# Pseudo-inverse reconstruction of bandlimited signals from nonuniform generalized samples with orthogonal kernels

Nguyen T. Thao, Member, IEEE, Dominik Rzepka, Member, IEEE, and Marek Miśkowicz, Senior Member, IEEE

Abstract-Contrary to the traditional pursuit of research on nonuniform sampling of bandlimited signals, the objective of the present paper is not to find sampling conditions that permit perfect reconstruction, but to perform the best possible signal recovery from any given set of nonuniform samples, whether it is finite as in practice, or infinite to achieve the possibility of unique reconstruction in  $L^2(\mathbb{R})$ . This leads us to consider the pseudoinverse of the whole sampling map as a linear operator of Hilbert spaces. We propose in this paper an iterative algorithm that systematically performs this pseudo-inversion under the following conditions: (i) the input lies in some closed space A (such as a space of bandlimited functions); (ii) the samples are formed by inner product of the input with given kernel functions; (iii) these functions are orthogonal at least in a Hilbert space  $\mathcal H$  that contains A. This situation turns out to appear in certain time encoders that are part of the increasingly important area of eventbased sampling. As a result of pseudo-inversion, we systematically achieve perfect reconstruction whenever the samples uniquely characterize the input, we obtain minimal-norm estimates when the sampling is insufficient, and the reconstruction errors are controlled in the case of noisy sampling. The algorithm consists in alternating two projections according to the general method of projections onto convex sets (POCS) and can be implemented by iterating time-varying discrete-time filtering. We finally show that our signal and sampling assumptions appear in a nontrivial manner in other existing problems of data acquisition. This includes multi-channel time encoding where  $\mathcal{H}$  is of the type  $L^2(\mathbb{R})^M$ , and traditional point sampling with the adoption of a Sobolev space  $\mathcal{H}$ . This thus uncovers the unexpected possibility of sampling pseudo-inversion in existing applications, while indicating the potential of our formalism to generate future sampling schemes with systematic pseudo-inverse reconstructions.

*Index Terms*—bandlimited signals, nonuniform sampling, generalized sampling, time encoding, bandlimited interpolation, pseudo-inverse, Kaczmarz method, POCS, frame algorithm, Sobolev spaces.

# I. INTRODUCTION

The reconstruction of bandlimited signals from nonuniform samples is a challenging topic that has been studied since the 50's [1], [2], [3], [4], [5], [6], [7], although its practical development has remained somewhat limited. But this subject is currently attracting new attention with the increasing trend of event-based sampling in data acquisition [8], [9], [10], [11], [12], [13], [14]. This approach to sampling has grown in an effort to simplify the complexity of the analog sampling circuits, lower their power consumption and simultaneously

ment and Electronics, AGH University of Kraków, Kraków, Poland, emails: drzepka@agh.edu.pl, miskow@agh.edu.pl increase their precision. This is made possible in particular by the replacement of amplitude encoding by time encoding, which takes advantage of the higher precision of solid-state circuits in time measurement [15]. Time encoding has however induced the use of nonuniform samples that were not commonly studied in the past literature. A well-known example is the time encoder introduced by Lazar and Tóth in [16] which acquires the integrals of an input signal over successive nonuniform intervals, based an asynchronous Sigma-Delta modulator (ASDM).

In their most general form, nonuniform samples of an input signal x in some Hilbert space  $(\mathcal{A}, \langle \cdot, \cdot \rangle)$  are scalars of the form

$$\mathbf{s}_k = \langle x, g_k \rangle, \qquad k \in \mathsf{Z}$$
 (1)

where Z is some set of indices and  $(g_k)_{k\in\mathbb{Z}}$  is a given family of functions in  $\mathcal{A}$ , which we call the sampling kernel functions [17]. Reconstructing x from  $\mathbf{s} = (\mathbf{s}_k)_{k\in\mathbb{Z}}$  can be presented as solving the linear equation  $Su = \mathbf{s}$  where S is the linear operator

$$S: \ u \in \mathcal{A} \ \mapsto \ \left( \langle u, g_k \rangle \right)_{k \in \mathbb{Z}}.$$
 (2)

When A is a set of bandlimited functions and Z is finite, a basic approach is to reduce S to a matrix and estimate x by  $S^+$ s, where  $S^+$  is the matrix pseudo-inverse of S [18]. This problem reduction to finite-dimensional linear algebra is however incompatible with the theoretically infinite time support of bandlimited signals. Even in practice where signals are always time limited, their time support are typically seen in signal processing as virtually infinite compared to the windows of operation. It was proposed to split the resolution of Su = s by finite blocks of signals, in which exact algebraic inversions are performed [19]. With bandlimited signals, block truncations however creates analytically uncontrolled distortions at the block boundaries and necessitates ad-hoc empirical methods of compensations. This also departs from traditional signal processing which on the contrary preserves input signals in their entirety, while performing finite-complexity approximations of the ideal operations by sliding-window processing (such as FIR filtering). In traditional LTI processing, errors of approximation are well controlled by Fourier analysis. In the context of nonuniform sampling, Fourier analysis is no longer applicable by loss of time invariance. However, linear operations are still rigorously analyzed by functional analysis. With such an approach, the first rigorous numerical method of bandlimited interpolation of nonuniform point samples, called the frame algorithm, was derived by Duffin and Schaeffer [1]. The method used by Lazar and Toth in [16] for input reconstruction from integrals has a similar structure, based on an algorithm by Feichtinger and Gröchenig [5]. It was shown

N. T. Thao is with the Department of Electrical Engineering, The City College of New York, CUNY, New York, USA, email: tnguyen@ccny.cuny.edu. D. Rzepka and M. Miśkowicz are with the Department of Measure-

This research was supported by National Center of Science of Poland under the grant DEC- 2018/31/B/ST7/03874.

in [20], [21] all these types of algorithm can be be converted into iterative time-varying linear filtering, with potential of sliding-window implementations.

Until recently, these iterative methods have been limited to cases where S is exactly invertible, based on some sufficient (and not always necessary) sampling conditions. The goal of this paper is to find iterative algorithms of similar type that converge more generally to the pseudo-inverse  $S^{\dagger}$  of S in the theoretical sense of linear operators in Hilbert spaces. The motivation is to achieve perfect reconstruction whenever S is effectively invertible, while proposing an optimal reconstruction whenever S is not invertible (in case of insufficient sampling) or the sampling sequence  $(s_k)_{k \in \mathbb{Z}}$  is corrupted by noise. In this way, a goal is to incorporate within a single method an algorithm whose behavior is consistent with the theoretical results of sampling from harmonic analysis while giving optimal solutions in real practical situations of sampling.

Solving this question in the most general case of nonorthogonal sampling kernels  $(g_k)_{k\in Z}$  is difficult. At the other extreme, reconstructing x from the samples of (1) is trivial when the family  $(g_k)_{k\in Z}$  is an orthogonal basis of the input space  $\mathcal{A}$ . This is the case of Shannon's sampling theorem. In this paper, we consider an intermediate situation where achieving the pseudo-inverse  $S^{\dagger}$  becomes possible by successive filtering, and which is described by the following condition:

$$(g_k)_{k \in \mathsf{Z}} \text{ is orthogonal in a Hilbert space } (\mathcal{H}, \langle \cdot, \cdot \rangle)$$
  
and x is in a closed subspace  $\mathcal{A} \subset \mathcal{H}.$  (3)

Although seemingly ideal, this condition turns out to be realized in the time-encoding system of Lazar and Toth [16], in the case where  $\mathcal{H} = L^2(\mathbb{R})$  and  $\mathcal{A}$  is a subspace of bandlimited signals. This was noticed and utilized in [20] to construct an algorithm achieving  $S^{\dagger}$  in the specific application of [16]. The algorithm was based on a particular application of the method of projection onto convex sets (POCS) [22], [23]. This was later generalized to integrate-and-fire encoding with leakage in [21]. The purpose of the present article is to extract from [20], [21] the most general framework of pseudo-inversion of S by successive filtering under the abstract assumption of (3). In this generalization, the content of these references is revisited and reformulated to reach its most fundamental ingredients and obtain a self-sufficient theory that is independent of the applications. Our formalism contains theoretical results as well as efficient techniques of practical implementations. A goal is to propose a new framework of nonuniform sampling schemes for which a pseudo-inverse input reconstruction method by successive filtering is readily available. While this objective is meant to influence the design of future sampling schemes, we also show in this paper an immediate impact of the proposed theory by applying it to two existing sampling/reconstruction schemes: one in multi-channel time encoding [24], [11] and one in the original case of nonuniform point sampling [4]. In these two sampling applications, the authors studied their proposed reconstruction algorithms under specific assumptions of unique reconstruction. In both cases, we show that their own algorithms coincide with our generic POCS algorithm. In this process, we end up pointing previously unknown properties that their algorithm possess, including their ability to achieve perfect reconstruction even in situations where proofs of unique reconstruction are not available, the characteristics of their limit when the sampling is insufficient, and their behavior towards sampling noise. But on the theoretical side, an important role of these two applications is to show that the abstract condition (3) can be found in unexpected situations, using some non-standard technique of signal analysis. In the first example, condition (3) is extracted after some non-trivial reduction of the complex algorithm of [24], [11]. Beyond pointing out the unknown properties of their method, our high-level formalization allows a concise reformulation of it together with an organized presentation of its implementation at the level of discrete-time filters. For a complementary demonstration, the difficulty of the second example is not in the complexity of the sampling system, but in the non-trivial signal theoretic approach that is required. For condition (3) to be realized in this case, the traditional Hilbert space  $L^2(\mathbb{R})$ needs to be replaced by the homogeneous Sobolev space  $H^1(\mathbb{R})$  [25], [26]. This only allows us to prove convergence up to a constant component but does lead for the first time to a result of pseudo-inversion of point sampling by successive filtering.

The paper is organized as follows. We start in Section II by reviewing the basic knowledge that samples of the form (1)bring about an input signal x in a general Hilbert space  $\mathcal{H}$ and without any assumption on the sampling kernels  $(g_k)_{k \in \mathbb{Z}}$ . We give the basic principle of the POCS algorithm for finding estimates that are consistent with the samples of (1). In Section III, we show how condition (3) leads a specific configuration of POCS algorithm that is more efficient and that will be later shown to have special connections with the pseudo-inversion of S. For that purpose, we devote Section IV to reviewing the notion of pseudo-inverse for linear operators in infinite dimension, which is not commonly used knowledge in signal processing. Section V then contains the major mathematical contribution of this article. By starting from a zero initial estimate, we prove that the POCS iteration tends to  $S^{\dagger}s$  by contraction whenever the sampling configuration theoretically allows a stable consistent reconstruction (which is systematically the case when Z is finite). When the initial estimate is a signal  $u^{(0)} \neq 0$ , we show that the POCS limit is more generally the signal of the type  $S^{\dagger}\mathbf{s} + v$  that is closest to  $u^{(0)}$  under the constraint  $v \in A$  and Sv = 0. This is of particular interest when consistent reconstruction is not unique due to insufficient sampling, and one wishes to pick a consistent estimate that is close to a signal guess of statistical or heuristic nature [9]. This reconstruction simultaneously takes care of sampling errors with an action of "noise shaping" in the case of oversampling [27]. In Section VI, we discuss some important aspects of practical implementations. We finally present in Sections VII and VIII the two mentioned examples of application.

# II. CONSISTENT RECONSTRUCTION FROM GENERALIZED SAMPLES

In this section, we review the background on the use of POCS for the reconstruction of bandlimited signals from nonuniform generalized samples. Without loss of generality, we assume that the considered bandlimited signals have Nyquist period 1 and call their space  $\mathcal{B}$ .

## A. Nonuniform generalized samples

To understand the specific contribution of POCS to signal reconstruction from samples, we first need to give the most general definition of signal sampling. Let x be a function in some Hilbert space  $\mathcal{H}$  equipped with an inner product  $\langle \cdot, \cdot \rangle$ . We call a generalized sample of  $x \in \mathcal{H}$  any scalar value of the type

$$\mathbf{s} = \langle x, g \rangle$$

where g is a known function of  $\mathcal{H}$  [28], which we call a sampling kernel function. In the most basic example of bandlimited signal sampling, the Hilbert space is  $\mathcal{H} = \mathcal{B}$ equipped with the canonical inner product  $\langle \cdot, \cdot \rangle_2$  of  $L^2(\mathbb{R})$ . The point sample of  $x \in \mathcal{B}$  at instant  $\tau$  yields the form

$$x(\tau) = \langle x, g \rangle$$
 with  $g(t) := \operatorname{sinc}(t - \tau)$  (4)

where

$$\operatorname{sinc}(t) := \sin(\pi t) / (\pi t).$$
(5)

The point sample of the *n*th derivative of  $x \in \mathcal{B}$  at  $\tau$  is also viewed as a generalized sample of x as it can be shown using integration by parts that

$$\frac{\mathrm{d}^n x}{\mathrm{d}t^n}(\tau) = \langle x, g \rangle \qquad \text{with} \qquad g(t) := (-1)^n \frac{\mathrm{d}^n \mathrm{sinc}}{\mathrm{d}t^n} (t - \tau).$$

We also trivially have generalized samples by integration as illustrated by the following example,

$$\int_{a}^{b} x(t) \, \mathrm{d}t = \langle x, g \rangle \qquad \text{with} \qquad g(t) := \mathbf{1}_{[a,b]}(t)$$

where  $1_{[a,b]}(t)$  is the indicator function of [a,b].

When  $\mathcal{H}$  is a space of functions on  $\mathbb{R}$ , we say that a set of samples  $(s_k)_{k\in\mathbb{Z}} = (\langle x, g_k \rangle)_{k\in\mathbb{Z}}$  is uniform when  $\mathbb{Z} = \mathbb{Z}$ and there exists a function  $g \in \mathcal{H}$  and a period T such that  $g_k(t) = g(t - kT)$  for all  $k \in \mathbb{Z}$ . Shannon's sampling theorem applies to the particular case where  $g(t) = \operatorname{sinc}(t)$  and T = 1in our bandlimited setting. Nonuniform sampling then consists in all of the other cases. This could be when  $g_k(t) = g(t - t_k)$ where the instants  $t_k$  are not regularly spaced, or when  $(g_k)_{k\in\mathbb{Z}}$ are just not the shifted versions of a single function.

## B. Set theoretic view of sampling

The output of each sample  $s_k = \langle x, g_k \rangle$  gives us the deterministic knowledge that x belongs to the hyperplane

$$\mathcal{H}_k := \{ v \in \mathcal{H} : \langle v, g_k \rangle = \mathbf{s}_k \}.$$
(6)

Thus, a full sampling sequence  $\mathbf{s} = (\mathbf{s}_k)_{k \in \mathbb{Z}}$  tells us that x belongs to the intersection

$$\mathcal{H}_{\mathbf{s}} := \bigcap_{k \in \mathsf{Z}} \mathcal{H}_k. \tag{7}$$

Perfect reconstruction is possible if and only if  $\mathcal{H}_{s}$  is limited to a singleton. When this is not the case, there is no deterministic knowledge from  $(s_{k})_{k\in\mathbb{Z}}$  to distinguish x from any other element of  $\mathcal{H}_{s}$ . We call the functions of  $\mathcal{H}_{s}$  the *consistent estimates* of x. The next proposition gives some general knowledge on the algebraic structure of  $\mathcal{H}_{s}$ . Proposition 2.1: Let  $\mathbf{s} = (\mathbf{s}_k)_{k \in \mathbb{Z}}$  be any sequence. If  $\mathcal{H}_{\mathbf{s}} \neq \emptyset$ , then, for any  $u \in \mathcal{H}_{\mathbf{s}}$ ,

$$\mathcal{H}_{\mathbf{s}} = u + \mathcal{G}^{\perp} \tag{8}$$

where  $\mathcal{G}$  is the closed linear span of  $(g_k)_{k \in \mathbb{Z}}$ 

$$\mathcal{G} := \overline{\operatorname{span}}(g_k)_{k \in \mathbf{Z}}$$

and  $\mathcal{G}^{\perp}$  is its orthogonal complement in  $\mathcal{H}$ .

