Particle approximations of Wigner distributions for *n* arbitrary observables

Ralph Sabbagh,[1] Olga Movilla Miangolarra,[1] Hamid Hezari,[2] and Tryphon T. Georgiou[1]
[1]*Department of Mechanical and Aerospace Engineering, University of California, Irvine, California 92697, USA*
[2]*Department of Mathematics, University of California, Irvine, California 92697, USA*

(Received 27 September 2024; accepted 6 January 2025; published 27 January 2025)

A class of signed joint probability measures for n arbitrary quantum observables is derived and studied based on quasicharacteristic functions with symmetrized operator orderings of Margenau-Hill type. It is shown that the Wigner distribution associated with these observables can be rigorously approximated by such measures. These measures are given by affine combinations of Dirac delta distributions supported over the finite spectral range of the quantum observables and give the correct probability marginals when coarse-grained along any principal axis. We specialize to bivariate quasiprobability distributions for the spin measurements of spin-1/2 particles and derive their closed-form expressions. As a side result, we point out a connection between the convergence of these particle approximations and the Mehler-Heine theorem. Finally, we interpret the supports of these quasiprobability distributions in terms of repeated thought experiments.

DOI: 10.1103/PhysRevResearch.7.013102

I. INTRODUCTION

In 1986, Cohen and Scully developed bivariate quasiprobability distributions for the spin measurements of spin-1/2 particles [1]. Therein, two quasiprobability distributions based on quasicharacteristic functions induced from symmetrized operator orderings were studied. The first, termed the Margenau-Hill distribution, is given by the Fourier transform of the quasicharacteristic function

$$f_{\mathrm{MH}_1}(\xi_1, \xi_2) = \mathrm{tr}\left(\rho \frac{e^{i\xi_1 \hat{S}_1} e^{i\xi_2 \hat{S}_2} + e^{i\xi_2 \hat{S}_2} e^{i\xi_1 \hat{S}_1}}{2}\right),$$

where ρ is a density matrix and $\hat{S}_1$ and $\hat{S}_2$ are the spin operators along two arbitrary directions in the Bloch sphere. The second, termed the Wigner distribution, is given by the Fourier transform of the quasicharacteristic function

$$f_{\mathrm{W}}(\xi_1, \xi_2) = \mathrm{tr}(\rho e^{i\xi_1 \hat{S}_1 + i\xi_2 \hat{S}_2}),$$

and is the analog of the standard Wigner distribution for spin observables $\hat{S}_1$ and $\hat{S}_2$. In 1992, Chandler *et al.* derived the trivariate counterparts, with spin observables along mutually orthogonal directions [2]. Therein, it was shown that the computation for the trivariate Wigner distribution is simpler than its bivariate counterpart.

In 2020, Schwonnek and Werner studied the Wigner distribution for an arbitrary tuple of bounded Hermitian operators $(\hat{A}_1, \dots, \hat{A}_n)$ on a finite-dimensional Hilbert space [3], and defined it as the Fourier transform of the quasicharacteristic function

$$f_{\mathrm{W}}(\xi) = \mathrm{tr}(\rho e^{i\xi \cdot \hat{A}}),$$

where $\xi \cdot \hat{A} = \sum_{k=0}^{n} \xi_k \hat{A}_k,\ \xi \in \mathbb{R}^n$. The distribution is termed "Wigner" because it recovers the standard Wigner distribution when specialized to the canonical pair $\hat{A}_1 = \hat{X}$ and $\hat{A}_2 = \hat{P}$. Many of its basic properties, such as the support, location of singularities, positivity, and behavior under symmetry groups, were studied and illustrated with examples.

In many aspects, the standard Wigner distribution, defined by

$$p_{\mathrm{W}} := \frac{1}{(2\pi)^2} \mathcal{F}(f_{\mathrm{W}}),$$

where $\mathcal{F}(\cdot)$ is the Fourier transform, has lent itself as a convenient choice for the phase-space representation of quantum states. This, in large part, is due to the Fourier duality of the canonical pair and their continuous spectra. It is often represented by a bounded and continuous function that integrates to 1 and is sign indefinite: a salient nonclassical feature. In contrast, when the same definition is applied to an n-tuple of noncommuting matrices such as in Refs. [1–3], the nonclassicality becomes threefold. Not only is the distribution sign indefinite, but it is no longer a measure to begin with, and the support need not be discrete. The last two features are nonclassical because the distribution is intended to be a joint law on the outcomes of simultaneously measuring the discrete observables. Indeed, when the observables do commute, p_{W} is a classical discrete law. In general, however, it is a distribution supported beyond its intended set, with a rich singularity structure that is intimately related to the eigenvalues of the associated observables [3]. For instance, while the Margenau-Hill distribution studied in Ref. [1] is a discrete measure on the set of spin measurement outcomes $(\pm\hbar/2, \pm\hbar/2)$, the Wigner distribution for the same operators is supported on a disk of radius $\hbar/2$, with a complicated singularity near the boundary.

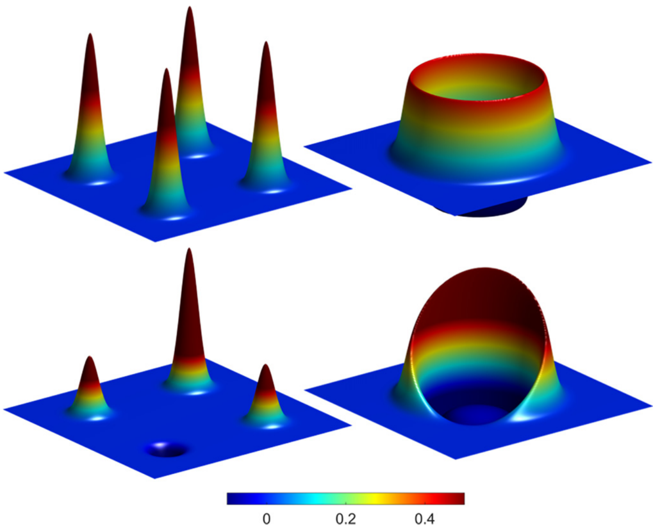

FIG. 1. Gaussian-regularized ($\varepsilon = 0.01$) bivariate quasiprobability distributions for the spin-1/2 observables along the x and y directions, at the state $\rho = \hat{I}/2$ (top) and $\rho = 0.5[1\ (1-i)/\sqrt{2};\ (1+i)/\sqrt{2}\ 1]$ (bottom) represented in the $\hat{S}_z$ eigenbasis. Left: Margenau-Hill quasiprobability distribution p_{MH_1} consisting of four Dirac delta distributions supported over the Cartesian product of spin-1/2 eigenvalues $(\pm\hbar/2) \times (\pm\hbar/2)$. Right: Wigner quasiprobability distribution p_{W} supported on a Disk of radius $\hbar/2$ with a complicated singularity near the boundary.

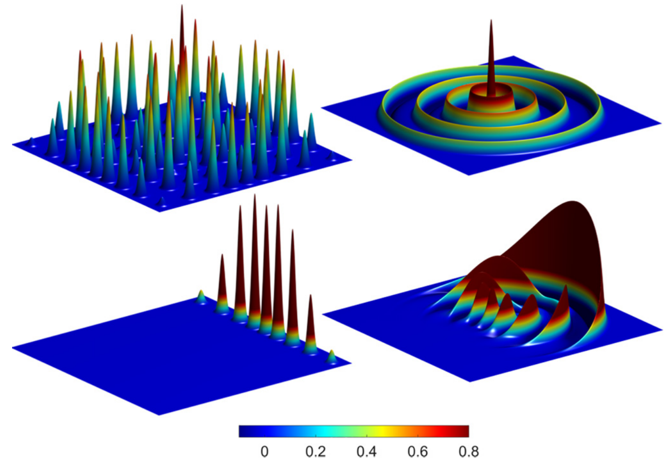

FIG. 2. Gaussian-regularized ($\varepsilon = 0.1$) bivariate quasiprobability distributions for the spin-4 observables along the x and y directions, at the maximally mixed state (top) and the $+4\hbar$ eigenstate along the y direction (bottom). Left: Margenau-Hill quasiprobability distribution p_{MH_1} consisting of 81 Dirac delta distributions, each of which is supported over a Cartesian product of spin eigenvalues, namely $(\pm k_1\hbar) \times (\pm k_2\hbar)$, where $k_1, k_2 \in \{0, \ldots, 4\}$. Right: Wigner quasiprobability distribution p_{W} supported on a disk of radius $4\hbar$ with singularities near concentric rings with radii $k\hbar$, $k = 0, \ldots, 4$.

Gaussian-regularized[1] plots for both distributions are shown in Fig. 1 and analogs for a spin-4 particle are shown in Fig. 2.

Although the Wigner distribution p_{W} lacks basic classical features, i.e., being a measure and having discrete support, its most remarkable classical feature still stands. It is the unique joint distribution for which the marginals of all linear combinations of the observables coincide with their quantum counterparts [3]. In contrast, the Margenau-Hill distribution in Refs. [1,2] is a discrete measure that is supported over the classical set of measurement outcomes of the observables, but does not give the correct probability marginals for all linear combinations like p_{W} does. Thus, each distribution possesses classical features expected from a joint probability distribution as well as nonclassical features arising from the noncommutativity of the observables. A study of these features began in Ref. [3] for the Wigner distribution p_{W}, and in this work we examine the features of a class of Margenau-Hill counterparts denoted by p_{MH_m}, where $m \in \mathbb{N}$.

Specifically, we introduce, analyze, and interpret the quasiprobability distributions

$$p_{\mathrm{MH}_m} = \frac{1}{(2\pi)^n}\mathcal{F}(f_{\mathrm{MH}_m}), \quad m \in \mathbb{N},$$

where f_{MH_m} are quasicharacteristic functions defined for an arbitrary tuple of Hermitian matrices $(\hat{A}_1, \ldots, \hat{A}_n)$ and a quantum state ρ; see Sec. III A, Eq. (3). It is shown that these distributions are real-valued, signed, and discrete probability measures given by affine combinations of Dirac delta distributions, and give the correct probability marginal when coarse-grained along any principal axis.

We shall refer to p_{MH_m} as the Margenau-Hill quasiprobability distribution of order m, and show that

$$\lim_{m\to\infty} p_{\mathrm{MH}_m} = p_{\mathrm{W}},$$

in a suitable topology that can be upgraded to that of uniform convergence if the distributions are smeared with an appropriate Schwarz function. And so, while the general Wigner distribution p_{W} is not a quasiprobability measure for the associated observables, it is not far from being one.

Lastly, we specialize to pairs of spin-1/2 observables along orthogonal directions and derive closed-form expressions for p_{MH_m} for any $m \in \mathbb{N}$. As a side result, we point out therein a connection between the convergence of the particle approximations p_{MH_m} to p_{W} and the Mehler-Heine theorem. We conclude by proposing an interpretation for the supports of the distributions p_{MH_m} and p_{W} in terms of repeated experiments. We emphasize that the numerical and analytical tractability showcased in the main results for general noncommuting matrices (see Sec. III Proposition 3) as well as for spin matrices in particular (see Sec. IV Proposition 12) exemplifies why it may be much more desirable to work with these particle approximations rather than the Wigner distribution, especially in high dimensions. Namely, the particle approximations p_{MH_m} are always signed probability measures, i.e., they can be used to assign mass on Borel sets of the generalized phase plane, and can be computed exactly and systematically. In contrast, the Wigner distribution corresponding to the same observables need not have a closed form expression, exhibits complicated singularities, and is never a measure unless all operators commute [4].

[1] See the end of Sec. III for details.

II. PRELIMINARIES

In this section, we establish notation and survey relevant facts related to the theory of distributions. Then, we introduce the Wigner distribution and state some of its properties. Finally, we introduce the Lie-Trotter product formula and the Mehler-Heine theorem. Throughout, we fix a tuple $(\hat{A}_1, \dots, \hat{A}_n)$ of self-adjoint operators on a finite-dimensional Hilbert space of dimension d and define for $\xi = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n$ the linear combination

$$\xi \cdot \hat{A} := \sum_{k=1}^{n} \xi_k \hat{A}_k.$$

Lastly, we fix a quantum state to be given by a density operator ρ, i.e., $\rho \in \mathbb{C}^{d\times d}$, with $\rho = \rho^\dagger \geqslant 0$ and $\mathrm{tr}(\rho) = 1$.

A. Distributions

Let $C_0^\infty(\mathbb{R}^n) \subseteq \mathcal{S}(\mathbb{R}^n) \subseteq C^\infty(\mathbb{R}^n)$ denote the spaces of compactly supported smooth functions, Schwarz functions, and smooth functions on $\mathbb{R}^n$, respectively, and $\mathcal{D}'(\mathbb{R}^n) \supseteq \mathcal{S}'(\mathbb{R}^n) \supseteq \mathcal{E}'(\mathbb{R}^n)$ the corresponding dual spaces of distributions, tempered distributions, and compactly supported distributions on $\mathbb{R}^n$, respectively. The support and singular support of $p \in \mathcal{D}'(\mathbb{R}^n)$ are denoted by $\mathrm{supp}(p)$ and $\mathrm{singsupp}(p)$, respectively. The n-dimensional Fourier transform

$$\mathcal{F}(f)(\xi) = \int_{\mathbb{R}^n} f(x) e^{-ix\cdot\xi} \mathrm{d}x, \quad \xi \in \mathbb{R}^n,$$

is an automorphism on $\mathcal{S}(\mathbb{R}^n)$ and it induces naturally an automorphism on the dual $\mathcal{S}'(\mathbb{R}^n)$. The inverse map is given by Fourier's inversion formula

$$f(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \mathcal{F}(f)(\xi) e^{ix\cdot\xi} \mathrm{d}\xi, \quad x \in \mathbb{R}^n.$$

Next, we state one direction of the Paley-Wiener-Schwartz theorem [5, Theorem 7.3.1], which relates the support properties of a function to analyticity properties of its Fourier transform. To this end, recall that the *supporting function* of a convex compact set $K \subseteq \mathbb{R}^n$ is

$$H_K(x) = \sup_{y\in K} \langle x, y\rangle, \quad x \in \mathbb{R}^n.$$

Theorem (Paley-Wiener-Schwartz): Let K be a convex compact set in $\mathbb{R}^n$. If f is analytic everywhere in $\mathbb{C}^n$ and satisfies

$$|f(z)| \leqslant C e^{H_K(\mathrm{Im}(z))}, \quad z \in \mathbb{C}^n,$$

where $\mathrm{Im}(z)$ denotes the imaginary part of z and $C > 0$, then the restriction of f to $\mathbb{R}^n$ is the Fourier transform of a distribution $p \in \mathcal{E}'(\mathbb{R}^n)$ with

$$\mathrm{supp}(p) \subseteq K.$$

This theorem will be used in Sec. III to study support properties of the distributions p_{MH_m}. A converse statement of the theorem also holds, see [5, Theorem 7.3.1], but will not be needed herein.

