Entanglement is a valuable resource for quantum technology. In metrology, entangled probes are capable of more accurate measurements than unentangled probes [1–6]. In addition to using entangled probes to enhance the measurement of a single parameter, using entanglement to estimate many parameters at once, or a function of those parameters, has recently been an area of interest due to potential applications in tasks such as nanoscale nuclear magnetic resonance imaging [7–15].

In this Letter, we are interested in generalizing the work of Ref. [15], which demonstrated a lower bound on the variance of an estimator of a linear combination of \( d \) parameters coupled to \( d \) qubits. We will generalize this approach to measuring an arbitrary real-valued, analytic function of \( d \) parameters and show that entanglement can reduce the variance of such an estimate by a factor of \( \mathcal{O}(d) \). Finally, we present a protocol which achieves optimal variance asymptotically in the limit of long measurement time. In addition, when the parameters are coupled to \( d \) interferometers or to a combination of interferometers and qubits, we propose an analogous Heisenberg-scaling protocol to improve measurement noise. However, in this case, we lack a proof of optimality. We also can use the protocol presented in Ref. [16] to couple the parameters to continuous variables detected by homodyne measurements.

We will also examine the application of such a protocol to field interpolation. Suppose sensors are placed at \( d \) spatially separated locations, but we wish to know the field at a point with no sensor. We may pick a reasonable ansatz for the field with no more than \( d \) parameters, use our \( d \) measurements to fix the degrees of freedom of that ansatz, and compute the field at our desired point. Because the field of interest is a function of the field at \( d \) other locations, our protocol offers reduced noise over performing the same procedure without using entanglement.

**Setup.**—In this Letter, bold is used to indicate vectors, hats (as in \( \hat{H} \)) indicate operators, and variables with a tilde (such as \( \tilde{f} \)) are estimators of the corresponding quantity with no tilde (such as \( f \)). The notation \( \mathbb{E}_Y[X] \) means the expected value of \( X \) over all possible \( Y \). If we merely write \( \mathbb{E}[X] \), then we average over all parameters required to define \( X \) (e.g. if \( Y \) depended on \( Z \), then \( \mathbb{E}_Z[\mathbb{E}_Y[X]] \)). We define the variance, \( \text{Var}_Y[X] \), similarly.

We consider a system with \( d \) sensor nodes, where node \( i \) consists of a single qubit coupled to a real parameter \( \theta_i \) (see Fig. 1), and suppose that the state evolves under the Hamiltonian

\[
\hat{H} = \hat{H}_c(t) + \frac{1}{2} \theta_i \hat{\sigma}_i^z,
\]

where \( \hat{\sigma}_i^{x,y,z} \) are the Pauli operators acting on qubit \( i \) and \( \hat{H}_c(t) \) is a time-dependent control Hamiltonian that we choose, which may include coupling to ancilla qubits. Here, and throughout the paper, repeated indices indicate summation. We want to measure an arbitrary real-

---

**Heisenberg-Scaling Measurement Protocol for Analytic Functions with Quantum Sensor Networks**

Kevin Qian, Zachary Eldredge, Wenchao Ge, Guido Pagano, Christopher Monroe, James V. Porto, and Alexey V. Gorshkov

1 Joint Quantum Institute, NIST/University of Maryland, College Park, MD 20742, USA
2 Joint Center for Quantum Information and Computer Science, NIST/University of Maryland, College Park, MD 20742, USA
3 Montgomery Blair High School, Silver Spring, MD 20901, USA
4 Institute for Quantum Science and Engineering (IQSE) and Department of Physics & Astronomy, Texas A&M University, College Station, Texas 77843, USA
5 United States Army Research Laboratory, Adelphi, MD 20783, USA
6 IonQ, Inc., College Park, MD 20740

We generalize past work on quantum sensor networks to show that, for \( d \) input parameters, entanglement can yield a factor \( \mathcal{O}(d) \) improvement in mean squared error when estimating an analytic function of these parameters. We show that the protocol is optimal for qubit sensors, and conjecture an optimal protocol for photons passing through interferometers. Our protocol is also applicable to continuous variable measurements, such as one quadrature of a field operator. We outline a few potential applications, including calibration of laser operations in trapped ion quantum computing.

![FIG. 1. An illustration of a quantum sensor network of spatially separated nodes. At each node, there is an unknown parameter \( \theta_i \) coupled to a qubit, which accumulates phase proportional to \( \theta_i \).](image-url)
valued, analytic function $f(\theta)$ of $d$ unknown parameters \( \theta = (\theta_1, \ldots, \theta_d) \) for time $t_{\text{total}}$. We would like to determine how well the quantity $f(\theta)$ can be estimated, and find a protocol for doing so. To specify a protocol, we choose an input state, a control Hamiltonian $\hat{H}_c(t)$, and a final measurement.

For a general estimator, we use the mean squared error (MSE) $M$ of our estimate $\hat{f}$ from the true value $f(\theta)$ as a figure of merit. Explicitly, 
\[
M = \mathbb{E}[(\hat{f} - f(\theta))^2] = \text{Var} \hat{f} + \mathbb{E}[(\hat{f} - f(\theta))^2] \geq \text{Var} \hat{f}. \quad (2)
\]

**Lower bound on error.**—We now identify the minimum possible error of an estimator of $f(\theta)$ which measures for time $t_{\text{total}}$. For any unbiased estimator $\hat{f}$ which uses samples from a probabilistic process (such as physical experiments) to estimate the value $f(\theta)$, the variance is bounded by [17]
\[
\text{Var} \hat{f} \geq \delta^T \cdot F^{-1} \cdot \delta, \quad (3)
\]
where $\delta := \nabla f(\theta)$ and $F$ is the classical Fisher information matrix. The Fisher information matrix measures the sensitivity of the sampled probability distribution to changes in the parameters $\theta$. Equation (3) is known as the transformed multivariate Cramér-Rao Bound. Defining the quantum Fisher information matrix, $F_Q$, as the maximization of the Fisher information matrix over all measurement schemes, the matrix inequality $F^{-1} \geq F_Q^{-1}$ holds [18]. While $F$ tells us something about a particular experimental setup, $F_Q$ is maximized over all possible experiments that could be performed on a state. By combining Eq. (3) with Eq. (2), we obtain the quantum Cramér-Rao bound for any estimator,
\[
M \geq \text{Var} \hat{f} \geq \delta^T \cdot F_Q^{-1} \cdot \delta. \quad (4)
\]

