A NEW APPROACH TO SPECTRAL ESTIMATION FROM IRREGULAR SAMPLING

David BONACCI and Bernard LACAZE

TESA Laboratory
Signal Processing Unit
14-16 Port Saint-Etienne - 31000 TOULOUSE - FRANCE

ABSTRACT

This article addresses the problem of signal reconstruction, spectral estimation and linear filtering directly from irregularly-spaced samples of a continuous signal (or autocorrelation function in the case of random signals) when signal spectrum is assumed to be bounded. The number $2L$ of samples is assumed to be large enough so that the variation of the spectrum on intervals of width π/L is small. Reconstruction formulas are based on PNS (Periodic Nonuniform Sampling) schemes. They allow for reconstruction schemes not requiring regular resampling and suppress two stages in classical computations. The presented method can also be easily generalized to spectra in symmetric frequency bands (bandpass signals).

***Index Terms*—** Periodic Nonuniform Sampling, Sampling theory, Signal reconstruction, Nonuniform filtering, Analytic signal

1. INTRODUCTION

There is a lot of literature about irregular sampling. Recent one include papers by Selva [1] and Eldar [2] for multiband signals or Aldroubi (compressed sampling [3], data smoothing and interpolation by cubic splines [4]) or Oppenheim [5] for sinc reconstruction of bandpass signal using digital filtering.

In these papers, only approximate reconstruction is achieved but in this paper, exact reconstruction formulas are derived. In the context of an irregular sampling where the sampling instants are not regularly spaced but assumed to be known without error, conditions ensuring that a band-limited signal can be reconstructed exactly from infinite irregular sampling exist [6]. They are often approximate as the Lomb-Scargle [7] or polynomials or spline interpolation methods (for example [8], [9], [10]) but exact formulas are difficult to find in the literature. Main theoretical reason is the difficulty to decide if a Lagrange interpolation formula is or is not convergent. Indeed, in the irregular case, there is generally no closed-form expression of the interpolation canonical product and it is then difficult to know if it is close to zero or not when the time is close to the infinite (actually, it must not be close to 0). When using such approximate formulas, the residual error degrades the computation of the Fourier transform providing the spectrum. Moreover, the analyzed signal spectrum must often be baseband.

Such a context is met in Fourier transform spectroscopy. A real example is the SIFTI (Static Infrared Fourier Transform Interferometer) instrument, supporting a mission for atmospheric pollution sounding from space, by providing high spectral resolution and high Signal to Noise Ratio spectra of the atmosphere. Another example is the TROPOspheric Monitoring Instrument (TROPOMI): it is a spaceborne nadir viewing spectrometer with bands in the ultraviolet, the visible, the near infrared and the shortwave infrared. In these instruments, a part of a light beam is delayed by a moving mirror and interacts with the direct beam. The interferogram depends on the mirrors positions and its Fourier transform is the beam spectrum. These devices can be used in infrared where emissions of many atmospheric components, like CO, CH_4, O_3, NO_2 are found. In embarked experiments it is difficult to use continuous moving mirrors due to mechanical constraints, shocks and accelerations. To mitigate these problems and obtain more compact devices, in static interferometers, the moving mirror is replaced by stepped fixed mirrors [11], [12]. Even if the mirrors are manufactured with high accuracy (for instance, in devices like TROPOMI of SIFTI, mirrors have to be manufactured with accuracy below one micrometer), the resulting sampling in the interferometer is not perfectly uniform. Nevertheless the distance between the mirrors faces can be measured with high accuracy.

In other papers, we already gave reconstruction formulas from irregular sampling for band-limited signals [13], [14]. A FFT (Fast Fourier Transform) was necessary after this interpolation to obtain the power spectrum. In this new paper, formulas giving at the same time good estimations of the continuous signal value $f(t)$ at any time instant t and of its Fourier transform $F(\omega)$ without any other computations are presented and proved. Results are presented in the context of baseband signals at the Nyquist rate, but it is possible also to deal with baseband signals at the Landau rate [15], [6]. Moreover, results can be generalized to the case when $F(\omega)$ is not baseband without need of spectral translation (suppressing an intermediary operation generally used in demodulation schemes).

