Classroom learning dynamics using a cellular automata spatiotemporal model comparing peer instruction and traditional instruction

Clarence Ioakim T. Sy and Johnrob Y. Bantang*

National Institute of Physics, University of the Philippines Diliman, Quezon City, Philippines

Ranzivelle Marianne L. Roxas-Villanueva

Institute of Physics, University of the Philippines Los Baños, Laguna, Philippines

Peer instruction (PI) has become one of the popular means of classroom instruction in Physics Education. Such instruction method is vastly different from how classes are traditionally handled where the instructor conducts a lecture for the entire duration of the class. In this study, we model the learning within classes of students as a probabilistic cellular automata model and investigate the effects of different factors such as seating arrangements, class size N, average learning rate λ, and heterogeneity $\delta\lambda$ on the class overall learning efficiency. The learning efficiency of the class is measured in terms of the overall class rate using the total learning time $t_{\max}$ for the whole class. PI mode appears to be more applicable with heterogeneous classes (in-class high learning rate max difference $\delta\lambda$). As commonly known TI mode is shown to be more advantageous for large N and homogenous class aptitude (low $\delta\lambda$). We show that the learning curves of highly heterogeneous classes with TI mode have two stages of learning: a fast initial stage and a slow final stage. The slow final state class learning rate is due to the late learners with lower λ-values and have general shape dependent on other model parameters as well. Such curves are absent with PI which consistently shifts depending on the other factors. Differing from a previous empirical data for the PI mode, classes perform best when the best in the class are situated within the inner corner. The model we used for this study currently ignores the actual non-isotropic seating orientation with which students do not share knowledge readily in all neighboring seating directions. Our model simplified the student capabilities to binary values and does not consider the positive effect of aptitude similarity during peer interactions as described in previous studies. We therefore find that PI performs better than, or similar to, traditional instruction. A mix of TI and PI would be the optimal mode of instruction taking advantage of TI for fast learning in the initial stage and the robust learning curve of PI in the later stages of class learning. Despite these simplifications, our model is able to provide general insights that are in agreement with previous studies and existing practices.

CONTENTS

I. INTRODUCTION

Peer instruction (PI)–an increasingly popular method of instruction–attempts to solve a disconnect between students' ability to answer quantitative and conceptual questions, particularly in physics subject matters [1, 2].

Such disconnect is said to be attributed to the typical approach of memorizing physics problem patterns and applying the same solution patterns regardless of their potential non-applicability. The PI method forces students to think critically and to engage with peers guided by the learning material as opposed to simply absorbing the information from lectures given by the instructor as in the traditional mode of instruction (TI). The method's effectivity has been shown not only in high school physics settings, but also in other disciplines and higher education settings [3–5].

The PI method as outlined by Mazur and Somers [1] typically involves an instructor giving a short lecture on the topic, followed by a multiple-choice question (MCQ) that the student answer individually. PI emphasizes on the next phases of instruction which involve repeated cycle of students sharing their answers to the MCQ and discussing them with peers, usually until they get the concepts right. This part makes the PI method student centric requiring the instructor to intervene with slightly modified steps to adjust to the needs of the class. As long as the same learning objectives are tackled, the instructor can choose variations in the method [6]. The instructor may often decide on: increasing the multiplicity of the peer sharing and discussion phase; having different kinds of requirements and/or levels of evaluation [7]; providing different MCQ sets; or giving another short lecture

———————

* Corresponding author: jybantang@up.edu.ph

to ensure clarity of concepts involved [8]. Modifications to the PI method have also been reported to minimize some undesired behavioral patterns such as just copying off each other [6]. The PI therefore contrasts significantly against the traditional method (TI) with which the instructor simply gives a lecture for the entire duration, typically with minimal peer interactions among the students during class time.

Despite the unique approach of de-emphasizing actual problem-solving during the PI lecture phase, quantitative problem-solving skills have been consistently shown to remain uncompromised, and in many cases even improved, relative to TI. Several advantages were pointed out in favor of the PI method in relation to the following dimensions: decrease in the number of students with extremely low scores [7–9], improved conceptual understanding [7], as well as being independent of the background knowledge [8, 10].

The dynamics of the learning process involved during PI is complex resulting from the forced interaction during the peer discussion stage. Existing mathematical models for this are either predictive such as with the case of the artificial neural network (ANN) models [11] or lack the spatial aspect of the process such as utilizing differential equations [12, 13]. Somewhat in between these is a generalized Ising model by Bordogna and Albano [14, 15] that considers three sources of information for the student to learn from: teacher instruction, peer interaction, and bibliographic materials (books, lecture notes, etc.). Their model shows that students learn more when they engage discussions with their peers than those who only listen to lectures. They also show that group structure affects student learning, and that low aptitude students may learn at the expense of high aptitude peers—a transmissionist view of PI. A survey of many other mathematical techniques are found in the work of Nitta [13]. We are yet to find mathematical models for learning that compares PI with TI.

