Operator sampling and MIMO channel identification

G. E. Pfander and D. F. Walnut *

ABSTRACT

The classical sampling theorem, attributed to Whittaker, Shannon, Nyquist, and Kotelnikov, states that a bandlimited function can be recovered from its samples, as long as we use a sufficiently dense sampling grid. Here, we review the recent development of an operator sampling theory which allows for a "widening" of the classical sampling theorem. In this realm, bandlimited functions are replaced by "bandlimited operators", that is, by pseudodifferential operators which have bandlimited Kohn–Nirenberg symbols.

Similar to the Nyquist sampling density condition, we discuss sufficient and necessary conditions on the bandlimitation of pseudodifferential operators to ensure that they can be recovered by their action on a single distribution. In fact, we show that an operator with Kohn–Nirenberg symbol bandlimited to a Jordan domain of measure less than one can be recovered through its action on a distribution defined on a appropriately chosen sampling grid. Further, an operator with bandlimitation to a Jordan domain of measure larger than one cannot be recovered through its action on any tempered distribution whatsoever, pointing towards a fundamental difference to the classical sampling theorem where a large bandwidth could always be compensated through a sufficiently fine sampling grid. The dichotomy depending on the size of the bandlimitation is related to Heisenberg's uncertainty principle.

Further, we discuss an application of this theory to the channel measurement problem for Multiple–Input Multiple–Output (MIMO) channels.

Keywords: B andlimited Kohn–Nirenberg symbols, channel measurement, Multiple–Input Multiple–Output channels, Gabor and time–frequency analysis, operator identification, spreading functions, underspread operators, matrix identification.

1. INTRODUCTION

To infer reliable information on, or, better, to identify only partially known objects from very limited data is a key task in the sciences. In communications engineering, for example, two basic theorems addressing this objective are folklore:

(a) The classical sampling theorem states that a bandlimited function can be recovered from its samples on a sufficiently dense sampling grid.

(b) Operators representing time–invariant channels as encountered in wired communications can be identified through a single input/output pair, namely, the channel is completely characterized by the channels response to Dirac's delta impulse.

In the 1960s, Thomas Kailath posed the question whether (b) could be extended to slowly time–varying channels.[1] He conjectured a sufficient and necessary condition for the measurability of such operators based on the product of maximum time–spread (corresponding to the longest path the signal travels) and maximum frequency–spread (Doppler effect caused by the movement of sender, receiver, and reflecting objects): if *a priori* knowledge indicates that this product is less than or equal to one, then the operator is identifiable, else, it is not. Using recent techniques from Gabor analysis, his conjecture was was recently proven[2] and this and follow up results lead to the formulation of an operator sampling theory which we shall review in this paper.

In fact, Kailath's result were extended to cover pseudodifferential operators of any degree and therefore also to large classes of time–invariant operators and multiplication operators[3].[?] As we shall see below, by

G. E. Pfander is with the Jacobs University Bremen, Germany, and D. F. Walnut is with George Mason University, USA.
†School of Engineering and Science, Jacobs University Bremen, 28759 Bremen, Germany.
‡Department of Mathematical Sciences George Mason University, Fairfax, VA 22030

associating a bandlimited multiplication operator to its bandlimited symbol, our results entail the classical sampling theorem described in (a). Furthermore, straightforward calculations lead to a reconstruction formula for bandlimited operators, that is, of their so–called time–varying impulse responses, from the corresponding channel outputs. This formula reduces to the standard reconstruction formula for bandlimited functions in the case of multiplication operator symbols.

Kailath's product condition reflects bandlimitations on the Kohn–Nirenberg symbols of pseudodifferential operators to a rectangular domain. This result was extended to higher dimensions and, by means of the representation theory of the Weyl–Heisenberg group and the metaplectic representation of the symplectic group, to pseudodifferential operators with symbols bandlimited to a fundamental domain of a symplectic lattice.

In the one–dimensional case, we give a refinement of the *product* condition of Kailath as it was foreseen by Bello,[4] that is, we show that also operators with Kohn–Nirenberg symbol bandlimited to a Jordan domain of Lebesgue measure less than one and arbitrary geometry can be recovered through its action on a distribution which is supported on an appropriately chosen sampling grid. As expected, an operator which can only be assumed to have a Kohn–Nirenberg symbol bandlimited to a Jordan domain of measure larger than one cannot be recovered by its action on any tempered distribution[3] (see figure on page **??**). The dichotomy depending on the size of the bandlimitation is related to Heisenberg's uncertainty principle as described in.[2]

The paper is organized as follows. In Section **??** we give vocabulary from time–frequency analysis that is at the core of our sampling theory for operators. Section ...

2. SAMPLING PRINCIPLES

Here, we briefly recall the rudimentary concept of the sampling concepts for functions and operators.

2.1 Sampling of functions

Sampling theory was developed in the 19th and early 20th century to analyze the capacity of bandlimited telephony channels.[5] Nyquist and Küpfmüller, for instance, realized that the capacity of such wired and therefore translation invariant channel is proportional to the product of bandwidth and signal duration.[?] In modern terminology, a signal with integer bandwidth Ω [Hz] has Ω degrees of freedom in a time unit [s]. Consequently, to fully determine a signal f with bandwidth Ω, knowledge of $f(\frac{n}{\Omega})$ for all integers n suffices.

Given a normed space $D(\mathbb{R})$ of real or complex valued on the real line, for example, the space of bandlimited and square integrable functions, a fundamental task in sampling theory is to find necessary and/or sufficient conditions on a set $\{x_j\}_{j\in\mathbb{Z}} \subset \mathbb{R}$ so that any $f \in D(\mathbb{R})$ can be recovered from its values on $\{x_j\}$, namely, from the sequence $\{f(x_j)\}_{j\in\mathbb{Z}}$. Furthermore, we generally require that the recovery process is stable with respect to perturbation of the measurements. That is, for a normed space $d(\mathbb{Z})$ of sequences, we require that there are positive constants A, B with

$$A\|f\|_D \leq \|\{f(x_j)\}\|_d \leq B\|f\|_D \quad \text{for all } f \in D(\mathbb{R}). \tag{1}$$

For brevity, we shall denote double norm inequalities such as given in (1) by the symbol $\asymp$, that is, (1) will be simply written as

$$\|f\|_D \asymp \|\{f(x_j)\}\|_d \quad \text{for all } f \in D(\mathbb{R}). \tag{2}$$

If (1) is satisfied, then we call $\{x_j\}$ a *set of sampling* for $D(\mathbb{R})$.

For background on sampling theory, see the overview articles[6,7] and references therein.