*Proof:* By assumption,  $\mathbf{s}_k = \langle u, g_k \rangle$  for each  $k \in \mathbb{Z}$ . Hence,  $v \in \mathcal{H}_k \Leftrightarrow \langle v, g_k \rangle = \mathbf{s}_k = \langle u, g_k \rangle \Leftrightarrow \langle v - u, g_k \rangle = 0$ . Thus,  $v \in \mathcal{H}_{\mathbf{s}}$  if and only if  $v - u \in \mathcal{G}^{\perp}$ . In other words,  $\mathcal{H}_{\mathbf{s}} = u + \mathcal{G}^{\perp}$ .

Since a linear subspace always contains the 0 vector, then  $\mathcal{H}_s$  is a singleton if and only if  $\mathcal{G}^{\perp} = \{0\}$ . This is equivalent to  $\mathcal{G} = \mathcal{H}$ , which means that the linear span of  $(g_k)_{k \in \mathbb{Z}}$  is dense in the whole space  $\mathcal{H}$ . Thus, the unique reconstruct of x solely depends on the sampling kernels  $(g_k)_{k \in \mathbb{Z}}$ , and not on the input x itself.

## C. Systematic estimation

As mentioned in the introduction, the objective of this paper is not to study the question of unique reconstruction of an input x from its samples. The goal is to perform the best possible approximation of x from given samples s, whatever they are. While the term of "best possible" would require some definition, it is at least intuitive that any reconstruction of xthat is not consistent with its samples cannot be optimal. This idea is in fact rationally supported by the following property.

*Proposition 2.2:* Let C be a closed affine subspace of H that contains x. Then,

$$\forall u \in \mathcal{H}, \qquad \|P_{\mathcal{C}}u - x\| \le \|u - x\| \tag{9}$$

where  $P_{\mathcal{C}}$  designates the orthogonal projection of  $\mathcal{H}$  onto  $\mathcal{C}$ and  $\|\cdot\|$  is the norm induced by the inner product  $\langle\cdot,\cdot\rangle$  of  $\mathcal{H}$ . Moreover, the inequality is *strict* whenever  $u \notin \mathcal{C}$ .

*Proof:* For  $u \in \mathcal{H}$  given,  $P_{\mathbb{C}}u$  is the unique element  $v \in \mathbb{C}$  such that  $u-v \perp v-w$  for all  $w \in \mathbb{C}$ . By the Pythagorean theorem,  $||u-w||^2 = ||u-v||^2 + ||v-w||^2 \ge ||v-w||^2$  with a strict inequality when  $u \neq v$ , which happens whenever  $u \notin \mathbb{C}$ . The result of (9) is the particular case w = x.

Then, a systematic reconstruction procedure is to pick the consistent estimate

$$\hat{x} := P_{\mathcal{H}_{\mathbf{s}}} u^{(0)} \tag{10}$$

where  $u^{(0)}$  is some initial estimate proposed by the user. While  $u^{(0)}$  can be obtained by heuristic or statistical means,  $\hat{x}$  is an estimate of x that is guaranteed to be better than  $u^{(0)}$  and cannot be further improved deterministically. It is also the consistent estimate that is closest to  $u^{(0)}$  with respect to  $\|\cdot\|$ . The strength of this procedure is that whenever uniqueness of reconstruction is effective, whether one is able to prove it or not,  $\hat{x}$  is guaranteed to be the perfect reconstruction of

x. In the case of non-unique reconstruction, the type (10) of reconstruction was first considered by Yen in [2] for the estimation of a bandlimited signal of  $L^2(\mathbb{R})$  from a finite number of point samples, with the specific choice of  $u^{(0)} = 0$ . In this case,  $\hat{x}$  is the consistent estimate that is closest to 0, and hence of minimum norm. It can be easily shown from the knowledge of (8) that this  $\hat{x}$  must be in  $\mathcal{G}$ . We note here that [17] studied the case where consistent reconstruction is constrained to a linear subspace that may be different from  $\mathcal{G}$ .

The next proposition gives a more analytical description of  $P_{\mathcal{H}_s}$ .

Proposition 2.3: If  $x \in \mathcal{H}_{\mathbf{s}}$ , then

$$\forall u \in \mathcal{H}, \qquad P_{\mathcal{H}_{\mathbf{s}}}u = u + P_{\mathcal{G}}(x - u) \tag{11}$$

where  $\mathcal{G}$  was defined in Proposition 2.1.

*Proof:* Let  $v := u - P_{\mathcal{G}}(x-u)$ . While  $v \in u + \mathcal{G}$ ,  $v = x + (u-x) - P_{\mathcal{G}}(u-x) \in x + \mathcal{G}^{\perp}$ . So v is the orthogonal projection of u onto  $x + \mathcal{G}^{\perp} = \mathcal{H}_{s}$ .

Note that while x is unknown,  $P_{\mathcal{G}}(x-u)$  can still be theoretically evaluated as it can be shown to only depend on u and  $(s_k)_{k \in \mathbb{Z}}$ .

# D. POCS algorithm

In the most general case, there is unfortunately no closed form expression for  $P_{\mathcal{G}}u$ . When Z is a small finite set,  $P_{\mathcal{G}}$ can only be obtained by inversion (or pseudo-inversion) of a matrix whose ill conditioning in nonuniform sampling rapidly grows with the size of Z [29]. The basic technique used in this paper to find  $P_{\mathcal{H}_s}u$  is the POCS method [22], [23] whose basic function is to retrieve an element in the intersection of N closed convex sets. As closed affine spaces, the sets  $\mathcal{H}_k$ of (6) are a particular case of closed convex sets. Assuming a finite set  $Z = \{1, \dots, N\}$ , the basic version of the POCS method applied to (7) consists in the iteration

$$u^{(n+1)} := P_{\mathcal{H}_N} \cdots P_{\mathcal{H}_1} u^{(n)}, \qquad n \ge 0.$$
(12)

Since x belongs to every set  $\mathcal{H}_k$ , we conclude from Proposition 2.2 that the estimate error  $||u^{(n)} - x||$  strictly decreases with n as long as  $u^{(n)}$  has not reached  $\mathcal{H}_s$ . It is actually known [22, §III.B] that not only  $u^{(n)}$  always eventually converges to an element  $u^{(\infty)}$  in  $\mathcal{H}_s$  in norm, but the limit is more precisely

$$u^{(\infty)} = P_{\mathcal{H}_{\mathbf{s}}} u^{(0)}.$$

As an intersection of closed affine subspaces, note that  $\mathcal{H}_{s}$  is itself a closed affine subspace. Thus,  $u^{(\infty)}$  is the consistent estimate of x that is closest to the initial iterate  $u^{(0)}$ .

#### E. Kaczmarz algorithm

The Kaczmarz algorithm is the specific name that is given to the iteration of (12) when the sets  $\mathcal{H}_k$  are simply hyperplanes, as is the case of (6). Given the explicit description of (6),  $P_{\mathcal{H}_k}$ yields the following simple expression [23, p.403]

$$\forall u \in \mathcal{H}, \qquad P_{\mathcal{H}_k} u = u + \frac{\mathbf{s}_k - \langle u, g_k \rangle}{\|g_k\|^2} g_k$$
(13)

(this can also be obtained from (11) in the case where  $\mathcal{G}$  is the linear span of  $g_k$  alone). The Kaczmarz algorithm was first used in sampling by Yeh and Stark in [30] for solving numerically the problem of Yen in [2]. Since then however, this method has not known much development in nonuniform sampling of bandlimited signals, primarily due to its slow convergence.

A randomized version of the Kaczmarz algorithm was later introduced in [31] for potential statistical accelerations of the convergence. A standard version of it consists in performing a random permutation of the N projections of (12) at each iteration. But this variant is till not suitable for real-time causal signal processing.

# III. ORTHOGONAL SAMPLING KERNELS

The slow convergence of the Kaczmarz method is primarily due to the non-orthogonality of the sampling kernels  $(g_k)_{k \in \mathbb{Z}}$ . Another fundamental shortcoming of this method is the impossibility to reduce it to an iteration of the type  $u^{(n+1)} = Ru^{(n)}$ with some fixed transformation R when Z is infinite. While the number of samples is always finite in practice, an infinite index set Z is always needed when dealing with the theoretical question of perfect reconstruction of a bandlimited signal in  $L^2(\mathbb{R})$ . Under the assumption of condition (3) where A is the input space, we show in this section that there is a way to reduce the POCS algorithm to alternating two projections only. This even allows Z to be infinite. After giving some practical examples of this situation in data acquisition, we present this special POCS algorithm and its properties in absence of noise.

#### A. Practical examples of orthogonal sampling kernels

The cases of orthogonal kernels that have appeared in the literature in nonuniform sampling until now [20], [21] take place in the space  $\mathcal{H} := L^2(\mathbb{R})$  with  $\mathcal{A} := \mathcal{B}$  as the closed subspace of inputs. Consider a sampling scheme where the samples of  $x \in \mathcal{B}$  are of the type

$$\mathbf{s}_k := \int_{t_{k-1}}^{t_k} f_k(t) \, x(t) \, \mathrm{d}t, \qquad k \in \mathsf{Z}$$
(14)

for some increasing sequence of time instants  $(t_k)_{k\in\mathbb{Z}}$  and some family of functions  $(f_k)_{k\in\mathbb{Z}}$  in  $L^2(\mathbb{R})$ . This takes the form of (1) with

$$g_k(t) := f_k(t) \mathbf{1}_{[t_{k-1}, t_k]}(t), \qquad k \in \mathsf{Z}.$$

Clearly, the functions  $(g_k)_{k\in\mathbb{Z}}$  are orthogonal since their supports do not overlap (up to discrete points). The samples of (14) are those of leaky integrate-and-fire encoding (LIF) when the functions  $(f_k)_{k\in\mathbb{Z}}$  are of the type

$$f_k(t) := e^{-\alpha(t-t_{k-1})}$$

for some constant  $\alpha > 0$  [21]. In the case  $\alpha = 0$ , the samples are of the simple form

$$\mathbf{s}_k := \int_{t_{k-1}}^{t_k} x(t) \,\mathrm{d}t, \qquad k \in \mathsf{Z}$$
(15)

which are also the type of samples that one extracts from integrate-and-fire [8], [10], [11], [12], [13], [32] or from an asynchronous Sigma-Delta modulator (ASDM) [16], [20].

#### B. Consistent reconstruction

Under the new assumptions, x belongs to both A and  $\mathcal{H}_s$ when  $s = (s_k)_{k \in Z}$  is obtained from (1). So the set of consistent estimates is

$$\mathcal{A}_{\mathbf{s}} := \mathcal{A} \cap \mathcal{H}_{\mathbf{s}}. \tag{16}$$

We are going to see that  $\mathcal{A}_{s}$  has a similar structure to  $\mathcal{H}_{s}$  within  $\mathcal{A}.$  Let

$$\widetilde{g}_k := P_{\mathcal{A}} g_k, \qquad k \in \mathsf{Z}. \tag{17}$$

Since  $g_k - \tilde{g}_k$  is orthogonal to  $\mathcal{A}$ , then  $\langle u, g_k - \tilde{g}_k \rangle = 0$  for any  $u \in \mathcal{A}$ . Thus,

$$\forall u \in \mathcal{A}, \ \forall v \in \mathcal{H}, \qquad \langle u, g_k \rangle = \langle u, \widetilde{g}_k \rangle.$$
 (18)

It then follows from (7) and (6) that

$$\mathcal{A}_{\mathbf{s}} := \bigcap_{k \in \mathsf{Z}} \mathcal{A}_k \quad \text{where} \quad \mathcal{A}_k := \mathcal{A} \cap \mathcal{H}_k \tag{19}$$
$$= \left\{ v \in \mathcal{A} : \langle v, \widetilde{g}_k \rangle = \mathbf{s}_k \right\}.$$

With a proof similar to that of Proposition 2.1, we have the following result.

Proposition 3.1: Let  $\mathbf{s} = (\mathbf{s}_k)_{k \in \mathsf{Z}}$  be any sequence. If  $\mathcal{A}_{\mathbf{s}} \neq \emptyset$ , then, for any  $u \in \mathcal{A}_{\mathbf{s}}$ ,

$$\mathcal{A}_{\mathbf{s}} = u + \mathcal{F}^{\perp} \tag{20}$$

where  $\mathfrak{F}$  is the closed linear span of  $(\widetilde{g}_k)_{k \in \mathsf{Z}}$ 

$$\mathcal{F} := \overline{\operatorname{span}}(\widetilde{g}_k)_{k \in \mathsf{Z}} \tag{21}$$

and  $\mathcal{F}^{\perp}$  is the orthogonal complement of  $\mathcal{F}$  in  $\mathcal{A}$ .

Again, the reconstruction of x is unique if and only if  $\mathcal{F}^{\perp} = \{0\}$ . As this is equivalent to  $\mathcal{F} = \mathcal{A}$ , this means that the linear span of  $(\tilde{g}_k)_{k \in \mathbb{Z}}$  is dense in the whole space  $\mathcal{A}$ .

# C. POCS for orthogonal sampling kernels

The main point of the previous section was to prepare the detailed notation of the new sampling framework and reformulate the theoretical condition of unique reconstruction. Now, the new contribution is that the POCS method can be directly applied to the 2-set decomposition of (16). This is because  $P_{\mathcal{H}_s}$  is now accessible via its expression of (11). Indeed, since  $(g_k)_{k\in\mathbb{Z}}$  is now an *orthogonal basis* of  $\mathcal{G}$ , we have the explicit expansion of  $P_{\mathcal{G}}$ ,

$$\forall u \in \mathcal{H}, \quad P_{\mathcal{G}}u = \sum_{k \in \mathsf{Z}} \left\langle u, \frac{g_k}{\|g_k\|} \right\rangle \frac{g_k}{\|g_k\|} = \sum_{k \in \mathsf{Z}} \frac{\langle u, g_k \rangle}{\|g_k\|^2} g_k.$$
(22)

It then results from (11) and (1) that

$$\forall u \in \mathcal{H}, \qquad P_{\mathcal{H}_{\mathbf{s}}}u = u + \sum_{k \in \mathbf{Z}} \frac{\mathbf{s}_k - \langle u, g_k \rangle}{\|g_k\|^2} g_k. \tag{23}$$

We can thus implement the POCS iteration

$$u^{(n+1)} = P_{\mathcal{A}} P_{\mathcal{H}_{\mathbf{s}}} u^{(n)}, \qquad n \ge 0.$$
 (24)

As a consequence of Section II-D, we know that  $||u^{(n)} - x||$ strictly decreases as long as  $u^{(n)} \notin A_s$ , and

$$u^{(\infty)} = P_{\mathcal{A}_{s}} u^{(0)}.$$
 (25)

But the new contribution is double: while the POCS iteration of (24) is reduced to two projections only, it simultaneously allows Z to be infinite! This is crucial for achieving perfect reconstruction in infinite-dimensional spaces such as  $L^2(\mathbb{R})$ .



Fig. 1. Illustration of  $P_{\mathbb{C}}u$  and  $P_{\mathbb{C}}^{\lambda}u$  for  $\lambda \in (1,2)$ .

## D. Relaxed projections

The estimate error reduction of Proposition 2.2 has in fact the following generalized version.

**Proposition 3.2:** [23] Let  $\mathcal{C}$  be a closed affine subspace of  $\mathcal{H}$  that contains x and

$$P_{\mathcal{C}}^{\lambda}u := u + \lambda(P_{\mathcal{C}}u - u) \tag{26}$$

for any  $\lambda \in \mathbb{R}$  and  $u \in \mathcal{H}$ . Then,

$$\forall \lambda \in [0,2], \ \forall u \in \mathcal{H}, \qquad \|P_{\mathcal{C}}^{\lambda}u - x\| \le \|u - x\|.$$
(27)

Moreover, the inequality is *strict* when  $\lambda \in (0,2)$  and  $u \notin \mathbb{C}$ .

