Neural network approach to nonequilibrium phase transitions in open quantum systems

Minjae Jo,1 K. Choi,1 and B. Kahng1

1CCSS, CTP and Department of Physics and Astronomy, Seoul National University, Seoul 08826, Korea

Recent advances in the neural network (NN) machine learning technique provide a remarkable ability to identify the phases of matter. Here, we use this NN approach and identify the transition point of the absorbing transition arising in the quantum contact process in one dimension (1D-QCP), which is potentially realized in cold Rydberg atom systems. The training datasets for supervised learning consist of snapshots taken far from the transition point in quantum jump Monte Carlo (QJMC) simulations. Using the convolutional NN approach and finite-size scaling analysis, we identify the transition point, denoted as \( \omega_c \), of the 1D-QCP. Moreover, by performing extensive QJMC simulations at \( \omega_c \), we successfully determine all the critical exponents of the 1D-QCP and conclude that the 1D-QCP belongs to the directed percolation class. By contrast, the mean-field solution of the QCP shows a different type of phase transition, tricritical phase transition.

Introduction. Quantum critical behavior in driven dissipative many-body systems has been a major topic of research in condensed matter physics [1–16]. Although many fascinating experimental results for ultracold atomic gases have been obtained as experimental tools have improved [17], quantum critical behavior has not been clearly established yet owing to the limitations of numerical results arising from the exponential complexity of the Hilbert space. To compensate for this difficulty, various approximations have been developed, including tensor networks and the density matrix renormalization group [18–19], which have been applied with success to systems with short-range quantum correlations. However, for systems with a highly entangled state near the quantum critical point, these approximation methods may not be appropriate for determining the critical exponents [20,22]. In a critical region near the transition point, the quantum jump Monte Carlo (QJMC) method, also called the Monte Carlo wave function method [23–29], would be more useful than the tensor network approach for determining the critical behavior [30,31]. However, it is still challenging to identify a critical point precisely using the QJMC approach owing to the limited system size. In fact, the parameter space in quantum systems increases exponentially as the system size is increased. Here, we show that the use of the neural network (NN) approach compensates for this drawback.

The NN approach has recently served as a powerful tool [32–33] for classifying the phases in classical systems [34], which exhibit patterns generated by many elements. Each element has one of two values, for instance, the up and down spin states in ferromagnetic systems. By contrast, each element of a quantum system has a real value, and thus the patterns are much more complex. Nevertheless, the NN approach has reportedly been successfully used to determine the phase transition points of closed quantum systems on the basis of simulation data [35–38] and experimental images [39,40]. The NN approach has recently been applied to find the configurations of open quantum systems in steady state using the restricted Boltzmann machine [41–45]. However, it is challenging to investigate the critical behaviors using unsupervised learning techniques such as the restricted Boltzmann machine.

Here, we consider the one dimensional quantum contact process (1D-QCP) [46–52] by applying the NN approach with supervised learning to snapshots generated by QJMC simulations of small sizes. First, labeled snapshots of the 1D-QCP generated by QJMC simulations of a given finite system are taken far from a transition point in both directions, and then organized in datasets; the well-optimized NNs then respond sensitively to the transition point. This supervised learning method can correctly identify the position of the transition point. Second, using the obtained critical points for given system sizes, we perform finite-size scaling (FSS) analysis and identify the transition point in the thermodynamic limit. Next, we determine the critical exponents of the 1D-QCP by performing extensive QJMC simulations for a large system at the transition point. We find that the critical behavior of the 1D-QCP belongs to the directed percolation (DP) class. To the best of our knowledge, this is the first work in which the NN approach is applied to a nonequilibrium phase transition in an open quantum system. Without using the NN approach, we find that the critical behavior appears in wide ranges of the control parameters. Thus the transition point is not clearly identified.