Although the bound in Eq. (4) cannot always be saturated, it can when the generators $\partial \hat{H} / \partial \theta_i$ commute, as in Eq. (1) [18]. For a particular function $f$ at a particular $\theta$, $\delta$ is a constant. By using Cauchy-Schwarz inequalities and properties of the quantum Fisher information shown in Ref. [19], the same approach taken in Ref. [15] can be used to simplify the right side of Eq. (4) (this is detailed in Sec. S1 of the Supplemental Material [20]), yielding
\[
M \geq \text{Var} \hat{f} \geq \frac{\max_i \delta_i^2}{t_{\text{total}}^2} = \frac{\max_i f_i(\theta)^2}{t_{\text{total}}^2}, \quad (5)
\]

where for only the function $f(\theta)$, we define $f_i(\theta) = \frac{\partial f(\theta)}{\partial \theta_i}$. This definition also generalizes to multiple partial derivatives (i.e. $f_{ij} = \frac{\partial^2 f(\theta)}{\partial \theta_i \partial \theta_j}$). We will show later that this inequality can be saturated at asymptotic time $t_{\text{total}}$.

Before moving on to the optimal protocol, we will consider a protocol which does not use entanglement and does not saturate Eq. (5) as a useful contrast to an entangled strategy. Suppose we estimate each parameter individually, without bias. Then the MSE $\mathbb{E}[(f(\theta) - \hat{f}(\theta))^2]$ can be written as
\[
M_{\text{unentangled}} = f_i(\theta)^2 \text{Var} \hat{\theta}_i. \quad (6)
\]

A measurement of a single parameter can be made in time $t$ with variance $\text{Var} \hat{\theta}_i = \frac{1}{t^2}$ [21]. Therefore, our entanglement-free figure of merit is
\[
M_{\text{unentangled}} = \frac{\| \nabla f(\theta) \|^2}{t_{\text{total}}^2}, \quad (7)
\]
where the $\| \cdot \|$ in Eq. (7) denotes the Euclidean norm. More generally, we use $\| v \|_p$ to denote the $p$-norm of vector $v$. Since Eq. (7) only saturates Eq. (5) in trivial cases where $\nabla f(\theta)$ is zero in all but one component, the unentangled protocol described is not optimal.

**Two-step Protocol.**—We now present a protocol which asymptotically saturates Eq. (5). Our protocol consists of two steps. First, we make an unbiased estimate $\hat{\theta}$ of $\theta$ for time $t_1$. Second, given our estimates $\hat{\theta}$, we make an unbiased measurement $\hat{q}$ of the quantity $q = \nabla f(\theta) \cdot (\theta - \hat{\theta})$ using the linear combination protocol in Ref. [15], which takes time $t_2$. Our final estimate is $\hat{f} = f(\theta) + \hat{q}$.

It can be shown that our protocol is optimal (in terms of scaling with the total time $t_1 + t_2$) provided that the individual estimations of the parameters satisfy $\mathbb{E}[(\hat{\theta}_i - \theta_i)^4] = \mathcal{O}(t_i^{-4})$ and that $t_1$ and $t_2$ are chosen properly. To simplify our computations, we will make the more concrete assumption that our initial estimates $\hat{\theta}$ are each normally distributed as $N(\theta_i, \text{Var} \hat{\theta}_i)$. Then as computed in Sec. S2 of the Supplemental Material [20], the figure of merit for this protocol is
\[
M = \mathbb{E}[(\hat{f}(\theta) + \hat{q} - f(\theta))^2] \quad (8) = \mathbb{E}[\text{Var} \hat{q}] + \frac{2f_{ij}(\hat{\theta}) + f_{ii}(\hat{\theta})f_{jj}(\hat{\theta})}{4} \text{Var} \hat{\theta}_i \text{Var} \hat{\theta}_j. \quad (9)
\]

In Eq. (9), the first term is the error resulting from the second phase of the protocol, estimating the linear combination. The second term is a residual error remaining from the first phase of the protocol after it is corrected by the linear combination measurement.

For our particular Hamiltonian $\hat{H} = \frac{i}{2} \theta_i \hat{\sigma}_i^z$, as per Ref. [15], we know that the minimum variance of an unbiased estimator of some linear combination $\alpha \cdot \theta$ given time $t$ is
\[
\text{Var} \alpha \cdot \theta \geq \frac{\max_i \alpha_i^2}{t^2}, \quad (10)
\]
which can be achieved with the entangled GHZ state $|\psi_{\text{spin}}\rangle = \frac{1}{\sqrt{2}}(|0\rangle^{\otimes d} + |1\rangle^{\otimes d})$. We can apply this linear combination protocol to the second phase of our protocol by setting $\alpha = \nabla f(\theta)$. For the individual estimators of the first phase, we use the fact that an individual estimation can be made in time $t$ with variance $1/t^2$ [1].
Using these results, we simplify Eq. (9):

\[
M = \mathbb{E} \left[ \max_i f_i(\tilde{\theta})^2 \right] + \frac{2f(\theta_1) + f_1(\theta)}{t_1^4} + \mathcal{O}(\theta^3),
\]

where we have absorbed the second derivatives into \( f_i(\tilde{\theta}) \), which does not depend on time. Without loss of generality, we designate \( f_1(\tilde{\theta}) \) as the largest \( f_i(\tilde{\theta}) \). We then expand \( \mathbb{E}[f_i(\tilde{\theta})^2] \) as

\[
f_i(\theta)^2 + \frac{f_1(\theta)f_{ii}(\theta)}{t_1^2} + \frac{f_{ii}(\theta)^2}{t_1^4} + \mathcal{O}(\theta^3),
\]

We may substitute Eq. (13) into Eq. (12) to obtain

\[
M = \frac{g_2(\theta)}{t_2^2} + \frac{g_1(\theta)}{t_1^2} + \frac{1}{t_1^4} + \mathcal{O}(\theta^3),
\]

where \( g_2(\theta) = f_i(\theta)^2 \) and \( g_1(\theta) \) have been introduced to absorb more time-independent factors.