We illustrate this result graphically in Fig. 1. The parameter  $\lambda$  is called a *relaxation coefficient*. By iterating

$$u^{(n+1)} = P_{\mathcal{A}} P_{\mathcal{H}_{\mathbf{s}}}^{\lambda_n} u^{(n)}, \qquad n \ge 0$$
(28)

where  $(\lambda_n)_{n\geq 0}$  is some sequence of coefficients in (0, 2), one draws the same conclusion as in Section III-C: assuming that  $x \in \mathcal{A}_s$ , the estimate error  $||u^{(n)} - x||$  strictly decreases with nas long as  $u^{(n)}$  has not reached  $\mathcal{A}_s$ . To the best of the authors knowledge, the actual convergence in norm of such a sequence  $u^{(n)}$  to a point of S is not explicitly discussed in the literature. But this can be at least deduced from results in [33], which leads to the following theorem.

Theorem 3.3: Assume that  $\mathcal{A}_{s} \neq \emptyset$  and  $(\lambda_{n})_{n\geq 0}$  is a sequence of coefficients in [0, 2] such that  $\sum_{n\geq 0} \lambda_{n}(2-\lambda_{n}) = +\infty$ . Then the relaxed POCS iteration of (28) starting with  $u^{(0)} \in \mathcal{A}$  yields the limit of (25).

We justify this in Appendix A. The additional relaxation freedom allows in practice to accelerate the convergence. There is no analytical result on the optimal relaxation coefficients, which in practice are typically found empirically.

### IV. SAMPLING OPERATOR AND PSEUDO-INVERSE

We have assumed until now that the sampling sequence  $\mathbf{s} = (\mathbf{s}_k)_{k \in \mathbb{Z}}$  is given exactly by (1). In practice however, samples are often corrupted by noise. The question is what happens to the POCS iteration of (24) when s is deviated by some error. The worst case is when s can no longer be the sampled version of any signal in  $\mathcal{A}$ . We will see that the POCS iteration is guaranteed to converge with any s (up to some theoretical condition on its norm) when the linear operator S of (2) that maps any  $x \in \mathcal{A}$  into the sequence  $(\langle x, g_k \rangle)_{k \in \mathbb{Z}}$  has a pseudo-inverse  $S^{\dagger}$ . In this case, the limit  $u^{(\infty)}$  will appear to be exactly  $S^{\dagger}$ s when  $u^{(0)} = 0$ . For the moment, the present

section reviews some necessary material on operator theory that will be needed in Section V to prove the above mentioned POCS behavior.

## A. Sampling operator S

In finite dimension, the above mentioned transformation S is typically formalized as a matrix. One would then naturally estimate x by applying the matrix pseudo-inverse  $S^+$  on s. Under certain conditions, this is also possible in infinite dimension, but the setting is substantially more involved. In this case, a linear transformation typically involves infinite summations whose convergence must be guaranteed with respect to some norm. With the needed tool of orthogonal projection and the objective to invert S, this requires both the domain and the destination of S to be Hilbert spaces. A rigorous way to construct our operator S is to present it as follows:

$$\begin{array}{rccc} S: \ \mathcal{A} & \to & \mathcal{D} \\ & u & \mapsto & Su := \left( \langle u, g_k \rangle \right)_{k \in \mathbb{Z}} \end{array} \tag{29}$$

where

$$\mathcal{D} := \left\{ \mathbf{c} : \left( \mathbf{c}_k / \|g_k\| \right)_{k \in \mathsf{Z}} \in \ell^2(\mathsf{Z}) \right\}$$
(30)

and  $\ell^2(Z)$  is the set of square-summable sequences indexed by Z. We call S the *sampling operator*. While A is by default seen as a Hilbert space with respect to the inner product  $\langle \cdot, \cdot \rangle$ of the ambient space  $\mathcal{H}$ ,  $\mathcal{D}$  is a Hilbert space with respect to the inner product  $\langle \cdot, \cdot \rangle_{\mathcal{D}}$  defined by

$$\forall \mathbf{c}, \mathbf{d} \in \mathcal{D}, \qquad \langle \mathbf{c}, \mathbf{d} \rangle_{\mathcal{D}} := \sum_{k \in \mathsf{Z}} \frac{c_k \, \mathrm{d}_k}{\|g_k\|^2}.$$
 (31)

Next, note that Su does belong to  $\mathcal{D}$  for all  $u \in \mathcal{A}$ . Writing  $Su = (c_k)_{k \in \mathbb{Z}}$ , this is because  $(c_k/||g_k||)_{k \in \mathbb{Z}} = (\langle u, g_k/||g_k|| \rangle)_{k \in \mathbb{Z}}$ , which belongs to  $\ell^2(\mathbb{Z})$  since  $(g_k/||g_k||)_{k \in \mathbb{Z}}$  is orthonormal. When  $\mathbb{Z}$  is finite, note also that  $\ell^2(\mathbb{Z})$  is just  $\mathbb{R}^{\mathbb{Z}}$ . Finally, when the sampling is uniform (implying that  $||g_k||$  is constant), note that  $\langle \cdot, \cdot \rangle_{\mathcal{D}}$  coincides with the canonical inner product  $\langle \cdot, \cdot \rangle_2$  of  $\ell^2(\mathbb{Z})$ , up to a scaling factor. Otherwise,  $\langle \cdot, \cdot \rangle_{\mathcal{D}}$  can be interpreted as a weighted version of  $\langle \cdot, \cdot \rangle_2$  to compensate for the non-uniformity.

The sampling operator leads to a more concise way to characterize the set of consistent estimates  $A_s$ . Indeed, it follows from (19) and (29) that

$$\mathcal{A}_{\mathbf{s}} = \left\{ v \in \mathcal{A} : Sv = \mathbf{s} \right\} = S^{-1}(\mathbf{s}). \tag{32}$$

With this presentation,

$$\mathcal{A}_{\mathbf{s}} \neq \emptyset \qquad \Leftrightarrow \qquad \mathbf{s} \in \operatorname{ran}(S)$$
(33)

where ran(S) denotes the range S. Finally, it will be useful to note that the null space null(S) of S is given by

$$\operatorname{null}(S) = \mathcal{F}^{\perp} \tag{34}$$

as a result of (21) and (18).

# B. Adjoint operator $S^*$

In finite dimension, the matrix pseudo-inverse  $S^+$  is fundamentally linked to the matrix transpose  $S^{\top}$  of S. Generalizing pseudo-inversion in infinite dimension will also be based on a generalization of matrix transpose, which is as follows: a linear operator  $S^*$  from  $\mathcal{D}$  back to  $\mathcal{A}$  is said to be adjoint to S when

$$\forall u \in \mathcal{A}, \ \forall \mathbf{c} \in \mathcal{D}, \qquad \langle u, S^* \mathbf{c} \rangle = \langle Su, \mathbf{c} \rangle_{\mathcal{D}}. \tag{35}$$

A difficulty is that this uniquely defines  $S^*$  only when S is bounded, i.e., when there exists  $\beta \ge 0$  such that

$$\forall u \in \mathcal{A}, \qquad \|Su\|_{\mathcal{D}} \le \beta \|u\| \tag{36}$$

where  $\|\cdot\|_{\mathcal{D}}$  is the norm induced in  $\mathcal{D}$  by  $\langle\cdot,\cdot\rangle_{\mathcal{D}}$ . Let us verify that this is indeed realized with the operator S of (29). As

$$\forall \mathbf{c} \in \mathcal{D}, \qquad \|\mathbf{c}\|_{\mathcal{D}}^2 = \langle \mathbf{c}, \mathbf{c} \rangle_{\mathcal{D}} = \sum_{k \in \mathbb{Z}} \frac{|\mathbf{c}_k|^2}{\|g_k\|^2}$$
(37)

from (31),

$$\|Su\|_{\mathcal{D}}^{2} = \sum_{k \in \mathsf{Z}} \frac{|\langle u, g_{k} \rangle|^{2}}{\|g_{k}\|^{2}} = \sum_{k \in \mathsf{Z}} \left| \langle u, \frac{g_{k}}{\|g_{k}\|} \rangle \right|^{2} = \|P_{\mathfrak{G}}u\|^{2} \quad (38)$$

from (22) and by orthonormality of  $(g_k/||g_k||)_{k\in\mathbb{Z}}$ . By Bessel's inequality,  $P_{\mathcal{G}}$  is non-expansive. So (36) is satisfied with  $\beta = 1$ . The next expression gives the explicit description of  $S^*$ .

*Proposition 4.1:* The adjoint  $S^*$  to the operator S of (29) is given by

$$S^*: \qquad \mathcal{D} \quad \to \quad \mathcal{A} \\ \mathbf{c} = (\mathbf{c}_k)_{k \in \mathsf{Z}} \quad \mapsto \quad S^* \mathbf{c} = \sum_{k \in \mathsf{Z}} \frac{\mathbf{c}_k}{\|g_k\|^2} \, \widetilde{g}_k$$
(39)

*Proof:* Using (39) as the definition of  $S^*$ , we have for all  $u \in \mathcal{A}$  and  $\mathbf{s} \in \mathcal{D}$ ,

$$\langle u, S^* \mathbf{c} \rangle = \left\langle u, \sum_{k \in \mathbf{Z}} \frac{\mathbf{c}_k}{\|g_k\|^2} \widetilde{g}_k \right\rangle = \sum_{k \in \mathbf{Z}} \frac{\langle u, \widetilde{g}_k \rangle \mathbf{c}_k}{\|g_k\|^2} = \langle Su, \mathbf{c} \rangle_{\mathcal{D}}$$

as a result of (18) and (31).

The well known basic properties of operator adjoint [34] are

$$\operatorname{null}(S) = \operatorname{ran}(S^*)^{\perp}$$
 and  $\operatorname{null}(S^*) = \operatorname{ran}(S)^{\perp}$ . (40)

# C. Pseudo-inverse $S^{\dagger}$

The Moore-Penrose pseudo-inverse of S is the linear operator  $S^{\dagger}$  from  $\mathcal{D}$  back to  $\mathcal{A}$  such that [35, §11]

$$\forall \mathbf{s} \in \mathcal{D}, \qquad S^{\dagger} \mathbf{s} = \underset{v \in \mathcal{M}_{\mathbf{s}}}{\operatorname{argmin}} \|v\|, \tag{41}$$

where

$$\mathcal{M}_{\mathbf{s}} := \{ v \in \mathcal{A} : \|Sv - \mathbf{s}\|_{\mathcal{D}} \text{ is minimized} \}.$$
(42)

Contrary to the finite dimensional case,  $S^{\dagger}$  does not always exist for the simple reason that  $||Sv - s||_{\mathcal{D}}$  does not always have a minimizer  $v \in \mathcal{A}$ . This minimizer systematically exists only when

$$\operatorname{ran}(S)$$
 is closed in  $\mathcal{D}$ . (43)

Assuming this condition, the following are some useful properties of  $S^{\dagger}$  that can be found in [35, §11]:

$$S^{\dagger}S = P_{\operatorname{ran}(S^*)},\tag{44a}$$

$$S S^{\dagger} = P_{\operatorname{ran}(S)}, \tag{44b}$$

$$\operatorname{ran}(S^{\dagger}) = \operatorname{ran}(S^{*}), \tag{44c}$$

$$\operatorname{null}(S^{\dagger}) = \operatorname{null}(S^{*}) \tag{44d}$$

((44d) is a consequence of Theorem 11.1.6 of [35]). By assumption of (43), note that  $ran(S^*)$  is also closed (see for example Lemma 2.5.2 of [36]). This allows the valid statement of (44a) and implies that

$$\operatorname{ran}(S^{\dagger}) = \operatorname{ran}(S^{\ast}) = \operatorname{null}(S)^{\perp} = \mathcal{F}$$
(45)

as a result of (44c), (40), (34) and the fact that  $\mathcal{F}$  is closed. Next is a result on the algebraic structure of the set  $\mathcal{M}_s$  of (42).

*Proposition 4.2:* Assuming that ran(S) is closed,

$$\forall \mathbf{s} \in \mathcal{D}, \qquad \mathcal{M}_{\mathbf{s}} = \mathcal{A}_{\mathbf{\bar{s}}} = S^{\dagger} \mathbf{s} + \mathcal{F}^{\perp}$$
(46)

where  $\bar{\mathbf{s}} := P_{\operatorname{ran}(S)}\mathbf{s}$ .

*Proof:* Let  $\mathbf{s} \in \mathcal{D}$ . According to Theorem 11.1.1 of [35],  $||Sv-\mathbf{s}||_{\mathcal{D}}$  is minimized if and only if  $Sv = \bar{\mathbf{s}}$ . This proves that  $\mathcal{M}_{\mathbf{s}} = \mathcal{A}_{\bar{\mathbf{s}}}$ . It follows from (44b) that  $S(S^{\dagger}\mathbf{s}) = P_{\operatorname{ran}(S)}\mathbf{s} = \bar{\mathbf{s}}$ . It then results from (32) that  $S^{\dagger}\mathbf{s} \in \mathcal{A}_{\bar{\mathbf{s}}}$ . The second equality of (46) is then obtained by applying (20) after replacing  $\mathbf{s}$  and u by  $\bar{\mathbf{s}}$  and  $S^{\dagger}\mathbf{s}$ , respectively.

Remark: When  $\mathbf{s} \in \operatorname{ran}(S)$ , one simply has  $\mathcal{M}_{\mathbf{s}} = \mathcal{A}_{\mathbf{s}}$  since  $\overline{\mathbf{s}} = \mathbf{s}$  in this case. Now, as a result of (33), while  $\mathcal{A}_{\mathbf{s}}$  becomes empty when  $\mathbf{s} \notin \operatorname{ran}(S)$ , note that  $\mathcal{M}_{\mathbf{s}} = \mathcal{A}_{\overline{\mathbf{s}}}$  is never empty since  $\overline{\mathbf{s}} \in \operatorname{ran}(S)$ .

# D. Stable sampling

While condition (43) may seem abstract, it is seen in this section as a necessary condition for stable reconstruction. The next proposition first gives some equivalent analytical properties to the fact that ran(S) is closed. One of them involves the *reduced minimum modulus* of S, which is defined by

$$\gamma(S) := \inf_{u \in \mathcal{F} \setminus \{0\}} \frac{\|Su\|_{\mathcal{D}}}{\|u\|} \tag{47}$$

given that  $\mathcal{F} = \operatorname{null}(S)^{\perp}$  from (34) as  $\mathcal{F}$  is closed.

Proposition 4.3: The following statements are equivalent:

(i) ran(S) is closed in  $\mathcal{D}$ .

- (ii)  $\gamma(S) > 0$ ,
- (iii)  $(f_k)_{k\in\mathbb{Z}}$  is a frame of  $\mathcal{F}$  with  $f_k := \widetilde{g}_k / ||g_k||, k \in \mathbb{Z}$ .

The definition of a frame and the proof of this result are in Appendix B. The connection of (ii) with reconstruction stability is as follows. Consider the estimation of  $x \in A$  from noisy samples  $\mathbf{s} = Sx + \mathbf{e}$  where  $\mathbf{e}$  is some error sequence. To simplify the problem, assume that  $\mathcal{F} = A$ , guaranteeing uniqueness of reconstruction, and  $\mathbf{e} \in \operatorname{ran}(S)$ . Then  $Su = \mathbf{s}$  will imply that u = x + e where e is in  $\mathcal{F}$  while satisfying  $Se = \mathbf{e}$ . As  $||e||/||\mathbf{e}||_{\mathcal{D}} = ||e||/||Se||_{\mathcal{D}}$ , this ratio has no upper bound if (ii) is not satisfied. In other words, for a given sampling error norm  $||\mathbf{e}||_{\mathcal{D}}$ , the reconstruction error norm ||e|| can be arbitrarily large. This prevents reconstruction stability. On the other hand, if (ii) is satisfied, we will automatically have  $||e|| \leq ||\mathbf{e}||_{\mathcal{D}}/\alpha$ . For this reason, (ii) is called the condition of "stable sampling"<sup>1</sup> [3], [37]. Note that this is indeed an intrinsic property of the sampling (and not of the chosen reconstruction method) since it is solely dependent on the sampling kernels  $(g_k)_{k\in \mathbb{Z}}$  as seen for example in the equivalent formulation of (iii).