Section 2 presents interpolation formulas and simulations are carried out in Section 3, demonstrating the accuracy of reconstruction compared to the Lomb-Scarle algorithm and an interesting application of numerical filtering directly from irregular samples (example of Hilbert filtering). Section 4 concludes the paper and proofs of formulas are developed in appendix (Section 5).

2. RECONSTRUCTION FORMULAS

A deterministic square-integrable signal $f(t)$ and its Fourier transform $F(\omega)$ (with bounded support, assumed to be 2π-length) are related by

$$f(t)=\int_{-\pi}^{\pi}F(\omega)\,e^{i\omega t}d\omega \tag{1}$$

Assume that $f(t)$ is sampled at $2L$ time instants a_m, $m\in\mathbb{K}$ with

$$\mathbb{K}=\{-L,-L+1,..,-1,1,..,L\}. \tag{2}$$

The only condition required on the sampling instants a_m is the Landau condition (mean distance between the a_m linked to the spectral support width). In traditional approaches, signal is resampled using for instance

$$f(a_m)=\sum_{n=-\infty}^{\infty}\frac{\sin\pi(a_m-n)}{\pi(a_m-n)}f(n). \tag{3}$$

Right-hand side of the formula is truncated to $2L$ terms (number of known samples) and, assuming an invertible system, inverted to compute an approximation of the

$$f(n),n=-L,-L+1,..,-1,0,1,..,L-1. \tag{4}$$

An estimation of $f(t)$ is then obtained, leading to an estimation of $F(\omega)$ using FFT (Fast Fourier Transform). Actually, it is possible to build a sampling scheme that includes the observations at $a_m, m\in\mathbb{K}$ and to complete it with a PNS$2L$ sampling scheme (Periodic Nonuniform Sampling of order $2L$) involving adding samples outside the observation interval of size $2L$. To get the estimation of $f(t)$, the a_m are assumed to be distinct (modulo 1) and $|a_m|<L,\forall m$. Let us define the sequence t_{mn} by

$$t_{mn}=a_m+2nL, m\in\mathbb{K}, n\in\mathbb{Z} \tag{5}$$

where $\mathbb{Z}$ is the set of integers. The sequence t_{mn} has a density 1 because $|\mathbb{K}|=2L$ (cardinal of the set). It is then possible to perform an errorless reconstruction of $F(\omega)$ from the the knowledge of the $f(t_{mn})$ and in some other circumstances when $F(\omega)$ is not baseband-type [15], [13]. Intervals $\boldsymbol{D}_k$ are defined by

$$\boldsymbol{D}_k=\begin{cases}\left[(k-1)\frac{\pi}{L},k\frac{\pi}{L}\right], k=1,2,..,L\\ \left[k\frac{\pi}{L},(k+1)\frac{\pi}{L}\right], k=-1,-2,..,-L\end{cases} \tag{6}$$

and $f_k(t)$ is the result of the bandpass filtering of $f(t)$ on $\boldsymbol{D}_k$ i.e.:

$$f_k(t)=\int_{\boldsymbol{D}_k}F(\omega)\,e^{i\omega t}d\omega. \tag{7}$$

Obviously $f(t)$ is the sum of the components $f_k(t)$

$$f(t)=\sum_{k\in\mathbb{K}}f_k(t). \tag{8}$$

Functions $g_k(t)$ are defined by

$$g_k(t)=\begin{cases}f_k(t)\,e^{-\left(k-\frac{1}{2}\right)\frac{i\pi t}{L}}, k>0\\ f_k(t)\,e^{-\left(k+\frac{1}{2}\right)\frac{i\pi t}{L}}, k<0\end{cases} \tag{9}$$

and $V_m(t), m\in\mathbb{K}$ by ($\mathrm{sinc}(t)=(\sin t)/t$):

$$\begin{aligned}V_m(t)=&\sum_{n\in\mathbb{Z}}f(a_m+2nL)\,...\\ &...(-1)^n\,\mathrm{sinc}\left(\tfrac{\pi}{2L}(t-a_m)-n\pi\right).\end{aligned} \tag{10}$$