Finally, we say that a sequence of distributions p_n converges to p in $\mathcal{D}'(\mathbb{R}^n)$ as $n \to \infty$, or simply, $p_n \to p \in \mathcal{D}'(\mathbb{R}^n)$ as $n \to \infty$, if for all $\phi \in C_0^\infty(\mathbb{R}^n)$,

$$\lim_{n\to\infty} \langle p_n, \phi\rangle = \langle p, \phi\rangle.$$

The same definition applies for $p_n, p \in \mathcal{S}'(\mathbb{R}^n)$ or $\mathcal{E}'(\mathbb{R}^n)$ with respect to test functions ϕ taken in $\mathcal{S}(\mathbb{R}^n)$ or $C^\infty(\mathbb{R}^n)$, respectively.

B. Wigner quasiprobability distribution p_{W}

The Wigner quasiprobability distribution p_{W} associated with the observables $\hat{A}_1, \dots, \hat{A}_n$ and the quantum state ρ is a real-valued distribution in $\mathcal{S}'(\mathbb{R}^n)$ given by

$$p_{\mathrm{W}} = \frac{1}{(2\pi)^n} \mathcal{F}(f_{\mathrm{W}}),$$

where f_{W} is the quasicharacteristic function

$$f_{\mathrm{W}}(\xi) = \mathrm{tr}\big(\rho e^{i\xi\cdot\hat{A}}\big), \quad \xi \in \mathbb{R}^n.$$

It was shown in Ref. [3] that p_{W} is compactly supported, i.e., $p_{\mathrm{W}} \in \mathcal{E}'(\mathbb{R}^n)$, and moreover that

$$\mathrm{supp}(p_{\mathrm{W}}) \subseteq R,$$

where the compact convex set $R \subseteq \mathbb{R}^n$ is the *joint numerical range* of the operators $\hat{A}_1, \dots, \hat{A}_n$. In other words, R is the set of all vectors $a \in \mathbb{R}^n$ with components $a_i = \mathrm{tr}(\sigma\hat{A}_i)$ for some density operator σ. Furthermore, it was shown in Ref. [3] that

$$\mathrm{singsupp}(p_{\mathrm{W}}) \subseteq S,$$

where S is the closure of the set of all vectors $a \in \mathbb{R}^n$ with components $a_i = \mathrm{tr}(\sigma\hat{A}_i)$ for a subset of density operators σ, namely, the ones that correspond to non-degenerate eigenstates of $\xi \cdot \hat{A}$. The set S is semialgebraic (algebraic if $n = 2$) and its convex hull is R.

C. Lie-Trotter product formula

Given any complex-valued matrices $\hat{A}_1, \dots, \hat{A}_n$, then

$$\lim_{k\to\infty} \left(\prod_{i=1}^{n} e^{\hat{A}_i/k}\right)^k = e^{\sum_i \hat{A}_i}. \tag{1}$$

The proof for the case of more than two matrices, i.e., $n > 2$, follows verbatim the proof given in [6, Theorem 2.10] for two matrices. The formula is also implied by the proof of Lemma 1 in Appendix B.

The convergence properties of the Lie-Trotter product formula are key to proving the convergence of the Margenau-Hill quasiprobability distributions p_{MH_m} to the Wigner distribution p_{W} as $m \to \infty$. These distributions will be defined and studied in detail in Sec. III.

D. Mehler-Heine theorem

The Mehler-Heine theorem describes the asymptotic behavior of the Jacobi polynomials as their degree tends to infinity. These polynomials arise by trotterizing exponentials of spin-1/2 operators along orthogonal directions; see Lemma 2 in Sec. IV. They will be used to elucidate the nature of the convergence of the quasiprobability distributions p_{MH_m} as $m \to \infty$. In particular, we show that the convergence of the quasicharacteristic functions f_{MH_m} to f_{W} is a special case of the Mehler-Heine theorem; see Sec. IV.

We first define the Jacobi polynomials

$$P_n^{(\alpha,\beta)}(z) = \frac{1}{n!}\sum_{k=0}^{n}\binom{n}{k}c_{n,k}^{\alpha,\beta}\left(\frac{z-1}{2}\right)^k, \quad z\in\mathbb{C}, \tag{2}$$

for $\alpha,\beta\in\mathbb{R}$ and $c_{n,k}^{\alpha,\beta} = (n+\alpha+\beta+1)^{(k)}(\alpha+k+1)^{(n-k)}$. Then, the Mehler-Heine theorem [7] states that

$$\lim_{n\to\infty} n^{-\alpha}P_n^{(\alpha,\beta)}\left(\cos\left(\frac{z}{n}\right)\right) = \left(\frac{z}{2}\right)^{-\alpha}J_\alpha(z),$$

uniformly on compact subsets of $\mathbb{C}$, where $J_\alpha(z)$ is the Bessel function of the first kind of order α.

III. RESULTS

In this section, we define the Margenau-Hill quasiprobability distributions p_{MH_m}, $m\in\mathbb{N}$, and study their properties. These are discrete signed probability measures associated with the observables $\hat{A}_1,\ldots,\hat{A}_n$ and the quantum state ρ. Their marginals along the ith coordinate coincide with the one induced by the spectral measure of $\hat{A}_i$ and ρ, and therefore give the correct probability law on the measurement outcomes of the observable $\hat{A}_i$. Our main result is that $p_{\mathrm{MH}_m}\to p_{\mathrm{W}}$ in $\mathcal{E}'(\mathbb{R}^n)$ as $m\to\infty$, and that the convergence can be upgraded to uniform convergence when the distributions are smeared with an appropriate Schwarz function.

A. Margenau-Hill quasiprobability distributions p_{MH_m}

We define the Margenau-Hill quasiprobability distribution p_{MH_m} of order $m\in\mathbb{N}$ associated with the observables $\hat{A}_1,\ldots,\hat{A}_n$ and the quantum state ρ to be

$$p_{\mathrm{MH}_m} = \frac{1}{(2\pi)^n}\mathcal{F}(f_{\mathrm{MH}_m}),$$

where f_{MH_m} is the quasicharacteristic function

$$f_{\mathrm{MH}_m}(\xi) = \frac{1}{n!}\,\mathrm{tr}\left(\rho\sum_{\pi\in S_n}\left(\prod_{k=1}^{n}e^{i\frac{\xi_{\pi(k)}}{m}\hat{A}_{\pi(k)}}\right)^m\right), \tag{3}$$

and S_n is the symmetric group. That is, S_n is the set of all permutations of n elements. The products inside the summation are ordered from left to right, starting from the lowest index $k=1$. That is,

$$\prod_{k=1}^{n}e^{i\frac{\xi_{\pi(k)}}{m}\hat{A}_{\pi(k)}} = e^{i\frac{\xi_{\pi(1)}}{m}\hat{A}_{\pi(1)}}e^{i\frac{\xi_{\pi(2)}}{m}\hat{A}_{\pi(2)}}\cdots e^{i\frac{\xi_{\pi(n)}}{m}\hat{A}_{\pi(n)}}.$$

The reason why the definition of f_{MH_m} includes a summation over all possible permutations $\pi\in S_n$ is to ensure that the distributions p_{MH_m} are real-valued for all $m\in\mathbb{N}$. For example, when $n=2$ and $m=1$, we recover the Margenau-Hill quasicharacteristic function

$$f_{\mathrm{MH}_m}(\xi_1,\xi_2) = \mathrm{tr}\left(\rho\frac{e^{i\xi_1\hat{A}_1}e^{i\xi_2\hat{A}_2}+e^{i\xi_2\hat{A}_2}e^{i\xi_1\hat{A}_1}}{2}\right),$$

studied in [1] for the spin-1/2 operators $\hat{A}_1=\hat{S}_1$ and $\hat{A}_2=\hat{S}_2$. In what follows, we state and prove properties of p_{MH_m}, starting with the most basic ones.

Proposition 1: The distributions p_{MH_m} are tempered and real-valued; i.e., for all $m\in\mathbb{N}$, $p_{\mathrm{MH}_m}\in\mathcal{S}'(\mathbb{R}^n)$ and

$$\langle\overline{p_{\mathrm{MH}_m}},\phi\rangle = \langle p_{\mathrm{MH}_m},\phi\rangle, \quad \forall\phi\in\mathcal{S}(\mathbb{R}^n).$$

Proof: The functions f_{MH_m} are continuous in $\xi\in\mathbb{R}^n$, and uniformly bounded since

$$\begin{aligned}|f_{\mathrm{MH}_m}(\xi)| &\leqslant \frac{1}{n!}\sum_{\pi\in S_n}\mathrm{tr}\left(\left|\rho\left(\prod_{k=1}^{n}e^{i\frac{\xi_{\pi(k)}}{m}\hat{A}_{\pi(k)}}\right)^m\right|\right),\\ &= \frac{1}{n!}\sum_{\pi\in S_n}\mathrm{tr}(|\rho|) = \mathrm{tr}(\rho) = 1, \quad \forall m\in\mathbb{N},\end{aligned}$$

where $|A| := \sqrt{AA^\dagger}$. Thus, $\forall m\in\mathbb{N}$, the maps

$$u_m:\phi(\xi)\mapsto\int_{\mathbb{R}^n}f_{\mathrm{MH}_m}(\xi)\phi(\xi)\mathrm{d}\xi, \quad \phi(\xi)\in\mathcal{S}(\mathbb{R}^n),$$

are continuous linear forms on $\mathcal{S}(\mathbb{R}^n)$. That is, the distributions u_m are tempered. Consequently, their images under the Fourier transform, and therefore

$$p_{\mathrm{MH}_m} = \frac{1}{(2\pi)^n}\mathcal{F}(u_m),$$

are also tempered. Finally, since

$$\overline{f_{\mathrm{MH}_m}(\xi)} = f_{\mathrm{MH}_m}(-\xi), \quad \forall m\in\mathbb{N},$$

we have

$$\begin{aligned}\langle\overline{p_{\mathrm{MH}_m}},\phi\rangle &:= \overline{\langle p_{\mathrm{MH}_m},\overline{\phi}\rangle} = \frac{1}{(2\pi)^n}\overline{\langle u_m,\mathcal{F}(\overline{\phi})\rangle}\\ &= \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}f_{\mathrm{MH}_m}(-\xi)\mathcal{F}(\phi)(-\xi)\mathrm{d}\xi\\ &= \langle p_{\mathrm{MH}_m},\phi\rangle,\end{aligned}$$

which completes the proof. ■

As done for the Wigner distribution p_{W} in Ref. [3], we demonstrate next how the Paley-Wiener-Schwarz theorem can be used to prove that the Margenau-Hill distributions p_{MH_m} are compactly supported for all $m\in\mathbb{N}$. To this end, define the set

$$\Lambda := \sigma(\hat{A}_1)\times\sigma(\hat{A}_2)\times\cdots\times\sigma(\hat{A}_n),$$

where $\sigma(\hat{A}_k)$ is the spectrum of $\hat{A}_k$, $k=1,\ldots,n$. The set Λ consists of all tuples of eigenvalues and is the classical support that is expected from a joint probability distribution for the simultaneous measurement outcomes of the observables $\hat{A}_1,\ldots,\hat{A}_n$. The convex hull of Λ, denoted by $\mathrm{conv}(\Lambda)$, is the *free numerical range* of the observables $\hat{A}_1,\ldots,\hat{A}_n$, which is the set of all vectors in $\mathbb{R}^n$ with the ith component being equal to $\mathrm{tr}(\rho_i\hat{A}_i)$ for some density matrix ρ_i. This set contains the *joint numerical range* R, which in turn contains the support of the Wigner distribution p_{W}.

Proposition 2: For all $m\in\mathbb{N}$, the distributions p_{MH_m} are compactly supported, i.e., $p_{\mathrm{MH}_m}\in\mathcal{E}'(\mathbb{R}^n)$, and moreover

$$\mathrm{supp}(p_{\mathrm{MH}_m})\subseteq\mathrm{conv}(\Lambda),$$

where $\mathrm{conv}(\Lambda)$ is the convex hull of Λ.

Proof: The set $K := \mathrm{conv}(\Lambda)$ is a convex solid in $\mathbb{R}^n$ with vertex set

$$\mathcal{V}(K) = \{v \in \mathbb{R}^n | v_k \in \{\lambda_{\min}(\hat{A}_k),\ \lambda_{\max}(\hat{A}_k)\},\ \forall k\},$$

where $\lambda_{\min}$ and $\lambda_{\max}$ denote the minimal and maximal eigenvalues. Thus, the supporting function of K is

$$H_K(x) = \max_{v\in\mathcal{V}(K)} \langle v, x\rangle.$$

By Proposition 1, $p_{\mathrm{MH}_m} \in \mathcal{S}'(\mathbb{R}^n)$ for all $m \in \mathbb{N}$. Then by Fourier's inversion formula, we have

$$\mathcal{F}(p_{\mathrm{MH}_m}) = \frac{1}{(2\pi)^n}\mathcal{F}\big(\mathcal{F}\big(f_{\mathrm{MH}_m}\big)\big) = g_{\mathrm{MH}_m},$$

where $g_{\mathrm{MH}_m}(\xi) = f_{\mathrm{MH}_m}(-\xi)$. Finally, observe that the function $g_{\mathrm{MH}_m}(z)$, $z \in \mathbb{C}^n$, is analytic everywhere in $\mathbb{C}^n$ and satisfies the estimate

$$\begin{aligned}\big|g_{\mathrm{MH}_m}(z)\big| &\leqslant \frac{1}{n!}\sum_{\pi\in S_n}\mathrm{tr}\left(\left|\rho\left(\prod_{k=1}^n e^{i\frac{-z_{\pi(k)}}{m}\hat{A}_{\pi(k)}}\right)^m\right|\right),\\ &\leqslant \frac{e^{\max_{v\in\mathcal{V}(K)}\langle v,\mathrm{Im}(z)\rangle}}{n!}\sum_{\pi\in S_n}\mathrm{tr}(\rho) = e^{H_K(\mathrm{Im}(z))},\end{aligned}$$

for all $m \in \mathbb{N}$. Detailed steps for the last inequality are given in Appendix A. By the Paley-Wiener-Schwarz theorem, this completes the proof. ■

The above is a powerful approach that can be extended straightforwardly to bounded self-adjoint operators on an infinite-dimensional Hilbert space. Therein, Λ need not be a finite set. In the current finite-dimensional setting, however, Λ is always a finite set. Because of this, Proposition 2 can be significantly refined, and the support can be characterized more precisely as follows.

Proposition 3: For all $m \in \mathbb{N}$, p_{MH_m} is a finite affine combination of Dirac delta distributions with

$$\mathrm{singsupp}\big(p_{\mathrm{MH}_m}\big) = \mathrm{supp}\big(p_{\mathrm{MH}_m}\big) \subseteq \frac{1}{m}\sum_{i=1}^m \Lambda,$$

where the summation is in the sense of Minkowski.