As a final remark, ran(S) is always closed when the number of samples is finite, as is always the case in practice. This is because ran(S) is of finite dimension. At the same time, (ii) is systematically satisfied in this case with  $\alpha$  equal to the smallest positive singular value of S.

# V. CONNECTION OF POCS ITERATION TO SAMPLING PSEUDO-INVERSION

In this section, we show that the POCS iteration  $u^{(n+1)} = P_{\mathcal{A}}P_{\mathcal{H}_s}^{\lambda_n}u^{(n)}$  of (28) belongs to a larger family of algorithms that lead to the pseudo-inversion of *S*, and that is simultaneously a generalization of the *frame algorithm* [1], [3], [5]. At the end of the section, this will allow us to deduce the theoretical effect of sampling noise on the limit of  $u^{(n)}$ .

# A. Generic form of iterated map

The first step is to connect  $P_{\mathcal{A}}P_{\mathcal{H}_s}^{\lambda}$  to the operators S and  $S^*$ . It follows from (11) and (26) that

$$u \in \mathcal{H}, \qquad P_{\mathcal{H}_{\mathbf{s}}}^{\lambda} u = u + \lambda P_{\mathcal{G}}(x - u).$$
 (48)

We then obtain from (51) and (22) that

$$\forall u \in \mathcal{A}, \qquad P_{\mathcal{A}} P_{\mathcal{H}_{s}}^{\lambda} u = u + \lambda \sum_{k \in \mathsf{Z}} \frac{\mathbf{s}_{k} - \langle u, g_{k} \rangle}{\|g_{k}\|^{2}} \widetilde{g}_{k}$$

where  $\tilde{g}_k$  was defined in (17). Since  $s - Su = (s_k - \langle u, g_k \rangle)_{k \in \mathbb{Z}}$ , it then follows from (39) that

$$\forall u \in \mathcal{A}, \qquad P_{\mathcal{A}} P_{\mathcal{H}_{\mathbf{s}}}^{\lambda} u = u + \lambda S^*(\mathbf{s} - Su).$$
(49)

#### B. Generalized frame algorithm

where

The previous section has shown that the iteration of (28) is of the type

$$u^{(n+1)} = R_{\mathbf{s}}^{\lambda_n} u^{(n)}, \quad n \ge 0$$
(50)

$$R_{\mathbf{s}}^{\lambda}u := u + \lambda S^{*}(\mathbf{s} - Su) \tag{51}$$

for any  $u \in A$  and  $\lambda \in \mathbb{R}$ . We are going to see a direct connection between the iteration of (50) and the pseudo-inversion of S under very general assumptions on S. Specifically, we will only assume in this subsection and the next that S is a bounded operator from  $(A, \langle \cdot, \cdot \rangle)$  to any other Hilbert space

<sup>&</sup>lt;sup>1</sup>In the literature, "stable sampling" also implies uniqueness of reconstruction, i.e.,  $\mathcal{F} = \mathcal{A}$ . But in this paper, we use this expression in the more general context where  $\mathcal{F} \subset \mathcal{A}$ . This refers more generally to the stability of reconstruction of a consistent estimate.

 $(\mathcal{D}, \langle \cdot, \cdot \rangle)_{\mathcal{D}}$  (not necessarily made of sequences). Section IV will still be applicable, except (29-31), (37-39) and Proposition 4.3 (iii) ((45) will be used as the definition of  $\mathcal{F}$ ).

Starting from  $u^{(0)} \in \mathcal{F}$ , note first from (50) and (51) that  $u^{(n)}$ must remain in  $\mathcal{F}$  since  $\operatorname{ran}(S^*) = \mathcal{F}$  from (45). Therefore, if  $u^{(n)}$  is convergent, its limit must be a fixed point of  $R_{s}^{\lambda}$  in  $\mathcal{F}$ . The next proposition then gives an immediate connection of (50) with the pseudo-inversion of S.

Proposition 5.1: Assume that S has closed range. For any  $s \in D$  and  $\lambda \in \mathbb{R}$ ,  $S^{\dagger}s$  is a fixed point of  $R_s^{\lambda}$  in  $\mathcal{F}$ .

*Proof:* Note first that  $S^{\dagger} \mathbf{s} \in \mathcal{F}$  from (45). Next, it is easy to see from (51) that any solution u to the equation  $S^*(\mathbf{s} - Su) = 0$  is a fixed point of  $R_{\mathbf{s}}^{\lambda}$ . It follows from (44b) that  $\mathbf{s} - SS^{\dagger}\mathbf{s} = \mathbf{s} - P_{\operatorname{ran}(S)}\mathbf{s}) \in \operatorname{ran}(S)^{\perp} = \operatorname{null}(S^*)$  due to (40). So  $S^*(\mathbf{s} - SS^{\dagger}\mathbf{s}) = 0$ .

What remains is to find a condition for (50) to be convergent. This question turns out to be resolved in the analysis of the *frame algorithm* [1], [3], [5] which falls in the particular case where  $\mathcal{D} = \ell^2(\mathbb{Z})$ ,  $\mathbf{s} \in \operatorname{ran}(S)$  and  $\mathcal{F} = \mathcal{A}$ . In the following proposition, we reproduce this analysis under the general Hilbert space assumptions of this section.

*Proposition 5.2:* For any  $s \in \mathcal{D}$  and  $\lambda \in \mathbb{R}$ ,

$$\forall v, w \in \mathcal{F}, \qquad \|R_{\mathbf{s}}^{\lambda}v - R_{\mathbf{s}}^{\lambda}w\| \le c_{\lambda} \|v - w\| \qquad (52)$$

where

$$e_{\lambda} = \max\left(|1 - \lambda \|S\|^2|, |1 - \lambda \gamma(S)^2|\right), \tag{53}$$

$$||S|| := \sup_{u \in \mathcal{F} \setminus \{0\}} ||Su||_{\mathcal{D}} / ||u||$$
(54)

and  $\gamma(S)$  is defined in (47).

(

*Proof:* It follows from (51) that

$$R_{\mathbf{s}}^{\lambda}u = Q^{\lambda}u + \lambda S^{*}\mathbf{s}$$
 where  $Q^{\lambda}u := u - \lambda S^{*}Su$ . (55)

Thus,  $R_{s}^{\lambda}v - R_{s}^{\lambda}w = Q^{\lambda}(v - w)$ . Then, (52) is satisfied with  $c_{\lambda} := \sup_{u \in \mathcal{F} \setminus \{0\}} ||Q^{\lambda}u|| / ||u||$ . Because  $Q^{\lambda}$  is self-adjoint, it is known (see for example [38, §2.13]) that  $c_{\lambda}$  is equivalently the supremum of  $|\langle u, Q^{\lambda}u \rangle|$  over the set  $U_{\mathcal{F}} := \{u \in \mathcal{F} : ||u|| = 1\}$ . We have

$$\langle u, Q^{\lambda}u \rangle = \|u\|^2 - \lambda \|Su\|_{\mathcal{D}}^2$$

since  $\langle u, Q^{\lambda}u \rangle = \langle u, u \rangle - \lambda \langle u, S^*Su \rangle = ||u||^2 - \lambda \langle Su, Su \rangle_{\mathcal{D}}$ due to (35). It then follows from (54) and (47) that the infimum and the supremum of  $\langle u, Q^{\lambda}u \rangle$  over  $U_{\mathcal{F}}$  are  $1 - \lambda ||S||^2$  and  $1 - \lambda \gamma(S)^2$ , respectively. This leads to (53).

If  $c_{\lambda} < 1$  in (52) for some given  $\lambda$ , then  $R_{s}^{\lambda}$  is a contraction within  $\mathcal{F}$  and  $u^{(n)}$  from (50) is guaranteed to converge starting from  $u^{(0)} \in \mathcal{F}$  with  $\lambda_{n} = \lambda$  for all  $n \geq 0$ . By property of contraction [39],  $S^{\dagger}s$  will have to be the unique fixed point of  $R_{s}^{\lambda}$  in  $\mathcal{F}$  and hence the limit of  $u^{(n)}$ . When S is bounded of closed range, it is shown in [1], [3] that  $c_{\lambda}$  is minimized with  $\lambda = 2/(\gamma(S)^{2} + ||S||^{2})$  of minimum value  $(||S||^{2} - \gamma(S)^{2})/(||S||^{2} + \gamma(S)^{2})$ . The interest of the present paper will be more generally to know the values of  $\lambda$  that lead to a coefficient  $c_{\lambda} < 1$ .

Proposition 5.3: Let 
$$c_{\lambda}$$
 be defined by (53). For  $\epsilon > 0$ ,  
 $\lambda \in [\epsilon, 2||S||^{-2} - \epsilon] \implies c_{\lambda} \le \rho_{\epsilon} := 1 - \epsilon \gamma(S)^2$ . (56)  
If S has closed range,  $\rho_{\epsilon} < 1$ .

*Proof:* By construction,  $\gamma(S) \leq ||S||$ . In this case, it can be verified that  $c_{\lambda}$  is more specifically equal to  $\max(\lambda ||S||^2 - 1, 1 - \lambda \gamma(S)^2)$ . Now, while  $\lambda ||S||^2 - 1 \leq (2||S||^{-2} - \epsilon) ||S||^2 - 1 = 1 - \epsilon ||S||^2$ , we have  $1 - \lambda \gamma(S)^2 \leq 1 - \epsilon \gamma(S)^2$ . The second upper bound is the larger one, which proves (56). When S has closed range,  $\gamma(S) > 0$  by Proposition 4.3, and hence  $\rho_{\epsilon} < 1$ .

## C. Generalized iteration

In the previous section, we focused on the case where  $u^{(0)} \in \mathcal{F}$  and  $(\lambda_n)_{n\geq 0}$  is constant. We now give a general result of convergence of  $u^{(n)}$  from (50) in absence of these two conditions. The case where  $u^{(0)} \notin \mathcal{F}$  will be resolved thanks to the following property.

Proposition 5.4: For all  $v \in \mathcal{F}$  and  $w \in \mathcal{F}^{\perp}$ ,

$$R_{\mathbf{s}}^{\lambda}(v+w) = R_{\mathbf{s}}^{\lambda}v + w.$$

*Proof:* It follows from (55) that  $R_{\mathbf{s}}^{\lambda}(v+w) = Q^{\lambda}(v+w) + \lambda S^* \mathbf{s} = R_{\mathbf{s}}^{\lambda}v + Q^{\lambda}w$  by linearity of  $Q^{\lambda}$ . But due to (34), Sw = 0. Therefore,  $Q^{\lambda}w = w$ .

By combining this with Propositions 5.1-5.3, we obtain the following result.

Theorem 5.5: Assuming that S is bounded of closed range, let  $(u^{(n)})_{n\geq 0}$  be a sequence satisfying the recursion of (50) for some  $\mathbf{s} \in \mathcal{D}$  and initial iterate  $u^{(0)} \in \mathcal{A}$ , with a sequence of coefficients  $(\lambda_n)_{n\geq 0}$  such that  $\lambda_n \in [\epsilon, 2||S||^{-2} - \epsilon]$  for all  $n \geq 0$  and some constant  $\epsilon > 0$ . Then,  $(u^{(n)})_{n\geq 0}$  is convergent of limit

$$u^{(\infty)} = S^{\dagger} \mathbf{s} + P_{\mathcal{F}^{\perp}} u^{(0)} = P_{\mathcal{M}_{\mathbf{s}}} u^{(0)}$$
(57)

and  $||u^{(n)} - u^{(\infty)}|| \le \rho_{\epsilon}^{n} ||u^{(0)} - u^{(\infty)}||, \quad \forall n \ge 0$  (58)

where  $\mathcal{M}_{s}$  and  $\rho_{\epsilon}$  are defined in (42) and (56).

*Proof*: Let  $u^{(0)} \in \mathcal{A}$ . We can write that  $u^{(0)} = v^{(0)} + w^{(0)}$ where  $v^{(0)} := P_{\mathcal{F}}u^{(0)}$  and  $w^{(0)} := P_{\mathcal{F}^{\perp}}u^{(0)}$ . Let  $v^{(n+1)} := R_{\mathbf{s}}^{\lambda_n}v^{(n)}$  for all  $n \geq 0$ . Since  $v^{(0)} \in \mathcal{F}$ , we saw in the previous section that  $v^{(n)}$  remains in  $\mathcal{F}$  for all  $n \geq 0$ . By assumption on  $\lambda_n$ , it follows from Proposition 5.3 that  $c_{\lambda_n} \leq \rho_{\epsilon}$ . Then, by applying this with Propositions 5.1 and 5.2 on the sequence  $(v^{(n)})_{n\geq 0}$ , we obtain that  $\|v^{(n+1)} - S^{\dagger}\mathbf{s}\| = \|R_{\mathbf{s}}^{\lambda_n}v^{(n)} - R_{\mathbf{s}}^{\lambda_n}(S^{\dagger}\mathbf{s})\| \leq \rho_{\epsilon}\|v^{(n)} - S^{\dagger}\mathbf{s}\|$ . Since  $\rho_{\epsilon} < 1$ , this proves that  $v^{(n)}$  is convergent of limit  $v^{(\infty)} = S^{\dagger}\mathbf{s}$  and that

$$\|v^{(n)} - v^{(\infty)}\| \le \rho_{\epsilon}^{n} \|v^{(0)} - v^{(\infty)}\|, \quad \forall n \ge 0.$$
 (59)

Meanwhile, one easily finds by induction from Proposition 5.4 that  $u^{(n)} = v^{(n)} + w^{(0)}$  for all  $n \ge 0$ . Therefore,  $u^{(n)}$ tends to  $u^{(\infty)} = v^{(\infty)} + w^{(0)} = S^{\dagger} \mathbf{s} + P_{\mathcal{F}^{\perp}} u^{(0)}$ . On the one hand, (58) is deduced from (59) by a mere space translation by  $w^{(0)}$ . On the other hand, it results from (46) that  $P_{\mathcal{M}_{\mathbf{s}}} u^{(0)} =$  $P_{(\mathcal{F}^{\perp} + S^{\dagger} \mathbf{s})} u^{(0)} = P_{\mathcal{F}^{\perp}} (u^{(0)} - S^{\dagger} \mathbf{s}) + S^{\dagger} \mathbf{s} = P_{\mathcal{F}^{\perp}} u^{(0)} + S^{\dagger} \mathbf{s}$ since  $S^{\dagger} \mathbf{s} \in \mathcal{F}$  from (45). This proves (57).

# D. Application to POCS iteration

A known shortcoming of the frame algorithm is that the range of admissible relaxation coefficients depends on ||S|| which is not necessarily accessible in practice. This problem is however less critical with the iteration of (28) which corresponds to  $R_s^{\lambda} = P_{\mathcal{A}} P_{\mathcal{H}_s}^{\lambda}$ . Returning to the beginning of Section V, this is the case of (51) where S is defined by (29) with (30) under the assumption of (3). Because of (38) and the fact that  $P_{\mathcal{G}}$  is non-expansive, we know that  $||S|| \leq 1$ . In this case, the interval  $[\epsilon, 2||S||^{-2} - \epsilon]$  includes  $[\epsilon, 2-\epsilon]$  as a subset. Theorem 5.5 then has the following consequence.

Corollary 5.6: Let  $(u^{(n)})_{n\geq 0}$  be recursively defined by (28) for some  $s \in \mathcal{D}$ , coefficients  $(\lambda_n)_{n\geq 0}$  and initial iterate  $u^{(0)} \in \mathcal{A}$  with the following assumptions: the operator S defined by (29) and (30) has closed range and  $\lambda_n \in [\epsilon, 2-\epsilon]$  for all  $n \geq 0$ and some  $\epsilon > 0$ . Then  $(u^{(n)})_{n\geq 0}$  is convergent of limit  $u^{(\infty)}$ satisfying (57) and (58).

