Quantum Science and Technology

PAPER

Experimental multi-state quantum discrimination through optical networks

Alessandro Laneve[1],* , Andrea Geraldi[1], Frenkli Hamiti[1], Paolo Mataloni[1] and Filippo Caruso[2],*

1 Dipartimento di Fisica, Sapienza Università di Roma, Piazzale Aldo Moro, 5, I-00185 Roma, Italy
2 Department of Physics and Astronomy & LENS, University of Florence, Via Carrara 1, I-50019 Sesto Fiorentino, Italy
* Authors to whom any correspondence should be addressed.

E-mail: alessandro.laneve@uniroma1.it and filippo.caruso@unifi.it

Abstract

Developing strategies to effectively discriminate between different quantum states is a fundamental issue in quantum information and communication. The actual realization of generally optimal protocols in this task is often limited by the need of supplemental resources and very complex receivers. We have experimentally implemented two discrimination schemes in a minimum-error scenario based on a receiver featured by a network structure and a dynamical processing of information. The first protocol implemented in our experiment, directly inspired to a recent theoretical proposal, achieves binary optimal discrimination, while the second one provides a novel approach to multi-state quantum discrimination, relying on the dynamical features of the network-like receiver. This strategy exploits the arrival time degree of freedom as an encoding variable, achieving optimal results, without the need for supplemental systems or devices. Our results further reveal the potential of dynamical approaches to quantum state discrimination tasks, providing a possible starting point for efficient alternatives to current experimental strategies.

1. Introduction

Quantum state discrimination (QSD) consists of the ability of distinguishing between the possible states of a quantum system. This task cannot be achieved in a perfect way in the case of non-orthogonal states. As a consequence, the quest for general strategies to achieve optimal QSD has become a main issue in Quantum Information topics, regarding its tight bond with the implementation of secure quantum communication protocols [1, 2], or its implications in Quantum Foundations [3, 4]. Different approaches can be adopted in QSD: to discriminate states with the minimum error probability [5], or in a non ambiguous way [6] but admitting unconclusive results. Other approaches are also possible [7–9].

The quest for effective protocols has led to many results concerning optimal theoretical bounds and optimal receiver models for binary discrimination in different contexts [5, 10–12]. On the other hand, regarding the general case of multiple states and dimension $D > 2$, definitive results have not been achieved yet; while the optimal one-shot bound has been set [11], many attempts have been done to develop new and more effective strategies using adaptive protocols [13–16] or auxiliary systems [17]. Other proposals involve the exploitation of an auxiliary system too [18], in order to increase the dimension of the system to be discriminated, or consist of encoding the states in a complex modal structure [19]. Since adaptive strategies have been proved to be effective, although quite expensive in terms of resources, recent theoretical efforts have been developed to apply neural networks models [20, 21] and machine learning (ML) protocols [22] to QSD problems. In addition to the attempts mentioned above, a dynamical approach based on information processing via quantum walks (QWs) has been fancied in the last years [23, 24]. The network depicted in [21], relying on a generalization of QWs, quantum stochastic walks (QSWs) [25], frames a very intuitive model of information processing as well as a wide applicability.
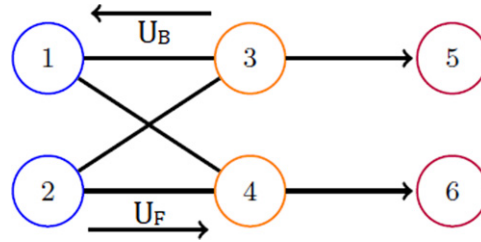
Figure 1. $2r-2r-2$ model of the quantum network. The system goes from a superposition state of input layer nodes $\{1, 2\}$ to one of $\{3, 4\}$, evolving according to the unitary evolution $\hat{U}_F$. Then, the population of the system which has not been trapped in the sinks evolves through $\hat{U}_B$ to a new general state of the input layer nodes, starting a new evolution loop.

In the present work, we implement the first experimental QSD protocol through an optical network, based on the one proposed in [21], but featuring important extensions. We demonstrate the effectiveness of a receiver relying on a QSW-like evolution and through the experimental analysis of the information dynamics in a time-binned extraction protocol, we are able to realize an optimal QSD protocol for a four-state alphabet in $D = 2$, exploiting a novel theoretical and experimental approach. Thanks to the dynamical features of the network, we are able to encode the quantum information about the four states in the classical observable of detection time, mapping the four different states in four different time-wise classical probability distributions, thus enhancing their distinguishability.

The discrimination method proposed here may prove useful in different QSD frameworks, from unambiguous state discrimination to minimum error discrimination (which is the one adopted in this work), and possibly in hybrid protocols, providing an essentially new approach to the problem, featured by an evident spare in terms of resources.
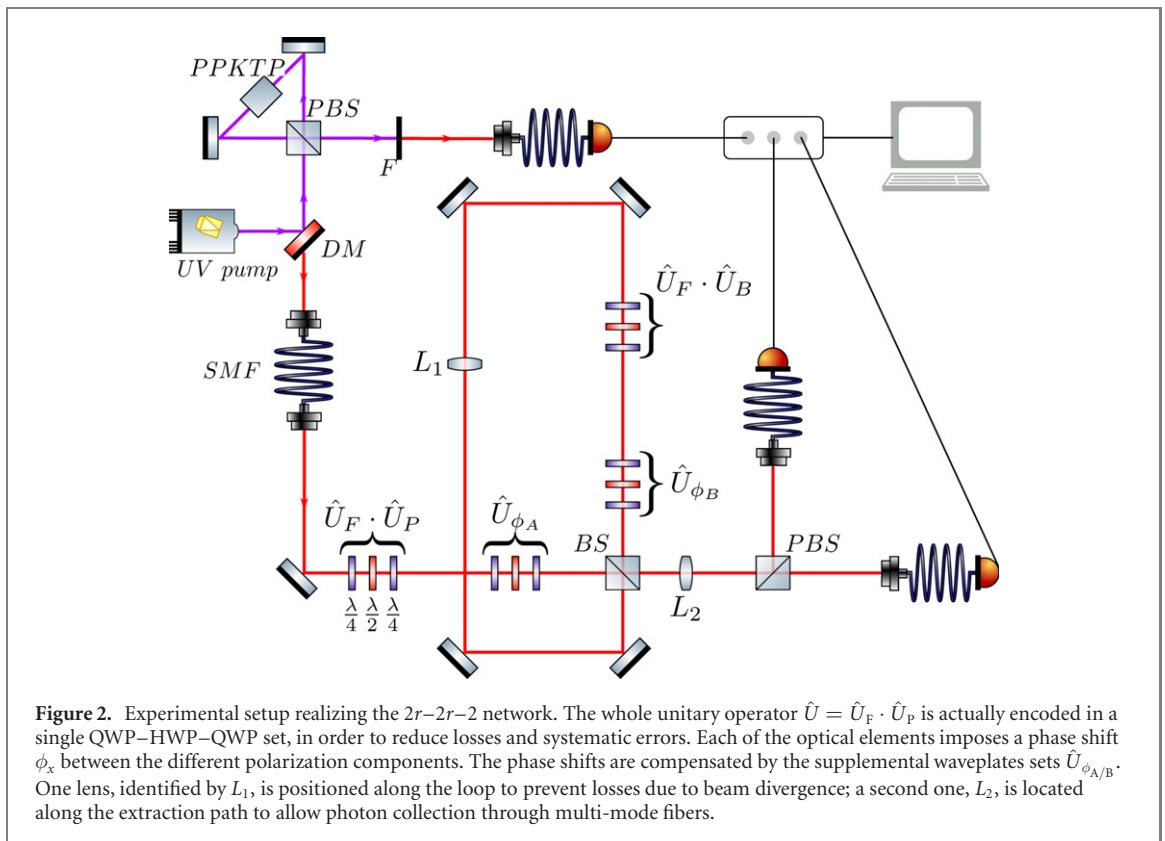
2. Theoretical and experimental framework

The first step of our work consists of experimentally reproducing a network featured by a $2r-2r-2$ topology (following to the nomenclature adopted in [21]) and depicted in figure 1: two input nodes, namely $\{1, 2\}$, are linked in an undirected way with two nodes in the intermediate layer, namely $\{3, 4\}$, which are in turn linked in a directed way with two output nodes, namely $\{5, 6\}$, acting as sink nodes.

