# Efficient quantum communications with coherent state fingerprints over multiple channels

Niraj Kumar,[1,2] Eleni Diamanti,[1] and Iordanis Kerenidis[2,3]
[1]*LIP6, CNRS, Université Pierre et Marie Curie, Sorbonne Universités, 75005 Paris, France*
[2]*IRIF, CNRS, Université Paris Diderot, Sorbonne Paris Cité, 75013 Paris, France*
[3]*CQT, National University of Singapore, Singapore*
(Received 3 October 2016; published 31 March 2017)

We provide an example of a communication model and a distributed task, for which there exists a realistic quantum protocol that is asymptotically more efficient than any classical protocol, both in the communication and the information resources. For this, we extend a recently proposed coherent state mapping for quantum communication protocols, study the use of coherent state fingerprints over multiple channels, and show their role in the design of an efficient quantum protocol for estimating the Euclidean distance of two real vectors within a constant factor.

DOI: 10.1103/PhysRevA.95.032337

## I. INTRODUCTION

Quantum information processing harnesses the power of quantum mechanics in order to enhance the efficiency and security of information and communication technologies. This is illustrated, for instance, in nonlocal games, where experiments have confirmed the violation of Bell inequalities that correspond to the CHSH and other games [1–4]. Another prominent example is quantum cryptography, where many protocols with unconditionally stronger security than classically possible have been demonstrated, including quantum key distribution, digital signatures, or coin flipping [5–7]. Unlike the above-mentioned cases, most quantum algorithms are far from implementable with current technologies, with the exception of nonuniversal boson sampling machines that have been realized for small inputs [8–10].

Communication complexity is an ideal model for testing quantum mechanics and for understanding the efficiency of quantum networks. This model studies the amount of communication required by separate parties to jointly compute a task. There are several examples where communicating quantum information can result in considerable savings in the communication overhead [11–17]. Nevertheless, it is in general difficult to test these results experimentally and demonstrate quantum superiority in practice since the quantum protocols typically necessitate large, highly entangled states, which are out of reach of current photonic technologies.

Recently, Arrazola and Lütkenhaus proposed a mapping for encoding quantum communication protocols involving pure states of many qubits, unitary operations, and projective measurements to protocols based on coherent states of light in a superposition of optical modes, linear optics operations, and single-photon detection [18]. This powerful model was used to propose the practical implementation of coherent state quantum fingerprints [19], leading to two experimental demonstrations: a proof-of-principle use of such fingerprints for solving the communication task of equality asymptotically better than the best known classical protocol with respect to the transmitted information [20], and a subsequent implementation beating the classical lower bound for the transmitted information [21]. Following these demonstrations that have focused on equality and on transmitted information, an important question remains: Is there a realistic model for proving and testing in practice that quantum information is asymptotically better than classical information for communication tasks with respect to *all* important communication and information resources?

We answer in the affirmative by proposing a communication model and a task for which we prove that quantum mechanics allows for a considerably more efficient protocol in all relevant resources. We do this by building upon the mapping of Ref. [18] to introduce coherent state fingerprints over multiple channels and show how to use them for solving efficiently a task that is at the foundation of many applications in machine learning, namely estimating the Euclidean distance of two real vectors within a constant factor. Our results show that, in principle, it is possible to demonstrate quantum superiority for advanced communication tasks in quantum networks using photonic technologies within experimental reach.

## II. COMMUNICATION RESOURCES

We start by defining the simultaneous message passing model and the resources that we are trying to optimize. In this model, two players, Alice and Bob, receive inputs $x$ and $y$ respectively from an input set $\mathcal{X} \times \mathcal{Y}$. Their task is to use some private coins and send a single and smallest possible message to a referee, who should be able to compute the value of a function $f(x,y)$ with a small error $\delta \in \{0,1\}$.

The communication cost of the protocol is the number of bits the two players have to send to the referee and the communication complexity of the task is the minimum communication cost over all protocols that solve the task. In real world communication networks, very often the cost is rather calculated as the time one uses the communication channel, for example, on the phone network. We note that these costs are interchangeable, provided that the communication channel has a specific maximum rate. We define the time unit as the time to send a single bit over the communication channel and then, in an optimal protocol, bits and communication time are equal, since one will always send one bit per time unit. Another resource one can study is the transmitted information, which, instead of the number of bits sent, calculates the real bits of information about the inputs that the messages carry. For example, if Alice always sends the same, long message, independent of her input, then the communication time will be

 

large, while the transmitted information will be zero, since no information about the input has been transmitted. Transmitted information is a resource that is important for privacy, when on top of having an efficient protocol, we want the referee to solve the task without learning much about the players' inputs. One can define the transmitted information as the mutual information between the messages and the inputs and can upper bound it with the logarithm of the number of different messages. The transmitted information is always at most the communication time, since one bit carries at most one bit of information, and hence the bottleneck is always the time.

We can similarly define the resources for quantum protocols. The communication time is again the number of time units the protocol takes, where in a time unit at most one qubit can be sent in expectation. Here, we have added "in expectation" since typically in quantum communications the qubits are realized by photons emitted by practical light sources and hence their mean number follows a Poisson distribution [5]. In the following, we also make this change to the classical model to make a more correct comparison; i.e., we allow one bit in expectation per time unit, which does not change the order of the communication time. We will also upper bound the transmitted information as the logarithm of the minimum dimension of the Hilbert space that contains all the possible quantum messages that are sent in the protocol. For example, if Alice has as input an $n$-bit string $x$ and sends a message that contains $n/2$ qubits of the form $|x_1 x_2..x_{n/2}\rangle$ and another $n/2$ qubits in the state $|0\rangle$, then the communication time is $n$ while the transmitted information is $n/2$.

