POLITECNICO DI TORINO
Repository ISTITUZIONALE

On the existence of optimizers for nonlinear time-frequency concentration problems: the Wigner distribution

Original
On the existence of optimizers for nonlinear time-frequency concentration problems: the Wigner distribution / Stra, F., Svela, E., Trapasso, S.I.. - In: JOURNAL DE MATHÉMATIQUES PURES ET APPLIQUÉES. - ISSN 0021-7824. - 212:(2026), pp. 1-29. [10.1016/j.matpur.2026.103919]

Availability:
This version is available at: 11583/3012091 since: 2026-06-15T22:28:41Z

Publisher:
Elsevier

Published
DOI:10.1016/j.matpur.2026.103919

Terms of use:

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

Publisher copyright

(Article begins on next page)

13 September 2026

J. Math. Pures Appl. 212 (2026) 103919

Contents lists available at ScienceDirect

Journal de Mathématiques Pures et Appliquées

journal homepage: www.elsevier.com/locate/matpur

On the existence of optimizers for nonlinear time-frequency concentration problems: the Wigner distribution

Federico Stra [a], Erling Svela [b,*], S. Ivan Trapasso [a]

[a] *Politecnico di Torino, Dipartimento di Scienze Matematiche "G. L. Lagrange", corso Duca degli Abruzzi 24, 10129 Torino, Italy*
[b] *Department of Mathematical Sciences, Norwegian University of Science and Technology, 7491 Trondheim, Norway*

ARTICLE INFO

Article history:
Received 14 November 2025
Available online 30 April 2026

MSC:
49Q10
49R05
42B10
94A12
81S30

Keywords:
Nonlinear time-frequency concentration
Wigner distribution
Born–Jordan distribution
Concentration compactness

ABSTRACT

We prove that, for any measurable phase space subset $\Omega \subset \mathbb{R}^{2d}$ with $0 < |\Omega| < \infty$ and any $1 \leq p < \infty$, the nonlinear concentration problem

$$\sup_{f \in L^2(\mathbb{R}^d)\setminus\{0\}} \frac{\|Wf\|_{L^p(\Omega)}}{\|f\|_{L^2}^2}$$

admits an optimizer, where Wf is the Wigner distribution of f. The main obstruction is that Wf is covariant (not invariant) under time-frequency shifts, which impedes weak upper semicontinuity, so the effects of constructive interference must be taken into account. We close this compactness gap via concentration compactness for Heisenberg-type dislocations, together with a new asymptotic formula that quantifies the limiting contribution to concentration over Ω from asymptotically separated wave packets. When $p = \infty$ we also identify the sharp constant 2^d and show that it is attained.
We also discuss some related extensions: for τ-Wigner distributions with $\tau \in (0, 1)$ we isolate a chain phenomenon that obstructs the same strategy beyond the Wigner case ($\tau = 1/2$), while for the Born–Jordan distribution in $d = 1$ we obtain weak continuity, and thus existence of concentration optimizers for all $1 \leq p < \infty$ (the $p = \infty$ supremum equals π but is not attained).

© 2026 The Authors. Published by Elsevier Masson SAS. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

RÉSUMÉ

Nous démontrons que, pour tout sous-ensemble mesurable de l'espace de phase $\Omega \subset \mathbb{R}^{2d}$ tel que $0 < |\Omega| < \infty$ et tout $1 \leq p < \infty$, le problème non linéaire de concentration temps-fréquence

$$\sup_{f \in L^2(\mathbb{R}^d)\setminus\{0\}} \frac{\|Wf\|_{L^p(\Omega)}}{\|f\|_{L^2}^2}$$

* Corresponding author.
E-mail addresses: federico.stra@polito.it (F. Stra), erling.a.t.svela@ntnu.no (E. Svela), salvatoreivan.trapasso@polito.it (S. Ivan Trapasso).

https://doi.org/10.1016/j.matpur.2026.103919
0021-7824/© 2026 The Authors. Published by Elsevier Masson SAS. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

admet une fonction extrémale, où Wf désigne la distribution de Wigner de f. La difficulté principale tient au fait que Wf est covariante, et non invariante, par translations temps-fréquence, ce qui empêche la semi-continuité supérieure pour la topologie faible ; il faut donc prendre en compte les effets d'interférence constructive. Nous surmontons ce défaut de compacité en combinant une méthode de concentration-compacité adaptée aux dislocations de type Heisenberg et une nouvelle formule asymptotique qui quantifie, dans la limite, la contribution de paquets d'ondes asymptotiquement séparés à la concentration sur Ω. Dans le cas $p = \infty$, nous déterminons en outre la meilleure constante, égale à 2^d, et montrons qu'elle est atteinte.

Nous discutons également quelques prolongements de ces résultats. Pour les distributions de τ-Wigner, avec $\tau \in (0, 1)$, nous mettons en évidence un phénomène en chaîne qui empêche la mise en oeuvre de la même stratégie au-delà du cas de Wigner ($\tau = 1/2$). En revanche, pour la distribution de Born–Jordan en dimension $d = 1$, nous établissons la continuité faible et en déduisons l'existence de fonctions extrémales pour tout $1 \leq p < \infty$ (la borne supérieure pour $p = \infty$ vaut π, sans être atteinte).

© 2026 The Authors. Published by Elsevier Masson SAS. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Though approaching its centennial, the Wigner distribution is still one of the most popular time-frequency representations in signal analysis, not to mention its prominent role in many problems and applications of quantum mechanics. Introduced somewhat enigmatically by Wigner in 1932 as a quasi-probability distribution on phase space to account for quantum corrections to classical statistical mechanics [36], it was later rediscovered by Ville in 1948 [35]. Since then, it has become a widely used time-frequency method in signal processing due to its robustness and desirable features.

To be specific, recall that the Wigner distribution of $f \in L^2(\mathbb{R}^d)$ is defined by

$$Wf(z) = \int_{\mathbb{R}^d} e^{-2\pi i \xi \cdot y} f\Big(x + \frac{y}{2}\Big) \overline{f\Big(x - \frac{y}{2}\Big)} \, \mathrm{d}y, \qquad z = (x, \xi) \in \mathbb{R}^{2d}.$$

Therefore, Wf coincides with the Fourier transform of the two-point autocorrelation of f. Informally, we may also view $Wf(x, \xi)$ as the joint time-frequency energy density of f localized near the phase space point (x, ξ). We note that the map $f \mapsto Wf$ is inherently quadratic: For every $\alpha, \beta \in \mathbb{C}$ and $f, g \in L^2(\mathbb{R}^d)$ we have

$$W(\alpha f + \beta g) = |\alpha|^2 Wf + \alpha\bar{\beta} W(f, g) + \bar{\alpha}\beta W(g, f) + |\beta|^2 Wg,$$

where $W(f, g)$ and $W(g, f)$ denote cross-Wigner distributions, cf. (2) below. The occurrence of these cross-interference terms, along with self-interference fringes, is a classical drawback in signal processing: Even for well-localized signals (such as Gaussian functions) the joint time-frequency energy exhibits non-trivial scattering phenomena in phase space [1,3,4]. While the uncertainty principle prevents Wf from being localized within a domain $\Omega \subset \mathbb{R}^{2d}$ of arbitrarily small measure [18], it is natural to investigate whether optimal phase space concentration on a given domain can be achieved despite these interferences and obstructions — a problem of both theoretical and practical interest, see for instance [24,31,14,2,21] and the references therein.

There are several possible choices for how to measure phase space concentration. A convenient one is the L^2 norm of Wf over Ω, which has to be interpreted as the fraction of time-frequency energy trapped within Ω. More generally, one can use the L^p norms with $p \neq 2$, which can be rightfully viewed as favoring

spreading or concentration/sparsity depending on the range of p. Globally, this heuristic point of view is reinforced by the L^p-norm inequalities due to Lieb [23]:

$$\begin{cases} \|Wf\|_{L^p} \geq C_{p,d}\|f\|_{L^2}^2 & (1 \leq p \leq 2) \\ \|Wf\|_{L^p} \leq C_{p,d}\|f\|_{L^2}^2 & (p \geq 2), \end{cases} \qquad C_{p,d} = \Big(\frac{2^{p-1}}{p}\Big)^{d/p}.$$

Nevertheless, when it comes to local versions, it has remained open whether L^p concentration can be optimized over a given domain of finite, positive measure. This is precisely the problem addressed in this note, and our main result settles it in the affirmative.

Theorem 1.1. *Let $\Omega \subset \mathbb{R}^{2d}$ be measurable with $0 < |\Omega| < \infty$ and let $p \in [1, \infty)$. The supremum*

$$L = \sup_{f \in L^2(\mathbb{R}^d)\setminus\{0\}} \frac{\big(\int_\Omega |Wf(z)|^p \,\mathrm{d}z\big)^{1/p}}{\|f\|_{L^2}^2} \tag{1}$$

is attained.

The same conclusion holds in the case where $p = \infty$, the proof being considerably easier — see Section 5 below, where we also find the optimal value $L = 2^d$.

Let us highlight that in the linear case (that is, optimizing the integral $\int_\Omega Wf(z)\,\mathrm{d}z$ — see the recent monograph [22] for interesting related problems) a minimax argument reduces the concentration problem to the analysis of a certain compact operator, see for instance [14,24,33]. On the other hand, the nonlinear nature of the problem (1) precludes such linear techniques and, in spite of its simple formulation, requires sophisticated tools. Indeed, our proof of Theorem 1.1 relies on the theory of *concentration compactness* [34,25,26], a powerful framework especially designed to handle loss of compactness due to invariance of the target functional under the action of a non-compact family of transformations (the so-called *dislocations*). Before giving more details, let us stress that this approach has been successfully used in time-frequency concentration problems only recently: The authors of [30] solved with this approach the exact analogue of the problem in this note for the *ambiguity function*, that is

$$Af(z) = \int_{\mathbb{R}^d} e^{-2\pi i \xi\cdot y} f\Big(y + \frac{x}{2}\Big)\overline{f\Big(y - \frac{x}{2}\Big)}\,\mathrm{d}y, \qquad z = (x,\xi) \in \mathbb{R}^{2d}.$$

While the optimization objective in (1) is closely related to the one considered in [30] (ultimately by symplectic Fourier transform), the Wigner distribution and the ambiguity function behave quite differently in some respects. Concretely, recall that the time-frequency shift of $f \in L^2(\mathbb{R}^d)$ by $z = (x,\xi) \in \mathbb{R}^{2d}$ is defined by

$$\pi(z)f(y) = e^{2\pi i \xi\cdot y} f(y - x).$$

The ambiguity function is completely invariant under time-frequency shifts, that is $|A(\pi(z)f)| = |Af|$, hence the corresponding concentration functional (1) is invariant as well. These remarks clearly motivate resorting to concentration compactness, as the main defect of compactness in this problem comes from the non-compact symmetry $f \mapsto \pi(z)f$, which allows phase-space mass to drift away without changing the objective. To keep track of the energy that might otherwise be lost at infinity we perform a *profile decomposition*: Up to subsequences, a maximizing sequence $f^{(n)}$ can be split into a finite sum of separated packets as

$$f^{(n)} = \sum_{j=1}^{k} \pi(z_j^{(n)}) f_j \ + \ w_k^{(n)}, \qquad k \in \mathbb{N} \setminus \{0\},$$

where $f_1, \dots, f_k$ are fixed *profiles* whose centers $z_j^{(n)}$ separate asymptotically (i.e., $|z_j^{(n)} - z_{j'}^{(n)}| \to \infty$ for $j \neq j'$) and the remainder $w_k^{(n)}$ is small in a suitable sense — see Lemma 2.3 below for a more precise account on the matter.

On the other hand, the behavior of the Wigner transform under time-frequency shift is subtler: it is only *covariant* under time-frequency shifts, that is $|W(\pi(z)f)(w)| = |Wf(w-z)|$ (cf. (4) below). It is therefore not evident whether a concentration compactness strategy could lead to useful conclusions in this context. To shed some light on the issue, we emphasize that the optimization is effectively translation-invariant, since

$$L = \sup_{f \in L^2(\mathbb{R}^d)} \frac{\|Wf\|_{L^p(\Omega)}}{\|f\|_{L^2}^2} = \sup_{z \in \mathbb{R}^{2d}} \sup_{f \in L^2(\mathbb{R}^d)} \frac{\|\, Wf(\cdot - z)\,\|_{L^p(\Omega)}}{\|f\|_{L^2}^2},$$

and that the target functional is invariant under antipodal pairing, that is

$$|W(\pi(z)f, \pi(-z)f)| = |Wf|.$$

Nevertheless, loss of compactness still stems from mass escaping to infinity in phase space as before, but the decisive difference is that, unlike the ambiguity case, Wigner cross-terms may survive in the limit when *midpoints* of profiles remain bounded, producing potentially constructive interference from antipodally shifted profiles. Moreover, as detailed below, this invariance causes the failure of semicontinuity in the optimization problem, which further justifies attempts beyond the direct methods, such as resorting to concentration compactness. A crucial component of our analysis consists in isolating these *surviving pairs* in a profile decomposition and quantifying their limiting contribution to the $L^p(\Omega)$ norm (see Section 3), which is the key extra ingredient (of independent interest) needed to make the concentration compactness approach work for the Wigner concentration problem.

This leaves open the possibility of *exotic optimizers* when surviving pairs, despite being arbitrarily far apart in phase space in the limit, interfere constructively to optimize concentration on Ω. Although we cannot rule out the occurrence of such a degenerate scenario, we show that it is exceptional in a sense detailed in Section 4.3.

The next natural step beyond the Wigner transform is considering the analogous nonlinear concentration problem for more general covariant time-frequency representations. A large and popular family of those stems from averaging the Wigner distribution with a suitable kernel: the so-called *Cohen class* consists of distributions of the form

$$Qf = Wf * \sigma, \qquad \sigma \in \mathcal{S}'(\mathbb{R}^{2d}).$$

Phase space averaging naturally promotes smoothing, which in turn is expected to better tame interferences (hence, degenerate scenarios) in concentration problems. We are able to confirm this intuition for several types of kernels, but proofs involve a completely different set of techniques than those used in this paper (that is, the recent advances in quantum harmonic analysis [28,29,27]), and this is the reason why we prefer to postpone these studies to a separate, forthcoming manuscript.