At variance with the theoretical scheme of [21], we propose a model featuring a discrete time evolution and preventing the permanence of the system in the same state after an evolution step. In conclusion, the state of the network switches from a superposition state of the input nodes to a superposition of the intermediate ones, but for the loss of total probability to the sinks, which eventually brings the network dynamics to vanish. This model shares many interesting features with [21]: in fact, the dynamical achievement of binary discrimination at the Helstrom bound level remains possible, as well as other implications related to the exploitation of the time degree of freedom.

The setup works considering the two input nodes $1, 2$ as the two basis states of the polarization of a photon, namely $1 \equiv |H\rangle$ and $2 \equiv |V\rangle$, while the photon itself corresponds to the walker of this quantum walk-like propagation model. The initial state of the system $|\psi\rangle$ is encoded in a superposition state of photon polarization.

In the experiment, a high brilliance SPDC source S is exploited to generate pairs of photons in a polarization state $|H\rangle_s \otimes |H\rangle_a \equiv |H, H\rangle$, where the subscript s indicates the *system* photon, which will evolve through the network, while the subscript a indicates the *ancilla* photon, which does not undergo the network evolution and is exploited as an external trigger for detection. Details about photon source as well as experimental setup shown in figure 2 are reported in appendix A.1.

Before starting the evolution in the network, the initial state $|\psi\rangle$ is prepared by a unitary preparation stage ($\hat{U}_P$). In our setup, unitary operations are applied through different waveplates sets composed by a sequence of a quarter waveplate (QWP), a half waveplate (HWP) and a further QWP. This sequence allows to apply any kind of unitary operator in the polarization degree of freedom, allowing to transform the state $|H\rangle_s$ into any desired pure $|\psi\rangle$ state. After the state preparation, the system can be considered to be in the input layer of the network. The photon is then subjected to the first evolution stage, applying the operator $\hat{U}_F$ for the evolution from nodes $\{1, 2\}$ to $\{3, 4\}$. The nodes $\{3, 4\}$, also referred to as *sinker nodes*, are identified with the basis states of the polarization, too. The time degree of freedom allows to discriminate between states in the nodes $\{1, 2\}$ and the ones in $\{3, 4\}$. A beam splitter (BS) is then placed along the path: its aim is to redirect the photon to the sink nodes $\{5, 6\}$ with a given probability p_s or to send it into the

Figure 2. Experimental setup realizing the $2r$–$2r$–2 network. The whole unitary operator $\hat{U} = \hat{U}_F \cdot \hat{U}_P$ is actually encoded in a single QWP–HWP–QWP set, in order to reduce losses and systematic errors. Each of the optical elements imposes a phase shift ϕ_x between the different polarization components. The phase shifts are compensated by the supplemental waveplates sets $\hat{U}_{\phi_{A/B}}$. One lens, identified by L_1, is positioned along the loop to prevent losses due to beam divergence; a second one, L_2, is located along the extraction path to allow photon collection through multi-mode fibers.
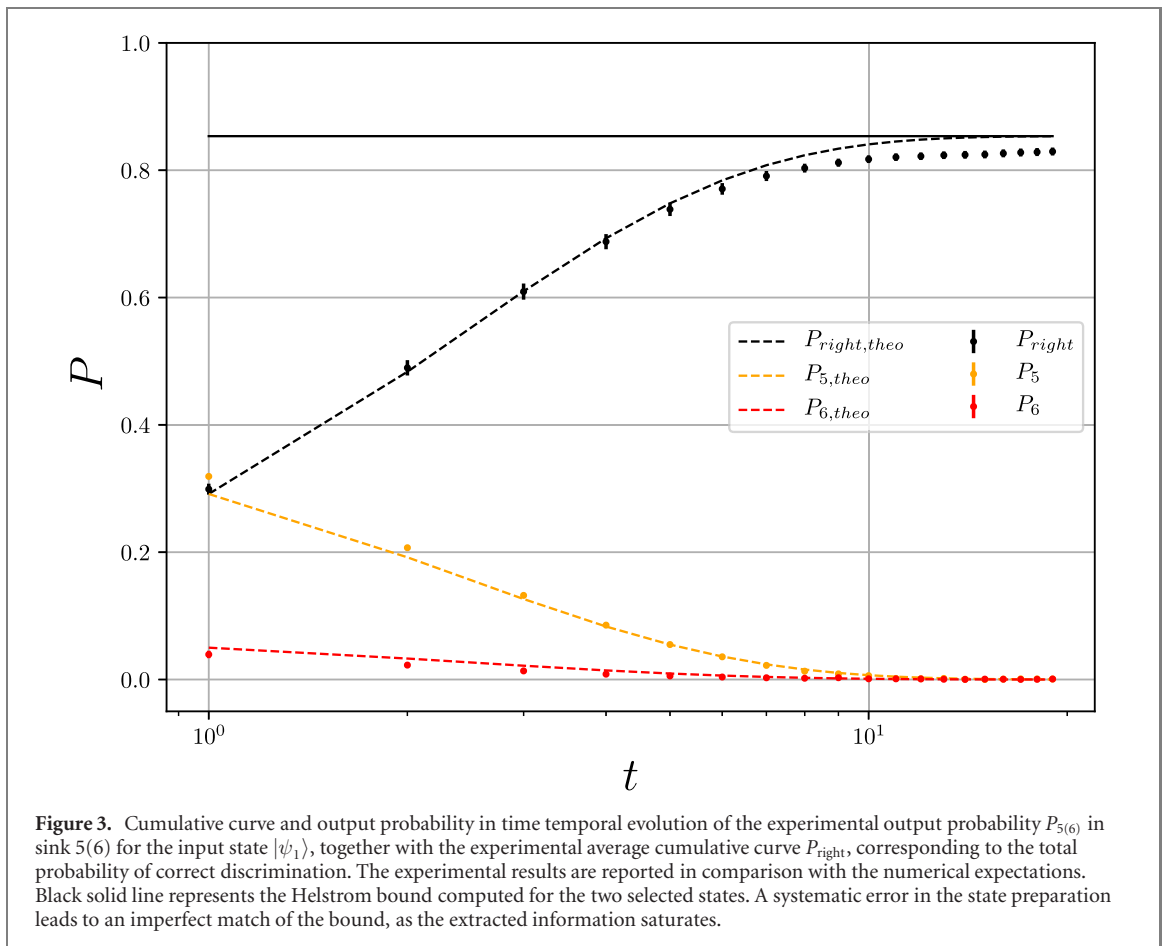
network with probability $1 - p_s$, in order to continue the evolution. In the first case, when the photon travels towards the sinks, it impinges on a polarizing beam splitter which separates the $|H\rangle$ and $|V\rangle$ components of the photon state, representing the sink nodes $\{5, 6\}$. The population of both sinks is then measured by single photon detectors. The detection of the *s* photon is performed in coincidence with the detection of the *a* photon, thanks to a preliminary synchronization process. In the second case, i.e. if the photon is redirected into the network, it passes through another set of QWP–HWP–QWP, which are set to apply the operator $\hat{U}_F \cdot \hat{U}_B$, that is the product of operators $\hat{U}_B$, the operator describing the propagation of the system from nodes $\{3, 4\}$ to $\{1, 2\}$, and $\hat{U}_F$. In this way the system evolves from a state of nodes $\{3, 4\}$ to $\{1, 2\}$ and then again to nodes $\{3, 4\}$.

