Optimal quantum states for frequency estimation

F Fröwis, M Skotiniotis, B Kraus and W Dür

1 Institut für Theoretische Physik, Universität Innsbruck, Technikerstraße 25, A-6020 Innsbruck, Austria
2 Group of Applied Physics, University of Geneva, CH-1211 Geneva 4, Switzerland
E-mail: florian.froewis@unige.ch

Received 10 March 2014, revised 19 May 2014
Accepted for publication 6 June 2014
Published 5 August 2014

Abstract
We investigate different quantum parameter estimation scenarios in the presence of noise, and identify optimal probe states. For frequency estimation of local Hamiltonians with dephasing noise, we determine optimal probe states for up to 70 qubits, and determine their key properties. We find that the so-called one-axis twisted spin-squeezed states are only almost optimal, and that optimal states need not to be spin-squeezed. For different kinds of noise models, we investigate whether optimal states in the noiseless case remain superior to product states also in the presence of noise. For certain spatially and temporally correlated noise, we find that product states no longer allow one to reach the standard quantum limit in precision, while certain entangled states do. Our conclusions are based on numerical evidence using efficient numerical algorithms which we developed in order to treat permutational invariant systems.

Keywords: quantum metrology, frequency estimation, qubits systems, dephasing and depolarising noise

1. Introduction

How can one determine a quantity or parameter with high precision? This is one of the central questions in physics throughout all fields. In addition to the numerous practical problems in this quest, one faces some fundamental limits, for example, set by measurement statistics.
Metrology as the science of measurement aims at identifying these limits, and at developing optimal schemes for estimating an unknown parameter in a given experimental setup. In the seminal paper by Caves [1] these questions were tackled within the quantum mechanical framework. It was found that quantum metrology offers a significant advantage as compared to classical strategies, where a quadratic improvement in achievable precision is obtained owing to the use of quantum entanglement. Nowadays, these insights find widespread applications in interferometry [2, 3], atomic clocks [4, 5], gravitational wave detectors [6, 7] and frequency estimation [8, 9].

In this paper, we consider frequency estimation where the strength, \( \omega \), of a local Hamiltonian, \( H \), describing the interaction of \( N \) qubits with an external field in \( z \) direction should be estimated. In the noiseless case one finds that for classical strategies the achievable precision, \( \delta \omega \), that measures the statistical deviation of the estimated parameter from the actual value, is given by the standard quantum limit (SQL), \( \delta \omega \geq O(1/\sqrt{N}) \) where \( N \) is the number of probe systems. In turn, entangled states allow one to achieve Heisenberg scaling where \( \delta \omega = O(1/N) \). Optimal states are readily identified as a coherent superposition of eigenstates of \( H \) corresponding to the minimal and maximal eigenvalue. The optimal state corresponds to the GHZ state (after Greenberger, Horne and Zeilinger [10]).

In the presence of noise and imperfections the situation changes drastically. For dephasing noise (what we will refer to as 'standard scenario' in the following), one finds that GHZ states do not offer any advantage as compared to classical strategies with noise [11]. Moreover, no state can reach the quadratic gain in precision and at most a constant gain factor can be achieved. This is expressed in terms of general bounds [12, 13], that show for generic kinds of noise the impossibility to reach Heisenberg scaling\(^3\). In the standard scenario so-called one-axis twisted spin-squeezed states (SSS) [21] were identified to reach the optimal constant gain factor [22] in the limit of large \( N \), and it is widely believed that such states are optimal even for finite \( N \).

However, the situation is far from being fully understood. On the one hand, optimal states for finite \( N \) in the standard scenario are unknown, and it is not known whether states other than the SSS are optimal in the asymptotic case. On the other hand, beyond the standard scenario where one considers different kinds of noise models (e.g., depolarizing rather than dephasing noise), or different Hamiltonians [23], basically nothing is known about optimal states be it for finite \( N \) or in the asymptotic limit (an exception is transversal noise considered in [14]).

In this paper, we are concerned with the question of identifying optimal states for frequency estimation in different noisy metrology scenarios. Our central results for the standard scenario are as follows:

- We determine optimal states for finite \( N \). To this aim we develop numerical tools to treat permutational invariant states that allow us to investigate systems of up to 70 qubits.
- We find that spin-squeezing is not necessary for optimality. For finite \( N \) one-axis twisted SSS are in fact only almost optimal.

\(^3\) Note that, for purely transversal noise and in the case of certain correlated noise, a scaling better than the SQL can be achieved [14–17]. In some cases, even Heisenberg scaling can be achieved using quantum error correction [18–20], or by other means [17].
• We identify a key feature for optimality which is a specific global distribution of coefficients of the state in the eigenbasis of $H$. This also implies a certain value of the variance of $H$.

Beyond the standard scenario, where we consider local depolarizing noise or spatially and temporally correlated noise, we find:

• Contrary to the standard scenario GHZ states may remain superior to product states also in the noisy case.

• For spatially and temporally correlated noise, $N$ qubit product states do not even reach the SQL, whereas GHZ states do. This opens again a gap in the scaling. In addition, the equivalence between parallel and sequential strategies [24] (i.e., to either use one particle $N$ times, or $N$ particles in parallel) does not hold in this case.

The paper is organized as follows. In section 2, we provide relevant background information, and describe the different scenarios we consider. There, we also discuss the ansatz space of qubit states we consider and outline the numerical method we develop. In section 3, we consider the standard scenario with local Hamiltonian and dephasing noise, and determine optimal states and their features. In section 4, we present results for frequency estimation scenarios with different kinds of noise models. We summarize and conclude in section 5. Some technical details, in particular regarding the numerical algorithm we develop here, as well as additional analytical results are presented in the appendices.

2. Preliminaries

In this section we briefly review the quantum metrology scenarios considered in this work, as well as the main theoretical tools behind our numerical routines. Specifically, in section 2.1 we review the main results of classical and quantum metrology and provide the mathematical descriptions for all the models we investigate. section 2.2 introduces the ansatz state space used in our numerical optimization routines, and section 2.3 provides a brief overview of the mathematical tools on which our numerical optimization is based on.

2.1. Dynamical evolution models for quantum frequency estimation

In a metrological scenario the goal is to estimate a parameter, $\omega \in \mathbb{R}$, from a probability distribution, $p(x|\omega)$, where $x \in \mathbb{R}$ denotes the outcomes of a suitable measurement. An unbiased estimate, $\hat{\omega}$, of $\omega$ is obtained by suitable post-processing of the measurement outcomes. The variance, $\delta \omega^2 = \langle (\omega - \hat{\omega})^2 \rangle$, of an unbiased estimator is lower-bounded via the well-known Cramér–Rao inequality [25]

$$\delta \omega^2 \geq \frac{1}{nF},$$

(1)

$^4$ An estimate is unbiased if its expected value, $\langle \hat{\omega} \rangle$, with respect to the probability distribution $p(x|\omega)$ is equal to $\omega$. 
where \( F \) is the Fisher information given by \([26]\)

\[
F = \int \left( \frac{\partial \ln p(x|\omega)}{\partial \omega} \right)^2 \text{d}x,
\]

and \( n \) is the number of measurements. It is known that the lower bound in equation (1) can be achieved in the limit \( n \to \infty \) by the maximum likelihood estimator \([27]\).

Assume that the parameter of interest is a quantity governing the evolution of a physical system, such as the frequency of rotation of a spin about a given axis. Then, the scenario described above is implemented by preparing the system in some initial known state, \( \rho \), and allowing it to evolve under the requisite dynamics for some time before measuring its final state, \( \rho(\omega, t) \). This process is repeated \( n \) times to obtain the measurement statistics, \( \omega\left|\rho_x\right| \), from which an estimate \( \omega^\hat{} \) can be extracted. If the physical system is quantum mechanical in nature, i.e., \( \rho(\omega, t) \in \mathbb{B}(\mathcal{H}) \), then the probability distribution, \( p(x|\omega) \), is given by \( p(x|\omega) = \text{Tr}(M_x \rho(\omega, t)) \), where the set of measurement operators \( \{M_x: \mathbb{B}(\mathcal{H}) \to \mathbb{B}(\mathcal{H})\} \) satisfies \( \sum_x M_x^\dagger M_x = I \).

As any set of measurement operators, \( \{M_x\} \), constitute an admissible measurement it is natural to ask which measurement minimizes equation (1) or, equivalently, maximizes the Fisher information. The quantum Fisher information (QFI), \( \mathcal{F} \), is defined as the maximal Fisher information over all allowable measurements, and is given by \([28-30]\)

\[
\mathcal{F} = \text{Tr}[\dot{\rho} L_{\omega}],
\]

where

\[
L_{\omega} = 2 \sum_{i,j} \frac{\langle \psi_i | \dot{\rho}(\omega, t) | \psi_j \rangle}{\lambda_i + \lambda_j} | \psi_i \rangle \langle \psi_j |
\]

is known as the symmetric logarithmic derivative. Here, \( \dot{\rho}(\omega, t) = \partial \rho(\omega, t)/\partial \omega \), \( \lambda_i \) are the eigenvalues of \( \rho(\omega, t) \), \( | \psi_i \rangle \) the corresponding eigenvectors, and the sum in equation (4) is over all \( i, j \) satisfying \( \lambda_i + \lambda_j \neq 0 \). The most informative measurement is the one whose elements are the projectors on the eigenspaces of the symmetric logarithmic derivative. Substituting equation (3) in place of the Fisher information in equation (1) yields the quantum Cramér–Rao inequality which provides the ultimate lower bound on precision achievable by a quantum mechanical strategy.

As the QFI already incorporates the optimization over all measurements, it remains to minimize the uncertainty in \( \omega \) with respect to all other available resources. In this work we focus solely on estimating the frequency of rotation of a spin around a given axis. For this task the two relevant resources are the number of probe systems, \( N \), in a given run of the experiment and the total running time, \( T = nt \), of the experiment. Here, \( t \) denotes the interrogation time, the time interval the \( N \) probes are subjected to the \( \omega \)-dependent dynamical evolution before they are measured. Note that in order to ensure that the Cramér–Rao bound is achievable, we require \( n \) in equation (1) to be very large which implies that \( T \) needs to be much larger than any other time scale. With respect to these resources the quantum Cramér–Rao bound for frequency estimation reads \([11]\)
\[ \delta \omega^2 \geq \frac{1}{T F(\rho, \omega, t)^2}, \]  

where \( F(\rho, \omega, t) \) denotes the QFI of the final state of the \( N \) probes. The interrogation time, \( t \), is a controllable parameter which needs to be optimized in order to maximize the precision in frequency estimation. Henceforth, for a given initial state, \( |\psi\rangle \), the maximal QFI is defined as \( \tilde{\mathcal{F}} := \max_t T F(\rho, \omega, t)/t \).