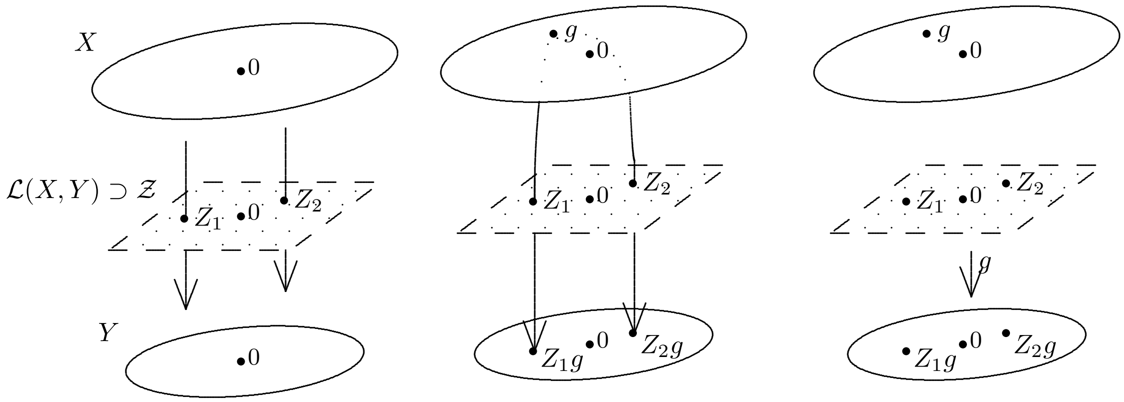

Figure 1. Illustration of the operator identification problem. One seeks a function g in the domain of the operator class $\mathcal{Z}$ which induces a map from $\mathcal{Z}$ into the range space which is bounded and bounded below.

2.2 Sampling of operators and operator identification

The goal of operator identification is to select, for given normed linear spaces $D(\mathbb{R}^d)$ and $Y(\mathbb{R}^d)$ of functions on $\mathbb{R}^d$, and a normed linear space of bounded linear operators $\mathcal{H} \subset \mathcal{L}(D(\mathbb{R}^d), Y(\mathbb{R}^d))$, an element $g \in D(\mathbb{R}^d)$ which induces a bounded and injective, or better, a map $\Phi_g : \mathcal{H} \longrightarrow Y(\mathbb{R}^d),\ H \mapsto Hg$ which is bounded and bounded below (see Figure 2.2).

Consequently, we call $\mathcal{H}$ *identifiable by* $g \in D(\mathbb{R}^d)$, if there exist $A, B > 0$ with

$$A\,\|H\|_{\mathcal{Z}} \leq \|Hg\|_Y \leq B\,\|H\|_{\mathcal{Z}} \quad \text{for all } H \in \mathcal{Z}, \tag{3}$$

that is

$$\|H\|_{\mathcal{Z}} \asymp \|Hg\|_Y. \tag{4}$$

If $D(\mathbb{R}^d)$ is a space which contains Dirac impulses $\delta_x : f \mapsto f(x)$ for $x \in \mathbb{R}^d$ and if we can choose an identifier $g \in D(\mathbb{R}^d)$ of the form $g = \sum_j c_j \delta_{x_j}$, $x_j \in \mathbb{R}^d$ and $c_j \in \mathbb{C}$ for $j \in Z^d$, then we call $\{x_j\}$ a set of sampling for $\mathcal{Z}$ and g a sampling function for the operator class $\mathcal{Z}$.
[2]

3. HILBERT SPACE THEORY

For simplicity and illustration, we shall first develop put side to side the sampling theory of square integrable functions and the corresponding sampling theory of Hilbert–Schmidt operators.

[2] One could include a paragraph here describing the fact that in function sampling, the data is discrete, the object to be identified defined on the continuum, while in operator sampling, the data might well be an object defined on the continuum, while the operator acts on distributions. As operators can be given by the corresponding matrix (kernel), one could picture the transition from samples of functions to functions by points going to the line, while the transition from Hg to H as going from the the line to the plane.

3.1 Preliminaries on square integrable functions

$L^2(\mathbb{R}^d)$ denotes the space of complex valued and Lebesgue measurable functions on Euclidean space $\mathbb{R}^d$ which satisfy[8]

$$\|f\|_{L^2} = \left(\int |f(x)|^p \, dx\right)^{\frac{1}{2}} < \infty.$$

The space $L^2(\mathbb{R}^d)$ is a Hilbert space with inner product $\langle f, g\rangle = \int f(x)\overline{g}(x)\, dx$.

The *Fourier transformation* is the unitary operator

$$\mathcal{F} : L^2(\mathbb{R}^d) \longrightarrow L^2(\mathbb{R}^d), \quad f \mapsto \widehat{f} = \mathcal{F}f,$$

which is, again, densely defined by

$$\widehat{f}(\gamma) = \int_{\mathbb{R}^d} f(x)\, e^{-2\pi i \gamma\cdot x}\, dx\,, \quad \gamma \in \mathbb{R}^d.$$

For functions in $L^2(\mathbb{R}^{2d})$, we shall at times employ the *symplectic Fourier transformation* $\mathcal{F}_s : L^2(\mathbb{R}^{2d}) \longrightarrow L^2(\mathbb{R}^{2d})$ which is, again, densely defined by

$$\mathcal{F}_s(t,\nu) = \iint_{\mathbb{R}^{2d}} f(x,\xi)\, e^{-2\pi i(\nu\cdot x - \xi\cdot t)}\, dx\, d\xi, \quad t,\nu \in \mathbb{R}^d.$$

3.2 Sampling of square integrable functions on $\mathbb{R}$

We shall start our discussion with the most classical sampling result. It addresses bandlimited and square integrable functions on the real line. In this realm, the Paley–Wiener space with bandwidth $\Omega > 0$ is given by

$$PW^2(S_\Omega) = \left\{ f \in L^2(\mathbb{R}) : \ \operatorname{supp} \widehat{f} \subseteq S_\Omega = [-\tfrac{\Omega}{2}, \tfrac{\Omega}{2}] \right\} . \tag{5}$$

Theorem 3.1. *For $\Omega, T > 0$ with $T\Omega \le 1$, we have*

$$\|f\|_{L^2} = T\|\{f(nT)\}\|_{l^2} \quad \textit{and} \quad f(x) = T\sum_{n\in\mathbb{Z}} f(nT)\,\frac{\sin(2\pi(x-nT)}{\pi(x-nT)} \quad \textit{for all } f \in PW^2(S_\Omega). \tag{6}$$

Extracting the key properties of the so-called sinc function $\operatorname{sinc}(x) = \frac{\sin 2\pi(x-nT)}{\pi(x-nT)}$ in (9), we obtain