Proof: Let $f^{*m} := f * f * \cdots * f$ denote the convolution of f with itself m times. Then, explicit formal computation shows that, for all $m \in \mathbb{N}$, $x \in \mathbb{R}^n$,

$$\begin{aligned}p_{\mathrm{MH}_m}(x) &= \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} f_{\mathrm{MH}_m}(\xi)e^{-ix\cdot\xi}\mathrm{d}\xi\\ &= \frac{1}{n!(2\pi)^n}\sum_{\pi\in S_n}\mathrm{tr}\left(\rho\int_{\mathbb{R}^n}\left(\prod_{k=1}^n e^{i\frac{\xi_{\pi(k)}}{m}\hat{A}_{\pi(k)}}\right)^m e^{-ix\cdot\xi}\mathrm{d}\xi\right),\end{aligned}$$

where

$$\begin{aligned}\int_{\mathbb{R}^n}\left(\prod_{k=1}^n e^{i\frac{\xi_{\pi(k)}}{m}\hat{A}_{\pi(k)}}\right)^m e^{-ix\cdot\xi}\mathrm{d}\xi &= \left(\int_{\mathbb{R}^n}\left(\prod_{k=1}^n e^{i\frac{\xi_{\pi(k)}}{m}\hat{A}_{\pi(k)}}\right)e^{-ix\cdot\xi}\mathrm{d}\xi\right)^{*m}\\ &= \left(\prod_{k=1}^n\left(\int_{\mathbb{R}} e^{i\frac{\xi_{\pi(k)}}{m}\hat{A}_{\pi(k)}}e^{-ix_{\pi(k)}\xi_{\pi(k)}}\mathrm{d}\xi_{\pi(k)}\right)\right)^{*m}\\ &= \left(\prod_{k=1}^n\hat{U}_{\pi(k)}\left(\int_{\mathbb{R}} e^{-i\xi_{\pi(k)}(x_{\pi(k)}-\frac{\hat{D}_{\pi(k)}}{m})}\mathrm{d}\xi_{\pi(k)}\right)\hat{U}^\dagger_{\pi(k)}\right)^{*m} = \left(\prod_{k=1}^n\hat{U}_{\pi(k)}\hat{E}_{\pi(k)}\hat{U}^\dagger_{\pi(k)}\right)^{*m},\end{aligned}$$

and $\hat{U}_{\pi(k)}\hat{D}_{\pi(k)}\hat{U}^\dagger_{\pi(k)}$ is the eigendecomposition of $\hat{A}_{\pi(k)}$. The diagonal matrix $\hat{E}_{\pi(k)}$, given by

$$\hat{E}_{\pi(k)} = \int_{\mathbb{R}} e^{-i\xi_{\pi(k)}(x_{\pi(k)}-\frac{\hat{D}_{\pi(k)}}{m})}\mathrm{d}\xi_{\pi(k)},$$

has in its ith diagonal entry the Dirac delta distribution

$$2\pi\cdot\delta\left(x_{\pi(k)} - \frac{[\hat{D}_{\pi(k)}]_{ii}}{m}\right),$$

where $[\hat{D}_{\pi(k)}]_{ii}$ is the ith diagonal entry of $\hat{D}_{\pi(k)}$. This implies that

$$\begin{aligned}&\mathrm{singsupp}\left(\prod_{k=1}^n\hat{U}_{\pi(k)}\hat{E}_{\pi(k)}\hat{U}^\dagger_{\pi(k)}\right)\\ &\quad= \mathrm{supp}\left(\prod_{k=1}^n\hat{U}_{\pi(k)}\hat{E}_{\pi(k)}\hat{U}^\dagger_{\pi(k)}\right) \subseteq \Lambda/m.\end{aligned}$$

By performing the convolution m times, the resulting support for p_{MH_m} is contained in the set Λ/m added to itself m times. Thus, the distribution p_{MH_m} is a linear combination of Dirac delta distributions supported in $(1/m)\sum_{i=1}^m \Lambda$. To show that this linear combination is affine, i.e., that the coefficients of the combination add up to 1, it is enough to note that

$$\begin{aligned}\langle p_{\mathrm{MH}_m}, 1\rangle &= \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} f_{\mathrm{MH}_m}(\xi)\mathcal{F}(1)\mathrm{d}\xi\\ &= f_{\mathrm{MH}_m}(0) = 1,\end{aligned}$$

which completes the proof. ■

Thus, the Margenau-Hill distribution of any order is a discrete signed probability measure associated with the observables $\hat{A}_1, \ldots, \hat{A}_n$ and the quantum state ρ. Finally, we verify that the marginals of p_{MH_m} along any principal axis give the correct probability law on the measurement outcomes of the corresponding observable.

Proposition 4: The jth marginal of p_{MH_m} is given by

$$\int_{\mathbb{R}^{n-1}} p_{\mathrm{MH}_m}\mathrm{d}x_{\setminus j} = \sum_{s=1}^d \langle\psi_j(s)|\rho|\psi_j(s)\rangle\delta(x_j - \lambda_j(s)),$$

where $\mathrm{d}x_{\setminus j} := \mathrm{d}x_1 \cdots \mathrm{d}x_{j-1}\mathrm{d}x_{j+1} \cdots \mathrm{d}x_n$ and $\lambda_j(s)$, $|\psi_j(s)\rangle$ are the corresponding eigenvalues and eigenvectors of $\hat{A}_j$, respectively, with $s = 1, \ldots, d$.

Proof: Starting from the left-hand side of the above,

$$\begin{aligned}
&\frac{1}{(2\pi)^n}\int_{\mathbb{R}^{n-1}} \mathcal{F}\big(f_{\mathrm{MH}_m}\big)e^{i\sum_{k\neq j}^n x_k\xi_k}\mathrm{d}x_{\setminus j}\Big|_{\xi_k=0,\,\forall k\neq j}\\
&= \frac{1}{2\pi}\int_{\mathbb{R}} f_{\mathrm{MH}_m}(\xi)e^{-ix_j\xi_j}\mathrm{d}\xi_j\Big|_{\xi_k=0,\,\forall k\neq j}\\
&= \frac{1}{2\pi}\int_{\mathbb{R}} \mathrm{tr}(\rho e^{i\xi_j\hat{A}_j})e^{-ix_j\xi_j}\mathrm{d}\xi_j\\
&= \mathrm{tr}\bigg(\rho\hat{U}_j\bigg(\frac{1}{2\pi}\int_{\mathbb{R}} e^{-i\xi_j(x_j-\hat{D}_j)}\mathrm{d}\xi_j\bigg)\hat{U}_j^\dagger\bigg)\\
&= \sum_{s=1}^{d}\langle\psi_j(s)|\rho|\psi_j(s)\rangle\delta(x_j-\lambda_j(s)),
\end{aligned}$$

where $\hat{A}_j = \hat{U}_j\hat{D}_j\hat{U}_j^\dagger$ is the eigendecomposition of $\hat{A}_j$, and $\lambda_j(s)$, $|\psi_j(s)\rangle$ as claimed. ■

B. Convergence to the Wigner distribution p_{W}

Herein, we prove that $p_{\mathrm{MH}_m} \to p_{\mathrm{W}}$ in $\mathcal{E}'(\mathbb{R}^n)$ as $m \to \infty$, and that the convergence can be upgraded to uniform convergence when the distributions are smeared with an appropriate Schwarz function.

Proposition 5: The Margenau-Hill quasiprobability distribution p_{MH_m} converges to the Wigner distribution p_{W} in $\mathcal{E}'(\mathbb{R}^n)$ as $m \to \infty$.

Proof: By definition, we must show that

$$\lim_{m\to\infty}\langle p_{\mathrm{MH}_m},\phi\rangle = \langle p_{\mathrm{W}},\phi\rangle, \quad \forall\phi \in C^\infty(\mathbb{R}^n).$$

To that end, let $\chi \in C_0^\infty(\mathbb{R}^n)$ be a cutoff function equal to 1 on a neighborhood of $K = \mathrm{conv}(\Lambda)$ and let

$$\psi = \chi\phi \in C_0^\infty(\mathbb{R}^n),$$

for any $\phi \in C^\infty(\mathbb{R}^n)$; then

$$\begin{aligned}
\big\langle p_{\mathrm{MH}_m},\,\phi\big\rangle &= \langle p_{\mathrm{MH}_m},\,\psi\rangle\\
&= \frac{1}{(2\pi)^n}\langle\mathcal{F}(f_{\mathrm{MH}_m}),\,\psi\rangle\\
&= \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} f_{\mathrm{MH}_m}(\xi)\mathcal{F}(\psi)(\xi)\mathrm{d}\xi. \qquad (4)
\end{aligned}$$

By the Lie-Trotter product formula (1), we have the pointwise convergence

$$\lim_{m\to\infty} f_{\mathrm{MH}_m}(\xi) = \mathrm{tr}(\rho e^{i\xi\cdot A}) = f_{\mathrm{W}}(\xi).$$

Thus, the integrand in Eq. (4) converges pointwise to $f_{\mathrm{W}}(\xi)\mathcal{F}(\psi)(\xi)$ and is bounded in absolute value by $|\mathcal{F}(\psi)(\xi)| \in \mathcal{S}(\mathbb{R}^n)$ for all $m \in \mathbb{N}$. By the dominated convergence theorem we get

$$\begin{aligned}
&\lim_{m\to\infty}\langle p_{\mathrm{MH}_m},\phi\rangle\\
&= \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} f_{\mathrm{W}}(\xi)\mathcal{F}(\psi)(\xi)\mathrm{d}\xi\\
&= \frac{1}{(2\pi)^n}\langle\mathcal{F}(f_{\mathrm{W}}),\psi\rangle = \langle p_{\mathrm{W}},\psi\rangle = \langle p_{\mathrm{W}},\phi\rangle,
\end{aligned}$$

where the last equality follows from the fact that $\mathrm{supp}(p_{\mathrm{W}}) \subseteq R \subseteq K$. ■

Thus, we have shown that, while the general Wigner distribution p_{W} need not be a quasiprobability measure for the associated observables, it is not far from being one. More precisely, Proposition 5 implies that, for any $\epsilon > 0$ and $\phi \in C^\infty(\mathbb{R}^n)$, there will exist a discrete signed quasiprobability measure μ such that

$$\left|\langle p_{\mathrm{W}},\phi\rangle - \int_{\mathbb{R}^n}\phi\,\mathrm{d}\mu\right| \leqslant \epsilon,$$

with the measure μ being a Margenau-Hill quasiprobability distribution p_{MH_m} for sufficiently large m.

Finally, the convergence result in Proposition 5 can be upgraded to uniform convergence if the distributions are smeared, i.e., convolved, with an appropriate Schwarz function. To prove this, the following lemma is needed.

Lemma 1: The Margenau-Hill quasicharacteristic function f_{MH_m} converges to the Wigner quasicharacteristic function f_{W} as $m \to \infty$ uniformly on compact subsets of $\mathbb{C}^n$.

Proof: The proof is given in Appendix B. ■

Evidently, Lemma 1 refers to the analytic extensions of f_{MH_m} and f_{W} to all $\mathbb{C}^n$. In particular, the original functions f_{MH_m} converge to f_{W} as $m \to \infty$ uniformly on compact subsets of $\mathbb{R}^n$.

Proposition 6: If $\psi \in \mathcal{S}(\mathbb{R}^n)$ satisfies $\mathcal{F}(\psi) \in C_0^\infty(\mathbb{R}^n)$, then $\psi * p_{\mathrm{MH}_m} \to \psi * p_{\mathrm{W}}$ uniformly as $m \to \infty$.

Proof: Recall that since ψ is smooth and the distributions p_{MH_m} and p_{W} are tempered, then $\psi * p_{\mathrm{MH}_m}$ and $\psi * p_{\mathrm{W}}$ are also smooth. Let $\mathrm{supp}(\mathcal{F}(\psi)) \subseteq K$ compact; then

$$\begin{aligned}
&|\psi * p_{\mathrm{W}} - \psi * p_{\mathrm{MH}_m}|(x)\\
&= \left|\int_{\mathbb{R}^n}\mathcal{F}(\psi)(\xi)\big[f_{\mathrm{W}}(-\xi) - f_{\mathrm{MH}_m}(-\xi)\big]e^{-i\xi\cdot x}\,\mathrm{d}\xi\right|\\
&\leqslant \int_K |\mathcal{F}(\psi)(\xi)|\big|f_{\mathrm{W}}(-\xi) - f_{\mathrm{MH}_m}(-\xi)\big|\mathrm{d}\xi\\
&\leqslant \left(\int_K |\mathcal{F}(\psi)(\xi)|\mathrm{d}\xi\right)\cdot\sup_{\xi\in K}\big|f_{\mathrm{W}}(-\xi) - f_{\mathrm{MH}_m}(-\xi)\big|\\
&\leqslant \left(\int_K |\mathcal{F}(\psi)(\xi)|\mathrm{d}\xi\right)\cdot\sup_{\xi\in -K}\big|f_{\mathrm{W}}(\xi) - f_{\mathrm{MH}_m}(\xi)\big|, \qquad (5)
\end{aligned}$$

where the last integral is finite as $\mathcal{F}(\psi) \in \mathcal{S}(\mathbb{R}^n)$ and therefore integrable. By Lemma 1, $f_{\mathrm{MH}_m} \to f_{\mathrm{W}}$ as $m \to \infty$ uniformly on compact subsets of $\mathbb{R}^n$. Thus,

$$\sup_{\xi\in -K}\big|f_{\mathrm{W}}(\xi) - f_{\mathrm{MH}_m}(\xi)\big| \to 0,$$

as $m \to \infty$. Taking the supremum over all $x \in \mathbb{R}^n$ on both sides of Eq. (5) implies that

$$\sup_{x\in\mathbb{R}^n}\big|\psi * p_{\mathrm{W}} - \psi * p_{\mathrm{MH}_m}\big| \to 0,$$

as $m \to \infty$, which completes the proof. ■

We illustrate in Figs. 3–7 the convergence of the Margenau-Hill quasiprobability distributions p_{MH_m} to the corresponding Wigner distribution p_{W} for pairs of spin operators along the x and y directions and various quantum states. The figures display smoothed versions of the distributions obtained by convolving with a Gaussian function. This is effected by

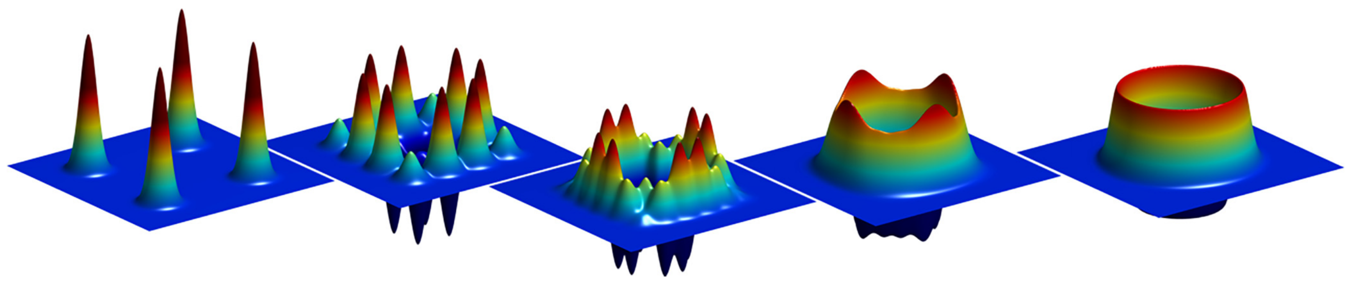

FIG. 3. Gaussian-regularized ($\varepsilon = 0.01$) bivariate quasiprobability distributions for the spin-1/2 observables along the x and y directions and the maximally mixed state $\rho = \hat{I}/2$. From left to right: p_{MH_1}, p_{MH_3}, p_{MH_5}, $p_{\mathrm{MH}_{10}}$, and p_{W}. Each p_{MH_m} consists of $(m+1)^2$ Dirac delta distributions supported over the grid of points $(1/m)\sum_{i=1}^{m} \Lambda$.

