



OPEN

## Speeding up quantum perceptron via shortcuts to adiabaticity

Yue Ban<sup>1,2</sup>✉, Xi Chen<sup>1,3</sup>, E. Torrontegui<sup>4,5</sup>, E. Solano<sup>1,3,6,7</sup> & J. Casanova<sup>1,6</sup>

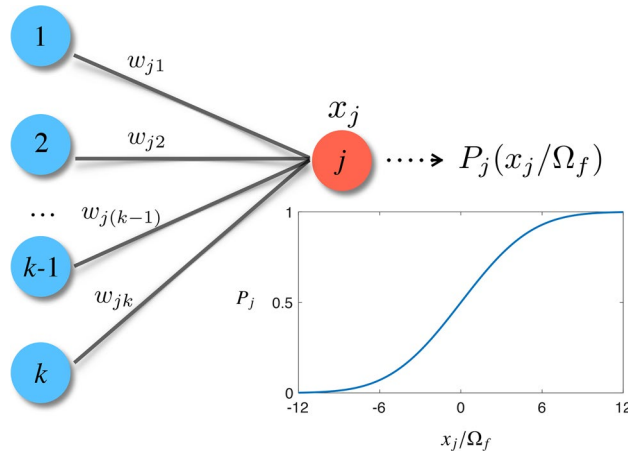
The quantum perceptron is a fundamental building block for quantum machine learning. This is a multidisciplinary field that incorporates abilities of quantum computing, such as state superposition and entanglement, to classical machine learning schemes. Motivated by the techniques of shortcuts to adiabaticity, we propose a speed-up quantum perceptron where a control field on the perceptron is inversely engineered leading to a rapid nonlinear response with a sigmoid activation function. This results in faster overall perceptron performance compared to quasi-adiabatic protocols, as well as in enhanced robustness against imperfections in the controls.

In the era of information expansion, the merge of quantum information and artificial intelligence will have a transformative impact in science, technology, and our societies<sup>1–3</sup>. In particular, classical networks of artificial neurons (or nodes) represent a successful framework for machine learning strategies, with the *perceptron* being the simplest example of a node<sup>4</sup>. The perceptron is based on the McCulloch-Pitts neuron<sup>5</sup>, and it was originally proposed by Rosenblatt in 1957 to create the first trained networks<sup>6</sup>. Nowadays, extensions of these original ideas such as multilayer perceptrons in networks with interlayer connectivity are exploited to deal with demanding computational tasks.

The emergence of quantum computing and machine learning has boosted the development of both fields<sup>7–13</sup>, giving rise to the field of quantum machine learning. In this context, quantum neural networks (QNNs) have attracted growing interest<sup>14,15</sup> since the seminal idea proposed by Kak<sup>16</sup>. In particular, the entering of classical machine learning techniques into the quantum domain has the potential to accelerate the performance of different applications such as classification and pattern recognition<sup>2,17–23</sup>. In addition, nowadays the excellent degree of quantum control over the registers in modern quantum platforms<sup>24–27</sup> allows the performance of quantum operations with high fidelity, which further feeds the idea of having reliable QNNs. However, the linear and unitary framework of quantum mechanics raises a serious dilemma, since neural networks present nonlinear and dissipative behaviours which are hard to reproduce at the quantum level. To address this challenge, many efforts have been attempted by exploiting quantum measurements<sup>16,28</sup>, the quadratic kinetic term to generate nonlinear behaviours<sup>29</sup>, dissipative<sup>16</sup> or repeat-until-success<sup>30</sup> quantum gates, and reversible circuits<sup>31</sup>. Among them, gate-based QNNs<sup>32</sup> with training optimization procedures<sup>33</sup> are feasible to implement by a set of unitary operations. Furthermore, gate-based QNNs can behave as variational quantum circuits that encode highly nonlinear transformations while remaining unitary<sup>20</sup>. Also, a quantum algorithm implementing the quantum version of a binary-valued perceptron was introduced in Ref.<sup>18</sup>, showing an exponential advantage in resources storage. Remarkably, a universal *quantum perceptron* has been proposed as an efficient approximator in Ref.<sup>34</sup>, where the quantum perceptron is encoded in an Ising model with a sigmoid activation function. In particular, the sigmoid nonlinear response is parametrized by the potential exerted by other neurons, and driven by adiabatic techniques.

In this article, motivated by the nonadiabatic control provided by shortcuts to adiabaticity (STA) techniques<sup>35,36</sup>, we design fast sigmoidal responses with the aid of the invariant-based inverse engineering (IE)<sup>37–39</sup>. The IE method is based on dynamical modes of Lewis-Riesenfeld invariant instead of one instantaneous eigenstates of the original reference Hamiltonian<sup>40,41</sup>. As IE directly imposes boundary conditions in the wave function evolution, the nonlinear activation function of the quantum perceptron encoded in the probability of the excited state can be achieved in a fast and robust way. In particular, an external control field on the perceptron is designed such that it leads to a fast nonlinear activation function with a wide tolerance window to the variation of the input potential induced by neurons in the previous layer. We demonstrate that our method produces

<sup>1</sup>Department of Physical Chemistry, University of the Basque Country UPV/EHU, Apartado 644, 48080 Bilbao, Spain. <sup>2</sup>School of Materials Science and Engineering, Shanghai University, 200444 Shanghai, People's Republic of China. <sup>3</sup>International Center of Quantum Artificial Intelligence for Science and Technology (QuArtist) and Department of Physics, Shanghai University, 200444 Shanghai, People's Republic of China. <sup>4</sup>Departamento de Física, Universidad Carlos III de Madrid, Avda. de la Universidad 30, 28911 Leganés (Madrid), Spain. <sup>5</sup>Instituto de Física Fundamental IFF-CSIC, Calle Serrano 113, 28006 Madrid, Spain. <sup>6</sup>IKERBASQUE, Basque Foundation for Science, Plaza Euskadi 5, 48009 Bilbao, Spain. <sup>7</sup>IQM, Nymphenburgerstr. 86, 80636 Munich, Germany. ✉email: ybanxc@gmail.com



**Figure 1.** Schematic configuration of a quantum perceptron. When it is integrated in a feed-forward neural network, the potential depends on neurons in earlier layers, e.g.,  $\hat{x}_j = \sum_{i=1}^k w_{ji} \hat{\sigma}_i^z - b_j$ , where the activation function of the quantum perceptron is the probability of the excited state  $P_j(x_j / \Omega_f)$  at the final time  $t = t_f$  in the form of sigmoid-shape, shown in the inset.

solutions that outperform those based on adiabatic techniques, which significantly facilitates the implementation of quantum perceptrons in modern platforms such as nitrogen vacancy (NV) centers in diamond. Note that, the latter are settings where external control fields can be introduced with extraordinary precision<sup>42</sup>.