Theorem 3.2. *For $\Omega, T > 0$ with $T\Omega < 1$, choose $s \in PW^2\big(S_{\frac{2}{T}-\Omega}\big)$ with $\widehat{s} = 1$ on $[-\frac{\Omega}{2}, \frac{\Omega}{2}]$*[3]*. Then*

$$\|\,\{f(nT)\}\,\|_{l^2} \asymp \|f\|_{L^2} \quad \textit{and} \quad f(x) = \sum_{n\in\mathbb{Z}} f(nT)s(x-nT) \quad \textit{for all } f \in PW^2(S_\Omega). \tag{7}$$

For results on irregular sampling sets see,[6,9] The following, for example, is well known.[10]

Theorem 3.3. *If $\{x_j\}$ satisfies $\liminf_{r\to\infty} \inf_{y\in\mathbb{R}} \dfrac{\text{cardinality of}(\{x_j\}\cap(y+[0,r]))}{r} > \dfrac{1}{\Omega}$, then $\{x_j\}$ is a set of sampling for $PW^2(S_\Omega)$.*

The results given above have straightforward generalizations to functions defined on higher dimensional Euclidean space.[6,7,9] For a bandwidth vector $\Omega = (\Omega_1, \ldots, \Omega_d)$, we can set

$$PW^2(S_\Omega) = \left\{ f \in L^2(\mathbb{R}^d) : \ \operatorname{supp} \widehat{f} \subseteq [-\tfrac{\Omega_1}{2}, \tfrac{\Omega_1}{2}] \times \ldots \times [\tfrac{\Omega_d}{2} \tfrac{\Omega_d}{2}] \right\} \tag{8}$$

[3]Note that $\Omega, T > 0$ with $T\Omega < 1$ implies $0 < \Omega < \frac{1}{T} < \frac{2}{T} - Omega$.

Then

THEOREM 3.4 (THE CLASSICAL SAMPLING THEOREM). *The set $T_1\mathbb{Z} \times T_2\mathbb{Z} \times \ldots \times T_d\mathbb{Z}$ is a set of sampling for $PW^2(\Omega)$ if and only if $T_k\Omega_k \leq 1$ for all $k = 1, \ldots, d$. Moreover, we have*

$$f(x) = \sum_{n_1\in\mathbb{Z}} \cdots \sum_{n_d\in\mathbb{Z}} f(T_1, \ldots, T_d) \frac{\sin(2\pi T_1(x_1 - n_1)}{\pi(x_1 - n_1)} \cdot \ldots \cdot \frac{\sin(2\pi T_d(x_d - n_d)}{\pi(x_d - n_d)} \tag{9}$$

3.3 Preliminaries on Hilbert–Schmidt operators

A *Hilbert–Schmidt operator* H is a bounded linear operator on $L^2(\mathbb{R}^d)$ which can be represented as an integral operator

$$\begin{aligned} Hf(x) &= \int \kappa_H(x,t) f(t)\, dt \\ &= \int \kappa_H(x, x-t) f(x-t)\, dt \quad (a.e.), \end{aligned} \tag{10}$$

with kernel $\kappa_H \in L^2(\mathbb{R}^2)$.[11,12] The linear space of Hilbert–Schmidt operators $HS(L^2(\mathbb{R}^d))$ is endowed with the Hilbert space structure of $L^2(\mathbb{R}^2)$ by setting

$$\langle H_1, H_2\rangle_{\mathrm{HS}} = \langle \kappa_{H_1}, \kappa_{H_2}\rangle_{L^2}.$$

The Kohn–Nirenberg symbol σ_H of a Hilbert–Schmidt operator H is given by[13,14]

$$\sigma_H(x,\xi) = \int \kappa_H(x, x-y)\, e^{-2\pi i y\xi}\, dy \quad (a.e.). \tag{11}$$

It leads to the operator representation

$$Hf(x) = \int \sigma(x,\xi) \widehat{f}(\xi) e^{2\pi i x\xi}\, d\xi. \tag{12}$$

The Kohn–Nirenberg symbol was originally defined to describe so–called pseudodifferential operators. In fact, the simple observation that the n-th derivative operator $D^{(n)} : f \mapsto f^{(n)}$ can be expressed by

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

leads to the representation

$$Df(x) = \int \left(\sum_{n=0}^{N} a_n(x)(2\pi i \xi)^n \right) \widehat{f}(\xi)\, e^{2\pi i x\xi}\, d\xi \tag{13}$$

for the differential operator $D : f \mapsto \sum_{n=0}^{N} a_n(x) f^{(n)}(x)$. Clearly, (13) is a special case of (12). Note that in this section, σ_H is a square integrable function while the symbol $\sigma_D(x,\xi) = \sum_{n=0}^{N} a_n(x)(2\pi i\xi)^n$ has polynomial growth in ξ.

Convolution operators $f \mapsto f * h = \int h(y) f(x-y)\, dy$ represent time–invariant channels in communications engineering. In communications, one therefore commonly chooses the so–called time–varying impulse response $h_H(x,y) = \kappa_H(x, x-y)$ to represent the general case of time–varying operators H by

$$Hf(x) = \int h_H(x,y) f(x-y)\, dy. \tag{14}$$

Additionally, in time–frequency analysis and in communications engineering, the spreading function η_H of a Hilbert–Schmidt operator H, given by

$$\eta_H(t,\nu) = \int \kappa_H(x, x-t) e^{-2\pi i \nu x}\, dx \quad (a.e.) \tag{15}$$

is commonly used to represent operators. It leads to a representation of H by means of

$$Hf(x) = \iint \eta_H(t,\nu)\, T_t M_\nu f(x)\, dx\, dt\, d\nu\,, \tag{16}$$

where the unitary time and frequency shift operators T_t and M_ν, $t,\nu \in \mathbb{R}$, on $L^2(\mathbb{R})$ are given by $\big(T_t f\big)(x) = f(t-x)$ and $\big(M_\nu f\big)(x) = e^{2\pi i x\cdot\nu} f(t)$. Note that (15) together with (11) implies that $\sigma_H = \mathcal{F}_s \eta_H$. [4]

Equation (16) illustrates that support restrictions on η_H reflect limitations on the maximal time and frequency shifts which the input signals undergo: Hf is a continuous superposition of time–frequency shifted versions of f with weighting function η_H.

A comparison of (14) to a time–invariant convolution operator together with (15) shows that the condition $\eta_H(t,\cdot) \subseteq [-\frac{b}{2}, \frac{b}{2}]$ for all $t \in \mathbb{R}$, excludes high frequencies and therefore rapid change of $\kappa(x, x-t)$ as a function of x. This illuminates the role of support constraints on spreading functions in the analysis of *slowly time–varying* communications channels.

