DONOHO-LOGAN LARGE SIEVE PRINCIPLES FOR MODULATION AND POLYANALYTIC FOCK SPACES

LUÍS DANIEL ABREU AND MICHAEL SPECKBACHER

To the memory of Kurt Gödel

Abstract. We obtain estimates for the $L^p$-norm of the short-time Fourier transform (STFT) for functions in modulation spaces, providing information about the concentration on a given subset of $\mathbb{R}^2$, leading to deterministic guarantees for perfect reconstruction using convex optimization methods. More precisely, we will obtain large sieve inequalities of the Donoho-Logan type, but instead of localizing the signals in regions $T \times W$ of the time-frequency plane using the Fourier transform to intertwine time and frequency, we will localize the representation of the signals in terms of the short-time Fourier transform in sets $\Delta$ with arbitrary geometry. At the technical level, since there is no proper analogue of Beurling’s extremal function in the STFT setting, we introduce a new method, which rests on a combination of an argument similar to Schur’s test with an extension of Seip’s local reproducing formula to general Hermite windows. When the windows are Hermite functions, we obtain local reproducing formulas for polyanalytic Fock spaces which lead to explicit large sieve constant estimates and, as a byproduct, to a reconstruction formula for $f \in L^2(\mathbb{R})$ from its STFT values on arbitrary discs. A discussion on optimality follows, along the lines of Donoho-Stark paper on uncertainty principles and signal recovery. We also consider the case of discrete Gabor systems, vector-valued STFT transforms and rephrase the results in terms of the polyanalytic Bargmann-Fock transforms.

MSC2010: 42C40, 46E15, 46E20, 42C15, 11N36

Keywords: Large Sieve, deterministic compressive sensing, short-time Fourier transform, modulation spaces, polyanalytic functions, concentration estimates, signal recovery, multiplexing, Banach frames, Hermite functions

1. Introduction

The large sieve principle covers a number of far reaching analysis techniques, mostly aimed at solving problems in analytic number theory, but which have also found applications in a number of other mathematical fields, like probability [47], numerical [11] and signal analysis [23, 24, Theorem 7], to name a few. The terminology stems from its number theory origins, which can be traced back to the sieve of Eratosthenes. In number theory, the large sieve principle is mostly concerned with asymptotic averages of arithmetic functions on integers constrained by congruences modulo sets of primes. A typical example of the large sieve
principle is the inequality for trigonometric polynomials:

\[
\left| \sum_{l=1}^{R} \sum_{k=m+1}^{m+n} a_k e^{2\pi i kx_l} \right|^2 \leq \Delta(n, \delta) \sum_{k=m+1}^{m+n} |a_k|^2,
\]

where the points \(x_1, \ldots, x_R\) are \(\delta\)-separated mod 1. By choosing them as fractions \(p/q\) with \(\gcd(p, q) = 1\), several applications in number theory follow \([47, 51]\).

According to inequality (1.1), only a small energy portion of the trigonometric polynomial is concentrated at the points \(\alpha_1, \ldots, \alpha_R\), with the constant \(\Delta(n, \delta)\) controlling the size of the fraction. Since the energy concentration inside a domain is important to find optimal approximation methods required in signal recovery, such an observation suggests applications in signal analysis, where one can find a rich setting. Donoho and Logan \([23]\), in a paper that, together with \([24]\), spearheaded the modern theory of Compressed Sensing \([19, 25, 33]\), introduced the concept of maximum Nyquist density, \(\rho(T, W)\), which measures the sparsity of a real band-limited signal on the time domain \(T \subset \mathbb{R}\) with band-size \(W\):

\[
\rho(T, W) := W \cdot \sup_{t \in \mathbb{R}} |T \cap [t, t + 1/W]| \leq W \cdot |T|.
\]

Let us motivate our work with some results from \([23, 24]\). If the set \(T\) has large area but small Lebesgue measure on any interval of length \(1/W\), then \(\rho(T, W)\) can be considerably small compared to the natural Nyquist density \(W \cdot |T|\). We will call such sets \(T\) sparse in the sense of Lebesgue measure. Throughout the paper we will write \(P_T f := \chi_T f\) to denote multiplication by the indicator function of \(T\).

While analytic number theory is mostly interested in Hilbert space large sieve inequalities, in signal analysis one finds remarkable applications of Banach space large sieve inequalities, with a special emphasis on \(L^1\)-normed spaces. In \([23\) Theorem 7], the authors considered the space

\[
B_1(W) := \left\{ f \in L^1(\mathbb{R}) : \text{supp}(\hat{f}) \subseteq [-\pi W, \pi W] \right\}
\]

and proved, for \(\theta < \frac{2\pi}{W}\), the inequality

\[
\|P_T f\|_1 \leq \frac{\pi W/2}{\sin(\pi W \theta/2)} \left( \sup_{t \in \mathbb{R}} |T \cap [t, t + \theta]| \right) \cdot \|f\|_1.
\]

An inequality similar as (1.4) also holds if \(\chi_T(x)\) is replaced by a positive \(\sigma\)-finite measure \(\mu\). It provides the concentration bound

\[
\delta_1(T) := \sup_{f \in B_1(W)} \frac{\|P_T f\|_1}{\|f\|_1} \leq \frac{\pi}{2} \rho(T, W),
\]

resulting in sufficient conditions for perfect reconstruction of a band-limited signal corrupted by sparse noise using \(L^1\)-norm minimization.
In this paper, we will obtain large sieve inequalities of the Donoho-Logan type, but instead of localizing the signals in regions $T \times W$ of the time-frequency plane using the Fourier transform to intertwine time and frequency, we will localize the representation of the signals in terms of the short-time Fourier transform (STFT)

$$V_gf(x, \xi) = \int_{\mathbb{R}} f(t)g(t-x)e^{-2\pi i \xi t}dt,$$

to general regions $\Delta$ of the time-frequency plane. Instead of the maximum Nyquist density (1.2) we will use the following concept of planar maximum Nyquist density introduced in [10]:

$$\rho(\Delta, R) := \sup_{z \in \mathbb{R}^2} |\Delta \cap (z + D_{1/R})| \leq |\Delta|,$$

where $D_{1/R} \subset \mathbb{R}^2$ is the disc of radius $1/R$ centered in the origin. If the set $\Delta$ is sparse in the sense of Lebesgue measure (small concentration in any disc of radius $1/R$), then $\rho(\Delta, R)$ can be considerably smaller than the natural Nyquist density $|\Delta|$ (see [21, 22, 35, 49, 50, 59] for natural Nyquist densities in the context of Fock and Gabor spaces). Instead of measuring the concentration of band-limited signals in a time-limited region $T \subset \mathbb{R}$, we will measure the joint time-frequency content on a region $\Delta \subset \mathbb{R}^2$. In its most general version, our results can be seen as estimates on the bounds of Bessel measures for the STFT [12, 52]. By selecting Hermite functions as windows in (1.5), good explicit estimates in terms of $\rho(\Delta, R)$ can be obtained. The following is a sample of our findings in the $L^1$-case (we will prove it for $1 \leq p < \infty$). The modulation space $M^1$, also known as Feichtinger’s algebra $S_0$ [27], will play the role of the space $B_1(W)$ in [23].

**Theorem.** Let $\Delta \subset \mathbb{R}^2$ be measurable and $f \in M^1$. Denote by $h_r$ the $r$th Hermite function. For every $0 < R < \infty$,

$$\|V_{h_r}f \cdot \chi_{\Delta}\|_1 \leq \frac{\rho(\Delta, R)}{C_r(R)} \|V_{h_r}f\|_1,$$

where the constant $C_r(R)$ is explicitly determined.

This will provide estimates for the concentration bound

$$\delta(\Delta) := \sup_{f \in M^1} \frac{\|V_{h_r}f \cdot \chi_{\Delta}\|_1}{\|V_{h_r}f\|_1},$$

and, consequently, conditions for perfect reconstruction of a $M^1$ function corrupted by sparse noise using $L^1$-norm minimization (precise statements are given in section 5.1).

The techniques of proof are new in sieve theory and, in particular, different from those in [23, 51], where Beurling’s extremal function [17] plays a key role. Since we are not aware of a proper analogue of extremal function theory in the STFT setting, we had to develop new methods, which essentially depend on combining an argument similar to Schur’s test with an extension of Seip’s local reproducing formula’s [59] to general Hermite windows.
The paper is organized as follows. In the preliminaries section we gather the essential background on time-frequency analysis. In Section 3 we formulate our results in a general Banach space setting, provide a general discussion about the signal recovering applications that motivate the results and extend Selberg-Bombieri inequality [16] to the continuous setting. In Section 4 we restrict to Hermite windows. Then, we extend Seip\'s local reproducing formulas [59] to polyanalytic Fock spaces associated with the Landau levels and use them to obtain explicit estimates for the maximum Nyquist density in the Hermite window case. Section 5 contains a discussion on optimality of the constants, including a phase-space versions of a theorem by Donoho-Stark [24, Theorem 10]. Moreover, we revisit the signal recovery problem in the context of reconstructing STFT data. In Section 6 we obtain large sieve inequalities for discrete and vector valued Gabor systems. Finally, we conclude with a section discussing some open problems.

2. Preliminaries

2.1. The short-time Fourier transform. Let \( z = (x, \xi), \ w = (y, \eta) \in \mathbb{R}^2 \) and \( g \in L^2(\mathbb{R}) \). A time-frequency shift of the function \( g \) is defined as

\[
\pi(z)g(t) := M_\xi T_x g(t) = e^{2\pi i \xi t} g(t - x),
\]

where \( T_x \) denotes the translation operator and \( M_\xi \) the modulation operator. The composition of two time-frequency shifts is given by

\[
\pi(z)\pi(w) = e^{-2\pi i x \eta} \pi(z + w)
\]

and the adjoint operator of \( \pi(z) \) is

\[
\pi(z)^* = e^{-2\pi i x \xi} \pi(-z).
\]

The short-time Fourier transform (STFT) or Gabor transform of a function \( f \) with window \( g \) is defined by

\[
V_g f(z) := \langle f, \pi(z)g \rangle = \int_{\mathbb{R}} f(t)g(t - x)e^{-2\pi i t \xi} dt.
\]

An important property of the STFT is the so called orthogonality relation

\[
\langle V_{g_1} f_1, V_{g_2} f_2 \rangle = \langle f_1, f_2 \rangle \langle g_2, g_1 \rangle.
\]

In particular, if \( \|g\|_2 = 1 \), then

\[
\|V_g f\|_2 = \|f\|_2
\]

and \( V_g : L^2(\mathbb{R}) \to L^2(\mathbb{R}^2) \) is an isometry mapping onto the reproducing kernel Hilbert space

\[
V_g := \{V_g f : f \in L^2(\mathbb{R})\} \subset L^2(\mathbb{R}^2).
\]

The corresponding reproducing equation is

\[
V_g f(z) = \langle V_g f, K_g(z, \cdot) \rangle,
\]
where $K_g(z, w) := \langle \pi(w)g, \pi(z)g \rangle$. Thus, one can write the orthogonal projection of $L^2(\mathbb{R}^2)$ on $\mathcal{V}_g$ as the integral operator with kernel $K_g$. We note in passing that $K_g(z, w)$ is the correlation kernel in the Weyl-Heisenberg ensemble \[8, 9].

