UNIVERSITY OF THE PHILIPPINES

Bachelor of Science in Applied Physics

Clarence Ioakim T. Sy

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

Thesis Adviser:

Johnrob Y. Bantang, Ph.D.
National Institute of Physics
University of the Philippines Diliman

Endorsed: 2024 11 11
JOHNROB Y. BANTANG

Date of Submission:
July 2025

Thesis Classification:
F

This thesis is available to the public

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

ENDORSEMENT

This is to certify that this thesis entitled **Classroom learning dynamics using a cellular automata spatiotemporal model comparing peer instruction and traditional instruction**, prepared and submitted by Clarence Ioakim T. Sy in partial fulfillment of the requirements for the degree of Bachelor of Science in Applied Physics, is hereby accepted.

JOHNROB Y. BANTANG, Ph.D.
Thesis Adviser

The National Institute of Physics endorses acceptance of this thesis in partial fulfillment of the requirements for the degree of Bachelor of Science in Applied Physics.

WILSON O. GARCIA, Ph.D.
Director
National Institute of Physics

To thinking complexly

ABSTRACT

CLASSROOM LEARNING DYNAMICS USING A CELLULAR AUTOMATA SPATIOTEMPORAL MODEL COMPARING PEER INSTRUCTION AND TRADITIONAL INSTRUCTION

Clarence Ioakim T. Sy
University of the Philippines (2025)

Adviser:
Johnrob Y. Bantang, Ph.D.

Peer instruction (PI) has recently 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 transfer of knowledge within the class as a probabilistic cellular automata model and investigate the effects of different factors such as seating arrangements, size, learning rate, and heterogeneity on the classes' overall learning efficiency in peer instruction. We compared the learning efficiency between the traditional learning model and the PI model. We found that larger class sizes sway the advantage towards traditional instruction, while increased learning rate heterogeneity favors PI. Additionally, an increase in the students' effective learning rate benefits both traditional instruction and PI but in different ways. Classes under traditional instruction were found to have two stages of learning when heterogeneity was introduced: a fast initial stage and a slow final stage. On the other hand, learning trends in PI were generally unaffected by heterogeneity, having similar effects with the other factors we considered. Among the seating arrangements (SAs) we considered for PI, the inner corner SA performed the best in terms of both the time it takes for all the students to learn and the classroom's learning rate. This result differs from previous studies where they found that the outer corner SA performed the best. The difference stems from the simplifications made in this model. We did not consider the orientation factor of each student, resulting in an isotropic system. Our model also uses binary values and does not consider the effect of aptitude similarity that have been described in previous studies. Despite these simplifications, our findings generally agree with previous studies and existing practices that PI performs similarly or better than traditional instruction and that a mix of traditional instruction and PI would be the optimal method of instruction. We offer insights based on the qualitative results of the model.

PACS: 01.40.Ha Learning in education, 05.45.Pq Numerical simulations, 89.75.-k Complex systems

Table of Contents

List of Figures

Chapter 1

Peer Instruction and Traditional Models of Teaching

Most of the literature on peer instruction (PI) for physics education references Eric Mazur's work *Peer Instruction: A User's Manual* [1, 2]. Since he started teaching physics at Harvard University in 1984, he has found that although students can memorize the laws and equations of physics and apply them in numerical problem-solving, there is a lack of understanding of the concepts behind the equations. Although his students can solve traditional quantitative problems and score high, they scored lower when given conceptual qualitative questions. Performing some statistical analysis on the scores of his students, he found that although 52% of his students did similarly well on conceptual and conventional problems, 39% of them did substantially worse on the conceptual problems than the conventional problems, while only 9% did substantially better on the conceptual problems. From this, he concluded the way students approached these physics courses was to memorize problem-solving algorithms that don't even work with all problems. He said that this explains why his bright students sometimes blundered and why students generally get frustrated with physics.

The rationale behind PI is that students must think for themselves during class discussions where they have to convince their seatmates that their answer is correct. In this way, students don't simply absorb the information given to them as they would during traditional lectures. When students engage learning materials actively and cooperatively, they develop good complex reasoning skills effectively — this applies to higher education students as well [3]. Despite having its roots in introductory physics courses, PI has also seen success in higher-level courses [4, 5].