The dynamical evolution of the \( N \) probe systems is described by a master equation of the Lindblad form

\[ \frac{d\rho}{dt} = -i\omega [h, \rho] + \mathcal{L}[\rho], \]  

where \( \omega \) is the frequency we are interested in estimating, \( \rho \) is the initial state of the \( N \) probes, and \( H = \omega h \) and \( \mathcal{L} \) are the Hamiltonian and Lindblad operators generating the unitary and non-unitary (noisy) part of the evolution respectively. In this work we consider the local generator

\[ h = S_z := \frac{1}{2} \sum_{i=1}^{N} \sigma_z^{(i)}. \]  

We note that for local Hamiltonians, and in the absence of noise, the optimal precision in estimating frequency using an initially pure product state scales at the SQL, \( \delta \omega^2 = O(N^{-4}) \). Contradistinctively, if the \( N \) probes are initially prepared in a GHZ state, i.e. an equal superposition of the eigenstates corresponding to the maximum and minimum eigenvalue of \( h \), then the precision in estimation scales at the Heisenberg limit, \( \delta \omega^2 = O(N^{-2}) \) [31].

The noise models we consider here are of two main types. The first type of noise we consider is local, uncorrelated noise described by the Lindblad operator \( \mathcal{L}[\rho] = \sum_{j=1}^{N} \mathcal{L}_j[\rho] \) with

\[ \mathcal{L}_j[\rho] = \frac{\gamma}{2} \left( -\rho + \mu_1 \sigma_z^{(j)} \rho \sigma_z^{(j)} + \mu_2 \sigma_x^{(j)} \rho \sigma_x^{(j)} + \mu_3 \sigma_y^{(j)} \rho \sigma_y^{(j)} \right), \]  

where \( \gamma \) denotes the strength of the noise and \( \sum \mu_j = 1 \). The case \( \mu_1 = \mu_2 = 0, \mu_3 = 1 \) corresponds to the case of local, uncorrelated dephasing noise and is the main type of noise considered in this work (section 3), whereas the case \( \mu_1 = \mu_3 = \mu_2 = 1/3 \) corresponds to local uncorrelated, depolarizing noise (section 4.1).

The second type of noise we consider is correlated dephasing noise, where correlations are both in space and time (section 4.2), with Lindblad operator given by

\[ \mathcal{L}[\rho] \equiv -\gamma f(t) \left[ \sigma_z, \left[ \sigma_z, \rho \right] \right]. \]  

Here, \( f(t) = 1 - \exp(-\gamma t) \) denotes the temporal profile of the noise, whereas the double-commutator in equation (9) governs the spatial correlations. Such type of noise is physically relevant in ion trap setups due to fluctuations in the phase reference beams addressing all \( N \) ions collectively, i.e., with infinite spatial correlation length, but finite memory [32–35]. It has been demonstrated that for local Hamiltonians temporal correlations alone allow for subSQL precision in optical interferometry [36], whereas spatially correlated noise alone, with finite or infinite correlation length, can even allow for Heisenberg scaling in precision using a suitably chosen higher dimensional entangled state [17] or considering a different Hamiltonian [37].

A key feature for all scenarios considered in this paper is that the unitary and noisy part in equation (6) commute. This means that the solution for the time-dependent density operator
reads

$$\rho (\omega, t) = e^{-i\omega t}\mathcal{E}[|\psi\rangle\langle\psi|]e^{i\omega t},$$

where $\mathcal{E}[\rho] = \exp(\mathcal{L}t)[\rho]$ denotes the noise channel and $|\psi\rangle$ the initial state. Writing $\mathcal{E}[|\psi\rangle\langle\psi|]$ in its spectral decomposition, $\mathcal{E}[|\psi\rangle\langle\psi|] = \sum\lambda_i |\psi_i\rangle\langle\psi_i|$, the QFI can be expressed as [28, 30]

$$\mathcal{F} = 4t^2 \sum_{i<j} \left( \frac{\lambda_i - \lambda_j}{\lambda_i + \lambda_j} \right)^2 \left| \langle \psi_i | h | \psi_j \rangle \right|^2.$$

Note that $\mathcal{F}$ is here independent of $\omega$. For the case of pure states and in the absence of noise it can be easily verified that $\mathcal{F} = 4t^2 V(h) := 4t^2 \left( \langle h^2 \rangle - \langle h \rangle^2 \right)$. However, even if we start with an initially pure state, the evolution of such a state under the full dynamics of equation (6) will in general yield a mixed state. We emphasize that maximizing $\mathcal{F}/t$ in time reduces the role of $\gamma$ to a proportionality constant, that is, $\mathcal{F}_\gamma = \mathcal{F}_\gamma |_{\gamma=t}$. Therefore, the findings presented in this paper are independent of a specific choice of the value of $\gamma$, which eases also the numerical effort. This is in contrast to the so-called phase estimation scenario, where the interrogation time is fixed. Then, there are typically two regimes, $N \ll \gamma$ and $N \gg \gamma$, where one can observe qualitatively different behaviour for the optimal initial states [12, 13].

### 2.2. Ansatz space

Our goal is to numerically determine the initial states leading to maximal sensitivity for the case of noisy frequency estimation and to identify the key properties of such states. Specifically, we are interested in determining the states that maximize the QFI. As the latter is convex [38] it suffices to consider only pure states of $N$ qubits. However, numerically searching for arbitrary pure states that maximize the QFI seems unfeasible for a large number of qubits. Thus, we restrict ourselves to a subspace of the total Hilbert space in order to reduce computational effort.

In order to pick the most suitable subspace we note that the dynamical evolutions we consider are symmetric under particle exchange. This observation has motivated several authors to consider pure initial states that are symmetric under particle permutations [11, 37, 39], i.e. states of the form

$$|\psi\rangle = \sum_{m_0=-j_{\text{max}}}^{j_{\text{max}}} c_m |j_{\text{max}}, m\rangle.$$

The states $|j_{\text{max}}, m\rangle$ are simultaneous eigenstates of the total angular momentum operator, $\hat{S}^2$, and its projection onto the $z$-axis, $\hat{S}_z$, with corresponding eigenvalues $j_{\text{max}}(j_{\text{max}} + 1) = N/2 (N/2 + 1)$ and $m$ respectively, and are the so-called Dicke states with $N/2 + m$ excitations [40]. Note that we define $|0\rangle$ as the excited state, so that for instance the Dicke state of 3 qubits, one of which is in the $|0\rangle$ state, is $|3/2, -1/2\rangle = 1/\sqrt{3} (|011\rangle + |101\rangle + |110\rangle)$.

We can reduce the number of parameters in equation (12) further by requiring that the coefficients $c_m$ are real and positive, as any complex phase can be taken into account by applying a phase gate that commutes with the dynamics (equation (6)). In addition, the dynamics are invariant under collective spin flips $\sigma_{z}^{\otimes N}$. Requiring the same symmetry for the states leads to $c_m = c_{-m}$. 

New J. Phys. 16 (2014) 083010 F Fröwis et al
With the exception of correlated noise (see equation (9)) [37], we are not aware of any proof that the state that maximizes the QFI in the presence of local uncorrelated noise belongs to our ansatz space. However, by extensive numerical studies for \(N = 2, 3\), and by comparing specific examples of asymmetric states and their symmetrized counterparts, it seems that the optimal state must exhibit the same symmetry as the dynamics. With all these considerations taken into account we choose our ansatz space as the space spanned by the states of equation (12) with \(c_m \in \mathbb{R}, \ c_m > 0\), and \(c_m = c_{-m}\). This subspace is denoted by \(S_N\) in the following. Note that with these restrictions, states in equation (12) point in \(x\) direction, i.e., with \(\sigma_{x} = \sum_{m=-j_{\max}}^{j_{\max}} \sqrt{\left(\frac{N}{2} + m\right)/2^N} | j_{\max}, m \rangle\). The second family of states we consider is the GHZ state

\[
|\text{GHZ} \rangle = \frac{1}{\sqrt{2}} (| N/2, -N/2 \rangle + | N/2, N/2 \rangle).
\]

As mentioned in section 2.1 these states achieve Heisenberg scaling in precision for noiseless frequency estimation using the Hamiltonian in equation (7).

The third family of states we consider are the one-axis twisted SSS [21]

\[
|\text{SSS} (\mu) \rangle = e^{-i S_x} e^{-i S_z} |\text{PS} \rangle.
\]

This state family is defined by two parameters: the squeezing parameter, \(\mu\), and a local rotation parameter \(\nu\) which serves to re-orient the squeezing axis so that the benefit for the specific Hamiltonian is optimized. As the value of \(\nu\) depends only on \(\mu\), the one-axis twisted SSS are essentially a one-parameter family. It was shown that for local Hamiltonians and for local uncorrelated dephasing noise (see section 3) one axis-twisted SSS are asymptotically optimal [22].

The fourth family of states we consider are the Dicke states in the \(x\)-basis,

\[
|D_k \rangle = \text{Had}^\otimes N | N/2, N/2 - k \rangle,
\]

where \(\text{Had}\) denotes the Hadamard operation

\[
\text{Had} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.
\]

Note that \(|\text{SSS} (0) \rangle = |D_0 \rangle = |\text{PS} \rangle|.

In addition to the four families of states above, we conduct a numerical search for the optimal state in the entire ansatz space \(S_N\). For this task, we choose the Nelder–Mead simplex algorithm [41] which is known to be successful for low-dimensional, unconditioned search problems [42]. This algorithm takes as input a state and time and optimizes the coefficients \(c_m\) of equation (12) by maximizing \(\mathcal{F}/t\). The input state is either a specific state or a randomly chosen one. The coefficients, \(c_m\), of a random state are stochastic variables which follow a normal
distribution. After all coefficients are chosen the random state is normalized. Such a random state is denoted by \( |\psi_{\text{rand}}\rangle \).

In the next subsection we describe how the QFI can be efficiently calculated for states in our ansatz space.

