Sampling and reconstruction of operators

Götz E. Pfander, *Member, IEEE,* and David Walnut

Abstract

We study the recovery of operators with bandlimited Kohn-Nirenberg symbol from the action of such operators on a weighted impulse train, a procedure we refer to as operator sampling. In previous work, Kozek and the authors have shown that operator sampling is possible if the symbol of the operator is bandlimited to a set with area less than one. In this paper we develop explicit reconstruction formulas for operator sampling that generalize reconstruction formulas for bandlimited functions. We give necessary and sufficient conditions on the sampling rate that depend on size and geometry of the bandlimiting set. Moreover, we show that under mild geometric conditions, classes of operators bandlimited to an unknown set of area less than one-half permit sampling and reconstruction. A similar result considering unknown sets of area less than one was independently achieved by Heckel and Boelcskei.

Operators with bandlimited symbols have been used to model doubly dispersive communication channels with slowly-time-varying impulse response. The results in this paper are rooted in work by Bello and Kailath in the 1960s.

Index Terms

Bandlimined Kohn-Nirenberg symbols, spreading function, operator Paley-Wiener space, channel measurement, channel identification, operator identification, operator sampling, Gabor analysis, symplectic matrices.

CONTENTS

Manuscript received XXX; revised YYY. This work was supported by the *Deutsche Forschungsgemeinschaft* (DFG) under grant 50292 DFG PF-4 Sampling Operators.

G.E. Pfander is with Jacobs University, Germany and D. Walnut is with George Mason University.

I. INTRODUCTION

In this paper we develop a sampling theory and reconstruction formulas for operators bandlimited to domains of small area. Analogously to the classical sampling theory of functions, the objective of operator sampling is to fully characterize an operator from at first sight insufficient information, specifically by observing an operator's action on a single discretely supported distribution, viz; a weighted sum of delta distributions. The theory developed herein applies to so-called bandlimited operators, defined as operators whose Kohn-Nirenberg symbol is bandlimited. The symplectic Fourier transform of the Kohn-Nirenberg symbol of an operator is referred to as its spreading function, so that we are considering operators whose spreading function is compactly supported. More generally, we extend reconstruction to operators in which the spreading function is supported in a fundamental domain of a lattice. In engineering terms, the operators considered are characterized by limited time-frequency dispersion.

A. Identification and sampling of operators

The operator identification problem addresses the question whether an operator from a given class can be recovered from its action on a single probing signal. That is, for a given class of operators $\mathcal{H}$, does there exist an input signal g so that Hg determines H. Mathematically speaking, we require that the map $\Phi_g : \ H \mapsto Hg$ be injective on $\mathcal{H}$. In order to be stable under noise introduced, for example, by physical considerations or digital processing, it is reasonable to require in addition that the map Φ_g have a bounded inverse.

Definition 1.1: Let $\mathcal{H}$ be a collection of linear operators mapping a space of functions or distributions $X(\mathbb{R})$ to a normed function space $Y(\mathbb{R})$. If for some $g \in X(\mathbb{R})$,

$$\Phi_g : \mathcal{H} \longrightarrow Y(\mathbb{R}),\ H \mapsto Hg$$

is bounded above and below, that is, if there are constants $0 < A \leq B$ such that

$$A\|H\|_{\mathcal{H}} \leq \|Hg\|_Y \leq B\,\|H\|_{\mathcal{H}} \text{ for all } H \in \mathcal{H}, \tag{1}$$

then we say that $\mathcal{H}$ is *identifiable with identifier* $g \in X(\mathbb{R})$. If $\mathcal{H}$ is not linear, then condition (1) is replaced by

$$A\|H_1 - H_2\|_{\mathcal{H}} \leq \|H_1 g - H_2 g\|_Y \leq B\,\|H_1 - H_2\|_{\mathcal{H}} \text{ for all } H_1, H_2 \in \mathcal{H}. \tag{2}$$

The sampling and reconstruction theory for operators developed here addresses identifiability of operator classes utilizing discretely supported distributions.

Definition 1.2: A strictly increasing sequence $\Lambda = \{\lambda_n\}_{n\in\mathbb{Z}}$ in $\mathbb{R}$ is a *set of sampling* for an operator class $\mathcal{H}$, if for some never-vanishing sequence $(c_n)_{n\in\mathbb{Z}}$, we have that $\sum_{n\in\mathbb{Z}} c_n \delta_{\lambda_n}$ identifies $\mathcal{H}$. We define the *sampling rate* of Λ by

$$D(\Lambda) = \lim_{r\to\infty} \frac{n^-(r)}{r}$$

where

$$n^-(r) = \inf_{x\in\mathbb{R}} \#\{n\colon \lambda_n \in [x, x+r]\}$$

Assuming that the limit exists, $D(\Lambda)$ can be interpreted as the average number of terms $c_n\delta_{\lambda_n}$ appearing in the identifier per unit time and corresponds to the *lower Beurling density* of Λ. The assumption that the sequence $c = (c_n)$ never vanishes ensures that the sampling rate depends only on Λ. In particular, we avoid the situation in which for some set $\Lambda' \supseteq \Lambda$, of higher density than Λ, $\sum_m d_m\delta_{\lambda'_m} = \sum_n c_n\lambda_n$ where $d_m = c_n$ whenever $\lambda'_m = \lambda_n$ and $d_m = 0$ otherwise.

In this paper, we will restrict our attention to sets of sampling that are periodic subsets of a lattice in $\mathbb{R}$, and moreover will focus on periodic weighting sequences $c = (c_n)$.

Definition 1.3: We say that an operator class $\mathcal{H}$ can be identified by *regular operator sampling* if there exists $T > 0$, $L \in \mathbb{N}$, and a period-L sequence $c = (c_n)$ such that $\sum_{n\in\mathbb{Z}} c_n\delta_{nT}$ identifies $\mathcal{H}$. In the language of Definition 1.3,

$$\Lambda = \{nT:\ c_n \neq 0\} \subseteq T\,\mathbb{Z}.$$

Moreover,

$$D(\Lambda) = \frac{1}{T}\,\frac{\|c\|_0}{L} \tag{3}$$

where

$$\|c\|_0 = \#\{n\colon 0 \le n \le L{-}1 \text{ and } c_n \neq 0\}$$

is the support size of the vector $(c_0, \ldots, c_{L-1})$.

Our work addresses the identifiability of classes of operators characterized by their Kohn-Nirenberg symbol being bandlimited to a set S. In [10], [21] (cf. [22] and [17]), the following result for the identifiability of operator Paley-Wiener spaces (see Definition 1.5 below) is given. Here and in the following, $|S|$ denotes the Lebesgue measure of the set S.

Theorem 1.4: $OPW^2(S)$ is identifiable by regular operator sampling if S is compact and $|S| < 1$, and not identifiable if S is open and $|S| > 1$.

B. Operator representations, bandlimited operators, and operator Paley-Wiener spaces

Similarly to linear operators on finite dimensional space being represented by matrices, the Schwartz kernel theorem implies that linear operators on any of the classical function spaces on $\mathbb{R}$ can be represented by their kernel, that is, formally, we have

$$Hf(x) = \int \kappa_H(x,y)f(y)\,dy, \tag{4}$$

for a unique kernel κ_H.[1]

As operators are in 1-1 correspondence with their kernels, they can also be formally represented by their time-varying impulse response h, their Kohn-Nirenberg symbol σ, and their spreading function η.

[1] In fact, with $\mathcal{S}(\mathbb{R}^d)$ denoting the space of Schwartz class functions and $\mathcal{S}'(\mathbb{R}^d)$ its dual, we can associate to any linear and continuous operator mapping $\mathcal{S}(\mathbb{R}^d)$ to $\mathcal{S}'(\mathbb{R}^d)$ a kernel $\kappa \in \mathcal{S}'(\mathbb{R}^{2d})$ so that (4) holds in a weak sense. Below, we shall consider operators acting boundedly on the space of square integrable functions $L^2(\mathbb{R})$ which fall in the framework outlined above. We refer to [17] for a more detailed functional analytic treatment of operator and function spaces involved.

In fact, formally,

$$Hf(x) = \int h_H(x,t)\, f(x-t)\, dt \tag{5}$$

$$= \iint \eta_H(t,\nu)\, e^{2\pi i\nu(x-t)}\, f(x-t)\, d\nu\, dt \tag{6}$$

$$= \int \sigma_H(x,\xi)\, e^{2\pi i x\xi}\, \widehat{f}(\xi)\, d\xi, \tag{7}$$

where

$$\int \eta_H(t,\nu)\, e^{2\pi i\nu(x-t)}\, d\nu = h_H(x,t) = \kappa_H(x,x-t) = \int \sigma_H(x,\xi)\, e^{2\pi i\xi t}\, d\xi, \tag{8}$$

and the Fourier transform in (7) is normalized as $\mathcal{F}f(\xi) = \widehat{f}(\xi) = \int f(x)\, e^{-2\pi i x\xi}\, dx$.

Operator representations such as those given in (5), (6), (7) are considered in the theory of so-called pseudodifferential operators where we write

$$\sigma(x,D)f(x) = \int \sigma(x,\xi)\, e^{2\pi i x\xi}\, \widehat{f}(\xi)\, d\xi.$$

Observing further, that with the so-called symplectic Fourier transform given by

$$\mathcal{F}_s F(t,\nu) = \iint F(x,\xi)\, e^{-2\pi i(x\nu - t\xi)}\, dx\, d\xi,$$

(8) implies $e^{-2\pi i t\nu}\, \eta_H(t,\nu) = \mathcal{F}_s\sigma_H(t,\nu)$. We say that the operator

$$H \text{ is } \textit{bandlimited} \text{ to the set } S \subseteq \mathbb{R}^2 \text{ if } \operatorname{supp}\eta_H = \operatorname{supp}\mathcal{F}_s\sigma_H \subseteq S.$$

Considering now spaces of such operators we arrive at the following definition.

Definition 1.5: Given a set $S \subseteq \mathbb{R}^2$, define the *operator Paley-Wiener space* $OPW(S)$ by

$$OPW(S) = \{H \in \mathcal{L}(L^2(\mathbb{R}), L^2(\mathbb{R})) \colon \ \operatorname{supp}\mathcal{F}_s\sigma_H = \operatorname{supp}\eta_H \subseteq S\}$$

where $\mathcal{L}(L^2(\mathbb{R}), L^2(\mathbb{R}))$ denotes bounded operators on $L^2(\mathbb{R})$. The space of Hilbert-Schmidt operators in $OPW(S)$ is

$$OPW^2(S) = OPW(S) \cap HS(L^2(\mathbb{R})) = \{H \in \mathcal{L}(L^2(\mathbb{R}), L^2(\mathbb{R})) \colon \ \operatorname{supp}\mathcal{F}_s\sigma_H \subseteq S, \quad \sigma_H \in L^2(\mathbb{R}^2)\}.$$

The reconstruction formulas presented in this paper for $OPW^2(S)$ hold formally for all of $OPW(S)$. Operator Paley-Wiener spaces defined by membership of the symbol in generic mixed L^p spaces is considered in [17]; see also Section II-B below for some expamles.

C. Physical relevance of bandlimited operators

In communications engineering, (5) and (6) are commonly used as models for linear (time-varying) communication channels. The *time-varying impulse response* of the channel $h_H(x,t)$ is interpreted as the response of the channel at time x to a unit impulse at time $x-t$, that is, originating t time units earlier. Hence, if $h_H(x,t) \neq 0$ only for $0 \leq t \leq T$, then H is causal with maximum time-dispersion T.

If $h_H(x,t) = h_H(t)$ then the characteristics of the channel are time-invariant and in this case the channel is a convolution operator. Such channels are identifiable since $h_H(t)$ is the response of the channel to the input signal $\delta_0(t)$, the unit-impulse at $t=0$.

A mobile communication channel has the property that $h_H(x,t)$ depends on x, but changes as a function of x rather slowly, since the change in the channel, for example, by movement of receiver, transmitter, or reflecting objects, is slow when compared with the speed of light at which information travels. This slow variance can be expressed through a bandlimitation of $h_H(x,t)$ as a function of x, that is, as a support constraint on the *spreading function* of H, $\eta_H(t,\nu) = \int h_H(x,t)\, e^{-2\pi i\nu(x-t)}\, dx$, as a function of ν. We conclude that a causal doubly dispersive communications channel with maximum time dispersion T, and $h_H(x,t)$ bandlimited in x to $[-\Omega,\Omega]$ is represented by a spreading function supported on the set $[0,T]\times[-\Omega,\Omega]$, that is, by operators in $OPW([0,T]\times[-\Omega,\Omega])$ since $\eta_H = \mathcal{F}_s\sigma_H$.

To substantiate this bandlimitation on $\sigma_H(x,t)$ further, we denote translation by t by $T_t : f(x) \mapsto f(x-t)$ and modulation by ν by $M_\nu : f(x) \mapsto e^{2\pi i\nu x}\, f(x)$. The latter is also referred to as frequency shift as $\widehat{M_\nu f} = T_\nu\, \widehat{f}$. Then (6) becomes the operator-valued integral

$$H = \iint \eta_H(t,\nu)\, T_t\, M_\nu\, d\nu\, dt = \int_0^T \int_{-\Omega}^{\Omega} \eta_H(t,\nu)\, T_t\, M_\nu\, d\nu\, dt\,,$$

that is, the spreading function is the coefficient vector of the time-frequency shifts that a communication channel carries out. Hence, $OPW([0,T]\times[-\Omega,\Omega])$ has *maximum time-delay* T and *maximum frequency shift* Ω.

D. Relation to other work

In 1963, T. Kailath [7], [8], [9] asserted that for time-variant communication channels to be identifiable it is necessary and sufficient that the maximum time-delay, a, and Doppler shift, b, satisfy $ab \leq 1$ and gave an argument for this assertion based on counting degrees of freedom. In the argument, Kailath looks at the response of the channel to a train of impulses separated by at least a time units, so that in this sense the channel is being "sampled" by a succession of evenly-spaced impulse responses. The condition $ab \leq 1$ allows for the recovery of sufficiently many samples of $h_H(x,t)$ to determine it uniquely.

Kailath's conjecture was given the precise mathematical framework described above and proved in [10].

In 1969, P. A. Bello [2] argued that what is important for channel identification is not the product ab of the maximum time-delay and Doppler shift of the channel but the area of the support of the spreading function. It is notable that Kailath also asserted something along these lines. This means that a time-variant channel whose spreading function has essentially arbitrary support is identifiable as long as the area of that support is smaller than one. Using ideas from [10], Bello's conjecture was confirmed in [22].

Building on the results from [10], [21], [22] a number of results have been established that are now part of the herein described sampling theory for operators. For example, the results in [21] were extended from the setting of Hilbert-Schmidt operators to a much wider class of pseudodifferential operators in [17]. In [6], the choice of non periodic (irregular/jitter) sampling locations for operator sampling was discussed. Necessary and sufficient conditions for the identifiability of bandlimited Multiple Input Multiple Output (MIMO) channels were given in [16].

More recently, sampling results for stochastic operators, that is, for operators with stochastic spreading functions, have been obtained [15], [24], [23]. Also, in applications, it is required to replace the identifier considered in this paper by finite time, finite bandwidth, that is, smooth, signals. Local recovery results in this setting, as well as a reconstruction formula that allows for the application of coarse quantization methods prior to the approximate recovery of the operator are given in [12]. Focusing on a parametric setup, the identification of bandlimited operators was analyzed with respect to applicability in super-resolution radar [1].

II. Main results

A. Sampling and reconstruction of operators

One of the goals of this paper is to give an explicit reconstruction formula for the impulse response of the channel operator from its response to the identifier. Such formulas illustrate a connection between operator identification and classical sampling theory and motivates the terminology of *operator sampling* given above.

The main result of this paper is the following.

Theorem 2.1: Let $S \subseteq \mathbb{R}^2$ satisfy $|S| < 1$ and suppose that for some $\Omega, T > 0$ with $T\Omega = 1/L$, L prime, S is contained in a fundamental domain of $1/\Omega\,\mathbb{Z} \times 1/T\,\mathbb{Z}$ (that is, the sets $S + (k/\Omega, \ell/T)$, $k, \ell \in \mathbb{Z}$, are pairwise disjoint). Further assume that there exist integers $0 \le q_j,\ m_j \le L{-}1$, $0 \le j \le L{-}1$, and $(t_0, \nu_0) \in \mathbb{R}^2$ such that with $R_{q,m} = [0,T]{\times}[0,\Omega]{+}(t_0,\nu_0) + (qT, m\Omega)$, $q,\ m \in \mathbb{Z}$,

$$S_{per} = [0,1/\Omega]{\times}[0,1/T] + (t_0,\nu_0) \cap \bigcup_{k,\ell\in\mathbb{Z}} (S + (k/\Omega, \ell/T)) \subseteq \bigcup_{j=0}^{L-1} R_{q_j,m_j}. \tag{9}$$

Then $OPW^2(S)$ can be identified by regular operator sampling with identifier $g = \sum_n c_n \delta_{nT}$, (c_n) a period-L sequence, and there exist period-L sequences $b_j = (b_{j,k})$ and functions $\Phi_j(t,\nu)$ for $0 \le j \le L{-}1$, such that

$$h(x,t) = e^{2\pi i(t+t_0)\nu_0} \sum_k \sum_{j=0}^{L-1} b_{j,k}\, Hg(t-(q_j-k)T)\, e^{-2\pi i m_j (q_j-k)/L}\, \Phi_j(t, x-(t+t_0)+(q_j-k)T). \tag{10}$$

where the sum converges unconditionally in $L^2(\mathbb{R}^2)$. Here

$$\Phi_j(t,s) = \int e^{2\pi i\nu s}\, \chi_{S_j}(t,\nu)\, d\nu$$

where

$$S_j = S \cap \bigcup_{k,\ell\in\mathbb{Z}} (R_{q_j,m_j} + (k/\Omega, \ell/T)).$$

B. Illustrations and Special Cases of Theorem 2.1

As a special case of Theorem 2.1, Shannon's sampling theorem can be extended to the following sampling theorem for operators. This result first appeared in [17].

