (a), and 10^{-3} for the approximate retrieval of (b).



(a) Retrieval of $K = 6$ Diracs. Default

(b) Approximation

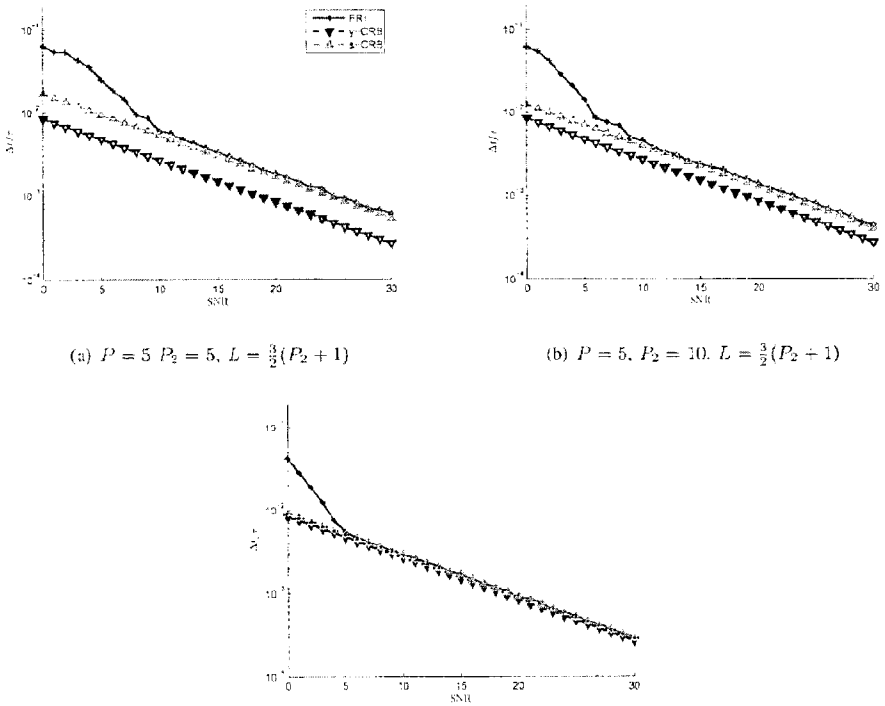
Figure 7. *B-Spline kernel behaviour.* We recover $K = 6$ Diracs from $N = 31$ noisy samples. (a) is for the default polynomial recovery of [1], enhanced using pre-whitening. We sample with a kernel of order $P + 1 = 26$ and use the same number of moments. (b) shows the results obtained using the approximated recovery with $\alpha_m = \frac{\pi}{2(P_2+1)}(2m - P_2)$ where $m = 0, \ldots, P_2$. Here we sample with a kernel of order $P + 1 = 6$ but also build $P_2 + 1 = 26$ moments. The SNR in both cases is 20dB.

We show further results when we use the approximate method to retrieve $K = 2$ Diracs from $N = 31$ samples taken by a B-Spline kernel of order $P + 1 = 6$. Even when we fix the order of the kernel, we can reconstruct an increasing numbers of moments $P_2 + 1$ to improve the performance. Figures 8 (a-d) are for parameters $\alpha_m = j\frac{\pi}{L}(2m - P_2)$ with $m = 0, \ldots, P_2$ and $L = \frac{3}{2}(P_2 + 1)$. As the number of moments $P_2 + 1$ increases, the performance is better and eventually reaches the sample-based CRB.

VII. Conclusions

In this paper we have considered the FRI reconstruction problem in the presence of noise. We have first revisited existing results in the noiseless setting, and the most effective treatment of noise in the current literature. Then, we have studied robust alternatives to the previous line of work.

More specifically, our contribution is twofold: We determined how to design optimal exponential reproducing kernels, in the sense that they lead to the most stable signal reconstruction. Moreover, we



(a) $P = 5$ $P_2 = 5$, $L = \frac{3}{2}(P_2 + 1)$

(b) $P = 5$, $P_2 = 10$, $L = \frac{3}{2}(P_2 + 1)$

(c) $P = 5$, $P_2 = 30$, $L = \frac{3}{2}(P_2 + 1)$

Figure 8. *Approximated retrieval using a B-Spline.* These figures show the error in the estimation of the first Dirac out of $K = 2$ retrieved using the approximated FRI recovery. We show how, even when we fix the order of the kernel $P + 1$, we can reconstruct any number of moments $P_2 + 1$ and improve the performance. In fact, with the appropriate choice $L = \frac{3}{2}(P_2 + 1)$ the performance improves until the sample-based CRB is reached.

departed from the ideal situation in which we have full control of the sampling kernel and considered the case where we are given corrupted samples taken with an arbitrary acquisition device. In this situation, we developed a universal FRI reconstruction strategy that works with any kernel. In contrast to the original FRI framework which tries to find the exact parameters of the input signal, we propose an approximate recovery of the input based on the approximate reproduction of exponentials. The advantage of our new approach is that it can be applied to any sampling kernel and provides a more stable and precise reconstruction than specific classes of kernels used in the past.

APPENDIX A

GENERALISED STRANG-FIX CONDITIONS

An exponential reproducing kernel is any function $\varphi(t)$ that, together with a linear combination of its shifted versions, can generate polynomial exponential functions of the form $t^r e^{\alpha_m t}$. The parameters α_m are complex, with $m = 0, \ldots, P$ and $r = 0, \ldots, R$. Exponential reproducing kernels are a generalisation of the family of polynomial reproducing kernels. The latter are characterised by the well-known Strang-Fix conditions [1], [33], which state that a kernel $\psi(t)$ is able to reproduce polynomials, i.e.:

$$t^r = \sum_{n \in \mathbb{Z}} c_r \psi(t - n),$$

if and only if

$$\hat{\psi}(0) \neq 0 \text{ and } \hat{\psi}^{(r)}(2\pi l) = 0 \text{ for } l \neq 0 \text{ and } r = 0, \ldots, R.$$

Here, $\hat{\psi}(\omega)$ is the Fourier transform of $\psi(t)$, and $\hat{\psi}^{(r)}(\omega)$ represents its rth order derivative.

