

L^2 -STABILITY FOR STFT PHASE RETRIEVAL

SUSANNA BERTOLINI, JAUME DE DIOS PONT, BEN PINEAU, JOÃO P. G. RAMOS,
AND MITCHELL A. TAYLOR

ABSTRACT. We prove that the short-time Fourier transform with Gaussian window performs L^2 -local stable phase retrieval at the constant function. The proof involved significant interplay between mathematicians and LLMs. An autoformalization in Lean 4 of an extension of our result to L^2 -local stable phase retrieval for all Hermite windows and all elements in the finite span of the canonical basis vectors is also presented.

1. INTRODUCTION

Short-time Fourier transform (STFT) phase retrieval refers to the problem of recovering an unknown function $f \in L^2(\mathbb{R}^d)$ from the phaseless measurement $|V_g f|$, where $V_g : L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^{2d})$ denotes the STFT with window function $g \in L^2(\mathbb{R}^d)$. This problem arises in ptychography, which is a widely used imaging technique that obtains information about an unknown object f by illuminating it with shifts of a probe g . Due to the high oscillation of the diffracted waves, practitioners can only record the far-field intensity $|V_g f|^2$, which leads to the phase retrieval problem described above.

The mathematics of STFT phase retrieval is a rich subject, with contributions coming from many different communities. Recall that, given a window function $g \in L^2(\mathbb{R}^d)$, the *short-time Fourier transform* $V_g : L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^{2d})$ is defined by

$$V_g f(x, \omega) = \mathcal{F}(T_x \bar{g} f)(\omega) = \int_{\mathbb{R}^d} f(t) \overline{g(t-x)} e^{-2\pi i t \cdot \omega} dt, \quad \text{for } x, \omega \in \mathbb{R}^d.$$

It is easy to see that

$$(1) \quad \|V_g f\|_{L^2(\mathbb{R}^{2d})} = \|g\|_{L^2(\mathbb{R}^d)} \|f\|_{L^2(\mathbb{R}^d)}$$

and the fundamental problem in STFT phase retrieval is to stably reconstruct $f \in L^2(\mathbb{R}^d)$ from the phaseless measurement $|V_g f|$, up to global phase ambiguity.

If stability issues are ignored, there is a standard method for determining whether the operator V_g does phase retrieval. Recall that the *ambiguity function* $\mathcal{A}f$ of a square-integrable function f is defined by $\mathcal{A}f(x, \omega) := e^{\pi i x \cdot \omega} V_f f(x, \omega)$. It is well-known that $\mathcal{A}f$ is a continuous function that determines every $f \in L^2(\mathbb{R})$ up to global phase. Moreover, we have the relation

$$(2) \quad \mathcal{F}(|V_g f|^2)(\omega, -x) = \mathcal{A}f(x, \omega) \overline{\mathcal{A}g(x, \omega)}.$$

For the Gaussian window g , it is easy to see that $\mathcal{A}g$ is again a Gaussian, so the above identity implies that V_g does phase retrieval. Similarly, since the ambiguity function of each Hermite window only vanishes on circles, these windows do STFT phase retrieval.

The stability problem for STFT phase retrieval has been studied by many authors [2, 4, 18, 21, 22, 23, 28]. The most natural notion of local stability is as follows.

Definition 1.1. Fix a window function $g \in L^2(\mathbb{R}^d)$. We say that V_g does local stable phase retrieval at $f \in L^2(\mathbb{R}^d)$ if there exists a constant $C(f) < \infty$ such that for all $h \in L^2(\mathbb{R}^d)$ we have

$$(3) \quad \inf_{|\lambda|=1} \|f - \lambda h\|_{L^2(\mathbb{R}^d)} \leq C(f) \||V_g f| - |V_g h|\|_{L^2(\mathbb{R}^{2d})}.$$

Despite extensive efforts on this problem, we are unaware of a single pair of functions $f, g \in L^2(\mathbb{R}^d) \setminus \{0\}$ for which V_g does local stable phase retrieval at f . On the other hand, for *any* window $g \in L^2(\mathbb{R}^d) \setminus \{0\}$, it is known [4, 5] that one can construct a dense set of functions $f \in L^2(\mathbb{R}^d)$ for which $C(f) = \infty$.

The objective of the present article is to show that there are many functions $f, g \in L^2(\mathbb{R}^d) \setminus \{0\}$ for which V_g does local stable phase retrieval at f in the sense of Definition 1.1. In fact, when g is the Gaussian or any other Hermite window, there is a canonical dense set of functions with $C(f) < \infty$.

1.1. History and preliminaries. Influential results of Cahill-Casazza-Daubechies [8] and Alaifari and Grohs [3] show that phase retrieval using frames cannot be uniformly stable, unless the Hilbert space \mathcal{H} is finite-dimensional. More precisely, whenever $T : \mathcal{H} \rightarrow L^2(\mu)$ is the analysis operator of a frame and $\dim \mathcal{H} = \infty$, we have $\sup_{f \in \mathcal{H}} C(f) = \infty$, where $C(f) \in [0, \infty]$ is the smallest constant satisfying

$$\inf_{|\lambda|=1} \|f - \lambda g\|_{\mathcal{H}} \leq C(f) \||Tf| - |Tg|\|_{L^2(\mu)} \text{ for all } g \in \mathcal{H}.$$

In fact, it was recently shown in [5] that we have $C(f) = \infty$ for a dense collection of $f \in \mathcal{H}$. In particular, in the natural setting of Definition 1.1, instabilities cannot be avoided.

In view of the negative results above, the question of *how* instabilities occur became fundamental. Importantly, the instabilities found in the aforementioned works can be viewed as mathematical artifacts rather than inherent issues with the experimental setup. Indeed, they all occur (in some sense) “at infinity”. This intuition has been made precise in the celebrated works of Alaifari-Daubechies-Grohs-Yin [2] and Grohs-Rathmair [22, 23], and more recently in [4]. However, all of these papers have the significant drawback that they only apply to the Gaussian window and (for technical reasons) they must replace the natural L^2 -norm on the right-hand side of (3) with a weighted first order Sobolev norm.

The lack of robustness when the window is changed is obviously an important issue, as the window is not precisely known in practice. However, the issue of changing the topology is equally problematic, as it makes the forward problem $f \mapsto |V_g f|$ ill-posed. In view of this, it has been an important open problem to understand when stability holds in the natural L^2 -norm and to what extent one may replace the Gaussian window by a more general window [29].

1.1.1. Reformulating the problem. Let $g \in L^2(\mathbb{R}^d)$ satisfy $\|g\|_{L^2(\mathbb{R}^d)} = 1$. In view of the identity (1), Definition 1.1 can be seen as a special case of the following unified problem by taking $E = V_g(L^2(\mathbb{R}^d)) \subseteq L^2(\mathbb{R}^{2d})$.

Definition 1.2. Let E be a subspace of $L^2(\Omega, \mu)$ for some measure space (Ω, μ) . We say that E does local stable phase retrieval at $f \in E$ if there exists a constant $C(f) < \infty$ such that for all $h \in E$ we have

$$\inf_{|\lambda|=1} \|f - \lambda h\|_{L^2(\mu)} \leq C(f) \||f| - |h|\|_{L^2(\mu)}.$$

In a pioneering work by Calderbank-Daubechies-Freeman-Freeman [9], infinite-dimensional subspaces E of real-valued $L^2(\mathbb{R}^d)$ have been constructed that satisfy $\sup_{f \in E} C(f) < \infty$. Similar examples over the complex field were later constructed in [11] and a general theory of stable phase retrieval in Banach lattices was initiated in [16]. This has led to a host of new and exciting results

[10, 19, 20] that lie in the intersection of various fields. When we state our main results below, we will use the language of Definition 1.2 rather than Definition 1.1.

1.1.2. *Function spaces.* For any measurable function $F: \mathbb{C} \rightarrow \mathbb{C}$, we define the *Fock norm*

$$\|F\|_{\mathcal{F}}^2 := \|F\|_{\gamma}^2 := \frac{1}{\pi} \int_{\mathbb{C}} |F(z)|^2 e^{-|z|^2} dm(z).$$

The *Bargmann–Fock space* $\mathcal{F}^2(\mathbb{C})$ is the closed subspace of entire functions with $\|F\|_{\mathcal{F}} < \infty$. The monomials $e_n(z) := z^n/\sqrt{n!}$ form an orthonormal basis of $\mathcal{F}^2(\mathbb{C})$, so every $F(z) = \sum_{n \geq 0} a_n z^n$ satisfies $\|F\|_{\mathcal{F}}^2 = \sum_{n \geq 0} |a_n|^2 n!$. With this convention, the Fock norm is defined for arbitrary measurable functions – not only holomorphic ones – and in particular we have $\| |F| \|_{\mathcal{F}} = \|F\|_{\mathcal{F}}$.

Up to universal factors, one may identify the Fock space with the image of the short-time Fourier transform with Gaussian window. Since elements of the Fock space are entire functions, it is obvious that phase retrieval holds in this subspace, i.e., for all $F, G \in \mathcal{F}^2(\mathbb{C})$ we have $|F| = |G|$ if and only if $F = \lambda G$ for some unimodular scalar λ . However, the question of stability is far more subtle. In fact, it was recently shown [4, 5] that there is a dense set of elements of $\mathcal{F}^2(\mathbb{C})$ that *fail* local stable phase retrieval (even with derivative loss). The main goal of this article is to construct a complementary dense set that *does* local stable phase retrieval in the natural \mathcal{F} -norm.

Our stability results apply not only to the Gaussian window, but to any Hermite window. In this case (again up to universal factors), the STFT maps $L^2(\mathbb{R})$ onto the *true polyanalytic Fock space*¹ $\mathcal{H}_k(\mathbb{C})$. Here, for any integer $k \geq 0$, we define $\mathcal{H}_k(\mathbb{C})$ as the $\|\cdot\|_{\gamma}$ -closed linear span of the basis

$$\Phi_{k,n}(z) = \frac{1}{\sqrt{k!n!}} \sum_{j=0}^{\min(k,n)} (-1)^j \binom{k}{j} \frac{n!}{(n-j)!} z^{n-j} \bar{z}^{k-j}, \quad n \geq 0.$$

Clearly, $\mathcal{H}_0(\mathbb{C}) = \mathcal{F}^2(\mathbb{C})$ is the Bargmann–Fock space and $\mathcal{H}_k(\mathbb{C})$ is a subspace of polyanalytic functions. By (2), $\mathcal{H}_k(\mathbb{C})$ performs phase retrieval. However, the major conceptual difference compared with the case $k = 0$ is that the ambiguity functions of the higher Hermite windows have non-trivial zero sets. A priori, this could affect the stability of the phase recovery process, since solving for $\mathcal{A}f$ in (2) would require division by the ambiguity function of the window.

1.1.3. *Higher dimensional generalizations.* The above definitions lift naturally to higher dimensions. Indeed, the d -dimensional basis consists of tensor products of the one-dimensional $\Phi_{k,n}$. To make this precise, we fix a dimension $d \geq 1$ and a multi-index \vec{k} . For each multi-index $\vec{\alpha}$, we define the function

$$\Phi_{\vec{k}, \vec{\alpha}}(\vec{z}) := \prod_{q=0}^{d-1} \Phi_{\vec{k}_q, \vec{\alpha}_q}(\vec{z}_q), \quad \vec{z} = (z_0, \dots, z_{d-1}).$$

The space $\mathcal{H}_{\vec{k}}(\mathbb{C}^d)$ is then defined as the closure of $\text{span}\{\Phi_{\vec{k}, \vec{\alpha}} : \vec{\alpha} \in \mathbb{N}^d\}$ in the norm

$$\|F\|_{\mathcal{F}_d}^2 := \|F\|_{\gamma_d}^2 := \frac{1}{\pi^d} \int_{\mathbb{C}^d} |F(\vec{z})|^2 e^{-\|\vec{z}\|^2} dm(\vec{z}).$$

1.2. **Main results.** The main results of this paper concern the L^2 -local stability of STFT phase retrieval for the canonical window functions at elements of the span of the canonical basis vectors. However, rather than stating one main result, we prefer to state the result in stages, following the path taken by the verification in the Lean 4 formal proof verification software [12].

The main result for which we will present a detailed proof is the following.

¹The full polyanalytic Fock space fails phase retrieval whenever $k \geq 1$, since $|\vec{z}| = |z|$. See, for example, [1] for a discussion on the true polyanalytic Fock space and its connections to time-frequency analysis.