After that, the walker impinges again on the BS, beginning another loop evolution or being detected into the sinks. Thus, the setup described above provides an experimental realization of a network featured by a $2r$–$2r$–2 topology. This experimental scheme was at first exploited to demonstrate the protocol proposed in [21].

## 3. Experimental realization of binary quantum state discrimination

Through the experimental apparatus described in the main text, it is possible to implement a discrimination protocol quite similar to the one proposed in [21], which we exploited as a test bed for the experimental scheme. Two non-orthogonal states $\{|\psi_1\rangle, |\psi_2\rangle\}$ are encoded in two states of the input layer; in order to discriminate them, the unitary evolution of the network must be tailored in such a way that the detection of the system in a sink reveals the presence of a given input state rather than the other one, with the minimum probability of wrong guess. In the case of binary QSD, the one-shot strategy consisting of projecting the two states to the orthogonal basis states produces the optimal performance, achieving the Helstrom bound. The natural generalization of this method to a multi-step strategy is a protocol performing a Helstrom bound level discrimination at any step, namely an optimal information extraction each time the system occurs in the sinker nodes (the intermediate layer). The most simple realization of such a protocol consists of generating an 'optimal' output state at the first extraction step, through the tuning of operator $\hat{U}_F$, and then tailoring $\hat{U}_B$ to produce the same state at any extraction step. In conclusion, $\hat{U}_F$ must be the optimally discriminating projector and $\hat{U}_B$ must be such that the product $\hat{U}_F \cdot \hat{U}_B = \hat{I}$. Such an optimal network will be able to discriminate states up to the Helstrom bound, by the computation of the cumulative population of the sinks in time. The states experimentally exploited for the protocol are $|\psi_1\rangle = \cos(\frac{\pi}{8})|H\rangle + \sin(\frac{\pi}{8})|V\rangle$ and $|\psi_2\rangle = \cos(\frac{\pi}{8})|H\rangle - \sin(\frac{\pi}{8})|V\rangle$, setting a strong similarity

Figure 3. Cumulative curve and output probability in time temporal evolution of the experimental output probability $P_{5(6)}$ in sink 5(6) for the input state $|\psi_1\rangle$, together with the experimental average cumulative curve P_{right}, corresponding to the total probability of correct discrimination. The experimental results are reported in comparison with the numerical expectations. Black solid line represents the Helstrom bound computed for the two selected states. A systematic error in the state preparation leads to an imperfect match of the bound, as the extracted information saturates.

with the case studied in [21]. The optimal unitary matrix for the realization of a suitable dynamical discrimination protocol was analytically found by maximizing or minimizing the output probability of each sink with respect to the corresponding input state. The result of this optimization lead to the evolution matrix $U = U_{\text{F}} = U_{\text{B}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, which has some analogy to the one computed in [21], besides some peculiar property. In fact, exploiting this evolution matrix, the discrimination protocol consists of a repeated optimal single-shot discrimination protocol: the network periodically brings the system in the state of the sinker nodes allowing optimal discrimination. In this way, at each loop completion, the information on the state is optimally extracted, leading to the output probabilities in time P_5 and P_6 reported in figure 3, where the sink node 5 is associated to the detection of state $|\psi_1\rangle$, while sink 6 to the detection of $|\psi_2\rangle$. It is now interesting to understand the dynamical behaviour of the probability of correct guess (as well as the dynamics of the extracted information). To this aim, the cumulative probability of correct discrimination in time is computed, for both states: $C_1 = \sum_\tau^t p_5^{(1)}(\tau)$, $C_2 = \sum_\tau^t p_6^{(2)}(\tau)$, where $p_k^{(i)}(\sigma)$ is the population of sink k, with input state $|\psi_i\rangle$, measured after the σth loop. The curve resulting from the average between C_1 and C_2, normalized to the total cumulative probability, is exhibited in figure 3 as P_{right}, in comparison with the numerical expectations. The experimental analysis proves the effectiveness of this protocol to achieve Helstrom level binary QSD in a dynamical context. These positive results pave the way towards the exploitation of these dynamical features in more complex and powerful protocols, fully taking advantage of a time-binned extraction of information. The following section clarifies this idea and reports on the results of the application of the optical network receiver to a four-state quantum discrimination problem in $D = 2$, in a minimum error probability scenario.

4. Experimental multi-state discrimination via time-binning protocol

The discrimination of more than two non-orthogonal states in a two-dimensional space is, in general, a task for which an absolutely optimal strategy is unknown [2]. We have developed a protocol based on time-binning that can be exploited to discriminate different sets of four states without accessing auxiliary systems or resources.
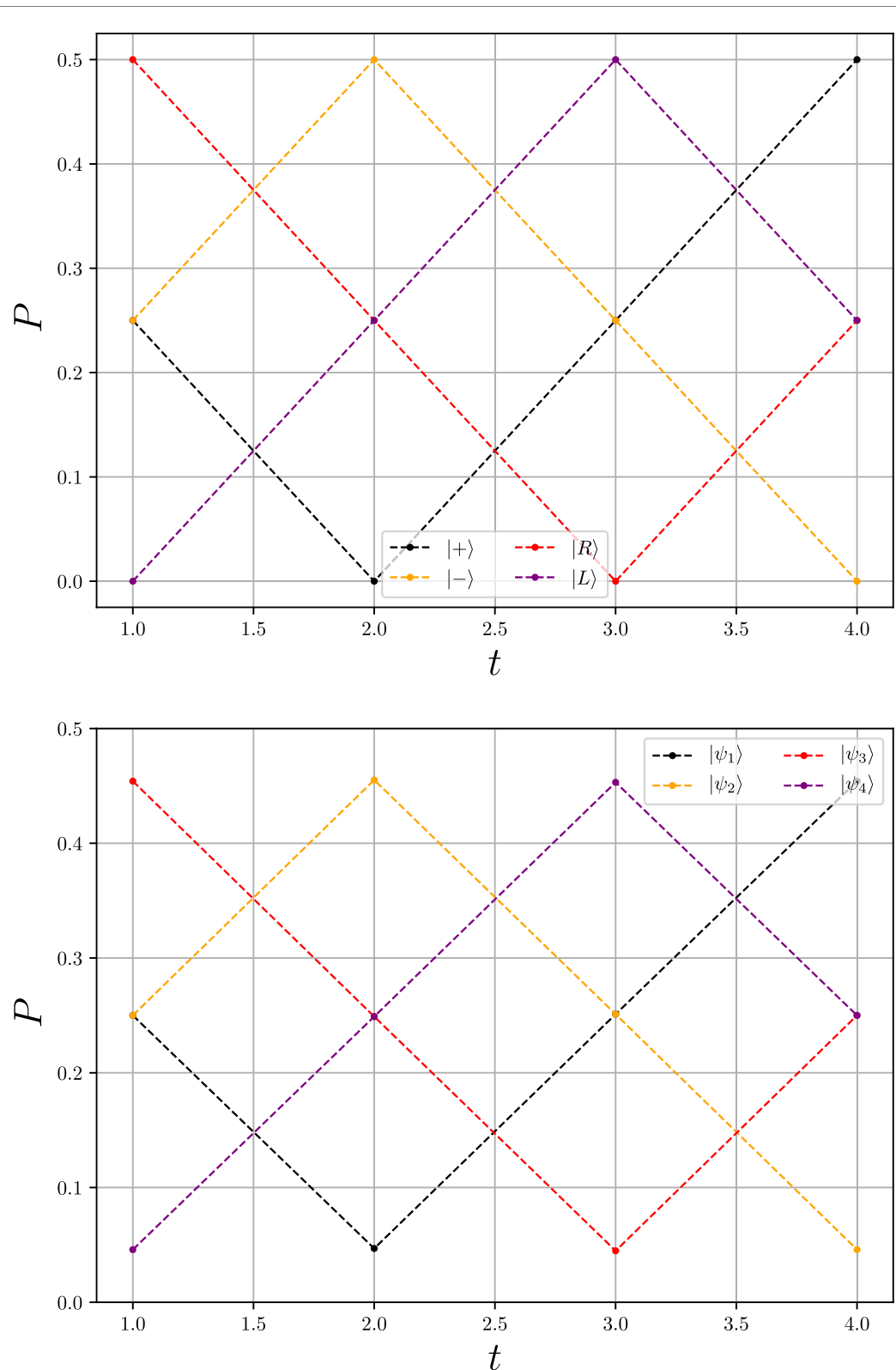
Figure 4. Output probability distributions for both states sets. Time-wise output probability of the first four extraction steps when the geometrically uniform states (top) and the Tetrad (bottom) states circulate in the network.