2.2. Hermite functions and complex Hermite polynomials. In time-frequency analysis, a particular interest is given to functions well concentrated in both time and frequency. A class of such functions is given by the Hermite functions $h_r$ defined as

$$h_r(t) = \frac{2^{1/4}}{\sqrt{\pi}} \left( \frac{-1}{2\sqrt{\pi}} \right)^r e^{\pi t^2} \frac{d^r}{dt^r} \left( e^{-2\pi t^2} \right), \quad r \geq 0.$$ 

The collection $\{h_r\}_{r \geq 0}$ forms an orthonormal basis for $L^2(\mathbb{R})$, minimizes the uncertainty principle \[18\] and optimizes the joint time-frequency concentration on discs \[21, 59\]. In \[38\], precise lattice conditions for vector valued frames with Hermite functions have been obtained, which, combined with Vasilevski’s work \[63\], inspired the study of sampling and interpolation problems in Fock spaces of polyanalytic functions \[2\], a hierarchy of function spaces which seems to be ubiquitous in several mathematical models \[4\]. The Hermite function $h_0$ is the Gaussian function explicitly given by

$$\varphi(t) = h_0(t) = \frac{1}{\sqrt{\pi}} e^{-\pi t^2}.$$ 

We will also use the so-called complex Hermite polynomials \[36, 46\]:

$$H_{j,r}(z, \overline{z}) = \begin{cases} \sqrt{\frac{2^j}{j!}} \pi^j z^{j-r} L_{j-r}^r \left( \pi |z|^2 \right), & j > r \geq 0, \\ (-1)^{r-j} \sqrt{\frac{2^j}{j!}} \pi^j \overline{z}^{j-r} L_{j-r}^r \left( \pi |z|^2 \right), & 0 \leq j \leq r, \end{cases}$$

where $L_{j-r}^r$ stands for the generalized Laguerre polynomials defined via the recurrence relation $L_0^0(x) = 1$, $L_1^0(x) = 1 + \alpha - x$ and

$$L_{j+1}^\alpha(x) = \frac{2j + 1 + \alpha - x}{j + 1} L_j^\alpha(x) - \frac{j + \alpha}{j + 1} L_{j-1}^\alpha(x), \quad j \geq 1.$$ 

If $\alpha \geq 0$, then $L_j^\alpha$ has the following closed form

$$L_j^\alpha(x) = \sum_{i=0}^{j} (-1)^i \binom{j + \alpha}{j - i} \frac{x^i}{i!}, \quad x \in \mathbb{R}, \quad j, \alpha \in \mathbb{N}_0.$$ 

Complex Hermite polynomials satisfy the doubly-indexed orthogonality

$$\int_{\mathbb{C}} H_{j,r}(z, \overline{z}) \overline{H_{j',r'}(z, \overline{z})} e^{-\pi |z|^2} dz = \delta_{jj'}\delta_{rr'},$$

and provide a basis for the space $L^2(\mathbb{C}, e^{-\pi |z|^2})$ \[7, 42\]. The relation between time-frequency analysis and polyanalytic functions \[1, 2\] can be understood in terms of the Laguerre connection \[32\], Chapter 1.9]

$$V_{h_r}h_j(x, -\xi) = e^{i\pi \xi \frac{r}{2}|z|^2} H_{j,r}(z, \xi).$$
The closed form of the reproducing kernels $K_{h_r}$ reads
\begin{equation*}
K_{h_r}(z,w) = \langle \pi(w)h_r, \pi(z)h_r \rangle = e^{\pi(x+y)(\xi-\eta)} L_r^0(\pi|z-w|^2)e^{-\pi|z-w|^2/2}.
\end{equation*}

Consequently,
\begin{equation}
\left|K_{h_r}(z,w)\right| = L_r^0(\pi|z-w|^2)e^{-\pi|z-w|^2/2}.
\end{equation}

The kernel $K_{h_r}$ describes the orthogonal projection onto the Bargmann-Fock space of pure polyanalytic functions of type $r$ (see Remark 1), which is precisely the $r$th-eigenspace of the Euclidean Landau operator with a constant magnetic field [5, 13, 42, 58].

2.3. Modulation spaces and (Banach-)frame theory. In order to quantitatively measure the behavior of a class of transformations generated by an integrable group representation, Feichtinger and Gröchenig developed the theory of coorbit spaces [28, 29, 30]. We will consider the particular instance of the coorbit spaces associated with the short-time Fourier transform, the so called modulation spaces [26]. Let $g$ be a window function satisfying $\|V_g g\|_1 < \infty$. The modulation space $M^p$ is defined as
\begin{equation}
M^p := \{f \in S'({\mathbb R}) : V_g f \in L^p({\mathbb R}^2)\},
\end{equation}
where $S'({\mathbb R})$ denotes the space of tempered distributions. The space $M^1$, also known as Feichtinger’s algebra $S_0$, will play the role of $B_1(W)$ in [23] as the fundamental space for applications in signal recovery using $L_1$-minimization. Since the reproducing kernel property extends to $M^p$, the range of the short-time Fourier transform on $M^p$ can be characterized in terms of projections on $L^p({\mathbb R}^2)$ as follows:
\begin{equation}
V_g^p := V_g(M^p) = \{F \in L^p({\mathbb R}^2) : \langle F, K_g(z, \cdot) \rangle = F(z)\}.
\end{equation}

Let $\Lambda \subset {\mathbb R}^2$ be a discrete set. The family $\{\pi(\lambda)g\}_{\lambda \in \Lambda}$ is a Gabor frame for $L^2({\mathbb R})$ if there exist positive constants $A, B$ such that
\begin{equation}
A\|f\|_2^2 \leq \sum_{\lambda \in \Lambda} |\langle f, \pi(\lambda)g \rangle|^2 \leq B\|f\|_2^2, \quad \forall f \in L^2({\mathbb R}).
\end{equation}

It follows from the theory of modulation spaces that a Gabor frame $G(g, \Lambda)$ with $g \in M^1$ is also a Gabor frame for every space $M^p$ with $p \geq 1$ [37, Theorem 13.6.1]. This is summarized in the following remark.

Remark 1. Assume that $g \in M^1$ and $G(g, \Lambda)$ is a Gabor frame for $L^2({\mathbb R})$. There exist two constants $A', B' > 0$ such that
\begin{equation}
A'\|f\|_{M^p}^p \leq \sum_{\lambda \in \Lambda} |\langle f, \pi(\lambda)g \rangle|^p \leq B'\|f\|_{M^p}^p, \quad \forall f \in M^p.
\end{equation}
Furthermore, there exists a dual window \( \gamma \in M^1 \) such that

\[
f = \sum_{\lambda \in \Lambda} \langle f, \pi(\lambda)g \rangle \pi(\lambda)\gamma, \quad \forall f \in M^p,
\]

with unconditional convergence in \( M^p \) if \( p < \infty \) and weak-* convergence in \( M^\infty \).

3. A General Large Sieve Principle

In this section we will first formulate our problem and applications in a very general setting. Later, particular situations with more structure will be considered, where explicit estimates for the large sieve constant can be obtained.

3.1. Sieving inequalities in Banach spaces. Let \((X, \mu), (X, \nu)\) be measure spaces and \( B_p \subset L^p(X, \nu) \) a Banach space. We derive a bound on the embedding \((B_p, \| \cdot \|_{L^p_\nu}) \hookrightarrow (B_p, \| \cdot \|_{L^p_\mu})\) using an argument similar to Schur’s test. Define the Banach algebra \( A_\mu \) of hermitian integral kernels, \( K(x, y) = \overline{K(y, x)} \), equipped with the norm

\[
\|K\|_{A_\mu} := \sup_{y \in X} \int_X |K(x, y)|d\mu(x)
\]

and multiplication rule

\[
K_1 \circ K_2(x, y) = \int_X K_1(x, z)K_2(z, y)d\mu(z).
\]

**Proposition 1.** Let \( \mu \) be a positive \( \sigma \)-finite measure on \( X \), \( B_1 \subset L^1(X, \nu) \) and \( K \in A_\mu \) be such that \( K : B_1 \to B_1 \),

\[
KF(x) := \int_X F(y)K(x, y)d\nu(y),
\]

is bounded and boundedly invertible on \( B_1 \). Then, for every \( F \in B_1 \), defined as

\[
\frac{\int_X |F|d\mu}{\int_X |F|d\nu} \leq \theta(K) \cdot \|K\|_{A_\mu},
\]

where

\[
\theta(K) := \sup_{\phi \in B_1} \left( \frac{\|\phi\|_{L^1_\nu}}{\|K\phi\|_{L^1_\mu}} \right).
\]
Proof: Let $F^* \in B_1$ be the unique function such that $KF^* = F$. Using Fubini’s theorem, we have

$$
\int_X |F(x)|d\mu(x) \leq \int_X \int_X |F^*(y)K(x,y)|d\nu(y)d\mu(x)
= \int_X |F^*(y)| \int_X |K(x,y)|d\mu(x)d\nu(y)
\leq \|F^*\|_{L^1} \cdot \|K\|_{A_\mu}
= \frac{\|F^*\|_{L^1}}{\|KF^*\|_1} \cdot \|K\|_{A_\mu} \cdot \|F\|_{L^1}
\leq \theta(K) \cdot \|K\|_{A_\mu} \cdot \|F\|_{L^1}.
$$

□

The following result follows immediately by complex interpolation.

Corollary 1. Let $\|F\|_{L^\infty} \leq C_\infty \|F\|_{L^\infty}$, $\forall F \in B_\infty$ and $1 \leq p < \infty$. Under the assumptions of Proposition 1, the following inequality holds

$$(3.2) \quad \frac{\int_X |F|^p d\mu}{\int_X |F|^p d\nu} \leq C_\infty^{p-1} \cdot \theta(K) \cdot \|K\|_{A_\mu}, \quad \forall F \in B_1 \cap B_\infty.$$

3.2. Localization and signal recovery. We now explain how the general estimates of the previous section can provide useful information for two signal recovery scenarios. Let $B_p \subset L^p(X)$ be a Banach space, where $\Delta \subset X \subset \mathbb{R}^d$ and define

$$
\delta_p(\Delta) := \sup_{F \in B_p} \frac{\int_\Delta |F(x)|^p dx}{\int_X |F(x)|^p dx}.
$$

3.2.1. Scenario 1: Perfect recovery of a signal corrupted by sparse noise by $L^1$-minimization. Let us assume that we observe a noisy version of a signal $F \in B_1$. In addition let us assume that the noise $N$ has arbitrary but finite $L^1$-norm and is supported on an unknown set $\Delta$. The following recovery result of Donoho and Stark [24] tells us that perfect reconstruction is possible if only less than $1/2$ of the mass of a function in $B_1$ can be concentrated on $\Delta$.