Theorem 2.2: For $H \in OPW^2(S)$, $S \subseteq [0,T){\times}[-\Omega/2, \Omega/2)$ compact and $T\Omega \le 1$,

$$h(x,t) = e^{-\pi i t/T} \sum_{n\in\mathbb{Z}} \big(H \sum_{k\in\mathbb{Z}} \delta_{kT}\big)(t+nT)\, \frac{\sin(\frac{\pi}{T}((x-t)-nT))}{\pi((x-t)-nT)} \chi_{[0,T]}(t). \tag{11}$$

where the sum converges in $L^2(\mathbb{R}^2)$ and for each t, uniformly in x.

Proof: By the assumption on S, we can take $T\Omega = 1$ and $(t_0,\nu_0) = (0, -\Omega/2)$ in Theorem 2.1 so that $L = 1$, $q_0 = m_0 = 0$, $c_0 = 1$ and hence $b_{0,0} = 1$. In this case

$$\Phi_0(t,s) = \chi_{[0,T]}(t) \int_{-1/2T}^{1/2T} e^{2\pi i\nu s}\, d\nu = \chi_{[0,T]}(t)\, \frac{\sin(\pi s/T)}{\pi s}$$

and (11) follows.

The most straightforward extension of Theorem 2.2 is to operator classes $OPW^2(S)$ where S is compact, $|S| < 1$, but S is not necessarily contained in a rectangle with area smaller than or equal to 1.

Theorem 2.3: If $S \subseteq (0,\infty)\times\mathbb{R}$ is compact with $|S| < 1$ then $OPW^2(S)$ is identifiable via regular operator sampling. Specifically, there exist $T > 0$ and $L \in \mathbb{N}$ such that $S \subseteq [0, LT]\times[-1/(2T), 1/(2T)]$, and a period-$L$ sequence $c = (c_n)$ such that $g = \sum_n c_n\, \delta_{nT}$ identifies $OPW^2(S)$. Moreover, there exist period-L sequences $b_j = (b_{j,k})$, and integers $0 \leq q_j,\, m_j \leq L-1$, for $0 \leq j \leq L-1$ such that

$$h(x,t) = e^{-\pi i t/T} \sum_k \sum_{j=0}^{L-1} b_{j,k}\, Hg(t-(q_j-k)T)\, e^{2\pi i m_j (x-t)/LT}\, \phi((x-t)+(q_j-k)T)\, r(t-q_jT) \quad (12)$$

where $r, \phi \in \mathcal{S}(\mathbb{R})$ satisfy

$$\sum_{k\in\mathbb{Z}} r(t+kT) = 1 = \sum_{n\in\mathbb{Z}} \widehat{\phi}(\gamma + n/LT), \quad (13)$$

where $r(t)\widehat{\phi}(\gamma)$ is supported in a neighborhood of $[0,T]\times[0, 1/LT]$, and where the sum in (12) converges unconditionally in L^2 and for each t uniformly in x.

Equation (12) is a direct generalization of (11) under the assumption that $r(t) = \chi_{[0,T]}(t)$ and $\widehat{\varphi}(\gamma) = \chi_{[0,\Omega]}(\gamma)$. The passage to smooth cut-off functions r and φ is enabled by the fact that $|S| < 1$ and allows for faster decay of the reconstruction functions, and for the validity and convergence of the reconstruction sums in more general function spaces. These matters have been studied extensively in [17].

By generalizing the setting to other function spaces, we can more precisely illustrate the connection between operator sampling and the classical sampling theorem attributed to Shannon, Whittaker, and Kotelnikov among others, and also the connection with the well-known fact that time-invariant operators can be identified by their impulse response.

Definition 2.4: We define the operator Paley-Wiener spaces $OPW^{\infty,2}(S)$ and $OPW^{2,\infty}(S)$ by

$$OPW^{\infty,2}(S) = \{H \in \mathcal{L}(L^2(\mathbb{R}), L^2(\mathbb{R}))\colon\ \operatorname{supp}\eta_H \subseteq S,\ \ \|\sigma_H\|_{L^{\infty,2}} = \Big\| \int |\sigma_H(\cdot,\xi)|^2 d\xi \Big\|_\infty^{1/2} < \infty\}$$

and

$$OPW^{2,\infty}(S) = \{H \in \mathcal{L}(L^2(\mathbb{R}), L^2(\mathbb{R}))\colon\ \operatorname{supp}\eta_H \subseteq S,\ \ \|\sigma_H\|_{L^{2,\infty}} = \Big(\int \|\sigma_H(x,\cdot)\|_\infty^2 dx \Big)^{1/2} < \infty\}$$

([17], Theorem 4.2). $OPW^{p,q}(S)$ is a Banach space with respect to the norm $\|H\|_{OPW^{p,q}} = \|\sigma_H\|_{L^{pq}}$.

Note that convolution with a compactly supported kernel whose Fourier transform is in L^2 is an operator in $OPW^{\infty,2}$ and multiplication by a bandlimited function in L^2 is an operator in $OPW^{2,\infty}$.

First, take H to be ordinary convolution by $h_H(t)$, this means that $h_H(x,t)$ depends only on t, that is, $h_H(x,t) = h_H(t)$. In this case H can be identified in principle by $g = \delta_0$, the unit impulse at the origin, since $Hg(x) = h_H(x)$. That is, $\Lambda = \{0\}$ is a sampling set for the class of convolution operators. Translating this into our operator sampling formalism results in something slightly different.

Assume that h is supported in the interval $[0, T']$, $\widehat{h} \in L^2$ and that $T > T'$, and $\Omega > 0$ are chosen so that $\Omega T < 1$. In this case, $\eta_H(t,\nu) = h(t)\,\delta_0(\nu)$ and $\sigma_H(x,\xi) = \widehat{h}(\xi)$. Therefore $\sigma_H \in L^{\infty,2}$ and $H \in OPW^{\infty,2}([0,T']\times[-\Omega/2, \Omega/2])$.

Applying Theorem 2.3 to this situation, note that if $g = \sum_n \delta_{nT}$ then Hg is simply the T–periodized impulse response $h(t)$, and it follows from the theorem (or by direct calculation) that with $r,\, \varphi \in \mathcal{S}(\mathbb{R})$, $r(t) = 1$ on $[0, T']$ and vanishing outside an interval of length T containing $[0, T']$, and with $\widehat{\varphi}(0) = 1$ and $\widehat{\varphi}$ vanishing outside $[-\Omega/2, \Omega/2]$,

$$\begin{aligned} r(t) \sum_{k\in\mathbb{Z}} (Hg)(t+kT)\varphi(x-t-kT) &= \sum_{k\in\mathbb{Z}} \sum_{n\in\mathbb{Z}} r(t)\, h(t+kT-nT)\, \varphi(x-t-kT) \\ &= \sum_{k\in\mathbb{Z}} h(t)\, \varphi(x-t-kT) = h(t). \end{aligned}$$

Here we have used the fact that $r(t) = 1$ on $[0, T']$ and vanishes outside a neighborhood of $[0, T']$ and that $\sum_k \varphi(x - t - kT) = 1$ by the Poisson Summation Formula and in consideration of the support constraints on $\widehat{\varphi}$. Indeed the theorem says that the sum $\sum_k \varphi(x - t - kT)$ converges to 1 in the L^∞ norm and in particular uniformly on compact sets.

To compare Theorem 2.3 with the classical sampling theorem, take H to be multiplication by some fixed function $m \in L^2$ with $\operatorname{supp} \widehat{m} \subseteq [-\Omega/2, \Omega/2]$ then $\eta_H(t, \nu) = \delta_0(t)\widehat{m}(\nu)$, $h(t, x) = \delta_0(t)\, m(x - t)$, and $\sigma_H(x, \xi) = m(x)$. Let $\Omega' > \Omega$ and $T > 0$ be such that $\Omega' T < 1$. Then $\sigma_H \in L^{2,\infty}$ and $H \in OPW^{2,\infty}([-T/2, T/2] \times [-\Omega/2, \Omega/2])$.

Choose r, $\varphi \in \mathcal{S}(\mathbb{R})$ such that $\operatorname{supp} r \subseteq [-T/2, T/2]$ and $r(0) = 1$ and $\operatorname{supp} \widehat{\varphi} \subseteq [-\Omega'/2, \Omega'/2]$ and $\widehat{\varphi}(\nu) = 1$ on $[-\Omega/2, \Omega/2]$. If $g = \sum_n \delta_{nT}$, then $Hg = \sum_n m(nT)\, \delta_{nT}$, and it follows from Theorem 2.3 (and by direct calculation) that

$$\begin{aligned} \delta_0(t)\, m(x - t) &= r(t) \sum_{k \in \mathbb{Z}} (Hg)(t + kT) \varphi(x - t - kT) \\ &= r(t) \sum_{k \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} m(nT)\, \delta_{(n-k)T}(t) \varphi(x - t - kT) \\ &= \sum_{n \in \mathbb{Z}} m(nT)\, \varphi(x - nT) \end{aligned}$$

by support considerations on the function $r(t)$. Therefore we have the summation formula

$$m(x) = \sum_{n \in \mathbb{Z}} m(nT)\, \varphi(x - nT)$$

where the sum converges unconditionally in L^2. This recovers the classical sampling formula when sampling above the Nyquist rate.

C. Necessary and sufficient conditions on the sampling rate in operator sampling

A second goal of this paper is to investigate efficient sampling rates for regular operator sampling. In the classical sampling theory of functions, the sampling rate must exceed the reciprocal of the area of the bandlimiting set; and regardless of the measure of the bandlimiting set, a (possibly high density) sampling set always exists. As mentioned above (Theorem 1.4), operator sampling of $OPW^2(S)$ is only possible if the measure of S satisfies $|S| \leq 1$. In addition, the sampling rates in operator sampling depend on the geometry of S in an intricate way. A necessary condition on the sampling rate is the following.

Theorem 2.5: If S is closed and Λ is a set of sampling for $OPW^2(S)$ with $\inf\{|\lambda - \mu| : \lambda, \mu \in \Lambda\} > 0$, then

$$D(\Lambda) \geq \Big\| \int_{\mathbb{R}} \chi_S(\cdot, \nu)\, d\nu \Big\|_\infty .$$

The quantity $\big\| \int_{\mathbb{R}} \chi_S(\cdot, \nu)\, d\nu \big\|_\infty$ can be interpreted as the maximum vertical extent of the set S.

To reduce the average rate at which we have to send Dirac impulses into a channel to apply regular operator sampling results, we seek to find

$$\min \Big\{ \frac{\|c\|_0}{TL} : \sum_n c_n \delta_{nT} \text{ identifies } OPW^2(S) \Big\} .$$

Clearly, the minimization problem is difficult since the choice of T and c is coupled. In general, we seek to choose T as large as possible, and L and c so that the relative support of c, $\|c\|_0/L$ is as small as possible. In the following, we address the problem of finding large T which allows for regular operator sampling and we establish a sufficient condition on $\|c\|_0/L$ based on the geometry and size of S.

We will first discuss the choice of T. Observe that our main result, Theorem 2.1, improves on Theorem 2.3 as the latter requires the region S be compact, and that for some $T > 0$ and $L \in \mathbb{N}$, with $\Omega = 1/LT$, $S \subseteq (0, 1/\Omega)\times(-1/2T, 1/2T)$ and that S be efficiently covered by rectangles of the form $[0,T]\times[0,\Omega] + (0, -1/2T) + (qT, m\Omega)$, $(q,m) \in \mathbb{Z}^2$, that is, that S intersect at most L such rectangles. This condition is weakened in Theorem 2.1 by requiring only that for some $T > 0$, $L \in \mathbb{N}$ and $\Omega = 1/LT$, S is a subset of a fundamental domain of $1/\Omega\,\mathbb{Z}\times 1/T\,\mathbb{Z}$ and that for some shift (t_0, ν_0),

$$S_{per} = [0, 1/\Omega]\times[0, 1/T] + (t_0, \nu_0) \cap \Big(\bigcup_{k,\ell\in\mathbb{Z}} S + (k/\Omega, \ell/T) \Big)$$

is efficiently covered by rectangles of the form $[0,T]\times[0,\Omega] + (t_0, \nu_0) + (qT, m\Omega)$, $(q,m) \in \mathbb{Z}^2$. In particular, S need no longer be closed nor bounded. Observe that in case S is compact, $T > 0$ and $L \in \mathbb{N}$ satisfying the hypotheses of Theorem 2.3 also satisfy the hypotheses of Theorem 2.1, and hence the latter theorem always allows a choice of T at least as large as the former.

To reduce the sampling rate further, we can consider S to be a subset of a fundamental domain of a general lattice, specifically replacing $T\mathbb{Z} \times \Omega\mathbb{Z}$ by a general lattice $A\mathbb{Z}^2$, and requiring S to be a subset of a fundamental domain of $LA\mathbb{Z}^2$. Our next theorem relies on basic insights on the role of symplectic geometry in time-frequency and generalizes Theorem 2.1. For simplicity, we state our result involving the covering of the periodization of S with respect to $LA\mathbb{Z}^2$ by fundamental domains of the lattice $A\mathbb{Z}^2$ for lower triangular matrices $A = \left(\begin{smallmatrix} T & 0 \\ a & \Omega \end{smallmatrix}\right)$ that are shifted in the ν direction. Since $a = 0$ puts us in the case of Theorem 2.1, letting a vary will always allow for choosing T at least as large as in Theorem 2.1.

In Section IV-B below we discuss the general case in detail and compute the quite involved resulting reconstruction formulas.

Theorem 2.6: Let $S \subseteq \mathbb{R}^2$ satisfy $|S| < 1$ and suppose that for some $A = \left(\begin{smallmatrix} T & 0 \\ a & \Omega \end{smallmatrix}\right)$ with $\det A = T\Omega = 1/L$, L prime, S is contained in a fundamental domain of the lattice $LA\mathbb{Z}^2$. Assume that for $\nu_0 \in \mathbb{R}$, integers $0 \le q_j,\, m_j \le L{-}1$, $0 \le j \le L{-}1$, and for the parallelograms $P_{q,m} = A\big([0,1]^2{+}(0,\nu_0) + (q,m)^T\big)$, $q,\, m \in \mathbb{Z}$, we have

$$A\big(L[0,1]^2 + (0,\nu_0)\big) \cap \bigcup_{k,\ell\in\mathbb{Z}} S + LA(k,\ell)^T \subseteq \bigcup_{j=0}^{L-1} P_{q_j,m_j}. \tag{14}$$

Then $OPW^2(S)$ can be identified by operator sampling. Namely, with the period-L sequence $c = (c_n)$ and the period-L sequences $b_j = (b_{j,k})_k$ from Theorem 2.1 and functions

$$\Phi_j(t,s) = \int e^{2\pi i\nu s}\, \chi_{S_j}(t,\nu)\, d\nu, \qquad S_j = S \cap \bigcup_{k,\ell\in\mathbb{Z}} (P_{q_j,m_j} + LA(k,\ell)^T),$$

we have

$$h(x,t) = e^{-\pi i a t^2/T} \sum_k \sum_{j=0}^{L-1} b_{j,k}\, e^{-\pi i a T(t/T-(q_j-k))^2}\; Hg(t-(q_j-k)T)$$
$$\Phi_j(t,\, x-(q_j-k)T)\, e^{2\pi i(q_j-k)at}\,.$$

where $g = \sum c_n e^{\pi i T a n^2}\delta_{nT}$ and the sum converges unconditionally in $L^2(\mathbb{R}^2)$. If the product Ta is rational, say $Ta/2 = p/q$ in lowest terms, then $(c_n e^{\pi i T a n^2})_n$ is periodic with period being the least common multiple of q and L. In particular, if $LTa/2$ is an integer, then the period is L as well.

The choice of the period-L sequence $c = (c_n)$ in Theorems 2.1, 2.3, and 2.6 is governed by the following considerations. Defining the translation operator T on $\mathbb{C}^L$ by $T(x_0,\, x_1,\, \ldots,\, x_{L-1}) = (x_{L-1},\, x_0,\, x_1,\, \ldots,\, x_{L-2})$

and the modulation operator M on $\mathbb{C}^L$ by $M(x_0, x_1, \ldots, x_{L-1}) = (\omega^0 x_0, \omega^1 x_1, \ldots, \omega^{L-1} x_{L-1})$ where $\omega = e^{2\pi i/L}$, we define the *finite Gabor system with window* c as the system of L^2 vectors in $\mathbb{C}^L$, $\{T^q M^m c\}_{q,m=0}^{L-1}$, where with a slight abuse of notation we think of c now as the vector $(c_0, \ldots, c_{L-1}) \in \mathbb{C}^L$ (for details on Gabor frames in finite dimensions, see [14], [13], [4] and the overview article [25]). Given a covering of S_{per} in (9) by at most L rectangles of the form $R_{q,m} = [0,T]\times[0,1/LT] + (t_0,\nu_0) + (qT, m/LT)$, specifically if

$$S_{per} \subseteq \bigcup_{j=0}^{L-1} R_{q_j,m_j},$$

then we require that c be chosen so that $\{T^{q_j} M^{m_j} c\}_{j=0}^{L-1}$ forms a basis for $\mathbb{C}^L$. It has been shown that if L is prime, then such a choice of c always exists [14], but such a choice may well exist even if L is not prime. Moreover, since the reconstruction formulas require that we invert the $L \times L$ matrix whose columns are given by $T^{q_j} M^{m_j} c$, choosing L as small as possible is also desirable.