1.1 Difference of Peer Instruction from Traditional Models

PI has been described as a student-centered approach to teaching. It aims to make the students actively engage with the material they are given more so than traditional lectures. Although there are many ways to implement PI, most methods involve shortening lectures delivered by the instructor, giving conceptual tests to gauge student understanding, and then allowing the students to interact and learn from each other [2, 6].

Considering that the goal of PI is to have more and deeper student engagement, the way it is implemented varies from traditional lectures. Most implementations of PI revolve around five things [2, 6].

First (1), students are expected to read the material before class. This pre-class learning opportunity allows the instructor to spend less time on delivering presentations to teach the material. It would also be beneficial if the materials that were given to students are written to be read before class. This is unfortunately not the case for most textbooks which are written mostly as something to be reflectively read after a lecture.

Second (2), short presentations that are usually given for each central point of the lesson. They are meant to contain little derivations, focus on the strategy on how to solve a problem from start to finish while highlighting the conceptual significance of each step. Because students are expected to have read up on the lesson, the instructor can also spend less time on definitions, especially those that are written in the book. In the case of physics education, example problems should ideally be quantitative to provide maximum physical insight while minimizing algebra.

Third (3) are qualitative and multiple choice concept tests. It is important for concept tests to be designed properly because they help the instructor gauge which parts of the lesson are hard for students to understand. They also serve as guides student discussions. In the paper written by Crouch, et al. [6], which had a modified version of Mazur's original peer instruction method [2], these concept tests are not graded, but students are given incentives when they participate regularly. Aside from incentives, the concept tests were included in the midterm and final exams, which gave more reasons for students to participate in them.

Fourth (4), peer discussion. After the concept test, students discuss their answers with those around them. They are encouraged to justify their answers by explaining the reason behind them. During this time, the instructor can walk around to listen in on student discussions to gain insights on students' understanding of the topic.

Fifth (5), at the end of peer discussion, another concept test on the same central point is given. The initial and final concept tests and insights from the peer discussions should guide the instructor on what to discuss and how much time to spend on the discussion. This discussion should focus mainly on explaining the correct answer to the concept test.

These parts or steps are not meant to be rigid. Since PI is supposed to be a student-centric approach, instructors should give considerations and make adjustments accordingly. The time spent on the concept tests and discussions should be varied based on the class as well as the difficulty of the topic. One can also vary the flow of discussion based on the percentage of correct answers. Lasry et al. [7] gave an example on how this can be done. If only a few students get the correct answer, the instructor can revisit the concept before giving another concept test. If most of the students get the right answer, the instructor can simply explain the correct answer before moving on to the next central point. In the case that the percentage of correct answers lies somewhere in the middle, the instructor can hold a peer discussion session before asking the students to revote on the concept test, which is then followed by a quick explanation from the instructor before moving on.

1.2 Benefits of Peer Instruction

Despite the de-emphasis on problem-solving in PI lectures, the students' learning in quantitative problem-solving was not compromised and was even improved compared to traditional lectures [6]. In this case, the students' learning comes primarily from discussions with their peers and homework assignments. They found that using PI significantly increased the students' scores on the quantitative problems. PI also decreased the number of students with extremely low scores on the quantitative problems. Their findings are consistent with the findings of other studies [7, 8]. Similarly, there was a significant increase in the students' conceptual understanding of the topics discussed in class [6]. This improvement holds true for both in-class concept tests and also for end-of-semester exams. Lasry et al. [7] affirms this finding by showing that PI provides better conceptual understanding to students while providing the same level of quantitative problem-solving skills as traditional instruction.

One concern for PI is that the student may not actually learn from peer interactions but instead choose answers based on those they believe are more knowledgeable. This hypothesis was tested by Smith et al. [9] by using isomorphic questions. Isomorphic questions are questions that have different "cover stories" but have the same underlying concept. By having a pair of isomorphic questions for the test-share-test process and by not showing the class's answer statistics, they found that the students actually learn from the peer interactions and don't just copy others' answers.