Finally, we must specify how the total time \( t_{\text{total}} \) is to be allocated between \( t_1 \) and \( t_2 \). In Sec. S3 of the Supplemental Material [20], we show that the best possible allocation satisfies

\[
t_1 = g(\theta)t_{\text{total}}^{3/5},
\]

where \( g \) is a function which depends only on \( f \) and \( \theta \). In particular, \( t_1 = \mathcal{O}(t_{\text{total}}^{3/5}) \), so the fraction of time spent on \( t_1 \) vanishes as \( t_{\text{total}} \to \infty \). Almost all of the time is spent on \( t_2 \), the linear combination step of the two-step protocol. It can readily be shown (as a special case of the proof in Sec. S4 of the Supplemental Material [20]) that Eq. (14) is asymptotically dominated by the first term, which (since \( t_2 \to t_{\text{total}} \)) is equal to the right-hand-side of the bound in Eq. (5). In other words, this distribution of time asymptotically achieves the optimal MSE.

The two-step protocol exhibits Heisenberg scaling as defined for distributed sensing [14, 15, 22]. Comparing Eq. (7) to Eq. (5) shows an improvement of \( \mathcal{O}(d) \), maximized when all components of \( \nabla f(\tilde{\theta}) \) are approximately equal. Intuitively, the advantage is maximal when all parameters contribute, but minimal (i.e. no advantage) when only one parameter affects the function value. Similarly, behavior was noted in the linear combination case [15].

Note that when actually implementing the protocol, the optimal \( t_1 \) is unknown since the function \( g \) that determines it depends on the true parameters \( \theta \). However, we do not need to use the optimal \( t_1 \) to saturate the bound in Eq. (5). If \( t_1 \) is a function \( c t_{\text{total}}^p \) of the total time where \( \frac{2}{3} < p < 1 \) and some constant \( c \), then the protocol will saturate Eq. (9). Although these different times do result in a higher MSE, the additional error is \( \mathcal{O}(t_{\text{total}}^{-4}) \), which is insignificant asymptotically. The two-step protocol will therefore be asymptotically optimal for a wide range of time allocations. A proof of this fact is provided in Sec. S4 of the Supplemental Material [20].

**Function Measurement in Other Physical Settings.**—We now consider a different physical setting for function estimation. Rather than \( d \) qubits which accumulate phase for some time \( t \), we instead pass \( n \) photons through \( d \) Mach-Zehnder interferometers and accumulate some fixed phase \( \theta_1 \) encoded into each interferometer (see Fig. 2). For single parameters, the use of entangled states to reduce noise in this setting has been explored in Refs. [23–27] with multiparameter cases explored in Refs. [14, 22]. In this setting, the relevant limitation is the total number of photons used in the measurement, rather than time. This constraint is particularly relevant when analyzing a biological or chemical sample which is sensitive to light, making it desirable to reduce noise with as few photons as possible. Similar biologically motivated situations are presented in Refs. [28–30].

For photons, a two-step protocol with similar structure to the protocol for qubits yields reduced noise compared to any estimate of \( f \) derived entirely from local measurements. Suppose we allot \( N_1 \) photons for the first step (individual measurement) and \( N_2 \) photons for the second step (linear combination), for a total of \( N_{\text{total}} = N_1 + N_2 \) photons. We again begin from the general result of Eq. (9). However, the use of photons which can be apportioned between modes introduces new structure to the problem. We need to partition the \( N_1 \) photons into \( N_1 = n_1 + n_2 + \cdots + n_d \), putting \( n_i \) photons into the \( i \)-th interferometer, as some parameters may affect our final result more than others. Thus, in the second term of Eq. (9), we replace \( \text{Var} \tilde{\theta}_i \) with \( \frac{1}{n_i^2} \) instead of \( \frac{1}{t_i} \) [23].

The optimal variance when measuring the linear combination \( \alpha \cdot \theta \) using \( N \) total photons is unknown. However, Ref. [14] conjectures the optimal variance to be

\[
\text{Var} \tilde{\alpha \cdot \theta} \geq \frac{\|\alpha\|^2}{N^2}.
\]
Furthermore, Ref. [14] provides a protocol achieving the bound in Eq. (16) using a proportionally weighted GHZ state: $|\psi_{\text{photon}}\rangle = \frac{1}{\sqrt{2}}(|n_1, 0, n_2, 0 \ldots + |0, n_1, 0, n_2, 0 \ldots \rangle$, where $n_i = N_{\text{total}} \sum \alpha_j$ and where, in reference to Fig. 2, the modes are listed from top to bottom. Note that this will only work for $\alpha$ proportional to some rational vector as photons are discrete. Since Eq. (16) is saturable, we may simplify the first term of Eq. (9) to obtain

$$M = \frac{E}{N_2^2} \left| \nabla f(\theta) \right|_{L_2}^2 + \frac{2f_{ij}(\theta)^2 + f_{ij}(\theta)\nabla f(\theta)}{n_i^2n_j^2}. \quad (17)$$

For fixed $f$ and $\theta$, the $\frac{1}{n_i n_j}$ terms in Eq. (17) are minimized for the same ratio of $n_1 : n_2 : \cdots : n_d$ regardless of the value of the total number of photons used, $N_1$. Each term is proportional to $N_1^{-4}$ multiplied by some function of $f, \theta$, and $d$. Therefore, the structure of Eq. (17) becomes identical to the structure of Eq. (14), with $N_1$ and $N_2$ replacing $t_1$ and $t_2$. As a result, the optimal allocation of photons between $N_1$ and $N_2$ will yield $N_1 = O(N_1^{3/5})$ and $N_2 = O(N_1^{2/5})$, meaning that the $N_2^{-2}$ term in Eq. (17) is dominant asymptotically. Therefore, for photons, we may asymptotically achieve

$$M = \left| \nabla f(\theta) \right|_{L_2}^2 + O\left(\frac{1}{N_{\text{total}}^{2/5}}\right). \quad (18)$$

This strategy is optimal if the linear combination estimation strategy presented in Ref. [14] is optimal, as conjectured in that work.

