Tunable Realizations of Correlated Quantum Walks using an Unsupervised Generative Model

Miri Kenig\(^{1}\) and Yoav Lahini\(^{1}\)

\(^{1}\)Department of Condensed Matter Physics, School of Physics and Astronomy, Tel-Aviv University, Tel Aviv 6997801, Israel

Quantum particles co-propagating on a lattice develop complex correlations due to an interplay between quantum statistics, interactions, and lattice disorder. Here we present a novel algorithm capable of learning these correlations and identifying the physical control parameters in a completely unsupervised manner. After training on a limited data set, the algorithm can generate a much larger number of new, unbiased, adjustable, yet physically correct instances. The knowledge encapsulated in the algorithm’s latent space allows tuning physical parameters to values the algorithm was not explicitly trained on and accelerate the learning of new, more complex problems. Our results demonstrate the ability of neural networks to learn quantum dynamics in an unsupervised manner and offer a route to their use in quantum simulations and computation.

The single-particle Quantum Walk, the quantum mechanical analog of the classical random walk, describes the evolution of an initially localized quantum particle on a lattice potential \([11–13]\). As a quantum particle can be in superposition, its evolution is governed by the interference of all the paths it can take on the lattice. As a result, quantum walkers exhibit several differences compared to their classical counterparts. On periodic potentials they propagate faster: ballistically rather than diffusively \([6]\). On disordered lattices, the quantum walker may come to a complete halt – a phenomenon with no classical counterpart known as Anderson Localization \([7–10]\).

Despite the name, single-particle quantum walks are essentially wave interference phenomena, and indeed they can, and have been, experimentally realized using classical waves \([10–13]\). Quantum effects emerge when several indistinguishable quantum particles propagate on the lattice simultaneously \([14–20]\). Here, the collective many-body evolution is governed by the interference of all possible multi-particle processes. As a result of quantum statistics, the propagating particles develop non-classical correlations, or dependencies, between their positions as they propagate across the lattice \([14–16, 21]\) (see Figure 1). While this is true even when the particles do not interact with each other, interactions introduce additional complexity, reflected in the growth of the Hilbert space required to account for all possible multi-particle processes and their quantum interferences \([17–20]\).

The ability to shape these high dimensional correlations by engineering the lattice potential, the interactions or other parameters generated broad interest in applying multi-particle quantum walks for various computational and information processing tasks. These include quantum computation \([18, 22]\), preparing and manipulating quantum states \([23, 24]\), devising search algorithm \([25, 26]\) implementing perfect state transfer \([27]\), exploring topological phases \([28, 31]\), evaluating Boolean formulas \([4, 32]\) and more. This interest is further motivated by the recent experimental realizations of pristine quantum-walk systems using photons \([11–14, 16, 33, 34]\), trapped ions \([35, 36]\), superconducting qubits \([37, 39]\), and ultracold atoms \([19, 20, 30, 31]\), which allow exquisite control of the number and initial positions of the particles, of their interactions and of the lattice potential. Still, the computational complexity of the multi-particle quantum walk poses a fundamental obstacle when attempting to inverse-design many-body operations \([22]\) or simulate them using classical computers, as even small quantum systems require significant computation time \([21]\).

Recently, it has been shown that deep learning algorithms can, in some cases, efficiently represent stationary quantum many-body states \([23, 42, 44]\) or learn hard
quantum distributions [45]. Here we explore the use of
deep learning for simulating correlated quantum dynamics,
 focusing on correlated few-body quantum walks. To
this end we employ a recent breakthrough in unsupervised
machine learning: generative adversarial networks (GANs) [46]. This exciting deep learning method pro-
cesses a dataset of examples in an unsupervised man-
er using two neural networks that compete against each
other to generate new, synthetic samples that can pass for
real data. Since invented by Goodfellow in 2014, GANs
demonstrated an outstanding capability to learn complex
images and then generate synthetic images that are hard
to distinguish from real ones. If GANs can learn in a
completely unsupervised manner the subtle many-pixel
correlations representing a generalized cat, and use that
information to generate synthetic yet authentic images of
non-existent cats – can they learn the physics of cor-
related quantum dynamics from examples only, and then
generate synthetic yet physically valid data samples?