**Results**

**Quantum perceptron.** The capacity of feed-forward neural networks to classify complex data relies in the “universal approximation theorem” proved by Cybenko<sup>43</sup>, claiming that any continuous function can be written as a linear combination of sigmoid functions. A QNN is also demonstrated as a universal approximator of continuous functions<sup>34</sup>. In a classical network, a perceptron (or neuron) generates the signal  $s_j = f(x_j)$  as a sigmoidal response to the weighted sum of the signals (or outputs) from the neurons in the previous layer. More specifically,  $x_j = \sum_{i=1}^k w_{ji} s_i - b_j$  with the neuron interconnectivities  $w_{ji}$ , the bias  $b_j$ , and  $s_i$  being the output of the  $i$ th neuron in the previous layer. In analogy with classical neurons, a quantum perceptron can be constructed as a qubit that encodes the nonlinear response to an input potential in the excitation probability, see Fig. 1. One possibility for the latter is the following gate<sup>34</sup>:

$$\hat{U}_j(\hat{x}_j; f) |0_j\rangle = \sqrt{1 - f(\hat{x}_j)} |0_j\rangle + \sqrt{f(\hat{x}_j)} |1_j\rangle, \tag{1}$$

where, in close similarity with the classical case, we have  $\hat{x}_j = \sum_{i=1}^k w_{ji} \hat{\sigma}_i^z - b_j$ , where  $\hat{\sigma}_i^z$  is the  $z$  Pauli matrix of the  $i$ th neuron (qubit),  $w_{ji}$  is interaction between the perceptron  $j$  and the  $i$ th neuron in the previous layer,  $b_j$  is the bias of the perceptron. The transformation in Eq. (1) can be engineered by evolving adiabatically the qubit with the Ising Hamiltonian ( $\hbar = 1$ )

$$\begin{aligned} \hat{H}(t) &= \frac{1}{2} \left[ \hat{x}_j \hat{\sigma}_j^z + \Omega(t) \hat{\sigma}_j^x \right] \\ &= \frac{1}{2} \left[ \sum_{i=1}^k (w_{ji} \hat{\sigma}_i^z \hat{\sigma}_j^z) - b_j \hat{\sigma}_j^z + \Omega(t) \hat{\sigma}_j^x \right], \end{aligned} \tag{2}$$

where the  $j$ th qubit (encoding the quantum perceptron) is controlled by an external field  $\Omega(t)$ , leading to a tunable energy gap in the dressed-state qubit basis  $|\pm\rangle$ , with  $\hat{\sigma}_j^x |\pm\rangle = \pm |\pm\rangle$ . When this perceptron is integrated in a feed-forward neural network, the potential depends on the neurons in earlier layers, as the perceptron interacts with other neurons in the previous layer (labeled by  $i = 1, \dots, k$ ) via the  $x_j$  potential, see Fig. 1. Therefore, the network is encoded in a Hilbert space via the external potential exerted by other neurons. The Ising Hamiltonian in Eq. (2) has the reduced eigenstate,

$$|\Phi(x_j / \Omega(t))\rangle = \sqrt{1 - f(x_j / \Omega(t))} |0\rangle + \sqrt{f(x_j / \Omega(t))} |1\rangle, \tag{3}$$

where  $x_j$  now represents the lowest eigenvalue of the operator  $\hat{x}_j$ , while  $f(x)$  corresponds to a sigmoid excitation probability

$$f(x) = \frac{1}{2} \left( 1 + \frac{x}{\sqrt{1 + x^2}} \right). \tag{4}$$

In order to generate the state on the right side of Eq. (1), we propose the following strategy: First, a Hadamard gate is applied to drive the state from  $|0\rangle$  to  $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ . Secondly, by appropriately tuning  $\Omega(t)$  according to inverse engineering (IE) techniques (to be explained later), the state  $|\Psi(0)\rangle = |+\rangle$  evolves to  $|\Psi(t_f)\rangle = |\Phi(x_j/\Omega_f)\rangle$  (up to some phase factor that can be eventually canceled by a phase gate), along with one eigenstate of the Lewis-Riesenfeld invariant of  $\hat{H}$ , with  $|\Phi(x_j/\Omega_f)\rangle$  being the instantaneous eigenstate of  $\hat{H}(t = t_f; \Omega_f)$ , and  $\Omega_f \equiv \Omega(t_f)$ . It is noteworthy to mention that, unlike the fast quasi-adiabatic passage (FAQUAD) approach<sup>34</sup>, our method based on IE does not need to achieve the initial condition  $\Omega(0) \gg |x_j|$ , as it is not required that the initial state meets one eigenstate of  $\hat{H}(0)$ . The latter results in a smooth control field  $\Omega(t)$  which is easy to be used in experiments.

Another possibility to achieve  $|\Psi(t_f)\rangle$  from  $|\Psi(0)\rangle$  is by an adiabatic driving in a Landau-Zener scheme. However, as it is discussed in Ref.<sup>34</sup>, this spends long time and may be unfeasible depending on the coherence time of the physical setup that implements the Hamiltonian in Eq. (2).

