



Approximation of quantum control correction scheme using deep neural networks

M. Ostaszewski¹ · J. A. Miszczak¹ · L. Banchi² · P. Sadowski¹

Received: 16 May 2018 / Accepted: 4 March 2019 / Published online: 22 March 2019
© The Author(s) 2019

Abstract

We study the functional relationship between quantum control pulses in the idealized case and the pulses in the presence of an unwanted drift. We show that a class of artificial neural networks called LSTM is able to model this functional relationship with high efficiency, and hence the correction scheme required to counterbalance the effect of the drift. Our solution allows studying the mapping from quantum control pulses to system dynamics and analysing its behaviour with respect to the local variations in the control profile.

Keywords Quantum dynamics · Quantum control · Deep learning · Recurrent neural network

1 Introduction

The main objective and motivation of quantum information processing is the development of new technologies based on principles of quantum mechanics such as superposition and entanglement [1]. Quantum technologies require the development of methods and principles of quantum control, the control theory of the quantum mechanical system [2]. Such methods have to be developed by taking into account the behaviour of quantum systems [3,4]. In particular, as quantum systems are very susceptible to noise, which may influence the results of the computation, the methods of quantum control have to include the means for counteracting the decoherence [5].

The presented work is focused on the development of tools suitable for analysing the relation between the control pulses used for idealized quantum systems, and the control pulses required to execute the quantum computation in the presence of undesirable

✉ M. Ostaszewski
mostaszewski@iitis.pl

¹ Institute of Theoretical and Applied Informatics, Polish Academy of Sciences, Bałtycka 5
44-100, Gliwice, Poland

² QOLS, Blackett Laboratory, Imperial College London, London SW7 2AZ, UK

dynamics. Modelling this relation is important to better understand the manifold of control pulses in the presence of noise, a case which is still poorly understood. We focus on quantum dynamics described by a quantum spin chain. We are interested in a method of approximating the correction function of normal control pulses (NCP), i.e. the function accepting control pulses corresponding to the system without the drift Hamiltonian and generating the de-noising control pulses (DCP) for the system with the drift Hamiltonian. The existence of this function is non-trivial, since there are infinitely many pulses that produce the same evolution. We propose an approximation that not only incorporates the global features of the function but also describes its local properties. This feature is in contrast to the available methods based on optimization which do not take into account the continuous behaviour of the map from NCP to the de-noising control pulses. Indeed, without further assumptions, in view of the non-injectivity of quantum control, we may expect no relation at all between control pulses obtained from the optimization of different problems. On the other hand, we show that machine learning methods can be used for this purpose.

Recently, significant research effort has been invested in the application of machine learning methods in quantum information processing [6–8]. In particular, optimization techniques borrowed from machine learning have been used to optimize the dynamics of quantum systems [9], either for quantum control [10,11] and simulation [12], or for implementing quantum gates with suitable time-independent Hamiltonians [13]. These techniques include also quantum control techniques from dynamic optimization [14] and reinforcement learning [15,16]. In the presence of noise, neural networks give tools for optimizing dynamical decoupling, which can be seen as a quantum control correction scheme as considered by us, in a special case of the target operation being identity [17]. On the level of gate decomposition, neural networks have also been applied to the problem of decomposing arbitrary operations as a sequence of elementary gate sets [18,19].

In this paper, we propose a method, based on an artificial neural network (ANN), to study the correction scheme between control pulses obtained in the ideal case, and those obtained when the system is subject to undesired dynamics. Moreover, we demonstrate that the utilized network has high efficiency and can be used to analysis of the properties of the model.

This paper is organized as follows. In Sect. 2, we introduce the model and describe the architecture of a deep neural network which will be used as an approximation function. In Sect. 3, we demonstrate that the proposed methods can be used for generating control pulses without the explicit information about the model of a quantum system. We also utilize it for the purpose of analysing the properties of the correction scheme. Section 5 contains summary of the presented results.

2 Methods: model and solution

In this section, we provide necessary notation and background information. We start by introducing a spin chain model and describe the problem of generating quantum control pulses that counteract the undesired dynamics present in the system. We also

introduce the architecture of the artificial neural network used to approximate the correction scheme.