In fact, note that if an operator H satisfies $\operatorname{supp} \eta_H(\cdot,\nu) \subseteq [0,a]$ for all $\nu \in \widehat{\mathbb{R}}$, then $\kappa_H(x, x-t)$ vanishes for $x \in \mathbb{R}$ and $t \notin [0,a]$, and for f with $\operatorname{supp} f \subseteq [0,T]$ we have $\operatorname{supp} Hf \subseteq [0, T+a]$. Similarly, if $\operatorname{supp} \eta_H(t,\cdot) \subseteq [-\frac{b}{2}, \frac{b}{2}]$ for all $t \in \mathbb{R}$, then for f with $\operatorname{supp} \widehat{f} \subseteq [-\Omega, \Omega]$ we have $\operatorname{supp} \widehat{Hf} \subseteq [-(\Omega+\frac{b}{2}), \Omega+\frac{b}{2}]$. Hence, the condition

$$\operatorname{supp} \eta_H \subseteq Q_{a,b} = [0,a] \times [-\tfrac{b}{2}, \tfrac{b}{2}] \tag{17}$$

for some $a, b > 0$, reflects a limitation on the maximal time delay a and the maximal frequency spread $\frac{b}{2}$ produced by H. An operator which satisfies (17) for $a, b > 0$ is called *underspread* if $ab \leq 1$ and *overspread* if $ab > 1$.

Note that

$$\|\eta_H\|_{L^2} = \|\sigma_H\|_{L^2} = \|h_H\|_{L^2} = \|\kappa_H\|_{L^2} = \|H\|_{HS},$$

a fact which will be used to obtain norm inequalities of the form (3) for Hilbert–Schmidt operators.

The previous paragraphs emphasize the usefulness of η_H in the time–frequency analysis of operators. Additional remarks on the use of Hilbert–Schmidt operators as model of physical time–varying linear systems, as they appear in radar and in mobile communications can be found in.[2]

3.4 Sampling of Hilbert–Schmidt operators

The classical sampling theorem assumes a band–limitation on the functions that are considered. The operator sampling theory discussed here considers in turn operators whose Kohn–Nirenberg symbol is band–limited to some region in the time–frequency plane, that is, whose spreading function is supported on a region in the (t,ν) plane. To avoid pathological counterexamples we shall only consider support conditions to so–called Jordan domains which are described in detail in the appendix and in.[15] [5]

For any Jordan domain $S \subseteq \mathbb{R}^2$, we refer to

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

as *operator Paley–Wiener space* with bandlimitation to S.

[4] time and frequency shifts not defined yet

[5] I would leave it in the appendix as engineers do not know the concept and the appendix might convince them that this is a technical assumption only. Maybe we should try to formulate this for open sets, but engineers would not know what that is either.

Note that operators in $OPW^2(S)$ with S compact can be extended to act on distribution spaces containing Dirac's δ as well as the so–called Shah distribution $⊥⊥⊥_T = \sum_n \delta_{nT}$, a fact that even though discussed in detail only in Section **??** will be used already here.

Theorem 3.2 has now the following analogon in the theory of operator sampling.[?, ?, 2]

THEOREM 3.5. *For $\Omega, T, T' > 0$ and $0 < \Omega T' < \Omega T < 1$, choose $s \in PW^2\big(S_{\frac{2}{T}-\Omega}\big)$ with $\widehat{s} = 1$ on $[-\frac{\Omega}{2}, \frac{\Omega}{2}]$ and $r \in \mathcal{S}$ with $\operatorname{supp} r \subset [-T + \frac{T'}{2}, T - \frac{T'}{2}]$ and $r = 1$ on $[-\frac{T'}{2}, \frac{T'}{2}]$. Then*

$$\|H⊥⊥⊥\|_{L^2} \asymp \|H\|_{HS} \quad \text{and} \quad h_H(t,x) = r(t) \sum_{k\in\mathbb{Z}} (H⊥⊥⊥_T)(t+kT)\, s(x-kT) \quad \text{for all } H \in OPW^2([-\tfrac{\Omega}{2}, \tfrac{\Omega}{2}]\times[-\tfrac{T'}{2}, \tfrac{T'}{2}])$$

The result discusses a characterization for rectangular bandlimitations on the Kohn–Nirenberg symbol. as alluded to by Kailath. Note that in operator sampling, the necessary sampling Frequency is determined by the width Ω of the set $[-\frac{\Omega}{2}, \frac{\Omega}{2}]\times[-\frac{T'}{2}, \frac{T'}{2}]$. Further, the height of the set is restricted by Ω as well, hence the area of the set cannot be compensate by sampling at a higher rate. In fact, not only sampling is impossible if the area is greater than one, but identification in general is impossible.[?, 2]

THEOREM 3.6. *If $T\Omega > 1$ then $OPW^2([0,T] \times [-\frac{\Omega}{2}, \frac{\Omega}{2}])$ is not identifiable.*

As mentioned in the introduction, Bello though believed that restriction to rectangular domains $[-\frac{\Omega}{2}, \frac{\Omega}{2}]\times[-\frac{T'}{2}, \frac{T'}{2}]$ is not needed and that the same phenomenon is present if we use domains of arbitrary shape.[?, 3, 4]

THEOREM 3.7. *For any Jordan domain $S \subset \mathbb{R} \times \widehat{\mathbb{R}}$, we have*

- *If $\operatorname{vol}(S) < 1$ then there exists a sampling set for $OPW^2(S)$.*
- *If $\operatorname{vol}(S) > 1$ then $OPW^2(S)$ is not identifiable.*

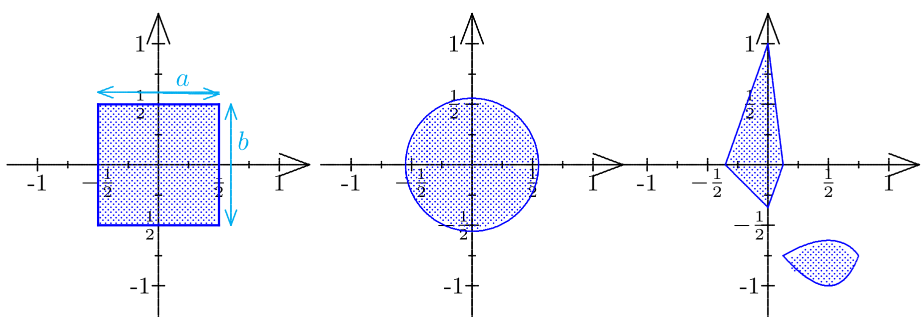

Figure 2. Spreading support regions of area less than or equal to one which characterize identifiable operator classes.