**Accelerating quantum perceptron by IE.** We adopt the IE method to achieve the  $|\Psi(0)\rangle \rightarrow |\Phi(x_j/\Omega_f)\rangle$  state transfer with shorter time than FAQUAD<sup>44</sup>. The control field  $\Omega(t)$  is then engineered to guarantee that at the final evolution time  $t = t_f$  the qubit excitation probability  $P_j(x_j/\Omega_f)$  corresponds to a sigmoid-like response, i.e. to a mono-valuate  $f$  function satisfying  $\lim_{x \rightarrow -\infty} f(x) \rightarrow 0$  and  $\lim_{x \rightarrow \infty} f(x) \rightarrow 1$ . Since the universality of neural networks does not rely on the specific shape of the sigmoid function<sup>43,45</sup>, e.g. Eq. (4), we quantify the performance of the control field  $\Omega(t)$  in the interval  $[-x^{\max}, x^{\max}]$  with the distance  $C = 2 - F_0 - F_1$ . Here  $F_0 = |\langle 0 | \Psi(t_f; x_j/\Omega_f = -x^{\max}) \rangle|^2$  and  $F_1 = |\langle 1 | \Psi(t_f; x_j/\Omega_f = x^{\max}) \rangle|^2$  characterize how the engineered states overlap with  $|0\rangle$  and  $|1\rangle$ , at  $x_j/\Omega_f = -x^{\max}$  and  $x_j/\Omega_f = x^{\max}$  respectively. Note that, for a sigmoid-like function,  $C \rightarrow 0$ , for  $x^{\max} \rightarrow \infty$ . Meanwhile, in all the numerical results, the activation function is found to be well-behaved, i.e., the function is monotonic and with a sigmoid-like behaviour,  $\lim_{x \rightarrow -\infty} f(x) \rightarrow 0$  and  $\lim_{x \rightarrow \infty} f(x) \rightarrow 1$ . As we will see later, our IE technique also provides with robustness with respect to timing errors.

Now we show the procedure to find the control  $\Omega(t)$ . To this end, we start with the parameterisation of the dynamical state

$$|\Psi(t)\rangle = \cos(\theta/2)e^{i\beta/2}|0\rangle + \sin(\theta/2)e^{-i\beta/2}|1\rangle, \quad (5)$$

with the two unknown polar and azimuthal angles,  $\theta \equiv \theta(t)$  and  $\beta \equiv \beta(t)$ , on the Bloch sphere. Having the state in Eq. (5) at hand, the corresponding orthogonal state  $|\Psi_{\perp}(t)\rangle$  gets completely determined and the Lewis-Riesenfeld invariant can be thus constructed with constant eigenvalues<sup>37,38</sup>. Substituting one of the states ( $|\Psi(t)\rangle$  or  $|\Psi_{\perp}(t)\rangle$ ) into the time-dependent Schrödinger equation driven by the Hamiltonian in Eq. (2), we obtain the following coupled differential equations (for more details see Methods.)

$$\Omega(t) = \dot{\theta} / \sin \beta, \quad (6)$$

$$x_j = \dot{\theta} \cot \theta \cot \beta - \dot{\beta}. \quad (7)$$

Setting the wavefunction  $|\Psi(0)\rangle = |+\rangle$  and  $|\Psi(t_f)\rangle = |\Phi(x_j/\Omega_f)\rangle$  at the initial and final times leads to the boundary conditions

$$\begin{aligned} \theta(0) &= 2 \sin^{-1} \left[ \sqrt{f(x_j/\kappa)} \right], \\ \theta(t_f) &= 2 \sin^{-1} \left[ \sqrt{f(x_j/\Omega_f)} \right], \end{aligned} \quad (8)$$

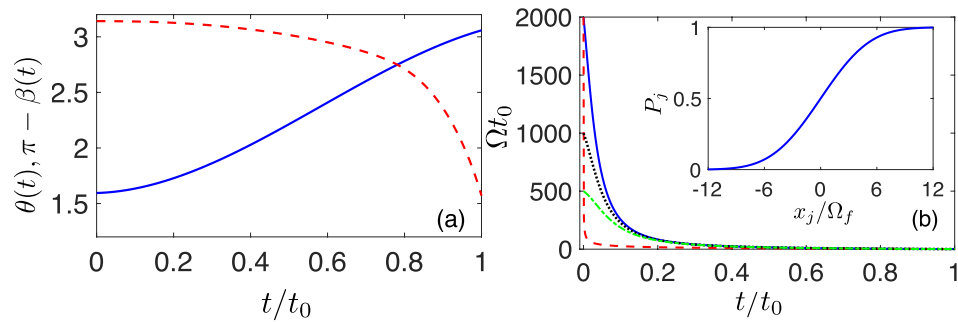
with the introduced  $\kappa$  parameter being infinitely large which results in  $|\Phi(x_j/\kappa)\rangle = |+\rangle$ . Also, it is important to remark that  $\kappa$  does not need to equal the value of our control  $\Omega(t)$  at  $t = 0$ , as  $|\Phi(x_j/\kappa)\rangle$  is not necessarily the eigenstate of  $\hat{H}[t = 0; \Omega(0)]$ . In addition, from Eq. (6) one can find the following conditions for the first derivatives of  $\theta$  at the boundaries

$$\dot{\theta}(0) = \Omega(0) \sin \beta(0), \quad \dot{\theta}(t_f) = \Omega_f \sin \beta(t_f). \quad (9)$$