2.3. Numerical methods

Evaluating the QFI (see equation (11)) is computationally hard as it requires full diagonalization of the density matrix \( \mathcal{E}[|\psi\rangle\langle\psi|] \), where \( \mathcal{E} \) denotes the noisy channel (see equation (10)). However, if the eigenvalues of \( \mathcal{E}[|\psi\rangle\langle\psi|] \) are highly degenerate, then the computational effort for calculating the QFI can be significantly reduced. A state that is symmetric under particle permutations exhibits such large degeneracies. In particular, for \( h = S_z \) and \( \mathcal{L} \) as in equations (8), (9), we are able to use a specific representation of \( \rho(t) \) and \( h \) that allows for an efficient calculation of \( \mathcal{F} \). Here, we briefly summarize this representation for the local uncorrelated dephasing scenario, equation (8), with \( \mu_x = \mu_y = 0, \mu_z = 1 \), and refer the reader to appendix A for more details.

Under the effect of local dephasing, any state in the ansatz space remains permutationally invariant. Hartmann showed in [43] that any permutationally invariant state can be represented by a weighted sum of \( O(N^3) \) specific operators. Thus, even though \( \mathcal{E}[|\psi\rangle\langle\psi|] \) may be of full rank, where \( |\psi\rangle \) is permutationally symmetric, its representation is always efficient. Moreover, the operators introduced in [43] are such that the action of local dephasing is easy to express. As shown in appendix A the spectral decomposition of \( \mathcal{E}[|\psi\rangle\langle\psi|] \) can then be efficiently computed. For an efficient computation of \( \mathcal{F} \), it is necessary that \( h \) obeys the same symmetries as the state \( \mathcal{E}[|\psi\rangle\langle\psi|] \), as \( \mathcal{F}(U\mathcal{E}[|\psi\rangle\langle\psi|]U^\dagger) = \mathcal{F}[|\psi\rangle\langle\psi|] \).

The general procedure for computing the QFI is hence the following. Express the initial state, equation (12), in terms of the ‘Hartmann operators’ introduced in [43]. Next, calculate the action of local dephasing on the basis operators. Then, express the operators in terms of joint eigenvectors of \( S^2 \) and \( S_z \), and numerically determine the spectral decomposition of \( \mathcal{E}[|\psi\rangle\langle\psi|] \) as a function of the coefficients, \( c_m \). One can then efficiently evaluate equation (11). We note that the same method can be applied for the case of depolarizing noise, even though the effect of this noise model when computed in terms of the ‘Hartmann operators’ is more cumbersome. Spatially correlated noise, on the other hand, is easier to treat since the states of equation (12) stay within the subspace spanned by \( \{|j, m\}\rangle \) under the action of the noise (see appendix A for more details).

3. Local Hamiltonian with dephasing noise

In this section, we consider the standard scenario where \( h \) is given by equation (7) and \( \mathcal{L} \) is given by equation (8) with \( \mu_x = \mu_y = 0, \mu_z = 1 \), respectively. Hence, the unitary and noisy evolutions in the standard scenario act parallel to each other if visualized on the Bloch sphere.
Let us first recall the known results pertaining to the standard scenario. It is easy to see that the product state, $|+\rangle^\otimes N$, leads to $\mathcal{F}_\gamma = N T e^{N \gamma} / 2$, which has the same scaling as in the absence of noise. The GHZ state, however, yields $\mathcal{F}_{\text{GHZ}} = \mathcal{F}_{\text{PS}}$, which is qualitatively different from the quadratic scaling in $N$ for $\gamma = 0.5$. Hence, the advantage of the GHZ state compared to the product state is lost as soon as $\gamma \neq 0$. As already mentioned in section 2.1, it is known that, asymptotically, the maximal QFI, $\mathcal{F}$, for any state scales at most linearly with the system size, and the maximal improvement, for any $|\psi\rangle$ and $N$, is bounded by $\mathcal{F}_\psi \leq N T / (2 \gamma) = e \mathcal{F}_{\text{PS}}$, where the bound can only be achieved for $N \to \infty$ [12]. Thus, the minimum error obtainable is given by $0.61 (\delta \omega)_{\text{PS}}$. It is also known that one-axis twisted SSS with a specific squeezing parameter $\mu$ (which depends on $N$ only) achieve this bound asymptotically [22].

Let us now consider the finite $N$ case. Our aim here is to numerically determine the optimal states and identify their properties. First, in section 3.1, we present numerical evidence that we are able to find the global optimal initial state, $|\psi_{\text{opt}}\rangle$, with very high accuracy. Then, we reveal that there exists a $\mu = \mu_{\text{opt}}$ such that $|\text{SSS}(\mu_{\text{opt}})\rangle$ is already almost optimal for finite system sizes. However, this implies that there exist states that perform slightly better than $|\text{SSS}(\mu_{\text{opt}})\rangle$. Next, in section 3.2, we investigate the squeezing strength and the variance $\mathcal{V}(h)$ of the optimal states. Surprisingly, our results indicate that squeezing, often considered as the key-property for high

5 There is an apparent contradiction in these statements: if $\gamma \to 0$, then $\mathcal{F}$ of any state is infinite, while for $\gamma = 0$, one often discusses the finite value (e.g., $\mathcal{F} = t^N N^2$ for the GHZ state). However, the optimal measurement time for any state is of the order of $1/\gamma$. That is, the measurement time goes to infinity for vanishing $\gamma$. For infinite measurement times and $\gamma = 0$, the QFI of any non-trivial state is infinite.
sensitivity experiments, is not necessary. In contrast, we find that for all optimal states it holds that the variance of $\hat{h}$ scales as $\mathcal{V}(\hat{h}) \propto N^{1.674}$. Finally, in section 3.3, we show that a randomly chosen state $|\psi_{\text{rand}}\rangle \in S_N$ leads to a $\mathcal{F}$ that is typically larger than $\mathcal{F}_{\text{PS}}$.

Before we go to the details, we point to figure 1(a), which intends to give the reader an overview of the relation between the maximal QFI, $\mathcal{F}$, and the variance, $\mathcal{V}(\hat{h})$, for $N = 10$. One observes that the performances $|\psi_{\text{opt}}\rangle$ and $|\psi_{\text{rand}}\rangle$ are comparable, while Dicke states are suboptimal. In addition, the state density of a random sampling indicates the distribution of $S_N$. Similar observations can be made for all $N$ considered here.

3.1. Optimal states

In this section, we present evidence that we are able to numerically determine the optimal initial states in the standard scenario.

The search for the optimal one-axis twisted SSS for a given $N$ is numerically relatively simple as one only needs to optimize over two parameters, the interrogation time, $t$, and the squeezing parameter, $\mu$. On the contrary, optimization within the ansatz class $S_N$ is much more demanding as one has to optimize over $[N/2] + 1$ parameters including time (see equation (12)). For each value of $N$, ten randomly chosen states $|\psi_{\text{rand}}\rangle \in S_N$ are chosen as input states and the optimization algorithm tries to come as close as possible to $\mathcal{F}_{\text{opt}} = \max_{\psi \in S_N} \mathcal{F}_{\psi}$. Similarly, an eleventh run is performed with $|\psi_{\text{opt}}\rangle$ as the input state for the optimization algorithm.

We first analyse the fluctuations of the results within these eleven runs for each $N$ (see figure 1(b), blue dashed line). One only finds small relative deviations from the mean result. We are thus confident that we found the global maximum within $S_N$ for $N \leq 70$ as, if the algorithm would end up in local minima, one would expect to obtain significantly different values for $\mathcal{F}$ from run to run. In the following, we refer to optimal states, $|\psi_{\text{opt}}\rangle$, as those states that are numerically found by this algorithm.

Next, we compare $\mathcal{F}_{\text{opt}}$ to $\mathcal{F}_{\text{SSS}(|\mu_{\text{opt}}\rangle)}$. For all $N \leq 70$, we found a relative improvement of approximately $10^{-5}$ to $10^{-3}$, which is significantly larger than the fluctuations of the optimization algorithm (see figure 1(b), green solid line). We therefore conclude that whereas one-axis twisted SSS are very close to the optimal states, there exist states which perform better.

3.2. Properties of optimal states

With the confidence of discussing the actual global optimal states, we can analyse their properties. Here, this is done for the squeezing strength, $\xi^2$, and for the variance $\mathcal{V}(\hat{h})$.

The squeezing strength [44, 45] measures how spin-squeezed an $N$-qubit state is. For states in the ansatz class (equation (12)), it is defined as

$$\xi^2 := \min_{\varphi} \frac{\mathcal{V}[\cos(\varphi)\hat{S}_z + \sin(\varphi)\hat{S}_y]}{N/4}. \quad (18)$$

A state is called spin-squeezed if $\xi < 1$. Let us remark that the optimal one-axis twisted SSS, $|\psi_{\text{opt}}\rangle$, is not the state with the smallest value of $\xi$.

We calculate $\xi$ for the optimal states, $|\psi_{\text{opt}}\rangle$, found by the optimization routine and observe an interesting phenomenon (see figure 2(a)). If the input state for the optimization algorithm is $|\psi_{\text{opt}}\rangle$, then the resultant state after the optimization is also spin-squeezed (squares in figure 2(a)). In contrast, optimal states found with random input states, $|\psi_{\text{rand}}\rangle$, are not...
necessarily spin-squeezed (dots in figure 2(a)). In particular, for $N > 50$, we found no such state which was spin-squeezed. We therefore conclude that, within $\mathcal{S}_N$, spin-squeezing is not a necessary property for optimal frequency estimation. We remark that we qualitatively found the same result for a more general inequality for spin squeezing along any direction [46]. This is to be expected since squeezing in any direction can be decomposed into squeezing component along the mean spin direction and a squeezing component in an orthogonal direction. As squeezing along the mean spin direction does not enhance the states performance for frequency estimation it follows that squeezing in any direction cannot provide better precision that squeezing in an orthogonal direction only.

We now turn to the second property, the variance with respect to the generator $h$. Whereas we find a large deviation of $\xi^2$ for the optimal states, one observes the contrary for $\mathcal{V}(h)$, which is the figure of merit in the noiseless case. In figure 2(b), the lower blue line represents the average variance of $h$ for the optimal states $|\psi_{\text{opt}}\rangle$. The standard deviation from this average is plotted as well for each $N$. However, the relative divergences from the average values are of the order of $10^{-6}$ to $10^{-3}$ and therefore hardly visible. Even though the random states used for the initialization of the optimization algorithm exhibit very different variances (scaling as $O(N^2)$), the optimal states show a reduced variance that is very well approximated by $\mathcal{V}(h) = 0.196N^{1.674}$. This gives strong evidence that the variance is a crucial parameter even in noisy metrology. More importantly, the variance does not scale with $N^2$, as it does for optimal states in the noiseless case.