Eq. (18) also exhibits Heisenberg scaling. Suppose we were to measure each parameter individually and then calculate the function. When measuring the parameters individually, we obtain the same error formula as Eq. (6), except now we set $\text{Var} \theta_i = \frac{1}{n_i}$ to get

$$M_{\text{untangled}} = \frac{f_i(\theta)^2}{n_i}. \quad (19)$$

The optimal distribution requires an $n_i$ proportional to the weight $f_i(\theta)^{2/3}$, yielding an entanglement-free error of

$$M_{\text{untangled}} = \frac{\left| \nabla f(\theta) \right|_{L_2}^2}{N_{\text{total}}^{2/3}}. \quad (20)$$

As with qubits, by comparing Eq. (18) with Eq. (20) in the case where all of the $f_i(\theta)$ are approximately equal, we find that the photonic two-step protocol yields a $O(d)$ improvement in error over measuring each parameter individually. This improvement when all quantities are equally important can also be seen in Ref. [22] for the special case of $f$ being a linear combination. As in the qubit case, the improvement in error is lessened when $\nabla f(\theta)$ is not approximately equal in all components.

In fact, this method can be extended still more generally. Rather than cases where the signal is imprinted on photons by a phase shift, we can consider the protocol developed in Ref. [16], which is capable of entanglement-enhanced distributed sensing of continuous variables by using homodyne measurements. Besides measuring parameters in different physical settings, we may also measure functions of variables coupled to spins, phase-shifts of photons, continuous variables, and any combination of these. The measurement protocol for parameters coupled to a combination of spins and photons is detailed in Sec. S5 of the Supplemental Material [20], but other combination states may be found analogously.

**Applications.**—As our protocol can measure any analytic function of $\theta$, it is widely applicable. In fact, there is no requirement that different $\theta_i$ have the same physical origin. For instance, a $\theta_1$ representing an electric field and $\theta_2$ measuring a magnetic field could be used to measure the Poynting vector.

One potential application of function measurements is the interpolation of non-linear functions. Suppose that an ansatz with $d$ tunable parameters is made for the strength of the field in a region. With readings from $\geq d$ different points, one could determine the parameters of the ansatz and therefore determine the value of the field at other points. Estimations of these ansatz parameters, which are functions of the measured fields, may potentially be improved using entangled states depending on the figure of merit [18, 31]. Note that this procedure can be carried out even if it is difficult to invert the ansatz in terms of the $d$ measurements, as described in Sec. S6 of the Supplemental Material [20].

Interpolation in this manner can proceed by two different schemes. We can either attempt to measure the ansatz parameters themselves, which allows computation of the field at all other points, or we can skip the final computation step by writing the field at a point of interest as a function of all the points that can be measured. This final function can then be directly measured using an entangled protocol, which will be more accurate. However, the first approach has the advantage that knowing the ansatz parameters allows estimation of all points in the space in question.

One particular interpolation of interest arises in ion trap quantum computing. In trapped ion chains, qubits are manipulated using Gaussian laser beams, and two primary sources of error are intensity and beam pointing fluctuations [32–34]. Our protocol offers better ways to characterize this noise. In order to detect the field error at a qubit’s position without disturbing the qubit, we can perform interpolation by measuring the field’s effect on other ions, possibly of a different atomic species, positioned nearby. Given the ansatz of the Gaussian beam profile, we are able to calculate the field at the qubit of interest and perhaps correct the error.

**Outlook.**—We have presented a Heisenberg-scaling
measurement protocol using quantum sensor networks for measuring any multivariate, real-valued, analytic function, and this protocol is consistent with the Heisenberg limit when measuring functions with comparably-sized gradients in each component. Future work may include proving the optimality of the two-step protocol when constrained by the number of photons, which would require extending the results of Ref. [14].

We specifically identified field interpolation as a promising application of our work, but we stress that our protocol can assist in the measurement of any analytic function. More work remains to determine when it is optimal to measure the coefficients of interpolation and when it is optimal to directly measure the final function. We are also interested in fleshing out possible intersections between quantum function estimation and machine learning. Supervised machine learning is a type of interpolation: estimating functional outputs for unknown inputs by extracting information from known input-output pairs [35]. It is possible our protocol could be used to improve the accuracy of training a machine learning model if the necessary quantity for training was a function of physical measurements. Additionally, the final output of many machine learning algorithms, such as neural networks, is a non-linear but infinitely differentiable function of the inputs [36]. Our work could aid in computing this complicated function for new input when making predictions.

We would like to thank M. Foss-Feig, S. Rolston, J. Gross, and S. Kimmel for helpful discussions. This work was supported by ARL CDQI, ARO MURI, ARO, NSF PFC at JQI, NSF Ideas Lab on Quantum Computing, the DoE ASCR Quantum Testbed Pathfinder program, and the DOE BES Materials and Chemical Sciences Research for Quantum Information Science program. Z.E. is supported in part by the ARCS Foundation Research for Quantum Information Science program, and the DOE BES Materials and Chemical Sciences Research, Volume I: Estimation Theory (Prentice Hall PTR, 1993) Chap. 3.

We would like to thank M. Foss-Feig, S. Rolston, J. Gross, and S. Kimmel for helpful discussions. This work was supported by ARL CDQI, ARO MURI, ARO, NSF PFC at JQI, NSF Ideas Lab on Quantum Computing, the DoE ASCR Quantum Testbed Pathfinder program, and the DOE BES Materials and Chemical Sciences Research for Quantum Information Science program. Z.E. is supported in part by the ARCS Foundation.