Actual assessment results by Roxas *et al.* [11] is used to train their ANN to predict expected second assessment scores of students depending on seating arrangements. Using this neural network, they were also able to characterize students' learning and investigate the effects of group homogeneity. Their study revealed that an optimal seating arrangement exists for students under PI methods based on their final aptitude. In their paper, the measure of each students' improvement was calculated via the Hake gain [16] as $g \equiv (S_2 - S_1)/(1 - S_1)$ where S_1 and S_2 are scores for respectively first and second assessment. They also used the output/input ratio ($O/I = S_2/S_1$) to gauge student improvement. Class improvement is the average of any of the scores.

The results of their study show that the outer corner seating arrangement (SA) performed the best using both the Hake gain and the O/I, followed by inner corner, then random, then center. The seating arrangements are illustrated in Fig. 1. In simulated classrooms (ANN system), each with 64 students and 10 classrooms in total, they found that homogenous classrooms with low apti-

tudes have significantly higher O/I-values under PI. This means that low aptitude students benefit the most from being grouped together.
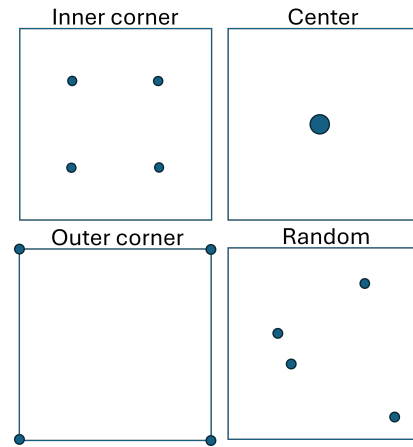

FIG. 1. Seating arrangements (SAs) considered for PI. Dots represent position of high aptitude students. The position of other students is typically arranged in rectangular pattern across the room (here square for brevity). The random case is just one of the many possible arrangements. Unless specified, we present the inner corner SA in our results.

Hake gain predicts the following expression for the expected score S_2 during the assessment after the first such that when done iteratively, every next assessment S_2 is dependent on the first assessment score S_1 as follows.

$$S_2 = S_1 + g \cdot (1 - S_1) \tag{1}$$

where g may be a function of S_1. Nitta [13] presented a model of PI learning that relates the probabilities that the students answers the MCQ correctly (effectively, post- and pre-test). Such probabilities would then imply that the Nitta Model predicts that the gain is dependent on the previous score such that:

$$g = g(S_1) = S_1. \tag{2}$$

Other models provide Hake gain expressions (e.g. Pritchard *et al.* [12] 3) resulting to more complex formulation for the function $g(S_1)$. This model considers other factors that may affect the learning process. K_{ext} is knowledge that students have that did were not part of the assessment and α_{memory} and $\alpha_{connected}$ are probabilities of learning and information being retained respectively.

$$g = 1 - \exp\left\{-[\alpha_0(1 - \beta) + \alpha_1 \beta K_{\text{ext}}]t\right\} \tag{3}$$

Despite the diversity of the models that we found, current models of PI lack some aspects that we would like to incorporate together like seating arrangements, students' learning rate, and aptitude heterogeneity with which the introduction of PI from TI are typically founded on.

II. CELLULAR AUTOMATA METHOD

We use a probabilistic cellular automata (pCA) model [17] (also known as stochastic cellular automata [18–21]) to study the spatio-temporal dynamics of learning for a more direct implementation of any instructional mode. For this pCA model, each cell in the automaton represents a student arranged in such a way each student is adjacent to (a CA *neighbor* of) four others resulting to the usual rectangular lattice configuration. For the rectangular lattice, we assign the unique position (i,j) to each student such that $i,j \in [1,L]$ and the classroom size is $N = L^2$. The students at the edges and corners have less number of neighbors as in the real situations.

Each student at every lattice position (i,j) is then assigned a learning rate $\lambda_{i,j} \in \mathcal{R}$ as a quantification of how fast they learn as an individual with $0 \le \lambda_{i,j} \le 1$ (or a probability to learn from any other neighbor). This parameter is directly related to the learning aptitude of the students such that we consider the fast learners to have *high aptitude score* ($\lambda \gg 0$). We also introduce the state $s_{i,j}$ representing degree of knowledge of the student such that in this work we use $s_{i,j} \in \{0,1\}$ as a two-level distinction between *learned* (or high-scoring student, $s_{i,j} = 1$) and *unlearned* (low-scoring, 0). The choice of two-level state should not prevent the possibility of having more levels as long as they can be quantitatively discriminated.