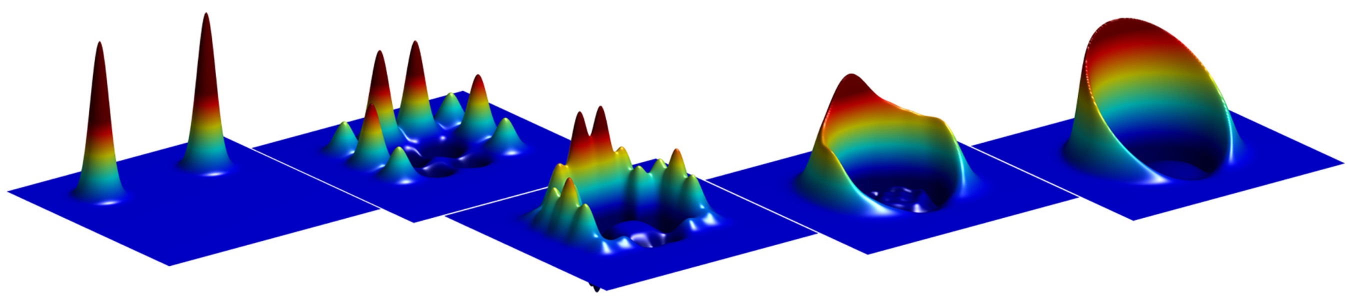

FIG. 4. Gaussian-regularized ($\varepsilon = 0.01$) bivariate quasiprobability distributions for the spin-1/2 observables along the x and y directions for the $-\hbar/2$ eigenstate in the x direction. From left to right: p_{MH_1}, p_{MH_3}, p_{MH_5}, $p_{\mathrm{MH}_{10}}$, and p_{W}. Each p_{MH_m} consists of $(m+1)^2$ Dirac delta distributions supported over the grid of points $(1/m)\sum_{i=1}^{m} \Lambda$.

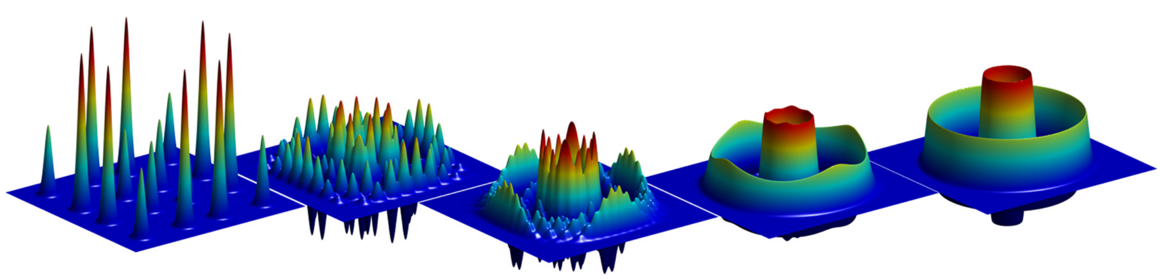

FIG. 5. Gaussian-regularized ($\varepsilon = 0.01$) bivariate quasiprobability distributions for the spin-3/2 observables along the x and y directions and the maximally mixed state $\rho = \hat{I}/4$. From left to right: p_{MH_1}, p_{MH_3}, p_{MH_5}, $p_{\mathrm{MH}_{15}}$, and p_{W}. Each p_{MH_m} consists of $(3m+1)^2$ Dirac delta distributions supported over the grid of points $(1/m)\sum_{i=1}^{m} \Lambda$.

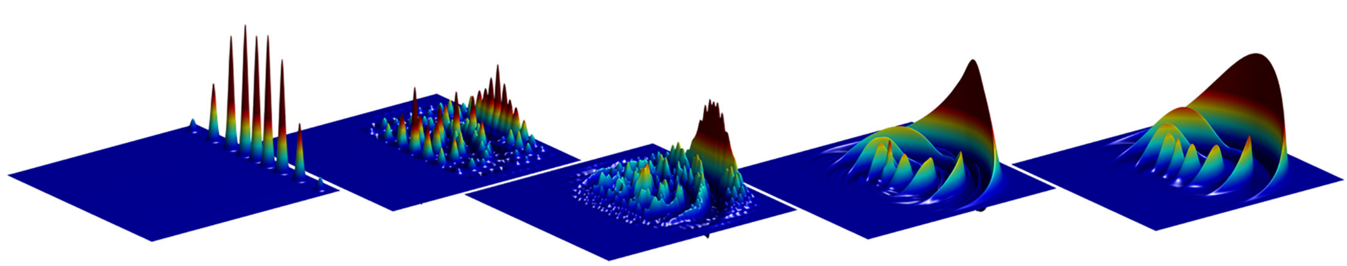

FIG. 6. Gaussian-regularized ($\varepsilon = 0.1$) bivariate quasiprobability distributions for the spin-4 observables along the x and y directions for the $+4\hbar$ eigenstate along the y direction. From left to right: p_{MH_1}, p_{MH_2}, p_{MH_3}, $p_{\mathrm{MH}_{10}}$, and p_{W}. Each p_{MH_m} consists of $(8m+1)^2$ Dirac delta distributions supported over the grid of points $(1/m)\sum_{i=1}^{m} \Lambda$.

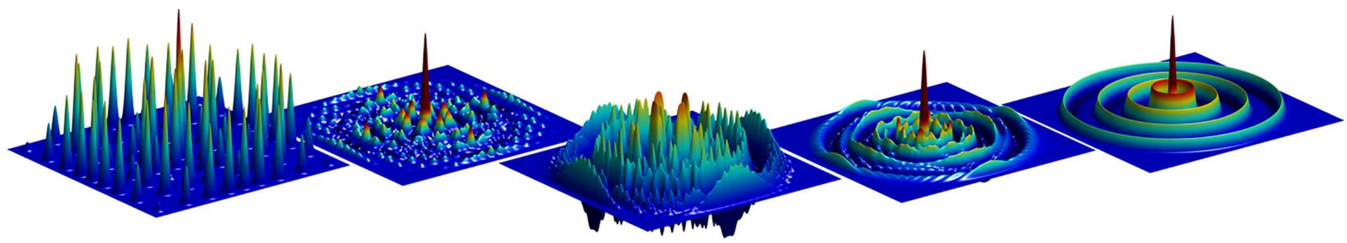

FIG. 7. Gaussian-regularized ($\varepsilon = 0.1$) bivariate quasiprobability distributions for the spin-4 observables along the x and y directions and the maximally mixed state $\rho = \hat{I}/9$. From left to right: p_{MH_1}, p_{MH_2}, p_{MH_3}, p_{MH_5}, and p_{W}. Each p_{MH_m} consists of $(8m+1)^2$ Dirac delta distributions supported over the grid of points $(1/m)\sum_{i=1}^{m}\Lambda$.

multiplying the quasicharacteristic function f by the decaying exponential $e^{-\varepsilon\xi^2}$, where the regularizing parameter ε is tuned for best visualization, as explained in Ref. [3]. Specifically, when ε is too large, the features of the distribution p are wiped out, and when ε is to small, singularities manifest with exceedingly large values. An example of the effect of the regularizing parameter ε on the resolution of the plots is shown in Fig. 8.

IV. SPIN 1/2

In this section, we specialize to quasiprobability distributions for a pair of spin-1/2 observables by setting

$$\hat{A}_1 = \hat{S}_1 := \hat{S}\cdot\hat{n}_1, \quad \hat{A}_2 = \hat{S}_2 := \hat{S}\cdot\hat{n}_2,$$

where $\hat{S}$ is the spin-1/2 operator and $\hat{n}_1$ and $\hat{n}_2$ are orthogonal directions in the Bloch sphere. The directions are given by the unit vectors

$$\hat{n}_i = \sin(\theta_i)\cos(\phi_i)\hat{x} + \sin(\theta_i)\sin(\phi_i)\hat{y} + \cos(\theta_i)\hat{z},$$

where $\hat{n}_1\cdot\hat{n}_2 = 0$, for some $\theta_i,\ \phi_i \in \mathbb{R},\ i \in \{1, 2\}$. First, we recall some properties of $\hat{S}_1$ and $\hat{S}_2$ as well as their Wigner distribution studied in Refs. [1,3]. Then we compute for all $m \in \mathbb{N}$ closed-form expressions for the Margenau-Hill quasicharacteristic functions f_{MH_m} and quasiprobability distributions p_{MH_m}. As a byproduct, we elucidate the nature of the convergence of f_{MH_m} to f_{W} as $m \to \infty$ by relating it to a special case of the Mehler-Heine theorem. For simplicity of exposition, we replace the spin-1/2 values $\pm\hbar/2$ by ± 1.

A. Properties of $\hat{S}_1$ and $\hat{S}_2$

The operators $\hat{S}_1$ and $\hat{S}_2$, represented in the $\hat{S}_z := \hat{S}\cdot\hat{z}$ eigenbasis, are given by

$$\hat{S}_j = \begin{bmatrix} \cos(\theta_j) & e^{-i\phi_j}\sin(\theta_j) \\ e^{i\phi_j}\sin(\theta_j) & -\cos(\theta_j) \end{bmatrix}, \quad j \in \{1, 2\},$$

and satisfy various properties summarized below.

Proposition 7: Let $\hat{I}$ denote the identity matrix. Then, for all $\xi_1,\ \xi_2 \in \mathbb{R},\ j, k \in \{1, 2\}$, the following properties hold:

$$\hat{S}_j^2 = \hat{I},\ \det(\hat{S}_j) = -1,\ \mathrm{tr}(\hat{S}_j) = 0,$$
$$[\hat{S}_j, \hat{S}_k] = 2i\hat{S}\cdot(\hat{n}_j \times \hat{n}_k), \quad \{\hat{S}_j, \hat{S}_k\} = 2(\hat{n}_j\cdot\hat{n}_k)\hat{I},$$
$$e^{i(\xi_1\hat{S}_1+\xi_2\hat{S}_2)} = \cos(\|\xi\|)\hat{I} + i(\xi_1\hat{S}_1 + \xi_2\hat{S}_2)\frac{\sin(\|\xi\|)}{\|\xi\|},$$

where $\|\xi\| = \sqrt{\xi_1^2 + \xi_2^2}$. The symbol $\times$ denotes the cross product, and $[\cdot,\cdot]$ and $\{\cdot,\cdot\}$ denote the commutator and anticommutator brackets, respectively.

Proof: The proof is given in Appendix C. ■

Next, recall that the quantum state ρ of a spin-1/2 particle can be represented in the $\hat{S}_z$ eigenbasis by

$$\rho = \frac{1}{2}\begin{bmatrix} 1+z & x-iy \\ x+iy & 1-z \end{bmatrix},$$

where $x := \mathrm{tr}(\rho\hat{S}_x)$, $y := \mathrm{tr}(\rho\hat{S}_y)$, and $z := \mathrm{tr}(\rho\hat{S}_z)$ are the Bloch vector coordinates satisfying $x^2 + y^2 + z^2 \leqslant 1$. Equivalently, one can consider any system of coordinates induced by the expectation values of three mutually orthogonal spin operators. In our case, one can take $\hat{S}_1$, $\hat{S}_2$, and $\hat{S}_3 := \hat{S}\cdot(\hat{n}_1 \times \hat{n}_2)$, so that

$$s_1^2 + s_2^2 + s_3^2 \leqslant 1, \tag{6}$$

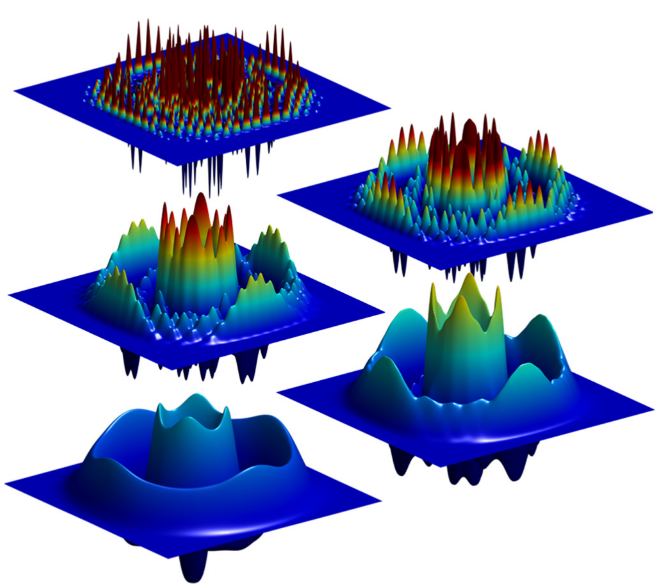

FIG. 8. Dependence of the resolution of the features of the Gaussian-regularized Margenau-Hill distribution p_{MH_5} on the regularizing parameter ε, for the spin-3/2 observables along the x and y directions for the maximally mixed state $\rho = \hat{I}/4$. From top to bottom: $\varepsilon = 0.001,\ 0.005,\ 0.01,\ 0.02,\ 0.05$. The distribution p_{MH_5} consists of 16^2 Dirac delta distributions supported over the grid of points $1/5\sum_{i=1}^{5}\Lambda$. They are best visualized for $\varepsilon = 0.001$.

where $s_1 := \mathrm{tr}(\rho\hat{S}_1)$, $s_2 := \mathrm{tr}(\rho\hat{S}_2)$, and $s_3 := \mathrm{tr}(\rho\hat{S}_3)$ are the expected values for the spin components of the particle along the directions $\hat{n}_1$, $\hat{n}_2$, and $\hat{n}_1 \times \hat{n}_2$, respectively.

B. Wigner distribution p_{W}

We recall some properties of the Wigner distribution p_{W} associated with the observables $\hat{S}_1$ and $\hat{S}_2$ and the quantum state ρ studied in Refs. [1,3].