Proposition 2. Let $\delta_1(\Delta) < 1/2$ and $G = F + N$, where $F \in B_1$ and $\text{supp}(N) \subset \Delta$. Then perfect reconstruction of $F$ is possible via

$$
F = \arg \min_{B \in B_1} \|B - G\|_1.
$$

3.2.2. Scenario 2: Approximated recovery of missing data by $L^1$-minimization. Let us now assume that the information of the signal on $\Delta$ is missing, so that one observes

$$
H(x) = \begin{cases} 
(F + N)(x), & \text{for } x \notin \Delta \\
0, & \text{for } x \in \Delta
\end{cases}.
$$
with the noise $N$ having small norm. The task of finding approximations of $F$ from $G$ is also known as the inpainting problem in signal processing [11, 54].

**Proposition 3.** Let $F \in B_1$, $\|N\|_1 \leq \varepsilon$ and $\delta_1(\Delta) < 1$. Set

$$\beta(H) := \arg \min_{S \in B_1} \|(I - P_\Delta)(H - S)\|_1.$$  

For any solution $\beta(H)$ we have

$$\|F - \beta(H)\|_1 \leq \frac{2\varepsilon}{1 - \delta_1(\Delta)}.$$  

**Proof:** The proof follows from the estimate

$$\|(I - P_\Delta)(H - \beta(H))\|_1 \leq \|(I - P_\Delta)(F + N - F)\|_1 \leq \|n\|_1 \leq \varepsilon,$$

since

$$\|F - \beta(H)\|_1 = \|(I - P_\Delta)(F - \beta(H))\|_1 + \|P_\Delta(F - \beta(H))\|_1$$

$$\leq \|(I - P_\Delta)(H - \beta(H))\|_1 + \|(I - P_\Delta)(F - H)\|_1 + \|\Delta\|\|F - \beta(H)\|_1$$

$$\leq 2\varepsilon + \delta_1(\Delta)\|F - \beta(H)\|_1.$$  

□

**Remark 2.** It is essential to assume the existence of a solution to the minimization problem of Proposition 3. Moreover, the solution is not necessarily unique.

A similar result can be shown using $L^2$-minimization:

**Proposition 4.** Let $F \in B_2$ and define

$$\gamma(H) := \arg \min_{S \in B_2} \|(I - P_\Delta)(H - S)\|_2.$$  

If $\delta_2(\Delta) < 1$, then for any solution $\gamma(H)$ it holds

$$\|F - \gamma(H)\|_2^2 \leq \frac{4\varepsilon^2}{1 - \delta_2(\Delta)}.$$  

**Proof:** Use the argument of Proposition 3 and the inequality $(a + b)^2 \leq 2a^2 + 2b^2$. □

3.3. **Selberg’s inequality.** A similar estimate for the case $p = 2$ can be formulated as a direct corollary of Selberg’s inequality, as stated by Bombieri in [16]. Since the proof follows the same structure of the proof in [16, Proposition 1] by replacing sums by integrals, we leave the details to the interested reader.

**Proposition 5.** Let $\mathcal{H}$ be a separable Hilbert space, $(X, \mu)$ a measure space, $\Delta \subseteq X$ and $\psi : X \to \mathcal{H}$. Then

$$\int_\Delta \frac{\left| \langle f, \psi_x \rangle \right|^2}{\int_\Delta \left| \langle \psi_x, \psi_y \rangle \right| \, d\mu(y)} \, d\mu(x) \leq \|f\|^2,$$

$\forall f \in \mathcal{H}$.  

Corollary 2. Under the same assumption as before,
\[(3.3) \quad \int_{\Delta} |\langle f, \psi_x \rangle|^2 d\mu(x) \leq \sup_{x \in \Delta} \int_{\Delta} |\langle \psi_x, \psi_y \rangle| d\mu(y), \quad \forall f \in \mathcal{H}.\]

If we take the $\psi_x$ to be time-frequency shifts of a function $g \in M^1$ with $\|g\|_2 = 1$, we obtain the following result.

Corollary 3. Let $\Delta \subset \mathbb{R}^2$ and $g \in M^1$. Then, for every $f \in L^2(\mathbb{R})$,
\[(3.4) \quad \int_{\Delta} |V_g f(z)|^2 dz \leq \sup_{z \in \Delta} \int_{\Delta} |K_g(z, w)| dw.\]

Note that this result yields a slight improvement to Proposition 6 of the next section in the case $p = 2$, since the supremum only needs to be taken over $\Delta$ and not over $\mathbb{R}^2$.

In the next sections, we will restrict the windows $g$ to the family of Hermite functions, where, besides a sharp off diagonal fall of the kernel, one can perform explicit computations and obtain workable explicit large sieve constants.

4. Local reproducing formulas and explicit maximum Nyquist density estimates

4.1. Sieving inequalities for general short-time Fourier transforms. Specializing the result for general Banach spaces from Corollary 1, the following inequality for the STFT in modulation spaces follows.

Proposition 6. Let $g \in M^1$ with $\|g\|_2 = 1$. For $f \in M^p$, $1 \leq p < \infty$, it holds
\[(4.1) \quad \frac{\|V_g f \cdot \chi_{\Delta}\|_p}{\|V_g f\|_p^p} \leq \sup_{z \in \mathbb{R}^2} \int_{\Delta} |K_g(z, w)| dw.\]

Despite the apparent simplicity of the above inequality, it is virtually of no use without information about the kernel $K_g(z, w)$. If the kernel has proper off-diagonal decay properties, one expects to obtain good large sieve constants and (4.1) can be simplified. Indeed, assume that $K_g$ is $\varepsilon$-concentrated on $\Omega \subset \mathbb{R}^2$, more precisely, that
\[\sup_{z \in \mathbb{R}^2} \int_{\mathbb{R}^2 \setminus (z + \Omega)} |K_g(z, w)| dw < \varepsilon\]
and $\|g\|_2 = 1$. Then,
\[(4.2) \quad \frac{\|V_g f \cdot \chi_{\Delta}\|_p}{\|V_g f\|_p^p} \leq \sup_{z \in \mathbb{R}^2} |\Delta \cap (z + \Omega)| + \varepsilon.\]

Depending on the window $g$, the set $\Omega$ may have to be chosen to be big, leading to bad estimates in (4.2). Using the local reproducing formulas of the next subsection, we will see that, for circular domains $\Omega \subset \mathbb{R}^2$ and choosing windows from the Hermite function sequence, one can obtain telling explicit estimates.
4.2. Local reproducing formulas for Hermite windows and circular domains. With a slight abuse of language, double orthogonality often refers to orthogonality in concentric domains. This is known to be the case for STFT’s of Hermite functions with Gaussian windows \[59\] which span the Bargmann-Fock space of entire functions. In this contribution we show that also the Hermite functions allow for local reproducing formulas, extending the results in \[9, Proposition 4.2\] to Bargmann-Fock spaces of polyanalytic functions. In fact we show even more: the reproducing kernel corresponding to the Hermite function \(h_r\) locally reproduces the short-time Fourier transform using the window function \(h_j\). At first sight, this may be perceived as a counter-intuitive result, since

\[
\int_{\mathbb{R}^2} V_{h_r}(z)V_{h_j}(z)dz = 0, \text{ for } r \neq j,
\]

by the orthogonality relation (4.3).

We could have used similar methods as in \[9, Proposition 4.2\], but we provide a more direct proof, based on the expression of the complex Hermite polynomials in terms of Laguerre functions (2.5).

**Proposition 7.** Denote by \(D_R\) the disc of radius \(R\) centered at 0. It then holds

\[
\int_{D_R} H_{j,r}(z,\overline{z})H_{j',r'}(z,\overline{z})e^{-\pi|z|^2}dz = C_{j,r,j',r'}(R) \cdot \delta_{j-r-j',r'},
\]

with

\[
C_{j,r,j',r'}(R) = \sqrt{\frac{r!r'}{j!j'}} \pi^{j-r} \int_0^{\pi R^2} L_r^{-r}(t)L_{r'}^{-r'}(t)e^{-t}dt.
\]

For \(j = r\) and \(j' = r'\) we obtain

\[
\int_{D_R} H_{r,r}(z,\overline{z})H_{r,r}(z,\overline{z})e^{-\pi|z|^2}dz = \int_0^{\pi R^2} L_r^0(t)L_{r'}^0(t)e^{-t}dt
\]

and, for \(j = j'\),

\[
\int_{D_R} H_{j,r}(z,\overline{z})H_{j,r}(z,\overline{z})e^{-\pi|z|^2}dz = C_{j,r,r}(R) \cdot \delta_{r-r}.
\]

**Proof:** To avoid dividing the proof in two cases, we will use the identity

\[
\frac{(-x)^j}{j!}L_r^{-r}(x) = \frac{(-x)^r}{r!}L_j^{-j}(x),
\]

to write (2.3) as

\[
H_{j,r}(z,\overline{z}) = \sqrt{\frac{r!}{j!}} \pi^{j-r} z^{j-r} L_r^{-r}(\pi|z|^2), \quad j, r \in \mathbb{N}_0.
\]
Thus,
\[
\int_{D_R} H_{j,r}(z, \bar{z}) H_{j',r'}(z, \bar{z}) e^{-\pi |z|^2} \, dz
= \int_{D_R} \sqrt{\frac{r!}{j! j'!}} i^{r+j-r'} |z|^{j+r-j'} L_r^{j-r'} \left( \pi |z|^2 \right) L_{r'}^{j'-r'} \left( \pi |z|^2 \right) e^{-\pi |z|^2} \, dz.
\]
Setting \( z = \rho e^{i\theta} \) we obtain
\[
\int_{D_R} H_{j,r}(z, \bar{z}) H_{j',r'}(z, \bar{z}) e^{-\pi |z|^2} \, dz
= \sqrt{\frac{r!}{j! j'!}} i^{r+j-r'+j+1} \int_0^R \int_0^{2\pi} \rho^{r-j-r'+j+1} e^{i\theta (j-j'-r+r')} L_r^{j-r'} \left( \pi \rho^2 \right) L_{r'}^{j'-r'} \left( \pi \rho^2 \right) e^{-\pi \rho^2} \, d\rho \, d\theta
= \delta_{j-r+r'} \sqrt{\frac{r!}{j! j'!}} \pi^{j-r} \int_0^R 2\pi \rho L_r^{j-r} \left( \pi \rho^2 \right) L_{r'}^{j'-r'} \left( \pi \rho^2 \right) e^{-\pi \rho^2} \, d\rho.
\]
Proposition \(^7\) now yields the local reproducing formula for Hermite windows.