For each time step, $t \leftarrow (t+1)$, each student's state $s_{i,j}$ is updated depending on the mode of instruction described in the following. We use an algorithm (see Fig. 2) general enough to contain various modes of instructions. The mode of instructions defines how $P_{i,j}$ is obtained.
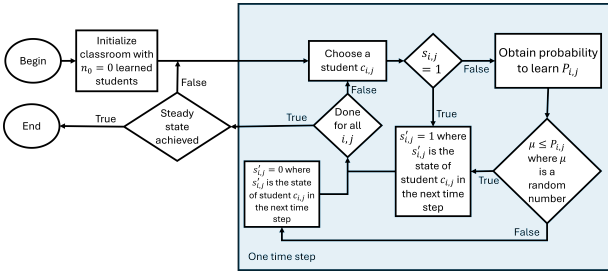

FIG. 2. Flowchart for the pCA algorithm. The PI and TI differ on how the probability $P_{i,j}$ is obtained as described in the text.

For traditional instruction (TI) mode, the probability of a student to learn in each time step ($P_{i,j}$) is

$$P_{i,j} = \lambda_{i,j}\ \rho_{\delta i,\delta j} \tag{4}$$

with $\rho_{\delta i,\delta j} \in [0,1]$ as the potential dependence of learning efficiency on the relative position $(\delta i, \delta j)$ of the source of learning (inspiration, input, interaction, discussion, etc.) with respect to the student position (i,j) [22]. The relative position of the student may matter since instructors in TI mode, for example, may be dependent on the row i since usually students near the front ($i = 1$) tend to

be more attentive and engaged in discussions than those in the last row ($i = L$). To futher simplify for TI, we fix $\rho_{\delta i,\delta j} = \rho_0$ as the probability of the students to learn from the teacher directly as if during an efficient plenary (or broadcast) lecture throughout the class such that we simply have $P_{i,j} = \lambda_{i,j}\rho_0$.

Different from TI, the peer instruction mode (PI) involves some learning during interaction with peers. Such interaction happens with the nearest neighbor unless specific instructions say different. We consider the simple case such that the probability $P_{i,j}$ is dependent also on the states of the nearest neighbor. In this case, we have a compounded probabilities resulting to

$$P_{i,j} = 1 - \prod_{\forall \delta i,\delta j} \left[1 - (\lambda_{i,j}\ \rho_{\delta i,\delta j})(s_{i+\delta i,j+\delta j})\right] \tag{5}$$

where we now have dependence on the state $s_{i',j'}$ ($i' = i + \delta i$ and $j' = j + \delta j$) factored in every probability of learning transmission. We also consider the simplified case such that $\rho_{\delta i,\delta j} = 0$ for $|\delta i| > 1$ and $|\delta j| > 1$ and $\rho_{\delta i,\delta j} = \rho_0$ (fixed, isotropic). This assumption translates to the class keeping a seat arrangement in a lattice pattern and that peer instruction can be modelled as essentially nearest-neighbor interaction only (we use Moore neighborhood) [20, 23, 24], although higher degree neighborhood interactions are possible [22].

The model allows incorporation of class heterogeneity due to variation in the learning rates of the students such as by specifying the distribution function for random assignment of λ-values. A variation about a fixed rate $\lambda_0 = 0.5$ is chosen for this work such that $\lambda_{i,j} = \lambda_0 \pm \delta\lambda$ (two-level heterogeneity). The value of $\delta\lambda$ acts as the parameter that determines the degree of class heterogeneity as it is tuned from 0 to less than 0.5 (we are limited by the probability range $0 < \lambda_{i,j} < 1$). We have to remove the case $\delta\lambda = 0.5$ that makes the model unrealistic (having $\lambda = 0$).

III. RESULTS AND DISCUSSION

All TI simulations start with no learned students, while PI simulations start with four (4) learned students. Class performance is measured as the speed of learning. Differences in the dynamical development over time of the class overall state is also observed. Here, we use the time $t_{\max}$ (in arbitrary units, a.u.) it takes for the entire class to reach a steady state. This steady state is considered reached during the numerical experiments when no new student is recorded to have transitioned from *unlearned* ($s = 0$) to *learned* ($s = 1$). This consideration prevents from having unnecessary very long waiting time for the student of very low learning rate $\lambda \approx 0$ to transition $s : 0 \rightarrow 1$. The average quantities and standard errors are based on twenty(20) independent randomized runs.

III.1. Dynamical difference between TI and PI

Figure 3 shows the typical monotonic increase in the fraction of learned students in classes starting with an unlearned state (all students in *unlearned* state, except for PI with which we start with four initially learned students). TI mode shows an early onset of high number of learning students in the first few iterations of lecture. The overall learning rate for TI tapers off as the entire class waits for the last few persons to actually learn. Due to the required percolation between learned and unlearned states, the PI mode shows a timid initial rise of the number of learned students only to suddenly rise towards the later part. PI mode catches up fastest in the later part of the iterations especially for low class size N. This observation for the PI is consistent independent of the seating arrangement (SA) considered. Class size does not affect TI mode as expected from the broadcasted mode of learning where each student independently and simultaneously learns from the teacher directly.