The sinc function converges slowly to 0 (slow decay) but function $f(t)$ is assumed to be also converging to 0 (finite-energy function) so that, when L is large enough, the added unknown samples have only a negligible influence and the truncation of the series in (10) has only a negligible effect. In the appendix (Section 5) following formulas are proved:

$$\begin{aligned}V_m(t)=&\sum_{k=1}^{L}g_k(t)\,e^{\left(k-\frac{1}{2}\right)\frac{i\pi a_m}{L}}\\ &+\sum_{k=-L}^{-1}g_k(t)\,e^{\left(k+\frac{1}{2}\right)\frac{i\pi a_m}{L}}.\end{aligned} \tag{11}$$

Provided that (11) is an invertible system:

$$g_k(t)=\frac{\det\Delta_k(t)}{\det\Delta} \tag{12}$$

where the $2L$x$2L$ Δ matrix is defined by

$$\Delta=[\delta_{mk}], \delta_{mk}=\begin{cases}e^{\left(k-\frac{1}{2}\right)\frac{i\pi a_m}{L}}, k>0\\ e^{\left(k+\frac{1}{2}\right)\frac{i\pi a_m}{L}}, k<0\end{cases} \tag{13}$$

and where $\Delta_k(t)$ is the matrix deduced from Δ replacing the column k by the $V_m(t)$ defined in (10).

In the framework of finite-energy functions, $f(t)$ converges to 0 and this is also the case for the sinc function in (10): the terms for $n\neq 0$ in (10) can then be neglected when L is large enough. $V_m(t)$ then becomes observable as it involves only the $f(a_m)$. The $a_m+2nL, n\neq 0$ are not sampling times and are introduced only to get exact reconstruction formulas and the knowledge of the $f(a_m+2nL), n\neq 0$ is not required.

If $F(\omega)$ has slow variations on intervals $\boldsymbol{D}_k$ (width π/L), the definition of $f_k(t)$ by (7) leads to

$$f_k(0)=g_k(0)=\int_{\boldsymbol{D}_k}F(\omega)\,d\omega\cong\frac{\pi}{L}F\left(\frac{k\pi}{L}\right). \tag{14}$$

Sampling formula (12) then gives at the same time estimations of $f(t)$ and $F(\omega)$ with no additional computation. Obviously, the accuracy depends on the variations of $F(\omega)$ and on the value of L. In previous papers [13], [14] this problem of irregular sampling has also been addressed by adding a periodic (or not) sequence of unobserved samples disappearing in computations and the mathematical proof of a good convergence has been given. But, in the present work, the decisive advantage comes from formula (14) allowing an easy and direct computation of $F(\omega)$.

3. SIMULATIONS

3.1. Signal reconstruction and spectral estimation

In order to validate the proposed method, a simple example is first built. The following signal $f(t)$ (15) and its Fourier transform $F(\omega)$ (16) is considered (signal is chosen even here as a real autocorrelation function but this is not a necessary condition):

$$\begin{aligned} f(t) = \tfrac{8}{16-t^2}\left(1+\cos\tfrac{\pi t}{4}\right) + \\ \tfrac{16}{64-t^2}\left(\cos\tfrac{\pi t}{4}+\cos\tfrac{3\pi t}{8}\right) + \\ \tfrac{16}{t^2}\left(2\cos\tfrac{5\pi t}{8}-\cos\tfrac{\pi t}{2}-\cos\tfrac{3\pi t}{4}\right). \end{aligned} \qquad (15)$$

$$F(\omega) = \begin{cases} \sin 4\omega, \omega \in \left[0, \frac{\pi}{4}\right] \\ \sin 8\omega, \omega \in \left[\frac{\pi}{4}, \frac{3\pi}{8}\right] \\ 8\left(\omega - \frac{\pi}{2}\right), \omega \in \left[\frac{\pi}{2}, \frac{5\pi}{8}\right] \\ -8\left(\omega - \frac{3\pi}{4}\right), \omega \in \left[\frac{5\pi}{8}, \frac{3\pi}{4}\right] \end{cases} \qquad (16)$$