Training the GAN algorithm

We start by generating the dataset used to train the
GAN algorithm. We consider two indistinguishable boson
particles, propagating on one-dimensional lattices
with ten sites, disordered on-site energies and uniform
nearest-neighbor hopping. The two-body Bose-Hubbard
Hamiltonian can be constructed from the lattice para-

ters:

\[ H = \sum_m E_m a_m^\dagger a_m + \sum_{\langle n,m \rangle} J_{n,m} a_n^\dagger a_m + \frac{\Gamma}{2} \sum_n (\hat{n}_m - 1) \]

where \( E_m \) is the on-site energy of site \( m \), \( a_m^\dagger \) \( a_m \) are the creation/annihilation operator for a particle at site \( m \),
and \( J_{n,m} \leq 0 \) is the hopping rate between neighboring
lattice sites. \( \hat{n}_m = a_m^\dagger a_m \) is the corresponding number
operator, and \( \Gamma \) is the on-site interaction energy between
the particles. In all the following cases, the parameters
\( J_{n,m} = -1 \) and \( E_m \) are randomly chosen from a flat
distribution \( 0 < E_m < 3 \). The disorder level is chosen
such that the localization length is of the order of the
lattice size, and therefore the propagating particles do not
remain tightly localized at their starting position [9].

We first consider a specific initial condition in which
the two particles are placed at two adjacent sites at the
center of the lattice \( |\psi_{initial}\rangle = a_4^\dagger a_5^\dagger |0\rangle \). To generate the
training samples, we numerically propagate the initial
state according to the Hamiltonian for a time \( t = 2 \), so
that the propagating particles do not reach and reflect
from the edges of the lattice. Next, we calculate the two-
particle correlation \( \langle q,r \rangle = \langle a_q^\dagger a_q^\dagger a_r a_r \rangle \) which represents the probability of finding exactly one particle at site \( q \)
and one particle at site \( r \) at the end of the evolution,
taking into account the interference of all possible two
particle processes, the quantum statistics, the interaction
between the particles and the effect of the disordered
lattice potential. Such propagations were used previously
by some of us for the study of correlated two-particle
quantum walks on periodic lattices, in both simulations
and experiments using photons [13, 16] and ultra-cold
atoms [17, 20]. It was found that the choice of initial
state and the level of interactions change the correlation
map significantly. Here we study the two-body and three-
body quantum walk on disordered lattices.

The GAN algorithm processes images, which here are
prepared in the following way: for each random real-
ization of the lattice parameters, we generate a single
training image, as shown in Fig 1B. Each image has two
parts. The left panel of the image contains a column
of ten pixels, representing the specific random realiza-
tion of the on-site energies \( E_m \), where the color code
represents the value of each entry in 8-bit RGB. The
right side of the training image represents the calculated
two-particle correlation at the end of the propagation,
as explained above. We generate three separate data sets
for three levels of interactions: \( \Gamma = 0 \) for the interact-
ion free case, \( \Gamma = 3 \) for intermediate interactions
and \( \Gamma = 1000 \) for strong interactions – corresponding to hard-
core bosons [17, 20]. Each training set contains 70,000
training images, each with a different realization of disor-
der. The performance during learning is evaluated using
the Frechet Inception Distance (FID) score [17], which is
often used to evaluate GAN images’ quality.