The main result in this paper relevant to finding a sufficient condition on the sampling rate for identification of $OPW^2(S)$ is the following.

Theorem 2.7: Let $S \subseteq \mathbb{R}^2$, $|S| < 1$, and suppose that for some $T > 0$ and $N \in \mathbb{N}$, S is a subset of a fundamental domain of the lattice $TN\mathbb{Z} \times 1/T\,\mathbb{Z}$, and that S_{per} can be covered by no more than N rectangles of the form $R_{q_j,m_j} = [0,T]\times[0,1/TN] + (q_jT, m_j/TN)$, $(q_j, m_j) \in \mathbb{Z}^2$. Then for every sufficiently large prime L, $OPW^2(S)$ can be identified via regular operator sampling by an identifier satisfying

$$\frac{\|c\|_0}{L} < \sum_j |R_{q_j,m_j}|.$$

Note that if $\sum_j |R_{q_j,m_j}|$ is close to 1 (that is, if the covering of S_{per} by rectangles is very coarse), then the conclusion of the theorem is quite weak. However, if $|S|$ is small, and if, for some $T > 0$ not too small, S_{per} can be covered by a union of rectangles whose total area is small, then Theorem 2.7 gives some hope of an efficient sampling scheme for $OPW^2(S)$.

D. Sampling and reconstruction of operators with small, but unknown support

Just as in classical sampling, operator sampling requires full knowledge of the bandlimitation we expect an operator to have, that is, the reconstruction formulas for $OPW^2(S)$ depend on knowing the region S. However, in some applications S may not be known precisely, but only some information on its size, geometry and location is given by physical considerations. In Theorem 2.8 we address the question whether such operator can be sampled and reconstructed in a stable matter. Independently of our work, Heckel and Boelcskei have analyzed the problem of sampling operators with unknown bandlimitation in greater detail [3]. In an analogous setup, they were able to prove identifiability for unknown support sets of area less than one, rather than less than $1/2$ achieved below.

Theorem 2.8: Fix $A, B, \epsilon, U > 0$ and $N \in \mathbb{N}$. Let $\mathcal{H}(A,B,U,N,\epsilon)$ contain all operators with $\operatorname{supp} \mathcal{F}_s\sigma_H = \operatorname{supp} \eta_H \subseteq [-A,A]\times[-B,B]$ and such that there exist N Jordan curves C_i with the property that

1) $\operatorname{supp} \mathcal{F}_s\sigma_H = \operatorname{supp} \eta_H$ is contained in the interior sets of the Jordan curves,
2) the sum of areas of the interior sets is less than $1/2 - \epsilon$, and
3) the sum of lengths of the Jordan curves is bounded by U.

Then there exists a prime L and an L-periodic sequence $\{c_n\}$ such that $g = \sum_n c_n \delta_{n/\sqrt{L}}$ identifies $\mathcal{H}(A,B,U,N,\epsilon)$.

The reconstruction of an operator $H \in \mathcal{H}(A,B,U,N,\epsilon)$ is then carried out in three steps: first, we find T, Ω with $L = 1/T\Omega$ prime which ensure that the "rectified" support R_H of H has area not greater than 1/2 (see the proof of Theorem 2.8 in Section III-D). Under this assumption, we determine R_H. In the final step, we apply the operator reconstruction formula developed in Theorem 2.1 to $OPW(R_H)$.

To determine the rectified support of η_H with $H \in \mathcal{H}(A, B, U, N, \epsilon)$, we will apply ideas from compressed sensing. Indeed, Lemma 3.9 below, shows that from $H \sum_n c_n \delta_{n\sqrt{L}}$, we can compute a length L vector $y(t, \nu)$ with $y(t, \nu) = G(c)x(t, \nu)$ and where the unknown discrete support of the length L^2 vector $x(t, \nu)$ encodes the support of the bivariate function $\eta_H(t, \nu)$. In fact, recovering the vector $x(t, \nu)$ for a single point (t, ν) provides us with the support structure of η_H. Note that the conditions given above imply that $x(t, \nu)$ has at most $L/2$ nonzero components.

The matrix $G(c)$ consists of time-frequency shifted copies of the vector $(c_0, \ldots, c_{L-1})$. This matrix plays the role of a so-called measurement matrix and has the ability to recover any $L/2$-sparse vector $x(t, \nu)$ [13], [14]. But finding an $L/2$-sparse vector requires consideration of every support structure out of $\binom{L^2}{L/2}$ possible ones, which is hardly possible for L not being of the order $2, 3, 5$. In addition, to check whether c is appropriate, we would have to compute $\binom{L^2}{L}$ determinants of $L \times L$ matrices, which is again only possible for very small L. (It is shown in [14], [13], that if each of the L entries are chosen randomly, for example choosing the entries independently according to a uniform distribution on the unit circle in the complex plane, then $G(c)$ has with probability 1 no zero minors.)

If we know that far fewer than $L/2$ cells are active, then we can try to apply compressed sensing algorithms such as Basis Pursuit or Orthogonal Matching Pursuit to recover x from $y = G(c)x$. Indeed, the matrix $G(c)$ with randomly chosen c has been established to be a good measurement matrix with high probability [19], [18], [20], [11].

The work of Boelcskei and Heckel improves on our results above. They show that if only $L - 1$ cells are active, these can be determined. Their analysis and derived recovery algorithms rely on the fact that by varying (t, ν) you obtain a family of equations $y(t, \nu) = G(c)x(t, \nu)$ where the vectors $x(t, \nu)$ have identical sparsity structure.

III. Sampling and reconstruction

The purpose of this section is to present a proof of Theorem 2.1. Before doing so, a few preliminaries on finite Gabor systems and the Zak transform are required.

A. Preliminaries

Definition 3.1: Given $L \in \mathbb{N}$, let $\omega = e^{2\pi i/L}$ and define the *translation operator* T on $(x_0, \ldots, x_{L-1}) \in \mathbb{C}^L$ by

$$Tx = (x_{L-1}, x_0, x_1, \ldots, x_{L-2}),$$

and the *modulation operator* M on $\mathbb{C}^L$ by

$$Mx = (\omega^0 x_0, \omega^1 x_1, \ldots, \omega^{L-1} x_{L-1}).$$

Given a vector $c \in \mathbb{C}^L$ the *finite Gabor system with window* c is the collection $\{T^q M^p c\}_{q,p=0}^{L-1}$.

Note that the discrete Gabor system defined above consists of L^2 vectors in $\mathbb{C}^L$ which form an overcomplete tight frame for $\mathbb{C}^L$ [14].

Definition 3.2: The non-normalized Zak Transform is defined for $f \in \mathcal{S}(\mathbb{R})$ by

$$Z_a f(t, \nu) = \sum_{n \in \mathbb{Z}} f(t - an)\, e^{2\pi i a n \nu}.$$

The normalized Zak Transform $Z_a f(t, \nu)$ satisfies the quasi-periodicity relations

$$Z_a f(t + a, \nu) = e^{2\pi i a \nu}\, Z_a f(t, \nu) \text{ and } Z_a f(t, \nu + 1/a) = Z_a f(t, \nu).$$

$\sqrt{a}\, Z_a$ can be extended to a unitary operator from $L^2(\mathbb{R})$ onto $L^2([0,a]\times[0,1/a])$.

The following Lemma connects the output $Hg(x)$ where g is a weighted delta-train, to the spreading function $\eta_H(t,\nu)$. From this a reconstruction formula can be derived.

Lemma 3.3: Let $a>0$ be given and let $g=\sum_n \delta_{na}$. Then for all $(t,\nu)\in\mathbb{R}^2$,

$$(Z_a\circ H)g(t,\nu)=a^{-1}\sum_k\sum_m \eta_H(t+ak,\nu+m/a)\, e^{-2\pi i\nu ka},$$

where η_H is the spreading function of the operator H.

Proof: For $f\in\mathcal{S}(\mathbb{R})$ and $\phi\in\mathcal{S}'(\mathbb{R})$, define the *short-time Fourier transform (STFT)* of f with respect to ϕ by $V_\phi f(t,\nu)=\langle f, T_t M_\nu \phi\rangle$. Straightforward calculations show that if $g=\sum_n \delta_{na}$ then

$$V_g f(t,\nu)=Z_a f(t,\nu),$$

and moreover that

$$\langle Hg,f\rangle=\langle \eta_H, V_g f\rangle=\langle \eta_H, Z_a f\rangle$$

where the bracket on the left is the L^2 inner product on $\mathbb{R}$ and that on the right the L^2 inner product on $\mathbb{R}^2$. Because Z_a is unitary up to a constant, it follows that

$$a\,\langle (Z_a\circ H)g, Z_a f\rangle=\langle Hg,f\rangle=\langle \eta_H, Z_a f\rangle$$

where this time the bracket on the left is the L^2 inner product on the rectangle $[0,a]\times[0,1/a]$. The inner product on the right can be rewritten as

$$\begin{aligned}
\langle \eta_H, Z_a f\rangle &= \iint \eta_H(t,\nu)\,\overline{Z_a f(t,\nu)}\,dt\,d\nu\\
&= \sum_k\sum_m \int_{m/a}^{(m+1)/a}\int_{ka}^{(k+1)a} \eta_H(t,\nu)\,\overline{Z_a f(t,\nu)}\,dt\,d\nu\\
&= \sum_k\sum_m \int_0^{1/a}\int_0^{a} \eta_H(t+ka,\nu+m/a)\,\overline{Z_a f(t+ka,\nu+m/a)}\,dt\,d\nu\\
&= \int_0^{1/a}\int_0^{a} \sum_k\sum_m \eta_H(t+ka,\nu+m/a)\, e^{-2\pi i\nu ka}\,\overline{Z_a f(t,\nu)}\,dt\,d\nu.
\end{aligned}$$

Since this holds for every $f\in\mathcal{S}(\mathbb{R})$, the result follows.

Lemma 3.4: Let $T,\Omega>0$ be given such that $T\Omega=1/L$ for some $L\in\mathbb{N}$, let (c_n) be a period-L sequence, and define $g=\sum_n c_n\,\delta_{nT}$. Then for $(t,\nu)\in\mathbb{R}\times\widehat{\mathbb{R}}$,

$$(Z_{1/\Omega}\circ H)g(t,\nu)=\Omega\sum_{q=0}^{L-1} c_{-q}\sum_k\sum_m \eta_H(t+k/\Omega+qT,\nu+m\Omega)\, e^{-2\pi i(\nu+m\Omega)qT}\, e^{-2\pi i\nu k/\Omega}. \quad (15)$$

Proof: A straightforward calculation shows that for any $\alpha\in\mathbb{R}$, the spreading function of the operator $H\circ T_\alpha$ is $\eta_H(t-\alpha,\nu)\,e^{2\pi i\nu\alpha}$ where η_H is the spreading function of H. Next note that writing uniquely

$j = nL - q$, $0 \le q \le L{-}1$,

$$\begin{aligned} g &= \sum_j c_j\, \delta_{nT} \\ &= \sum_{q=0}^{L-1} \sum_{n\in\mathbb{Z}} c_{nL-q}\, \delta_{nLT-qT} \\ &= \sum_{q=0}^{L-1} c_{-q} \sum_{n\in\mathbb{Z}} \delta_{nLT-q/L\Omega} \\ &= \sum_{q=0}^{L-1} c_{-q} T_{-q/L\Omega} \sum_{n\in\mathbb{Z}} \delta_{n/\Omega}. \end{aligned}$$

Therefore, by Lemma 3.3,

$$\begin{aligned} (Z_{1/\Omega}\circ H)g(t,\nu) &= (Z_{1/\Omega}\circ H)\Big(\sum_{q=0}^{L-1} c_{-q} T_{-q/L\Omega} \sum_{n\in\mathbb{Z}} \delta_{n/\Omega}\Big)(t,\nu) \\ &= \sum_{q=0}^{L-1} c_{-q} (Z_{1/\Omega}\circ H\circ T_{-q/L\Omega})\Big(\sum_{n\in\mathbb{Z}} \delta_{n/\Omega}\Big)(t,\nu) \\ &= \Omega \sum_{q=0}^{L-1} c_{-q} \sum_k \sum_m \eta_H(t + k\Omega + q/L\Omega, \nu + m\Omega)\, e^{2\pi i(\nu+m\Omega)(q/L\Omega)}\, e^{-2\pi i \nu k/\Omega}\,. \end{aligned}$$

Definition 3.5: Given a bivariate function $f(t,\nu)$ and parameters $T,\Omega > 0$, define the *(*$1/\Omega$, $1/T$*)–quasiperiodization* of f, denoted f^{QP}, by

$$f^{QP}(t,\nu) = \sum_k \sum_\ell f(t + k/\Omega, \nu + \ell/T)\, e^{-2\pi i\nu k/\Omega} \tag{16}$$

whenever the sum is defined.

Remark 3.6: (a) Note that f^{QP} satisfies $f^{QP}(t,\nu + 1/T) = f^{QP}(t,\nu)$ and $f^{QP}(t + 1/\Omega,\nu) = e^{2\pi i\nu/\Omega} f^{QP}(t,\nu)$ for all $(t,\nu)\in\mathbb{R}^2$. These are similar to the quasiperiodicity conditions satisfied by the Zak transform.

(b) Under the assumption that the support of f is contained in a fundamental domain of $1/\Omega\,\mathbb{Z}\times 1/T\,\mathbb{Z}$, the following lemma shows that f can be easily recovered from the function

$$f^{QP}(t,\nu)\, \chi_{[0,1/\Omega]}(t)\, \chi_{[0,1/T]}(\nu).$$

Lemma 3.7: Suppose that $\mathrm{supp}(f) \subseteq S$ and that S is contained in a fundamental domain of $1/\Omega\,\mathbb{Z}\times 1/T\,\mathbb{Z}$. Then

$$f(t,\nu) = \sum_k \sum_\ell f^{QP}(t - k/\Omega, \nu - \ell/T)\, \chi_{[0,1/\Omega]}(t - k/\Omega)\, \chi_{[0,1/T]}(\nu - \ell/T)\, e^{2\pi i k\nu/\Omega}\, \chi_S(t,\nu) \tag{17}$$

where if $f \in L^2(\mathbb{R}^2)$, the sum converges in L^2 and uniformly on compact sets.

Proof: First note that under the given assumptions, the functions being summed in (17) have pairwise disjoint supports. Since $|S| < 1$, it follows that the sum converges in L^2 if $f \in L^2(\mathbb{R}^2)$. Moreover, since on each compact set, the sum reduces to a finite sum, we get uniform convergence on compact sets.

To complete the proof, we show that (17) holds pointwise. Note first that for all (t,ν),

$$f^{QP}(t,\nu)\,\chi_S(t,\nu) = f(t,\nu)$$

since if $(t,\nu) \in S$ then because S is a fundamental domain, only the $(k,\ell) = (0,0)$ term survives in (16). It remains to show that

$$\sum_k \sum_\ell f^{QP}(t - k/\Omega, \nu - \ell/T)\ \chi_{[0,1/\Omega]}(t - k/\Omega)\ \chi_{[0,1/T]}(\nu - \ell/T)\, e^{2\pi i k\nu/\Omega} = f^{QP}(t,\nu).$$

To see this, note that

$$\begin{aligned}
&\sum_{k,\ell} f^{QP}(t - k/\Omega, \nu - \ell/T)\,\chi_{[0,1/\Omega]}(t - k/\Omega)\,\chi_{[0,1/T]}(\nu - \ell/T)\, e^{2\pi i k\nu/\Omega} \\
&= \sum_{k,\ell}\sum_{j,m} f(t + (j-k)/\Omega, \nu + (m-\ell)/T)\, e^{-2\pi i \nu j/\Omega}\,\chi_{[0,1/\Omega]}(t - k/\Omega)\,\chi_{[0,1/T]}(\nu - \ell/T)\, e^{2\pi i k\nu/\Omega}.
\end{aligned}$$

Now suppose that $(t,\nu) \in [p/\Omega, (p+1)/\Omega) \times [q/T, (q+1)/T)$ for some $(p,q) \in \mathbb{Z}^2$. Then in the above sum, only the $(k,\ell) = (p,q)$ term survives, and we arrive at

$$\begin{aligned}
&\sum_{k,\ell}\sum_{j,m} f(t + (j-k)/\Omega, \nu + (m-\ell)/T)\, e^{-2\pi i \nu j/\Omega}\,\chi_{[0,1/\Omega]}(t - k/\Omega)\,\chi_{[0,1/T]}(\nu - \ell/T)\, e^{2\pi i k\nu/\Omega} \\
&= \sum_{j,m} f(t + (j-p)/\Omega, \nu + (m-q)/T)\, e^{-2\pi i \nu (j-p)/\Omega}\,\chi_{[0,1/\Omega]}(t - p/\Omega)\,\chi_{[0,1/T]}(\nu - q/T) \\
&= f^{QP}(t,\nu).
\end{aligned}$$

Lemma 3.8: Let $T, \Omega > 0$ be given such that $T\Omega = 1/L$ for some $L \in \mathbb{N}$, let (c_n) be a period-L sequence. Then with $g = \sum_n c_n\,\delta_{nT}$,

$$(Z_{1/\Omega} \circ H)g(t,\nu) = \Omega \sum_{q=0}^{L-1} c_{-q} \sum_{m=0}^{L-1} \sum_k \eta_H^{QP}(t + qT, \nu + m\Omega)\, e^{-2\pi i \nu q T}\, e^{-2\pi i \nu m q/L} \tag{18}$$

for $(t,\nu) \in \mathbb{R}^2$.

