Experimental quantum-enhanced kernels on a photonic processor

Iris Agresti,[1,*] Zhenghao Yin,[1] Giovanni de Felice,[2] Douglas Brown,[2] Alexis Toumi,[2] Ciro Pentangelo,[3] Simone Piacentini,[3] Andrea Crespi,[3] Francesco Ceccarelli,[3] Roberto Osellame,[3] Bob Coecke,[2] and Philip Walther[1,4]

[1] *University of Vienna, Faculty of Physics, Vienna Center for Quantum Science and Technology (VCQ), Boltzmanngasse 5, Vienna A-1090, Austria*
[2] *Quantinuum, 17 Beaumont Street, Oxford OX1 2NA, UK*
[3] *Istituto di Fotonica e Nanotecnologie, Consiglio Nazionale delle Ricerche (IFN-CNR), piazza L. Da Vinci 32, 20133 Milano, Italy*
[4] *Christian Doppler Laboratory for Photonic Quantum Computer, Faculty of Physics, University of Vienna, 1090 Vienna, Austria*

In recent years, machine learning and quantum computing have revolutionized the field of computer science. In particular, the first allowed for a wide range of possibilities, from research to everyday life scopes. Despite the generality of these models, the amount of computational resources grows with the complexity of the task and quantum (or quantum inspired) protocols can ease such requirements. Here, we demonstrate a kernel method on a photonic integrated platform to perform a binary classification task, exploiting quantum interference. We benchmark our protocol against standard algorithms and show that it outperforms them for the considered task. This result opens the way to more efficient computing algorithms and for the formulation of tasks where quantum effects boost the performance of well-established classical methods.

In the last decades, a flurry of interest has been devoted to the development of technologies based on quantum mechanical phenomena, which have promised to outperform their classical counterpart. However, a clear advantage of quantum computation has been experimentally demonstrated only recently and on computational tasks with no practical applications [1–3]. In this context, it is crucial to investigate what kind of tasks can be effectively enhanced by quantum computing with respect to the operation of classical computers, within state-of-art technologies. For instance, an open question is whether quantum computing can have an impact on machine learning.

Here, we investigate a classical-quantum hybrid machine learning model in the form of a Kernel method, i.e. an algorithm that maps data points to a (high-dimensional) *feature space*. This non-linear transformation enhances the identification of patterns, which are not evident in the initial space. Indeed, once a suitable mapping is performed, it is possible to search for the hyperplane which best groups the data into distinct classes, through a support vector machine (SVM) [4], according to the inner product of the mapped data. Hence, an interesting question is whether using a quantum apparatus to perform the data mapping and evaluate the inner products among the resulting feature points can lead to an enhanced performance. Such a hybrid classical-quantum model would benefit, on one hand, from the quantum feature maps that result from the evolution of quantum systems (which can be hard to simulate on classical computers) and, on the other, it would outsource the hardest part of the computation to the quantum hardware. Our experimental demonstration relies on mapping data points in the feature space through the unitary evolution of two-boson Fock states (see Fig. 1). To experimentally demonstrate this method, we exploit a photonic platform and, in particular, an integrated photonic processor [5] where we inject two-boson Fock states to map the data to be classified. Such an encoding allows us to arbitrarily tune the dimension of the feature space and, even for relatively small dimensions, it provides a strong non-linearity to achieve a high classification accuracy with data which is non-linearly separable.

A kernel method relies on a function that maps N input data points x_i, that need to be binarily classified, from a space $\mathcal{X} \subseteq \mathbb{R}^d$ into a feature space $\mathcal{H}$, through a *feature map* $\Phi : \mathcal{X} \to \mathcal{H}$. Then, a SVM can be used to produce a *prediction* function $f_K : \mathcal{X} \to \mathbb{R}$ as $f_K(x) = \sum_i \alpha_i K(x, x_i)$, where these α_i coefficients are obtained by solving a linear optimization problem. Hence, the inputs of the optimization are the labels y and the matrix obtained by computing the pairwise distances between data points $K_{i,j} = K(x_i, x_j) = \langle \Phi(x_j) | \Phi(x_i) \rangle$, the so-called *Gram matrix*.

