POLITECNICO DI TORINO
Repository ISTITUZIONALE

On the existence of optimizers for time–frequency concentration problems

*Original*
On the existence of optimizers for time–frequency concentration problems / Nicola, F.; Romero, J. L.; Trapasso, S. I.. - In: CALCULUS OF VARIATIONS AND PARTIAL DIFFERENTIAL EQUATIONS. - ISSN 0944-2669. - 62:1(2023). [10.1007/s00526-022-02358-6]

*Availability:*
This version is available at: 11583/2973560 since: 2022-12-02T09:40:03Z

*Publisher:*
Springer

*Published*
DOI:10.1007/s00526-022-02358-6

*Terms of use:*

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

*Publisher copyright*
Springer postprint/Author's Accepted Manuscript

This version of the article has been accepted for publication, after peer review (when applicable) and is subject to Springer Nature's AM terms of use, but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: http://dx.doi.org/10.1007/s00526-022-02358-6

(Article begins on next page)

12 October 2024

# ON THE EXISTENCE OF OPTIMIZERS FOR TIME-FREQUENCY CONCENTRATION PROBLEMS

FABIO NICOLA, JOSÉ LUIS ROMERO, AND S. IVAN TRAPASSO

Abstract. We consider the problem of the maximum concentration in a fixed measurable subset $\Omega \subset \mathbb{R}^{2d}$ of the time-frequency space for functions $f \in L^2(\mathbb{R}^d)$. The notion of concentration can be made mathematically precise by considering the $L^p$-norm on $\Omega$ of some time-frequency distribution of $f$ such as the ambiguity function $A(f)$. We provide a positive answer to an open maximization problem, by showing that for every subset $\Omega \subset \mathbb{R}^{2d}$ of finite measure and every $1 \leq p < \infty$, there exists an optimizer for

$$\sup\{\|A(f)\|_{L^p(\Omega)} : \ f \in L^2(\mathbb{R}^d), \ \|f\|_{L^2} = 1\}.$$

The lack of weak upper semicontinuity and the invariance under time-frequency shifts make the problem challenging. The proof is based on concentration compactness with time-frequency shifts as dislocations, and certain integral bounds and asymptotic decoupling estimates for the ambiguity function. We also discuss the case $p = \infty$ and related optimization problems for the time correlation function, the cross-ambiguity function with a fixed window, and for functions in the modulation spaces $M^q(\mathbb{R}^d)$, $0 < q < 2$, equipped with continuous or discrete-type (quasi-)norms.

*This is an Accepted Manuscript of an article published by Springer in Calculus of Variations and Partial Differential Equations on 9 November 2022, available at: https://link.springer.com/article/10.1007/s00526-022-02358-6, DOI: 10.1007/s00526-022-02358-6.*

## 1. Introduction and discussion of the main results

The notion of concentration of a function $f \in L^2(\mathbb{R}^d)$ in a measurable subset $\Omega \subset \mathbb{R}^{2d}$ of the time-frequency space is central in harmonic analysis and is also at the core signal processing [19, 36, 25, 49]. From a mathematical point of view, the study of this issue represents a fascinating and multifaceted challenge, with a longstanding and distinguished tradition [22, 28, 29, 44, 12], and it ultimately reduces to one of the different, subtle manifestations of the uncertainty principle [15, 21, 41].

2010 *Mathematics Subject Classification.* 49Q10, 49R05, 42B10, 94A12, 81S30.

*Key words and phrases.* Time-frequency concentration, optimization, ambiguity function, concentration compactness.

A natural family of phase-space concentration measures is given by the $L^p$ norms on $\Omega$ of some *time-frequency distribution* of $f$, such as the short-time Fourier transform (see below), or the *ambiguity function*

$$A(f)(x,\omega) = \int_{\mathbb{R}^d} f\Big(t+\frac{x}{2}\Big)\overline{f\Big(t-\frac{x}{2}\Big)}e^{-2\pi i t\cdot\omega}\,dt, \quad x,\omega\in\mathbb{R}^d,$$

which is a quadratic time-frequency representation popular in engineering and radar applications [8, 51, 37].

The design of maximally concentrated waveforms is of great theoretical and practical interest, as these provide compact elementary blocks tailored to a given tiling of the time-frequency space, according to a paradigm that dates back to the pioneering work by Gabor [23] at least. The companion problem of designing pulses with a peaky ambiguity function is of particular relevance in radar signal analysis [41, Section 3.4.3], wireless communications [2, 38, 42], and signal recovery.

In spite of the importance of the problem and the extensive numerical experimentation (see e.g. [16, 40]), the existence of an optimizer for the functional $\|A(f)\|_{L^p(\Omega)}$, $1\le p<\infty$, among the functions $f\in L^2(\mathbb{R}^d)$, $\|f\|_{L^2}=1$, is still open. While this fact can seem surprising given the maturity of the field of time-frequency analysis, close inspection of the problem soon reveals a number of technical difficulties, including the lack of weak upper semicontinuity of the involved functional and its invariance with respect to a non-compact group of *time-frequency shifts*:

$$\pi(z)f(t) = e^{2\pi i t\cdot\omega}f(t-x), \qquad z=(x,\omega)\in\mathbb{R}^d\times\mathbb{R}^d. \tag{1.1}$$

Our main result establishes the existence of optimizers for the $L^p$-norm of the ambiguity function on a domain.

**Theorem 1.1.** *Let $\Omega\subset\mathbb{R}^{2d}$ be a measurable subset of finite, positive measure, and $1\le p<\infty$. Then the supremum*

$$\sup_{f\in L^2(\mathbb{R}^d)\setminus\{0\}} \frac{\big(\int_\Omega |A(f)(x,\omega)|^p dx d\omega\big)^{1/p}}{\|f\|_{L^2}^2} \tag{1.2}$$

*is attained.*

*Moreover, for $1<p<\infty$, if $f^{(n)}$ is any maximizing sequence normalized in $L^2(\mathbb{R}^d)$, then there exists a subsequence (still denoted by $f^{(n)}$) and $z^{(n)}\in\mathbb{R}^{2d}$ such that $\pi(-z^{(n)})f^{(n)}$ converges in $L^2$ to a maximizer.*

The optimization objective (1.2) is invariant under time-frequency shifts, since $|A(\pi(x,\omega)f)| = |A(f)|$ for $x,\omega\in\mathbb{R}^d$, $f\in L^2(\mathbb{R}^d)$. The first step towards Theorem 1.1 is to account for such symmetries. At the outset, our proof is based on a *concentration compactness* strategy [6, 43, 45, 35, 34, 48], where the time-frequency shifts $\{\pi(x,\omega)\colon x,\omega\in\mathbb{R}^d\}$ serve as dislocation operators. The corresponding profile decompositions of maximizing sequences are then leveraged by means of certain integral estimates for the ambiguity function from [10] — expressing continuity at an intermediate level between the "dislocation topology" and the weak topology — and an asymptotic decoupling property in $L^p(\Omega)$ for sums of functions asymptotically separated *in the Fourier domain*. The latter can be aptly regarded as an asymptotic version of a known almost-orthogonality principle, cf. [47, Lemma 6.1]. Our method yields not only the existence of optimizers, but, for $1<p<\infty$, also implies that

every normalized maximizing sequence is relatively compact in $L^2(\mathbb{R}^d)$, up to time-frequency shifts. This stronger conclusion is consistent with numerical practices that seek to optimize (1.2) by fixing a time-frequency center of gravity [40].

The attainability of (1.2) in the whole time-frequency space ($\Omega = \mathbb{R}^{2d}$) was studied in the celebrated article [31], under the assumption $p \geq 2$ — which is a necessary restriction in that case — and with very different techniques, in particular exploiting the explicit expression of the candidate maximizers. Indeed, the value of (1.2) was exactly calculated and maximizers were characterized as Gaussian functions. For domains $\Omega$ with a special geometry, such as a ball, a similar characterization could be expected. Theorem 1.1 is a first step in that direction, as it implies that maximizers exist and therefore satisfy a certain variational equation. The analysis of such equation is however challenging and we postpone it to a subsequent contribution (in preparation) — cf. [39] for a related problem.

We stress that the conclusion of Theorem 1.1 does not extend to the case $p = \infty$. Instead, we have the following characterization.

**Proposition 1.2.** *Let $\Omega \subset \mathbb{R}^{2d}$ be a measurable subset of finite, positive measure. Then*

$$\sup_{f \in L^2(\mathbb{R}^d)\setminus\{0\}} \frac{\|A(f)\|_{L^\infty(\Omega)}}{\|f\|_{L^2}^2} = 1 \tag{1.3}$$

*and the supremum is attained if and only if $|\Omega \cap B_r| > 0$ for every $r > 0$, where $B_r = \{z \in \mathbb{R}^{2d} : |z| < r\}$. In this case, every $f \in L^2(\mathbb{R}^d) \setminus \{0\}$ is a maximizer.*

The magnitude of the ambiguity function $|A(f)(x,\omega)| = |\langle f, \pi(x,\omega) f\rangle|$ is a *time-frequency auto-correlation function.* To better appreciate the subtleties involved in its optimization, we show that a result similar to Theorem 1.1 fails for time or frequency correlations considered individually. Indeed, denote the translation and modulation operators on $L^2(\mathbb{R}^d)$ by

$$T_x f(t) = f(t-x), \quad M_\omega f(t) = e^{2\pi i t\cdot\omega} f(t), \quad x, \omega \in \mathbb{R}^d,\ f \in L^2(\mathbb{R}^d), \tag{1.4}$$

so that $\pi(x,\omega) = M_\omega T_x$. The following result is in stark contrast with Theorem 1.1.