2.1 Model of quantum system

Let us consider a system of two interacting qubits. The evolution of the system is described by GKSL master equation

$$\frac{d\rho}{dt} = -i[H(t), \rho] + \sum_j \gamma_j \left(L_j \rho L_j^\dagger - \frac{1}{2} \{L_j^\dagger L_j, \rho\} \right). \tag{1}$$

where the Hamiltonian has three components

$$H(t) = H_c(t) + H_0 + H_\gamma. \tag{2}$$

We consider the control Hamiltonian of the form

$$H_c(t) = h_x(t)\sigma_x \otimes \mathbf{1} + h_z(t)\sigma_z \otimes \mathbf{1}, \tag{3}$$

and the base Hamiltonian

$$H_0 = \sigma_x \otimes \sigma_x + \sigma_y \otimes \sigma_y + \sigma_z \otimes \sigma_z. \tag{4}$$

The last element in Eq. (2) is the drift Hamiltonian, which can be an arbitrary two-qubit Hamiltonian multiplied by real parameter $\gamma > 0$. Incoherent part of Eq. 1 models interaction with an environment with strengths γ_j . It should be noted that in this paper, we will never consider the case when γ and γ_j both are different from zero.

Quantum optimal control refers to the task of executing a given unitary operation via the evolution of the system, in our case described by Eq. (2). To achieve this, one has to properly choose the coefficients $h(t) = (h_x(t), h_z(t))$ in Eq. (3). The set of reachable unitaries can be characterized [2] by studying the Lie algebra generated by the terms in Eq. (2). For $H_\gamma = 0$, our system is *fully controllable*, so any target can be obtained with suitable choice of $h(t)$, with no restriction on the control pulses. We assume that function $h(t)$ is piecewise constant in time slots $\Delta t_i = [t_i, t_{i+1}]$, which are the equal partition of evolution time interval $T = \bigcup_{i=0}^{n-1} \Delta t_i$. We also assume that $h_x(t)$ and $h_z(t)$ have values from interval $[-1, 1]$.

Function $h(t)$ will be represented as vectors of values of $h_x(t)$ and $h_z(t)$. For the case $\gamma = 0$, we say that it represents NCP—*normal control pulses*. Alternatively, for $\gamma \neq 0$, we say that $h(t)$ represents DCP—*de-noising control pulses*. Since $h(t)$ is piecewise constant, both NCP and DCP have two indices, with first index corresponding to time slots $\{0, \dots, n-1\}$, and the second index corresponding to the direction $\{x, z\}$, namely

$$\begin{aligned} NCP_{i,j} &= h_j(t), \quad \gamma = 0, \\ DCP_{i,j} &= h_j(t), \quad \gamma \neq 0, \end{aligned} \tag{5}$$

where

$$t \in [t_i, t_{i+1}] \quad i \in \{0, \dots, n-1\}, \quad j \in \{x, z\}. \quad (6)$$

It is worth noting that the mapping from the set of control operations to the unitary operator is not injective. Namely, the same unitary can be obtained using different choices of $h(t)$. To study the relationship between NCP and DCP, we need to select the DCP which is more closely related to the NCP. Because of continuity, we do that numerically by using the NCP as starting guess of the DCP. The final optimal DCP is then found using a local optimization around the initial NCP.

The figure of merit in our problem is the fidelity distance between superoperators, defined as [20]

$$F = 1 - F_{\text{err}}, \quad (7)$$

with

$$F_{\text{err}} = \frac{1}{2N^2} (\text{Tr}(Y - X(T))^\dagger (Y - X(T))), \quad (8)$$

where N is the dimension of the system in question, Y is superoperator of the fixed target operator, and $X(T)$ is evolution superoperator of operator resulting from the numerical integration of Eq. (1) with given controls. In particular, for a target unitary operator U , its superoperator Y is given by the formula

$$Y = U \otimes \bar{U}. \quad (9)$$

Superoperator $X(T)$ is obtained from the unitary operator resulting from the integration of Eq. (1).

2.2 Architecture of artificial neural network

The control pulses used to drive the quantum system with Hamiltonian from Eq. 3 are formally a time series. This aspect suggests that one may study their properties using methods from pattern recognition and machine learning [21,22] that have been successfully applied to process data with similar characteristics. The mapping from NCP to DCP shares similar mathematical properties with that of statistical machine translation [23], a problem which is successfully modelled with artificial neural networks (ANN) [24]. Because of this analogy, we use ANN as the approximation function to learn the correction scheme for control pulses. A trained artificial neural network will be used as a map from NCP to DCP

$$\text{ANN(NCP)} = \text{nnDCP}, \quad (10)$$

where nnDCP, neural network DCP, is an approximation of DCP obtained by using the neural network.