Note that the above assumption on  $\lambda_n$  implies that  $\sum_{n\geq 0} \lambda_n (2-\lambda_n) = +\infty$ . Theorem 3.3 is then applicable without the condition that  $\operatorname{ran}(S)$  is closed, and hence with no stable sampling condition. The requirement of this theorem, however, is that  $\mathcal{A}_s$  must be non-empty. Now when  $\operatorname{ran}(S)$  is closed, (57) can be seen as an extension of (25) allowing  $\mathcal{A}_s$  to be empty (due to sampling noise for example) following the remark at the end of Section IV-C. Also, we have from (58) that the convergence of  $u^{(n)}$  is *linear* (referring to the exponent of  $\gamma^n$ ), which is not guaranteed by Theorem 3.3.

## E. Consequence on noisy sampling

In absence of sampling noise, we already know the signal implications of the POCS limit  $P_{\mathcal{A}_s} u^{(0)}$  of (25). In practice however, the sampling sequence that is injected into (25) is often not s, but

$$\hat{\mathbf{s}} := \mathbf{s} + \mathbf{e}$$
 where  $\mathbf{s} := Sx$  (60)

and  $\mathbf{e} = (\mathbf{e}_k)_{k \in \mathbb{Z}}$  is some unknown error sequence. It then follows from (57) and (46) that

$$u^{(\infty)} = P_{\mathcal{F}^{\perp}} u^{(0)} + S^{\dagger} \hat{\mathbf{s}} = \left( P_{\mathcal{F}^{\perp}} u^{(0)} + S^{\dagger} \mathbf{s} \right) + S^{\dagger} \mathbf{e}$$
$$= P_{\mathcal{A}_{-}} u^{(0)} + S^{\dagger} \mathbf{e}$$

since  $\mathbf{s} \in \operatorname{ran}(S)$  and hence  $\overline{\mathbf{s}} = \mathbf{s}$ . While  $P_{\mathcal{A}_s} u^{(0)}$  is the noise-free POCS iteration limit,  $S^{\dagger}\mathbf{e}$  is the deviation of this limit due to the sampling error sequence  $\mathbf{e}$ .

Proposition 5.7: Let 
$$\mathbf{\bar{e}} := P_{\operatorname{ran}(S)}\mathbf{e}$$
. Then,  
 $S^{\dagger}\mathbf{e} = S^{\dagger}\mathbf{\bar{e}}$  and  $\|\mathbf{\bar{e}}\|_{\mathcal{D}} \leq \|\mathbf{e}\|_{\mathcal{D}}$ 

*Proof:* It follows from (44d) and (40) that  $\operatorname{null}(S^{\dagger}) = \operatorname{null}(S^*) = \operatorname{ran}(S)^{\perp}$ . By construction,  $\overline{\mathbf{e}} - \mathbf{e} \in \operatorname{ran}(S)^{\perp} = \operatorname{null}(S^{\dagger})$ . Thus  $S^{\dagger}\mathbf{e} = S^{\dagger}\overline{\mathbf{e}}$ . Meanwhile,  $\|\overline{\mathbf{e}}\|_{\mathcal{D}} \leq \|\mathbf{e}\|_{\mathcal{D}}$  is from Bessel's inequality.

Hence, only the component  $\overline{\mathbf{e}}$  of  $\mathbf{e}$  in  $\operatorname{ran}(S)$  contributes to the reconstruction deviation. Thus, the POCS iteration has a filtering effect on the noise sequence  $\mathbf{e}$ . Meanwhile, the error

To have a strict inequality  $\|\bar{\mathbf{e}}\|_{\mathcal{D}} < \|\mathbf{e}\|_{\mathcal{D}}$ , note that ran(S) must be a proper subspace of  $\mathcal{D}$ . This necessitates some oversampling. The higher the oversampling ratio is, the smaller ran(S) is compared to  $\mathcal{D}$ , and the smaller  $\|\bar{\mathbf{e}}\|_{\mathcal{D}}$  is compared to  $\|\mathbf{e}\|_{\mathcal{D}}$ . This corresponds to the noise-shaping effect of oversampling in uniform sampling [27].

## VI. PRACTICAL IMPLEMENTATION ASPECTS

We present a rigorous way to implement the iterative part of (50) in discrete-time, without the need to involve generic discrete-time decompositions of the signals of  $\mathcal{A}$  (such as sincbasis decompositions for bandlimited signals). Although not studied in this paper, we will then touch on the issue of finitecomplexity implementations.

# A. Discrete-time implementation of iteration

Using (49), (50) takes the form

$$u^{(n+1)} = u^{(n)} + \lambda_n S^*(\mathbf{s} - Su^{(n)}), \qquad n \ge 0.$$
 (61)

This is in practice an iteration of continuous-time functions. For digital signal processing, one expects this iteration to be discretized. Now, whether the space  $\mathcal{A}$  has a countable basis or not, there is a way to obtain  $u^{(n)}$  by a pure discrete-time iteration in  $\mathcal{D}$ . The principle is as follows. Note from (61) that  $u^{(n+1)} - u^{(n)} \in \operatorname{ran}(S^*)$  for all  $n \ge 0$ . Hence,  $u^{(n)} - u^{(0)}$  must be in  $\operatorname{ran}(S^*)$  as well. This implies that there exists some discrete-time sequence  $\mathbf{c}^{(n)} \in \mathcal{D}$  such that  $u^{(n)} = u^{(0)} + S^* \mathbf{c}^{(n)}$ . The next proposition implies a way to construct  $\mathbf{c}^{(n)}$  recursively.

*Proposition 6.1:* For any given initial estimate  $u^{(0)}$ , the iterate  $u^{(n)}$  of (50) is equivalently obtained by iterating the system

$$\mathbf{c}^{(n+1)} = \mathbf{c}^{(n)} + \lambda_n (\mathbf{s}_0 - SS^* \mathbf{c}^{(n)}) \tag{62a}$$

$$u^{(n)} = u^{(0)} + S^* \mathbf{c}^{(n)} \tag{62b}$$

for  $n \ge 0$ , where  $\mathbf{s}_0 := \mathbf{s} - Su^{(0)}$ , starting with  $\mathbf{c}^{(0)} = \mathbf{0}$ .

*Proof:* It follows from (62b) and (62a) that

$$u^{(n+1)} = u^{(0)} + S^* \mathbf{c}^{(n+1)} = u^{(0)} + S^* (\mathbf{c}^{(n)} + \lambda_n (\mathbf{s}_0 - SS^* \mathbf{c}^{(n)}))$$
  
=  $(u^{(0)} + S^* \mathbf{c}^{(n)}) + \lambda_n S^* (\mathbf{s} - Su^{(0)} - SS^* \mathbf{c}^{(n)})$   
=  $u^{(n)} + \lambda_n S^* (\mathbf{s} - Su^{(n)})$ 

which leads to (61).

The outstanding contribution of (62) compared to (50) is that the pure discrete-time operation of (62a) can be iterated *alone* until the targeted iteration number m. Then, the continuoustime operation of (62b) just needs to be executed *once* at n = m. The operator  $SS^*$  involved in (62a) can be seen as a square matrix of coefficients

$$SS^* = \left[\frac{\langle \tilde{g}_{k'}, g_k \rangle}{\|g_{k'}\|^2}\right]_{(k,k') \in \mathsf{Z} \times \mathsf{Z}}$$
(63)

which needs to be predetermined before the iteration. In general, the coefficients can only be obtained numerically. It was proposed in [20], [21] to obtain them from a *single-argument* lookup table.

## B. Finite-complexity implementation

A remaining issue is the finite-complexity implementation of (62). The most critical part is the recurrent operation of  $SS^*$  in (62a). For any  $\mathbf{c} = (\mathbf{s}_k)_{k \in \mathbb{Z}} \in \mathcal{D}$ , the *k*th component of  $SS^*\mathbf{c}$  is

$$SS^*\mathbf{c}|_k = \sum_{k' \in \mathsf{Z}} \mathbf{h}_{k,k'} \,\mathbf{s}_{k'} \tag{64}$$

where  $h_{k,k'} := \langle \tilde{g}_{k'}, g_k \rangle_2 / ||g_{k'}||_2^2$ . For concrete analysis, let us assume the traditional case where  $\mathcal{H} = L^2(\mathbb{R}), \mathcal{A} = \mathcal{B}$  and  $Z \subset \mathbb{Z}$ . When the sampling is uniform, i.e.,  $g_k(t) = g(t-kT)$ for some  $g(t) \in L^2(\mathbb{R})$ ,  $\mathbf{h}_{k,k'}$  reduces to  $\mathbf{h}_{k'-k}$  where  $h_k := \langle \tilde{g}, g_k \rangle_2 / \|g\|_2^2$ . So  $SS^*c$  is just a convolution operation. Note that  $(h_k)_{k \in \mathbb{Z}}$  are simply the squared uniform samples of  $\widetilde{g}(t)$  with some scaling factor. With the bandlimitation,  $h_k$ decays towards infinity in a sinc-like manner. For operation of finite complexity, it is necessary to approximate  $h_k$  using standard FIR windowing methods, which creates reconstruction distortions. When the sampling is nonuniform,  $SS^*$  loses its time invariance. But it can still be seen as a linear filter of timevarying impulse response  $h_{k,k'}$ , with expected decays when k'gets far away from k. However, the windowing of time-varying filters plus the effect of filter distortions to the POCS iteration remain virgin topics that require new substantial investigations not tackled in this article and to be addressed in the future. Some preliminary experiments can been found in [20].

## VII. MULTI-CHANNEL ORTHOGONAL SAMPLING

In this section, we illustrate the theoretical power of our formalism by revisiting the sampling/reconstruction system designed in [32], [11] for multi-channel time encoding. A basic POCS algorithm was used to reconstruct a multi-channel signal from the elaborate sampling system shown in Fig. 2<sup>2</sup>. We show that this encoding system turns out to satisfy the abstract conditions of (1) and (3). While the reconstruction iterates of [32], [11] coincide with those of (84) under our formalism, our theory uncovers the full pseudo-inversion property of this method, which was only studied in a noise-free case of perfect reconstruction in these references. Moreover, while the reconstruction method was mostly presented at a conceptual level in [32], [11], our abstract reformulation simultaneously allows a more explicit descriptions of practical implementations.

## A. System description

The time-encoding system of [32], [11] assumes that the source signals are multidimensional bandlimited functions

$$\boldsymbol{y}(t) = (y^1(t), \cdots, y^N(t)) \in \mathcal{B}^N$$

<sup>2</sup>The letters 'x' and 'y' from [32], [11] have been interchanged in Fig. 2 to be compatible with the notation of the present article.



Fig. 2. Multi-channel system of time-encoding machines (TEM) from [32], [11].

Next, instead of sampling the functions  $y^i(t)$  individually, the system first expands y(t) into a redundant representation

$$\boldsymbol{x}(t) := \mathbf{A}\boldsymbol{y}(t), \qquad t \in \mathbb{R}$$

where A is an  $M \times N$  matrix that is assumed in [32], [11] to be full rank with  $M \ge N$ . In our present analysis, we will not necessarily assume so. The signal x(t) is thus of the form

$$\boldsymbol{x}(t) = (x^1(t), \cdots, x^M(t)), \qquad t \in \mathbb{R}.$$

Each component  $x^i(t)$  is then processed through an integrateand-fire encoding machine, which outputs a sequence of spikes located at some increasing time instants  $(t_j^i)_{j \in Z_i}$ , where  $Z_i$  is some index set of consecutive integers. From the derivations of [16], this provides the knowledge of successive integral values

$$\mathbf{s}_{i,j} := \int_{t_{j-1}^i}^{t_j^i} x^i(t) \, \mathrm{d}t, \qquad j \in \mathsf{Z}_i.$$
(65)

The work of [32], [11] uses a POCS iteration to retrieve x(t), before y(t) is recovered with the relation

$$\forall t \in \mathbb{R}, \qquad \boldsymbol{y}(t) = \mathbf{A}^+ \boldsymbol{x}(t)$$
 (66)

where  $\mathbf{A}^+$  is the matrix pseudo-inverse of  $\mathbf{A}$ .

#### B. Signal and system formalization

We now show the existence of spaces  $\mathcal{H}$  and  $\mathcal{A}$  that allow to present the samples  $s_{i,j}$  of (65) in the form of (1) with condition (3). Let  $\mathcal{H} := (L^2(\mathbb{R}))^M$ . Each element  $u \in \mathcal{H}$  is a function of time  $u(t) = (u^1(t), \dots, u^M(t))$ . The canonical inner product of  $\mathcal{H}$  is defined by

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle := \sum_{i=1}^{M} \langle u^i, v^i \rangle_2 = \int_{\mathbb{R}} \boldsymbol{u}(t)^{\mathsf{T}} \boldsymbol{v}(t) \, \mathrm{d}t, \qquad \boldsymbol{u}, \boldsymbol{v} \in \mathcal{H}$$
(67)

where in the last expression, u(t) and v(t) are seen for each  $t \in \mathbb{R}$  as *M*-dimensional column vectors. The signal x(t) to be retrieved is an element  $x \in \mathcal{H}$  that specifically lies in the closed subspace

$$\mathcal{A} := \left\{ \boldsymbol{v} \in \mathcal{B}^M : \forall t \in \mathbb{R}, \ \boldsymbol{v}(t) \in \operatorname{ran}(\mathbf{A}) \right\}$$

Next, let us show that  $s_{i,j}$  of (65) can be formalized as

$$\mathbf{s}_{i,j} = \langle \boldsymbol{x}, \boldsymbol{g}_{i,j} \rangle, \qquad (i,j) \in \mathsf{Z}.$$
 (68)

Naturally,

$$\mathsf{Z} := \{(i,j) : i \in \mathsf{M} \text{ and } j \in \mathsf{Z}_i\} \text{ where } \mathsf{M} := \{1, \cdots, M\}.$$

Then, (68) clearly coincides with (65) by taking

$$\boldsymbol{g}_{i,j}(t) := \left(0, \cdots, 0, g_j^i(t), 0, \cdots, 0\right) \tag{69}$$

where  $g_i^i(t)$  is at the *i*th coordinate position and is equal to

$$g_j^i(t) = \mathbf{1}_{I_j^i}(t)$$
 where  $I_i^j := [t_{j-1}^i, t_j^i].$  (70)

It is clear as a result that  $(g_{i,j})_{(i,j)\in Z}$  is an orthogonal family of H. Thus, condition (3) is realized. In fact, to obtain this property, it is sufficient to have

$$(g_j^i)_{j \in \mathsf{Z}_i}$$
 orthogonal in  $L^2(\mathbb{R})$  for each  $i \in \mathsf{M}$ . (71)

We will just assume this condition for now. and will apply the explicit assumption of (70) only later on in Section VII-F.