Theorem 3.5 and Theorem 3.6 generalize to higher dimensions using simple tensor product arguments. Moreover, symplectic geometry can be used to obtain[2]

THEOREM 3.8. *Let A be a symplectic matrix and $S = \alpha A[-\frac{1}{2}, \frac{1}{2}] + (t_0, \nu_0) \subset \mathbb{R}^{2d}$ with $\alpha > 0$ and $(t_0, \nu_0) \in \mathbb{R}^{2d}$, we have*

- *If $\alpha < 1$, then there exists a sampling set for $OPW^2(S)$.*
- *If $\alpha > 1$ then $OPW^2(S)$ is not identifiable.*

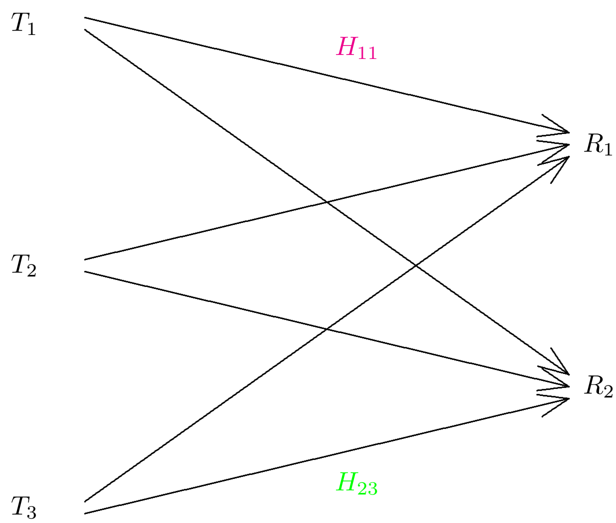

Figure 3. Multiple–Input Multiple–Output (MIMO) channels with $N = 3$ transmitters, $M = 2$ receivers

3.5 Multiple–Input Multiple–Output

Multiple transmit and receive antenna methods have been developed to obtain high capacity wireless channels (see[16–19] and references within). Methods which achieve high capacities often rely on the precise knowledge of the channel at the receiver and/or the transmitter (see,[16] pp 298).

In such MIMO channel setups, N signals are transmitted by N antennas simultaneously. On the receiver side, M antennas record channel output signals that represent the superposition of the N input signals, each individually distorted depending on the path the signal has travelled from its transmitting antenna to the receiving antenna. Consequently, a linear MIMO channel operator can be modelled by a matrix of $N{\cdot}M$ SISO channel operators. It maps a vector of N transmission signals to M channel output signals.

We denote by $HS(L^2(\mathbb{R}))^{M\times N}$ the space of N-input, M-output MIMO channels whose $N{\cdot}M$ subchannels are Hilbert–Schmidt operators on $L^2(\mathbb{R})$.[20] The operator space $HS(L^2(\mathbb{R}))^{M\times N}$ is equipped with norm[6]

$$\|\boldsymbol{H}\|_{HS} = \sqrt{\sum_{m=1}^{M}\sum_{n=1}^{N}\|H_{mn}\|_{HS}^2}, \quad \boldsymbol{H} = \begin{pmatrix} H_{11} & \cdots & H_{1N} \\ \vdots & & \vdots \\ H_{M1} & \cdots & H_{MN} \end{pmatrix} \in HS(L^2(\mathbb{R}))^{M\times N}.$$

Further, the spreading function $\boldsymbol{\eta_H} = \boldsymbol{\eta}(\boldsymbol{H})$ and the *spreading support* of $\boldsymbol{H} = \begin{pmatrix} H_{11} & \cdots & H_{1N} \\ \vdots & & \vdots \\ H_{M1} & \cdots & H_{MN} \end{pmatrix} \in HS(L^2(\mathbb{R}))^{M\times N}$ are defined componentwise, that is, we have

$$\boldsymbol{\eta}(\boldsymbol{H}) = \begin{pmatrix} \eta(H_{11}) & \cdots & \eta(H_{1N}) \\ \vdots & & \vdots \\ \eta(H_{M1}) & \cdots & \eta(H_{MN}) \end{pmatrix}, \quad \operatorname{supp} \boldsymbol{\eta}(\boldsymbol{H}) = \begin{pmatrix} \operatorname{supp} \eta(H_{11}) & \cdots & \operatorname{supp} \eta(H_{1N}) \\ \vdots & & \vdots \\ \operatorname{supp} \eta(H_{M1}) & \cdots & \operatorname{supp} \eta(H_{MN}) \end{pmatrix}.$$

Our identifiability result for MIMO channels considers operator classes of the form

$$\boldsymbol{O}PW^2(\boldsymbol{S}) = \Big\{\boldsymbol{H} \in HS(L^2(\mathbb{R}))^{M\times N} : \operatorname{supp} \boldsymbol{\eta}(\boldsymbol{H}) \subseteq \boldsymbol{S}\Big\}, \quad \boldsymbol{S} \subseteq (\mathbb{R}\times\widehat{\mathbb{R}})^{M\times N}.$$

To avoid pathological counterexamples of our main result Theorem 3.9, we shall only consider $\mathcal{H}_{\boldsymbol{S}}$ where $\boldsymbol{S}$ is the cartesian products of so called Jordan domains.

[6] It is easy to see that $HS(L^2(\mathbb{R}))^{N\times N} = HS(L^2(\mathbb{R})^N)$.

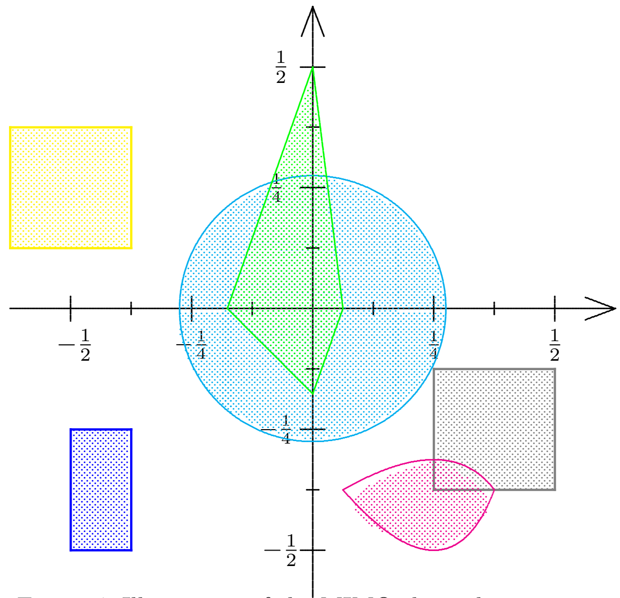

Figure 4. Illustration of the MIMO channel setup.