Theorem 1.3 (Local SPR at $\mathbf{1}$ in $\mathcal{F}^2(\mathbb{C})$). *There exists an absolute constant $M > 0$ such that for any $F \in \mathcal{F}^2(\mathbb{C})$ we have*

$$\inf_{|\lambda|=1} \|F - \lambda\|_{\mathcal{F}} \leq M \| |F| - 1 \|_{\mathcal{F}}.$$

Theorem 1.3 follows from Theorem 1.4, a standard density argument, and a general fact about reproducing kernel Hilbert spaces that we will discuss below.

Theorem 1.4 (Local-local SPR at $\mathbf{1}$ in $\mathcal{F}^2(\mathbb{C})$). *There exist absolute constants $\delta > 0$ and $\widetilde{M} > 0$ such that the following holds. Let P be a polynomial with $\|P\|_{\mathcal{F}} \leq \delta$. Then there exists $w \in \mathbb{C}$ with $|w| = 1$ such that*

$$\|w(1 + P) - 1\|_{\mathcal{F}} \leq \widetilde{M} \| |1 + P| - 1 \|_{\mathcal{F}}.$$

One may take $\delta = 1/4601$ and $\widetilde{M} = 23003$.

We remark that the assumption that P is a polynomial in Theorem 1.4 is for convenience, as a standard density argument allows one to pass from a polynomial P to a general element $F \in \mathcal{F}^2(\mathbb{C})$. The reduction from Theorem 1.4 to Theorem 1.3 is a general RKHS fact that does not rely on any further structure of the Fock space. We refer the reader to Section 2.2 for this general result.

Theorem 1.3 can be vastly generalized using similar methods. Since the proof becomes much more technical but not more instructive, we will only sketch the main changes and provide a link to the full Lean verification. Our most general result establishes local stability for every Hermite window at every element in the linear span of the canonical basis vectors.

Theorem 1.5 (General stability for Hermite phase retrieval). *Fix $d \geq 1$ and a multi-index $\vec{\kappa}$. Let $P_0 \in \text{span}\{\Phi_{\vec{\kappa}, \vec{\alpha}}\}_{\vec{\alpha}}$ be an element of the finite linear span of the canonical basis vectors of $\mathcal{H}_{\vec{\kappa}}(\mathbb{C}^d)$. There exists a constant $M = M(d, \vec{\kappa}, P_0) > 0$ such that for any $F \in \mathcal{H}_{\vec{\kappa}}(\mathbb{C}^d)$ we have*

$$\inf_{|\lambda|=1} \|F - \lambda P_0\|_{\gamma_d} \leq M \| |F| - |P_0| \|_{\gamma_d}.$$

Remark 1.6. Theorem 1.5 has been autoformalized in Lean 4; we refer the reader to Sections 1.3 and 4 for a discussion of the Lean formalization and to [6, 7, 25, 26, 27] for recent examples of formalizations involving AI.

Remark 1.7. We emphasize that the assumption that $P_0 \in \text{span}\{\Phi_{\vec{\kappa}, \vec{\alpha}}\}_{\vec{\alpha}}$ is absolutely essential, as there is a dense set of $F_0 \in \mathcal{H}_{\vec{\kappa}}(\mathbb{C}^d)$ for which Theorem 1.5 is false. We also remark that although each basis element of $\text{span}\{\Phi_{\vec{\kappa}, \vec{\alpha}}\}_{\vec{\alpha}}$ factors as a tensor product, the proof of Theorem 1.5 is not obtained by a formal tensorization of the one-dimensional statements.

As a direct consequence of Theorem 1.3, we are able to prove a *sharp stability version* of the log-Sobolev inequality for entire functions. More specifically, if $F \in \mathcal{F}^2(\mathbb{C})$ with $\|F\|_{\gamma} = 1$, the logarithmic Sobolev inequality for entire functions states that

$$\int_{\mathbb{C}} |F'(z)|^2 d\gamma(z) \geq \int_{\mathbb{C}} |F(z)|^2 \ln |F(z)|^2 d\gamma(z),$$

where $d\gamma$ denotes the Gaussian measure. This estimate is directly tied to the classical logarithmic Sobolev inequality (see, e.g., [24, 30, 32] for the first works dealing with such an estimate), and its stability has been a recent subject of interest. Indeed, we first highlight [13], where the authors prove asymptotically sharp stability estimates for the Sobolev inequality, which, by a limiting argument, yields stability estimates for the classical log-Sobolev inequality. More recently, R. Frank, F. Nicola and P. Tilli [15] showed that the same estimate holds also in the aforementioned *entire* case, obtaining the following result.

Theorem 1.8 (Frank–Nicola–Tilli, [15, Theorem 14]). *There exists a constant $c_* > 0$ such that, for every $F \in \mathcal{F}^2(\mathbb{C})$ with $\|F\|_\gamma = 1$, we have*

$$\int_{\mathbb{C}} |F'(z)|^2 d\gamma(z) \geq \int_{\mathbb{C}} |F(z)|^2 \ln |F(z)|^2 d\gamma(z) + c_* \inf_{\substack{\beta \in \mathbb{C} \\ |c|=1}} \|F - ce^{\beta z - |\beta|^2/2}\|_\gamma^2.$$

As we shall see below, Theorem 1.8 follows almost directly from Theorem 1.3 and [13, Corollary 1.2].

1.3. Explanation of LLM usage and Lean. Here we explain the events that led to the results in this paper.

LLM Usage. The problem under consideration had previously been studied by several of the authors over the past years, with varying degrees of success. On 28 March 2026, the authors were experimenting with ChatGPT Plus (GPT 5.4) and, in a multi-turn conversation, became intrigued by a (false) reduction of the problem to a circle estimate (which one may view as a primitive version of the results in Section 2.4). The authors then tried to understand this reduction, further prompting GPT until they realized that a modification of the proposed argument could be salvaged and would likely lead to a proof. We then sketched the full proof of Theorem 2.1 by ourselves, which naively lost a log-factor in the stability estimate, and asked GPT 5.4 in Codex to refine it. This ultimately resulted in a complete proof of Theorem 2.1, which we knew implied Theorem 1.3. We then autoformalized the polynomial version of this result in Lean 4 using Claude Opus 4.6.

In hindsight, the fact that the structure of the problem is more easily visible in polar coordinates should have been evident. In fact, spiral sampling is commonly used in experimental applications. Moreover, the final proof is not particularly difficult, and once the structure was made clear, the authors could have completed the proof of Theorem 1.3 without any additional assistance. What is interesting, however, is that the LLMs were able to incrementally generalize Theorem 1.3 to Theorem 1.5 with very little human mathematical input. Of course, it should be emphasized that the authors strongly believed that the method would generalize. Moreover, as we will see below, significant challenges arose on the Lean side when verifying the full result.

As alluded to above, several of the intermediate results in this paper were known to the authors for a long time. For example, the “local-local to local” reduction in Proposition 2.3 is relatively standard. The orthogonal reduction in Theorem 2.4 is also just a natural variant of [16, Theorem 3.10] and was not produced by an LLM.

Lean. The main formal content of the results in this paper has been verified in Lean 4 [12] using the Mathlib library [31]. Here, we point out some interesting features of the formalization that may be useful for other problems. We refer the reader to Section 4 for a formalized statement of our main theorem.

In the beginning, we carried out the Lean verification only for finite linear combinations of the basis elements. Our justification for this was threefold. First, the density argument is completely standard – see Section 2.2. Second, the density argument is a general fact that is easy to formalize in a systematic way. Finally, we found that working directly with arbitrary polyanalytic functions rather than “finite” polynomials posed unnecessary challenges in the Lean formalization.

The formalization process was completed with extensive assistance from LLMs; namely, the first result was formalized with Claude Opus 4.6 in Claude Code and later generalizations with GPT 5.4 and 5.5 through Codex, in all cases using Oliver Dressler’s MCP [14] under the supervision of the authors. The authors checked the statements of the main theorems and provided support whenever the formal development encountered blockers, such as statement mismatches between the blueprint and the corresponding results in the Lean files, or overly strong statements that made the proof

substantially harder to fill in. We found, for example, that it was worthwhile to make all statements as quantitative as possible and to avoid compactness arguments – remnants of this insistence can be (intentionally) witnessed by the explicit choice of constants in the proof, which we did not try to optimize. For clarity, the reviewer facing statement of the main theorem in Section 4 was rewritten to assert the existence of a constant rather than to choose an explicit constant. However, some of the less general statements remain quantitative, namely the one for the Fock space case, which is an exact copy of Theorem 1.4.

The full Lean code is available on the GitHub repository, which can be accessed from here: <https://github.com/susannabertolini/PhaseRetrieval>. We remark that we put substantial effort into properly presenting the *statement* of the main theorem in Lean (see Section 4) as well as in decoupling the Lean statement of the theorem from the LLM generated proof (which in itself is of low quality). The main difficulties of the formalization arose from incorrect translations from the blueprint to the scaffolding: this led to incorrect or imprecise scaffolding files even with an extremely detailed blueprint. Another recurring issue was the tendency of the LLMs to incorporate unnecessary Hilbert-space structure whenever the blueprint mentioned it for context, even if the main theorem was clearly stated for finite linear combinations only. This introduced many density arguments that did not play a role in the proof of the main theorem but created blockers that had to be recognized as irrelevant and removed before being able to complete the formalization.

1.4. Acknowledgments. J.P.G.R. was supported by the FCT through project SHADE (project 2023.17881.ICDT, DOI 10.54499 / 2023.17881.ICDT) and by FAPERJ through the JCNE grant no. SEI-260003/020475/2025.

The authors thank Alessio Figalli, Javier Gomez-Serrano and Joaquim Serra for their support and guidance during the preparation of this article. J.D. and M.T. also thank Manuel Guizar-Sicairos and Abraham Levitan for hosting them at PSI and for the instructive discussions on the experimental aspects of phase retrieval. Finally, we thank the Banach lattice and phase retrieval communities for their encouragement to pursue this line of research.

The latter parts of this work were conducted while J.D. and M.T. were in residence at the Institute for Computational and Experimental Research in Mathematics in Providence, RI, during the *Techniques and Tools for the Formalization of Analysis* program. This program was supported by the National Science Foundation under Grant No. DMS-2424556. The authors are grateful to the organizers as well as to several participants for their helpful comments on the exposition of the formalized statement in Section 4.

2. PROOF OF THEOREM 1.3

Here, we present the proof of Theorem 1.3. We begin by giving an outline of the proof. After this, we present some general reductions. Finally, we tackle the main estimate specific to Fock space phase retrieval.

2.1. Outline of the proof. Below, we give an outline of the logical structure of the proof of Theorem 1.3 as well as a brief overview of the key ideas. We will elaborate on each point in more detail in the corresponding subsection.

The local reduction (Section 2.2). The main goal of this subsection will be to show that Theorem 1.3 is implied by Theorem 1.4. This will follow from a more general argument that holds in any reproducing kernel Hilbert space (which the Fock space is a particular example of). Essentially, we will show that if one can prove phase retrieval at a function F_0 and establish local stability with respect to very small perturbations of F_0 , then one can prove local stability for general perturbations of F_0 . In the context of Theorem 1.3, one takes $F_0 = 1$, but we prefer to state this

argument in a more abstract form so that it also applies (for instance) in the context of Theorem 1.5.

The orthogonal reduction (Section 2.3). Another general argument will be used to show that if orthogonal perturbations of a unit vector F_0 satisfy a certain coercivity estimate with constant M , then all sufficiently small perturbations (after optimizing over a global phase) satisfy the same coercivity estimate with a comparable constant $\widetilde{M} = O(M)$. We leave the precise statement and proof of this reduction to Section 2.3, but we emphasize that it is just the natural local variant of [16, Theorem 3.10]. In the special case of the Fock space $\mathcal{F}^2(\mathbb{C})$, this observation will enable us to reduce the proof of Theorem 1.4 to the following ‘‘orthogonal coercivity’’ estimate, whose proof is the technical heart of the paper.

Theorem 2.1 (Fock-space coercivity). *There exists an absolute constant C such that for every $D \geq 1$ and every polynomial $U(z) = \sum_{n=1}^D a_n z^n$ with $U(0) = 0$, we have*

$$\|U\|_{\mathcal{F}} \leq C \|\rho(U)\|_{\mathcal{F}}.$$

Here, $\rho(w) := \|1 + w\| - 1$ and one can take $C = 4600$.