Proposition 8: The Wigner distribution p_{W} associated with the observables $\hat{S}_1$ and $\hat{S}_2$ and the quantum state ρ is formally given by

$$p_{\mathrm{W}}(x_1, x_2) = (1 + s_1x_1 + s_2x_2)p^0_{\mathrm{W}}(x_1, x_2),$$

where

$$p^0_{\mathrm{W}}(x_1, x_2) = \frac{1}{(2\pi)^2}\mathcal{F}(\cos(\|\cdot\|))(x_1, x_2)$$

is the Wigner distribution for any state with Bloch vector normal to the plane defined by $\hat{n}_1$ and $\hat{n}_2$, and

$$\mathcal{F}(\cos(\|\cdot\|))(x_1, x_2) = \int_{\mathbb{R}^2}\cos(\|\xi\|)e^{-i(x_1\xi_1+x_2\xi_2)}\mathrm{d}\xi_1\mathrm{d}\xi_2.$$

Proof: By definition, the Wigner distribution p_{W} is given by

$$p_{\mathrm{W}}(x_1, x_2) = \frac{1}{(2\pi)^2}\int_{\mathbb{R}^2} f_{\mathrm{W}}(\xi_1, \xi_2)e^{-i(x_1\xi_1+x_2\xi_2)}\mathrm{d}\xi_1\mathrm{d}\xi_2,$$

where f_{W} is given by

$$\begin{aligned} f_{\mathrm{W}}(\xi_1, \xi_2) &= \mathrm{tr}(\rho e^{i(\xi_1\hat{S}_1+\xi_2\hat{S}_2)}) \\ &= \cos(\|\xi\|) + i(s_1\xi_1 + s_2\xi_2)\frac{\sin(\|\xi\|)}{\|\xi\|} \\ &= \left(1 - is_1\frac{\partial}{\partial\xi_1} - is_2\frac{\partial}{\partial\xi_2}\right)\cos(\|\xi\|). \end{aligned}$$

Performing integration by parts with respect to ξ_1 and ξ_2 yields

$$p_{\mathrm{W}}(x_1, x_2) = \frac{(1 + s_1x_1 + s_2x_2)}{(2\pi)^2}\mathcal{F}(\cos(\|\cdot\|))(x_1, x_2).$$

If a state ρ^0 corresponds to a Bloch vector normal to the plane defined by $\hat{n}_1$ and $\hat{n}_2$, then the expectations $s^0_1 = \mathrm{tr}(\rho^0\hat{S}_1)$ and $s^0_2 = \mathrm{tr}(\rho^0\hat{S}_2)$ are zero, so that

$$\begin{aligned} p^0_{\mathrm{W}}(x_1, x_2) &= \frac{1}{(2\pi)^2}\int_{\mathbb{R}^2}\cos(\|\xi\|)e^{-i(x_1\xi_1+x_2\xi_2)}\mathrm{d}\xi_1\mathrm{d}\xi_2 \\ &= \frac{1}{(2\pi)^2}\mathcal{F}(\cos(\|\cdot\|))(x_1, x_2), \end{aligned}$$

which yields the desired result. ■.

Next, using the Paley-Wiener-Schwarz theorem, we prove that p_{W} is always supported in the unit disk.

Proposition 9: The Wigner quasiprobability distribution p_{W} associated with the observables $\hat{S}_1$ and $\hat{S}_2$ and the quantum state ρ satisfies

$$\mathrm{supp}(p_{\mathrm{W}}) \subseteq \mathbb{D},$$

where $\mathbb{D} \subseteq \mathbb{R}^2$ is the closed unit disk.

Proof: Let $\|A\|_{\mathrm{F}} := \sqrt{\mathrm{tr}(AA^\dagger)}$ denote the Frobenius norm. By the Cauchy-Schwarz inequality, $|\mathrm{tr}(AB)| \leqslant \|A\|_{\mathrm{F}}\|B\|_{\mathrm{F}}$. Then, the function

$$f_{\mathrm{W}}(z_1, z_2) = \mathrm{tr}(\rho e^{i(z_1\hat{S}_1+z_2\hat{S}_2)}), \quad z := (z_1, z_2) \in \mathbb{C}^2,$$

which is analytic everywhere in $\mathbb{C}^2$, satisfies the estimate

$$\begin{aligned} |f_{\mathrm{W}}(z_1, z_2)| &\leqslant \|\rho\|_{\mathrm{F}}\|e^{i(z_1\hat{S}_1+z_2\hat{S}_2)}\|_{\mathrm{F}} \\ &= \|\rho\|_{\mathrm{F}}\|e^{-[\mathrm{Im}(z_1)\hat{S}_1+\mathrm{Im}(z_2)\hat{S}_2]}\|_{\mathrm{F}} \\ &= \|\rho\|_{\mathrm{F}}\left\|\sum_{k=0}^{\infty}\frac{(-1)^k[\mathrm{Im}(z_1)\hat{S}_1 + \mathrm{Im}(z_2)\hat{S}_2]^k}{k!}\right\|_{\mathrm{F}}. \end{aligned}$$

By noting that $[\mathrm{Im}(z_1)\hat{S}_1 + \mathrm{Im}(z_2)\hat{S}_2]^2 = \|\mathrm{Im}(z)\|_2^2\hat{I}$, we can split the above sum into even and odd parts to get

$$\left\|\sum_{k=0}^{\infty}\frac{(-1)^k[\mathrm{Im}(z_1)\hat{S}_1 + \mathrm{Im}(z_2)\hat{S}_2]^k}{k!}\right\|_{\mathrm{F}} \leqslant \sum_{k=0}^{\infty}\frac{\|\mathrm{Im}(z)\|_2^{2k}}{(2k)!}\|\hat{I}\|_{\mathrm{F}} + \|\mathrm{Im}(z_1)\hat{S}_1 + \mathrm{Im}(z_2)\hat{S}_2\|_{\mathrm{F}}\sum_{k=0}^{\infty}\frac{\|\mathrm{Im}(z)\|_2^{2k}}{(2k+1)!} \leqslant \sqrt{2}\sum_{k=0}^{\infty}\frac{\|\mathrm{Im}(z)\|_2^k}{k!}.$$

Thus, we obtain the estimate

$$\begin{aligned} |f_{\mathrm{W}}(z_1, z_2)| &\leqslant \sqrt{2}\|\rho\|_{\mathrm{F}}\sum_{k=0}^{\infty}\frac{\|\mathrm{Im}(z)\|_2^k}{k!} \\ &= \sqrt{2}\|\rho\|_{\mathrm{F}}e^{\|\mathrm{Im}(z)\|_2} = \sqrt{2}\|\rho\|_{\mathrm{F}}e^{H_{\mathbb{D}}(\mathrm{Im}(z))}. \end{aligned}$$

By the Paley-Wiener-Schwarz theorem, $\mathrm{supp}(p_{\mathrm{W}}) \subseteq \mathbb{D}$ as desired. ■

Remark: Since p^0_{W} is radially symmetric, the Wigner distribution p_{W} can be expressed in polar coordinates via the change of variables $x_1 = r\cos(\theta)$ and $x_2 = r\sin(\theta)$, where $r = \|x\|$, as

$$p_{\mathrm{W}}(r, \theta) = [1 + s_1r\cos(\theta) + s_2r\sin(\theta)]p^0_{\mathrm{W}}(r).$$

Thus, the Wigner distribution for any state ρ can always be written as the Wigner distribution p^0_{W} corresponding to, for instance, the maximally mixed state $\rho = I/2$, multiplied by the function

$$h(r, \theta) = 1 + s_1r\cos(\theta) + s_2r\sin(\theta).$$

Compare, for example, the two Wigner distributions for the observables $\hat{S}_1 = \hat{S}_x$ and $\hat{S}_2 = \hat{S}_y$ shown in the two right subplots of Fig. 1. Modulo the Gaussian regularization, the bottom Wigner distribution can be obtained by multiplying the top Wigner distribution by

$$h(r, \theta) = 1 + \frac{r}{\sqrt{2}}[\cos(\theta) + \sin(\theta)] = 1 + r\sin\left(\theta + \frac{\pi}{4}\right).$$

That is because the Wigner distribution for the state

$$\rho = \frac{1}{2}\begin{bmatrix} 1 & \frac{1-i}{\sqrt{2}} \\ \frac{1+i}{\sqrt{2}} & 1 \end{bmatrix},$$

corresponds to the Bloch vector coordinates $x = 1/\sqrt{2}$, $y = 1/\sqrt{2}$, and $z = 0$. Thus, $s_1 = s_2 = 1/\sqrt{2}$.

Lastly, notice that the function h is always non-negative in $\mathbb{D}$ where p_{W} is supported. To see this, apply the reverse triangle inequality, followed by the Cauchy-Schwarz inequality, to get

$$\begin{aligned} h(r,\theta) &\geqslant 1 - |s_1 r\cos(\theta) + s_2 r\sin(\theta)| \\ &\geqslant 1 - |r|\sqrt{s_1^2 + s_2^2} \geqslant 0, \end{aligned}$$

where the last inequality follows from Eq. (6). This implies that the sign of the Wigner distribution for $\hat{S}_1$ and $\hat{S}_2$, which reflects the quantum nature of the representation, is independent of the state ρ.

C. Margenau-Hill quasicharacteristic functions f_{MH_m} and the Mehler-Heine theorem

In what follows, we compute closed-form expressions for the Margenau-Hill quasicharacteristic functions f_{MH_m} and show that their convergence to f_{W}, which is uniform on compact subsets of $\mathbb{C}^2$ by Lemma 1, is a special case of the Mehler-Heine theorem. We start by proving the following result.

Lemma 2: Let $m \in \mathbb{N}$ and $\xi_1, \xi_2 \in \mathbb{R}$. Then,

$$\frac{\left(e^{i\frac{\xi_1}{m}\hat{S}_1}e^{i\frac{\xi_2}{m}\hat{S}_2}\right)^m + \left(e^{i\frac{\xi_2}{m}\hat{S}_2}e^{i\frac{\xi_1}{m}\hat{S}_1}\right)^m}{2} = T_m(a_m) + ib_m U_{m-1}(a_m), \tag{7}$$

and

$$\frac{\left(e^{i\frac{\xi_1}{m}\hat{S}_1}e^{i\frac{\xi_2}{m}\hat{S}_2}\right)^m - \left(e^{i\frac{\xi_2}{m}\hat{S}_2}e^{i\frac{\xi_1}{m}\hat{S}_1}\right)^m}{2} = U_{m-1}(a_m)\sin\left(\frac{\xi_1}{m}\right)\sin\left(\frac{\xi_2}{m}\right)\hat{S}_2\hat{S}_1, \tag{8}$$

where

$$\begin{aligned} a_m &:= \cos\left(\frac{\xi_1}{m}\right)\cos\left(\frac{\xi_2}{m}\right)\hat{I}, \\ b_m &:= \sin\left(\frac{\xi_1}{m}\right)\cos\left(\frac{\xi_2}{m}\right)\hat{S}_1 + \cos\left(\frac{\xi_1}{m}\right)\sin\left(\frac{\xi_2}{m}\right)\hat{S}_2, \end{aligned}$$

and $T_m(\cdot)$ and $U_m(\cdot)$ are the mth degree Chebyshev polynomials of the first and second kinds, respectively.

Proof: The proof is given in Appendix D. ■

The resemblance of Eq. (7) with the mth power

$$z^m = T_m(a) + ibU_{m-1}(a),$$

of a complex number $z = a + ib$ of modulus 1, $a, b \in \mathbb{R}$, is inescapable. Then, Eqs. (7) and (8) together may be seen as a noncommutative version of de Moivre's formula. Using the above lemma, we can derive closed-form expressions for the Margenau-Hill quasicharacteristic functions f_{MH_m} for all $m \in \mathbb{N}$, as explained next.

Proposition 10: The Margenau-Hill quasicharacteristic functions $f_{\mathrm{MH}_m}(\xi_1, \xi_2)$ associated with the spin operators $\hat{S}_1$, $\hat{S}_2$ and the state ρ are given by

$$\left(1 - is_1\frac{\partial}{\partial\xi_1} - is_2\frac{\partial}{\partial\xi_2}\right)T_m\left(\cos\left(\frac{\xi_1}{m}\right)\cos\left(\frac{\xi_2}{m}\right)\right),$$

where $T_m(\cdot)$ is the mth degree Chebyshev polynomial of the first kind, $m \in \mathbb{N}$.

Proof: Equation (7) implies that for all $m \in \mathbb{N}$,

$$f_{\mathrm{MH}_m}(\xi_1, \xi_2) = \mathrm{tr}[\rho(T_m(a_m) + ib_m U_{m-1}(a_m))].$$

Since $\frac{\mathrm{d}}{\mathrm{d}x}T_m(x) = mU_{m-1}(x)$, the right-hand side equals

$$\begin{aligned} &\mathrm{tr}\left[\rho\left(\hat{I} - i\hat{S}_1\frac{\partial}{\partial\xi_1} - i\hat{S}_2\frac{\partial}{\partial\xi_2}\right)T_m\left(\cos\left(\frac{\xi_1}{m}\right)\cos\left(\frac{\xi_2}{m}\right)\right)\right] \\ &= \left(1 - is_1\frac{\partial}{\partial\xi_1} - is_2\frac{\partial}{\partial\xi_2}\right)T_m\left(\cos\left(\frac{\xi_1}{m}\right)\cos\left(\frac{\xi_2}{m}\right)\right), \end{aligned}$$

which completes the proof. ■

Remark: So far, we have seen that both the Wigner quasicharacteristic function f_{W} as well as the Margenau-Hill quasicharacteristic functions f_{MH_m} can be derived by applying the operator

$$1 - is_1\frac{\partial}{\partial\xi_1} - is_2\frac{\partial}{\partial\xi_2}$$

to the quasicharacteristic function corresponding to any state ρ_0 satisfying $\mathrm{tr}(\rho^0\hat{S}_1) = \mathrm{tr}(\rho^0\hat{S}_2) = 0$. Thus, knowledge of the quasicharacteristic function for any such ρ_0 is enough to construct the quasicharacteristic function for any other state ρ.

In light of Proposition 10, we conclude by noting the connection between the Mehler-Heine theorem and the convergence property of the Margenau-Hill quasicharacteristic functions f_{MH_m} to f_{W}.

Proposition 11: For the special case of spin observables $\hat{S}_1$ and $\hat{S}_2$, Lemma 1 follows directly from the Mehler-Heine theorem.

Proof: The proof is given in Appendix E. ■

It is unclear whether this connection between the Lie-Trotter product formula and the Mehler-Heine theorem for spin-1/2 observables is simply a mathematical coincidence or the manifestation of a deeper fact.

D. Margenau-Hill quasiprobability distributions p_{MH_m}

We now derive closed-form expressions for the Margenau-Hill quasiprobability distributions p_{MH_m} for all $m \in \mathbb{N}$. This is done by taking the Fourier transform of the corresponding expressions for f_{MH_m} in Proposition 10.

Proposition 12: The Margenau-Hill quasiprobability distribution of order $m \in \mathbb{N}$ associated with the spin observables $\hat{S}_1$, $\hat{S}_2$ and a quantum state ρ is given by

$$p_{\mathrm{MH}_m} = (1 + s_1x_1 + s_2x_2)p^0_{\mathrm{MH}_m},$$

where $s_i = \mathrm{tr}(\rho\hat{S}_i)$, $i \in \{1, 2\}$,

$$p^0_{\mathrm{MH}_m} = \sum_{n=0}^{m}\frac{a_{mn}}{4^n}\left[\sum_{p,q=0}^{n}\binom{n}{p}\binom{n}{q}\delta_{\left(x_1 - \frac{n-2p}{m}, x_2 - \frac{n-2q}{m}\right)}\right],$$

and a_{mn} are the coefficients of the Chebyshev polynomial of the first kind of order m, that is,

$$T_m(x) = \sum_{n=0}^{m} a_{mn}x^n.$$

Proof: By Proposition 10, the Margenau-Hill quasicharacteristic function $f_{\mathrm{MH}_m}(\xi_1, \xi_2)$ is given by

$$\left(1 - is_1\frac{\partial}{\partial\xi_1} - is_2\frac{\partial}{\partial\xi_2}\right)T_m\left(\cos\left(\frac{\xi_1}{m}\right)\cos\left(\frac{\xi_2}{m}\right)\right),$$

for all $m \in \mathbb{N}$. It follows that

$$p_{\mathrm{MH}_m} = \frac{1}{(2\pi)^2}\mathcal{F}(f_{\mathrm{MH}_m}) = (1 + s_1x_1 + s_2x_2)p^0_{\mathrm{MH}_m},$$

where

$$\begin{aligned} p^0_{\mathrm{MH}_m}(x_1, x_2) &= \sum_{n=0}^{m}\frac{a_{mn}}{(2\pi)^2}\mathcal{F}\left(\cos^n\left(\frac{x_1}{m}\right)\cos^n\left(\frac{x_2}{m}\right)\right),\\ &= \sum_{n=0}^{m}\frac{a_{mn}}{4^n}\left[\sum_{p,q=0}^{n}\binom{n}{p}\binom{n}{q}\delta_{\left(x_1-\frac{n-2p}{m},\, x_2-\frac{n-2q}{m}\right)}\right]. \end{aligned}$$

The last equality follows from the fact that

$$\mathcal{F}\left(\cos^n\left(\frac{\cdot}{m}\right)\right) = \frac{2\pi}{2^n}\left[\sum_{p=0}^{n}\binom{n}{p}\delta\left(\cdot - \frac{n-2p}{m}\right)\right],$$

which completes the proof. ∎

V. REPEATED EXPERIMENTS

In this section, we discuss the Wigner distribution p_{W} and its particle approximations p_{MH_m} in the context of repeated experiments.