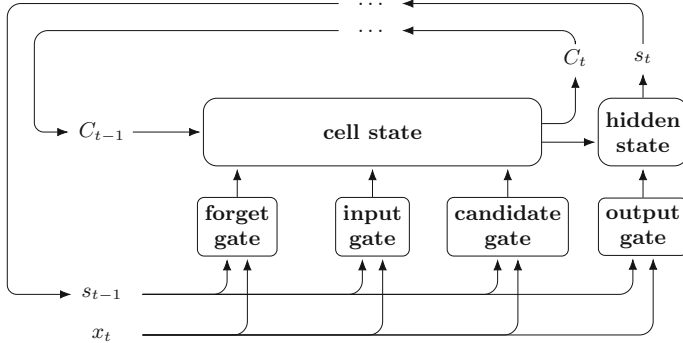


Fig. 1 Structure of LSTM network used in the experiments

Because of time series character of control sequences, we utilize bidirectional long short-term memory (LSTM) networks [25]. The long short-term memory block is a special kind of recurrent neural network (RNN), a type of neural network with directed cycles between units. These cycles allow the RNN to keep track of the inputs received during the former times. In other words, the output at given time depends not only on current input, but also on the history of earlier inputs. This kind of neural network is applicable in situations with time series with long-range correlations, such as in natural language processing where the next word depends on the previous sentence but also on the context (Fig. 1).

LSTM block consists of two cells

cell state: $c_t = c_{t-1} \circ f_t + \tilde{c}_t \circ i_t,$

hidden state: $s_t = \tanh(c_t) \circ o_t,$

which are constructed from following gates

input gate: $i_t = \text{sigma}(x_t V^i + s_{t-1} W^i),$

forget gate: $f_t = \text{sigma}(x_t V^f + s_{t-1} W^f),$

output gate $o_t = \text{sigma}(x_t V^o + s_{t-1} W^o),$

candidate: $\tilde{c}_t = \tanh(x_t V^c + s_{t-1} W^c),$

where \circ is element-wise multiplication of two vectors, s_{t-1} is previous hidden state, and x_t is an actual input state. Matrices V and W are the weight matrices of each gate. As one can see, there is only one gate with hyperbolic tangent as activation function—candidate gate. Rest of the gates have sigmoid activation function which has values from $[0, 1]$ interval. Using this function, neural networks decide which values are worth to keep and which should be forgotten. These gates maintain the memory of the network.

Basic architectures of RNN are not suitable for maintaining long-time dependences, due to the so-called *vanishing/exploding gradient problem*—the gradient of the cost function may either exponentially decay or explode as a function of the hidden units. Thanks to the structure of LSTM, the exploding gradient problem is reduced. The bidirectional version of LSTM is characterized by the fact that it analyses the input sequence/time series forwards and backwards. Thanks to this, it uses not only

information from the past but also from the future [26,27]. For this purpose, the bidirectional LSTM unit consists of two LSTM units. One of them takes as an input vector $[h(t_1), h(t_2), \dots, h(t_n)]$, and the second takes $[h(t_n), h(t_{n-1}), \dots, h(t_1)]$.

As in Eq. (10), the result of the network is a vector of control pulses nnDCP. To evaluate the quality of nnDCP, we apply loss function described in Eq. 8. To achieve this, we integrate a new superoperator $X(T)$ from nnDCP and measure how far it is from the target superoperator Y .

For two-qubit systems, we found that three stacked bidirectional LSTM layers are sufficient for obtaining the high value of fidelity. Moreover, at the end of the network we use one dense layer which processes the output of stacked LSTM to obtain our nnDCP. Experiments are performed using TENSORFLOW library [28,29].

3 Results: experiments

For the purpose of testing the proposed method, we use a sample of Haar random unitary matrices [30,31]. Pseudo-uniform random unitaries can be obtained also in the quantum control setting via random functions $h(t)$, provided that the control time T is long enough [32]. The exact implementation of sampling random Haar unitary matrices is available at [33]. In our experiments, we use QuTIP [34–36] to generate control pulses to training and testing data. First, we construct the target operators

$$U_{\text{target}} = U \otimes \mathbb{1}, \quad (11)$$

where $U \in \mathbf{U}(2)$ is a random matrix. Next, using QuTIP we generate NCP corresponding to the target operators. In the case of our set-up, the parameters are fixed as follows:

- time of evolution $T = 6$,
- number of intervals $n = 32$,
- control pulses in $[-1, 1]$.

We train the network using a subset of the generated pairs $\{(\text{NCP}, Y_{\text{target}})\}$, where Y_{target} is a target superoperator obtained from U_{target} . The training process network takes NCP as input and generates the nnDCP. Next, this nnDCP is an input to the loss function. For this purpose, we construct superoperator $X(T)$ resulting from integration of Eq. (1) using nnDCP. Next, we calculate error function between $X(T)$ and Y_{target} as in Eq. (8). In this section, we analyse only coherent drift, i.e. in Eq. (1) $\gamma_j = 0$.

Source code-implementing experiments described in this paper are available at [33].