Remark 2.2. The condition $U(0) = 0$ means that U is orthogonal to the constant one function in $\mathcal{F}^2(\mathbb{C})$. Thus, the orthogonal reduction tells us that Theorem 2.1 is strong enough to recover our main results. We remark that the constants in Theorem 1.4 are obtained by inserting $C = 4600$ into Theorem 2.4.

We now outline the steps needed to prove Theorem 2.1.

Polar coordinates. A simple but important idea in our proof below is to represent the polynomial U in Theorem 2.1 in polar coordinates. This will enable us to rephrase many of our estimates in terms of estimates for trigonometric polynomials (where we have more direct access to Fourier analytic tools). We further note that the Gaussian integral decomposes as

$$(4) \quad \frac{1}{\pi} \int_{\mathbb{C}} F(z) e^{-|z|^2} dm(z) = 2 \int_0^\infty r e^{-r^2} \left(\int_{S^1} F(re^{it}) d\sigma(t) \right) dr,$$

where $d\sigma = dt/(2\pi)$ is the probability Haar measure on the circle. Moreover, both norms in Theorem 2.1 decompose in this way, so it suffices to compare the inner circle integrals for each radius r . This simple observation will be very useful in the forthcoming analysis.

Circle estimates (Section 2.4). On a circle of radius r , the function $t \mapsto U(re^{it})$ is a trigonometric polynomial with strictly positive frequencies (since $U(0) = 0$). We prove two estimates relating the L^2 norm of such a polynomial P to that of $\rho \circ P$:

- (i) A *local circle estimate* that works for any finite set of positive frequencies, at the cost of a constant depending on how many frequencies are present.
- (ii) A *high-frequency estimate* that applies when the lowest frequency is much larger than the bandwidth of frequencies, giving a universal constant independent of the number of frequencies.

Frequency localization (Section 2.5). The polynomial U may have up to D frequencies, but the monomial z^n concentrates its Gaussian mass near the annulus $|z| \approx \sqrt{n}$. We exploit this by partitioning the frequencies into blocks, each localized to a specific annulus. For each annulus, only a bounded window of nearby blocks contributes significantly, and the energy leaking from distant blocks decays as a Gaussian in the block distance.

Finishing the proof (Section 2.6). Here, we assemble the estimates of the previous two subsections to complete the proof of the orthogonal coercivity bound in Theorem 2.1. The observation is that on each annulus, the nearby blocks form a trigonometric polynomial to which one of the

two circle estimates applies. This gives a uniform bound on the ρ -norm of the “local” polynomial. We then use the 1-Lipschitz property of ρ to pass from the local polynomial to the full U , integrate over all annuli, and absorb the exponentially small leakage.

2.2. The local reduction. In this subsection, we reduce Theorem 1.3 to Theorem 1.4. The reduction is abstract and only relies on the RKHS structure of the Fock space. The first step is to show that local stable phase retrieval is implied by an even more local variant. The second step is a simple density argument.

Proposition 2.3 (Local reduction). *Let \mathcal{H} be a reproducing kernel Hilbert space of real or complex-valued functions isometrically embedded in $L^2(\Omega, \mu)$ where (Ω, μ) is some measure space. Fix $F \in \mathcal{H}$ and $\epsilon > 0$. Suppose that there exists a constant $C > 0$ such that for all $G \in \mathcal{H}$ satisfying*

$$\inf_{|\lambda|=1} \|F - \lambda G\|_{L^2(\Omega)} < \epsilon$$

we have

$$\inf_{|\lambda|=1} \|F - \lambda G\|_{L^2(\Omega)} \leq C \| |F| - |G| \|_{L^2(\Omega)}.$$

Suppose also that phase retrieval holds at F , i.e., if $G \in \mathcal{H}$ and $|F| = |G|$ a.e., then there exists $\lambda \in \mathbb{C}$ with $|\lambda| = 1$ so that $F = \lambda G$. Then there exists a constant $C' > 0$ such that for all $G \in \mathcal{H}$,

$$\inf_{|\lambda|=1} \|F - \lambda G\|_{L^2(\Omega)} \leq C' \| |F| - |G| \|_{L^2(\Omega)}.$$

Proof. We need to verify that there exists a constant C' such that for all $G \in \mathcal{H}$,

$$\inf_{|\lambda|=1} \|F - \lambda G\|_{L^2(\Omega)} \leq C' \| |F| - |G| \|_{L^2(\Omega)}.$$

First suppose that $\|G\|_{L^2(\Omega)} \geq 4\|F\|_{L^2(\Omega)}$. Then

$$\| |F| - |G| \|_{L^2(\Omega)} \geq \|G\|_{L^2(\Omega)} - \|F\|_{L^2(\Omega)} \geq \frac{3}{4} \|G\|_{L^2(\Omega)},$$

while

$$\inf_{|\lambda|=1} \|F - \lambda G\|_{L^2(\Omega)} \leq \|F\|_{L^2(\Omega)} + \|G\|_{L^2(\Omega)} \leq \frac{5}{4} \|G\|_{L^2(\Omega)}.$$

Thus the desired inequality holds in this case with any $C' \geq 5/3$. Therefore, we may restrict attention to G with $\|G\|_{L^2(\Omega)} < 4\|F\|_{L^2(\Omega)}$.

Now suppose that the desired inequality is true for all G satisfying

$$\inf_{|\lambda|=1} \|F - \lambda G\|_{L^2(\Omega)} < \epsilon,$$

but is not true in general. Then for every sufficiently large n we can find $G_n \in \mathcal{H}$ with

$$\inf_{|\lambda|=1} \|F - \lambda G_n\|_{L^2(\Omega)} \geq \epsilon,$$

$$\inf_{|\lambda|=1} \|F - \lambda G_n\|_{L^2(\Omega)} > n \| |F| - |G_n| \|_{L^2(\Omega)},$$

and also $\|G_n\|_{L^2(\Omega)} < 4\|F\|_{L^2(\Omega)}$. This latter inequality guarantees that $\inf_{|\lambda|=1} \|F - \lambda G_n\|_{L^2(\Omega)}$ is uniformly bounded, and hence

$$\| |F| - |G_n| \|_{L^2(\Omega)} \rightarrow 0.$$

In particular, after passing to a subsequence, we may assume that

$$|G_{n_k}(x)| \rightarrow |F(x)| \quad \text{for almost every } x \in \Omega.$$

Since $\|G_{n_k}\|_{L^2(\Omega)}$ is bounded independently of k and the embedding of \mathcal{H} into $L^2(\Omega, \mu)$ is isometric, the sequence (G_{n_k}) is bounded in \mathcal{H} . Passing to a further subsequence if necessary, we may assume that $G_{n_{k_\ell}} \rightharpoonup G$ weakly in \mathcal{H} . Since point evaluations are bounded on \mathcal{H} , we have

$$G_{n_{k_\ell}}(x) \rightarrow G(x)$$

for every $x \in \Omega$. Hence $|G_{n_{k_\ell}}(x)| \rightarrow |G(x)|$ for every $x \in \Omega$. This forces $|G| = |F|$ almost everywhere. By the phase-retrieval hypothesis, $G = \lambda F$ for some unimodular scalar λ .

Collecting what we know, $G_{n_{k_\ell}}$ converges weakly in \mathcal{H} to λF . Moreover, from $\||F| - |G_n|\|_{L^2(\Omega)} \rightarrow 0$, we see that

$$\|G_n\|_{L^2(\Omega)} = \||G_n|\|_{L^2(\Omega)} \rightarrow \||F|\|_{L^2(\Omega)} = \|\lambda F\|_{L^2(\Omega)}.$$

Because the embedding is isometric, weak convergence in \mathcal{H} together with convergence of the L^2 -norms to the norm of the weak limit implies strong convergence in L^2 . Thus $G_{n_{k_\ell}} \rightarrow \lambda F$ in $L^2(\Omega)$. For all sufficiently large ℓ , this contradicts (after relabeling the unimodular scalar)

$$\inf_{|\lambda|=1} \|F - \lambda G_{n_{k_\ell}}\|_{L^2(\Omega)} \geq \epsilon.$$

□

2.2.1. Proof of Theorem 1.3. Using the above reduction, we immediately deduce Theorem 1.3 from Theorem 1.4 by density and by applying Proposition 2.3 with $F = 1$ and $\mathcal{H} = \mathcal{F}^2(\mathbb{C})$. Indeed, Theorem 1.4 immediately extends from polynomial perturbations P to arbitrary elements $\tilde{P} \in \mathcal{F}^2(\mathbb{C})$ with $\|\tilde{P}\|_{\mathcal{F}} < \delta$ by density. In that density passage, we choose polynomials $P_n \rightarrow \tilde{P}$, apply Theorem 1.4, pass to a subsequence of phases $w_n \rightarrow w$, and use continuity of $t \mapsto |1+t|$ in L^2 . Phase retrieval at the constant function in the Fock space holds because $|G| = 1$ a.e. forces the entire function G to be a unimodular constant. Hence, it is easy to see that Proposition 2.3 is applicable.

2.3. The orthogonal reduction. In this subsection, we work in the generality of a complex Hilbert space \mathcal{H} isometrically embedded in $L^2(\mu)$ for some measure space (Ω, μ) . We show that *orthogonal coercivity* – a stability estimate for perturbations perpendicular to a fixed unit vector – implies *local stability* for all sufficiently small perturbations, after optimizing over a global phase. The argument is based on function space theory and uses no structure specific to the Fock space.

Theorem 2.4 (Orthogonal reduction). *Let \mathcal{H} be a complex Hilbert space with norm $\|\cdot\|$, isometrically embedded in $L^2(\Omega, \mu)$. Let $f_0 \in \mathcal{H}$ with $\|f_0\| = 1$ and suppose that there exists $M > 0$ such that*

$$(5) \quad \|g\| \leq M \||f_0 + g| - |f_0|\|_{L^2} \quad \text{for all } g \perp f_0.$$

Then there exist $\delta > 0$ and $\tilde{M} > 0$, depending only on M , such that for every $h \in \mathcal{H}$ with $\|h\| \leq \delta$ and $\langle h, f_0 \rangle \in \mathbb{R}$, we have

$$\|h\| \leq \tilde{M} \||f_0 + h| - |f_0|\|_{L^2}.$$

One may take $\delta = 1/(M+1)$ and $\tilde{M} = 5M+3$.

Remark 2.5. The condition $\langle h, f_0 \rangle \in \mathbb{R}$ is a phase normalization: it says that h has been rotated so that its component along f_0 is real. In the application to Theorem 1.4, this is achieved by choosing the optimal phase w so that $\langle w(f_0 + p) - f_0, f_0 \rangle$ is real, and aligning the phase of the inner product $\langle f_0 + p, f_0 \rangle$.

Proof. Write $m = \||f_0 + h| - |f_0|\|_{L^2}$ for the right-hand side of our inequality. We break the proof into four simple steps.