## III. EUCLIDEAN DISTANCE OF REAL VECTORS

We now describe a fundamental communication task. Alice and Bob possess large data sets $x$ and $y$ respectively, which are unit vectors in $\mathcal{R}^n$. They would like to allow a referee to check how similar their data are by estimating the Euclidean distance (for simplicity we define its square), $||x-y||_2^2 = \sum_{j=1}^n (x_j - y_j)^2$ (or equivalently the inner product, since $\langle x,y\rangle = 1 - ||x-y||_2^2/2$). We call this problem Euclidean distance or ED.

Alice and Bob can transmit their entire data to the referee, but this is nonoptimal. The idea is to send fingerprints of the data, which are much shorter but still allow the referee to approximate their Euclidean distance within some additive constant. Classically, this problem requires Alice and Bob to send fingerprints of size $\Omega(\sqrt{n})$ [22–25]. We consider here that Alice and Bob do not have access to any shared randomness, otherwise the problem can be solved with only constant communication [26]. It is natural that parties do not *a priori* have such shared resources, especially in large networks where the communication is between many different pairs of parties.

Quantum fingerprints can be exponentially shorter than the classical ones in this case. In particular, Alice and Bob create and send quantum fingerprints, namely $|\text{fin}_x\rangle = \sum_{j=1}^n x_j |j\rangle$ and $|\text{fin}_y\rangle = \sum_{j=1}^n y_j |j\rangle$, respectively. The referee then estimates the Euclidean distance by performing a controlled swap operation on the fingerprints, which outputs "1" with probability $1/2 + |\langle x,y\rangle|^2/2$ [11]. The referee estimates this probability, and hence the Euclidean distance, within an additive constant $\epsilon$ with probability at least $1-\delta$, using a constant number of fingerprints equal to $O[\log(1/\delta)/\epsilon^2]$. The communication time is $O(\log n)$, since Alice and Bob send a constant number of fingerprints and each fingerprint consists of $\log n$ qubits and can be sent in $\log n$ time units. The transmitted information is also $O(\log n)$, equal to the communication time. Hence, if we look at the ratio of the quantum over the classical resources, then both resources asymptotically go to zero as $n$ grows. Unfortunately, implementing these fingerprints with qubit systems is out of reach for current technologies for large $n$.

The notion of quantum fingerprints has been used in practice for the equality problem [19–21], where the inputs are binary strings and the referee checks whether they are exactly the same. The equality problem can be reduced with the help of error correcting codes to approximating the Euclidean distance between the two vectors within a constant factor and hence the previous protocol solves equality with the same resources. Since most real data are represented as real-valued vectors, the Euclidean distance problem is more pertinent than equality, since it is rather improbable that two different sets of real-valued data will be exactly equal. Hence, here, we extend the use of the term *quantum fingerprints* to real-valued inputs, where we check that fingerprints can be used to approximate the distance of the inputs and do not check whether they are exactly equal.

## IV. COHERENT STATE FINGERPRINTS FOR EUCLIDEAN DISTANCE

The coherent state mapping of Ref. [18] led to a protocol for equality with communication time $O(n)$ and transmitted information $O(\log n)$. This protocol therefore provides an exponential advantage in the transmitted information, at the expense of a quadratically worse performance in communication time compared to the classical protocol, for which the order of both resources is $\Omega(\sqrt{n})$.

A schematic of the corresponding protocol for Euclidean distance is shown in Fig. 1. Alice and Bob's fingerprints are trains of $n$ coherent states sent to the referee. Alice's state, $|\alpha_x\rangle$, is prepared by the displacement operator $\hat{D}_x(\alpha) = \exp(\alpha \hat{a}_x^\dagger - \alpha^* \hat{a}_x)$ applied to the vacuum state, where $\hat{a}_x = \sum_{j=1}^n x_j \hat{b}_j$ is the annihilation operator of the fingerprint mode [19], and $\hat{b}_j$ is the photon annihilation operator of the $j$th time mode. Hence,

$$|\alpha_x\rangle = \hat{D}_x(\alpha)|0\rangle = \otimes_{j=1}^n |x_j \alpha\rangle_j, \tag{1}$$

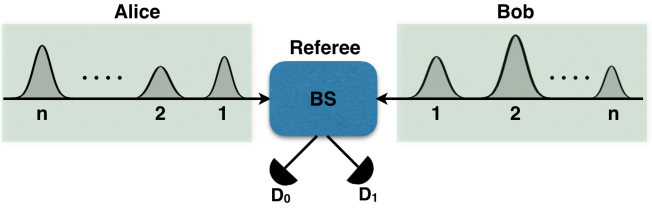


FIG. 1. Alice and Bob send $n$ coherent pulses, with the $j$th pulse's amplitude determined by $x_j\alpha$ and $y_j\alpha$, respectively. The referee interferes their states in a 50/50 BS and detects the output signals using single-photon detectors $D_0$ and $D_1$.

where $|x_j\alpha\rangle_j$ is a coherent state with amplitude $x_j\alpha$ occupying the $j$th mode. The mean photon number for the state $|\alpha_x\rangle$ is $\mu = \sum_j |x_j\alpha|^2 = |\alpha|^2$, independent of the input size. Bob similarly creates the fingerprint $|\alpha_y\rangle$.