and $F(\omega) = 0$ for other positive values of ω.
The sampling instants a_m are chosen uniformly distributed on $\left[m - \frac{1}{10}, m + \frac{1}{10}\right]$, so that the minimum distance between them is different from 0 (actually equal to 0.8). This is a realistic sufficient condition for system (11) to be invertible and this is also a realistic assumption for stepped interferometer mirrors' positions. Note that even if the instants are chosen randomly for simulations, this is only a realization of the uniform distribution. The instants are then sorted and assumed to be known: we remain in the framework of deterministic sampling. The total number $2L$ of samples is equal to 60. Figures 1 and 2 illustrate this example with noisy samples (centered gaussian white noise) with 30 dB SNR (Signal to Noise Ratio).

Figure 1 shows that even with a small number of non-uniform observed samples ($L = 30$ and sampling instants covering the interval $[-L, L]$) the time function $f(t)$ is reconstructed with good accuracy on $[-L, L]$ with (11), as well as its Fourier transform $F(\omega)$ (Figure 2) through (14). When L increases, the accuracy of the interpolation also increases, allowing at the same time to derive spectrograms with fast variations (which happens when characteristic frequencies are searched). With a larger number of observed samples ($L = 60$), the time function $f(t)$ is reconstructed on a larger temporal interval and Figure 3 shows that its Fourier transform $F(\omega)$ is improved (especially in fast spectral variation areas). Indeed, only the mean value of $F(\omega)$ on intervals of length $\frac{\pi}{L}$ is estimated through (11), which explains why increasing the number of observed samples improves the Fourier transform.

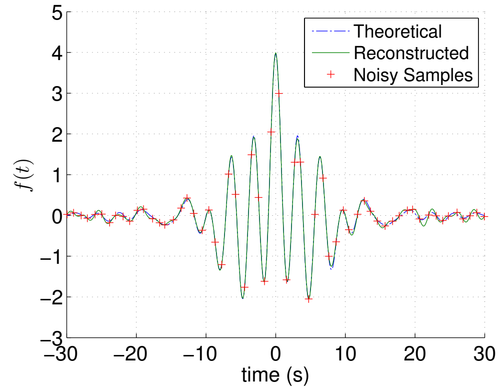

Fig. 1. Reconstructed and theoretical time function $f(t)$ for $L = 30$.

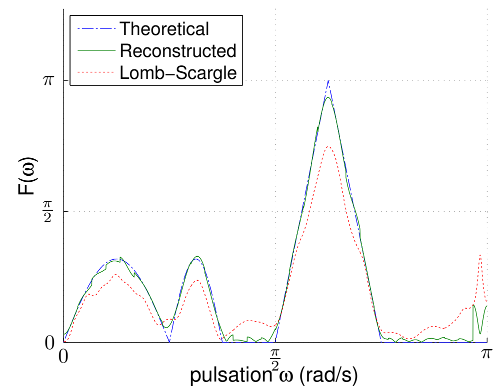

Fig. 2. Reconstructed and theoretical $F(\omega)$ function for $L = 30$ and comparison to the Lomb-Scargle spectral estimator with same data.

A comparison to the Lomb-Scargle spectral estimator is also displayed for $L = 30$ and $L = 60$. The unnormalized version of the Lomb-Scargle spectral estimator is used so that the comparison is fair (the 2 estimators are able to estimate the power of a white noise). Only the positive part of the spectrum is displayed in Figures 2 and 3 (assuming real signal). In these two situations, with the same irregular data, our estimate has much less ripple and amplitudes are also estimated with better accuracy. For this kind of spectral shape (continuous spectrum), the Lomb-Scargle spectral estimator underestimates the spectrum amplitude for $L = 30$ and it is worse for $L = 60$. This is not the case for our spectral estimate as perfect reconstruction is achieved for signal with decaying amplitude. Note that the conclusion would be inverted (in favor of Lomb-Scargle) for signals with peak spectra like sinusoids.