Figure 3(b) shows the effect of SAs for the PI mode. The inner corner SA outperforms all the other SAs for PI mode since such SA always has the highest rate of and earliest onset of rise in the number of learned students in the class compared to other SAs. When the students are randomly arranged, such SA performs as average of the two extreme types. The center and the outer corners perform similarly due to the idealized assumptions of the model (e.g. isotropic co-learning).

III.2. Transmission and heterogeneity

The learning rate ρ is expected to affect the class learning curve. The level of heterogeneity however affects the traditional mode of learning. Figure 4(a) shows that the class learning curves belong to the sigmoidal plots except for characteristic slopes and shifts. As one would expect, high ρ_0-values increase the initial number of learned students in both modes resulting to a relative shift in the onset of class overall learning. Such onset may be measured as the iteration number at which the fraction of learned students has reached 50% of the entire class. Opposed to TI, the shape of the learning curve for PI mode is more stable for different values of ρ_0 since only delay is introduced. TI's advantage for quick onset of learning however, is noticeable in all plots for the same ρ_0-values.

The effect of varying ρ_0 is nonlinear. For both modes, a change of students' average learning rate from $\rho_0 = 0.1$ to 0.5 translates to a greater shift in the class learning curve, as opposed the same shift due to the change from $\rho_0 = 0.5$ to 0.9. Class-level improvements can therefore be affected by the initial average class performance.

The effect of class heterogeneity for both modes can be shown using a simple model with mixed student learning rates λ such that some are low $(\lambda - \delta\lambda)$ and others are high $(\lambda + \delta\lambda)$. Figure 4(b) shows how this level of heterogeneity $\delta\lambda$ affects the resulting class learning curves. For the TI mode, the high level of heterogeity at $\delta\lambda = 0.4$
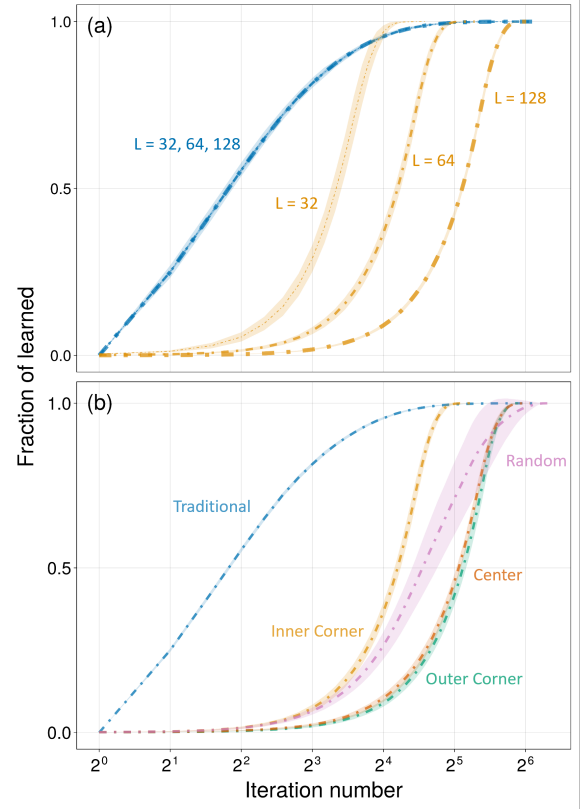

FIG. 3. Different monotonic increase of the total fraction of students in the learned state over iteration number. The blue colored plots are for TI while other colors are for PI. The same color corresponds to the seating arrangement (SA) used. Each PI's SA corresponds to a different color: orange (for inner corner), green (outer corner), orange (center), and pink (random). The linewidth is proportional to the N used. Other pertinent parameters fixed: $\rho_0 = 0.5$ and $\delta\lambda = 0.2$. The bands around each line show the standard deviation of the data after 20 independent random trials. Higher fraction of learned students indicate better overall learning. (a) Plot for different class sizes N. The best performant SA (Inner Corner) is used for comparison. (b) Plot for different seating arrangements (SAs).

(students have either $\lambda = 0.1$ or $\lambda = 0.9$ within the class) introduces a significant slow down midpoint of the initial rise of class learning curve. Such slow down is not significant for low heterogeneity. This two-stage phenomenon of class learning is attributed to the early switching of the fast learners (Stage 1) leaving the slow learners struggling until the later lecture iterations (Stage 2). The PI mode is robust to heterogeneity showing only slight uniform delay in class learning curve.