In this work, we implement a quantum version of the kernel method, in which the pairwise distances between data points are estimated by sampling from the output probability distribution arising from the unitary evolution of a Fock input state. Therefore, our feature map plugs the data that needs to be classified into the free parameters defining a unitary evolution applied to a fixed indistinguishable photon Fock state of dimension $m = 6$ and whose sum of occupational numbers is $n = 2$, as follows: $x \mapsto |\Phi(x)\rangle = U_x|\psi\rangle$. Then, to select a classification task that would

* iris.agresti@univie.ac.at

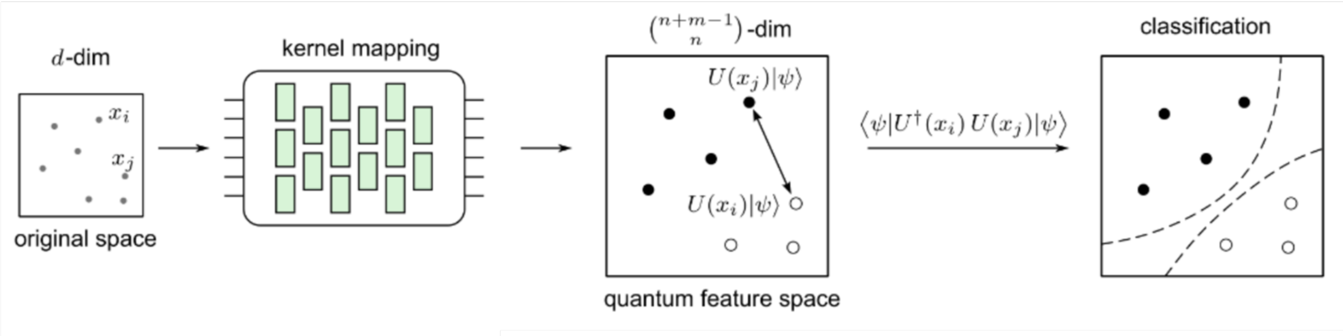

FIG. 1. **Photonic quantum kernel estimation. a**. The photonic quantum kernel maps each data point x_i to be classified from a d-dimensional space into a quantum state $|\Phi\rangle_i$, living in a Hilbert feature space. In detail, the classical data x_i is encoded as follows: $|\Phi\rangle_i = U(x_i)|\psi\rangle$. Then, from the inner pairwise products, we perform the classification finding the hyperplane best separating the classes, i.e. through a classical support vector machine (SVM).

benefit from the described model, we use a quantifier called the *geometric difference* [6]. This allows to compare the performance of two kernels K_a and K_b and therefore to pick the task that is most suitable for either of them. In our case, we want to select a task where our photonic kernel, displaying the interference of indistinguishable photons, performs better than its counterpart produced by distinguishable photons (we will refer to the latter as *coherent kernel*).