**Proposition 1.3.** *Let $\Omega \subset \mathbb{R}^d$ be a measurable subset of finite, positive measure, and $1 \leq p < \infty$. Then*

$$\sup_{f \in L^2(\mathbb{R}^d)\setminus\{0\}} \frac{\left(\int_\Omega |\langle f, T_x f\rangle|^p dx\right)^{1/p}}{\|f\|_{L^2}^2} = |\Omega|^{1/p} \tag{1.5}$$

*and the supremum is not attained.*

Of course, a similar negative result holds true for the frequency correlation function $\langle f, M_\omega f\rangle = \langle \hat{f}, T_\omega \hat{f}\rangle$. As $\langle f, T_x f\rangle = (f * f^\vee)(x)$, with $f^\vee(x) = \overline{f(-x)}$, Proposition 1.3 could be rephrased as an optimization problem for positive definite functions and it is related, at least in spirit, to the optimization of the constants in Young's inequality; see, e.g., [32, Chapter 4].

Optimization problems analogous to (1.2) can be considered also for linear time-frequency representations, such as the *short-time Fourier transform* $V_g f(x,\omega) =$

$\langle f, \pi(x,\omega)g\rangle$, where $g \in L^2(\mathbb{R}^d) \setminus \{0\}$ is a fixed window function. While the short-time Fourier transform is not intrinsically associated with the function $f$, as it requires the introduction of an additional parameter $g$, it is a popular tool in signal analysis, in part because it is mathematically simpler than the ambiguity function. For example, the existence of a maximizer for $\|V_g f\|_{L^p(\Omega)}$ is much easier to establish than Theorem 1.1, because the introduction of the window function $g$ weakens the nonlinearity of the optimization objective, replacing the quadratic term $|A(f)|$ with the so-called *cross-ambiguity* $|A(f,g)| = |V_g f|$. A proof of the existence of optimizers for $\|V_g f\|_{L^p(\Omega)}$ and a technical comparison to Theorem 1.1 is presented in Section 5. (The case $p = 2$ is straightforward, as it corresponds to the maximization of the eigenvalues of a so-called *localization operator* [12, 50]; finer questions such as optimal domains of prescribed measure and characterization of extremizers for Gaussian windows are studied in [39].)

We also point out that different but related optimization problems have been considered in the literature over the years, such as maximizing the *integral* on a subset $\Omega \subset \mathbb{R}^{2d}$ of time-frequency distributions in the Cohen class [7] (as opposed to their $L^p$-norms); see for instance [17, 33]. For this kind of optimization, we refer the reader to the comprehensive recent survey [30], where deep connections with the spectral theory of pseudo-differential operators are discussed.

Finally, we stress once again that $|A(f)(x,\omega)| = |\langle f, \pi(x,\omega)f\rangle|$, so that (1.1) can be also regarded as an optimization problem for the diagonal matrix coefficients of the Schrödinger representation of the reduced Heisenberg group. This point of view encourages us to investigate for which other groups and unitary representations a similar property holds – an interesting question that appears to be largely open at the time of writing. Indeed, the matrix coefficients encode the properties of the corresponding representation and their study has a well-established tradition [1, 27, 11, 26, 14], focused on proving refined estimates on the whole group – as opposite to a subset $\Omega$.

The article is organized as follows. In Section 2 we provide brief background on time-frequency analysis and concentration compactness. Section 3 is devoted to the proof of Theorem 1.1, whereas in Section 4 we prove Propositions 1.2 and 1.3. In Section 5 we consider the optimization problem for $\|A(f,g)\|_{L^p(\Omega)}$ and discuss technical differences with respect to Theorem 1.1. Finally, Section 6 provides two variants of Theorem 1.1. There, we replace $L^2(\mathbb{R}^d)$ by the *modulation spaces* $M^q(\mathbb{R}^d)$, $0 < q < 2$, which are (quasi-)Banach spaces defined by imposing certain integrability requirements to the short-time Fourier transform, widely used in time-frequency analysis [3]. More precisely, we incorporate modulation-space norms into the optimization objective (1.2), and also consider their often preferred discrete counterparts.

## 2. Notation and preliminary results

2.1. **General notation.** The inner product in $L^2(\mathbb{R}^d)$ is denoted by $\langle \cdot, \cdot \rangle$. The space of Schwartz functions in $\mathbb{R}^d$ is denoted by $\mathcal{S}(\mathbb{R}^d)$, while $\mathcal{S}'(\mathbb{R}^d)$ stands for the space of temperate distributions. We write $A \lesssim B$ if $A \leq CB$ for some absolute constant $C > 0$, whereas $A \lesssim_k B$ means that such a constant depends on the parameter

$k$. The Lebesgue measure of a subset $E$ of $\mathbb{R}^d$ (or $\mathbb{R}^{2d}$) is denoted by $|E|$ while $\chi_E$ stands for its characteristic function.

2.2. **Tools from time-frequency analysis.** We recall some definitions and facts from time-frequency analysis; see [25, 20, 18] for extensive background. The main objects are the time-frequency shifts (1.1), which define a *unitary projective representation* $z \mapsto \pi(z)$ of $\mathbb{R}^{2d}$ on $L^2(\mathbb{R}^d)$. In particular, for all $z, z' \in \mathbb{R}^{2d}$,

$$\pi(z)\pi(z') = \sigma(z, z')\pi(z + z'), \tag{2.1}$$

$$\pi(z)^* = \overline{\sigma(z, -z)}\pi(-z), \tag{2.2}$$

where $\sigma(z, z')$ is an adequate unimodular complex number, called *cocycle*.

*Function spaces.* We now fix a *window function* $g \in \mathcal{S}(\mathbb{R}^d) \setminus \{0\}$. Most definitions below depend (albeit non-essentially) on such choice. The short-time Fourier transform (STFT) of a temperate distribution $f \in \mathcal{S}'(\mathbb{R}^d)$ is defined by

$$V_g f(z) = \langle f, \pi(z)g\rangle, \quad z \in \mathbb{R}^{2d}. \tag{2.3}$$

By considering the $L^p$ norm of $V_g f$ in $\mathbb{R}^{2d}$ one can naturally measure the time-frequency content of a distribution and introduce corresponding families of function spaces. For example, the *modulation space* $M^\infty(\mathbb{R}^d)$ consists of temperate distributions $f \in \mathcal{S}'(\mathbb{R}^d)$ such that

$$\|f\|_{M^\infty} := \sup_{z\in\mathbb{R}^{2d}} |V_g f(z)| = \sup_{z\in\mathbb{R}^{2d}} |\langle f, \pi(z)g\rangle| < \infty.$$

Different choices of the window produce equivalent norms, and $L^2(\mathbb{R}^d) \hookrightarrow M^\infty(\mathbb{R}^d)$.

The Wiener amalgam space $\mathcal{W}(L^2, L^\infty)$ consists of all measurable functions $f : \mathbb{R}^d \to \mathbb{C}$ such that

$$\|f\|_{\mathcal{W}(L^2,L^\infty)} := \operatorname*{ess\,sup}_{y\in\mathbb{R}^d} \|f\, T_y g\|_{L^2} < \infty,$$

where $T_y$ is the translation (1.4). Again, different windows give rise to equivalent norms and $L^2(\mathbb{R}^d) \hookrightarrow \mathcal{W}(L^2, L^\infty)$. Notice that the reverse inclusion holds locally, namely

$$\|f\|_{L^2(K)} \lesssim_K \|f\|_{\mathcal{W}(L^2,L^\infty)}, \tag{2.4}$$

for every compact subset $K \subset \mathbb{R}^d$. This follows immediately if the window $g$ is chosen so that $g = 1$ in a sufficiently large ball, so that $f = f\, T_y g$ on $K$ for sufficiently small $y$.

*Cross-ambiguity function.* The cross-ambiguity function of $f, g \in L^2(\mathbb{R}^d)$ is

$$A(f, g)(x, \omega) = \int_{\mathbb{R}^d} f\Big(t + \frac{x}{2}\Big)\overline{g\Big(t - \frac{x}{2}\Big)} e^{-2\pi i t\cdot\omega}\, dt = e^{\pi i x\cdot\omega} V_g f(x, \omega), \qquad x, \omega \in \mathbb{R}^d.$$

Hence $A(f, f) = A(f)$ is the ambiguity function of $f$. It is easy to see that $A(f, g)$ is a continuous function in $\mathbb{R}^{2d}$ and vanishes at infinity. Moreover, by the Cauchy-Schwarz inequality,

$$|A(f, g)| \le \|f\|_{L^2}\|g\|_{L^2}.$$

The following estimate from [10, Corollary 4.2] will play a crucial role:

$$\|A(f, g)\|_{\mathcal{W}(L^2,L^\infty)} \lesssim \|f\|_{L^2}\|g\|_{M^\infty}, \tag{2.5}$$

where of course the space $\mathcal{W}(L^2, L^\infty)$ is understood in $\mathbb{R}^{2d}$. While we only need (2.5) for $f, g \in L^2(\mathbb{R}^d)$, the formula is still valid for $f \in L^2(\mathbb{R}^d)$ and $g \in M^\infty(\mathbb{R}^d)$. (In that case, $A(f, g)$ is a priori defined only as a temperate distribution, and part of the content of (2.5) is that $A(f, g)$ is in fact locally in $L^2$ when the right-hand side is finite). Thus, as in many other parts of the article, we are concerned with $L^2(\mathbb{R}^d)$ equipped with the $M^\infty$ (norm) topology, but not with genuine distributions in $M^\infty(\mathbb{R}^d)$.

We finally recall from [20, Chapter 4] or [13, Proposition 175 and Corollary 217] the covariance property for the cross-ambiguity function: if $\mathcal{A}$ is a (real $2d \times 2d$) symplectic matrix we have

$$A(U_{\mathcal{A}} f, U_{\mathcal{A}} g)(z) = A(f, g)(\mathcal{A}z) \quad z \in \mathbb{R}^{2d},\ f, g \in L^2(\mathbb{R}^d) \tag{2.6}$$

for a suitable unitary operator $U_{\mathcal{A}}$ on $L^2(\mathbb{R}^d)$ (called metaplectic operator).