    The protocol consists of a redistribution of probability in time by means of the dynamical features of the network. In order to clarify the exploitation of the time degree of freedom as a supplemental resource for our QSD protocol, it is necessary to provide a formal discussion of our framework. The $2r$−$2r$−2 network can be unambiguously described in terms of two subspaces: the first one represents the two indirectly connected layers, the state of which can be written as a four-dimensional vector, where the first two dimensions represents the input layer and the second pair represents the intermediate layer, i.e. the sinker nodes. The state of this subspace evolves in time via unitary operations. The subspace of the sink nodes,

which we will refer to as the sink space, on the contrary, undergoes measurements and does not feature a unitary evolution in time. At first, we focus on the evolution of a photon in a state of the first subspace, which we indicate as the network space. Since in our model a superposition state between the input layer and the sinker nodes is not allowed, the only possible states have the form $\begin{pmatrix} \alpha \\ \beta \\ 0 \\ 0 \end{pmatrix}$ or $\begin{pmatrix} 0 \\ 0 \\ \gamma \\ \delta \end{pmatrix}$. Therefore, the one-step evolution of the system is described by a 4×4 unitary matrix, composed of two antidiagonal blocks $U = \begin{pmatrix} 0 & 0 & a & b \\ 0 & 0 & c & d \\ a' & b' & 0 & 0 \\ c' & d' & 0 & 0 \end{pmatrix}$. Indeed, also the permanence of the system in the same layer after an evolution step is not allowed by our model. The left-bottom block, which we refer to as U_F, represents the forward evolution of the system from input layer nodes $\{1, 2\}$ towards the sinker nodes $\{3, 4\}$, while the right-top block, which we refer to as U_B, represents the backward evolution of the system from sinker nodes $\{3, 4\}$ to the input layer nodes $\{1, 2\}$. In the actual implementation, both layers are encoded in the polarization degree of freedom, while they have a different temporal and spatial location. Therefore, we can practically describe the evolution as an alternate application of operators $\hat{U}_F$ and $\hat{U}_B$ to the same 2D vector. These two matrices can be identically or differently set, producing a symmetric or asymmetric network. In our framework, we are also able to discriminate the extraction time of the photon, i.e. the amount of times the photon has completed a loop evolution before being sent to the sinks. This supplemental degree of freedom, which is crucial in our protocol, can be formally addressed as one auxiliary system, featuring two subspaces: one for the photon travelling through the network, representing the evolution step of the system, and one for the sinks, representing the extraction time. In the following discussion, we indicate with the subscript n or s the corresponding subspace of the considered vector, both for the nodes and the time degree of freedom. The unitary evolution operators $\hat{U}_F$ and $\hat{U}_B$, which only act on the network subspace, will now represent the joint operator $\hat{U}_{F/B} \otimes \hat{I}_s$, being $\hat{I}_s$ the identity over the sink subspace. Moreover, we will define as $\hat{I}_N$ the identity over the joint space of network and sinks, while the identity over the time auxiliary system will be I_t. In conclusion, the evolution of the system in terms of the state in the network subspace n and in the sink subspace s, besides the time degree of freedom, can be segmented in three different phases, given the initial state $|\psi_0\rangle_{n,s} = (\alpha_0|1\rangle_n + \beta_0|2\rangle_n)|t_0\rangle_n$. In the first one, the initial state evolves from input to intermediate layer through $\hat{U}_F$:

$$|\psi_0^{(1)}\rangle_{n,s} = (\hat{U}_F \otimes \hat{I}_t)|\psi_0\rangle_{n,s} = (\gamma_0|3\rangle_n + \delta_0|4\rangle_n)|t_0\rangle_n. \tag{1}$$

The second step consists of the extraction to the sink nodes, which corresponds to the generation of a superposition state of the system being in the network or in the sinks at a certain step time, through a projector from the network to the sinks:

$$|\psi_0^{(2)}\rangle_{n,s} = \Big[\sqrt{T} \left(\hat{I}_N \otimes \hat{I}_t \right) +$$
$$\sqrt{1-T} \left(|5\rangle_s\langle 3|_n + |6\rangle_s\langle 4|_n \right) \otimes \left(\sum_{k=0}^{\infty} |t_k\rangle_s\langle t_k|_n \right) \Big] |\psi_0^{(1)}\rangle = \tag{2}$$
$$\sqrt{T} \left(\gamma_0|3\rangle_n + \delta_0|4\rangle_n \right) |t_0\rangle_n + \sqrt{1-T} \left(\gamma_0|5\rangle_s + \delta_0|6\rangle_s \right) |t_0\rangle_s,$$

where T is the probability of the system to stay in the network and $1 - T$ the probability of being extracted to the sinks. The last step brings again the network in a state of the input layer, while the sinks do not evolve at all:

$$|\psi_1\rangle_{n,s} = \left(\hat{U}_B \otimes \sum_{k=0}^{\infty} |t_{k+1}\rangle_n\langle t_k|_n \right) |\psi_0^{(2)}\rangle_{n,s} =$$
$$\sqrt{T} \left(\alpha_1|1\rangle_n + \beta_1|2\rangle_n \right) |t_1\rangle_n + \sqrt{1-T} \left(\gamma_0|5\rangle_s + \delta_0|6\rangle_s \right) |t_0\rangle_s, \tag{3}$$

where the 'time state' of the network has been updated, because a forward and backward evolution has been completed. The evolution continues as a repetition of these three steps, leading to a general state, after M completed loops:

$$|\psi_M\rangle = \sqrt{T}^M \left(\alpha_M|1\rangle_n + \beta_M|2\rangle_n\right)|t_M\rangle_n$$

$$+ \sum_{k=1}^{M} \sqrt{T}^{k-1}\sqrt{1-T}\left(\gamma_{k-1}|5\rangle_s + \delta_{k-1}|6\rangle_s\right)|t_{k-1}\rangle_s, \tag{4}$$

where $\alpha_k, \beta_k, \gamma_k, \delta_k$ with $k = 0, \ldots, M$ are the coefficients for each basis state of the network after a k steps evolution. The measurement we are able to perform, in the end, corresponds to a projection on both the sink state and the corresponding step time, which determines the probability of finding the system in a given sink σ at a certain time t_k:

$$P_\sigma(t_k) = |\langle\sigma|_s\langle t_k|_s|\phi\rangle_s|^2 \tag{5}$$

where $|\phi\rangle_s$ is a general joint state of the sinks and the associated extraction time. This formalism allows us to intuitively describe the capability of our setup to exploit time as a further degree of freedom for discrimination, while the actual realization is implemented through a set of subsequent post-selection procedures.