Clearly, the variance does not fully determine the usefulness of a quantum state for noisy frequency estimation. In fact, a state that exhibits the same variance as the optimal states does

---

**Figure 2.** (a) Squeezing strength $\xi^2$ as defined in equation (18) for $|\psi_{\text{opt}}\rangle$. The blue line represents the border below which states are spin-squeezed. In case $\mathcal{S}_N(\mu_{\text{SSS}})$ was used as an initial state in the search algorithm, the resulting state, after optimization, was found to be spin-squeezed as well (squares). However, if the initialization was done with a random initial state, $|\psi_{\text{rand}}\rangle$, this is, for larger $N$, typically not the case. (b) Variance of $h$ for two sets of states is compared for different $N$. The upper, red curve is the average variance of the uniformly sampled states from the ansatz class $\mathcal{S}_N$. For each $N \in \{10, 20, \ldots, 70\}$, $10^6$ states are sampled. The average value is very well fitted by $N(N + 2)/12$, which is the variance of the uniform distribution. The lower, blue curve is the average variance from the eleven optimal states found with random starts (section 3.2). Numerically, we find that the variance is excellently fitted by $0.196N^{1.674}$. The standard deviation (bars) from the average for these states is comparatively small.
not necessarily have to have an optimal QFI. For instance, the Dicke states $|D_1\rangle$ and $|D_2\rangle$ in figure 1(a) are close to the optimal variance, yet one finds $\mathcal{F}_{\text{opt}} < \mathcal{F}_{\text{SSS opt}}$. To understand this observation, we study the distributions of the coefficients, $c_m$, of $|\mu\rangle_{\text{SSS opt}}$ and Dicke states $|D_k\rangle$ (equation (16)) where we optimize over $k$; for the optimal states $|\psi_{\text{opt}}\rangle$ (where the values of $\mathcal{F}_{\text{opt}}$ and $\mathcal{F}_{\text{SSS(\mu)}}$ are in this plot indistinguishable); and for the unconditioned as well as the conditioned random states, $|\psi_{\text{rand}}\rangle$. For $N = 70$, the maximal improvement is about $1.69 \mathcal{F}_{\text{PS}}$, which corresponds to a minimal uncertainty of about $0.76 (\Delta \omega)_{\text{PS}}$. The standard deviations of the results for the random states are indicated as shaded areas of the same colour.

In order to discover which of the two differences above has the largest impact on $\mathcal{F}$ we perform the following numerical test. We sample a large number of random states ($10^4$ for each $N \in \{10, 20, ..., 70\}$). Then, the $c_m$ of the random states are multiplied by $\cos(\theta m)$, with $\theta$ chosen such that the variance of the random state is equal to the variance of the optimal states.\footnote{In the very unlikely case where no such $\theta$ exists, the state is discarded.} Then, $\mathcal{F}$ is calculated. We find that for increasing $N$, this class of modified random states typically performs much better than the unmodified random states (see section 3.3 and figure 3(b)). What is more, the relative difference between $\mathcal{F}_{\text{opt}}$ and $\mathcal{F}$ of the modified random states decreases with $N$, in contrast to the native random states. This indicates that, for larger $N$, the detailed distribution of $c_m$ is not so important, as these modified random states are very different on the level of single coefficients. In particular, $\mathcal{F}$ apparently does not depend on a specific kind of ‘smoothness’, which is present, e.g., for $|\text{SSS(\mu opt)}\rangle$ (compare also to the sharpness quantity in [47, 48]). However, the global structure of a relatively broad distribution that is suppressed at the boundary seems to play a major role.
Finally, we numerically determined that the optimal interrogation time for all optimal states can be very well approximated as $t_{\text{opt}} \approx 0.4885N^{-0.261}$. In particular, this implies that the optimal interrogation time vanishes asymptotically, albeit much slower than for the GHZ state ($t_{\text{GHZ}} = 1/(2N)$).

### 3.3. Typical random states $|\psi_{\text{rand}}\rangle$ perform better than product states

As a side result, we remark that random states, $|\psi_{\text{rand}}\rangle$, typically perform better than product states. For each $N \in \{10, 20, ..., 70\}$, we generate $10^6$ random states as described in section 2.2, and for each state $\mathcal{F}$ and $\mathcal{W} (h)$ are calculated. We find that for larger $N$ almost all random states perform better than both product states and the GHZ state (see figure 3(b)), while their mean variance is about $N (N + 2)/6$ (see figure 2(b)). This highlights the importance of $\mathcal{S}_N$, since this result is not expected for states randomly chosen from the entire Hilbert space. It also shows that a quantum improvement is nothing extraordinary and can be achieved by many different states (within $\mathcal{S}_N$).

### 4. Beyond the standard scenario

The scenario with local Hamiltonian and dephasing noise discussed in section 3 is certainly an important instance in frequency estimation theory. However, many experimental setups have to be described by different unitary evolutions and/or different noise models. Hence, the investigation of scenarios different than the standard scenario is of fundamental and practical relevance. This section is devoted to two such scenarios. In section 4.1 we consider the situation where local dephasing noise is replaced by local depolarization noise (equation (8) with $\mu_1 = \mu_2 = \mu_3 = 1/3$), which can be viewed as a combination of local dephasing and transversal bit-flip noise, and show that this change does not lead to a dramatically different picture. In section 4.2 we consider spatial and temporal correlated dephasing noise, described by Lindblad operator given in equation (9) and show that the GHZ state remains superior to product states. In fact, a favourable scaling of the QFI can be achieved for the GHZ state.

#### 4.1. Local depolarization

In [14], it was shown that the SQL can be overcome if dephasing noise is replaced by transversal bit-flip noise. There, it was shown that $\mathcal{F}$ reaches a scaling of $O(N^{5/3})$ for the GHZ state, which is close to Heisenberg scaling. Indeed, we recently showed that the Heisenberg limit can be achieved in this scenario by using error correction techniques [18]. It is natural to ask what happens if there are deviations from this perfectly directed noise if one does not make use of error correction techniques. The answer was already given in [14], where it was shown that in the asymptotic limit one obtains again the SQL. However, it is unknown which states are optimal in such a scenario. In this section, we study the local depolarization noise, equation (8) with $\mu_1 = \mu_2 = \mu_3 = 1/3$. Physically, this model describes uncorrelated interactions between the system qubits and a bath at infinite temperature. As can be easily seen, the minimal error achievable with product states as initial states is the same as for dephasing noise, i.e., $(\delta \omega)_{PS} = \sqrt{2pe/(NT)}$. Recently, an upper bound on $\mathcal{F}$ was obtained [13], stating that $\mathcal{F}_{\text{opt}} \leq 4e/3\mathcal{F}_{PS}$. Thus the minimal uncertainty $\delta \omega$ for any initial state scales with the SQL and is at most $0.53 (\delta \omega)_{PS}$, which is smaller than for local dephasing noise. It is not known, however,
whether this bound is tight. As mentioned before, the one axis-twisted SSS were shown to asymptotically reach the (optimal) minimal error. This was done by calculating a lower bound $G \leq \bar{G}$ (see appendix B) that reaches the ultimate precision limit in the standard scenario. As the replacement of phase noise by depolarization noise does not change $G$, it follows that, asymptotically, $|\text{SSS}(\mu_{opt})|$ reaches at least the same precision with depolarization noise. To our knowledge there exists no example of a state that overcomes this limit. Therefore, it is not clear whether the bound of [13] can actually be achieved.

Let us now come to the results of our numerical studies. The algorithm to compute the QFI in case of this noise model is still efficient for symmetric ansatz states of equation (12). However, in practice it is much more demanding (see appendix A) and we are therefore limited to system sizes up to $N = 30$ (see figure 4). Whereas for $N = 2, 3$, the GHZ state is the optimal initial state, this is no longer the case for $N \geq 4$. However, the GHZ state remains superior to the product state. The QFI for the GHZ state reads (see appendix C)

$$F_{\text{GHZ}} = \frac{2^N e^{-2\gamma t^2}}{(1 + e^{-\gamma})^N + (1 - e^{-\gamma})^N}. \quad (19)$$

For large $N$, the maximization of $F/t$ over time leads to $\bar{F}_{\text{GHZ}} = 2NT/(3\gamma e) \approx 1.33 \bar{F}_{\text{PS}}$. We remark that the superior performance of the GHZ state in this scenario can be understood from the results of [49] for the case where the noise has components parallel and perpendicular to the direction of the unitary. Note that this improvement is already achieved for a relatively small $N$. The optimal states are found using the algorithm described in section 3.2. One observes that the relative difference between the optimal states and $|\text{SSS}(\mu_{opt})|$ is larger than in the standard scenario. For $N \leq 30$, it is in the range of a few percent. However, this gap seems to vanish as $N$ increases such that one could expect that SSS become optimal for larger $N$ (see figure 4).

Compared to the standard scenario of section 3, the best Dicke states are closer to the optimal states for small $N$. However, the difference in their performance as compared to the standard scenario seems to vanish for larger $N$ (see figures 3(b) and 4). Also the average improvement of the randomly chosen initial states (10^6 random states for each $N \in \{10, 15, ..., 30\}$) is compatible with that of the phase noise scenario. Altogether, these results are qualitatively very similar to those obtained in the standard scenario. It is worth
noting, however, that there are certain differences in the asymptotic case, e.g., that the GHZ state is superior to the product state in case of depolarizing noise.

4.2. Correlated dephasing

We now discuss our findings for the case of correlated dephasing noise, a dominant source of noise in experiments based on trapped ions [32–35] and on the stability of atomic clocks [50–53] that arises mainly due to fluctuations in the phase reference that collectively affect all the ions. It was shown that optimal states for the standard scenario offer no quantum advantage in this case as spatially correlated noise causes quantum coherences to vanish even faster than in the case of local dephasing noise [37]. However, clever use of decoherence-free subspaces [37], or higher-dimensional probe systems, i.e., qutrits [17], allows for higher precision compared to uncorrelated noise and can even restore Heisenberg scaling in precision. On the other hand, temporal correlations can lead to an improved performance of highly entangled states such as the GHZ state [15, 16]. The reason is that the optimal measurement time $t_d$ decreases with $N$, eventually entering regime where $\gamma t \ll 1$. Then (short), temporal correlations lead to a quadratic suppression of the noise strength.