As shown in Fig. 1, the referee uses a 50/50 beam splitter (BS) to interfere the incoming coherent states. This yields the output state for the $j$th time unit

$$\left|\frac{(x_j+y_j)}{\sqrt{2}}\alpha\right\rangle_{j,D_0}\otimes\left|\frac{(x_j-y_j)}{\sqrt{2}}\alpha\right\rangle_{j,D_1}, \tag{2}$$

where the subscripts $D_0$ and $D_1$ denote the single-photon detectors placed at the output arms of the BS.

Previously, for equality, only the clicks of $D_1$ have been used for the estimation. Then, since the expected number of clicks of the detector depends directly on its dark count probability, it is crucial to keep this probability very low. Here, we try to deal with this problem, by using the clicks from both detectors to construct a more robust estimator for the ED that can also be used for equality.

More precisely, let $Z_j^0$ and $Z_j^1$ be the binary random variables that are 1 with the probability with which $D_0$ and $D_1$ clicks respectively at the $j$th time unit, namely $p_j^0 = 1 - \exp(-\frac{\mu(x_j+y_j)^2}{2}) \approx \mu\frac{(x_j+y_j)^2}{2}$, and $p_j^1 = 1 - \exp(-\frac{\mu(x_j-y_j)^2}{2}) \approx \mu\frac{(x_j-y_j)^2}{2}$. Here, the approximation holds because we take $\mu$ to be typically small, and $x$ and $y$ are unit vectors in $\mathcal{R}^n$ and for large $n$ the terms $(x_j+y_j)^2$ and $(x_j-y_j)^2$ are typically in the order of $1/n$. The Euclidean distance ($\tilde{E}$) is equal to

$$\tilde{E} = 2 - \frac{1}{\mu}\mathbb{E}\left[\sum_{j=1}^{n}\left(Z_j^0 - Z_j^1\right)\right]. \tag{3}$$

The advantage of using statistics from both detectors comes from the fact that the Euclidean distance estimator depends now on the difference of the clicks of the detectors, and hence on expectation the number of dark counts cancels out, when we assume the dark count probabilities are the same for both detectors. We remark that this can be enforced by symmetrization procedures [7], although in practice, since the symmetrization will not be perfect, the estimator will in fact depend on the square of the dark count probability, which is easier to keep low.

By a Chernoff type argument [27], we can optimize the experimental parameters so that we can estimate $\sum_{j=1}^{n}(Z_j^0 - Z_j^1)$ within a factor $\epsilon$ with probability at least $1-\delta$. For constant $\epsilon, \delta$, the overall communication time is $O(n)$ while the transmitted information is $O(\mu \log n)$. Note that in each time unit, $\mu/n \ll 1$ photons are sent in expectation, thus satisfying our model's criterion of no more than one photon in each time unit.

The first two rows in Table I summarize the resources of the two protocols for ED. The performance achieved with the coherent state fingerprint protocol is the same as for equality, i.e., exponentially better in transmitted information but quadratically worse in communication time. We describe now a quantum protocol that can outperform any classical protocol in both resources.

TABLE I. The order of the communication time and transmitted information for all classical and quantum protocols for Euclidean distance described in this work.

| | Comm. Time | Trans. Info. |
|---|---|---|
| Classical | $\Omega(\sqrt{n})$ | $\Omega(\sqrt{n})$ |
| Coherent | $O(n)$ | $O(\mu \log n)$ |
| Multiple-channel classical | $\Omega(\frac{\sqrt{n}}{\log k})$ | $\Omega(\frac{\sqrt{n}}{\log k})$ |
| Multiple-channel coherent | $\frac{n}{k}$ | $O(\mu \log n)$ |

## V. COHERENT STATE FINGERPRINTS OVER MULTIPLE CHANNELS

We extend both the classical and quantum communication models to allow Alice and Bob to have multiple channels with the referee. Our model is not the usual multiplexing model where the parties can send $t$ bits or photons in parallel through $t$ channels and the referee can process all $t$ bits or photons in one time step. This model in fact would not make any difference in the relative power of the classical and quantum case, since the communication time will just be divided by $t$ in both cases. The important distinction we make in our model concerns the difference between the number of channels used by Alice and Bob for sending their bits or photons, and the number of bits or photons that we allow the referee to be able to process in parallel in one time unit. As we said, if these two parameters are the same, then both classical and quantum protocols save the same factor. Hence, in our model, we need to have more channels than the number of bits or photons that can be processed in order to take advantage of the fact that the quantum protocol sends mostly empty pulses. Moreover, since we want to examine the relative power of communication, we make sure that all other parameters are the same in the classical and quantum setting: We allow the same number of channels between parties, the same number of bits or photons that can be processed in parallel, and the number of channels could potentially be larger than the number of bits or photons that can be processed in parallel.

More concretely, we consider a model where we allow the referee to have the ability in one time unit to actually process (meaning to receive, perform some computation, and write to memory) $t$ bits of information that come from $t$ bits or $t$ photons that have been received in parallel. For example, the referee can have a multicore processor and parallel access to memory. We also allow the sender and the referee to communicate by using multiple channels to send bits or pulses. Let us assume that there are $m$ such channels and that for each channel the sender can prepare and send a bit or pulse and the receiver can detect a bit or a photon (if there exists one in the pulse). These two parameters, $t$ and $m$, may not necessarily be the same. In fact, in our model we are interested in the case where $m$ is bigger than $t$, and so we can write $m = kt$. Note that even though we have more than $t$ channels, there is no point in sending more than $t$ bits and photons in parallel through these channels, since the referee will not be able to process all the bits and photons. In the rest of the paper, we take $t = 1$ and hence $m = k$, because in fact our analysis shows that the ratio between the classical and quantum time is independent of $t$.

