Detecting Coherence via Spectrum Estimation

Xiao-Dong Yu∗ and Otfried Gühne†

Naturwissenschaftlich-Technische Fakultät, Universität Siegen, Walter-Flex-Straße 3, 57068 Siegen, Germany

(Dated: August 28, 2018)

Coherence is a basic phenomenon in quantum mechanics and considered to be an essential resource in quantum information processing. Although the quantification of coherence has attracted a lot of interest, the lack of efficient methods to measure the coherence in experiments limits the applications of coherence measures. In this work, we address this problem by developing an efficient method to witness and estimate the coherence of quantum systems based on spectrum estimation of the density matrix. With this method, we can not only witness the presence of coherence, but also obtain a good estimation of coherence of quantum systems with only few measurements. As an illustration, we show how to use our method to characterize the freezing phenomenon of coherence with only two local measurements for any $N$-qubit quantum systems. Our approach also has other applications in quantum information processing, such as the characterization of distillability and entanglement transformations.

Introduction.—Quantum coherence is a fundamental feature of quantum mechanics, describing the capability of a quantum state to exhibit quantum interference phenomena. Consequently, it is an essential ingredient in quantum information processing [1], and plays a central role in emergent fields, such as quantum metrology [2, 3] and quantum thermodynamics [4].

The notion of quantum coherence was already early developed in quantum optics [5–7], but only in recent years the quantification of coherence has been treated rigorously and frameworks for quantifying coherence were proposed [8–13]. Notably, in Ref. [11] a rigorous framework based on the notion of a general resource theory was introduced. In this framework, the free states are the incoherent states which are diagonal in the incoherent basis, and the free operations are incoherent operations whose Kraus operators map the incoherent states to incoherent states. Other frameworks were also proposed to make the quantification of coherence applicable to various physical situations [9, 12, 14–17]. The main difference between these frameworks is that they have different notions of free operations.

Based on these frameworks, several coherence measures have been proposed, such as the relative entropy of coherence [9, 11], the $l_1$-norm of coherence [11], the geometric measure of coherence [18], the robustness of coherence [19, 20], and others [21–24]. These coherence measures make it possible to quantitatively study the role of coherence in various physical contexts. Further important properties of quantum coherence, such as the distillation of coherence [22, 25], the relation between coherence and quantum correlations [18, 26–31], and the freezing phenomenon of coherence [32, 33], can be studied based on the coherence measures that have been proposed.

While many theoretical works have been devoted to the quantification of coherence, only few quantities have been examined in experiments [34–36]. One important reason for this situation is the fact that few methods are known to obtain the coherence measures in experiments except full state tomography [37–39]. The lack of efficient and scalable methods for coherence detection severely limits the applications of coherence measures.

To overcome this situation and to make the quantification of coherence a common tool for quantum information processing, it is of paramount importance to improve the evaluation of coherence in experiments. In this work, we address this problem by developing an efficient method to witness and estimate the coherence of quantum systems based on spectrum estimation of the density matrix. With this method, we can not only witness the presence of coherence, but also obtain a good estimation of coherence of quantum systems with only few measurements. As an illustration, we show how to use our method to characterize the freezing phenomenon of coherence with only two local measurements for any $N$-qubit quantum systems. Our approach relies on the mathematical theory of the majorization lattice, and can, as we explain, also be used for other problems in quantum information processing.

Resource theory of coherence.—In the resource theory of coherence, the free states are incoherent states, defined as $\delta = \sum_p p|l\rangle\langle l|$, where $|l\rangle$ represents a fixed reference basis, known as the incoherent basis. The set of all incoherent states is denoted by $I$. All other states which are not diagonal in the incoherent basis contain some coherence.

The definition of free operations within the resource theory of coherence is not unique, and several approaches have been proposed based on different physical or mathematical considerations [8]. The most widely used definition is the incoherent operations proposed by Baumgratz et al. [11], which are defined as $\Lambda(\rho) = \sum_n K_n\rho K_n^\dagger$, where the Kraus operators $K_n$ satisfy not only $\sum_n K_n^\dagger K_n = 1$, but also $K_n^\dagger K_n \subseteq I$ for each $K_n$, i.e., each $K_n$ maps incoherent states to incoherent states.

With the definition of free states and free operations, the frameworks for quantifying quantum coherence are constructed from the general resource theory [40, 41]. Basically, it requires any coherence measure satisfies the non-negativity $C(\rho) \geq 0$ and $C(\rho) = 0$ if and only if $\rho \in I$ and monotonicity $C(\rho) \geq C(\Lambda(\rho))$ for any incoherent operation $\Lambda$. Some other constraints may also be included based on different physical considerations [8, 11, 17, 42].

In this work, we focus on the estimation of relative entropy of coherence, which is defined as $C_\rho(\rho) = \min_{\delta \in I} S(\rho\|\delta)$, where $S(\rho\|\delta) = \Tr(\rho \log \rho - \rho \log \sigma)$ is the relative entropy. This is not only because the relative entropy of coherence is a legitimate coherence measure in all of the proposed frame-
works and plays crucial roles in many physical processes such as coherence distillation and coherence freezing. More importantly, the estimation of relative entropy of coherence can also provide a lower bound for many other coherence measures, such as coherence of formation, robustness of coherence, and $l_1$-norm of coherence [22, 23]. Mathematically, the relative entropy of coherence also admits the following closed form,

$$C_r(\rho) = S(\rho_d) - S(\rho),$$  \hspace{1cm} (1)

where $S$ is von Neumann entropy and $\rho_d$ is the diagonal part of $\rho$ in the incoherent basis.