Proof: The proof follows immediately from Lemma 3.4. Letting $m = nL + \ell$, $n \in \mathbb{Z}$ and $0 \le \ell \le L-1$, in (15) gives

$$\begin{aligned}
(Z_{1/\Omega} \circ H)g(t,\nu) &= \Omega \sum_{q=0}^{L-1} c_{-q} \sum_{\ell=0}^{L-1} \sum_k \sum_n \eta_H(t + k/\Omega + qT, \nu + nL\Omega + \ell\Omega)\, e^{-2\pi i(\nu + nL\Omega + \ell\Omega)qT}\, e^{-2\pi i \nu k/\Omega} \\
&= \Omega \sum_{q=0}^{L-1} c_{-q} \sum_{\ell=0}^{L-1} \left[\sum_{k,n} \eta_H(t + k/\Omega + qT, \nu + nL\Omega + \ell\Omega)\, e^{-2\pi i(\nu + \ell\Omega)k/\Omega}\right] e^{-2\pi i \nu q T}\, e^{-2\pi i \ell q/L} \\
&= \Omega \sum_{q=0}^{L-1} c_{-q} \sum_{\ell=0}^{L-1} \eta_H^{QP}(t + qT, \nu + \ell\Omega)\, e^{-2\pi i \nu q T}\, e^{-2\pi i \ell q/L}
\end{aligned}$$

where we have used the fact that $L\Omega = 1/T$.

Lemma 3.9: Let $T, \Omega > 0$ be given such that $T\Omega = 1/L$ for some $L \in \mathbb{N}$, let (c_n) be a period-L

sequence. Then with $g = \sum_n c_n\, \delta_{nT}$, $(t,\nu) \in \mathbb{R}^2$, and $p = 0,\, 1,\, \ldots,\, L{-}1$,

$$
\begin{aligned}
&e^{-2\pi i\nu Tp}\,(Z_{1/\Omega} \circ H)g(t+Tp,\nu) \\
&\qquad = \Omega \sum_{q=0}^{L-1}\sum_{m=0}^{L-1} (T^q\, M^m c)_p\, e^{-2\pi i\nu Tq}\, \eta_H^{QP}(t+Tq, \nu+\Omega m).
\end{aligned}
\tag{19}
$$

By (18),

$$
\begin{aligned}
(Z_{1/\Omega} \circ H)g(t+pT,\nu) &= \Omega \sum_{q=0}^{L-1} c_{-q} \sum_{\ell=0}^{L-1} \eta_H^{QP}(t+(q+p)T, \nu+m\Omega)\, e^{-2\pi i\nu qT}\, e^{-2\pi i\nu mq/L} \\
&= \Omega \sum_{q=p}^{L-1+p} c_{-(q-p)} \sum_{\ell=0}^{L-1} \eta_H^{QP}(t+qT, \nu+m\Omega)\, e^{-2\pi i\nu(q-p)T}\, e^{-2\pi i\nu m(q-p)/L} \\
&= \Omega \left(\sum_{q=p}^{L-1} + \sum_{q=L}^{L-1+p}\right) c_{-(q-p)} \sum_{\ell=0}^{L-1} \eta_H^{QP}(t+qT, \nu+m\Omega)\, e^{-2\pi i\nu(q-p)T}\, e^{-2\pi i\nu m(q-p)/L}.
\end{aligned}
$$

Using the periodicity of (c_n), the quasiperiodicity of η_H^{QP}, and the fact that $LT = 1/\Omega$, we continue with

$$
\begin{aligned}
&\Omega \sum_{q=L}^{L-1+p} c_{-(q-p)} \sum_{\ell=0}^{L-1} \eta_H^{QP}(t+qT, \nu+m\Omega)\, e^{-2\pi i\nu(q-p)T}\, e^{-2\pi i\nu m(q-p)/L} \\
&= \Omega \sum_{q=L}^{L-1+p} c_{-((q-L)-p)} \sum_{\ell=0}^{L-1} \eta_H^{QP}(t+LT+(q-L)T, \nu+m\Omega)\, e^{-2\pi i\nu((q-L)-p)T}\, e^{-2\pi i\nu m((q-L)-p)/L}\, e^{-2\pi i\nu LT} \\
&= \Omega \sum_{q=L}^{L-1+p} c_{-((q-L)-p)} \sum_{\ell=0}^{L-1} e^{-2\pi i\nu/\Omega}\, \eta_H^{QP}(t+1/\Omega+(q-L)T, \nu+m\Omega)\, e^{-2\pi i\nu((q-L)-p)T}\, e^{-2\pi i\nu m((q-L)-p)/L} \\
&= \Omega \sum_{q=L}^{L-1+p} c_{-((q-L)-p)} \sum_{\ell=0}^{L-1} \eta_H^{QP}(t+(q-L)T, \nu+m\Omega)\, e^{-2\pi i\nu((q-L)-p)T}\, e^{-2\pi i\nu m((q-L)-p)/L} \\
&= \Omega \sum_{q=0}^{p-1} c_{-(q-p)} \sum_{\ell=0}^{L-1} \eta_H^{QP}(t+qT, \nu+m\Omega)\, e^{-2\pi i\nu(q-p)T}\, e^{-2\pi i\nu m(q-p)/L}.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
(Z_{1/\Omega} \circ H)g(t+Tp,\nu) &= \Omega \left(\sum_{q=p}^{L-1} + \sum_{q=0}^{p-1}\right) c_{-(q-p)} \sum_{m=0}^{L-1} e^{-2\pi i m(q-p)/L} \eta_H^{QP}(t+qT, \nu+m\Omega)\, e^{-2\pi i\nu(q-p)T} \\
&= \Omega \sum_{q=0}^{L-1}\sum_{m=0}^{L-1} c_{-(q-p)}\, e^{-2\pi i m(q-p)/L}\, \eta_H^{QP}(t+qT, \nu+m\Omega)\, e^{-2\pi i\nu(q-p)T}.
\end{aligned}
$$

Since $(T^q\, M^m c)_p = c_{p-q}\, e^{2\pi i m(p-q)/L}$, the result follows.

B. Proof of Theorem 2.1

Suppose that S satisfies all the hypotheses of Theorem 2.1 with $(t_0, \nu_0) = (0, 0)$. Starting with (19), we have that for $0 \le p \le L{-}1$, and all $(t, \nu) \in \mathbb{R}^2$,

$$\begin{aligned} &e^{-2\pi i\nu Tp}\,(Z_{1/\Omega} \circ H)g(t + Tp, \nu) \\ &\qquad = \Omega \sum_{q=0}^{L-1}\sum_{m=0}^{L-1} (T^q\, M^m c)_p\, e^{-2\pi i \nu Tq}\, \eta_H^{QP}(t + Tq, \nu + \Omega m). \end{aligned}$$

Under the assumption (9), it follows that

$$\begin{aligned} &(Z_{1/\Omega} \circ H)g(t + Tp, \nu)\, e^{-2\pi i\nu Tp} \\ &\qquad = \Omega \sum_{j=0}^{L-1} (T^{q_j}\, M^{m_j} c)_p\, \eta_H^{QP}(t + q_jT, \nu + m_j\Omega)\, e^{-2\pi i\nu q_j T} \\ &\qquad = \sum_{j=0}^{L-1} a_{j,p}\left(\eta_H^{QP}(t + q_jT, \nu + m_j\Omega)\, e^{-2\pi i\nu q_j T}\right) \end{aligned}$$

where $[a_{j,p}]_{j,p=0}^{L-1}$ is an $L \times L$ matrix whose j^{th} column is $\Omega\,(T^{q_j}\, M^{m_j} c) \in \mathbb{C}^L$. Assuming that L is prime, we can choose a period-L sequence $c = (c_n)$ such that the matrix $[a_{j,p}]$ is invertible. In fact, the set of such $c \in \mathbb{C}^L$ is a dense open subset of $\mathbb{C}^L$ (see [13]). Let $[a_{j,p}]^{-1} = [b_{j,p}]$.

Again by (9), $\eta_H \in OPW^2(S)$ satisfies

$$\eta_H^{QP}(t, \nu)\chi_{[0,1/\Omega]}(t)\chi_{[0,1/T]}(\nu) = \sum_{j=0}^{L-1} \eta_H^{QP}(t, \nu)\chi_{[0,T]}(t - q_jT)\chi_{[0,\Omega]}(\nu - m_j\Omega),$$

and for each $0 \le j \le L{-}1$,

$$\eta_H^{QP}(t{+}q_jT, \nu{+}m_j\Omega)\chi_{[0,T]}(t)\chi_{[0,\Omega]}(\nu) = \sum_{p=0}^{L-1} b_{j,p}\, \chi_{[0,T]}(t)\chi_{[0,\Omega]}(\nu)\, e^{2\pi i\nu(q_j - p)T}\,(Z_{1/\Omega}{\circ} H)g(t{+}pT, \nu).$$

Therefore, by the quasiperiodicity of the Zak transform,

$$\begin{aligned} &\eta_H^{QP}(t, \nu)\chi_{[0,1/\Omega]}(t)\chi_{[0,1/T]}(\nu) \\ &\qquad = \sum_{j=0}^{L-1}\sum_{p=0}^{L-1} b_{j,p}\, \chi_{[0,T]}(t - q_jT)\chi_{[0,\Omega]}(\nu - m_j\Omega) \\ &\qquad\qquad e^{2\pi i(\nu - m_j\Omega)(q_j - p)T}\,(Z_{1/\Omega} \circ H)g(t - (q_j - p)T, \nu). \end{aligned}$$

Applying (17),

$$\begin{aligned}
\eta_H(t,\nu) &= \eta_H^{QP}(t,\nu)\,\chi_S(t,\nu)\\
&= \sum_{k,\ell} \eta_H^{QP}(t-k/\Omega,\nu-\ell/T)\,\chi_{[0,1/\Omega]}(t-k/\Omega)\chi_{[0,1/T]}(\nu-\ell/T)\,e^{2\pi i k\nu/\Omega}\,\chi_S(t,\nu)\\
&= \sum_{k,\ell} e^{2\pi i k\nu/\Omega} \sum_{j,p=0}^{L-1} b_{j,p}\,\chi_{[0,T]}(t-k/\Omega-q_jT)\chi_{[0,\Omega]}(\nu-\ell/T-m_j\Omega)\\
&\qquad e^{2\pi i(\nu-\ell/T-m_j\Omega)(q_j-p)T}\,(Z_{1/\Omega}\circ H)g(t-k/\Omega-(q_j-p)T,\nu-\ell/T)\,\chi_S(t,\nu)\\
&= \sum_{k,\ell} e^{2\pi i k\nu/\Omega} \sum_{j,p=0}^{L-1} b_{j,p}\,\chi_{[0,T]}(t-k/\Omega-q_jT)\chi_{[0,\Omega]}(\nu-\ell/T-m_j\Omega)\\
&\qquad e^{2\pi i(\nu-m_j\Omega)(q_j-p)T}\,e^{-2\pi i\nu k\Omega}\,(Z_{1/\Omega}\circ H)g(t-(q_j-p)T,\nu)\,\chi_S(t,\nu)\\
&= \sum_{j,p=0}^{L-1} b_{j,p}\,e^{2\pi i(\nu-m_j\Omega)(q_j-p)T}\,(Z_{1/\Omega}\circ H)g(t-(q_j-p)T,\nu)\\
&\qquad \sum_{k,\ell}\chi_{[0,T]}(t-k/\Omega-q_jT)\chi_{[0,\Omega]}(\nu-\ell/T-m_j\Omega)\,\chi_S(t,\nu).
\end{aligned}$$

Defining

$$S_j = S\cap\Big(\bigcup_{k,\ell} R_{q_j,m_j}+(k/\Omega,\ell/T)\Big),$$

it follows that $S=\bigcup_{j=0}^{L-1} S_j$, that the union is disjoint, and that

$$\sum_{k,\ell}\chi_{[0,T]}(t-k/\Omega-q_jT)\chi_{[0,\Omega]}(\nu-\ell/T-m_j\Omega)\,\chi_S(t,\nu)=\chi_{S_j}(t,\nu).$$

Therefore,

$$\begin{aligned}
\eta_H(t,\nu) &= \sum_{j,p=0}^{L-1} b_{j,p}\,e^{-2\pi i(\nu-m_j\Omega)(q_j-p)T}\sum_{n\in\mathbb{Z}} Hg(t-n/\Omega-(q_j-p)T)\,e^{2\pi i\nu n/\Omega}\,\chi_{S_j}(t,\nu)\\
&= \sum_{j=0}^{L-1}\sum_{p=0}^{L-1}\sum_{n\in\mathbb{Z}} b_{j,p}\,e^{-2\pi i(\nu-m_j\Omega)(q_j-p)T}\,Hg(t-nLT-(q_j-p)T)\,e^{2\pi i\nu nLT}\,\chi_{S_j}(t,\nu).
\end{aligned}$$

Extending $b_{j,p}$ to have period L in p, it follows that

$$\begin{aligned}
\eta_H(t,\nu) &= \sum_{j=0}^{L-1}\sum_{p=0}^{L-1}\sum_{n\in\mathbb{Z}} b_{j,p-nL}\,e^{-2\pi i(\nu-m_j\Omega)(q_j-(p-nL))T}\,Hg(t-(q_j-(p-nL))T)\,\chi_{S_j}(t,\nu)\\
&= \sum_{j=0}^{L-1}\sum_{k} b_{j,k}\,e^{-2\pi i(\nu-m_j\Omega)(q_j-k)T}\,Hg(t-(q_j-k)T)\,\chi_{S_j}(t,\nu).
\end{aligned}$$

Finally

$$
\begin{aligned}
h(x,t) &= \int \eta(t,\nu)\, e^{2\pi i(x-t)\nu}\, d\nu \\
&= \sum_{j=0}^{L-1} \sum_k b_{j,k}\, e^{-2\pi i m_j (q_j-k)/L}\, Hg(t-(q_j-k)T) \int e^{-2\pi i\nu((x-t)+(q_j-k)T)} \chi_{S_j}(t,\nu) \\
&= \sum_{j=0}^{L-1} \sum_k b_{j,k}\, e^{-2\pi i m_j (q_j-k)/L}\, Hg(t-(q_j-k)T)\, \Phi_j(t,(x-t)+(q_j-k)T)
\end{aligned}
\tag{20}
$$

where

$$\Phi_j(t,s) = \int e^{2\pi i\nu s}\, \chi_{S_j}(t,\nu)\, d\nu.$$

To complete the proof, note that for almost every t, the set, $\{\nu\colon (t,\nu) \in S_j\}$ is contained in a fundamental domain of the lattice $T\mathbb{Z}$ of $\mathbb{R}$. This implies that the measure of each such section is no more than $1/T$, and in particular that for almost every t, $\chi_{S_j}(t,\cdot) \in L^2(\mathbb{R})$. Therefore, by Plancherel's Formula,

$$\iint |\Phi_j(t,s)|^2\, dt\, ds = \iint \left| \int e^{2\pi i\nu s}\, \chi_{S_j}(t,\nu)\, d\nu \right|^2 ds\, dt = \iint |\chi_{S_j}(t,\nu)|^2\, d\nu\, dt = |S_j|^2 < \infty$$

and for almost every (t,s),

$$|\Phi_j(t,s)| \le \int \chi_{S_j}(t,\nu)\, d\nu \le 1/T.$$

Hence $\Phi_j \in L^2 \cap L^\infty(\mathbb{R}^2)$. Convergence of the reconstruction sum in $L^2(\mathbb{R}^2)$ follows from the observation that $Hg \in L^2(\mathbb{R})$ (see Lemma 3.3) and basic properties of the Zak Transform (see e.g., [5], Section 8.2).