In a similar fashion to polynomial reproducing kernels, it is possible to derive a more general form of the Strang-Fix conditions for exponential reproducing kernels. More specifically, a kernel $\varphi(t)$ is able to reproduce exponential polynomials, i.e.:

$$t^r e^{\alpha_m t} = \sum_{n \in \mathbb{Z}} c_{m,n,r} \varphi(t - n),$$

if and only if

$$\hat{\varphi}^{(r)}(\alpha_m) \neq 0 \text{ and } \hat{\varphi}^{(r)}(\alpha_m + 2j\pi l) = 0 \text{ for } l \neq 0 \text{ and } r = 0, \ldots, R.$$

Here, $\hat{\varphi}^{(r)}(s)$ represents the rth order derivative of the double-sided Laplace transform of $\varphi(t)$. The proof is obtained from the Strang-Fix conditions for polynomial reproducing kernels, by considering the function $\psi(t) = e^{-\alpha_m t}\varphi(t)$ that clearly reproduces t^r. Therefore, $\psi(t)$ satisfies the Strang-Fix conditions because it is a polynomial reproducing kernel. In addition, the Fourier transform of $\psi(t)$ is related to the Laplace transform of $\varphi(t)$ as: $\hat{\psi}(\omega) = \hat{\varphi}(\alpha_m + j\omega)$. By simply applying the Strang-Fix conditions to $\hat{\psi}(\omega)$ we arrive at the generalised Strang-Fix conditions for $\hat{\varphi}(\omega)$, which completes the proof.

APPENDIX B

THE FAMILY OF EMOMS INCLUDES THE SOS AND DIRICHLET KERNELS

Let us consider the exponential reproducing kernel $\varphi_0(t) = \varphi\left(t + \frac{P+1}{2}\right)$ of support $P+1$ and centred in zero, with $\varphi(t) = \gamma(t) * \beta_{\vec{\alpha}}(t)$, where $\beta_{\vec{\alpha}}(t)$ has Fourier transform given by (11). We restrict our analysis to P being even and we use exponential parameters

$$\alpha_m = j\omega_m = j\frac{\pi}{P+1}(2m - P), \tag{30}$$

where $m = 0, \ldots, P$. We next use the periodic extension of $\varphi_0(t)$, that is $\varphi_{P+1}(t) = \sum_{l \in \mathbb{Z}} \varphi_0(t + l(P+1))$, which can be written as follows:

$$\varphi_{P+1}(t) = \frac{1}{P+1} \sum_{k \in \mathbb{Z}} \hat{\varphi}_0 \left(\frac{2\pi k}{P+1} \right) e^{j \frac{2\pi k}{P+1} t}, \tag{31}$$

where the last term follows from the application of Poisson summation formula. The case of P being odd can be derived likewise, but by periodising over $2(P+1)$. Also note that the Fourier transform of the shifted kernel $\varphi_0(t)$ can be written as:

$$\hat{\varphi}_0(\omega) = \gamma(\omega) \prod_{m=0}^{P} \operatorname{sinc} \left(\frac{\omega - \omega_m}{2} \right). \tag{32}$$

If we denote

$$\hat{\varphi}_0(\omega_m) = |\hat{\varphi}(\omega_m)| = |\hat{\gamma}(\omega_m) \hat{\beta}_{\vec{\alpha}}(\omega_m)| = \eta_m, \tag{33}$$

we design exponential reproducing kernels of maximum order and minimum support (eMOMS), different from those of Section IV-B, which still correspond to a specific subfamily of the generalised exponential reproducing kernels of [30].

In (31) the Fourier transform $\hat{\varphi}_0(\omega)$ is evaluated at $\omega_k = \frac{2\pi k}{P+1}$. Taking into account (33), we know that $\hat{\varphi}_0(\omega_k) = \eta_k$ for the range $k = -\frac{P}{2}, \ldots, \frac{P}{2}$. We also have that $\hat{\varphi}_0(\omega_k) = 0$ for any other k, because we can find a term in the product (32) equal to $\operatorname{sinc}(\ell\pi) = 0$. Therefore, Equation (31) can be reduced to:

$$\varphi_{P+1}(t) = \frac{1}{P+1} \sum_{k=-\frac{P}{2}}^{\frac{P}{2}} \eta_k e^{j \frac{2\pi k}{P+1} t}. \tag{34}$$

Note that when the values $\eta_k = 1$ for all k, then equation (34) reduces to one period of the Dirichlet kernel of period $P+1$:

$$\varphi_{P+1}(t) = \frac{1}{P+1} \sum_{k=-\frac{P}{2}}^{\frac{P}{2}} e^{j \frac{2\pi k}{P+1} t} = \frac{1}{P+1} \frac{\sin(\pi t)}{\sin(\frac{\pi t}{P+1})}.$$

And this is precisely the family of eMOMS kernels of Section IV-B.

To end, we now consider one period of (34) and we use $t = \frac{x}{T}$. Moreover, we restrict the number of samples $N = P+1$ and denote $\tau = NT = (P+1)T$. Then we get the time domain definition of the SoS kernel [3]:

$$\varphi_{P+1} \left(\frac{x}{T} \right) = g(x) = \operatorname{rect} \left(\frac{x}{\tau} \right) \frac{1}{N} \sum_{k \in \mathcal{K}} \eta_k e^{\frac{2\pi k}{\tau} x}.$$

Here, the number of samples N needs to be odd and the set of indices $\mathcal{K} = \{-\frac{N-1}{2}, \ldots, \frac{N-1}{2}\}$.

REFERENCES

[1] P. L. Dragotti, M. Vetterli, and T. Blu, "Sampling Moments and Reconstructing Signals of Finite Rate of Innovation: Shannon Meets Strang-Fix," *IEEE Transactions on Signal Processing*, vol. 55, pp. 1741–1757, May 2007.

[2] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *IEEE Transactions on Signal Processing*, vol. 50, pp. 1417–1428, June 2002.

[3] R. Tur, Y. C. Eldar, and Z. Friedman, "Innovation Rate Sampling of Pulse Streams with Application to Ultrasound Imaging," *IEEE Transactions on Signal Processing*, vol. 59, pp. 1827–1842, April 2011.