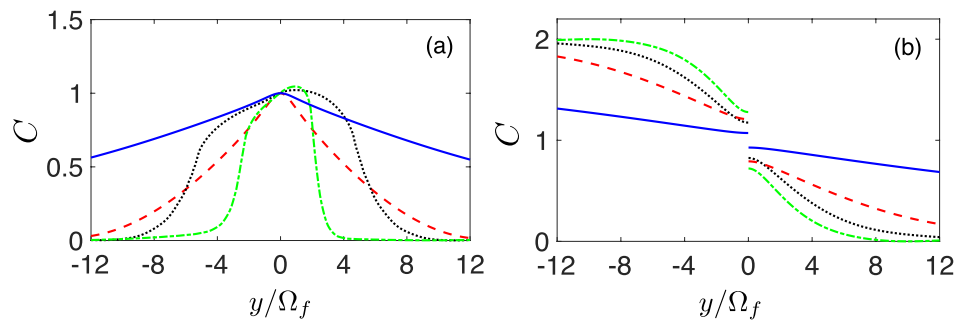
We can interpolate  $\theta$  by choosing a simple polynomial function  $\theta = \sum_{i=0}^N a_i t^i$  and a trigonometric function  $\theta = a_0 + a_1 t + \sum_{i=2}^N a_i \sin[(i-1)\pi t/t_f]$  with less coefficients required for matching the same boundary conditions<sup>46</sup>. The appropriate adoptions on the coefficients can make the solution approach the one gained from optimal control theory<sup>47</sup>. We present the comparison of the performance of activation function by using IE with these two ansatzes and exponential functions inspired by regularized optimal solutions in Supplementary Information. We stress that, unlike the method in Ref.<sup>38,48</sup>, in our case  $\theta$  and  $\beta$  are correlated. We impose  $\beta(t_f) = \pi/2$  and  $\beta(0) = \pi - \epsilon$  (note that we will allow a certain deviation by introducing the  $\epsilon$  parameter, see later). Once we construct  $\theta$ , the function  $\beta$  can be obtained by solving Eq. (7) with the boundary condition  $\beta(t_f) = \pi/2$ . After the functions  $\theta$  and  $\beta$  are obtained, the control field  $\Omega(t)$  is deduced using Eq. (6).

The solution to  $\beta$  from Eq. (7) depends on  $x_j$  leading to a set of  $\Omega \equiv \Omega(t, x_j)$ . However, in order to make the control independent of the input potential, we set  $\Omega(t) = \Omega(t, x_j = y)$  where the value of  $y$  is chosen such that it minimizes the  $C$  distance for different  $x_j$  in a certain interval (see next section).

**IE performance.** As the state evolves from  $|\Psi(0)\rangle = |+\rangle$ , the  $\kappa$  parameter should be a large number compared to the input potential  $x_j$ . We numerically study situations where  $\kappa = 2000$  and explored the range



**Figure 2.** (a) The functions of  $\theta$  (solid-blue) and  $\pi - \beta$  (dashed-red), where  $\theta$  is interpolated by a polynomial ansatz  $\theta = \sum_{i=0}^3 a_i t^i$ , and  $\beta$  is solved from Eq. (7) for  $t_f = 1$  and  $\kappa = 2000$ , with  $\epsilon = 2 \times 10^{-5}$ . (b) The control fields  $\Omega(t)$  designed from IE (solid-blue) with the help of  $\theta, \beta$  and from FAQUAD (dashed-red), where  $\kappa = 2000$ . We also show the control fields derived from  $\kappa = 1000$  (dotted-black) and  $\kappa = 500$  (dot-dashed-green). The inset in (b) displays the corresponding activation functions for different  $\kappa$ , which coincide with each other. In both plots,  $y/\Omega_f = 12$ .



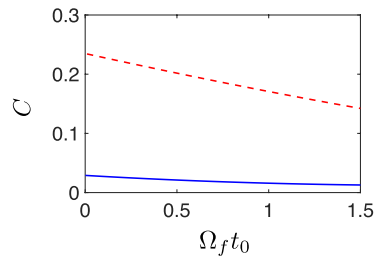
**Figure 3.** The dependence of  $C$  value on  $y$  with the application of IE  $\theta = \sum_{i=0}^3 a_i t^i$  (a) and FAQUAD (b) for different operation times  $t_f = 0.1$  (solid-blue),  $t_f = 0.2$  (dashed-red),  $t_f = 0.5$  (dotted-black), and  $t_f = 1$  (dot-dashed-green).

$|x_j|/\Omega_f \in [-x^{\max}, x^{\max}]$ , with  $x^{\max} = 12$ . Note that, we consider the situation where  $x^{\max} = 12$ , although our results are not limited to the specific number. We use dimensionless units, by setting the unit of time  $t_0$  such that the control field  $\Omega(t)$  is given in terms of  $1/t_0$ . In addition, we consider an unbiased perceptron with  $b_j = 0$ .

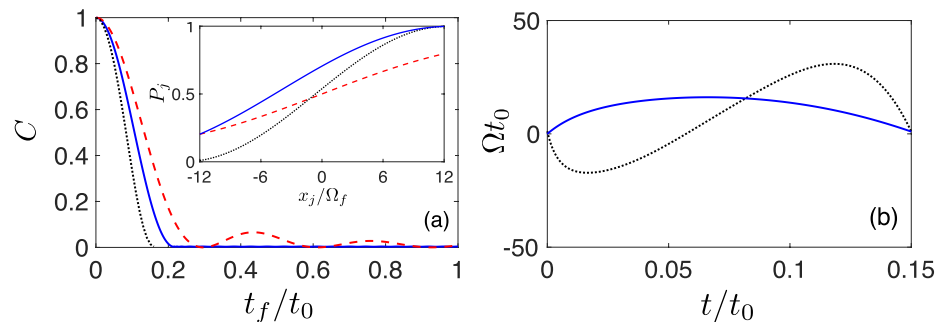
Not limited to a fixed large number of  $\kappa$ , our method shows the flexibility and the feasibility of the control field. For a case in which we impose  $\Omega_f = 1$  and solve Eq. (7) with a fixed value for  $x_j/\Omega_f = y/\Omega_f = 12$ , we find  $\theta(0) = 1.576 \simeq \pi/2$ . Figure 2a indicates the obtained solutions for  $\theta$  and  $\beta$  for this case in which we have also selected the operation time  $t_f = 1$ . We find that the boundary condition for  $\beta(0)$  is also satisfied with a tiny error of  $\epsilon = 2 \times 10^{-5}$ . In this specific case, we find that the designed control  $\Omega(t)$  at  $t = 0$  is  $\Omega(0) = 1999.6 \approx \kappa$  when  $\kappa = 2000$ , the initial state corresponds to the eigenstate state of the Hamiltonian. Also, we observed that  $\beta(0)$  tends to  $\pi$  when  $t_f$  gets larger. In Fig. 2b, the control field  $\Omega(t)$  obtained with our method is illustrated. This  $\Omega(t)$  leads to an excitation probability such that it arrives at  $P_j(x^{\max}) = 0.998$ . Using the same control field  $\Omega(t)$ , we find that the probability of the state  $|1\rangle$  for other input neural potentials  $x_j/\Omega_f \in [-x^{\max}, x^{\max}]$  is in the form of a sigmoid-like response ranging from 0 to 1 during the interval, as shown in the inset of Fig. 2b. This proves the successful construction of a sigmoid-shape transfer function, which is a crucial factor for a quantum perceptron. The fields calculated from  $\kappa = 1000$ ,  $\kappa = 500$  lead to the same sigmoid activation function which, as shown in the inset of Fig. 2b, cannot be distinguished to the one derived from  $\kappa = 2000$ .