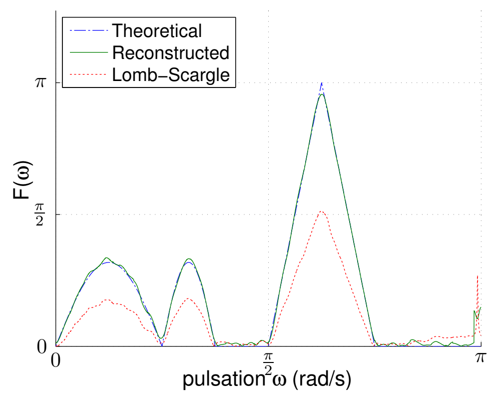

Fig. 3. Reconstructed and theoretical $F(\omega)$ function for $L = 60$ and comparison to the Lomb-Scargle spectral estimator with same data.

3.2. Digital filtering directly from irregular samples

The method also allows to process linear filtering directly from irregular samples. For instance, let $g(t)$ be defined by

$$g(t) = 2\int_0^\pi e^{i\omega t} F(\omega)\, d\omega. \tag{17}$$

$g(t)$ is the analytic signal of the signal $f(t)$ defined above.

$$\begin{aligned} g(t) = &\tfrac{16}{16-t^2} e^{i\pi t/8} \cos\tfrac{\pi t}{8} \\ &+\tfrac{32}{64-t^2} e^{5i\pi t/16} \cos\tfrac{\pi t}{16} \\ &+\tfrac{32}{t^2} e^{5i\pi t/8} \sin^2\tfrac{\pi t}{16}. \end{aligned} \tag{18}$$

Clearly, from (7) and (8) it follows:

$$g(t) = 2\sum_{k\in\mathbf{K}, k>0} f_k(t). \tag{19}$$

Here, digital filtering is applied using only terms $k > 0$ in the final reconstruction step (8) (filter frequency response is null for $f \in D_k, k < 0$ and 1 for $f \in D_k, k \geqslant 0$). Any linear filter response expressed as a combination of $\mathbf{D}_k$ intervals can be obtained. The bounds of the intervals have to be integer multiples of $\frac{\pi}{L}$, but this tends to become a weaker condition as L increases. The reconstructed analytic signal $g(t)$ is displayed in Figures 4 and 5. Same noisy samples (with 30 dB SNR) as in Figure 1 were used ($L = 30$). For better viewing, Figure 5 displays the phase of $g(t)e^{-i\frac{\pi}{2}t}$ rather than directly the phase of the analytic signal $g(t)$ which contains a strong linear variation versus ω. These figures illustrate the ability to process linear filtering with the presented reconstruction formulas directly from the non uniform samples, without need for intermediate reconstruction of the signal, which could be of great practical interest for a lot of applications using nonuniform samples (reduced computational load).

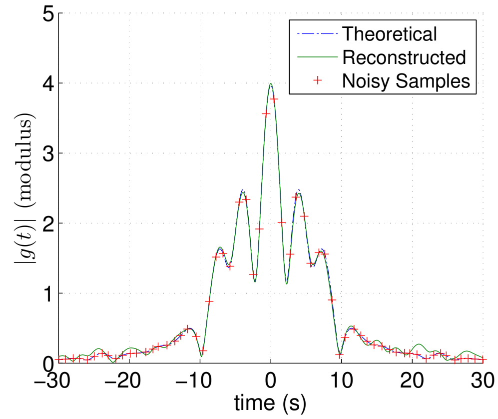

Fig. 4. Analytic signal for $L = 30$. Modulus of $g(t)$.