3.1 Performance of the neural network

The first experiment is designed to analyse the efficiency of the trained network in terms of fidelity error of generated nnDCP control pulses. Trained LSTM has mean fidelity on the test set as presented in Table 1. It should be noted that, despite the fact that the mean fidelity on the test set is high, the trained network sometimes has outlier results, i.e. it returns nnDCP which corresponds to the operator with high fidelity error.

Table 1 The average of the mean fidelity scores for one qubit drifts

| H_γ | γ | | | |
|---|----------|-----|-----|-----|
| | 0.2 | 0.4 | 0.6 | 0.8 |
| $\gamma(\sigma_y \otimes \mathbf{1})$ | .96 | .92 | .95 | .96 |
| $\gamma(0.2\sigma_x \otimes \mathbf{1} + 0.8\sigma_y \otimes \mathbf{1})$ | .97 | .94 | .97 | .97 |
| $\gamma(0.5\sigma_x \otimes \mathbf{1} + 0.5\sigma_y \otimes \mathbf{1})$ | .98 | .96 | .98 | .97 |
| $\gamma(0.8\sigma_x \otimes \mathbf{1} + 0.2\sigma_y \otimes \mathbf{1})$ | .99 | .97 | .98 | .99 |

For each Hamiltonian and each parameter γ , 10 independent experiments have been performed. It should be noted that first two columns with results, i.e. columns corresponding to $\gamma = 0.2$ and 0.4 , are obtained from network which takes mini batches equal to 5 and 4000 training samples and 1000 testing samples. Next, two columns, i.e. $\gamma = 0.6$ and 0.8 , are obtained from network with mini batches 50 and 8000 training samples and 4000 testing samples

Table 2 The average of the mean fidelity scores for spin chain drifts

| H_γ | γ | |
|--|----------|-----|
| | 0.6 | 0.8 |
| $\gamma(\sigma_y \otimes \sigma_y)$ | .98 | .90 |
| $\gamma(0.1\sigma_x \otimes \sigma_x + 0.8\sigma_y \otimes \sigma_y + 0.1\sigma_z \otimes \sigma_z)$ | .98 | .98 |
| $\gamma(0.1\sigma_x \otimes \sigma_x + 0.6\sigma_y \otimes \sigma_y + 0.3\sigma_z \otimes \sigma_z)$ | .98 | .98 |
| $\gamma(0.2\sigma_x \otimes \sigma_x + 0.6\sigma_y \otimes \sigma_y + 0.2\sigma_z \otimes \sigma_z)$ | .96 | .97 |
| $\gamma(0.3\sigma_x \otimes \sigma_x + 0.6\sigma_y \otimes \sigma_y + 0.1\sigma_z \otimes \sigma_z)$ | .98 | .98 |

These results are obtained for network with mini batch equal to 10 and 8000 training samples and 4000 testing samples

The performed experiments show that it is possible to train LSTM networks for a given system with high efficiency. Results from Table 1 are obtained by trained artificial neural networks on different kinds of drifts, i.e. $\alpha\sigma_x \otimes \mathbf{1} + (1 - \alpha)\sigma_y \otimes \mathbf{1}$ for $\alpha \in \{0, 0.2, 0.5, 0.8\}$, with different values of γ . The experiment with this kind of drift is representative because σ_y is orthogonal to the controls σ_x, σ_z . The performed experiment shows that it is possible to train LSTM to have the efficiency on the test set above 90% for chosen gammas. Some of the results from Table 1 have average score lower than 90%. This is caused by the outlier cases, when network performs with very low efficiency. Such efficiency for chosen gammas is sufficient to our goal, which is to study the relations between the system and control pulses for relatively small disturbances.

In Table 2, we present average score for model with spin chain drift. As one can see, these results are much better than in Table 1. This can be explained by the fact that this type of drift is similar to the base Hamiltonian H_0 which is also a spin chain.

Results from above tables may suggest the answer for non-trivial questions: For what kind of drift, map between NCP and DCP is easier to learn? In particular, consider two cases: one qubit drift operating on the same qubit as the control Hamiltonian or a drift similar to the base Hamiltonian. In which case the mapping is easier to approximate with the LSTM network? Obtained results suggest that, when the drift is

asymmetric with respect to the base Hamiltonian, the neural network requires larger mini batch sizes to achieve similar efficiency.

As one can see, the efficiency of the artificial neural networks depends on the choice of the hyperparameters. In our case, for some values of parameter γ one needs to use bigger batch size and the larger training set to obtain satisfactory results. This behaviour is incomprehensible because, for $\gamma = 0.6$ and batch size equal to 5, RNN has problems with convergence. However, by increasing the batch size we were able to improve the performance. We would like to stress that our aim was to show that it is possible to obtain mean efficiency greater than 90%, not to examine what is the best possible efficiency of the network.