Our IE method provides a wider range of  $y/\Omega_f$  than FAQUAD to construct sigmoid transfer functions. In Fig. 3a the value of the distance  $C$  obtained with the IE method, as a function of  $y/\Omega_f$  for various operation times  $t_f$ , is shown. It can be observed that a low value for  $C$  appears with large values for  $|y|$  and  $t_f$ . We have checked (also for  $t_f = 1$ ) the appearance of nonlinear perceptron responses that connect 0 and 1 with a sigmoid shape. In particular, these lead to  $C < 10^{-2}$  in the range  $y/\Omega_f \in [5, 12]$  with control fields  $\Omega(t)$  for  $t_f = 1$  similar to the one in Fig. 2b. In contrast,  $C$  goes to almost 2 at  $y/\Omega_f = -x^{\max}$  by FAQUAD techniques<sup>44</sup>, in which only for long  $t_f$  and in the regime  $y/\Omega_f \rightarrow x^{\max}$  the transfer function can be produced, see Fig. 3b.

The target state  $|\Psi(t_f)\rangle = |\Phi(x_j/\Omega_f)\rangle$  depends on the value of the driving field at the final time, see Eq. (3). In general we observe that, with our IE method, a larger value of the control field at  $t = t_f$  (i.e.  $\Omega_f$ ) offers higher fidelity. As an example of the latter, in Fig. 4 we show the value of  $C$  as a function of  $\Omega_f$  for  $t_f = 0.2$  with the



**Figure 4.** Dependence of  $C$  value on  $\Omega_f$  is shown for IE  $\theta = \sum_{i=0}^3 a_i t^i$  (solid-blue) and FAQUAD (red-dashed) protocols, when  $t_f = 0.2$ ,  $y/\Omega_f = 12$ .



**Figure 5.** (a) Dependence of  $C$  as a function of the final time  $t_f$ , using IE in the cases of  $\theta = \sum_{i=0}^3 a_i t^i$  (solid-blue),  $\theta = \sum_{i=0}^5 a_i t^i$  (dotted-black) and FAQUAD (dashed-red). The inset of (a) shows the corresponding transfer functions for  $t_f = 0.15$ , where the dotted-black curve represents the quasi-optimal-time solution with  $a_2 = -50$ ,  $a_3 = -3980$ . (b) For  $t_f = 0.15$ , the driving field  $\Omega(t)$  designed from IE in the cases of using  $\theta = \sum_{i=0}^3 a_i t^i$  (solid-blue), using  $\theta = \sum_{i=0}^5 a_i t^i$  with the optimal parameters  $a_2 = -50$ ,  $a_3 = -3980$  (dotted-black), and  $y/\Omega_f = 12$ .

application of IE (solid-blue) and FAQUAD (dashed-red). In this figure one can observe the improved performance of our IE method. Actually, every point of the lower value  $C$  by IE implies the successful discovery of sigmoid-shape transfer function and driving field  $\Omega(t)$ .

**Quasi-optimal-time solution.** As the activation function  $P(x_j/\Omega_f)$  connects 0 and 1 at  $-x^{\max}$  and  $x^{\max}$ , we set  $C < 0.01$  as the criteria of successful construction of a quantum perceptron. In Fig. 5a, we illustrate the dependence of  $C$  value on  $t_f$  by using the polynomial ansatz  $\theta = \sum_{i=0}^N a_i t^i$  with  $N = 3$  and  $N = 5$  of IE as well as FAQUAD<sup>34</sup>. When  $N = 3$ , the smallest  $t_f$ , such that  $C < 0.01$  is satisfied, is 0.2, while employing techniques based on FAQUAD, this is at  $t_f = 0.3$ . The further reduction of the smallest  $t_f$ , such that  $C < 0.01$  is satisfied, can be improved since IE method allows to approach the quasi-optimal-time solution by introducing more degrees of freedom in the ansatz of  $\theta$ <sup>47</sup>, leading to faster quantum perceptrons. With  $N = 5$  (i.e. a solution with two additional parameters, namely  $a_4$  and  $a_5$ ), see Fig. 5a (dotted-black curve) we get a speed up of 2 with respect to FAQUAD method, leading to the minimal operation time  $t_f^{\min} = 0.15$ . The values of the transfer function at  $-x^{\max}$  and  $x^{\max}$  and  $C$  value with the application of IE strategies in polynomial, trigonometric and exponential functions as well as FAQUAD can be seen in Supplementary Information, showing that high-order polynomial ansatz can give a quasi-optimal-time solution.

Moreover, we find that the IE method is robust with respect to timing errors, i.e. variations on the operation time  $t_f$ . More specifically, once the minimal value of  $C$  is reached for solid-blue in Fig. 5a,  $C$  does not show any appreciable oscillation for  $t > t_f^{\min}$ . Conversely, the FAQUAD driving leads to the dashed-red curve in Fig. 5a that shows an oscillatory behavior of  $C$ , indicating that only at some specific  $t_f$  the sigmoid transfer function can be constructed.