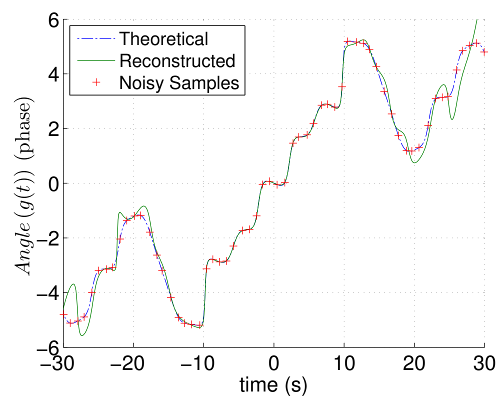

Fig. 5. Analytic signal for $L = 30$. Phase of $g(t)e^{-i\frac{\pi}{2}t}$.

4. CONCLUSION

In this paper we have given and proved formulas allowing to obtain at the same time good estimations of reconstructed signal and corresponding power spectra when the sampling instants are not regularly spaced but assumed to be known. Formulas obtained by using PNS2L (Periodic Nonuniform Sampling) scheme are shown to be valid for L large enough and their accuracy is compared to the Lomb-Scargle, showing better results. Increasing the value of L allows taking into account larger variations of spectra. Moreover, an original solution for performing numerical filtering directly from non uniform samples is derived from formulas and illustrated. Finally, it is easily possible to generalize given formulas to the case when spectral support is divided into two symmetric intervals (this is the case for example when using SIFTI, TROPOMI or others real equipments) using the general concept of Landau rate rather than Nyquist rate to highlight the real width of the spectrum.

5. APPENDIX

In this appendix, formula (11) is proved. First, let us consider function $e^{i\omega t}$ for $\omega \in \left]\alpha, \alpha + \pi/L\right[$, t being fixed. The Fourier series development on this interval is

$$\begin{gathered} h_t(\omega) = \sum_{n\in\mathbb{Z}} c_n(t)\, e^{2inL\omega} \\ c_n(t) = e^{i\alpha(t-2nL)+\frac{i\pi t}{2L}-in\pi}\operatorname{sinc}\left(\frac{\pi t}{2L} - n\pi\right). \end{gathered} \tag{20}$$

$h_t(\omega)$ is equal to $e^{i\omega t}$ on $\left]\alpha, \alpha + \pi/L\right[$ but not outside this interval except when $\pi t/L$ is an even integer. For instance:

$$\begin{gathered} h_t(\omega) = e^{i\left(\omega-\frac{\pi}{L}\right)t} \\ \omega \in \left]\alpha + \pi/L, \alpha + 2\pi/L\right[. \end{gathered} \tag{21}$$

As (20) is true $\forall t$:

$$\begin{gathered} e^{i\omega t} = e^{i\omega\tau} \sum_{n\in\mathbb{Z}} c_n(t-\tau)\, e^{2inL\omega} \\ \omega \in \left]\alpha, \alpha + \pi/L\right[. \end{gathered} \tag{22}$$

This variant of (20) stands for any real τ. The sum at the right is uniformly bounded, allowing changes in the order of summations. Therefore, when brought in (7) with $\tau = a_m$ and for the correct value of α ($(k-1)\frac{\pi}{L}$ or $k\frac{\pi}{L}$ following the sign of k):

$$\begin{gathered} f_k(t) = e^{\frac{i\pi}{L}\left(k-\frac{1}{2}\right)(t-a_m)} .. \\ .. \sum_{n\in\mathbb{Z}} f_k(a_m + 2nL)(-1)^n \operatorname{sinc}\left(\pi\left(\frac{t-a_m}{2L} - n\right)\right) \end{gathered} \tag{23}$$

for $k > 0$ and, for $k < 0$:

$$\begin{gathered} f_k(t) = e^{\frac{i\pi}{L}\left(k+\frac{1}{2}\right)(t-a_m)} .. \\ .. \sum_{n\in\mathbb{Z}} f_k(a_m + 2nL)(-1)^n \operatorname{sinc}\left(\pi\left(\frac{t-a_m}{2L} - n\right)\right). \end{gathered} \tag{24}$$