On the other hand, such heuristics do not make a general rule: In Section 6 we discuss some findings for a simple one-parameter generalization of the Wigner distribution, that is the τ-Wigner distribution:

$$W_\tau f(z) = \int_{\mathbb{R}^d} e^{-2\pi i \xi \cdot y} f(x + \tau y) \overline{f(x - (1-\tau)y)} \ \mathrm{d}y, \qquad \tau \in (0,1).$$

These are members of the Cohen class, since for $\tau \neq 1/2$ they arise from averaging the Wigner distribution with a chirp-like kernel in the Fourier domain:

$$W_\tau f = Wf * \sigma_\tau, \qquad \widehat{\sigma_\tau}(x,\xi) = e^{-2\pi i(\tau-1/2)x\cdot\xi}.$$

Despite appearances, when $\tau \neq 1/2$ the structure of surviving cross terms is entirely different from that of the Wigner case. We still attempt to push the same strategy as far as possible, and then pinpoint the currently open gap — which is again related to quantification of asymptotic contributions, where now pairs of escaping profiles may also exhibit infinite chaining phenomena.

To mitigate the asymmetry with respect to τ beyond the Wigner case we employ an additional averaging strategy, originating in signal processing to further tame the effect of interferences for Wigner-type distributions [20]. Indeed, the *Born–Jordan distribution* can be defined equivalently by

$$W_{\mathrm{BJ}} f(z) = \int_0^1 W_\tau f(z)\, \mathrm{d}\tau = Wf * \Theta_\sigma, \qquad \Theta(\zeta) = \frac{\sin(\pi\, \zeta_1\cdot\zeta_2)}{\pi\, \zeta_1\cdot\zeta_2}, \quad \zeta = (\zeta_1,\zeta_2) \in \mathbb{R}^{2d},$$

where Θ_σ stands for the symplectic Fourier transform of Θ. The Born–Jordan kernel is of very low regularity, as it fails to be locally in any L^p space — see for instance [5,9] and the references therein, also for framing its relevance in quantum mechanics. Nevertheless, at least in the planar case (i.e., $d = 1$) we are able to confirm the smoothing effect of integration over τ and obtain a full affirmative answer to the concentration problem (see Section 6.2), although for $p = \infty$ the supremum is not attained. In fact, the profile decomposition machinery actually shows that the concentration functional corresponding to the Born–Jordan distribution is weakly continuous on $L^2(\mathbb{R}^d)$. This fact is particularly interesting since it is a nontrivial[1] instance of a $L^p(\Omega)$-concentration functional associated with a quadratic time-frequency distribution where (semi)continuity does not fail. Moreover, this novel evidence is in line with known heuristics from signal processing about the strong smoothing effect of the Born–Jordan kernel, but here it is rigorously confirmed at a deeper topological level in terms of a regularity upgrade with respect to the analogous problem for the Wigner distribution. The current techniques do not allow us to extend the conclusion beyond the one-dimensional case, but it is actually unclear whether the supremum is finite when $d > 1$. Indeed, the nodal set of the kernel is a considerably larger manifold in that case. Perhaps restricting the optimization environment to a subspace of $L^2(\mathbb{R}^d)$ consisting of more regular functions could help (e.g., the Feichtinger algebra $M^1(\mathbb{R}^d)$ or other modulation spaces [10,11]), but these questions are left to future investigations.

The remainder of the paper is organized as follows. We collect the required preliminaries in Section 2, while the technical results needed to prove Theorem 1.1 are proven in Section 3, where we also discuss some consequences, such as failure of semicontinuity. In Section 4 we give the proof of Theorem 1.1. The concentration problem in the case $p = \infty$ is settled in Section 5. We conclude with the analysis of the concentration problem for the τ-Wigner and Born–Jordan distributions in Section 6.

2. Preliminaries

2.1. Notation

We write $a \lesssim b$ if there is a universal constant $C > 0$ such that $a \leq C\, b$. If the constant depends on an auxiliary parameter λ, that is $C = C(\lambda)$, then we write $a \lesssim_\lambda b$.

[1] For a fixed $g \in L^2(\mathbb{R}^d)$ the quadratic map $f \mapsto |A(f,g)|^2 = Wf * W\check{g}$, known as the spectrogram of f, is also sequentially weakly continuous from $L^2(\mathbb{R}^d)$ to $L^p(\Omega)$, for $1 \leq p < \infty$. However, its continuity is a direct consequence of the analogous sequential weak continuity of the linear map $f \mapsto A(f,g)$ computing the ambiguity transform, which was proved in [30, Proposition 5.1].

The inner product on $L^2(\mathbb{R}^d)$ is assumed to be conjugate-linear in the second argument, that is $\langle f, g\rangle = \int_{\mathbb{R}^d} f(y)\overline{g(y)}\,\mathrm{d}y$. The same brackets are used to denote the duality between a temperate distribution $f \in \mathcal{S}'(\mathbb{R}^d)$ and a function g in the Schwartz class.

If a sequence $\{f^{(n)}\}$ in $L^2(\mathbb{R}^d)$ converges weakly to f we write $f^{(n)} \rightharpoonup f$.

The Fourier transform of $f\colon \mathbb{R}^d \to \mathbb{C}$, whenever meaningful, is $\widehat{f}(\omega) = \int_{\mathbb{R}^d} e^{-2\pi i\omega\cdot y} f(y)\,\mathrm{d}y$.

For $\Omega \subset \mathbb{R}^{2d}$ we denote by $\|\cdot\|_{L^p(\Omega)}$ the local p-norm on Ω, while $\|\cdot\|_{L^p}$ always denotes the global p-norm over $\mathbb{R}^{2d}$.

2.2. Some results in time-frequency analysis

A recurring theme in the treatment of quadratic time-frequency representations is the (unavoidable) appearance of cross terms. In our context, this crucial role is played by the *cross-Wigner distribution*: Given $f, g \in L^2(\mathbb{R}^d)$, we set

$$W(f,g)(z) = \int_{\mathbb{R}^d} e^{-2\pi i\xi\cdot y} f\Big(x + \frac{y}{2}\Big)\overline{g\Big(x - \frac{y}{2}\Big)}\,\mathrm{d}y, \qquad z = (x,\xi) \in \mathbb{R}^{2d}. \tag{2}$$

Let us now collect some important properties of the cross-Wigner distribution, see [17, Section 4] for details. Letting $g = f$ recovers $Wf = W(f,f)$. We have that $\overline{W(f,g)} = W(g,f)$, and in particular Wf is real.

A simple density argument, taking $g \in \mathcal{S}(\mathbb{R}^d)$ first, shows that $W(f,g) \in L^2(\mathbb{R}^{2d})$ for all $f, g \in L^2(\mathbb{R}^d)$, and the following orthogonality relation (Moyal's identity) holds for $f_1, f_2, g_1, g_2 \in L^2(\mathbb{R}^d)$:

$$\langle W(f_1,g_1), W(f_2,g_2)\rangle_{L^2(\mathbb{R}^{2d})} = \langle f_1, f_2\rangle_{L^2(\mathbb{R}^d)}\overline{\langle g_1, g_2\rangle_{L^2(\mathbb{R}^d)}}.$$

As a result, and by a straightforward application of Cauchy-Schwarz's inequality, we obtain the norm bounds

$$\|W(f,g)\|_{L^2} = \|f\|_{L^2}\|g\|_{L^2}, \qquad \|W(f,g)\|_{L^\infty} \le 2^d\|f\|_{L^2}\|g\|_{L^2}. \tag{3}$$

Furthermore, it can be proved that $W(f,g)$ belongs to the space $C_0(\mathbb{R}^{2d})$ of continuous functions on $\mathbb{R}^{2d}$ that vanish at infinity.

A crucial property for our purposes is the *covariance under time-frequency shifts*. Recall that, given $z = (x,\xi) \in \mathbb{R}^{2d}$, the time-frequency shift $\pi(z)$ acts on $f \in L^2(\mathbb{R}^d)$ by $\pi(z)f(y) = e^{2\pi i\xi\cdot y} f(y - x)$. Setting $[z,u]$ for the symplectic inner product of z and u, that is

$$[z,u] = x_z\cdot\xi_u - \xi_z\cdot x_u = z\cdot Ju, \qquad J = \begin{pmatrix} O & I \\ -I & O \end{pmatrix} \in \mathbb{R}^{2d\times 2d},$$

the behavior of the Wigner distribution under time-frequency shifts reads as follows.

Lemma 2.1 *([6, Proposition 3.1]). Let $a = (x_a,\xi_a)$, $b = (x_b,\xi_b) \in \mathbb{R}^{2d}$ and set*

$$c(a,b) \coloneqq \frac{1}{2}\big(x_a + x_b\,,\, \xi_a + \xi_b\big).$$

Then, for all $f, g \in L^2(\mathbb{R}^d)$ and $z \in \mathbb{R}^{2d}$,

$$W(\pi(a)f, \pi(b)g)(z) = e^{\pi i(\xi_a+\xi_b)\cdot(x_a-x_b)} e^{2\pi i[z,a-b]}\, W(f,g)\big(z - c(a,b)\big). \tag{4}$$

Refined continuity properties of the mapping $f \to Wf$ can be given for functions with refined time-frequency regularity, such as those belonging to modulation and Wiener amalgam spaces. To avoid introducing auxiliary notions we have chosen to follow the approach of [8, Section 17], which is equivalent to the

standard one via short-time Fourier transform up to harmless global constants. For a thorough treatment of modulation spaces, as well as their standard definition, we direct the reader to [11,12].

We say that a temperate distribution $f \in \mathcal{S}'(\mathbb{R}^d)$ belongs to the modulation space $M^\infty(\mathbb{R}^d)$ if for some (in fact, any) $g \in \mathcal{S}(\mathbb{R}^d) \setminus \{0\}$ we have

$$\|f\|_{M^\infty} \coloneqq \sup_{z \in \mathbb{R}^{2d}} |W(f,g)(z)| < \infty.$$

Different choices of g result in equivalent norms on the Banach space $M^\infty(\mathbb{R}^d)$. We also stress the continuous inclusion $L^2(\mathbb{R}^d) \hookrightarrow M^\infty(\mathbb{R}^d)$ in view of (3).

We also need to deal with a special Wiener amalgam space [16]: We say that a measurable function $f\colon \mathbb{R}^d \to \mathbb{C}$ belongs to the Wiener amalgam space $W(L^2, L^\infty)(\mathbb{R}^d)$ if for some (hence any) function $g \in C^\infty(\mathbb{R}^d) \setminus \{0\}$ with compact support we have

$$\|f\|_{W(L^2,L^\infty)} \coloneqq \Big\| \|f(y)g(y-x)\|_{L^2_y} \Big\|_{L^\infty_x} < \infty.$$

See also [13,15,19]. Intuitively speaking, we may think of a function in $W(L^2, L^\infty)(\mathbb{R}^d)$ as having simultaneously local L^2 and global L^∞ regularity. As a concrete instance of this viewpoint, it is not difficult to show that if K is a compact subset of $\mathbb{R}^d$ then

$$\|f\|_{L^2(K)} \lesssim_K \|f\|_{W(L^2,L^\infty)} \lesssim \|f\|_{L^2}, \tag{5}$$

and therefore the continuous inclusions $L^2(\mathbb{R}^d) \hookrightarrow W(L^2, L^\infty)(\mathbb{R}^d) \hookrightarrow L^2(K)$ hold.

A key part of our argument involves refining the bounds (3) for the cross-Wigner distribution $W(f,g)$ in terms of the norms just introduced.

Theorem 2.2 *([7, Theorem 4.12]). Let $f \in L^2(\mathbb{R}^d)$ and $g \in M^\infty(\mathbb{R}^d)$. Then $W(f,g) \in W(L^2, L^\infty)(\mathbb{R}^{2d})$, with*

$$\|W(f,g)\|_{W(L^2,L^\infty)} \lesssim_d \|f\|_{L^2} \|g\|_{M^\infty}.$$

2.3. Profile decomposition

We will use profile decomposition in $L^2(\mathbb{R}^d)$, using the time-frequency shifts $\{\pi(z)\}_{z \in \mathbb{R}^{2d}}$ as the set of dislocations. The details of this decomposition can be found in [30], where this technique was introduced in order to study a similar problem. The defining properties are recalled below. For an account of the general theory, see [25,26,34]. Here we confine ourselves to collecting the relevant facts in the following result.

Lemma 2.3. *Let $f^{(n)}$ be a sequence in $L^2(\mathbb{R}^d)$ such that $\limsup_{n\to\infty} \|f^{(n)}\|_{L^2} \leq 1$. There exists a subsequence (still denoted with $f^{(n)}$), profiles $f_j \in L^2(\mathbb{R}^d)$ and points $z_j^{(n)} \in \mathbb{R}^{2d}$, $j \in \mathbb{N}$, such that for every integer $k \geq 1$ we have*

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

where $w_k^{(n)} \in L^2(\mathbb{R}^d)$ and the following conditions hold:

- $\lim_{n\to\infty} |z_j^{(n)} - z_{j'}^{(n)}| = +\infty$ *whenever $j \neq j'$, $f_j \neq 0$ and $f_{j'} \neq 0$.*

- $\sum_{j=1}^{k} \|f_j\|_{L^2}^2 + \limsup_{n\to\infty} \|w_k^{(n)}\|_{L^2}^2 \leq 1$.
- $\lim_{k\to\infty} \limsup_{n\to\infty} \|w_k^{(n)}\|_{M^\infty} = 0$.
- *For each* $k \geq 1$ *and* $1 \leq j \leq k$ *we have* $\lim_{n\to\infty} \pi(z_j^{(n)})^* w_k^{(n)} = 0$ *in the weak topology on* $L^2(\mathbb{R}^d)$.
- $\lim_{n\to\infty} \Big\| \sum_{j=1}^{k} \pi(z_j^{(n)}) f_j \Big\|_{L^2}^2 = \sum_{j=1}^{k} \|f_j\|_{L^2}^2$.

3. Controlling cross terms

3.1. An asymptotic formula

The following result will play a central role in taming the cross terms arising from profile decompositions for our optimization problem. It is in fact a result of independent interest, since it provides a quantitative control on the L^p concentration on a given phase space domain asymptotically generated by constructive interference of antipodally escaping wave packets.