2.3. **Tools from concentraction compactness.** Concentration compactness is a general paradigm to study optimization problems when compactness arguments fail due to invariance under the action of a non-compact group (see e.g. [34, 35, 48, 46]). We recall some basic facts specialized to the (projective) representation given by the time-frequency shifts. The main conclusion is that any bounded sequence in $L^2(\mathbb{R}^d)$ has a subsequence with a special profile decomposition.

*Time-frequency shifts as dislocations in $L^2$.* It is easily checked that time-frequency shifts satisfy the following important *dislocation property*:

(2.7) If $f \in L^2(\mathbb{R}^d)$ and $z_n \in \mathbb{C}$ is a sequence with $|z_n| \to +\infty$, then $\pi(z_n)f$ converges weakly to 0 in $L^2(\mathbb{R}^d)$.

The dislocation property allows us to apply the theory of concentration compactness, because it implies that the time-frequency shifts $\{\pi(z) : z \in \mathbb{R}^{2d}\}$ define a so-called *dislocation set* of unitary operators in $L^2(\mathbb{R}^d)$ ([48, Definition 3.2]). Whereas we do not need to recall the general (technical) definition here, we observe that, according to [48, Proposition 3.1], it is sufficient to check that

(2.8) If for some sequences $z_n, z_n' \in \mathbb{R}^{2d}$, the sequence of operators $\pi(z_n')^*\pi(z_n)$ does not converge weakly to zero (as bounded operators on $L^2(\mathbb{R}^d)$), then it has a strongly convergent subsequence.

To see that property (2.8) holds, note first that, by (2.1) and (2.2), $\pi(z_n')^*\pi(z_n) = c(z_n, z_n')\pi(z_n - z_n')$, with $|c(z_n, z_n')| = 1$. Hence, if $\pi(z_n')^*\pi(z_n)$, does not converge weakly to zero, by (2.7), $|z_n - z_n'|$ does not tend to $+\infty$, and therefore has a convergent subsequence. By passing to a further subsequence, the phase factors $c(z_n, z_n')$ will also converge, and the conclusion follows from the strong continuity of the representation $\mathbb{R}^{2d} \ni z \mapsto \pi(z)$.

*D-weak convergence.* Associated with the set of dislocations $\{\pi(z) : z \in \mathbb{R}^{2d}\}$, there is a corresponding notion of *weak dislocation convergence — D-weak convergence* for short [48, Definition 3.1]): a sequence $f_n$ in $L^2(\mathbb{R}^d)$ $D$-weakly converges to $f \in L^2(\mathbb{R}^d)$ if for every $g \in L^2(\mathbb{R}^d)$:

$$\sup_{z \in \mathbb{R}^{2d}} |\langle f_n - f, \pi(z)g\rangle| \to 0.$$

Letting $g$ be a window function for the short-time Fourier transform (2.3) one sees that $D$-weak convergence implies convergence in $M^\infty(\mathbb{R}^d)$.

*Profile decomposition.* The general theory of concentration compactness in Hilbert spaces (see e.g. [48, Theorem 3.1 and its proof] or [46, Theorem 4.5.3]) now yields the following: Let $f^{(n)}$ be a sequence in $L^2(\mathbb{R}^d)$ with $\limsup_{n\to\infty} \|f^{(n)}\|_{L^2} \le 1$, then there exists a subsequence (that we still denote $f^{(n)}$) and *profiles* $f_j \in L^2(\mathbb{R}^d)$, $j = 1, 2, \dots$, such that the following *profile decomposition* holds for $k = 1, 2, \dots$:

$$f^{(n)} = \sum_{j=1}^{k} \pi\big(z_j^{(n)}\big) f_j + w_k^{(n)} \tag{2.9}$$

for suitable $z_j^{(n)} \in \mathbb{R}^{2d}$, $w_k^{(n)} \in L^2(\mathbb{R}^d)$, where

$$|z_j^{(n)} - z_{j'}^{(n)}| \to +\infty \quad \text{as } n \to \infty, \text{ if } j \neq j' \text{ and } \ f_j \neq 0, f_{j'} \neq 0, \tag{2.10}$$

$$\sum_{j=1}^{k} \|f_j\|_{L^2}^2 + \limsup_{n\to\infty} \|w_k^{(n)}\|_{L^2}^2 \le 1, \tag{2.11}$$

$$\lim_{k\to\infty} \limsup_{n\to\infty} \|w_k^{(n)}\|_{M^\infty} = 0. \tag{2.12}$$

Moreover,

$$\pi\big(z_j^{(n)}\big)^* w_k^{(n)} \to 0 \text{ weakly in } L^2, \text{ as } n \to \infty, \text{ for each } k \ge 1 \text{ and } 1 \le j \le k, \tag{2.13}$$

and

$$\lim_{n\to\infty} \Big\| \sum_{j=1}^{k} \pi(z_j^{(n)}) f_j \Big\|_{L^2}^2 = \sum_{j=1}^{k} \|f_j\|_{L^2}^2. \tag{2.14}$$

**Remark 2.1.** *As a consequence of the decomposition, we see that $\pi\big(z_k^{(n)}\big)^* f^{(n)}$ converges weakly to $f_k$ in $L^2$, as $n \to \infty$, for each $k \ge 1$.*

For simplicity, we have introduced the time-frequency profile decomposition as an application of the abstract theory of dislocation sets [48]. Alternatively, it would have also been possible to derive the decomposition from the theory of *dislocation groups* [46, Section 4.5.2], by considering the *reduced Weyl-Heisenberg group* $\{\lambda\pi(z) : z \in \mathbb{R}^{2d}, |\lambda| = 1\}$ and by exploiting the compactness of the set of phase factors $\lambda$.

## 3. Proof of the main result (Theorem 1.1)

In the following lemmas we deal with $k$-tuples $h = (h_1, \dots, h_k)$ of functions on a measure space endowed with a $\sigma$-finite measure and we use the notation $\ell^r(L^p)$ for the corresponding vector-valued norm:

$$\|h\|_{\ell^r(L^p)} = \left( \sum_{j=1}^{k} \|h_j\|_{L^p}^r \right)^{1/r}. \tag{3.1}$$

The following result provides a version of the classical Riesz-Thorin interpolation theorem for linear operators defined only on some finite dimensional subspace of simple functions.

**Lemma 3.1.** *Let $N$ be a positive integer and let $E_1, \dots E_N$ be disjoint measurable sets of finite measure. Let*

$$X = X_{E_1,\dots,E_N} = \mathrm{span}_{\mathbb{C}}\{\chi_{E_1}, \dots, \chi_{E_N}\}.$$

*Let $1 \le p_0, q_0, p_1, q_1, r_0, r_1, p, q, r \le \infty$ and $0 < \theta < 1$, with $1/p = (1-\theta)/p_0 + \theta/p_1$, $1/q = (1-\theta)/q_0 + \theta/q_1$, $1/r = (1-\theta)/r_0 + \theta/r_1$, $p, r < \infty$.*

*Let $T : X^k \to L^{q_0} + L^{q_1}$ be a linear operator and $M_0, M_1 > 0$ such that*

$$\|Th\|_{L^{q_0}} \le M_0 \|h\|_{\ell^{r_0}(L^{p_0})} \tag{3.2}$$

*and*

$$\|Th\|_{L^{q_1}} \le M_1 \|h\|_{\ell^{r_1}(L^{p_1})} \tag{3.3}$$

*for every $h \in X^k$. Then*

$$\|Th\|_{L^{q}} \le M_0^{1-\theta} M_1^{\theta} \|h\|_{\ell^{r}(L^{p})} \tag{3.4}$$

*for every $h \in X^k$.*

*Proof.* A function $h = (h_1, \dots, h_k) \in X^k$ can be written uniquely as

$$h_j = \sum_{i=1}^{N} a_j^i \, \chi_{E_i},$$

with $a_j^i \in \mathbb{C}$. In terms of the coefficients, the norm (3.1) reads

$$\|h\|_{\ell^r(L^p)} = \Big( \sum_{j=1}^{k} \Big( \sum_{i=1}^{N} |a_j^i|^p |E_i| \Big)^{r/p} \Big)^{1/r}.$$

Therefore, the conclusion follows from standard interpolation results with respect to weighted sequences spaces; see, e.g., [4, Theorems 5.1.1, 5.1.2, and 5.6.3]. Alternatively, one can see that the standard proof of the Riesz-Thorin theorem can be carried out within the space $X^k$. □

**Remark 3.2.** *Lemma 3.1 fails if the sets $E_1, \dots, E_n$ are not disjoint. In particular, the conclusion does not generalize to all finite dimensional subspaces $X$ of simple functions. For example, let $k = 1$ (scalar case), the measure space $\{0, 1, 2\}$ be endowed with the counting measure, $E_1 = \{0, 1\}$, $E_2 = \{1, 2\}$ and $X = \mathrm{span}_{\mathbb{C}}\{\chi_{E_1}, \chi_{E_2}\}$, $Th = (h(0)+h(1)+h(2))\chi_{\{0\}}$ for $h \in X$, $1 \le q_0, q_1, r_0, r_1 \le \infty$ arbitrary, $p_0 = 1$, $p_1 = \infty$, $p = 1/\theta = 2$. Then,* (3.2) *holds with $M_0 = 1$ and* (3.3) *holds with $M_1 = 2$, while the function $h = \chi_{E_1} + \chi_{E_2}$ provides a counterexample to the corresponding estimate* (3.4).

As an application of Lemma 3.1 we obtain the following asymptotic interpolation estimate for sequences of operators.