The key of the problem is then to exhibit the observed data:

$$f(a_m) = \sum_{k\in\mathbb{K}} f_k(a_m). \tag{25}$$

Multiplying (23), (24) by $e^{\frac{i\pi}{L}\left(k-\frac{1}{2}\right)(t-a_m)}$ and $e^{\frac{i\pi}{L}\left(k+\frac{1}{2}\right)(t-a_m)}$ leads to (10) and (11). Considering that the value of f becomes negligible above the largest a_m and below the smallest for L large enough, the unobserved terms $f(a_m + 2nL)$, $n \neq 0$ can be suppressed as $|a_m| < L$. In a realistic set of measurements, this can be justified considering that measurement errors are likely to be larger than the true values for $t < -L$ and $t > L$. The $2L$x$2L$ system of equations (11) is then entirely determined, allowing to derive estimations of $f(t)$ and $F(\omega)$ using (12).

REFERENCES

[1] J. Selva, "Regularized Sampling of Multiband Signals," *Signal Processing, IEEE Transactions on*, vol. 58, no. 11, pp. 5624–5638, Nov 2010.

[2] M. Mishali and Y.C. Eldar, "From Theory to Practice: Sub-Nyquist Sampling of Sparse Wideband Analog Signals," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 4, no. 2, pp. 375–391, April 2010.

[3] A. Aldroubi and Haichao Wang, "An Adaptive and Information Theoretic Method for Compressed Sampling," *CAMSAP, 2009 3rd IEEE International Workshop on*, pp. 193–196, Dec 2009.

[4] A. Aldroubi and H. Feichtinger, "Complete iterative reconstruction algorithms for irregularly sampled data in spline-like spaces," *ICASSP., IEEE International Conference on*, vol. 3, pp. 1857–1860, Apr 1997.

[5] S. Maymon and A.V. Oppenheim, "Sinc Interpolation of Nonuniform Samples," *IEEE Transactions on Signal Processing*, vol. 59, pp. 4745–4758, 2011.

[6] F. A. Marvasti, *Nonuniform sampling, theory and practice*, Kluwer Ac. Press, 2002.

[7] J. Scargle, "Statistical Aspects of Spectral Analysis on Unevenly Spaced Data," *The Astronomical Journal*, vol. 263, pp. 835–853, 1982.

[8] S. Signall Y. R. Sun, "Algorithms for nonuniform bandpass sampling in radio receiver," *Proc. ISACS*, vol. 1, pp. 25–28, 2003.

[9] P. P. Vaidyanathan, "Generalizations of the sampling theorem : seven decades after Nyquist," *IEEE Trans on Circuits and Systems*, vol. 9, pp. 1094–1109, 2001.

[10] M. Unser P. Thevenaz, T. Blu, "Interpolation revisited," *IEEE Trans on Med. Imag.*, vol. 7, pp. 739–758, 2000.

[11] P. F. Levelt, G. H. J. van den Oord, M. Dobber, H. Eskes, M. van Weele, P. Veefkind, R. van Oss, I. Aben, R.T. Jongma, J. Landgraf, J. de Vries, and H. Visser, "TROPOMI and TROPI: UV/VIS/NIR/SWIR instruments," *Proc. SPIE*, vol. 6296, 2006.

[12] P. J. Hébert, E. Cansot, C. Pierangelo, C. Buil, F. Bernard, J. Loesel, T. Trémas, L. Perrin, E. Courau, C. Casteras, I. Maussang, and D. Simeoni, "Instrumental concept and preliminary performances of SIFTI : static infrared Fourier interferometer," *Proc. ICSO*, 2008.

[13] B. Lacaze, "Reconstruction formula for irregular sampling," *Sampling Th. in Sign. Image Proc.*, vol. 1, pp. 33–43, 2005.

[14] B. Lacaze, "The Ghost Sampling Sequence Method," *Sampling Th. in Sign. Image Proc.*, vol. 1, pp. 13–21, 2009.

[15] J. R. Higgins, "Some Gap Sampling Series for Multiband Signals," *Signal Processing*, pp. 303–309, 1987.