Theorem 3.1. *Let* $f, g \in L^2(\mathbb{R}^d)$*, and let* $a^{(n)}, b^{(n)} \in \mathbb{R}^{2d}$ *satisfy*

$$d^{(n)} \coloneqq a^{(n)} - b^{(n)}, \quad |d^{(n)}| \to \infty, \qquad c^{(n)} \coloneqq \frac{a^{(n)} + b^{(n)}}{2} \to c \in \mathbb{R}^{2d}.$$

Define

$$\mathcal{S}_n(z) \coloneqq W(\pi(a^{(n)})f, \pi(b^{(n)})g)(z) + W(\pi(b^{(n)})g, \pi(a^{(n)})f)(z).$$

For every measurable $\Omega \subset \mathbb{R}^{2d}$ *with* $0 < |\Omega| < \infty$ *and* $1 \leq p < \infty$,

$$\lim_{n\to\infty} \|\mathcal{S}_n\|_{L^p(\Omega)} = 2\,\mathcal{C}_p\, \|W(f,g)\|_{L^p(\Omega - c)} \tag{6}$$

and

$$2\,\mathcal{C}_p\, \|W(f,g)\|_{L^p(\Omega-c)} \leq L(\|f\|_{L^2}^2 + \|g\|_{L^2}^2), \tag{7}$$

where L *is defined in* (1) *and*

$$\mathcal{C}_p = \left(\frac{1}{2\pi} \int_0^{2\pi} |\cos(\theta)|^p \, \mathrm{d}\theta \right)^{1/p} = \Big(\frac{1}{\pi} \mathrm{B}\Big(\frac{p+1}{2}, \frac{1}{2} \Big) \Big)^{1/p},$$

where B *is Euler's beta function.*

Remark 3.2. When $p < 2$, $\mathcal{S}_n$ need not be in $L^p(\mathbb{R}^{2d})$. However, each cross-Wigner term belongs to $L^2(\mathbb{R}^{2d})$, hence also to $L^2(\Omega) \hookrightarrow L^p(\Omega)$ for every $1 \leq p \leq 2$. As a result, by triangle inequality we have $\mathcal{S}_n \in L^p(\Omega)$.

Before delving into the details of the proof, some comments on the heuristics behind this result could be helpful. Setting $a^{(n)} = (x_a^{(n)}, \xi_a^{(n)})$ and $b^{(n)} = (x_b^{(n)}, \xi_b^{(n)})$, we have by covariance (Lemma 2.1):

$$W(\pi(a^{(n)})f, \pi(b^{(n)})g)(z) = e^{\pi i(\xi_a^{(n)} + \xi_b^{(n)}) \cdot (x_a^{(n)} - x_b^{(n)})} e^{2\pi i [z, d^{(n)}]} W(f,g)(z - c^{(n)}).$$

Since $W(g,f)=\overline{W(f,g)}$, after setting $\delta^{(n)}=2\pi d^{(n)}$ and $\phi^{(n)}=\pi(\xi_a^{(n)}+\xi_b^{(n)})\cdot(x_a^{(n)}-x_b^{(n)})+[c^{(n)},\delta^{(n)}]$ we get

$$\mathcal{S}_n(z)=2\operatorname{Re}(e^{i(\phi^{(n)}+[z-c^{(n)},\delta^{(n)}])}W(f,g)(z-c^{(n)})).$$

In Equation (6) the symmetric block $\mathcal{S}_n(z)$ accounting for the cross-Wigner interaction between two wave packets can then be viewed as the (the real part of the) product of a carrier plane wave with frequency $\delta^{(n)}$ and an envelope $W(f,g)$, both recentered at $c^{(n)}$. In fact, expanding the real part after letting $\alpha(z)$ denote the phase of $W(f,g)(z)$, we obtain an explicit cosine grating of the envelope:

$$\mathcal{S}_n(z)=2\Big|W(f,g)(z-c^{(n)})\Big|\cos\Big(\phi^{(n)}+[z-c^{(n)},\delta^{(n)}]+\alpha(z-c^{(n)})\Big).$$

The picture can be easily reinforced with a toy example in $d=1$, taking Gaussian wave packets $f(t)=g(t)=e^{-\pi|t|^2}$ antipodally travelling with $a^{(n)}=(r^{(n)},0)$ and $b^{(n)}=(-r^{(n)},0)$ for a sequence $\big|r^{(n)}\big|\to\infty$, so that $c^{(n)}=c=0$ and

$$\mathcal{S}_n(x,\xi)=2\sqrt{2}e^{-2\pi(|x|^2+|\xi|^2)}\cos\Big(-4\pi r^{(n)}\xi\Big).$$

Measuring the L^p-energy of $\mathcal{S}_n$ within an observation window Ω thus exhibits two effects: As $c^{(n)}\to c$ the envelope remains within $\Omega-c$, while intensity fringes separated by $\sim 1/|\delta^{(n)}|$ become closer and closer as $\big|d^{(n)}\big|\to\infty$, in fact leading to equidistribution of linear phases in the asymptotic regime (ultimately by the Riemann-Lebesgue lemma). Intuitively speaking, the separation length goes beyond the resolution power of any possible measuring procedure, which in fact eventually records the average L^p-brightness, explaining the occurrence of the visibility factor $\mathcal{C}_p$. This matches the parallelogram-law intuition when $p=2$ and also the expected L^∞ behavior since $\mathcal{C}_p\to 1$ as $p\to\infty$.

The second part, Equation (7), gives an inequality in terms of L, hence requires to control the cross-interference energy by a single signal in light of (1). This is done in the proof by introducing the functions $h_\theta^\pm=f\pm e^{i\theta}g$ for every $\theta\in[0,2\pi]$, which should not come as a surprise in view of the previous discussion. Indeed, an easy polarization argument yields

$$W(h_\theta^+)-W(h_\theta^-)=4\operatorname{Re}\left(e^{-i\theta}W(f,g)\right)=4|W(f,g)|\cos(\alpha-\theta),$$

and the claim follows by measuring the L^p-energy (i.e., taking the L^p norm) *after* averaging over θ. This is needed to remove the dependence on the (unknown) local phase α, concretely, by replacing $|\cos(\alpha-\theta)|^p$ with its angular mean $\mathcal{C}_p^p$. In addition, it is done to make the right-hand side independent of θ, indeed, $\|h_\theta^+\|_2^2+\|h_\theta^-\|_2^2=2(\|f\|_2^2+\|g\|_2^2)$.

To summarize these remarks, Theorem 3.1 shows that two asymptotically separated packets generally interfere via Wigner distribution to contribute to the L^p norm over a given phase space region by an amount that can be determined (roughly speaking: $\sim$ cross-Wigner envelope mass $\times$ averaged visibility of oscillating carrier), and which cannot exceed the optimal value L.

Proof of Theorem 3.1. With the notation introduced in the previous discussion, the change of variable $z'=z-c^{(n)}$ yields

$$\begin{aligned}&\lim_{n\to\infty}\|\operatorname{Re}\Big(W(\pi(a^{(n)})f,\pi(b^{(n)})g)\Big)\|_{L^p(\Omega)}^p\\&\quad=\lim_{n\to\infty}\int_\Omega\Big|\operatorname{Re}\Big(W(\pi(a^{(n)})f,\pi(b^{(n)})g)(z)\Big)\Big|^p\,\mathrm{d}z\end{aligned}$$

$$= \lim_{n\to\infty} \int_{\Omega_n} \left|\mathrm{Re}\Big(e^{i(\phi^{(n)}+[z',\delta^{(n)}])}W(f,g)(z')\Big)\right|^p \mathrm{d}z'$$
$$= \lim_{n\to\infty} \int_{\Omega_n} |W(f,g)(z')|^p \left|\cos\Big([z',\delta^{(n)}]+\phi^{(n)}+\alpha(z')\Big)\right|^p \mathrm{d}z',$$

where we set $\Omega_n = \Omega - c^{(n)}$. Therefore, after setting $F(z) = |W(f,g)(z)|^p$ and $H(\theta) = |\cos(\theta)|^p$, our claim reads

$$\lim_{n\to\infty} \int_{\Omega_n} F(z)H([z,\delta^{(n)}]+\phi^{(n)}+\alpha(z))\,\mathrm{d}z = \Big(\frac{1}{2\pi}\int_0^{2\pi} H(\theta)\,\mathrm{d}\theta\Big) \int_{\Omega-c} F(z)\,\mathrm{d}z.$$

We argue by approximation with trigonometric polynomials as detailed below.

Step 1. We first prove that the claim holds when $H(\theta)$ is replaced by a complex exponential $P(\theta) = e^{ik\theta}$ with $k \in \mathbb{Z}$.

For $k = 0$ the problem boils down to showing that

$$\lim_{n\to\infty} \int_{\Omega_n} F(z)\,\mathrm{d}z = \int_{\Omega-c} F(z)\,\mathrm{d}z,$$

which follows at once using continuity on L^1 of translates of the indicator function χ_Ω of Ω and the bound in (3):

$$\left|\int_{\mathbb{R}^{2d}} F(z)(\chi_\Omega(z+c^{(n)}) - \chi_\Omega(z+c))\,\mathrm{d}z\right| \le \|F\|_{L^\infty} \Big\|\chi_\Omega(z+c^{(n)}) - \chi_\Omega(z+c)\Big\|_{L^1_z} \to 0.$$

For $k \neq 0$ the right-hand side vanishes, so we must show that

$$\lim_{n\to\infty} \int_{\Omega_n} F(z)e^{ik([z,\delta^{(n)}]+\phi^{(n)}+\alpha(z))}\,\mathrm{d}z = 0.$$

If we now let $G_n(z) = F(z)e^{ik\alpha(z)}\chi_\Omega(z+c^{(n)})$, the integral on the left-hand side is

$$\int_{\Omega_n} F(z)e^{ik([z,\delta^{(n)}]+\phi^{(n)}+\alpha(z))}\,\mathrm{d}z = e^{ik\phi^{(n)}} \int_{\mathbb{R}^{2d}} G_n(z)e^{ik[z,\delta^{(n)}]}\,\mathrm{d}z$$
$$= e^{ik\phi^{(n)}}\widehat{G_n}(-kJd^{(n)}).$$

Since $G_n \in L^1(\mathbb{R}^{2d})$ and $\left|d^{(n)}\right| \to \infty$ with n, we aim to invoke the Riemann-Lebesgue lemma to conclude that this integral vanishes in the limit. This is not possible at once, since G_n also depends on n. However, it is enough to introduce the function $G(z) = F(z)e^{ik\alpha(z)}\chi_\Omega(z+c)$ and arguing as before we have $\|G_n - G\|_{L^1} \to 0$. As a result, we obtain

$$\left|\int_{\Omega_n} F(z)e^{ik([z,\delta^{(n)}]+\phi^{(n)}+\alpha(z))}\,\mathrm{d}z\right| = \left|e^{ik\phi^{(n)}}\widehat{G_n}(-kJd^{(n)})\right|$$
$$\le |\widehat{G_n}(-kJd^{(n)}) - \widehat{G}(-kJd^{(n)})| + |\widehat{G}(-kJd^{(n)})|$$
$$\le \|G_n - G\|_{L^1} + |\widehat{G}(-kJd^{(n)})| \to 0,$$

by virtue of the previous remark and the Riemann-Lebesgue lemma.

Therefore, by linearity we infer for any trigonometric polynomial $P(\theta) = \sum_{m=-N}^{N} c_m e^{im\theta}$ that

$$\lim_{n\to\infty} \int_{\Omega_n} F(z)P([z,\delta^{(n)}] + \phi^{(n)} + \alpha(z))\,\mathrm{d}z = \Big(\frac{1}{2\pi}\int_0^{2\pi} P(\theta)\,\mathrm{d}\theta\Big) \int_{\Omega-c} F(z)\,\mathrm{d}z.$$

Step 2. Given an $\varepsilon > 0$ we let $P_\varepsilon(\theta)$ be a trigonometric polynomial such that $\|H - P_\varepsilon\|_{L^\infty(\mathbb{T})} < \varepsilon$. For simplicity, given a function G on the circle we will write $I_n(G) = \int_{\Omega_n} F(z)G([z,\delta^{(n)}] + \phi^{(n)} + \alpha(z))\,\mathrm{d}z$ and $\overline{G} = \frac{1}{2\pi}\int_0^{2\pi} G(\theta)\,\mathrm{d}\theta$. We now get

$$\Big|I_n(H) - \overline{H}\|F\|_{L^1(\Omega-c)}\Big| \le |I_n(H) - I_n(P_\varepsilon)| + \Big|I_n(P_\varepsilon) - \overline{P_\varepsilon}\|F\|_{L^1(\Omega-c)}\Big| + \|F\|_{L^1(\Omega-c)}\big|\overline{P_\varepsilon} - \overline{H}\big|.$$

For what concerns the first term we have

$$|I_n(H) - I_n(P_\varepsilon)| \le \|H - P_\varepsilon\|_{L^\infty}\|F\|_{L^1(\Omega-c^{(n)})} < \varepsilon\|F\|_{L^1(\Omega-c^{(n)})}.$$

Since $\|F\|_{L^1(\Omega-c^{(n)})} \to \|F\|_{L^1(\Omega-c)}$ (cf. Step 1), there exists $C > 0$ such that $|I_n(H) - I_n(P_\varepsilon)| < C\varepsilon$. Moreover, as proved in Step 1, the second term vanishes in the limit and can therefore be made smaller than ε by picking n large enough. The desired result then follows after bounding the third term with $\big|\overline{P_\varepsilon} - \overline{H}\big| \le \varepsilon$.