**Lemma 3.3.** *With the same notation of Lemma 3.1, let $T_n : \ell^{r_0}(L^{p_0}) + \ell^{r_1}(L^{p_1}) \to L^{q_0} + L^{q_1}$, $n \in \mathbb{N}$, be a sequence of linear operators.*

*Suppose that $T_n$ is bounded $\ell^{r_0}(L^{p_0}) \to L^{q_0}$ for every $n \in \mathbb{N}$, with*

$$\limsup_{n\to\infty} \|T_n h\|_{L^{q_0}} \le M_0 \|h\|_{\ell^{r_0}(L^{p_0})}, \tag{3.5}$$

*and, for every* $n \in \mathbb{N}$,

$$\|T_n h\|_{L^{q_1}} \leq M_1 \|h\|_{\ell^{r_1}(L^{p_1})}. \tag{3.6}$$

*Then*

$$\limsup_{n\to\infty} \|T_n h\|_{L^q} \leq M_0^{1-\theta} M_1^{\theta} \|h\|_{\ell^r(L^p)}. \tag{3.7}$$

*Proof.* First, we prove that (3.7) holds when $h$ is a $k$-tuple of simple functions; hence $h$ belongs to some space $X^k$, with $X = X_{E_1,\dots,E_N}$ and $E_1, \dots, E_N$ of finite measure and pairwise disjoint, as in Lemma 3.1.

The family of operators $T_n : \ell^{r_0}(L^{p_0}) \to L^{q_0}$ is equicontinuous by the uniform boundedness principle, hence the estimate (3.5) holds uniformly with respect to $h$ when $h$ belongs to a compact subset of $\ell^{r_0}(L^{p_0})$. Precisely, for every $\epsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that

$$\|T_n h\|_{L^{q_0}} \leq M_0 \|h\|_{\ell^{r_0}(L^{p_0})} + \epsilon$$

for every $n \geq n_0$ and every $h$ in a compact subset of $\ell^{r_0}(L^{p_0})$. In particular, this holds for functions $h$ normalized in $\ell^{r_0}(L^{p_0})$ and in the finite dimensional space $X^k$, with $X = X_{E_1,\dots E_N}$ as above. By homogeneity, we deduce that for every $\epsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that

$$\|T_n h\|_{L^{q_0}} \leq (M_0 + \epsilon) \|h\|_{\ell^{r_0}(L^{p_0})}$$

for $h \in X^k$, $n \geq n_0$. By (3.6) and Lemma 3.1 we obtain

$$\|T_n h\|_{L^q} \leq (M_0 + \epsilon)^{1-\theta} M_1^{\theta} \|h\|_{\ell^r(L^p)}$$

for $h \in X^k$, $n \geq n_0$, which implies (3.7) for $h \in X^k$.

Since $p$ is assumed to be finite, the set of simple functions is dense in $L^p$ and the family of operators $T_n : \ell^r(L^p) \to L^q$ is equicontinuous (by the assumptions and complex interpolation), so that (3.7) holds for every $h \in \ell^r(L^p)$. □

**Remark 3.4.** *It is is easy to see that the conclusion of Lemma 3.3 still holds if* (3.6) *is replaced by the assumption that* $T_n$ *is bounded* $\ell^{r_1}(L^{p_1}) \to L^{q_1}$ *for every* $n \in \mathbb{N}$, *with*

$$\limsup_{n\to\infty} \|T_n h\|_{L^{q_1}} \leq M_1 \|h\|_{\ell^{r_1}(L^{p_1})},$$

*but we will not need this fact.*

We are now ready to prove our main result.

*Proof of Theorem 1.1.* **Step 1.** *Profile decomposition.*

Let $L$ be the supremum in (1.2). Since $\Omega$ has positive measure, we have $L > 0$. (Indeed, it is easy to see that there exists $f \in L^2(\mathbb{R}^d)$ such that $A(f) \neq 0$ on $\Omega$; see, for example, the proof of Proposition 1.2 below for details.) Moreover, from the pointwise estimate $|Af| \leq \|f\|_{L^2}^2$ we obtain $L \leq |\Omega|^{1/p}$, hence $L$ is finite.

Let $f^{(n)}$ be a maximizing sequence, that we can assume normalized without loss of generality, that is $\|f^{(n)}\|_{L^2} = 1$. After passing to a subsequence, we apply the profile decomposition described in Section 2.3. We use the notation introduced there and

we also set $F_k^{(n)} := \sum_{j=1}^k \pi\big(z_j^{(n)}\big) f_j$, so that $f^{(n)} = F_k^{(n)} + w_k^{(n)}$. Observe for future reference that by (2.11) and (2.14) we have

$$\|F_k^{(n)}\|_{L^2}^2 + \|w_k^{(n)}\|_{L^2}^2 \leq 2, \quad n \geq n_k, \tag{3.8}$$

for some $n_k \in \mathbb{N}$ depending on $k$.

Using the sesquilinearity of the cross-ambiguity distribution, we can write

$$\begin{aligned} A(f^{(n)}) = \sum_{j=1}^k A(\pi(z_j^{(n)})f_j) + \sum_{\substack{1\leq j,j'\leq k\\ j\neq j'}} A(\pi(z_j^{(n)})f_j, \pi(z_{j'}^{(n)})f_{j'}) \\ + A(F_k^{(n)}, w_k^{(n)}) + A(w_k^{(n)}, F_k^{(n)}) + A(w_k^{(n)}). \end{aligned} \tag{3.9}$$

We now study the asymptotic behavior, as $n \to \infty$, of the $L^p(\Omega)$-norm of the terms on the right-hand side of (3.9).

**Step 2.** *Asymptotic decoupling.* A simple computation gives

$$A(\pi(z_j^{(n)})f_j) = M_{Jz_j^{(n)}} A(f_j) \tag{3.10}$$

where $J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$ is the canonical symplectic matrix and $M$ stands for the modulation operator (here in $\mathbb{R}^{2d}$).

For fixed $k \geq 1$ and $n \in \mathbb{N}$, consider the operator

$$T_k^{(n)} : L^2(\Omega)^k \to L^2(\Omega), \qquad T_k^{(n)}(h_1, \dots, h_k) = \sum_{j=1}^k M_{Jz_j^{(n)}} h_j,$$

where the $h_j$'s are understood extended by zero on $\mathbb{R}^{2d} \setminus \Omega$. By (2.10) and the Riemann-Lebesgue lemma

$$\begin{aligned} \lim_{n\to\infty} \Big\| T_k^{(n)}(h_1, \dots, h_k) \Big\|_{L^2(\Omega)}^2 &= \lim_{n\to\infty} \sum_{j,j'=1}^k \langle h_j \overline{h_{j'}}, M_{J\left(z_j^{(n)} - z_{j'}^{(n)}\right)} \chi_\Omega \rangle_{L^2(\mathbb{R}^{2d})} \\ &= \sum_{j=1}^k \|h_j\|_{L^2(\Omega)}^2. \end{aligned} \tag{3.11}$$

In addition, for any $1 \leq p_1 \leq \infty$,

$$\Big\| T_k^{(n)}(h_1, \dots, h_k) \Big\|_{L^{p_1}(\Omega)} = \Big\| \sum_{j=1}^k M_{Jz_j^{(n)}} h_j \Big\|_{L^{p_1}(\Omega)} \leq \sum_{j=1}^k \|h_j\|_{L^{p_1}(\Omega)}. \tag{3.12}$$

We now use Lemma 3.3 to interpolate between (3.12) with $p_1 = 1, \infty$ and (3.11), and obtain

$$\limsup_{n\to\infty} \Big\| \sum_{j=1}^k M_{Jz_j^{(n)}} h_j \Big\|_{L^p(\Omega)} \leq \left( \sum_{j=1}^k \|h_j\|_{L^p(\Omega)}^{p^*} \right)^{1/p^*},$$

with

$$p^* = \min\{p, p'\}.$$

Applying this estimate to $h_j = A(f_j)$ and combining it with (3.10) we obtain the following *asymptotic decoupling estimate*:

$$\limsup_{n\to\infty} \Big\| \sum_{j=1}^{k} A(\pi(z_j^{(n)})f_j)\Big\|_{L^p(\Omega)} \leq \left(\sum_{j=1}^{k} \|A(f_j)\|_{L^p(\Omega)}^{p^*}\right)^{1/p^*}. \tag{3.13}$$

**Step 3.** *The error terms in* (3.9).

Proceeding with the analysis of the other terms in (3.9), we have

$$|A(\pi(z_j^{(n)})f_j, \pi(z_{j'}^{(n)})f_{j'})| = |A(f_j, \pi(z_{j'}^{(n)} - z_j^{(n)})f_{j'})|. \tag{3.14}$$

Since $A(f,g) \to 0$ at infinity if $f, g \in L^2(\mathbb{R}^d)$, we see by (2.10) and the dominated convergence theorem (the expression in (3.14) is $\leq 1$ in $\mathbb{R}^{2d}$), that

$$\|A(\pi(z_j^{(n)})f_j, \pi(z_{j'}^{(n)})f_{j'})\|_{L^p(\Omega)} \to 0 \quad \text{as } n \to \infty, \qquad (j \neq j'). \tag{3.15}$$

Second, concerning the term $A(F_k^{(n)}, w_k^{(n)})$ in (3.9), in view of the embedding (2.4) and the estimate (2.5) we deduce that, for every compact $K \subset \mathbb{R}^{2d}$,

$$\begin{aligned}\|A(F_k^{(n)}, w_k^{(n)})\|_{L^2(K)} &\lesssim_K \|A(F_k^{(n)}, w_k^{(n)})\|_{\mathcal{W}(L^2,L^\infty)} \\ &\lesssim \|F_k^{(n)}\|_{L^2}\|w_k^{(n)}\|_{M^\infty},\end{aligned}$$

where the implied constant is independent of $k, n$. Using (3.8) we see that, for $n \geq n_k$,

$$\|A(F_k^{(n)}, w_k^{(n)})\|_{L^2(K)} \lesssim_K \|w_k^{(n)}\|_{M^\infty}$$

and

$$\|A(F_k^{(n)}, w_k^{(n)})\|_{L^\infty(\mathbb{R}^{2d})} \leq \|F_k^{(n)}\|_{L^2}\|w_k^{(n)}\|_{L^2} \leq 1. \tag{3.16}$$

Hence, by Hölder's inequality, we conclude that, for $n \geq n_k$,

$$\|A(F_k^{(n)}, w_k^{(n)})\|_{L^p(K)} \lesssim_K \|w_k^{(n)}\|_{M^\infty}^{\theta_p},$$

where $\theta_p = 1$ for $1 \leq p \leq 2$, and $\theta_p = 2/p$ if $2 < p < \infty$. We claim that this implies

$$\lim_{k\to\infty} \limsup_{n\to\infty} \|A(F_k^{(n)}, w_k^{(n)})\|_{L^p(\Omega)} = 0. \tag{3.17}$$

Indeed, this is clear from (2.12) if $\Omega$ is compact. When $\Omega$ is merely a measurable set with finite measure we can choose a compact subset $K \subset \Omega$ with $|\Omega \setminus K|$ arbitrarily small and use again the uniform bound (3.16) on $\Omega \setminus K$.