**Theorem 1.** For every \( R > 0 \) and every \( r, j \in \mathbb{N}_0 \) one has

\[
V_{h_r} f(z) = C_{j,r}(R)^{-1} \int_{z+D_R} V_{h_r} f(w) K_{h_j}(z, w) \, dw,
\]
with
\[
C_{j,r}(R) := \langle \chi_{D_R} \cdot V_{h_r} h_r, V_{h_j} h_j \rangle = \int_0^{\pi R} L_r^0(t) L_j^0(t) e^{-t} \, dt.
\]

**Proof:** Rewriting Proposition \(^7\) using
\[
V_{h_r} h_j(x, -\xi) = e^{i\pi x \xi - \frac{\pi}{2} |\xi|^2} H_{j,r}(z, \bar{z}),
\]
leads to
\[
\int_{D_R} V_{h_r} h_j(x, -\xi) \bar{V}_{h_r} h_j(x, -\xi) \, dz = C_{j,r,j',r'}(R) \cdot \delta_{j-r-j'+r}.
\]
In the case \( j = r \) we obtain by a change of variables
\[
\int_{D_R} V_{h_r} h_j(z) \bar{V}_{h_r} h_r(z) \, dz = \int_{D_R} H_{j,j}(z, \bar{z}) H_{j,r,r}(z, \bar{z}) e^{-\pi |z|^2} \, dz = \delta_{j,r} \cdot C_{j,j}(R),
\]
which implies that the following holds weakly
\[
\int_{D_R} V_{h_r} h_j(w) \pi(w) h_r \, dw = C_{j,r}(R) \cdot h_r.
\]
Using (2.2) and (2.3) it thus follows

\[
V_h f(z) = \langle f, \pi(z) h_r \rangle = C_{j,r}(R)^{-1} \int_{D_R} \langle f, \pi(z) \pi(w) h_r \rangle \langle \pi(w) h_j, h_j \rangle dw
\]

\[
= C_{j,r}(R)^{-1} \int_{D_R} e^{2\pi i x \eta} \langle f, \pi(z + w) h_r \rangle \langle \pi(w) h_j, h_j \rangle dw
\]

\[
= C_{j,r}(R)^{-1} \int_{z+D_R} e^{2\pi i (\eta - \xi)} \langle f, \pi(w) h_r \rangle \langle \pi(w - z) h_j, h_j \rangle dw
\]

\[
= C_{j,r}(R)^{-1} \int_{z+D_R} \langle f, \pi(w) h_r \rangle \langle \pi(w) h_j, \pi(z) h_j \rangle dw.
\]

□

Another consequence of Proposition 7 is the following local inversion formula, which allows to reconstruct \( f \) from the values of the STFT on arbitrary discs:

**Theorem 2.** For every \( R > 0, r \in \mathbb{N}_0 \) and \( z \in \mathbb{R}^2 \) we have

(4.6) \( f = \sum_{j \in \mathbb{N}_0} \left( C_{j,r}(R)^{-1} \int_{z+D_R} V_h f(w) \langle \pi(w) h_r, \pi(z) h_j \rangle dw \right) \pi(z) h_j, \quad \forall f \in L^2(\mathbb{R}). \)

**Proof:** Write \( f \) with respect to the orthonormal basis \( \{h_j\}_{j \in \mathbb{N}_0} \)

\[
f = \sum_{j \in \mathbb{N}_0} a_j h_j.
\]

By linearity of the STFT, one has

\[
V_h f(w) = \sum_{j \in \mathbb{N}_0} a_j \langle h_j, \pi(w) h_r \rangle.
\]

Now, Proposition 7 gives

\[
\int_{D_R} \langle f, \pi(w) h_r \rangle \langle \pi(w) h_r, h_k \rangle dw = \sum_{j \in \mathbb{N}_0} a_j \int_{D_R} \langle h_j, \pi(w) h_r \rangle \langle \pi(w) h_r, h_k \rangle dw
\]

\[
= \sum_{j \in \mathbb{N}_0} a_j \int_{D_R} H_{j,r}(w, \overline{w}) \overline{H_{k,r}(w, \overline{w})} e^{-\pi |w|^2} dw
\]

\[
= a_k C_{k,r}(R).
\]

Thus,

\[
f = \sum_{j \in \mathbb{N}_0} a_j h_j = \sum_{j \in \mathbb{N}_0} C_{j,r}(R)^{-1} h_j \int_{D_R} \langle f, \pi(w) h_r \rangle \langle \pi(w) h_r, h_j \rangle dw.
\]
Applying this equality to \( \pi(z)^* f \) and using the same argument as in the proof of Theorem 1 yields

\[
\pi(z)^* f = \sum_{j \in \mathbb{N}_0} C_{j,r}(R)^{-1} h_j \int_{D_R} \langle \pi(z)^* f, \pi(w)h_r \rangle \langle \pi(w)h_r, h_j \rangle dw
\]

\[
= \sum_{j \in \mathbb{N}_0} C_{j,r}(R)^{-1} h_j \int_{z+D_R} \langle f, \pi(w)h_r \rangle \langle \pi(w)h_r, \pi(z)h_j \rangle dw.
\]

Now apply \( \pi(z) \) on both sides to conclude the proof.

Remark 3. If \( j = r \), then \( C_r(R) = 1 - e^{-\pi R^2} P_r(\pi R^2) \), where \( P_r \) is a polynomial of degree \( 2r \), \( P_r(0) = 1 \) and \( P_0 \equiv 1 \). See the appendix of [45] for detailed calculations.

Remark 4. The so-called true (or pure, according to [42, 43]) polyanalytic Fock space \( F_j(\mathbb{C}) \), which can be defined as the span of \( \{ H_{j,r}(z, z) \}_{r \in \mathbb{N}} \) in \( L_2(\mathbb{C}, e^{-\pi|z|^2}) \) (see Sections 6.3 and 6.4 for more details) or, equivalently, as the subspace of \( L_2(\mathbb{C}) \) whose elements satisfy the reproducing formula

\[
F(z) = \int_{\mathbb{C}} F(w)L_0^0(\pi |z - w|^2)e^{\pi z \overline{w}}e^{-\pi |w|^2} dw,
\]

From Theorem 1 it follows that (4.3) implies the following local reproducing formula for \( F \in F_j(\mathbb{C}) \):

\[
F(z) = C_{j,r}(R)^{-1} \int_{z+D_R} F(w)L_0^0(\pi |z - w|^2)e^{\pi z \overline{w}}e^{-\pi |w|^2} dw.
\]

For \( j = r \) this is what one would expect as a local version of (4.7) and as an extension of the following local reproducing formula for functions in the analytic Fock space \( F(\mathbb{C}) = F_0(\mathbb{C}) \), obtained by Seip in [59] (this is also implicit in [21]):

\[
F(z) = (1 - e^{-\pi R^2})^{-1} \int_{z+D_R} F(w)e^{\pi z \overline{w}}e^{-\pi |w|^2} dw.
\]

However, for \( j \neq r \) the spaces \( F_j(\mathbb{C}) \) and \( F_r(\mathbb{C}) \) are orthogonal and one could hardly expect (4.8) to be true since, for every \( F \in F_j(\mathbb{C}) \),

\[
\int_{\mathbb{C}} F(w)L_0^0(\pi |z - w|^2)e^{\pi z \overline{w}}e^{-\pi |w|^2} dw = 0, \quad \forall z \in \mathbb{C}.
\]

Consequently, for \( j \neq r \), the formula (4.8) only holds for finite \( R \).

Remark 5. It is not clear to us whether there exist other window functions that allow for double orthogonality in sequences of concentric domains other than the disc.
4.3. **Estimates with explicit constants.** We are now ready to formulate our main localization result for the Hermite functions. Recall the maximum Nyquist density

\[ \rho(\Delta, R) := \sup_{z \in \mathbb{R}^2} |\Delta \cap (z + D_R)|. \]

We will also make use of the following notion of density

\[ A_r(\Delta, R) := \sup_{z \in \mathbb{R}^2} \int_{\Delta \cap z + D_R} |L^0_\pi(||z - w||^2)|e^{-\pi||z - w||^2/2}dw. \]

**Theorem 3.** Let \( \Delta \subset \mathbb{R}^2 \) and \( f \in M^p, \ 1 \leq p < \infty \). For every \( 0 < R < \infty \), it holds

\[ \|V_{\mu} f \chi_{\Delta}\|_p \leq \frac{A_r(\Delta, R)}{C_r(R)} \leq \frac{\rho(\Delta, R)}{C_r(R)}. \]

**Proof:** In Proposition take \( K := K_{hr} \cdot \Omega_R \), where \( \Omega_R(z, w) := \chi_{D_R}(z - w) \). Then \( \theta(K) = 1/C_r(R) \). Thus, if \( d\mu(z) = \chi_\Delta dz \), we have

\[ \|V_{\mu} f \chi_{\Delta}\|_1 \leq \frac{1}{C_r(R)} \|K_{hr} \cdot \Omega_R\|_{A_{\chi_\Delta} dz}. \]

Using the explicit formula (2.7),

\[ \|K_{hr} \cdot \Omega_R\|_{A_{\chi_\Delta} dz} = \sup_{z \in \mathbb{R}^2} \int_{\Delta} |K_{hr}(z, w)|\chi_{D_R}(z - w)dw \]

\[ = \sup_{z \in \mathbb{R}^2} \int_{\Delta \cap z + D_R} |L^0_\pi(||z - w||^2)|e^{-\pi||z - w||^2/2}dw \]

\[ = A_r(\Delta, R) \leq \rho(\Delta, R). \]

Hence, the result holds for \( p = 1 \). As \( M^1 \cap M^\infty = M^1 \) is dense in \( M^p \) and

\[ \sup_{\|f\|_{M^\infty} = 1} \|V_{\mu} f \chi_{\Delta}\| = 1, \]

the result for \( 1 < p < \infty \) follows from Corollary

**Remark 6.** Results in the spirit of Theorem 3 can be found for example in [31, Section 4] or [12]. The estimates there are however only given for sets with particular geometry, e.g. sets that are thin at infinity or have finite Lebesgue measure, or without explicit constants.

An immediate consequence of Theorem 3 is the following refined (local) \( L^p \)-uncertainty principle for the short-time Fourier transform (see [37, Proposition 3.3.1] and [18, 39, 57] for other uncertainty principles for the STFT).