To conclude, we need to prove the inequality

$$2\mathcal{C}_p\|W(f,g)\|_{L^p(\Omega-c)} \le L(\|f\|_{L^2}^2 + \|g\|_{L^2}^2).$$

For $\theta \in [0, 2\pi]$ we write $h_\theta^\pm = f \pm e^{i\theta}g$, hence

$$W(h_\theta^\pm) = Wf + Wg \pm 2\,\mathrm{Re}\big(e^{-i\theta}W(f,g)\big),$$

and thus

$$W(h_\theta^+) - W(h_\theta^-) = 4\,\mathrm{Re}\big(e^{-i\theta}W(f,g)\big).$$

After raising to the power p and integrating in z and θ

$$\frac{1}{2\pi}\int_0^{2\pi} \|W(h_\theta^+) - W(h_\theta^-)\|_{L^p(\Omega-c)}^p\,\mathrm{d}\theta = \frac{1}{2\pi}\int_0^{2\pi} \|4\,\mathrm{Re}\big(e^{-i\theta}W(f,g)\big)\|_{L^p(\Omega-c)}^p\,\mathrm{d}\theta.$$

Let us now treat each side of this equation separately. If $\alpha(z)$ denotes the phase of the complex number $W(f,g)(z)$, so that we can write $W(f,g)(z) = |W(f,g)(z)|e^{i\alpha(z)}$, the right-hand side equals (by Fubini's theorem)

$$\frac{4^p}{2\pi}\int_{\Omega-c}\int_0^{2\pi} |\cos(\alpha(z) - \theta)|^p |W(f,g)(z)|^p\,\mathrm{d}\theta\,\mathrm{d}z = 4^p\mathcal{C}_p^p\|W(f,g)\|_{L^p(\Omega-c)}^p.$$

On the other hand, to estimate the left-hand side we use Minkowski's inequality to get

$$\begin{aligned}\frac{1}{2\pi}\int_0^{2\pi}\|W(h_\theta^+)-W(h_\theta^-)\|_{L^p(\Omega-c)}^p\,\mathrm{d}\theta &\le \frac{1}{2\pi}\int_0^{2\pi}\big(\|W(h_\theta^+)\|_{L^p(\Omega-c)}+\|W(h_\theta^-)\|_{L^p(\Omega-c)}\big)^p\,\mathrm{d}\theta\\ &\le \frac{1}{2\pi}\int_0^{2\pi}L^p\big(\|h_\theta^+\|_{L^2}^2+\|h_\theta^-\|_{L^2}^2\big)^p\,\mathrm{d}\theta\\ &= \frac{1}{2\pi}\int_0^{2\pi}2^pL^p\big(\|f\|_{L^2}^2+\|g\|_{L^2}^2\big)^p\,\mathrm{d}\theta\\ &= 2^pL^p\big(\|f\|_{L^2}^2+\|g\|_{L^2}^2\big)^p,\end{aligned}$$

where we also used the invariance of L under shifts in the Problem (1), since by covariance (4) we have

$$\|Wf\|_{L^p(\Omega-c)}=\|Wf(\cdot-c)\|_{L^p(\Omega)}=\|W(\pi(c)f)\|_{L^p(\Omega)}\le L\|\pi(c)f\|_{L^2}^2=L\|f\|_{L^2}^2.$$

The claim thus follows. □

3.2. Properties of inequality optimizers

By inspecting the proof of Theorem 3.1, we obtain the following characterization of the inequality's (non-trivial) optimizers.

Corollary 3.3. *Let $f,g\in L^2(\mathbb{R}^d)\setminus\{0\}$. The pair (f,g) achieves equality in the corresponding cross-term inequality from Theorem 3.1, i.e.,*

$$2\,\mathcal{C}_p\,\|W(f,g)\|_{L^p(\Omega)}=L(\|f\|_{L^2}^2+\|g\|_{L^2}^2)$$

if and only if the following conditions hold:

(i) The functions $h_\theta^\pm=f\pm e^{i\theta}g$ are optimizers for the problem (1) for every $\theta\in[0,2\pi]$.
(ii) For every $\theta\in[0,2\pi]$ we have

$$\|W(h_\theta^+)-W(h_\theta^-)\|_{L^p(\Omega)}=\|W(h_\theta^+)\|_{L^p(\Omega)}+\|W(h_\theta^-)\|_{L^p(\Omega)}. \tag{8}$$

It is clear that simultaneous occurrence of these conditions place strong constraints on the structure of non-trivial inequality extremizers. Indeed, we illustrate below some of those that can be easily inferred.

Corollary 3.4. *Let $f,g\in L^2(\mathbb{R}^d)\setminus\{0\}$ be optimizers for the inequality (7). Then (8) implies $Wf(z)=-Wg(z)$ for almost every $z\in\Omega$.*

Proof. Recall that $h_\theta^\pm=f\pm e^{i\theta}g$, $\theta\in[0,2\pi]$. We write $X=Wf+Wg$ and $Y_\theta=2\,\mathrm{Re}(e^{-i\theta}W(f,g))$, so that $W(h_\theta^\pm)=X\pm Y_\theta$, and the claim amounts to showing that $X=0$ a.e. on Ω.

Equality in Minkowski's inequality as in (8) for $1\le p<\infty$ implies that, for every $\theta\in[0,2\pi]$,

$$\|W(h_\theta^+)-W(h_\theta^-)\|_{L^p(\Omega)}=\||W(h_\theta^+)|+|W(h_\theta^-)|\|_{L^p(\Omega)},$$

from which we infer the pointwise identity

$$|W(h_\theta^+)(z)-W(h_\theta^-)(z)|=|W(h_\theta^+)(z)|+|W(h_\theta^-)(z)|\qquad\text{for a.e. } z\in\Omega.$$

Let $D \subset [0, 2\pi]$ be a countable dense subset and for each $\theta \in D$ let $N_\theta \subset \Omega$ be a null set outside which the above identity holds. We also introduce the full-measure subset $E = \Omega \setminus \bigcup_{\theta \in D} N_\theta$ so that for every $z \in E$ the identity holds for every $\theta \in D$.

Expanding the Wigner distributions and since X, Y_θ are real-valued, this amounts to

$$2|Y_\theta(z)| = |X(z) + Y_\theta(z)| + |X(z) - Y_\theta(z)| = 2\max\{|X(z)|, |Y_\theta(z)|\} \qquad \forall\, z \in E, \forall\, \theta \in D,$$

from which we deduce that $|Y_\theta(z)| \geq |X(z)|$ for all $z \in E$ and $\theta \in D$. On the other hand, setting $W(f, g)(z) = |W(f, g)(z)|e^{i\alpha(z)}$, by continuity we have

$$|X(z)| \leq \inf_{\theta \in D} |Y_\theta(z)| = 2|W(f, g)(z)| \inf_{\theta \in D} |\cos(\alpha(z) - \theta)| = 0, \quad \forall z \in E,$$

hence we infer $X(z) = 0$ for a.e. $z \in \Omega$ as claimed. □

Corollary 3.5. *Let $f, g \in L^2(\mathbb{R}^d) \setminus \{0\}$ be optimizers for the inequality (7). Then $\langle f, g\rangle = 0$.*

Proof. From Corollary 3.4 we know that $Wf = -Wg$ almost everywhere on Ω, hence for every $\theta \in [0, 2\pi]$ we have

$$\begin{aligned}\|W(f + e^{i\theta}g)\|_{L^p(\Omega)} &= \|Wf + Wg + 2\,\mathrm{Re}(e^{-i\theta}W(f, g))\|_{L^p(\Omega)} \\ &= \|2\,\mathrm{Re}(e^{-i\theta}W(f, g))\|_{L^p(\Omega)} \\ &= \|-2\,\mathrm{Re}(e^{-i\theta}W(f, g))\|_{L^p(\Omega)} = \|W(f - e^{i\theta}g)\|_{L^p(\Omega)}.\end{aligned}$$

Furthermore since both $f + e^{i\theta}g$ and $f + e^{i(\theta+\pi)}g = f - e^{i\theta}g$ are optimizers, this yields

$$L\|f + e^{i\theta}g\|_{L^2}^2 = \|W(f + e^{i\theta}g)\|_{L^p(\Omega)} = \|W(f - e^{i\theta}g)\|_{L^p(\Omega)} = L\|f - e^{i\theta}g\|_{L^2}^2.$$

Since $L > 0$ we conclude that $\|f + e^{i\theta}g\|_{L^2}^2 = \|f - e^{i\theta}g\|_{L^2}^2$ for every $\theta \in [0, 2\pi]$. By expanding, we see that

$$\|f \pm e^{i\theta}g\|_{L^2}^2 = \|f\|_{L^2}^2 + \|g\|_{L^2}^2 \pm 2\,\mathrm{Re}\langle f, e^{i\theta}g\rangle,$$

so it must be that

$$\mathrm{Re}(e^{-i\theta}\langle f, g\rangle) = \mathrm{Re}(\langle f, g\rangle)\cos\theta + \mathrm{Im}(\langle f, g\rangle)\sin\theta = 0, \qquad \forall \theta \in [0, 2\pi],$$

which clearly forces $\langle f, g\rangle = 0$. □

3.3. Failure of semicontinuity

The persistence of the cross-Wigner block $\mathcal{S}_n$ when the two packets $\pi(a^{(n)})f$ and $\pi(b^{(n)})g$ separate antipodally in phase space while their midpoint remains bounded is a major obstruction to a direct-method proof of the existence of an optimizer. Indeed, this causes both the lack of sequential weak upper semicontinuity, as we now prove more explicitly, and the non relative compactness of a maximizing sequence.

Proposition 3.6. *Let $\Omega \subset \mathbb{R}^{2d}$ be measurable with $0 < |\Omega| < \infty$, and $1 \leq p < \infty$. The functional*

$$J(f) \coloneqq \|Wf\|_{L^p(\Omega)}, \qquad f \in L^2(\mathbb{R}^d),$$

is not sequentially weakly upper semicontinuous on $L^2(\mathbb{R}^d)$ at any point.

Proof. Consider $g(y) = e^{-\pi|y|^2}$, so that $Wg > 0$ on $\mathbb{R}^{2d}$. Let $(\xi_n) \subset \mathbb{R}^d$ be a sequence such that $|\xi_n| \to \infty$ and set

$$a^{(n)} \coloneqq (0, \xi_n), \qquad b^{(n)} \coloneqq (0, -\xi_n), \qquad v_n \coloneqq \pi(a^{(n)})g + \pi(b^{(n)})g.$$

As a straightforward consequence of the Riemann-Lebesgue lemma we have $v_n \rightharpoonup 0$ in $L^2(\mathbb{R}^d)$ and

$$\|v_n\|_{L^2}^2 = 2\|g\|_{L^2}^2 + 2\,\mathrm{Re}\int_{\mathbb{R}^d} |g(y)|^2 e^{4\pi i \xi_n \cdot y}\,dy \to 2\|g\|_{L^2}^2.$$

By sesquilinearity of W we also have

$$W(v_n) = W(\pi(a^{(n)})g) + W(\pi(b^{(n)})g) + \mathcal{S}_n,$$

with the symmetric cross-block

$$\mathcal{S}_n \coloneqq W\big(\pi(a^{(n)})g, \pi(b^{(n)})g\big) + W\big(\pi(b^{(n)})g, \pi(a^{(n)})g\big).$$

We now claim that we have

$$\|W(\pi(a^{(n)})g)\|_{L^p(\Omega)} \to 0, \qquad \|W(\pi(b^{(n)})g)\|_{L^p(\Omega)} \to 0.$$

Indeed, if Ω is compact then by covariance (Lemma 2.1) we have

$$\|W(\pi(a^{(n)})g)\|_{L^p(\Omega)} = \|Wg\|_{L^p(\Omega - a^{(n)})} \leq |\Omega|^{1/p} \sup_{z \in \Omega - a^{(n)}} |Wg(z)|.$$

Since Ω is compact, thus bounded, and Wg has C_0-decay, the supremum in the above line vanishes as $n \to \infty$, since $|a^{(n)}| \to \infty$, and likewise for $b^{(n)}$. For a general finite measure set Ω we use inner regularity to pick a compact set $K \subset \Omega$ such that $|\Omega \setminus K| < \varepsilon$ and combine the argument for compact sets with the triangle inequality.

Now, if we set

$$\varepsilon_n \coloneqq \|W(\pi(a^{(n)})g)\|_{L^p(\Omega)} + \|W(\pi(b^{(n)})g)\|_{L^p(\Omega)} \to 0,$$

the triangle inequality gives

$$\|\mathcal{S}_n\|_{L^p(\Omega)} - \varepsilon_n \leq \|W(v_n)\|_{L^p(\Omega)} \leq \|\mathcal{S}_n\|_{L^p(\Omega)} + \varepsilon_n.$$

By Theorem 3.1 with $d^{(n)} = (0, 2\xi_n)$ and $c^{(n)} = 0$, we have $\|\mathcal{S}_n\|_{L^p(\Omega)} \longrightarrow 2\,\mathcal{C}_p\,\|Wg\|_{L^p(\Omega)}$, hence

$$\|W(v_n)\|_{L^p(\Omega)} \longrightarrow B \coloneqq 2\,\mathcal{C}_p\,\|Wg\|_{L^p(\Omega)} > 0 = \|W(0)\|_{L^p(\Omega)} = J(0),$$

which proves that $J(f) = \|Wf\|_{L^p(\Omega)}$ is not sequentially weakly upper semicontinuous at 0.

Let us now consider an arbitrarily chosen $f \in L^2(\mathbb{R}^d)$, and for $\lambda > 0$ set $g_{n,\lambda} = f + \lambda v_n$, so that $g_{n,\lambda} \rightharpoonup f$. We have

$$W(g_{n,\lambda}) = Wf + \lambda^2 W(v_n) + 2\lambda\,\mathrm{Re}\,W(f, v_n). \tag{9}$$

Again by covariance, C_0-decay and regularity we have

$$\|W(f,v_n)\|_{L^p(\Omega)} \leq \|W(f,g)\|_{L^p(\Omega - a^{(n)}/2)} + \|W(f,g)\|_{L^p(\Omega - b^{(n)}/2)} \to 0.$$

By (9) $\|W(v_n)\|_{L^p(\Omega)} \to B > 0$, hence by the triangle inequality we obtain

$$\liminf_{n\to\infty} \|W(g_{n,\lambda})\|_{L^p(\Omega)} \;\geq\; \lambda^2 B - \|Wf\|_{L^p(\Omega)}.$$

Choosing λ large enough so that $\lambda^2 B > 2\|Wf\|_{L^p(\Omega)}$ yields $\limsup_n \|W(g_{n,\lambda})\|_{L^p(\Omega)} > \|Wf\|_{L^p(\Omega)}$, proving failure of sequential weak upper semicontinuity at f. □

4. Proof of Theorem 1.1

First of all, it is clear that since $|\Omega| > 0$ we have $L > 0$. Moreover, by (3) we obtain $\left(\int_\Omega |Wf(z)|^p \,\mathrm{d}z\right)^{1/p} \leq 2^d |\Omega|^{1/p} \|f\|_{L^2}^2$, from which we infer that $L < \infty$.

4.1. Profile decomposition

Let now $f^{(n)}$ be a maximizing sequence for the problem (1), and assume without loss of generality that $\|f^{(n)}\|_{L^2} = 1$ for all n. We perform profile decomposition as detailed in Lemma 2.3, passing to subsequences wherever needed, and write for simplicity $F_k^{(n)} = \sum_{j=1}^k \pi(z_j^{(n)}) f_j$, so that $f^{(n)} = F_k^{(n)} + w_k^{(n)}$. By sesquilinearity and covariance of the Wigner distribution we have

$$\begin{aligned} W(f^{(n)}) &= \sum_{j=1}^k W(\pi(z_j^{(n)})f_j) + \sum_{1\leq j\neq j'\leq k} W(\pi(z_j^{(n)})f_j, \pi(z_{j'}^{(n)})f_{j'}) \\ &\quad + W(F_k^{(n)}, w_k^{(n)}) + W(w_k^{(n)}, F_k^{(n)}) + W(w_k^{(n)}) \\ &= \sum_{j=1}^k W(f_j)(\cdot - z_j^{(n)}) + 2\sum_{1\leq j<j'\leq k} \mathrm{Re}\Big(W(\pi(z_j^{(n)})f_j, \pi(z_{j'}^{(n)})f_{j'})\Big) \\ &\quad + W(F_k^{(n)}, w_k^{(n)}) + W(w_k^{(n)}, F_k^{(n)}) + W(w_k^{(n)}). \end{aligned} \tag{10}$$

Let us analyze this expression term by term.