Here, we will focus on qubits and study the effect of temporal as well as spatial correlated noise as defined in equation (9). For spatially correlated noise, the master equation reads as in equation (6). An additional (time-dependent) factor multiplying the super-operator $\mathcal{L}$ (see equation (9)) models the time-evolution assumed in [35]. For $\gamma t \gg 1$, we are in the Markovian regime $f(t) \approx 1$, while for $\gamma t \ll 1$, the effect of noise is suppressed through the temporal correlations by $f(t) \approx \gamma t$.

For the GHZ state, pure spatial correlations lead to $\mathcal{F}$ constant in $N$ [37], whereas pure temporal correlations give rise to an improved scaling of the optimal QFI of $O(N^{3/2})$ [15]. With $h = S_z$, the QFI for the GHZ states with both kinds of correlations equals

$$\mathcal{F}_{\text{GHZ}} = 2^N N^2 \exp \left[ -2 (\gamma t + e^{-\gamma t} - 1) N^2 \right].$$  \hspace{1cm} (20)

For $N \gg 1$ the optimization over time can be easily performed and leads to $\mathcal{F}_{\text{GHZ}} = NT/(\sqrt{2e} \gamma)$, i.e. the SQL.

We now compare this result to the one obtained for product states as initial states. Note that, in order to avoid the effect of the correlated noise, one might use a sequential approach, where one qubit after the other is exposed to the evolution. In this case the total time is given by $NT$, where $T$ denotes the evolution time for a single experiment. Then, $\mathcal{F}_{\text{PS}}$ coincides with the one obtained for local dephasing. The improvement of using a GHZ state is hence limited to a constant factor $\mathcal{F}_{\text{GHZ}} \approx 2.37 \mathcal{F}_{\text{PS}}$. If, however, a parallel scheme is employed, where a $N$-qubit product state is used as initial state, the effect of noise changes the scaling of $\mathcal{F}$. Indeed, numerically we find that then $\mathcal{F}_{\text{PS}} = O(N^{1/2})$, which is worse than the SQL. Hence, for a parallel scheme, we indeed encounter a quantum improvement in scaling. This is because, for the GHZ state, the time-like and space-like correlations partially compensate each other. In contrast, the product state, whose optimal measurement time is relatively long (and constant in $N$), cannot benefit from short correlations in time as much as it is impaired by spatial correlations. It is worth mentioning that, in this scenario, the equivalence between the sequential and parallel strategies for product states, discussed in [24], is not valid.

Numerically we searched for the optimal states within the ansatz class of equation (12) for $N \leq 75$ using the optimization algorithm [41] as explained in section 2.2. We find that the GHZ
states are the optimal states. We believe that this is due to the fact that the GHZ state has the shortest optimal measurement time and, for large enough $N$, its entire time evolution occurs in the regime where the strength of the noise is quadratically suppressed.

5. Summary and conclusions

In this paper, we investigated different frequency estimation scenarios mainly by numerical means.

In the scenario of local dephasing, we numerically identified optimal states within a system size up to 70 qubits. We found that one-axis twisted SSS are almost optimal for a certain squeezing parameter. However, we also gave strong evidence that the property of being spin-squeezed is not necessary by presenting optimal states that are not spin-squeezed. Moreover, we identified a key property all optimal states have in common, namely the variance with respect to the Hamiltonian as a consequence of a specific global distribution of the coefficients $c_m$ in equation (12). Interestingly, the variance of the optimal states scales as $N^p$ with $p$ strictly smaller than two.

Furthermore, we showed that similar results apply to the case of local depolarization noise. However, there the GHZ state outperforms product states. In contrast to that, we demonstrated that more drastic changes of the dynamics give qualitative different results. In the case of temporal and spatial correlation of noise, the scaling of $\mathfrak{F}$ was shown to be better for the GHZ state than for the product state. Note that recently phase estimation in the presence of local or spatially correlated noise has been investigated, and a similar family of states as presented here have been identified as optimal states for parameter estimation for large $N$ [54].

The asymptotic behaviour is in general an interesting question. Concerning the standard scenario, for example, one could ask whether the gap in $\mathfrak{F}$ between the optimal states $|\psi_{opt}\rangle$ and other state families like Dicke states or random states persists or vanishes in the limit of large $N$. In appendix B, a lower bound on $\mathfrak{F}$ is investigated. While, for specific SSS, this bound asymptotically reaches the maximal value $\max_{n \to \infty} \mathfrak{F}_{SSS}(\mu_n) = NT/(2\gamma)$, we analytically find that, for a state family with $\mathcal{V}(h) = O(N^2)$, this quantity is strictly below $NT/(2\gamma)$. This underlines the importance of the variance (or more generally, the global spectral distribution) and could indicate that a too large variance prohibits optimal performance.

Beyond this, there are further big open questions. For example, why is there asymptotically no improvement in scaling for the standard scenario? Clearly, $\mathfrak{F}$ scales with $N^{e(g)}$, where $g(N) > 1$ for finite $N$ but is strictly one for $N \to \infty$. Furthermore, the numerical findings indicate which variance $\mathcal{V}(h)$ is optimal, but it is not clear why it has this particular value.

From a more practical point of view our results show that whilst one-axis twisted SSS are nearly optimal, there are a plethora of other states from $\mathcal{S}_N$ that perform equally well, or better than the one-axis twisted SSS. Thus, if spin-squeezing is an easily-implementable operation in a given practical realization of frequency estimation, it is a good resource for frequency estimation. Moreover, for such states a comparatively simple measurement is known to be asymptotically optimal [22]. However, in experimental realizations where spin-squeezing is not available, or is an expensive operation to implement, our results show that optimal precision might still be possible as there exist non-squeezed states within $\mathcal{S}_N$ that will be just as good as one-axis twisted SSS. Hence, our findings open the way to identify high performance states for...
frequency estimation that may be easily prepared in certain set-ups. Which states are easiest to prepare, and at what cost, for a certain experimental set-up is an interesting question that warrants further investigation. Moreover, we showed that for ion trap experiments with spatially and temporally correlated noise, the GHZ state is not only superior to product states used in parallel, but also achieves a constant improvement even compared to the sequential strategy with single ions.

An open problem that is also relevant from the experimental point of view is the choice of the optimal measurement. By calculating the QFI, the optimization of the measurement is intrinsically performed. In fact, for states other than SSS, where it is known that the optimal measurement is local, nothing substantial is known about the optimal measurement basis (apart from the fact that it is the projection onto the eigenspaces of the symmetric logarithmic derivative).

Acknowledgments

This work was supported by the Austrian Science Fund (FWF): grant numbers P24273-N16, Y535-N16, SFB F40-FoQus F4012-N16 and J3462.

Appendix A. Details on the numerical method

Here, we present some details of the numerical methods for an efficient calculation of the QFI in the subspace of permutationally invariant density matrices.

We first treat the case of local phase noise with local Hamiltonian (section 3) and comment later on other scenarios. In the presence of local phase noise, which is described by $E^{\otimes N}$, where $E(\rho) = e^{-i\theta} \rho + (1 - e^{-i\theta}) \sigma_i \rho \sigma_i$, the density matrix $\rho = E^{\otimes N} (|\psi\rangle \langle \psi|)$ is no longer permutationally symmetric, but it is still permutationally invariant. Therefore, there exists an efficient representation of $\rho$ and, as we will see, also an efficient way to calculate the eigenvalues and eigenstates, even though the rank of $\rho$ is in general exponential in $N$.

In this appendix, we use the common orthonormal eigenbasis of $S^2 = S_x^2 + S_y^2 + S_z^2$ and $S_z$, which we denote by $|j, m, \alpha\rangle$. One has that $S^2 |j, m, \alpha\rangle = j (j + 1) |j, m, \alpha\rangle$ and $S_z |j, m, \alpha\rangle = m |j, m, \alpha\rangle$ for any $\alpha$, which labels the degeneracies. All the states in the ansatz class (see equation (12)) satisfy $j = j_{\max} = N/2$. As there is no degeneracy in the common eigenstates of $S^2$ and $S_z$ for $j = j_{\max}$ we skip the index $\alpha$ in the following.

A.1. Permutationally invariant operator basis

In this subappendix we introduce the permutationally invariant operator basis of Hartmann [43] and determine the action of these operators on the basis states $\{|j, m, \alpha\rangle\}$.

Let $\mathcal{H}^{\otimes N}$ be the state space of $N$ spin-1/2 systems. Defining $|0\rangle \equiv |1/2, 1/2\rangle, |1\rangle \equiv |1/2, -1/2\rangle$, any $\rho \in \mathcal{B}(\mathcal{H})$ can be written as

$$\rho = \sum_{i_1, i_2, \ldots, i_N} \rho_{(i_1, i_2, \ldots, i_N)} X_{(i_1, i_2, \ldots, i_N)^*} \quad \text{(A.1)}$$
where

\[ X(i_1, i_2, i_3, i_4) = |0\rangle \langle 0|^{\otimes 4} \otimes |1\rangle \langle 1|^{\otimes 4}, \]  

(A.2)

form a basis of \( \mathcal{B}(\mathcal{H}) \) and the sum in equation (A.1) runs over all \( i_1, i_2, i_3, i_4 \) satisfying \( \sum_{j=1}^{4} i_j = N \).