**Corollary 4.** Suppose that \( f \in M^p, \ 1 \leq p < \infty, \) satisfies \( \|V_{\mu} f\| = 1 \) and that \( \Delta \subset \mathbb{R}^2 \) and \( \varepsilon \geq 0 \) are such that

\[ 1 - \varepsilon \leq \int_{\Delta} |V_{\mu} f(z)|^p dz. \]
Then

\[ 1 - \varepsilon \leq \inf_{R > 0} \left( \frac{\rho(\Delta, R)}{C_r(R)} \right) \leq |\Delta|. \]

**Remark 7.** Essentially, Corollary 4 statements that the short-time Fourier transform of a function in $M^p$ (using a Hermite window) cannot be well concentrated on sets that are locally small over the entire time-frequency plane. For an explicit example set $R = 1$ and $r = 0$. Then $C_0(R) = 1 - e^{-\pi} \approx 0.96$, which implies that

\[ \rho(\Delta, 1) \geq 0.95(1 - \varepsilon). \]

Moreover, if we choose $\varepsilon = 0.01$, there exists a subset of $\Delta$ contained in a disc of radius one, covering at least approximately $3/10$ of the area of that disc.

**Remark 8.** If $K_g$ in (4.1) shows sufficient off diagonal decay, then the bound (4.1) behaves in a similar way as the local integral $A_r(\Delta, R)$ for the Hermite functions.

5. Optimality and sparse sets

Donoho and Logan [23, Chapter 5] discussed optimality of the constant in (1.4) as well as their $L^2$-estimate. As it turns out, using extremal functions like the Beurling-Selberg function [17] gives optimal constants within their method. In the STFT setup considered in this paper, as far our knowledge goes, there is no theory of extremal functions available. In the case of Gaussian window, we believe that our local reproducing kernel is at least optimal among all kernels obtained from truncating functions in $V_\varphi$ on $D_R$ as $V_\varphi \varphi$ optimizes the concentration problem on the disc for any $p \geq 1$.

Large sieve inequalities are particularly powerful if the localization domain is sparse. It is nevertheless interesting to test the estimates in cases where the solution of the localization problem is known. Note that neither Donoho Logan’s result applied to localization on an interval nor Theorem 3 applied to a disc achieve the actual solution. But this is to be expected as the estimates hold for general sets.

With a view to comparing the estimated and actual values in cases where the exact solution is known, consider $\Delta = D_R$ and the Gaussian window $g = \varphi = h_0$. In this case it is well known that the Gaussian maximizes the concentration of the short-time Fourier transform in $D_R$ [21, 59]. The $p$-norm can be explicitly evaluated as follows:

\[
\int_{D_R} |V_\varphi \varphi(z)|^p dz = \int_{D_R} e^{-\pi \rho |z|^2/2} dz = 2\pi \int_0^R \rho e^{-\pi \rho^2/2} d\rho = \frac{2}{p} (1 - e^{-\pi p R^2/2}).
\]

Therefore, $\|V_\varphi \varphi\|_p^p = \frac{2}{p}$ and the optimal solution of the concentration problem on $D_\rho$ is given by

\[
(5.1) \quad \sup_{f \in M^p} \frac{\int_{D_R} |V_\varphi f(z)|^p dz}{\|V_\varphi f\|_p} = \frac{\int_{D_R} |V_\varphi \varphi(z)|^p dz}{\|V_\varphi \varphi\|_p} = (1 - e^{-\pi p R^2/2}).
\]
Moreover,

\[
A_0(D_{\rho}, R) = \sup_{z \in \mathbb{R}^2} \int_{D_{\rho} \cap z + D_R} e^{-\pi|z-w|^2/2} dw = \int_{D_{\tau}} e^{-\pi|w|^2/2} dw = 2(1 - e^{-\pi T^2/2}),
\]

with \( T = \min\{\rho, R\} \) and \( C_0(R) = 1 - e^{-\pi R^2} \). Using Theorem 3 we obtain the concentration estimate

\[
\inf_{R > 0} A_0(D_{\rho}, R) = \inf_{R > 0} \frac{2(1 - e^{-\pi T^2/2})}{1 - e^{-\pi R^2/2}} = 2(1 - e^{-\pi T^2/2}).
\]

Comparing our general estimate with the actual optimal value from (5.1) we observe that the estimate is not optimal for any \( p \geq 1 \). Let for example \( p = 1 \). Then

\[
\frac{2(1 - e^{-\pi T^2/2})}{1 - e^{-\pi T^2/2}} = 2.
\]

For \( p \in [1, \infty] \) write \( L^p(\mathbb{C}) \) to denote the Banach space of all measurable functions equipped with the norm

\[
\|F\|_{L^p(\mathbb{C})} = \left( \int_{\mathbb{C}} |F(z)|^p e^{-\pi p|z|^2/2} dz \right)^{1/p}.
\]

Now we turn our focus to the asymptotics of the concentration problem. Define the distance of two sets in a standard way via

\[
\text{dist}(A, B) := \inf\{|x - y| : x \in A, y \in B\}.
\]

Let us consider the case where \( \Delta \) is given by a finite union of sets \( \Delta_k \) with increasing separation

\[
d := \min_{k \neq \ell} \text{dist}(\Delta_k, \Delta_\ell).
\]

It is easy to see that \( A_r(\Delta, R) \to \max_k A_r(\Delta_k, R) \) as \( d \to \infty \). As we will show below, the concentration problem is accurately described by this observation: it is decoupled. A related result for the case of one dimensional band-limited functions was derived in [24, Theorem 10].

**Proposition 8.** Let \( g \in M^1 \) and \( \Delta \) be the union of \( N \) disjoint, compact sets \( \Delta_1, \ldots, \Delta_N \). If \( d \) tends to infinity, then

\[
\sup_{f \in L^2(\mathbb{R})} \frac{\|V_g f \cdot \chi_\Delta\|_2}{\|V_g f\|_2} \to \max_{k=1,\ldots,N} \sup_{f \in L^2(\mathbb{R})} \frac{\|V_g f \cdot \chi_{\Delta_k}\|_2}{\|V_g f\|_2}.
\]

**Proof:** First, it trivially holds that

\[
\max_{k=1,\ldots,N} \sup_{f \in L^2(\mathbb{R})} \|V_g f \cdot \chi_{\Delta_k}\|_2 \leq \sup_{f \in L^2(\mathbb{R})} \|V_g f \cdot \chi_\Delta\|_2.
\]

We can restrict ourselves to the case \( \Delta = \Delta_1 \cup \Delta_2 \). The general result then follows by induction. Now, for simplicity assume that \( \|g\|_2 = 1 \) and define \( f_k, k = 1, 2 \), via

\[
f_k := V_g^*(V_g f \cdot \chi_{\Delta_k}),
\]
where \( U_d(\Delta) := \{ z \in \mathbb{R}^2 : \text{dist}(z, \Delta) \leq d \} \). Let \( \varepsilon > 0 \) and choose \( d = d(\varepsilon) \) large enough such that for all \( f \in L^2(\mathbb{R}) \) and \( k \in \{1, 2\} \)

\[
\| (V_g f - V_g f_k) \cdot \chi_{\Delta_k} \|_2 \leq \varepsilon \|V_g f\|_2, \tag{5.3}
\]

and

\[
\| V_g f_k \cdot \chi_{U_{d/2}(\Delta_k)^c} \|_2 \leq \varepsilon \|V_g f\|_2. \tag{5.4}
\]

It is indeed possible to choose \( d \) accordingly, since

\[
\| (V_g f - V_g f_k) \cdot \chi_{\Delta_k} \|_2^2 = \int_{\Delta_k} \left| V_g f(z) - \int_{U_{d/3}(\Delta_k)} V_g f(w) K_g(z, w) dw \right|^2 dz
\]

\[
= \int_{\Delta_k} \left| \int_{\mathbb{R}^2 \setminus U_{d/3}(\Delta_k)} V_g f(w) K_g(z, w) dw \right|^2 dz
\]

\[
\leq \int_{\Delta_k} \int_{\mathbb{R}^2 \setminus U_{d/3}(\Delta_k)} |K_g(z, w)|^2 dw dz \|V_g f\|_2^2
\]

\[
\leq |\Delta_k| \sup_{z \in \Delta_k} \int_{\mathbb{R}^2 \setminus U_{d/3}(\Delta_k)} |\langle g, \pi(z - w) \rangle|^2 dw \|V_g f\|_2^2
\]

\[
\leq |\Delta_k| \int_{\mathbb{R}^2 \setminus D_{d/3}} |\langle g, \pi(w) \rangle| dw \|V_g f\|_2^2
\]

\[
= C(k, d) \|V_g f\|_2^2,
\]

where the last inequality follows from \( |z - w| \geq d/3 \), for \( z \in \Delta_k \) and \( w \in \mathbb{R}^2 \setminus U_{d/3} \). From the STFT being an isometry it now follows that \( C(k, d) \to 0 \) as \( d \to \infty \). To show that (5.4) is satisfied if \( d \) is chosen big enough, observe at first that

\[
\sup_{z \in U_{d/2}(\Delta_k)^c} |V_g f_k(z)| \leq \sup_{z \in U_{d/2}(\Delta_k)^c} \int_{U_{d/3}(\Delta_k)} |V_g f(w) K_g(z, w)| dw \leq \|V_g f\|_1 \|V_g g\|_1.
\]

The \( L^1 \)-norm on the other hand can be estimated as

\[
\|V_g f_k \cdot \chi_{U_{d/2}(\Delta_k)^c} \|_1 \leq \int_{U_{d/2}(\Delta_k)^c} \int_{U_{d/3}(\Delta_k)} |V_g f(w) K_g(z, w)| dw dz
\]

\[
\leq \sup_{w \in U_{d/3}(\Delta_k)} \int_{U_{d/2}(\Delta_k)^c} |K_g(z, w)| dz \|V_g f\|_1
\]

\[
\leq \int_{\mathbb{R}^2 \setminus D_{d/3}} |\langle g, \pi(z) \rangle| dz \|V_g f\|_1
\]

\[
= \tilde{C}(k, d) \|V_g f\|_1.
\]

Hence, (5.4) follows by interpolation. As \( \|V_g f_k\|_2 \leq \|V_g f\|_2 \) we deduce from (5.3) that

\[
\|V_g f \cdot \chi_{\Delta}\|_2^2 = \|V_g f \cdot \chi_{\Delta_1}\|_2^2 + \|V_g f \cdot \chi_{\Delta_2}\|_2^2
\]

\[
\leq \|V_g f_1 \cdot \chi_{\Delta_1}\|_2^2 + \|V_g f_2 \cdot \chi_{\Delta_2}\|_2^2 + C\varepsilon \|V_g f\|_2^2.
\]
Moreover, by (5.4), we have the following almost orthogonality relation for \( V_gf_1 \) and \( V_gf_2 \):

\[
\|V_gf_1 + V_gf_2\|^2 \geq \|V_gf_1\|^2 + \|V_gf_2\|^2 - 2\langle V_gf_1, V_gf_2 \rangle
\]

\[
\geq \|V_gf_1\|^2 + \|V_gf_2\|^2 - 2\langle V_gf_1 \cdot \chi_{U_{d/2}(\Delta_1)c^2}, V_gf_2 \rangle - 2\langle V_gf_1, V_gf_2 \cdot \chi_{U_{d/2}(\Delta_2)c^2} \rangle
\]

\[
\geq \|V_gf_1\|^2 + \|V_gf_2\|^2 - 4\varepsilon\|V_gf_1\|\|V_gf_2\|
\]

\[
\geq \|V_gf_1\|^2 + \|V_gf_2\|^2 - 4\varepsilon\|V_gf\|^2.
\]

