LOCAL SAMPLING AND APPROXIMATION OF OPERATORS WITH BANDLIMITED KOHN-NIRENBERG SYMBOLS

FELIX KRAHMER AND GÖTZ E. PFANDER

ABSTRACT. Recent sampling theorems allow for the recovery of operators with bandlimited Kohn-Nirenberg symbols from their response to a single discretely supported identifier signal. The available results are inherently non-local. For example, we show that in order to recover a bandlimited operator precisely, the identifier cannot decay in time nor in frequency. Moreover, a concept of local and discrete representation is missing from the theory. In this paper, we develop tools that address these shortcomings.

We show that to obtain a local approximation of an operator, it is sufficient to test the operator on a truncated and mollified delta train, that is, on a compactly supported Schwarz class function. To compute the operator numerically, discrete measurements can be obtained from the response function which are localized in the sense that a local selection of the values yields a local approximation of the operator.

Central to our analysis is to conceptualize the meaning of localization for operators with bandlimited Kohn-Nirenberg symbol.

Keywords. Operator identification, pseudodifferential operators, Kohn-Nirenberg symbol, time frequency localization, local approximation, tight Gabor frames.

2010 Mathematics Subject Classification. Primary 41A35, 94A20; Secondary 42B35, 47B35, 47G30, 94A20.

1. INTRODUCTION

In *communications engineering*, the effect of a slowly time-varying communication channel is commonly modeled as superposition of translations (time shifts due to multipath propagation) and modulations (frequency shifts caused by Doppler effects). In order to recover transmitted signals from their channel outputs, precise knowledge of the nature of the channel is required. A common procedure for channel identification in this sense is to periodically send short duration test signals. The resulting outputs are then used to estimate channel parameters which allow for an inversion of the operator [15, 2, 16, 27, 1, 14].

Kailath [15] and Bello [2] analyzed the identifiability of such channels. In mathematical terms, the channels considered are characterized by bandlimited Kohn-Nirenberg symbols and the channel identification problem becomes an operator identification problem: can an operator with bandlimited Kohn-Nirenberg symbol be identified from the output corresponding to a given test input signal?

Kozek and Pfander [16], and Pfander and Walnut [27] gave mathematical proof of the assertions by Kailath and Bello that there exists a suitable test signal as long as the band support of the symbol of the operator has outer Jordan content less than one. The suggested test signals are periodically weighted regularly spaced Dirac-delta distributions as introduced in [27]. In [24], Pfander coined the term *operator sampling* as the resulting theory has many direct parallels to the sampling theory for bandlimited functions. For example, an operator sampling reconstruction

Felix Krahmer is with the University of Göttingen, Institute for Numerical and Applied Mathematics, Lotzestraße 16-18, 37083 Göttingen, Germany, Tel.: +49 551 39 10584, Fax: +49 551 39 3944, f.krahmer@math.uni-goettingen.de.

Götz E. Pfander is with Jacobs University Bremen, School of Engineering and Science, Campus Ring 12, 28759 Bremen, Germany. Tel.: +49 421 200 3211, Fax: +49 421 200 49 3211, g.pfander@jacobs-university.de.

Date: October 19, 2013.

formula was established which generalizes the reconstruction formula in the classical sampling theorem for bandlimited functions (see [24] and Theorem 2.2 below).

As the test signals which appear in the results of [16, 27, 24, 26] decay neither in time nor in frequency, they cannot be realized in practice. In this paper, we show that indeed, for stable identification of operator classes defined by a bandlimitation of the Kohn-Nirenberg symbol, test signals that lack decay in time and frequency are necessary. When seeking to recover only the operator's action on a time-frequency localized subspace, however, this ideal but impractical signal can be replaced with a mollified and truncated copy; the test signal can thereby be chosen to be a compactly supported Schwartz function as shown below.

Furthermore, an important difference to the sampling theory for bandlimited functions is that the response to a test signal in operator sampling is a square-integrable function rather than a discrete set of sample values. While there are many ways to discretely represent the response function, the question remains which of the multitude of commonly considered representations allow to recover the operator most efficiently. In the case of a bandlimited function, one feature that distinguishes the representation by samples is locality: a sample is the function value at a given location; due to the smoothness of bandlimited functions it represents the function in the neighborhood of the sampling point. A key consequence of this feature is that it allows to approximate the function in a given region using only samples taken in a fixed-size neighborhood of it.

In this paper we develop discrete representations of operators with bandlimited Kohn-Nirenberg symbols that, on the one hand, can be computed in a direct and simple way from the output corresponding to a test signal and, on the other hand, have locality properties analogous to those we appreciate in the classical sampling theory. We work with the same concept of locality as in the localized sampling results mentioned above, namely, locality will be defined through the action of the operator on time-frequency localized functions. Combining the two parts, we obtain that time-frequency measurements of the output corresponding to a truncated and mollified weighted sum of Dirac delta distributions yield a local discrete representation of a bandlimited operator.

The paper is organized as follows. In Section 2 we recall operator sampling terminology in some detail and discuss previous results. We then summarize our main results in Section 3. Section 4 provides results on local approximations of operators; in Section 5 we discuss identification using smooth and finite duration test signals, and Section 6 uses Gabor frames to derive our novel discretization scheme for operators with bandlimited Kohn-Nirenberg symbols.

2. Background

2.1. **Symbolic calculus.** The Schwartz kernel theorem states that every continuous linear operator $H : \mathcal{S}(\mathbb{R}) \to \mathcal{S}'(\mathbb{R})$ is of the form

$$Hf(x) = \int \kappa(x,t)f(t)dt$$

for a unique *kernel* $\kappa \in \mathcal{S}'(\mathbb{R}^2)$, where $\mathcal{S}(\mathbb{R}^d)$ is the Schwartz space, and $\mathcal{S}'(\mathbb{R}^d)$ is its dual, the space of tempered distributions [10]. This integral representation is understood in the weak sense, that is,

$$\langle Hf, g \rangle = \langle \kappa, \overline{f} \otimes g \rangle$$

for all $f, g \in \mathcal{S}(\mathbb{R})$, where $f \otimes g(x,y) = f(x)g(y)$ and $\langle \cdot, \cdot \rangle$ is the sesquilinear pairing between $\mathcal{S}(\mathbb{R})$ and $\mathcal{S}'(\mathbb{R})$, and as $\mathcal{S}(\mathbb{R})$ continuously embeds into $L^2(\mathbb{R})$, Schwartz kernel representations exist in particular for bounded operators on $L^2(\mathbb{R}^d)$.

Each such operator has consequently a *spreading function* representation

(2.1) $$Hf(x) = \iint \eta(t,\gamma)e^{2\pi i \gamma x} f(x-t) \ dt \, d\gamma,$$

a *time-varying impulse response* representation

$$Hf(x) = \int h(x,t)f(x-t)dt,$$

and a *Kohn-Nirenberg symbol* representation

(2.2)
$$Hf(x) = \int \sigma(x,\xi)\widehat{f}(\xi)\,e^{2\pi i x\xi}\,d\xi,$$

where the Fourier transform $\widehat{f}$ is normalized by

$$\mathcal{F}f(\xi) = \widehat{f}(\xi) = \int f(t)\,e^{-2\pi it\xi}\,dt$$

for integrable f.

We write H_σ and σ_H, η_H, κ_H when it is necessary to emphasize the correspondence between H and σ, η, κ. The symbols σ and η are related via the *symplectic Fourier transform* $\mathcal{F}_s$ which is defined densely by

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

that is, $\sigma = \mathcal{F}_s\eta$.

For convenience, we use the notation $\boldsymbol{\eta}_H(t,\gamma) = e^{2\pi i\gamma t}\eta_H(t,\gamma)$ and $\boldsymbol{\sigma}_H$ for its symplectic Fourier transform. A straightforward computation shows $\sigma_{H^*} = \overline{\boldsymbol{\sigma}_H}$, where H^* denotes the adjoint of H. In our proofs, we shall frequently transition from σ to $\boldsymbol{\sigma}$. This does not cause a problem in our analysis since inequality (2.4) below combined with $\|H\|_{\mathcal{L}(L^2(\mathbb{R}))} = \|H^*\|_{\mathcal{L}(L^2(\mathbb{R}))}$ shows that for $M \subseteq \mathbb{R}^2$ compact there exist $A, B > 0$ with

(2.3)
$$A\|\sigma\|_{L^\infty(\mathbb{R}^2)} \leq \|\boldsymbol{\sigma}\|_{L^\infty(\mathbb{R}^2)} \leq B\|\sigma\|_{L^\infty(\mathbb{R}^2)}$$

for all $H_\sigma \in OPW(M)$.

2.2. Sampling in operator Paley Wiener spaces.

Following [15, 2, 16, 26, 24], the operators considered in this paper are assumed to have strictly bandlimited Kohn-Nirenberg symbols, that is, they have compactly supported spreading function. Slowly time-varying mobile communications channels may violate this assumption [29]; a more refined model is that the spreading function exhibits rapid decay. Still this suffices to guarantee that truncating the spreading function introduces a global error that can be controlled. For example, applying Theorem 2.2 to an operator H whose spreading function has L^2 distance at most ϵ to a function supported on a rectangle of area one results in an operator which differs from H differs by at most 2ϵ in operator norm (see [17] for further details). This justifies to restrict to the simpler model of strictly bandlimited symbols.

The space of bounded operators whose Kohn-Nirenberg symbols are bandlimited to a given set M — we will also use the shorthand terminology *bandlimited operators* — is called *operator Paley-Wiener space*[1]; it is denoted by

$$OPW(M) = \{H \in \mathcal{L}(L^2(\mathbb{R})) : \operatorname{supp}\mathcal{F}_s\sigma_H \subseteq M\}.$$

The Kohn-Nirenberg symbol of an L^2-bounded operator with $\operatorname{supp}\mathcal{F}_s\sigma_H$ compact is bounded. In fact, for some $A, B > 0$ we have,

(2.4)
$$A\|\sigma_H\|_{L^\infty(\mathbb{R})} \leq \|H\|_{\mathcal{L}(L^2(\mathbb{R}))} \leq B\|\sigma_H\|_{L^\infty(\mathbb{R})}, \quad H \in OPW(M),$$

where $\|H\|_{\mathcal{L}(L^2(\mathbb{R}))}$ is the operator norm of H (Proposition 4.1 below).

Certainly, if we have direct access to σ_H, then some of our approximation theoretic goals can be accomplished using classical two-dimensional sampling results applied to σ_H. In the model considered here, however, we do not have access to any of the values of the symbol σ_H of the

---

[1]In general terms, operator Paley-Wiener spaces are defined by requiring its members to have bandlimited Kohn-Nirenberg symbols which are in a prescribed weighted and mixed L^p space [24]. For example, to restrict the attention to bandlimited Hilbert-Schmidt operators, we would consider only operators with square integrable symbols. These form a subset of the operators considered in this paper.

operator H directly, but we must rely on the operator output Hw which results from applying H to a single test input w. Due to stability consideration, we say that the linear space $OPW(M)$ is identifiable by w if for $A, B > 0$ we have

$$(2.5) \qquad A\|H\|_{\mathcal{L}(L^2(\mathbb{R}))} \leq \|Hw\|_{L^2(\mathbb{R})} \leq B\|H\|_{\mathcal{L}(L^2(\mathbb{R}))},$$

for all $H \in OPW(M)$ [16]. "Sampling" the operator means that the identifier w in (2.5) is a weighted sequence of Dirac delta distributions, that is,

$$w = \sum_{k \in \mathbb{Z}} c_k \delta_{kT},$$

where c_k is an appropriately chosen periodic sequence [20, 27, 24].

A guiding paradigm in the sampling theory of operators is the direct analogy to sampling of bandlimited functions. To illustrate this analogy, we compare the classical sampling theorem (often credited to Cauchy, Kotelnikov, Shannon, and Whittaker, among others), Theorem 2.1, with the corresponding result for operators, Theorem 2.2 [24]. Note that Theorem 2.1 formally follows from Theorem 2.2 by choosing the operator H in Theorem 2.2 to be the pointwise multiplication operator $f \mapsto \sigma \cdot f$ [24].

The engineering intuition underlying sampling theorems is that reducing a function to periodic samples at a rate of $1/T$ samples per unit interval corresponds to a periodization with shift $1/T$ in frequency space [22]. Thus, as long as $T\Omega \leq 1$, a function bandlimited to $[-\frac{\Omega}{2}, \frac{\Omega}{2}]$ can be recovered via a convolution with a low-pass kernel, that is, a function ϕ that satisfies

$$(2.6) \qquad \widehat{\phi}(\xi) = \begin{cases} 1/\Omega, & \text{if } |\xi| \leq \frac{\Omega}{2}, \\ 0, & \text{if } |\xi| \geq \frac{1}{2T}. \end{cases}$$

If $T\Omega = 1$, the only such function is the sinc kernel $\phi(t) = \mathrm{sinc}(\pi t/T) = \frac{\sin(\pi t/T)}{\pi t/T}$. For $T\Omega < 1$, there are many such functions; in particular ϕ in the Schwartz class is possible. With this notion, the classical sampling theorem reads as follows.

Theorem 2.1. *For $g \in L^2(\mathbb{R})$ with $\mathrm{supp}\,\mathcal{F}g \subseteq [-\frac{\Omega}{2}, \frac{\Omega}{2}]$ and $T\Omega \leq 1$, we have*

$$(2.7) \qquad g(x) = \sum_{n \in \mathbb{Z}} g(nT)\,\phi(x - nT)$$

with uniform convergence and convergence in $L^2(\mathbb{R})$. Here, ϕ is any low-pass kernel satisfying (2.6).

Recall that every operator H on $L^2(\mathbb{R})$ is in one-to-one correspondence with its kernel κ_H, that is, for a unique tempered distribution κ_H, we have $Hf(x) = \int \kappa_H(x, y)\,f(y)\,dy$ weakly. In the following, χ_A denotes the characteristic function of a set A.

Theorem 2.2. [24] *For $H : L^2(\mathbb{R}) \longrightarrow L^2(\mathbb{R})$ with $\sigma_H \in L^2(\mathbb{R}^2)$, $\mathrm{supp}\,\mathcal{F}_s\sigma_H \subseteq [0, T] \times [-\frac{\Omega}{2}, \frac{\Omega}{2}]$, and $T\Omega \leq 1$, we have*

$$(2.8) \qquad \kappa_H(x + t, x) = \chi_{[0,T]}(t) \sum_{k \in \mathbb{Z}} \left(H \sum_{n \in \mathbb{Z}} \delta_{nT} \right)(t + kT)\,\phi(x - kT),$$

with convergence in $L^2(\mathbb{R}^2)$ and uniform convergence in x for each t. Again, ϕ is any low-pass kernel satisfying (2.6).

We point to an important difference between the applicability of Theorems 2.1 and 2.2: in Theorem 2.1, a bandlimitation to a large set $[-\frac{\Omega}{2}, \frac{\Omega}{2}]$ can be resolved by choosing a small T; on the other side, Theorem 2.2 is not applicable if the bandlimiting set $[0, T] \times [-\frac{\Omega}{2}, \frac{\Omega}{2}]$ has area greater than one. Indeed, the the following is known.