Furthermore, they found that for groups of students who did not get the initial question right (naïve), a significant number got the second question right despite not knowing the answer to the first question. This is in line with students' opinions wherein they say that a student who knows the answer to the initial question was not required for the discussions to be productive. In fact, some even shared that they learned better when no one in the group that knew the answer because in this way, they are forced to work through and discuss the problem. This finding contradicts the view that the value of PI comes from being "transmissionist" where students only learn from learned individuals, either students or instructors. The study done by Smith et al. [9] provides evidence that PI is more constructivist rather than transmissionist - meaning that the students are learning on their own through discussion and not from simply hearing the correct answer.

Besides those mentioned above, Lasry et al. [7] tackles two other things: the dependence of PI on student background knowledge and the effect of PI on student attrition. They found that PI performs as well as or better the traditional instruction, even when students under traditional instruction had more background knowledge. In a two-year college, they found that PI is more effective than traditional instruction for both high and low background knowledge students. In a four-year college like Harvard University, they found that there was no correlation between pre-test scores and the gain in the post-test scores, unlike in the two-year college. This difference in dependence tells us that the gains of PI is not universal and is hypothesized to be dependent on students' high reasoning abilities, as suggested by previous studies [10]. In the same study, it was found that PI significantly reduced the number of students who dropped out of the course. PI helps reduce student attrition by shifting the focus and method of instruction of the courses. A paper by Tobias [11] suggests that the competitive nature and the focus on skill performance are possible reasons for students dropping out in science courses. PI addresses both by shifting the focus to conceptual understanding and cooperative learning. As a consequence of the reduction of students dropping out, the failure rate of students also dropped significantly.

1.3 Drawbacks of Peer Instruction

As mentioned in Section 1.1, prior to coming to the classroom, students are expected to have read up on the topic to help PI methods proceed smoother [2, 6]. This can prove to be a disadvantage when instructors must rely on students' own discipline to do the tasks. Although there are some things that can be done to help students, like providing guide questions or learning objective, instructors will still have to ultimately trust their students to do the work outside the classroom.

Another draw back of PI that was outlined by Crouch et al. [6] is that concept tests take up a lot of time. Because of the amount of time it takes up, instructors will have to adjust by either (1) removing some topics from the course or (2) not having class sessions for some topics, leaving it to students to learn on their own via reading and problem sets. A similar problem arises when we consider the classes' focus on conceptual understanding (which is part of classroom time management for PI). It will also be up to the students to study the things that were not tackled in-depth in class such as derivation and practice problems.

In addition to dependence on students' self-discipline and a different classroom time management style, the effectiveness of PI can also depend on students' background knowledge and reasoning capabilities [7]. While this dependence can be advantageous, it can also be a disadvantage. Although previous studies show that PI does not really perform worse than traditional instruction [6, 7], we should consider the time and resources needed to transition from the latter to the former and decide whether that is worth the non-uniform development.

1.4 Existing mathematical models of PI

Although there are some mathematical models for learning, the ones that describe learning in the classroom are few and far in between - even more so for those that model PI.

Roxas et al. [12] used actual assessment results to train a neural network to map student interactions in PI classrooms. Using this neural network, they were able to characterize information transfer and investigate the effects of group homogeneity. Their study also investigated the optimal seating arrangement for students under PI methods based on their aptitude. In their paper, the measure of students' improvement was calculated via the Hake gain as shown in Equation 1.1 [13]. They also used the output/input ratio (O/I), which was the ratio of second assessment scores vs first assessment scores, to gauge student improvement. However, it should be noted that O/I values tend to be biased towards low-scoring students.

$$\langle g \rangle = \frac{\langle 2^{\text{nd}} \text{ assessment} - 1^{\text{st}} \text{ assessment} \rangle}{\langle 1 - 1^{\text{st}} \text{ assessment} \rangle} \tag{1.1}$$

The results of their study show that the outer corner seating arrangement (SA) performed the best, followed by inner corner, then random, then center (see Figure 2.4 for SA visualizations). In simulated classrooms, each with 64 students and 10 classrooms in total, they found that homogenous classrooms with low aptitudes have significantly higher O/I values. This means that low aptitude students benefit the most from being grouped together.