Now, as \( \|V_gf_1 + V_gf_2\|^2 \leq \|V_gf \cdot \chi_{\Delta_1 \cup \Delta_2}\|^2 \leq \|V_gf\|^2 \) it follows that

\[
\frac{\|V_gf \cdot \chi_{\Delta}\|^2}{\|V_gf\|^2} \leq (1 + C\varepsilon) \frac{\|V_gf_1 \cdot \chi_{\Delta_1}\|^2 + \|V_gf_2 \cdot \chi_{\Delta_2}\|^2}{\|V_gf_1\|^2 + \|V_gf_2\|^2} + C\varepsilon
\]

\[
\leq (1 + C\varepsilon) \max_{k=1,2} \frac{\|V_gf_k \cdot \chi_{\Delta_k}\|^2}{\|V_gf_k\|^2} + C\varepsilon
\]

\[
\leq (1 + C\varepsilon) \max_{k=1,2} \sup_{f \in L^2(\mathbb{R})} \frac{\|V_gf \cdot \chi_{\Delta_k}\|^2}{\|V_gf\|^2} + C\varepsilon,
\]

which concludes the proof if we take the supremum over \( L^2(\mathbb{R}) \) on the left hand side. \( \square \)

**Remark 9.** Although we expect a similar result to hold for the concentration problem in \( M^p \) we were not able to prove it. The main problem is that our argument relies on \( \|V_gV_g^*\|_{2 \to 2} = 1 \) which is not true on \( M^p, p \neq 2 \).

Finally, we present a conjecture on an extremal problem of localization of the STFT with Gaussian window which is the joint time-frequency analogue of [24, Conjecture 1].

**Conjecture 1.** Let \( \Delta \subset \mathbb{R}^2 \) be a set of finite measure and \( \varphi = h_0 \) be the Gaussian. Then

\[
\sup_{\|\Delta\| = A} \sup_{f \in M^p} \frac{\|V_{\varphi}f \cdot \chi_{\Delta}\|_p^p}{\|V_{\varphi}f\|_p^p}
\]

is attained if and only if \( \Delta = z + D \sqrt{A} / \pi \) for some \( z \in \mathbb{R}^2 \), up to perturbations of Lebesgue measure zero.

The next Proposition provides extra support for the conjecture, by showing that the disc is the unique solution (up to perturbations of Lebesgue measure zero) of a certain extremal problem. This will in turn imply that the disc maximizes \( A_0(\Delta, R) \) for all \( R > 0 \), where the area of \( \Delta \) is fixed. Consequently, Conjecture 1 is backed by the estimates of Theorem 3.

**Proposition 9.** Let \( \alpha > 0 \). The disc \( D_R, R = \sqrt{A / \pi} \) is the unique (up to perturbations of Lebesgue measure zero) minimizer of the following extremal problem:

\[
\sup_{\Omega \subset \mathbb{R}^n} \int_{\Omega} e^{-\alpha |z|^2} dz, \quad \text{subject to} \ |\Omega| = A.
\]
Proof: Let us assume to the contrary that there exists \( \Omega \subset \mathbb{R}^n \), such that \( |\Omega| = C \) and \( |\Omega \setminus D_R| > 0 \) which maximizes (5.5). Define \( \Omega_e := \Omega \setminus D_R \). Then there exists \( \varepsilon = \varepsilon(R, |\Omega \setminus D_R|) > 0 \) such that \( |\Omega_{R+\varepsilon}| \geq \frac{|\Omega_R|}{2} > 0 \). Let \( I \subset D_R \setminus \Omega \) be any set that satisfies \( |I| = |\Omega_{R+\varepsilon}| \). (Such a set exists as \( \Omega \) contains a set of size \( |\Omega_{R+\varepsilon}| \) outside the disc \( D_R \) and has the same size as the disc.) Define another set \( \Omega^* := \Omega \setminus \Omega_{R+\varepsilon} \cup I \). It then holds that

\[
\int_{\Omega^*} e^{-\alpha|z|^2} dz = \int_{\Omega \setminus \Omega_{R+\varepsilon}} e^{-\alpha|z|^2} dz + \int_{I} e^{-\alpha|z|^2} dz \\
\geq \int_{\Omega \setminus \Omega_{R+\varepsilon}} e^{-\alpha|z|^2} dz + e^{-\alpha R^2} |I| \\
\geq \int_{\Omega \setminus \Omega_{R+\varepsilon}} e^{-\alpha|z|^2} dz + e^{-\alpha(R+\varepsilon)^2} |\Omega_{R+\varepsilon}| \\
\geq \int_{\Omega \setminus \Omega_{R+\varepsilon}} e^{-\alpha|z|^2} dz + \int_{\Omega_{R+\varepsilon}} e^{-\alpha|z|^2} dz \\
= \int_{\Omega} e^{-\alpha|z|^2} dz,
\]

which contradicts the assumption that \( \Omega \) maximizes (5.5).

\[\square\]

5.1. Recovery of STFT measurements. Now we will rephrase Proposition 2 and 3 in the context of reconstructing STFT-data using Theorem 3. Let \( B_1 = V_{h_r} = V_{h_r}(M^1) \subset L^1(\mathbb{R}^2) \). By the correspondence principle (2.9) we can replace minimization on \( B_1 \) by minimization on \( M^1 \) (which is independent of the particular choice of the order of the Hermite window).

Corollary 5. Suppose that \( G = V_{h_r} f + N \) is observed, where \( f \in M^1 \), \( N \in L^1(\mathbb{R}^2) \) and that the unknown support \( \Delta \) of \( N \) satisfies

\[
(5.6) \quad A_r(\Delta, R) < \frac{C_r(R)}{2},
\]

for some \( R > 0 \). Then \( \delta(\Delta) < \frac{1}{2} \) and the solution of the minimization problem

\[
\beta(G) = \arg \min_{g \in M^1} \|G - V_{h_r} g\|_1
\]

is unique and recovers the signal \( f \) perfectly \((\beta(G) = f)\).

Corollary 6. Let \( f \in M^1 \) and suppose that one observes \( H = P_{\Delta^c}(V_{h_r} f + N) \), where \( \|N\|_1 \leq \varepsilon \) and that the domain \( \Delta^c \) of missing data satisfies

\[
(5.7) \quad A_r(\Delta^c, R) < C_r(R),
\]

for some \( R > 0 \). Then any solution of

\[
\sigma(H) = \arg \min_{g \in M^1} \|P_{\Delta^c}(H - V_{h_r} g)\|_1
\]
satisfies
\[ \| V_{h_r}(f - \sigma(H)) \|_1 \leq \frac{2\varepsilon \cdot C_r(R)}{C_r(R) - A_r(\Delta, R)}. \]

For other approaches to the recovery of sparse time-frequency representations which concentrate on the set-up of finite sparse time-frequency representations, see [55, 56].

6. Extensions to other settings

6.1. Discrete Gabor systems. In this section, we will apply Corollary 1 to discrete Gabor systems. Let \( \Lambda \subset \mathbb{R}^2 \) be discrete and \( \Delta \subset \Lambda \). We define the discrete maximum Nyquist density by
\[ \rho^d(\Delta, R) := \sup_{z \in \mathbb{R}^2} \# \{ \Delta \cap z + D_R \} \]
and consider also
\[ A^d_r(\Delta, R) := \sup_{z \in \mathbb{R}^2} \sum_{\lambda \in \Delta \cap z + D_R} L_r^0(\pi \| z - \lambda \|^2) e^{-\pi |z - \lambda|^2/2}. \]

We define the measure \( \mu \) to be
\[ d\mu(\lambda) := \delta_{\Lambda}(\lambda) \cdot \chi_{\Delta}(\lambda) d\lambda. \]
First, observe that
\[ \sup_{f \in M^\infty} \| V_{h_r} f \|_\Delta = 1. \]
Applying Corollary 1 yields
\[ \sum_{\lambda \in \Delta} |V_{h_r} f(\lambda)|^p \| V_{h_r} f \|_p^p \leq \rho^d(\Delta, R) \leq A^d_r(\Delta, R) \cdot C_r(\Lambda). \]
If we also assume that \( \{ \pi(\lambda) h_r \}_{\lambda \in \Lambda} \) is a frame for \( L^2(\mathbb{R}) \), then, by Remark 1 there exists a lower Banach frame bound \( m = m(r, \Lambda, p) \) such that
\[ m\| f \|_{M^p}^p = m\| V_{h_r} f \|_p^p \leq \sum_{\lambda \in \Lambda} |V_{h_r} f(\lambda)|^p. \]
It then holds
\[ (6.1) \sum_{\lambda \in \Delta} \frac{|V_{h_r} f(\lambda)|^p}{\| V_{h_r} f \|_p^p} \leq \frac{A^d_r(\Delta, R)}{C_r(\Lambda)} \leq \frac{\rho^d(\Delta, R)}{m \cdot C_r(\Lambda)}. \]
If we would like to use this estimate for discrete signal recovery, then the bound \( \frac{A^d_r(\Delta, R)}{m \cdot C_r(\Lambda)} \) should be small, or at least less than one half. If \( \Lambda \) is a lattice in \( \mathbb{R}^2 \) which does not deviate too much from the square lattice then \( m \) scales with the density of \( \Lambda \). Since also \( \rho^d(\Delta, R) \) and \( A^d_r(\Delta, R) \) show similar behavior, it is still possible to get small concentration bounds if the density of \( \Lambda \) is increased. An interesting direction for further research could also be to study the concentration problem for complete Gabor systems that are not frames (this is the case of several lattice configurations for Gabor systems with Hermite functions, which are known to be complete [40] but not frames [48]), but still allow for reconstruction using dual systems [62].
6.2. Vector-valued STFT transforms. Vector-valued time-frequency analysis \cite{44} is motivated by the problem of multiplexing of signals, where one wants to transmit several signals over a single channel followed by separating and recovering the signals at the receiver \cite{14}. A classical way to do this is to store the information for every function in mutually orthogonal subspaces. The orthogonality relation \eqref{eq:2.4} for the short-time Fourier transform suggested a vector-valued version of the STFT using mutually orthogonal windows, called the super Gabor transform \cite{1}, in a reference to the connection with \cite{34,38}. In the case of a vector constituted by Hermite functions, this reads

\[
V_h f(z) = V_{(h_0, \ldots, h_n)}(f_0, \ldots, f_n)(z) := \sum_{k=0}^n V_{h_k} f_k(z),
\]

where \( f := (f_0, f_1, \ldots, f_n) \in L^2(\mathbb{R})^{n+1} \) and \( h_n := (h_0, h_1, \ldots, h_n) \) is the vector of the first \( n+1 \) Hermite functions. This is the continuous transform associated to Gabor superframes with Hermite windows \cite{34,38} and to sampling in polyanalytic Fock spaces \cite{2}. The function \( f_k \) can then be reconstructed by

\[
f_k = V_{h_k}^* V_h f.
\]

The range of this transform is a Hilbert space with reproducing kernel given by

\[
K_{h_n}(z, w) = e^{i\pi(x+y)(\omega-\eta)} L_n^1(\pi|z-w|^2) e^{-\pi|z-w|^2/2},
\]

(this follows from \eqref{eq:4.7} and the summation relation \( \sum_{k=0}^n L_k^0 = L_{n+1}^0 \) of the Laguerre functions). For the basis functions of the reproducing kernel \( K_{h_n} \), double orthogonality is lost, since the cross terms are not zero (the Laguerre functions are not orthogonal on any interval \([0, R]\), for \( R < \infty \)). We can however still define a local kernel that yields an estimate in terms of \( A_0(\Delta, R) \).