Generating new correlated quantum walks

After training is complete, the trained networks are
used to generate new, synthetic, high-resolution samples,
which look similar to the training images. To test if
the synthetic images represent new and valid physical re-
results, we extract from each generated image the values of
the on-site energies \( E_m \), calculate the expected two-body
correlation after propagation using exact diagonalization,
and compared the resulting correlation matrix to the gen-
erated one. The similarity between the generated and the
exact correlations is evaluated using the Kullback–Leibler
derivative \( KL = \sum P(x_1, x_2) \log \left( \frac{P(x_1, x_2)}{Q(x_1, x_2)} \right) \) were
the reference distribution \( P(x_1, x_2) \) is the exact normalized
correlation function and the tested distribution \( Q(x_1, x_2) \)
is the normalized generated distribution. Figure 2 shows
the quality of generated samples as the learning pro-
gresses. Initially the GAN generates blurred images
with corresponding high \( KL \) score, yet after learning the trained network generates synthetic data of ran-
don-on-site energies and two-particle correlation with
\( KL < 0.025 \pm 0.015 \). In comparison, random normal-
ized test distributions yield \( KL > 0.5 \pm 0.09 \). This result
means that after training, the GAN network can produce,
FIG. 2. Learning correlated two-particle quantum walk on disordered lattices. (A) Improvement in generating physically correct correlations as learning progresses, compared to exact calculation (last column) for: (i) $\Gamma = 0$, no interactions; (ii) $\Gamma = 3$, intermediate interactions; (iii) $\Gamma = 1000$, strong interactions. Results are for $|\psi_{\text{initial}}\rangle = a_5^\dagger a_6^\dagger |0\rangle$, averaged over 1000 realizations of disorder. (B) The physical validity of the generated correlation functions as a function of the learning length, quantified using the Kullback-Leibler divergence. Blue overlay represents standard deviation across disorder realizations.

without further calculations, new pairs of random lattice parameters and resulting two-particle correlations with the correct physics. The number of possible generated realizations is only limited by the random seed size, $2^{32}$ in our case.

Latent space truncation tunes disorder level

Next, we show that even though our algorithm was trained on a fixed level of disorder, the algorithm can be tuned, post-training, to generate on demand synthetic samples with different levels of disorder. To this end we use the truncation trick [48] used in the StyleGAN algorithm [49, 50]. In this procedure the intermediate latent space of the generator is truncated, usually to deal with areas that are poorly represented in the training data. The training set used here has a fixed level of disorder, i.e. the values of the lattice on-site energies $E_m$ are randomly chosen from the interval $0 < E_m < \eta$ with $\eta = 3$. We find that by controlling the truncation threshold $\delta$, the generator can be made to produce synthetic samples with a controlled level of disorder, i.e. with $\eta$ between 0 and 3. This is demonstrated in figure 3A, which displays the histogram of generated $E_m$ values over 1000 generated samples (total of 10,000 values), as a function of the truncation parameter $\delta$. As $\delta$ is varied between 0 and 1, the spread of the histogram grows, reaching full spread (similar to the training set) at $\delta = 1$. Thus, as $\delta$ is increased the level of disorder in the generated samples increases as well. As a result, the correlation matrices and particle densities exhibit a crossover from free propagation to Anderson localization. Notably, this procedure does not degrade the fidelity of the generated correlation matrices and density distribution, i.e. the results remain physically valid with Kullback-Leibler divergence below 0.03 in all cases.

Generating unitary transformations

So far, we have restricted ourselves to unsupervised learning of two-particle correlation functions and to a particular initial state – the two particles are initially on adjacent sites at the center of the lattice. Taking this technique a step forward, we consider now the full two-particle problem, i.e. considering all possible initial states and calculating the resulting complex-valued two-particle wave function. To this end, we construct the two-body Hamiltonian in the occupation number basis. The size of this real and symmetric matrix, $55 \times 55$, reflects the size of the Hilbert space, or the number of ways in which two particles can be arranged on ten sites. The unitary propagator, $U = e^{-iHt}$ is a complex valued and symmetric matrix of similar size, whose columns represent the resulting two-particle wavefunction, in the number basis, for each possible initial arrangement. Here, the training images contain three parts as shown in Figure 4. The left panel represents the ten value of the random on-site energies $E_m$ as before. The right panel of the
Latent-space truncation trick tunes the level of disorder in the generated samples. (A) histograms of generated on-site energies from 1000 samples as a function of the truncation parameter $\delta$. While the original network was trained on disorder level corresponding to the rightmost histogram, truncating the latent space produces different levels of disorder. (B) Disorder-averaged correlation functions (top) and densities (bottom) for two particles initially on the same site and zero interaction. (C) Disorder-averaged correlation functions (top) and densities (bottom) for two particles initially on adjacent site an $\Gamma = 3$. All results are physically valid with Kullback-Leibler divergence below 0.03.