Nitta [14] gives us a few existing models that model PI. One of the models that was presented is a generalized Ising Model by Bordogna and Albano [15, 16] where they consider 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.

Nitta also presents a model by Pritchard et al [17] where PI is modeled as a set differential equations that is dependent on the probability of students learning to stick (memory model, Equation 1.2a) and the ability for students to associate new learnings from old knowledge via logistic differential equation (connectedness model, Equation 1.2b.)

Memory model:

$$\frac{dU_T(t)}{dt} = -\alpha_m U_T(t) \tag{1.2a}$$

Connectedness model:

$$\frac{dU_T(t)}{dt} = -\alpha_c U_T(t) K_T(t) \tag{1.2b}$$

where knowledge is taken to grow at a uniform rate, as in the tutoring model:

$$K_T(t) = a_{tu} t + K_T(0) \tag{1.2c}$$

$$U(t) + K(T) = 1 \tag{1.2d}$$

In these equations $U(T)$ and $K(T)$ are the unknown and known knowledge domains respectively. α_m, α_c, and α_{tu} are the corresponding rates for the memory model, connectedness model, and tutoring model

In deriving their own equations to model PI, Nitta arrived at equations simialr to Hake gain to evaluate the effectiveness of PI for a concept test question and Pritchard's connectedness model to model the classes' learning processes. Comparing their equations to data, they concluded that these metrics and equations roughly agree with the data and could give us insights on the learning dynamics of the classroom.

1.5 Problem statement

The process of PI is complex, with many interacting components. Existing models are either predictive, as in the case of the neural network modeling of Roxas et al. [12] or lack the spatial aspect of the process as with the differential equations of Pritchard [17] and Nitta [14]. While Bordogna et al. [15, 16] present to us a dynamical model in their generalized Ising model, it lacks some of the aspects of PI we'd like to consider like seating arrangements, students' learning rate, and heterogeneity.

We propose that a probabilistic cellular automata model can be used to study the spatiotemporal dynamics of both PI and traditional instruction when incorporating these different aspects into the model.

Chapter 2

The Classroom as a PCA

The classroom is a complex system that can be modeled as a probabilistic cellular automata (PCA). This chapter will discuss the classroom as a complex system and the probabilistic cellular automata model. The chapter will also discuss the implications of the model on the classroom and the teaching-learning process.

2.1 Cellular Automata and its use in modelling complex systems

A two-dimensional (2D) rectangular cellular automata can be defined by a five-tuple [18, 19]:

$$\text{CA} = \{\mathcal{S}, \mathcal{C}, \mathcal{L}, \mathcal{N}, \mathcal{R}\} \tag{2.1}$$

where

$\mathcal{S} =$ is the set of possible states that each cell can assume. The state s can use any kind of representation such as the set of integers $\{0, \ldots, n-1\}$ with n as the total number of possible states.

$\mathcal{C} = \{c = i, j \mid i \in \{1, 2, 3, \ldots, L_1\}, j \in \{1, 2, 3, \ldots, L_2\} \text{s.t. } L_1 \times L_2 = N\}$ is the set of identifiers for each cell in the automaton where N is the total number of cells and L_1 and L_2 are the lengths of each side of the automaton space. The cells can then be identified by their position in the automaton (i, j). So, the state of cell $c \in \mathcal{C}$ can be written as $s_c = s_{i,j} \in \mathcal{S}$

$\mathcal{L} =$ defines the lattice neighborhood which is generally a mapping $f : \mathcal{C} \to C^M$ where M is the number of neighbors of a cell $c \in \mathcal{C}$. Any given cell c is mapped to another tuple of cells: $L_{i,j} = \{(i-1, j-1), (i-1, j), (i-1, j+1), (i, j-1), \ldots, (i, j)$. Where r is the radius of the Moore neighborhood. We then say that $\mathcal{L}_{i,j}$ contains the set of neighboring cell for $c_{i,j}$.