[1] J. J. Bollinger, W. M. Itano, D. J. Wineland, and D. J. Heinzen, Phys. Rev. A 54, R4649 (1996).
[2] S. F. Huelga, C. Macchiavello, T. Pellizzari, A. K. Ekert, M. B. Plenio, and J. I. Cirac, Phys. Rev. Lett. 79, 3865 (1997).
[3] M. G. A. Paris, Int. J. Quantum Inf. Suppl. 7, 125 (2008).
[4] L. Pezz and A. Smerzi, Phys. Rev. Lett. 102, 100401 (2009).
[5] G. Tóth, Phys. Rev. A 85, 022322 (2012).
[6] Z. Zhang and L. M. Duan, New J. Phys. 16, 103037 (2014).
[7] M. G. Genoni, M. G. A. Paris, G. Adesso, H. Nha, P. L. Knight, and M. S. Kim, Phys. Rev. A 87, 012107 (2013).
[8] P. C. Humphreys, M. Barbieri, A. Datta, and I. A. Walmsley, Phys. Rev. Lett. 111, 070403 (2013).
[9] Y. Gao and H. Lee, Eur. Phys. J. D 68, 347 (2014).
[10] M. D. Vidrighin, G. Donati, M. G. Genoni, X.-M. Jin, W. S. Kolthammer, M. S. Kim, A. Datta, M. Barbieri, and I. A. Walmsley, Nat. Commun. 5, 3532 (2014).
[11] J.-D. Yue, Y.-R. Zhang, and H. Fan, Sci. Rep. 4, 5933 (2014).
[12] Y.-R. Zhang and H. Fan, Phys. Rev. A 90, 043818 (2014).
[13] P. Kok, J. Dunningham, and J. F. Ralph, Phys. Rev. A 95, 012326 (2017).
[14] T. J. Proctor, P. A. Knott, and J. A. Dunningham, Phys. Rev. Lett. 120, 080501 (2018).
[15] Z. Eldredge, M. Foss-Feig, J. Gross, S. Rolston, and A. Gorshkov, Phys. Rev. A 97 (2018).
[16] Q. Zhuang, Z. Zhang, and J. H. Shapiro, Phys. Rev. A 97, 032329 (2018).
[17] S. M. Kay, in Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory (Prentice Hall PTR, 1993) Chap. 3.
[18] T. Baumgratz and A. Datta, Phys. Rev. Lett. 116, 9 (2015).
[19] S. Boixo, S. T. Flammia, C. M. Caves, and J. M. Geremia, Phys. Rev. Lett. 98 (2007).
[20] See the Supplemental Material at http://link.aps.org/ for more details on derivations used throughout this work. [21] D. J. Wineland, J. J. Bollinger, W. M. Itano, F. L. Moore, and D. J. Heinzen, Phys. Rev. A 46, R6797 (1992).
[22] W. G. K. Jacobs, Z. Eldredge, A. V. Gorshkov, and M. Foss-Feig, Phys. Rev. Lett. 121, 043604 (2018).
[23] M. Holland and K. Burnett, Phys. Rev. Lett. 71, 1355 (1993).
[24] T. Kim, O. Pfeifer, M. J. Holland, J. Noh, and J. L. Hall, Phys. Rev. A 57, 4004 (1998).
[25] A. R. Usha Devi and A. K. Rajagopal, Phys. Rev. A 79, 062320 (2009).
[26] R. Demkowicz-Dobrzański, M. Jarzyna, and J. Kołodyński (Elsevier, 2015) pp. 345–435.
[27] H. T. Dinani, M. K. Gupta, J. P. Dowling, and D. W. Berry, Phys. Rev. A 93, 063804 (2016).
[28] T. W. Kee and M. T. Cicerone, Opt. Lett. 29, 2701 (2004).
[29] O. Alem, T. H. Sander, R. Mhaskar, J. LeBlanc, H. Eswaran, U. Steinhoff, Y. Okada, J. Kitching, L. Trahms, and S. Knappe, Phys. Med. Biol. 60, 4797 (2015), 26041047.
[30] K. Jensen, R. Budvytyte, R. A. Thomas, T. Wang, A. Fuchs, M. V. Balabas, G. Vasilakis, L. Mosgaard, T. Heinburg, S.-P. Olesen, and E. S. Polzik, Sci. Rep. 6, 29638 (2016).
[31] G. Chiribella, G. M. D’Ariano, and M. F. Sacchi, Phys. Rev. A 72, 042348 (2005).
[32] J. I. Cirac and P. Zoller, Phys. Rev. Lett. 74, 4901 (1995).
[33] H. Häffner, C. F. Roos, and R. Blatt, Physics Reports 469, 155 (2008).
[34] K. R. Brown, A. C. Wilson, Y. Colombe, C. Ospelkaus, A. M. Meier, E. Knill, D. Leibfried, and D. J. Wineland, Phys. Rev. A 84, 030303 (2011).
[35] S. Russell and P. Norvig, in Artificial Intelligence, A Modern Approach (Prentice Hall, 2010).
[36] J. Schmidhuber, Neural Networks 61, 85 (2015).

We specifically identified field interpolation as a promising application of our work, but we stress that our protocol can assist in the measurement of any analytic function. More work remains to determine when it is optimal to measure the coefficients of interpolation and when it is optimal to directly measure the final function. We are also interested in fleshing out possible intersections between quantum function estimation and machine learning. Supervised machine learning is a type of interpolation: estimating functional outputs for unknown inputs by extracting information from known input-output pairs [35]. It is possible our protocol could be used to improve the accuracy of training a machine learning model if the necessary quantity for training was a function of physical measurements. Additionally, the final output of many machine learning algorithms, such as neural networks, is a non-linear but infinitely differentiable function of the inputs [36]. Our work could aid in computing this complicated function for new input when making predictions.
Supplemental Material for ”Heisenberg-Scaling Measurement Protocol for Analytic Functions with Quantum Sensor Networks”

Kevin Qian,1,2,3 Zachary Eldredge,1,2 Wenchao Ge,4 Guido Pagano,1,2,5 Christopher Monroe,1,2,6 James V. Porto,3 and Alexey V. Gorshkov1,2

1Joint Quantum Institute, NIST/University of Maryland, College Park, MD 20742, USA
2Joint Center for Quantum Information and Computer Science, NIST/University of Maryland, College Park, MD 20742, USA
3Montgomery Blair High School, Silver Spring, MD 20901, USA
4Institute for Quantum Science and Engineering (IQSE) and Department of Physics & Astronomy, Texas A&M University, College Station, Texas 77843, USA
5United States Army Research Laboratory, Adelphi, MD 20783, USA
6IonQ, Inc., College Park, MD 20740

In this Supplemental Material, we present detailed derivations for some of the results which appear in the main text. Sec. S1 derives the Quantum Cramér-Rao bound appropriate for our situation. Sec. S2 describes the detailed derivation of the figure of merit which can be used to evaluate the two-step protocol. Secs. S3 and S4 concern the optimal allocation of time between the two steps of the protocol. Sec S5 describes a measurement protocol and entangled state to measure a function of parameters coupled to a combination of spins and photons more accurately. Finally, Sec. S6 shows how our protocol can be used to assist in the estimation of ansatz parameters even if closed-form inverses for the ansatz variables are not known.