Step 1: Orthogonal decomposition. Set $a = \operatorname{Re}\langle h, f_0 \rangle = \langle h, f_0 \rangle$ and $g = h - af_0$. Then $g \perp f_0$, $h = af_0 + g$ and $\|h\|^2 = a^2 + \|g\|^2$, so

$$(6) \quad \|h\| \leq |a| + \|g\|.$$

Step 2: Control of a . The pointwise identity $|f_0 + h|^2 - |f_0|^2 = 2 \operatorname{Re}(h \overline{f_0}) + |h|^2$ integrates to

$$\int_{\Omega} (|f_0 + h|^2 - |f_0|^2) \, d\mu = 2a + \|h\|^2.$$

On the other hand, the factorization $|f_0 + h|^2 - |f_0|^2 = (|f_0 + h| - |f_0|)(|f_0 + h| + |f_0|)$ and Cauchy–Schwarz give

$$|2a + \|h\|^2| \leq m \cdot \| |f_0 + h| + |f_0| \|_{L^2} \leq m(2 + \|h\|).$$

Hence,

$$(7) \quad |a| \leq \frac{2 + \|h\|}{2} m + \frac{\|h\|^2}{2}.$$

Step 3: Control of g . The reverse triangle inequality $||f_0 + g| - |f_0 + h|| \leq |g - h| = |a| |f_0|$ gives $\||f_0 + g| - |f_0|\|_{L^2} \leq |a| + m$. The hypothesis (5) then yields

$$(8) \quad \|g\| \leq M(|a| + m).$$

Step 4: Completing the proof. Substituting (7) and (8) into (6), we see that

$$\|h\| \leq |a| + M(|a| + m) = (M + 1)|a| + Mm \leq \underbrace{\left(M + \frac{(M + 1)(2 + \|h\|)}{2} \right)}_{=: C(\|h\|)} m + \frac{M + 1}{2} \|h\|^2.$$

Since $\|h\| \leq \delta$, we have $\|h\|^2 \leq \delta \|h\|$, so

$$\|h\| \leq C(\delta) m + \frac{(M + 1)\delta}{2} \|h\|.$$

Choosing $\delta = 1/(M + 1)$ makes $(M + 1)\delta/2 = 1/2$, giving $\|h\|/2 \leq C(\delta) m$ and therefore

$$\|h\| \leq 2C(\delta) m.$$

A short computation shows that $C(\delta) = M + (M + 1)(2 + \delta)/2 < (5M + 3)/2$, so one may take $\widetilde{M} = 5M + 3$. \square

Corollary 2.6 (Fock-space application). *Taking $f_0 = 1$ (the constant function), $\mathcal{H} = \mathcal{F}^2(\mathbb{C})$, and $M = 4600$ (from Theorem 2.1), the orthogonal reduction gives $\delta = 1/4601$ and $\widetilde{M} = 23003$. This is Theorem 1.4.*

Proof. The condition $g \perp 1$ in $\mathcal{F}^2(\mathbb{C})$ means precisely that $g(0) = 0$, i.e., g has no constant term. Theorem 2.1 provides the orthogonal coercivity hypothesis (5) with $M = 4600$. The optimal phase w is chosen so that $h = w(1 + P) - 1$ satisfies $\langle h, 1 \rangle \in \mathbb{R}$ and $\|h\|_{\mathcal{F}} \leq \delta$; the existence of such a w is straightforward. Theorem 2.4 then yields the conclusion of Theorem 1.4. \square

Our goal for the remaining subsections will be to establish the key bound in Theorem 2.1.

2.4. Circle estimates. In this section we prove two estimates controlling the L^2 norm of a trigonometric polynomial P on the circle in terms of the L^2 norm of $\rho \circ P$. Both exploit the fact that all frequencies of P are strictly positive (which is a consequence of our reduction to the case $U(0) = 0$) but they apply in complementary regimes. The *local circle estimate* (Proposition 2.8) works for any frequency set and gives a constant proportional to \sqrt{L} , where L is the number of frequencies. The *high-frequency estimate* (Proposition 2.10) applies when the lowest frequency is much larger than the bandwidth of frequencies, and gives an absolute constant. Together they will provide a uniform bound on every annulus.

Throughout this section, P denotes a trigonometric polynomial $P(t) = \sum_{n \in E} b_n e^{int}$ for a finite set $E \subset \mathbb{Z}_{>0}$ and all norms are taken in $L^2(S^1, d\sigma)$. Recall as well that the function $\rho(w) = ||1+w| - 1|$ is 1-Lipschitz, and in particular we have $\rho(w) \leq |w|$ and

$$(9) \quad \rho(w) \geq \rho(z) - |w - z|.$$

2.4.1. *The local circle estimate.* The following elementary inequality will be used several times in the forthcoming calculations, so we record it here for convenience.

Lemma 2.7. *Let $a, b, c \geq 0$ with $a \geq b - c$. Then $a^2 \geq \frac{1}{2} b^2 - c^2$.*

Proof. If $b \leq c$ the right-hand side is nonpositive. If $b > c$ then $a \geq b - c > 0$ gives $a^2 \geq (b - c)^2 \geq \frac{1}{2} b^2 - c^2$, where the last step is the identity $(b - c)^2 - \frac{1}{2} b^2 + c^2 = \frac{1}{2} (b - 2c)^2 \geq 0$. \square

Proposition 2.8 (Local circle estimate). *Let $E \subset \mathbb{Z}_{>0}$ with $|E| = L \geq 1$, let $b_n \in \mathbb{C}$ for $n \in E$, and set $P(t) = \sum_{n \in E} b_n e^{int}$. Then*

$$\|P\|_{L^2}^2 \leq 144 L \cdot \|\rho \circ P\|_{L^2}^2.$$

Proof. Write $\alpha = \|P\|_{L^2}$ and $\beta = \|\rho \circ P\|_{L^2}$. We show that $\beta \geq \alpha / (12\sqrt{L})$, which gives the result after squaring. To prove this estimate, we will consider two different regimes; namely, the one where α is sufficiently small and its complement. The threshold between the two regimes will be taken to be $\alpha = \alpha_0 = 1 / (4\sqrt{L})$.

Two structural properties of P underpin both cases. First, by Cauchy–Schwarz and Parseval we have the pointwise bound $|P(t)| \leq \sqrt{L} \alpha$ for all t . Second, since every $n \in E$ satisfies $n \geq 1$, the zeroth Fourier coefficient of P vanishes, i.e., $\int_{S^1} P d\sigma = 0$. More importantly, we have $n + m \geq 2$ for all $n, m \in E$, so $\int_{S^1} P^2 d\sigma = 0$ as well. Writing $P = u + iv$ with u, v real-valued, this identity splits as

$$(10) \quad \int_{S^1} u^2 d\sigma = \int_{S^1} v^2 d\sigma = \frac{\alpha^2}{2} \quad \text{and} \quad \int_{S^1} uv d\sigma = 0.$$

The equipartition of mass between the real and imaginary parts of P given by (10) will be crucially used in the estimates below.

Small amplitudes ($\alpha \leq \alpha_0$). Here, we have $\|P\|_\infty \leq \sqrt{L} \alpha \leq 1/4$. For $|w| \leq 1/4$ we also have $|1 + w| \geq 3/4$ and $\text{Re}(w) \geq -1/4$, and therefore $|1 + w| + 1 + \text{Re}(w) \geq 3/2 > 0$. The algebraic identity

$$|1 + w| - 1 - \text{Re}(w) = \frac{\text{Im}(w)^2}{|1 + w| + 1 + \text{Re}(w)}$$

shows that $|1 + w| - 1 = \text{Re}(w) + \varepsilon$ with $0 \leq \varepsilon \leq \text{Im}(w)^2 / (3/2) \leq |w|^2$. In particular, $\rho(w) = |\text{Re}(w) + \varepsilon| \geq |\text{Re}(w)| - \varepsilon$, so by Lemma 2.7 we have

$$\rho(P)^2 \geq \frac{1}{2} u^2 - |P|^4.$$

In other words, when P is small in amplitude, most of the mass of $\rho(P)$ comes from the real part of P . One might a priori worry about the hypothetical possibility that if $P(t)$ were allowed to

be purely imaginary for most t , the nonlinear function $\rho(P(t))$ might be small relative to $P(t)$. However, integrating the above and using property (10), we see that we indeed have the coercive bound $\beta^2 \geq \alpha^2/4 - \|P\|_\infty^2 \alpha^2 \geq \alpha^2/4 - L\alpha^4$. Since $L\alpha^2 \leq 1/16$, we get $\beta^2 \geq 3\alpha^2/16$, so $\beta \geq \sqrt{3}\alpha/4 \geq \alpha/(12\sqrt{L})$.

Large amplitudes ($\alpha > \alpha_0$). The fact that P has mean zero (since it is supported on positive frequencies) also gives $\int_{S^1} |1 + P|^2 d\sigma = 1 + \alpha^2$, so

$$\alpha^2 = \int_{S^1} (|1 + P|^2 - 1) d\sigma \leq \int_{S^1} \rho(P) (|1 + P| + 1) d\sigma,$$

where we used the identity $||1 + P|^2 - 1| = (|1 + P| + 1)\rho(P)$. By Cauchy–Schwarz,

$$\alpha^2 \leq \beta \cdot (\sqrt{1 + \alpha^2} + 1).$$

If $\alpha \leq 1$, then $\beta \geq \alpha^2/3$ and $\alpha > \alpha_0 = 1/(4\sqrt{L})$ gives $\beta \geq \alpha/(12\sqrt{L})$. If $\alpha > 1$, then $\beta \geq \alpha/3 \geq \alpha/(12\sqrt{L})$. In both regimes, we have $\beta \geq \alpha/(12\sqrt{L})$. \square

2.4.2. Averaging over circles. The local circle estimate grows with the number of frequencies L . To eliminate this dependency for high-frequency polynomials, we will need to use the fact that the ρ -image of the circle sees a quantitative fraction of its radius. More precisely, we have the following proposition.

Proposition 2.9. *For all $r \geq 0$, we have*

$$\int_{S^1} \rho(re^{i\theta})^2 d\sigma(\theta) \geq \frac{r^2}{8}.$$

Proof. We claim that on the arc $|\theta| \leq \pi/4$ we have $\rho(re^{i\theta}) \geq r/\sqrt{2}$. Indeed, $\cos \theta \geq 1/\sqrt{2}$ gives

$$|1 + re^{i\theta}| - 1 = \frac{r^2 + 2r \cos \theta}{|1 + re^{i\theta}| + 1} \geq \frac{r(r + \sqrt{2})}{r + 2} \geq \frac{r}{\sqrt{2}}.$$

Since the numerator is non-negative, we have $|1 + re^{i\theta}| \geq 1$ and hence ρ equals $|1 + re^{i\theta}| - 1$ without the absolute value. Restricting the integral to this arc, we have $\int \rho^2 d\sigma \geq \frac{1}{4}(r/\sqrt{2})^2 = r^2/8$. \square

2.4.3. The high-frequency estimate. When the polynomial oscillates rapidly, the Lipschitz property of ρ and Proposition 2.9 combine to give a much stronger bound. The idea is to factor $P(t) = e^{iNt} Q(t)$ where Q is a slowly varying envelope. On short intervals, the phase e^{iNt} sweeps through a full rotation while Q barely changes. Proposition 2.9 then ensures that ρ captures a positive fraction of $|Q|^2$. We make this intuition precise in the estimate below.

Proposition 2.10 (High-frequency estimate). *Let $N \geq 1$ and $L \geq 1$ satisfy $1343 L^2 \leq N^2$. For any $b_0, \dots, b_{L-1} \in \mathbb{C}$, the polynomial $P(t) = \sum_{m=0}^{L-1} b_m e^{i(N+m)t}$ satisfies*

$$\|P\|_{L^2}^2 \leq 32 \|\rho \circ P\|_{L^2}^2.$$

Proof. Write $P(t) = e^{iNt} Q(t)$ with $Q(t) = \sum_{m=0}^{L-1} b_m e^{imt}$, so $\|P\| = \|Q\|$. Partition the circle into N equal arcs $J_k = [2\pi k/N, 2\pi(k+1)/N)$, and let $q_k = N \int_{J_k} Q d\sigma$ be the average of Q on J_k .

On each interval J_k of length $h = 2\pi/N$, the Poincaré inequality gives $\int_{J_k} |Q - q_k|^2 d\sigma \leq h^2 \int_{J_k} |Q'|^2 d\sigma$.² Summing over k and using Bernstein’s inequality $\|Q'\|^2 \leq L^2 \|Q\|^2$ (since every

²We use the inequality $\int_0^h |f - \bar{f}|^2 dx \leq h^2 \int_0^h |f'|^2 dx$ with the non-sharp constant 1; here \bar{f} denotes the mean of f on $[0, h]$. The sharp constant is $1/\pi^2$. After the conversion of the Lean code to natural language, we saw that the LLMs emphasized that this better estimate would have improved the threshold from 1343 to 136 but they did not use it as we insisted that the proof be as elementary as possible.

frequency of Q has modulus at most $L-1 < L$) we see that

$$(11) \quad \delta := \sum_k \int_{J_k} |Q - q_k|^2 d\sigma \leq \frac{4\pi^2 L^2}{N^2} \|Q\|^2.$$

Note that

$$(12) \quad \|Q\|^2 = \frac{1}{N} \sum_k |q_k|^2 + \delta.$$

As t traverses J_k , the argument Nt traverses one full period. Substituting $s = Nt$ and applying Proposition 2.9 with $r = |q_k|$ we obtain

$$\int_{J_k} \rho(e^{iNt} q_k)^2 d\sigma = \frac{1}{N} \int_{S^1} \rho(|q_k| e^{is})^2 d\sigma(s) \geq \frac{|q_k|^2}{8N}.$$

On J_k , the Lipschitz bound (9) gives $\rho(P) = \rho(e^{iNt} Q) \geq \rho(e^{iNt} q_k) - |Q - q_k|$. By Lemma 2.7, $\rho(P)^2 \geq \frac{1}{2} \rho(e^{iNt} q_k)^2 - |Q - q_k|^2$. Integrating over J_k and using the above we obtain

$$\int_{J_k} \rho(P)^2 d\sigma \geq \frac{|q_k|^2}{16N} - \int_{J_k} |Q - q_k|^2 d\sigma.$$

Summing over k and applying (12), it follows that

$$\|\rho \circ P\|^2 \geq \frac{1}{16} (\|Q\|^2 - \delta) - \delta = \frac{\|Q\|^2}{16} - \frac{17\delta}{16}.$$

Substituting the bound (11), we arrive at

$$\|\rho \circ P\|^2 \geq \frac{1}{16} \left(1 - \frac{68\pi^2 L^2}{N^2}\right) \|Q\|^2.$$

The condition $1343 L^2 \leq N^2$ gives $68\pi^2 L^2/N^2 \leq 1/2$ (since $136\pi^2 < 1343$), and therefore $\|\rho \circ P\|^2 \geq \|P\|^2/32$. \square

2.5. Frequency localization. The circle estimates from Section 2.4 require the trigonometric polynomial to have a controlled number of frequencies, yet U may have arbitrarily many. The Gaussian weight is what resolves this issue, as the radial factor $r \mapsto r^{2n+1} e^{-r^2}$ concentrates sharply near $r = \sqrt{n+1/2}$. This will allow us to group frequencies of U into “blocks” approximately supported on disjoint annuli.