If $(t_0,\nu_0) \neq (0,0)$, we formally compute

$$
\begin{aligned}
H &= \iint_S \eta_H(t,\nu)\, M_\nu T_t\ dt\, d\nu \\
&= \iint_{S-(t_0,\nu_0)} \eta_H(t+t_0,\nu+\nu_0)\ T_{t+t_0} M_{\nu+\nu_0}\ dt\, d\nu \\
&= \iint_{S-(t_0,\nu_0)} \eta_H(t+t_0,\nu+\nu_0)\ T_{t_0} T_t M_{\nu_0} M_\nu\ dt\, d\nu \\
&= \iint_{S-(t_0,\nu_0)} \eta_H(t+t_0,\nu+\nu_0)\, e^{-2\pi i t\nu_0}\ T_{t_0} M_{\nu_0} T_t M_\nu\ dt\, d\nu \\
&= T_{t_0} M_{\nu_0} \iint_{S-(t_0,\nu_0)} \eta_H(t+t_0,\nu+\nu_0)\, e^{-2\pi i t\nu_0}\ T_t M_\nu\ dt\, d\nu \\
&= T_{t_0} M_{\nu_0} \widetilde{H},
\end{aligned}
$$

where $\eta_{\widetilde{H}}(t,\nu) = \widetilde{\eta}(t,\nu) = e^{-2\pi i t\nu_0}\, \eta_H(t+t_0,\nu+\nu_0)$. Taking inverse Fourier transforms $\nu \to x$ on both sides, we obtain $\widetilde{h}(t,x) = e^{-2\pi i t\nu_0}\, h_H(t+t_0,x)\, e^{-2\pi i\nu_0 x}$ which is

$$h_H(t,x) = e^{2\pi i(x+t-t_0)\nu_0}\, \widetilde{h}(t-t_0,x).$$

With $\widetilde{S} = S-(t_0,\nu_0)$, we can apply (20) to reconstruct $\widetilde{h}$ from $\widetilde{H}g$ with the constructed $g = \sum c_n \delta_{nT}$, that is,

$$\widetilde{h}(x,t) = \sum_k \sum_{j=0}^{L-1} b_{j,k}\, \widetilde{H}g(t-(q_j-k)T)\, e^{-2\pi i m_j(q_j-k)/L}\, \widetilde{\Phi}_j(t,(x-t)+(q_j-k)T). \tag{21}$$

where

$$\widetilde{\Phi}_j(t,s) = \int e^{2\pi i\nu s}\, \chi_{\widetilde{S}_j}(t,\nu)\, d\nu$$

and

$$\widetilde{S}_j = \widetilde{S} \cap \bigcup_{k,\ell\in\mathbb{Z}} (\widetilde{R}_{q_j,m_j} + (k/\Omega, \ell/T)).$$

Shifting the set equation above by (t_0,ν_0), we recover the definition of S_j, namely

$$\begin{aligned} S_j = \widetilde{S}_j + (t_0,\nu_0) &= \widetilde{S} + (t_0,\nu_0) \cap \bigcup_{k,\ell\in\mathbb{Z}} (\widetilde{R}_{q_j,m_j} + (k/\Omega,\ell/T) + (t_0,\nu_0)) \\ &= S \cap \bigcup_{k,\ell\in\mathbb{Z}} (R_{q_j,m_j} + (k/\Omega,\ell/T)) \end{aligned}$$

and

$$\begin{aligned} \Phi_j(t,s) &= \int e^{2\pi i\nu s}\, \chi_{S_j}(t,\nu)\, d\nu = \int e^{2\pi i\nu s}\, \chi_{\widetilde{S}_j}(t-t_0,\nu-\nu_0)\, d\nu \\ &= e^{2\pi i s\nu_0} \int e^{2\pi i\nu s}\, \chi_{\widetilde{S}_j}(t-t_0,\nu)\, d\nu = e^{2\pi i s\nu_0}\, \widetilde{\Phi}_j(t-t_0,s) \end{aligned}$$

This translates to

$$\begin{aligned} h(x,t) &= e^{2\pi i(x+t-t_0)\nu_0} \sum_k \sum_{j=0}^{L-1} b_{j,k}\, \widetilde{H}g(t-t_0-(q_j-k)T)\, e^{-2\pi i m_j(q_j-k)/L}\, \widetilde{\Phi}_j(t-t_0,(x-t-t_0)+(q_j-k)T) \\ &= e^{2\pi i(x+t-t_0)\nu_0} \sum_k \sum_{j=0}^{L-1} b_{j,k}\, (M_{-\nu_0}T_{-t_0}H)g(t-t_0-(q_j-k)T)\, e^{-2\pi i m_j(q_j-k)/L} \\ &\qquad\qquad e^{-2\pi i(x-t-t_0+(q_j-k)T)\nu_0}\, \Phi_j(t,(x-t-t_0)+(q_j-k)T) \\ &= e^{2\pi i(x+t-t_0)\nu_0} \sum_k \sum_{j=0}^{L-1} b_{j,k}\, e^{-2\pi i(t-t_0-(q_j-k)T)\nu_0} Hg(t-(q_j-k)T)\, e^{-2\pi i m_j(q_j-k)/L} \\ &\qquad\qquad e^{-2\pi i(x-t-t_0+(q_j-k)T)\nu_0}\, \Phi_j(t,(x-t-t_0)+(q_j-k)T) \\ &= e^{2\pi i(t+t_0)\nu_0} \sum_k \sum_{j=0}^{L-1} b_{j,k}\, Hg(t-(q_j-k)T)\, e^{-2\pi i m_j(q_j-k)/L} \\ &\qquad\qquad \Phi_j(t,(x-(t+t_0)+(q_j-k)T)\,. \end{aligned}$$

C. *Outline of Proof of Theorem 2.3.*

Suppose that $S \subseteq (0,\infty)\times\mathbb{R}$ is compact with $|S| < 1$. Then for $\delta > 0$ sufficiently small, the set $S_\delta = S + [-\delta,\delta]^2$ also satisfies $S_\delta \subseteq (0,\infty)\times\mathbb{R}$ and $|S_\delta| < 1$. Since $|S_\delta| < 1$, then for any $L \in \mathbb{N}$ there exists $T > 0$ such that with $\Omega T = 1/L$,

$$S_\delta \subseteq (0, 1/\Omega)\times(-1/2T, 1/2T)$$

and S_δ is contained in at most L rectangles of the form

$$R_{q,m} = [0,T]\times[0,\Omega] - (0,1/2T) + (qT, m\Omega)$$

$q,\, m \in \mathbb{Z}$. Specifically, for some $0 \le q_j,\, m_j \le L-1$, $0 \le j \le L-1$,

$$S_\delta \subseteq \bigcup_{j=0}^{L-1} R_{q_j,m_j} = R.$$

Since $S \subseteq R$, it is sufficient to prove the theorem with $OPW^2(S)$ replaced by $OPW^2(R)$.

By Lemma 3.9, given $H \in OPW^2(R)$ with spreading function $\eta_H(t,\nu)$, and given any weighted delta train of the form $g = \sum_n c_n\, \delta_{nT}$ where $c = (c_n)$ is a period-L sequence, (19) holds with η_H^{QP} replaced by η_H for all (t,ν) in a neighborhood of $[0,T]\times[0,\Omega]$. For specificity, call this neighborhood $R_{0,0}^\epsilon = ([0,T]\times[0,\Omega] - (0,1/2T)) + [-\epsilon,\epsilon]^2$.

Let $r, \varphi \in \mathcal{S}(\mathbb{R})$ satisfy

$$\begin{aligned} \operatorname{supp} r &\subseteq [-\epsilon/2, T+\epsilon/2], \\ \operatorname{supp} \widehat{\varphi} &\subseteq [-\epsilon/2, \Omega+\epsilon/2], \end{aligned} \tag{22}$$

so that $\operatorname{supp} r(t)\widehat{\varphi}(\nu) \subseteq R_{0,0}^\epsilon$, and

$$\sum_{k\in\mathbb{Z}} r(t+kT) = 1 = \sum_{n\in\mathbb{Z}} \widehat{\varphi}(\nu + n\Omega), \tag{23}$$

for all $(t,\nu) \in \mathbb{R}^2$. For $\epsilon < \delta$, it is not hard to show that if $R_{q,m} \not\subseteq R$ then

$$\eta_H(t,\gamma) r(t-qT)\widehat{\varphi}(\nu - m\Omega) = 0. \tag{24}$$

Therefore,

$$\eta_H(t,\nu) = \sum_{j=0}^{L-1} \eta_H^{QP}(t,\nu)\, r(t-q_jT)\, \widehat{\varphi}(\nu - m_j\Omega).$$

Following precisely the proof of Theorem 2.1, with $r(t)$ replacing $\chi_{[0,T]}(t)$ and $\widehat{\varphi}(\nu)$ replacing $\chi_{[0,\Omega]}(\nu)$,

$$\eta_H(t,\nu) = \sum_{j=0}^{L-1} \sum_k b_{j,k}\, e^{-2\pi i(\nu - m_j\Omega)(q_j-k)T}\, Hg(t-(q_j-k)T)\, R_j(t,\nu)$$

where

$$\begin{aligned} R_j(t,\nu) &= \sum_{k,\ell} r(t-k/\Omega - q_jT)\widehat{\varphi}(\nu - \ell/T - m_j\Omega)\, \chi_R(t,\nu) \\ &= r(t-q_jT)\, \widehat{\varphi}(\nu - m_j\Omega). \end{aligned}$$

Finally, taking $t_0 = 0$ and $\nu_0 = -1/2T$,

$$h(x,t) = e^{\pi i t/T} \sum_{j=0}^{L-1} \sum_k b_{j,k}\, e^{2\pi i m_j(q_j-k)/L}\, Hg(t-(q_j-k)T)\, \Phi_j(t,(x-t)+(q_j-k)T)$$

where here

$$\begin{aligned}\Phi_j(t,s) &= \int e^{2\pi i\nu s}\, R_j(t,\nu)\, d\nu \\ &= r(t-q_jT) \int e^{2\pi i\nu s}\widehat{\varphi}(\nu - m_j\Omega)\, d\nu \\ &= r(t-q_jT)\, e^{2\pi i s m_j\Omega}\varphi(s).\end{aligned}$$

Plugging this into (10) gives the result.

D. Proof of Theorem 2.8.

Choose L prime with $A, B \leq (L-1)/2$ and $4(U/\sqrt{L}+N/L) \leq \epsilon$. We will first show that any operator in $\mathcal{H}(A,B,U,N,\epsilon)$ has the property that $supp\,\eta$ touches at most $L/2$ sets of the form

$$R_{q,m} = [0, 1/\sqrt{L}] \times [0, 1/\sqrt{L}] + (q/\sqrt{L}, m/\sqrt{L}), \quad q, m = -(L-1)/2, -(L-1)/2+1, \ldots, (L-1)/2. \tag{25}$$

To this end, note that a Jordan curve C_i with length $u_i \in ((k_i-1)/\sqrt{L}, k_i/\sqrt{L})$, $k_i \in \mathbb{N}$, touches at most $4k_i$ boxes, in fact, this bound is rather pessimistic and only sharp for $k_i = 1$. Note that

$$\sqrt{L}U \geq \sqrt{L}\sum_{i=1}^{N} u_i \geq \sqrt{L}\sum_{i=1}^{N}(k_i-1)/\sqrt{L} = \big(\sum_{i=1}^{N} k_i\big) - N,$$

and, hence, the number of boxes $B(\partial S)$ needed to cover the boundary ∂S of S satisfies

$$B(\partial S) \leq \sum_{i=1}^{N} B(C_i) \leq \sum_{i=1}^{N} 4k_i \leq 4(\sqrt{L}U + N).$$

We conclude that the "fat" boundary, that is, the $1/\sqrt{L} \times 1/\sqrt{L}$ rectification of the boundary has area bounded above by

$$4(\sqrt{L}U+N)/(\sqrt{L})^2 = 4(U/\sqrt{L}+N/L) \leq \epsilon.$$

It follows immediately, that at most $L/2$ sets $R_{q,m}$ from (25) are needed to cover S.

Now, let $\{S_m:\ m = 1, \ldots, \binom{L^2}{L}\}$ be the collection of area 1 sets that are formed by exactly L subsets of the form $R_{q,m}$ in (25). Since $OPW(S_m)$ is identifiable, there exist L-periodic sequences $(c_n^m)_n$ and $A_m, B_m > 0$ with the property that

$$A_m\|H\|_{HS} \leq \|H\sum_{n\in\mathbb{Z}} c_n^m \delta_{n/\sqrt{L}}\|_{L^2} \leq B_m\|H\|_{HS}, \quad H \in OPW^2(S_m), \quad m = 1, \ldots, \tbinom{L^2}{L}.$$

In [14] it is shown, that indeed, we can choose a single sequence $(c_n)_n = (c_n^m)_n$ so that the above holds. For this choice of (c_n), set $A = \max\{A_m,\ m = 1, \ldots, \binom{L^2}{L}\}$ and $B = \min\{B_m,\ m = 1, \ldots, \binom{L^2}{L}\}$. With this choice, we have

$$A\|H\|_{HS} \leq \|H\sum_{n\in\mathbb{Z}} c_n \delta_{n/\sqrt{L}}\|_{L^2} \leq B\|H\|_{HS}, \quad H \in \bigcup_{m=1,\ldots,\binom{L^2}{L}} OPW^2(S_m).$$

The proof is complete by observing that for $H_1, H_2 \in \mathcal{H}(A,B,U,N,\epsilon)$ (which is not a linear space), we have $H_1 - H_2 \in OPW^2(S_m)$ for some m, and, hence,

$$A\|H_1-H_2\|_{HS} \leq \|(H_1-H_2)\sum_{n\in\mathbb{Z}} c_n\delta_{n/\sqrt{L}}\|_{L^2} \leq B\|H_1-H_2\|_{HS}, \quad H_1, H_2 \in \mathcal{H}(A,B,U,N,\epsilon).$$

Clearly, this leads also to the weaker statement $(H_1 - H_2)\sum_n c_n \delta_{n/\sqrt{L}} = 0$ implies $H_1 = H_2$.

IV. Necessary and sufficient conditions on the sampling rate in operator sampling

The goal of this section is to prove Theorems 2.5 and 2.7 giving necessary and sufficient conditions on the sampling rate for operator sampling in $OPW^2(S)$.

A. *Proof of Theorem 2.5*

Since S is closed, each t-section S_t of S is closed and, hence, measurable. Therefore, $\chi_S(t,\cdot)$ is a nonnegative measurable function and $\int_{\mathbb{R}} \chi_S(t,\nu)\,d\nu \in [0,\infty]$ is well defined for all $t\in\mathbb{R}$. It suffices to show the result for $A_\infty = \big\| \int_{\mathbb{R}} \chi_S(\cdot,\nu)\,d\nu \big\|_\infty$ finite, the infinite case then follows from this.

Assume that Λ is a set of sampling with $D(\Lambda) < a_\infty < A_\infty$.

Then, we can choose a set P with positive measure and $\int_{\mathbb{R}} \chi_S(t,\nu)\,d\nu \geq a_\infty$ for all $t\in P$. Assume without loss of generality $P \subseteq [0,1]$. For any ϵ, there exist $m_t \in PW(S_t)$ with $\|m_t\|_{L^2} = 1$ and $\|m_t|_\Lambda\|_{\ell^2} \leq \epsilon$, $t\in P$. Define $\kappa_H(x,y) = m_{x-y}(y)$ for $x-y\in P$, and 0 otherwise. Then $h_H(x,t) = \kappa_H(x,x-t) = m_t(x-t)$ and $\eta_H(t,\nu) = \widehat{m}_t(\nu)$ for $t\in P$, and 0 otherwise, so $H\in OPW^2(S)$. Observe that $\|\sigma_H\|_{L^2} = \sqrt{|P|}$.

Note that it is easily seen that if $\sum_{\lambda\in\Lambda} c_\lambda\delta_\lambda$ identifies $OPW^2(S)$, then (c_λ) is bounded. Also, by hypothesis, there exists $K\in\mathbb{N}$ which bounds the cardinality of $\Lambda\cap[x,x+1]$ above for all $x\in\mathbb{R}$. We compute

$$
\begin{aligned}
\big\|H\sum_{\lambda\in\Lambda} c_\lambda\delta_\lambda\big\|_{L^2}^2 &= \int \big|\sum_{\lambda\in\Lambda} c_\lambda \kappa_H(x,\lambda)\big|^2\,dx \\
&= \int \big|\sum_{\lambda\in\Lambda} c_\lambda m_{x-\lambda}(\lambda)\big|^2\,dx \\
&\leq \|(c_\lambda)\|_{\ell^\infty}^2 \int \big|\sum_{\lambda\in\Lambda} m_{x-\lambda}(\lambda)\big|^2\,dx \\
&\leq \|(c_\lambda)\|_{\ell^\infty}^2\, K \int \sum_{\lambda\in\Lambda} |m_{x-\lambda}(\lambda)|^2\,dx \\
&= \|(c_\lambda)\|_{\ell^\infty}^2\, K \sum_{\lambda\in\Lambda} \int |m_{x-\lambda}(\lambda)|^2\,dx \\
&= \|(c_\lambda)\|_{\ell^\infty}^2\, K \sum_{\lambda\in\Lambda} \int_\lambda^{\lambda+1} |m_{x-\lambda}(\lambda)|^2\,dx \\
&= \|(c_\lambda)\|_{\ell^\infty}^2\, K \sum_{\lambda\in\Lambda} \int_0^1 |m_t(\lambda)|^2\,dt \\
&= \|(c_\lambda)\|_{\ell^\infty}^2\, K \int_0^1 \sum_{\lambda\in\Lambda} |m_t(\lambda)|^2\,dt \\
&\leq \|(c_\lambda)\|_{\ell^\infty}^2\, K \int_0^1 \epsilon^2\,dt \\
&= \|(c_\lambda)\|_{\ell^\infty}^2\, K\,\epsilon^2\,.
\end{aligned}
$$

B. Lattice tilings and proof of Theorem 2.6

In this section we will prove Theorem 2.6, but also derive results where the tiling of S is defined by arbitrary full rank lattices in $\mathbb{R}^2$. The reconstruction formulas use results from representation theory; these carry over to the higher dimensional setting if the lattice is symplectic.