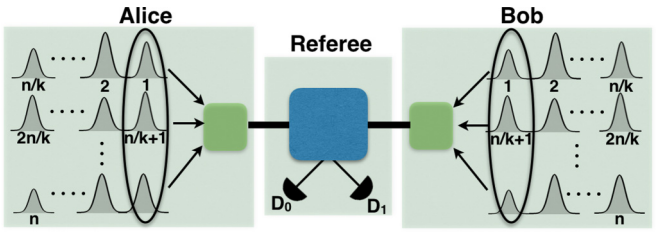


FIG. 2. Our multiple-channel protocol. Alice and Bob create coherent state fingerprints for each of the $k$ substrings and at each time unit they send the corresponding set of $k$ pulses to the referee. The referee interferes all $k$ pairs of pulses through a BS and processes the resulting information (at most 1 bit/photon per time unit). The protocol proceeds similarly for all $n/k$ time units and the referee estimates the ED.

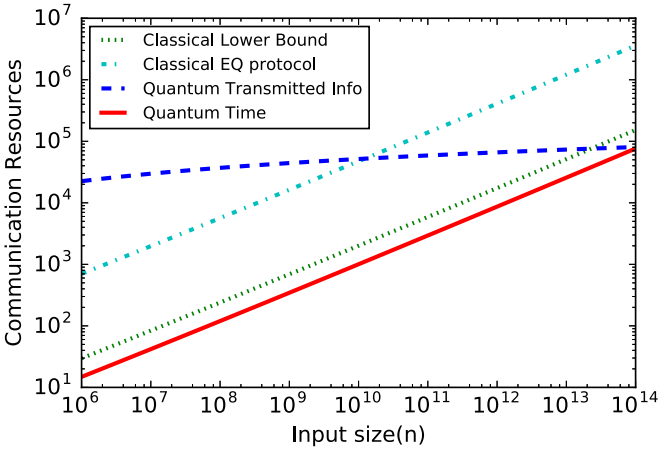


FIG. 3. Log-log plot for communication resources vs input size ($n$) for solving ED within $\epsilon = 0.1$ with error $\delta \leqslant 10^{-6}$. We compare the classical lower bound, the best known classical protocol for equality and our multiple-channel coherent state ED protocol. The size of the error correcting code used for comparing classical ED with classical EQ is $m = 3.6n$. We set the number of channels $k = O(\sqrt{n})$ and use the optimized value $\mu = 1969.45$ (see Appendix C).

In summary, our model is the following: Alice and Bob can use $k$ channels, where in every communication time unit, they can send in expectation at most one bit or one photon in total over all $k$ channels. Or in other words, the number of available channels exceeds the number of bits or photons that can be processed by the referee by a factor of $k$.

In this model, in the classical case, the use of multiple channels reduces the communication by at most a $\log k$ factor, since we can simulate any multiple channel protocol with a single channel one with a $\log k$ overhead: For every bit sent through one of the $k$ channels, we send the same bit and the index of the channel in $\log k$ bits through the single channel.

In the quantum case, we can take better advantage of the multiple channels and have a protocol with communication time of order $n/k$, while the transmitted information remains of order $\log n$. The reason is that most of the pulses sent are empty of photons and hence we can use the multiple channels to send in parallel many pulses, without sending more than one photon in expectation per time unit. More precisely, Alice and Bob divide their $n$ bit input into $k$ substrings, each of length $n/k$. They create coherent state fingerprints for each of the $k$ substrings and at each time unit they send $k$ pulses through the channels, one from each of the $k$ fingerprints. The referee interferes the corresponding pulses as in the initial protocol, either by using $k$ sets of BS and detectors, or by time ordering the pulses and using a single set of BS and detectors. The communication time is now reduced by a factor of $k$. By choosing $k$ to be $\omega(\sqrt{n})$, we can make both resources of the quantum protocol asymptotically smaller than the best classical protocol. The expected number of photons in each time unit is $\mu k/n$, which for large enough $n$ and since $k$ is asymptotically smaller than $n$ can be made $< 1$, hence satisfying the no more than one photon per time unit constraint.

As we mentioned above, if the referee can treat more than one bit or photon per time unit, say $t$, then by assuming we have $k \times t$ available channels, both the classical and quantum protocols save an additional factor $t$ and so the ratio of the classical and quantum communication time remains unchanged. Figure 2 illustrates our model for $t = 1$. In Appendix A we describe a possible realization of our protocol using multiplexing techniques.

The last two rows of Table I compare the classical and quantum multiple-channel protocols for ED. This is the first example of a model and a task, for which an in principle realistic quantum protocol is asymptotically more efficient both in the communication time and the transmitted information than any classical protocol.

## VI. PERFORMANCE ANALYSIS

We consider some standard experimental imperfections in the setting of Fig. 2, including the BS interferometer visibility $\nu$ and the detector dark count probability $p_d$. Errors may be due to the use of multiple channels as well but as these depend on the specific implementation we do not consider them here. We assume that all of referee's detectors have the same parameters, e.g., dark count probabilities.

With the above imperfections we recalculate the probability of a click in $D_0$ and $D_1$ and show in Appendix B that

$$\tilde{E} = 2 - \frac{1}{\mu(2\nu - 1)} \mathbb{E}\left[\sum_{j=1}^{n} \left(Z_j^0 - Z_j^1\right)\right]. \tag{4}$$

The main error comes from estimating the expectation of the detectors clicking. This error is bounded using the Chernoff-Hoeffding bounds and optimizing the experimental parameters. Another source of error comes from estimating the experimental parameters, i.e., the BS visibility $\nu$ and $\mu$. Note that the expectation of the difference in counts of the two detectors does not depend on the dark counts, provided it is the same for 0 and 1.