We begin with some general notation. For $\ell \geq 1$, we define

$$I_\ell = \{n \in \mathbb{N} : \ell^2 \leq n < (\ell+1)^2\},$$

which has $|I_\ell| = 2\ell + 1$ elements. We then define $U_\ell(z) = \sum_{n \in I_\ell, n \leq D} a_n z^n$. The blocks I_1, I_2, \dots are pairwise disjoint and cover $\{1, \dots, D\}$ up to index $\Lambda = \lfloor \sqrt{D} \rfloor$. Clearly, we have

$$(13) \quad \|U\|_{\mathcal{F}}^2 = \sum_{\ell=1}^{\Lambda} \|U_\ell\|_{\mathcal{F}}^2.$$

Similarly, on each circle of radius r , blocks with distinct frequency supports are orthogonal, so we have $\int_{S^1} U_{\ell_1}(re^{it}) \overline{U_{\ell_2}(re^{it})} d\sigma(t) = 0$ for $\ell_1 \neq \ell_2$. Given $n \in I_\ell$ with $\ell \geq 1$, we define the radius $r_n = \sqrt{n+1/2} \in (\ell, \ell+1)$.

Fix $M \geq 1$. For each index $j \geq 0$ we define

$$V_j = \sum_{\substack{1 \leq \ell \leq \Lambda \\ |\ell-j| \leq M}} U_\ell \quad \text{and} \quad R_j = U - V_j.$$

Note that V_j gathers together the blocks whose peaks fall within M steps of annulus j . For $j \geq M+1$, the frequency support of V_j lies in $[(j-M)^2, (j+M+1)^2)$, and the number of frequencies is $L_j = (2M+1)(2j+1)$.

The mass that a block U_ℓ contributes to a distant annulus decays like $e^{-(|j-\ell|-1)^2}$. Indeed, the function $\varphi_n(r) = (2n+1) \log r - r^2$ (defined so that $r^{2n+1}e^{-r^2} = e^{\varphi_n(r)}$) has a unique maximum at r_n and satisfies $\varphi_n''(r) \leq -2$, which gives the bound

$$(14) \quad e^{\varphi_n(r)} \leq e^{\varphi_n(r_n)} e^{-(r-r_n)^2}.$$

For $n \geq 1$, it is easy to see by using Stirling's approximation that $e^{\varphi_n(r_n)} \leq e^{1/4} n!/2$. Combining this with (14), we see that for $r \in [j, j+1]$ and $n \in I_\ell$, we have

$$(15) \quad r^{2n+1}e^{-r^2} \leq \frac{e^{1/4} n!}{2} e^{-\text{dist}(r_n, [j, j+1])^2},$$

where $\text{dist}(r_n, [j, j+1]) = \max(j - r_n, r_n - (j+1), 0)$. By definition, $r_n \in (\ell, \ell+1)$, so this distance is at least $(|j - \ell| - 1)_+$. Our next objective is to quantify the "leakage" between the blocks.

Proposition 2.11. *For $\ell \geq 1$ and $j \geq 0$, we have*

$$\frac{1}{\pi} \int_{\{j \leq |z| < j+1\}} |U_\ell(z)|^2 e^{-|z|^2} dm(z) \leq e^{1/4} e^{-(|j-\ell|-1)_+^2} \|U_\ell\|_{\mathcal{F}}^2.$$

Proof. Passing to polar coordinates and using the orthogonality of distinct monomials on each circle, the left-hand side equals $2 \sum_{n \in I_\ell} |a_n|^2 \int_j^{j+1} r^{2n+1} e^{-r^2} dr$. Using (15) to bound each radial integral and summing gives the result. \square

By summing over all blocks, we arrive at the following corollary.

Corollary 2.12. *For $M \geq 2$, define $\eta_M = 2e^{1/4} \sum_{m \geq M} e^{-m^2}$. Then*

$$\sum_{j \geq 0} \frac{1}{\pi} \int_{\{j \leq |z| < j+1\}} |R_j(z)|^2 e^{-|z|^2} dm(z) \leq \eta_M \cdot \|U\|_{\mathcal{F}}^2.$$

Proof. By the orthogonality on each annulus, $\int |R_j|^2 = \sum_{|\ell-j| > M} \int |U_\ell|^2$. Now apply Proposition 2.11, swap the order of summation, and note that for $|j - \ell| > M \geq 2$ the decay factor is $e^{-(|j-\ell|-1)^2}$. The inner sum over j contributes at most $2 \sum_{m \geq M} e^{-m^2}$ and multiplying by $e^{1/4}$ and using (13) gives the result. \square

We can now state the uniform estimate that holds on every annulus. The function $\zeta \mapsto V_j(r\zeta)$ is a trigonometric polynomial on S^1 with all frequencies in $\mathbb{Z}_{>0}$, and we can apply either Proposition 2.8 or Proposition 2.10 depending on j .

Proposition 2.13. *With $M = 5$, we have for every $j \geq 0$ and every $r \in [j, j+1]$,*

$$\int_{S^1} |V_j(r\zeta)|^2 d\sigma(\zeta) \leq 1620^2 \cdot \int_{S^1} \rho(V_j(r\zeta))^2 d\sigma(\zeta).$$

Proof. When $r = 0$ both sides vanish (all monomials have positive degree). For $r > 0$, the function $\zeta \mapsto V_j(r\zeta)$ is a trigonometric polynomial with positive frequencies and coefficients $a_n r^n$. We keep the peculiar case distinction made during the autoformalization.

Case $j \leq 817$. With $M = 5$, the number of frequencies satisfies $L_j \leq 11(2 \cdot 817 + 1) < 18,000$. By Proposition 2.8, $\|V_j(r \cdot)\|^2 \leq 144 L_j \cdot \|\rho \circ V_j(r \cdot)\|^2$, and $144 \cdot 18,000 < 1620^2$.

Case $j \geq 818$. The lowest frequency is $N_j = (j-5)^2$ and we have $L_j = 11(2j+1)$. The ratio L_j^2/N_j^2 is decreasing in j (for $j \geq 6$), so it suffices to verify the condition $1343 L_j^2 \leq N_j^2$ of Proposition 2.10

at $j = 818$. A direct computation confirms that $1343 \cdot (11 \cdot 1637)^2 < 813^4$. The high-frequency estimate then gives $\|V_j(r \cdot)\|^2 \leq 32 \|\rho \circ V_j(r \cdot)\|^2 \leq 1620^2 \|\rho \circ V_j(r \cdot)\|^2$. \square

2.6. Proof of the orthogonal coercivity estimate. With the estimates from Sections 2.4 and 2.5 in hand, the proof of Theorem 2.1 is short.

Proof of Theorem 2.1. Fix $M = 5$ throughout, and write $B = 1620$ for the constant from Proposition 2.13. Let V_j and R_j be as above.

For $r \in [j, j+1]$, Proposition 2.13 gives

$$(16) \quad \int_{S^1} |V_j(r\zeta)|^2 d\sigma \leq B^2 \int_{S^1} \rho(V_j(r\zeta))^2 d\sigma.$$

Since ρ is 1-Lipschitz and $U = V_j + R_j$, we have

$$\rho(V_j(z)) \leq \rho(U(z)) + |R_j(z)|.$$

Using the inequality $(a + b)^2 \leq 2a^2 + 2b^2$ and substituting into (16), we see that

$$(17) \quad \int_{S^1} |V_j(r\zeta)|^2 d\sigma \leq 2B^2 \int_{S^1} \rho(U(r\zeta))^2 d\sigma + 2B^2 \int_{S^1} |R_j(r\zeta)|^2 d\sigma.$$

Applying $|U|^2 \leq 2|V_j|^2 + 2|R_j|^2$ and (17), we arrive at

$$(18) \quad \int_{S^1} |U(r\zeta)|^2 d\sigma \leq 4B^2 \int_{S^1} \rho(U)^2 d\sigma + (4B^2 + 2) \int_{S^1} |R_j|^2 d\sigma.$$

We now multiply (18) by $2re^{-r^2}$, integrate r over $[j, j+1]$, and use the polar decomposition (4). The $\rho(U)$ term on the right-hand side does not depend on j , while the R_j term does. Summing over $j \geq 0$, we conclude that

$$\|U\|_{\mathcal{F}}^2 \leq 4B^2 \|\rho(U)\|_{\mathcal{F}}^2 + (4B^2 + 2) \sum_{j \geq 0} \frac{1}{\pi} \int_{\{j \leq |z| < j+1\}} |R_j|^2 e^{-|z|^2} dm.$$

By Corollary 2.12 with $M = 5$, the above sum is at most $\eta_5 \|U\|_{\mathcal{F}}^2$. Hence

$$(19) \quad \|U\|_{\mathcal{F}}^2 \leq 4B^2 \|\rho(U)\|_{\mathcal{F}}^2 + (4B^2 + 2) \eta_5 \|U\|_{\mathcal{F}}^2.$$

We now absorb the additional term in the above inequality. Since $4B^2 + 2 < 10^7$ and $\eta_5 < 4 \times 10^{-11}$, the coefficient of $\|U\|_{\mathcal{F}}^2$ on the right-hand side of (19) satisfies $(4B^2 + 2) \eta_5 < 4 \times 10^{-4} < \frac{1}{2}$. Moving this term to the left-hand side and using $1 - (4B^2 + 2) \eta_5 > 1/2$, we conclude that

$$\|U\|_{\mathcal{F}}^2 \leq 8B^2 \|\rho(U)\|_{\mathcal{F}}^2 = 8 \cdot 1620^2 \|\rho(U)\|_{\mathcal{F}}^2 \leq 4600^2 \|\rho(U)\|_{\mathcal{F}}^2. \quad \square$$

3. COMMENTS, GENERALIZATIONS AND EXTENSIONS

3.1. Generalizations of Theorem 1.3. Here, we briefly discuss some of the ideas that allow one to generalize Theorem 1.3. To motivate the structure of this subsection, it is instructive to explain the story of how we came to Theorem 1.3 and later Theorem 1.5.

The collaboration on this project began when the fourth author privately shared an unpublished proof of a weaker version of Theorem 1.3; namely, with F replaced by F^2 in every instance. The proof of this weaker result is instructive and will be published elsewhere. However, despite extensive efforts, we were unable to extend the method to prove Theorem 1.3. This necessitated the new method that we presented above.

After establishing Theorem 1.3, we asked the LLMs to generalize the method further, which eventually resulted in Theorem 1.5. Importantly, however, we did not immediately move from Theorem 1.3 to Theorem 1.5, but instead proceeded incrementally.

The first step of the generalization was to prove the analogue of Theorem 1.3 for the first Hermite window at the first canonical basis vector. After this, we extended the result to all Hermite windows at the first canonical basis vector, and later to all Hermite windows at all elements in the finite span of the canonical basis vectors. Finally, we repeated all of the above steps in higher dimensions. Along the way, conclusions from the previous step were Lean verified and then utilized as inspiration for the LLMs to proceed to the next generalization.