The exploitation of a globally asymmetrical network unitary, with $U_F \neq U_B$, allows for advanced applications, such as four state discrimination. In particular, this method is effective for a set of geometrically uniform states [2], since the network parameters are fixed in time. Indeed, the first observed application involves a set of geometrically uniform states, which we consider equally likely to be received:

$$|+\rangle = 1/\sqrt{2}\left(|H\rangle + |V\rangle\right)$$
$$|-\rangle = 1/\sqrt{2}\left(|H\rangle - |V\rangle\right)$$
$$|R\rangle = 1/\sqrt{2}\left(|H\rangle + \mathrm{i}|V\rangle\right) \tag{6}$$
$$|L\rangle = 1/\sqrt{2}\left(|H\rangle - \mathrm{i}|V\rangle\right).$$

The network unitary is tailored with the aim of producing the most different output probability distributions as the four different states occur at the input layer. That has been achieved selecting one of the two sinks and requesting that in each of the first four time bins a different input state had the relative maximum probability of being extracted. The evolution matrices produced following this method are

$U_F = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ and $U_B = \frac{1}{\sqrt{2}}\begin{pmatrix} \frac{1+\mathrm{i}}{2} & \frac{1+\mathrm{i}}{2} \\ \frac{1-\mathrm{i}}{2} & \frac{-1+\mathrm{i}}{2} \end{pmatrix}$. The product matrix, representing the loop evolution

of the system between two distinct extraction steps, results then $U_L = U_F \circ U_B = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & \mathrm{i} \\ \mathrm{i} & 1 \end{pmatrix}$. It is worth

noting that $U_L^4 = I$, from which the four-periodicity of the output probability distributions derives. Therefore, the first four time bins contain all of the information on the input state, while the experimental exploitation of the further ones is only useful for collecting a greater amount of meaningful signal. From a more analytical point of view, if we consider only one of the two sinks, e.g. the one corresponding to horizontal light polarization, the measurement at each time step t represents the application of the POVM $U_L^t \circ |H\rangle\langle H|$. Hence, our setup is capable of reproducing the families of POVM identified by the global unitary of the loops and depending on the step evolution parameter $\{U_L^t \circ |H\rangle\langle H| \| t \in \mathbb{N}\}$ and $\{U_L^t \circ |V\rangle\langle V| \| t \in \mathbb{N}\}$, although the actually distinct POVMs for each set are only four. Since the states are geometrically homogeneous, separated by a $\frac{\pi}{2}$ rotation around the same axis, this kind of procedure produces an optimal output. In fact, without changing the network parameters in time, it is possible to achieve a $P_{err} = 0.5$, which is the analytical bound for a set of four geometrically homogeneous states [11]. This is understandable by looking at figure 4 (top), where the numerical probability of detection in the first four time bins is displayed for the geometrically uniform set and the normalization is performed without considering the total signal decrease which the system experiences after every extraction step. Moreover, the probability distributions are normalized in such a way that, given a certain input state, the total probability equals to 1. In this way, the error probability of guess can be straightforwardly computed, given the input state.

The main feature of these distributions, each of which provides an unambiguous time-signature of a given input state, is the four-periodicity of the probability trajectory in time; it relies on the network
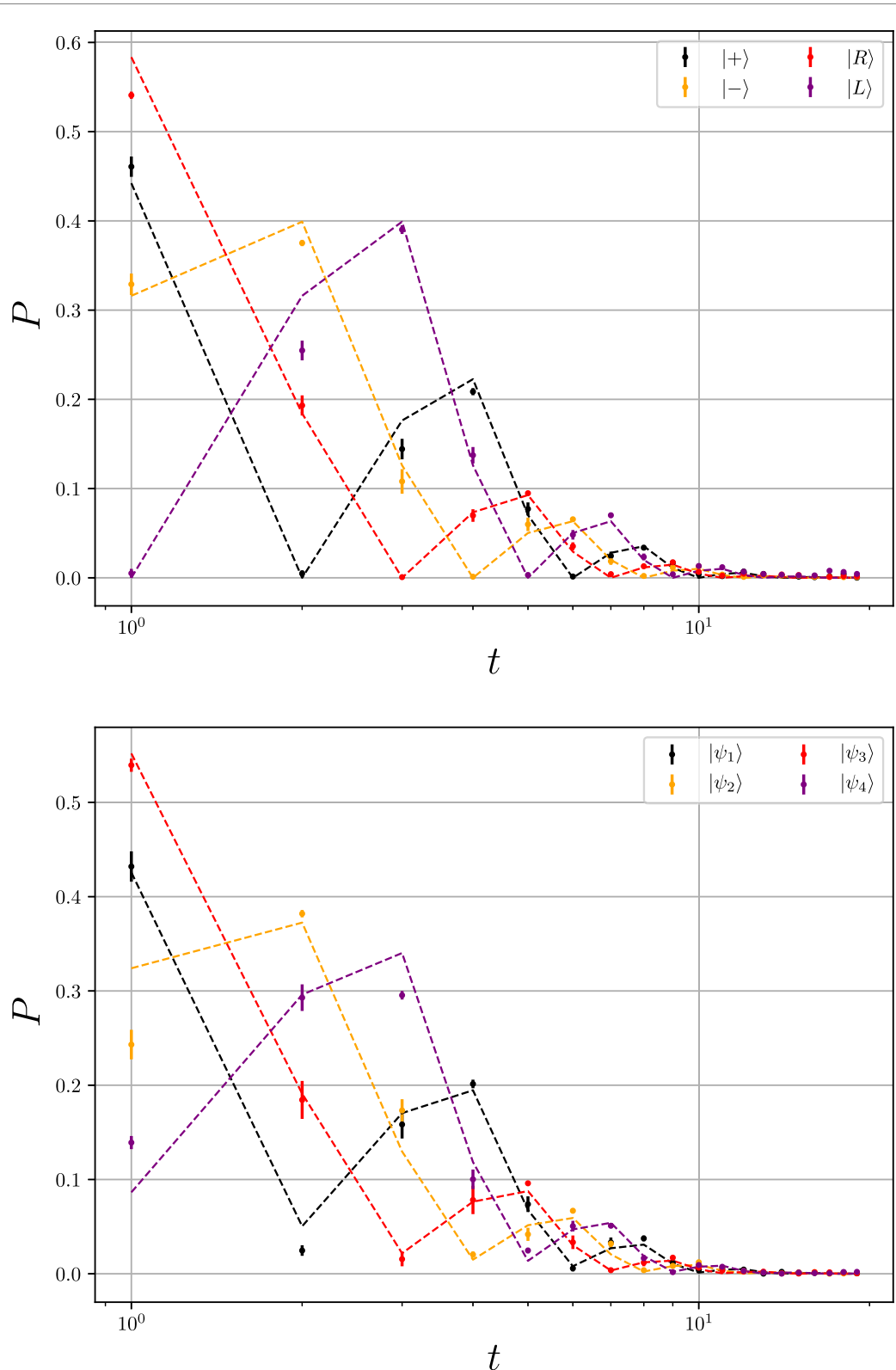
Figure 5. Step-wise output probability for geometrically uniform states and the Tetrad set. Experimental output probability as a function of the number of round trips travelled by photons when the geometrically uniform states (top) and the Tetrad (bottom) states circulate in the network. Normalization of each distribution is performed summing over the total output probability for each single state, in order to account for the experimental signal decrease as the observed time bin grows. Data are reported in comparison with corresponding numerical results (dashed lines).