3.2 Utilization of the approximation

In this section, we show the results of an experiment, which allows checking what is the behaviour of the ANN() function when we perform local disturbances on a set of random NCP, and we check what is the deviation of the new nnDCP from the original DCP. The original DCP is obtained from GRAPE algorithm initialized by NCP.

Let us suppose that we have a trained LSTM. The procedure of checking its sensitivity on variations of $h_j(t)$, for $j \in \{x, z\}$ and t in the i th time slot, is as follows.

Step 1 Select a NCP vector from the testing set and generate the corresponding nnDCP.

Step 2 If the fidelity between target operator and operator resulting from nnDCP is lower than 90%, return to **Step 1**.

Step 3 Select $i \in \{0, \dots, n-1\}$, $j \in \{x, z\}$, change (i, j) coordinate of NCP by fixed ε

- if $\text{NCP}_{i,j} < 1 - \varepsilon$, then $\text{NCP}_{i,j} + \varepsilon$,
- if $\text{NCP}_{i,j} > 1 - \varepsilon$, then $\text{NCP}_{i,j} - \varepsilon$, and calculate new collection of nnDCP vectors with elements $\text{nnDCP}^{i,j}$, denoting that the element was obtained by perturbing j th component at the i th time slot.

Step 4 If the fidelity between target operator and operator resulting from $\text{nnDCP}^{i,j}$ is lower than 90%, return to **Step 3**.

Step 5 Calculate norm l_2 (Euclidean norm) of difference between nnDCP and $\text{nnDCP}^{i,j}$.

Because of outlier results of the network, there are additional conditions on generated DCP in the above algorithm. Applying the above algorithm for each NCP from the testing set, we can obtain a sample of variations. Next, we can analyse empirical distributions of these variations.

As an example of using the above method, we consider the following example. Let us suppose that the drift operator is of the form

$$H_\gamma = \gamma \sigma_y \otimes \mathbf{1}. \quad (12)$$

Then, for $\varepsilon = 0.1$ and $\gamma = 0.2$, the exemplary variation histograms are presented in Fig. 2. Thanks to trained network, we are able to analyse the impact of small changes in the input on the output (**Step 5**).

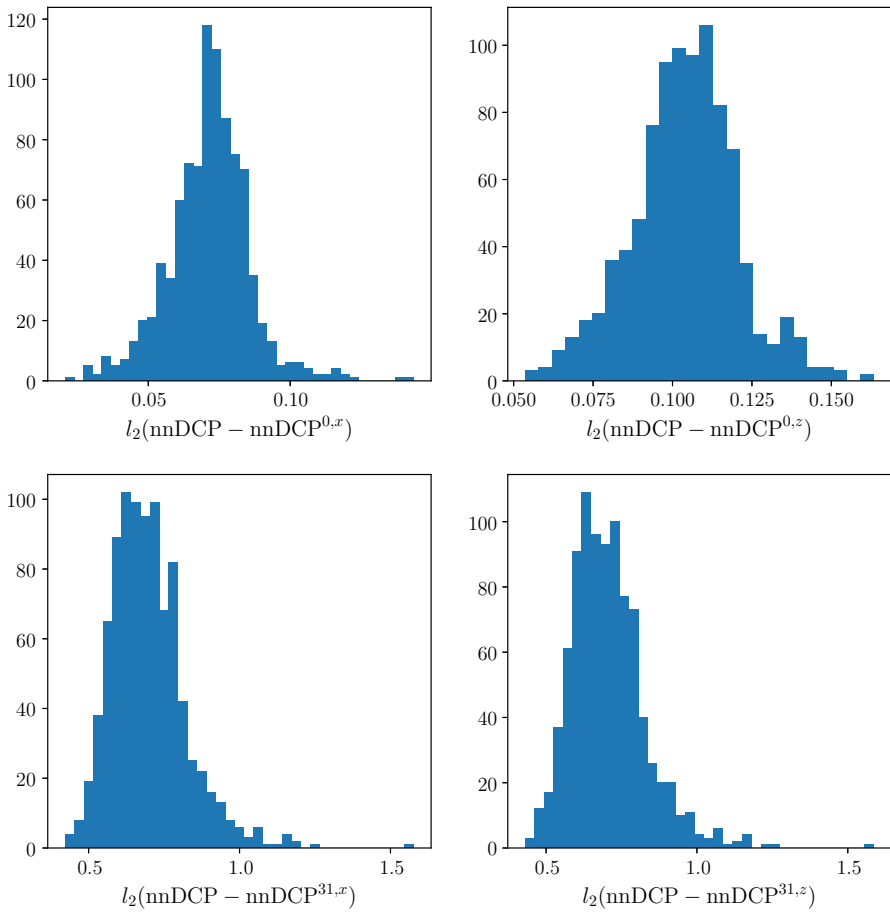


Fig. 2 Exemplary histograms performing variation of ANN approximation. Histograms are generated from disturbances on set of 1000 NCP corresponding to random matrices, for fixed system $H_\gamma = \gamma\sigma_y \otimes \mathbf{1}$.