As before, we assume that $S \subseteq \mathbb{R}^2$ satisfies $|S| < 1$. Suppose that for some $A = \left(\begin{smallmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{smallmatrix}\right)$ with $\det A = 1/L$, L prime, S is contained in a fundamental domain of the lattice The lattice $LA\mathbb{Z}^2$ is the so-called adjoint lattice A° of A. Indeed, $A^\circ = (1/\sqrt{L})\ (\sqrt{L}\, A)^\circ = \sqrt{L}\sqrt{L}A = LA$ (see [5] for details). We shall assume without loss of generality that $a_{11} \neq 0$. Otherwise, we could replace the first column with the second and the second with the negative of the first, leading to a different parametrization of the same lattice. Further assume that there exist t_0, ν_0, and integers $0 \leq q_j,\, m_j \leq L-1$, $0 \leq j \leq L-1$ such that with the parallelograms $P_{q,m} = A\big([0,1]^2 + (t_0, \nu_0) + (q,m)^T\big)$, $q,\, m \in \mathbb{Z}$, replacing rectangles in Theorem 2.1, we have

$$LA[0,1]^2 \cap \bigcup_{k,\ell \in \mathbb{Z}} S + LA(k,\ell)^T \subseteq \bigcup_{j=0}^{L-1} P_{q_j, m_j}. \tag{26}$$

As before, we will set

$$\Phi_j(t,s) = \int e^{2\pi i \nu s}\, \chi_{S_j}(t,\nu)\, d\nu$$

where

$$S_j = S \cap \bigcup_{k,\ell \in \mathbb{Z}} (P_{q_j, m_j} + LA(k,\ell)^T).$$

We will derive reconstruction formulas and show that if a_{12}/a_{11} is rational, then $OPW^2(S)$ can be identified with a weighted delta train and if $a_{21}a_{11}$ is rational as well, then we are assured that the coefficient sequence $(\widetilde{c}_n)$ is periodic, that is, we are in the framework of regular operator sampling.

We shall assign to each operator $H \in OPW^2(S)$ an operator in $\widetilde{H} \in OPW^2(L^{-1/2}A^{-1}S)$ and then apply the reconstruction formula in Theorem 2.1 to reconstruct $\widetilde{h} = h_{\widetilde{H}}$ of $\widetilde{H} \in OPW^2(L^{-1/2}A^{-1}S)$. From this, we will construct $h = h_H$ and therefore H.

The result is based on the existence of the operators $\mu(\sqrt{L}A)$ that appear in the following computation. The existence follows from the representation theory of the Weyl-Heisenberg group and is discussed in this setting in [10], [17]. Let $\rho(t,\nu) = e^{\pi i t\nu} T_t M_\nu$, $\eta^\#(t,\nu) = e^{-\pi i t \nu}\eta(t,\nu)$, and $B = \sqrt{L}A$. Then

$$\begin{aligned}
H &= \iint \eta(t,\nu)\, T_t M_\nu \; dt\, d\nu \\
&= \iint_S \eta(t,\nu) e^{-\pi i t\nu}\, e^{\pi i t \nu} T_t M_\nu \; dt\, d\nu \\
&= \iint_S \eta^\#(t,\nu)\; \rho(t,\nu)\, dt\, d\nu \\
&= \iint_{B^{-1}(S)} \eta^\#(B(t,\nu))\; \rho(B(t,\nu))\; dt\, d\nu \\
&= \iint \eta^\#(B(t,\nu))\; \mu(B)\rho(t,\nu)\, \mu(B)^* \; dt\, d\nu \\
&= \mu(B) \iint \eta^\#(B(t,\nu))\; \rho(t,\nu)\; dt\, d\nu\; \mu(B)^* \\
&= \mu(B)\; \widetilde{H}\; \mu(B)^*\,,
\end{aligned}$$

with $\widetilde{\eta}^{\#}(t,\nu)=\eta^{\#}(B(t,\nu))$. Setting $Q_1(t,\nu)=t$ and $Q_2(t,\nu)=\nu$ we have

$$\widetilde{\eta}(t,\nu)=e^{\pi i(t\nu-Q_1B(t,\nu)\cdot Q_2B(t,\nu))}\eta(B(t,\nu)).$$

Moreover, observe that $\widetilde{S}=B^{-1}S$ satisfies the hypothesis of Theorem 2.1 with $T=\Omega=1/\sqrt{L}$. We have therefore with an L periodic sequence $(\widetilde{c}_n)$, $\widetilde{g}=\sum\widetilde{c}_n\delta_{n\sqrt{L}}$, and $B^{-1}=\left(\begin{smallmatrix} b_{22} & -b_{12}\\ -b_{21} & b_{11}\end{smallmatrix}\right)$ the reconstruction formulas

$$\widetilde{h}(x,t)=e^{2\pi i(t+t_0)\nu_0}\sum_k\sum_{j=0}^{L-1}b_{j,k}\,\widetilde{H}\widetilde{g}(t-(q_j-k)/\sqrt{L})\,e^{-2\pi i m_j(q_j-k)/L}\,\widetilde{\Phi}_j(t{+}t_0,x-(t{+}t_0)+(q_j-k)/\sqrt{L}),$$

$$\begin{aligned}
\widetilde{\eta}(t,\nu)&=e^{2\pi i(t+t_0)\nu_0}\sum_k\sum_{j=0}^{L-1}b_{j,k}\,\widetilde{H}\widetilde{g}(t-(q_j-k)/\sqrt{L})\,e^{-2\pi i m_j(q_j-k)/L}\,\chi_{B^{-1}S_j}(t{+}t_0,\nu)\,e^{2\pi i(t+t_0-(q_j-k)/\sqrt{L})\nu}\\
&=e^{\pi i(t\nu-Q_1B(t,\nu)\cdot Q_2B(t,\nu))}\eta(B(t,\nu))
\end{aligned}$$

$$\begin{aligned}
\eta(t,\nu)&=e^{2\pi i(Q_1B^{-1}(t,\nu)+t_0)\nu_0}\,e^{-\pi i(Q_1B^{-1}(t,\nu)\cdot Q_2B^{-1}(t,\nu)-t\nu)}\sum_k\sum_{j=0}^{L-1}b_{j,k}\,\widetilde{H}\widetilde{g}(Q_1B^{-1}(t,\nu)-(q_j-k)/\sqrt{L})\\
&\qquad e^{-2\pi i m_j(q_j-k)/L}\chi_{S_j}\big((t,\nu){+}B(t_0,0)\big)\;e^{2\pi i(Q_1B^{-1}(t,\nu)+t_0-(q_j-k)/\sqrt{L})Q_2B^{-1}(t,\nu)}\\
&=e^{2\pi i(Q_1B^{-1}(t,\nu)\nu_0+t_0Q_2B^{-1}(t,\nu)+t_0\nu_0)}\,e^{\pi i(Q_1B^{-1}(t,\nu)\cdot Q_2B^{-1}(t,\nu)-t\nu)}\\
&\qquad\sum_k\sum_{j=0}^{L-1}b_{j,k}\,\widetilde{H}\widetilde{g}(Q_1B^{-1}(t,\nu)-(q_j-k)/\sqrt{L})\\
&\qquad e^{-2\pi i m_j(q_j-k)/L}\chi_{S_j}\big((t,\nu){+}B(t_0,0)\big)\;e^{-2\pi i(q_j-k)/\sqrt{L})Q_2B^{-1}(t,\nu)}\\
&=e^{2\pi i((b_{22}t-b_{12}\nu)\nu_0+t_0(b_{11}\nu-b_{21}t)+t_0\nu_0)}\,e^{\pi i((b_{22}t-b_{12}\nu)\cdot(b_{11}\nu-b_{21}t)-t\nu)}\\
&\qquad\sum_k\sum_{j=0}^{L-1}b_{j,k}\,\widetilde{H}\widetilde{g}((b_{22}t-b_{12}\nu)-(q_j-k)/\sqrt{L})\\
&\qquad e^{-2\pi i m_j(q_j-k)/L}\chi_{S_j}\big((t,\nu){+}(b_{11}t_0,b_{21}t_0)\big)\;e^{-2\pi i(q_j-k)/\sqrt{L})(b_{11}\nu-b_{21}t)}\\
&=e^{2\pi i\big(\big((a_{22}t-a_{12}\nu)\nu_0+t_0(a_{11}\nu-a_{21}t)\big)\sqrt{L}+t_0\nu_0\big)}\,e^{\pi i\big(L(a_{22}t-a_{12}\nu)\cdot(a_{11}\nu-a_{21}t)-t\nu\big)}\\
&\qquad\sum_k\sum_{j=0}^{L-1}b_{j,k}\,\widetilde{H}\widetilde{g}((a_{22}t-a_{12}\nu)\sqrt{L}-(q_j-k)/\sqrt{L})\\
&\qquad e^{-2\pi i m_j(q_j-k)/L}\chi_{S_j}(t+\sqrt{L}a_{11}t_0,\,\nu+\sqrt{L}a_{21}t_0)\;e^{-2\pi i(q_j-k)(a_{11}\nu-a_{21}t)}. \qquad (27)
\end{aligned}$$

Taking inverse Fourier transforms $\nu\to x$ on both sides gives us a formula for h, but as the right hand side contains the product of three functions in ν, the resulting formula for h does not give much insight in general. If $a_{12} = 0$ though, the above simplifies (using $a_{11}a_{22} = 1/L$) to

$$\eta(t,\nu) = \sum_k \sum_{j=0}^{L-1} b_{j,k}\, \widetilde{H}\widetilde{g}(a_{22}\sqrt{L}t - (q_j - k)/\sqrt{L})\, e^{-2\pi i m_j (q_j-k)/L} \chi_{S_j}(t + \sqrt{L}a_{11}t_0,\, \nu + \sqrt{L}a_{21}t_0)$$
$$e^{2\pi i (t_0\nu_0 a_{11}\sqrt{L} + L/2\, a_{11}a_{22}t - t/2 - (q_j-k)a_{11})\nu}\, e^{2\pi i(-\sqrt{L}a_{22}a_{21}t_0\nu_0 t^2 + t_0\nu_0 - L/2\, a_{21}a_{22}t^2 + (q_j-k)a_{21}t)}$$
$$= \sum_k \sum_{j=0}^{L-1} b_{j,k}\, \widetilde{H}\widetilde{g}(a_{22}\sqrt{L}t - (q_j - k)/\sqrt{L})\, e^{-2\pi i m_j (q_j-k)/L} \chi_{S_j}(t + \sqrt{L}a_{11}t_0,\, \nu + \sqrt{L}a_{21}t_0)$$
$$e^{2\pi i (t_0\nu_0 a_{11}\sqrt{L} - (q_j-k)a_{11})\nu}\, e^{-2\pi i(\sqrt{L}t_0\nu_0 + L/2)a_{22}a_{21}t^2}\, e^{2\pi i t_0\nu_0}\, e^{2\pi i (q_j-k)a_{21}t}$$

which leads to

$$h(x,t) = e^{-2\pi i(\sqrt{L}t_0\nu_0 + L/2)a_{22}a_{21}t^2}\, e^{2\pi i t_0\nu_0} \sum_k \sum_{j=0}^{L-1} b_{j,k}\, \widetilde{H}\widetilde{g}(\sqrt{L}(a_{22}t - (q_j-k)/L))\, e^{-2\pi i m_j(q_j-k)/L}$$
$$\Phi_j(t + \sqrt{L}a_{11}t_0 t,\, x + t_0\nu_0 a_{11}\sqrt{L} - (q_j - k)a_{11})\, e^{-2\pi i \sqrt{L}a_{21}t_0(x + t_0\nu_0 a_{11}\sqrt{L} - (q_j-k)a_{11})}\, e^{2\pi i(q_j-k)a_{21}t}$$

and, if $t_0 = 0$,

$$h(x,t) = e^{-\pi i L a_{22}a_{21}t^2} \sum_k \sum_{j=0}^{L-1} b_{j,k}\, \widetilde{H}\widetilde{g}(\sqrt{L}(a_{22}t - (q_j-k)/L))\, e^{-2\pi i m_j(q_j-k)/L}$$
$$\Phi_j(t,\, x - (q_j-k)a_{11})\, e^{2\pi i (q_j-k)a_{21}t}$$

By construction, we have $\widetilde{H}\widetilde{g} = \mu(B)^* H \mu(B)\widetilde{g}$ with $\widetilde{g} = \sum \widetilde{c}_n \delta_{n/\sqrt{L}}$. Hence, we can replace $\widetilde{H}$ in (27) by $\mu(B)^* H$ and $\widetilde{g}$ by g where $g = \mu(B)\widetilde{g}$. In the following, we will give explicit representation of $\mu(B)$ and examine $g = \mu(B)\widetilde{g}$. Note that the given reconstruction formulas hold true for any tempered distribution $g = \mu(B)\widetilde{g}$, but we are mainly interested in the case that $\mu(B)\widetilde{g}$ is discretely supported, or, better, $g = \mu(B)\widetilde{g} = \sum \widetilde{c}_n \delta_{nT}$ for some T and a periodic sequence $c = (c_n)$. In applications, this would allow us to use any hardware developed to excite an operator described in Theorem 2.1.

Recall that $B = \sqrt{L}A$, so $\det B = 1$ and we assume $b_{11} \neq 0$. We have

$$\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ b_{21}/b_{11} & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -b_{11}b_{12} & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} b_{11} & 0 \\ 0 & 1/b_{11} \end{pmatrix} \tag{28}$$

Using notation from [5], we have

$$\mu_1(\alpha) = \mu\begin{pmatrix} 1 & 0 \\ \alpha & 1 \end{pmatrix} : f \mapsto e^{\pi i \alpha (\cdot)^2} f\,,$$
$$\mathcal{F} = \mu\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} : f \mapsto \widehat{f}\,,$$
$$\mu_2(\alpha) = \mu\begin{pmatrix} \alpha & 0 \\ 0 & 1/\alpha \end{pmatrix} : f \mapsto \alpha^{-1/2} f(\,\cdot\,/\alpha),$$

hence,

$$\mu\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \mu_1(b_{21}/b_{11})\,\mathcal{F}^*\mu_1(-b_{11}b_{12})\,\mathcal{F}\,\mu_2(b_{11}) = \mu_1(a_{21}/a_{11})\,\mathcal{F}^*\mu_1(-La_{11}a_{12})\,\mathcal{F}\,\mu_2(\sqrt{L}a_{11}).$$

This leads to

$$\begin{aligned}
\mu(B)\widetilde{g} &= \mu\left(\begin{smallmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{smallmatrix}\right) \sum c_n \delta_{n/\sqrt{L}} \\
&= \mu_1(a_{21}/a_{11})\, \mathcal{F}^* \mu_1(-La_{11}a_{12})\, \mathcal{F}\, \mu_2(\sqrt{L}a_{11}) \sum c_n \delta_{n/\sqrt{L}} \\
&= (\sqrt{L}a_{11})^{-1/2}\; \mu_1(a_{21}/a_{11})\, \mathcal{F}^* \mu_1(-La_{11}a_{12})\, \mathcal{F} \sum c_n \delta_{na_{11}} \\
&= (\sqrt{L}a_{11})^{-1/2}\; \mu_1(a_{21}/a_{11})\, \mathcal{F}^* \mu_1(-La_{11}a_{12}) \sum \widehat{c}_m \delta_{m/(La_{11})} \\
&= (\sqrt{L}a_{11})^{-1/2}\; \mu_1(a_{21}/a_{11})\, \mathcal{F}^* \sum \widehat{c}_m\, e^{-\pi i La_{11}a_{12}(m/(La_{11}))^2} \delta_{m/(La_{11})} \\
&= (\sqrt{L}a_{11})^{-1/2}\; \mu_1(a_{21}/a_{11})\, \mathcal{F}^* \sum \widehat{c}_m\, e^{-2\pi i m^2 a_{12}/(2La_{11})} \delta_{m/(La_{11})}
\end{aligned}$$

where we have used the fact that the Fourier transform of a delta train of the form $\sum_{n\in\mathbb{Z}} c_n \delta_{nT}$, where $c = (c_n)$ has period L is another delta train of the same form. Specifically,

$$\mathcal{F} \sum_{n\in\mathbb{Z}} c_n \delta_{nT} = \frac{1}{LT} \sum_{m\in\mathbb{Z}} \widehat{c}_m\, \delta_{m/LT} \tag{29}$$

where $\widehat{c}$ denotes the Discrete Fourier Transform of c, that is

$$\widehat{c}_m = \sum_{k=0}^{L-1} c_k\, e^{-2\pi i k m/L}.$$

Equation (29) is a simple consequence of the fact that

$$\mathcal{F} \sum_{n\in\mathbb{Z}} \delta_{nW} = \frac{1}{W} \sum_{m} \delta_{m/W}.$$

The sequence $e^{-2\pi i m^2 a_{12}/(2La_{11})}$ is periodic in m if $e^{-2\pi i m a_{12}/(2La_{11})}$ is, that is, if a_{12}/a_{11} is rational. In the following, LCM refers to least common multiples of natural numbers, and for a rational number a, $\mathrm{q}[a]$ denotes the smallest natural number q such that qa is an integer. With this notation, $(\widehat{c'})_m = \widehat{c}_m\, e^{-2\pi i m a_{12}/(2La_{11})}$ forms a sequence with period $L' = \mathrm{LCM}\{\mathrm{q}[a_{12}/(2La_{11})],\, L\}$. Once again employing (29),

$$\begin{aligned}
\mu(B)\widetilde{g} &= (\sqrt{L}a_{11})^{-1/2}\; \mu_1(a_{21}/a_{11})\, \mathcal{F}^* \sum (\widehat{c'})_m \delta_{m/(La_{11})} \\
&= (\sqrt{L}a_{11})^{-1/2}\; \mu_1(a_{21}/a_{11}) \sum c'_n \delta_{na_{11}L/L'} \\
&= (\sqrt{L}a_{11})^{-1/2}\; \sum c'_n\, e^{\pi i a_{21}/a_{11}(na_{11}L/L')^2} \delta_{na_{11}L/L'} \\
&= (\sqrt{L}a_{11})^{-1/2}\; \sum c'_n\, e^{2\pi i n^2 a_{21}a_{11}(L/L')^2/2} \delta_{na_{11}L/L'}.
\end{aligned}$$

We conclude that $\mu(B)g = \sum \widetilde{c}_n \delta_{nT}$ with $T = a_{11}L/\mathrm{q}[a_{12}/(2La_{11})]$ if a_{12}/a_{11} is rational. Moreover, if $a_{21}a_{11}$ is rational as well, then we are assured that the coefficient sequence $(\widetilde{c}_n)$ has period

$$L'' = \mathrm{LCM}\{\mathrm{q}[a_{21}a_{11}(L/L')^2/2],\, L'\} = \mathrm{LCM}\{\mathrm{q}[a_{21}a_{11}(L/\mathrm{q}[a_{12}/(2La_{11})])^2/2],\, \mathrm{q}[a_{12}/(2La_{11})],\, L\},$$

that is, we are in the framework of regular operator sampling.