structure of our receiver and is intimately connected to the number of intermediate layers of it. This particular case of study features two pairs of orthogonal states, providing an intrinsic advantage in producing maximally different time dependent distributions. The same protocol was also applied to the set known as Tetrad set [18], consisting of four maximally equidistant states on the Poincaré sphere:

$$|\psi_1\rangle = 1/\sqrt{3}\,(-|H\rangle + \sqrt{2}\,e^{-2\pi i/3}|V\rangle)$$

$$|\psi_2\rangle = 1/\sqrt{3}\,(-|H\rangle + \sqrt{2}\,e^{2\pi i/3}|V\rangle)$$

$$|\psi_3\rangle = 1/\sqrt{3}\,(-|H\rangle + \sqrt{2}|V\rangle)$$

$$|\psi_4\rangle = |H\rangle$$

(7)

featuring several interesting properties for quantum communication and cryptography. The results of the theoretical analysis are shown in figure 4 for the Tetrad set. For this latter set, the network optimization was carried according to the same method, aiming at maximizing the discrimination probability between different pairs of states at each extraction step. The corresponding evolution matrices, resulting from a numerical maximization, were
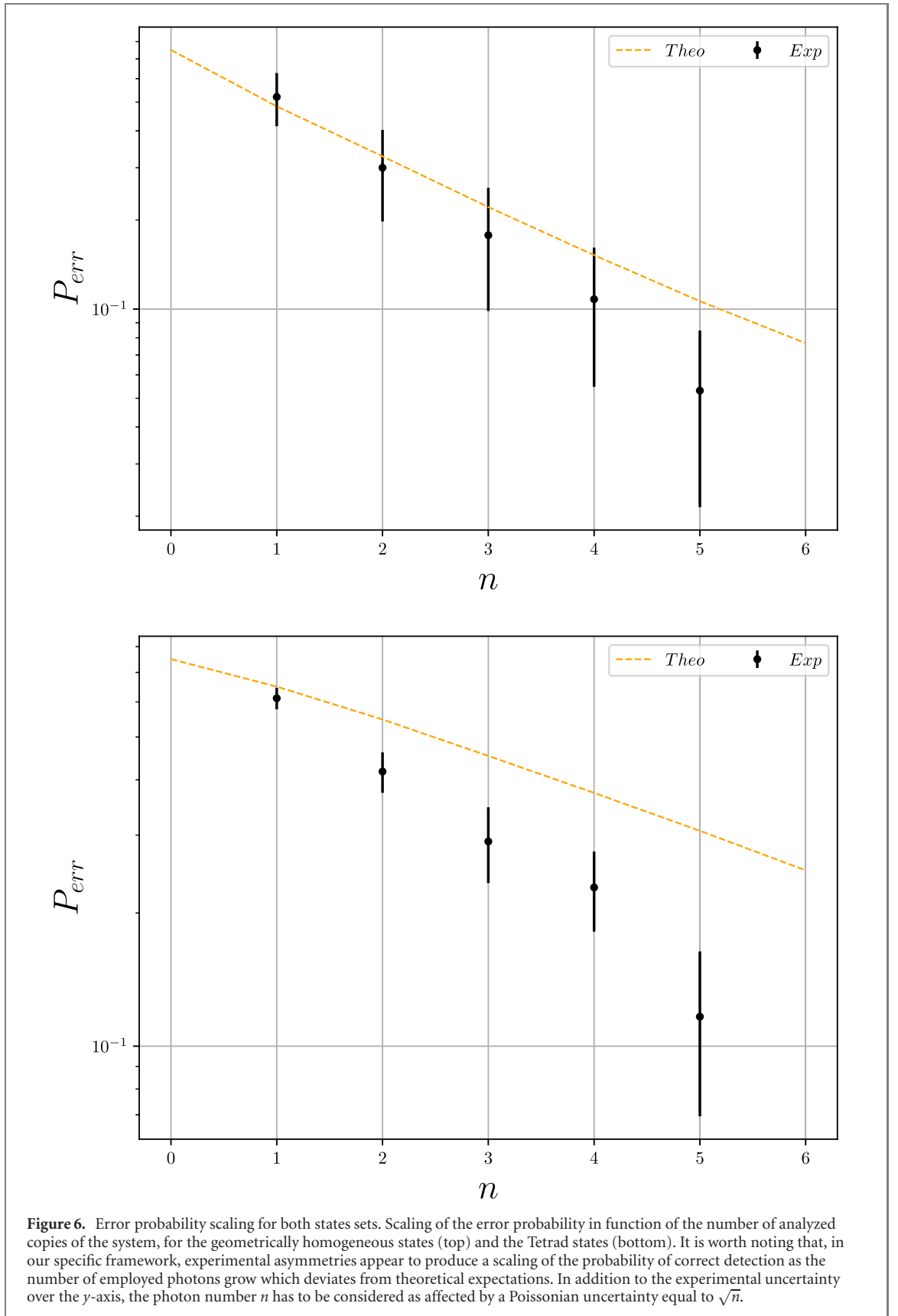
$$U_F = \frac{1}{\sqrt{2}}\begin{pmatrix} 0.953\,021 & 0.302\,905 \\ -0.302\,905 & 0.953\,021 \end{pmatrix}$$

$$U_B = \frac{1}{\sqrt{2}}\begin{pmatrix} -0.674\,645 + 0.216\,571i & 0.211\,8 - 0.673\,121i \\ -0.211\,8 - 0.673\,121i & -0.674\,645 - 0.216\,571i \end{pmatrix}$$

which, again, produce a loop evolution $U_L \sim \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}$. In this case, the method produces a $P_{err} > 0.5$, in this case, the protocol is slightly suboptimal, since the distribution minima are not vanishing (see figure 4 (bottom)), but a more powerful optimization procedure may return more effective results. Nevertheless, as discussed below, as the number of available copies increases, the error probability scales down exponentially.

The resulting experimental detection probability distributions in time for both the sets are reported in figure 5, in comparison with the numerical ones, computed accounting for signal decrease after each evolution step and experimental parameters. The time dependent distributions show a good agreement with the numerical expectations computed accounting for actual parameters, demonstrating a reliable procedure of multi-state discrimination, without the exploitation of any supplementary system, or spreading the states in multi-mode configurations which require an abundance of detecting devices. Specifically, the probability of error which can be computed by these distributions are $P_{compass} = 0.564$ and $P_{tetrad} = 0.591$. The discrepancy with the ideal case values can be quite understood in terms of some experimental asymmetries and non perfect phase-compensation, which produce small but significant polarization rotations, hence slightly lower maxima and higher minima in the distributions. Another very important source of noise consists of the background coincidence counts, which are only partially removed by the procedure described in appendix B. Indeed, according to this scheme, only a single photon-counting detector is needed (the second one provides complementary information), together with the capability of discriminating the arrival time of photons with quite loosen precision. In this case, that capability is indeed provided by a supplemental photon detector which performs coincidence measurements, but, in realistic communication protocols, the second trigger photon can be replaced by more sophisticated synchronization techniques.