Recall from Proposition 3 that the Margenau-Hill quasiprobability distribution p_{MH_m} of order m is supported on the grid of points

$$\frac{1}{m}\sum_{i=1}^{m}\Lambda = \frac{\Lambda + \Lambda + \cdots + \Lambda}{m},$$

where the summation is in the sense of Minkowski, and

$$\Lambda := \sigma(\hat{A}_1) \times \sigma(\hat{A}_2) \times \cdots \times \sigma(\hat{A}_n),$$

with $\sigma(\hat{A}_k)$ the spectrum of $\hat{A}_k$, $k = 1, \ldots, n$. The set Λ consists of all tuples of eigenvalues and is the support that is expected of a joint probability law on the classical measurement outcomes of the observables $\hat{A}_1, \ldots, \hat{A}_n$. For instance, in the case of spin-1/2 observables $\hat{S}_1$ and $\hat{S}_2$ discussed in Sec. IV, the set Λ is given by

$$\begin{aligned} \Lambda &= (-1, +1) \times (-1, +1),\\ &= \{(-1, -1), (-1, +1), (+1, -1), (+1, +1)\},\\ &= \{(\pm1, \pm1), (\pm1, \mp1)\}, \end{aligned}$$

which is the support for the Margenau-Hill quasiprobability distribution p_{MH_m} of order $m = 1$.

Since the probability measures p_{MH_m} are sign-indefinite, the underlying experiments are not realizable, and will hence be referred to as *thought experiments*. For instance, the thought experiment associated with the signed measure p_{MH_1} is the *simultaneous* measurement of the spin components of a spin-1/2 particle in state ρ along the directions $\hat{n}_1$ and $\hat{n}_2$; see Ref. [8].

Likewise, the support for p_{MH_2}, which is

$$\begin{aligned} \frac{\Lambda + \Lambda}{2} &= \frac{[(-1, +1) \times (-1, +1)] + [(-1, +1) \times (-1, +1)]}{2},\\ &= \{(\pm1, \pm1), (\pm1, \mp1), (\pm1, 0), (0, \pm1), (0, 0)\}, \end{aligned}$$

suggests that the associated thought experiment for the case $m = 2$ would be the *instantaneous repetition of the thought experiment in* $m = 1$ *twice*, for the same ρ, with the average of the two results recorded as the outcome. The simultaneous measurements performed in these thought experiments are not quantum measurements because $\hat{S}_1$ and $\hat{S}_2$ do not commute. As a consequence, the axiom of repetition need not apply, and nine possible outcomes are present. For instance, the outcome $(0, +1)$ may arise as a result of $(+1, +1)$ instantiating in the first simultaneous measurement and $(-1, +1)$ in the second.

In a similar manner, the support of p_{MH_N} for any $N \in \mathbb{N}$ becomes the set of outcomes obtained by *instantaneously repeating the thought experiment for* $m = 1$*,* N *times, and recording the average.* Thus, the Wigner quasiprobability distribution p_{W} corresponds to this limiting thought experiment that involves an *infinite* instantaneous repetition of the thought experiment in $m = 1$.

Since this limit process averages out the results from theoretically sampling the system in state ρ infinitely many times, the support of the Wigner distribution is still confined in the unit square. See, for instance, Fig. 9, which displays the evolution of the support of the Margenau-Hill distributions p_{MH_N} for $\hat{S}_1$ and $\hat{S}_2$ as $N \to \infty$. The fact, however, that the support is always inside the *unit Disk*, which is the joint numerical range of $\hat{S}_1$ and $\hat{S}_2$, suggests that with every run of this thought experiment, the resulting average must correspond to the spin components of the particle in some state σ, as

$$\mathrm{tr}(\sigma\hat{S}_1)^2 + \mathrm{tr}(\sigma\hat{S}_2)^2 \leqslant 1.$$

We believe that a time-resolved version of this thought experiment could be linked to the continuous monitoring of non-commuting observables [9]. Such an interpretation could provide insight on certain features of p_{W}, such as its regions of positivity as well as its shape, from a physical perspective. The above discussion is no different for a general tuple of operators $\hat{A}_1, \ldots, \hat{A}_n$, and potential links to the theory of continuous measurement are of great interest.

Finally, when all of the observables $\hat{A}_1, \ldots, \hat{A}_n$ commute, all successive supports collapse to that of p_{MH_1} and the sequence of thought experiments, which are now realizable, must lead to outcomes confined to the classical sample space Λ. In the context of our proposed thought experiments, this implies that making multiple repetitions of the same measurement does not alter the average when the observables commute; i.e., subsequent measurements are identical to the outcomes obtained in the first measurement. This is consistent with the axiom of repetition, which asserts that performing the same measurement on a quantum system will not change the resulting outcome.

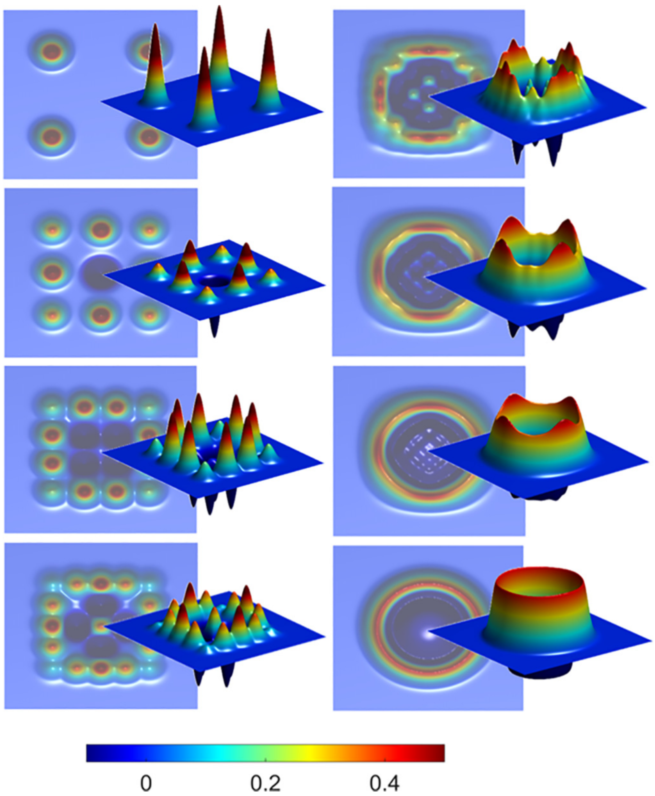

FIG. 9. Evolution of the support of the Margenau-Hill distribution p_{MH_m} for the spin-1/2 observables along the x and y directions and the maximally mixed state $\rho = \hat{I}/2$. From top to bottom: (left) p_{MH_1}, p_{MH_2}, p_{MH_3}, p_{MH_4}, (right) p_{MH_6}, p_{MH_8}, $p_{\mathrm{MH}_{11}}$, p_{W}. The distribution p_{MH_m} consists of $(m+1)^2$ Dirac delta distributions supported over the grid of points $(1/m)\sum_{i=1}^{m} \Lambda$.

VI. CONCLUSION

In this work, a class of real-valued signed discrete probability measures given by

$$p_{\mathrm{MH}_m} = \frac{1}{(2\pi)^n}\mathcal{F}(f_{\mathrm{MH}_m}),$$

for n arbitrary quantum observables is derived and studied based on quasicharacteristic functions f_{MH_m} with symmetrized operator orderings of Margenau-Hill type. These measures are given by affine combinations of Dirac delta distributions supported over the finite spectral range of the quantum observables, and give the correct probability marginals when coarse-grained along any principal axis. We showed that these particle approximations converge weakly to their corresponding Wigner distribution, and the convergence can be upgraded if they are smeared with an appropriate Schwarz function. Closed-form expressions in the case of bivariate quasiprobability distributions for the spin measurements of spin-1/2 particles are provided. As a side result, the convergence of the approximants in this case follows from the Mehler-Heine theorem. Finally, we discussed the Wigner distribution and its particle approximations in the context of repeated thought experiments. Namely, the supports of p_{W} and p_{MH_m} point towards thought experiments involving repeated simultaneous measurements on the state ρ. When the operators mutually commute, these supports reduce to the classical grid of eigenvalues, in agreement with the axiom of repetition.

ACKNOWLEDGMENTS

This research has been supported in part by the NSF under Grant No. ECCS-2347357, AFOSR under Grant No. FA9550-24-1-0278, and ARO under Grant No. W911NF-22-1-0292.

APPENDIX A: PROOF OF PROPOSITION 2

We provide herein details for the inequality

$$\sum_{\pi\in S_n} \mathrm{tr}\left(\left|\rho\left(\prod_{k=1}^{n} e^{i\frac{-z_{\pi(k)}}{m}\hat{A}_{\pi(k)}}\right)^{m}\right|\right) \leqslant e^{\max_{v\in\mathcal{V}(K)}\langle v,\mathrm{Im}(z)\rangle}\sum_{\pi\in S_n}\mathrm{tr}(\rho).$$

By definition,

$$\left|\rho\left(\prod_{k=1}^{n} e^{i\frac{-z_{\pi(k)}}{m}\hat{A}_{\pi(k)}}\right)^{m}\right| = \sqrt{\rho\left(\prod_{k=1}^{n} e^{i\frac{-z_{\pi(k)}}{m}\hat{A}_{\pi(k)}}\right)^{m}\left(\prod_{k=1}^{n} e^{i\frac{\bar{z}_{\pi(n-k+1)}}{m}\hat{A}_{\pi(n-k+1)}}\right)^{m}\rho}.$$

Notice that the innermost pair of factors simplify and are bounded as follows:

$$\begin{aligned}\left(\prod_{k=1}^{n} e^{i\frac{-z_{\pi(k)}}{m}\hat{A}_{\pi(k)}}\right)\left(\prod_{k=1}^{n} e^{i\frac{\bar{z}_{\pi(n-k+1)}}{m}\hat{A}_{\pi(n-k+1)}}\right) &= \left(\prod_{k=1}^{n-1} e^{i\frac{-z_{\pi(k)}}{m}\hat{A}_{\pi(k)}}\right)e^{-2\frac{\mathrm{Im}(z_{\pi(n)})}{m}\hat{A}_{\pi(n)}}\left(\prod_{k=2}^{n} e^{i\frac{\bar{z}_{\pi(n-k+1)}}{m}\hat{A}_{\pi(n-k+1)}}\right)\\ &\leqslant \left(\prod_{k=1}^{n-1} e^{i\frac{-z_{\pi(k)}}{m}\hat{A}_{\pi(k)}}\right)\left(\prod_{k=2}^{n} e^{i\frac{\bar{z}_{\pi(n-k+1)}}{m}\hat{A}_{\pi(n-k+1)}}\right)e^{2\frac{v_n\,\mathrm{Im}(z_{\pi(n)})}{m}}\hat{I},\end{aligned}$$

where the last inequality follows because the positive operator $e^{-2\frac{\mathrm{Im}(z_{\pi(n)})}{m}\hat{A}_{\pi(n)}}$ is less than $e^{2\frac{v_n\,\mathrm{Im}(z_{\pi(n)})}{m}}\hat{I}$ for some $v_n \in \{\lambda_{\min}(\hat{A}_{\pi(n)}),\ \lambda_{\max}(\hat{A}_{\pi(n)})\}$, and applying a congruence transformation will not change this fact; $e^{2\frac{v_n\,\mathrm{Im}(z_{\pi(n)})}{m}}\hat{I}$ is a scalar multiple of the identity and can be moved to the right. Repeating this process for the subsequent pairs of innermost factors, we obtain the final estimate

$$\left(\prod_{k=1}^{n} e^{i\frac{-z_{\pi(k)}}{m}\hat{A}_{\pi(k)}}\right)\left(\prod_{k=1}^{n} e^{i\frac{\bar{z}_{\pi(n-k+1)}}{m}\hat{A}_{\pi(n-k+1)}}\right) \leqslant e^{2\frac{\max_{v\in\mathcal{V}(K)}\langle v,\mathrm{Im}(z)\rangle}{m}}\hat{I}.$$

Thus,

$$\left(\prod_{k=1}^{n} e^{i\frac{-z_{\pi(k)}}{m}\hat{A}_{\pi(k)}}\right)^{m}\left(\prod_{k=1}^{n} e^{i\frac{\bar{z}_{\pi(n-k+1)}}{m}\hat{A}_{\pi(n-k+1)}}\right)^{m} \leqslant e^{2\max_{v\in\mathcal{V}(K)}\langle v,\mathrm{Im}(z)\rangle}\hat{I}.$$

Multiplying from the left and right by ρ and recalling that the $\sqrt{\cdot}$ function is operator monotone yields

$$\left|\rho\left(\prod_{k=1}^{n} e^{i\frac{-z_{\pi(k)}}{m}\hat{A}_{\pi(k)}}\right)^{m}\right| \leqslant \sqrt{\rho e^{2\max_{v\in\mathcal{V}(K)}\langle v,\mathrm{Im}(z)\rangle}\hat{I}\rho} = e^{\max_{v\in\mathcal{V}(K)}\langle v,\mathrm{Im}(z)\rangle}\rho.$$

Taking the trace on both sides and summing over all $\pi \in S_n$ yields the desired inequality.