Majorization and the majorization lattice.—Majorization is a mathematical tool that is widely used in quantum theory, such as quantum entanglement and quantum thermodynamics [43, 44], and also in other fields, such as statistics and economics [45]. A probability distribution $a = (a_1, a_2, \ldots, a_n)$ is said to majorize a probability distribution $b = (b_1, b_2, \ldots, b_n)$ (written as $a \succ b$), if they satisfy $\sum_{i=1}^{k} a_i \geq \sum_{i=1}^{k} b_i$ for all $k = 1, 2, \ldots, n$, where $a^1(b^1)$ is the probability distribution with the same components as $a(b)$, but sorted in descending order. Hereafter, we will assume the components of a probability distribution are already sorted in descending order and all vectors are column vectors, unless stated otherwise.

Majorization imposes an important constraint between the measurement result and the spectrum of the quantum state, which is shown in the following well known fact [43].

Lemma 1. Let $\rho$ be a quantum state in an $n$-dimensional Hilbert space with spectrum $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$.

(1) If we perform a von Neumann measurement on the quantum state and get the probability distribution of measurement outcomes $p = (p_1, p_2, \ldots, p_n)$, then $p \prec \lambda$.

(2) The above condition is also sufficient in the sense that if $p = (p_1, p_2, \ldots, p_n)$ is some probability distribution such that $p \prec \lambda$, then there exists a von Neumann measurement $\{P_i\}_{i=1}^{n}$ such that $p_i = \text{Tr}(P_i\rho)$.

For the estimation of coherence, we also take advantage of the strict Schur-concavity of Shannon entropy $S$. Mathematically, it says that if $a < b$ and $a \neq b$, then $S(a) > S(b)$.

Compared with majorization, the theory of the majorization lattice is much less-known. Given two distributions $a$ and $b$ which are not comparable in the sense of majorization, one may ask whether there is a “smallest” distribution majorizing both of them. In fact, such a distribution can be defined and is called the majorization join. Similarly, the “largest” distribution that is majorized by both $a$ and $b$ is called the majorization meet. More formally, a probability distribution $c$ is called the majorization join (meet) of $a$ and $b$ if it satisfies the following two conditions: (1) $c > a, b$ ($c < a, b$); (2) $c < \tilde{c}$ ($c > \tilde{c}$) for any $\tilde{c}$ that satisfies $a, b < \tilde{c}$ ($a, b > \tilde{c}$) [46].

The majorization join and meet of $a$ and $b$ are usually denoted as $a \vee b$ and $a \wedge b$ respectively. The basic process for the construction of majorization join (meet) is quite simple. The necessary and sufficient condition for $\tilde{c} < a, b$ is $\sum_{i=1}^{k} c_i \leq \min\{\sum_{i=1}^{k} a_i, \sum_{i=1}^{k} b_i\}$. Hence if $\sum_{i=1}^{k} c_i = \min\{\sum_{i=1}^{k} a_i, \sum_{i=1}^{k} b_i\}$ and $c_i$ are in descending order, then $c < a, b$ and $c > \tilde{c}$, i.e., $c = a \wedge b$. For the case of majorization meet, this definition can be directly used. For majorization join, the vector $c$ constructed from $\sum_{i=1}^{k} c_i = \max\{\sum_{i=1}^{k} a_i, \sum_{i=1}^{k} b_i\}$ may not be in descending order. Some further flattening operations, as shown in Ref. [46], may be needed. In Appendix A, we present a more detailed description of majorization join and meet, and further show that they can be naturally generalized to the case where an infinite number of probability distributions is considered.

Single-partite systems.—We first consider single-partite quantum systems to explain the basic idea of our method. From Eq. (1), we can easily see that the coherence of the quantum system can be revealed by the diagonal part of the quantum state $\rho$, i.e., the vector $d = (d_1, d_2, \ldots, d_n)$, and the spectrum of $\rho$, i.e., the vector $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$. In an experiment, one can easily determine the diagonal part of the quantum state $\rho$ by just measuring in the incoherent basis. It remains to determine the spectrum of $\rho$, although the eigenbasis is unknown.

Suppose that $\rho$ is incoherent, then we have $d = \lambda$. According to Lemma 1, the probability distribution $p = (p_1, p_2, \ldots, p_n)$ from any von Neumann measurement must satisfy $p \prec \lambda = d$. So, if we perform a von Neumann measurement and get the probability distribution $p$ such that $p \prec d$, then we can immediately assert that $\rho$ is coherent. Conversely, if $\rho$ is coherent, there always exists a von Neumann measurement $\{P_i\}_{i=1}^{n}$ such that the probability distribution of the outcomes $p$ satisfies that $p \prec d$. As a special case, we can just choose $P_i = |\varphi_i\rangle\langle\varphi_i|$ for $i = 1, 2, \ldots, n$, where $|\varphi_i\rangle$ are the eigenvectors of $\rho$.

The previous discussion implies that we can always prove the presence of coherence by showing the probability distribution of the outcomes of some measurement is not majorized by the probability distribution of the measurement outcomes in the incoherent basis. This provides a qualitative method for witnessing quantum coherence. However, contrary to the case of entanglement and separability, the set of incoherent states is of measure zero in the state space, so a mere statement about the presence of coherence is of limited value. Instead, a quantitative method, giving an estimate on the amount of coherence, is desirable. So the problem arises what can one say about the amount of coherence in a quantum system, if the probability distributions $p$ and $d$ with $p \prec d$ are known.