# C. POCS iteration implementation

All of Sections III and V is applicable. In this process, we are thus uncovering the general pseudo-inversion properties of the POCS iteration limit of (28), which was only studied in a noise-free case of perfect reconstruction with  $\lambda_n = 1$  in [24], [11]. But the more specific contribution of interest here is the application of Section VI-A on a practical implementation of the POCS iteration. We recall that the relaxed POCS algorithm of (61) is efficiently implemented by iterating the system (62). Using the specific function notation of Section VII-B, this system is

$$\mathbf{c}^{(n+1)} = \mathbf{c}^{(n)} - \lambda_n S S^* \mathbf{c}^{(n)} + \mathbf{s}_0 \tag{72a}$$

$$u^{(n)} = u^{(0)} + S^* \mathbf{c}^{(n)}$$
(72b)

starting from  $\mathbf{c}^{(0)} = \mathbf{0}$ , where

$$\mathbf{s}_0 := \lambda_n (\mathbf{s} - S \boldsymbol{u}^{(0)}), \tag{73a}$$

$$SS^* = \left[\frac{\langle \widetilde{\boldsymbol{g}}_{i'j'}, \boldsymbol{g}_{i,j} \rangle}{\|\boldsymbol{g}_{i'j'}\|^2}\right]_{((i,j),(i',j'))\in\mathsf{Z}\times\mathsf{Z}},\tag{73b}$$

$$S^* \mathbf{c} = \sum_{(i,j)\in\mathsf{Z}} \frac{\mathbf{s}_{i,j}}{\|\boldsymbol{g}_{ij}\|^2} \,\widetilde{\boldsymbol{g}}_{i,j}, \quad \mathbf{c} = (\mathbf{s}_{i,j})_{(i,j)\in\mathsf{Z}} \in \mathcal{D} \quad (73c)$$

and 
$$\widetilde{\boldsymbol{g}}_{i,j} = P_{\mathcal{A}} \boldsymbol{g}_{i,j}, \quad (i,j) \in \mathsf{Z}.$$
 (73d)

In the next two subsections, we derive the expressions of  $SS^*$ and  $S^*\mathbf{c}$  n terms of the family of scalar functions  $(g_i^j)_{(i,j)\in \mathbb{Z}}$ .

#### D. Discrete-time iteration

The derivation of  $SS^*$  needed in (72a) starts with the following preliminary result.

Proposition 7.1: If u(t) = u(t) u where  $u(t) \in L^2(\mathbb{R})$  and  $\mathbf{u} \in \mathbb{R}^M$ , then,

$$P_{\mathcal{A}}\boldsymbol{u}(t) = \tilde{\boldsymbol{u}}(t) \mathbf{A}\mathbf{A}^{+}\mathbf{u}.$$

We prove this in Appendix C. Then, the coefficients of the operator  $SS^*$  described in (73b) are given as follows.

Proposition 7.2: For all 
$$(i, j), (i', j') \in \mathbb{Z}$$
,  
 $\|\boldsymbol{g}_{i,j}\|^2 = \|g_j^i\|_2^2$  and  $\langle \tilde{\boldsymbol{g}}_{i',j'}, \boldsymbol{g}_{i,j} \rangle = \langle \tilde{g}_{j'}^{i'}, g_j^i \rangle_2 \mathbf{a}_{ii'}$  (74)  
where  $\tilde{\boldsymbol{u}} := P_{\mathcal{B}} \boldsymbol{u}$  for any  $\boldsymbol{u} \in L^2(\mathbb{R})$  and  $\mathbf{a}_{ii'}$  is the entry of

 $P_{\mathcal{B}} u$  for any matrix  $\mathbf{A}\mathbf{A}^+$  at index  $(i, i') \in \mathsf{M} \times \mathsf{M}$ .

*Proof:* The first equality of (74) is clear from (69) and (67). Let  $\mathbf{e}_i$  designate the *i*th coordinate vector of  $\mathbb{R}^M$ . It follows from (69), (73c) and Proposition 7.1 that

$$\boldsymbol{g}_{i,j}(t) = g_j^i(t) \, \mathbf{e}_i$$
 and  $\widetilde{\boldsymbol{g}}_{i,j}(t) = \widetilde{g}_j^i(t) \, \mathbf{A} \mathbf{A}^+ \mathbf{e}_i$  (75)

for all  $(i, j) \in Z$ . Thus, both  $g_{i,j}(t)$  and  $\tilde{g}_{i,j}(t)$  are of the form u(t) u. The general identity

$$\left\langle u(t)\mathbf{u}, v(t)\mathbf{v} \right\rangle = \left\langle u, v \right\rangle_2 \mathbf{u}^{\mathsf{T}} \mathbf{v}$$

that results from (67) then implies the second equality of (74)with  $a_{ii'} := (\mathbf{A}\mathbf{A}^+\mathbf{e}_i)^\top \mathbf{e}_{i'} = \mathbf{e}_i^\top (\mathbf{A}\mathbf{A}^+)\mathbf{e}_{i'}$  since  $\mathbf{A}\mathbf{A}^+$  is an orthogonal projection and hence symmetric.

#### E. Final continuous-time output

Once (72a) has been iterated the desired number of times m, one can output the continuous-time multi-channel signal  $\boldsymbol{u}^{(m)}(t)$  from (72b). For that purpose, we need to know the explicit expression of  $S^*\mathbf{c}$  in terms of  $\mathbf{c} = (\mathbf{s}_{i,j})_{(i,j)\in \mathsf{Z}} \in \mathcal{D}$ . It follows from (73c), (74) and (75) that  $S^*c$  is the continuoustime function

$$S^{*}\mathbf{c} = \sum_{(i,j)\in\mathbf{Z}} \frac{\mathbf{s}_{i,j}}{\|g_{j}^{i}\|_{2}^{2}} \tilde{g}_{j}^{i}(t) \mathbf{A} \mathbf{A}^{+} \mathbf{e}_{i} = \sum_{i\in\mathbf{M}} P_{\mathcal{B}}(c^{i}(t)) \mathbf{A} \mathbf{A}^{+} \mathbf{e}_{i}$$
  
here  $c^{i}(t) := \sum_{j\in\mathbf{Z}_{i}} \frac{\mathbf{s}_{i,j}}{\|g_{j}^{i}\|_{2}^{2}} g_{j}^{i}(t).$  (76)

where

If one needs to provide an estimate of the source signal y(t), the relation (66) naturally leads us to consider the estimate

$$\boldsymbol{v}^{(m)}(t) := \mathbf{A}^+ \boldsymbol{u}^{(m)}(t) = \mathbf{A}^+ S^* \mathbf{c}^{(m)} \in \mathcal{B}^N.$$

For any  $c \in D$ , it results from (76) that

$$\mathbf{A}^{+}S^{*}\mathbf{c} = \sum_{i \in \mathsf{M}} P_{\mathcal{B}}(c^{i}(t)) \mathbf{a}_{i}^{+}$$

where  $\mathbf{a}_i^+ := \mathbf{A}^+ \mathbf{A} \mathbf{A}^+ \mathbf{e}_i = \mathbf{A}^+ \mathbf{e}_i$  as a general result of matrix pseudo-inverse. Then,  $\mathbf{a}_i^+$  is nothing but the *i*th column vector of  $A^+$ .

# F. Explicit case of samples of (65)

We now look at the more specific implementation for the samples  $s_{ij}$  considered in (65). As shown in Section VII-B, this corresponds to the case where  $g_i^i(t)$  is given by (70). The quantities that need to be derived more explicitly are

$$\left\|g_{j}^{i}\right\|_{2}, \qquad \left<\tilde{g}_{j'}^{i'}, g_{j}^{i}\right>_{2} \qquad \text{and} \qquad c^{i}(t)$$

for (74) and (76). It is first clear that

$$\|g_j^i\|_2^2 = t_j^i - t_{j-1}^i$$

As a generalization of a derivation from [20] in the case of a single channel, it can be derived that

$$\begin{array}{l} \langle \tilde{g}_{j'}^{i'}, g_{j}^{i} \rangle_{2} = f(T_{j,j'-1}^{i,i'}) - f(T_{j-1,j'-1}^{i,i'}) - f(T_{j,j'}^{i,i'}) + f(T_{j-1,j'}^{i,i'}) \\ (77) \\ \text{where } T_{j,j'}^{i,i'} := t_{j}^{i} - t_{j'}^{i'} \text{ and } f(t) := \int_{0}^{t} (t-\tau) \operatorname{sinc}(\tau) \, \mathrm{d}\tau. \\ \text{Although } \langle \tilde{g}_{j'}^{i'}, g_{j}^{i} \rangle_{2} \text{ depends on the four time instants} \\ (t_{i}^{i}, t_{i-1}^{i}, t_{i'+1}^{i'}, t_{i'-1}^{i'}) \text{ and is composed of four terms, it is the} \end{array}$$

same single-argument numerical function f(t) that is used.

The values of this function can be stored in a lookup table, so that the required computation is just limited to a few additions and subtractions.

Finally

$$c^{i}(t) = \sum_{j \in \mathsf{Z}_{i}} \mathrm{s}^{i}_{j} \ \mathbf{1}_{I^{i}_{j}}(t) \qquad ext{where} \qquad \mathrm{s}^{i}_{j} := rac{\mathrm{s}_{i,j}}{t^{i}_{j} - t^{i}_{j-1}}$$

This is nothing but the piecewise constant function equal to  $s_j^i$  in  $I_j^i$  for each  $j \in Z_i$ . This is produced by analog circuits using a zero-order hold.

## VIII. BANDLIMITED INTERPOLATION BY ITERATIVE PIECEWISE-LINEAR CORRECTIONS

Until now, the examples of orthogonal sampling kernels we have provided all result from sampling by integration. An ultimate goal would be to have such a type of kernels for point sampling. For x(t) that is at least continuous on  $\mathbb{R}$ , one wishes to have  $x(t_k) = \langle x, h_k \rangle_2$  where the functions  $(h_k)_{k \in \mathbb{Z}}$  have non-overlapping supports so that they are orthogonal in  $L^2(\mathbb{R})$ . The only way to do that would be to take  $h_k(t) := \delta(t-t_k)$ where  $\delta(t)$  is the Dirac impulse. However, this is not a function of  $L^2(\mathbb{R})$ . We show in this section that orthogonal kernels can actually be obtained for point sampling by choosing as ambient space H a homogeneous Sobolev space. The resulting POCS method turns out to coincide with an existing algorithm by Grochenig [4, §4.1]. However, as in the example of Section VII, this algorithm was only analyzed in a case of noise-free perfect reconstruction, while our analysis leads to a complete set of pseudo-inversion properties.

## A. Initial idea

Assume that a bandlimited function x is given by point samples  $(x(t_k))_{k\in\mathbb{Z}}$  at some known increasing sequence of instants  $(t_k)_{k\in\mathbb{Z}}$ . For functions u(t) of sufficient regularity, the point samples  $(u(t_k))_{k\in\mathbb{Z}}$  yield the relation

$$u(t_k) - u(t_{k-1}) = \int_{t_{k-1}}^{t_k} u'(t) \, \mathrm{d}t = \left\langle u', \mathbf{1}_{[t_{k-1}, t_k]} \right\rangle_2 = \langle u', g'_k \rangle_2$$
(78)

where

$$g_k(t) := \begin{cases} 0 & , \quad t \in (-\infty, t_{k-1}) \\ t - t_{k-1} & , \quad t \in [t_{k-1}, t_k) \\ t_k - t_{k-1} & , \quad t \in [t_k, \infty) \end{cases} , \quad k \in \mathbb{Z}.$$
(79)

Using the notation

$$\Delta a_k := a_k - a_{k-1} \quad \text{and} \quad \langle u, v \rangle := \langle u', v' \rangle_2, \quad (80)$$

we obtain from (78) that

$$\Delta u(t_k) = \langle u, g_k \rangle, \qquad k \in \mathsf{Z}.$$
(81)

From the point samples  $(x(t_k))_{k\in\mathbb{Z}}$ , we can then form the generalized samples

$$\mathbf{s}_k = \langle x, g_k \rangle$$
 where  $\mathbf{s}_k := \Delta x(t_k), \quad k \in \mathsf{Z}.$  (82)

Moreover, since  $g'_k = 1_{[t_{k-1},t_k]}$ ,  $(g'_k)_{k\in\mathbb{Z}}$  is an orthogonal family in  $L^2(\mathbb{R})$ . So  $(g_k)_{k\in\mathbb{Z}}$  is orthogonal with respect to  $\langle \cdot, \cdot \rangle$ . The main issue is to find a Hilbert space  $\mathcal{H}$  in which  $\langle \cdot, \cdot \rangle$  is a well defined inner product.

#### B. Ambient Hilbert space

A suitable candidate for  $\mathcal{H}$  is the homogeneous Sobolev space [25], [26] defined by

$$\dot{H}^{1}(\mathbb{R}) = \left\{ u(t) : u \text{ is absolutely continuous on } \mathbb{R} \\ \text{and } u' \in L^{2}(\mathbb{R}) \right\}$$

where absolute continuity is here in the local sense. In this case, u is absolutely continuous on  $\mathbb{R}$  if and only if it is differentiable almost everywhere of locally integrable derivative u' such that  $u(b) = u(a) + \int_a^b u'(t) dt$  for any  $a \leq b$  [40, §11.4.6]. Since  $u' \in L^2(\mathbb{R})$  for any  $u \in \dot{H}^1(\mathbb{R})$ , then the function  $\langle \cdot, \cdot \rangle$  of (80) is well defined in  $\dot{H}^1(\mathbb{R})$ . The remaining issue is that the induced function

$$||u|| := \langle u, u \rangle^{1/2} = ||u'||_2 \tag{83}$$

is only a seminorm, as ||u|| = 0 only implies that u(t) is a constant function. In the construction of  $\dot{H}^1(\mathbb{R})$ , it is in fact implied that its functions are uniquely defined up to a constant component (similarly to the functions of  $L^2(\mathbb{R})$  that are uniquely defined pointwise only up to a set of measure 0). Under this setting,  $||\cdot||$  is a norm and  $(\dot{H}^1(\mathbb{R}), \langle \cdot, \cdot \rangle)$  is rigorously a Hilbert space. Qualitatively,  $||u||^2$  is the total slope energy of u(t). We have the following list of properties.

Proposition 8.1:

- (i)  $(g_k)_{k \in \mathbb{Z}}$  is an orthogonal family of  $\dot{H}^1(\mathbb{R})$ .
- (ii) For all  $u \in H^1(\mathbb{R})$ , u satisfies (81).
- (iii) The subspace  $\mathcal{A}$  of all bandlimited functions of  $\dot{H}^1(\mathbb{R})$  of Nyquist period 1 is closed.

*Proof:* (i) As we already saw the orthogonality of  $(g_k)_{k\in\mathbb{Z}}$  with respect to  $\langle \cdot, \cdot \rangle$ , we just need to verify that  $g_k \in \dot{H}^1(\mathbb{R})$ . For each k,  $g_k$  is easily seen to be absolutely continuous. Meanwhile,  $\|g'_k\|_2^2 = \|\mathbf{1}_{[t_{k-1},t_k]}\|_2^2 = \Delta t_k$  so that  $g'_k \in L^2(\mathbb{R})$ . (ii) By absolute continuity, every function  $u \in \dot{H}^1(\mathbb{R})$  satisfies (78) and hence (81).

(iii) Let  $U(\omega)$  be the Fourier transform of a function u(t). For any interval I, let  $u_I(t)$  be the function whose Fourier transform is  $1_I(\omega)U(\omega)$ . For any given  $u \in \dot{H}^1(\mathbb{R})$ , we can write  $u = u_B + u_{\bar{B}}$  where  $B := [-\pi.\pi]$  and  $\bar{B} := \mathbb{R} \setminus B$ . Since  $||v'||_2^2 = \frac{1}{2\pi} \int_{\mathbb{R}} \omega^2 |V(\omega)|^2 d\omega$ , it is clear that  $u'_B, u'_{\bar{B}} \in L^2(\mathbb{R})$ . Meanwhile, they are both absolutely continuous since  $u_B$  is infinitely differentiable and  $u_{\bar{B}} = u - u_B$ . So  $u_B, u_{\bar{B}} \in \dot{H}^1(\mathbb{R})$ . Since  $\langle v, w \rangle = \frac{1}{2\pi} \int_{\mathbb{R}} \omega^2 V(\omega) W^*(\omega) d\omega$ , then  $u_B \perp u_{\bar{B}}$  in  $\dot{H}^1(\mathbb{R})$ . While  $\mathcal{A}$  is the subset of  $\dot{H}^1(\mathbb{R})$  of functions whose Fourier transforms are supported by  $B := [-\pi, \pi]$ , let  $\bar{\mathcal{A}}$  be the subset of  $\dot{H}^1(\mathbb{R})$  of functions whose Fourier transforms are supported by  $\bar{B}$ . We have just proved that  $\mathcal{A} \oplus \bar{\mathcal{A}}$  is an orthogonal decomposition of  $\dot{H}^1(\mathbb{R})$ . This proves that  $\mathcal{A}$  is closed in  $\dot{H}^1(\mathbb{R})$ .