APPENDIX B: PROOF OF LEMMA 1

Let K be any compact set in $\mathbb{C}^n$ and let $\|\cdot\|_{\mathrm{F}}$ denote the Frobenius norm

$$\|A\|_{\mathrm{F}} := \sqrt{\mathrm{tr}(AA^{\dagger})}, \quad A \in \mathbb{C}^{d\times d}.$$

Recall that $|\mathrm{tr}(AB)| \leqslant \|A\|_{\mathrm{F}}\|B\|_{\mathrm{F}}$ and that the functions f_{W} and f_{MH_m} can be analytically extended to $\mathbb{C}^n$ for all $m \in \mathbb{N}$. Then,

$$\begin{aligned}\sup_{\xi\in K}|f_{\mathrm{W}}(\xi) - f_{\mathrm{MH}_m}(\xi)| &= \sup_{\xi\in K}\left|\mathrm{tr}\left(\rho e^{i\xi\cdot\hat{A}}\right) - \frac{1}{n!}\mathrm{tr}\left(\rho\sum_{\pi\in S_n}\left(\prod_{k=1}^{n} e^{i\frac{\xi_{\pi(k)}}{m}\hat{A}_{\pi(k)}}\right)^{m}\right)\right| \\ &\leqslant \frac{1}{n!}\sum_{\pi\in S_n}\sup_{\xi\in K}\left|\mathrm{tr}\left[\rho\left(e^{i\xi\cdot\hat{A}} - \left(\prod_{k=1}^{n} e^{i\frac{\xi_{\pi(k)}}{m}\hat{A}_{\pi(k)}}\right)^{m}\right)\right]\right| \\ &\leqslant \frac{\|\rho\|_{\mathrm{F}}}{n!}\sum_{\pi\in S_n}\sup_{\xi\in K}\left\|e^{i\xi\cdot\hat{A}} - \left(\prod_{k=1}^{n} e^{i\frac{\xi_{\pi(k)}}{m}\hat{A}_{\pi(k)}}\right)^{m}\right\|_{\mathrm{F}}.\end{aligned}$$

Thus, it is enough to show that

$$\sup_{\xi\in K}\left\|e^{i\xi\cdot\hat{A}} - \left(\prod_{k=1}^{n} e^{i\frac{\xi_{\pi(k)}}{m}\hat{A}_{\pi(k)}}\right)^{m}\right\|_{\mathrm{F}} \to 0 \text{ as } m\to\infty,$$

for all permutations $\pi \in S_n$. By symmetry, it is sufficient to consider the trivial permutation only. To that end, define

$$C := e^{i\frac{\xi}{m}\cdot\hat{A}} \text{ and } D := \prod_{k=1}^{n} e^{i\frac{\xi_k}{m}\hat{A}_k}.$$

Then, applying the Cauchy product formula on D yields

$$D = \prod_{k=1}^{n}\left(\sum_{j_k=0}^{\infty}\frac{(i\frac{\xi_k}{m}\hat{A}_k)^{j_k}}{j_k!}\right) = \sum_{|j|=0}^{\infty}\left(\prod_{k=1}^{n}\frac{1}{m^{j_k}}\frac{(i\xi_k\hat{A}_k)^{j_k}}{j_k!}\right) = \sum_{|j|=0}^{\infty}\frac{1}{m^{|j|}}\left(\prod_{k=1}^{n}\frac{(i\xi_k\hat{A}_k)^{j_k}}{j_k!}\right),$$

where $|j| := j_1 + \cdots + j_n$. Thus, we get

$$\begin{aligned}\|C-D\|_{\mathrm{F}} &= \left\|\sum_{|j|=0}^{\infty}\frac{1}{m^{|j|}}\frac{(i\xi\cdot\hat{A})^{|j|}}{|j|!} - \sum_{|j|=0}^{\infty}\frac{1}{m^{|j|}}\left(\prod_{k=1}^{n}\frac{(i\xi_k\hat{A}_k)^{j_k}}{j_k!}\right)\right\|_{\mathrm{F}} = \left\|\sum_{|j|=2}^{\infty}\frac{1}{m^{|j|}}\left(\frac{(i\xi\cdot\hat{A})^{|j|}}{|j|!} - \prod_{k=1}^{n}\frac{(i\xi_k\hat{A}_k)^{j_k}}{j_k!}\right)\right\|_{\mathrm{F}} \\ &\leqslant \frac{1}{m^2}\sum_{|j|=0}^{\infty}\left\|\frac{(i\xi\cdot\hat{A})^{|j|}}{|j|!} - \prod_{k=1}^{n}\frac{(i\xi_k\hat{A}_k)^{j_k}}{j_k!}\right\|_{\mathrm{F}} \leqslant \frac{1}{m^2}\sum_{|j|=0}^{\infty}\left(\frac{(|\xi|\cdot\|\hat{A}\|_{\mathrm{F}})^{|j|}}{|j|!} + \prod_{k=1}^{n}\frac{(|\xi_k|\|\hat{A}_k\|_{\mathrm{F}})^{j_k}}{j_k!}\right) \\ &= \frac{1}{m^2}\left(e^{|\xi|\cdot\|\hat{A}\|_{\mathrm{F}}} + \sum_{|j|=0}^{\infty}\left(\prod_{k=1}^{n}\frac{(|\xi_k|\|\hat{A}_k\|_{\mathrm{F}})^{j_k}}{j_k!}\right)\right) = \frac{2}{m^2}e^{|\xi|\cdot\|\hat{A}\|_{\mathrm{F}}},\end{aligned}$$

where $|\xi| \cdot \|\hat{A}\|_{\mathrm{F}} := |\xi_1|\|\hat{A}_1\|_{\mathrm{F}} + \cdots + |\xi_n|\|\hat{A}_n\|_{\mathrm{F}}$. By noting that $\|C\|_{\mathrm{F}}, \|D\|_{\mathrm{F}} \leqslant 1$, we get

$$\|C^m - D^m\|_{\mathrm{F}} = \left\| \sum_{k=0}^{m-1} C^k (C-D) D^{m-k-1} \right\|_{\mathrm{F}} \leqslant m\|C-D\|_{\mathrm{F}} \leqslant \frac{2}{m} e^{|\xi| \cdot \|\hat{A}\|_{\mathrm{F}}}.$$

Taking the supremum of both sides over $\xi \in K$ and letting $m \to \infty$ yields the desired result. ■

APPENDIX C: PROOF OF PROPOSITION 6

(1) Proof that $\hat{S}_j^2 = \hat{I}$:

$$\hat{S}_j^2 = \begin{bmatrix} \cos(\theta_j) & e^{-i\phi_j}\sin(\theta_j) \\ e^{i\phi_j}\sin(\theta_j) & -\cos(\theta_j) \end{bmatrix} \begin{bmatrix} \cos(\theta_j) & e^{-i\phi_j}\sin(\theta_j) \\ e^{i\phi_j}\sin(\theta_j) & -\cos(\theta_j) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \hat{I}.$$

(2) Proof that $\det(\hat{S}_j) = -1$:

$$\det(\hat{S}_j) = \begin{vmatrix} \cos(\theta_j) & e^{-i\phi_j}\sin(\theta_j) \\ e^{i\phi_j}\sin(\theta_j) & -\cos(\theta_j) \end{vmatrix} = -\cos^2(\theta_j) - \sin^2(\theta_j) = -1.$$

(3) Proof that $\mathrm{tr}(\hat{S}_j) = 0$:

$$\mathrm{tr}(\hat{S}_j) = \mathrm{tr}\left(\begin{bmatrix} \cos(\theta_j) & e^{-i\phi_j}\sin(\theta_j) \\ e^{i\phi_j}\sin(\theta_j) & -\cos(\theta_j) \end{bmatrix} \right) = \cos(\theta_j) - \cos(\theta_j) = 0.$$

(4) Proof that $[\hat{S}_j, \hat{S}_k] = 2i\hat{S} \cdot (\hat{n}_j \times \hat{n}_k)$: If $j = k$, then

$$[\hat{S}_j, \hat{S}_j] = \hat{S}_j^2 - \hat{S}_j^2 = 0 = 2i\hat{S} \cdot (\hat{n}_j \times \hat{n}_j).$$

If $j \neq k$, recall that the vector $\hat{n}_j \times \hat{n}_k$ is given by

$$\begin{aligned}
\hat{n}_j \times \hat{n}_k &= \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \sin(\theta_j)\cos(\phi_j) & \sin(\theta_j)\sin(\phi_j) & \cos(\theta_j) \\ \sin(\theta_k)\cos(\phi_k) & \sin(\theta_k)\sin(\phi_k) & \cos(\theta_k) \end{vmatrix} \\
&= [\sin(\theta_j)\sin(\phi_j)\cos(\theta_k) - \cos(\theta_j)\sin(\theta_k)\sin(\phi_k)]\hat{x} - [\sin(\theta_j)\cos(\phi_j)\cos(\theta_k) - \cos(\theta_j)\sin(\theta_k)\cos(\phi_k)]\hat{y} \\
&\quad + (\sin(\phi_k - \phi_j)\sin(\theta_j)\sin(\theta_k))\hat{z} \\
&:= (\hat{n}_j \times \hat{n}_k)_x\hat{x} + (\hat{n}_j \times \hat{n}_k)_y\hat{y} + (\hat{n}_j \times \hat{n}_k)_z\hat{z}.
\end{aligned}$$

Then,

$$\begin{aligned}
[\hat{S}_j, \hat{S}_k] &= \hat{S}_j\hat{S}_k - \hat{S}_k\hat{S}_j \\
&= \begin{bmatrix} \cos(\theta_j) & e^{-i\phi_j}\sin(\theta_j) \\ e^{i\phi_j}\sin(\theta_j) & -\cos(\theta_j) \end{bmatrix} \begin{bmatrix} \cos(\theta_k) & e^{-i\phi_k}\sin(\theta_k) \\ e^{i\phi_k}\sin(\theta_k) & -\cos(\theta_k) \end{bmatrix} \\
&\quad - \begin{bmatrix} \cos(\theta_k) & e^{-i\phi_k}\sin(\theta_k) \\ e^{i\phi_k}\sin(\theta_k) & -\cos(\theta_k) \end{bmatrix} \begin{bmatrix} \cos(\theta_j) & e^{-i\phi_j}\sin(\theta_j) \\ e^{i\phi_j}\sin(\theta_j) & -\cos(\theta_j) \end{bmatrix} \\
&= \begin{bmatrix} \cos(\theta_j)\cos(\theta_k) + e^{i(\phi_k - \phi_j)}\sin(\theta_j)\sin(\theta_k) & e^{-i\phi_k}\cos(\theta_j)\sin(\theta_k) - e^{-i\phi_j}\sin(\theta_j)\cos(\theta_k) \\ e^{i\phi_j}\sin(\theta_j)\cos(\theta_k) - e^{i\phi_k}\cos(\theta_j)\sin(\theta_k) & \cos(\theta_j)\cos(\theta_k) + e^{i(\phi_j - \phi_k)}\sin(\theta_j)\sin(\theta_k) \end{bmatrix} \\
&\quad - \begin{bmatrix} \cos(\theta_j)\cos(\theta_k) + e^{-i(\phi_k - \phi_j)}\sin(\theta_j)\sin(\theta_k) & e^{-i\phi_j}\cos(\theta_k)\sin(\theta_j) - e^{-i\phi_k}\sin(\theta_k)\cos(\theta_j) \\ e^{i\phi_k}\sin(\theta_k)\cos(\theta_j) - e^{i\phi_j}\cos(\theta_k)\sin(\theta_j) & \cos(\theta_j)\cos(\theta_k) + e^{-i(\phi_j - \phi_k)}\sin(\theta_j)\sin(\theta_k) \end{bmatrix} \\
&= 2i \begin{bmatrix} \sin(\phi_k - \phi_j)\sin(\theta_j)\sin(\theta_k) & ie^{-i\phi_j}\sin(\theta_j)\cos(\theta_k) - ie^{-i\phi_k}\cos(\theta_j)\sin(\theta_k) \\ -ie^{i\phi_j}\sin(\theta_j)\cos(\theta_k) + ie^{i\phi_k}\cos(\theta_j)\sin(\theta_k) & \sin(\phi_j - \phi_k)\sin(\theta_j)\sin(\theta_k) \end{bmatrix} \\
&= 2i(\hat{S}_{\hat{x}}(\hat{n}_j \times \hat{n}_k)_{\hat{x}} + \hat{S}_{\hat{y}}(\hat{n}_j \times \hat{n}_k)_{\hat{y}} + \hat{S}_z(\hat{n}_j \times \hat{n}_k)_{\hat{z}}) = 2i\hat{S} \cdot (\hat{n}_j \times \hat{n}_k),
\end{aligned}$$

where

$$\hat{S}_{\hat{x}} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \hat{S}_{\hat{y}} = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \hat{S}_{\hat{z}} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

(5) Proof that $\{\hat{S}_j, \hat{S}_k\} = 2(\hat{n}_j \cdot \hat{n}_k)\hat{I}$: If $j \neq k$, then

$$\{\hat{S}_j, \hat{S}_k\} = \hat{S}_j\hat{S}_k + \hat{S}_k\hat{S}_j = \hat{S}_j\hat{S}_k - \hat{S}_j\hat{S}_k = 0 = 2(\hat{n}_j \cdot \hat{n}_k)\hat{I}.$$

If $j = k$, then

$$\{\hat{S}_j, \hat{S}_j\} = 2\hat{S}_j^2 = 2\hat{I} = 2(\hat{n}_j \cdot \hat{n}_j)\hat{I}.$$

(6) Proof that $e^{i(\xi_1\hat{S}_1 + \xi_2\hat{S}_2)} = \cos(\|\xi\|)\hat{I} + i(\xi_1\hat{S}_1 + \xi_2\hat{S}_2)\frac{\sin(\|\xi\|)}{\|\xi\|}$: Since

$$(\xi_1\hat{S}_1 + \xi_2\hat{S}_2)^2 = \xi_1^2\hat{I} + \xi_1\xi_2\{\hat{S}_1, \hat{S}_2\} + \xi_2^2\hat{I} = \|\xi\|^2\hat{I},$$

Then,

$$\begin{aligned} e^{i(\xi_1\hat{S}_1 + \xi_2\hat{S}_2)} &= \sum_{k=0}^{\infty} \frac{[i(\xi_1\hat{S}_1 + \xi_2\hat{S}_2)]^k}{k!} = \sum_{k=0}^{\infty} \frac{[i(\xi_1\hat{S}_1 + \xi_2\hat{S}_2)]^{2k}}{(2k)!} + \sum_{k=0}^{\infty} \frac{[i(\xi_1\hat{S}_1 + \xi_2\hat{S}_2)]^{2k+1}}{(2k+1)!} \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k\|\xi\|^{2k}\hat{I}}{(2k)!} + i(\xi_1\hat{S}_1 + \xi_2\hat{S}_2)\sum_{k=0}^{\infty} \frac{(-1)^k\|\xi\|^{2k}}{(2k+1)!} \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k\|\xi\|^{2k}}{(2k)!}\hat{I} + i(\xi_1\hat{S}_1 + \xi_2\hat{S}_2)\frac{1}{\|\xi\|}\sum_{k=0}^{\infty} \frac{(-1)^k\|\xi\|^{2k+1}}{(2k+1)!} = \cos(\|\xi\|)\hat{I} + i(\xi_1\hat{S}_1 + \xi_2\hat{S}_2)\frac{\sin(\|\xi\|)}{\|\xi\|}. \end{aligned}$$

If either ξ_1 or ξ_2 is set to 0, we get the familiar identity $e^{i\xi_j\hat{S}_j} = \cos(\xi_j)\hat{I} + i\hat{S}_j\sin(\xi_j)$, $j \in \{1, 2\}$.