The same argument shows that

$$\lim_{k\to\infty} \limsup_{n\to\infty} \|A(w_k^{(n)}, F_k^{(n)})\|_{L^p(\Omega)} = 0, \tag{3.18}$$

and

$$\lim_{k\to\infty} \limsup_{n\to\infty} \|A(w_k^{(n)})\|_{L^p(\Omega)} = 0. \tag{3.19}$$

**Step 4.** *Existence of optimizers.*

By the very definition of $f^{(n)}$, (3.9), the triangle inequality and (3.13), (3.15), (3.17), (3.18), (3.19), letting first $n \to \infty$ and then $k \to \infty$, we obtain

$$L = \lim_{n\to\infty} \|A(f^{(n)})\|_{L^p(\Omega)} \leq \left( \sum_{j=1}^{\infty} \|A(f_j)\|_{L^p(\Omega)}^{p^*} \right)^{1/p^*}.$$

On the other hand, by the very definition of $L$, we have

$$\|A(f_j)\|_{L^p(\Omega)} \leq L\|f_j\|_{L^2}^2 \tag{3.20}$$

so that from (2.11) we obtain

$$\begin{aligned} L = \lim_{n\to\infty} \|A(f^{(n)})\|_{L^p(\Omega)} &\leq \left( \sum_{j=1}^{\infty} \|A(f_j)\|_{L^p(\Omega)}^{p^*} \right)^{1/p^*} \\ &\leq L\left( \sum_{j=1}^{\infty} \|f_j\|_{L^2}^{2p^*} \right)^{1/p^*} \leq L\sum_{j=1}^{\infty} \|f_j\|_{L^2}^2 \leq L. \end{aligned}$$

We then see that all these inequalities must be equalities. If $p > 1$, so that $p^* > 1$, this is possible only if $f_j = 0$ except for one $j$, say $j = 1$, and $\|f_1\|_{L^2} = 1$. Hence $f_1$ is a maximizer.

Moreover, concerning the claim in the statement for $1 < p < \infty$, from (2.9) we have $f^{(n)} = \pi(z_1^{(n)})f_1 + w_1^{(n)}$. Since, by Remark 2.1, $\pi\big(z_1^{(n)}\big)^* f^{(n)}$ converges weakly to $f_1$ in $L^2$, and $\|f^{(n)}\|_{L^2} = \|f_1\|_{L^2} = 1$, we deduce that $\pi\big(z_1^{(n)}\big)^* f^{(n)}$ converges to $f_1$ in $L^2$. Finally, by passing to a subsequence, the cocycles in (2.2) can be assumed to converge, $\sigma(z^{(n)}, -z^{(n)}) \to c$, and therefore $\pi\big(-z_1^{(n)}\big)f^{(n)}$ converges to the maximizer $cf_1$.

If $p = 1$, then $p^* = 1$, and the same chain of equalities implies, together with (3.20), that equality holds in (3.20) for all $j$. Since $f_j \neq 0$ for at least one $j$, such a $f_j$ will be a maximizer for the problem (1.2). The proof is then concluded. □

**Remark 3.5.** *If one is only interested in the existence of a maximizer, even in the case $p > 1$, the conclusion would follow as in the case $p = 1$, that is, by applying the triangle inequality to $\sum_{j=1}^{k} A(\pi(z_j^{(n)})f_j)$ - hence ignoring the oscillations and without using Lemma 3.3. The more elaborated argument given above is rewarded with the stronger conclusion for $1 < p < \infty$.*

## 4. Proofs of Propositions 1.2 and 1.3

*Proof of Proposition 1.2.* It is clear that $\|A(f)\|_{L^\infty(\Omega)} \leq \|f\|_{L^2}^2$. On the other hand, consider a point $(x_0, \omega_0) \in \Omega$ of positive Lebesgue density for $\Omega$ (which exists since $|\Omega| > 0$). Using the covariance property of $A(f)$ under symplectic transformations recalled in (2.6) and the transitivity of the linear symplectic group on $\mathbb{R}^{2d} \setminus \{0\}$, we can suppose $\omega_0 = 0$. Let $f(x) = 2^{d/4}e^{-\pi|x|^2}$ and $f_\lambda(x) = \lambda^{-d/2}f(x/\lambda)$, $\lambda > 0$. Explicit computations show that $\|f_\lambda\|_{L^2} = 1$ and

$$|A(f_\lambda)(x,\omega)| = |V_{f_\lambda}f_\lambda(x,\omega)| = e^{-\frac{\pi}{2\lambda^2}|x|^2 - \frac{\pi\lambda^2}{2}|\omega|^2},$$

see, e.g., [20, Proposition 1.48 and Appendix A]. Since $A(f_\lambda)$ is continuous at the point $(x_0, 0)$, of positive Lebesgue density for $\Omega$,

$$\|A(f_\lambda)\|_{L^\infty(\Omega)} \geq e^{-\frac{\pi}{2\lambda^2}|x_0|^2},$$

which implies

$$\liminf_{\lambda\to+\infty} \|A(f_\lambda)\|_{L^\infty(\Omega)} \geq 1.$$

Concerning the existence and characterization of maximizers, we invoke the following *radar correlation estimate*:

$$|A(f)(x,\omega)| < A(f)(0,0) = 1, \quad \text{if } \|f\|_{L^2} = 1; \tag{4.1}$$

see, e.g., [25, Lemma 4.2.1]. Hence, if $|\Omega \cap B_r| > 0$ for every $r > 0$, it follows from the continuity of $A(f)$ that every $f \in L^2(\mathbb{R}^d) \setminus \{0\}$ is a maximizer.

If instead there exists $r_0 > 0$ such that $|\Omega \cap B_{r_0}| = 0$, then

$$\|A(f)\|_{L^\infty(\Omega)} = \|A(f)\|_{L^\infty(\Omega\setminus B_{r_0})} \leq \sup_{(x,\omega)\in\overline{\Omega\setminus B_{r_0}}} |A(f)(x,\omega)|.$$

Since $A(f)$ vanishes at infinity, in view of (4.1), this last supremum is still $< 1$, because it is attained at some point of the closed set $\overline{\Omega \setminus B_{r_0}}$, which does not contain the origin. □

*Proof of Proposition 1.3.* Let us first prove (1.5). From the trivial pointwise estimate $|\langle f, T_x f\rangle| \leq \|f\|_{L^2}^2$ it is clear that the supremum in (1.5) is $\leq |\Omega|^{1/p}$. On the other hand, for $\lambda > 0$ let $B_\lambda$ be the open ball in $\mathbb{R}^d$ with center 0 and radius $\lambda$ and denote by $\chi_\lambda$ its characteristic function. For $K \subset \Omega$ compact, set $M_K = \max\{|x| : x \in K\}$.

Then, for $x \in K$ and $\lambda \geq M_K$ we have

$$\langle \chi_\lambda, T_x\chi_\lambda\rangle \geq |B_{\lambda - M_K}|$$

so that, for $\lambda \geq M_K$,

$$\begin{aligned}\frac{\left(\int_\Omega |\langle \chi_\lambda, T_x\chi_\lambda\rangle|^p dx\right)^{1/p}}{\|\chi_\lambda\|_{L^2}^2} &\geq \frac{\left(\int_K |\langle \chi_\lambda, T_x\chi_\lambda\rangle|^p dx\right)^{1/p}}{\|\chi_\lambda\|_{L^2}^2} \\ &\geq \frac{|B_{\lambda-M_K}||K|^{1/p}}{|B_\lambda|},\end{aligned}$$

which implies

$$\liminf_{\lambda\to+\infty} \frac{\left(\int_\Omega |\langle \chi_\lambda, T_x\chi_\lambda\rangle|^p dx\right)^{1/p}}{\|\chi_\lambda\|_{L^2}^2} \geq |K|^{1/p}$$

as $\lambda \to +\infty$. Since $|\Omega \setminus K|$ can be arbitrarily small, (1.5) is proved.

Let us now prove that there is no extremal function. Suppose on the contrary that $f \in L^2(\mathbb{R}^d) \setminus \{0\}$ is such an extremal function, which we can further assume to be normalized in $L^2$: $\|f\|_{L^2} = 1$. Then

$$\int_\Omega |\langle f, T_x f\rangle|^p dx = |\Omega|,$$

which together with the estimate $|\langle f, T_x f\rangle| \leq 1$ implies that

$$|\langle f, T_x f\rangle| = 1$$

for almost every $x \in \Omega$. Hence, since $|\Omega| > 0$, there exists $x_0 \in \mathbb{R}^d$, $x_0 \neq 0$, $c \in \mathbb{C}$, $|c| = 1$, such that

$$T_{x_0} f = cf.$$

Taking the Fourier transform we obtain that $f = 0$, hence a contradiction. □

## 5. Optimization with fixed window

To put our main result into context, we now mention the problem of the optimization of the cross-ambiguity when one of the arguments is kept fixed (or, equivalently, the optimization of the short-time Fourier transform with a fixed window). As we show below, the existence of optimizers is in this case much easier to prove — while the characterization of such extremizers with, for example, the Gaussian window, is a challenging subject [39].