As before, we generate three separate training sets for three levels of interactions $\Gamma = 0, 3, 1000$, each containing 70,000 training images with different disorder realizations. The GAN network is then trained on each dataset, while the performance during learning is evaluated using the Frechet Inception Distance score. The training here takes longer, as expected, but by the end the trained network can generate synthetic novel samples. The physical validity of the synthetic samples was tested by extracting the generated on-site energies $E_m$, calculating the two-body Hamiltonian, exponentiating to obtain the expected unitary transform $U_e$, and comparing it to the generated unitary $U_g$. The fidelity was calculated using the standard form, $Fidelity = \frac{\text{trace}(U_g^\dagger U_e)}{N}$. Figure 4 shows the quality of generated samples as the learning progresses. At the end of training the network generates synthetic data of random on-site energies and unitary evolution with fidelity surpassing 92%.

In addition we tested the performance of the GAN algorithm in learning the correlated quantum walk of three particles on a 6-site random lattice. These results are presented in Figure 4, the fidelity here is 83%.

Training using a pre-trained network

Next, we show that the training time can be significantly shortened by taking as the starting point a network that completed training on different regime of parameters. Figure 4B shows the training dynamics of
FIG. 4. Learning full unitary transformations of correlated quantum walks on disordered lattices. (A) A training image, including the lattice parameters (left) and magnitude and phase of the resulting unitary. (B) Fidelity of the generated samples during learning, for a GAN network starting from scratch (circles) and for a pre-trained network (squares) - see text. (C) Improvement in generating physically correct unitaries as learning progresses, compared to exact calculation for (i) two particles and $\Gamma = 3$; (ii) two particles and $\Gamma = 1000$; (iii) three particles, $\Gamma = 1000$. Images show the magnitude of the generated unitary transformations, results are averaged over 1000 realization of disorder.

FIG. 5. Latent-space truncation used to tune disorder level in the generated two-particle unitaries. (A) histograms of generated on-site energies from 1000 samples as a function of the truncation parameter $\delta$. (B) Disorder-averaged two particles unitaries (magnitude). Fidelity in all cases is above 85%.

Discussion

In this work we have shown that a machine learning algorithm based on GAN can learn the physical rules of correlated few-body quantum dynamics on disordered lattices in a completely unsupervised manner. The trained network can produce, without further calculation, new and physically valid samples containing both lattice parameters and the resulting correlation functions or full high dimensional unitary evolution.

Remarkably, the GAN algorithm automatically identifies the relevant physical control parameters (disorder in our case). This information is encapsulated in the algorithm’s latent space. Truncating it results in tuning of the disorder level in the generated correlated quantum walks, even though the network was not trained on these ranges explicitly. This feature holds great potential for computation-free exploration of physical problems. In the application of GANs to images, (e.g. faces), this procedure enabled continuous tuning of different semantic features, such as age, gender, hair length etc. Here, these axes correspond to real physical parameters. This suggests that carefully choosing the training set will allow separate tuning of different physical parameters, such as level of interaction, initial condition, noise or even number of particles. Such trained networks could become a valuable tool in the fast exploration of complex physical problems.

While here we focused on quantum dynamics problems...
and simulated training datasets, our algorithm could be used on a range of other physical problems and experimentally obtained data. This approach could become a powerful tool in helping researchers to identify the relevant control parameters in complex data sets. It could also be used to interpolate sparse data sets correctly and quickly with generated data.