Spatial configurations of the class over iteration numbers provide insight on how learning spreads to the entire class. Figure 4(c) shows the comparison of such configurations of the two modes over different iteration times but for smaller class size $N = 32$ for brevity. The first to learn with the TI mode are those who are fast learners (dark pixels with higher λ-values) and then spreads over the whole class later when those with lower λ-values
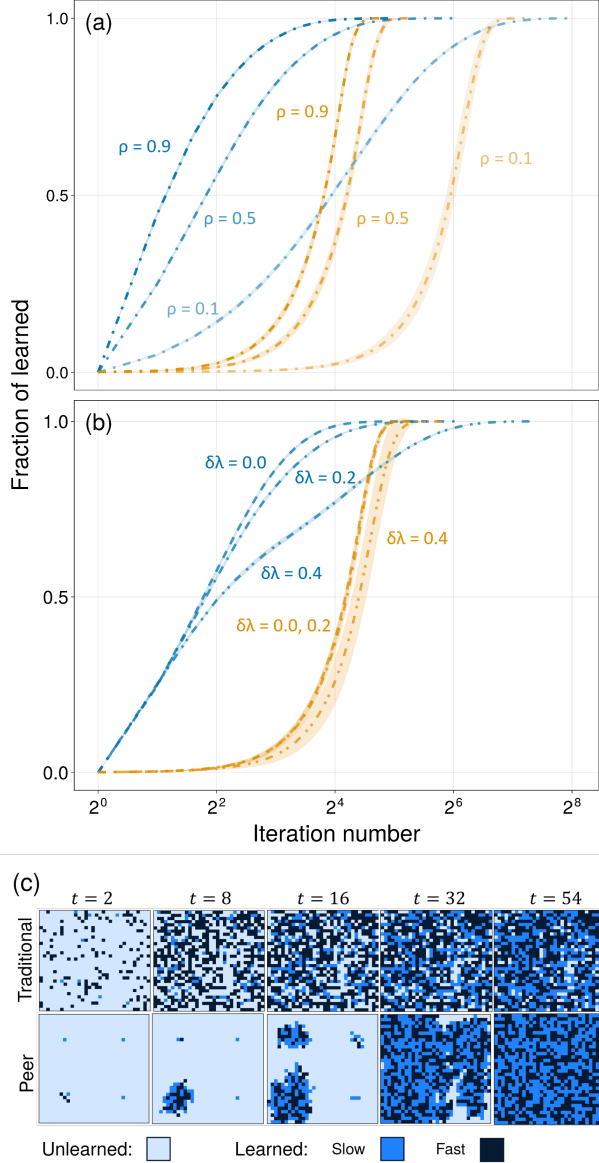
FIG. 4. The same plot as in Fig. 3, but for ranges of values for different chosen parameter. Unless indicated, the parameters are fixed such that $N = 64$, $\rho_0 = 0.5$, $\delta\lambda = 0.2$ (fairly homogenous) and an Inner Center SA for PI mode. (a) Plots for varied ρ_0-values, introducing varying degrees of shifts along the iteration (time) axis. (b) Plots for varied $\delta\lambda$-values. (c) Representative development of classroom spatial configuration ($L = 32$, $\rho_0 = 0.3$) observed for different iteration numbers from $t = 2$ (each column doubles the previous t) for TI (first row) and PI (second row) modes. The learned and unlearned states are differentiated by the color intensity, with $\lambda = 0.9$ (darker) and 0.1 (lighter) shades.

eventually learns. A slow pickup in the learning curve is observable with PI mode. This phenomenon is due to the percolation-based spread of learned states within the class. The spread through all learners however appears to average out resulting to a sudden increase of class learning rate in the later iterations due to larger number of learning sites analogous to increased active

sites in percolation phenomenon such as those in other CA-based models[25]. From this observations, we can also conclude that a mode using a combination of both TI and PI will benefit from these two phenomenon of innoculation of learning (via broadcast in TI step) and co-learning from peers (via percolation in the PI step). Current practices [1, 6, 8, 11] indeed also employ reading materials in advance and the instructor gives a short lecture (TI mode) before the students engage in peer discussion (PI mode).

III.3. Combined effects

The combined effects of average transmission probability (learning rate) ρ_0 and of the level of heterogeneity $\delta\lambda$ is shown in Figure 5 for two class sizes $N = 32^2$ (small) and 128^2 (large). The plots show the total time to learn $t_{\max}$ for the entire class. This time $t_{\max}$ is computed as the total number of iterations before the last cell (student) in the model has switched to learned state, or when the system does not change after about 20 rounds of random updates. For this plot, higher $t_{\max}$ is disadvantageous since it takes more iterations (hence time) before the entire class is expected to have learned.

Both Figures 5(a) and 5(b) show that as the level of heterogeneity increases, the traditional approach (TI) require longer $t_{\max}$ than the peer instruction (PI) especially for small classes. The difference becomes smaller for higher ρ_0 or for higher average class aptitude (fast learners). For large classes, this difference can actually reversed, but only for cases of homogenous, fast learning, large classes. Classically, the utilization of class grouping ("sections" based on aptitude such as honors class) is thus best suited for TI. Instructional strategies would actually prefer PI only for cases of inhomogenous class.