Let us consider the special case that $a_{12}/(2a_{11})$ is an integer (for example, if $a_{12} = 0$ as in Theorem 2.6), then $\mathrm{q}[a_{12}/(2La_{11})] \in \{1, L\}$, so $L' = L$ and $L'' = \mathrm{LCM}\{\mathrm{q}[a_{21}a_{11}/2],\, L\}$. If in addition $La_{21}a_{11}/2$ is an integer, then $\mathrm{q}[a_{21}a_{11}/2] \in \{1, L\}$ and $L'' = L$.

To complete the proof of Theorem 2.6, observe first that $L = L'$, and indeed $(c_n) = (c'_n)$. Consequently

$$g = \mu(B)\widetilde{g} = \sum c_n\, e^{\pi i n^2 a_{21} a_{11}} \delta_{n a_{11}}.$$

Further, observe that

$$\begin{aligned}\mu\left(\begin{smallmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{smallmatrix}\right)^* &= \mu_2(\sqrt{L}a_{11})^* \mathcal{F}^* \mu_1(-La_{11}a_{12})^* \mathcal{F}\, \mu_1(a_{21}/a_{11})^* \\ &= \mu_2(1/(\sqrt{L}a_{11}))\, \mathcal{F}^* \mu_1(La_{11}a_{12}) \mathcal{F}\, \mu_1(-a_{21}/a_{11}).\end{aligned}$$

Hence, if $a_{12} = 0$, then

$$\begin{aligned}\mu\left(\begin{smallmatrix} b_{11} & 0 \\ b_{21} & b_{22} \end{smallmatrix}\right)^* f(x) &= \mu_2(1/(\sqrt{L}a_{11}))\, \mu_1(-a_{21}/a_{11}) f(x) = (\sqrt{L}a_{11})^{1/2}\, e^{-\pi i a_{21}/a_{11} (\sqrt{L}a_{11}x)^2}\, f(\sqrt{L}a_{11}x) \\ &= (\sqrt{L}a_{11})^{1/2}\, e^{-\pi i L a_{21} a_{11} x^2}\, f(\sqrt{L}a_{11}x)\end{aligned}$$

and

$$\begin{aligned}&\mu(B)^* Hg(\sqrt{L}(a_{22}t - (q_j - k)/L)) \\ &\quad = (\sqrt{L}a_{11})^{1/2}\, e^{-\pi i L a_{21} a_{11} (\sqrt{L}(a_{22}t-(q_j-k)/L))^2}\ Hg(\sqrt{L}a_{11}\sqrt{L}(a_{22}t - (q_j - k)/L)) \\ &\quad = (\sqrt{L}a_{11})^{1/2}\, e^{-\pi i a_{21} a_{11} (La_{22}t-(q_j-k))^2}\ Hg(t - a_{11}(q_j - k))\end{aligned}$$

We conclude that

$$\begin{aligned}h(x,t) = {} & (\sqrt{L}a_{11})^{1/2} e^{-2\pi i(\sqrt{L}t_0\nu_0 + L/2)a_{22}a_{21}t^2}\, e^{2\pi i t_0 \nu_0} \\ & \sum_k \sum_{j=0}^{L-1} b_{j,k}\, e^{-\pi i a_{21} a_{11} (La_{22}t-(q_j-k))^2}\ Hg(t - a_{11}(q_j - k))\, e^{-2\pi i m_j (q_j-k)/L} \\ & \Phi_j(t + \sqrt{L}a_{11}t_0 t,\, x + t_0\nu_0 a_{11}\sqrt{L} - (q_j - k)a_{11})\, e^{-2\pi i \sqrt{L} a_{21} t_0 (x + t_0\nu_0 a_{11}\sqrt{L} - (q_j-k)a_{11})}\, e^{2\pi i (q_j-k)a_{21}t}\end{aligned}$$

and, if $t_0 = 0$,

$$\begin{aligned}h(x,t) = {} & (\sqrt{L}a_{11})^{1/2} e^{-\pi i L a_{22} a_{21} t^2} \sum_k \sum_{j=0}^{L-1} b_{j,k}\, e^{-\pi i a_{21} a_{11} (La_{22}t-(q_j-k))^2}\ Hg(t - a_{11}(q_j - k)) \\ & \Phi_j(t,\, x - (q_j - k)a_{11})\, e^{2\pi i (q_j-k)a_{21}t}\,.\end{aligned}$$

C. Gabor matrices and proof of Theorem 2.7.

Preliminary to the proof of Theorem 2.7, we present some results on finite Gabor systems and their associated matrix representations.

Definition 4.1: Let $L \in \mathbb{N}$, and $c = (c_k)_{k\in\mathbb{Z}}$ a period-L sequence be given. Define the *full Gabor system matrix* $G(c)$ to be the $L \times L^2$ matrix given by

$$G(c) = [\quad D_0\, W_L \quad | \quad D_1\, W_L \quad | \quad \cdots \quad | \quad D_{L-1}\, W_L \quad]$$

where D_k is the diagonal matrix with diagonal $T^k c = (c_{L-k},\, \ldots,\, c_{L-1},\, c_0,\, \ldots,\, c_{L-k-1})$, and where W_L is the $L \times L$ Fourier matrix $W_L = (e^{2\pi i nm/L})_{n,m=0}^{L-1}$.

Note that for $0 \le q,\, p \le L-1$, the $(q+1)$st column of the submatrix $D_p W_L$ is the vector $M^p T^q c$ where the operators M and T are as in Definition 3.1, and where $c = (c_0,\, \ldots,\, c_{L-1})$. This means that each column of the matrix $G(c)$ is a unimodular constant multiple of an element of the finite Gabor system with window c, $\{T^q M^p c\}_{q,p=0}^{L-1}$, defined in Definition 3.1.

Suppose now that we are given a particular Gabor system matrix $G(c)$ and a collection of columns from this matrix is chosen. We associate to that choice the L-tuple $\tau = (\tau_0, \tau_1, \ldots, \tau_{L-1})$, where τ_k is the number of columns chosen from the submatrix $D_k W_L$. The total number of columns chosen is given by $\|\tau\|_1$, the number of submatrices $D_k W_L$ from which any columns are chosen by $\|\tau\|_0$ (the support size of τ), and the largest number of columns chosen from any submatrix $D_k W_L$ by $\|\tau\|_\infty$. Denote by $G_0(c)$ the $L \times \|\tau\|_1$ submatrix of $G(c)$ defined by this choice of columns, and denote by $M(c)$ the collection of all $\|\tau\|_1 \times \|\tau\|_1$ submatrices of $G_0(c)$. In other words, each matrix $M \in M(c)$ corresponds to some choice of $\|\tau\|_1$ rows of $G_0(c)$. Finally, recalling that $\|c\|_0$ denotes the number of nonzero elements of the vector $(c_0, c_1, \ldots, c_{L-1})$, let

$$\mu = \min\{\|c\|_0 \colon \exists M \in M(c),\ \det M \neq 0\}.$$

In other words, given a collection of columns of $G(c)$ with associated vector τ, μ is the minimum support length of a period-L sequence c such that for some choice of $\|\tau\|_1$ rows, the resulting square matrix in $M(c)$ is nonsingular.

For example, if we take $L = 7$, and fix some sequence c of period 7, then the matrix $G(c)$ is 7×49, and each submatrix $D_k W_7$ is 7×7. A choice of 6 columns from $G(c)$ might look like this.

$$M = \left(\begin{array}{cc|ccc|c}
\omega^0 c_2 & \omega^0 c_2 & \omega^0 c_3 & \omega^0 c_3 & \omega^0 c_3 & \omega^0 c_6 \\
\omega^1 c_3 & \omega^2 c_3 & \omega^0 c_4 & \omega^2 c_4 & \omega^5 c_4 & \omega^3 c_0 \\
\omega^2 c_4 & \omega^4 c_4 & \omega^0 c_5 & \omega^4 c_5 & \omega^3 c_5 & \omega^6 c_1 \\
\omega^3 c_5 & \omega^6 c_5 & \omega^0 c_6 & \omega^6 c_6 & \omega^1 c_6 & \omega^2 c_2 \\
\omega^4 c_6 & \omega^7 c_6 & \omega^0 c_0 & \omega^1 c_0 & \omega^6 c_0 & \omega^5 c_3 \\
\omega^5 c_0 & \omega^3 c_0 & \omega^0 c_1 & \omega^3 c_1 & \omega^4 c_1 & \omega^1 c_4 \\
\omega^6 c_1 & \omega^5 c_1 & \omega^0 c_2 & \omega^5 c_2 & \omega^2 c_2 & \omega^4 c_5
\end{array}\right). \tag{30}$$

Here 2 columns have been chosen from the submatrix $D_2 W_7$, 3 from $D_3 W_7$ and 1 from $D_6 W_7$, and this choice corresponds to the vector $\tau = (0, 0, 2, 3, 0, 0, 1)$. For this example, there are 7 ways to choose 6 rows of the matrix and so the set $M(c)$ consists of 7 6×6 matrices.

In our analysis below, we shall use the following shorthand notation for a matrix structured as the one above. Namely, we will write

$$\begin{array}{ccccccc}
 & & \downarrow\downarrow & \downarrow\downarrow\downarrow & & & \downarrow \\
0 & 1 & 2 & 3 & 4 & 5 & 6 \\
1 & 2 & 3 & 4 & 5 & 6 & 0 \\
2 & 3 & 4 & 5 & 6 & 0 & 1 \\
3 & 4 & 5 & 6 & 0 & 1 & 2 \\
4 & 5 & 6 & 0 & 1 & 2 & 3 \\
5 & 6 & 0 & 1 & 2 & 3 & 4 \\
6 & 0 & 1 & 2 & 3 & 4 & 5
\end{array} \tag{31}$$

which can be further simplified by removing the irrelevant columns

$$\begin{array}{ccc}
\downarrow\downarrow & \downarrow\downarrow\downarrow & \downarrow \\
2 & 3 & 6 \\
3 & 4 & 0 \\
4 & 5 & 1 \\
5 & 6 & 2 \\
6 & 0 & 3 \\
0 & 1 & 4 \\
1 & 2 & 5
\end{array} \tag{32}$$

Our goal in this subsection is to prove the following theorem.

Theorem 4.2: Suppose that the L-vector τ describes a collection of columns chosen from a full Gabor matrix.

(1) If L is prime then $\mu \leq (\|\tau\|_1 - \|\tau\|_0) + 1$.

(2) For any $L \in \mathbb{N}$, $\mu \geq \|\tau\|_\infty$.

Remark 4.3: (a) Note that the vector τ does not completely determine the columns chosen from $G(c)$ but only their distribution within $G(c)$ and hence that the conclusions of Theorem 4.2 do not depend on the actual collection of columns chosen.

(b) It is not hard to see that the estimates in Theorem 4.2 cannot be improved. For example, if one column is chosen from distinct submatrices $D_k W_L$, then the vector τ will have $\|\tau\|_1$ non-zero entries each of which is 1. Hence $\|\tau\|_1 = \|\tau\|_0$, and $\|\tau\|_\infty = 1$. Choosing $c_0 = 1$, $c_1 = c_2 = \cdots = c_{L-1} = 0$, and choosing those rows of $G_0(c)$ in which c_0 appears, it follows that the resulting $\|\tau\|_1 \times \|\tau\|_1$ matrix M is a nonsingular diagonal matrix and hence that

$$\mu = \|\tau\|_\infty = (\|\tau\|_1 - \|\tau\|_0) + 1.$$

At the other extreme, if we choose all $\|\tau\|_1$ columns from one submatrix $D_k W_L$, then we would have $\|\tau\|_0 = 1$ and $\|\tau\|_1 = \|\tau\|_\infty$. If fewer than $\|\tau\|_1$ of the c_k are nonzero, then any choice of $\|\tau\|_1$ rows of $G_0(c)$ will contain at least one identically zero row, and hence the corresponding square matrix M would be singular. This means that

$$\mu \geq (\|\tau\|_1 - \|\tau\|_0) + 1 = \|\tau\|_1 = \|\tau\|_\infty.$$

Moreover, if L is prime we once again have equality ([14]).

In order to prove Theorem 4.2 we must recall the main result and proof from [14], namely

Theorem 4.4: If L is prime, then there exists a period-L sequence c such that every minor of the full Gabor system matrix $G(c)$ is nonzero.

The proof of the theorem involved the following steps.

1. Given any square submatrix of $G(c)$, call it M, $\det(M)$ is a homogeneous polynomial of degree L in the variables $c_0, c_1, \ldots, c_{L-1}$.

2. In order to show that this polynomial does not vanish identically, it suffices to show that there is at least one monomial in $\det(M)$ with a nonzero coefficient.

3. Such a monomial, p_M is defined recursively as follows. If $\|\tau\|_1 = 1$ then M is simply a multiple of a single variable c_j and we define $p_M = c_j$. If $\|\tau\|_1 > 1$, let c_j be the variable of lowest index appearing in M. Choose any entry of M in which c_j appears, eliminate from M the row and column containing that entry, and call the remaining matrix M'. Define $p_M = c_j\, p_{M'}$. It can be shown that p_M is independent of which entry in M is chosen at each step and only depends on the variable c_j chosen at that step.

4. The remainder of the proof of Theorem 4.4 consists of showing that the coefficient of p_M is a product of minors of W_L. Since L is prime, a classical result asserts that such minors never vanish.

Proof of Theorem 4.2: (1). Let L be prime, and assume that columns are chosen from $G(c)$ according to the vector τ. By definition, there will be at least one column chosen from $\|\tau\|_0$ distinct submatrices $D_k W_L$ of $G(c)$. This means that there are exactly $\|\tau\|_0$ distinct rows in which the variable c_0 formally appears. Choose those rows and the remaining $\|\tau\|_1 - \|\tau\|_0$ rows arbitrarily, and let M be the resulting $\|\tau\|_1 \times \|\tau\|_1$ submatrix. Proceeding now with the construction of the monomial p_M defined above, it follows that p_M will contain exactly $\|\tau\|_0$ factors of c_0 plus at most $\|\tau\|_1 - \|\tau\|_0$ other distinct factors. Hence p_M will be a monomial with at most $\|\tau\|_1 - \|\tau\|_0 + 1$ distinct variables appearing. Since L is prime, the argument of [14] shows that the coefficient on this monomial is nonzero so that $\det(M)$ is not identically zero if the remaining c_k are all set to zero. Hence

$$\mu \leq \|\tau\|_1 - \|\tau\|_0 + 1.$$

(2). Let $L \in \mathbb{N}$ be given and suppose that columns are chosen from $G(c)$ according to the vector τ. Let $\|\tau\|_1$ rows be chosen from the submatrix $G_0(c)$, and call the resulting $\|\tau\|_1 \times \|\tau\|_1$ matrix M. Any diagonal of M must have τ_k entries chosen from τ_k distinct rows of each submatrix $D_k W_L$. Hence every term in the expansion of $\det(M)$ is a multiple of a monomial with at least τ_k distinct variables appearing in it. Therefore, if fewer than $\|\tau\|_\infty$ of the c_k are non-zero, then the polynomial $\det(M)$ will vanish identically. Hence $\mu \geq \|\tau\|_\infty$.

Example 4.5: The following example will show that for arbitrarily large L there are vectors τ such that for any choice of submatrix $G_0(c)$, $\|\tau\|_\infty < \mu < \|\tau\|_1 - \|\tau\|_0 + 1$. More specifically, the following theorem holds.

Theorem 4.6: For every $L \in \mathbb{N}$ large enough, there is an L-vector τ describing a choice of columns of a full Gabor matrix $G(c)$ such that $\|\tau\|_\infty < \mu$. Moreover, if L is prime, then also $\mu < \|\tau\|_1 - \|\tau\|_0 + 1$.