The QCP is a quantum generalization of the classical contact process belonging to the DP universality class [53–59]. Each site is a qubit system that can be in the active or inactive state. A QCP consists of three incoherent and two coherent processes: i) decay, in which an active site is incoherently inactivated spontaneously at a rate \( \gamma \); ii) incoherent branching or coagulation, in which an active particle incoherently activates or inactivates an inactive particle at the nearest-neighbor site at a rate \( \kappa \), respectively; and iii) coherent branching or coagulation, which is a quantum counterpart of process ii) driven by a Hamiltonian at a rate \( \omega \), respectively. The coherent processes will be described in detail later. The classical ii) and quantum iii) rules are in competition, which may change the transition behavior. Previous studies using the mean-field and semiclassical approaches revealed that the competition results in a tricritical point [46–48]. However, a recent study using the tensor network approach in one dimension suggests that the tricritical point is shifted onto the quantum axis and a continuous transition line remains [50]. Thus, on the quantum axis (\( \kappa = 0 \)), the 1D-QCP model exhibits a continuous absorbing phase transition, and a question arises as which class
The Hamiltonian and dissipative terms \[60\]:

\[ \hat{\rho} = -i \left[ \hat{H}_S, \hat{\rho} \right] + \sum_{\ell=1}^{N} \left[ \hat{L}^{(d)}_{\ell} \hat{\rho} \hat{L}^{(d)\dagger}_{\ell} - \frac{1}{2} \left( \hat{L}^{(d)}_{\ell} \hat{L}^{(d)\dagger}_{\ell} \right) \hat{\rho} \right]. \tag{1} \]

The Hamiltonian \( \hat{H}_S \), which governs the branching and coagulation processes and represents coherent interactions, is expressed as

\[ \hat{H}_S = \omega \sum_{\ell=1}^{N} \left[ (\hat{n}_{\ell-1} + \hat{n}_{\ell+1}) \sigma^z_{\ell} \right]. \tag{2} \]

The Lindblad operators of decay, branching, and coagulation are given by

\[ \begin{align*}
L^{(d)}_{\ell} &= \sqrt{\kappa} \sigma^-_{\ell}, \\
L^{(b)}_{\ell} &= \sqrt{\kappa} \sigma^+_{\ell}, \\
L^{(c)}_{\ell} &= \sqrt{\kappa} (\hat{n}_{\ell-1} + \hat{n}_{\ell+1}) \sigma^-_{\ell},
\end{align*} \tag{3} \]

respectively. \( \sigma^+ \) and \( \sigma^- \) are the raising and lowering operators of the spin at site \( \ell \), respectively; they are defined in terms of the spin basis as \( \sigma^+ = |\uparrow\rangle \langle \downarrow| \) and \( \sigma^- = |\downarrow\rangle \langle \uparrow| \). In addition, \( \hat{n} = \sigma^+ \sigma^- \) and \( \sigma^x = \sigma^+ + \sigma^- \) are the number operator and spin flip operator, respectively.

Quantum branching and coagulation processes occur at a rate \( \omega \), and the corresponding classical processes occur at a rate \( \kappa \). When \( \omega \to 0 \), the model is reduced to the classical contact process, which belongs to the DP class. Hereafter, we consider the quantum limit of the model in the limit \( \kappa \to 0 \). In addition, we rescale time and the quantum control parameter \( \omega \) in units of \( \gamma \); therefore, we set \( \gamma = 1 \).

When \( \omega \) is small, inactive particles become more abundant with time, and eventually the system is fully occupied by inactive particles. Then, the system is no longer dynamic and falls into an absorbing state, which is represented by \( \hat{\rho}_{ab} = |\downarrow\rangle \langle \downarrow| \). When \( \omega \) is large, the system remains in an active state with a finite density of active particles [Fig. 1(a)]. Thus, the QCP exhibits a phase transition from an active to an absorbing state as the control parameter \( \omega \) is decreased. In Fig. 1(b), the phase transition seems to be continuous. In fact, it was conjectured that the 1D-QCP exhibits a continuous transition \[49\]. The critical point and spatial correlation exponent are obtained numerically as \( \nu = 6.0 \pm 0.05 \) and \( \nu_{\perp} = 0.5 \pm 0.2 \), respectively, for the tensor network \[50\].

**NN approach.** For classical systems, the transition point of a continuous absorbing transition is normally indicated by the presence of power-law behavior of the order parameter with respect to time \[53, 58, 61\]. Consequently, a large system size is required to identify the transition point. Moreover, the critical exponents are sensitive depending on the transition point. Accurately identifying the transition point using QJMC simulations of the QCP is even more difficult and is thus a challenging problem.

To overcome this difficulty, we notice that the system is in the absorbing state for \( \omega \ll \omega_c \) and in the active state for \( \omega \gg \omega_c \). Combining this observation with a recently proposed NN supervised learning concept, we identify the transition point as follows.