S1. Error Bound Derivation

In this section, we derive Eq. (5) in the main text. Specifically, we bound the accuracy of any measurement scheme’s estimate of \( f(\theta) \), starting with the Quantum Cramér-Rao Bound. The Cramér-Rao Bound is only applicable if \( F_Q \) is invertible. However, \( F_Q \) is only guaranteed to be positive semi-definite, and not necessarily invertible. We can resolve this issue by projecting \( F_Q \) onto its own image [S1], and letting this new matrix be denoted \( \tilde{F}_Q \).

Let \( b \) be an arbitrary vector, so that

\[
\delta^T F_Q^{-1} \delta = \frac{\| \sqrt{\tilde{F}_Q^{-1}} \delta \|_2^2}{b^T \tilde{F}_Q b}.
\]  

(S1)

Then by the Cauchy-Schwartz Inequality,

\[
\delta^T F_Q^{-1} \delta \geq \frac{\| \delta^T \sqrt{\tilde{F}_Q^{-1}} \sqrt{\tilde{F}_Q b} \|_2^2}{b^T \tilde{F}_Q b}.
\]  

(S2)

\[
\geq \frac{\| \delta^T b \|_2^2}{b^T \tilde{F}_Q b}.
\]  

(S3)

Taking \( b \) to be the \( b \)-th element of the standard basis gives

\[
M = \mathbb{E}[(\hat{f} - f(\theta))^2] \geq \text{Var} \hat{f} \geq \frac{\delta_b^2}{(\tilde{F}_Q)_{bb}}.
\]  

(S4)

Notice that \( (\tilde{F}_Q)_{bb} \) is equivalent to the quantum Fisher information for a single parameter. In Ref. [S2], it was shown that for any Hamiltonian coupled to parameters, including those with a time-dependent control Hamiltonian \( \hat{H}_c(t) \) and ancilla bits,

\[
(\tilde{F}_Q)_{bb} \leq t_{\text{total}}^2 \left\| \frac{\partial H}{\partial \theta_b} \right\|_s^2.
\]  

(S5)
where \( \| \hat{g} \|_s \) denotes the seminorm of the operator, the difference between the maximal and minimal eigenvalues of \( g \). For the Hamiltonian presented in this work, \( \frac{\partial \hat{H}}{\partial \theta_k} = \frac{1}{2} \hat{\sigma}_k \) and its seminorm is 1. Since Eq. (S4) must hold for all values of \( b \), then combining Eq. (S4) and Eq. (S5) yields Eq. (5) in the main text:

\[
M \geq \text{Var} \hat{f} \geq \frac{\max_i \delta_i^2}{\sum_{i=1}^{n} \delta_i^2} = \frac{\max_i f_i(\theta)^2}{\sum_{i=1}^{n} f_i(\theta)^2_{\text{total}}} .
\]

(S6)

### S2. Figure of Merit for Two-step Protocol

In this section, we derive Eq. (9) in the main text. Specifically, we derive the figure of merit for the two-step protocol in terms of the measurement accuracy of the independent parameters and the measurement accuracy of the linear combination, yielding a general formula which applies to any physical realization.

For the sake of concision, let \( \Delta = \hat{\theta} - \theta \) which satisfies \( \mathbb{E}[\Delta] = 0 \). Furthermore, let \( T_k \) be \( k! \) times the \( k \)-th term of the Taylor expansion of \( f \) (so \( T_1 = f_i(\theta) \Delta_i, T_2 = f_{ij}(\theta) \Delta_i \Delta_j, T_3 = f_{ijk}(\theta) \Delta_i \Delta_j \Delta_k, \) etc.). Thus, the Taylor expansion of \( f(\bar{\theta}) \) would be

\[
f(\bar{\theta}) = f(\theta) + T_1 + \frac{T_2}{2} + \frac{T_3}{6} + \ldots
\]

(S7)

We compute our figure of merit:

\[
M = \mathbb{E}[ (f(\bar{\theta}) + \bar{q} - f(\theta))^2 ]
\]

\[
= \mathbb{E}[ (f(\bar{\theta}) - f(\theta))^2 ] + \mathbb{E}[\bar{q}^2] + 2\mathbb{E}[f(\bar{\theta})\bar{q}] - 2f(\theta)\mathbb{E}[\bar{q}]
\]

\[
= \left( \mathbb{E}[T_1^2] + \mathbb{E}[T_1 T_2] + \frac{1}{3} \mathbb{E}[T_1 T_3] + \frac{1}{4} \mathbb{E}[T_2^2] + \mathcal{O}(\Delta^5) \right) + \left( \mathbb{E}[\text{Var}(\bar{q})] + \mathbb{E}[\bar{q}^2] \right)
\]

\[
+ 2 \left( f(\theta)\mathbb{E}[q] + \mathbb{E}[T_1 q] + \frac{1}{2} \mathbb{E}[T_2 q] + \frac{1}{6} \mathbb{E}[T_3 q] + \mathcal{O}(\Delta^5) \right) - 2f(\theta)\mathbb{E}[q]
\]

\[
= \mathbb{E}[\text{Var}(\bar{q})] + \mathbb{E}[(q + T_1)^2] + \mathbb{E}[(q + T_1) T_2] + \frac{1}{3} \mathbb{E}[(q + T_1) T_3] + \frac{1}{4} \mathbb{E}[T_2^2] + \mathcal{O}(\Delta^5).
\]

(S10)

(S11)

The actual computation of the labeled terms is rather involved and space consuming, so it is presented in a separate subsection of this Supplemental Material (see Sec. S2 A). Notice that we may simplify

\[
q + T_1 = \Delta_i(f_i(\theta) - f_i(\bar{\theta}))
\]

\[
= -\Delta_i(f_{ij}(\theta) \Delta_j + \mathcal{O}(\Delta^2))
\]

\[
= -T_2 + \mathcal{O}(\Delta^3),
\]

(S12)

(S13)

(S14)

so Eq. (S11) evaluates to

\[
M = \mathbb{E}[\text{Var}(\bar{q})] + \mathbb{E}[T_2^2] - \mathbb{E}[T_2^2] - \frac{1}{3} \mathbb{E}[T_2 T_3] + \frac{1}{4} \mathbb{E}[T_2^2] + \mathcal{O}(\Delta^5)
\]

\[
= \mathbb{E}[\text{Var}(\bar{q})] + \frac{1}{4} \mathbb{E}[T_2^2] + \mathcal{O}(\Delta^5)
\]