The objective of this subsection will be to present some of the key ideas that arose while completing the above steps. These ideas mainly occurred during the first few generalizations, specifically when passing to the first Hermite window at the first canonical basis vector and to the Gaussian window in higher dimensions. The final proof of Theorem 1.5 is a technical modification of variants of these ideas, and we have not been able to find deeper meaning in it. For this reason, we have not presented the full proof of Theorem 1.5 in detail, as it does not seem instructive.

With the above being said, it may simply be that we have not yet found the *right* proof of Theorem 1.5, or perhaps even the right statement. Indeed, given that we now have a verification of correctness of Theorem 1.5, it is very natural to ask the following.

- (1) Is there an analogue for the wavelet transform (where now the geometry is the upper half space) or on bounded domains (cf. [2, 4])?
- (2) What is the true role of the Hermite windows; can the result be generalized to other windows which do STFT phase retrieval?
- (3) What is the true role of the Gaussian measure; can the result be generalized to other (radial) log-concave measures?
- (4) There are dense sets where stability holds and dense sets where it does not. Can we quantify which set is bigger? In particular, how does the stability constant in Theorem 1.5 depend on P_0 , at least in the case of the Fock space? A related question is to quantify the “size” of the set of pairs $(f, g) \in L^2(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$ so that $V_g f$ does local stable phase retrieval in $L^2(\mathbb{R}^d)$.
- (5) Are there other applications of Theorem 1.5? Are there other perspectives on the Fock space (e.g., related to ancient solutions to heat flows, coherent states, etc.) that could provide alternative methods to prove or extend our results?

3.1.1. *A special case of Theorem 1.5.* We now discuss some of the ideas used to generalize Theorem 1.3. Importantly, we note that independent iterations of the same prompts resulted in different strategies to generalize the proof, which may be a fact of independent interest.

For simplicity, we discuss the special case of Theorem 1.5 where we take $d = k = 1$ and $F_0 = \Phi_{1,0} := \Phi_0 = \bar{z}$. The general case is more technical, but follows a similar line of reasoning to what is explained below. In this special case, the result is the analogue of Theorem 1.3 for the first true polyanalytic Fock space $\mathcal{H}_1(\mathbb{C})$, where Φ_0 plays the role of the constant function. The proof proceeds by using the orthogonal reduction to reduce matters to proving the orthogonal coercivity bound

$$(20) \quad \|U\|_\gamma \lesssim \| |\Phi_0 + U| - |\Phi_0| \|_\gamma \quad (U \in \mathcal{H}_1 \cap \Phi_0^\perp).$$

This is the analogue of Theorem 2.1. Once we have this estimate, the overall architecture of the proof is the same as in the Fock space case, and with a bit of work, one can essentially reduce all of the estimates to small variations of the ones used in this case. However, there are two minor but important points when carrying out these reductions.

- (1) The circle estimates in Section 2.4 only apply to trigonometric polynomials supported on positive frequencies, and use the defect function $\rho(w) = ||1+w|-1|$ and not $\| |\Phi_0 + w| - |\Phi_0| \|$. Moreover, U is now allowed to have a non-trivial zeroth order Fourier mode because Φ_0 is supported at frequency -1 . The fix to this is simple: after writing everything in polar

coordinates, one simply factors (for $r > 0$) out the basis vector Φ_0 on both sides of the estimate (20). This shifts U to positive frequencies and normalizes the defect function to the one used in the Fock case, where the circle estimates in Section 2.4 apply.

- (2) The second minor difference is in the annulus localization carried out in Section 2.5. Here, instead of the monomials z^n (in the Fock case), we have to understand the mass concentration properties of the basis vectors $\Phi_n := \Phi_{1,n}$ to obtain the analogue of the bound (15).

Once these two points are addressed, the proof proceeds in an almost identical fashion to the Fock space case. The proof of the general result in Theorem 1.5 in one dimension is much more technical, but can ultimately be deduced by following a parallel line of reasoning.

3.1.2. *Higher dimensional analogues.* We now discuss higher dimensional analogues in \mathbb{C}^d . We mention here the simple analogue of the Fock space case $\mathcal{F}_d = \mathcal{F}^2(\mathbb{C}^d)$.

Theorem 3.1 (Local SPR at $\mathbf{1}$ in \mathcal{F}_d). *There exists a constant $M = M_d > 0$ such that for any $F \in \mathcal{F}_d$ we have*

$$\inf_{|\lambda|=1} \|F - \lambda\|_{\mathcal{F}_d} \leq M \| |F| - 1 \|_{\mathcal{F}_d}.$$

As in the one-dimensional case, Theorem 3.1 can be reduced to proving the following orthogonal coercivity estimate.

Theorem 3.2 (Orthogonal coercivity on \mathbb{C}^d). *Fix $d \geq 1$. There exists a constant $C_d > 0$ such that every $U \in \mathcal{F}_d$ with $U(0) = 0$ satisfies*

$$\|U\|_{\mathcal{F}_d}^2 \leq C_d^2 \|\rho \circ U\|_{\mathcal{F}_d}^2.$$

This is the analogue of Theorem 2.1 in the higher dimensional setting. As in the one-dimensional Fock space, we still have an orthonormal basis of monomials, except here they are of the form

$$e_\alpha(z) := \frac{z^\alpha}{\sqrt{\alpha!}}$$

where $\alpha \in \mathbb{N}^d$. As before, an important point in proving the orthogonal coercivity bound is in understanding the regions of space where e_α concentrates most of its mass. This allows one to carry out the analogue of the frequency localization in Section 2.5. This part is relatively similar to the one-dimensional case.

The part of the proof that requires a relatively new idea is the analogue of the circle estimates in Section 2.4. This is clarified by writing the $\|\cdot\|_{\mathcal{F}_d}$ norm in polar form

$$\|F\|_{\mathcal{F}_d}^2 = \frac{2}{(d-1)!} \int_0^\infty r^{2d-1} e^{-r^2} \left(\int_{S^{2d-1}} |F(r\omega)|^2 d\sigma_d(\omega) \right) dr,$$

and so, one needs analogues (which hold for functions on S^{2d-1}) of the following estimates:

- (1) the local circle estimate $\|P\|_{L^2(S^1)}^2 \leq 144L \|\rho \circ P\|_{L^2(S^1)}^2$ for trigonometric polynomials with L positive frequencies;
- (2) the high-frequency estimate $\|P\|_{L^2(S^1)}^2 \leq 32 \|\rho \circ P\|_{L^2(S^1)}^2$ for bands $\{N, \dots, N+L-1\}$ with $1343L^2 \leq N^2$.

Fortunately, by a simple averaging argument, both of these one-dimensional estimates can be lifted to the higher dimensional sphere S^{2d-1} . We state below the lifted estimates. The analogues of the local estimates and high frequency estimates are given, respectively, as follows.

Proposition 3.3 (Lifted local estimate). *Let $E \subset \mathbb{N}_{\geq 1}$ be finite with $|E| = L$, and let*

$$P(\omega) = \sum_{n \in E} G_n(\omega),$$

where each G_n is the restriction to S^{2d-1} of a homogeneous holomorphic polynomial of degree n . Then

$$\|P\|_{L^2(S^{2d-1})}^2 \leq 144L \|\rho \circ P\|_{L^2(S^{2d-1})}^2.$$

Proposition 3.4 (Lifted high-frequency estimate). *Let $N, L \geq 1$ satisfy $1343L^2 \leq N^2$, and let*

$$P(\omega) = \sum_{m=0}^{L-1} G_{N+m}(\omega),$$

where each G_{N+m} is the restriction to S^{2d-1} of a homogeneous holomorphic polynomial of degree $N+m$. Then

$$\|P\|_{L^2(S^{2d-1})}^2 \leq 32 \|\rho \circ P\|_{L^2(S^{2d-1})}^2.$$

With this in hand, the proof of Theorem 3.1 follows a similar path to the one-dimensional case (albeit with some added technical, but routine, details). To obtain Theorem 1.5, one may combine the above ideas with the ingredients described above for the one-dimensional Hermite case. While these are the core additional ingredients, the execution is rather technical, which is why we have omitted it.

3.2. Proof of Theorem 1.8. We now show how Theorem 1.8 follows from Theorem 1.3 and [13, Corollary 1.2]. The argument has three main ingredients. The first is an identity between the norm of the gradient of $|F|$ and $|F'|$, which converts the log-Sobolev deficit for entire functions into a Gaussian log-Sobolev deficit for the modulus $|F|$. The second is the dimension-free Gaussian log-Sobolev inequality due to Dolbeault–Esteban–Figalli–Frank–Loss [13]. The final ingredient is the standard covariance property of the Fock space, which will allow us to transfer the local stability at the constant function provided by Theorem 1.3 into local stability at any other extremizer.

Before moving on with the proof of Theorem 1.8, we state some preliminary facts. For $\beta \in \mathbb{C}$, let

$$G_\beta(z) := e^{\beta z - |\beta|^2/2}, \quad z \in \mathbb{C}.$$

Apart from being the class of extremizers to the logarithmic Sobolev inequality, these functions also constitute the family of normalized reproducing kernels for the Bargmann-Fock space. In particular, we have $G_\beta \in \mathcal{F}^2(\mathbb{C})$ with $\|G_\beta\|_{\mathcal{F}} = 1$, and $|G_\beta(z)| = e^{\operatorname{Re}(\beta z) - |\beta|^2/2}$. We define the *Bargmann shift* $W_\beta : \mathcal{F}^2(\mathbb{C}) \rightarrow \mathcal{F}^2(\mathbb{C})$ by

$$(W_\beta H)(z) := e^{\beta z - |\beta|^2/2} H(z - \bar{\beta}), \quad H \in \mathcal{F}^2(\mathbb{C}).$$

When computing the Fock norm of $W_\beta H$, a change of variables $w = z - \bar{\beta}$ together with the identity $|z|^2 = |w|^2 + 2\operatorname{Re}(\beta w) + |\beta|^2$ makes the weight $e^{2\operatorname{Re}(\beta z) - |\beta|^2 - |z|^2}$ become $e^{-|w|^2}$. Hence, W_β is an isometry. A direct computation yields $W_\beta W_{-\beta} = I = W_{-\beta} W_\beta$, so W_β is unitary on $\mathcal{F}^2(\mathbb{C})$ with $W_\beta^{-1} = W_{-\beta}$. Note also that $W_\beta \mathbf{1} = G_\beta$ and $W_{-\beta} G_\beta = \mathbf{1}$.

Next, we claim that for every $F \in \mathcal{F}^2(\mathbb{C})$ and every $\beta \in \mathbb{C}$,

$$(21) \quad \| |W_{-\beta} F| - \mathbf{1} \|_{\mathcal{F}} = \| |F| - |G_\beta| \|_{\mathcal{F}}.$$

Since $W_{-\beta}$ is unitary on $\mathcal{F}^2(\mathbb{C})$ and $\|\mathbf{1}\|_{\mathcal{F}} = \|G_\beta\|_{\mathcal{F}} = 1$, expanding both sides reduces (21) to the identity

$$\langle |W_{-\beta} F|, \mathbf{1} \rangle_{\mathcal{F}} = \langle |F|, |G_\beta| \rangle_{\mathcal{F}}.$$

This latter identity follows from a direct change of variables. Indeed, by writing

$$|W_{-\beta} F(z)| = e^{-\operatorname{Re}(\beta z) - |\beta|^2/2} |F(z + \bar{\beta})|$$

and substituting $w = z + \bar{\beta}$, the exponent $-\operatorname{Re}(\beta z) - |\beta|^2/2 - |z|^2$ becomes

$$\operatorname{Re}(\beta w) - |\beta|^2/2 - |w|^2 = \log |G_\beta(w)| - |w|^2.$$

Thus,

$$\begin{aligned} \frac{1}{\pi} \int_{\mathbb{C}} |W_{-\beta} F(z)| e^{-|z|^2} dm(z) &= \frac{1}{\pi} \int_{\mathbb{C}} |F(w)| |G_{\beta}(w)| e^{-|w|^2} dm(w) \\ &= \langle |F|, |G_{\beta}| \rangle_{\mathcal{F}}. \end{aligned}$$

We are now ready to prove Theorem 1.8.

Proof of Theorem 1.8. Let $F \in \mathcal{F}^2(\mathbb{C})$ with $\|F\|_{\mathcal{F}} = 1$. We may assume $\|F'\|_{\mathcal{F}} < \infty$, since otherwise the left-hand side of Theorem 1.8 is infinite and there is nothing to prove.