Clearly, our restriction to Jordan domains is not relevant to applications such as those in communications engineering. The following useful characterization of Jordan domains is well known. It is discussed in detail in.[15]

THEOREM 3.9. *Let $\boldsymbol{S} = (S_{mn}) \subseteq (\mathbb{R}\times\widehat{\mathbb{R}})^{M\times N}$ be the cartesian product of Jordan domains.*

1. *If $\sum_{n=1}^{N} \mu(S_{mn}) < 1$ for all $m \in \{1, \dots, M\}$, then $\boldsymbol{O}PW^2(\boldsymbol{S})$ is identifiable.*
2. *If $\sum_{n=1}^{N} \mu(S_{mn}) > 1$ for some $m \in \{1, \dots, M\}$, then $\boldsymbol{O}PW^2(\boldsymbol{S})$ is not identifiable.*

4. GENERAL OPERATOR PALEY–WIENER SPACES

As noted earlier, the restriction to bandlimited Hilbert–Schmidt operators is too restrictive as it excludes, for example, the identity operator, convolution operators, and multiplication operators as we shall see below.

Let us first consider the identity operator on $L^2(\mathbb{R})$. Formally, we have

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

and we observe that $\sigma_{Id}(x,\xi) = 1$. Clearly, $\sigma_{Id} \notin L^2(\mathbb{R}^2)$ which confirms that the identity operator is not Hilbert–Schmidt. Similarly, let H be a convolution operator with impulse response h. Then Parseval–Plancherell implies that

$$Hf(x) = \int f(x-y)h(y)\, dy = \int \widehat{h}(\xi)\widehat{f}(\xi)e^{2\pi i x\xi}\, d\xi.$$

Clearly $\sigma_H(x,\xi) = \widehat{h}(\xi)$ is not square integrable as a function on $\mathbb{R}^2$. Similarly, the multiplication operator H given by $Hf(x) = m(x)f(x) = \int m(x)\widehat{f}(\xi)e^{2\pi i\xi x}\, d\xi$ satisfies $\sigma(x,\xi) = m(x)$ which does not fit into the framework of Hilbert–Schmidt operators.

Not only are the symbol classes that are needed to describe convolution and multiplication operators not square integrable, but they are anisotropic. That is, to recover their symbol requires us once to ignore the x direction, while at the other time, only the x direction carries the information. To address this phenomenon, we will resort to mixed norm, and, in fact, to so–called modulation spaces.

4.1 Preliminaries on modulation spaces

For $1 \le p < \infty$, $L^p(\mathbb{R}^d)$ denotes the Banach space of complex valued and Lebesgue measurable functions which satisfy $\|f\|_{L^p} = \int |f(x)|^p\, dx < \infty$.[8] As is customary, $L^\infty(\mathbb{R}^d)$ denotes the space of essentially bounded functions with norm $\|f\|_{L^\infty} = \operatorname{ess\,sup}|f(x)|$. Analogously, a sequence $c = \{c_j\}_j \in \mathbb{Z}^d$ belongs to $l^p(\mathbb{Z}^d$, if $\|c\|_{l^p} = \sum |c_j|^p < \infty$ for $1 \le p < \infty$ or $\|c\|_{l^\infty} = \sup |c_j| < \infty$.

Here and in the following, $\mathcal{S}(\mathbb{R}^d)$ denotes the space of *Schwartz functions* on $\mathbb{R}^d$ and $\mathcal{S}'(\mathbb{R}^d)$ its dual of *tempered distributions*. The usefulness of $\mathcal{S}(\mathbb{R}^d)$ and its dual $\mathcal{S}'(\mathbb{R}^d)$ of tempered distributions in harmonic analysis stems in part from the fact that the Fourier transform defines a bijective isomorphisms on $\mathcal{S}(\mathbb{R}^d)$ and on $\mathcal{S}'(\mathbb{R}^d)$ (equipped with the weak-$*$ topology). Also, $\mathcal{S}'(\mathbb{R}^d)$ contains constant functions, *Dirac's delta* $\delta : f \mapsto f(0)$, and *Shah distributions* $⧢_a = \sum_{n\in\mathbb{Z}} \delta_{an}$, where $\delta_{na} = T_{na}\delta$, $a > 0$, which are used below.

Similarly to the *Fourier transformation*, the *time shift operator* T_t, $t \in \mathbb{R}$, given by $T_t f(x) = f(x-t)$ and the *modulation operator* M_ω, $\omega \in \widehat{\mathbb{R}}$, $M_\omega f(x) = e^{2\pi i \omega \cdot x} f(x)$ act as unitary operators on $L^2(\mathbb{R})$ and bijective isomorphism on $\mathcal{S}(\mathbb{R})$ and $\mathcal{S}'(\mathbb{R})$. Note that M_ω is also called *frequency shift operator* since $\widehat{M_\omega f} = T_\omega \widehat{f}$. Further, we refer to $\pi(\lambda) = \pi(t,\nu) = T_t M_\nu$ for $\lambda = (t,\nu) \in \mathbb{R}\times\widehat{\mathbb{R}}$ as *time–frequency shift operator*.

The first ingredient in defining modulation spaces are short–time Fourier transformations. The short time Fourier transform of a tempered distribution $f \in S'(\mathbb{R}^d)$ with respect to the Gaussian $\varphi_0(x) = e^{-\|x\|^2}$, $x \in \mathbb{R}^d$ is given by[21]

$$V_{\varphi_0} f(t,\nu) = \langle f, T_t M_\nu \varphi_0 \rangle, \quad (t,\nu) \in \mathbb{R}^d \times \widehat{\mathbb{R}}^d,$$

where here and in the following the dual pairing $\langle\cdot,\cdot\rangle$ is considered to be linear in the first component and antilinear in the second component. Note that in some cases, we shall use different window functions than φ_0. Regardless which nontrivial Schwartz function we choose, the modulation spaces defined below are identical with equivalent norms.[22]

The second ingredient in defining modulation spaces are mixed L^p spaces which we shall describe now. Namely, for a measurable function f on $\mathbb{R}^d$ and $p = (p_1, \ldots, p_d)$, $1 \le p_1, \ldots, p_d < \infty$, we set

$$\|f\|_{L^p} = \left(\int_{\mathbb{R}} \left(\ldots \left(\left(\int_{\mathbb{R}} |f(x_1,\ldots,x_d)|^{p_1} dx_1 \right)^{p_2/p_1} dx_2 \right)^{p_3/p_2} \ldots dx_{d-1} \right)^{p_d/p_{d-1}} dx_d \right)^{1/p_d},$$

with the customary adjustments if some $p_k = \infty$.