**Proposition 5.1.** *Let $g \in L^2(\mathbb{R}^d) \setminus \{0\}$ and $\Omega \subset \mathbb{R}^{2d}$ be a measurable subset of finite, positive measure. Let $1 \leq p < \infty$. Then the supremum*

$$\sup_{f \in L^2(\mathbb{R}^d)\setminus\{0\}} \frac{\left(\int_\Omega |A(f,g)(x,\omega)|^p dx d\omega\right)^{1/p}}{\|f\|_{L^2}} \tag{5.1}$$

*is attained. Moreover, any maximizing sequence that is normalized in $L^2(\mathbb{R}^d)$ has a subsequence that converges in $L^2$ to a maximizer.*

*Proof of Proposition 5.1.* We claim that the functional $f \mapsto \|A(f,g)\|_{L^p(\Omega)}$ is sequentially weakly continuous on $L^2(\mathbb{R}^d)$. Indeed, if $f^{(n)}$ converges weakly to $f \in L^2(\mathbb{R}^d)$, it follows at once from the definition of the cross-ambiguity function that $A(f^{(n)}, g) \to A(f,g)$ pointwise in $\mathbb{R}^{2d}$, and moreover $|A(f^{(n)}, g)| \leq \|f^{(n)}\|_{L^2}\|g\|_{L^2} \lesssim 1$ on $\mathbb{R}^{2d}$, so that the claim follows from the dominated convergence theorem.

Let $f^{(n)}$ be a maximizing sequence with $\|f^{(n)}\|_{L^2} = 1$, and let $L$ be the supremum in (5.1). Since $|\Omega| > 0$ and $g \not\equiv 0$, it follows that $L > 0$. Indeed, it is sufficient to consider a point $z_0 \in \Omega$ of positive Lebesgue density for $\Omega$ and observe that the function $A(\pi(z_0)g, g)$ is continuous and $|A(\pi(z_0)g, g)(z_0)| = \|g\|_{L^2}^2 > 0$.

Then $f^{(n)}$ has a subsequence, that we still denote by $f^{(n)}$, weakly convergent to some $f \in L^2(\mathbb{R}^d)$, and, by the above mentioned sequential weak continuity, $\|A(f,g)\|_{L^p(\Omega)} = L$. Since $L > 0$, $f \neq 0$. In addition, $\|f\|_{L^2} \leq \liminf_{n\to\infty} \|f^{(n)}\|_{L^2} \leq 1$, so that $f$ is a maximizer and $\|f\|_{L^2} = 1$. As a consequence, $f^{(n)} \to f$ in $L^2$. □

**Remark 5.2.** *For $p = 2$, the existence of a maximizer for the problem* (5.1) *also follows from the spectral properties of the non-negative bounded operator $V_g^* \chi_\Omega V_g$ on $L^2(\mathbb{R}^d)$. Indeed, $|A(f,g)| = |V_g f|$, so that*

$$\int_\Omega |A(f,g)(x,\omega)|^2 \, dx d\omega = \langle V_g^* \chi_\Omega V_g f, f \rangle.$$

*Since $|\Omega| < \infty$, the operator $V_g^* \chi_\Omega V_g$ is compact (in fact, trace class* [9, 50]*), so that any eigenfunction corresponding to the maximum eigenvalue is a maximizer for the problem* (5.1) *(with $p = 2$).*

We emphasize that for the optimization problem in Theorem 1.1 we could not have argued as in the proof of Proposition 5.1, because of the lack of sequential weak upper semicontinuity of the corresponding functional, as shown below.

**Proposition 5.3.** *Let $\Omega \subset \mathbb{R}^{2d}$ be a measurable subset of finite, positive measure. The functional $\|A(f)\|_{L^2(\Omega)}$ on $L^2(\mathbb{R}^d)$ is not sequentially weakly upper semicontinuous at any point.*

*Proof.* The computations in the proof of Proposition 1.2 show that there exists $g \in L^2(\mathbb{R}^d)$ - in fact, a Gaussian function - such that $\|A(g)\|_{L^2(\Omega)} > 0$, since $|\Omega| > 0$.

Let now $f \in L^2(\mathbb{R}^d)$. Then $f + \pi(z)g$ converges weakly to $f$ as $|z| \to +\infty$. On the other hand,

$$\|A(f + \pi(z)g)\|^2_{L^2(\Omega)} = \|A(f) + A(f, \pi(z)g) + A(\pi(z)g, f) + M_{Jz}A(g)\|^2_{L^2(\Omega)},$$

where we used (3.10).

Since $A(f, \pi(z)g) + A(\pi(z)g, f) \to 0$ in $L^2(\Omega)$ as $|z| \to +\infty$, by arguing as in (3.11) we obtain

$$\lim_{|z|\to+\infty} \|A(f + \pi(z)g)\|^2_{L^2(\Omega)} = \|A(f)\|^2_{L^2(\Omega)} + \|A(g)\|^2_{L^2(\Omega)} > \|A(f)\|^2_{L^2(\Omega)},$$

which gives the desired conclusion. □

## 6. Variations on the main result

6.1. **The optimization problem in modulation spaces.** We now derive a variant of Theorem 1.1, where the function is optimized over the modulation space $M^q(\mathbb{R}^d)$, $0 < q < 2$. For the precise formulation, fix a window function $g \in \mathcal{S}(\mathbb{R}^d) \setminus \{0\}$ and $0 < q \leq \infty$; then $M^q(\mathbb{R}^d)$ is defined as the space of temperate distributions $f \in \mathcal{S}'(\mathbb{R}^d)$ such that

$$\|f\|_{M^q} := \|V_g f\|_{L^q(\mathbb{R}^{2d})} < \infty.$$

Different windows $g$ give rise to the same space with equivalent norms. Moreover, $M^2(\mathbb{R}^d) = L^2(\mathbb{R}^d)$ with equivalent norms, and $M^{q_1}(\mathbb{R}^d) \hookrightarrow M^{q_2}(\mathbb{R}^d)$ if and only if $0 < q_1 \leq q_2 \leq \infty$; see [3, 24] and [25, Chapter 10] for background.

Thus, a modulation-space norm estimate $\|f\|_{M^q} \leq 1$ prescribes a certain integrability and decay for a function $f$. The next result allows one to incorporate such constraints into the optimization of the ambiguity function.

**Theorem 6.1.** *Let $\Omega \subset \mathbb{R}^{2d}$ be a measurable subset of finite, positive measure, and $1 \leq p < \infty$, $0 < q < 2$. Then the supremum*

$$\sup_{f \in M^q(\mathbb{R}^d)\setminus\{0\}} \frac{\left(\int_\Omega |A(f)(x,\omega)|^p dx d\omega\right)^{1/p}}{\|f\|^2_{M^q}} \tag{6.1}$$

*is attained. Moreover if $f^{(n)}$ is any maximizing sequence normalized in $M^q(\mathbb{R}^d)$, then there exists a subsequence (still denoted by $f^{(n)}$) and $z^{(n)} \in \mathbb{R}^{2d}$ such that $\pi(-z^{(n)})f^{(n)}$ converges in $M^q$ to a maximizer.*

*Proof of Theorem 6.1.* **Step 1**. *Profile decomposition in $M^q$.*

The first part of the proof is similar to (in fact, simpler than) the one of Theorem 1.1 and it will only be sketched. Let $L$ be the supremum in (6.1); as in the proof of Theorem 1.1, we note that $L > 0$ since $|\Omega| > 0$.

Consider a maximizing sequence $f^{(n)}$, now normalized in $M^q$: $\|f^{(n)}\|_{M^q} = 1$. Since $0 < q < 2$, we have $M^q \hookrightarrow L^2$, so that the sequence $f^{(n)}$ is bounded in $L^2$ and we can apply (after passing to a suitable subsequence) the profile decomposition in $L^2$ as described in Section 2.3, albeit with minor modifications; cf. [46, Theorem 4.5.3] or [48, Theorem 3.1 and its proof]. The formulas (2.9), (2.10), (2.12) and (2.14) hold, whereas (2.11) is now replaced by

$$\sum_{j=1}^{k} \|f_j\|_{L^2}^2 + \limsup_{n\to\infty} \|w_k^{(n)}\|_{L^2}^2 \le C$$

for some $C > 0$, since $\limsup_{n\to\infty} \|f^{(n)}\|_{L^2}$ is still finite but no longer necessarily $\le 1$. While this is sufficient to prove (3.15), (3.17), (3.18), (3.19), Step 4 of the proof of Theorem 1.1 requires some modifications. To complete the proof, we will prove that the profiles $f_j$ are not merely in $L^2(\mathbb{R}^d)$ but actually belong to $M^q(\mathbb{R}^d)$, and, moreover, satisfy the following precise norm estimate:

$$\sum_{j=1}^{\infty} \|f_j\|_{M^q}^q \le 1. \tag{6.2}$$

Postponing the proof of this fact, let us see how to deduce the existence of optimizers. We start from the expansion (3.9) for $A(f^{(n)})$. By the triangle inequality and (3.10), (3.15), (3.17), (3.18), (3.19) we obtain

$$L = \lim_{n\to\infty} \|A(f^{(n)})\|_{L^p(\Omega)} \le \sum_{j=1}^{\infty} \|A(f_j)\|_{L^p(\Omega)}.$$

By the definition of $L$,

$$\|A(f_j)\|_{L^p(\Omega)} \le L\|f_j\|_{M^q}^2,$$

and, since $q < 2$,

$$\begin{aligned} L = \lim_{n\to\infty} \|A(f^{(n)})\|_{L^p(\Omega)} \le \sum_{j=1}^{\infty} \|A(f_j)\|_{L^p(\Omega)} &\le L \sum_{j=1}^{\infty} \|f_j\|_{M^q}^2 \\ &\le L \left( \sum_{j=1}^{\infty} \|f_j\|_{M^q}^q \right)^{2/q} \le L. \end{aligned}$$

This implies that all $f_j$ are zero except one, say $f_1$, and $\|f_1\|_{M^q} = 1$. Hence $f_1$ is a maximizer.