In Fig. 3 we show the transmitted information and communication time as a function of $n$ for the multiple-channel protocol with $k = O(\sqrt{n})$, and compare its performance with the classical lower bound and with the best-known classical protocol for equality. The analytical expressions for all protocols are provided in Appendix C. We see first that if we are only interested in the communication time, which is often the case, then our protocol outperforms the classical

limit even for small $n$ and by consequence for small number of channels which can be feasible in practice. Moreover, for large $n$, our protocol outperforms the classical limit for both resources. For current parameters, the number of channels needed is about $10^5$, which may not be realistic. By improving the experimental parameters one may decrease this number.

## VII. DISCUSSION

A noteworthy feature of the Euclidean distance protocol studied in our work is that Alice and Bob do not need a memory to store their inputs and they do not perform global operations on them. In other words, this protocol works also in the *streaming* scenario, where Alice and Bob receive their inputs one bit at a time [28]. We note that this is not the case either for the equality protocol, where an error correcting code needs to be applied to the entire input string, or for the qubit protocol where the fingerprint is encoded in a superposition of $\log n$ qubits. It will be interesting to further explore this scenario for efficient quantum communications. More generally, expanding the family of distributed tasks in the coherent state communication model studied in this work is important for demonstrating in practice quantum superiority in a network setting.

## ACKNOWLEDGMENTS

We thank Juan-Miguel Arrazola for useful discussions. We acknowledge financial support from the European Research Council project QCC, the Agence Nationale de la Recherche France projects QRYPTOS and COMB, the Ile-de-France Region project QUIN, and the France-USA Partner University Fund project CRYSP.

## APPENDIX A: DESCRIPTION OF OFDM APPROACH FOR MULTIPLE-CHANNEL COHERENT STATE ED

One way to implement the multiple-channel protocol proposed in our work is by using $k$ physical channels. This would be the case for backbone communication networks, where nodes are connected via a large number of channels. Another way could be to employ all-optical orthogonal frequency division multiplexing (OFDM), an advanced classical multiplexing technique that has recently been adapted for performing high-rate quantum key distribution [29]. In this case, Alice and Bob create coherent state fingerprints for each of their $k$ substrings in orthogonal frequency subcarrier modes that are generated using frequency offset locked laser diodes (alternatively, a pulsed laser source such as a mode-locked laser can be used for this purpose, as shown in Ref. [29]). At every time unit, the corresponding pulse from each of the $k$ fingerprints is multiplexed in a $k \times 1$ OFDM encoder and the output signal is sent to the referee.

The frequency separation between any two adjacent subcarriers in the OFDM scheme is $\omega_{j+1} - \omega_j = \Delta f$ and the encoded OFDM signal has a pulse width of $T = 1/\Delta f$. At the first time unit $t_1$, Alice's coherent pulses for each of the $k$ substrings are $\{|x_1\alpha\rangle_1, |x_{\frac{n}{k}+1}\alpha\rangle_1 \ldots, |x_{\frac{(k-1)n}{k}+1}\alpha\rangle_1\}$. They get encoded in the OFDM signal $\hat{E}_1(t)$ (subscript denotes time):

$$\hat{E}_1(t) = \sum_{j=1}^{k} e^{-\frac{|x_{\frac{(j-1)n}{k}+1}|^2\mu}{2}} e^{x_{\frac{(j-1)n}{k}+1}\alpha\hat{a}_j^\dagger} e^{i\omega_j t} = \sum_{j=1}^{k} \hat{A}_1^j e^{i\omega_j t}$$

for $0 < t < \mathrm{T}$, with the $j$th subcarrier frequency given by $\omega_j = \omega_0 + 2\pi j\Delta f$. Bob employs the same multiplexing technique to prepare his OFDM signal.

Once the OFDM signals from Alice and Bob for the time step $t_1$ reaches the referee, he decodes them via an optical discrete Fourier transform (ODFT) on the input signal $\hat{E}_1(t)$. The output circuit to decode the $q$th subcarrier signal (for $q = 1, \ldots, k$) at time step $t_1$ is

$$\hat{D}_1^q(t) = \frac{1}{k}\sum_{j=1}^{k} \hat{E}_1(t-(j-1)T_c)e^{i2\pi(j-1)(q-1)/k} = \hat{A}_1^q e^{i\omega_q t},$$

where $T_c = T/k$ and we used the orthogonality condition $\Delta f = 1/T$. A typical duration for the OFDM signal is $T = 100$ ps [29].

The advantage of this technique is that because of the orthogonality of the employed subcarriers, these do not interfere with each other despite overlapping sidebands between adjacent carriers, leading to an efficient demultiplexing; however, the number of supported subcarriers in practice currently remains quite low.