Remarkably, for short times, e.g.  $t_f = 0.15$ , the transfer functions and driving fields are completely different for IE and FAQUAD protocols. In the inset of Fig. 5a,b, we give the detailed demonstration of transfer functions and driving fields designed from IE. On the one hand, FAQUAD protocol cannot produce the sigmoid function, by connecting from 0 to 1 at the edges, see the inset of Fig. 5a dashed-red curve. On the other hand, we find that the case of IE with the polynomial ansatz of  $N = 3$  fails to connect the state  $|0\rangle$  presenting  $P(-x^{\max}) = 0.2$  (solid-blue curve). However, We can overcome this limitation by increasing the order of the polynomial ansatz to  $N = 5$ . Here, we compare the activation functions achieved by different strategies at the same value of  $y/\Omega_f = 12$ . It is worth mentioning that by increasing the value of  $x^{\max}$  which means more energy is supplied to the system, we can recover a more stretched sigmoid with the FAQUAD protocol or IE with the polynomial ansatz of  $N = 3$ .

However, in this work, we find the external driving  $\Omega(t)$  by which the perceptron can have a sigmoidal response in a fixed Hamiltonian configuration with the range  $[-x^{\max}, x^{\max}]$ .

In addition, the derived controls  $\Omega(t)$  from IE methods are smooth and present values close to zero at  $t = 0$ , see Fig. 5b. Compared to the case of  $t_f = 1$ , shorter operation time leads to larger  $\epsilon$  so that  $\Omega(0)$  is farther away from  $\kappa$ . This is in contrast with the control  $\Omega(t)$  derived from FAQUAD techniques that demands an abrupt change from  $\Omega(0) = 2000$  to  $\Omega(t_f) = 1$ , see Fig. 2b. This demonstrates the appropriateness of our IE derived controls to be implemented experimentally. In this regard, in the next section we give estimations based on state of the art experimental parameters in NV centers in diamond that demonstrates the suitability of an implementation of our method in such quantum platform.

## Discussions

We have demonstrated that the enhanced performance of our method using IE techniques leads to sigmoid activation functions within a minimal operation time of  $t_f^{\min} = 0.15 t_0$ . If, for instance, one selects  $t_0 = 500$  ns, the maximum value for the control  $\Omega(t)$  amounts to  $|\Omega_{\max}| \approx 50$  MHz for the kind of solutions presented in Fig. 5b (see horizontal axis limits in that figure). This permits the application of our controls in modern quantum platforms such as NV centers in diamond that present coherence times much longer than  $0.15 t_0 = 0.15 \times 500 = 75$  ns even at room temperature<sup>49,50</sup>. In addition, current arbitrary waveform generators allow to change the amplitude of the delivered microwave field (and consequently of the Rabi frequency  $\Omega$ ) in time-scales significantly smaller than 1 ns<sup>42,51</sup>. Then, one can easily introduce the controls in Fig. 5b to produce nonlinear sigmoid responses in NV centers. IE is also helpful to achieve the robust control in a specific physical setup<sup>52–54</sup> when one considers the Ising model with unwanted transitions between the target two-level system and other levels. In this manner one could envision a diamond chip with several NVs, each of them with available nearby nuclear spin qubits, as a quantum hardware to construct QNN using IE methods.

## Methods

**Inverse engineering and derivation of auxiliary differential equations.** The quantum perceptron gate evolves a qubit with the general Hamiltonian (Eq. (2)) which has the instantaneous ground state (Eq. (3)) with the basis  $|0\rangle = (0, 1)^T$  and  $|1\rangle = (1, 0)^T$  and a sigmoid excitation probability (Eq. (4)). Therefore, we need to control the final state exactly as  $|\Psi(t_f)\rangle = |\Phi(x_j/\Omega(t_f))\rangle$  in the form of Eq. (3). Inverse engineering by parameterizing the Bloch sphere angles  $\theta$  and  $\beta$  can manipulate the dynamical state evolution in a fast way. After substituting the wave function  $|\Psi(t)\rangle$  (Eq. (5)) or the orthogonal state  $|\Psi_{\perp}(t)\rangle$  into Schrödinger equation, we can obtain two equations

$$-i\dot{\theta} \sin \frac{\theta}{2} - \dot{\beta} \cos \frac{\theta}{2} = x_j \cos \frac{\theta}{2} + \Omega(t) \sin \frac{\theta}{2} e^{-i\beta}, \quad (10)$$

$$i\dot{\theta} \cos \frac{\theta}{2} + \dot{\beta} \sin \frac{\theta}{2} = -x_j \sin \frac{\theta}{2} + \Omega(t) \cos \frac{\theta}{2} e^{i\beta}. \quad (11)$$

Eq. (10)  $\times \sin(\theta/2)$  + Eq. (11)  $\times \cos(\theta/2)$  and Eq. (10)  $\times \sin(\theta/2)$  – Eq. (11)  $\times \cos(\theta/2)$ , respectively, result in the analytical expressions of  $\Omega(t)$  (Eq. (6)) and  $\beta$  (Eq. (7)). Once setting the operation time  $t_f$  and the dynamics of the polar angle  $\theta$ , we can obtain the function  $\beta$  by solving Eq. (7) with the boundary condition  $\beta(t_f) = \pi/2$ . Hence, from Eq. (6), we derive the applied field  $\Omega(t)$ .

**Fast quasi-adiabatic method.** Another protocol to construct a quantum perceptron by controlling the qubit gate is to use FAQUAD strategy<sup>34,44</sup>, which can achieve the fast and adiabatic-like procedure. The adiabatic parameter

$$\mu(t) = \hbar \left| \frac{\langle \phi_0(t) | \partial_t \phi_1(t) \rangle}{E_1(t) - E_0(t)} \right| \quad (12)$$

is kept as a constant  $\mu(t) = c$  during the whole control process, where the instantaneous eigenstates for the Hamiltonian (Eq. 2) are

$$|\phi_l\rangle = \cos(\alpha/2)|1\rangle + (-1)^l \sin(\alpha/2)|0\rangle \quad (13)$$