Proof: In order to construct this vector τ, first choose P, $R \in \mathbb{N}$ such that $P \leq R$ and

$$\frac{R+P-1}{RP} < \frac{1}{2}.$$

Note that these imply that at least $R \geq P \geq 3$. Given $L \in \mathbb{N}$ with $L \geq 9$, we can write $L = PR + j$ uniquely for some $0 \leq j \leq R-1$. Define the L-vector τ as follows. Let $\tau_k = 2$ for $0 \leq k \leq R-1$, and for $k = mR-1$, $2 \leq m \leq P$, and let $\tau_k = 0$ otherwise. Then $\|\tau\|_0 = R+P-1$, $\|\tau\|_\infty = 2$, and $\|\tau\|_1 = 2(R+P-1)$. We will show that $\|\tau\|_\infty = 2 < 3 \leq \mu$ and that in case L is also prime, $\mu \leq R < R+P = \|\tau\|_1 - \|\tau\|_0 + 1$. In our shorthand notation, the matrix $G_0(c)$ chosen has the following form.

$$\begin{array}{ccccccccccc}
\downarrow\downarrow & \downarrow\downarrow & \downarrow\downarrow & \downarrow\downarrow & \cdots & \downarrow\downarrow & \downarrow\downarrow & \downarrow\downarrow & \downarrow\downarrow & \cdots & \downarrow\downarrow \\
0 & N-1 & N-2 & N-3 & \cdots & N-R+2 & N-R+1 & N-2R+1 & N-3R+1 & \cdots & N-PR+1 \\
1 & 0 & N-1 & N-2 & \cdots & N-R+3 & N-R+2 & N-2R+2 & N-3R+2 & \cdots & N-PR+2 \\
2 & 1 & 0 & N-1 & \cdots & N-R+4 & N-R+3 & N-2R+3 & N-3R+3 & \cdots & N-PR+3 \\
3 & 2 & 1 & 0 & \cdots & N-R+5 & N-R+4 & N-2R+4 & N-3R+4 & \cdots & N-PR+4 \\
\vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots \\
R-2 & R-3 & R-4 & R-5 & \cdots & 0 & N-1 & N-R-1 & N-2R-1 & \cdots & N-(P-1)R-1 \\
R-1 & R-2 & R-3 & R-4 & \cdots & 1 & 0 & N-1 & N-R-1 & \cdots & N-(P-2)R-1 \\
\vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots \\
2R-1 & 2R-2 & 2R-3 & 2R-4 & \cdots & R+1 & R & 0 & N-1 & \cdots & N-(P-3)R-1 \\
\vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots \\
3R-1 & 3R-2 & 3R-3 & 3R-4 & \cdots & 2R+1 & 2R & R & 0 & \cdots & N-(P-4)R-1 \\
\vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots \\
PR-1 & PR-2 & PR-3 & PR-4 & \cdots & (P-1)R+1 & (P-1)R & (P-2)R & (P-3)R & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots \\
N-1 & N-2 & N-3 & N-4 & \cdots & N-R+1 & N-R & N-2R & N-3R & \cdots & N-PR
\end{array},$$

Note that a matrix in $M(c)$ corresponds to a choice of two entries in each column of the above matrix such that no choice appears in more than one row.

In order to see the first inequality, let $c = (c_n)_{n\in\mathbb{Z}}$ be a period-L sequence, and define $G_0(c)$ to be the matrix formed by choosing $2(R+P-1)$ columns of $G(c)$ according to τ. Specifically, we choose 2 columns from each submatrix $D_k W_L$ of $G(c)$ for all those k for which $\tau_k = 2$. Now suppose that $\|c\|_0 = 2$, that is, that there are exactly two non-zero terms in the vector $(c_0, \ldots, c_{L-1})$. We will show that any choice of $2(R+P-1)$ rows of $G_0(c)$ will contain a zero row, which will imply that $\mu \geq 3$. In order to simplify the argument, let us assume without loss of generality that $c_0 \neq 0$. If not then we could replace τ by a circular shift of τ in the argument that follows. Therefore, let us assume that c_0 and c_{k_0} are the only non-zero entries of c.

Note that each variable c_j appears exactly twice in each row of $G_0(c)$ that it appears in at all, and hence that each c_j appears in at most $R+P-1$ rows of $G_0(c)$. In order for a row of $G_0(c)$ to not vanish identically, at least one of the variables c_0 or c_{k_0} must appear in that row. Since c_0 and c_{k_0} appear in at most $R+P-1$ rows, in order to choose $2(R+P-1)$ non-vanishing rows of $G_0(c)$ we must be able to choose $R+P-1$ rows containing c_0 and an additional $R+P-1$ rows containing c_{k_0}. We will show that this is not possible by showing that there must be at least one row of $G_0(c)$ in which both c_0 and c_{k_0} appear. Specifically, we will show that all of the variables $c_1, c_2, \ldots, c_{L-1}$ appear at least once in the first R rows of $G_0(c)$. Clearly, c_0 also appears in each of these rows.

In the pair of columns of $G_0(c)$ chosen from the matrix $D_0 W_L$, the variables $c_1, \ldots, c_{R-1}$ appear in the first R rows. Given $1 \leq m \leq P$, consider the pair of columns of $G_0(c)$ chosen from the matrix $D_{mR-1} W_L$. It is not hard to see that in the first R rows of these columns, the variables $c_{(P-m)R+j+1}, \ldots, c_{P-(m-1))R+j}$ appear. Consequently, as m runs from 1 through P, all of the variables $c_{j+1}, \ldots, c_{PR+j}$ will appear in the first R rows of $G_0(c)$. This completes the first part of the proof.

Now suppose that L is prime. We will show that $\mu \leq R$ by showing that we can choose $2(R+P-1)$ rows of $G_0(c)$ in such a way that the monomial p_M of the resulting square matrix M, as described in the remark following the statement of Theorem 4.4, will have no more than R distinct variables c_j appearing in it.

First, choose the $R+P-1$ rows of $G_0(c)$ in which c_0 appears. For all $1 \leq m \leq P-1$, note that c_1 appears in row $mR+1$, c_2 appears in row $mR+2$ and in general, c_k appears in row $mR+k$ for $k = 1, 2, \ldots, R-1$. Note also that c_0 does not appear in these rows. Therefore, choose those $(P-1)(R-1)$ rows of $G_0(c)$. Note that $(R+P-1)+(P-1)(R-1) = RP > 2(R+P-1)$ by our assumption at the beginning of the proof. This means that by choosing rows in this way, and eliminating some if necessary, we arrive at a square sub-matrix M of $G_0(c)$. The corresponding monomial p_M will have $R+P-1$ factors of c_0 and at most $P-1$ factors of $c_1, c_2, \ldots, c_{R-1}$, resulting in no more than R distinct variables appearing in p_M. Hence $\mu \leq R < R+P = \|\tau\|_1 - \|\tau\|_0 + 1$.

Proof of Theorem 2.7: Suppose that $S \subseteq \mathbb{R}^2$ and that for some $T > 0$ and $N \in \mathbb{N}$, S satisfies the hypotheses of Theorem 2.7. We may assume without loss of generality that in fact S_{per} can be covered by fewer than L rectangles and in fact that

$$\sum_{\{(q,m):\ R_{q,m} \cap S_{per} \neq \emptyset\}} |R_{q,m}| + \frac{2}{N} < 1.$$

If not then we may replace T by $T' = T/k$ and N by $N' = mN$ for some $k, m \in \mathbb{N}$. This leads to a finer rectification of the set S_{per}, and since $|S_{per}| = |S| < 1$, we can approximate S in such a way that the needed inequality is satisfied.

Let $L \geq N^2$ be prime and let $\Omega = 1/TL$. Then with

$$R'_{q,m} = [0,T] \times [0,\Omega] + (qT, m\Omega),$$

$q, m \in \mathbb{Z}$, it follows that each rectangle $R_{q,m}$ in the original rectification of S is covered by a collection of rectangles $R'_{q',m'}$ satisfying

$$\sum_{\{(q',m'):\ R_{q',m'} \cap R_{q,m} \neq \emptyset\}} |R_{q',m'}| \leq |R_{q,m}| + \frac{2}{L}.$$

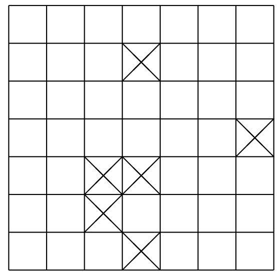

Fig. 1. Rectification of a region R with $L = 7$ and $\tau^R = (0, 0, 2, 3, 0, 0, 1)$. A corresponding collection of columns of $G(c)$ is given by 30.

Consequently the rectification of S by the rectangles $R'_{q,m}$ satisfies

$$\begin{aligned}\sum_{\{(q',m')\colon R'_{q',m'}\cap S_{per}\neq\emptyset\}} |R_{q',m'}| &\leq \sum_{\{(q,m)\colon R_{q,m}\cap S_{per}\neq\emptyset\}} \sum_{\{(q',m')\colon R'_{q',m'}\cap R_{q,m}\neq\emptyset\}} |R'_{q',m'}| \\ &\leq \sum_{\{(q,m)\colon R_{q,m}\cap S_{per}\neq\emptyset\}} \left(|R_{q,m}| + \frac{2}{L}\right) \\ &\leq \sum_{\{(q,m)\colon R_{q,m}\cap S_{per}\neq\emptyset\}} |R_{q,m}| + \frac{2N}{L} < 1.\end{aligned}$$

Therefore, since $|R'_{q,m}| = 1/L$ it follows that S_{per} is rectified by no more than L rectangles of the form $R'_{q,m}$, and in particular we can write

$$S_{per} \subseteq \bigcap_{\{j\colon R'_{q_j,m_j}\cap S_{per}\neq\emptyset\}} R'_{q_j,m_j} = R$$

for some integers $0 \leq q_j, m_j \leq L-1$. Define the L–vector $\tau^R = (\tau^R_0, \tau^R_1, \ldots, \tau^R_{L-1})$ by $\tau^R_k = \#\{j\colon m_j = k\}$. In other words, τ^R_k is the number of boxes in R of the form $R'_{q_j,k}$ (see Figure 1) and

$$\|\tau^R\|_1 = \#\{j\colon R'_{q_j,m_j} \subseteq R\}$$

is the total number of boxes $R'_{q,m}$ in R.

Since $T\Omega = 1/L = |R'_{q_j,m_j}|$,

$$\frac{\|\tau^R\|_1}{L} = \sum_{\{j\colon R'_{q_j,m_j}\subseteq R\}} |R'_{q_j,m_j}|.$$

Since $S_{per} \subseteq R$, any identifier of $OPW^2(R)$ is also an identifier of $OPW^2(S)$. Let $H \in OPW^2(S)$, and assume that L is prime. By Lemma 3.9, (19) holds for any period-L sequence $c = (c_n)$ and for all $(t,\nu) \in [0,T]\times[0,\Omega]$. By our assumptions on S, all but $\|\tau^R\|_1 \leq L$ of the entries on the right side vanish so that (19) reduces to

$$e^{-2\pi i\nu Tp}\,(Z_{1/\Omega}\circ H)g(t+Tp,\nu)$$
$$=\Omega\sum_{j=0}^{\|\tau^S\|_1-1}(T^{q_j}\,M^{m_j}c)_p\,e^{-2\pi i\nu Tq_j}\,\eta_H^{QP}(t+Tq_j,\nu+\Omega m_j), \tag{33}$$

$p=0,\,1,\,\ldots,\,L-1$, where for $0\leq j\leq\|\tau^R\|_1-1$, $R_{q_j,m_j}\subseteq R$. It is clear that the reduced system (33) is the result of choosing $\|\tau^R\|_1$ columns of the full Gabor system matrix $G(c)$ according to the L-vector τ^R. Since L is prime, Theorem 4.2 implies that there is a period-L sequence $c=(c_n)$ such that the reduced Gabor matrix $G(c)$ has full rank and

$$\|c\|_0\leq\|\tau^R\|_1-\|\tau^R\|_0+1\leq\|\tau^R\|_1.$$

Therefore,

$$\frac{\|c\|_0}{L}\leq\frac{\|\tau^R\|_1}{L}=\sum_{\{j:\,R_{q_j,m_j}\subseteq R\}}|R_{q_j,m_j}|.$$

Remark 4.7: a) Note that Theorem 2.7 does not give a sufficient sampling rate required to identify $OPW^2(S)$ but only on the relative support of the weighting sequence $c=(c_n)$. The sampling rate will of course depend on the parameter T.

(b) It is clear that the sampling rate cannot be bounded by the area of S alone. For example, if $a>0$ and $S=[0,a)\times[0,1)$, then $|S|=a$ but since $\left\|\int\chi_S(\cdot,\nu)\,d\nu\right\|=1$, Theorem 2.5 implies that any delta train identifying $OPW(S)$ must have a sampling rate of at least one sample per unit.

V. Conclusion

This paper contains results relevant to two questions on the identification and recovery of operators with bandlimited symbols from the response of the operator to a regular delta train with periodic weights. Such operators model time-variant linear communication channels. The identification and recovery procedure studied here is referred to as operator sampling. The procedure is a generalization of classical sampling results for bandlimited functions, and provides a rigorous interpretation of the determination of a time-invariant communication channel by measuring its response to a unit impulse.

We first obtain explicit reconstruction formulas in several cases: when the spreading support of the operator is compact, when it is a subset of a fundamental domain of a rectangular lattice, and when it is a subset of a fundamental domain of a general symplectic lattice. In all cases, the spreading support is required to have measure less than one, and the precise formulas depend on covering the support region efficiently by rectangles or parallelograms. For these results it is required that the support set be known. We also obtain a result showing that, under mild geometric conditions, recovery is possible when the support set is unknown but has area smaller than $1/2$. A similar result for unknown support sets of area smaller than one was proved independently in [3].

Next, we give a necessary condition on the rate of sampling, that is, the average number of deltas in the identifying weighted delta train per unit time, required to identify an operator with bandlimited symbol. The necessary rate depends on the geometry of the spreading support. Several considerations relevant to finding a sufficient condition on the sampling rate are given. Separate consideration is given to the spacing between successive deltas in the identifying delta train, which we seek to maximize, and the relative support of the weighting sequence, which we seek to minimize. We present a qualitative discussion related to maximizing the former in terms of finding the most efficient possible covering of the spreading support with rectangles or parallelograms. An asymptotic result bounding the relative support of the weighting sequence above by the area of the support set is given.

REFERENCES

[1] W.U. Bajwa, K. Gedalyahu, and Y.C. Eldar. Identification of parametric underspread linear systems and super-resolution radar. *IEEE Trans. Signal Process.*, 59(6):2548–2561, 2011.

[2] P.A. Bello. Measurement of random time-variant linear channels. *IEEE Trans. Comm.*, 15:469–475, 1969.

[3] H. Bölcskei and R. Heckel. personal communication.

[4] H. G. Feichtinger, W. Kozek, and F. Luef. Gabor analysis over finite abelian groups. *Appl. Comput. Harmon. Anal.*, 26(2):230–248, 2009.

[5] K. Gröchenig. *Foundations of Time-Frequency Analysis*. Applied and Numerical Harmonic Analysis. Birkhäuser, Boston, MA, 2001.

[6] Y. M. Hong and G. E. Pfander. Irregular and multi-channel sampling of operators. 2009. Preprint.

[7] T. Kailath. Sampling models for linear time-variant filters. Technical Report 352, Massachusetts Institute of Technology, Research Laboratory of Electronics, 1959.

[8] T. Kailath. Measurements on time–variant communication channels. *IEEE Trans. Inform. Theory*, 8(5):229–236, Sept. 1962.

[9] T. Kailath. Time–variant communication channels. *IEEE Trans. Inform. Theory: Inform. Theory. Progress Report 1960–1963*, pages 233–237, Oct. 1963.

[10] W. Kozek and G. E. Pfander. Identification of operators with bandlimited symbols. *SIAM J. Math. Anal.*, 37(3):867–888, 2006.

[11] F. Krahmer, S. Mendelson, and H. Rauhut. Suprema of chaos processes and the restricted isometry property. *Preprint*, 2012.

[12] F. Krahmer and G. E. Pfander. Localized sampling and local approximation of operators with bandlimited kohn-nirenberg symbols. preprint.

[13] F. Krahmer, G. E. Pfander, and P. Rashkov. Uncertainty in time-frequency representations on finite abelian groups and applications. *Appl. Comput. Harmon. Anal.*, 25(2):209–225, 2008.

[14] J. Lawrence, G. E. Pfander, and D. F. Walnut. Linear independence of Gabor systems in finite dimensional vector spaces. *J. Fourier Anal. Appl.*, 11(6):715–726, 2005.

[15] O Oktay, G. E. Pfander, and P. Zheltov. Scattering function estimation for overspread radar targets ,. preprint.

[16] G. E. Pfander. Measurement of time–varying Multiple–Input Multiple–Output channels. *Appl. Comp. Harm. Anal.*, 24:393–401, 2008.

[17] G. E. Pfander. Sampling of operators. http://arxiv.org/abs/1010.6165, preprint, 2010.

[18] G. E. Pfander and H. Rauhut. Sparsity in time–frequency representations. *J. Fourier Anal. Appl.*, 16(2):233–260, 2010.

[19] G. E. Pfander, H. Rauhut, and J. Tanner. Identification of matrices having a sparse representation. *IEEE Trans. Signal Process.*, 56(11):5376–5388, 2008.

[20] G. E. Pfander, H. Rauhut, and J. A. Tropp. The restricted isometry property for time-frequency structured random matrices. preprint.

[21] G. E. Pfander and D. Walnut. Measurement of time–variant channels. *IEEE Trans. Info. Theory*, 52(11):4808–4820, 2006.

[22] G. E. Pfander and D. Walnut. Operator identifcation and Feichtinger's algebra. *Sampl. Theory Signal Image Process.*, 5(2):151–168, 2006.

[23] G. E. Pfander and P. Zheltov. Identification of stochastic operators. preprint.

[24] G. E. Pfander and P. Zheltov. Sampling of stochastic operators. preprint.

[25] G.E. Pfander. Gabor frames in finite dimensions. In P.G. Casazza and G. Kutyniok, editors, *Finite Frames: Theory and Applications*. Springer, to appear.