The same figure however also shows us that whenever we find TI advantageous, the difference in the $t_{\max}$ is not that large. In contrast, when PI is advantageous, the difference is very large. This suggests that PI can be a safe option — never performing much worse than TI, but can perform much better in many cases. The only disadvantage for PI is the huge delay in the rise of class learning curve. Overall, a mix of the two can be used to harness the advantages of both instruction modes.

IV. COMPARISON WITH AVAILABLE DATA

Our simulation design in its simplicity approximates well other qualities of classroom learning by comparing our results with real-world data as follows. Class-level learning process is usually assessed via the Hake gain g (cf. Eqn. 1). One can consider the Hake gain as a recurrence relation between successive evaluations S_1 and S_2. Since f can be used as an approximation for the probability S that the evaluation items are answered correctly by any member of the class, the Hake gain g can be used as a measure of how our CA model qualitatively
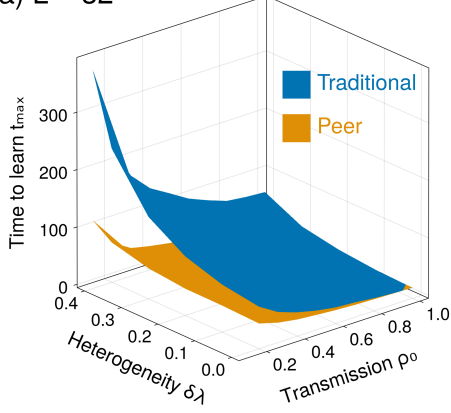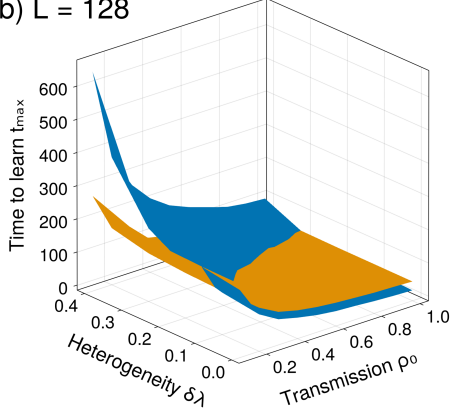
(a) L = 32

(b) L = 128

FIG. 5. Time to learn $t_{\max}$ as a function of positional learning factor ρ_0 and heterogeneity $\delta\lambda$ of PI (inner corner SA) and TI for two different class sizes $N = L^2$ ($L = 32$ and 128). The blue surface represent TI and the orange surface represents PI. Lower time to learn indicates better learning.

compares well with real-world data. Consequently our model determines how g is dependent on S_1 by mapping the evaluation score to the fraction of learned students: $S \to f$, and that we have $S_2 = f_t$ for $S_1 = f_{t-1}$.

Modify from hereon. Comparing our data to Nitta's model (Equation 1) [13], as in Figure 6, only our data for PI approaches their model — this is expected since their model is for PI and not TI. Furthermore, PI only approaches their model when the conditions are favorable for PI (Figure 6(b)). Considering that their model has been shown to agree with experimental data, it suggests that real world students' behavior and learning dynamics are better represented by the values of high positional learning factor ρ_0 and small class sizes N. It should also be noted that Nitta's model and data tracks individual students' learning while our model tracks the class's learning as a whole. We chose to track learning this way because there is not much insight to be gained from tracking individual students' learning in a binary-state model. Hence, we used the fraction of learned students f_t in place of Nitta's usage of students' test scores for S_t. Nevertheless, the agreement of our model to theirs

and to experimental data suggests that our model can be a good approximation of the real world with further improvements.
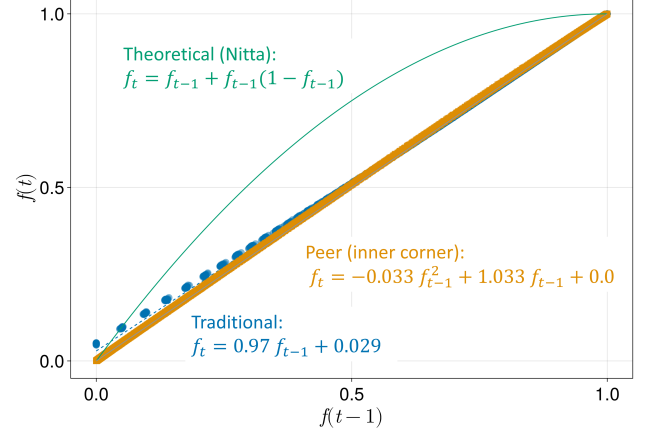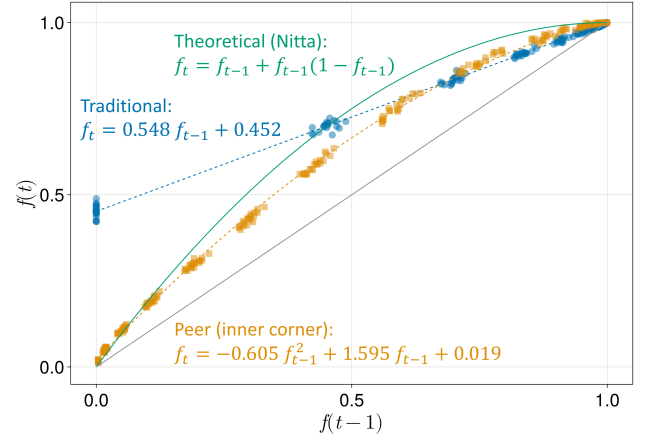