To implement the NN approach, we first organize a dataset of the occupation probability of site \( \ell \), which is denoted as \( p_\ell(t) = \text{Tr} [\hat{\rho}(t) \hat{n}_\ell] \). Using the QJMC method, we generate a steady-state configuration and obtain the occupation probabilities of each site, \( \{ p_\ell \} \). We collect 5000 configurations in

![FIG. 1. (a) Trajectory of the 1D-QCP with \( \kappa = 0 \) and \( \omega > \omega_c \) from a single active site at the center. Branching and coagulation processes are coherently driven by a Hamiltonian. (b) Histogram of the densities of active sites in steady states as a function of \( \omega \) for system size \( N = 20 \). The data are obtained using QJMC simulations.](image)
the training sets used in the convolutional NN (CNN) analysis. (b) Plot of \( \omega \) versus \( N \), where \( \omega_c \) is chosen so as to yield power-law behavior, which is typical near the transition point \( \omega_c \). The slope represents the value of the critical exponent \( -1/\nu_\perp \). (c) Scaling plot of the output versus \( (\omega - \omega_c)N^{1/\nu_\perp} \). For the obtained numerical values of \( \nu_\perp \) and \( \omega_c \), the data collapse well for system sizes \( N = 10, 12, 14, 16, \) and 18. From (b) and (c), we obtain \( \omega_c \approx 6.04 \) and \( \nu_\perp = 1.06 \pm 0.04 \).

Next, to train the machine, we construct the hidden layers of the NN, including one-dimensional convolutional layers \([63]\), batch normalization layers \([64]\), and fully connected layers, as shown in Fig. 3. We employ the framework of TensorFlow \([65]\) and use ReLU and tanh for the activation function in the hidden layer. Two neurons in the output layer are used, and a softmax function is used as the activation function in the output layer. We employ the cross-entropy or the mean-square error function as the cost (error) function of the NN, which is then optimized using Adam \([66]\) or RMSProp \([67]\). We change the architecture and optimization algorithms in various ways. Regardless of these changes, the well-trained machines produce consistent results. Once the NN is well-trained with the labeled training dataset in the two regions \( \omega \in [0, 4] \) and \( \omega \in [8, 12] \), we obtain the outputs for the entire \( \omega \) region.

Finite-size scaling. In Fig. 2(a), we plot the output averaged over \( 5 \times 10^3 \) configurations for system sizes \( N = 10, 12, 14, 16, \) and 18. The two outputs provide the probabilities that the system will fall into the absorbing state and remain in the active state, respectively. The crossing point of the two outputs indicates a critical point \( \omega_c(N) \) for a given system size \( N \) [Fig. 2(a)]. Using the obtained \( \omega_c(N) \) for different system sizes, we determine \( \omega_c \) in the thermodynamic limit by plotting \( \omega_c - \omega_c(N) \) versus \( N \) [Fig. 2(b)], which is expected to behave as \( \omega_c - \omega_c(N) \sim N^{-1/\nu_\perp} \). Indeed, the plot exhibits power-law decay behavior when an appropriate value of \( \omega_c \) is chosen, and the critical exponent \( \nu_\perp \) is obtained as the slope of the power-law behavior. We obtain \( \omega_c \approx 6.04 \) and \( \nu_\perp = 1.06 \pm 0.04 \); the latter is in agreement with \( \nu_\perp \approx 1.096 \) for the DP class in one dimension. Finally, an FSS plot is drawn in the form of the output versus \( (\omega - \omega_c)N^{1/\nu_\perp} \) for different \( \nu \) values [Fig. 2(c)].

Next, we measure the values of the other critical exponents using the numerical data obtained by the QJMC method in the critical region around \( \omega_c \). To do this, we first consider a QCP starting from a single active seed at \( \ell = 0 \), where the remaining sites are inactive. This configuration is expressed as \( \hat{\rho}(0) = \hat{\sigma}_0^a \hat{\rho}_{ab} \hat{\sigma}_0^b \). We measure the following quantities to characterize the criticality of the QCP: i) the survival probability, i.e., the probability that the system does not fall into an absorbing state, \( P(t) = 1 - \text{Tr}[\hat{\rho}(t)\hat{\rho}_s] \); ii) the number of active sites, \( N_a(t) = \sum_i \text{Tr}[\hat{\rho}(t)\hat{h}_i] \); and iii) the mean square distance of the active sites from the origin, \( R^2(t) = \sum_i \text{Tr}[\hat{\rho}(t)\hat{h}_i^2]/N_a(t) \). In addition, the initial configuration may be taken as that in which the entire system is occupied by active sites, which is represented by \( \hat{\rho}(0) = |\uparrow \cdots \uparrow \rangle \langle \uparrow \cdots \uparrow | \). From this initial configuration, we measure iv) the density \( n(t) \) of active sites at
time $t$ averaged over all runs. This quantity is formulated as $n(t) = (\sum_t \text{Tr}[\hat{\rho}(t)\hat{n}_r])/N$, $v$ the density $n_{\text{sur}}(t)$ of active sites at time $t$ averaged over the surviving runs may be measured.