# C. POCS algorithm

After forming the generalized samples  $s_k := \Delta x(t_k)$  from the point samples of x, we have shown in (82) that the generic sampling form of (1) is achieved with  $\mathcal{H} = \dot{H}^1(\mathbb{R})$ ,  $\mathcal{A} = \mathcal{B}$ and the family  $(g_k)_{k \in \mathbb{Z}}$  defined in (79) which is orthogonal in  $\mathcal{H}$ . All the conditions of Section III are thus realized. The signal x can then be estimated by iterating

$$u^{(n+1)} = P_{\mathcal{A}} P_{\mathcal{H}_{\mathbf{s}}} u^{(n)}, \qquad n \ge 0 \tag{84}$$

which we have simply repeated from (24) for convenience, where  $\mathcal{H}_{\mathbf{s}}$  is defined in (7) and  $\mathbf{s} := (\Delta x(t_k))_{k \in \mathbb{Z}}$ .

Proposition 8.2: With  $\mathbf{s} := (\Delta x(t_k))_{k \in \mathbb{Z}}$ ,  $\mathcal{H}_{\mathbf{s}}$  is the set of functions  $u(t) \in \dot{H}^1(\mathbb{R})$  such that  $u(t_k) - x(t_k)$  is constant for  $k \in \mathbb{Z}$ .

*Proof:* From (81),  $\langle u, g_k \rangle = \Delta u(t_k)$  for any  $u \in H^1(\mathbb{R})$ . So, from (7),  $u \in \mathcal{H}_s$  if and only if  $\Delta u(t_k) = s_k = \Delta x(t_k)$  for all  $k \in \mathbb{Z}$ . This is equivalent to  $u(t_k) - x(t_k) = u(t_{k-1}) - x(t_{k-1})$  for all  $k \in \mathbb{Z}$ . This proves the proposition.

By applying the POCS results of Section III-C to (84), we obtain the following result.

Proposition 8.3: Assume that  $(t_k)_{k\in\mathbb{Z}}$  is any increasing sequence of time instants. Let  $(u^{(n)})_{n\geq 0}$  be recursively defined by (84) starting from some  $u^{(0)} \in \mathcal{A}$ , and  $u^{(\infty)}$  be the function of  $\mathcal{A}$  that interpolates the points  $(t_k, x(t_k))_{k\in\mathbb{Z}}$  while minimizing  $\|(u^{(\infty)} - u^{(0)})'\|_2$ . Then,  $\|(u^{(n)} - u^{(\infty)})'\|_2$  monotonically tends to 0 with n.

**Proof:** Given that  $\mathcal{H} = \dot{H}^1(\mathbb{R})$  with the norm defined in (83), we know from Section III-C that  $||u^{(n)} - \hat{u}|| =$  $||(u^{(n)} - \hat{u})'||_2$  monotonically tends to 0 with  $\hat{u} := P_{\mathcal{A}_s} u^{(0)}$ . Since  $\hat{u} \in \mathcal{A}_s = \mathcal{A} \cap \mathcal{H}_s$ , we know by Proposition 8.2 that  $\hat{u}$ is an element of  $\mathcal{A}$  that interpolates the points  $(t_k, x(t_k))_{k \in \mathbb{Z}}$ up to a constant component. But since  $\hat{u} = P_{\mathcal{A}_s} u^{(0)}$ , it also minimizes  $||\hat{u} - \hat{u}^{(0)}|| = ||(\hat{u} - \hat{u}^{(0)})'||_2$ . Thus  $\hat{u}$  and  $u^{(\infty)}$  differ by just a constant. Then,  $\hat{u}$  can be replaced by  $u^{(\infty)}$  in all the above norm expressions.

#### D. Conincidence with Grochenig's algorithm

As  $\mathbf{s} = (\Delta x(t_k))_{k \in \mathbb{Z}}$  is equivalent to  $x \in \mathcal{H}_s$ , we recall from Proposition 2.3 that

$$P_{\mathcal{H}_{\mathbf{s}}}u := u + P_{\mathcal{G}}(x - u) \tag{85}$$

where  $\mathcal{G}$  is the closed linear span of  $(g_k)_{k \in \mathbb{Z}}$ . Given the definition of this family in (79),  $P_{\mathcal{G}}$  is characterized as follows.

Proposition 8.4: For any  $u \in \dot{H}^1(\mathbb{R})$ ,  $P_{\mathcal{G}}u$  is the function that linearly interpolates the points  $(t_k, u(t_k))_{k \in \mathbb{Z}}$  up to a constant component.

*Proof:* Let  $u \in \dot{H}^1(\mathbb{R})$ . Since  $(g_k)_{k \in \mathbb{Z}}$  is an orthogonal basis of  $\mathcal{G}$ , then

$$P_{\mathcal{G}}u = \sum_{k \in \mathsf{Z}} \frac{\langle u, g_k \rangle}{\|g_k\|^2} g_k(t) = \sum_{k \in \mathsf{Z}} \frac{\Delta u(t_k)}{\Delta t_k} g_k(t) \tag{86}$$

using (81) and the result  $||g_k||^2 = ||1_{[t_{k-1},t_k]}||_2^2 = \Delta t_k$ . Let  $\hat{u}$  be the linear interpolant of the points  $(t_k, u(t_k))_{k \in \mathbb{Z}}$ . For each  $k \in \mathbb{Z}$  and every  $t \in (t_{k-1}, t_k)$ , it is easy to see that  $u'(t) = \Delta u(t_k) / \Delta t_k = \hat{u}'(t)$ . As u and  $\hat{u}$  are both absolutely continuous, then  $u - \hat{u}$  is a constant function.

Grochenig previously introduced in [4, \$4.1] the following iteration<sup>3</sup>

$$u^{(n+1)} = P_{\mathcal{B}}\left(u^{(n)} + L(x - u^{(n)})\right), \qquad n \ge 0 \qquad (87)$$

where  $u^{(n)}$  is in the subspace  $\mathcal{B} \subset L^2(\mathbb{R})$  of bandlimited functions, which we assume of Nyquist period 1, and Luis the exact linear interpolation of the points  $(t_k, u(t_k))_{k \in \mathbb{Z}}$ . When thinking of  $u^{(n)}$  as elements of  $\mathcal{A} \subset \dot{H}^1(\mathbb{R})$ , then (87) coincides with (84) given (85).

#### E. Analysis comparison

It is interesting to compare the convergence analysis of (87) from [4] with the result of Proposition 8.3. Under the condition that

$$\lim_{k \to \pm \infty} t_k = \pm \infty \quad \text{and} \quad \Delta < 1, \tag{88}$$

where  $\Delta := \sup_{k \in \mathbb{Z}} \Delta t_k$ , it was shown in [4] that the transformation of (87) is a contraction with respect to the  $L^2$ -norm. As x is a fixed point of (87), this proves that  $u^{(n)}$  linearly converges in  $L^2$ -norm to x. Meanwhile, Proposition 8.3 analyzes the convergence of  $u^{(n)}$  in terms of the  $L^2$ -norm of its derivative. A shortcoming is the loss of information on its constant component. However, this proposition contains a lot more results, as follows.

Note first that none of the conditions of (88) are assumed in Proposition 8.3. In the case where  $t_k$  has a finite limit  $t_{\infty}$ (resp.  $t_{-\infty}$ ) when t goes to  $\infty$  (resp.  $-\infty$ ), then we just need to assume that Lu is constant and equal to  $u(t_{\infty})$  (resp.  $u(t_{-\infty})$ ) in  $[t_{\infty}, \infty)$  (resp.  $(-\infty, t_{\infty}]$ ) to be consistent with  $P_{\mathcal{G}}u$  as a result of (86) with (79). We can even include the case where Z is a finite set  $\{1, \dots, N\}$ , which corresponds to the case of N+1 sampling instants  $t_0, \dots, t_N$  ( $t_0$  and  $t_N$  playing the roles of  $t_{-\infty}$  and  $t_{\infty}$  in the construction of L).

Next, we can see that perfect reconstruction is achieved (up to a constant component) whenever the samples uniquely define x(t) as a bandlimited signal. There are known cases where this is realized without the constraint that  $\Delta < 1$ . For example, uniqueness of reconstruction is known from [1] to be realized when  $(|t_k - kT|)_{k \in \mathbb{Z}}$  is upper bounded for some  $T \in [0, 1)$  and  $\Delta t_k$  has a positive lower bound. This condition can lead to arbitrarily large  $\Delta$ .

When consistent reconstruction is not unique, Proposition 8.3 implies that (87) is still convergent (up to a constant component) and the limit  $u^{(\infty)}$  is the bandlimited interpolator that is closest to  $u^{(0)}$  in terms of the norm of (83). If  $u^{(0)} = 0$ ,  $u^{(\infty)}$  can be interpreted as the bandlimited interpolator of minimum slope energy. But choosing a nonzero function  $u^{(0)}$  can be useful to "attract" the interpolator  $u^{(\infty)}$  towards some signal guess obtained by other means (see experiment in Section VIII-G).

On top the uncertainty on constant component, a shortcoming of Proposition 8.3 is the absence of linear convergence. This would require to know the condition on  $(t_k)_{k\in\mathbb{Z}}$  for  $\operatorname{ran}(S)$  to be closed. This is a difficult problem that cannot be resolved here. However, we know by default that linear

<sup>&</sup>lt;sup>3</sup>This iteration appears in eq.(24) of [4] in the equivalent form of  $u^{(n+1)} = u^{(n)} + P_{\mathcal{B}}P_{\mathcal{G}}(x-u^{(n)})$  since  $u^{(n)} \in \mathcal{B}$ .

convergence is achieved whenever Z is finite, which is the case of interest in practice. Finally, our framework includes the additional option of relaxation and insight on the POCS limit under sampling noise.

The lack of knowledge on the constant components is mostly due to the analysis in  $\dot{H}^1(\mathbb{R})$  which cannot incorporate the fact that L performs in (87) an exact interpolation. It is actually possible to combine the consequences of this analysis with the latter fact and eliminate the constant-component uncertainties. As this is beyond the main scope of the present paper, this will be reserved for a future publication.

#### F. Numerical experiments in oversampling situation

We plot in Fig. 3 the MSE performance of a number of iterative bandlimited reconstruction algorithms from nonuniform point samples of similar complexity per iteration, including the original frame algorithm introduced in the pioneering paper on nonuniform sampling [1], the basic Kaczmarz method and its randomized version presented in Section II-E, and Grochenig's algorithm without and with relaxation. For each algorithm, the MSE of the *n*th iterate  $u^{(n)}$ , reported in solid lines, is measured by averaging the relative error  $||u^{(n)} - x||_2^2/||x||_2^2$  over 100 randomly generated bandlimited inputs x that are periodic of period 315 (assuming a Nyquist period 1). Even though our analysis of Grochenig's algorithm has been constructed in the Sobolev space  $H^1(\mathbb{R})$ , we have maintained the  $L^2$ -norm in the error measurements as this is the standard reference of MSE in signal processing. We have however superimposed in mixed lines the MSE obtained by averaging  $||u^{(n)} - x||^2 / ||x||^2$ where  $\|\cdot\|$  is the norm of  $\dot{H}^1(\mathbb{R})$  given in (83), specifically for the results of Grochenig's algorithm in curves (iv) and (v). Even though the two norms are not equivalent, we observe that they yield similar results. So, while we do not provide analytical justifications for this similarity, we see that the Sobolev norm remains an adequate tool of error predictions in these experiments. In (v), the relaxation coefficient has been optimized empirically.

In Fig. 3(a,b), we compare the behavior of the algorithms with respect to two types of sampling nonuniformity. In (a),  $(\Delta t_k)_{k \in \mathbb{Z}}$  is generated as an i.i.d sequence that is uniformly distributed in [0.3, 1] (in Nyquist period unit), leading to an oversampling ratio of about 1.54. Meanwhile, the sampling instants in (b) are grouped into clusters that are nonuniformly spaced between each other. Each cluster is made of 3 sampling instants equally spaced by 1/4 (as opposed to 1 in Nyquist rate sampling), and the overall density of of clusters is such that the oversampling ratio is 2. The two figures show that the randomized Kaczmarz method appears to be well tuned for the random nonuniformity of (a), but not at all for the clustered type of nonuniformity in (b), where it barely improves the basic cyclic version of the Kaczmarz method. Meanwhile, Grochenig's algorithm shows its systematic superiority to the cyclic Kaczmarz method in MSE. This is particularly remarkable for the clustered type of nonuniformity, which is known to be a challenge for bandlimited interpolation. Now, the inclusion of relaxation in (v) is also of outstanding impact: it substantially improves the unrelaxed version in a way that could not have been predicted or justified in the the general framework of contractions, like in the analysis of [4].

Fig. 3 (c) goes back to a nonuniformity of the type of (a) at however a higher oversampling ratio to highlight the noise-shaping effect of the POCS algorithm in the presence of sampling errors. Here, the sequence  $(\Delta t_k)_{k\in\mathbb{Z}}$  is uniformly distributed in [0, 0.5], and the sample errors are Gaussian random variables that are 45 dB's below the input in variance. While the randomized Kaczmarz method exhibits the same type of fast convergence as in (a), it shows inferior capabilities of noise filtering compared to Grochenig's method. This is a consequence of the pseudo-inversion property of the latter method when formalized as a POCS algorithm. As an extra result, the figure shows the particularly poor behavior of the cyclic Kaczmarz method to high oversampling.

#### G. Numerical experiments in sub-Nyquist situation

We show in Fig. 4 an example of POCS iteration limit in a case of sub-Nyquist sampling. In this experiment, the samples of the bandlimited input x(t) are obtained by levelcrossing sampling [6], [8], [9], that is, from the crossings of x(t) with fixed levels represented by horizontal grey lines [6], [8], [9]. The resulting sampling ratio is 0.77 (the time unit is the Nyquist period of x(t)), which prevents uniqueness of reconstruction. While the blue curves result from a zero initial estimate  $u^{(0)}(t)$ , the red curves are obtained by choosing for  $u^{(0)}(t)$  the bandlimited version of the piecewise constant function shown in black dotted curve. This stair case function can be generated from the mere knowledge of the level crossings. In each of the two cases of initial estimate, the result of 3 iteration numbers is plotted, using the following line types in increasing order of iteration: dotted line, mixed line, solid line. The exact iteration numbers of the plots are indicated in the legend. To validate our convergence analysis, we have actually plotted the result of infinite iteration by employing the theoretical limit formula of (57) and computing  $S^{\dagger}$  by matrix pseudo-inverse, given the low dimensionality of the experiment. The figure shows the good convergence of  $u^{(n)}(t)$  to this ideal limit. Now, the main point of this figure is to show an example of reconstruction improvement using a heuristically designed nonzero initial estimate  $u^{(0)}(t)$ .

## IX. SUMMARY

We have introduced in this paper an abstract framework where a bandlimited input signal can be reconstructed/estimated from nonuniform generalized samples by pseudo-inversion of the sampling operation, using an algorithm that consists of iterated time-varying filters. This requires that the sampling kernel functions be orthogonal at least in a Hilbert space that is larger than the input space. We prove in this paper the full pseudo-inversion properties of our algorithm. They include perfect reconstruction whenever the samples uniquely define the input, minimum-norm reconstruction when the sampling is insufficient, and a noise-shaping effect on sampling errors. While the required condition on the sampling kernels was previously observed in the time encoder of Lazar and Tóth, this condition is shown in this



Fig. 3. MSE results of iterative bandlimited reconstructions from nonuniform point samples at randomly generated instants  $(t_k)_{k \in \mathbb{Z}}$  with the following statistics: (a) uniform distribution of  $\Delta t_k$  in [0.3, 1] (1.3 relaxation for (v)); (b) randomly positioned clusters of sampling instants (1.45 relaxation for (v)); (c) uniform distribution of  $\Delta t_k$  in [0, 0.5] with sampling noise (1.05 relaxation for (v)). The solid lines correspond to MSE based on the  $L^2$ -norm (as reported in the ordinate description), while the mixed lines correspond to MSE based the Sobolev norm.