As \( \{ |j, m, \alpha\rangle \} \) is an orthonormal basis, the operators of equation (A.2) can be written as

\[ X(i_1, i_2, i_3, i_4) = \sum_{j, j', \alpha, \alpha'} c_{i_1, i_2, i_3, i_4}^{(j, j', \alpha, \alpha')} |j, m, \alpha\rangle \langle j', m', \alpha'|, \]  

(A.3)

where \( m = i_1 + i_2 - N/2 \), \( m' = i_3 + i_4 - N/2 \), and \( 2d \equiv i_1 + i_3 \) can be interpreted as the number of places where the symbols in the corresponding bra and ket in equation (A.2) are different [43]. Let \( \pi: S_N \to \mathcal{H}^{\otimes N} \) be a unitary representation of the permutation group of \( N \) objects, \( S_N \). Using Schur’s lemmas [55], the unitary representation \( \{ \pi_g; g \in S_N \} \) can be decomposed into irreducible representations of \( S_N \). Indeed, in the \( \{ |j, m, \alpha\rangle \} \) basis all \( \{ \pi_g; g \in S_N \} \) take the form

\[ \pi_g = \sum_j n_{g}^{(j)} \otimes 1_{d_j}. \]  

(A.4)

Here, \( j \) labels the inequivalent irreducible representations on \( S_N \), \( \pi^{(j)} \), present in the unitary representation \( \pi \), and \( d_j \) denotes the multiplicity of the corresponding \( \pi^{(j)} \). Consequently, the total Hilbert space, \( \mathcal{H}^{\otimes N} \), can be conveniently written as

\[ \mathcal{H}^{\otimes N} \cong \bigoplus \mathcal{M}^{(j)} \otimes \mathcal{N}^{(j)}, \]  

(A.5)

where the space \( \mathcal{M}^{(j)} \equiv \text{span} \{ |j, \alpha\rangle_{\alpha=1}^{d_j} \} \) of dimension \( \Delta_j \equiv \dim(\pi^{(j)}) \) is the space upon which the irreducible representation \( \pi^{(j)} \) acts non-trivially, and the space \( \mathcal{N}^{(j)} \equiv \text{span} \{ |j, m\rangle_{m=1}^{d_j} \} \) of dimension \( d_j = 2j + 1 \), is the space upon which \( \{ \pi_g; g \in S_N \} \) act trivially [56]. The tensor product in equation (A.5) is not a tensor product between real physical systems, but rather a tensor product between two virtual systems described by the state spaces \( \mathcal{M}^{(j)} \) and \( \mathcal{N}^{(j)} \) respectively. The states of these virtual systems can be defined via the isomorphism \( |j, m, \alpha\rangle \equiv |j, \alpha\rangle \otimes |j, m\rangle \).

Hartmann showed that an operator basis for the permutationally symmetric states is given by [43]

\[ K_{m, m', d}^{\dagger} = \frac{1}{i_1! i_2! i_3! i_4!} \sum_{g \in S_N} \pi_g \left( X(i_1, i_2, i_3, i_4) \right) \pi_g^{\dagger}. \]  

(A.6)

Note that the definition of the basis operators in equation (A.6) differ from that of [43] by a factor \( N! / (i_1! i_2! i_3! i_4!) \). Using Schur’s lemmas it can be shown that equation (A.6) can be written with respect to the \( \{ |j, m, \alpha\rangle \} \) basis as

\[ K_{m, m', d} = \frac{N!}{i_1! i_2! i_3! i_4!} \sum_{j} (\mathcal{D} \mathcal{M}^{(j)} \otimes \mathcal{I} \mathcal{N}^{(j)}) \mathcal{P}^{(j)} \left[ X(i_1, i_2, i_3, i_4) \right], \]  

(A.7)

where \( \mathcal{D} \) is the completely depolarizing map, \( \mathcal{D}[A] = \frac{\text{tr}(A)}{\text{dim}(\mathcal{H})} 1 \), \( \forall A \in \mathcal{B}(\mathcal{H}) \), \( \mathcal{I} \) is the identity map, and \( \mathcal{P}^{(j)}[A] = \Pi_j A \Pi_j \), where \( \Pi_j \) is the projector onto the space \( \mathcal{M}^{(j)} \otimes \mathcal{N}^{(j)} \) [56]. Careful
counting gives a total number of $1/6 (N + 1) (N + 2) (N + 3)$ different basis operators for the space of permutationally invariant matrices [43].

Plugging equation (A.3) into equation (A.7) yields

$$K_{m,m',d} = \frac{N!}{i_1 i_2 i_3} \sum_j \left( D_M^{\alpha} \otimes I_{N^0} \right) \left[ c_{m,m',d}^{j,\alpha} |j, \alpha\rangle \otimes |j, m\rangle \langle j, m'| \right]$$

$$= \frac{N!}{i_1 i_2 i_3} \sum_j \sum_{\alpha} \Delta_j \sum_{\alpha'} c_{m,m',d}^{j,\alpha,\alpha'} |j, \alpha\rangle \langle j, m| \langle j, m'| \right].$$

(A.8)

Defining

$$\frac{N!}{i_1 i_2 i_3} \sum_{\alpha} c_{m,m',d}^{j,\alpha,\alpha'} \Delta_j = \mu_{m,m',d}^{(j)} \alpha\alpha$$

and noting that

$$I_M^{\alpha} = \sum_{\alpha} |j, \alpha\rangle \langle j, \alpha| \right),$$

equation (A.8) becomes

$$K_{m,m',d} = \sum_j \sum_{\alpha} \mu_{m,m',d}^{(j)} |j, \alpha\rangle \langle j, \alpha| \otimes |j, m\rangle \langle j, m'| \right]$$

$$= \sum_j \sum_{\alpha} \mu_{m,m',d}^{(j)} |j, m, \alpha\rangle \langle j, m', \alpha| \right]$$

(A.9)

where we have made use of the isomorphism $|j, \alpha\rangle \otimes |j, m\rangle$ in the last line of equation (A.9). Finally, defining

$$P_{m,m}^{(j)} \alpha\alpha \equiv \sum_{\alpha} |j, \alpha\rangle \otimes |j, m\rangle$$

gives

$$K_{m,m',d} = \sum_{j,m} P_{m,m}^{(j)} \alpha\alpha K_{m,m',d}$$

(A.10)

where $j_{\min} = 0 (1/2)$ for $N$ even (odd).

Thus, the action of the permutationally symmetric operators, equation (A.6) on the basis states $\{ |j, m, \alpha\rangle \}$ is given by equation (A.10). The action of phase noise $E$ is particularly easy to express in this basis. Since

$$E(\{0\} \langle 0 |) = |0\rangle \langle 0|, \quad E(\{1\} \langle 1 |) = |1\rangle \langle 1|$$

and $E(0 \langle 1 |) = e^{-\gamma t} |0\rangle \langle 0|$, one finds that

$$E^{\otimes N} K_{m,m',d} = e^{-2\gamma t} K_{m,m',d}.$$ 

(A.11)

A2. Efficient representation of permutationally invariant states under phase noise

In this subappendix we seek an efficient description of states from our ansatz space (see equation (12)) in terms of the Hartmann operators $K_{m,m',d}$ of equation (A.6). Consider an arbitrary pure state, $\rho = \sum_{m,m'} c_{m,m'} |N/2, m\rangle \langle N/2, m'|$ where $c_{m,m'} \in [0, 1]$, belonging to our ansatz space. For $m \geq m'$, and $m + m' \leq 0$ the matrix elements $|N/2, m\rangle \langle N/2, m'|$ can be expressed as

$$\sqrt{\left( \frac{N}{N/2 + m} \right) \left( \frac{N}{N/2 + m'} \right)} |N/2, m\rangle \langle N/2, m'| = \sum_{g \in S_N} \pi_g \sum_{l_1 l_2 l_3 l_4} \frac{1}{i_1 i_2 i_3 i_4} X_{i_1 i_2 i_3 i_4} \pi_g$$

$$= \sum_{i_1 i_2 i_3 i_4} \left( \frac{1}{i_1 i_2 i_3 i_4} \sum_{g \in S_N} \pi_g X_{i_1 i_2 i_3 i_4} \pi_g \right)$$

(A.12)
where the second sum in the first line of equation (A.12) runs over all \(i_1, i_2, i_3, i_4\) satisfying \(i_1 + i_2 = N/2 + m, i_1 + i_3 = N/2 + m'\). Using equation (A.7) it follows that the expression in brackets in the second line of equation (A.12) is equal to \(K_{m,m',d}\). In addition, as \(m, m'\) are fixed the sum over \(i_1, i_2, i_3, i_4\) must yield operators \(K_{m,m',d}\) with different \(d\). Recalling that \(d^2\) counts the number of non-coincidences between the strings in the ket and bra of \(X_{i_1,i_2,i_3,i_4}\) it is easy to determine that, for \(m + m' \geq 0\), \(- \frac{m - m'}{2} \leq d \leq \frac{N - |m + m'|}{2}\). Similar arguments hold for all other cases so that generally one finds
\[
\sum_{d} = \frac{1}{\sqrt{N(N/2 + m)(N/2 + m')}} \sum_{d=|m-m'|/2}^{(N-|m+m'|)/2} K_{m,m',d}. \tag{A.13}
\]

Using equations (A.11), (A.13) the action of phase noise on the state \(\rho\) can be easily evaluated to be
\[
\rho = \mathcal{E}^{SN}(|\psi\rangle\langle\psi|) = \sum_{m,m'} \frac{c_m c_{m'}}{\sqrt{N(N/2 + m)(N/2 + m')}} \sum_{d=|m-m'|/2}^{(N-|m+m'|)/2} e^{-2d^n} K_{m,m',d}. \tag{A.14}
\]

Note that in order to compute the QFI, we do not need to apply the unitary operator, \(U = e^{iHt}\), since it commutes with the Hamiltonian, i.e., \(F(\rho) = F(U\rho U^d)\). Defining
\[
\lambda_{j,m,m'}^{(j)} = \sum_{m,m'} c_m c_{m'} \sqrt{N(N/2 + m)(N/2 + m')} e^{-2d^n} K_{m,m',d}, \tag{A.15}
\]
we have
\[
\rho = \sum_{j,m,m'} \lambda_{j,m,m'}^{(j)} P_{m,m'}. \tag{A.15}
\]

Therefore, we find an efficient description of the noisy state \(\rho\) in terms of the operators \(P_{m,m'}\). As the states \(|j, m, \alpha\rangle\) are all orthogonal, the eigenvalues of \(\rho\) are obtained by diagonalizing the matrices \(\Lambda^{(j)} = \{\lambda_{j,m,m'}^{(j)}\}_{m,m'} \in \mathbb{C}^{(2^j+1) \times (2^j+1)}\), which can be done efficiently in \(N\).