Theorem 2.3. [27, 25] *$OPW(M)$ is identifiable in the sense of* (2.5) *with appropriate $w = \sum_{n \in \mathbb{Z}} c_n \delta_{nT}$ if M is compact with measure less than 1. If M is open and has area greater than 1, then there exists no tempered distribution w identifying $OPW(M)$.*

4

Hence, it is necessary to restrict ourselves to operator Paley-Wiener spaces defined by compact sets M with Lebesgue measure one. For such spaces, one can extend Theorem 2.2 to the following.

Theorem 2.4. [26] *Let M be compact with Lebesgue measure less than one. Then there exists $T, \Omega > 0$ with $T\Omega = \frac{1}{L}$, L prime, $\delta > 0$, and L-periodic sequences $\{c_n\}_n$, $\{b_{jq}\}_q$, $j = 0, 1, \ldots, n-1$, so that*

$$(2.9) \quad \kappa_H(x+t, x) = LT \sum_{j=0}^{L-1} r(t - k_j T) e^{2\pi i n_j \Omega x} \sum_{q \in \mathbb{Z}} b_{jq}$$
$$\left(H \sum_n c_n \delta_{nT} \right)(t - (k_j - q)T)\, \phi(x + (k_j - q)T), \quad H \in OPW(M),$$

where r, ϕ are Schwartz class functions that satisfy

$$r(t)\widehat{\phi}(\gamma) = 0 \text{ if } (t, \gamma) \notin (-\delta, T + \delta) \times (-\delta, \Omega + \delta),$$

and

$$(2.10) \qquad\qquad \sum_{k \in \mathbb{Z}} r(t - kT) \equiv 1 \equiv \sum_{n \in \mathbb{Z}} \widehat{\phi}(\gamma - n\Omega).$$

Moreover, (2.9) converges in $L^2(\mathbb{R})$ and uniformly in x for each t.

3. Main Results

3.1. Local representations of operators.
In classical as well as in operator sampling, working with Schwartz class kernels r, ϕ is of advantage. Indeed, in the classical sampling theorem, the slow decay of the sinc kernel in (2.7) implies that a small perturbation of just a few coefficients $g(nT)$ can lead to significant deviations of all values $g(t)$ outside of the sampling grid $T\mathbb{Z}$; this includes values achieved at locations far from the sampling points nT. Hence to approximately recover the function values locally, that is, on an compact interval, it does not suffice to know the function samples in a constant size neighborhood of that interval. When working with Schwartz class kernels, in contrast, such a local approximate reconstruction is possible; one can achieve

$$(3.1) \qquad\qquad \left| g(x) - \sum_{nT \in [a,b]} g(nT)\, \phi(x - nT) \right| < \epsilon,$$

for all $x \in [a + d(\epsilon), b - d(\epsilon)]$ where the neighborhood size $d(\epsilon)$ depends on the approximation level ϵ but not on the interval $[a, b]$.

A corresponding possibility of using local information for local reconstruction is not given in Theorem 2.2. Moreover, the identifier $w = \sum_{n \in \mathbb{Z}} \delta_{nT}$ neither decays in time or in frequency, clearly showing that in practice this input signal is not usable. However, in the framework of Theorem 2.2, this is unavoidable, as we show in the following theorem.

Theorem 3.1. *If the tempered distribution w identifies $OPW([0, T] \times [-\Omega/2, \Omega/2])$, $T\Omega > 0$, then w decays weakly neither in time nor in frequency, that is, we have neither*

$$\langle w, \varphi(\cdot - x) \rangle \overset{x \to \pm\infty}{\longrightarrow} 0 \quad nor \quad \langle \widehat{w}, \varphi(\cdot - \xi) \rangle \overset{\xi \to \pm\infty}{\longrightarrow} 0$$

for all Schwartz class functions φ.

We address this problem by developing a concept of "local recovery" of an operator, in analogy to the local recovery of a function in (3.1). Indeed, the key to most results presented in this paper is to aim only for the recovery of the operator restricted to a set of functions "localized" on a prescribed set S in the time-frequency plane. This is indeed reasonable in communications where band and time constraints on transmitted signals are frequently present. In [14], for example, operators that map bandlimited input signals to finite duration output signals are considered. Bivariate Fourier series expansions of such an operator's compactly supported Kohn-Nirenberg symbol allow the authors to discretize the a-priori continuous input-output relations (2.2) and (2.1).

Our definition of function localization in time and frequency is based on Gabor frames. It involves translation and modulation operators,

$$\mathcal{T}_t f : f \mapsto f(\cdot - t) \text{ and } \mathcal{M}_\nu : f \mapsto e^{2\pi i \nu(\cdot)} f.$$

These operators are unitary on $L^2(\mathbb{R})$ and isomorphisms on all function and distribution spaces considered in this paper.

For any $g \in L^2(\mathbb{R})$ and $a, b > 0$, we say that the Gabor system

$$(g, a\mathbb{Z} \times b\mathbb{Z}) = \{\mathcal{T}_{ka}\mathcal{M}_{\ell b}g\}_{k,\ell \in \mathbb{Z}}$$

is a tight frame for $L^2(\mathbb{R}^d)$ if for some $A > 0$, the so-called frame bound, we have

$$f = A \sum_{k,\ell \in \mathbb{Z}} \langle f, \mathcal{T}_{ka}\mathcal{M}_{\ell b}g \rangle \, \mathcal{T}_{ka}\mathcal{M}_{\ell b}g$$

for all $f \in L^2(\mathbb{R}^d)$. Each coefficient in this expansion can be interpreted to reflect the local behavior of the function near the indexing point in time-frequency space. Hence, a natural way to define time-frequency localized functions is that all but certain expansion coefficients are small.

Definition 3.2. *Let $(g, a\mathbb{Z} \times b\mathbb{Z})$, $g \in \mathcal{S}(\mathbb{R})$, be a tight frame for $L^2(\mathbb{R})$ with frame bound 1. We say that $f \in L^2(\mathbb{R})$ is ϵ–time-frequency localized on the set S if*

$$\sum_{(ka,\ell b) \in S} |\langle f, \mathcal{M}_{\ell b}\mathcal{T}_{ka}\, g \rangle|^2 \geq (1 - \epsilon^2) \sum_{(ka,\ell b) \in \mathbb{R}^2} |\langle f, \mathcal{M}_{\ell b}\mathcal{T}_{ka}\, g \rangle|^2 .$$

Our next result states that a sufficient condition for two operators to approximately agree on functions ϵ–time-frequency localized on a set S is that their Kohn-Nirenberg symbols almost agree on a neighborhood of S. Below, $B(r)$ denotes the Euclidean unit ball with radius r and center 0; the dimension will always be clear from the context. For brevity of notation, we set $S - B(r) = \left(S^c + B(r)\right)^c$ for $S \subseteq \mathbb{R}^2$.

Theorem 3.3. *Fix M compact and let $(g, a\mathbb{Z} \times b\mathbb{Z})$, $g \in \mathcal{S}(\mathbb{R})$, be a tight frame for $L^2(\mathbb{R})$ with frame bound 1. Then any pair of operators $H, \widetilde{H} \in OPW(M)$ for which one has*

$$\|\sigma_H\|_{L^\infty(\mathbb{R}^2)}, \ \|\sigma_{\widetilde{H}}\|_{L^\infty(\mathbb{R}^2)} \leq \mu \quad \text{and} \quad \|\sigma_H - \sigma_{\widetilde{H}}\|_{L^\infty(S)} \leq \epsilon \mu$$

on a set $S \subseteq \mathbb{R}^2$ satisfy

$$\|Hf - \widetilde{H}f\|_{L^2(\mathbb{R})} \leq C\,\epsilon\,\mu\,\|f\|_{L^2(\mathbb{R})}$$

for all $f \in L^2(\mathbb{R})$ that are ϵ–time-frequency localized on $S - B\big(d(\epsilon)\big)$ in the sense of Definition 3.2. Here $C > 0$ is an absolute constant and $d : (0,1) \longrightarrow \mathbb{R}^+$ satisfies $d(\epsilon) = o(\sqrt[k]{1/\epsilon})$ for all $k \in \mathbb{N}$ as $\epsilon \to 0$.

A generalization of Theorem 3.3 – labeled Theorem 4.2 – is proven in Section 4.

Our next main result concerns truncated and mollified versions of the identifier $\sum_n c_n \delta_{nT}$ and provides localized versions of Theorems 2.2 and 2.4. For $S = \mathbb{R}^2$, it reduces to Theorems 2.2 and 2.4.

Theorem 3.4. *Fix M compact with Lebesgue measure $\mu(M) < 1$ and let $(g, a\mathbb{Z} \times b\mathbb{Z})$, $g \in \mathcal{S}(\mathbb{R})$, be a tight frame for $L^2(\mathbb{R})$ with frame bound 1. Let $S \subseteq I_1 \times I_2 \subseteq \mathbb{R}^2$, where I_1 and I_2 may coincide with $\mathbb{R}$. Furthermore, choose the tempered distribution φ such that $\varphi \geq 0$ and $\widehat{\varphi} \equiv 1$ on I_2 and let*

$$\widetilde{w} = \sum_{nT \in I_1} c_n \varphi(\cdot - nT).$$

Then for any $H \in OPW(M)$ with $\|\sigma_H\|_{L^\infty(\mathbb{R}^2)} \leq \mu$ and $\widetilde{H} \in OPW(M)$ defined via

$$(3.2) \quad \kappa_{\widetilde{H}}(x + t, x) = LT \sum_{j=0}^{L-1} r(t - k_j T)\Big(\sum_{q \in \mathbb{Z}} b_{jq} H\widetilde{w}(t - (k_j - q)T)\phi(x + (k_j - q)T)\Big) e^{2\pi i n_j \Omega x},$$

6

one has

$$\|Hf - \widetilde{H}f\|_{L^2(\mathbb{R})} \le C\,\epsilon\,\mu\,\|f\|_{L^2(\mathbb{R})}$$

for all $f \in L^2(\mathbb{R})$ that are ϵ–time-frequency localized on $S - B\big(d(\epsilon)\big)$ in the sense of Definition 3.2. Here $C > 0$ is an absolute constant, r and ϕ are Schwartz class functions defined as in Theorem 2.4, but for $\delta > 0$ such that $\mu(M + [-3\delta, 3\delta]^2) < 1$, and $d : (0,1) \longrightarrow \mathbb{R}^+$ is a function independent of S which satisfies $d(\epsilon) = o(\sqrt[k]{1/\epsilon})$ for all $k \in \mathbb{N}$ as $\epsilon \to 0$.

For rectangular bandlimitation domains $M = [0,T] \times [-\frac{\Omega}{2}, \frac{\Omega}{2}]$ one can choose the identifier $\sum\limits_{nT \in I_1} \varphi(\cdot - nT)$ and define $\widetilde{H}$ via the formula

$$\kappa_{\widetilde{H}}(x + t, x) = T \sum_{n \in \mathbb{Z}} \big(H \sum_{nT \in I_1} \varphi(\cdot - nT)\big)(t + nT)\,\phi(x - nT)\,.$$

Note that this theorem is completely analogous to the condition (3.1) for localized function sampling. Due to the two-dimensional nature of the operator, however, localization is an issue in both time (restricting to a finite number of deltas) and frequency (replacing the deltas by approximate identities). If one is interested in localization only in time or only in frequency, one can choose one of the I_i to be $\mathbb{R}$ and thus consider

$$\widetilde{w} = \sum_{nT \in I_1} c_n \delta_{nT} \quad \text{or} \quad \widetilde{w} = \sum_n c_n \varphi(\cdot - nT),$$

again with $(c_n) \equiv 1$ in case of rectangular domains M.

3.2. **Local sampling of operators.** An additional important structural difference between classical sampling and operator sampling remains: in Theorems 2.2 and 3.4, the reconstruction formulas (2.8) and (2.9) involve as "coefficients" functions, not scalars. Among the many possibilities to discretely represent the operator's response to the identifier w, we consider Gabor representations of this sample function. A time-frequency localized subset of the coefficients will then yield a corresponding local approximation of the operator. Theorem 3.5 below establishes a reconstruction formula based on Gabor coefficients that allows for the exact recovery of the operator; Theorem 3.6 shows that a local subset of the coefficients yields a local approximation of the operator. Again, one obtains considerably simpler formulas for rectangular domains, but for reasons of brevity, we focus on the comprehensive setup of arbitrary domains.

For a Schwartz class function ϕ and a tempered distribution f on $\mathbb{R}$ we call

$$V_\phi f(x, \xi) = \langle f, \mathcal{M}_\xi \mathcal{T}_x \phi \rangle, \quad x, \xi \in \mathbb{R},$$

the *short-time Fourier transform* of f with respect to the window function ϕ. Throughout this paper, all pairings $\langle \cdot, \cdot \rangle$ are taken to be linear in the first component and antilinear in the second.

Theorem 3.5. *For M compact with Lebesgue measure $\mu(M) < 1$ there exists L prime, $\delta > 0$, $T, \Omega > 0$ with $T\Omega = 1/L$, and L-periodic sequences $\{c_n\}_n$, $\{b_{jq}\}_q$, $j = 0, \dots, L-1$, so that for $H \in OPW(M)$,*

(3.3) $$\sigma_H(x, \xi) = \frac{LT}{\beta_1 \beta_2} \sum_{j=0}^{L-1} e^{-2\pi i (x n_j \Omega + \xi k_j T)} e^{2\pi i n_j \Omega k_j T}$$

$$\sum_{m,\ell \in \mathbb{Z}} \sigma_{m,\ell}^{(j)} \, V_\phi r\big(x - \big(\tfrac{mL}{\beta_1} + k_j\big)T,\ \xi - \big(\tfrac{\ell L}{\beta_2} + n_j\big)\Omega\big),$$

where

$$\sigma_{m,\ell}^{(j)} = \sum_{q \in \mathbb{Z}} b_{jq}\ \phi\big((-q - k_j - mL/\beta_1)T\big)\ \langle H \sum_n c_n \delta_{nT},\ \mathcal{T}_{qT} \mathcal{M}_{\ell \Omega L / \beta_2} r \rangle,$$

and r, ϕ are Schwartz class functions such that r and $\widehat{\phi}$ are real valued and satisfy[2]

(3.4) $\qquad r(t) = 0$ if $t \notin (-\delta, \delta + T)$, $\quad \widehat{\phi}(\gamma) = 0$ if $\gamma \notin (-\delta - \Omega/2, \delta + \Omega/2)$,

and

(3.5) $$\sum_{k \in \mathbb{Z}} |r(t + kT)|^2 \equiv 1 \equiv \sum_{n \in \mathbb{Z}} |\widehat{\phi}(\gamma + n\Omega)|^2,$$

with oversampling rates $\beta_2 \geq 1 + 2\delta/T$ and $\beta_1 \geq 1 + 2\delta/\Omega$.[3]

Observe that the reconstruction formulas given in Theorems 2.4 and 3.4 require the functions r and $\widehat{\phi}$ to generate partitions of unity (3.5), while (2.10) above requires that the functions obtained by taking the square of the modulus form partitions of unity.