Step 1: A gradient identity. A direct application of the Cauchy–Riemann equations gives $|\nabla F|^2 = 2|F'|^2$. We also have, for entire F , the standard identity

$$(22) \quad |\nabla|F||^2 = |F'|^2 \quad \text{a.e. on } \mathbb{C}.$$

To see this, write $F = u + iv$ with u, v real-valued and apply the Cauchy–Riemann equations $u_{x_1} = v_{x_2}$, $u_{x_2} = -v_{x_1}$. On $\{F \neq 0\}$, $\nabla|F| = (u\nabla u + v\nabla v)/|F|$, and a short computation gives

$$(uu_{x_1} + vv_{x_1})^2 + (uu_{x_2} + vv_{x_2})^2 = (u^2 + v^2)(u_{x_1}^2 + v_{x_1}^2),$$

so that $|\nabla|F||^2 = u_{x_1}^2 + v_{x_1}^2 = |F'|^2$. As the zero set of an entire function on \mathbb{C} is discrete, (22) follows. In particular, $|F| \in H^1(d\gamma)$ with

$$(23) \quad \int_{\mathbb{C}} |\nabla|F||^2 d\gamma = \|F'\|_{\mathcal{F}}^2.$$

Step 2: The application of [13, Corollary 1.2]. Let $d\tilde{\gamma}(x) = e^{-\pi|x|^2} dx$ on \mathbb{R}^2 denote the normalized Gaussian measure. We recall that [13, Corollary 1.2] states that, for every $u \in H^1(d\tilde{\gamma})$,

$$(24) \quad \int_{\mathbb{R}^2} |\nabla u|^2 d\tilde{\gamma} - \pi \int_{\mathbb{R}^2} u^2 \ln \left(\frac{u^2}{\|u\|_{L^2(\tilde{\gamma})}^2} \right) d\tilde{\gamma} \geq \frac{\beta_* \pi}{2} \inf_{b \in \mathbb{R}^2, a \in \mathbb{R}} \int_{\mathbb{R}^2} (u - a e^{b \cdot x})^2 d\tilde{\gamma},$$

where $\beta_* > 0$ is an absolute constant. We now apply (24) to $u(x) := |F|(\sqrt{\pi}x)$. Since the change of variables $z = \sqrt{\pi}x$ sends $\tilde{\gamma}$ to γ , tracking each term in (24) and using (23) for the gradient contribution, we obtain

$$\|u\|_{L^2(d\tilde{\gamma})} = \|F\|_{\mathcal{F}} = 1, \quad \int_{\mathbb{R}^2} |\nabla u|^2 d\tilde{\gamma} = \pi \int_{\mathbb{C}} |\nabla|F||^2 d\gamma = \pi \|F'\|_{\mathcal{F}}^2,$$

$$\pi \int_{\mathbb{R}^2} u^2 \ln u^2 d\tilde{\gamma} = \pi \int_{\mathbb{C}} |F|^2 \ln |F|^2 d\gamma.$$

For the deficit term on the right-hand side of (24), we write $b = (b_1, b_2) \in \mathbb{R}^2$ and set $\beta := (b_1 - ib_2)/\sqrt{\pi} \in \mathbb{C}$, so that $b \cdot x = \text{Re}(\beta z)$. Then

$$\int_{\mathbb{R}^2} (u - a e^{b \cdot x})^2 d\tilde{\gamma} = \int_{\mathbb{C}} (|F(z)| - a e^{\text{Re}(\beta z)})^2 d\gamma(z).$$

Substituting all of the above into (24) and dividing by π , we obtain

$$(25) \quad \|F'\|_{\mathcal{F}}^2 - \int_{\mathbb{C}} |F|^2 \ln |F|^2 d\gamma \geq \frac{\beta_*}{2} \inf_{\beta \in \mathbb{C}, a \in \mathbb{R}} \| |F| - a e^{\text{Re}(\beta \cdot)} \|_{\mathcal{F}}^2.$$

Step 3: Rewriting the deficit term. Since $e^{\text{Re}(\beta z)} = e^{|\beta|^2/2} |G_{\beta}(z)|$ and $e^{|\beta|^2/2} \in (0, \infty)$, the substitution $a \mapsto a e^{|\beta|^2/2}$ shows that

$$\inf_{\beta \in \mathbb{C}, a \in \mathbb{R}} \| |F| - a e^{\text{Re}(\beta \cdot)} \|_{\mathcal{F}}^2 = \inf_{\beta \in \mathbb{C}, a \in \mathbb{R}} \| |F| - a |G_{\beta}| \|_{\mathcal{F}}^2.$$

For fixed $\beta \in \mathbb{C}$, set $t_\beta := \langle |F|, |G_\beta| \rangle_{\mathcal{F}}$. By Cauchy–Schwarz, $t_\beta \in [0, 1]$, and the optimal $a \in \mathbb{R}$ is $a = t_\beta$, with minimum value $1 - t_\beta^2$. Since $\| |F| - |G_\beta| \|_{\mathcal{F}}^2 = 2(1 - t_\beta)$ and $1 - t_\beta^2 = (1 - t_\beta)(1 + t_\beta) \geq \frac{1}{2} \cdot 2(1 - t_\beta)$, we obtain

$$(26) \quad \inf_{\beta \in \mathbb{C}, a \in \mathbb{R}} \| |F| - a |G_\beta| \|_{\mathcal{F}}^2 \geq \frac{1}{2} \inf_{\beta \in \mathbb{C}} \| |F| - |G_\beta| \|_{\mathcal{F}}^2.$$

Step 4: The application of Theorem 1.3. Fix any $\beta \in \mathbb{C}$ and set $H := W_{-\beta}F \in \mathcal{F}^2(\mathbb{C})$. By (21), $\| |H| - 1 \|_{\mathcal{F}} = \| |F| - |G_\beta| \|_{\mathcal{F}}$. By unitarity of $W_{-\beta}$ and the identity $W_{-\beta}G_\beta = \mathbf{1}$, we have

$$\| |H| - c \|_{\mathcal{F}} = \| W_{-\beta}(F - cG_\beta) \|_{\mathcal{F}} = \| F - cG_\beta \|_{\mathcal{F}} \quad \text{for every } c \in \mathbb{C}.$$

Theorem 1.3 applied to H therefore yields

$$\inf_{|c|=1} \| F - cG_\beta \|_{\mathcal{F}} \leq M \| |F| - |G_\beta| \|_{\mathcal{F}}.$$

Squaring and taking the infimum over $\beta \in \mathbb{C}$, we obtain

$$(27) \quad \inf_{\substack{\beta \in \mathbb{C} \\ |c|=1}} \| F - cG_\beta \|_{\mathcal{F}}^2 \leq M^2 \inf_{\beta \in \mathbb{C}} \| |F| - |G_\beta| \|_{\mathcal{F}}^2.$$

Combining (25), (26) and (27), it follows that

$$\| F' \|_{\mathcal{F}}^2 - \int_{\mathbb{C}} |F|^2 \ln |F|^2 d\gamma \geq \frac{\beta_\star}{4M^2} \inf_{\substack{\beta \in \mathbb{C} \\ |c|=1}} \| F - c e^{\beta z - |\beta|^2/2} \|_{\mathcal{F}}^2.$$

This proves Theorem 1.8. □

4. DISCUSSIONS ON THE LEAN VERIFICATION

This section discusses the formalization of Theorem 1.5. The file `DimdPoly.lean` provides a self-contained version of the main statement that was Lean verified³. When writing these statements, our goal was to maximize intelligibility for mathematicians who are not well-versed in Lean, even if it sometimes resulted in *less idiomatic* Lean code.

The full Lean code is available at <https://github.com/susannabertolini/PhaseRetrieval>. We refer the reader to the READMEs for an explanation of the structure of the repository. We emphasize that although the reviewer facing statements have been carefully curated by the authors with the objective of making them understandable to a broad audience, the rest of the Lean code is of poor quality and is simply there to verify correctness – it is not meant to be digested by the reader.

This autoformalization, and the trust we placed in it, is possible only because of the enormous efforts by the Mathlib community [31] in carefully formalizing mathematical objects in Lean and ensuring that their meaning is what the mathematical community expects. In particular, our formalization contains plenty of mathematical statements such as `MeasureTheory.volume.withDensity` or `Real.sqrt` whose construction is not detailed in this section, but which have been thoughtfully designed and very carefully reviewed by the Mathlib community.

We now attempt to give an explanation of the content of the `DimdPoly.lean` file. The formalized theorem discussed here is Theorem 1.5.

Throughout the discussion, we fix a positive dimension d . In the Lean file, the dimension $d \in \mathbb{N}$ is declared once as a file-level variable, with the hypothesis (named `hd`) that $0 < d$:

```
variable {d : ℕ} (hd : 0 < d)
```

³This is in the spirit of the comparator approach [17] to safely verify Lean proofs. Indeed, the repository also contains an appropriate `Challenge.lean` file to be verified.

4.1. **The canonical basis.** We start by defining the one-dimensional basis. For $k, n \in \mathbb{N}$ and $z \in \mathbb{C}$, we use

$$\Phi_{k,n}(z) = \frac{1}{\sqrt{n!k!}} \sum_{j=0}^{\min\{n,k\}} (-1)^j j! \binom{k}{j} \binom{n}{j} z^{n-j} \bar{z}^{k-j}.$$

In Lean, everything has a type. Types are a primitive notion, in the spirit of sets. For example, 0 has type \mathbb{N} , expressed as $(0 : \mathbb{N})$. The function $f(x) := x^2$ has type $\mathbb{R} \rightarrow \mathbb{R}$. One can define new functions, that take various inputs of various types, as follows:

```
def DefinitionName (variable0 : Type0) (variable1 variable2 : Type1) : OutputType :=
  Insert your definition
```

With this in mind, $\Phi_{k,n}(z)$ is expressed as a function `HermitePoly` that takes two integers k n , one complex number z and returns a complex number, as follows

```
def HermitePoly (k n : ℕ) (z : ℂ) : ℂ :=
  (Real.sqrt ((Nat.factorial n) * (Nat.factorial k)))-1 *
  ∑ j ∈ Finset.range (min n k + 1),
  ((-1) ^ j) * (Nat.factorial j) *
  (Nat.choose n j) * (Nat.choose k j) *
  z ^ (n - j) * (star z) ^ (k - j)
```

Here, `Finset.range N` is the finite set $\{0, \dots, N - 1\}$. Thus, `∑ j ∈ Finset.range (min n k + 1)` expresses the sum over $j = 0, \dots, \min\{n, k\}$.

The d -dimensional Hermite polynomials are tensor products of the $\Phi_{k,n}$:

$$\Phi_{\vec{\kappa}, \vec{\alpha}}(\vec{z}) := \prod_{q=0}^{d-1} \Phi_{\kappa_q, \alpha_q}(z_q).$$

In Lean, we define points $\kappa \in \mathbb{N}^d$ as maps from `Fin d` to \mathbb{N} , written as $\kappa : \text{Fin } d \rightarrow \mathbb{N}$. Here, `Fin d` is the finite type with elements $\{0, 1, \dots, d - 1\}$. In other words, the vector $\vec{z} := (z_0, \dots, z_{d-1})$ is defined as the map that sends $q \in \{0, \dots, d - 1\}$ to z_q . In Lean, this is written as follows:

```
def Φ (κ α : Fin d → ℕ) (z : Fin d → ℂ) : ℂ :=
  ∏ q : Fin d, HermitePoly (κ q) (α q) (z q)
```

The notation is meant to be read literally: κ and α are multi-indices, and z is a point of \mathbb{C}^d . The notation `∏ q : Fin d, ...` denotes the finite product over all coordinates.

We next define finite linear combinations of the basis functions, indexed by their coefficients. Given $\vec{\kappa} \in \mathbb{N}^d$, and a finitely supported coefficient family $\{F(\vec{\alpha})\}_{\vec{\alpha} \in \mathbb{N}^d}$, with $F(\vec{\alpha}) \in \mathbb{C}$, we define

$$P_F(\vec{z}) = \sum_{\vec{\alpha} \in \text{Supp } F} F(\vec{\alpha}) \Phi_{\vec{\kappa}, \vec{\alpha}}(\vec{z}).$$

The Lean definition is:

```
def TrueHermitePoly (κ : Fin d → ℕ)
  (F : Finsupp (Fin d → ℕ) ℂ) : (Fin d → ℂ) → ℂ :=
  fun z ↦ F.sum (fun α Fα ↦ Fα * Φ κ α z)
```

Let F be a finitely supported map, which we denote as `Finsupp`. The expression `F.sum` is the finite sum over the support of F . The definition `TrueHermitePoly κ F` is a function $\mathbb{C}^d \rightarrow \mathbb{C}$.