### A.3. Determination of \(\mu_{j,m,m',d}^{(j)}\)

The crucial part is now to find the transition values \(\mu_{j,m,m',d}^{(j)}\). First, consider the case \(m = m'\), which renders the operators \(K_{m,m,d}\) self-adjoint and the transition values become eigenvalues. Note that \(K_{m,m,0}\) are projectors on the eigenspace of \(S_z\) with eigenvalue \(m\), which implies that \(\mu_{j,m,0}^{(j)} = 1\) for all \(j, m\). Hence,
\[
S_z = \sum_{m=-N/2}^{N/2} m K_{m,m,0}. \tag{A.16}
\]

In addition, one finds
\[
S^2 = \frac{N}{2} N_{id} + \sum_{i \neq j} |101\rangle\langle10| + h. c. = \sum_{m} \sum_{n} N_2 K_{m,m,0} + K_{m,m,1}. \tag{A.17}
\]

Combining equations (A.16), (A.17) gives
\[
S^2 = \sum_{m} \sum_{n} (N/2 + m^2) K_{m,m,0} + K_{m,m,1}. \tag{A.18}
\]
As \( (S^2)_{m,m} = j (j + 1) \), one finds that \( m^2 + N/2 + \mu_{m,m}^{(i)} = j (j + 1) \). Hence, one can calculate the values of \( \mu_{m,m}^{(i)} \) for all \( j, m \).

Multiplying \( X_{(N/2+m-d,d,d,N/2-m-d)} \) by \( \pi_i X_{(N/2+m-1,1,1,N/2-m+1)} \pi_i^\dagger \), for some \( g \in S_N \) and \( d > 1 \), results in one out of four possibilities, depending on the choice of \( g \).

\[
\begin{align*}
X_{(N/2+m-d,d,d,N/2-m-d)} \pi_i X_{(N/2+m-1,1,1,N/2-m+1)} \pi_i^\dagger & \rightarrow \\
& \begin{cases} \\
X_{(N/2+m-d-1,d+1,d+1,N/2-m-d-1)} \\
X_{(N/2+m-d,d,d,N/2-m-d)} \\
X_{(N/2+m-d+1,d-1,d-1,N/2-m-d+1)} \\
0.
\end{cases}
\end{align*}
\] (A.19)

To carry over the results for the respective symmetrized versions, one has to carefully count the number of non-trivial permutations of \( K_{m,m,d} \) and \( K_{m,m,1} \) and the number of instances for the four cases in equation (A.19). For example, the third possibility in equation (A.19) occurs \( N![(N/2 + m - d)!d!d!(N/2 - m - d)!] (N/2 + m - d) (N/2 - m - d) \) times, which is \((d+1)^2\) times more often than the total number of non-trivial permutations for \( K_{m,m,d+1} \).

Similarly, one has to reason for the other two non-trivial cases. Then, one has

\[
K_{m,m,d} K_{m,m,1} = \left[ \left( \frac{N}{2} - d + 1 \right)^2 - m^2 \right] K_{m,m,d-1} + d (N - 2d) K_{m,m,d} + (d + 1)^2 K_{m,m,d+1}.
\] (A.20)

Hence by taking successive powers of \( S^2 \),

\[
S^{2n} = \sum_{m=-N/2}^{N/2} \sum_{d=0}^{n} k_{m,m,d}^{(n,0)} K_{m,m,d}
\] (A.21)

(the index 0 will be used later) one obtains for any \( n \)

\[
[j (j + 1)]^n = \sum_{m=-N/2}^{N/2} \sum_{d=0}^{n} \kappa_{m,m,d}^{(n,0)} \mu_{m,m,d}^{(j)}
\] (A.22)

Starting with \( n = 1 \), one calculates \( \mu_{m,m,n}^{(j)} \) recursively by inserting the eigenvalues \( \mu_{m,m,d}^{(j)} \) with \( d < n \).

In the case of \( m \neq m' \), observe that \( S_- = S_x - i S_y = \sum_{m=-N/2}^{N/2-1} K_{m,m+1/2} \). Similar arguments as for equation (A.20) lead to

\[
K_{m,m',d} K_{m',m'+1/2} = \left[ \frac{1}{2} (N + m + m') - d + 1 \right] K_{m,m'+1,d-1/2} + \left[ \frac{1}{2} (m + m') + d + 1 \right] K_{m,m'+1,d+1/2}.
\] (A.23)

Hence, one iteratively finds the expression of \( S^2 S_+^i \) in terms of the Hartmann basis, equation (A.6),

\[
S^{2n} S_+^i = \sum_{m=-N/2}^{N/2-i} \sum_{d=i/2}^{n+i/2} k_{m,m,d}^{(n,i)} K_{m,m+i,d}.
\] (A.24)

As \( \langle j, m, \alpha | S^2 S_+^i | j, m + 1, \alpha \rangle = [j (j + 1)]^n \prod_{i=1}^{\infty} C_{j,m+i}^- \) with \( C_{j,m}^- = \sqrt{j (j + 1) - m (m - 1)} \), one has linear equations for any \( (n, i) \) which can be solved recursively using the solution for \( (n - 1, i) \):
\[ \sum_{d=1}^{n+1/2} \kappa_{m,m,d}^{(n,i)} \mu_{m,m+i,d}^{(j)} = \left( j (j + 1) \right)^{i} \prod_{k=1}^{i} C_{j,m+k}^- . \] (A.25)

In this way, one finds all \( \mu_{m,m+i,d}^{(j)} \) and therefore an efficient way to express \( \rho \) in the basis of \( P_{m,m'} \). The actual implementation of the effective density operator description is done with four-dimensional tensors. Their processing is facilitated by using the Tensor Toolbox [57]. The QFI is then calculated via the numerical diagonalizing of \( \rho \).

**A.4. Remarks on other scenarios**

We shortly note how the method changes if one considers different scenarios.

**Depolarization noise.** In section 4.1, local depolarization noise is considered. The action of depolarizing noise on diagonal elements is:

\[ E(\{0\}) = \frac{1}{2} \{0\} \{0\} + \frac{1}{2} \{1\} \{1\} \] and

\[ E(\{1\}) = \frac{1}{2} \{1\} \{1\} + \frac{1}{2} \{0\} \{0\} \] using this fact it is straightforward to determine the action of depolarizing noise on the Hartmann operators as

\[ \sum_{\nu} \gamma_{\nu} \mu_{\nu}^m = \sum_{\nu} \gamma_{\nu} \mu_{\nu}^{m'} = \sum_{\nu} \gamma_{\nu} \mu_{\nu}^m \] (A.26)

The values of \( i_k \) are determined by \( (m, m', d) \) (see equation (A.6)) and the sets for the summation index read

\[ X^+ = \{ 0, \ldots, \min(i_1, i_4 - \epsilon) \} \quad \text{and} \quad X^- = \{ 0, \ldots, \min(i_2, i_3 + \epsilon) \} \] (A.28)

The effort in calculating the action of depolarization noise is thus higher than that of local phase noise. For this reason the maximal system size is restricted to \( N = 30 \).

**Correlated noise.** In the presence of correlated noise (see section 4.2), the numerical treatment is straightforward. The action onto an off-diagonal element simply reads

\[ E \left[ \{N/2, m\} \{N/2, m'\} \right] = p^{(m-m')} \{N/2, m\} \{N/2, m'\} \] with \( p = \exp(-\gamma t - \epsilon^2 + 1) \). This means that the noisy state stays within the subspace spanned by \{ \{N/2, m\} \}. Therefore, a more complicated treatment using the Hartmann basis is not necessary.

**Appendix B. Asymptotic behaviour**

As discussed in section 3.2, the numerical results for the standard scenario suggest that the variance as well as a specific behaviour of the coefficients, \( c_m \), of the state, in the eigenbasis of the Hamiltonian, play a key role in the metrological performance of a quantum state in the standard scenario. An interesting question is what happens to the average improvement of \( F \) for
certain state families as $N$ becomes large. For instance, although the relative difference between $\xi_{\text{opt}}$ and $\xi$ of Dicke states increases in the investigated range of $N$, this gap could stabilize to a finite value, or even close in the limit $N \to \infty$. Unfortunately, in this limit, numerical results do not provide definitive answers as calculation of the QFI requires the spectral decomposition of the time-evolved quantum state. Even for simple state families, this is asymptotically not feasible.

In this appendix, we consider a particular measurement that clearly leads to a lower bound on the QFI. The measurement is assumed to be in the eigenbasis of $S_y = \exp(-i\varphi S_y) S_x \exp(i\varphi S_y)$, i.e., a measurement in the $x$–$y$ plane. The quantity

$$G = \max_{\varphi} \left| \frac{d}{dt} \langle S_y \rangle_{\rho(t)} \right|^2,$$

which has also been considered in [11, 22], is equivalent to the QFI for any initial state if $N = 1$, and for the product state for any $N$. It was recently shown [58] that $G$ is in general a lower bound on the QFI, i.e. $G \leq F$.

Optimizing over $\varphi$ and inserting the time-evolved state (see equation (6)) for a symmetric state, $|\psi\rangle$, we obtain

$$G = \frac{Nt^2 s^2}{e^{2\pi} - 1 + 4V_T(S_y)/N},$$

with $s = 2 \langle S_y \rangle / N$.

If $s \to 1$ and $V_T(S_y)/N \to 0$ for $N \to \infty$, $\max_{t} TG/t \leq \xi$ reaches the optimal value $NT/(2\gamma)$ for $t \to 0$. We now show that states $|\psi\rangle$ whose distribution of coefficients, $\{c_m\}$, is sufficiently smooth and for which $V(h) = O(N^2)$, cannot reach $s = 1$ asymptotically. Therefore, such a state can not achieve the bound $NT/(2\gamma)$, and is suboptimal with respect to the quantity $G$. Note that if a similar statement was to hold true for $\xi$, one could conclude that states satisfying $V(h) = O(N^2)$ are asymptotically less useful than the optimal states.

Let us now come to the lemma.