## APPENDIX B: ANALYSIS OF EUCLIDEAN DISTANCE PROTOCOL WITH IMPERFECTIONS

In the presence of the limited beam splitter visibility $\nu$, the output states for the referee in the detectors $D_0$ and $D_1$ at time step $j$ is given by

$$\left|\frac{\alpha}{\sqrt{2}}[\sqrt{\nu}(x_j + y_j) + \sqrt{1-\nu}(x_j - y_j)]\right\rangle_{j,D_0} \otimes \left|\frac{\alpha}{\sqrt{2}}[\sqrt{\nu}(x_j - y_j) + \sqrt{1-\nu}(x_j + y_j)]\right\rangle_{j,D_1} \quad \text{(B1)}$$

Let $Z_j^0$ and $Z_j^1$ be the binary random variables with value 1 with probability $\Pr[\text{click in } j, D_0]$ and $\Pr[\text{click in } j, D_1]$ respectively, for the $j$th time unit and value 0 otherwise. These probabilities are

$$\Pr[\text{click in } j, D_0] = p_j^{D_0} \approx \frac{\mu}{2}\left[\nu(x_j + y_j)^2 + (1-\nu)(x_j - y_j)^2 + 2\sqrt{\nu(1-\nu)}\left(x_j^2 - y_j^2\right)\right] + p_d, \quad \text{(B2)}$$

$$\Pr[\text{click in } j, D_1] = p_j^{D_1} \approx \frac{\mu}{2}\left[\nu(x_j - y_j)^2 + (1-\nu)(x_j + y_j)^2 + 2\sqrt{\nu(1-\nu)}\left(x_j^2 - y_j^2\right)\right] + p_d, \quad \text{(B3)}$$

where $p_d$ is the dark count probability for both detectors. We define $\Delta Z_j = Z_j^0 - Z_j^1$ as the difference in the clicks from $D_0$ and $D_1$ at $j$th step. Note that $\Delta Z_j$ is a random variable that takes values $\{-1, 0, 1\}$. We define $\Delta Z = \sum_j \Delta Z_j$ and we

have

$$\mathbb{E}[\Delta Z] = \frac{\mu}{2}(2\nu - 1)(||x + y||^2 - ||x - y||^2). \quad \text{(B4)}$$

Using Eq. (B4) and the fact that $||x + y||^2 - ||x - y||^2 = 4(1 - ||x - y||^2/2)$, we obtain the Euclidean distance between the data sets $x$ and $y$ as

$$\tilde{E} = ||x - y||^2 = 2 - \frac{1}{\mu(2\nu - 1)}\mathbb{E}[\Delta Z]. \quad \text{(B5)}$$

The error in the estimation of the Euclidean distance comes from two different sources: (i) the estimation of the mean value of $\Delta Z$ and the (ii) error in parameter estimation of $\mu$ and $\nu$, which in general depends on the experimental setup but can be considered very small.

Let us deal with the first source of error, i.e., the estimation of the mean value of $\Delta Z$. Basically, we can bound the probability that $\Delta Z$ is far from its mean value, by using the Chernoff-Hoeffding bounds to get a statement of the form: $\Pr(|\Delta Z - \mathbb{E}[\Delta Z]| \geqslant \epsilon\mathbb{E}[\Delta Z]) \leqslant \delta$. We prove this bound by the Theorem Theorem 1 given below.

*Theorem 1.* Let $X = \sum_j X_j$ be sum of $n$ random variables with each variable $X_j$ taking values $\{1, -1, 0\}$ with probability $\{p_j^{\{1\}}, p_j^{\{-1\}}, 1 - p_j^{\{1\}} - p_j^{\{-1\}}\}$ respectively. For any $a \in \mathbb{R}$,

$$\Pr(X \geqslant a) \leqslant \left(\frac{2P^{\{1\}}}{a + \sqrt{a^2 + 4P^{\{1\}}P^{\{-1\}}}}\right)^a \times e^{-(P^{\{1\}} + P^{\{-1\}} - \sqrt{a^2 + 4P^{\{1\}}P^{\{-1\}}})}, \quad \text{(B6)}$$

where $P^{\{1\}} = \sum_j p_j^{\{1\}}$ and $P^{\{-1\}} = \sum_j p_j^{\{-1\}}$.

*Proof.* Markov's inequality on $X = \sum_j X_j$ gives us for any $s > 0$,

$$\Pr(X \geqslant a) = \Pr(e^{sX} \geqslant e^{sa}) \leqslant \frac{\mathbb{E}(e^{sX})}{e^{sa}}. \quad \text{(B7)}$$

Note that $\mathbb{E}(e^{sX}) = \mathbb{E}(e^{s\sum_j X_j}) = \prod_j \mathbb{E}(e^{sX_j})$, where we used the property of independence of variables $X_j$. This allows us to prove the Chernoff bound by bounding each individual $\mathbb{E}(e^{sX_j})$.

*Lemma 1.* Let Y be a random variable that takes value 1 with probability $p_1$, $-1$ with probability $p_{-1}$, and 0 otherwise, then for all $s \in \mathbb{R}$,

$$\mathbb{E}(e^{sY}) \leqslant e^{p_1(e^s - 1) + p_{-1}(e^{-s} - 1)}.$$

*Proof.* We have

$$\begin{aligned}\mathbb{E}(e^{sY}) &= p_1 e^s + p_{-1}e^{-s} + [1 - (p_1 + p_{-1})]1 \\ &= 1 + [p_1(e^s - 1) + p_{-1}(e^{-s} - 1)] \\ &\leqslant e^{p_1(e^s - 1) + p_{-1}(e^{-s} - 1)}.\end{aligned}$$

Applying Lemma 1, we obtain

$$\begin{aligned}\mathbb{E}(e^{sX}) &\leqslant \prod_j e^{p_j^{\{1\}}(e^s - 1) + p_j^{\{-1\}}(e^{-s} - 1)} \\ &= e^{\sum_j p_j^{\{1\}}(e^s - 1) + p_j^{\{-1\}}(e^{-s} - 1)}. \quad \text{(B8)}\end{aligned}$$

We denote $P^{\{1\}} = \sum_j p_j^{\{1\}}$ and $P^{\{-1\}} = \sum_j p_j^{\{-1\}}$, and by applying the result of Eq. (B8) on Markov's inequality Eq. (B7) we obtain

$$Pr(X \geqslant a) \leqslant e^{P^{\{1\}}(e^s - 1) + P^{\{-1\}}(e^{-s} - 1) - sa}. \quad \text{(B9)}$$