Theorem 3.6. *Fix M compact with $\mu(M) < 1$, let T, Ω, L and w, r, ϕ be defined in Theorem 2.4, and let $(g, a\mathbb{Z} \times b\mathbb{Z})$, $g \in \mathcal{S}(\mathbb{R})$, be a tight frame for $L^2(\mathbb{R})$ with frame bound 1.*

Then $H \in OPW(M)$ with $\|\sigma_H\|_{L^\infty(\mathbb{R}^2)} \leq \mu$, and $\widetilde{H} \in OPW(M)$ defined via its symbol

$$\widetilde{\sigma}(x, \xi) = \frac{LT}{\beta_1 \beta_2} \sum_{j=0}^{L-1} e^{-2\pi i (xn_j \Omega + \xi k_j T)} e^{2\pi i n_j \Omega k_j T} \sum_{(mLT/\beta_1, \ell L\Omega/\beta_2) \in S} \widetilde{\sigma}_{m,\ell}^{(j)} V_\phi r\left(x - \left(\frac{mL}{\beta_1} + k_j\right)T, \ \xi - \left(\frac{\ell L}{\beta_2} + n_j\right)\Omega\right),$$

where

$$\widetilde{\sigma}_{m,\ell}^{(j)} = \sum_{q \in \mathbb{Z}} b_{jq} \ \phi\big((-q - k_j - mL/\beta_1)T\big) \ \langle H\widetilde{w}, \ \mathcal{T}_{qT} \mathcal{M}_{\ell\Omega L/\beta_2} r\rangle,$$

satisfy

$$\|Hf - \widetilde{H}f\|_{L^2(\mathbb{R})} \leq C \, \varepsilon \, \mu \, \|f\|_{L^2(\mathbb{R})}$$

for all $f \in L^2(\mathbb{R})$ which are ϵ–time-frequency localized on $S - B(d(\epsilon))$ with respect to $(g, a\mathbb{Z} \times b\mathbb{Z})$ in the sense of Definition 3.2. Again $S \subseteq I_1 \times I_2 \subset \mathbb{R}^2$ is given, φ and $\widetilde{w}$ are defined as in Theorem 3.4, $C > 0$, and d can again be chosen independent of S with $d(\epsilon) = o(\sqrt[k]{1/\epsilon})$ for all $k \in \mathbb{N}$ as $\epsilon \to 0$.

The discrete representations introduced in Theorems 3.5 and 3.6 resolve a fundamental conceptual difference between classical sampling and operator sampling. In contrast to classical sampling, where the sampling values can be extracted individually, the contributions of the different Dirac-deltas in the operator sampling formula are combined in a single function and cannot easily be separated. Hence, while choosing a higher sampling rate in the function case yields more information, in the operator case, this additional information is mixed in an inseparable way. These aliasing effects [16] make it impossible to obtain redundant representations merely by oversampling in Theorem 2.2 or Theorem 2.4. In reconstruction formula (3.3), however, the oversampling parameters β_i can be chosen arbitrarily, allowing for representations of arbitrarily large redundancy.

This interplay of large redundancy and good local representation properties of the discrete coefficients can be exploited to coarsely quantize bandlimited operators, i.e., to represent these samples by values from a finite alphabet which allow for approximate recovery via the same reconstruction formulas as in Theorems 3.5 and 3.6. For such methods, as they have been studied in the mathematical literature for frame expansions over $\mathbb{R}^n$ [3, 4, 19] or the space of bounded bandlimited functions on $\mathbb{R}$ [6, 13, 7], the possibility to oversample is of crucial importance. We will, however, leave this to future work.

---

[2] For example, we can choose $r = \chi_{[0,T)} * \varphi_\delta$, where φ_δ is an approximate identity, that is, a non-negative function with $\varphi_\delta \in \mathcal{S}(\mathbb{R})$, $\operatorname{supp} \varphi_\delta \subseteq [-\delta/2, \delta/2]$, and $\int \varphi_\delta = 1$.

[3] Then the Gabor systems $\{r_{k,l} = \mathcal{T}_{kT} \mathcal{M}_{\ell/\beta_2 T} r\}_{k,\ell \in \mathbb{Z}}$, $\{\mathcal{T}_{n\Omega} \mathcal{M}_{m/\beta_1 \Omega} \widehat{\phi}\}_{m,n \in \mathbb{Z}}$, and $\{\Phi_{m,-n,l,-k} = \mathcal{T}_{(mTL/\beta_1, \ell L\Omega)} \mathcal{M}_{(n\Omega, /\beta_2, kT)}\}_{m,n,k,\ell \in \mathbb{Z}}$ are tight Gabor frames with $A = \beta_2/T$, $A = \beta_1/\Omega$, and $A = \beta_1\beta_2/(T\Omega) = \beta_1\beta_2 L$, respectively, whenever $\beta_2 \geq 1 + 2\delta/T$ and $\beta_1 \geq 1 + 2\delta/\Omega$.

4. Local approximation of bandlimited operators

In this section we show that a local approximation of an operator's symbol always yields a local approximation of the operator in the sense of Definition 3.2. The given results are of general interest and will be stated in more general terms than other results in this paper. This does not increase the difficulty of proof, but necessitates to recall additional terminology from time-frequency analysis.

For that, recall that for any full rank lattice $\Lambda = A\mathbb{Z}^{2d} \subseteq \mathbb{R}^{2d}$, $\det A \neq 0$, $\ell_s^p(\Lambda)$ denotes the set of sequences $(c_\lambda)_{\lambda \in \Lambda}$ for which

$$\|c\|_{\ell_s^p(\Lambda)} = \left(\sum_{\lambda \in \Lambda} \left| (\|\lambda\|_\infty^s + 1) c_\lambda \right|^p \right)^{1/p} < \infty.$$

A time-frequency shift by $\lambda = (t, \nu) \in \Lambda$ is denoted by $\pi(\lambda) = \mathcal{M}_\nu \mathcal{T}_t$ and in the following we will consider Gabor systems of the form $(g, \Lambda) = \{\pi(\lambda)g\}_{\lambda \in \Lambda}$.

Among the many equivalent definitions of modulation spaces, we choose the following. Let $g_0(x) = e^{-\|x\|}$, $1 \leq p \leq \infty$ and $s \in \mathbb{R}$. Then

(4.1) $\qquad M_s^p(\mathbb{R}^d) = \{ f \in \mathcal{S}'(\mathbb{R}^d) : \|f\|_{M_s^p(\mathbb{R}^d)} = \|(\langle f, \pi(\lambda)g_0 \rangle)_\lambda\|_{\ell_s^p(\frac{1}{2}\mathbb{Z}^{2d})} < \infty \},$

where we generally omit the subscript $s = 0$. For details on modulations paces, see, for example, [11, 8]. In the following we shall use the fact that whenever (g, Λ) is a tight L^2-Gabor frame (see below for a precise definition) with $g \in M^1(\mathbb{R}^d)$ then replacing the L^2-Gabor frame $(g_0, \frac{1}{2}\mathbb{Z}^{2d})$ in (4.1) with (g, Λ) leads to an equivalent norm on $M^p(\mathbb{R}^d)$ [11]. That is, there exist positive constants A and B with

(4.2) $\qquad A\|f\|_{M^p(\mathbb{R}^d)}^p \leq \sum_{\lambda \in \Lambda} |\langle f, \pi(\lambda)g \rangle|^p \leq B\|f\|_{M^p(\mathbb{R}^d)}^p, \quad f \in M^p(\mathbb{R}^d)$

if $1 \leq p < \infty$ and

$$A\|f\|_{M^\infty(\mathbb{R}^d)} \leq \sup_{\lambda \in \Lambda} |\langle f, \pi(\lambda)g \rangle| \leq B\|f\|_{M^\infty(\mathbb{R}^d)}, \quad f \in M^\infty(\mathbb{R}^d)$$

if $p = \infty$. In either case, we call (g, Λ) an ℓ^p-frame with lower frame bound A and upper frame bound B. If we can choose $A = B$ in case of $p = 2$ then we call (g, Λ) a tight Gabor frame.

The norm equivalence (2.4) follows from the following result since $M^2(\mathbb{R}) = L^2(\mathbb{R})$.

Theorem 4.1. *Let $1 \leq p \leq \infty$ and M compact. Then there exist positiv constants $A = A(M, p)$ and $B = B(M, p)$ with*

$$A\|\sigma_H\|_{L^\infty(\mathbb{R}^2)} \leq \|H\|_{\mathcal{L}(M^p(\mathbb{R}))} \leq B\|\sigma_H\|_{L^\infty(\mathbb{R}^2)}, \quad H \in OPW(M).$$

Proof. Theorem 2.7 in [24] (see for example the proof of Theorem 3.3 in [24]) provides $C = C(M, p)$ with

$$\|Hf\|_{M^p(\mathbb{R})} \leq C\|\sigma_H\|_{L^\infty(\mathbb{R}^2)}\|f\|_{M^p(\mathbb{R})}$$

for all $H \in OPW(M)$. This establishes the existence of $B = B(M, p)$ above.

In addition, we shall use the following facts. In [9, 11] it is shown that the operator norm of an operator mapping the modulation space $M^1(\mathbb{R})$ into its dual $M^\infty(\mathbb{R})$ is equivalent to the $M^\infty(\mathbb{R}^2)$ norm of its kernel κ, which can easily shown to be equivalent to the $M^\infty(\mathbb{R}^2)$ norm of the time-varying impulse response h. Moreover, we use the fact that $M^\infty(\mathbb{R}^2)$ is invariant under Fourier transforms (in some or all variables) and that the $M^\infty(\mathbb{R}^2)$ norm can be replaced by the $L^\infty(\mathbb{R}^2)$ norm if we restrict ourselves to functions bandlimited to a fixed set M [21, 24]. Last but not least, we use that the identity map embedding $M^p(\mathbb{R})$ into $M^q(\mathbb{R})$, $p \leq q$, is bounded.

Writing $\lesssim$ to express that $A \leq CB$ for some constant C depending only on the support M and $A \asymp B$ to denote equivalence in norms, i.e., $A \lesssim B$ and $B \lesssim A$, we obtain for all $H \in OPW(M)$

$$\|\sigma_H\|_{L^\infty(\mathbb{R}^2)} \asymp \|\sigma_H\|_{M^\infty(\mathbb{R}^2)} \asymp \|h_H\|_{M^\infty(\mathbb{R}^2)} \asymp \|\kappa_H\|_{M^\infty(\mathbb{R}^2)} \asymp \|H\|_{\mathcal{L}(M^1(\mathbb{R}), M^\infty(\mathbb{R}))} \lesssim \|H\|_{\mathcal{L}(M^p(\mathbb{R}))}$$

and the result follows. $\qquad \square$

We proceed to prove the following generalization of Theorem 3.3. Indeed, the earlier stated result follows again from the fact that $L^2(\mathbb{R}) = M^2(\mathbb{R})$ and $g \in \mathcal{S}(\mathbb{R})$ implies $g \in M_s^1(\mathbb{R})$ for all $s \geq 1$. We focus on the case of arbitrary domains; a simpler proof for rectangular domains can be obtained using Theorem 2.2 instead of Theorem 2.4.

Theorem 4.2. *Fix M compact and $p \in [1, \infty]$. Let (g, Λ), $g \in M_s^1(\mathbb{R})$, $s \geq 1$, be a tight frame for $L^2(\mathbb{R})$ with frame bound 1. Then any $H \in OPW(M)$ with*

$$\|\sigma_H\|_{L^\infty(\mathbb{R}^2)} \leq \mu \quad and \quad \|\sigma_H\|_{L^\infty(S)} \leq \epsilon \mu,$$

satisfies

$$\|Hf\|_{M^p(\mathbb{R})} \leq C \epsilon \mu \|f\|_{M^p(\mathbb{R})}$$

for all $f \in M^p(\mathbb{R})$ time-frequency localized on $S - B(d(\epsilon)) = \left(S^c + B(d(\epsilon)) \right)^c$ in the sense that, for $p < \infty$,

$$\sum_{\lambda \in \Lambda \cap (S - B(d(\epsilon)))} |\langle f, \pi(\lambda)g \rangle|^p \geq (1 - \epsilon^p) \sum_{\lambda \in \Lambda} |\langle f, \pi(\lambda)g \rangle|^p,$$

or, for $p = \infty$,

$$\sup\left\{ |\langle f, \pi(\lambda)g \rangle|, \ \lambda \in \Lambda \cap \left(S - B(d(\epsilon)) \right) \right\} \geq (1 - \epsilon) \sup\left\{ |\langle f, \pi(\lambda)g \rangle|, \ \lambda \in \Lambda \right\}.$$

Here $C > 0$ is an absolute constant and $d : (0,1) \longrightarrow \mathbb{R}^+$ is a function independent of S which satisfies $d(\epsilon) = o(\epsilon^{-1/s})$ as $\epsilon \to 0$.

Proof. Step 1. Preliminary observations and choice of auxiliary objects. Choose a nonnegative $\phi \in \mathcal{S}(\mathbb{R}^2)$ with $\int \phi(x)\, dx = 1$ and $\mathrm{supp}\, \phi \subseteq [-\frac{1}{2}, \frac{1}{2}]^2$. Recall that

$$\Lambda^\perp = \{\mu \in \mathbb{R}^2 : \ e^{2\pi i \langle \mu, \lambda \rangle} = 1 \text{ for all } \lambda \in \Lambda\}$$

is called dual lattice of the lattice Λ in $\mathbb{R}^2$. Let $\widetilde{\Lambda}$ be a lattice containing Λ with the property that there exists a compact and convex fundamental domain D of $\widetilde{\Lambda}^\perp$ which contains $M + [-\frac{1}{2}, \frac{1}{2}]^2$. Set

$$\sigma_P = \|\chi_D * \phi\|_{L^2(\mathbb{R}^2)}^{-1} \mathcal{F}(\chi_D * \phi)$$

and, using the sampling theorem for lattices in $\mathbb{R}^n$ [23, 11], we obtain for all $H \in OPW(M)$

$$\sigma_H = \sum_{\lambda \in \widetilde{\Lambda}} \sigma_H(\lambda)\, \mathcal{T}_\lambda \sigma_P$$

and hence

(4.3) $$H = \sum_{\lambda \in \widetilde{\Lambda}} \sigma_H(\lambda)\, \pi(\lambda) P \pi(\lambda)^*.$$

As explained above, the fact that (g, Λ) is a Gabor frame in $L^2(\mathbb{R})$ with $g \in M^1(\mathbb{R})$, implies that it is also an ℓ^p-frame for $M^p(\mathbb{R})$ and there exists $C_1, C_2 > 0$ with

(4.4) $$\|f\|_{M^p(\mathbb{R})} \leq C_1 \|\{\langle f, \pi(\lambda)g \rangle\}_{\lambda \in \Lambda}\|_{\ell^p(\Lambda)} \leq C_1 C_2 \|f\|_{M^p(\mathbb{R})}, \quad f \in M^p(\mathbb{R}).$$

As the synthesis map is the adjoint of the analysis map, we also have

(4.5) $$\Big\| \sum_{\lambda \in \Lambda} c_\lambda \pi(\lambda)g \Big\|_{M^p(\mathbb{R})} \leq C_2 \|\{c_\lambda\}_{\lambda \in \Lambda}\|_{\ell^p(\Lambda)}.$$