One can consider the median of the distribution of variations as the quantitative measure of the influence of the changes in the input signal on the resulting DCP. To check whether medians of distributions of changes are similar, we perform Kruskal–Wallis statistical test for each pair of the changed coordinates (see Fig. 3a). Results presented in Fig. 3a show that most of the distributions received are statistically different regarding the median, i.e. most of the p values is less than 0.05. The values on the main diagonals, where we compared disturbances on the same controls in the same time slots, are equal to 1. This observation confirms that the test behaves appropriately in this case. On the other hand, one can observe that for time slots $10 \leq i \leq 30$ Kruskal–Wallis test for distributions obtained for $NCP_{i+1,x} \pm \varepsilon$ and $NCP_{i,z} \pm \varepsilon$ gives p values greater than 0.05. Thus, the disturbances introduced in this time slots on h_x and h_z coordinates of NCP results in similar variations of the resulting DCP. The situation is different for time slots $i \leq 10$, where one can see that the variation in DCP signal depends on which coordinate of the NCP signal is disturbed.

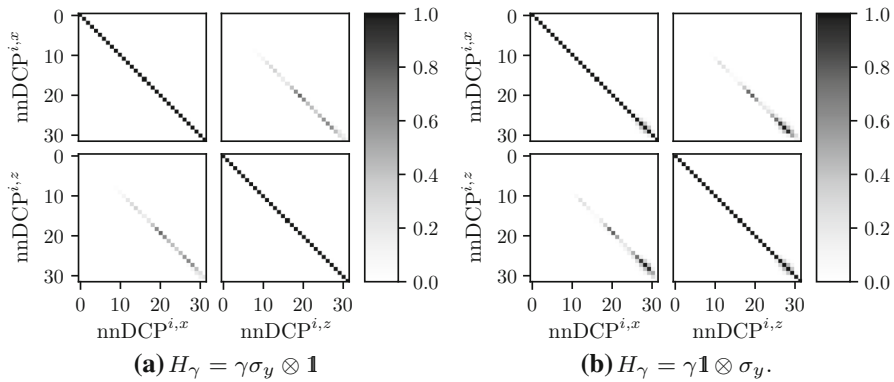


Fig. 3 Plots of p values of Kruskal–Wallis test for tested drift Hamiltonians. Each element (l, k) of above matrix plots represents a p value of the test between empirical distributions of changes implied by disturbances on l th time slot and k th time slot. The horizontal axis of the left column and the vertical axis of top row plots correspond to disturbances on σ_x control, while horizontal axis of the right column and the vertical axis of the bottom plots correspond to disturbances on σ_x control

The similar effect can be observed if the drift Hamiltonian acts on the second qubit only. In this case, we can consider experiment where the drift Hamiltonian is of the form

$$H_\gamma = \gamma \mathbb{1} \otimes \sigma_y, \tag{13}$$

with $\varepsilon = 0.1$ and $\gamma = 0.2$.

The results of Kruskal–Wallis test for this situation are presented in Fig. 3b. One can see that our approach suggests that there are similarities in distributions of variations implied by disturbances on different controls and near time slots. This suggests that local disturbances in control signals have a similar effect in the case of drift on the target system and drift on the auxiliary system.

Moreover, one can see that the constructed approximation exhibits symmetry of the model. This effect can be observed by analysing the disturbances of h_x and h_z controls in the same time slot. From the performed experiments, one can see they give similar variations quantified by the distribution of DCP changes.

One should note that on plots in Fig. 3, where we compare variations of $\text{nnDCP}^{i,x}$ and $\text{nnDCP}^{i,z}$, p values greater than 0.05 are focused along the diagonal. This symmetry is not perfect, but one should note that training data are generated from random unitary matrices. Moreover, we do not impose any restrictions on the training set to ensure the uniqueness of the correction scheme.