**Lemma 1.** We consider states $|\psi\rangle \in S_N$ given in equation (12), where the coefficients $c_m = c_{N-m} \in [0, 1]$ only depend on $N$. We assume that the coefficients are sufficiently ‘smooth’ in the sense that, $c_{m+1} = c_m \pm \epsilon_m(N)$ with $\lim_{N \to \infty} \epsilon_m(N) = 0$ for all $m \in \{-N/2, \ldots, N/2 - 1\}$. In addition, be $|\psi\rangle$ such that $V_T(h) = \sum_{m=0}^{N} m^2 c_m^2 = O(N^2)$. Then it holds that

$$\lim_{N \to \infty} \langle S_z \rangle_{\psi} / N < 1.$$

**Proof.** We first compute $\langle S_z \rangle_{\psi}$. From the theory of angular momentum $S_z = 1/2(S_+ + S_-)$ where the operators $S_{\pm}$ are the standard raising and lower operators whose action on the basis states $\{|j, m\rangle\}$ is given by

$$S_{\pm} |j, m\rangle = \sqrt{(j \mp m)(j \pm m + 1)} |j, m \pm 1\rangle$$

(B.3)
respectively. Computing \( \langle S_x \rangle_w \) yields
\[
\langle S_x \rangle_w = \frac{1}{N/2-1} \sum_{m=-N/2}^{N/2-1} \sqrt{(N/2 - m)(N/2 + m + 1)} c_m c_{m+1}
\]
\[
\langle S_y \rangle_w = \frac{1}{N/2+1} \sum_{m=-N/2+1}^{N/2} \sqrt{(N/2 + m)(N/2 - m + 1)} c_m c_{m-1}.
\]
(B.4)

However, as \( c_m = c_{N-m} \in \mathbb{R} \) it is easy to see that \( \langle S_x \rangle = \langle S_y \rangle \). Hence
\[
2 \frac{\langle S_x \rangle_w}{N} = 2 \sum_{m=-N/2}^{N/2-1} \frac{\sqrt{(N - m)(N/2 + m + 1)}}{N} c_m c_{m+1}
\]
\[
= \sum_{m=-N/2}^{N/2-1} \sqrt{(1 - \frac{2m}{N})(1 + \frac{2m}{N} + \frac{2}{N})} c_m \left[ c_m - c_m(N) \right],
\]
(B.5)

where we have used the smoothness of the coefficients \( c_m \) in the second line of equation (B.5). In order to convert the sum into an integral, let \( x = 2m/N \). Taking the limit \( N \to \infty \) one obtains
\[
\lim_{N \to \infty} 2 \frac{\langle S_x \rangle_w}{N} = \int_{-1}^{1} \sqrt{1 - x^2} c(x) dx,
\]
where \( c(x) = \lim_{N \to \infty} c_{N/2} \in [0, 1] \) with \( \int_{-1}^{1} c(x) dx = 1 \). Note that \( \sqrt{1 - x^2} \leq 1 \) and equals unity if and only if \( x = 0 \). Due to the normalization of \( c(x) \), we find that \( \lim_{N \to \infty} 2 \frac{\langle S_x \rangle_w}{N} = 1 \) only if \( c(x) = \delta(x) \), the Dirac delta function. However, as \( V(h) = O(N^2) \) by assumption, \( c(x)^2 \) cannot be the Delta function. Hence, \( \lim_{N \to \infty} 2 \frac{\langle S_x \rangle_w}{N} < 1 \). This completes the proof. \(\square\)

Appendix C. Maximal QFI for GHZ under depolarizing noise

We consider here the depolarizing channel (see equation (8)). In [18] (appendix A), we have shown that the QFI of for the GHZ state as initial state is given by
\[
\mathcal{F} = t^2 N^2 \frac{2^N e^{-2Ny}}{(1 + e^{-\gamma t})^N + (1 - e^{-\gamma t})^N}.
\]
(C.1)

For the time optimization, we assume that, for large \( N \), one can restrict oneself to \( \gamma t \ll 1 \), which is a posteriori justified. Then, one approximates \( (1 + e^{-\gamma t})/2 \) by \( e^{-\gamma t/2} \) and neglect \( \left( (1 - e^{-\gamma t})/2 \right)^N \). Hence, the QFI equals \( \mathcal{F} \approx t^2 N^2 \exp(-3/2\gamma t) \). The small difference in the power then gives rise to a modified optimal time \( t_{\text{opt}} \approx 2/(3\gamma e) \) which leads to \( \mathcal{F}_{\text{GHZ}} = 2NT/(3\gamma e) \).

References

[1] Caves C M 1981 *Phys. Rev. D* **23** 1693
[2] Holland M J and Burnett K 1993 *Phys. Rev. Lett.* **71** 1355–8
[3] Lee H, Kok P and Dowling J P 2002 *J. Mod. Opt.* **49** 2325–38
[4] Valencia A, Scarcelli G and Shih Y 2004 *Appl. Phys. Lett.* **85** 2655–7
[5] de Burgh M and Bartlett S D 2005 *Phys. Rev. A* **72** 042301
[6] McKenzie K, Shaddock D A, McClelland D E, Buchler B C and Lam P K 2002 *Phys. Rev. Lett.* **88** 231102
[7] Collaboration T L S 2011 *Nat. Phys.* **7** 962–5
[8] Wineland D J, Bollinger J J, Itano W M, Moore F L and Heinzen D J 1992 *Phys. Rev. A* **46** 6797–800
[9] Bollinger J J, Itano W M, Wineland D J and Heinzen D J 1996 *Phys. Rev. A* **54** 4649–52
[10] Greenberger D M, Horne M A and Zeilinger A 1989 Going beyond Bell’s theorem *Bell’s Theorem, Quantum Theory and Conceptions of the Universe* (Berlin: Springer) pp 69–72
[11] Huelga S F, Macchiavello C, Pellizzari T, Ekert A K, Plenio M B and Cirac J I 1997 *Phys. Rev. A* **54** 4649–52
[12] Escher B M, de Matos Filho R L and Davidovich L 2011 *Nat. Phys.* **7** 406–11
[13] Demkowicz-Dobrzański R, Kołodyński J and Guţă M 2012 *Nat. Commun.* **3** 1063
[14] Chaves R, Brask J B, Markiewicz M, Kołodyński J and Acín A 2013 *Phys. Rev. Lett.* **111** 120401
[15] Chin A W, Huelga S F and Plenio M B 2012 *Phys. Rev. Lett.* **109** 233601
[16] Matsuzaki Y, Benjamin S C and Fitzsimons J 2011 *Phys. Rev. A* **84** 012103
[17] Chaves R, Brask J B, Markiewicz M, Kołodyński J and Acín A 2013 *Phys. Rev. Lett.* **111** 120401
[18] Kessler E M, Lovchinsky I, Sushkov A O and Lukin M D 2014 *Phys. Rev. Lett.* **112** 150802
[19] Kitagawa M and Ueda M 1993 *Phys. Rev. A* **47** 5138
[20] Ulam-Orgikh D and Kitagawa M 2001 *Phys. Rev. A* **64** 052106
[21] Rosenkranz M and Jaksch D 2009 *Phys. Rev. A* **79** 022103
[22] Maccone L 2013 *Phys. Rev. A* **88** 042109
[23] Cramér H 1961 *Mathematical Methods of Statistics* (Princeton, NJ: Princeton University Press)
[24] Fisher R A 1922 *Phil. Trans. R. Soc. Lond. Ser. A* **222** 309–68
[25] Fisher R A 1925 *Math. Proc. Camb. Phil. Soc.* **22** 700–25
[26] Helstrom C 1976 *Quantum Detection and Estimation Theory* (New York: Academic)
[27] Holevo A S 1980 Probabilistic and Statistical Aspects of Quantum Theory (North-Holland Series in Statistics and Probability) (Amsterdam: North-Holland)
[28] Braunstein S L and Caves C M 1996 *Phys. Rev. Lett.* **72** 3439–43
[29] Giovannetti V, Lloyd S and Maccone L 2004 *Science* **306** 1330–6
[30] Kielpinski D, Meyer V, Rowe M A, Sackett C A, Itano W M, Monroe C and Wineland D J 2001 *Science* **291** 1013–5
[31] Langr C et al 2005 *Phys. Rev. Lett.* **95** 060502
[32] Roos C F, Chwalla M, Kim K, Riebe M and Blatt R 2006 *Nature* **443** 316–9
[33] Monz T, Schindler P, Barreiro J T, Chwalla M, Nigg D, Coish W A, Harlander M, Hänsel W, Heinrich M and Blatt R 2011 *Phys. Rev. Lett.* **106** 130506
[34] Szankowski P, Chwedenczuk J and Trippehach M 2012 arXiv:1212.2528
[35] Dorner U 2012 *New J. Phys.* **14** 043011
[36] Braunstein S L, Caves C M and Milburn G J 1996 *Ann. Phys.* **247** 135–73
[37] Demkowicz-Dobrzański R, Dorner U, Smith B J, Lundeen J S, Wasilewski W, Banaszek K and Walmsley I A 2009 *Phys. Rev. A* **80** 013825
[38] Dicke R H 1954 *Phys. Rev.* **93** 99–110
[39] Nelder J A and Mead R 1965 *Comput. J.* **7** 308–13
[40] Lagarias J C, Reeds J A, Wright M H and Wright P E 1998 *SIAM J. Optim.* **9** 112–47
[41] Hartmann S 2012 arXiv:1201.1732
[42] Wang X and Sanders B C 2003 *Phys. Rev. A* **68** 012101
[43] Sorensen A, Duan L M, Cirac J I and Zoller P 2001 *Nature* **409** 63–66
[44] Korbicz J K, Cirac J I and Lewenstein M 2005 *Phys. Rev. Lett.* **95** 120502
[45] Berry D W and Wiseman H M 2000 *Phys. Rev. Lett.* **85** 5098–101
[46] Combes J and Wiseman H M 2005 J. Opt. B: Quantum Semiclass. Opt. **7** 14
[47] Chaves R, Brask J B, Markiewicz M, Kołodyński J and Acín A 2013 *Phys. Rev. Lett.* **111** 120401
[48] Bužek V, Derka R and Massar S 1999 *Phys. Rev. Lett.* **82** 2207–10
[51] André A, Sørensen A S and Lukin M D 2004 Phys. Rev. Lett. 92 230801
[52] Borregaard J and Sørensen A S 2013 Phys. Rev. Lett. 111 090801
[53] Kessler E M, Kómár P, Bishof M, Jiang L, Sørensen A S, Ye J and Lukin M D 2013 arXiv:1310.6043[quant-ph]
[54] Knysh S I, Chen E H and Durkin G A 2014 arXiv:1402.0495
[55] Chen J Q 1989 Group Representation Theory for Physicists (Singapore: World Scientific)
[56] Bartlett S D, Rudolph T and Spekkens R W 2007 Rev. Mod. Phys. 79 555
[57] Bader B W, Kolda T G et al 2012 Matlab Tensor Toolbox version 2.5 www.sandia.gov/~tgkolda/
        TensorToolbox
[58] Zhong W, Lu X M, Jing X and Wang X 2013 arXiv:1311.6600