Finally, since, by Remark 2.1, $\pi\big(z_1^{(n)}\big)^* f^{(n)}$ converges weakly (in $L^2$) to $f_1$, it turns out that $V_g(\pi\big(z_1^{(n)}\big)^* f^{(n)}) \to V_g f_1$ pointwise. Moreover $\|V_g(\pi\big(z_1^{(n)}\big)^* f^{(n)})\|_{L^q} = \|f^{(n)}\|_{M^q} = 1 = \|f_1\|_{M^q} = \|V_g f_1\|_{L^q}$, so that $V_g(\pi\big(z_1^{(n)}\big)^* f^{(n)})$ tends to $V_g f_1$ in $L^q$ by the Brézis-Lieb Lemma [5, 32], i.e. $\pi\big(z_1^{(n)}\big)^* f^{(n)} \to f_1$ in $M^q$. We now invoke (2.2) and eliminate the cocycles as in the proof of Theorem 1.1.

**Step 2**. *Precise norm estimate for the profiles.*

We now prove (6.2). As noted in Remark 2.1, each $f_j$ is indeed the weak limit (in $L^2$) of (adjoint) time-frequency shifts of $f^{(n)}$, which are assumed to be normalized in $M^q(\mathbb{R}^d)$. Since $V_g f_j$ is then the pointwise limit of the corresponding short-time Fourier transforms, we see that $f_j \in M^q(\mathbb{R}^d)$ by Fatou's lemma.

Moreover, (2.13) implies that $V_g(\pi(-z_j^{(n)})w_k^{(n)})(z) = V_g(w_k^{(n)})(z + z_j^{(n)})$ tends to zero uniformly on compact subsets of $\mathbb{R}^{2d}$ as $n \to \infty$ — due to the strong continuity of time-frequency shifts.

Suppose first that $1 \le q < 2$. For fixed $k \ge 1$, given $\epsilon > 0$ there exist therefore compact subsets $K_j \subset \mathbb{R}^{2d}$, $j = 1, \dots, k$, and $n_k \in \mathbb{N}$ such that ($L^q$ standing for $L^q(\mathbb{R}^{2d})$)

$$\|V_g(\pi(z_j^{(n)})f_j)\chi_{\mathbb{R}^{2d}\setminus(z_j^{(n)}+K_j)}\|_{L^q} = \|V_g(f_j)\chi_{\mathbb{R}^{2d}\setminus K_j}\|_{L^q} < \epsilon$$

and

$$\sum_{j=1}^{k} \|V_g(w_k^{(n)})\chi_{z_j^{(n)}+K_j}\|_{L^q} < \epsilon$$

for $n \ge n_k$. For each such $n$, by (2.9) and the triangle inequality,

$$\begin{aligned} 1 &= \|V_g f^{(n)}\|_{L^q} \\ &\ge \Big\| \sum_{j=1}^{k} V_g(\pi(z_j^{(n)})f_j)\chi_{z_j^{(n)}+K_j} + V_g(w_k^{(n)})\chi_{\mathbb{R}^{2d}\setminus\cup_{j=1}^{k}(z_j^{(n)}+K_j)} \Big\|_{L^q} - (k+1)\epsilon. \end{aligned}$$

On the other hand, by (2.10), if $n$ is large enough the compact subsets $z_j^{(n)} + K_j$, $j = 1, \dots, k$, are pairwise disjoint (in the last summation we can consider just the indices $j$ such that $f_j \neq 0$), so that

$$\begin{aligned} \Big\| \sum_{j=1}^{k} V_g(\pi(z_j^{(n)})f_j)\chi_{z_j^{(n)}+K_j} + V_g(w_k^{(n)})\chi_{\mathbb{R}^{2d}\setminus\cup_{j=1}^{k}(z_j^{(n)}+K_j)} \Big\|_{L^q}^q \\ &\ge \sum_{j=1}^{k} \|V_g(\pi(z_j^{(n)})f_j)\chi_{z_j^{(n)}+K_j}\|_{L^q}^q \\ &= \sum_{j=1}^{k} \|V_g(f_j)\chi_{K_j}\|_{L^q}^q \\ &\ge \sum_{j=1}^{k} (\|f_j\|_{M^q} - \epsilon)_+^q, \end{aligned}$$

where $(\cdot)_+$ denotes the positive part function.

In conclusion we have

$$\sum_{j=1}^{k} (\|f_j\|_{M^q} - \epsilon)_+^q \le (1 + (k+1)\epsilon)^q.$$

Since $\epsilon$ and $k$ are arbitrary, we have proved (6.2) in the case $1 \le q < 2$.

The argument needs to be slightly adapted for $0 < q < 1$. In this case we choose the compact subsets $K_j \subset \mathbb{R}^{2d}$, $j = 1, \dots, k$, and $n_k \in \mathbb{N}$ so that

$$\|V_g(\pi(z_j^{(n)})f_j)\chi_{\mathbb{R}^{2d}\setminus(z_j^{(n)}+K_j)}\|_{L^q}^q = \|V_g(f_j)\chi_{\mathbb{R}^{2d}\setminus K_j}\|_{L^q}^q < \epsilon$$

and

$$\sum_{j=1}^{k} \|V_g(w_k^{(n)})\chi_{z_j^{(n)}+K_j}\|_{L^q}^q < \epsilon$$

for $n \geq n_k$. Again, by (2.9) and the triangle inequality, now for $\|\cdot\|_{L^q}^q$,

$$\begin{aligned} 1 &= \|V_g f^{(n)}\|_{L^q}^q \\ &\geq \Big\| \sum_{j=1}^{k} V_g(\pi(z_j^{(n)})f_j)\chi_{z_j^{(n)}+K_j} + V_g(w_k^{(n)})\chi_{\mathbb{R}^{2d}\setminus\cup_{j=1}^{k}(z_j^{(n)}+K_j)} \Big\|_{L^q}^q - (k+1)\epsilon. \end{aligned}$$

An argument similar to that used in the previous case now gives

$$\sum_{j=1}^{k} (\|f_j\|_{M^q}^q - \epsilon) \leq 1 + (k+1)\epsilon,$$

which implies (6.2) also for $0 < q < 1$. □

6.2. **Optimization with respect to Gabor systems.** While the constraint $f \in M^q(\mathbb{R}^d)$ in Theorem 6.1 is independent of the choice of the window $g \in \mathcal{S}(\mathbb{R}^d)$, the functional optimized in (6.1) does depend on $g$ because it involves the window-dependent (quasi-)norm $\|f\|_{M^q}$. In practice, such norms are often replaced by certain discrete counterparts computed in terms of so-called *Gabor systems*.

Precisely, consider a full-rank lattice $\Lambda \subset \mathbb{R}^{2d}$ and $g \in \mathcal{S}(\mathbb{R}^d)$ such that the set of functions $\{\pi(\lambda)g\}_{\lambda\in\Lambda}$ is a *frame for* $L^2(\mathbb{R}^d)$, i.e.,

$$\|f\|_{L^2}^2 \lesssim \sum_{\lambda\in\Lambda} |\langle f, \pi(\lambda)g\rangle|^2 \lesssim \|f\|_{L^2}^2.$$

Then it turns out that the quantity

$$|f|_{M^q} := \left( \sum_{\lambda\in\Lambda} |\langle f, \pi(\lambda)g\rangle|^q \right)^{1/q} \tag{6.3}$$

(with obvious changes if $q = \infty$) gives an equivalent (quasi-)norm in $M^q(\mathbb{R}^d)$, $0 < q \leq \infty$ [3, 24], [25, Chapter 10]. The next result is an analog of Theorem 6.1 for the discrete (quasi-)norm (6.3).

**Theorem 6.2.** *The statement in Theorem 6.1 is still valid if* (6.1) *is replaced by*

$$\sup_{f\in M^q(\mathbb{R}^d)\setminus\{0\}} \frac{\left(\int_\Omega |A(f)(x,\omega)|^p dx d\omega\right)^{1/p}}{|f|_{M^q}^2}.$$

The derivation of Theorem 6.2 requires minimal adaptations. Indeed, the map $\Lambda \ni \lambda \to \pi(\lambda)$ is still a projective unitary representation on $L^2(\mathbb{R}^d)$, and the corresponding operators $\{\pi(\lambda)\}_{\lambda\in\Lambda}$ still define a dislocation set. The corresponding notion of $D$-weak convergence reads

$$\sup_{\lambda\in\Lambda} |\langle f_n - f, \pi(\lambda)h\rangle| \to 0$$

for every $h \in L^2(\mathbb{R}^d)$, and still implies convergence in $M^\infty$, due to the equivalence of the $|\cdot|_{M^\infty}$ and $\|\cdot\|_{M^\infty}$ norms. Thus, profile decompositions as in Section 2.3 exist, now with $z_j^{(n)} \in \Lambda$. The proof of Theorem 6.1 adapts almost verbatim — even

in notation, by replacing the $L^q$ (quasi-)norm in $\mathbb{R}^{2d}$ with respect to the Lebesgue measure by the $L^q$ (quasi-)pseudo-norm in $\mathbb{R}^{2d}$ with respect to the Radon measure $\sum_{\lambda\in\Lambda}\delta_\lambda$. The key point is that such a measure is invariant under the translations $z \mapsto z + z_j^{(n)}$, because $z_j^{(n)} \in \Lambda$.

## Acknowledgments

The authors are very grateful to Karlheinz Gröchenig for bringing to their attention the problem solved here, in connection to an unpublished manuscript of his and Markus Neuhauser.