At the critical point, the above physical quantities exhibit the following power-law behaviors:

$$P(t) \propto t^{-\delta'}, \quad N_d(t) \propto t^\eta, \quad R^2(t) \propto t^{2z}, \quad \text{and} \quad n_{\text{sur}}(t) \propto t^{-\delta'}.$$  

Thus, the density behaves as $n(t) \sim n_{\text{sur}}(t)P(t) \propto t^{-\delta'-\delta'}$.

We characterize the critical behavior in finite systems using the FSS theory [59]. The critical exponents are determined using the data collapse technique for scaling functions. For instance, at the critical point, the average density $n(t)$ of active sites behaves as $n(t) = t^{\delta'-\delta} f_{\delta}(tN^{-\zeta})$. Similarly, the other quantities are reduced as

$$P(t) = t^{-\delta'} f_p(tN^{-\zeta}), \quad N_d(t) = t^\eta f_\eta(tN^{-\zeta}),$$  

where $f_{\delta}$, $f_p$, and $f_\eta$ are scaling functions.

At the critical point, we perform the FSS analysis to obtain the critical exponents. First, we obtain the exponents $\delta$ and $\delta'$, $\eta$, $\delta$, and $z$ by measuring the slopes in the double-logarithmic plots shown in Fig. 4. Then, we collapse the data by using the obtained exponents to compute the dynamic exponent $z$. Specifically, we plot $n(t)^{\delta+\delta'}$ versus $tN^{-\zeta}$ in Fig. 4(a), $N_d(t)^{-\eta}$ versus $tN^{-\zeta}$ in Fig. 4(b), and $P(t)/t^{\delta'}$ versus $tN^{-\zeta}$ in Fig. 4(c) for different system sizes $N$. We measure the exponent $z$ directly using the plot of $R^2(t)$ versus $t$ in Fig. 4(d).

The critical exponents are obtained as $\delta+\delta' = 0.32 \pm 0.01$, $\eta = 0.30 \pm 0.05$, $\delta' = 0.16 \pm 0.05$, and $z = 1.55 \pm 0.06$. Finally, according to the NN results, $v_\perp = 1.06 \pm 0.04$. All the critical exponents are in good agreement with the DP values within the error bars, as shown in Table I.

### Summary and Conclusions

The absorbing transition in the QCP was first introduced several years ago, motivated by experimental realizations using cold Rydberg atoms. As the first step, a mean-field solution that exhibits tricritical transition behavior was obtained. This was counterintuitive, because the classical contact process exhibits a continuous transition alone. Naturally, interest turned to lower-dimensional cases, for instance, the 1D-QCP. Here, using the NN approach with supervised learning, we successfully identified the transition point of the 1D-QCP. This finding enables us to perform FSS analysis and obtain various critical exponents. On the basis of these results, we confirms that the 1D-QCP belongs to the classical DP class.

We remark that the transition point $\omega_c = 6.0 \pm 0.05$ obtained by using the tensor network approach [50, 51] is close to the value we obtained (within the error bars) using the NN approach. However, different values of the critical exponents were obtained. This suggests that the exponent values may be very sensitive to the transition point. Another possible cause would be the simulation methodology. The QIMC method would be better for estimating the exponent values, because the 1D-QCP would produce a highly entangled state near the transition point, and the tensor network approach might therefore be inappropriate.

The NN approach is frequently used to classify experimentally obtained snapshots of ultracold quantum gases in closed quantum systems [39, 40]. Along this line, our result for cold Rydberg atoms may be regarded as another example to which the NN approach is relevant.

### Acknowledgments

This research was supported by the NRF, Grant No. NRF-2014R1A3A2069005 (BK).