In order to verify the effectiveness of the multi-state discrimination protocol in an actual scenario, we tuned the photon source depicted above to a low average photon number regime: through this source, it is not possible to deterministically generate photons, but it is rather possible to set an average rate of generation of single photons. Nevertheless, through this procedure, it is possible to test the average quality of the protocol by computing the average error probability P_{err} as a single copy of the system is available and how this quantity scales as the average available number of copies grows. The experimental strategy to compute the average P_{err} starts by setting the average rate of photon generation per second of the source. An average rate of 1–2 total observed coincidences per second was set, and a higher average photon number was straightforwardly obtained by measuring for a longer time. In this regime, the same time-binned measurements exploited to evaluate the step-wise output distributions can be performed, registering *n*-seconds events for different integration time intervals *n*. We address as *n-seconds event* a time-binned measurement performed for *n* seconds. The events measured in this way have still to be cleaned off the accidental coincidences, as mentioned in the previous section. Hence, it is not possible to consider single instances of *n*-seconds events, but rather an average *n*-seconds event has to be taken into account, to get the chance of subtracting background noise (which is only meaningful as an average quantity).

We consider the temporal probability distributions sampled in the previous section: for each of the two sets, we have four time-wise sampled probability distributions $P_i(t)$, one for each of the states in the set, describing the conditional probability of a photon being extracted at time t, given a certain input state $|\psi_i\rangle$. We can normalize these probability distributions to 1 without particular consequences, since we consider

Figure 6. Error probability scaling for both states sets. Scaling of the error probability in function of the number of analyzed copies of the system, for the geometrically homogeneous states (top) and the Tetrad states (bottom). It is worth noting that, in our specific framework, experimental asymmetries appear to produce a scaling of the probability of correct detection as the number of employed photons grow which deviates from theoretical expectations. In addition to the experimental uncertainty over the y-axis, the photon number n has to be considered as affected by a Poissonian uncertainty equal to $\sqrt{n}$.

the case of equally likely states and we limit observation to only one sink. Hence, if we have an average m-photon event $\bar{E}_m$, measured as described above and cleaned off the noise, we can compute the probability of this event, given the input state we set, as $P(\bar{E}_m|\psi_i) = \prod_t P_i(n_t^{(m)})$ where $n_t^{(m)}$ is the number of photon detected at time t for the event $\bar{E}_m$. As a consequence, the probability of error in guessing the input state can be computed as:

$$P_{\text{err}}(\psi_i, m) = 1 - P(\psi_i|\bar{E}_m) = 1 - \frac{P(\bar{E}_m|\psi_i)}{\sum_{\psi_j} P(\bar{E}_m|\psi_j)} \tag{8}$$

exploiting the Bayes' rule. The resulting trends of P_{err} in function of the average number of detected photons, which act as the available amount of copies of the system, are shown in figure 6, for both sets of states.

5. Discussion and conclusions

In the last two decades genuine quantum features have been often exploited as a catalyst to improve the efficiency of a number of practical tasks. In this scenario, optimal strategies for QSD of actual quantum states have a great deal of relevance and a wide range of applications: quantum communication [26], quantum key distribution [27] and also quantum sensing, in the case of distinguishing different external fields affecting the system dynamics (such as in NV-center noise spectroscopy [28] or avian magnetoreception [29]). In recent years the development of actual single-photon protocols for QSD has been quite rare, while a lot of effort has been spent in developing protocols requiring the exploitation of coherent states (adaptive strategies, quantum phase shift keying [13–15, 17]). However, it is well known that the use of coherent states in quantum communication protocols does not grant equal security as compared to the exploitation of actual quantum states.

In this work, we exploited actual single-photon states and we achieved a nearly optimal protocol featuring a clear spare of resources, in terms of auxiliary systems and physical measurement devices. The spare of auxiliary physical systems is achieved thanks to the network structure of our receiver, which allows us to implement strategies based on time-binning of information extraction, in the case of binary and multi-state discrimination, without the need for supplemental devices; that is a relevant property when regarding for applications. This scenario shares some similarities with the weak-measurement framework developed in [30, 31]; in stark contrast with that, we were able to develop a protocol applying to a higher number of states, to the problem of minimum error probability and with a completely novel theoretical and experimental quantum network approach. The exploitation of actual quantum states in this protocol makes it quite interesting for the application to secure quantum communication tasks and quantum key distribution, tightly relying on the quantum nature of the implementing platforms for their effective realization. Our results represent on one side a basic proof of the protocols effectiveness, but they pave the way to further extensions, such as the adoption of adaptive methods for more general tasks, or the implementation of more complex networks, in order to increase the maximum possible number of states which could be discriminated.

Through these improvements, we will be able to turn our simple model in a proper quantum neural network [32], with a potential which is yet to be uncovered in this field. Noise-robustness of our protocols needs still to be studied, consistently with the approach of [21], in order to reveal possible usefulness for quantum computing in NISQ devices. Even ML techniques need to be applied to our framework, aiming at the development of quantum ML protocols in which quantum information is classified as in classical supervised deep learning schemes. In addition to that, we are currently working on new experimental platforms, which may be suitable to extend our approach to systems characterized by a higher dimensionality. That could be realized exploiting the orbital angular momentum of light [33] or implementing larger optical networks.

In conclusion, a new approach in experimental realizations of QSD protocols has been reported in this work, featuring some practical and theoretical advantages for the applications and leaving a lot of room for improvement and generalization. It may represent a valuable alternative to established strategies, once its potential is fully understood.

Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

Appendix A. Experimental methods

A.1. Setup details

Photon pairs are generated by a high brilliance SPDC source realized according to the model described in [34]: a PPKTP crystal, embedded in a Sagnac interferometer, pumped by a single mode CW laser radiation ($\lambda_{\mathrm{p}} = 405$ nm), which generates collinear pair of photons (*system* and *ancilla*) with opposite polarization at a wavelength $\lambda_{i,s} = 2\lambda_{\mathrm{p}} = 810$ nm. The flux rate of pairs of photons generated by the source in one second is N_{p} s$^{-1} \sim 100\,000$ s^{-1}. This means that the mean time between the generation of one photon and the subsequent one is $t = 1$ s$/100\,000 = 10^{-5}$ s, corresponding to a space separation between two successive photon pairs of $l = t \cdot c \sim 3000$ m. In this conditions, for a certain length x of the setup, which we set to $x \sim 2$ m, there are on average 1500 steps before another photon passes through the loop. We decided to observe 20 steps only of the evolution, which, considering a net loss of signal per loop completion $L = 0.057\%$, we estimated being a good compromise between measurement feasibility and significance of the data. We can estimate the probability that two photon pairs are generated in the time interval needed to perform 20 steps of the evolution, which corresponds to $(x*20/c)/t \sim 0.02$. This means that the probability of having two photon pairs circulating in the setup at the same time is around 2% for all the evolution, a value low enough to perform the experiment also with our CW source, since the measurements are barely affected by the presence of spurious coincidences deriving from photons of other generations.

The photons generated by SPDC are both coupled to a pair of optical fibers, but headed to different outputs: the *system* photon is injected into the setup and actually undergoes the network evolution, while the *ancilla* photon is directly sent to a single-photon detector, acting as a trigger for coincidences. After collection of both photons, their detection is processed by ID-Quantique time tagger ID800: this device features the possibility of setting a narrow coincidences window (up to 81 ps), the capability of recording the relative detection time of a photon, hence the possibility of electronically setting any delay between two photon counters with the aim of computing coincidences and also to perform high-resolved time scanning of the delay between two detecting channels. Thanks to these features, we were able to implement the time-binning discrimination strategy, which is theoretically fancied above. It is worth mentioning that the first considered extraction step is the second, because of the different extraction probability featured by the first one: in fact, an unbalanced beamsplitter (BS) was exploited for the extraction step, featuring a transmittivity $T \sim 70\%$ and a reflectivity $R \sim 30\%$. Therefore, as understandable from figure 2 of the main text, the first extraction step features T as extraction probability, while the remaining ones are characterized by R. Because of that, we chose to neglect the coincidences of photons extracted at the first step.