The present research has been partially supported by the MIUR grant Dipartimenti di Eccellenza 2018-2022, CUP: E11G18000350001, DISMA, Politecnico di Torino. J. L. R. gratefully acknowledges support from the Austrian Science Fund (FWF): Y 1199.

S. I. T. is member of the Machine Learning Genoa (MaLGa) Center, Università di Genova. F. N. and S. I. T. are members of the Gruppo Nazionale per l'Analisi Matematica, la Probabilità e le loro Applicazioni (GNAMPA) of the Istituto Nazionale di Alta Matematica (INdAM).

## Statements and Declarations

The authors declare no competing interests. Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

## References

[1] V. Bargmann. Irreducible unitary representations of the Lorentz group. *Ann. of Math. (2)*, 48:568–640, 1947.

[2] J. J. Benedetto, R. L. Benedetto, and J. T. Woodworth. Optimal ambiguity functions and Weil's exponential sum bound. *J. Fourier Anal. Appl.*, 18(3):471–487, 2012.

[3] Á. Bényi and K. A. Okoudjou. *Modulation Spaces: With Applications to Pseudodifferential Operators and Nonlinear Schrödinger Equations*. Applied and Numerical Harmonic Analysis. Birkhäuser Basel, 2020.

[4] J. Bergh and J. Löfström. *Interpolation Spaces. An Introduction*. Grundlehren der Mathematischen Wissenschaften, No. 223. Springer-Verlag, Berlin-New York, 1976.

[5] H. Brézis and E. Lieb. A relation between pointwise convergence of functions and convergence of functionals. *Proc. Amer. Math. Soc.*, 88(3):486–490, 1983.

[6] H. Brézis and L. Nirenberg. Positive solutions of nonlinear elliptic equations involving critical Sobolev exponents. *Comm. Pure Appl. Math.*, 36(4):437–477, 1983.

[7] L. Cohen. *Time-Frequency Analysis*. Prentice-Hall, New York, 1995.

[8] C. E. Cook and M. Bernfeld. *Radar Signals - An Introduction to Theory and Applications*. Academic Press, New York, 1967.

[9] E. Cordero and K. Gröchenig. Time-frequency analysis of localization operators. *J. Funct. Anal.*, 205(1):107–131, 2003.

[10] E. Cordero and F. Nicola. Sharp integral bounds for Wigner distributions. *Int. Math. Res. Not. IMRN*, (6):1779–1807, 2018.

[11] M. Cowling. The Kunze-Stein phenomenon. *Ann. of Math. (2)*, 107(2):209–234, 1978.

[12] I. Daubechies. Time-frequency localization operators: a geometric phase space approach. *IEEE Trans. Inform. Theory*, 34(4):605–612, 1988.

[13] M. A. de Gosson. *Symplectic Methods in Harmonic Analysis and in Mathematical Physics*, volume 7. Birkhäuser/Springer Basel AG, Basel, 2011.

[14] L. Ehrenpreis and F. Mautner. Uniformly bounded representations of groups. *Proc. Nat. Acad. Sci. U.S.A.*, 41:231–233, 1955.

[15] C. L. Fefferman. The uncertainty principle. *Bull. Amer. Math. Soc. (N.S.)*, 9(2):129–206, 1983.

[16] H. G. Feichtinger, D. Onchis-Moaca, B. Ricaud, B. Torrésani, and C. Wiesmeyr. A method for optimizing the ambiguity function concentration. In *2012 Proceedings of the 20th European Signal Processing Conference (EUSIPCO)*, pages 804–808, 2012.

[17] P. Flandrin. Maximum signal energy concentration in a time-frequency domain. In *ICASSP-88., International Conference on Acoustics, Speech, and Signal Processing*, pages 2176–2179 vol.4, 1988.

[18] P. Flandrin. *Time-Frequency/Time-Scale Analysis*, volume 10 of *Wavelet Analysis and its Applications*. Academic Press, Inc., San Diego, CA, 1999. With a preface by Yves Meyer, Translated from the French by Joachim Stöckler.

[19] P. Flandrin. *Explorations in Time-Frequency Analysis*. Cambridge University Press, 2018.

[20] G. B. Folland. *Harmonic Analysis in Phase Space*, volume 122 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 1989.

[21] G. B. Folland and A. Sitaram. The uncertainty principle: a mathematical survey. *J. Fourier Anal. Appl.*, 3(3):207–238, 1997.

[22] W. H. J. Fuchs. On the magnitude of fourier transforms. In *Proc. Intern. Congress Math.*, volume II, pages 106–107. North-Holland, Amsterdam, 1954.

[23] D. Gabor. Theory of communication. *J. IEE*, 93(III):429–457, 1946.

[24] Y. V. Galperin and S. Samarah. Time-frequency analysis on modulation spaces $M_m^{p,q}$, $0 < p,\ q \leq \infty$. *Appl. Comput. Harmon. Anal.*, 16(1):1–18, 2004.

[25] K. Gröchenig. *Foundations of Time-Frequency Analysis*. Birkhäuser Boston, Inc., Boston, MA, 2001.

[26] A. W. Knapp. *Representation theory of semisimple groups*. Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 2001. An overview based on examples, Reprint of the 1986 original.

[27] R. A. Kunze and E. M. Stein. Uniformly bounded representations and harmonic analysis of the $2 \times 2$ real unimodular group. *Amer. J. Math.*, 82:1–62, 1960.

[28] H. J. Landau. An overview of time and frequency limiting. In *Fourier techniques and applications (Kensington, 1983)*, pages 201–220. Plenum, New York, 1985.

[29] H. J. Landau and H. O. Pollak. Prolate spheroidal wave functions, Fourier analysis and uncertainty. II. *Bell System Tech. J.*, 40:65–84, 1961.

[30] N. Lerner. Integrating the wigner distribution on subsets of the phase space, a survey, arXiv:2102.08090, 2021.

[31] E. H. Lieb. Integral bounds for radar ambiguity functions and Wigner distributions. *J. Math. Phys.*, 31(3):594–599, 1990.

[32] E. H. Lieb and M. Loss. *Analysis*. American Mathematical Society, Providence, RI, 1997.

[33] E. H. Lieb and Y. Ostrover. Localization of multidimensional Wigner distributions. *J. Math. Phys.*, 51(10):102101, 6, 2010.

[34] P.-L. Lions. The concentration-compactness principle in the calculus of variations. The limit case. I. *Rev. Mat. Iberoamericana*, 1(1):145–201, 1985.

[35] P.-L. Lions. The concentration-compactness principle in the calculus of variations. The limit case. II. *Rev. Mat. Iberoamericana*, 1(2):45–121, 1985.

[36] S. Mallat. *A Wavelet Your of Signal Processing - The Sparse Way*. Elsevier/Academic Press, Amsterdam, 2009.

[37] G. Matz, H. Bölcskei, and F. Hlawatsch. Time-frequency foundations of communications: Concepts and tools. *IEEE Signal Processing Magazine*, 30(6):87–96, 2013.

[38] G. Matz, D. Schafhuber, K. Gröchenig, M. Hartmann, and F. Hlawatsch. Analysis, optimization, and implementation of low-interference wireless multicarrier systems. *IEEE Transactions on Wireless Communications*, 6(5):1921–1931, 2007.

[39] F. Nicola and P. Tilli. The faber-krahn inequality for the short-time fourier transform. *Invent. Math.* (in press), DOI: 10.1007/s00222-022-01119-8, 2022.

[40] B. Ricaud, G. Stempfel, B. Torrésani, C. Wiesmeyr, H. Lachambre, and D. Onchis. An optimally concentrated Gabor transform for localized time-frequency components. *Adv. Comput. Math.*, 40(3):683–702, 2014.
[41] B. Ricaud and B. Torrésani. A survey of uncertainty principles and some signal processing applications. *Adv. Comput. Math.*, 40(3):629–650, 2014.
[42] A. W. Rihaczek. *Principles of High-Resolution Radar*. Artech House, Boston, 1996.
[43] J. Sacks and K. Uhlenbeck. The existence of minimal immersions of 2-spheres. *Ann. of Math. (2)*, 113(1):1–24, 1981.
[44] D. Slepian. Some comments on Fourier analysis, uncertainty and modeling. *SIAM Rev.*, 25(3):379–393, 1983.
[45] M. Struwe. A global compactness result for elliptic boundary value problems involving limiting nonlinearities. *Math. Z.*, 187(4):511–517, 1984.
[46] T. Tao. *Compactness and Contradiction*. American Mathematical Society, Providence, RI, 2013.
[47] T. Tao, A. Vargas, and L. Vega. A bilinear approach to the restriction and Kakeya conjectures. *J. Amer. Math. Soc.*, 11(4):967–1000, 1998.
[48] K. Tintarev and K.-H. Fieseler. *Concentration Compactness*. Imperial College Press, London, 2007.
[49] M. Vetterli, J. Kovačević, and V. K. Goyal. *Foundations of Signal Processing*. Cambridge University Press, 2014.
[50] M. W. Wong. *Wavelet Transforms and Localization Operators*. Birkhäuser Verlag, Basel, 2002.
[51] P. M. Woodward. *Probability and Information theory, with Applications to Radar*. Pergamon Press, Oxford-Edinburgh-New York-Paris-Frankfurt, 1964.

Dipartimento di Scienze Matematiche, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy

*Email address*: fabio.nicola@polito.it

Faculty of Mathematics, University of Vienna, Oskar-Morgenstern-Platz 1, A-1090 Vienna, Austria, and, Acoustics Research Institute, Austrian Academy of Sciences, Wohllebengasse 12-14 A-1040, Vienna, Austria

*Email address*: jose.luis.romero@univie.ac.at

MaLGa Center - Department of Mathematics, University of Genova, via Dodecaneso 35, 16146 Genova, Italy

*Email address*: salvatoreivan.trapasso@unige.it