Remainder terms

We first deal with the three remainder terms containing $w_k^{(n)}$, and show that the term $W(F_k^{(n)}, w_k^{(n)})$ vanishes in the limit by arguing as in [30]. Similar arguments apply to the other remainder terms with obvious modifications.

We resort to (5) and Theorem 2.2 to obtain, for any compact $K \subset \mathbb{R}^{2d}$,

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

We emphasize that the implied constants are independent of n and k, but in general may depend on K. Thus, since by Lemma 2.3 we can choose n sufficiently large to ensure $\left\|F_k^{(n)}\right\|_{L^2}^2 + \left\|w_k^{(n)}\right\|_{L^2}^2 \leq 2$, we conclude that

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

Combining with the trivial bound

$$\|W(F_k^{(n)}, w_k^{(n)})\|_{L^\infty} \leq 2^d \|F_k^{(n)}\|_{L^2} \|w_k^{(n)}\|_{L^2} \leq 2^{d+1}, \tag{11}$$

by Hölder's inequality we get

$$\|W(F_k^{(n)}, w_k^{(n)})\|_{L^p(K)} \lesssim_K \|w_k^{(n)}\|_{M^\infty}^{\theta_p}, \qquad \theta_p = \begin{cases} 1 & 1 \leq p \leq 2 \\ \frac{2}{p} & p \geq 2. \end{cases}$$

This bound suffices to prove that

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

Indeed, if Ω is compact the conclusion is immediate from the previous argument. If not, it is enough to exhaust by a compact $K \subset \Omega$ with $|\Omega \setminus K|$ arbitrarily small and resort to (11).

Structure of profile decomposition

Let us now examine the main terms appearing in the profile decomposition (10).

A single compact profile

Since $|z_j^{(n)} - z_{j'}^{(n)}| \to \infty$ for all $j \neq j'$ by construction, it is easy to argue by contradiction that $|z_j^{(n)}| \to \infty$ for all but at most one j, which we denote by j_0 in the case it does exist. Since $Wf \in C_0(\mathbb{R}^{2d})$ for all $f \in L^2(\mathbb{R}^d)$, we have, by covariance and inner regularity, $\|W(f_j)(\cdot - z_j^{(n)})\|_{L^p(\Omega)} \to 0$ as $n \to \infty$ for every $j \neq j_0$. So, there is at most one potentially surviving pure profile $f_{j_0} \in L^2(\mathbb{R}^d) \setminus \{0\}$, from now on referred to as the *compact profile*, corresponding (up to subsequences) to the bounded, converging sequence $z_{j_0}^{(n)} \to z_{j_0} \in \mathbb{R}^{2d}$. We conveniently set $f_{j_0} = 0$ and $z_{j_0}^{(n)} \equiv 0$ in the case where the index j_0 with the aforementioned properties does not exist.

Cross terms and their interactions

Note that the Wigner covariance relation implies

$$\left| W(\pi(z_j^{(n)}) f_j, \pi(z_{j'}^{(n)}) f_{j'})(z) \right| = \left| W(f_j, f_{j'}) \left(z - \frac{z_j^{(n)} + z_{j'}^{(n)}}{2} \right) \right|.$$

Just like for the pure terms, only the profile pairs for which $z_j^{(n)} + z_{j'}^{(n)}$ remains bounded as $n \to \infty$ avoid vanishing in the limit. We will refer to those as *surviving pairs*.

In fact, if $z_j^{(n)} + z_{j'}^{(n)}$ remains bounded, then neither j nor j' can associate with a the compact profile. Indeed, neither $z_j^{(n)}$ nor $z_{j'}^{(n)}$ can remain bounded with n, since in that case $|z_j^{(n)} - z_{j'}^{(n)}| \leq |z_j^{(n)} + z_{j'}^{(n)}| + 2\min\{|z_j^{(n)}|, |z_{j'}^{(n)}|\}$ would remain bounded as well. Similarly, neither j nor j' can appear in a different surviving pair: for instance, if we also had that $z_j^{(n)} + z_\ell^{(n)}$ remained bounded, then we would have $|z_{j'}^{(n)} - z_\ell^{(n)}| \leq |z_{j'}^{(n)} + z_j^{(n)}| + |z_j^{(n)} + z_\ell^{(n)}| < \infty$ in the limit, another contradiction.

It is thus natural to introduce the set $\mathcal{J}$ consisting of pairs of indices (j, j') with $j < j'$ in the profile decomposition of $f^{(n)}$ such that $|z_j^{(n)} + z_{j'}^{(n)}|$ remains bounded, and therefore converges up to subsequence.

4.2. The main argument

The previous discussion shows that when we take the limit in n in (10) the only profiles that eventually contribute are the compact profile f_{j_0}, if it exists, and surviving pairs $(f_j, f_{j'})$ with $(j, j') \in \mathcal{J}$. Since profile

decomposition holds for any k we can take the limit, first in n, then in k. By combining our observations about the surviving terms in the expansion of the Wigner distribution, after passing to a common subsequence over $\mathcal{J}$ via a diagonal argument we obtain

$$L \leq \lim_{n\to\infty} \left\| W(\pi(z_{j_0}^{(n)}) f_{j_0}) \right\|_{L^p(\Omega)} + \lim_{n\to\infty} \sum_{(j,j')\in\mathcal{J}} \left\| \mathcal{S}_{j,j'}^{(n)} \right\|_{L^p(\Omega)},$$

where we set $\mathcal{S}_{j,j'}^{(n)} = 2\,\mathrm{Re}\Big(W(\pi(z_j^{(n)}) f_j, \pi(z_{j'}^{(n)}) f_{j'}) \Big)$ (cf. notation in Theorem 3.1). It is clear that

$$\lim_{n\to\infty} \|W(\pi(z_{j_0}^{(n)}) f_{j_0})\|_{L^p(\Omega)} \leq \lim_{n\to\infty} L \|\pi(z_{j_0}^{(n)}) f_{j_0}\|_{L^2}^2 = L\|f_{j_0}\|_{L^2}^2,$$

while invoking Theorem 3.1 yields

$$\lim_{n\to\infty} \left\| \mathcal{S}_{j,j'}^{(n)} \right\|_{L^p(\Omega)} = 2\mathcal{C}_p \|W(f_j, f_{j'})\|_{L^p(\Omega - c_{j,j'})} \leq L(\|f_j\|_{L^2}^2 + \|f_{j'}\|_{L^2}^2),$$

where $c_{j,j'} \in \mathbb{R}^{2d}$ is the limit of (a subsequence of) $c_{j,j'}^{(n)}$, which is bounded since $(j,j') \in \mathcal{J}$. To sum up, we have obtained

$$\begin{aligned} L &\leq \lim_{n\to\infty} \left\| W(\pi(z_{j_0}^{(n)}) f_{j_0}) \right\|_{L^p(\Omega)} + \sum_{(j,j')\in\mathcal{J}} \lim_{n\to\infty} \left\| \mathcal{S}_{j,j'}^{(n)} \right\|_{L^p(\Omega)} \\ &\leq L\Big(\|f_{j_0}\|_{L^2}^2 + \sum_{(j,j')\in\mathcal{J}} \|f_j\|_{L^2}^2 + \|f_{j'}\|_{L^2}^2 \Big) \\ &\leq L \sum_{j=1}^{\infty} \|f_j\|_{L^2}^2 \\ &\leq L. \end{aligned}$$

Note that in the second inequality we crucially used the matching structure of the surviving pairs, which avoids double counting. It is thus clear that every inequality in this chain must be in fact an equality. In particular, we get

$$\|f_{j_0}\|_{L^2}^2 + \sum_{(j,j')\in\mathcal{J}} \Big(\|f_j\|_{L^2}^2 + \|f_{j'}\|_{L^2}^2 \Big) = 1.$$

Now if $f_{j_0} \neq 0$ then $\lim_{n\to\infty} \|W(\pi(z_{j_0}^{(n)}) f_{j_0})\|_{L^p(\Omega)} = L\|f_{j_0}\|_{L^2}^2$. Since $z_{j_0}^{(n)}$ converges (up to subsequence) to $z_0 \in \mathbb{R}^{2d}$ we get, by employing the same argument as in the proof of Theorem 3.1,

$$\lim_{n\to\infty} \|W(\pi(z_{j_0}^{(n)}) f_{j_0})\|_{L^p(\Omega)} = \|W f_{j_0}\|_{L^p(\Omega - z_0)} = \|W(\pi(z_0) f_{j_0})\|_{L^p(\Omega)},$$

hence showing that $\pi(z_0) f_{j_0}$ is an optimizer.

On the other hand, if no compact profile exists we must examine the cross terms. Since every inequality is an equality in the above chain, for each pair $(j,j') \in \mathcal{J}$ we must have

$$2\mathcal{C}_p \|W(f_j, f_{j'})\|_{L^p(\Omega - c_{j,j'})} = L(\|f_j\|_{L^2}^2 + \|f_{j'}\|_{L^2}^2).$$

If one of f_j or $f_{j'}$ is zero, then they must both be zero. Otherwise, in view of Corollary 3.3, after setting $h_{j,\theta}^{\pm} = f_j \pm e^{i\theta} f_{j'}$ for $\theta \in [0, 2\pi]$, by covariance we conclude that

$$L\|h^{\pm}_{j,\theta}\|^2_{L^2} = \|W(h^{\pm}_{j,\theta})\|_{L^p(\Omega - c_{j,j'})} = \|W(\pi(c_{j,j'})h^{\pm}_{j,\theta})\|_{L^p(\Omega)}.$$

Observe that by Corollary 3.5 we have $\|h^{\pm}_{j,\theta}\|^2_{L^2} = \|f_j\|^2_{L^2} + \|f_{j'}\|^2_{L^2} \neq 0$. In particular, this means that there exists a one-parameter family of optimizers for the problem (1), namely $\{\pi(c_{j,j'})(f_j + e^{i\theta} f_{j'}) : \theta \in [0, 2\pi]\}$.

To summarize, if $\|f_{j_0}\|^2_{L^2} \neq 0$, then $\pi(z_0)f_{j_0}$ is an optimizer. If no compact profile exists then there is at least one surviving pair $(f_j, f_{j'})$ of non-trivial profiles which generates a family $\{\pi(c_{j,j'})(f_j + e^{i\theta} f_{j'})\}_{\theta\in[0,2\pi]}$ of optimizers. In any case, we can conclude that the supremum (1) is attained. □

4.3. Comments on exotic optimizers

Theorem 1.1 establishes the existence of an optimizer for the Problem (1). However, the proof reveals two very different situations: up to shifts, the optimizer may be the single compact profile extracted from a profile decomposition or any of those belonging to the one-parameter family associated with a surviving pair.

It must be stressed that, in the second (degenerate) case, the surviving pair $(f_j, f_{j'})$ happens to be an optimizer of the inequality (7), hence it comes along with the additional constraints highlighted in Corollary 3.3, some of which have been deduced in Corollaries 3.4 and 3.5. In view of these results, it is clear that if $(f_j, f_{j'})$ is a surviving pair in the profile decomposition (10), then it comes along with a family of optimizers $\pi(c_{j,j'})(f_j + e^{i\theta} f_{j'})$ with $\theta \in [0, 2\pi]$ satisfying the following constraints:

$$Wf_j(z) = -Wf_{j'}(z) \quad \text{for a.e. } z \in \Omega - c_{j,j'}, \qquad \langle f_j, f_{j'}\rangle = 0.$$

The optimizing contribution on Ω is then purely due to constructive interferences, since

$$\big\|W(\pi(c_{j,j'})(f_j + e^{i\theta} f_{j'}))\big\|_{L^p(\Omega)} = \big\|2\,\mathrm{Re}(e^{-i\theta} W(f_j, f_{j'}))\big\|_{L^p(\Omega - c_{j,j'})}.$$

While these conditions seem not to be a priori incompatible, it is evident that they place extremely rigid constraints on the exotic optimizers. Still, they are not strong enough to completely rule out the occurrence of such optimizers and have so far eluded our attempts to exhibit examples of fitting functions. Nevertheless, we can claim that the occurrence of exotic optimizers is rare in the grand scheme. More precisely, the subset of pairs of exotic optimizers is meagre in $(L^2(\mathbb{R}^d))^2$, being a subset of the set $\mathcal{N}$ of optimizers for the inequality (7), which is in turn a subset of the set $\mathcal{O}$ of pairs of orthogonal functions in $(L^2(\mathbb{R}^d))^2$ as a consequence of Corollary 3.5. It is straightforward to prove that the latter is a nowhere dense subset, being indeed closed with empty interior. In particular, for every $(f, g) \in \mathcal{O}$ one can find a pair of non-zero directions $h, k \in L^2(\mathbb{R}^d)$ such that the family $(f + th, g + tk)$, $t > 0$, belongs to $\mathcal{O}^c$ while converging to (f, g).

Remark 4.1. In fact, by refining the previous argument it is possible to show that $\mathcal{O}$, hence the exceptional set $\mathcal{N}$, satisfies pretty much any other meaningful notion of negligibility, such as being directionally porous, which in turn implies being shy (or equivalently Haar null) and Gaussian null. We avoid such a technical discussion and address the interested reader to [37] for further details.

We are thus left with some open issues. Theorem 1.1 establishes the existence of an optimizer, but examining the proof reveals several candidates for the optimizer. As such, it is natural to ask the following questions related to uniqueness of the optimizer.

Question 1. Are we guaranteed that a compact profile f_{j_0} exists? Are the properties provided in Section 4.3 actually incompatible?

5. The L^∞ concentration problem

In the case $p = \infty$, the existence of an optimizer is significantly easier to establish.

Proposition 5.1. *Let $\Omega \subset \mathbb{R}^{2d}$ be measurable such that $0 < |\Omega| < \infty$. We have*

$$\sup_{f \in L^2(\mathbb{R}^d)\setminus\{0\}} \frac{\|Wf\|_{L^\infty(\Omega)}}{\|f\|_{L^2}^2} = 2^d, \tag{12}$$

and the supremum is attained.