To answer this question, we use the strict Schur-concavity of Shannon entropy, which implies that $S(\rho) = S(\lambda) \leq S(p)$, where $\lambda$ is the spectrum of $\rho$, and $p$ is the probability distribution of measurement outcomes of any von Neumann measurement. Consequently, we have a lower bound of the coherence,

$$C_r(\rho) \geq \max\{0, S(d) - S(p)\},$$  \hspace{1cm} (2)

Still, this bound may not be strictly positive, even if we can conclude from $p \prec d$ that the quantum system contains coher-
ence. As an example, consider the state $|\psi\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ in a qutrit system, then $d = (1/2, 1/2, 0)$. If we perform a measurement in the basis $\{|0\rangle + |1\rangle + |2\rangle\}/\sqrt{3}, (|0\rangle + \omega|1\rangle + \omega^2|2\rangle)/\sqrt{3}, (|0\rangle + \omega^2|1\rangle + \omega|2\rangle)/\sqrt{3}$, where $\omega = e^{2\pi i/3}$, then the probability distribution of the measurement outcomes is $p = (2/3, 1/3, 1/3)$. Then we have $p \neq d$, but the lower bound from Eq. (2) is still zero, as $S(d) - S(p) < 0$.

The main idea of solving the problem is that we can construct the “smallest” probability distribution that majorizes both $p$ and $d$, i.e., the majorization join of $p$ and $d$. According to Lemma 1, we have that $d < \lambda$ and $p < \lambda$. From the definition of majorization join, we get that $d \vee p < \lambda$. Furthermore, $p \neq d$ implies that $d \vee p$ strictly majorizes $d$, i.e., $d < d \vee p$ and $d \neq d \vee p$. Hence, the strict Schur-concavity of Shannon entropy implies that $C_r(\rho) = S(d) - S(\lambda) \geq S(d) - S(d \vee p) > 0$. These results are summarized in the following proposition.

**Proposition 2.** If the probability distribution of the outcomes of a von Neumann measurement $p$ is not majorized by diagonal entries of the quantum state $d$, i.e., $p \neq d$, then the quantum system contains coherence. Furthermore, a nonzero lower bound of the relative entropy of coherence is given by

$$C_r(\rho) \geq S(d) - S(d \vee p).$$

(3)

The lower bound provided by Proposition 2 is device-independent in the sense that we do not resort to the actual form of the measurement basis for measuring the probability distribution $p$. The benefit of this feature is that the bound in Eq. (3) is robust to the errors in implementing the measurement. In addition, the lower bound provided by Eq. (3) is also tight in this sense, i.e., it is the best lower bound we can get without resorting to the actual form of the measurement. See Appendix B for more details.

The method represented in Eq. (3) can also be naturally generalized to the case that many measurements are performed. Suppose that $p_1, p_2, \ldots, p_k$ are the probability distributions of measurement outcomes of $k$ different measurements then we can take advantage of the majorization join of all $p_1, p_2, \ldots, p_k$ and $d$ to estimate the lower bound of the relative entropy of coherence, i.e.,

$$C_r(\rho) \geq S(d) - S(d \vee p_1 \vee \cdots \vee p_k).$$

(4)

As $d \vee p_1 \vee \cdots \vee p_k > d \vee p_1 \vee \cdots \vee p_{k-1} \cdots > d \vee p_1$, then the Schur-concavity of Shannon entropy implies that the more measurements are performed, the better the lower bound is. This allows one to successively improve the estimate of the coherence of the quantum system. In actual experiments, one usually has some expectations or predictions concerning the state of the quantum system. This may also be used to choose the measurement and may provide better bounds on the coherence.

**Multi-partite systems.**—Now we extend our method to the more challenging case of the multi-partite quantum systems. The difference between the multi-partite case and the single-partite case is that usually only local measurements are allowed for multi-partite quantum systems in experiments. Hence, for an entangled basis, we cannot get the probability distribution $p$ efficiently in experiments. As a compromise, we resort to the estimation method of $p$. In entanglement detection theory, a lot of efficient methods have been developed to estimate the fidelity $\langle \psi|\phi\rangle$ with local measurements [47]. Applying these methods to the entangled basis $\{|\psi\rangle\}$, we can get the estimations of the probability distribution $p = (p_1, p_2, \ldots, p_n)$, where $p_i = \langle \psi|\phi_i\rangle$. Here the components may not be in descending order. Usually, the estimations can be expressed as linear constraints $Ap \geq \alpha$ and $Bp = \beta$, where $A$ and $B$ are matrices, $\alpha$ and $\beta$ are vectors, and “$\geq$” denotes the component-wise comparison. For example, in the theory of witnesses for graph states this is the case [48].

Let $X$ denote the feasible set, i.e., $X = \{p \mid Ap \geq \alpha, Bp = \beta\}$, then, using the transitivity of the majorization relation, it directly follows that $\lambda > d \vee (\bigwedge_{p \in X} p)$, where $\lambda$ is the spectrum of $\rho$, $d$ is the diagonal part of $\rho$, and $\bigwedge_{p \in X} p$ is the majorization meet of all probability distributions in $X$. Then the Schur-concavity of Shannon entropy implies that

$$C_r(\rho) = S(d) - S(d \vee (\bigwedge_{p \in X} p)).$$

(5)

The main difficulty for calculating the bound in Eq. (5) is the majorization meet of infinite number of probability distributions in $X$. In the following, we show that this problem can be converted into a linear program, for which efficient algorithms exist [49].