APPENDIX D: PROOF OF LEMMA 2

Note first that Eq. (8) can be rewritten as

$$\frac{\left(e^{i\frac{\xi_1}{m}\hat{S}_1}e^{i\frac{\xi_2}{m}\hat{S}_2}\right)^m\hat{S}_1\hat{S}_2 + \left(e^{i\frac{\xi_2}{m}\hat{S}_2}e^{i\frac{\xi_1}{m}\hat{S}_1}\right)^m\hat{S}_2\hat{S}_1}{2} = U_{m-1}(a_m)\sin\left(\frac{\xi_1}{m}\right)\sin\left(\frac{\xi_2}{m}\right), \tag{D1}$$

by simply multiplying Eq. (8) from the right by $\hat{S}_1\hat{S}_2$. We want to show that Eqs. (7) and (D1) hold for all $m \in \mathbb{N}$. To that end, the proof will proceed by induction on m. For $m = 1$, we have

$$a_1 = \cos(\xi_1)\cos(\xi_2)\hat{I}, \quad b_1 = \hat{S}_1\sin(\xi_1)\cos(\xi_2) + \hat{S}_2\sin(\xi_2)\cos(\xi_1).$$

Then, by recalling that

$$\begin{aligned} e^{i\xi_1\hat{S}_1}e^{i\xi_2\hat{S}_2} &= (\cos(\xi_1)\hat{I} + i\hat{S}_1\sin(\xi_1))(\cos(\xi_2)\hat{I} + i\hat{S}_2\sin(\xi_2)) \\ &= a_1 + ib_1 - \sin(\xi_1)\sin(\xi_2)\hat{S}_1\hat{S}_2, \end{aligned}$$

we find that Eqs. (7) and (D1) hold in this case, namely,

$$\begin{aligned} \frac{e^{i\xi_1\hat{S}_1}e^{i\xi_2\hat{S}_2} + e^{i\xi_2\hat{S}_2}e^{i\xi_1\hat{S}_1}}{2} &= a_1 + ib_1 = T_1(a_1) + ib_1U_0(a_1), \\ \frac{e^{i\xi_1\hat{S}_1}e^{i\xi_2\hat{S}_2}\hat{S}_1\hat{S}_2 + e^{i\xi_2\hat{S}_2}e^{i\xi_1\hat{S}_1}\hat{S}_2\hat{S}_1}{2} &= U_0(a_1)\sin(\xi_1)\sin(\xi_2). \end{aligned}$$

Next, suppose that Eqs. (7) and (D1) are true for some $m \in \mathbb{N}$, and define

$$\xi_1' := \frac{\xi_1 m}{m+1}, \quad \xi_2' := \frac{\xi_2 m}{m+1}, \quad s_1 := \sin\left(\frac{\xi_1}{m+1}\right), \quad s_2 := \sin\left(\frac{\xi_2}{m+1}\right).$$

Then, by recalling that

$$\begin{aligned} e^{i\frac{\xi_1'}{m}\hat{S}_1}e^{i\frac{\xi_2'}{m}\hat{S}_2} &= \left[\cos\left(\frac{\xi_1'}{m}\right)\hat{I} + i\hat{S}_1\sin\left(\frac{\xi_1'}{m}\right)\right]\left[\cos\left(\frac{\xi_2'}{m}\right)\hat{I} + i\hat{S}_2\sin\left(\frac{\xi_2'}{m}\right)\right] \\ &= \left[\cos\left(\frac{\xi_1}{m+1}\right)\hat{I} + i\hat{S}_1\sin\left(\frac{\xi_1}{m+1}\right)\right]\left[\cos\left(\frac{\xi_2}{m+1}\right)\hat{I} + i\hat{S}_2\sin\left(\frac{\xi_2}{m+1}\right)\right] \\ &= a_{m+1} + ib_{m+1} - s_1s_2\hat{S}_1\hat{S}_2, \end{aligned}$$

we find that

$$
\begin{aligned}
&\frac{\left(e^{i\frac{\xi_1}{m+1}\hat{S}_1}e^{i\frac{\xi_2}{m+1}\hat{S}_2}\right)^{m+1}+\left(e^{i\frac{\xi_2}{m+1}\hat{S}_2}e^{i\frac{\xi_1}{m+1}\hat{S}_1}\right)^{m+1}}{2}\\
&=\frac{\left(e^{i\frac{\xi'_1}{m}\hat{S}_1}e^{i\frac{\xi'_2}{m}\hat{S}_2}\right)^{m}e^{i\frac{\xi'_1}{m}\hat{S}_1}e^{i\frac{\xi'_2}{m}\hat{S}_2}+\left(e^{i\frac{\xi'_2}{m}\hat{S}_2}e^{i\frac{\xi'_1}{m}\hat{S}_1}\right)^{m}e^{i\frac{\xi'_2}{m}\hat{S}_2}e^{i\frac{\xi'_1}{m}\hat{S}_1}}{2}\\
&=\underbrace{\frac{\left(e^{i\frac{\xi'_1}{m}\hat{S}_1}e^{i\frac{\xi'_2}{m}\hat{S}_2}\right)^{m}+\left(e^{i\frac{\xi'_2}{m}\hat{S}_2}e^{i\frac{\xi'_1}{m}\hat{S}_1}\right)^{m}}{2}}_{\text{use Eq. (7)}}(a_{m+1}+ib_{m+1})-s_1s_2\underbrace{\frac{\left(e^{i\frac{\xi'_1}{m}\hat{S}_1}e^{i\frac{\xi'_2}{m}\hat{S}_2}\right)^{m}\hat{S}_1\hat{S}_2+\left(e^{i\frac{\xi'_2}{n}\hat{S}_2}e^{i\frac{\xi'_1}{m}\hat{S}_1}\right)^{m}\hat{S}_2\hat{S}_1}{2}}_{\text{use Eq. (D1)}}\\
&=[T_m(a_{m+1})+ib_{m+1}U_{m-1}(a_{m+1})](a_{m+1}+ib_{m+1})-s_1s_2[U_{m-1}(a_{m+1})s_1s_2]\\
&=a_{m+1}T_m(a_{m+1})-\left(b_{m+1}^2+s_1^2s_2^2\right)U_{m-1}(a_{m+1})+ib_{m+1}[a_{m+1}U_{m-1}(a_{m+1})+T_m(a_{m+1})]\\
&=a_{m+1}T_m(a_{m+1})-[\hat{I}-(a_{m+1})^2]U_{m-1}(a_{m+1})+ib_{m+1}[a_{m+1}U_{m-1}(a_{m+1})+T_m(a_{m+1})]\\
&=T_{m+1}(a_{m+1})+ib_{m+1}U_m(a_{m+1}),
\end{aligned}
$$

which proves Eq. (7) for $m+1$. The last equality follows from the fact that the Chebyshev polynomials satisfy

$$T_{m+1}(x)=xT_m(x)-(1-x^2)U_{m-1}(x)\quad\text{and}\quad U_m(x)=xU_{m-1}(x)+T_m(x),\quad\forall x\in\mathbb{R}.$$

Likewise, we have

$$
\begin{aligned}
&\frac{\left(e^{i\frac{\xi_1}{m+1}\hat{S}_1}e^{i\frac{\xi_2}{m+1}\hat{S}_2}\right)^{m+1}\hat{S}_1\hat{S}_2+\left(e^{i\frac{\xi_2}{m+1}\hat{S}_2}e^{i\frac{\xi_1}{m+1}\hat{S}_1}\right)^{m+1}\hat{S}_2\hat{S}_1}{2}\\
&=\frac{e^{i\frac{\xi'_1}{m}\hat{S}_1}e^{i\frac{\xi'_2}{m}\hat{S}_2}\left(e^{i\frac{\xi'_1}{m}\hat{S}_1}e^{i\frac{\xi'_2}{m}\hat{S}_2}\right)^{m}\hat{S}_1\hat{S}_2+e^{i\frac{\xi'_2}{m}\hat{S}_2}e^{i\frac{\xi'_1}{m}\hat{S}_1}\left(e^{i\frac{\xi'_2}{m}\hat{S}_2}e^{i\frac{\xi'_1}{m}\hat{S}_1}\right)^{m}\hat{S}_2\hat{S}_1}{2}\\
&=-s_1s_2\underbrace{\frac{\hat{S}_1\hat{S}_2\left(e^{i\frac{\xi'_1}{m}\hat{S}_1}e^{i\frac{\xi'_2}{m}\hat{S}_2}\right)^{m}\hat{S}_1\hat{S}_2+\hat{S}_2\hat{S}_1\left(e^{i\frac{\xi'_2}{m}\hat{S}_2}e^{i\frac{\xi'_1}{m}\hat{S}_1}\right)^{m}\hat{S}_2\hat{S}_1}{2}}_{\text{use Eq. (7)}}+(a_{m+1}+ib_{m+1})\underbrace{\frac{\left(e^{i\frac{\xi'_1}{m}\hat{S}_1}e^{i\frac{\xi'_2}{m}\hat{S}_2}\right)^{m}\hat{S}_1\hat{S}_2+\left(e^{i\frac{\xi'_2}{m}\hat{S}_2}e^{i\frac{\xi'_1}{m}\hat{S}_1}\right)^{m}\hat{S}_2\hat{S}_1}{2}}_{\text{use Eq. (D1)}}\\
&=s_1s_2[T_m(a_{m+1})-ib_{m+1}U_{m-1}(a_{m+1})]+(a_{m+1}+ib_{m+1})[U_{m-1}(a_{m+1})s_1s_2]\\
&=s_1s_2[a_{m+1}U_{m-1}(a_{m+1})+T_m(a_{m+1})]=s_1s_2U_m(a_{m+1}),
\end{aligned}
$$

which proves Eq. (D1) for $m+1$. By induction, (7) and (D1) hold for all $m\in\mathbb{N}$.

APPENDIX E: ALTERNATE PROOF FOR LEMMA 1 WHEN $\hat{A}_1=\hat{S}_1$ AND $\hat{A}_2=\hat{S}_2$

Recall that the Mehler-Heine theorem (2) states that

$$\lim_{m\to\infty}m^{-\alpha}P_m^{(\alpha,\beta)}\left(\cos\left(\frac{z}{m}\right)\right)=\left(\frac{z}{2}\right)^{-\alpha}J_\alpha(z),$$

uniformly on compact subsets of $\mathbb{C}$. As pointed out in Ref. [10], Szegő's proof of the theorem establishes that

$$\lim_{m\to\infty}m^{-\alpha}P_m^{(\alpha,\beta)}\left(1-\frac{z^2}{2m^2}+o(m^{-2})\right)=\left(\frac{z}{2}\right)^{-\alpha}J_\alpha(z).$$

Next, since

$$\cos\left(\frac{z_1}{m}\right)\cos\left(\frac{z_2}{m}\right)=1-\frac{z_1^2+z_2^2}{2m^2}+o(m^{-2}),$$

where the last equality follows by substituting each factor with its Maclaurin series, we get

$$
\begin{aligned}
\lim_{m\to\infty}m^{-\alpha}P_m^{(\alpha,\beta)}\left(\cos\left(\frac{z_1}{m}\right)\cos\left(\frac{z_2}{m}\right)\right)&=\lim_{m\to\infty}m^{-\alpha}P_m^{(\alpha,\beta)}\left(1-\frac{z_1^2+z_2^2}{2m^2}+o(m^{-2})\right)\\
&=\left(\frac{\sqrt{z_1^2+z_2^2}}{2}\right)^{-\alpha}J_\alpha\left(\sqrt{z_1^2+z_2^2}\right),
\end{aligned}\tag{E1}
$$

uniformly on compact subsets of $\mathbb{C}^2$. The analog of (E1) was established in Ref. [10] for a ratio of cosines, instead of a product. Setting $\alpha = \beta = -1/2$,

$$J_{-1/2}(z) = \sqrt{\frac{2}{\pi z}}\cos(z), \quad P_m^{(-1/2,-1/2)}(z) = \frac{(2m)!}{2^{2m}(m!)^2}T_m(z),$$

where $T_m(\cdot)$ denotes the m^{th} degree Chebyshev polynomial of the first kind. Then, Eq. (E1) becomes

$$\lim_{m\to\infty}\frac{\sqrt{m\pi}(2m)!}{2^{2m}(m!)^2}T_m\left(\cos\left(\frac{z_1}{m}\right)\cos\left(\frac{z_2}{m}\right)\right) = \cos\left(\sqrt{z_1^2+z_2^2}\right). \tag{E2}$$

Stirling's formula gives that

$$\lim_{m\to\infty}\frac{\sqrt{m\pi}(2m)!}{2^{2m}(m!)^2} = 1,$$

and hence Eq. (E2) reduces to

$$\lim_{m\to\infty}T_m\left(\cos\left(\frac{z_1}{m}\right)\cos\left(\frac{z_2}{m}\right)\right) = \cos\left(\sqrt{z_1^2+z_2^2}\right),$$

Since the uniform convergence of a sequence of analytic functions f_m to f on compact subsets of $\mathbb{C}^2$ implies the same type of convergence for their derivatives to $\frac{\partial}{\partial z}f$,

$$\lim_{m\to\infty}\frac{\partial}{\partial z_i}T_m\left(\cos\left(\frac{z_1}{m}\right)\cos\left(\frac{z_2}{m}\right)\right) = \frac{\partial}{\partial z_i}\cos\left(\sqrt{z_1^2+z_2^2}\right),$$

uniformly on compact subsets of $\mathbb{C}^2$, $i \in \{1, 2\}$. Hence,

$$\lim_{m\to\infty}\left(1 - is_1\frac{\partial}{\partial z_1} - is_2\frac{\partial}{\partial z_2}\right)T_m\left(\cos\left(\frac{z_1}{m}\right)\cos\left(\frac{z_2}{m}\right)\right) = \left(1 - is_1\frac{\partial}{\partial z_1} - is_2\frac{\partial}{\partial z_2}\right)\cos\left(\sqrt{z_1^2+z_2^2}\right),$$

uniformly on compact subsets of $\mathbb{C}^2$, which is exactly the statement in Lemma 1 for the spin operators $\hat{S}_1$ and $\hat{S}_2$.

[1] L. Cohen and M. O. Scully, Joint Wigner distribution for spin-1/2 particles, Found. Phys. **16**, 295 (1986).
[2] C. Chandler, L. Cohen, C. Lee, M. Scully, and K. Wodkiewicz, Quasi-probability distribution for spin-1/2 particles, Found. Phys. **22**, 867 (1992).
[3] R. Schwonnek and R. F. Werner, The Wigner distribution of n arbitrary observables, J. Math. Phys. **61**, 082103 (2020).
[4] R. F. V. Anderson, On the Weyl functional calculus, J. Funct. Anal. **6**, 110 (1970).
[5] L. Hörmander, *The Analysis of Partial Differential Operators I*, Classics in Mathematics (Springer, Berlin, 1983), Vol. 256.
[6] B. C. Hall, *Lie Groups, Lie Algebras, and Representations* (Springer, Berlin, 2003).
[7] G. Szegő, *Orthogonal Polynomials*, Colloquium Publications (American Mathematical Society, Providence, 1975), Vol. 23.
[8] M. O. Scully, H. Walther, and W. Schleich, Feynman's approach to negative probability in quantum mechanics, Phys. Rev. A **49**, 1562 (1994).
[9] K. Jacobs and D. A. Steck, A straightforward introduction to continuous quantum measurement, Contemp. Phys. **47**, 279 (2006).
[10] R. L. Streit, Limits of Chebyshev polynomials when the argument is a ratio of cosines, J. Approx. Theory **40**, 393 (1984).