We define the set of all (finite) linear combinations of the basis vectors by

$$\mathcal{H}_{\vec{\kappa}, \text{fin}}(\mathbb{C}^d) := \text{span}\{\Phi_{\vec{\kappa}, \vec{\alpha}} : \vec{\alpha} \in \mathbb{N}^d\},$$

which in Lean is the range of all the elements obtained with the construction above.

```
def TrueHermitePolys (κ : Fin d → ℕ) :
  Set ((Fin d → ℂ) → ℂ) :=
  Set.range (TrueHermitePoly κ)
```

4.2. **The Gaussian measure.** The Gaussian measure has density

$$\frac{1}{\pi^d} e^{-\sum_{q=0}^{d-1} |z_q|^2}$$

In Lean, we write

```
def GaussianDensity (z : Fin d → ℂ) : ℝ :=
  (1 / Real.pi ^ d) *
  Real.exp (-Finset.sum Finset.univ fun q : Fin d ↦ ‖z q‖ ^ 2)
```

From this Gaussian density, which is a function, we can define a measure. In Lean, measures are not built from densities taking values in the reals, but from densities taking values in the non-negative reals and possibly $+\infty$, which are denoted by `ENNReal`. With this in mind, one builds the Gaussian measure as follows:

```
def γ : MeasureTheory.Measure (Fin d → ℂ) :=
  MeasureTheory.volume.withDensity fun z ↦ ENNReal.ofReal (GaussianDensity z)
```

We can now define the complex-valued $L^2(d\gamma)$ as `(MeasureTheory.Lp ℂ 2 (γ (d := d)))`. In this case, Lean was not able to infer the value of `d` automatically from the measure $d\gamma$ and it must be provided by us. We define `TrueHermiteClosure κ` as $\mathcal{H}_{\vec{\kappa}}(\mathbb{C}^d) := \overline{\mathcal{H}_{\vec{\kappa}, \text{fin}}(\mathbb{C}^d)}$, which is the closure of the set of L^2 functions that are γ -almost-everywhere equal to an element in `TrueHermitePolys`.

```
def TrueHermiteFunctions (κ : Fin d → ℕ) :
  Set (MeasureTheory.Lp ℂ 2 (γ (d := d))) :=
  closure { f | ∃ P ∈ TrueHermitePolys κ, f =m[γ] P }
```

In Lean, L^p functions are defined as equivalence classes of functions which are a priori known to be L^p -integrable. Therefore, we can not directly take the closure of `TrueHermitePolys`.

4.3. **Statement of the main result.** The theorem in `DimdPoly.lean` is the following stability statement:

Theorem. *Let $d > 0$ and $\vec{\kappa} \in \mathbb{N}^d$. For every $P_0 \in \mathcal{H}_{\vec{\kappa}, \text{fin}}(\mathbb{C}^d)$ there exists a constant $M = M(\vec{\kappa}, P_0) > 0$ such that for every $F \in \mathcal{H}_{\vec{\kappa}}(\mathbb{C}^d)$ we have*

$$\inf_{|\theta|=1} \|P_0 - \theta F\|_{\gamma_d}^2 \leq M^2 \| |P_0| - |F| \|_{\gamma_d}^2.$$

In order to state this theorem in Lean, we rewrite it in several key ways. First, in the statement of the theorem, we must separate the hypothesis that P_0 is a function `(Fin d → ℂ) → ℂ` from the statement that P_0 belongs to `TrueHermitePolys κ`. Second, in order to avoid the heavy set/inf notation, we prove the equivalent statement

$$\exists \theta \in \mathbb{C} \text{ such that } |\theta| = 1 \text{ and } \|P_0 - \theta F\|_{\gamma_d}^2 \leq M^2 \| |P_0| - |F| \|_{\gamma_d}^2.$$

With these modifications, the Lean statement of the stability theorem reads as follows:

```
theorem stable_phase_retrieval
  (hd : 0 < d) (κ : Fin d → ℕ)
  {P : (Fin d → ℂ) → ℂ} (hP : P ∈ TrueHermitePolys κ) :
  ∃ M : ℝ, 0 < M ∧
  ∀ Q ∈ TrueHermiteFunctions κ,
  ∃ θ : ℂ, ‖θ‖ = 1 ∧
  ∫ z, ‖P z - θ * Q z‖ ^ 2 ∂ γ
  ≤ M ^ 2 * ∫ z, (‖P z‖ - ‖Q z‖) ^ 2 ∂ γ := by
```

Remark 4.1. The Lean statement utilizes θ instead of λ . The discrepancy in notation is due to the fact that, in Lean, the constant λ is reserved notation.

The reader may explore the statements (but not the proof) in an interactive Lean 4 environment on this live.lean-lang.org link. The service live.lean-lang.org is provided by the Lean Focused Research Organization.

REFERENCES

- [1] Luis Daniel Abreu and Hans G. Feichtinger. Function spaces of polyanalytic functions. In *Harmonic and complex analysis and its applications*, Trends Math., pages 1–38. Birkhäuser/Springer, Cham, 2014.
- [2] Rima Alaifari, Ingrid Daubechies, Philipp Grohs, and Rujie Yin. Stable phase retrieval in infinite dimensions. *Found. Comput. Math.*, 19(4):869–900, 2019.
- [3] Rima Alaifari and Philipp Grohs. Phase retrieval in the general setting of continuous frames for Banach spaces. *SIAM J. Math. Anal.*, 49(3):1895–1911, 2017.
- [4] Rima Alaifari, Ben Pineau, Mitchell A. Taylor, and Matthias Wellershoff. Cheeger’s constant for the Gabor transform and ripples. *arXiv preprint arXiv:2512.18058*, 2025.
- [5] Wedad Alharbi, Salah Alshabhi, Daniel Freeman, and Dorsa Ghoreishi. Locality and stability for phase retrieval. *Sampl. Theory Signal Process. Data Anal.*, 22(1):Paper No. 10, 16, 2024.
- [6] David Kurniadi Angdinata, Evan Chen, Chris Cummins, Ben Eltschig, Dejan Grubisic, Leopold Haller, Letong Hong, Andranik Kurghinyan, Kenny Lau, Hugh Leather, et al. ABC implies that Ramanujan’s tau function misses almost all primes. *arXiv preprint arXiv:2603.29970*, 2026.
- [7] Scott Armstrong and Julia Kempe. Formalization of De Giorgi–Nash–Moser Theory in Lean. *arXiv preprint arXiv:2604.05984*, 2026.
- [8] Jameson Cahill, Peter G. Casazza, and Ingrid Daubechies. Phase retrieval in infinite-dimensional Hilbert spaces. *Trans. Amer. Math. Soc. Ser. B*, 3:63–76, 2016.
- [9] Robert Calderbank, Ingrid Daubechies, Daniel Freeman, and Nikki Freeman. Stable phase retrieval for infinite dimensional subspaces of $L_2(\mathbb{R})$. *arXiv preprint arXiv:2203.03135*, 2022.
- [10] Manuel Camúñez, Enrique García-Sánchez, and David de Hevia. A characterization of complex stable phase retrieval in Banach lattices. *Proc. Amer. Math. Soc.*, 154(5):2083–2096, 2026.
- [11] Michael Christ, Ben Pineau, and Mitchell A. Taylor. Examples of Hölder-stable phase retrieval. *Math. Res. Lett.*, 31(5):1339–1352, 2024.
- [12] Leonardo de Moura and Sebastian Ullrich. The Lean 4 Theorem Prover and Programming Language. In *Automated Deduction – CADE 28: 28th International Conference on Automated Deduction, Virtual Event, July 12–15, 2021, Proceedings*, page 625–635, Berlin, Heidelberg, 2021. Springer-Verlag.
- [13] Jean Dolbeault, Maria J. Esteban, Alessio Figalli, Rupert L. Frank, and Michael Loss. Sharp stability for Sobolev and log-Sobolev inequalities, with optimal dimensional dependence. *Camb. J. Math.*, 13(2):359–430, 2025.
- [14] Oliver Dressler. Lean Theorem Prover MCP. <https://github.com/o0o0o0o/lean-lsp-mcp>, 2026.
- [15] Rupert L. Frank, Fabio Nicola, and Paolo Tilli. The generalized Wehrl entropy bound in quantitative form. *J. Eur. Math. Soc. (JEMS)*, 2025. Published online first.
- [16] D. Freeman, T. Oikhberg, B. Pineau, and M. A. Taylor. Stable phase retrieval in function spaces. *Math. Ann.*, 390(1):1–43, 2024.
- [17] Lean FRO. Comparator. <https://github.com/leanprover/comparator>, 2026.
- [18] Hartmut Führ and Max Getter. Cheeger Bounds for Stable Phase Retrieval in Reproducing Kernel Hilbert Spaces. *arXiv preprint arXiv:2512.24169*, 2025.
- [19] Enrique García-Sánchez and David de Hevia. Isometric embeddings into $C(K)$ -spaces doing stable phase retrieval. *arXiv preprint arXiv:2512.08110*, 2025.
- [20] Enrique García-Sánchez, David de Hevia, and Mitchell A. Taylor. On the existence of large subspaces of $C(K)$ that perform stable phase retrieval. *arXiv preprint arXiv:2512.08114*, 2025.
- [21] Philipp Grohs and Lukas Liehr. Stable Gabor phase retrieval in Gaussian shift-invariant spaces via biorthogonality. *Constr. Approx.*, 59(1):61–111, 2024.
- [22] Philipp Grohs and Martin Rathmair. Stable Gabor phase retrieval and spectral clustering. *Comm. Pure Appl. Math.*, 72(5):981–1043, 2019.
- [23] Philipp Grohs and Martin Rathmair. Stable Gabor phase retrieval for multivariate functions. *J. Eur. Math. Soc. (JEMS)*, 24(5):1593–1615, 2022.
- [24] Leonard Gross. Logarithmic Sobolev inequalities. *Amer. J. Math.*, 97(4):1061–1083, 1975.
- [25] Sidharth Hariharan, Christopher Birkbeck, Seewoo Lee, Ho Kiu Gareth Ma, Bhavik Mehta, Auguste Poiroux, and Maryna Viazovska. A Milestone in Formalization: The Sphere Packing Problem in Dimension 8. *arXiv preprint arXiv:2604.23468*, 2026.

- [26] Boon Suan Ho. A generalization of Boppa's entropy inequality. *arXiv preprint arXiv:2601.19327*, 2026.
- [27] Vasily Ilin. Semi-autonomous formalization of the Vlasov-Maxwell-Landau equilibrium. *arXiv preprint arXiv:2603.15929*, 2026.
- [28] Philippe Jaming and Martin Rathmair. Gabor phase retrieval via semidefinite programming. *Found. Comput. Math.*, 26(1):245–311, 2026.
- [29] Martin Rathmair. Stable STFT Phase Retrieval and Poincaré Inequalities. *International Mathematics Research Notices*, 2024(22):14094–14114, 2024.
- [30] A. J. Stam. Some inequalities satisfied by the quantities of information of Fisher and Shannon. *Information and Control*, 2:101–112, 1959.
- [31] The mathlib Community. The Lean Mathematical Library. In *Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs, CPP 2020, New Orleans, LA, USA, January 2020*. ACM.
- [32] Fred B. Weissler. Logarithmic Sobolev inequalities for the heat-diffusion semigroup. *Trans. Amer. Math. Soc.*, 237:255–269, 1978.

DEPARTMENT OF MATHEMATICS, ETH ZÜRICH, RAMISTRASSE 101, 8092 ZÜRICH, SWITZERLAND.

Email address: susanna.bertolini@math.ethz.ch

CENTER FOR DATA SCIENCE, NEW YORK UNIVERSITY, NEW YORK, NEW YORK 10011, USA

Email address: jdedios@nyu.edu

COURANT INSTITUTE FOR MATHEMATICAL SCIENCES, NEW YORK UNIVERSITY

Email address: brp305@nyu.edu

INSTITUTO DE MATEMÁTICA PURA E APLICADA (IMPA), RIO DE JANEIRO, RJ, BRAZIL

Email address: joao.ramos@impa.br

DEPARTMENT OF MATHEMATICS, ETH ZÜRICH, RAMISTRASSE 101, 8092 ZÜRICH, SWITZERLAND.

Email address: mitchell.taylor@math.ethz.ch