[1] I. Carusotto and C. Ciuti, Rev. Mod. Phys. 85, 299 (2013).
[2] C. Noh and D. G. Angelakis, Rep. Prog. Phys. 80, 016401 (2017).
[3] H. J. Carmichael, Phys. Rev. X 5, 031028 (2015).
[4] J. M. Fink, A. Dombi, A. Vukics, A. Wallraff, and P. Domokos, Phys. Rev. X 7, 011012 (2017).

---

* bkahng@snu.ac.kr
[39] B. S. Rem, N. Käming, M. Tarnowski, L. Asteria, N. Fläschner, J. Venderley, V. Khemani, and E.-A. Kim, Phys. Rev. Lett. 116, 257204 (2016).

[40] A. Bohrdt, C. S. Chiu, G. Ji, M. Xu, D. Greif, M. Greiner, E. Demler, F. Rusdt, and M. Knap, Nat. Phys. 15, 921 (2019).

[41] M. Schulz, I. Sinayskiy, and F. Petruccione, Physics 12, 74 (2019).

[42] M. J. Hartmann and G. Carleo, Phys. Rev. Lett. 122, 250502 (2019).

[43] N. Yoshioka and R. Hamazaki, Phys. Rev. B 99, 214306 (2019).

[44] A. Nagy and V. Savona, Phys. Rev. Lett. 122, 250501 (2019).

[45] F. Vicentini, A. Biella, N. Regnault, and C. Ciuti, Phys. Rev. Lett. 122, 250503 (2019).

[46] M. Marcuzzi, M. Buchhold, S. Diehl, and I. Lesanovsky, Phys. Rev. Lett. 116, 245701 (2016).

[47] M. Buchhold, B. Everest, M. Marcuzzi, I. Lesanovsky, and S. Diehl, Phys. Rev. B 95, 014308 (2017).

[48] M. Jo, J. Um, and B. Kang, Phys. Rev. E 99, 032131 (2019).

[49] D. Roscher, S. Diehl, and M. Buchhold, Phys. Rev. A 98, 062117 (2018).

[50] F. Carollo, E. Gillman, H. Weimer, and I. Lesanovsky, Phys. Rev. Lett. 123, 100604 (2019).

[51] E. Gillman, F. Carollo, and I. Lesanovsky, New J. Phys. 21, 093064 (2019).

[52] E. Gillman, F. Carollo, and I. Lesanovsky, arXiv:2002.09238 [quant-ph] (2020).

[53] J. Marro and R. Dickman, Nonequilibrium Phase Transition in Lattice Models, (Cambridge University Press, Cambridge 2005).

[54] T. E. Harris, Ann. Probab. 2, 969 (1974).

[55] G. Odor, Rev. Mod. Phys. 76, 663 (2004).

[56] R. M. Ziff, E. Galur, and Y. Barshad, Phys. Rev. Lett. 56, 2553 (1986).

[57] R. Dickman and I. Jensen, Phys. Rev. Lett. 67, 2391 (1991).

[58] M. Henkel, H. Hinrichsen, and S. Lübeck, Non-Equilibrium Phase Transitions, Theoretical and Mathematical Physics, Vol. 1 (Springer, Netherlands 2009).

[59] H. Hinrichsen, Adv. Phys. 49, 815 (2000).

[60] H.-P. Breuer and F. Petruccione, The Theory of Open Quantum Systems, (Oxford University Press, Oxford 2007).

[61] M. Jo and B. Kang, Phys. Rev. E 101, 022121 (2020).

[62] D. Harris and S. Harris, Digital Design and Computer Architecture, 2nd ed. (Morgan Kaufmann Publishers Inc., San Francisco 2012).

[63] S. Kiranyaz, T. Ince, R. Hamila, and M. Gabbouj, 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan, 2015 (2015).

[64] S. Ioffe and C. Szegedy, arXiv:1502.03167v3 [cs.LG] (2015).

[65] M. Abadi et al. TensorFlow: Large-scale machine learning on heterogeneous systems, arXiv:1603.04467 [cs.DC] (2015).

[66] D. P. Kingma and J. Ba, arXiv:1412.6980 (2014).

[67] G. Hinton, “Neural Networks for Machine Learning - Lecture 6a - Overview of mini-batch gradient descent.” (2012) http://www.cs.toronto.edu/~hinton/coursera/lecture6/lec6.pdf