Suppose $\hat{p} = \bigwedge_{p \in X} p = (\hat{p}_1, \hat{p}_2, \ldots, \hat{p}_n)$, then $\hat{p}_k = s_k - s_{k-1}$, where $s_0 = 0$ and $s_k$ can be written as the convex optimization problem, $\min_{p \in X} \sum_{i=1}^k p_i^j$, for $k = 1, 2, \ldots, n$. Note that the components of $p$ in $X$ may not be in descending order, so the “$\vee$” is necessary for $p_i^j$. By combining the minimax theorem with duality techniques, we convert the optimization $s_k = \min_{p \in X} \sum_{i=1}^k p_i^j$, to the following linear program,

$$\begin{align*}
\text{maximize} & \quad \alpha^T \mu + \beta^T \nu \\
\text{subject to} & \quad 0 \leq A^T \mu + B^T \nu \leq 1 \\
& \quad 1^T A^T \mu + 1^T B^T \nu = k \\
& \quad \mu \geq 0,
\end{align*}$$

(6)

where $\mu$ and $\nu$ are vectors that have the same dimension as $\alpha$ and $\beta$ respectively, and 0 and 1 are vectors with all components being 0 and 1 respectively. The proof is shown in Appendix C.

The estimation method based in Eq. (5) is quite efficient in practice. For example, we have tested the optimization problem for up to 10-qubit systems with 513 equality constraints and 1024 inequality constraints using the CVXPY package [50]. The vector $\hat{p} = \bigwedge_{p \in X} p$ with 1024 components can be determined within 30 minutes in a common laptop.

Before discussing examples, we note that our method also has other applications in quantum physics. First, it may be applied to the majorization criterion for distillability of two-party quantum states [51, 52]. This criterion states that for an
undistillable state the eigenvalues of the global state $\rho_{AB}$ are majorized by the eigenvalues of the reduced state $\rho_A$. At first sight, it seems that state tomography is required for checking this relation, but our methods provide a way to circumvent this. As already mentioned, from some local measurements one can typically obtain linear constraints on the eigenvalues of the global state and reduced state. Then, using Eq. (6) we can compute the majorization meet of the possible global eigenvalues and the majorization join of the local ones. If they violate the relation mentioned above, the state must be distillable. This concept may be generalized to other separability criteria based on majorization [53].

Second, majorization of the Schmidt coefficients of a pure state provides a necessary and sufficient condition for the pure state transformations under local operations and classical communication in the resource theory of entanglement [54–56]. Thus, our methods can be used to obtain the common resource states that can generate a whole subclass of entangled states.

**Characterizing the freezing of coherence.**—We first briefly review some basic results in the freezing of coherence. The freezing of coherence means that the coherence of the quantum system (quantified by some coherence measure) is not affected by noise. When the freezing of coherence is independent of the choice of coherence measures, it is called universal freezing of coherence. Universal freezing of coherence was described in Refs. [32, 33]. Especially, in Ref. [33], the authors prove a universal freezing theorem, which says that under a strictly incoherent channel, the universal freezing of coherence occurs if and only if the relative entropy of coherence is frozen. This implies that if we can witness the freezing of relative entropy of coherence, then we can assure that the coherence of the quantum system is completely unaffected by noise.

One of the most important examples of universal freezing of coherence is the N-qubit GHZ state $\rho_0 = ((0)^{\otimes N} + |1)^{\otimes N})/\sqrt{2}$, in the local bit flip channel $\Lambda^{\otimes N}$, where $\Lambda_\gamma(\rho) = (1/2 + e^{-\gamma/2})\rho + (1/2 - e^{-\gamma/2})|0\rangle\langle 0|\rho|0\rangle\langle 0|$, and $\gamma$ is a parameter that represents the strength of the noise. Direct calculations show that the state at time $t$ is of the form

$$\rho_t = \sum_{j} p_j^t |\varphi_j^t\rangle\langle \varphi_j^t|,$$

where $|\varphi_j^t\rangle$ is the GHZ basis, i.e.,

$$|\varphi_j^t\rangle = |\varphi_{i_1,t_1,...,i_N,t_N}\rangle = |i_{12}...i_{NK}\rangle \pm |\bar{i}_{12}...\bar{i}_{NK}\rangle, \tag{8}$$

with $l_1 = 0$, $l_{i\neq 1} = 0,1$, and $\bar{l}_i = 1 - l_i$, and $p = (p_j^t)$ are the corresponding time-dependent probabilities. It is easy to prove that $C_t(\rho_0) = C(\rho_0)$, then according to the universal freezing theorem, all coherence measures are frozen, i.e., $C_t(\rho_1) = C(\rho_0)$ for all coherence measures $C$

In order to get an accurate estimation of the coherence, we would like to perform a GHZ-basis measurement to estimate the coherence of the quantum system. Since the GHZ-basis measurement is highly entangled, it is not easy to get the exact probability distribution in experiments. As a compromise, we choose to estimate the probability distribution with the following two local measurement settings,

$$X = \bigotimes_{i=1}^N \sigma_{x_i}^{(j)}, \quad Z = \bigotimes_{i=1}^N \sigma_{z_i}^{(j)}. \tag{9}$$

An advantage of this measurement setting is that we can not only get the probabilities $\langle X \rangle$ and $\langle Z \rangle$, but also all $\langle Z_S \rangle$ with $Z_S = \bigotimes_{i\in S} \sigma_{z_i}^{(j)}$, for any subset $S$ of $\{1,2,\ldots,N\}$. Furthermore, all $X$ and $Z_E$ are diagonal in the GHZ basis, when $E$ is a nonempty subset of $\{1,2,\ldots,N\}$ with an even number of elements. Hence, from $\langle X \rangle$ and $\langle Z_E \rangle$, together with the condition that $p$ is a probability distribution, we can get $2^{n-1} + 1$ linear equalities and $2^n$ linear inequalities for the estimation of the probability distribution $p$. As all the constraints are linear equalities or inequalities, our method can be applied immediately to get the estimation of relative entropy of coherence as expressed in Eq. (5). For convenience, we split $p$ into two parts, $p = p^x \oplus p^z = (p_0^x) \oplus (p_0^z)$. In the ideal case, if the fidelity of the initial GHZ state is one, we have that $\langle X \rangle = 1$ at any time. Then the $2^{n-1} - 1$ independent equalities from $\langle X_E \rangle$ completely determine the probability distribution $p^x$ which also has $2^{n-1} - 1$ independent parameters. Hence, we get $\land_{p \in X} p = p^x \oplus 0$. Additionally, the measurement $Z$ will also give us the diagonal part of the quantum state $d = \frac{1}{2} p^z \oplus \frac{1}{2} p^x$. Then Eq. (5) implies that $C_t(\rho_0) \geq 1 = C(\rho_0)$. Thus, we prove the freezing of coherence with only two local measurements given in Eq. (9).