Modulation spaces are defined for $p = (p_1, \ldots, p_d)$ and $q = (q_1, \ldots, q_d)$, $1 \le p_k, q_k \le \infty$, by setting[21,23]

$$M^{p,q} = \left\{ f \in \mathcal{S}'(\mathbb{R}^d) : \ V_{\varphi_0} f \in L^{pq} \right\}.$$

Any modulation space $M^{p,q}$ is a Banach space with norm $\|f\|_{M^{p,q}} = \|V_{\varphi_0} f\|_{L^{pq}}$. Note that we do not use the following convention of writing $M^p = M^{p,p}$ if $p = q$, since this notation somewhat conflicts with PW^p which is defined below and for which $PW^p \subseteq M^{p,q}$ for all $q \ge 1$. Further, we shall always separate decay condition in the time variable from decay conditions on the Fourier side by a comma. To illustrate the ordering of the indices in the modulation spaces $M^{p,q}$ for $d > 1$ and $p \ne q$, we state exemplary that $f \in M^{23,45}$ if and only if

$$\int \Big(\int \Big(\int \Big(\int |V_{\varphi_0} f(t_1, t_2, \nu_1, \nu_2)|^2\, dt_1 \Big)^{\frac{3}{2}} dt_2 \Big)^{\frac{4}{3}} d\nu_1 \Big)^{\frac{5}{4}} d\nu_2 \le \infty.$$

Clearly $f \otimes g \in M^{p_1 p_2, q_1 q_2}$ if and only if $f \in M^{p_1,q_1}$ and $g \in M^{p_2,q_2}$. Further, in this case $\|f \otimes g\|_{M^{p_1 p_2, q_1 q_2}} = \|f\|_{M^{p_1,q_1}} \|g\|_{M^{p_2,q_2}}$.

Let us also note that $M^{1,1}$ is the Feichtinger algebra, often denoted by S_0, and $M^{\infty,\infty}$ is its dual S_0'.

In addition to the modulation spaces given above, we shall consider the weighted modulation spaces

$$M^{p,q}_{\mathrm{s}} = \left\{ f \in \mathcal{S}'(\mathbb{R}^d) : \ V_{\varphi_0} f \, m_{\mathrm{s}} \in L^{pq} \right\},$$

where $m_s(x,\xi) = (1+\xi^2)^{\frac{s}{2}}$ for $s \in \mathbb{R}$. Note that the spaces $M_s^{2,2}$ are also known as Bessel potential spaces.

In the following, we shall concentrate on *Paley–Wiener modulation spaces*, that is, we set for $1 \le p \le \infty$

$$PWs^p(\Omega) = \Big\{ f \in \mathcal{M}_s^{p,\infty}(\mathbb{R}^d) : \ \operatorname{supp} \widehat{f} \subseteq \prod [-\tfrac{\Omega_k}{2}, \tfrac{\Omega_k}{2}] \Big\} \subseteq M^{p,q}(\mathbb{R}^d) \text{ for all } q \ge 1.$$

Note that for $f \in PW^p(\Omega)$, we have $\|f\|_{M^{p,q}} \asymp \|f\|_{M^{p,r}}$ for $1 \le p, q \le r$, and we define $\|\cdot\|_{PW^p} = \|\cdot,\|_{M^{p,1}}$.[7]

4.2 Sampling in Modulation spaces

Possibly include Classical sampling theorem in an L^p or M^p format. Not sure.

4.3 Operator Paley–Wiener spaces

To formulate a sampling theory for operators, we first observe a well known correspondence between a large class of linear operators ... can be written as integral operator with distributional kernel.

THEOREM 4.1. *For any bounded and linear operator $H : \mathcal{S}(\mathbb{R}^d) \longrightarrow \mathcal{S}'(\mathbb{R}^d)$ exists κ_H in$\mathcal{S}'(\mathbb{R}^{2d})$ with $\langle Hf, g\rangle = \langle \kappa, f \otimes g\rangle$.*

Clearly, Theorem 4.1 shows that a large class of bounded and linear operators can be represented by

In the theory of pseudodifferential operators, one commonly describes operators using their so–called Kohn–Nirenberg symbol which is given by. In fact, we have

Equation (16) implies that for given functions $f,\, g \in L^2(\mathbb{R}^d)$,

$$\begin{aligned}\langle Hg, f\rangle &= \iiint \eta_H(t,\nu) T_t M_\nu g(x) \overline{f(x)}\, dt\, d\nu\ dx \\ &= \iint \eta_H(t,\nu) \overline{\int f(x) \overline{T_t M_\nu g(x)}\, dx}\ dt\, d\nu \\ &= \langle \eta_H, V_g f\rangle \end{aligned} \tag{18}$$

where $V_g f(t,\nu) = \langle f, T_t M_\nu g\rangle$, $t \in \mathbb{R}^d$ and $\nu \in \widehat{\mathbb{R}}^d$ is the *short-time Fourier transform (STFT)* of f with respect to the *window function* g. It is clear then that the STFT is a natural tool to study the connection between the properties of the operator H and its spreading function η_H. If $\|g\|_{L^2(\mathbb{R}^d)} = 1$ then the STFT is an isometric isomorphism of $L^2(\mathbb{R}^d)$ onto a closed subspace of $L^2(\mathbb{R}^d \times \widehat{\mathbb{R}}^d)$. In this case the function f can be recovered by

$$f(x) = \iint V_g f(t,\nu)\, T_t\, M_\nu g(x)\, dt\, d\nu$$

whenever the integral makes sense.

- Here we should just state which operator classes act on which modulation spaces. Not much detail, nothing new .

THEOREM 4.2. *Sampling theorem with rectangle but general p, q, weight s.*

THEOREM 4.3. *$OPW^\infty(\Omega)$ is identifiable if $S_\Omega < 1$, not identifiable if $S_\Omega > 1$*

Note that if $\mathrm{s} \le -n$, $n \in \mathbb{N}$, then the operators corresponding to $PW_{\mathrm{s}}^{\infty,\infty}(S)$ include linear differential operators $\sum_{k=0}^n a_k(x) \frac{\partial^k}{\partial x^k}$ with bounded a_k and $\{0\} \times \bigcup_k \operatorname{supp} \widehat{a}_k \subset S$, and any pseudodifferential operators K of order m for which σ_K satisfies $\operatorname{supp} \widehat{\sigma_K} \subseteq S$.[24]

[7] This should be right.