[4] P. Stoica and R. L. Moses, *Introduction to Spectral Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 2000.

[5] A. Hormati and M. Vetterli, "Compressive Sampling of Multiple Sparse Signals Having Common Support Using Finite Rate of Innovation Principles," *IEEE Signal Processing Letters*, vol. 18, pp. 331–334, May 2011.

[6] J. Berent, P. L. Dragotti, and T. Blu, "Sampling Piecewise Sinusoidal Signals With Finite Rate of Innovation Methods," *IEEE Transactions on Signal Processing*, vol. 58, pp. 613–625, February 2010.

[7] I. Maravic and M. Vetterli, "Exact Sampling Results for Some Classes of Parametric Non-Bandlimited 2-D Signals," *IEEE Transactions on Signal Processing*, vol. 52, pp. 175–189, January 2004.

[8] P. Shukla and P. L. Dragotti, "Sampling Schemes for Multidimensional Signals with Finite Rate of Innovation," *IEEE Transactions on Signal Processing*, vol. 55, pp. 3670–3686, July 2007.

[9] H. Pan, T. Blu, and P. L. Dragotti, "Sampling Curves with Finite Rate of Innovation," *Proceedings of the Ninth International Workshop on Sampling Theory and Applications (SampTA'11), Singapore*, May 2011.

[10] C. Chen, P. Marziliano, and A. C. Kot, "2D Finite Rate of Innovation Reconstruction Method for Step Edge and Polygon Signals in the Presence of Noise ," *IEEE Transactions on Signal Processing*, vol. 60, pp. 2851–2859, June 2012.

[11] I. Maravic and M. Vetterli, "Sampling and reconstruction of signals with finite rate of innovation in the presence of noise," *IEEE Transactions on Signal Processing*, vol. 53, pp. 2788–2805, August 2005.

[12] T. Blu, P. L. Dragotti, M. Vetterli, P. Marziliano, and L. Coulot, "Sparse Sampling of Signal Innovations," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 31–40, 2008.

[13] V. Y. F. Tan and V. K. Goyal, "Estimating signals with finite rate of innovation from noisy samples: A stochastic algorithm," *IEEE Transactions on Signal Processing*, vol. 56, pp. 5135–5146, October 2008.

[14] A. Erdozain and P. M. Crespo, "A new stochastic algorithm inspired on genetic algorithms to estimate signals with finite rate of innovation from noisy samples," *Signal Processing*, vol. 90, pp. 134–144, January 2010.

[15] A. Erdozain and P. M. Crespo, "Reconstruction of aperiodic FRI signals and estimation of the rate of innovation based on the state space method," *Signal Processing*, vol. 91, no. 8, pp. 1709–1718, 2011.

[16] J. A. Urigüen, P. L. Dragotti, and T. Blu, "On the exponential reproducing kernels for sampling signals with finite rate of innovation," *Proceedings of the Ninth International Workshop on Sampling Theory and Applications (SampTA'11), Singapore*, May 2011.

[17] I. Khalidov, T. Blu, and M. Unser, "Generalized L-Spline Wavelet Bases," in *Proceedings of the SPIE Conference on Mathematical Imaging: Wavelet XI*, vol. 5914, (San Diego CA, USA), pp. 59140F–1–59140F–8, July 31-August 3, 2005.

[18] M. Unser and T. Blu, "Cardinal Exponential Splines: Part I—Theory and Filtering Algorithms," *IEEE Transactions on Signal Processing*, vol. 53, pp. 1425–1438, April 2005.

[19] A. Ron, "Factorization theorems for univariate splines on regular grids," *Israel J. Math.*, vol. 70, no. 1, pp. 48–68, 1990.

[20] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice Hall, 1993.

[21] J. A. Cadzow, "Signal Enhancement - A Composite Property Mapping Algorithm," *Acoustics, Speech and IEEE Transactions on Signal Processing*, vol. 36, pp. 49–62, January 1988.

[22] Y. Hua and T. K. Sakar, "Matrix Pencil Method for Estimating Parameters of Exponentially Damped Undamped Sinusoids in Noise," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 38, pp. 814–824, May 1990.

[23] B. D. Rao, "Model based processing of signals: A state space approach," *Proceedings of the IEEE*, vol. 3980, pp. 283–309, February 1992.

[24] B. De Moor, "The Singular Value Decomposition and Long and Short Spaces of Noisy Matrices," *IEEE Transactions on Signal Processing*, vol. 41, pp. 2826–2838, September 1993.

[25] Y. C. Eldar and A. V. Oppenheim, "MMSE Whitening and Subspace Whitening," *IEEE Trans. Signal Processing*, vol. 49, pp. 1846–1851, July 2003.

[26] E. Ollila, "On the Cramér-Rao bound for the constrained and unconstrained complex parameters," *Sensor Array and Multichannel Signal Processing Workshop*, pp. 414–418, July 2008.

[27] L. Berman and A. Feuer, "On perfect conditioning of vandermonde matrices on the unit circle," *Electronic Journal of Linear Algebra*, vol. 16, pp. 157–161, July 2007.

[28] G. H. Golub, P. Milanfar, and J. Varah, "A stable numerical method for inverting shape from moments," *SIAM Journal on Scientific Computing (SISC)*, vol. 21, pp. 1222–1243, December 1999.

[29] T. Blu, P. Thevenaz, and M. Unser, "MOMS: maximal-order interpolation of minimal support," *Image Processing, IEEE Transactions on*, vol. 10, pp. 1069–1080, July 2001.

[30] M. Unser, "Cardinal Exponential Splines: Part II—Think Analog, Act Digital," *IEEE Transactions on Signal Processing*, vol. 53, pp. 1439–1449, April 2005.

[31] M. Unser, A. Aldroubi, and M. Eden, "Polynomial Spline Signal Approximations: Filter Design and Asymptotic Equivalence with Shannon's Sampling Theorem," *IEEE Transactions on Information Theory*, vol. 38, pp. 95–103, January 1992.

[32] M. Unser, "Sampling-50 years after Shannon," in *Proceedings of the IEEE*, pp. 569–587, April 2000.