(S15)

(S16)

since \( \mathbb{E}[T_2 T_3] \) is \( \mathcal{O}(\Delta^5) \). Now, this simplifies further as

\[
M = \mathbb{E}[\text{Var}(\bar{q})] + \frac{1}{4} \mathbb{E}[T_2^2]
\]

\[
= \mathbb{E}[\text{Var}(\bar{q})] + \frac{1}{4} \mathbb{E}[(f_{ij}(\theta) \Delta_i \Delta_j)^2]
\]

\[
= \mathbb{E}[\text{Var}(\bar{q})] + \frac{1}{4} \mathbb{E} \left[ \sum_{i<j} f_{ij}(\theta)^2 \Delta_i^2 \Delta_j^2 + 2 \sum_{i<j} f_{ij}(\theta) f_{jj}(\theta) \Delta_i^2 \Delta_j^2 + \sum_i f_{ii}(\theta)^2 \Delta_i^2 \right]
\]

(S17)

(S18)

(S19)
since all terms with some $\Delta_i$ to a single power will factor out as $E[\Delta_i] = 0$. We will assume that $\Delta_i \sim \mathcal{N}(0, \frac{1}{T_i})$ is normally distributed. This is not strictly necessary as long as the distribution of errors satisfies $E[\Delta_i^2] \leq \mathcal{O}(t_i^{-1})$, a condition that is satisfied by phase estimation procedures like those in Ref. [S3]. However, assuming normality allows the calculation to proceed easily, as we will be able to simplify $E[\Delta_i^2] = 3\text{Var}\theta_i$. Thus, we arrive at

$$M = E[\text{Var}q]\hat{q} + \frac{1}{4}(4 \sum_{i<j} f_{ij}(\theta)^2 \text{Var} \hat{\theta}_i \text{Var} \hat{\theta}_j + 2 \sum_{i<j} f_{ii}(\theta)f_{jj}(\theta) \text{Var} \hat{\theta}_i \text{Var} \hat{\theta}_j + \sum_i 3f_{ii}(\theta)^2 \text{Var} \theta_i^2)$$

(S20)

$$= E[\text{Var}q]\hat{q} + \sum_{i,j} \frac{2f_{ij}(\theta) + f_{ii}(\theta)f_{jj}(\theta)}{4} \text{Var} \hat{\theta}_i \text{Var} \hat{\theta}_j.$$  

(S21)

### A. Simplification of labeled terms

In this subsection, we present the simplification of the labeled terms from Eqs. (S9-S11) in full detail.

Term 2 is simplified by using the definition of $\text{Var}q\hat{q}$. One needs to be careful as there are two layers of expected values - one for the values of $q$ and one for the estimator $q$:  

$$E[q]^2 = E_\theta[E_q[q]^2]$$

(S22)

$$= E_\theta[\text{Var}q\hat{q} + E_q[\hat{q}]^2]$$

(S23)

$$= E_\theta[\text{Var}q\hat{q} + q^2]$$

(S24)

$$= E[\text{Var}q\hat{q}] + E[q^2].$$

(S25)

Terms 1 and 3 are simplified by expanding the Taylor series for $f(\hat{\theta})$ up to $\Delta^4$ terms; note that $q = \mathcal{O}(\Delta)$, so we only need to expand the Taylor series up to $\mathcal{O}(\Delta^3)$ terms:

$$E\left[\left(f(\hat{\theta}) - f(\theta)\right)^2\right] = E[f(\hat{\theta})^2] - 2f(\theta)E[f(\hat{\theta})] + f(\theta)^2$$  

(S26)

$$= f(\theta)^2 + E[T_1^2] + f(\theta)E[T_2] + E[T_1T_2] + \frac{1}{3}f(\theta)E[T_3]$$

$$+ \frac{1}{12}f(\theta)E[T_4] + \frac{1}{3}E[T_1T_3] + \frac{1}{4}E[T_2^2] + \mathcal{O}(\Delta^5)$$

(S27)

$$- 2f(\theta)\left(f(\theta) + \frac{1}{2}E[T_2] + \frac{1}{6}E[T_3] + \frac{1}{24}E[T_4] + \mathcal{O}(\Delta^5)\right) + f(\theta)^2$$

$$= E[T_1^2] + E[T_1T_2] + \frac{1}{3}E[T_1T_3] + \frac{1}{4}E[T_2^2] + \mathcal{O}(\Delta^5).$$

(S28)

$$E[f(\hat{\theta})\hat{q}] = E_\theta[E_q[f(\hat{\theta})\hat{q}]]$$

(S29)

$$= E_\theta[f(\hat{\theta})q]$$

(S30)

$$= E\left[\left(f(\theta) + T_1 + \frac{T_2}{2} + \frac{T_3}{6} + \mathcal{O}(\Delta^4)\right)q\right]$$

(S31)

$$= f(\theta)E[q] + E[T_1q] + \frac{E[T_2q]}{2} + \frac{E[T_3q]}{6} + \mathcal{O}(\Delta^5).$$

(S32)

### S3. Optimal allocation of time for the two steps

In this section, we derive Eq. (15) in the main text, which specifies the optimal allocation of time in the limit $t_{\text{total}} \to \infty$. We want to choose the $t_1, t_2$, under the constraint that $t_1 + t_2 = t_{\text{total}}$, which minimize the MSE

$$M = \frac{g_2(\theta)}{t_2^2} + \frac{g_3(\theta)}{t_1^2t_2^2} + \frac{g_1(\theta)}{t_1^2}.$$  

(S33)
Notice that the $g_1, g_2, g_3$ functions are only dependent on $\theta$ and not $t_1$, so we may set the derivative of $M$ with respect to $t_1$ equal to 0 and obtain
\[
\frac{2g_2(\theta)}{t_2^3} + \frac{2g_3(\theta)}{t_1^2t_2^2} = \frac{2g_3(\theta)}{t_2^3t_1^2} + \frac{4g_1(\theta)}{t_1^3}.
\] (S34)

Let $r = t_1/t_2$. Then we may rearrange to obtain
\[
g_2(\theta)t_2^3 = \frac{g_3(\theta)}{r} + \frac{2g_1(\theta)}{r^3} - g_3(\theta).
\] (S35)