with the eigenenergies are  $E_l = -(-1)^l \hbar \sqrt{\Omega^2 + x_j^2}/2$ ,  $\alpha = \arccos[-x_j/\sqrt{\Omega^2 + x_j^2}]$  and  $l \in \{0, 1\}$ . In order to construct a universal quantum gate, a single control should not depend on the neuron potential  $x_j$ . The largest value  $|\mu|$  occurs at  $|x_j/\Omega_f| \approx 1.272$ . We take this  $\mu$  value as an optimal condition that works for all input neuron configurations. As the relation between the field and time is invertible, we can apply the chain rule to Eq. (12) and obtain

$$\frac{d\Omega}{dt} = -\frac{c}{\hbar} \left| \frac{E_1(\Omega) - E_0(\Omega)}{\langle \phi_0(\Omega) | \partial_{\Omega} \phi_1(\Omega) \rangle} \right|, \quad (14)$$

where the negative sign represents  $\Omega(t)$  monotonously decreases from  $\Omega(0)$  to  $\Omega(t_f)$ . The total duration time is rescaled as  $s = t/t_f$  so that  $\tilde{\Omega}(s) := \Omega(s t_f)$  and  $d\Omega/dt = t_f^{-1} d\tilde{\Omega}/ds$ . As a result, we have

$$\frac{d\tilde{\Omega}}{ds} = -\frac{\tilde{c}}{\hbar} \left| \frac{E_1 - E_0}{\langle \phi_0 | \partial_{\tilde{\Omega}} \phi_1 \rangle} \right|_{\tilde{\Omega}}, \quad (15)$$

$$\tilde{c} = ct_f = -\hbar \int_{\tilde{\Omega}(0)}^{\tilde{\Omega}(1)} \frac{d\tilde{\Omega}}{\left| \frac{E_1 - E_0}{\langle \phi_0 | \partial_{\tilde{\Omega}} \phi_1 \rangle} \right|_{\tilde{\Omega}}}. \quad (16)$$

A selection of  $t_f$  corresponds to different scaling of  $\tilde{c}$  and  $\Omega(t = st_f) = \tilde{\Omega}(s)$ . Consequently, we can derive  $\Omega(t)$  from  $\tilde{\Omega}(s)$  by solving the differential equation (Eq. (15)).

Received: 22 January 2021; Accepted: 26 February 2021

Published online: 11 March 2021

## References

- Arute, F. *et al.* Quantum supremacy using a programmable superconducting processor. *Nature* **574**, 505–510 (2019).
- Biamonte, J. *et al.* Quantum machine learning. *Nature* **549**, 195–202 (2017).
- Schuld, M., Sinayskiy, I. & Petruccione, F. An introduction to quantum machine learning. *Contemp. Phys.* **56**, 172–185 (2015).
- Minsky, M. & Papert, S. *A. Perceptrons: An Introduction to Computational Geometry Text* (MIT Press, Cambridge, 2017).
- McCulloch, W. S. & Pitts, W. A logical calculus of the ideas immanent in nervous activities. *Bull. Math. Biophys.* **5**, 115 (1949).
- Rosenblatt, F. Tech. Rep. Inc. Report No. 85-460-1 (Cornell Aeronautical Laboratory, 1957).
- Gyongyosi, L. & Imre, S. A survey on quantum computing technology. *Comput. Sci. Rev.* **31**, 51–71 (2019).
- Gyongyosi, L., Imre, S. & Nguyen, H. V. A survey on quantum channel capacities. *IEEE Commun. Surv. Tutor.* **20**, 2 (2018).
- Gyongyosi, L. Unsupervised quantum gate control for gate-model quantum computers. *Sci. Rep.* **10**, 10701 (2020).
- Gyongyosi, L. & Imre, S. Optimizing high-efficiency quantum memory with quantum machine learning for near-term quantum devices. *Sci. Rep.* **10**, 135 (2020).
- Gyongyosi, L. & Imre, S. Circuit depth reduction for gate-model quantum computers. *Sci. Rep.* **10**, 11229 (2020).
- Gyongyosi, L. Quantum state optimization and computational pathway evaluation for gate-model quantum computers. *Sci. Rep.* **10**, 4543 (2020).
- Gyongyosi, L. & Imre, S. Dense quantum measurement theory. *Sci. Rep.* **9**, 6755 (2019).
- Schuld, M., Sinayskiy, I. & Petruccione, F. The quest for a quantum neural network. *Quant. Inf. Process.* **13**, 25672586 (2014).
- Ciliberto, C. *et al.* Quantum machine learning: A classical perspective. *Proc. R. Soc. A* **474**, 20170551 (2018).
- Kak, S. On quantum neural computing. *Inform. Sci.* **83**, 143–160 (2015).
- Dunjko, V. & Briegel, H. J. Machine learning and artificial intelligence in the quantum domain: A review of recent progress Rep. *Prog. Phys.* **81**, 074001 (2018).
- Tacchino, F. *et al.* An artificial neuron implemented on an actual quantum processor. *NPJ Quantum Inf.* **5**, 26 (2019).
- Ouyang, X. *et al.* Experimental demonstration of quantum-enhanced machine learning in a nitrogen-vacancy-center system. *Phys. Rev. A* **101**, 012307 (2020).
- Killoran, N. *et al.* Continuous-variable quantum neural networks. *Phys. Rev. Res.* **1**, 033063 (2019).
- Rebentrost, P., Bromley, T. R., Weedbrook, C. & Lloyd, S. Quantum Hopfield neural network. *Phys. Rev. A* **98**, 042308 (2018).
- Zhao, J. *et al.* Building quantum neural networks based on a swap test. *Phys. Rev. A* **100**, 012334 (2019).
- Li, Y., Zhou, R., Xu, R., Luo, J. & Hu, W. Quantum deep convolutional neural network for image recognition. *Quantum Sci. Technol.* **5**, 044003 (2020).
- Leibfried, D., Blatt, R., Monroe, C. & Wineland, D. Quantum dynamics of single trapped ions. *Rev. Mod. Phys.* **75**, 281 (2003).
- Devoret, M. H. & Schoelkopf, R. J. Superconducting circuits for quantum information: An outlook. *Science* **339**, 1169–1174 (2013).
- Bloch, I. Ultracold quantum gases in optical lattices. *Nat. Phys.* **1**, 23–30 (2005).
- O'Brien, J. L., Furusawa, A. & Vučković, J. Photonic quantum technologies. *Nat. Photon.* **3**, 687 (2009).
- Zak, M. & Williams, C. P. Quantum neural nets. *Int. J. Theor. Phys.* **37**, 651–684 (1998).
- Behrman, E. C., Nash, L. R., Steck, J. E., Chandrashekar, V. G. & Skinner, S. R. Simulations of quantum neural networks. *Inf. Sci.* **128**, 257–269 (2000).
- Cao, Y., Guerreschi, G. G. & Aspuru-Guzik, A. Quantum Neuron: An elementary building block for machine learning on quantum computers. arXiv:1711.11240 (2017).
- Wan, K. H., Dahlsten, O., Kristjánsson, H., Gardner, R. & Kim, M. S. Quantum generalisation of feedforward neural networks. *NPJ Quant. Inf.* **3**, 36 (2017).
- Farhi, E. & Neven, H. Classification with Quantum Neural Networks on Near Term Processors. arXiv:1802.06002 (2018).
- Gyongyosi, L. & Imre, S. Training optimization for gate-model quantum neural networks. *Sci. Rep.* **9**, 12679 (2019).
- Torrontegui, E. & Garcia-Ripoll, J. J. Unitary quantum perceptron as efficient universal approximator. *Europhys. Lett.* **125**, 30004 (2019).
- Torrontegui, E. *et al.* Shortcuts to adiabaticity. *Adv. At. Mol. Opt. Phys.* **62**, 117 (2013).
- Guéry-Odelin, D. *et al.* Shortcuts to adiabaticity: Concepts, methods, and applications. *Rev. Mod. Phys.* **91**, 045001 (2019).
- Chen, X. *et al.* Fast optimal frictionless atom cooling in harmonic traps: Shortcut to adiabaticity. *Phys. Rev. Lett.* **104**, 063002 (2010).
- Chen, X., Torrontegui, E. & Muga, J. G. Lewis-Riesenfeld invariants and transitionless quantum driving. *Phys. Rev. A* **83**, 062116 (2011).
- Shen, C. P., Wu, J. L., Su, S. L. & Liang, E. Construction of robust Rydberg controlled-phase gates. *Opt. Lett.* **44**, 2036 (2019).
- Berry, M. V. Transitionless quantum driving. *J. Phys. A Math. Theor.* **42**, 365303 (2009).
- Wu, J. L. & Su, S. L. Universal speeded-up adiabatic geometric quantum computation in three-level systems via counterdiabatic driving. *J. Phys. A Math. Theor.* **52**, 335301 (2019).
- Zopes, J. *et al.* High-resolution quantum sensing with shaped control pulses. *Phys. Rev. Lett.* **119**, 260501 (2017).
- Cybenko, G. Approximation by superpositions of a sigmoidal function. *Math. Control Signals Syst.* **2**, 303–314 (1989).
- Martínez-Garaot, S., Ruschhaupt, A., Gillet, J., Busch, T. & Muga, J. G. Fast quasiadiabatic dynamics. *Phys. Rev. A* **92**, 043406 (2015).
- Hornik, K., Stinchcombe, M. & White, H. Multilayer feedforward networks are universal approximators. *Neural Netw.* **2**, 359–366 (1989).
- Biagiotti, L. & Melchiorri, C. *Trajectory Planning for Automatic Machines and Robots* (Springer, Berlin, 2008).
- Martikyan, V., Guéry-Odelin, D. & Sugny, D. Comparison between optimal control and shortcut to adiabaticity protocols in a linear control system. *Phys. Rev. A* **101**, 013423 (2020).