[33] G. Strang and G. Fix, "Fourier analysis of the finite element variational method," *Constructive Aspect of Functional Analysis*, pp. 796–830, 1971.

Claims:

1. A signal processing method for estimating a frequency domain representation of a signal from a series of samples distorted by an instrument function, the method comprising:

obtaining the series of samples;

obtaining a set of coefficients that fit a set of basis functions to a complex exponential function wherein the set of basis functions comprises a plurality of basis functions each defined by a shifted version of the instrument function in a signal domain;

estimating the frequency domain representation of the signal based on the series of samples and the coefficients;

wherein the estimate of the instrument function is based on a characterisation of the instrument function in the frequency domain at frequencies associated with the complex exponential function.

2. The method of claim 1 in which the samples are separated in the signal domain by an inter-sample interval, and each basis function is shifted in the signal domain by an integer number of inter-sample intervals, and the fit comprises a projection into the space defined by the basis functions.

3. The method of claim 1 in which the samples are separated in the signal domain by an inter-sample interval, and each basis function is shifted in the signal domain by a non-integer number of inter-sample intervals, and the fit is performed by reducing a merit function.

4. The method of claim 1, 2, or 3 wherein the coefficients comprise a frequency domain representation of a respective corresponding one of the basis functions.

5. The signal processing method of any preceding claim in which the complex exponential function comprises signal domain oscillatory components, and the frequencies at which the instrument function is characterised comprise the frequencies of the oscillatory components.

6. The signal processing method of claim 5 in which the frequencies are evenly spaced in the frequency domain.

7. The signal processing method of claim 5 or 6, in which the frequencies are selected to reduce the variance in the magnitude of the instrument function at the frequencies whilst increasing the spacing between the frequencies.

8. The signal processing method of claim 5, 6 or 7 in which the signal belongs to a class of finite rate innovation, FRI, signal, and the frequencies are selected based on the performance of a maximum likelihood estimator of innovation parameters of that class of signal.

9. The signal processing method of any preceding claim further comprising providing an estimate of the signal, based on the frequency domain representation.

10. The signal processing method of any preceding claim comprising:

obtaining an error in the representation provided by the coefficients and the basis functions of the complex exponential function;

and, based on this error, determining whether to re-estimate the frequency domain representation of the signal.

11. The signal processing method of claim 10, in which determining whether to re-estimate the frequency domain representation of the signal comprises obtaining an estimate of noise in the original signal, and performing a comparison based on the estimated noise and the error in the representation of the complex exponential function.

12. The signal processing method of claim 10 or 11 in which re-estimating the frequency domain representation of the signal comprises determining a filter configured to minimise the energy of the frequency domain estimation representation of the signal, and adjusting the frequency domain representation based on the filter.

13. The signal processing method of claim 12 in which adjusting based on the filter comprises obtaining an estimate of parameters describing the location, in the signal domain, of the signal using the roots of the filter.

14. The signal processing method of claim 10 or 11 in which re-estimating the frequency domain representation of the signal comprises obtaining an estimate of parameters describing the location, in the signal domain, of the signal by deriving eigenvalues from a matrix pencil problem.

15. The method of any preceding claim, in which the signal comprises an image, and further comprising scaling the frequency domain representation of the image based on the frequencies associated with the complex exponential function, determining a set of filter coefficients, b_{kj}, configured to minimise the energy of the frequency domain representation of the image, and determining a boundary in the image from the filter coefficients.

16. An image coregistration method comprising the method of claim 15, and further comprising:

determining a transform for coregistering the image with a second image based on the boundary and based on a boundary obtained for the second image.

17. A method of enhancing image resolution, comprising identifying, according to the method of claim 15, at least one boundary for each of a plurality of images, and determining, for each boundary, a transform for coregistering each boundary with a selected reference boundary, and determining a combined image based on the plurality of images and the transforms.

18. The method of claim 17 in which determining the combined image determining the value of pixels in the combined image based on samples from at least one of the images, and the transform associated with that image.

19. A method of calibrating a signal acquisition device comprising a signal acquisition instrument, the method comprising:

obtaining an estimate of the transfer function of the instrument;

selecting a set of frequencies that are evenly spaced in the frequency domain;

and storing, in a memory of the device, coefficients based on samples of the transfer function at the selected frequencies.

20. The method of claim 19, in which the frequencies are selected to reduce the variance of the magnitude of the instrument function at the frequencies, whilst increasing the spacing between the frequencies.

21. The method of claim 19 or 20 in which the signal acquisition device is adapted for acquiring finite rate innovation, FRI, signals, and the frequencies are selected based on the performance of a maximum likelihood estimator of innovation parameters of that class of signal.

22. The method of any of claims 19 to 22 wherein the frequencies are selected to reduce a Cramer-Rao bound determined based on those frequencies.

23. The method of any of claims 19 to 22 wherein the signal acquisition device is configured to determine a frequency domain representation of an acquired signal based on the coefficients and signal domain versions of the transfer function.

24. A digital signal processing apparatus comprising:

a data obtainer configured to obtain a series of samples, of a signal, that corresponds to a signal distorted by an instrument function, and

a data scaler, configured to scale the series of samples, using coefficients wherein the coefficients are selected to approximate a complex exponential function from a set of basis functions, defined by the instrument function shifted by a number of samples;

a data provider for providing a frequency domain representation, of the signal, based on the scaled samples;

wherein the instrument function is based on a characterisation of the instrument function in the frequency domain at frequencies associated with the complex exponential function.

25. The digital signal processing apparatus of claim 24 configured to perform the method of any of claims 1 to 16.

26. The digital signal processing apparatus of claim 24 or 25 configured according to the method of any of claims 19 to 23.

27. A computer program product, or computer readable medium comprising program instructions operable to program a processor to perform the method of any of claims 1 to 18.