Since Λ is a subgroup of $\widetilde{\Lambda}$, we have $\widetilde{\Lambda} = \bigcup_{\ell=1}^n (\Lambda + \mu_\ell)$ for some $\mu_1, \mu_2, \ldots, \mu_n$. Here n is finite, as otherwise the set would be dense, hence not a discrete lattice, and depends only on M and (g, Λ). It is easily seen that $(g, \Lambda + \mu_\ell)$, $\ell = 1, \ldots, n$, also satisfies (4.4) and (4.5). Setting $\widetilde{g} = n^{-1/2} g \in M_s^1(\mathbb{R})$, we conclude that the Gabor system $(\widetilde{g}, \widetilde{\Lambda})$ is a tight frame for $L^2(\mathbb{R})$ with frame bounds equal 1 and an ℓ^p-frame with for $M^p(\mathbb{R})$ with

$$\|f\|_{M^p(\mathbb{R})} \leq C_1 n^{\frac{1}{2} - \frac{1}{p}} \|\{\langle f, \pi(\widetilde{\lambda})\widetilde{g} \rangle\}_{\widetilde{\lambda} \in \widetilde{\Lambda}}\|_{\ell^p(\widetilde{\Lambda})}$$

10

$$\leq C_1 \, n^{\frac{1}{2}-\frac{1}{p}} \, C_2 \, n^{\frac{1}{p}-\frac{1}{2}} \, \|f\|_{M^p(\mathbb{R})} = C_1 C_2 \, \|f\|_{M^p(\mathbb{R})}, \quad f \in M^p(\mathbb{R}).$$

We claim that

(4.6)
$$\left\{ \langle P\pi(\lambda)\widetilde{g}, \pi(\widetilde{\lambda})\widetilde{g} \rangle \right\} \in \ell_s^1(\widetilde{\Lambda} \times \widetilde{\Lambda}).$$

To see this, recall that $\sigma_P \in \mathcal{S}(\mathbb{R}^2) \subseteq M_s^1(\mathbb{R}^2)$, and, hence, $\widetilde{\sigma}_P$ given by $\sigma_P(x,\xi)\, e^{2\pi i x \xi}$ is in $M_s^1(\mathbb{R}^2)$ as $e^{2\pi i x \xi}$ is a Fourier multiplier and hence also a time multiplier for $M_s^1(\mathbb{R}^2)$ (Lemma 2.1 in [12], related results can be found in [5, 30, 31, 32]). A direct computation implies that for $\lambda = (t,\nu)$ and $\widetilde{\lambda} = (\widetilde{t}, \widetilde{\nu})$ we have

$$|\langle P\pi(t,\nu)\widetilde{g}, \pi(\widetilde{t},\widetilde{\nu})\widetilde{g} \rangle| = \Big| \iint \sigma_P(x,\xi)\, e^{2\pi i x \xi} \, \widehat{\mathcal{M}_\nu \mathcal{T}_t \widetilde{g}}(\xi) \overline{\mathcal{M}_{\widetilde{\nu}} \mathcal{T}_{\widetilde{t}} \widetilde{g}(x)} \, d\xi \, dx \Big|$$

$$= \Big| \iint \sigma_P(x,\xi)\, e^{2\pi i x \xi} \, \mathcal{M}_{-t}\mathcal{T}_\nu \widehat{\widetilde{g}}(\xi) \overline{\mathcal{M}_{\widetilde{\nu}} \mathcal{T}_{\widetilde{t}} \widetilde{g}(x)} \, d\xi \, dx \Big|$$

$$= \big| \langle \widetilde{\sigma}_P, \mathcal{M}_{(\widetilde{\nu},t)} \mathcal{T}_{(\widetilde{t},\nu)} \widetilde{g} \otimes \overline{\widehat{\widetilde{g}}} \rangle \big|.$$

Equation (4.2) implies that the right hand side defines an $\ell_s^1(\widetilde{\Lambda} \times \widetilde{\Lambda})$ sequence since $\widetilde{\sigma_P} \in M_s^1(\mathbb{R}^2)$ and $(g \otimes \overline{\widehat{\widetilde{g}}}, \widetilde{\Lambda} \times \widetilde{\Lambda})$ is a Gabor frame with window $\widetilde{g} \otimes \overline{\widehat{\widetilde{g}}}$ in $M_s^1(\mathbb{R}^2)$. Hence, (4.6) holds and for $\{S_k\}_{k \in \mathbb{N}_0}$ defined by

$$S_k = \sum_{\|(\lambda,\widetilde{\lambda})\|_\infty = k} |\langle P\pi(\lambda)\widetilde{g}, \pi(\widetilde{\lambda})\widetilde{g} \rangle|,$$

we have $\{(k+1)^s S_k\} \in \ell^1(\mathbb{N})$. That is, $\{S_k\} = o(k^{-(s+1)})$ and for some $C > 0$ we have $\sum_{k=K}^{\infty} S_k \leq C\, K^{-s}$, $K \in \mathbb{N}$.

For $\epsilon > 0$ set $d(\epsilon) = (C/\epsilon)^{1/s}$ and observe that then

$$\sum_{\widetilde{\lambda} \in \widetilde{\Lambda}} \sum_{\lambda \in \widetilde{\Lambda} \cap B(d(\epsilon))^c} \left| \langle P\pi(\lambda)\widetilde{g}, \pi(\widetilde{\lambda})\widetilde{g} \rangle \right| \leq \sum_{k=d(\epsilon)}^{\infty} S_k \leq C\big((C/\epsilon)^{1/s}\big)^{-s} = \epsilon.$$

Now, set $A(\lambda, \widetilde{\lambda}) = \langle P\pi(\lambda)\widetilde{g}, \pi(\widetilde{\lambda})\widetilde{g} \rangle$ if $\lambda \in \widetilde{\Lambda} \cap B(d(\epsilon))^c$ and 0 else.

Step 2. Decomposing Hf as $Hf = H_{\mathrm{in}} f_{\mathrm{in}} + H_{\mathrm{out}} f_{\mathrm{in}} + H f_{\mathrm{out}}$. We set

$$\Lambda_{\mathrm{in}} = \Lambda \cap \big(S - B(d(\epsilon))\big), \quad \widetilde{\Lambda}_{\mathrm{in}} = \widetilde{\Lambda} \cap \big(S - B(d(\epsilon))\big), \quad \Lambda_{\mathrm{out}} = \Lambda \setminus \Lambda_{\mathrm{in}}, \quad \widetilde{\Lambda}_{\mathrm{out}} = \widetilde{\Lambda} \setminus \widetilde{\Lambda}_{\mathrm{in}}.$$

and

$$f_{\mathrm{in}} = \sum_{\lambda \in \Lambda_{\mathrm{in}}} \langle f, \pi(\lambda)g \rangle \, \pi(\lambda)g = \sum_{\lambda \in \widetilde{\Lambda}_{\mathrm{in}}} c_\lambda \, \pi(\lambda)\widetilde{g}, \quad f_{\mathrm{out}} = f - f_{\mathrm{in}},$$

where $c_\lambda = \sqrt{n}\, \langle f, \pi(\lambda)g \rangle$ if $\lambda \in \Lambda$ and 0 else. Similarly, inspired by (4.3), we set for $H \in OPW(M)$

$$H_{\mathrm{in}} = \sum_{\lambda \in \widetilde{\Lambda} \cap S} \sigma_H(\lambda)\, \pi(\lambda) P \pi(\lambda)^*, \quad H_{\mathrm{out}} = H - H_{\mathrm{in}}$$

and note that $H_{\mathrm{in}}, H_{\mathrm{out}} \in OPW(D + [-\frac{1}{2}, \frac{1}{2}]^2)$.

Step 3. Bounding $\|H_{\mathrm{out}} f_{\mathrm{in}}\|_{M^p(\mathbb{R})}$. We use the separation of $\widetilde{\Lambda}_{\mathrm{in}}$ and $\widetilde{\Lambda} \cap S^c$ by $d(\epsilon)$ to compute

$$|\langle H_{\mathrm{out}} f_{\mathrm{in}}, \pi(\widetilde{\lambda})\widetilde{g} \rangle| = \Big| \big\langle \sum_{\nu \in \widetilde{\Lambda} \cap S^c} \sigma_H(\nu)\pi(\nu)P\pi(\nu)^* \sum_{\lambda \in \widetilde{\Lambda}_{\mathrm{in}}} c_\lambda \pi(\lambda)\widetilde{g}, \, \pi(\widetilde{\lambda})\widetilde{g} \big\rangle \Big|$$

$$\leq \sum_{\nu \in \widetilde{\Lambda} \cap S^c} |\sigma_H(\nu)| \sum_{\lambda \in \widetilde{\Lambda}_{\mathrm{in}}} |c_\lambda| \, \Big| \langle \pi(\nu)P\pi(\nu)^*\pi(\lambda)\widetilde{g}, \, \pi(\widetilde{\lambda})\widetilde{g} \rangle \Big|$$

$$\leq \sum_{\nu \in \widetilde{\Lambda} \cap S^c} |\sigma_H(\nu)| \sum_{\lambda \in \widetilde{\Lambda}_{\mathrm{in}}} |c_\lambda| \, \Big| \langle P\pi(\lambda - \nu)\widetilde{g}, \, \pi(\widetilde{\lambda} - \nu)\widetilde{g} \rangle \Big|$$

11

$$\leq \sum_{\nu \in \widetilde{\Lambda} \cap S^c} |\sigma_H(\nu)| \sum_{\lambda \in \widetilde{\Lambda}_{\mathrm{in}}} |c_\lambda| \, A(\lambda - \nu, \widetilde{\lambda} - \nu)$$

$$\leq \|\sigma_H\|_{L^\infty(\mathbb{R}^2)} \sum_{\nu \in \widetilde{\Lambda}} \sum_{\lambda \in \widetilde{\Lambda}} |c_\lambda| \, A(\lambda - \nu, \widetilde{\lambda} - \nu).$$

For every sequence $\{d_\lambda\} \in \ell^q(\widetilde{\Lambda})$, $1/p + 1/q = 1$, we conclude

$$\left| \langle \{ \langle H_{\mathrm{out}} f_{\mathrm{in}}, \pi(\widetilde{\lambda}) \widetilde{g} \rangle \}_{\widetilde{\lambda} \in \widetilde{\Lambda}}, \{d_{\widetilde{\lambda}}\}_{\widetilde{\lambda} \in \widetilde{\Lambda}} \rangle \right|$$

$$\leq \|\sigma_H\|_{L^\infty(\mathbb{R}^2)} \sum_{\widetilde{\lambda} \in \widetilde{\Lambda}} \sum_{\nu \in \widetilde{\Lambda}} \sum_{\lambda \in \widetilde{\Lambda}} |c_\lambda| \, A(\lambda - \nu, \widetilde{\lambda} - \nu) \, |d_{\widetilde{\lambda}}|$$

$$= \|\sigma_H\|_{L^\infty(\mathbb{R}^2)} \sum_{\lambda \in \widetilde{\Lambda}} \sum_{\widetilde{\lambda} \in \widetilde{\Lambda}} \sum_{\nu \in \widetilde{\Lambda}} |c_{\lambda+\nu}| \, A(\lambda, \widetilde{\lambda}) \, |d_{\widetilde{\lambda}+\nu}|$$

$$\leq \|\sigma_H\|_{L^\infty(\mathbb{R}^2)} \|\{c_\lambda\}\|_{\ell^p(\widetilde{\Lambda})} \|\{d_\lambda\}\|_{\ell^q(\widetilde{\Lambda})} \sum_{\lambda \in \widetilde{\Lambda}} \sum_{\widetilde{\lambda} \in \widetilde{\Lambda}} A(\lambda, \widetilde{\lambda})$$

and

$$\|H_{\mathrm{out}} f_{\mathrm{in}}\|_{M^p(\mathbb{R})} \leq n^{\frac{1}{2} - \frac{1}{p}} \, C_1 \|\{ |\langle H_{\mathrm{out}} f_{\mathrm{in}}, \pi(\widetilde{\lambda}) \widetilde{g} \rangle| \}\|_{\ell^p(\widetilde{\Lambda})}$$

$$\leq n^{\frac{1}{2} - \frac{1}{p}} \, C_1 \|\sigma_H\|_{L^\infty(\mathbb{R}^2)} \|\{c_\lambda\}\|_{\ell^p(\widetilde{\Lambda})} \sum_{\lambda \in \widetilde{\Lambda}} \sum_{\widetilde{\lambda} \in \widetilde{\Lambda}} A(\lambda, \widetilde{\lambda})$$

$$\leq n^{\frac{1}{2} - \frac{1}{p}} \, C_1 \mu \, \|\{ n^{\frac{1}{2}} \, \langle f, \pi(\lambda) g \rangle \}\|_{\ell^p(\Lambda)} \, \epsilon \leq n^{1 - \frac{1}{p}} \, C_1 \, C_2 \, \epsilon \, \mu \, \|f\|_{M^p(\mathbb{R})}.$$

Step 4. Bounding $\|H f_{\mathrm{out}}\|_{M^p(\mathbb{R})}$. By Proposition 4.1 we have

$$\|H f_{\mathrm{out}}\|_{M^p(\mathbb{R})} \leq B(M, p) \, \|\sigma_H\|_{L^\infty(\mathbb{R}^2)} \|f_{\mathrm{out}}\|_{M^p(\mathbb{R})}.$$

By hypothesis, for $p < \infty$ we have

$$\|f_{\mathrm{out}}\|_{M^p(\mathbb{R})}^p = \| \sum_{\lambda \in \Lambda_{\mathrm{out}}} \langle f, \pi(\lambda) g \rangle \pi(\lambda) g \|_{M^p(\mathbb{R})}^p \leq C_2^p \sum_{\lambda \in \Lambda_{\mathrm{out}}} |\langle f, \pi(\lambda) g \rangle|^p$$

$$\leq C_2^p \, \epsilon^p \sum_{\lambda \in \Lambda} |\langle f, \pi(\lambda) g \rangle|^p \qquad \leq C_2^{2p} \, \epsilon^p \|f\|_{M^p(\mathbb{R})}^p \,,$$

and for $p = \infty$ we have

$$\|f_{\mathrm{out}}\|_{M^\infty(\mathbb{R})} = \| \sum_{\lambda \in \Lambda_{\mathrm{out}}} \langle f, \pi(\lambda) g \rangle \pi(\lambda) g \|_{M^\infty(\mathbb{R})} \leq C_2 \, \|\langle f, \pi(\lambda) g \rangle\|_{\ell^\infty(\Lambda_{\mathrm{out}})}$$

$$\leq C_2 \, \epsilon \, \|\langle f, \pi(\lambda) g \rangle\|_{\ell^\infty(\Lambda)} \qquad \leq C_2^2 \, \epsilon \, \|f\|_{M^\infty(\mathbb{R})} \,.$$