Fig. 4. Estimation of bandlimited input x(t) from level crossing samples in sub-Nyquist situation (sampling ratio of 0.77).

paper to appear in a non-trivial manner in a recent multichannel time-encoding system as well as in traditional point sampling, Our resulting algorithms turn out to coincide with existing reconstruction methods in these cases, But our framework reveals the pseudo-inversion properties of these methods, while proposing efficient discrete-time implementations.

## APPENDIX

#### A. Proof of Theorem 3.3

For any closed affine subspace  $\mathcal{C}$  and  $\lambda \in \mathbb{R}$ ,  $P_{\mathcal{C}}^{\lambda}u - u = \lambda(P_{\mathcal{C}}u - u) \in \overrightarrow{\mathcal{C}}^{\perp}$  where  $\overrightarrow{\mathcal{C}}$  is the linear subspace associated with  $\mathcal{C}$ . Since  $P_{\mathcal{H}_{\mathbf{s}}}^2 u - u = 2(P_{\mathcal{H}_{\mathbf{s}}}u - u)$ ,  $P_{\mathcal{H}_{\mathbf{s}}}^2$  is the symmetry with respect to the affine space  $\mathcal{H}_{\mathbf{s}}$ , and hence is non-expansive. Using the relation,  $P_{\mathcal{C}}^{\lambda}u - u = \frac{\lambda}{2}(P_{\mathcal{C}}^2u - u)$ , we have  $P_{\mathcal{A}}P_{\mathcal{C}}^{\lambda}u - u = \frac{\lambda}{2}(P_{\mathcal{A}}P_{\mathcal{C}}^2u - u)$  for all  $u \in \mathcal{A}$ . By

applying this with  $\mathcal{C} = \mathcal{H}_{s}$ ,  $\lambda = \lambda_{n}$  and  $u = u^{(n)}$ , the iterate  $u^{(n)}$  of (28) satisfies the recursion

$$u^{(n+1)} = u^{(n)} + \frac{\lambda_n}{2} (Ru^{(n)} - u^{(n)}), \qquad n \ge 0$$

where  $R := P_{\mathcal{A}} P_{\mathcal{H}_{s}}^{2}$ . This operator is non-expansive and the set of its fixed points can be verified to be  $\mathcal{A} \cap \mathcal{H}_{s} = \mathcal{A}_{s}$ . As  $\sum_{n \geq 0} \frac{\lambda_{n}}{2} (1 - \frac{\lambda_{n}}{2}) = +\infty$  with  $\frac{\lambda_{n}}{2} \in [0, 1]$ , we can apply Theorems 5.14 and 5.13 of [33], and conclude that  $u^{(n)}$  has a limit  $u^{(\infty)} \in \mathcal{A}_{s}$ . We have  $P_{\mathcal{H}_{s}}^{\lambda_{n}} u^{(n)} - u^{(n)} \in \mathcal{H}_{s}^{\perp}$  while  $P_{\mathcal{H}_{s}}^{\lambda_{n}} u^{(n)} - P_{\mathcal{H}_{s}}^{\lambda_{n}} u^{(n)} \in \mathcal{A}^{\perp}$ . Therefore,  $u^{(n+1)} - u^{(n)} \in \mathcal{A}_{s}^{\perp}$ . This proves that  $u^{(\infty)} = P_{\mathcal{A}_{s}} u^{(0)}$ .

## B. Proof of Proposition 4.3

(i)  $\Leftrightarrow$  (ii): Given (29), it is known from [41] that ran(S) is closed in  $\mathcal{D}$  if and only if there exists a constant  $\alpha > 0$  such that  $||Su||_{\mathcal{D}} \ge \alpha ||u||$  for all  $u \in \operatorname{null}(S)^{\perp}$ . This amounts to (ii) with the identity  $\operatorname{null}(S)^{\perp} = \mathcal{F}$  from (34).

(ii)  $\Leftrightarrow$  (iii): By definition,  $(f_k)_{k \in \mathbb{Z}}$  is a frame of  $\mathcal{F}$  if and only if there exist constants  $0 < A \leq B$  such that

$$\forall u \in \mathcal{F}, \qquad A \|u\|^2 \le \sum_{k \in \mathsf{Z}} |\langle u, f_k \rangle|^2 \le B \|u\|^2 \tag{89}$$

Now, it follows from (18) and (38) that

$$\forall u \in \mathcal{A}, \quad \sum_{k \in \mathsf{Z}} |\langle u, f_k \rangle|^2 = \sum_{k \in \mathsf{Z}} \frac{|\langle u, \tilde{g}_k \rangle|^2}{\|g_k\|^2} = \|Su\|_{\mathcal{D}}^2. \tag{90}$$

We see from (38) that the upper bound of (89) is by default satisfied with B := 1. Meanwhile, the lower bound A > 0 exists if and only if (ii) is satisfied.

# C. Proof of Proposition 7.1

*Lemma A.1:* Let  $u(t) := u(t) \mathbf{u}$  where  $u(t) \in L^2(\mathbb{R})$  and  $\mathbf{u} \in \mathbb{R}^M$ . If  $u(t) \in \mathcal{B}^{\perp}$  or  $\mathbf{u} \in \operatorname{ran}(\mathbf{A})^{\perp}$ , then  $u(t) \in \mathcal{A}^{\perp}$ .

*Proof:* Let 
$$\boldsymbol{v}(t) \in \mathcal{A}$$
. It follows from (67) that  
 $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \langle u(t)\mathbf{u}, \boldsymbol{v}(t) \rangle = \langle u(t), \mathbf{u}^{\top} \boldsymbol{v}(t) \rangle_{2}$ .

Clearly,  $\mathbf{u}^{\top} \boldsymbol{v}(t) \in \mathcal{B}$ . So if  $u(t) \in \mathcal{B}^{\perp}$ , then  $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = 0$ . Meanwhile, if  $\mathbf{u} \in \operatorname{ran}(\mathbf{A})^{\perp}$ , then  $\mathbf{u}^{\top} \boldsymbol{v}(t) = 0$  for each single  $t \in \mathbb{R}$ . Then  $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = 0$  regardless of u(t). This proves the lemma.

We now proceed with the proof of Proposition 7.1. It will be convenient to define  $\mathbf{P} := \mathbf{A}\mathbf{A}^+$ . Let  $\boldsymbol{w}(t) := \tilde{u}(t)\mathbf{P}\mathbf{u}$ . Its *i*th component is  $w^i(t) = \tilde{u}(t)q^i \in \mathcal{B}$ , where  $q^i$  is the *i*th coordinate of  $\mathbf{P}\mathbf{u}$ . So  $\boldsymbol{w}(t) \in \mathcal{B}^M$ . Meanwhile,  $\mathbf{P}\mathbf{u} \in \operatorname{ran}(\mathbf{A})$ , so  $\boldsymbol{w}(t) \in \operatorname{ran}(\mathbf{A})$  for each  $t \in \mathbb{R}$ . Then,  $\boldsymbol{w}(t) \in \mathcal{A}$ . Next, we can write

$$\boldsymbol{u}(t) - \boldsymbol{w}(t) = (\boldsymbol{u}(t) - \tilde{\boldsymbol{u}}(t)) \boldsymbol{u} + \tilde{\boldsymbol{u}}(t) (\boldsymbol{u} - \boldsymbol{P}\boldsymbol{u}).$$

While  $u(t) - \tilde{u}(t) \in \mathcal{B}^{\perp}$ ,  $\mathbf{u} - \mathbf{Pu} \in \operatorname{ran}(\mathbf{A})^{\perp}$  because  $\mathbf{P}$  is known to be the orthogonal projection of  $\mathbb{R}^{M}$  onto  $\operatorname{ran}(\mathbf{A})$ . So  $u(t) - w(t) \in \mathcal{A}^{\perp}$  according to the above lemma. Thus,  $w(t) = P_{\mathcal{A}} u(t)$ .

#### ACKNOWLEDGMENT

The authors would like to thank Eva Kopecká and Patrick Combettes for their indications on the convergence of relaxed POCS, and Sinan Güntürk for his help on Sobolev spaces.

#### REFERENCES

- R. Duffin and A. Schaeffer, "A class of nonharmonic Fourier series," *Transactions of the American Mathematical Society*, vol. 72, pp. 341– 366, Mar. 1952.
- [2] J. L. Yen, "On nonuniform sampling of bandwidth-limited signals," *IRE Trans. Circ. Theory*, vol. CT-3, pp. 251–257, Dec. 1956.
- [3] J. Benedetto, "Irregular sampling and frames," in *Wavelets: A Tutorial in Theory and Applications* (C. K. Chui, ed.), pp. 445–507, Boston, MA: Academic Press, 1992.
- [4] K. Gröchenig, "Reconstruction algorithms in irregular sampling," *Math. Comp.*, vol. 59, no. 199, pp. 181–194, 1992.
- [5] H. G. Feichtinger and K. Gröchenig, "Theory and practice of irregular sampling," in *Wavelets: Mathematics and Applications* (J. Benedetto, ed.), pp. 318–324, Boca Raton, FL: CRC Press, 1994.
- [6] F. Marvasti, Nonuniform Sampling: Theory and Practice. New York: Kluwer, 2001.
- [7] J. J. Benedetto and P. J. Ferreira, Modern sampling theory: mathematics and applications. Springer Science & Business Media, 2001.
- [8] M. Miśkowicz, ed., Event-Based Control and Signal Processing. Embedded Systems, Boca Raton, FL: CRC Press, 2018.
- [9] D. Rzepka, M. Miśkowicz, D. Kościelnik, and N. T. Thao, "Reconstruction of signals from level-crossing samples using implicit information," *IEEE Access*, vol. 6, pp. 35001–35011, 2018.
- [10] R. Alexandru and P. L. Dragotti, "Reconstructing classes of nonbandlimited signals from time encoded information," *IEEE Transactions* on Signal Processing, vol. 68, pp. 747–763, 2020.
- [11] K. Adam, A. Scholefield, and M. Vetterli, "Asynchrony increases efficiency: Time encoding of videos and low-rank signals," *IEEE Transactions on Signal Processing*, pp. 1–1, 2021.
- [12] D. Florescu and D. Coca, "A Novel Reconstruction Framework for Time-Encoded Signals with Integrate-and-Fire Neurons," *Neural Computation*, vol. 27, pp. 1872–1898, 09 2015.
- [13] D. Gontier and M. Vetterli, "Sampling based on timing: Time encoding machines on shift-invariant subspaces," *Applied and Computational Harmonic Analysis*, vol. 36, no. 1, pp. 63–78, 2014.
- [14] S. Rudresh, A. J. Kamath, and C. Sekhar Seelamantula, "A time-based sampling framework for finite-rate-of-innovation signals," in *ICASSP* 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 5585–5589, 2020.
- [15] J. Szyduczyński, D. Kościelnik, and M. Miśkowicz, "Time-to-digital conversion techniques: a survey of recent developments," *Measurement*, vol. 214, p. 112762, 2023.
- [16] A. Lazar and L. T. Tóth, "Perfect recovery and sensitivity analysis of time encoded bandlimited signals," *IEEE Trans. Circ. and Syst.-I*, vol. 51, pp. 2060–2073, Oct. 2004.

- [17] Y. C. Eldar and T. Werther, "General framework for consistent sampling in Hilbert spaces," *International Journal of Wavelets, Multiresolution* and Information Processing, vol. 3, no. 04, pp. 497–509, 2005.
- [18] D. Wei and J. G. Harris, "Signal reconstruction from spiking neuron models," in 2004 IEEE International Symposium on Circuits and Systems (ISCAS), vol. 5, pp. V–V, IEEE, 2004.
- [19] A. A. Lazar, E. K. Simonyi, and L. T. Toth, "An overcomplete stitching algorithm for time decoding machines," *IEEE Transactions on Circuits* and Systems I: Regular Papers, vol. 55, pp. 2619–2630, Oct 2008.
- [20] N. T. Thao and D. Rzepka, "Time encoding of bandlimited signals: Reconstruction by pseudo-inversion and time-varying multiplierless FIR filtering," *IEEE Transactions on Signal Processing*, vol. 69, pp. 341–356, 2021.
- [21] N. T. Thao, D. Rzepka, and M. Miśkowicz, "Bandlimited signal reconstruction from leaky integrate-and-fire encoding using POCS," *IEEE Transactions on Signal Processing*, vol. 71, pp. 1464–1479, 2023.
- [22] P. L. Combettes, "The foundations of set theoretic estimation," Proceedings of the IEEE, vol. 81, pp. 182–208, Feb 1993.
- [23] H. H. Bauschke and J. M. Borwein, "On projection algorithms for solving convex feasibility problems," *SIAM Rev.*, vol. 38, no. 3, pp. 367– 426, 1996.
- [24] K. Adam, A. Scholefield, and M. Vetterli, "Sampling and reconstruction of bandlimited signals with multi-channel time encoding," *IEEE Transactions on Signal Processing*, vol. 68, pp. 1105–1119, 2020.
- [25] L. Grafakos, Modern Fourier Analysis. Graduate Texts in Mathematics, Springer New York, 2014.
- [26] R. DeVore and G. Lorentz, *Constructive Approximation*. Grundlehren der mathematischen Wissenschaften, Springer Berlin Heidelberg, 1993.
- [27] S. Norsworthy, R. Schreier, G. Temes, and I. C. S. Society, *Delta-Sigma Data Converters: Theory, Design, and Simulation.* Wiley, 1997.
- [28] T. G. Dvorkind and Y. C. Eldar, "Robust and consistent sampling," *IEEE Signal Processing Letters*, vol. 16, no. 9, pp. 739–742, 2009.
- [29] H. Choi and D. C. Munson, "Analysis and design of minimax-optimal interpolators," *IEEE Transactions on Signal Processing*, vol. 46, pp. 1571– 1579, Jun 1998.
- [30] S.-J. Yeh and H. Stark, "Iterative and one-step reconstruction from nonuniform samples by convex projections," J. Opt. Soc. Am. A, vol. 7, pp. 491–499, Mar 1990.
- [31] T. Strohmer and R. Vershynin, "A randomized Kaczmarz algorithm with exponential convergence," *Journal of Fourier Analysis and Applications*, vol. 15, no. 2, pp. 262–278, 2009.
- [32] K. Adam, A. Scholefield, and M. Vetterli, "Encoding and decoding mixed bandlimited signals using spiking integrate-and-fire neurons," in ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 9264–9268, 2020.
- [33] H. Bauschke and P. Combettes, Convex Analysis and Monotone Operator Theory in Hilbert Spaces. CMS Books in Mathematics, Springer International Publishing, 2017.
- [34] D. G. Luenberger, Optimization by vector space methods. John Wiley & Sons, Inc., New York-London-Sydney, 1969.
- [35] G. Wang, Y. Wei, and S. Qiao, *Generalized Inverses: Theory and Computations*. Developments in Mathematics, Springer Singapore, 2018.
- [36] O. Christensen, *Frames and bases*. Applied and Numerical Harmonic Analysis, Birkhäuser Boston, Inc., Boston, MA, 2008. An introductory course.
- [37] Y. C. Eldar, Sampling theory: Beyond bandlimited systems. Cambridge University Press, 2015.
- [38] J. B. Conway, A course in functional analysis, vol. 96 of Graduate Texts in Mathematics. Springer-Verlag, New York, second ed., 1990.
- [39] D. Smart, Fixed Point Theorems. Cambridge Tracts in Mathematics, Cambridge, UK: Cambridge University Press, 1980.
- [40] S. Berberian, A First Course in Real Analysis. Undergraduate Texts in Mathematics, Springer New York, 2012.
- [41] T. Kato, Perturbation theory for linear operators. Classics in Mathematics, Springer-Verlag, Berlin, 1995. Reprint of the 1980 edition.