28. A computer readable medium configured according to the method of any of claims 19 to 23.

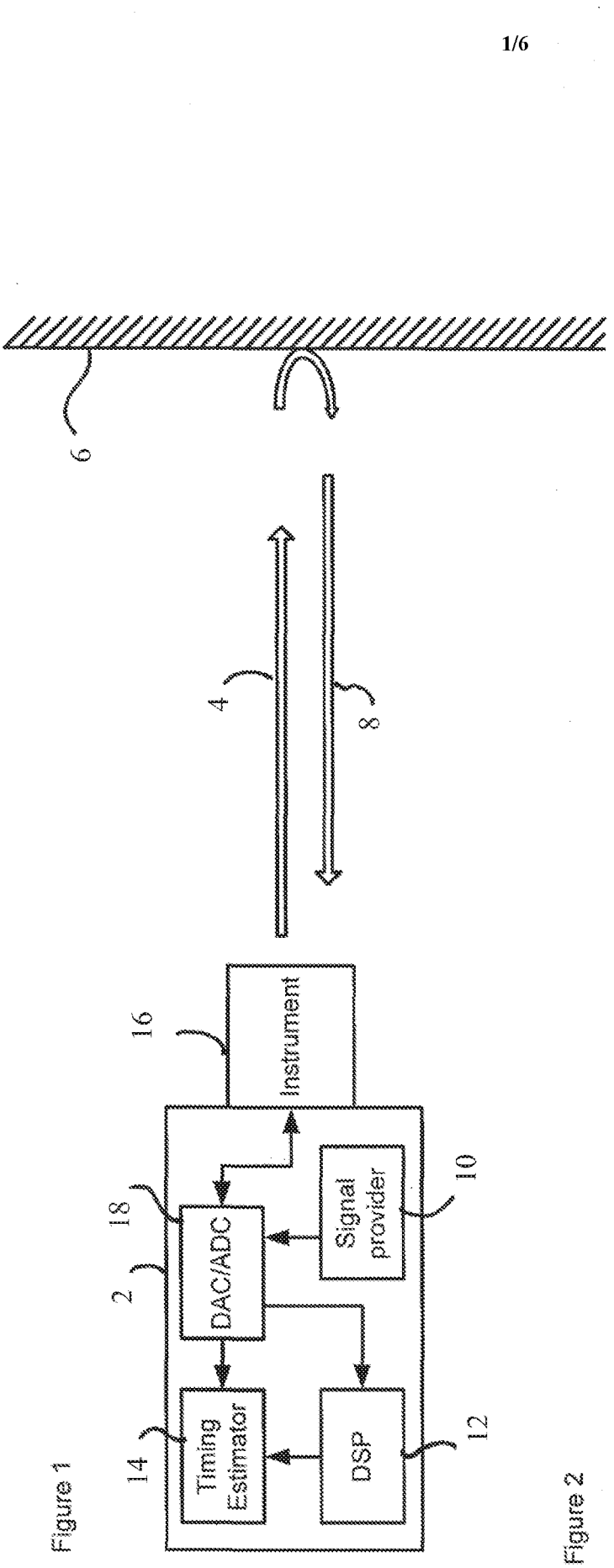


Figure 1

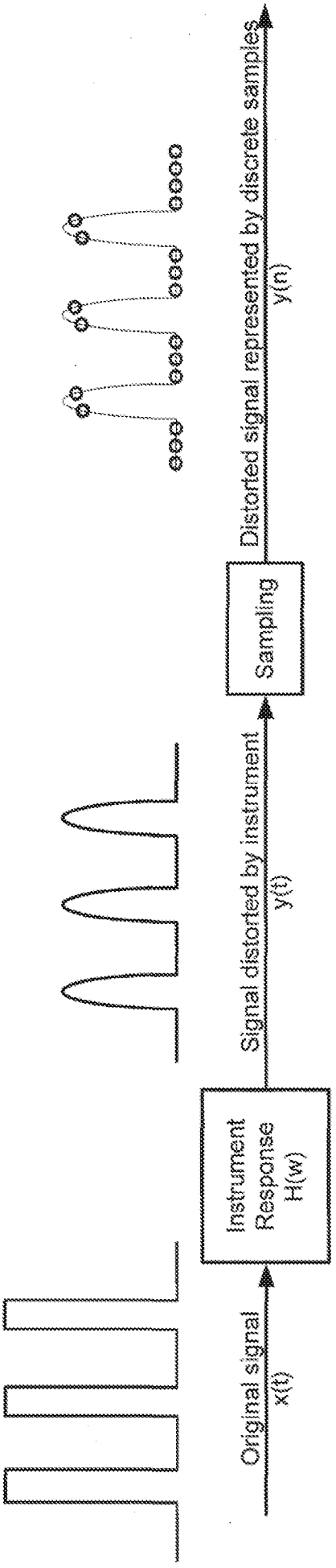


Figure 2

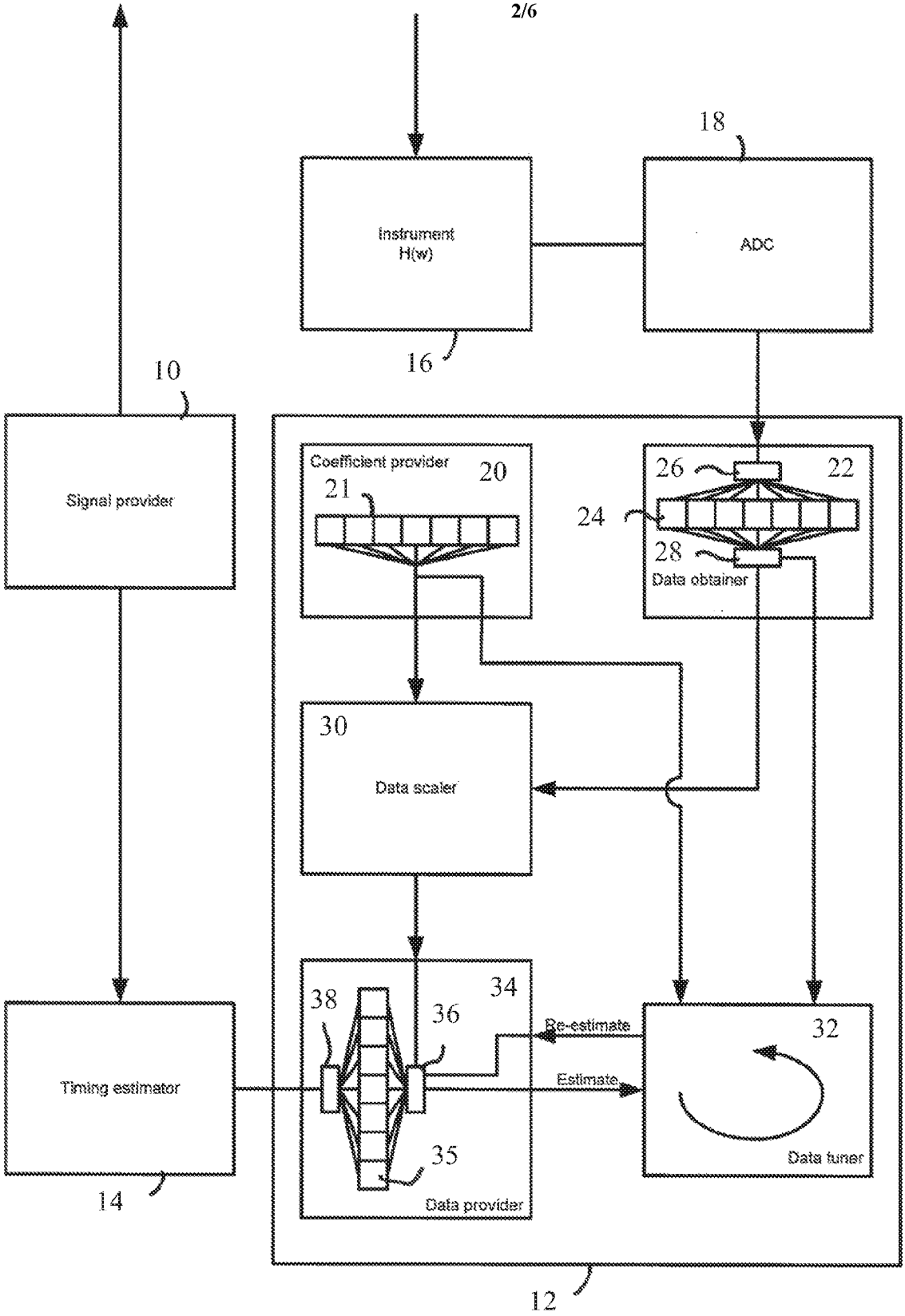

18
Instrument
H(w)
ADC
16
10
Signal provider
Coefficient provider
20
21
26
22
24
28
Data obtainer
30
Data scaler
34
38
36
Re-estimate
32
Timing estimator
Estimate
35
Data provider