An exciting direction that requires further work is to enable query of trained networks with partial data (say the desired unitary transformation) and have it immediately output the conjugate information (the lattice parameters in this case). Such modules could solve both forward and inverse problems much faster than other exact or approximate calculations. This could be useful, for example, in augmenting quantum hardware with fast control modules.

Acknowledgments

This research was supported by the Israel Science Foundation (ISF) and the Directorate for Defense Research and Development (DDR&D) grant No. 3427/21 and by the Israel Ministry of Science and Technology Grant No. 3-15671.

[1] Y. Aharonov, L. Davidovich, and N. Zagury, Physical Review A 48, 1687 (1993).
[2] S. E. Venegas-Andraca, Quantum Information Processing 11, 1015 (2012).
[3] J. Kempe, Contemporary Physics 44, 307 (2003).
[4] E. Farhi and S. Gutmann, Physical Review A 58, 915 (1998).
[5] O. Mülken and A. Blumen, Physics Reports 502, 37 (2011).
[6] M. Varbanov, H. Krovi, and T. A. Brun, Physical Review A 78, 022324 (2008).
[7] P. W. Anderson, Phys. Rev. 109, 1492 (1958).
[8] T. Schwartz, G. Barthal, S. Fishman, and M. Segev, Nature 446, 52 (2007).
[9] Y. Lahini, A. Avidan, F. Pozzi, M. Sorel, R. Morandotti, D. N. Christodoulides, and Y. Silberberg, Physical Review Letters 100, 013906 (2008).
[10] N. C. Harris, G. R. Steinbrecher, M. Prabhu, Y. Lahini, J. Mower, D. Bunandar, C. Chen, F. N. Wong, T. Baehr-Jones, M. Hochberg, et al., Nature Photonics 11, 447 (2017).
[11] B. Do, N. L. Stohler, S. Balasubramanian, D. S. Elliott, C. Eash, E. Fischbach, M. A. Fischbach, A. Mills, and B. Zwickl, JOSA B 22, 499 (2005).
[12] H. B. Perets, Y. Lahini, F. Pozzi, M. Sorel, R. Morandotti, and Y. Silberberg, Physical review letters 100, 170506 (2008).
[13] M. A. Broome, A. Fedrizzi, B. P. Lanyon, I. Kassal, A. Aspuru-Guzik, and A. G. White, Physical review letters 104, 153602 (2010).
[14] Y. Bromberg, Y. Lahini, R. Morandotti, and Y. Silberberg, Physical review letters 102, 253904 (2009).
[15] Y. Lahini, Y. Bromberg, D. N. Christodoulides, and Y. Silberberg, Phys. Rev. Lett. 105, 163905 (2010).
[16] A. Peruzzo, M. Lobino, J. C. Matthews, N. Matsuda, A. Politi, K. Poulios, X.-Q. Zhou, Y. Lahini, N. Ismail, K. Wörhoff, et al., Science 329, 1500 (2010).
[17] Y. Lahini, M. Verbin, S. D. Huber, Y. Bromberg, R. Pugatch, and Y. Silberberg, Phys. Rev. A 86, 011603 (2012).
[18] A. M. Childs, D. Gosset, and Z. Webb, Science 339, 791 (2013).
[19] C. Weitenberg, M. Endres, J. F. Sherson, M. Cheneau, P. Schauß, T. Fukuhara, I. Bloch, and S. Kuhr, Nature 471, 319 (2011).
[20] P. M. Preiss, R. Ma, M. E. Tai, A. Lukin, M. Rispoli, P. Zupancic, Y. Lahini, R. Islam, and M. Greiner, Science 347, 1229 (2015).
[21] S. Aaronson and A. Arkhipov, in Proceedings of the forty-third annual ACM symposium on Theory of computing (2011) pp. 333–342.