Set \( \mathcal{V}_g^p := \{ V_g f : f \in M^p \} \) then the orthogonal decomposition extends to the modulation spaces \( M^p \), see \cite{7}. For \( 1 \leq p < \infty \), we have

\[
V_h \left( \prod_{k=0}^n M^p \right) = \mathcal{V}_{h_0}^p \oplus \mathcal{V}_{h_1}^p \oplus \cdots \oplus \mathcal{V}_{h_n}^p.
\]

Therefore,

\[
\| V_h f \|_p = \left\| \sum_{k=0}^n V_{h_k} f_k \right\|_p \asymp \sum_{k=0}^n \| V_{h_k} f_k \|_p.
\]

It follows from Theorem \cite{11} that

\[
V_h f(z) = \sum_{k=0}^n V_{h_k} f_k(z) = \sum_{k=0}^n C_{k,0}(R)^{-1} \int_{z+D_R} V_{h_k} f_k(w) K_h(z, w) dw,
\]

where

\[
V_p g := \{ V_p f : f \in M^p \}
\]

and

\[
\mathcal{V}_g^p := \{ V_g f : f \in M^p \}.
\]
which yields

\[
\|V_{h_n} f \cdot \chi_\Delta\|_1 \leq \max_{0 \leq m \leq n} C_{m,0}(R)^{-1} \int \sum_{k=0}^{n} \int_{\Delta} |V_{h_k} f_k(w) K_{h_0}(z, w)|dwdz \\
\leq \max_{0 \leq m \leq n} C_{m,0}(R)^{-1} \cdot A_0(\Delta, R) \cdot \sum_{k=0}^{n} \|V_{h_k} f_k\|_1 \\
\leq \tilde{C} \cdot \max_{0 \leq m \leq n} C_{m,0}(R)^{-1} \cdot A_0(\Delta, R) \cdot \left\| \sum_{k=0}^{n} V_{h_k} f_k \right\|_1 \\
= \tilde{C} \cdot \max_{0 \leq m \leq n} C_{m,0}(R)^{-1} \cdot A_0(\Delta, R) \cdot \|V_{h_n} f\|_1.
\]

Hence, we have shown that the concentration operator of a multiplexed short-time Fourier transform can also be estimated in terms of \(A_0(\Delta, R)\) and \(\rho(\Delta, R)\) at the cost of a larger normalization constant and the additional factor \(\tilde{C}\).

6.3. True polyanalytic Fock spaces. The Bargmann transform \(B\), defined as

\[
Bf(z) = 2^{\frac{1}{4}} \int_{\mathbb{R}} f(t)e^{2\pi t z - \pi |z|^2 - \pi t^2} dt,
\]

is an isomorphism \(B : L^2(\mathbb{R}) \to F_2(\mathbb{C})\), where \(F_2(\mathbb{C})\) is the classical Bargmann-Fock space of entire functions. One can define a sequence of transforms \(B^{r+1} : L^2(\mathbb{R}) \to F_2^{r+1}(\mathbb{C})\) as a Hilbert space isomorphism mapping onto true polyanalytic Fock spaces \([2, 63]\) as follows:

\[
B^{r+1} f(z) = \left(\frac{\pi^r}{r!}\right)^{\frac{1}{2}} e^{\pi |z|^2} (\partial_z)^r \left[ e^{-\pi |z|^2} Bf(z) \right]
\]

The relation between Gabor transforms with Hermite functions and true polyanalytic Bargmann transforms of general order \(r\) reads \([2]\):

\[
e^{-i\pi x \xi + \frac{\pi |\xi|^2}{2}} V_{h_r} f(x, -\xi) = B^{r+1} f(z).
\]

The \(L^p\) version of the polyanalytic Bargmann-Fock spaces has been introduced in \([7]\), where the link to Gabor analysis has been particularly useful. For \(p \in [1, \infty]\) write \(L_p(\mathbb{C})\) to denote the Banach space of all measurable functions equipped with the norm

\[
\|F\|_{L_p(\mathbb{C})} = \left( \int_{\mathbb{C}} |F(z)|^p e^{-\pi |z|^2} \frac{1}{2} dz \right)^{1/p}.
\]

As a corollary of Theorem , we thus obtain the inequality

\[
\frac{\|F \cdot \chi_\Delta\|^p_{L_p(\mathbb{C})}}{\|F\|^p_{L_p(\mathbb{C})}} \leq \frac{\rho(\Delta, R)}{C'_{r}(R)}, \quad \forall F \in F^{r+1}.
\]
6.4. Polyanalytic Fock spaces. A function \( F(z, \bar{z}) \), defined on a subset of \( \mathbb{C} \), and satisfying the generalized Cauchy-Riemann equations

\[
(\partial_{\bar{z}})^n F(z, \bar{z}) = \frac{1}{2^n} \left( \frac{\partial}{\partial x} + i \frac{\partial}{\partial \xi} \right)^n F(x + i\xi, x - i\xi) = 0,
\]
is said to be polyanalytic of order \( n - 1 \).

**Definition 1.** We say that a function \( F \) belongs to the polyanalytic Fock space \( F_{n+1}^2(\mathbb{C}) \), if \( \|F\|_{L^2(\mathbb{C})} < \infty \) and \( F \) is polyanalytic of order \( n \).

Polyanalytic Fock spaces seem to have been first considered by Balk [15, pag. 170]. Vasilevski [63] obtained the following decompositions in terms of the spaces \( F_r^2(\mathbb{C}) \):

\[
F_n^2(\mathbb{C}) = F_1^2(\mathbb{C}) \oplus \ldots \oplus F_n^2(\mathbb{C})
\]

and

\[
L_2(\mathbb{C}) = \bigoplus_{n=1}^{\infty} F_n^2(\mathbb{C}).
\]

We can rewrite the transform of the previous section as a transform \( B^n : L^2(\mathbb{R}, \mathbb{C}^n) \to F^n(\mathbb{C}) \) mapping each vector \( f = (f_1, \ldots, f_n) \in L^2(\mathbb{R}, \mathbb{C}^n) \) to

\[
B^n f = e^{-i\pi x\xi + \pi |z|^2} V_{h_{n-1}} f(\lambda).
\]

Since the multiplier \( e^{-i\pi x\xi + \pi |z|^2} \) in (6.2) is the same for every \( n \), we have:

\[
B^n f = B^1 f_1 + \ldots + B^n f_n.
\]

This map is again a Hilbert space isomorphism and is called the polyanalytic Bargmann transform [2]. The identity

\[
V_{h_n} f(z) = \sum_{k=0}^{n} C_{k,0}(R)^{-1} \int_{z+D_R} V_{h_k} f_k(w) K_{h_0}(z, w) dw
\]

can be written as

\[
B^n f = \sum_{k=0}^{n} C_{k,0}(R)^{-1} \int_{z+D_R} B^k f_k(w) K_{h_0}(z, w) e^{-\pi |w|^2} dw.
\]

Rephrasing the discussion in the end of section 6.2, leads to the inequality

\[
\|B^n f \cdot \chi_\Delta\|_{L^1(\mathbb{C})} \leq \tilde{C} \cdot \max_{0 \leq m \leq n} C_{m,0}(R)^{-1} \cdot A_0(\Delta, R) \cdot \left\|B^n f\right\|_{L^1(\mathbb{C})}.
\]
7. Further questions

(1) If there exists a function $g \in L^2(\mathbb{R})$ that allows for a local reproducing formula on all discs of radius $R > 0$, i.e., if

$$V_g f(z) = C_g(R)^{-1} \int_{z+D_R} \langle f, \pi(w)g \rangle \langle \pi(w)g, \pi(z)g \rangle dw, \quad \forall f \in L^2(\mathbb{R}),$$

does it follow that $g$ is necessarily a Hermite function?

(2) This problem concerns a generalization of the main result in [6] using Hermite window instead of Gaussian window. If $\Omega$ is simply connected and $h_j$ is an eigenfunction of the following localization operator

$$H^r_{\Omega} f := \int_{\Omega} \langle f, \pi(w)h_r \rangle \pi(w)h_r dw,$$

does it follow that $\Omega$ is a disc centered at the origin?

(3) Is it possible to find a window $g$ such that double orthogonality holds in a sequence of non-circular domains $\Omega_1 \subset \Omega_2 \subset \ldots \Omega_\infty = \mathbb{R}^2$?

(4) Due to the orthogonality in concentric domains, the analysis in the case of Hermite windows avoided the use of the extremal functions required, for instance in [23]. However, if one aims to extend the results of [23] to the challenging setup of general de Branges spaces, used in the characterization of Fourier frames in [53], such a simplification is unlikely to occur. It is thus a natural question to ask if the results in [23] can be used for this purpose. A related setup where one can expect the aid of explicit formulas is the one of the band-limited multidimensional Fourier transform of radial functions [61], which essentially boils down to the band-limited Hankel transform, where the localization operators and the Nyquist rate have been studied in detail [3].

(5) Prove or disprove Conjecture 1

Acknowledgement

This work was funded by the Austrian Science Fund (FWF) START-project FLAME (‘Frames and Linear Operators for Acoustical Modeling and Parameter Estimation’; Y 551-N13) and FWF project ‘Operators and Time-Frequency Analysis’ P 31225-N32.