4 Discussions

In a case where Hamiltonian drift is absent $\gamma \neq 0$ and we want to find proper DCP for system with incoherent noise, things get more complicated. In our experiments, we tested three kinds of Lindblad noise, namely

1. with one Lindblad operator $L = |0\rangle\langle 1| \otimes \mathbb{1}$,
2. with two Lindblad operators $L_1 = \sigma_z \otimes \mathbb{1}$ and $L_2 = \mathbb{1} \otimes \sigma_z$,
3. and mixed two Lindblad operators $L_1 = \sigma_z \otimes \mathbb{1}$ and $L_2 = |0\rangle\langle 1| \otimes \mathbb{1}$.

For all these open system noise models, nnDCP obtains efficiency above 90% for $\gamma_1 = \gamma_2 \approx 0.01$. However, this strength of noise is so small that also NCP obtains similar efficiency. Therefore, the proposed model of the neural network does not correct NCP effectively, for incoherent types of noise.

Another important case in the context of quantum control is the robustness on the random fluctuations. In our experiments, we tried to generate nnDCP which will be robust for Gaussian fluctuations, i.e.

$$\forall i \in \{0, \dots, n-1\}, j \in \{x, z\} \text{ nnDCP}_{i,j} + \delta_{i,j}, \quad (14)$$

where $\delta_{i,j} \sim \mathcal{N}(0, \sigma)$. The standard deviation we tested in two cases $\sigma = 0.1$ and 0.2 , while the model of Hamiltonian drift was $0.4(\sigma_y \otimes \mathbb{1})$.

During the training process, returned by ANN control pulses were copied ten times and to each copy we added Gaussian fluctuations. Such disturbed nnDCP were applied to cost function Eq. (8), and next the average fidelity was calculated. Unfortunately, after this training process artificial neural network does not produce nnDCP which are more robust for random fluctuations. This failure might be caused by the fact that training of artificial neural network is based on gradient descent. As we know from [16], gradient-based algorithm does not give very good results. On the other hand, task with which we try to face is slightly different from the one in [16], because we want to generalize correction scheme over all target unitaries.

5 Concluding remarks

The primary objective of the presented work is to use artificial neural networks for the purpose of approximating the structure of quantum systems. We propose to use a bidirectional LSTM neural network. We argue that this type of artificial neural network is suitable to capture time dependences present in quantum control pulses. We have developed a method of reconstructing the relation between control pulses in an idealized case and control pulses required to implement quantum computation in the presence of undesirable dynamics. We argue that the proposed method can be a useful tool to study the manifold of quantum control pulses in the noisy regime, and to define new theoretical approaches to control noisy quantum systems.

Acknowledgements LB acknowledges support from the UK EPSRC Grant EP/K034480/1. MO acknowledges support from Polish National Science Center scholarship 2018/28/T/ST6/00429. JAM acknowledges support from Polish National Science Center Grant 2014/15/B/ST6/05204. Authors would like to thank Daniel Burgarth for discussions about quantum control, Bartosz Grabowski and Wojciech Masarczyk for discussions concerning the details of LSTM architecture, and Izabela Miszczak for reviewing the manuscript. Numerical calculations were possible thanks to the support of PL-Grid Infrastructure.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution,