It can be easily verified that $s' = \ln(\frac{a + \sqrt{a^2 + 4P^{\{1\}}P^{\{-1\}}}}{2P^{\{1\}}})$ minimizes the upper bound in Eq. (B9). This concludes our proof:

$$\Pr(X \geqslant a) \leqslant \left(\frac{2P^{\{1\}}}{a + \sqrt{a^2 + 4P^{\{1\}}P^{\{-1\}}}}\right)^a \times e^{-(P^{\{1\}} + P^{\{-1\}} - \sqrt{a^2 + 4P^{\{1\}}P^{\{-1\}}})}. \quad \text{(B10)}$$

To use the above theorem, let us define $\Lambda = P^{\{1\}} - P^{\{-1\}}$ and $a = \Lambda(1 + \epsilon)$, with $0 < \epsilon < 1$. Using the inequality $\sqrt{(\Lambda(1 + \epsilon))^2 + 4P^{\{1\}}P^{\{-1\}}} \leqslant (P^{\{1\}} + P^{\{-1\}}) + \Lambda\epsilon$ we can relax Eq. (B10) further to

$$\Pr[X \geqslant \Lambda(1 + \epsilon)] \leqslant \left(\frac{e^\epsilon}{\left(1 + \frac{\Lambda}{P^{\{1\}}}\epsilon\right)^{1+\epsilon}}\right)^\Lambda. \quad \text{(B11)}$$

Last, using the following inequalities, which are derived from Eqs. (B2), (B3), and (B4), $\mathbb{E}[\Delta Z] \leqslant 2\mu(2\nu - 1)$ and $\sum_j p_j^{D_0} \leqslant 2\mu\nu + (n - 2\mu)p_d$ we have

$$\begin{aligned}&\Pr(|\Delta Z - \mathbb{E}[\Delta Z]| \geqslant \epsilon\mathbb{E}[\Delta Z]) \\ &\leqslant 2\left[\frac{e^\epsilon}{\left(1 + \frac{2\nu - 1}{\nu + \frac{n - 2\mu}{2\mu}p_d}\epsilon\right)^{1+\epsilon}}\right]^{2\mu(2\nu - 1)} = \delta. \quad \text{(B12)}\end{aligned}$$

Equation (B12) highlights the contributions of $n, \mu, \nu$, and $p_d$ in the estimation of Euclidean distance, which can be estimated within any $0 < \delta < 1$ by controlling the mean photon number $\mu$, since $\nu, p_d$ are fixed by the experimental setup used.

## APPENDIX C: RESOURCE PERFORMANCE ANALYSIS

The plot in Fig. 3 of the main text compares the transmitted information ($I$) and communication time ($T$) vs the data input size $n$ for the classical lower bound, the best classical protocol for equality, and the quantum multiple-channel coherent state protocol. We fix that the protocol estimates ED within a constant factor $\epsilon = 0.1$ with error probability $\delta \leqslant 10^{-6}$.

*Classical lower bound.* We use the lower bounds for the equality problem [22,23,25] to provide a lower bound for ED.

Suppose we have a classical ED protocol for input size $n$ that approximates the distance within a fixed $\epsilon$ with probability at least $1 - \delta$. To construct a protocol for equality, we choose the error-correcting code (ECC) that amplifies the $n$-bit inputs $x$ and $y$ to $m$-bit codewords $E(x)$ and $E(y)$ respectively, with the minimum distance across being $d > 2\epsilon$. Then, we use the ED protocol on the codewords $E(x)$ and $E(y)$, and have

$$\tilde{E}\begin{cases}\leqslant \epsilon & \text{if } x = y \\ \geqslant d - \epsilon > \epsilon & \text{if } x \neq y\end{cases}. \quad \text{(C1)}$$

This guarantees solving the equality on $x$ and $y$ with probability $\geqslant 1 - \delta$.

This reduction implies that we can get a lower bound for ED for input size $n$ by a lower bound for equality with input size $rn$, where $r(< 1)$ is the rate of the ECC. For

$\epsilon = 0.1$ (i.e., $d > 0.2$), we can get $r = 0.28$ [30]. The time and transmitted information resource lower bound would then be $T_{cl} = I_{cl} = [(1 - 2\sqrt{\delta})\sqrt{\frac{rn}{2\log 2}} - 1]/\log k$ [21,23], which is shown in Fig. 3.

*Classical protocol for equality*. We also plot the best classical protocol that solves equality, which uses $2\sqrt{n} + 1$ bits and succeeds with probability $1 - \delta' = 3/4$. To get the desired $\delta \leqslant 10^{-6}$, the protocol is repeated 10 times. Thus to solve ED, the resources are $T_{cp} = I_{cp} = (20\sqrt{r \cdot n} + 10)/\log k$ [23].

*Quantum multiple-channel coherent state protocol*. For the quantum multiple-channel protocol that computes ED within $\epsilon = 0.1$ with error probability $\delta \leqslant 10^{-6}$, we optimize $\mu$ to be $\mu = 1969.45$ and set the other parameters to be $\nu = 0.99 \pm 0.005$ and $p_d = (1.0 \pm 0.1) \times 10^{-8}$ [20].

We use the results of Ref. [19] to precisely upper bound the transmitted information of the protocol (asymptotically it behaves as $I_{qp} = O(\mu \log n)$) and we also know that the communication time is $T_{qp} = n/k$, where $k$ is the number of channels used in the protocol. The communication time can be made less that the classical lower bound as long as $k > k_{crit}$, where $k_{crit}$ makes the quantum protocol time equal to the time in the classical lower bound and is of the order of $O(\sqrt{n})$.