![Figure 1](image.png)

**FIG. 1.** Characterizing the freezing of coherence. We consider the situation that the preparation of the initial state is affected by dephasing and depolarizing noise in three-qubit (Figs. (a) and (b) respectively) and four-qubit (Figs. (c), and (d) respectively) cases. The dashed line is the actual value of coherence $C(\rho)$ and the solid line is the estimation of coherence with our method. Lines with different colors represent the initial states with different fidelities.
the detailed noise models are shown in Appendix D. Note that in the case of the depolarizing noise, the freezing of coherence no longer occurs, but the coherence is still very resistant to noise. In both cases, our method can still prove the resistance of coherence to noise, as illustrated in Fig. 1.

Conclusions.—In this work, we propose an efficient method for coherence detection via spectrum estimation. This method is efficient in two senses: first, the number of measurements needed is quite small, which is friendly to experiments; second, the optimization process is only linear programming, which can be efficiently solved in practice. As an illustration, we show that we can witness the freezing phenomenon of coherence with only two local measurements. We hope this work can not only promote verification of various results on the quantification of coherence, but also promote the application of coherence measures to quantum information experiments.

We would like to thank Barbara Kraus, Katharina Schwaiger, and Nikolai Wyderka for discussions. This work was supported by the DFG and the ERC (Consolidator Grant 683107/TempoQ). X.D.Y. acknowledges funding from a CSC-DAAD scholarship.

Appendix A: Majorization lattice

For simplicity, we will restrict our discussion to probability distributions. All the discussions can be naturally generalized to the general case. We first recall the definition of majorization. A probability distribution \( \mathbf{a} = (a_1, a_2, \ldots, a_n) \) is said to majorize probability distribution \( \mathbf{b} = (b_1, b_2, \ldots, b_n) \) written as \( \mathbf{a} \succ \mathbf{b} \), if they satisfy \( \sum_{i=1}^{k} a_i^k \geq \sum_{i=1}^{k} b_i^k \) for all \( k = 1, 2, \ldots, n \), where \( \mathbf{a}^k \) (\( \mathbf{b}^k \)) is the vector with the same components as \( \mathbf{a} \) (\( \mathbf{b} \)), but sorted in descending order. As in the main text, we will assume the components of a probability distribution are already sorted in descending order, unless otherwise stated.

Majorization lattice deals with the “smallest” (“largest”) probability distribution that majorize (is majorized by) two probability distributions \( \mathbf{a} \) and \( \mathbf{b} \). More formally, a probability distribution \( \mathbf{c} \) is called the majorization join of \( \mathbf{a} \) and \( \mathbf{b} \), if it satisfies that

1. \( \mathbf{c} \succ \mathbf{a}, \mathbf{b} \);
2. \( \mathbf{c} \preceq \mathbf{c} \) for any \( \mathbf{c} \) that satisfies \( \mathbf{a}, \mathbf{b} \preceq \mathbf{c} \).

Similarly, a probability distribution \( \mathbf{c} \) is called the majorization meet of \( \mathbf{a} \) and \( \mathbf{b} \), if it satisfies that

1. \( \mathbf{c} \prec \mathbf{a}, \mathbf{b} \);
2. \( \mathbf{c} \succeq \mathbf{c} \) for any \( \mathbf{c} \) that satisfies \( \mathbf{a}, \mathbf{b} \succeq \mathbf{c} \).

From the definition, we can easily check that the majorization join and meet are unique. We will denote the majorization join and meet of \( \mathbf{a} \) and \( \mathbf{b} \) as \( \mathbf{a} \lor \mathbf{b} \) and \( \mathbf{a} \land \mathbf{b} \) respectively. From the definition, we can also easily prove that the majorization join and meet satisfy the commutativity and associativity, e.g.

1. \( \mathbf{p}_1 \lor \mathbf{p}_2 = \mathbf{p}_2 \lor \mathbf{p}_1; \)
2. \( (\mathbf{p}_1 \lor \mathbf{p}_2) \lor \mathbf{p}_3 = \mathbf{p}_1 \lor (\mathbf{p}_2 \lor \mathbf{p}_3). \)

Hence, we can simply denote the majorization join of \( m \) probability distributions \( \mathbf{p}_i, \ i = 1, 2, \ldots, m \) as \( \mathbf{p}_1 \lor \mathbf{p}_2 \lor \cdots \lor \mathbf{p}_m \) or \( \lor_{i=1}^{m} \mathbf{p}_i \) and the majorization join of all probability distributions in a set \( \mathcal{X} \) (finite or infinite) as \( \lor_{\mathbf{p} \in \mathcal{X}} \mathbf{p} \). Similar notations can also be used for majorization meet.