Proof. Recall from (3) the L^∞ estimate

$$\|Wg\|_{L^\infty(\Omega)} \leq \|Wg\|_{L^\infty} \leq 2^d \|g\|_{L^2}^2.$$

It is clear that $\|Wg\|_{L^\infty}$ is attained at some $w_0 \in \mathbb{R}^{2d}$, since $Wg \in C_0(\mathbb{R}^{2d})$ for $g \in L^2(\mathbb{R}^d)\setminus\{0\}$. Pick a point of positive Lebesgue density $c \in \Omega$ (recall that Ω has positive measure) and set $f = \pi(c - w_0)g$. Therefore, $\|Wf\|_{L^\infty(\Omega)} \geq |Wf(c)| = |Wg(w_0)| = \|Wg\|_{L^\infty}$, and since by covariance, $|Wg(w_0)| = |W(\pi(-w_0)g)(0)|$, we have

$$\sup_{f \in L^2(\mathbb{R}^d)\setminus\{0\}} \frac{\|Wf\|_{L^\infty(\Omega)}}{\|f\|_{L^2}^2} = \sup_{g \in L^2(\mathbb{R}^d)\setminus\{0\}} \frac{\|Wg\|_{L^\infty}}{\|g\|_{L^2}^2} = \sup_{g \in L^2(\mathbb{R}^d)\setminus\{0\}} \frac{|Wg(0)|}{\|g\|_{L^2}^2}.$$

If we let g be an even function, then the claim follows, since

$$\begin{aligned} Wg(0) &= \int_{\mathbb{R}^d} g(y/2)\overline{g(-y/2)}\,\mathrm{d}y \\ &= \int_{\mathbb{R}^d} |g(y/2)|^2\,\mathrm{d}y \\ &= 2^d \int_{\mathbb{R}^d} |g(u)|^2\,\mathrm{d}u = 2^d \|g\|_{L^2}^2. \quad \square \end{aligned}$$

Remark 5.2. Notice that the same argument works also when g is odd, since $Wg(0) = -2^d\|g\|_{L^2}^2$, hence the set of optimizers for (12) is quite large. On the other hand, there exist non-zero functions whose concentration functional vanishes. For instance, assuming that Ω is bounded with $R \coloneqq \operatorname{diam}(\Omega)$, if f is supported on $[R+1, R+2]$ then the support property of the Wigner distribution in $d = 1$ [17, Lemma 4.3.5] implies that $Wf(x,\xi) = 0$ whenever $|x| \leq R$. Therefore, for any $c = (c_x, c_\xi) \in \Omega$ and any $z = (x,\xi) \in \Omega$ we have $|x - c_x| \leq |z - c| \leq R$, hence $W(\pi(c)f)(z) = Wf(z - c) = 0$ for all $z \in \Omega$. As a result, $\|W(\pi(c)f)\|_{L^\infty(\Omega)} = 0$.

Remark 5.3. Compared to the case $p \in [1,\infty)$, where the conclusion is the same for the Wigner distribution as for the ambiguity function, the $p = \infty$ case is strikingly different. For the ambiguity function, the supremum is only attained if $0 \in \Omega$, but if this is the case, every function will be an optimizer [30, Proposition 1.2]. On the other hand, the supremum is always attained for the Wigner distribution, but for a function to be an optimizer it has to satisfy very specific symmetry conditions. This further illustrates the difference between the time-frequency covariant Wigner concentration and the time-frequency invariant ambiguity concentration problems.

6. Generalizations and open problems

In this section we discuss some variations on the main problem arising from considering different yet related time-frequency distributions, namely the τ-Wigner distributions and the Born–Jordan distribution. As in the proof of Theorem 1.1, we approach the corresponding problems via profile decomposition. We carry the strategy as far as possible, but certain crucial steps differ from the Wigner case and call for suitable adjustments, leading to very different outcomes: Indeed, we cannot conclude the existence of an optimizer (which remains an open problem for the time being) because the structure of maximizing profiles is significantly more complex (see Lemma 6.2) when $\tau \neq 1/2$, whereas for the Born–Jordan problem (roughly speaking, after averaging over τ) we are even able to unlock the direct method by proving full weak continuity.

6.1. Concentration problems for τ-Wigner distributions

Recall that the cross-τ-*Wigner distributions* [1,7] are defined for $f, g \in L^2(\mathbb{R}^d)$, $\tau \in (0,1)$ by

$$W_\tau(f,g)(z) = \int_{\mathbb{R}^d} e^{-2\pi i \xi\cdot y} f(x+\tau y)\overline{g(x-(1-\tau)y)}\,\mathrm{d}y.$$

Letting $\tau = 1/2$ recovers the Wigner distribution. Because of the similar structure, the τ-Wigner distributions inherit many of the important properties of the Wigner distribution, for instance covariance under time-frequency shifts: this reads (cf. [6, Proposition 3.1])

$$W_\tau(\pi(a)f, \pi(b)g)(z) = e^{2\pi i\,[z,a-b]} e^{2\pi i (x_a - x_b)\cdot(\tau\xi_a + (1-\tau)\xi_b)} W_\tau(f,g)\big(z - c_\tau(a,b)\big), \tag{13}$$

where we set

$$c_\tau(a,b) \coloneqq \big((1-\tau)x_a + \tau x_b\,,\, \tau\xi_a + (1-\tau)\xi_b\big), \qquad a = (x_a,\xi_a),\ b = (x_b,\xi_b).$$

Moreover, Theorem 2.2 also extends to W_τ.

Theorem 6.1 *([7, Theorem 4.12]). Let $f \in L^2(\mathbb{R}^d)$, $g \in M^\infty(\mathbb{R}^d)$. Then $W_\tau(f,g) \in W(L^2, L^\infty)(\mathbb{R}^{2d})$ for every $\tau \in (0,1)$, with*

$$\|W_\tau(f,g)\|_{W(L^2,L^\infty)} \lesssim_d C(\tau)\|f\|_{L^2}\|g\|_{M^\infty},$$

where $C(\tau) = \dfrac{2^d}{(\tau(1-\tau))^{d/2}}$.

It is therefore natural to ask whether the conclusion of Theorem 1.1 holds for the other τ-Wigner distributions as well. By the elementary L^∞-bound

$$\|W_\tau(f,g)\|_\infty \leq (\tau(1-\tau))^{-d/2}\|f\|_{L^2}\|g\|_{L^2}$$

it is clear that the supremum

$$L_\tau = \sup_{f\in L^2(\mathbb{R}^d)\setminus\{0\}} \frac{\big(\int_\Omega |W_\tau f(z)|^p\,\mathrm{d}z\big)^{1/p}}{\|f\|_{L^2}^2}$$

is finite (and of course positive). In addition, the same argument as in the proof of Proposition 3.6 shows that the corresponding concentration functional is not weakly upper semicontinuous at any point of $L^2(\mathbb{R}^d)$. We therefore again consider profile decomposition as in Lemma 2.3, leading to a similar expansion as in (10), albeit with W_τ in place of W. In light of Theorem 6.1 the error terms can be controlled using exactly the same argument, and the covariance property of W_τ shows that all pure terms vanish, possibly except for the compact profile f_{j_0}.

Due to the asymmetry of W_τ, though, the cross terms exhibit a less controlled behavior. Indeed, the expansion of $W_\tau(f^{(n)})$ in (10) naturally leads to grouping cross terms into symmetrized blocks: for each pair of distinct profile indices (j, j'), we set

$$\mathcal{S}^{(n)}_{j,j'}(z) \coloneqq W_\tau(\pi(z^{(n)}_j)f_j, \pi(z^{(n)}_{j'})f_{j'})(z) + W_\tau(\pi(z^{(n)}_{j'})f_{j'}, \pi(z^{(n)}_j)f_j)(z).$$

The asymptotic behavior of this block depends critically on the parameter τ. By virtue of covariance, setting $z^{(n)}_j = (x^{(n)}_j, \xi^{(n)}_j)$ and $z^{(n)}_{j'} = (x^{(n)}_{j'}, \xi^{(n)}_{j'})$ we have:

- $W_\tau(\pi(z^{(n)}_j)f_j, \pi(z^{(n)}_{j'})f_{j'})$ is centered at

$$c^{(n)}_{j,j'} \coloneqq c_\tau(z^{(n)}_j, z^{(n)}_{j'}) = \big((1-\tau)\, x^{(n)}_j + \tau\, x^{(n)}_{j'},\ \ \tau\, \xi^{(n)}_j + (1-\tau)\, \xi^{(n)}_{j'}\big).$$

- $W_\tau(\pi(z^{(n)}_{j'})f_{j'}, \pi(z^{(n)}_j)f_j)$ is centered at

$$c^{(n)}_{j',j} \coloneqq c_\tau(z^{(n)}_{j'}, z^{(n)}_j) = \big((1-\tau)\, x^{(n)}_{j'} + \tau\, x^{(n)}_j,\ \ \tau\, \xi^{(n)}_{j'} + (1-\tau)\, \xi^{(n)}_j\big).$$

Hence

$$c^{(n)}_{j,j'} - c^{(n)}_{j',j} = (1-2\tau)\,\big(x^{(n)}_j - x^{(n)}_{j'},\ -(\xi^{(n)}_j - \xi^{(n)}_{j'})\big).$$

Setting $d^{(n)}_{j,j'} \coloneqq z^{(n)}_j - z^{(n)}_{j'}$, we thus have

$$\big|c_\tau(z^{(n)}_j, z^{(n)}_{j'}) - c_\tau(z^{(n)}_{j'}, z^{(n)}_j)\big| = |1-2\tau|\,|d^{(n)}_{j,j'}|,$$

and since $|d^{(n)}_{j,j'}| \to \infty$ by construction we face a significant dichotomy:

- If $\tau = 1/2$ the centers coincide and the block coalesces into an oscillatory function as before:

$$\mathcal{S}^{(n)}_{j,j'} = 2\,\mathrm{Re}(W(\pi(z^{(n)}_j)f_j, \pi(z^{(n)}_{j'})f_{j'})).$$

- If $\tau \neq 1/2$ the centers separate along $|d^{(n)}_{j,j'}|$ and the two contributions decouple in phase space.

In any case, for the block $\mathcal{S}^{(n)}_{j,j'}$ to have a non-vanishing contribution to the $L^p(\Omega)$ norm in the limit, at least one of its two centers, $c^{(n)}_{j,j'}$ or $c^{(n)}_{j',j}$, must remain bounded as $n \to \infty$. Just like for the pure terms, only those blocks $\mathcal{S}^{(n)}_{j,j'}$ for which at least one of the two centers $c^{(n)}_{j,j'}, c^{(n)}_{j',j}$ remains bounded as $n \to \infty$ can contribute in the limit, and in that case we refer to $\mathcal{S}^{(n)}_{j,j'}$ as a *surviving block*. More precisely, we say the ordered pair (j,k) is *surviving* if the center sequence $\{c^{(n)}_{j,k}\}_n$ is bounded in $\mathbb{R}^{2d}$, while a *chain* is a (finite or infinite) sequence $(j_1, j_2, \dots)$ of distinct indices such that each consecutive ordered pair (j_r, j_{r+1}) is surviving. We stress that when $\tau \neq 1/2$ the centers $c^{(n)}_{j,j'}$ and $c^{(n)}_{j',j}$ diverge from one another, hence for any surviving block exactly one of these two center sequences can be bounded, while for $\tau = 1/2$ the two centers coincide.

These findings show that the structure of the graph of surviving indices is richer when $\tau \neq 1/2$. Indeed, simple generalizations of the arguments in the case $\tau = 1/2$ provide the following conclusions.

Lemma 6.2. *Fix $\tau \in (0,1)$ and consider the directed graph on profile indices where a directed edge (j,k) is placed for each and only surviving pair. Then:*

1. *A member of a surviving pair cannot be a compact profile: if (j,k) is surviving then $z_j^{(n)}$ and $z_k^{(n)}$ are unbounded sequences.*
2. *Each node has in-degree and out-degree ≤ 1: for each j there is at most one k with (j,k) surviving, and at most one ℓ with (ℓ,j) surviving. In particular, when $\tau = 1/2$ an index can belong to at most one surviving pair.*
3. *The graph is acyclic if $\tau \neq 1/2$.*

Therefore, the surviving-edge graph is a disjoint union of chains, that for $\tau = 1/2$ reduces to a disjoint union of undirected edges (a matching).

These findings illustrate that when $\tau \neq 1/2$ and chains do appear in a profile decomposition, the main obstruction in the main chaining argument $L_\tau \leq \cdots \leq L_\tau$, is avoiding energy double-counting. This issue can be ultimately fixed by combining the definition of L_τ with the asymptotic orthogonality of the extracted profiles: putting technicalities aside, if $\Pi = (j_1, j_2, \dots)$ is a (finite or infinite) chain from the profile decomposition then one has

$$\limsup_{n\to\infty} \Big\| W_\tau\Big(\sum_{j\in\Pi} \pi(z_j^{(n)}) f_j\Big)\Big\|_{L^p(\Omega)} \;\leq\; L_\tau \sum_{j\in\Pi} \|f_j\|_{L^2}^2.$$

Therefore, the same argument used for the Wigner distribution shows that if the compact profile $f_{j_0} \neq 0$, then an optimizer exists, but when $f_{j_0} = 0$ we cannot just claim that a surviving block is an asymptotic optimizer. Indeed, the Wigner proof crucially relied on Theorem 3.1, where the limit of the block contribution was explicitly computed and bounded. We do not have a sufficiently strong parallel result for chains, leaving open the existence of optimizers under this strategy.

Question 2. Is there an analogue of Theorem 3.1 for chains? Can we somehow guarantee that a compact profile exists?

On the other hand, we are able to obtain a full answer for the L^∞ optimization problem, marking once again the exceptional nature of the case $\tau = 1/2$.

Proposition 6.3. *Let $\Omega \subset \mathbb{R}^{2d}$ be measurable such that $0 < |\Omega| < \infty$, and $\tau \in (0,1) \setminus \{1/2\}$. We have*

$$\sup_{f\in L^2(\mathbb{R}^d)\setminus\{0\}} \frac{\|W_\tau f\|_{L^\infty(\Omega)}}{\|f\|_{L^2}^2} = \left(\frac{1}{\tau(1-\tau)}\right)^{d/2},$$

but the supremum is not attained.