[22] Y. Lahini, G. R. Steinbrecher, A. D. Bookatz, and D. Englund, npj Quantum Information 4, 1 (2018).
[23] X. Gao and L.-M. Duan, Nature communications 8, 1 (2017).
[24] Y. Huang and J. E. Moore, arXiv preprint arXiv:1701.06246 (2017).
[25] N. Shenvi, J. Kempe, and K. B. Whaley, Physical Review A 67, 052307 (2003).
[26] A. M. Childs and J. Goldstone, Physical Review A 70, 022314 (2004).
[27] A. Kay, Physical Review A 73, 032306 (2006).
[28] T. Kitagawa, M. S. Rudner, E. Berg, and E. Demler, Physical Review A 82, 033429 (2010).
[29] S. Panahiyan and S. Fritzschke, Physical Review A 100, 062115 (2019).
[30] Y. E. Kraus, Y. Lahini, Z. Ringel, M. Verbin, and O. Zilberberg, Physical review letters 109, 106402 (2012).
[31] C. Chen, X. Ding, J. Qin, Y. He, Y.-H. Luo, M.-C. Chen, C. Liu, X.-L. Wang, W.-J. Zhang, H. Li, L.-X. You, Z. Wang, D.-W. Wang, B. C. Sanders, C.-Y. Lu, and J.-W. Pan, Phys. Rev. Lett. 121, 100502 (2018).
[32] A. Ambainis, A. M. Childs, B. W. Reichardt, R. Špalek, and S. Zhang, SIAM Journal on Computing 39, 2513 (2010).
[33] P. P. Rohde, A. Schreiber, M. Štefaňák, I. Jex, and C. Silberhorn, New Journal of Physics 13, 013001 (2011).
[34] A. Aspuru-Guzik and P. Walthier, Nature physics 8, 285 (2012).
[35] F. Zähringer, G. Kirchmair, R. Gerritsma, E. Solano, R. Blatt, and C. Roos, Physical review letters 105, 100503 (2010).
[36] H. Schmitz, R. Matjeschk, C. Schneider, J. Glueckert, M. Enderlein, T. Huber, and T. Schaetz, Physical review letters 103, 090504 (2009).
[37] E. Flurin, V. V. Ramasesh, S. Hacohen-Gourgy, L. S. Martin, N. Y. Yao, and I. Siddiqi, Physical Review X 7, 031023 (2017).
[38] Z. Yan, Y.-R. Zhang, M. Gong, Y. Wu, Y. Zheng, S. Li, C. Wang, F. Liang, J. Lin, Y. Xu et al., Science 364, 753 (2019).
[39] M. Gong, S. Wang, C. Zha, M.-C. Chen, H.-L. Huang, Y. Wu, Q. Zhu, Y. Zhao, S. Li, S. Guo et al., Science (2021).
[40] M. Karski, L. Förster, J.-M. Choi, A. Steffen, W. Alt, D. Meschede, and A. Widera, Science 325, 174 (2009).
[41] T. Fukuhara, P. Schauß, M. Endres, S. Hild, M. Cheneau, I. Bloch, and C. Gross, Nature 502, 76 (2013).
[42] Y. Levine, O. Sharir, N. Cohen, and A. Shashua, Physical review letters 122, 065301 (2019).
[43] G. Carleo and M. Troyer, Science 355, 602 (2017).
[44] F. Noé, S. Olsson, J. Köhler, and H. Wu, Science 365, eaaw1147 (2019).
[45] A. Rocchetto, E. Grant, S. Strelchuk, G. Carleo, and S. Severini, npj Quantum Information 4, 1 (2018).
[46] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, in Advances in neural information processing systems (2014) pp. 2672–2680.
[47] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter, arXiv preprint arXiv:1706.08500 (2017).
[48] A. Brock, J. Donahue, and K. Simonyan, arXiv preprint arXiv:1809.11096 (2018).
[49] T. Karras, S. Laine, and T. Aila, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2019) pp. 4401–4410.
[50] T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen, and T. Aila, arXiv preprint arXiv:1912.04958 (2019).