Since $t_1 \gg 1$, then $r \ll 1$, so the $r^{-3}$ term dominates the RHS. Thus, $g_2(\theta) t_2^3 \approx \frac{2g_1(\theta)}{r}$, which implies
\[
t_1 \approx \left( \frac{2g_1(\theta)}{g_2(\theta)} \right)^{1/5} t_2^{3/5} \approx \left( \frac{2g_1(\theta)}{g_2(\theta)} \right)^{1/5} t_{\text{total}}^{3/5}.
\] (S36)

### S4. Proof that $t_1 = c t_{\text{total}}^p$ saturates bound

In this section, we prove that a broad class of time allocations saturate our lower variance bound (and hence make our protocol optimal). Recall that the MSE derived was
\[
M = \frac{g_2(\theta)}{t_2^3} + \frac{g_3(\theta)}{t_1^2t_2^2} + \frac{g_1(\theta)}{t_1^3},
\] (S37)

Suppose that $t_1 = c t_{\text{total}}^p$ for some $\frac{1}{2} < p < 1$ and some constant $c$. Since $p < 1$, we see that $\lim_{t_{\text{total}} \to \infty} \frac{t_2}{t_{\text{total}}} = 1$. Therefore, we may substitute our $t_1$ into the MSE formula in Eq. (S37) and simplify:
\[
\lim_{t_{\text{total}} \to \infty} M = \lim_{t_{\text{total}} \to \infty} \frac{g_2(\theta)}{t_2^3} + \frac{g_3(\theta)}{c^2t_{\text{total}}^2} + \frac{g_1(\theta)}{c^3t_{\text{total}}^3}.
\] (S38)

Since $p > \frac{1}{2}$, the $t_{\text{total}}^2$ term is dominant. Thus, as we defined $g_2 := f_1(\theta)^2 = \max_i f_i(\theta)^2$ under the assumption that $f_1(\theta)^2$ was maximal in Sec. S3, our asymptotic error is
\[
M = \max_i f_i(\theta)^2,
\] (S39)

which saturates the bound of Eq. (5) in the main text.

Other possible power-law scalings for $t_1$ fail. If $p \leq \frac{1}{2}$, the last term in Eq. (S38) becomes significant asymptotically and prevents the protocol from saturating Eq. (5) in the main text. If $p = 1$, then of course $c \leq 1$ or $t_1 > t_{\text{total}}$. In this case, we can no longer claim that $t_2$ approaches $t_1$. Even though the $g_2(\theta)^2$ would remain dominant, it would be scaled larger by $\frac{t_2^3}{(1-c)^3}$, which is always larger than 1. Hence, the protocol would no longer saturate the bound in Eq. (5) in the main text.

### S5. Function Measurement with Spins and Photons

In the main text, we considered scenarios in which parameters are either all coupled to qubits (spins) or all coupled to photons (interferometer modes). In this section, we discuss how to measure a function parameters coupled to a combination of spins and interferometers.

In such a hybrid scenario, we can still make use of the two-step protocol. The first step, obtaining initial estimates for the individual parameters, proceeds equivalently, since the measurements of the spins and of the photons can be viewed as occurring in parallel. For the linear combination case, we can assume that the optimal spin and photon input states can be entangled as follows:
\[
|\psi_{\text{spin-photon}}\rangle = \frac{1}{\sqrt{2}} \left( |n_1, 0, n_2, 0, \ldots \rangle \otimes |1, 1, 1, \ldots \rangle + |0, n_1, 0, n_2, \ldots \rangle \otimes |0, 0, 0, \ldots \rangle \right).
\] (S40)
Here, \( n_i = N_{\text{total}} \sum \alpha_i \), where the sum runs over only the \( j \) corresponding to photonic modes, denotes the number of photons which pass through the arms of the \( i \)-th interferometer. The state in Eq. (S40) is designed in such a way that the two branches of the overall wavefunction accumulate relative to each other a phase equal to the total linear combination we are interested in. In order to extract this final phase, the state can be unitarily mapped onto a qubit, which contains all of the accumulated phase and is then measured.

One caveat is that the linear combination protocol will accumulate phase proportional to time for the qubits and phase proportional to the number of photons for interferometers. For instance, if \( \theta_1 \) is coupled to a qubit (and therefore has units of frequency) and \( \theta_2 \) is coupled to an interferometer (and is therefore unitless), then the two branches of our state accumulate a relative phase \( \theta_1 t + \theta_2 n \). Therefore, one may have to adjust \( t \) or \( n \) in order to get the desired linear combination.

**S6. Interpolation of Functions Without a Closed-form Inverse**

In this section, we show how our protocol can be used to assist in the estimation of ansatz parameters even if closed-form inverses for the ansatz variables are not known.

Suppose that \( \theta = f(c, x) \) and that \( c = f^{-1}(\theta, x) \) exists, but has no closed-form solution which can be easily evaluated. First, we make measurements \( \hat{\theta} \). To create an initial estimate of the values \( c \), we use a numerical root-finder to find estimates \( \hat{c} \). We can now implement the second step of our protocol by finding the first derivatives \( \partial c_i / \partial \theta_j \) using the matrix identity \( \partial \theta / \partial c \cdot \partial c / \partial \theta = I \). Since \( f \) is known, \( \partial \theta / \partial c \) can be inverted to yield the \( \partial c / \partial \theta \) needed to estimate \( \hat{q} = \partial c / \partial \theta |_{\theta = \hat{\theta}} \cdot (\theta - \hat{\theta}) \). Our final estimate is \( \hat{c} + \hat{q} \), which was obtained without having to compute \( f^{-1} \) in general.

---

[S1] Timothy J. Proctor, Paul A. Knott, and Jacob A. Dunningham, “Multiparameter Estimation in Networked Quantum Sensors,” Phys. Rev. Lett. **120**, 80501 (2018).

[S2] Sergio Boixo, Steven T Flammia, Carlton M Caves, and J M Geremia, “Generalized limits for single-parameter quantum estimation.” Phys. Rev. Lett. **98** (2007).

[S3] Shelby Kimmel, Guang Hao Low, and Theodore J. Yoder, “Robust Calibration of a Universal Single-Qubit Gate-Set via Robust Phase Estimation,” (2015).