(a) $L = 128, \rho_0 = 0.1, \lambda = 0.5 \pm 0.4$

(b) $L = 32, \rho_0 = 0.9, \lambda = 0.5 \pm 0.0$

FIG. 6. Return map for selected parameter sets for both TI and PI. Blue circles represent the data points for TI and orange squares represent the data points for PI. The corresponding best-fit quadratic function obtained using the Levenberg-Marquardt algorithm for each set of data points are shown in blue and orange dashed lines respectively. The green line represents Nitta's theoretical model (Equation 1) and the gray line represents $y = x$.

Following Nitta's model and equation for gain (Equation 1 and 2), we can see for our PI model where

$$f_{t+1} \approx a + bf_t + cf_t^2 \tag{6}$$

b is expected to be $b = 1 - c$ and the gain g is

$$g \approx a + cf_t \tag{7}$$

Doing the same for our TI model which was best modeled by a linear function

$$f_{t+1} \approx f_t + g(1 - f_t) \approx m_0 + m_1 f_t \tag{8}$$

and noting that the coefficients m_0 and m_1 approximately sums to 1 (Figure 6) gives us the equation for

gain:

$$g \approx m_0 \qquad (9)$$

These values for gain g shows us that gains for TI are constant while gains for PI are linearly proportional to the fraction of learned students f_t. We find these values reasonable since TI is set up such that all students learn at the same time from the teacher while PI is set up such that students learn from each other.

Compared to Roxas' work [11] which found the outer corner SA to perform the best, we found the inner corner SA to perform the best. This difference can be attributed to the simplifications made in our model. We were not able to consider the aptitude similarity effect that happens where students of similar aptitude learn better when grouped together regardless of their actual aptitude level [6]. The implementation of this phenomenon is better suited for a continuous-state model rather than a binary-state model. Additionally, we did not consider the orientation of the students in the classroom, resulting in an isotropic system. In addition to the findings regarding seating arrangements, our model also agrees with their findings that homogenous classes find better improvements compared to heterogeneous classes.

V. SUMMARY AND CONCLUSION

We propose a probabilistic cellular automata model as a new way of investigating learning dynamics in the classroom. Through this model, we can investigate how different factors can affect students' learning in the classroom for both traditional instruction (TI) and peer instruction (PI).

We found that TI performs better in larger classes and when the students have low learning rate heterogeneity. On the other hand, PI performs better in smaller classes and when the students have high learning rate heterogeneity. To be more precise, class size, heterogeneity, and low positional learning factor affect both TI and PI negatively, but TI is not as affected by class size compared to PI and PI is not as affected by heterogeneity compared to TI.

We also found that when a highly heterogeneous class is under TI, the class exhibits a two-stage learning process where majority of the fast students learn in the first few time steps and the rest of the time is spent waiting for slower students to learn. Our findings regarding class heterogeneity are in line with those of Roxas, et al. (2010) [11]. However, our findings regarding seating arrangement differ from theirs where they found that the outer corner seating arrangement perform best while we found that the inner corner seating arrangement perform best. This difference is likely because of the simplifications made in our model — namely learning isotropy and non-implementation of an aptitude similarity effect. These simplifications are from the limitations of having a binary-state model.

Another limitation of having a binary-state model is that there are not much insight to be gained from tracking individual students' learning. Hence, we chose to track the class's learning as a whole. Despite having a different way to assess the methods of instruction effectiveness, our findings are consistent with the existing works like that of Nitta, et al. [13] which is backed by real-world data.

While our model can be a good approximation of the real world, there are still improvements that can be made. The current model only considers the peer discussion part of a whole PI session. A mixed model that considers the whole session with a mix of TI and TI would be a better representation of the real world. Additionally, transitioning to a continuous-state model would allow us to incorporate more real-world phenomena like aptitude similarity and orientation effects.