Proof. Arguing as in the Wigner case, the problem boils down to computing

$$\sup_{g\in L^2(\mathbb{R}^d)\setminus\{0\}} \frac{|W_\tau g(0)|}{\|g\|_{L^2}^2}.$$

Since

$$W_\tau g(0) = \int_{\mathbb{R}^d} g(\tau y)\overline{g(-(1-\tau)y)}\,\mathrm{d}y = (\tau(1-\tau))^{-d/2}\langle D_\tau g,\, D_{-(1-\tau)}g\rangle,$$

where $D_a g(y) = |a|^{d/2} g(ay)$ is the unitary dilation by $a \neq 0$, we infer

$$\sup_{g\in L^2(\mathbb{R}^d)\setminus\{0\}} \frac{|\langle D_\tau g, D_{-(1-\tau)}g\rangle|}{\|g\|_{L^2}^2} = \sup_{\|g\|_{L^2}=1} |\langle D_{-\frac{\tau}{1-\tau}} g, g\rangle| = \|D_{-\frac{\tau}{1-\tau}}\|_{L^2\to L^2} = 1.$$

We have invoked [32, Theorem 12.25] in the second equality. On the other hand, if the supremum were attained it would force $D_{-\frac{\tau}{1-\tau}} g = \lambda g$ a.e. for some $\lambda \in \mathbb{C}$ with $|\lambda| = 1$, which is impossible for $g \in L^2(\mathbb{R}^d) \setminus \{0\}$ unless $\left|\frac{\tau}{1-\tau}\right| = 1$, that is $\tau = 1/2$. □

6.2. The Born–Jordan concentration problem

Recall that for $f, g \in L^2(\mathbb{R}^d)$ the Born–Jordan distribution is defined as the τ-average of the τ-Wigner distributions:

$$W_{\mathrm{BJ}}(f,g) \;=\; \int_0^1 W_\tau(f,g)\;\mathrm{d}\tau.$$

Equivalently, it can be viewed as the Cohen class distribution with kernel

$$W_{\mathrm{BJ}}(f,g) = W(f,g) * \Theta_\sigma, \qquad \Theta(\zeta) = \frac{\sin(\pi\,\zeta_1\cdot\zeta_2)}{\pi\,\zeta_1\cdot\zeta_2}.$$

While the Born–Jordan distribution comes with its own set of problems stemming from the bad regularity/decay properties of Θ, it is intuitively clear that symmetry over τ should lead to a better behavior when it comes to phase space concentration compared to the τ-Wigner distributions. In fact, this is consistent with the findings on the reduction of cross- and self-interference fringes compared to the Wigner distribution, cf. again [5]. Furthermore, using the profile decomposition approach we are able to put these heuristics on a rigorous ground by showing a regularity gain of the Born–Jordan distribution's concentration functional compared to the Wigner distribution's. In particular, we show the following continuity result for $d = 1$.

Proposition 6.4. *Let $d = 1$, $1 \leq p < \infty$, and $\Omega \subset \mathbb{R}^2$ be measurable with $0 < |\Omega| < \infty$. The functional $J_{\mathrm{BJ}}(f) \coloneqq \|W_{\mathrm{BJ}} f\|_{L^p(\Omega)}$ is sequentially weakly continuous in $L^2(\mathbb{R})$.*

Although it is well known that the averaging process giving rise to the Born–Jordan distribution smooths in the sense that cross- and self-interference fringes are reduced, Proposition 6.4 shows that this averaging process even improves regularity on a topological level. This observation could potentially have much broader implications in time-frequency analysis, which we plan to explore in future work, but as far as the scope of this paper goes, Proposition 6.4 shows that when $d = 1$ the Born–Jordan concentration problem has a positive answer.

Theorem 6.5. *Let $\Omega \subset \mathbb{R}^2$ be measurable with $0 < |\Omega| < \infty$ and $1 \leq p < \infty$. Then*

$$L_{\mathrm{BJ}} = \sup_{f\in L^2(\mathbb{R})\setminus\{0\}} \frac{\left(\int_\Omega |W_{\mathrm{BJ}} f(z)|^p\,\mathrm{d}z\right)^{1/p}}{\|f\|_{L^2}^2}$$

is attained by an optimizer.

Let us note right away that by Cauchy-Schwarz's inequality we have

$$\|W_{\mathrm{BJ}}f\|_{L^\infty(\mathbb{R}^2)} \leq \int_0^1 \|W_\tau f\|_{L^\infty(\mathbb{R}^2)}\ \mathrm{d}\tau \leq \Big(\int_0^1 \frac{1}{\sqrt{\tau(1-\tau)}}\ \mathrm{d}\tau\Big)\|f\|^2_{L^2(\mathbb{R})} = \pi\|f\|^2_{L^2(\mathbb{R})},$$

so L_{BJ} is finite (and of course positive). Thus, by the direct method, Theorem 6.5 follows in a straightforward way once Proposition 6.4 is proved. To this aim, we will rely on several properties of the Born–Jordan distribution that are now outlined. We start with the covariance formula, which follows by straightforward integration of the corresponding one for τ-Wigner distributions.

Proposition 6.6. *For all $a = (x_a, \xi_a), b = (x_b, \xi_b), z \in \mathbb{R}^2$ and $f, g \in L^2(\mathbb{R})$,*

$$\begin{aligned} W_{\mathrm{BJ}}(\pi(a)f, \pi(b)g)(z) &= \int_0^1 W_\tau(\pi(a)f, \pi(b)g)(z)\ \mathrm{d}\tau \\ &= e^{2\pi i[z, a-b]} \int_0^1 e^{2\pi i(x_a - x_b)\cdot(\tau\xi_a + (1-\tau)\xi_b)}\, W_\tau(f,g)\big(z - c_\tau(a,b)\big)\ \mathrm{d}\tau, \end{aligned}$$

where $c_\tau(a,b) = ((1-\tau)x_a + \tau x_b, \tau\xi_a + (1-\tau)\xi_b)$. In particular, for joint shifts ($a = b$) we have the usual translation covariance

$$W_{\mathrm{BJ}}(\pi(a)f, \pi(a)g)(z) = W_{\mathrm{BJ}}(f,g)(z-a), \qquad W_{\mathrm{BJ}}(\pi(a)f)(z) = W_{\mathrm{BJ}}f(z-a).$$

Let us now substantiate the reduction of interferences by proving that cross-terms do not contribute to the concentration problem.

Lemma 6.7. *Let $\Omega \subset \mathbb{R}^2$ have finite measure and $1 \leq p < \infty$. Fix $f, g \in L^2(\mathbb{R})$ and sequences $a^{(n)}, b^{(n)} \in \mathbb{R}^2$ with $|a^{(n)} - b^{(n)}| \to \infty$. Then*

$$\lim_{n\to\infty} \|W_{\mathrm{BJ}}(\pi(a^{(n)})f, \pi(b^{(n)})g)\|_{L^p(\Omega)} = 0, \qquad \lim_{n\to\infty} \|W_{\mathrm{BJ}}(\pi(b^{(n)})g, \pi(a^{(n)})f)\|_{L^p(\Omega)} = 0.$$

Proof. Let $\mathcal{B}_n(\tau;\cdot) \coloneqq W_\tau(\pi(a^{(n)})f, \pi(b^{(n)})g)(\cdot)$. By (13),

$$\|\mathcal{B}_n(\tau;\cdot)\|_{L^p(\Omega)} = \|W_\tau(f,g)(\cdot - c_\tau(a^{(n)}, b^{(n)}))\|_{L^p(\Omega)}.$$

We claim that there is at most one $\tau_* \in [0,1]$ such that $\{c_{\tau_*}(a^{(n)}, b^{(n)})\}_n$ is bounded — as a result, for any fixed $\tau \neq \tau_*$ we have $|c_\tau(a^{(n)}, b^{(n)})| \to \infty$ as $n \to \infty$. Indeed, the following identity follows directly from the definition of c_τ:

$$c_{\tau_1}(a,b) - c_{\tau_2}(a,b) = (\tau_1 - \tau_2)\,\big(x_b - x_a,\ -(\xi_b - \xi_a)\big) = (\tau_1 - \tau_2)\,R\,(b-a), \tag{14}$$

with $R = \mathrm{diag}(I_x, -I_\xi)$ invertible. Consequently, if two distinct $\tau_1 \neq \tau_2$ both yield bounded center sequences along a subsequence, then (14) implies $(\tau_1 - \tau_2)\,R\,(b^{(n)} - a^{(n)})$ is bounded, contradicting $|a^{(n)} - b^{(n)}| \to \infty$.

Therefore, for any fixed $\tau \neq \tau_*$ we have $|c_\tau(a^{(n)}, b^{(n)})| \to \infty$. The standard argument from the proof of Theorem 3.1 thus shows that

$$\|\mathcal{B}_n(\tau;\cdot)\|_{L^p(\Omega)} \xrightarrow[n\to\infty]{} 0 \qquad \text{for each fixed } \tau \neq \tau_*.$$

We wish to invoke dominated convergence with respect to τ, so that

$$\|W_{\mathrm{BJ}}(\pi(a^{(n)})f, \pi(b^{(n)})g)\|_{L^p(\Omega)} \leq \int_0^1 \|\mathcal{B}_n(\tau;\cdot)\|_{L^p(\Omega)}\, \mathrm{d}\tau \xrightarrow[n\to\infty]{} 0.$$

All we need is to find $h \in L^1(0,1)$ such that $\|\mathcal{B}_n(\tau;\cdot)\|_{L^p(\Omega)} \lesssim h(\tau)$ for all $\tau \in (0,1)$ and all n. To this aim, when $1 \leq p \leq 2$ we can resort to the embedding $L^2(\Omega) \subset L^p(\Omega)$, Hölder's inequality and [7, Theorem 4.14] to obtain

$$\|\mathcal{B}_n(\tau;\cdot)\|_{L^p(\Omega)} \lesssim |\Omega|^{1/p-1/2}\|\mathcal{B}_n(\tau;\cdot)\|_{L^2} \lesssim |\Omega|^{1/p-1/2}\|f\|_{L^2}\|g\|_{L^2}.$$

When $p > 2$ we simply interpolate:

$$\begin{aligned}\|\mathcal{B}_n(\tau;\cdot)\|_{L^p(\Omega)} &\leq \|\mathcal{B}_n(\tau;\cdot)\|_{L^2(\Omega)}^{2/p}\|\mathcal{B}_n(\tau;\cdot)\|_{L^\infty(\Omega)}^{1-2/p} \\ &\lesssim (\tau(1-\tau))^{-\frac{1}{2}(1-\frac{2}{p})}\|f\|_{L^2}\|g\|_{L^2},\end{aligned}$$

and the resulting function is summable on $(0,1)$ for every $p > 2$. The analogous inequality follows from the identity $W_{\mathrm{BJ}}(f,g) = \overline{W_{\mathrm{BJ}}(g,f)}$. □

We can now prove the continuity result for W_{BJ}.

Proof of Proposition 6.4. Set $h^{(n)} := f^{(n)} - f$, so $h^{(n)} \rightharpoonup 0$ in $L^2(\mathbb{R})$. By bilinearity,

$$W_{\mathrm{BJ}}(f^{(n)}) = W_{\mathrm{BJ}}(f) + W_{\mathrm{BJ}}(h^{(n)}) + 2\,\mathrm{Re}(W_{\mathrm{BJ}}(f, h^{(n)})).$$

It is therefore enough to prove that

$$\|W_{\mathrm{BJ}}(h^{(n)})\|_{L^p(\Omega)} \to 0 \qquad \text{and} \qquad \|W_{\mathrm{BJ}}(f, h^{(n)})\|_{L^p(\Omega)} \to 0.$$

Since $h^{(n)}$ is a bounded sequence, we can without loss of generality assume $\|h^{(n)}\|_{L^2} \leq 1$ for all n.

Step 1. *Profile decomposition of $h^{(n)}$ and absence of compact profiles.* We invoke Lemma 2.3 and apply it to the bounded sequence $h^{(n)}$: For each fixed $k \in \mathbb{N}$ we can write

$$h^{(n)} = \sum_{j=1}^{k} \pi(z_j^{(n)})\phi_j + w_k^{(n)},$$

with

$$|z_j^{(n)} - z_{j'}^{(n)}| \to \infty \ \ (j \neq j'), \qquad \sum_{j=1}^{\infty} \|\phi_j\|_{L^2}^2 \leq 1, \qquad \lim_{k\to\infty} \limsup_{n\to\infty} \|w_k^{(n)}\|_{M^\infty} = 0.$$

Since $h^{(n)} \rightharpoonup 0$, no compact profile can occur: If some $z_{j_0}^{(n)}$ were bounded in $\mathbb{R}^{2d}$, then $\pi(z_{j_0}^{(n)})^* h^{(n)} \rightharpoonup \phi_{j_0} \neq 0$, contradicting $h^{(n)} \rightharpoonup 0$. Hence

$$|z_j^{(n)}| \to \infty \quad \text{for every } j \geq 1.$$

Step 2. *Vanishing of the profile contributions.* By covariance, inner regularity and the fact that $W_{\mathrm{BJ}}\phi_j \in C_0(\mathbb{R}^2)$ for $d = 1$ (see, e.g., [29, Example 6.1]), for each fixed j we infer

$$\|W_{\mathrm{BJ}}(\pi(z_j^{(n)})\phi_j)\|_{L^p(\Omega)} = \|W_{\mathrm{BJ}}\phi_j(\cdot - z_j^{(n)})\|_{L^p(\Omega)} \to 0.$$

Moreover, for $j \neq j'$ we have $|z_j^{(n)} - z_{j'}^{(n)}| \to \infty$, hence Lemma 6.7 yields for every $j \neq j'$

$$\lim_{n\to\infty} \|W_{\mathrm{BJ}}(\pi(z_j^{(n)})\phi_j, \pi(z_{j'}^{(n)})\phi_{j'})\|_{L^p(\Omega)} = 0.$$

Summing over $j = 1, \dots, k$ and then letting k be fixed, we therefore get

$$\lim_{n\to\infty} \Big\| W_{\mathrm{BJ}}\Big(\sum_{j=1}^{k} \pi(z_j^{(n)})\phi_j \Big) \Big\|_{L^p(\Omega)} = 0.$$

Step 3. *Remainder terms.* Fix a compact set $K \subset \mathbb{R}^2$ with $|\Omega \setminus K|$ arbitrarily small. The terms accounting for the remainders $w_k^{(n)}$ vanish by an argument similar to the one already used before, since by Theorem 6.1 we have

$$\begin{aligned} \|W_{\mathrm{BJ}}(f,g)\|_{L^2(K)} &\lesssim_K \int_0^1 \|W_\tau(f,g)\|_{W(L^2,L^\infty)}\,\mathrm{d}\tau \\ &\lesssim \Big(\int_0^1 (\tau(1-\tau))^{-1/2}\,\mathrm{d}\tau \Big) \|f\|_{L^2}\|g\|_{M^\infty} \\ &\lesssim \|f\|_{L^2}\|g\|_{M^\infty}. \end{aligned}$$

As such, arguing as in the proof of Theorem 1.1 this bound is sufficient to show that

$$\lim_{k\to\infty} \limsup_{n\to\infty} \|W_{\mathrm{BJ}}(F_k^{(n)}, w_k^{(n)})\|_{L^p(\Omega)} = 0,$$

and analogously for $W_{\mathrm{BJ}}(w_k^{(n)}, F_k^{(n)})$ and $W_{\mathrm{BJ}}(w_k^{(n)})$.