The basic process for the construction of the majorization meet \( \land_{\mathbf{p} \in \mathcal{X}} \mathbf{p} \) is quite simple. The necessary and sufficient condition for \( \mathbf{c} \prec \mathbf{p} \) for all \( \mathbf{p} \in \mathcal{X} \) is \( \sum_{i=1}^{k} c_i \leq \inf_{\mathbf{p} \in \mathcal{X}} \sum_{i=1}^{k} p_i \). Hence if \( \sum_{i=1}^{k} c_i = \inf_{\mathbf{p} \in \mathcal{X}} \sum_{i=1}^{k} p_i \) and \( c_k \) are in descending order, then \( \mathbf{c} \prec \mathbf{p} \) and \( \mathbf{c} \succeq \mathbf{c} \), i.e., \( \mathbf{c} = \land_{\mathbf{p} \in \mathcal{X}} \mathbf{p} \). In the case of majorization meet, this is always possible. Just let

\[
\mathbf{c}_k = s_k - s_{k-1},
\]

for \( k = 1, 2, \ldots, n \), where

\[
s_0 = 0 \text{ and } s_k = \inf_{\mathbf{p} \in \mathcal{X}} \sum_{i=1}^{k} p_i,
\]

Then for any \( \mathbf{p} \in \mathcal{X} \), we have \( \sum_{i=1}^{k} \mathbf{p}_i + \sum_{i=1}^{k-1} \mathbf{p}_i \leq 2 \sum_{i=1}^{k} \mathbf{p}_i \), as \( \mathbf{p}_k \) are in descending order. This implies that \( \inf_{\mathbf{p} \in \mathcal{X}} \sum_{i=1}^{k} \mathbf{p}_i + \inf_{\mathbf{p} \in \mathcal{X}} \sum_{i=1}^{k-1} \mathbf{p}_i \leq 2 \inf_{\mathbf{p} \in \mathcal{X}} \sum_{i=1}^{k} \mathbf{p}_i \), i.e., \( s_{k+1} + s_{k-1} \leq 2 s_k \). Thus \( c_{k+1} = s_{k+1} - s_k \leq s_k - s_{k-1} = c_k \), i.e., components of \( \mathbf{c} \) are in descending order.

For the majorization join \( \lor_{\mathbf{p} \in \mathcal{X}} \mathbf{p} \), the construction is a little bit more complicated. The probability distribution obtained from Eq. (10) with

\[
s_0 = 0 \text{ and } s_k = \sup_{\mathbf{p} \in \mathcal{X}} \sum_{i=1}^{k} p_i,
\]

may not be in decreasing order. Some further flattening operations may be needed as shown in the following algorithm (Steps 2-7). The main idea is that, for all \( k = 1, 2, \ldots, n \), the flattening operation never decreases \( c_k \) and always preserves the relation

\[
\sum_{i=1}^{k} c_i \leq \sum_{i=1}^{k} \bar{c}_i,
\]

where \( \bar{c} \) is any probability distribution such that Eq. (13) holds initially, i.e., \( \sup_{\mathbf{p} \in \mathcal{X}} \sum_{i=1}^{k} \mathbf{p}_i \leq \sum_{i=1}^{k} \bar{c}_i \). This property of the flattening operation can be easily checked with some basic calculations [46].

**Algorithm: Majorization join**

1. Let \( s_0 = 0 \) and \( s_k = \sup_{\mathbf{p} \in \mathcal{X}} \sum_{i=1}^{k} p_i \), for \( k = 1, 2, \ldots, n \);
2. Let \( c_k = s_k - s_{k-1} \), for \( k = 1, 2, \ldots, n \);
3. for \( k = 3, \ldots, n \), do
4. \( \text{if } c_k > c_{k-1}, \text{ then} \)
5. \( \text{Find the largest } l < k \text{ such that } \frac{1}{k-l} \sum_{i=l}^{k} c_i \leq c_l; \)
6. \( \text{Update each of } c_{l+1}, c_{l+2}, \ldots, c_k \text{ to } \frac{1}{k-l} \sum_{i=l}^{k} c_i; \)
7. \( \text{end if} \)
8. \( \text{end for} \)

**Examples:**
1. Consider the probability distributions \( p = \left( \frac{2}{5}, \frac{1}{5}, \frac{1}{5} \right) \) and \( d = \left( \frac{1}{2}, \frac{1}{2}, 0 \right) \), which is studied in the main text. Then \( s = \left( \frac{2}{3}, 1, 1 \right) \) in Step 1 and \( c = \left( \frac{2}{3}, \frac{1}{3}, 0 \right) \) in Step 2. In this case we do not need to do the flattening operation, because \( c_3 \leq c_2 \). Hence, we get \( p \vee d = \left( \frac{2}{3}, 1, 0 \right) \).

2. Let \( a = \left( \frac{2}{3}, \frac{1}{3}, \frac{1}{5} \right) \) and \( b = \left( \frac{1}{2}, \frac{1}{2}, \frac{1}{5} \right) \), then \( s = \left( \frac{2}{3}, \frac{1}{3}, \frac{1}{5}, 0 \right) \). In Step 2, we get that \( c = \left( \frac{2}{5}, \frac{1}{5}, 1, 0 \right) \). In this case, we have \( a \leq c \). Hence, we need to find the largest \( l < 2 \) such that \( \sum_{k=1}^{l} c_k \leq c_{l-1} \), which gives \( l = 2 \). Then update both \( c_2 \) and \( c_1 \) to \( \frac{1}{2}(c_2 + c_3) = \frac{1}{5} \), which gives \( c = \left( \frac{2}{5}, \frac{1}{5}, 1, \frac{1}{5} \right) \). No further operation is needed, because \( c_4 \leq c_3 \). Hence, the final result is \( a \wedge b = \left( \frac{2}{5}, \frac{1}{5}, 1, 0 \right) \).