We conclude

$$\|H f_{\mathrm{out}}\|_{M^p(\mathbb{R})} \leq B(M, p) \, C_2 \, \epsilon \, \|\sigma_H\|_{L^\infty} \|f\|_{M^p(\mathbb{R})} \leq B(M, p) \, C_2 \, \epsilon \, \mu \, \|f\|_{M^p(\mathbb{R})}.$$

Step 5. Bounding $\|H_{\mathrm{in}} f_{\mathrm{in}}\|_{M^p(\mathbb{R})}$. Since $\sigma_P \in \mathcal{S}(\mathbb{R}^2)$, the operator

$$\ell^\infty(\Lambda) \to L^\infty(\mathbb{R}^2), \quad \{c_\lambda\} \mapsto \sum_{\lambda \in \widetilde{\Lambda}} c_\lambda \, \mathcal{T}_\lambda \sigma_P$$

is bounded, say with operator norm bound C_3. Then, Proposition 4.1 implies

$$\|H_{\mathrm{in}} f_{\mathrm{in}}\|_{M^p(\mathbb{R})} \leq B(D + [-\tfrac{1}{2}, \tfrac{1}{2}]^2, p) \, \|\sigma_{H_{\mathrm{in}}}\|_{L^\infty(\mathbb{R}^2)} \|f_{\mathrm{in}}\|_{M^p(\mathbb{R})}$$

$$\leq B(D + [-\tfrac{1}{2}, \tfrac{1}{2}]^2, p) \, C_3 \, \|\{\sigma_H(\lambda)\}\|_{\ell^\infty(\widetilde{\Lambda} \cap S)} (1 + \epsilon) \|f\|_{M^p(\mathbb{R})}$$

$$\leq 2 \, B(D + [-\tfrac{1}{2}, \tfrac{1}{2}]^2, p) \, C_3 \, \epsilon \, \mu \, \|f\|_{M^p(\mathbb{R})}.$$

Since all constants are independent of ϵ, μ, H, and f, we summarize
$$\|Hf\|_{M^p(\mathbb{R})} = \|H_{\text{in}}f_{\text{in}} + H_{\text{out}}f_{\text{in}} + Hf_{\text{out}}\|_{M^p(\mathbb{R})} \le C\epsilon\,\mu\,\|f\|_{M^p(\mathbb{R})}. \qquad \square$$

5. Operator identification using localized identifiers

This section analyzes identifiers that are localized in time and frequency. Theorem 3.1 shows that such functions cannot serve as an identifier for the entire Paley-Wiener.

Proof of Theorem 3.1. Let $r \ne 0$ be a Schwartz function with $\operatorname{supp} r \subseteq [0, T]$ and $\phi \ne 0$ be a Schwartz function with $\operatorname{supp}\widehat{\phi} \subseteq [-\Omega/2, \Omega/2]$. Let H_n be defined via its kernel $\kappa_n(x, y) = \phi(x - n)r(x - y)$, so $h_n(x, t) = \phi(x - n)r(t)$ and $\eta_n(t, \nu) = \int h_n(x, t)e^{-2\pi i x\nu}dx = r(t)e^{2\pi i n\nu}\widehat{\phi}(\nu)$, so $H_n \in OPW([0, T] \times [-\Omega/2, \Omega/2])$ with $\|\sigma_{H_n}\|_{L^\infty(\mathbb{R}^2)} = \|\widehat{r}\|_{L^\infty(\mathbb{R})}\|\phi\|_{L^\infty(\mathbb{R})}$.

If w identifies $OPW([0, T] \times [-\Omega/2, \Omega/2])$, then by definition $H_n w \in L^2(\mathbb{R})$. Then

$$\int |H_n w(x)|^2\,dx = \int |\langle \kappa_n(x, y),\ w(y)\rangle_y|^2\,dx = \int |\phi(x - n)|^2\,|\langle r(x - y),\ w(y)\rangle_y|^2\,dx.$$

Clearly, $\langle r(x - y),\ w(y)\rangle_y \overset{x \to \pm\infty}{\longrightarrow} 0$ would imply $\|H_n w\|_{L^2(\mathbb{R})} \overset{n \to \pm\infty}{\longrightarrow} 0$ and contradict identifiability (2.5) since by (2.4) we have $\|H_n\|_{\mathcal{L}(L^2(\mathbb{R}))} \ge A\|\sigma_{H_n}\|_{L^\infty(\mathbb{R}^2)} = A\|\widehat{r}\|_{L^\infty(\mathbb{R})}\|\phi\|_{L^\infty(\mathbb{R})}$ for all $n \in \mathbb{Z}$.

To show that an identifier w cannot decay in frequency, we choose $H_n \in OPW([0, T] \times [-\frac{\Omega}{2}, \frac{\Omega}{2}])$ to have spreading functions $\eta_n(t, \nu) = r(t)e^{2\pi i n t}\widehat{\phi}(\nu)e^{-2\pi i t\nu}$. Let g be a Schwartz function and compute using Fubini's Theorem and, for notational simplicity, using bilinear pairings in place of sesquilinear ones,

$$\begin{aligned}
\langle H_n w(x), g(x)\rangle_x &= \langle \eta_n(t, \nu), \langle e^{2\pi i x\nu}w(x - t), g(x)\rangle_x\rangle_{t,\nu} \\
&= \langle r(t)e^{2\pi i n t}\widehat{\phi}(\nu), \langle e^{2\pi i(x-t)\nu}w(x - t), g(x)\rangle_x\rangle_{t,\nu} \\
&= \langle r(t)e^{2\pi i n t}\,w(x - t)\,g(x), \langle \widehat{\phi}(\nu), e^{2\pi i(x-t)\nu}\rangle_\nu\rangle_{t,x} \\
&= \langle \langle r(t)e^{2\pi i n t},\ w(x - t)\,\phi(x - t)\rangle_t,\ g(x)\rangle_x \\
&= \langle \langle \widehat{r}(\xi - n),\ e^{-2\pi i x\xi}\,\widehat{w} * \widehat{\phi}(\xi)\rangle_\xi,\ g(x)\rangle_x = \langle \widehat{r}(\xi - n)\,\widehat{w} * \widehat{\phi}(\xi),\ \widehat{g}(\xi)\rangle_\xi.
\end{aligned}$$

Hence,

$$\|H_n w\|_{L^2(\mathbb{R})}^2 = \|\widehat{H_n w}\|_{L^2(\mathbb{R})}^2 = \int |\widehat{r}(\xi - n)|^2\,|\langle \widehat{w}(\xi - \nu),\ \widehat{\phi}(\nu)\rangle_\nu|^2\,d\xi\,,$$

and we can conclude as above. $\qquad \square$

We proceed by showing that local identification of operators is possible with identifiers localized both in time and frequency, Theorem 3.4.

Proof of Theorem 3.4. The proof proceeds in two steps. First we show that replacing each Dirac-delta by a suitable smoothed out version locally introduces only a small error and identification using the resulting smooth identifier can be interpreted as sampling a modified bandlimited operator. Second we show that reducing to a finite number of samples also locally yields only a small error. Applying this to the modified operator arising in the first part proves that both reductions together also yield only a small error.

For the first part, choose $\varphi \in \mathcal{S}$ with $\operatorname{supp}\varphi \subseteq [-\delta, \delta]$, $\|\widehat{\varphi}\|_{L^\infty(\mathbb{R})} = 1$, and $|\widehat{\varphi}(\xi) - 1| \le \epsilon$ for $\xi \in I_2$. Define $C_\varphi : f \mapsto f * \varphi$ and set $H_C = H \circ C_\varphi$. Observe that

$$\begin{aligned}
H_C f(x) &= \iint \eta_H(t, \nu)e^{2\pi i x\nu}f * \varphi(x - t)\,dt\,d\nu \\
&= \iiint \eta_H(t, \nu)e^{2\pi i x\nu}f(x - t - y)\varphi(y)\,dy\,dt\,d\nu \\
&= \iiint \eta_H(t - y, \nu)e^{2\pi i x\nu}f(x - t)\varphi(y)\,dy\,dt\,d\nu \\
&= \iint \left(\int \eta_H(t - y, \nu)\varphi(y)\,dy\right)e^{2\pi i x\nu}f(x - t)dt\,d\nu\,,
\end{aligned}$$

that is, $\eta_{H_C}(t, \nu) = \eta_H(\cdot, \nu) * \varphi(t)$ and

$$\operatorname{supp} \eta_{H_C} \subseteq \operatorname{supp} \eta_H + [-\delta, \delta] \times \{0\}.$$

We can apply Theorem 2.4 for the operator H_C with $M_1 := M + [-\delta, \delta] \times \{0\}$ in place of M. As by assumption $M_1 + [-\delta, \delta]^2$ still has measure less than one, this can be done with δ, r and ϕ as given in the theorem. Defining $w_1 := \varphi * w$, we obtain

$$(5.1) \quad \kappa_{H_C}(x + t, x) = LT \sum_{j=0}^{L-1} r(t - k_j T) \Big(\sum_{q \in \mathbb{Z}} b_{jq} H w_1(t - (k_j - q)T) \phi(x + (k_j - q)T) \Big) e^{2\pi i n_j \Omega x}.$$

Observe that

$$\sigma_{H_C}(x, \xi) = \mathcal{F}_s \eta_{H_C}(x, \xi) = \sigma_H(x, \xi) \, \widehat{\varphi}(\xi),$$

and, by hypothesis, we have $\|\sigma_{H_C}\|_{L^\infty(\mathbb{R}^2)} \leq \|\sigma_H\|_{L^\infty(\mathbb{R}^2)} \leq \mu$ and $\|\sigma_H - \sigma_{H_C}\|_{L^\infty(S)} \leq \epsilon\mu$.

Note that for $I_1 = \mathbb{R}$, (5.1) agrees with (3.2) and we have $H_C = \widetilde{H}$, so this establishes the result.

For the second part, let us assume $S \subseteq I_1 \times \mathbb{R}$ and $M_1 \subset [c, d] \times \mathbb{R}$. Let $\psi \in \mathcal{S}(\mathbb{R})$ be nonnegative and satisfy $\sum_n \psi(x - nT) = 1$ and $\operatorname{supp} \widehat{\psi} \subset [-1/T, 1/T]$. Such a function can be obtained by choosing an arbitrary bandlimited, nonnegative $\psi_0 \in \mathcal{S}$ with $\|\psi_0\|_{L^1} = 1$ and defining $\psi = \chi_{[0, T]} * \psi_0$.

Set $P_A(x) = \sum_{nT \in A} \psi(x - nT)$, so $P_{[-N, N]} \to 1$ and $P_{[-N, N]^c} \to 0$ uniformly on compact subsets as $N \to \infty$. Moreover, $|P_A(x)| \leq 1$ for all A. Choose $N(\epsilon)$ so that $|P_{I_1 + [-N(\epsilon), N(\epsilon)]}(x) - 1| \leq \epsilon$ for $x \in I_1 + [c, d]$ and choose $R(\epsilon)$ with

$$(5.2) \quad \sum_{qT \notin I_1 + [-R(\epsilon), R(\epsilon)]} \|P_{I_1 + [-N(\epsilon), N(\epsilon)]}(x) \, V_{\phi^*} r(x - q, \xi)\|_{L^1(\mathbb{R}^2)} < \epsilon(1 - \epsilon)D.$$

where the nature of D is derived by the computations below. The existence of such $R(\epsilon)$ follows from the fact that $P_{I_1 + [-N(\epsilon), N(\epsilon)]}(x)$ and $V_{\phi^*} r$ decay faster than any polynomial. Furthermore, as $\psi \in \mathcal{S}$, a similar argument to the one given in the proof of 4.2 shows that for both $R(\epsilon)$ and $N(\epsilon)$, the growth rate is again bounded by $o(\sqrt[k]{1/\epsilon})$ for arbitrarily large $k\mathbb{N}$.

Let $w_2 = \sum_{kT \in I_1 + [-R(\epsilon), R(\epsilon)] + [-\delta, T + \delta]} c_k \delta_{kT}$ and observe that $\widetilde{H}$ as defined in the theorem satisfies

$$h_{\widetilde{H}}(x + t, t) = \kappa_{\widetilde{H}}(x + t, x)$$
$$= LT \sum_{j=0}^{L-1} r(t - k_j T) \Big(\sum_{q \in \mathbb{Z}} b_{jq} H_C w_2(t - (k_j - q)T) \phi(x + (k_j - q)T) \Big) e^{2\pi i n_j \Omega x}.$$

Since $M_1 \subset [c, d] \times \mathbb{R}$, we have $\operatorname{supp} H_C \delta_y \subseteq [c + y, d + y]$, and therefore,

$$H_C w(x) = H_C \sum_{k \in \mathbb{Z}} c_k \delta_{kT}(x) \;=\; H_C \sum_{kT \in I_1 + [-R(\epsilon), R(\epsilon)] + [-T - \delta, \delta] + [c, d]} c_k \delta_{kT}(x) = H_C w_2(x),$$
$$x \in K \equiv I_1 + [-R(\epsilon), R(\epsilon)] + [-T - \delta, \delta].$$

Note that $\widetilde{H} \in OPW(M_2)$, where $M_2 = M_1 + [-\delta, \delta]^2$ (for details, see, for example, [26]). As $M_2 + [-\delta, \delta]^2$ still has measure less than one, this implies that we can apply Theorem 2.4 again with the same δ. We obtain

$$h_{H_C}(x + t, t) - h_{\widetilde{H}}(x + t, t)$$
$$= LT \sum_{j=0}^{L-1} r(t - k_j T) \Big(\sum_{q \in \mathbb{Z}} b_{jq} H_C(w - w_2)(t - (k_j - q)T) \, \phi(x + (k_j - q)T) \Big) e^{2\pi i n_j \Omega x}$$
$$= LT \sum_{j=0}^{L-1} r(t - k_j T) \Big(\sum_{qT \notin K - (t - k_j T)} b_{jq} H_C(w - w_2)(t - (k_j - q)T) \, \phi(x + (k_j - q)T) \Big) e^{2\pi i n_j \Omega x}$$

14

$$=LT\sum_{j=0}^{L-1} r(t-k_jT)\Big(\sum_{qT\notin I_1+[-R(\epsilon),R(\epsilon)]} b_{jq}H_C\big(w-w_2\big)(t-(k_j-q)T)\,\phi(x+(k_j-q)T)\Big)e^{2\pi i n_j\Omega x}.$$