Data tuner
14
12

Figure 3.

Figure 4.

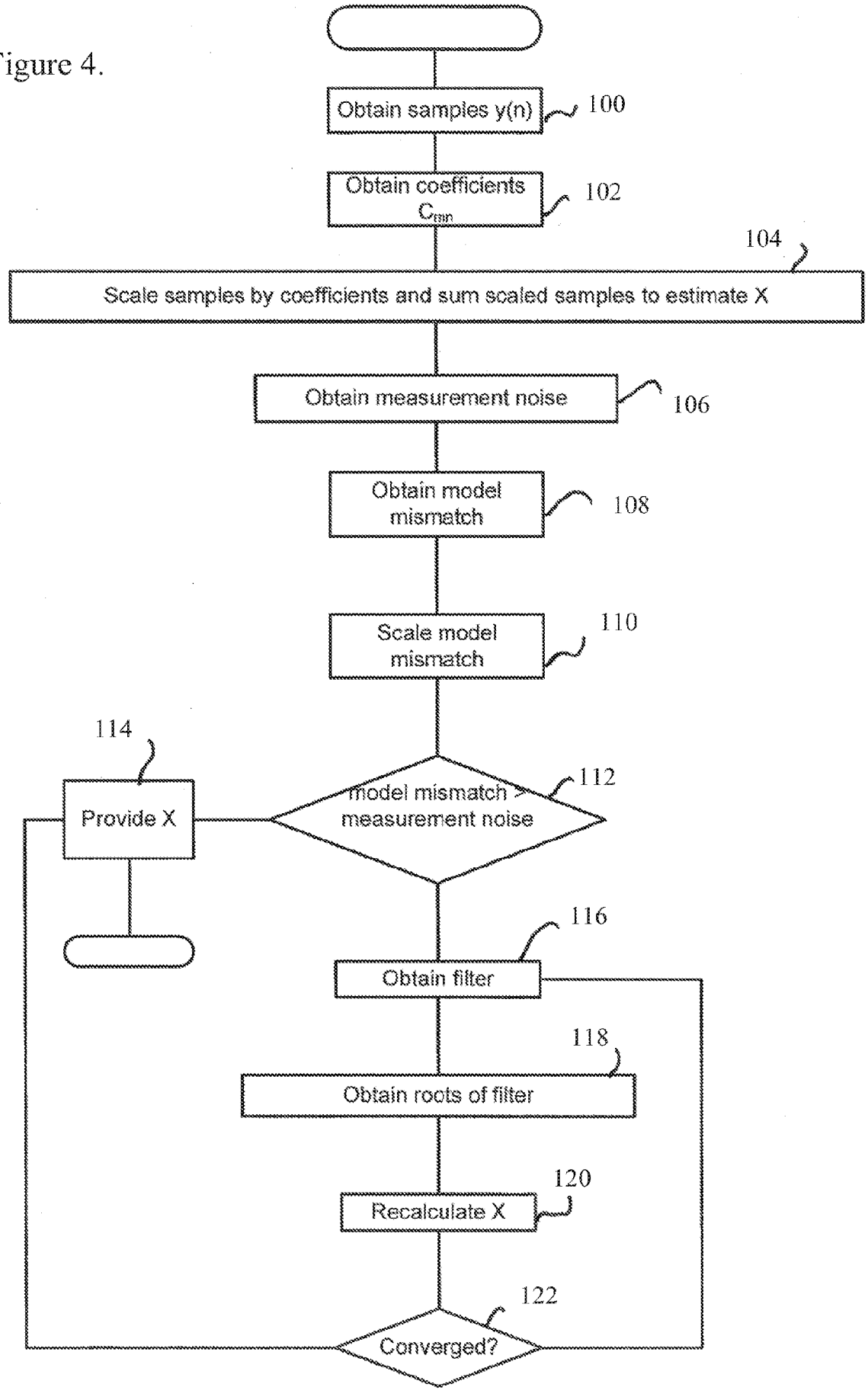

Obtain samples y(n)
100
Obtain coefficients C_{min}
102
104
Scale samples by coefficients and sum scaled samples to estimate X
Obtain measurement noise
106
Obtain model mismatch
108
Scale model mismatch
110
114
Provide X
model mismatch > measurement noise
112
116
Obtain filter
118
Obtain roots of filter
120
Recalculate X
122
Converged?