[1] B. Hensen, H. Bernien, A. Dréau, A. Reiserer, N. Kalb, M. Blok, J. Ruitenberg, R. Vermeulen, R. Schouten, C. Abellán *et al.*, Nature (London) **526**, 682 (2015).

[2] D. N. Matsukevich, P. Maunz, D. L. Moehring, S. Olmschenk, and C. Monroe, Phys. Rev. Lett. **100**, 150404 (2008).

[3] M. Ansmann, H. Wang, R. C. Bialczak, M. Hofheinz, E. Lucero, M. Neeley, A. O'Connell, D. Sank, M. Weides, J. Wenner *et al.*, Nature (London) **461**, 504 (2009).

[4] A. Pappa, N. Kumar, T. Lawson, M. Santha, S. Zhang, E. Diamanti, and I. Kerenidis, Phys. Rev. Lett. **114**, 020401 (2015).

[5] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev, Rev. Mod. Phys. **81**, 1301 (2009).

[6] R. J. Donaldson, R. J. Collins, K. Kleczkowska, R. Amiri, P. Wallden, V. Dunjko, J. Jeffers, E. Andersson, and G. S. Buller, Phys. Rev. A **93**, 012329 (2016).

[7] A. Pappa, P. Jouguet, T. Lawson, A. Chailloux, M. Legré, P. Trinkler, I. Kerenidis, and E. Diamanti, Nat. Commun. **5**, 3717 (2014).

[8] M. Tillmann, B. Dakić, R. Heilmann, S. Nolte, A. Szameit, and P. Walther, Nat. Photon. **7**, 540 (2013).

[9] A. Crespi, R. Osellame, R. Ramponi, D. J. Brod, E. F. Galvão, N. Spagnolo, C. Vitelli, E. Maiorino, P. Mataloni, and F. Sciarrino, Nat. Photon. **7**, 545 (2013).

[10] N. Spagnolo, C. Vitelli, M. Bentivegna, D. J. Brod, A. Crespi, F. Flamini, S. Giacomini, G. Milani, R. Ramponi, P. Mataloni *et al.*, Nat. Photon. **8**, 615 (2014).

[11] H. Buhrman, R. Cleve, J. Watrous, and R. de Wolf, Phys. Rev. Lett. **87**, 167902 (2001).

[12] H. Buhrman, R. Cleve, and A. Wigderson, in *Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing* (ACM Press, New York, NY, 1998), pp. 63–68.

[13] R. Raz, in *Proceedings of the Thirty-First Annual ACM Symposium on Theory of Computing* (ACM Press, New York, NY, 1999), pp. 358–367.

[14] Z. Bar-Yossef, T. S. Jayram, and I. Kerenidis, in *Proceedings of the Thirty-Sixth Annual ACM Symposium on Theory of Computing* (ACM Press, New York, NY, 2004), pp. 128–137.

[15] D. Gavinsky, J. Kempe, I. Kerenidis, R. Raz, and R. De Wolf, in *Proceedings of the Thirty-Ninth Annual ACM Symposium on Theory of Computing* (ACM Press, New York, NY, 2007), pp. 516–525.

[16] D. Gavinsky, in *Proceedings of the Forty-Eighth Annual ACM Symposium on Theory of Computing* (ACM Press, New York, NY, 2016), pp. 877–884.

[17] O. Regev and B. Klartag, in *Proceedings of the Forty-Third Annual ACM Symposium on Theory of Computing* (ACM Press, New York, NY, 2011), pp. 31–40.

[18] J. M. Arrazola and N. Lütkenhaus, Phys. Rev. A **90**, 042335 (2014).

[19] J. M. Arrazola and N. Lütkenhaus, Phys. Rev. A **89**, 062305 (2014).

[20] F. Xu, J. M. Arrazola, K. Wei, W. Wang, P. Palacios-Avila, C. Feng, S. Sajeed, N. Lütkenhaus, and H.-K. Lo, Nat. Commun. **6**, 8735 (2015).

[21] J.-Y. Guan, F. Xu, H.-L. Yin, Y. Li, W.-J. Zhang, S.-J. Chen, X.-Y. Yang, L. Li, L.-X. You, T.-Y. Chen *et al.*, Phys. Rev. Lett. **116**, 240502 (2016).

[22] A. Ambainis, Algorithmica **16**, 298 (1996).

[23] L. Babai and P. G. Kimmel, in *Proceedings of the Twelfth Annual IEEE Conference on Communication Complexity* (IEEE, Ulm, Germany, 1997), pp. 239–246.

[24] I. Newman, Infor. Process. Lett. **39**, 67 (1991).

[25] I. Newman and M. Szegedy, in *Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing* (ACM Press, New York, NY, 1996), pp. 561–570.

[26] I. Kremer, N. Nisan, and D. Ron, Comput. Complex. **8**, 21 (1999).

[27] E. Upfal and M. Mitzenmacher, *Probability and Computing: Randomized Algorithms and Probabilistic Analysis* (Cambridge University Press, New York, NY, 2005).

[28] Y. M. Noga Alon and M. Szegedy, J. Comput. Syst. Sci. **58**(1), 137 (1999).

[29] S. Bahrani, M. Razavi, and J. A. Salehi, J. Lightwave Technol. **33**, 4687 (2015).

[30] W. W. Peterson and E. J. Weldon, *Error-Correcting Codes* (MIT Press, Cambridge, MA, 1972).