Setting $\widetilde{K}=K^c+[-\delta,T+\delta]$ and using that $(\sigma_{H_C}(x,\xi)-\sigma_{\widetilde H}(x,\xi))\,P_{I_1+[-N(\epsilon),N(\epsilon)]}(x)$ is bandlimited to $M+\{0\}\times[-1/T,1/T]$), we compute

$$\|\sigma_{H_C}-\sigma_{\widetilde H}\|_{L^\infty(S)}\le 1/(1-\epsilon)\,\|(\sigma_{H_C}(x,\xi)-\sigma_{\widetilde H}(x,\xi))\,P_{I_1+[-N(\epsilon),N(\epsilon)]}(x)\|_{L^\infty(\mathbb{R}^2)}$$
$$\asymp 1/(1-\epsilon)\,\|(\sigma_{H_C}(x,\xi)-\sigma_{\widetilde H}(x,\xi))\,P_{I_1+[-N(\epsilon),N(\epsilon)]}(x)\|_{M^\infty(\mathbb{R}^2)}$$
$$\asymp 1/(1-\epsilon)\,\|(h_{H_C}(x,t)-h_{\widetilde H}(x,t))\,P_{I_1+[-N(\epsilon),N(\epsilon)]}(x)\|_{M^\infty(\mathbb{R}^2)}$$
$$\asymp 1/(1-\epsilon)\,\Big\|LT\,P_{I_1+[-N(\epsilon),N(\epsilon)]}(x)\sum_{j=0}^{L-1}r(t-k_jT)\,e^{2\pi i n_j\Omega(x-t)}$$
$$\sum_{qT\notin I_1+[-R(\epsilon),R(\epsilon)]} b_{jq}H_C(w-w_2)(t-(k_j-q)T)\,\phi(x-t+(k_j-q)T)\Big\|_{M^\infty(\mathbb{R}^2)}$$
$$\le LT/(1-\epsilon)\sum_{j=0}^{L-1}\sum_{qT\notin I_1+[-R(\epsilon),R(\epsilon)]}\Big\|\,P_{I_1+[-N(\epsilon),N(\epsilon)]}(x)\,r(t-k_jT)\,e^{2\pi i n_j\Omega(x-t)}$$
$$b_{jq}H_C(w-w_2)(t-(k_j-q)T)\,\phi(x-t+(k_j-q)T)\Big\|_{M^\infty(\mathbb{R}^2)}$$
$$\le LT/(1-\epsilon)\sum_{j=0}^{L-1}\sum_{qT\notin I_1+[-R(\epsilon),R(\epsilon)]}\Big\|H_C(w-w_2)(t-(k_j-q)T)\Big\|_{M^\infty(\mathbb{R}^2)}$$
$$\Big\|\,P_{I_1+[-N(\epsilon),N(\epsilon)]}(x)\,r(t-k_jT)\,e^{2\pi i n_j\Omega(x-t)}\,b_{jq}\,\phi(x-t+(k_j-q)T)\Big\|_{M^1(\mathbb{R}^2)}$$
$$\le \|H_C\|_{\mathcal{L}(M^\infty(\mathbb{R}))}\,\|w-w_2\|_{M^\infty(\mathbb{R})}\frac{LT}{1-\epsilon}$$
$$\sum_{j=0}^{L-1}\sum_{qT\notin I_1+[-R(\epsilon),R(\epsilon)]}|b_{jq}|\,\Big\|P_{I_1+[-N(\epsilon),N(\epsilon)]}(x)\,r(t)\,\phi(x-t-qT)\Big\|_{M^1(\mathbb{R}^2)},$$

where we used the invariance of the M^∞ and M^1 norm under translation and modulation and, for the last inequality, Theorem 4.1 – noting that, for functions constant in one of the coordinate directions, the $M^\infty(\mathbb{R})$ and $M^\infty(\mathbb{R}^2)$ norms agree. The second to last inequality is based on $M^1(\mathbb{R}^2)$ being a Banach algebra, namely on $\|g_1 g_2\|_{M^1(\mathbb{R}^2)}\le\|g_1\|_{M^1(\mathbb{R}^2)}\|g_2\|_{M^1(\mathbb{R}^2)}$ for $g_1,g_2\in M^1(\mathbb{R}^2)$. Indeed, for $f\in M^\infty(\mathbb{R}^2)$ and $g\in M^1(\mathbb{R}^2)$, we have

$$\|fg\|_{M^\infty(\mathbb{R}^2)}=\sup_{\|\widetilde f\|_{M^1(\mathbb{R}^2)}=1}|\langle fg,\widetilde f\rangle|=\sup_{\|\widetilde f\|_{M^1(\mathbb{R}^2)}=1}|\langle f,\widetilde f\overline g\rangle|\le\sup_{\|\widetilde f\|_{M^1(\mathbb{R}^2)}=1}\|f\|_{M^\infty(\mathbb{R})}\|\widetilde f\overline g\|_{M^1(\mathbb{R}^2)}$$
$$\le\sup_{\|\widetilde f\|_{M^1(\mathbb{R}^2)}=1}\|f\|_{M^\infty(\mathbb{R})}\|\widetilde f\|_{M^1(\mathbb{R}^2)}\|\overline g\|_{M^1(\mathbb{R}^2)}=\|f\|_{M^\infty(\mathbb{R}^2)}\|g\|_{M^1(\mathbb{R}^2)}.$$

Note that with $\phi^*(t)=\overline{\phi(-t)}$, we have

$$\int r(t)\phi(x-t)e^{-2\pi it\xi}dt=V_{\phi^*}r(x,\xi),$$

which is a bandlimited function since

$$\iint V_{\phi^*}r(x,\xi)e^{2\pi it\xi-x\nu}\,dx\,d\xi=\int r(t)\phi(x-t)e^{-2\pi ix\nu}dx=r(t)\widehat\varphi(\nu)\,e^{-2\pi it\nu}.$$

Using that the M^1-norm is invariant under partial Fourier transforms and the equivalence between the M^1 and L^1 norms which is implied by the bandlimitation of $P_{I_1+[-N(\epsilon),N(\epsilon)]}(x+q)\,V_{\phi^*}r(x,\xi)$

to $(-1/T, 1/T) \times \{0\} + (-\delta, \Omega + \delta) \times (-\delta, T + \delta)$, we obtain

$$\left\| P_{I_1 + [-N(\epsilon), N(\epsilon)]}(x+q)\, r(t)\, \phi(x-t) \right\|_{M^1(\mathbb{R}^2)} \asymp \left\| P_{I_1 + [-N(\epsilon), N(\epsilon)]}(x+q)\, V_{\phi^*} r(x, \xi) \right\|_{M^1(\mathbb{R}^2)}$$

$$\asymp \left\| P_{I_1 + [-N(\epsilon), N(\epsilon)]}(x+q)\, V_{\phi^*} r(x, \xi) \right\|_{L^1(\mathbb{R}^2)}.$$

Fix $g \in \mathcal{S}(\mathbb{R})$ and observe that $\|V_g f\|_{L^p(\mathbb{R}^2)}$ defines a norm on $M^p(\mathbb{R})$ equivalent to the $M^p(\mathbb{R})$ norm given in (4.1) [11]. For any $A \subset \mathbb{R}$ we obtain the uniform bound

$$\| \sum_{nT \in A} c_n \delta_{nT} \|_{M^\infty(\mathbb{R})} \asymp \| V_g \sum_{nT \in A} c_n \delta_{nT} \|_{L^\infty(\mathbb{R})} = \| \sum_{nT \in A} c_n g(nT - t) e^{2\pi i \nu nT} \|_{L^\infty(\mathbb{R})}$$

$$\leq \| \sum_{nT \in A} |c_n|\, |g(nT - t)| \|_{L^\infty(\mathbb{R})} \leq \| \sum_{n \in \mathbb{Z}} |c_n|\, |g(nT - t)| \|_{L^\infty(\mathbb{R})} < \infty.$$

The first norm inequality stems from the fact that for all $g \in M^1(\mathbb{R})$, $\|V_g f\|_{L^p(\mathbb{R}^2)}$ defines a norm on $M^p(\mathbb{R})$ equivalent to the $M^p(\mathbb{R})$ norm given in (4.1).

Combining this upper bound on $\|w - w_2\|_{M^\infty(\mathbb{R})}$ with the above estimate for $\|\sigma_{H_C} - \sigma_{\widetilde{H}}\|_{L^\infty(S)}$ and (5.2), we conclude

$$\|\sigma_{H_C} - \sigma_{\widetilde{H}}\|_{L^\infty(S)}$$

$$\lesssim D \|H_C\|_{\mathcal{L}(M^\infty(\mathbb{R}))} \frac{L^2 T}{1 - \epsilon} \|b_{jq}\|_{\ell^\infty} \sum_{qT \notin I_1 + [-R(\epsilon), R(\epsilon)]} \left\| P_{I_1 + [-N(\epsilon), N(\epsilon)]}(x+q)\, V_{\phi^*} r(x, \xi) \right\|_{L^1(\mathbb{R}^2)}$$

$$\leq D\epsilon \|H_C\|_{\mathcal{L}(M^\infty(\mathbb{R}))} \asymp D\epsilon \|\sigma_{H_C}\|_{L^\infty(\mathbb{R}^2)} \leq D\epsilon \|\sigma_H\|_{L^\infty(\mathbb{R}^2)} \leq D\epsilon\mu.$$

Choosing $R(\epsilon)$ above large to yield D small enough to compensate all the multiplicative constants, we obtain

$$\|\sigma_{H_C} - \sigma_{\widetilde{H}}\|_{L^\infty(S)} \leq \epsilon\mu.$$

As a meaningful statement is only obtained for $\epsilon < 1$, this bound directly implies that

$$\|\sigma_{\widetilde{H}}\|_{L^\infty(\mathbb{R}^2)} \leq 2\mu.$$

Combining this with the bound

$$\|\sigma_H - \sigma_{\widetilde{H}}\|_{L^\infty(\mathbb{R}^2)} \leq \|\sigma_H - \sigma_{H_C}\|_{L^\infty(\mathbb{R}^2)} + \|\sigma_{H_C} - \sigma_{\widetilde{H}}\|_{L^\infty(\mathbb{R}^2)} \leq 2\epsilon\mu,$$

Theorem 3.3 directly yields the result with a constant of twice the size as in Theorem 3.3. □

6. Reconstruction of bandlimited operators from discrete measurements

This section concerns the discrete representation given in Theorem 3.5. First, we prove this theorem, hence establishing that indeed this representation is globally exact.

Proof of Theorem 3.5: The proof is similar to the proof of Theorem 2.4 given in [26]. The main idea is to use a Jordan domain argument to cover a fixed compact set M of size less than one by shifts of a rectangle that still have combined area less than one and then to combine identifiability results for each of them to obtain identifiability for the whole set. Indeed, there exist L prime and $T, \Omega > 0$ with $T\Omega = \frac{1}{L}$ such that

$$\text{supp}(\eta) \subseteq \bigcup_{j=0}^{L-1} R + (k_j T, n_j \Omega) \subseteq [-(L-1)T/2, (L+1)T/2] \times [-L\Omega/2, L\Omega/2]$$

$$= [-1/(2\Omega) + T/2, 1/(2\Omega) + T/2] \times [-1/(2T), 1/(2T)]$$

where $R = [0, T) \times [-\Omega/2, \Omega/2)$, and the sequence $(k_j, n_j) \in \mathbb{Z}^2$ consists of distinct pairs. For $\delta > 0$ small enough (and possibly slightly smaller T, Ω, and a larger prime L), one can even achieve

$$M_\delta \subseteq \bigcup_{j=0}^{L-1} R + (k_j T, n_j \Omega) \subseteq [-(L-1)T/2, (L+1)T/2] \times [-L\Omega/2, L\Omega/2]$$

where M_δ is the δ-neighborhood of M.

Fix such δ and let $r, \phi \in \mathcal{S}(\mathbb{R})$ satisfy (3.4) and (3.5) for this δ. Clearly,

$$(6.1) \quad (k,n) \neq (k_j, n_j) \text{ for all } j \text{ implies } S_\delta \cap \Big(R + (kT, n\Omega)\Big) = \emptyset \text{ and } \eta(t,\gamma) r(t-kT)\widehat{\phi}(\gamma - n\Omega) = 0,$$

a fact that we shall use below.

Define the identifier $w = \sum_{n \in \mathbb{Z}} c_n \delta_{nT}$, where $\{c_n\}$ is L-periodic and observe that

$$Hw(x) = \iint \eta(t,\gamma)\, e^{2\pi i \gamma x} w(x-t)\, dt\, d\gamma = \iint \boldsymbol{\eta}(t,\gamma)\, e^{2\pi i \gamma(x-t)} \sum_{k \in \mathbb{Z}} c_k \delta_{kT}(x-t)\, dt\, d\gamma$$

$$= \sum_{k \in \mathbb{Z}} c_k \int \boldsymbol{\eta}(x - kT, \gamma)\, e^{2\pi i \gamma kT} d\gamma$$

$$= \sum_{m \in \mathbb{Z}} \sum_{k=0}^{L-1} c_{k+p} \int \boldsymbol{\eta}(x - (mL+k+p)T, \gamma)\, e^{2\pi i \gamma(mL+k+p)T} d\gamma$$

for any $p \in \mathbb{Z}$. We shall use the non-normalized Zak transform $Z_{LT} : L^2(\mathbb{R}) \longrightarrow L^2\big([0, LT) \times [-\Omega/2, \Omega/2)\big)$ defined by

$$Z_{LT} f(t,\gamma) = \sum_{n \in \mathbb{Z}} f(t - nLT)\, e^{2\pi i n LT \gamma}.$$

We compute using the Poisson summation formula and the fact that $\Omega = 1/LT$

$$(Z_{LT} \circ H)w(t, \nu) = \sum_{n \in \mathbb{Z}} Hw(t - nLT)\, e^{2\pi i n LT \nu}$$

$$= \sum_{m,n \in \mathbb{Z}} e^{2\pi i T n L \nu} \sum_{k=0}^{L-1} c_{k+p} \int \boldsymbol{\eta}(t - (nL + mL + k + p)T, \gamma)\, e^{2\pi i \gamma(mL+k+p)T} d\gamma$$

$$= \sum_{k=0}^{L-1} c_{k+p} \sum_{m,n \in \mathbb{Z}} e^{2\pi i T n L \nu} \int \boldsymbol{\eta}(t - (mL + k + p)T, \gamma)\, e^{2\pi i \gamma T((m-n)L+k+p)} d\gamma$$

$$= \sum_{k=0}^{L-1} c_{k+p} \sum_{m \in \mathbb{Z}} \int \boldsymbol{\eta}(t - (mL + k + p)T, \gamma)\, e^{2\pi i \gamma(mL+k+p)T} \sum_{n \in \mathbb{Z}} e^{2\pi i n L (\nu - \gamma)T} d\gamma$$

$$= \sum_{k=0}^{L-1} c_{k+p} \sum_{m \in \mathbb{Z}} \int \boldsymbol{\eta}(t - (mL + k + p)T, \gamma)\, e^{2\pi i \gamma(mL+k+p)T} \frac{1}{LT} \sum_{n \in \mathbb{Z}} \delta_{n/LT}(\nu - \gamma) d\gamma$$

$$= \Omega \sum_{k=0}^{L-1} c_{k+p} \sum_{m,n \in \mathbb{Z}} \boldsymbol{\eta}(t - (mL + k + p)T, \nu + n\Omega)\, e^{2\pi i(\nu + \Omega n)(mL+k+p)T}$$