Our experimental setup consists of two parts, a single-photon source generating the input states and a programmable integrated photonic processor. First, to generate the input state, we use a type II spontaneous parametric down-conversion source, which generates frequency degenerate single-photon pairs at 1546 nm in a periodically poled K-titanyl phosphate crystal. The two photons are then made indistinguishable in their polarization and arrival time, respectively, via wave retarders and a delay line, which we also use to tune the degree of indistinguishability of the generated photons. As mentioned before, our feature map sends each data point x_i onto the state resulting from the evolution $U(x_i)$ of a fixed input Fock state $|\psi\rangle$. Then, for the application of the SVM, which finds the best hyperplane separating the data, we need to evaluate the inner products between all of the points x_i, x_j in the feature space, which amounts to $\langle\psi|U(x_i)^\dagger U(x_j)|\psi\rangle$. This implies that, if we take $|\psi\rangle$ as a Fock state of 2 photons over 6 modes, the inner product $\langle\Phi(x_i)|\Phi(x_j)\rangle$ is given by projecting the evolved state $U(x_i)^\dagger U(x_j)\,|\psi\rangle$ onto $|\psi\rangle$. To this aim, we employ an integrated photonic processor [5] fabricated on a borosilicate glass substrate, in which optical waveguides are inscribed through femtosecond laser writing. The circuit features six input/output modes and it is based on a rectangular mesh of 15 programmable Mach-Zehnder interferometers [7] (see Fig. 2). Each interferometer is equipped with two thermal phase shifters, allowing for tunable reflectivity and phase. By properly choosing the values of the phase shifters, such arrangement allows us to perform any unitary transformation on the input photon states. At the output, detection is performed by superconducting nanowire single-photon detectors (SNSPDs). Due to the fact that these detectors are not photon-number resolving, we post-select the output events to those featuring two detectors clicking, i.e. the *collision-free* events.

We test the performance of two photonic kernels injecting the following input: $|\psi\rangle = |1,1,0,0,0,0\rangle$. Moreover, we are able to tune the indistinguishability to implement the quantum kernel and the coherent kernel. We test datasets consisting in a different number of data points: 40, 60, 80 and 100. For each of those sets, we use 2/3 of the data points for the training of the SVM, and the remaining 1/3 as a test set. We then estimate all of the pairwise products between the unitaries $U(x_i)^\dagger U(x_j)$. Hence, $|\langle\psi|U(x_i)^\dagger U(x_j)|\psi\rangle|^2$ is given by the probability of detecting the photons on the same modes from which they were injected. This implies that, for a dataset consisting of N data points, we perform $N(N-1)/2$ unitaries to compute all of the inner products. The accuracy is defined as the number of correctly classified points over the total size of the test set. Let us note that values lower than 0.5 indicate that the model was not able to learn the features of the training set and generalize to unknown data.

In Fig. 2b, we report the average test accuracy obtained for five different datasets with the same size, varying the dataset sizes from 40 to 100 as well. Moreover, the results obtained with the quantum kernel (blue curves) and the coherent kernel (orange curve) are compared with the following numerical kernels: Gaussian (grey curve), polynomial (yellow) and linear (purple). For the latter three kernel methods, we consider the maximal accuracy obtained by optimizing their hyperparameters. Although the task is built only comparing the performance of the kernels based on indistinguishable and distinguishable photons, the obtained accuracy is higher also than commonly used classical