Figure 5

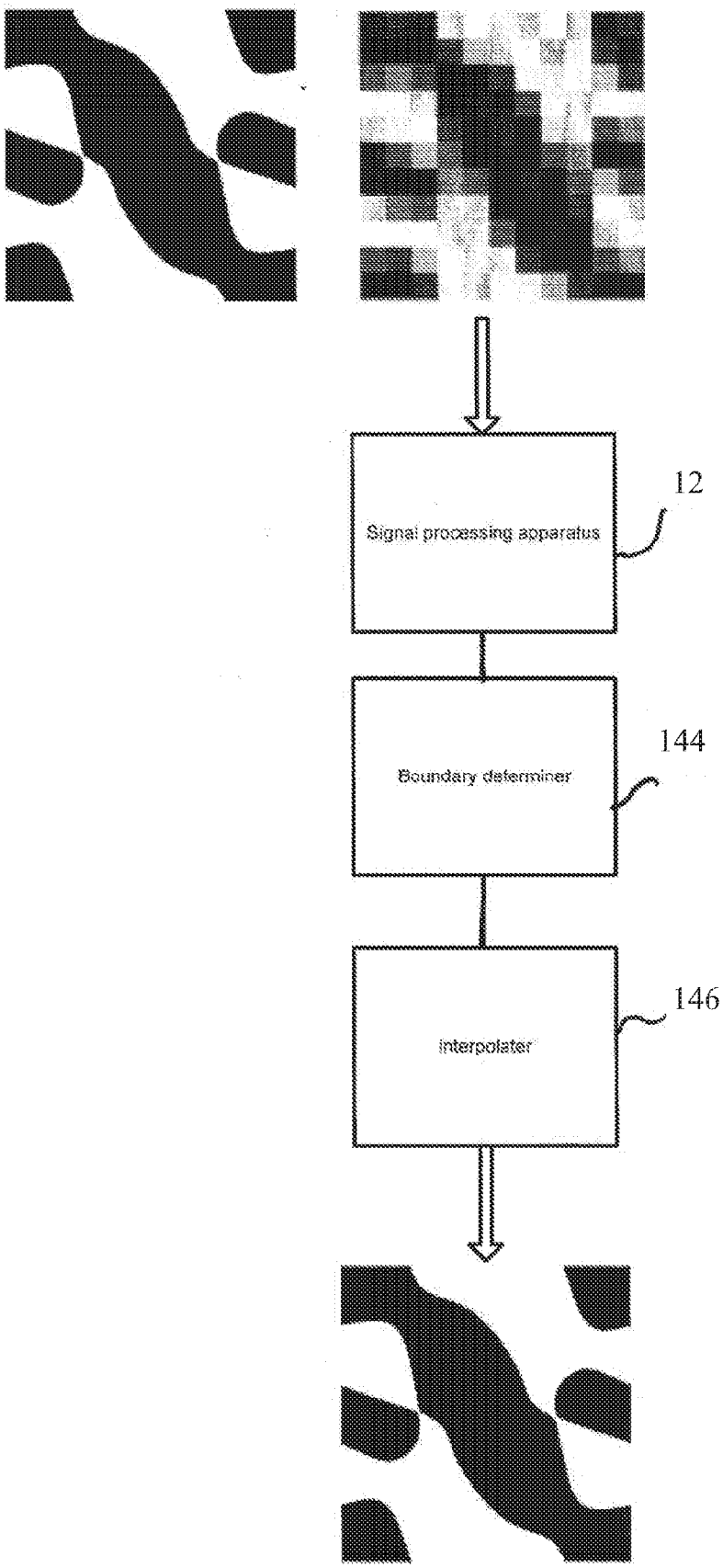

Signal processing apparatus
12
Boundary determiner
144
interpolater
146

Figure 6

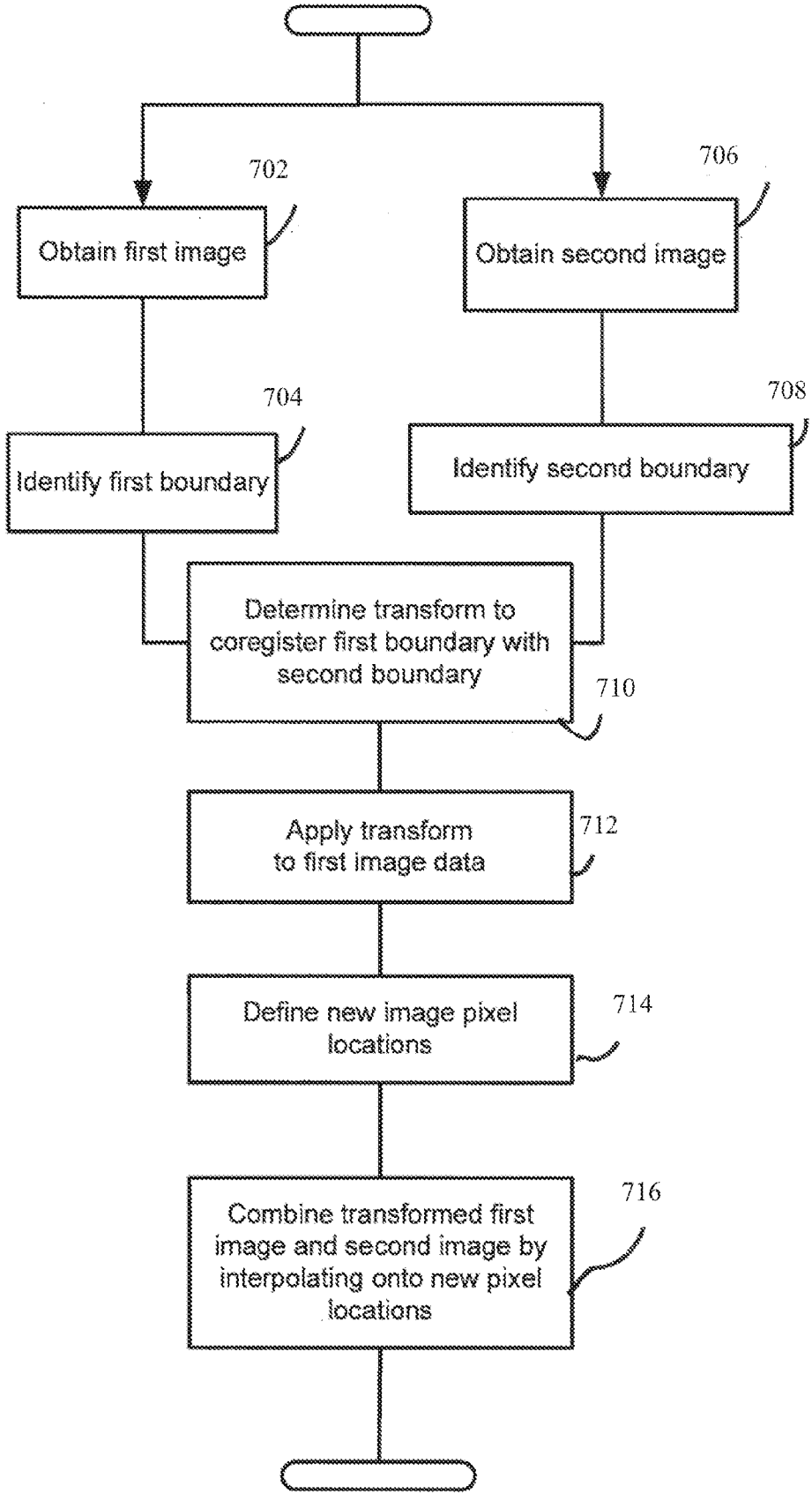
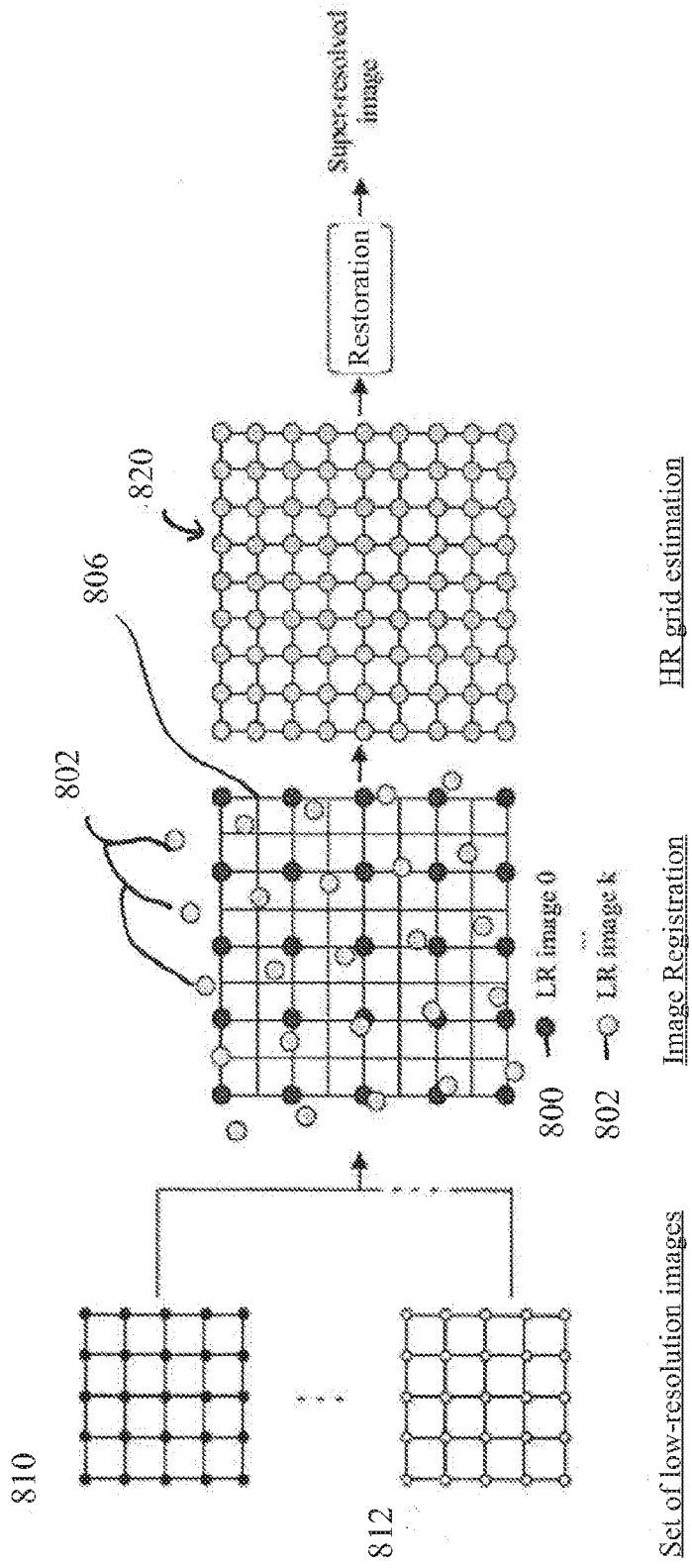

Obtain first image
702
Obtain second image
706
Identify first boundary
704
Identify second boundary
708
Determine transform to coregister first boundary with second boundary
710
Apply transform to first image data
712
Define new image pixel locations
714
Combine transformed first image and second image by interpolating onto new pixel locations
716

Figure 7.


Super-resolved image
Restoration
820
806
802
810
812
LR image 0
LR image k
800
802
Set of low-resolution images
Image Registration
HR grid estimation