A.2. Data analysis

The experimental verification of the effectiveness of the time-binning strategy for multi-state discrimination consisted of the experimental reconstruction of the expected output probability distributions. In our framework, the reconstruction consisted of recording the amount of photons being extracted in one of the two sinks after any amount of travelled loops. Since the source generates photon pairs in a non-deterministic fashion, the only way to be sure the counted photons had travelled the right distance in the setup consisted of exploiting a coincidence measurement. Thanks to the ID800 by ID-Quantique, it is possible to count coincidences at any delay between the system and ancilla photon. Therefore, it was possible to observe the amount of photons extracted at each step by suitably tuning the considered delay, depending on the path difference between photons generated within the same temporal window. The measured coincidences had then to be cleaned off the background noise: as described above, the first actual detection step, not taken into account in the analysis, features a much higher signal with respect to the subsequent ones. Therefore, it causes a high amount of accidental coincidences for any set delay. A measurement of the background noise due to the first extraction photons is performed for each delay and directly subtracted to the time-binned measurements, in order to get time-wise coincidences profiles which only display proper coincidences. In conclusion, clean step-wise coincidences counts were obtained, for each input state of both the considered sets. Because of the high amount of detected coincidences, the resulting output was considered as an average result per se. Therefore, the output probability distributions for each input states, displayed in the main text, were directly deduced by the normalization of the clean coincidences profiles.

ORCID iDs

Alessandro Laneve 🄳 https://orcid.org/0000-0002-0416-5131

References

[1] Chefles A 2000 Quantum state discrimination *Contemp. Phys.* **41** 401–24
[2] Bae J and Kwek L-C 2015 Quantum state discrimination and its applications *J. Phys. A: Math. Theor.* **48** 083001
[3] Pusey M F, Barrett J and Rudolph T 2012 On the reality of the quantum state *Nat. Phys.* **8** 475–8
[4] Schmid D and Spekkens R W 2018 Contextual advantage for state discrimination *Phys. Rev. X* **8** 011015
[5] Helstrom C W 1976 *Quantum Detection and Estimation Theory* vol 84 (New York: Academic)
[6] Ivanovic I D 1987 How to differentiate between non-orthogonal states *Phys. Lett.* A **123** 257–9
[7] Croke S *et al* 2006 Maximum confidence quantum measurements *Phys. Rev. Lett.* **96** 070401
[8] Kholevo A S 1973 Bounds for the quantity of information transmitted by a quantum communication channel *Prob. Inf. Transm.* **9** 177–83
    Kholevo A S 1973 *Probl. Pereda. Inf.* **9** 3–11 (Engl. transl.)
[9] Slussarenko S, Weston M M, Li J-G, Campbell N, Wiseman H M and Pryde G J 2017 Quantum state discrimination using the minimum average number of copies *Phys. Rev. Lett.* **118** 030502
[10] Kholevo A S 1974 Remarks on optimal quantum measurements *Probl. Inf. Transm.* **10** 317–20
    Kholevo A S *Probl. Pereda. Inf.* **10** 51–5 (Engl. transl.)
[11] Yuen H, Kennedy R and Lax M 1975 Optimum testing of multiple hypotheses in quantum detection theory *IEEE Trans. Inf. Theory* **21** 125–34
[12] Acín A, Bagan E, Baig M, Masanes L and Muñoz-Tapia R 2005 Multiple-copy two-state discrimination with individual measurements *Phys. Rev.* A **71** 032338
[13] Becerra F E, Fan J, Baumgartner G, Goldhar J, Kosloski J T and Migdall A 2013 Experimental demonstration of a receiver beating the standard quantum limit for multiple nonorthogonal state discrimination *Nat. Photon.* **7** 147–52
[14] Becerra F E, Fan J and Migdall A 2013 Implementation of generalized quantum measurements for unambiguous discrimination of multiple non-orthogonal coherent states *Nat. Commun.* **4** 2028
[15] Ferdinand A R, DiMario M T and Becerra F E 2017 Multi-state discrimination below the quantum noise limit at the single-photon level *npj Quantum Inf.* **3** 43
[16] Burenkov I A *et al* 2020 Time-resolving quantum measurement enables energy-efficient, large-alphabet communication *PRX Quantum* **1** 010308
[17] Sidhu J S *et al* 2021 Quantum receiver for phase-shift keying at the single-photon level *PRX Quantum* **2** 010332
[18] Clarke R B M *et al* 2001 Experimental demonstration of optimal unambiguous state discrimination *Phys. Rev.* A **63** 040305
[19] Solís-Prosser M A, Fernandes M F, Jiménez O, Delgado A and Neves L 2017 Experimental minimum-error quantum-state discrimination in high dimensions *Phys. Rev. Lett.* **118** 100501
[20] Patterson A, Chen H, Wossnig L, Severini S, Browne D and Rungger I 2021 Quantum state discrimination using noisy quantum neural networks *Phys. Rev. Res.* **3** 013063
[21] Pozza N D and Caruso F 2020 Quantum state discrimination on reconfigurable noise-robust quantum networks *Phys. Rev. Res.* **2** 043011
[22] Fanizza M, Mari A and Giovannetti V 2019 Optimal universal learning machines for quantum state discrimination *IEEE Trans. Inf. Theory* **65** 5931–44
[23] Kurzyński P and Wójcik A 2013 Quantum walk as a generalized measuring device *Phys. Rev. Lett.* **110** 200404
[24] Li Z, Zhang H and Zhu H 2019 Implementation of generalized measurements on a qudit via quantum walks *Phys. Rev.* A **99** 062342
[25] Whitfield J D, Rodríguez-Rosario C A and Aspuru-Guzik A 2010 Quantum stochastic walks: a generalization of classical random walks and quantum walks *Phys. Rev.* A **81** 022323
[26] Fujiwara M, Takeoka M, Mizuno J and Sasaki M 2003 Exceeding the classical capacity limit in a quantum optical channel *Phys. Rev. Lett.* **90** 167906
[27] Cerf N J, Bourennane M, Karlsson A and Gisin N 2002 Security of quantum key distribution using d-level systems *Phys. Rev. Lett.* **88** 127902
[28] Degen C L, Reinhard F and Cappellaro P 2017 Quantum sensing *Rev. Mod. Phys.* **89** 035002
[29] Lambert N, Chen Y-N, Cheng Y-C, Li C-M, Chen G-Y and Nori F 2013 Quantum biology *Nat. Phys.* **9** 10–8
[30] Bergou J, Feldman E and Hillery M 2013 Extracting information from a qubit by multiple observers: toward a theory of sequential state discrimination *Phys. Rev. Lett.* **111** 100501
[31] Namkung M and Kwon Y 2018 Analysis of optimal sequential state discrimination for linearly independent pure quantum states *Sci. Rep.* **8** 6515
[32] Schuld M, Sinayskiy I and Petruccione F 2014 The quest for a quantum neural network *Quantum Inf. Process.* **13** 2567–86
[33] Nagali E *et al* 2009 Quantum information transfer from spin to orbital angular momentum of photons *Phys. Rev. Lett.* **103** 013601
[34] Fedrizzi A, Herbst T, Poppe A, Jennewein T and Zeilinger A 2007 A wavelength-tunable fiber-coupled source of narrowband entangled photons *Opt. Express* **15** 15377–86