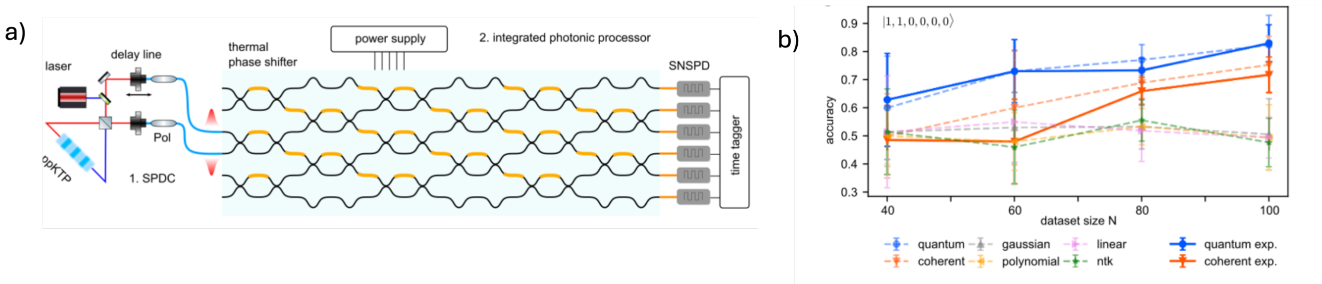

FIG. 2. **Implementation of photonic quantum kernel estimation. a.** The frequency degenerate photons are generated by a type-II spontaneous parametric down-conversion source. Afterwards, they are made indistinguishable in their polarization and arrival time and they are injected in two modes of an integrated photonic processor with 6 input/output modes [5]. The degree of indistinguishability can then be tuned through a delay line, changing their relative temporal delay. **b.** The average classification accuracies on 5 different sets for the quantum kernel (blue curves) and the coherent kernel (orange curve), along with the following classical kernels: Gaussian (grey), polynomial (yellow) and linear (purple). The input state considered is $|1,1,0,0,0,0\rangle$, i.e. where 2 photons are injected from the first and second mode of the circuit. The dashed line indicates the results of numerical simulations, while the solid ones the experimental results. The error bar shows the standard deviation of the classification accuracies on 5 datasets for all the kernels.

kernels [8, 9]. The dashed lines indicate the results of numerical simulations, while the solid lines indicate experimental results.

Despite the fact that we selected the classification task, to show a better performance of the quantum kernel (exhibiting quantum interference), with respect to the coherent one (displaying no quantum interference), both these photonic kernels perform the selected tasks significantly better than commonly used kernels, such as the Gaussian, polyonomial and linear kernels. These results open the way for further investigations related to the non-linearities that can be achieved through photonic platforms [10], which are crucial elements for machine learning purposes and, in particular, for neuromorphic computation models, such as *reservoir computing* [11]. This may be of remarkable importance when considering difficulties related to energy consumption, as it was shown that partially optical networks can be adopted to reduce the overall energy requirements with respect to electronic ones [12].

[1] Zhong, H.-S. *et al.* Quantum computational advantage using photons. *Science* **370**, 1460–1463 (2020).
[2] Madsen, L. S. *et al.* Quantum computational advantage with a programmable photonic processor. *Nature* **606**, 75–81 (2022).
[3] Arute, F. *et al.* Quantum supremacy using a programmable superconducting processor. *Nature* **574**, 505–510 (2019).
[4] Cortes, C. & Vapnik, V. Support-vector networks. *Machine learning* **20**, 273–297 (1995).
[5] Pentangelo, C. *et al.* High-fidelity and polarization-insensitive universal photonic processors fabricated by femtosecond laser writing. *Nanophotonics* (2024). URL https://www.degruyter.com/document/doi/10.1515/nanoph-2023-0636/html. Publisher: De Gruyter.
[6] Huang, H. Y. *et al.* Power of data in quantum machine learning. *Nature Communications* **12** (2021). URL http://dx.doi.org/10.1038/s41467-021-22539-9. ArXiv: 2011.01938 Publisher: Springer US.
[7] Clements, W. R., Humphreys, P. C., Metcalf, B. J., Kolthammer, W. S. & Walsmley, I. A. Optimal design for universal multiport interferometers. *Optica* **3**, 1460 (2016). ArXiv: 1603.08788.
[8] Shawe-Taylor, J. & Cristianini, N. *Kernel methods for pattern analysis* (Cambridge university press, 2004).
[9] Hofmann, T., Schölkopf, B. & Smola, A. J. Kernel methods in machine learning (2008).
[10] Spagnolo, M. *et al.* Experimental photonic quantum memristor. *Nature Photonics* **16**, 318–323 (2022).
[11] Govia, L., Ribeill, G., Rowlands, G. & Ohki, T. Nonlinear input transformations are ubiquitous in quantum reservoir computing. *Neuromorphic Computing and Engineering* **2**, 014008 (2022).
[12] Hamerly, R., Bernstein, L., Sludds, A., Soljačić, M. & Englund, D. Large-scale optical neural networks based on photoelectric multiplication. *Physical Review X* **9**, 021032 (2019).