Step 4. *Conclusion.* Expanding $W_{\mathrm{BJ}}(h^{(n)})$ by sesquilinearity into pure, cross and remainder terms, then combining the results in the previous steps, after letting $n \to \infty$ first and then $k \to \infty$ (which we can do by the same reasoning as in the proof of Theorem 1.1), we obtain

$$\lim_{n\to\infty} \|W_{\mathrm{BJ}}(h^{(n)})\|_{L^p(\Omega)} = 0.$$

The same argument allows us to deal with $W_{\mathrm{BJ}}(f, h^{(n)})$, where one can resort to Lemma 6.7 with the fixed packet f (i.e., take $a^{(n)} = z_j^{(n)}$, $b^{(n)} = 0$ so that $|a^{(n)} - b^{(n)}| \to \infty$) to get

$$\lim_{n\to\infty} \|W_{\mathrm{BJ}}(f, \pi(z_j^{(n)})\phi_j)\|_{L^p(\Omega)} = 0 \qquad \text{for each fixed } j.$$

Therefore, the claim $\|W_{\mathrm{BJ}} f^{(n)} - W_{\mathrm{BJ}} f\|_{L^p(\Omega)} \to 0$ is proved. □

We conclude this section by detailing the behavior of the L^∞ concentration problem.

Proposition 6.8. *Let $\Omega \subset \mathbb{R}^2$ be measurable such that $0 < |\Omega| < \infty$. Then*

$$\sup_{f \in L^2(\mathbb{R})\setminus\{0\}} \frac{\|W_{\mathrm{BJ}} f\|_{L^\infty(\Omega)}}{\|f\|_{L^2}^2} = \pi,$$

but the supremum is not attained.

Proof. The Born–Jordan distribution is covariant and when $d = 1$ we have $W_{\mathrm{BJ}} f \in C_0(\mathbb{R}^2)$ for all $f \in L^2(\mathbb{R})$ (see [29, Example 6.1]). Therefore, we mimic the proof (and notation) of Proposition 6.3, so without loss of generality we can assume that the ∞-norm is attained at $z = 0$ and we need to compute

$$\sup_{f \in L^2(\mathbb{R}) \setminus \{0\}} \frac{|W_{\mathrm{BJ}} f(0)|}{\|f\|_{L^2}^2}.$$

Now, with the substitution $s = \frac{\tau}{1-\tau}$ we have

$$W_{\mathrm{BJ}} f(0) = \int_0^1 (\tau(1-\tau))^{-1/2} \left\langle D_{-\frac{\tau}{1-\tau}} f, f \right\rangle \, \mathrm{d}\tau = \int_0^\infty \frac{s^{-1/2}}{1+s} \langle D_{-s} f, f \rangle \, \mathrm{d}s.$$

The upper bound $|W_{\mathrm{BJ}} f(0)| \leq \pi \|f\|_{L^2}^2$ is then clear.

If $Rf(y) = f(-y)$ denotes the reflection operator, after setting $s = e^t$ with $t \in \mathbb{R}$ and introducing the strongly continuous unitary group $G_t = D_{e^t}$, we can write

$$W_{\mathrm{BJ}} f(0) = \int_{\mathbb{R}} \frac{1}{2\cosh(t/2)} \langle G_t R f, f \rangle \, \mathrm{d}t = \langle Sf, f \rangle, \qquad f \in L^2(\mathbb{R}),$$

where we introduced the self-adjoint operator-valued integral

$$S \coloneqq \int_{\mathbb{R}} k(t) G_t R \, \mathrm{d}t, \qquad k(t) = \frac{1}{2\cosh(t/2)},$$

which is well defined since $\|S\|_{L^2 \to L^2} \leq \|k\|_{L^1} < \infty$.

Since R commutes with G_t, S leaves the subspaces of odd and even functions invariant. If T denotes the restriction to $L^2_e(\mathbb{R})$, the restriction to odd functions coincides with $-T$, hence $\|S\|_{L^2 \to L^2} = \max\{\|T\|_{L^2_e \to L^2}, \|-T\|_{L^2_o \to L^2}\} = \|T\|_{L^2_e \to L^2}$. In particular, we have

$$W_{\mathrm{BJ}} f(0) = \int_{\mathbb{R}} \frac{1}{2\cosh(t/2)} \langle G_t f, f \rangle \, \mathrm{d}t = \langle Tf, f \rangle, \qquad f \in L^2_e(\mathbb{R}).$$

Let us now introduce the unitary operator $U \colon L^2_e(\mathbb{R}) \to L^2(\mathbb{R})$ to be

$$(Uf)(y) = \sqrt{2}\, e^{y/2} f(e^y),$$

satisfying $U G_t U^{-1} = T_t$, that is G_t is unitarily equivalent to the translation group $(T_t f)(y) = f(y + t)$. Since U is unitary, G_t strongly continuous and $k \in L^1(\mathbb{R})$, we may pass U through the Bochner integral and obtain

$$UTU^{-1} f = \int_{\mathbb{R}} k(t) T_t f \, \mathrm{d}t = (k * f),$$

hence UTU^{-1} is the convolution operator with kernel k on $L^2(\mathbb{R})$. As such, we obtain by Plancherel's theorem

$$\|UTU^{-1} f\|_{L^2} = \|k * f\|_{L^2} = \|\widehat{k}\widehat{f}\|_{L^2} \leq \|\widehat{k}\|_{L^\infty} \|f\|_{L^2}.$$

An easy computation shows that $\widehat{k}(\xi) = \pi \operatorname{sech}(2\pi^2 \xi)$, whence

$$\|T\|_{L^2_e \to L^2} = \|UTU^{-1}\|_{L^2 \to L^2} = \|\widehat{k}\|_\infty = \widehat{k}(0) = \pi.$$

Non-attainment on even functions follows since equality in Hölder's inequality $\|\widehat{k}\widehat{f}\|_2 \leq \|\widehat{k}\|_\infty \|\widehat{f}\|_2$ with $f \in L^2_e(\mathbb{R}) \setminus \{0\}$ would force $|\widehat{k}| = \|\widehat{k}\|_\infty$ a.e. on the support of $\widehat{f}$, or equivalently $\widehat{f}$ to be supported where $|\widehat{k}| = \pi$, i.e. at $\{\xi = 0\}$, a null set, leading to $\widehat{f} = 0$. Since $S = -T$ on odd functions, non-attainment on the whole $L^2(\mathbb{R})$ follows. □

We would ideally like to extend our existence result for the Born–Jordan distribution to any dimension d. However, the Born–Jordan distribution decays significantly worse when $d > 1$. For instance, the integral $\int_0^1 C(\tau)\,\mathrm{d}\tau$ (cf. Theorem 6.1) needed to control the error terms in the profile decomposition diverges when $d > 1$. Similarly, the proof of Lemma 6.7 extends to higher dimension with a tradeoff between p and d: The claim still holds for all $1 \leq p < p_*(d)$, where

$$p_*(d) \coloneqq \begin{cases} \infty & (d = 1, 2) \\ \frac{2d}{d-2} & (d \geq 3). \end{cases}$$

Therefore, the contribution from the cross terms is 0 when $d < 3$, but for $d \geq 3$ the cross-term contribution is currently guaranteed to be 0 only provided that $p < p_*(d)$. As such, there is reason to believe that the Born–Jordan problem is only well-posed when $d = 1$.

Question 3. Is $L_{\mathrm{BJ}} = \infty$ when $d > 1$? What happens if we restrict to $f \in M^q(\mathbb{R}^d)$ with $1 \leq q < 2$ instead?

Acknowledgements

It is a pleasure to thank Franz Luef for discussions on the topics of this note. We also thank José Luis Romero for stimulating remarks on a first draft which led us to better highlight some results (namely, Propositions 3.6 and 6.4).

The second author thanks Politecnico di Torino for its hospitality while on a visit in Spring 2025 during which the work started.

The third author is a member of Gruppo Nazionale per l'Analisi Matematica, la Probabilità e le loro Applicazioni (GNAMPA) — Istituto Nazionale di Alta Matematica (INdAM).

References

[1] P. Boggiatto, G. De Donno, A. Oliaro, Time-frequency representations of Wigner type and pseudo-differential operators, Trans. Am. Math. Soc. 362 (9) (2010).
[2] A.J. Bracken, H.-D. Doebner, J.G. Wood, Bounds on integrals of the Wigner function, Phys. Rev. Lett. 83 (1999).
[3] L. Cohen, Time-Frequency Analysis, Prentice-Hall, 1995.
[4] L. Cohen, Uncertainty Principles of the Short-Time Fourier Transform, Advanced Signal Processing Algorithms, vol. 2563, SPIE, 1995, pp. 80–90.
[5] E. Cordero, M. de Gosson, F. Nicola, On the reduction of the interferences in the Born–Jordan distribution, Appl. Comput. Harmon. Anal. 44 (2) (2018) 230–245.
[6] E. Cordero, F. Nicola, S.I. Trapasso, Almost diagonalization of τ-pseudodifferential operators with symbols in Wiener amalgam and modulation spaces, J. Fourier Anal. Appl. 25 (4) (2019).
[7] E. Cordero, S.I. Trapasso, Linear perturbations of the Wigner distribution and the Cohen class, Anal. Appl. 18 (2020).
[8] M.A. de Gosson, Symplectic Methods in Harmonic Analysis and in Mathematical Physics, Birkhäuser/Springer Basel AG, Basel, 2011.
[9] M.A. de Gosson, Born–Jordan Quantization. Theory and Applications, Springer, Cham, 2016.
[10] H.G. Feichtinger, On a new Segal algebra, Monatshefte Math. 92 (4) (1981) 269–289.
[11] H.G. Feichtinger, Modulation spaces on locally compact abelian groups, Technical Report, University of Vienna, 1983.
[12] H.G. Feichtinger, Modulation spaces: looking back and ahead, Sampl. Theory Signal Image Process., Int. J. 5 (2) (2006) 109–140.
[13] H.G. Feichtinger, F. Luef, Wiener amalgam spaces for the fundamental identity of Gabor analysis, Collect. Math. 233 (253) (2006).

[14] P. Flandrin, Maximum signal energy concentration in a time-frequency domain, in: ICASSP-88, International Conference on Acoustics, Speech, and Signal Processing, vol. 4, 1988, pp. 2176–2179.
[15] J.J.F. Fournier, On the Hausdorff-Young theorem for amalgams, Monatshefte Math. 95 (2) (1983) 117–135.
[16] J.J.F. Fournier, J. Stewart, Amalgams of L^p and l^q, Bull. Am. Math. Soc. (N.S.) 13 (1) (1985) 1–21.
[17] K. Gröchenig, Foundations of Time-Frequency Analysis, Birkhäuser Boston, Inc., Boston, MA, 2001.
[18] K. Gröchenig, Uncertainty principles for time-frequency representations, in: Advances in Gabor Analysis, in: Appl. Numer. Harmon. Anal., Birkhäuser Boston, Boston, MA, 2003.
[19] C. Heil, An introduction to weighted Wiener amalgams, in: Wavelets and Their Applications, Allied Publishers, New Delhi, 2003, pp. 183–216.
[20] F. Hlawatsch, F. Auger, Time-Frequency Analysis: Concepts and Methods, Wiley, 2008.
[21] A.J.E.M. Janssen, Positivity and spread of bilinear time-frequency distributions, in: The Wigner Distribution, Elsevier Sci. B. V., Amsterdam, 1997, pp. 1–58.
[22] N. Lerner, Integrating the Wigner Distribution on Subsets of the Phase Space, a Survey, Memoirs of the European Mathematical Society, vol. 12, EMS Press, Berlin, 2024.
[23] E.H. Lieb, Integral bounds for radar ambiguity functions and Wigner distributions, J. Math. Phys. 31 (3) (1990).
[24] E.H. Lieb, Y. Ostrover, Localization of multidimensional Wigner distributions, J. Math. Phys. 51 (10) (2010).
[25] P.-L. Lions, The concentration-compactness principle in the calculus of variations. The limit case. I, Rev. Mat. Iberoam. 1 (1) (1985) 145–201.
[26] P.-L. Lions, The concentration-compactness principle in the calculus of variations. The limit case. II, Rev. Mat. Iberoam. 1 (2) (1985) 45–121.
[27] F. Luef, H. McNulty, Quantum time-frequency analysis and pseudodifferential operators, Indiana Univ. Math. J. (2026) 2026, https://doi.org/10.48550/arXiv.2403.00576, in press.
[28] F. Luef, E. Skrettingland, Convolutions for localization operators, J. Math. Pures Appl. 9 (2018) 118.
[29] F. Luef, E. Skrettingland, A Wiener Tauberian theorem for operators and functions, J. Funct. Anal. 280 (6) (2021) 108883.
[30] F. Nicola, J.L. Romero, S.I. Trapasso, On the existence of optimizers for time-frequency concentration problems, Calc. Var. Partial Differ. Equ. 62 (1) (2023).
[31] J. Ramanathan, P. Topiwala, Time-frequency localization via the Weyl correspondence, SIAM J. Math. Anal. 24 (5) (1993).
[32] W. Rudin, Functional Analysis, second edition, McGraw-Hill, New York, 1991.
[33] E.A.T. Svela, Extensions of Daubechies' theorem: Reinhardt domains, Hagedorn wavepackets and mixed-state localization operators, J. Funct. Anal. 290 (10) (2026) 111387.
[34] K. Tintarev, K.-H. Fieseler, Concentration Compactness, Imperial College Press, London, 2007.
[35] J. Ville, Théorie et applications de la notion de signal analytique, Câbles Transm. 2 (1948).
[36] E. Wigner, On the quantum correction for thermodynamic equilibrium, Phys. Rev. 40 (1932).
[37] L. Zajíček, On σ-porous sets in abstract spaces, Abstr. Appl. Anal. 5 (2005).