[1] E. Mazur, *Peer instruction: A User's Manual* (Benjamin Cummings, 1997).
[2] E. Mazur and M. D. Somers, Peer Instruction: A User's Manual, American Journal of Physics **67**, 359 (1999), https://pubs.aip.org/aapt/ajp/article-pdf/67/4/359/7528058/359_2_online.pdf.
[3] R. T. Johnson and D. W. Johnson, Active learning: Cooperation in the classroom, The annual report of educational psychology in Japan **47**, 29 (2008).
[4] A. P. Fagen, C. H. Crouch, T. Yang, and E. Mazur, Factors that make peer instruction work: A 700-user survey, in *talk given at the 2000 AAPT Winter Meeting, Kissimmee, FL* (2000).
[5] A. P. Fagen, C. H. Crouch, and E. Mazur, Peer instruction: Results from a range of classrooms, The physics teacher **40**, 206 (2002).
[6] M. K. Smith, W. B. Wood, W. K. Adams, C. Wieman,

J. K. Knight, N. Guild, and T. T. Su, Why peer discussion improves student performance on in-class concept questions, Science **323**, 122 (2009).
[7] C. H. Crouch and E. Mazur, Peer instruction: Ten years of experience and results, American journal of physics **69**, 970 (2001).
[8] N. Lasry, E. Mazur, and J. Watkins, Peer instruction: From harvard to the two-year college, American journal of Physics **76**, 1066 (2008).
[9] B. Thacker, E. Kim, K. Trefz, and S. M. Lea, Comparing problem solving performance of physics students in inquiry-based and traditional introductory physics courses, American Journal of Physics **62**, 627 (1994).
[10] S. Tobias, They're not dumb. they're different.: A new "tier of talent" for science, Change: The Magazine of Higher Learning **22**, 11 (1990).
[11] R. Roxas, S. Carreon-Monterola, and C. Monterola, Seat-

ing arrangement, group composition and competition-driven interaction: Effects on students' performance in physics, in *AIP Conference Proceedings*, Vol. 1263 (American Institute of Physics, 2010) pp. 155–158.

[12] D. E. Pritchard, Y.-J. Lee, and L. Bao, Mathematical learning models that depend on prior knowledge and instructional strategies, Physical Review Special Topics—Physics Education Research **4**, 010109 (2008).

[13] H. Nitta, A mathematical model of peer instruction and its applications, Upgrading Physics Education to Meet the Needs of Society , 87 (2019).

[14] C. M. Bordogna and E. V. Albano, Theoretical description of teaching-learning processes: A multidisciplinary approach, Physical Review Letters **87**, 118701 (2001).

[15] C. M. Bordogna and E. V. Albano, Simulation of social processes: application to social learning, Physica A: Statistical Mechanics and its Applications **329**, 281 (2003).

[16] R. R. Hake, Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses, American Journal of Physics **66**, 64 (1998), https://pubs.aip.org/aapt/ajp/article-pdf/66/1/64/12040304/64_1_online.pdf.

[17] C. I. T. Sy, R. X. A. Ramos, and J. Y. Bantang, Learning dynamics in a cellular automata model of classroom peer-to-peer interactions, in *Proceedings of the Samahang Pisika ng Pilipinas*, Vol. 42 (2024) pp. SPP–2024–1E–04.

[18] M. Arciaga, M. Pastor, R. Batac, J. Bantang, and C. Monterola, Experimental observation and an empirical model of enhanced heap stability resulting from the mixing of granular materials, Journal of Statistical Mechanics: Theory and Experiment **2009**, P07040 (2009).

[19] R. Batac, M. Pastor, M. Arciaga, J. Bantang, and C. Monterola, Kinks, logarithmic tails, and superstability in bi-disperse granular media, Physica A: Statistical Mechanics and its Applications **388**, 3072 (2009).

[20] P.-Y. Louis and F. R. Nardi, *Probabilistic Cellular Automata: Theory, Applications and Future Perspectives*, Vol. 27 (Springer, 2018).

[21] R. Fernández, P.-Y. Louis, and F. R. Nardi, Overview: Pca models and issues, in *Probabilistic Cellular Automata: Theory, Applications and Future Perspectives*, edited by P.-Y. Louis and F. R. Nardi (Springer International Publishing, Cham, 2018) pp. 1–30.

[22] Z. Dong, H. Liu, and X. Zheng, The influence of teacher-student proximity, teacher feedback, and near-seated peer groups on classroom engagement: An agent-based modeling approach, Plos one **16**, e0244935 (2021).

[23] J. Kari, Theory of cellular automata: A survey, Theoretical computer science **334**, 3 (2005).

[24] D. A. Zaitsev, A generalized neighborhood for cellular automata, Theoretical Computer Science **666**, 21 (2017).

[25] C. Beauchemin, J. Samuel, and J. Tuszynski, A simple cellular automaton model for influenza a viral infections, Journal of Theoretical Biology **232**, 223 (2005).