Appendix B: Tightness of the bound in Proposition 2

To show the tightness of the bound in Proposition 2, we only need to prove that there is a quantum state \( \rho \) such that: (1) \( d \) is its diagonal part; (2) \( p \) is the probability distribution of outcomes of some von Neumann measurement \( P \); (3) \( d \wedge p \) is its spectrum. Suppose \( d \wedge p = (c_1, c_2, \ldots, c_n) \) and let \( \tilde{\rho} = \sum_{i=1}^{n} c_i |i\rangle\langle i| \). According to Lemma 1, there are two bases \( \{|\psi_i\rangle\}_{i=1}^{n} \) and \( \{|\varphi_i\rangle\}_{i=1}^{n} \) such that \( \langle \psi_i | \tilde{\rho} | \varphi_i \rangle = d_i \) and \( \langle \psi_i | \tilde{\rho} | \psi_i \rangle = p_i \), as \( d < d \wedge p \) and \( p < d \wedge p \). Then we can choose the state \( \rho = U \tilde{\rho} U^\dagger \) and the von Neumann measurement \( P = \{U | \psi_i \rangle \langle \psi_i | U^\dagger \}^{n}_{i=1} \), where \( U = \sum_{i=1}^{n} |i\rangle \langle \varphi_i| \). We can easily verify that \( \rho \) and \( P \) fulfill the three conditions above, and hence the lower bound in Eq. (3) is tight.

Appendix C: Majorization join and meet over linear constraints

In this appendix, we consider a special case of majorization join and meet of infinite number of probability distributions, the majorization join and meet over linear constraints.

We first consider the case of majorization meet, i.e., \( \bigwedge_{p \in X} p \), where \( X = \{p \mid Ap \geq \alpha, Bp = \beta \} \). Here \( A \) and \( B \) are matrices, \( \alpha \) and \( \beta \) are vectors, and \( \geq \) denotes the component-wise comparison. Note that the components of \( p \) may not be in descending order in this case. The conditions that \( p \) is a probability distribution, i.e., \( p \geq 0 \) and \( 1^T p = 1 \), where \( 0 \) and \( 1 \) are vectors with all components being 0 and 1 respectively, are already included in the constraints. According to Appendix A, \( \bigwedge_{p \in X} p \equiv (p_1, p_2, \ldots, p_n) \) is given by \( \tilde{p}_k = s_k - s_{k-1} \), where \( s_0 = 0 \) and \( s_k \) can be written as the following convex optimization problem,

\[
\begin{align*}
\text{minimize} & \quad \sum_{i=1}^{k} p_i^2 \\
\text{subject to} & \quad Ap \geq \alpha \\
& \quad Bp = \beta,
\end{align*}
\]

for \( k = 1, 2, \ldots, n \). As the components of \( p \) in \( X \) may not be in descending order, the “\( \geq \)” is necessary for \( p_k \) in Eq. (14).

Now, we prove that the convex optimization is equivalent to the following linear program,

\[
\begin{align*}
\text{maximize} & \quad \alpha^T \mu + \beta^T \nu \\
\text{subject to} & \quad 0 \leq A^T \mu + B^T \nu \leq 1 \\
& \quad 1^T A^T \mu + 1^T B^T \nu = k \\
& \quad \mu \geq 0.
\end{align*}
\]

where \( \lambda \) and \( \nu \) are vectors whose dimensions dependent on the numbers of inequality and equality constraints in Eq. (14). Let \( Y_k = \{y \mid 0 \leq y \leq 1, 1^T y = k \} \), then \( \sum_{i=1}^{k} p_i^2 = \max_{y \in Y_k} y^T p \). Thus we get \( \min_{p \in X} \sum_{i=1}^{k} p_i^2 = \min_{p \in X} \max_{y \in Y_k} y^T p \). As both \( X \) and \( Y_k \) are compact and convex, and \( y^T p \) is a continuous convex-concave (actually bilinear) function on \( (p, y) \), the von Neumann minimax theorem implies that we can exchange optimizations min and max, i.e.,

\[
\begin{align*}
\min_{p \in X} \sum_{i=1}^{k} p_i^2 &= \max_{y \in Y_k} \min_{p \in X} y^T p = \min_{y \in Y_k} \max_{p \in X} y^T p \tag{16}
\end{align*}
\]

Now we consider the first optimization \( \min_{p \in X} y^T p \), i.e.,

\[
\begin{align*}
\text{minimize} & \quad y^T p \\
\text{subject to} & \quad Ap \geq \alpha \\
& \quad Bp = \beta \tag{17}
\end{align*}
\]

which is a linear program and hence the strong duality always holds. Thus \( \min_{p \in X} y^T p \) equals to

\[
\begin{align*}
\text{maximize} & \quad \alpha^T \mu + \beta^T \nu \\
\text{subject to} & \quad A^T \mu + B^T \nu = y \\
& \quad \mu \geq 0 \tag{18}
\end{align*}
\]

Combining Eqs. (16) and (18), we get the final linear program in Eq. (15).

At last, we briefly discuss how to calculate the majorization join over linear constraints i.e., \( \vee_{p \in X} p \), where \( X = \{p \mid Ap \geq \alpha, Bp = \beta \} \). According to Appendix A, the main process is to maximize \( \sum_{i=1}^{k} p_i^2 \) over \( X \). As \( \sum_{i=1}^{k} p_i^2 \) is a convex function on \( p \), the maximization is always achieved on the extreme points. As all the constraints are linear, finding the extreme points is equivalent to finding all the vertices of the polytope expressed by \( X \), for which efficient algorithms exist [57]. Thus, we transform the majorization join over linear constraints to the majorization join of finite probability distributions, which can be directly solved by the method shown in Appendix A.