By (6.1) we get for $p = 0, \dots, L-1$,

$$r(t)\widehat{\phi}(\nu)(Z_{LT} \circ H)w(t + pT, \nu)$$

$$= \Omega \sum_{k=0}^{L-1} c_{k+p} \sum_{m,n \in \mathbb{Z}} r(t)\widehat{\phi}(\nu)\boldsymbol{\eta}(t - (mL + k)T, \nu + n\Omega)e^{2\pi i T(\nu + n\Omega)(mL+k+p)}$$

$$= \Omega \sum_{j=0}^{L-1} c_{p+k_j} r(t)\widehat{\phi}(\nu)\boldsymbol{\eta}(t + k_j T, \nu + n_j \Omega)\, e^{2\pi i(\nu + n_j \Omega)T(p+k_j)}.$$

$$= \Omega e^{2\pi i \nu p T} \sum_{j=0}^{L-1} (T^{k_j} M^{n_j} c)_p \Big(e^{2\pi i \nu k_j T} r(t)\widehat{\phi}(\nu)\, \boldsymbol{\eta}(t + k_j T, \nu + n_j \Omega) \Big),$$

where here and in the following, $T : (c_0, c_1, \ldots, c_{L-2}, c_{L-1}) \mapsto (c_{L-1}, c_0, \ldots, c_{L-3}, c_{L-2})$ and $M : (c_0, c_1, \ldots, c_{L-2}, c_{L-1}) \mapsto (e^{2\pi i 0/L} c_0, e^{2\pi i 1/L} c_1, \ldots, e^{2\pi i (L-2)/L} c_{L-2}, e^{2\pi i (L-1)/L} c_{L-1})$, that is, $(T^{k_j} M^{n_j} c)_p = e^{2\pi i \frac{n_j(p+k_j)}{L}} c_{p+k_j}$. Equivalently, we obtain the matrix equation

$$(6.2) \quad [e^{-2\pi i \nu p T} r(t) \widehat{\phi}(\nu)(Z_{LT} \circ H) w(t + pT, \nu)]_{p=0}^{L-1}$$
$$= \Omega \mathbf{A} [e^{2\pi i \nu k_j T} r(t) \widehat{\phi}(\nu) \boldsymbol{\eta}(t + k_j T, \nu + n_j \Omega)]_{j=0}^{L-1}$$

where $\mathbf{A}$ is a $L \times L$ matrix, whose jth column is $T^{k_j} M^{n_j} c \in \mathbb{C}^L$. $\mathbf{A}$ is a submatrix of the $L \times L^2$ marix $\mathbf{G}$, whose columns are $\{T^k M^l c\}_{k,l=0}^{L-1}$. It was shown in [18] that if L is prime, then we can choose $c \in \mathbb{C}^L$ such that every $L \times L$ submatrix of $\mathbf{G}$ is invertible. In fact, the set of such $c \in \mathbb{C}^L$ is a dense open subset of $\mathbb{C}^L$ [18]. Hence we can apply the matrix $\mathbf{A}^{-1} =: [b_{jp}]_{j,p=1}^L$ on both sides of Equation (6.2) to obtain

$$(6.3) \quad e^{2\pi i \nu k_j T} r(t) \widehat{\phi}(\nu) \boldsymbol{\eta}(t + k_j T, \nu + n_j \Omega) = LT \sum_{p=0}^{L-1} b_{jp} e^{-2\pi i \nu p T} r(t) \widehat{\phi}(\nu)(Z_{LT} \circ H) w(t + pT, \nu)$$

for every $j = 0, 1, \ldots, L-1$.

In fact, until this point the proof agrees with the proof of (2.9) in Theorem 2.4. Indeed, if we extend $\{b_{jp}\}_p$ to a L-periodic sequence by setting $b_{j,p+mL} = b_{jp}$, replace the so far unused property (3.5) by (2.10) then further computations [24] give

$$h(x, t) = LT \sum_{j=0}^{L-1} r(t - k_j T) \Big(\sum_{q \in \mathbb{Z}} b_{jq} H w(t - (k_j + q)T) \phi(x - t + (k_j + q)T) \Big) e^{2\pi i n_j \Omega (x - t)}.$$

Observe that (3.5) implies that $(r, T\mathbb{Z} \times \frac{\Omega L}{\beta_2} \mathbb{Z}) = \{\mathcal{T}_{kT} \mathcal{M}_{\ell L \Omega/\beta_2} r\}_{k,\ell \in \mathbb{Z}}$ is a tight Gabor frame whenever $\beta_2 \geq 1 + 2\delta/T$ as, in this case, $(r, \frac{\beta_2}{\Omega L} \mathbb{Z} \times \frac{1}{T} \mathbb{Z}) = (r, \beta_2 T \mathbb{Z} \times \Omega L \mathbb{Z})$ is an orthogonal sequence and the Ron-Shen criterion applies [11, 28]. The same arguments imply that $(\widehat{\phi}, \Omega \mathbb{Z} \times \frac{LT}{\beta_1} \mathbb{Z})$ is a tight Gabor frame. Using a simple tensor argument, we obtain that $\{\Psi_{m,n,l,k}\}_{m,n,l,k \in \mathbb{Z}}$ forms a tight Gabor frame where

$$\Psi_{m,n,l,k}(t, \nu) = \mathcal{T}_{(kT, n\Omega)} \mathcal{M}_{L(\ell \Omega \beta_2 T, T/\beta_1)} \, r \otimes \widehat{\phi}(t, \nu)$$
$$= e^{2\pi i L \left(\frac{mT(\nu - n\Omega)}{\beta_1} + \frac{\ell \Omega (t - kT)}{\beta_2} \right)} \, r(t - kT) \, \widehat{\phi}(\nu - n\Omega).$$

The frame bound is $T\Omega L^2 T\Omega/(\beta_1 \beta_2) = 1/(\beta_1 \beta_2)$. We set $\Phi_{m,-n,l,-k} = \mathcal{F}_s \Psi_{m,n,l,k}$. Clearly, as $\mathcal{F}_s$ is unitary, we have that $\{\Phi_{m,n,l,k}\}_{m,n,l,k \in \mathbb{Z}}$ forms a tight frame with frame bound $1/(\beta_1 \beta_2)$, in fact, a tight Gabor frame as

$$\Phi_{m,n,l,k}(x, \xi) = \mathcal{F}_s \Psi_{m,-n,l,-k}(x, \xi) = (\mathcal{F}\mathcal{T}_{-kT} \mathcal{M}_{\ell L \Omega/\beta_2} r)(\xi) \, (\mathcal{F}^{-1} \mathcal{T}_{-n\Omega} \mathcal{M}_{mTL/\beta_1} \widehat{\phi})(x)$$
$$= (\mathcal{M}_{kT} \mathcal{T}_{\ell L \Omega/\beta_2} \widehat{r})(\xi) \, (\mathcal{M}_{n\Omega} \mathcal{T}_{mTL/\beta_1} \phi)(x)$$
$$= e^{2\pi i (nm + kl)/\lambda} (\mathcal{T}_{\ell L \Omega/\beta_2} \mathcal{M}_{kT} \widehat{r})(\xi) \, (\mathcal{T}_{mTL/\beta_1} \mathcal{M}_{n\Omega} \phi)(x).$$

Note that (6.1) together with the fact that the symplectic Fourier transform is unitary implies that the coefficients in the Gabor frame expansion of $\boldsymbol{\sigma}$ satisfy

$$\langle \boldsymbol{\sigma}, \Phi_{m,-n_j,l,-k_j} \rangle = \langle \boldsymbol{\eta}, \Psi_{m,n,l,k} \rangle = 0 \text{ unless } (n, k) = (n_j, k_j) \text{ for some } j.$$

Hence we need to estimate $\sigma_{m,\ell}^{(j)} = \langle \boldsymbol{\sigma}, \Phi_{m,-n_j,l,-k_j} \rangle$ for $j = 0, 1, \ldots, L-1$. We obtain by (6.3)

$$\sigma_{m,\ell}^{(j)} = \langle \boldsymbol{\sigma}, \Phi_{m,-n_j,l,-k_j} \rangle = \langle \boldsymbol{\eta}, \Psi_{m,n_j,l,k_j} \rangle$$
$$= \iint \boldsymbol{\eta}(t, \nu) e^{-2\pi i L \left(\frac{Tm(\nu - n_j \Omega)}{\beta_1} + \frac{\ell \Omega (t - k_j T)}{\beta_2} \right)} \, r(t - k_j T) \, \widehat{\phi}(\nu - n_j \Omega) dt d\nu$$
$$= \iint r(t) \, \widehat{\phi}(\nu) \boldsymbol{\eta}(t + k_j T, \nu + n_j \Omega) e^{2\pi i \nu k_j T} \, e^{-2\pi i (L(\frac{m\nu T}{\beta_1} + \frac{\ell t \Omega}{\beta_2}) + \nu k_j T)} dt d\nu$$

$$= \iint LT \sum_{p=0}^{L-1} b_{jp} e^{-2\pi i \nu p T} r(t) \widehat{\phi}(\nu) (Z_{LT} \circ H) w(t+pT, \nu) \, e^{-2\pi i (L(\frac{m\nu T}{\beta_1} + \frac{\ell t \Omega}{\beta_2}) + \nu k_j T)} dt d\nu$$

$$= LT \sum_{p=0}^{L-1} b_{jp} \iint r(t) \widehat{\phi}(\nu) \, e^{-2\pi i \nu p T} (Z_{LT} \circ H) w(t+pT, \nu) \, e^{-2\pi i (L(\frac{m\nu T}{\beta_1} + \frac{\ell t \Omega}{\beta_2}) + \nu k_j T)} dt d\nu$$

$$= LT \sum_{p=0}^{L-1} b_{jp} \iint r(t) \widehat{\phi}(\nu) \, e^{-2\pi i \nu p T} \sum_{q \in \mathbb{Z}} H w(t+pT-qLT) \, e^{2\pi i \nu q LT} e^{-2\pi i (L(\frac{m\nu T}{\beta_1} + \frac{\ell t \Omega}{\beta_2}) + \nu k_j T)} dt d\nu$$

$$= LT \sum_{p=0}^{L-1} b_{jp} \sum_{q \in \mathbb{Z}} \Big(\int r(t) H w(t+pT-qLT) e^{-2\pi i L \frac{\ell t \Omega}{\beta_2}} dt \Big) \Big(\int \widehat{\phi}(\nu) \, e^{2\pi i \nu T (qL - p - k_j - mL/\beta_1)} \, d\nu \Big)$$

$$= LT \sum_{q \in \mathbb{Z}} b_{jq} \Big(\int r(t) H w(t+qT) e^{-2\pi i L \frac{\ell t \Omega}{\beta_2}} dt \Big) \Big(\int \widehat{\phi}(\nu) \, e^{2\pi i \nu T (-q - k_j - mL/\beta_1)} \, d\nu \Big)$$

$$= LT \sum_{q \in \mathbb{Z}} b_{jq} \phi(T(-q - k_j - mL/\beta_1)) \Big(\int H w(t) e^{-2\pi i L \frac{\ell \Omega (t - qT)}{\beta_2}} r(t - qT) dt \Big)$$

$$= LT \sum_{q \in \mathbb{Z}} b_{jq} \phi(T(-q - k_j - mL/\beta_1)) \, \langle Hw, \mathcal{T}_{qT} \mathcal{M}_{\ell L \Omega / \beta_2} r \rangle,$$

where $b_{jq} = b_{jq'}$ for $q = mL + q'$ with $q' = 0, 1, \ldots, L-1$. We can hence set

$$C_{q,l}(Hw) = \langle Hw, \mathcal{T}_{qT} \mathcal{M}_{\ell L \Omega / \beta_2} r \rangle.$$

To sum up,

$$\boldsymbol{\sigma}(x, \xi) = \frac{1}{\beta_1 \beta_2} \sum_{j=0}^{L-1} \sum_{m, \ell \in \mathbb{Z}} \langle \boldsymbol{\sigma}, \Phi_{m, -n_j, l, -k_j} \rangle \Phi_{m, -n_j, l, -k_j}(x, \xi)$$

(6.4)
$$= \frac{LT}{\beta_1 \beta_2} \sum_{j=0}^{L-1} e^{-2\pi i (x n_j \Omega + \xi k_j T)} \sum_{m, \ell \in \mathbb{Z}} \sigma_{m, \ell}^{(j)} \, \widehat{r}\Big(\xi - \frac{\ell L \Omega}{\beta_2}\Big) \, \phi\Big(x - \frac{mTL}{\beta_1}\Big),$$

where

$$\sigma_{m, \ell}^{(j)} = \sum_{q \in \mathbb{Z}} b_{jq} \, \phi(a(-q - k_j - mL/\beta_1)) \, C_{q,l}(Hw).$$

Applying the symplectic Fourier transform to (6.4) yields

$$\eta(t, \nu) = e^{-2\pi i \nu t} \boldsymbol{\eta}(t, \nu) = e^{-2\pi i \nu t} \frac{LT}{\beta_1 \beta_2} \sum_{j=0}^{L-1} \sum_{m, \ell \in \mathbb{Z}} \sigma_{m, \ell}^{(j)} \, \mathcal{F}_s \left(\mathcal{M}_{(-n_j \Omega, -k_j T)} \, \mathcal{T}_{(\frac{mTL}{\beta_1}, \frac{\ell L \Omega}{\beta_2})} \phi \otimes \widehat{r} \right)(t, \nu)$$

$$= e^{-2\pi i \nu t} \frac{LT}{\beta_1 \beta_2} \sum_{j=0}^{L-1} \sum_{m, \ell \in \mathbb{Z}} \sigma_{m, \ell}^{(j)} \, \mathcal{T}_{(k_j T, -n_j \Omega)} \, \mathcal{M}_{(\frac{\ell L \Omega}{\beta_2}, -\frac{mTL}{\beta_1})} \, r \otimes \widehat{\phi}(t, \nu)$$

$$= \frac{LT}{\beta_1 \beta_2} \sum_{j=0}^{L-1} \sum_{m, \ell \in \mathbb{Z}} \sigma_{m, \ell}^{(j)} \, \mathcal{T}_{(k_j T, -n_j \Omega)} \, \mathcal{M}_{(\frac{\ell L \Omega}{\beta_2}, -\frac{mTL}{\beta_1})} \left(r \otimes \widehat{\phi}(t, \nu) \, e^{-2\pi i (\nu + n_j \Omega)(t - k_j T)} \right),$$

$$= \frac{LT}{\beta_1 \beta_2} \sum_{j=0}^{L-1} \sum_{m, \ell \in \mathbb{Z}} \sigma_{m, \ell}^{(j)} e^{2\pi i n_j \Omega k_j T} \, \mathcal{T}_{(k_j T, -n_j \Omega)} \, \mathcal{M}_{(\frac{\ell L \Omega}{\beta_2} - n_j \Omega, \, k_j T - \frac{mTL}{\beta_1})} \left(r \otimes \widehat{\phi}(t, \nu) \, e^{-2\pi i \nu t} \right).$$

For $U(t, \nu) = r \otimes \widehat{\phi}(t, \nu) e^{-2\pi i \nu t}$, we have

$$\mathcal{F}_s U(x, \xi) = \iint r(t) \widehat{\phi}(\nu) e^{-2\pi i \nu t} e^{-2\pi i (\xi t - \nu x)} \, d\nu \, dt$$

$$= \int r(t)\phi(x-t)e^{-2\pi i \xi t}\,dt = \int r(t)\overline{\phi}(t-x)e^{-2\pi i \xi t}\,dt = V_\phi r(x,\xi),$$

where we used that $\widehat{\phi}$ real valued implies $\phi(y) = \overline{\phi}(-y)$. Now, we compute

$$\sigma(x,\xi) = \mathcal{F}_s \eta\,(x,\xi)$$

$$= \frac{LT}{\beta_1 \beta_2} \sum_{j=0}^{L-1} \sum_{m,\ell \in \mathbb{Z}} \sigma_{m,\ell}^{(j)} e^{2\pi i n_j \Omega k_j T}\, \mathcal{F}_s \left(\mathcal{T}_{(k_j T, -n_j \Omega)}\, \mathcal{M}_{(\frac{\ell L \Omega}{\beta_2} - n_j \Omega,\, k_j T - \frac{mTL}{\beta_1})}\, U\right)(x,\xi),$$

$$= \frac{LT}{\beta_1 \beta_2} \sum_{j=0}^{L-1} \sum_{m,\ell \in \mathbb{Z}} \sigma_{m,\ell}^{(j)} e^{2\pi i n_j \Omega k_j T}\, \mathcal{M}_{(-n_j \Omega, -k_j T)}\, \mathcal{T}_{(\frac{mTL}{\beta_1} - k_j T,\, \frac{\ell L \Omega}{\beta_2} - n_j \Omega)}\, V_\phi r\,(x,\xi),$$

$$(6.5) \quad = \frac{LT}{\beta_1 \beta_2} \sum_{j=0}^{L-1} e^{-2\pi i (x n_j \Omega + \xi k_j T)} e^{2\pi i n_j \Omega k_j T} \sum_{m,\ell \in \mathbb{Z}} \sigma_{m,\ell}^{(j)}\, V_\phi r(x - \frac{mTL}{\beta_1} + k_j T,\ \xi - \frac{\ell L \Omega + n_j \Omega}{\beta_2}).$$

The convergence in (6.4) and (6.5) is defined in the weak sense, but can be shown to converge absolutely and uniformly on compact subsets. $\qquad\square$

Next we prove Theorem 3.6, that is, the direct local correspondence between the discretization values and the operator action.