and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Dowling, J., Milburn, G.: Quantum technology: the second quantum revolution. *Phil. Trans. R. Soc. A* **361**, 1655 (2003)
2. d'Alessandro, D.: Introduction to Quantum Control and Dynamics. CRC Press, Boca Raton (2007)
3. Gough, J.E., Belavkin, V.P.: Quantum control and information processing. *Quantum Inf. Process.* **12**, 1397 (2013)
4. Pawela, Ł., Puchała, Z.: Quantum control with spectral constraints. *Quantum Inf. Process.* **13**, 227 (2014)
5. Viola, L., Lloyd, S.: Dynamical suppression of decoherence in two-state quantum systems. *Phys. Rev. A* **58**, 2733 (1998)
6. Ciliberto, C., Herbster, M., Ialongo, A.D., Pontil, M., Rocchetto, A., Severini, S., Wossnig, L.: Quantum machine learning: a classical perspective. In: *Proc. R. Soc. A*, Vol. 474. The Royal Society, p. 20170551 (2018)
7. Dunjko, V., Briegel, H.J.: Machine learning & artificial intelligence in the quantum domain: a review of recent progress *Rep. Prog. Phys.* **81**, 074001 (2018). <https://doi.org/10.1088/1361-6633/aab406>
8. Ostaszewski, M., Miszczak, J., Sadowski, P.: Geometrical versus time-series representation of data in learning quantum control, [arXiv:1803.05169](https://arxiv.org/abs/1803.05169)
9. van Nieuwenburg, E., Bairey, E., Refael, G.: Learning phase transitions from dynamics. *Phys. Rev. B* **98**, 060301 (2018)
10. Zahedinejad, E., Schirmer, S., Sanders, B.: Evolutionary algorithms for hard quantum control. *Phys. Rev. A* **90**, 032310 (2014)
11. August, M., Hernández-Lobato, J. M.: Taking gradients through experiments: LSTMs and memory proximal policy optimization for black-box quantum control. *arXiv preprint arXiv:1802.04063* (2018)
12. Las Heras, U., Alvarez-Rodriguez, U., Solano, E., Sanz, M.: Genetic algorithms for digital quantum simulations. *Phys. Rev. Lett.* **116**, 230504 (2016)
13. Banchi, L., Pancotti, N., Bose, S.: Quantum gate learning in qubit networks: Toffoli gate without time-dependent control. *NPJ Quantum Inf.* **2**, 16019 (2016). <https://doi.org/10.1038/nnpjqi.2016.19>
14. Sridharan, S., Gu, M., James, M.: Gate complexity using dynamic programming. *Phys. Rev. A* **78**, 052327 (2008)
15. Bukov, M., Day, A., Sels, D., Weinberg, P., Polkovnikov, A., Mehta, P.: Machine learning meets quantum state preparation. *The Phase Diagram of Quantum Control* (2017). [arXiv:1705.00565](https://arxiv.org/abs/1705.00565)
16. Niu, M. Y., Boixo, S., Smelyanskiy, V., Neven, H.: Universal quantum control through deep reinforcement learning. *arXiv preprint arXiv:1803.01857* (2018)
17. August, M., Ni, X.: Using recurrent neural networks to optimize dynamical decoupling for quantum memory. *Phys. Rev. A* **95**, 012335 (2017)
18. Swaddle, M., Noakes, L., Smallbone, H., Salter, L., Wang, J.: Generating three-qubit quantum circuits with neural networks. *Phys. Lett. A* **381**, 3391 (2017)
19. Fösel, T., Tighineanu, P., Weiss, T., Marquardt, F.: Reinforcement learning with neural networks for quantum feedback. *arXiv preprint arXiv:1802.05267* (2018)
20. Floether, F., de Fouquieres, P., Schirmer, S.: Robust quantum gates for open systems via optimal control: Markovian versus non-Markovian dynamics. *New J. Phys.* **14**(7), 073023 (2012)
21. Bishop, C.: *Neural Networks for Pattern Recognition*. Oxford University Press, Oxford (1995)
22. Goodfellow, I., Bengio, Y., Courville, A., Bengio, Y.: *Deep Learning*, vol. 1. MIT Press, Cambridge (2016)
23. Koehn, P.: *Statistical Machine Translation*. Cambridge University Press, Cambridge (2009)
24. Bahdanau, D., Cho, K., Bengio, Y.: Neural machine translation by jointly learning to align and translate (2014). [arXiv:1409.0473](https://arxiv.org/abs/1409.0473)
25. Hochreiter, S., Schmidhuber, J.: Long short-term memory. *Neural Comput.* **9**, 1735 (1997)
26. Schuster, M., Paliwal, K.K.: Bidirectional recurrent neural networks. *IEEE Trans. Signal Process.* **45**(11), 2673–2681 (1997)

27. Graves, A., Fernández, S., Schmidhuber, J.: Bidirectional LSTM networks for improved phoneme classification and recognition. In: International Conference on Artificial Neural Networks organization. Springer, Berlin, pp. 799–804 (2005)
28. Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., Wicke, M., Yu, Y., Zheng, X.: In: 12th USENIX Symposium on Operating Systems Design and Implementation, Vol. 16, pp. 265–283 (2016)
29. TensorFlow: An open-source machine learning framework for everyone. <https://www.tensorflow.org/>
30. Mezzadri, F.: How to generate random matrices from the classical compact groups. *Not. AMS* **54**, 592 (2007)
31. Miszczak, J.: Generating and using truly random quantum states in Mathematica. *Comput. Phys. Commun.* **183**, 118 (2012)
32. Banchi, L., Burgarth, D., Kastoryano, M.J.: Driven quantum dynamics: will it blend? *Phys. Rev. X* **7**, 041015 (2017)
33. Approximation of quantum control using lstm. https://github.com/ZKSI/qcontrol_lstm_approx
34. QuTiP—Quantum Toolbox in Python. <http://qutip.org/> (2012)
35. Johansson, J., Nation, P., Nori, F.: QuTiP: an open-source Python framework for the dynamics of open quantum systems. *Comput. Phys. Commun.* **183**, 1760 (2012)
36. Johansson, J., Nation, P., Nori, F.: QuTiP 2: a Python framework for the dynamics of open quantum systems. *Comput. Phys. Commun.* **184**, 1234 (2013)

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