References

[1] L. D. Abreu, On the structure of Gabor and super Gabor spaces. Monatsh. Math., 161, 237-253 (2010).
[2] L. D. Abreu, Sampling and interpolation in Bargmann-Fock spaces of polyanalytic functions. Appl. Comp. Harm. Anal., 29, 287-302 (2010).
[3] L. D. Abreu, A. S. Bandeira, Landau’s necessary density conditions for the Hankel transform. J. Funct. Anal., 162, 1845-1866 (2012).
[4] L. D. Abreu, H. G. Feichtinger, *Function Spaces of Polyanalytic Functions*, Harmonic and Complex Analysis and its Applications. Birkhäuser, 1-38 (2014).

[5] L. D. Abreu, P. Balazs, M. de Gosson, Z. Mouayn, *Discrete coherent states for higher Landau levels*. Ann. of Phys., 363, 337-353 (2015).

[6] L. D. Abreu, M. Dörfler, *An inverse problem for localization operators*. Inverse Problems, 28 (11), (2012).

[7] L. D. Abreu, K. Gröchenig, *Banach Gabor frames with Hermite functions: polyanalytic spaces from the Heisenberg group*. Appl. Anal., 91, 1981-1997 (2012).

[8] L. D. Abreu, J. L. Romero, J. M. Pereira, S. Torquato, *The Weyl-Heisenberg ensembles: hyperuniformity and higher Landau levels*. J. Stat. Mech.: Theory and Experiment, 4, 043103, (2017).

[9] L. D. Abreu, K. Gröchenig, J. L. Romero, *Harmonic analysis in phase space and finite Weyl-Heisenberg ensembles*. ArXiv:170403042v1, (2017).

[10] L. D. Abreu, M. Speckbacher, A planar large sieve and sparsity of time-frequency representations. IEEE Proc. SampTA 2017, Tallinn, Estonia, 283-287 (2017).

[11] C. Aubel, H. Bölcskei, Vandermonde matrices with nodes in the unit disk and the large sieve, Appl. Comp. Harm. Anal., https://doi.org/10.1016/j.jacha.2017.07.006, (2018).

[12] G. Ascensi, *Sampling measures for the Gabor transform*. J. Approx. Theory, 200, 40-67 (2015).

[13] N. Askour, A. Intissar, Z. Mouayn, *Espaces de Bargmann généralisés et formules explicites pour leurs noyaux reproduisants*. C.R. Acad. Sci. Paris, 325, Série I, 707-712 (1997).

[14] R. Balan, *Multiplexing of signals using superframes*. In SPIE Wavelets applications. Signal and Image processing XIII, 4119, 118-129 (2000).

[15] M. B. Balk, Polyanalytic Functions. Akad. Verlag, Berlin (1991).

[16] E. Bombieri, *A note on the large sieve*. Acta Arith., 18 (1), 401-404 (1971).

[17] A. Beurling, *On functions with a spectral gap*, Published in ”Collected Works of Arne Beurling: Volume II Harmonic Analysis”, Birkhäuser, Boston, 370-372 (1989).

[18] A. Bonami, B. Demange, P. Jaming, *Hermite functions and uncertainty principles for the Fourier and the windowed Fourier transforms*. Rev. Mat. Iberoamericana, 19, 23-55 (2003).

[19] E. J. Candés, J. Romberg, T. Tao, *Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information*. IEEE Trans. Inf. Theor., 52, 489-509 (2006).

[20] E. Carneiro, F. Littmann, *Extremal functions in de Branges and Euclidean spaces*. Adv. Math., 260, 281-349 (2014).

[21] I. Daubechies, *Time-frequency localization operators: A geometric phase space approach*. IEEE Trans. Inform. Theory, 34, 605-612 (1988).

[22] F. DeMari, H. G. Feichtinger, and K. Nowak. *Uniform eigenvalue estimates for time-frequency localization operators*. J. London Math. Soc., 65(3), 720-732 (2002).

[23] D. L. Donoho, B. F. Logan, *Signal recovery and the large sieve*. SIAM J. Appl. Math., 52, 577-59 (1992).

[24] D. L. Donoho, B. P. Stark, *Uncertainty principles and signal recovery*. SIAM J. Appl. Math., 49, 906-931 (1989).

[25] D. L. Donoho, *Compressed sensing*. IEEE Trans. Infor. Theory, 52(4), 1289-1306 (2006).

[26] H. G. Feichtinger, *Modulation spaces on locally compact abelian groups*. Technical report, University of Vienna, (1983).

[27] H. G. Feichtinger, *On a new Segal algebra*. Monatsh. Math., 92 (4), 269-289 (1981).
[28] H. G. Feichtinger, K. Gröchenig, A unified approach to atomic decomposition via integrable group representations. Proc. Conference on Functions, Spaces and Applications, Lund, Springer Lect. Notes Math., 52-73 (1988).

[29] H. G. Feichtinger, K. Gröchenig, Banach Spaces Related to Integrable Group Representations and Their Atomic Decompositions I. J. Funct. Anal., 86, 307-340 (1989).

[30] H. G. Feichtinger, K. Gröchenig, Banach Spaces Related to Integrable Group Representations and Their Atomic Decompositions II. Monatsh. Math., 108, 129-148 (1989).

[31] C. Fernández, A. Galbis. Annihilating sets for the short time Fourier transform. Adv. Math., 224, 1904-1926 (2010).

[32] G. B. Folland, Harmonic Analysis in Phase Space. Ann. Math. Stud., Vol.122, Princeton Univ. Press, Princeton, N.J. (1989).

[33] S. Foucart, H. Rauhut, A mathematical introduction to compressive sensing. Birkhäuser 1 (3), (2013).

[34] H. Führ, Simultaneous estimates for vector-valued Gabor frames of Hermite functions. Adv. Comput. Math., 29 (4), 357-373, (2008).

[35] H. Führ, K. Gröchenig, A. Haimi, A. Klotz, J. L. Romero, Density of sampling and interpolation in reproducing kernel Hilbert spaces. J. London Math. Soc., 96 (3), 663-686 (2016).

[36] A. Ghanmi, A class of generalized complex Hermite polynomials. J. Math. Anal. Appl., 340 (2), 1395-1406 (2008).

[37] K. Gröchenig, Foundations of Time-Frequency Analysis. Birkhäuser, Boston, (2001).

[38] K. Gröchenig, Y. Lyubarskii, Gabor (Super)Frames with Hermite Functions. Math. Ann., 345, 267-286 (2009).

[39] K. Gröchenig, E. Malinnikova, Phase space localization of Riesz basis for $L^2(\mathbb{R}^d)$. Rev. Math. Iberoamericana, 29, 1003-1031 (2013).

[40] K. Gröchenig, A. Haimi, J. L Romero, Completeness of Gabor systems. J. Aprox. Theor., 207, 283-300 (2016).

[41] C. Guillemot and O. Le Meur, Image inpainting: Overview and recent advances. IEEE Signal Process. Mag., 31(1), 127-144 (2013).

[42] A. Haimi, H. Hedenmalm, The polyanalytic Ginibre ensembles. J. Stat. Phys., 153 (1), 10-47 (2013).

[43] A. Haimi, H. Hedenmalm, Asymptotic expansions of polyanalytic Bergman kernels. J. Funct. Anal., 267, 4667-4731 (2014).

[44] D. Han, D. R. Larson, Frames, Bases and Group Representations. Mem. Amer. Math. Soc. 147, no. 697 (2000).

[45] O. Hutnık, Wavelets from Laguerre Polynomials and Toeplitz-Type Operators. Integral Equations and Operator Theory, 71, 357 (2011).

[46] M. E. H. Ismail, Analytic properties of complex Hermite polynomials. Trans. Amer. Math. Soc., 368, 1189-1210 (2016).

[47] E. Kowalski, The Large Sieve and its Applications: Arithmetic Geometry, Random Walks and Discrete Groups (Vol. 175). Cambridge University Press, (2008).

[48] J. Lemvig, On some Hermite series identities and their applications to Gabor analysis. Monatsh. Math. 182, 899–912 (2017).

[49] Y. Lyubarskii, Frames in the Bargmann space of entire functions, Entire and subharmonic functions. 167-180, Adv. Soviet Math., 11, Amer. Math. Soc., Providence, RI (1992).

[50] M. Mitkovski, A. Ramirez, Density results for continuous frames. arXiv:1702.05285 (2017).
[51] H. L. Montgomery, *The analytic principle of the large sieve.* Bull. Amer. Math. Soc., 84, 547-546 (1978).

[52] J. Ortega-Cerdà, *Sampling measures.* Publ. Mat., 42, 559-566 (1998).

[53] J. Ortega-Cerdà, K. Seip, *Fourier frames.* Ann. of Math., 155 (3), 789-806 (2002).

[54] N. Perraudin, N. Holighaus, P. Majdak, and P. Balazs, *Inpainting of long audio segments with similarity graphs.* IEEE Trans. Audio, Speech Lang. Proc., (2018).

[55] G. E. Pfander, H. Rauhut, *Sparsity in time-frequency representations.* J. Four. Anal. Appl. 16, 233-260 (2010).

[56] G. E. Pfander, H. Rauhut, J. E Tropp, *The restricted isometry property for time–frequency structured random matrices.* Probab. Theory Related Fields, 156 (3-4), 707-737 (2013).

[57] B. Ricaud, B. Torrésani, *A survey of uncertainty principles and some signal processing applications.* Adv. Comput. Math., 40, 629-650 (2014).

[58] M. Ruzhansky, N. Tokmagambetov, *Very weak solutions of wave equation for Landau Hamiltonian with irregular electromagnetic field.* Lett. Math. Phys., 107, 591-618 (2017).

[59] K. Seip, *Reproducing formulas and double orthogonality in Bargmann and Bergman spaces.* SIAM J. Math. Anal., 22, 856-876 (1991).

[60] K. Seip, R. Wallstén, *Density Theorems for sampling and interpolation in the Bargmann-Fock space I,* J. Reine Angew. Math., 429, 91-106 (1992).

[61] D. Slepian, *Prolate spheroidal wave functions, Fourier analysis and uncertainty-IV: Extensions to many dimensions; generalized prolate spheroidal functions.* Bell Syst. Tech. J., 36, 3009-3057 (1964).

[62] M. Speckbacher, P. Balazs, *Reproducing pairs and Gabor systems at critical density.* J. Math. Anal. Appl., 455(2), 1072-1087 (2017).

[63] N. L. Vasilevski, *Poly-Fock spaces.* Differential operators and related topics, Vol. I (Odessa, 1997), 371–386, Oper. Theory Adv. Appl., 117, Birkhäuser, Basel, (2000).

Acoustics Research Institute, Austrian Academy of Sciences,, Wohllebengasse 12-14, 1040 Vienna, Austria