Proof of Theorem 3.6. We intend to apply Theorems 3.3 and 3.5. We assume that the set M as well as its enclosing rectangular grid are fixed, hence also the parameters T, Ω, and L. The dependence of the constants, auxiliary functions, etc., in the following derivations on these parameters will be suppressed for notational convenience; this should be seen as analogue to the one-dimensional scenario where the arising constants also depend on the shape and not just the size of the frequency support. Furthermore, set $Q = \max(LT, L\Omega)$.

We can bound using (2.3)

$$(6.6) \qquad |\sigma_{m,\ell}^{(j)}| = |\langle \boldsymbol{\sigma}, \Phi_{m,-n_j,\ell,-k_j} \rangle| \le \|\boldsymbol{\sigma}\|_\infty \|\Phi_{m,-n_j,\ell,-k_j}\|_1 \le B\|\sigma\|_\infty \|\hat{r} \otimes \phi\|_1 \le \tilde{B}\mu$$

For the second inequality, we used that the L_1-norm is invariant under translations and modulations.

Furthermore, note that $V_\phi r \in \mathcal{S}(\mathbb{R}^2)$, so there is a decreasing positive function $\rho \in \mathcal{S}([0,\infty))$ such that for $\widetilde{\rho}(x,\xi) = \rho(|x|)\rho(|\xi|)$ one has $|V_\phi r| \le \frac{1}{8\tilde{C}T}\widetilde{\rho}$ pointwise.

Now observe that, as ρ is decreasing,

$$\sum_{j=0}^{\infty} \alpha\rho(\alpha j) \le \rho(0) + \sum_{j=1}^{\infty} \int_{\alpha(j-1)}^{\alpha j} \rho(t)\,dt = \|\rho\|_1 + \|\rho\|_\infty.$$

We use this estimate to bound for arbitrary (x,ξ)

$$|\tilde{\sigma}(x,\xi)|$$

$$= \left| \frac{LT}{\beta_1 \beta_2} \sum_{j=0}^{L-1} e^{-2\pi i (x n_j \Omega + \xi k_j T)} e^{2\pi i n_j \Omega k_j T} \sum_{(mLT/\beta_1, \ell L\Omega/\beta_2) \in S} \sigma_{m,\ell}^{(j)} V_\phi r(x - (\frac{mL}{\beta_1} + k_j)T,\ \xi - (\frac{\ell L}{\beta_2} + n_j)\Omega) \right|$$

$$\le \frac{LT}{\beta_1 \beta_2} \sum_{(mLT/\beta_1, \ell L\Omega/\beta_2) \in S} |\sigma_{m,\ell}^{(j)}| \left| V_\phi r(x - (\frac{mL}{\beta_1} + k_j)T,\ \xi - (\frac{\ell L}{\beta_2} + n_j)\Omega) \right|$$

$$\le \frac{LT}{\beta_1 \beta_2} \sum_{j=0}^{L-1} \sum_{(mLT/\beta_1, \ell L\Omega/\beta_2) \in S} \frac{\tilde{C}\mu}{8\tilde{C}T}\, \rho\left(\left|x - (\frac{mL}{\beta_1} + k_j)T\right|\right) \rho\left(\left|\xi - (\frac{\ell L}{\beta_2} + n_j)\Omega\right|\right)$$

$$\le \frac{\mu}{8L\Omega T} \sum_{j=0}^{L-1} 4 \sum_{m,\ell=0}^{\infty} \frac{LT}{\beta_1} \rho\left(\frac{mL}{\beta_1}T\right) \frac{L\Omega}{\beta_2} \rho\left(\frac{\ell L}{\beta_2}\Omega\right) \le \frac{\mu}{2}\left(\|\rho\|_1 + \|\rho\|_\infty\right)^2$$

20

and hence
$$\|\sigma - \tilde{\sigma}\|_\infty \le \|\sigma\|_\infty + \|\tilde{\sigma}\|_\infty \le \mu + \frac{\mu}{2}\Big(\|\rho\|_1 + \|\rho\|_\infty\Big)^2 =: C_1\mu.$$

By the definition of $\mathcal{S}$, for every $\delta > 0$, there is a constant $C(\delta)$ such that for any fixed $0 \le j < L$,

(6.7) $$\delta \ge 8\|\rho\|_{L^1(\mathbb{R}^+)}\|\rho\|_{L^1[C(\delta)-2Q,\infty)} \ge 8\|\tilde{\rho}\|_{L_1(([-C(\delta)+Q,C(\delta)-Q]^2)^c)},$$

and hence, for $(x,\xi) \in S - B(C(\delta))$,

(6.8) $$\delta \ge \frac{L^2}{\beta_1\beta_2} \sum_{\substack{\ell,m\in\mathbb{Z} \\ \left(x-\frac{mL}{\beta_1}T,\xi-\frac{\ell L}{\beta_2}\Omega\right)\notin[-C(\delta),C(\delta)]^2}} \rho\Big(\Big|x - \Big(\frac{mL}{\beta_1} + k_j\Big)T\Big|\Big)\rho\Big(\Big|\xi - \Big(\frac{\ell L}{\beta_2} + n_j\Big)\Omega\Big|\Big).$$

To obtain (6.8) from (6.7), the boundary term in the discretization of the integral and the shifts by k_j and n_j, respectively, are each compensated by increasing the dimensions of the integration/summation domain by LT and $L\Omega$ in time and frequency, respectively, both of which are bounded by Q.

Note furthermore that, as $(x,\xi) \in S - B(C(\delta))$, a necessary condition for
$$\Big(x - \frac{mL}{\beta_1}T, \xi - \frac{\ell L}{\beta_2}\Omega\Big) \notin [-C(\delta), C(\delta)]^2$$

is that
$$(mLT/\beta_1, \ell L\Omega/\beta_2) \notin S.$$

Thus, using (6.6) and the triangle inequality, we can bound (6.8) from below obtaining

(6.9) $$\delta\mu \ge \Big|\frac{L^2T}{\beta_1\beta_2} \sum_{(mLT/\beta_1,\ell L\Omega/\beta_2)\notin S} \sigma^{(j)}_{m,\ell}V_\phi r\Big(x - \Big(\frac{mL}{\beta_1} + k_j\Big)T, \ \xi - \Big(\frac{\ell L}{\beta_2} + n_j\Big)\Omega\Big)\Big|.$$

Hence forming a weighted average (with complex weighting factors of modulus one) of Equation (6.9) over the L choices of j, we obtain

$$\delta\mu \ge \Big|\frac{LT}{\beta_1\beta_2}\sum_{j=0}^{L-1} e^{-2\pi i(xn_j\Omega + \xi k_j T)}e^{2\pi i n_j\Omega k_j T}\sum_{(\frac{mLT}{\beta_1},\frac{\ell L\Omega}{\beta_2})\notin S}\sigma^{(j)}_{m,\ell}V_\phi r\Big(x - \Big(\frac{mL}{\beta_1}+k_j\Big)T, \ \xi - \Big(\frac{\ell L}{\beta_2}+n_j\Big)\Omega\Big)\Big|$$

$$= |\sigma(x,\xi) - \tilde{\sigma}(x,\xi)|.$$

This yields $\|\sigma - \tilde{\sigma}\|_{L_\infty(S-B(C(\delta)))} \le \delta\mu$. Hence by Theorem 3.3, we conclude that
$$\|Hf - \tilde{H}f\|_2 \le C\frac{\delta}{C_1}\mu$$

for all functions f which are $\frac{\delta}{C_1}$-time-frequency-localized to $S - B(C(\delta)) - B(d(\epsilon))$. The result follows by choosing $\delta = \min\Big(\frac{C_1}{C}\epsilon, C_1\epsilon\Big)$ and $D(\epsilon) = C(\delta) + d(\epsilon)$. $\qquad\square$

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** (2011), no. 6, 2548–2561.
2. P.A. Bello, *Measurement of random time-variant linear channels*, IEEE Trans. Comm. **15** (1969), 469–475.
3. J. J. Benedetto, A. M. Powell, and Ö. Yılmaz, *Sigma-Delta quantization and finite frames*, IEEE Trans. Inform. Theory **52** (2006), 1990–2005.
4. J.J. Benedetto and O. Oktay, *Pointwise comparison of PCM and $\Sigma\Delta$ quantization*, Constr. Approx. **32** (2010), no. 1, 131–158.
5. Á. Bényi, K. Gröchenig, K. A. Okoudjou, and L. G. Rogers, *Unimodular Fourier multipliers for modulation spaces*, J. Funct. Anal. **246** (2007), no. 2, 366–384.
6. I. Daubechies and R. DeVore, *Reconstructing a bandlimited function from very coarsely quantized data: A family of stable sigma-delta modulators of arbitrary order*, Ann. Math. **158** (2003), 679–710.
7. P. Deift, C. S. Güntürk, and F. Krahmer, *An optimal family of exponentially accurate one-bit sigma-delta quantization schemes*, Comm. Pure Appl. Math. **64** (2011), no. 7, 883–919.
8. H.G. Feichtinger, *Atomic characterizations of modulation spaces through Gabor-type representations*, Rocky Mountain J. Math. **19** (1989), 113–126.
9. H.G. Feichtinger and K. Gröchenig, *Gabor wavelets and the Heisenberg group: Gabor expansions and short time Fourier transform from the group theoretical point of view*, Wavelets, Wavelet Anal. Appl., vol. 2, Academic Press, Boston, MA, 1992, pp. 359–397.
10. G.B. Folland, *Harmonic analysis in phase space*, Annals of mathematics studies, vol. 122, Princeton University Press, 1989.
11. K. Gröchenig, *Foundations of Time-Frequency Analysis*, Birkhäuser, Boston, 2001.
12. K. Gröchenig and C. Heil, *Modulation spaces and pseudodifferential operators*, Integral Equations Operator Theory **34** (1999), no. 4, 439–457. MR 1702232 (2001a:47051)
13. C. S. Güntürk, *One-bit Sigma-Delta quantization with exponential accuracy*, Comm. Pure Appl. Math. **56** (2003), 1608–1630.
14. R. Heckel and H. Boelcskei, *Identification of sparse linear operators*, preprint.
15. T. Kailath, *Measurements on time-variant communication channels.*, IEEE Trans. Inform. Theory **8** (1962), no. 5, 229– 236.
16. W. Kozek and G.E. Pfander, *Identification of operators with bandlimited symbols*, SIAM J. Math. Anal. **37** (2005), no. 3, 867–888.
17. F. Krahmer and G.E. Pfander, *Sampling and quantization of approximately bandlimited operators*, in preparation, 2013.
18. F. Krahmer, G.E. Pfander, and P. Rashkov, *Uncertainty principles for timefrequency representations on finite Abelian groups*, Appl. Comput. Harmon. Anal. **25** (2008), 209–225.
19. M. Lammers, A.M. Powell, and Ö Yılmaz, *Alternative dual frames for digital-to-analog conversion in sigma-delta quantization*, Adv. Comput. Math. **32** (2010), no. 1, 73–102.
20. J. Lawrence, G.E. Pfander, and D. Walnut, *Linear independence of Gabor systems in finite dimensional vector spaces*, J. Fourier Anal. Appl. **11** (2005), no. 6, 715–726.
21. K.A. Okoudjou, *A Beurling-Helson type theorem for modulation spaces*, J. Funct. Spaces Appl. **7** (2009), no. 1, 33–41.
22. A.V. Oppenheim, R.W. Schafer, and J.R. Buck, *Discrete-time signal processing*, 2nd ed., Prentice-Hall signal processing, Prentice-Hall, Upper Saddle River, NJ, 1999.
23. D. P. Petersen and D. Middleton, *Sampling and reconstruction of wave-number-limited functions in N-dimensional Euclidean spaces*, Information and Control **5** (1962), 279–323. MR 0151331 (27 #1317)
24. G.E. Pfander, *Sampling of operators*, to appear in J. Four. Anal. Appl.
25. ———, *Measurement of time–varying Multiple–Input Multiple–Output channels*, Appl. Comp. Harm. Anal. **24** (2008), 393–401.
26. G.E. Pfander and D. Walnut, *Sampling and reconstruction of operators*, preprint.
27. G.E. Pfander and D.F. Walnut, *Measurement of time-variant linear channels*, IEEE Trans. Inform. Theory **52** (2006), no. 11, 4808–4820.
28. A. Ron and Z. Shen, *Frames and stable bases for shift–invariant subspaces $L_2(\mathbf{R}^d)$*, Canadian Journal of Mathematics **47** (1995), no. 5, 1051–1094.
29. T. Strohmer, *Pseudodifferential operators and Banach algebras in mobile communications*, Appl. Comput. Harmon. Anal. **20** (2006), no. 2, 237–249.
30. J. Toft, *Continuity properties for modulation spaces, with applications to pseudo-differential calculus – I*, Journal of Functional Analysis **207** (2004), no. 2, 399–429.
31. ———, *Continuity properties for modulation spaces, with applications to pseudo-differential calculus – II*, Annals of Global Analysis and Geometry **26** (2004), no. 1, 73–106.
32. ———, *Continuity and schatten properties for pseudo-differential operators on modulation spaces*, Modern trends in pseudo-differential operators, Springer, 2007, pp. 173–206.