48. Takahashi, K. Shortcuts to adiabaticity for quantum annealing. *Phys. Rev. A* **95**, 012309 (2017).
49. Doherty, M. W. *et al.* The nitrogen-vacancy colour centre in diamond. *Phys. Rep.* **528**, 1–46 (2013).
50. Dobrovitski, V. V., Fuchs, G. D., Falk, A. L., Santori, C. & Awschalom, D. D. Quantum control over single spins in diamond. *Annu. Rev. Condens. Matter Phys.* **4**, 23 (2013).
51. Boris Naydenov, private communication.
52. Kiely, A. & Ruschhaupt, A. Inhibiting unwanted transitions in population transfer in two- and three-level quantum systems. *J. Phys. B At. Mol. Opt. Phys.* **47**, 115501 (2014).
53. Yu, X., Zhang, Q., Ban, Y. & Chen, X. Fast and robust control of two interacting spins *Phys. Rev. A* **97**, 062317 (2018).
54. Yan, Y. *et al.* Inverse engineering of shortcut pulses for high fidelity initialization on qubits closely spaced in frequency. *Opt. Express* **27**, 8267–8282 (2019).

## Acknowledgements

We acknowledge financial support from Spanish Government via PGC2018-095113-B-I00 (MCIU/AEI/FEDER, UE), Basque Government via IT986-16, as well as from QMiCS (820505) and OpenSuperQ (820363) of the EU Flagship on Quantum Technologies, and the EU FET Open Grant Quomorphic (828826). J. C. acknowledges the Ramón y Cajal program (RYC2018-025197-I) and the EUR2020-112117 Project of the Spanish MICINN, as well as support from the UPV/EHU through the Grant EHUrOPE. X. C. acknowledges NSFC (12075145), SMSTC (2019SHZDZX01-ZX04, 18010500400 and 18ZR1415500), the Program for Eastern Scholar and the Ramón y Cajal program of the Spanish MICINN (RYC-2017-22482). E. T. acknowledges support from Project PGC2018-094792-B-I00 (MCIU/AEI/FEDER,UE), CSIC Research Platform PTI-001, and CAM/FEDER Project No. S2018/TCS-4342 (QUITEMAD-CM).

## Author contributions

Y.B. developed the theoretical formalism, performed the analytic calculations and performed the numerical simulations. X.C. and E.T. verified the analytical method. J.C. supervised the project. All the authors contributed to the final version of the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-85208-3>.

**Correspondence** and requests for materials should be addressed to Y.B.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021