Appendix D: Characterizing the freezing of coherence

In this appendix, we discuss some details about the model for characterizing the freezing of coherence. The channel we consider is the local bit flip channel \( \Lambda^{bn} \), where \( A \) satisfy the Lindblad equation,

\[
\frac{d}{dt} \Lambda(\rho) = \gamma \left( (\sigma_x \rho \sigma_x - \rho) \right),
\]

(19)
and $\gamma$ is a parameter that represents the strength of the noise. The solution of Eq. (19) is given by

$$\Lambda(\rho) = \frac{1}{2} (1 + e^{-\gamma}) \rho + \frac{1}{2} (1 - e^{-\gamma}) \sigma_x \rho \sigma_x.$$  (20)

The initial state we consider is the $N$-qubit GHZ state, where the preparation of the initial state is inevitably affected by noise in experiments. We consider the two most common noise in experiments, the dephasing noise $\Delta_\epsilon$ and the depolarizing noise $D_{\epsilon}$,

Dephasing: \[ \Delta_\epsilon(\rho) = (1 - \epsilon)\rho + \epsilon \rho_d. \]

Depolarizing: \[ D_\epsilon(\rho) = (1 - \epsilon)\rho + \epsilon \frac{1}{2} I. \]  (21)

For the $N$-qubit GHZ state, the initial states that affected by dephasing noise is given by

$$\rho_\Delta^0 = \frac{1}{2} (1 + (1 - \epsilon)^N) |\varphi_{0...0}^+\rangle \langle \varphi_{0...0}^+| + \frac{1}{2} (1 - (1 - \epsilon)^N) |\varphi_{0...0}^-\rangle \langle \varphi_{0...0}^-|,$$

and the initial states that affected by depolarized noise is given by

$$\rho_D^0 = \frac{1}{2} (1 - \epsilon) (|\varphi_{0...0}^+\rangle \langle \varphi_{0...0}^+| + (1 - \epsilon)^N |\varphi_{0...0}^0\rangle \langle \varphi_{0...0}^0|)$$

$$+ \frac{1}{2} (1 - \epsilon) (|\varphi_{0...0}^-\rangle \langle \varphi_{0...0}^-| + (1 - \epsilon)^N |\varphi_{0...0}^-\rangle \langle \varphi_{0...0}^-|)$$

$$+ \sum_{w(l) \neq 0, N} \frac{1}{2} (1 - \epsilon) w(l) |\varphi_{N-w(l)}^0\rangle \langle \varphi_{N-w(l)}^0| + \frac{1}{2} w(l) (1 - \epsilon)^N |\varphi_{N-w(l)}^0\rangle \langle \varphi_{N-w(l)}^0|),$$

where $w(l)$ is the Hamming weight of the binary string $l = l_1 l_2 ... l_N$, and

$$|\varphi_{0...0}^\pm\rangle = \frac{1}{\sqrt{2}} (|0...0\rangle \pm |1...1\rangle).$$  (24)

The fidelities of the initial states the affected by depolarization and depolarized noise are given by

$$F^\Delta = \frac{1}{2} (1 + (1 - \epsilon)^N),$$

$$F^D = \frac{1}{2} ((1 - \epsilon)^N + (1 - \epsilon)^N),$$  (25)

respectively.
[38] G. Smith, J. A. Smolin, X. Yuan, Q. Zhao, D. Girolami, and X. Ma, arXiv:1707.09928.
[39] D.-J. Zhang, C. L. Liu, X.-D. Yu, and D. M. Tong, Phys. Rev. Lett. 120, 170501 (2018).
[40] F. G. S. L. Brandão and G. Gour, Phys. Rev. Lett. 115, 070503 (2015).
[41] B. Coecke, T. Fritz, and R. W. Spekkens, Inf. Comput. 250, 59 (2016).
[42] Y. Peng, Y. Jiang, and H. Fan, Phys. Rev. A 93, 032326 (2016).
[43] I. Bengtsson and K. Życzkowski, Geometry of quantum states: an introduction to quantum entanglement (Cambridge University Press, Cambridge, UK, 2017).
[44] M. Horodecki and J. Oppenheim, Nat. Commun. 4 (2013), 10.1038/ncomms3059.
[45] A. W. Marshall, I. Olkin, and B. C. Arnold, Inequalities: Theory of Majorization and its Applications, 2nd ed., Vol. 143 (Springer, New York, USA, 2011).
[46] F. Cicalese and U. Vaccaro, IEEE Trans. Inf. Theory 48, 933 (2002).
[47] O. Gühne and G. Tóth, Physics Reports 474, 1 (2009).
[48] G. Tóth and O. Gühne, Phys. Rev. A 72, 022340 (2005).
[49] S. Boyd and L. Vandenberghe, Convex optimization (Cambridge University Press, 2004).
[50] S. Diamond and S. Boyd, J. Mach. Learn. Res. 17, 1 (2016).
[51] M. A. Nielsen and J. Kempe, Phys. Rev. Lett. 86, 5184 (2001).
[52] T. Hiroshima, Phys. Rev. Lett. 91, 057902 (2003).
[53] R. Augusiak and J. Stasinska, New J. Phys. 11, 053018 (2009).
[54] M. A. Nielsen, Phys. Rev. Lett. 83, 436 (1999).
[55] S. Du, Z. Bai, and Y. Guo, Phys. Rev. A 91, 052120 (2015).
[56] H. Zhu, Z. Ma, Z. Cao, S.-M. Fei, and V. Vedral, Phys. Rev. A 96, 032316 (2017).
[57] D. Avis and K. Fukuda, Discrete Comput. Geom. 8, 295 (1992).