5. APPENDIX: JORDAN DOMAINS AND JORDAN CONTENT

The stated goal of this paper is to extend the proof of Kailath's conjecture which is proved for rectangles and parallelograms in[2] to "essentially arbitrary" regions. In this section we describe more precisely what is meant by "essentially arbitrary" from a mathematical point of view. Taking into account the requirements of the proof of Theorem **??**, we are led naturally to the notion of Jordan content and Jordan domains. The definition we use for convenience differs from but is equivalent to those found in most textbooks, for example, see.[8]

DEFINITION 5.1. *For* $K, L \in \mathbb{N}$ *set* $R_{K,L} = [0, \frac{1}{K}] \times [0, \frac{K}{L}]$ *and*

$$\mathcal{U}_{K,L} = \left\{ \bigcup_{j=1}^{J} \Big(R_{K,L} + (\tfrac{k_j}{K}, \tfrac{p_j K}{L}) \Big) : \; k_j, p_j \in \mathbb{Z}, J \in \mathbb{N} \right\}.$$

Let $M \subset \mathbb{R}\times\widehat{\mathbb{R}}$ *be bounded and let* μ *be the Lebesgue measure on* $\mathbb{R}\times\widehat{\mathbb{R}}$*. The* inner content *of* M *is defined as*

$$\begin{aligned}&\mathrm{vol}^-(M) = \\ &\quad \sup\{\mu(U) : U{\subset}M \text{ and } U{\in}\mathcal{U}_{K,L} \text{ for some } K{,}L{\in}\mathbb{N}\}\end{aligned} \tag{19}$$

and the outer content *of* M *is given by*

$$\begin{aligned}&\mathrm{vol}^+(M) = \\ &\quad \inf\{\mu(U) : U{\supset}M \text{ and } U{\in}\mathcal{U}_{K,L} \text{ for some } K{,}L{\in}\mathbb{N}\}.\end{aligned} \tag{20}$$

Clearly, we have $\mathrm{vol}^-(M) \leq \mathrm{vol}^+(M)$ *and if* $\mathrm{vol}^-(M) = \mathrm{vol}^+(M)$*, then we say that* M *is a* Jordan domain *with* Jordan content $\mathrm{vol}(M) = \mathrm{vol}^-(M) = \mathrm{vol}^+(M)$.

In the following proposition we collect some relevant facts on Jordan content (see for example,[8]).

PROPOSITION 5.2. *Let* $M \subset \mathbb{R}\times\widehat{\mathbb{R}}$.

1. *If* M *is a Jordan domain, then* M *is Lebesgue measurable with* $\mu(M) = \mathrm{vol}(M)$.
2. *If* M *is Lebesgue measurable and bounded and its boundary* ∂M *is a Lebesgue zero set, that is,* $\mu(\partial M) = 0$*, then* M *is a Jordan domain.*
3. *If* M *is open, then* $\mathrm{vol}^-(M) = \mu(M)$ *and if* M *is compact, then* $\mathrm{vol}^+(M) = \mu(M)$.
4. *If* $\mathcal{P} \subset \mathbb{N}$ *is unbounded, then replacing the quantifier "*for some $L \in \mathbb{N}$*" with "*for some $L \in \mathcal{P}$*" in (19) and in (20) leads to equivalent definitions of inner and outer Jordan content.*

REFERENCES

1. T. Kailath, "Measurements on time–variant communication channels," *IEEE Trans. Inform. Theory* **8**, pp. 229–236, Sept. 1962.
2. W. Kozek and G. Pfander, "Identification of operators with bandlimited symbols," *SIAM J. Math. Anal.* **37**(3), pp. 867–888, 2006.
3. G. Pfander and D. Walnut, "Measurement of time–variant channels," *IEEE Trans. Info. Theory* **52**(11), pp. 4808–4820, 2006.
4. P. Bello, "Measurement of random time-variant linear channels," *IEEE Trans. Comm.* **15**, pp. 469–475, 1969.
5. H. Lüke, "The origins of the sampling theorem," *IEEE Commun. Mag.* **37**(4), pp. 106–108, 1999.
6. A. Aldroubi and K. Gröchenig, "Nonuniform sampling and reconstruction in shift-invariant spaces," *SIAM Rev.* **43**(4), pp. 585–620 (electronic), 2001.
7. A. Jerri, "The shannon sampling theoremits various extensions and applications: A tutorial review," *IEEE Proceedings* **65**(11), pp. 1565–1596, 1977.

8. G. B. Folland, *Real analysis*, Pure and Applied Mathematics (New York), John Wiley & Sons Inc., New York, second ed., 1999. Modern techniques and their applications, A Wiley-Interscience Publication.
9. H. G. Feichtinger and K. Gröchenig, "Theory and practice of irregular sampling," in *Wavelets: mathematics and applications*, *Stud. Adv. Math.*, pp. 305–363, CRC, Boca Raton, FL, 1994.
10. H. Landau, "Necessary density conditions for sampling an interpolation of certain entire functions.," *Acta Math.* **117**, pp. 37–52, 1967.
11. J. Dieudonné, *Treatise on analysis. Vol. II*, Translated from the French by I. G. MacDonald. Pure and Applied Mathematics, Vol. 10-II, Academic Press, New York, 1970.
12. S. Gaal, *Linear Analysis and Representation Theory*, Springer–Verlag, New York, 1973.
13. J. Kohn and L. Nirenberg, "An algebra of pseudo–differential operators," *Comm. Pure Appl. Math.* **18**, pp. 269–305, 1965.
14. G. Folland, *Harmonic Analysis in Phase Space*, Annals of Math. Studies, Princeton Univ. Press, Princeton, NJ, 1989.
15. J. Lawrence, G. Pfander, and D. Walnut, "Linear independence of Gabor systems in finite dimensional vector spaces," *J. Fourier Anal. Appl.* **11**(6), pp. 715–726, 2005.
16. A. Goldsmith, *Wireless Communications*, Cambridge University Press, Cambridge, 2005.
17. H. Malvar, *Wireless Communications over MIMO Channels*, John Wiley & Sons, Chichester, 2006.
18. A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*, Cambridge University Press, Cambridge, 2003.
19. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. trans. telecommun.* **10**, pp. 585–959, 1999.
20. N. Grip and G. Pfander, "A discrete model for the efficient analysis of time–varying narrowband communications channels.".
21. K. Gröchenig, *Foundations of Time-Frequency Analysis*, Applied and Numerical Harmonic Analysis, Birkhäuser, Boston, MA, 2001.
22. H. Feichtinger and G. Zimmermann, "A Banach space of test functions for Gabor analysis," in *Gabor Analysis and Algorithms: Theory and Applications*, H. Feichtinger and T. Strohmer, eds., pp. 123–170, Birkhäuser, Boston, MA, 1998.
23. H. Feichtinger, "On a new segal algebra," *Monatsh. Math.* **92**, pp. 269–289, 1981.
24. M. E. Taylor, *Pseudodifferential operators*, vol. 34 of *Princeton Mathematical Series*, Princeton University Press, Princeton, N.J., 1981.
25. B. Fleury and P. Leuthold, "Radiowave propagation in mobile communications: an overview of european research," *IEEE Commun. Mag.* **34**(2), pp. 70–81, 1996.
26. D.-S. Yoo, *Channel characterization and system design in wireless communications.* PhD thesis, Univ. of Michigan, Ann Arbor, 2002.
27. T. Kailath, "Sampling models for linear time-variant filters," Tech. Rep. 352, Massachusetts Institute of Technology, Research Laboratory of Electronics, 1959.