$\mathcal{N} =$ $\mathcal{S}^M$, the set of neighborhood states. Thus, $N_c = N_{i,j} \in \mathcal{N}$ such that each $\mathcal{N}$ is in the form of the M-tuple $\{s_{i-1,j-1}, s_{i-1,j}, s_{i-1.j+1}, s_{i,j-1}, \ldots, s_{i,j}\}$.

$\mathcal{R} =$ defines the set of rules implemented in the CA with $g : s_{i,j} \mid \mathcal{L} \to \mathcal{S}$ as the mapping of any neighborhood state N_c to a new state $s'_{i,j}$ of the cell c. At the next time step, $s'_{i,j}$ replaces the original state $s_{i,j}$.

$\mathcal{N}$ can vary with the neighborhood structure and the boundary conditions of the automaton. The neighborhood structure dictates the shape of the neighborhood in the lattice. Common neighborhood structures include the von Neumann (diamond) and Moore (square) neighborhoods. Boundary conditions dictate how the automaton treats cells at the edge of the lattice when determining the neighborhood. Common boundary conditions include toroidal, spherical, and fixed boundary conditions.

$\mathcal{R}$ can also be affected by other factors such as whether the rules are deterministic or probabilistic and whether they are implemented synchronously or asynchronously. An automaton with deterministic rules will always produce the same output given the same input, while an automaton with probabilistic rules will produce different outputs given the same input. In Conway's Game of Life, a cell dies when it has three live neighbors, while a cell is born when it has two or three live neighbors. This is an example of a deterministic rule. An example of a probabilistic rule would be a cell dying with a probability of 0.25 when it has three live neighbors. An automaton with synchronous rules will update all cells simultaneously, while an automaton with asynchronous rules will update cells one at a time. (something explanation something about sync vs async)

Due to the flexibility of cellular automata, they can be used to model a wide variety of complex systems. Cellular automata have been used to model physical systems such as fluid dynamics, biological systems such as the spread of diseases, and social systems such as traffic flow [20]. Their discreteness and locality make it a good model for systems that are composed of many interacting parts. Thus, we have chosen to use a two-state probabilistic cellular automata to simulate the learning process for students in the classroom

2.2 PCA model for classroom dynamics

We used a two-dimensional binary probabilistic cellular automata (PCA) model to simulate the learning process in a classroom. In this PCA model, each cell in the automaton represents a student and the state of each cell represents their aptitude $S = \{\text{unlearned, learned}\} = \{0, 1\}$. We assign the neighborhood to be an outer-totalistic Moore neighborhood of radius $r = 1$ and define the boundary conditions to be fixed wherein the grid does not wrap around itself and $s_{i,j} = 0$ for $i, j \notin [1, L]$. How the automaton updates depends on the learning setup of the classroom. In this study we tackle two learning setups: traditional and peer instruction (PI). We take the traditional instruction to be the case where the teacher is the only source of information for the students — which is typical for classes where the time is mostly spent on lectures. In the PI setup, we consider that the students learn from each other but only when the students of interest already know the lesson or is in the learned state.

2.2.1 Traditional Instruction

To model the learning process in a classroom with the traditional instruction setup, we set the probability for the student to learn in each time P_{ij} to be dependent only on their individual learning rates $\lambda_{i,j}$, which describes how receptive they are to learning, and the probability for the student to learn from the teacher ρ_0. ρ_0 can be varied per individual based on where they are seated, but for simplicity, we set it to be the same for all students.

$$P_{ij} = \lambda_{ij}\rho_0 \tag{2.2}$$

where

$P_{i,j} \in [0, 1]$ is the probability of student $c_{i,j}$ to learn in each time step,

$\lambda_{i,j} = 1$ is the learning rate of student $c_{i,j}$

$\rho_0 \in [0, 1]$ is the probability of $c_{i,j}$ to learn from the teacher

In five-tuple form, the traditional PCA model for the classroom can be written as:

$$\mathcal{S} = \{\text{learned, unlearned}\} = \{0, 1\}$$

$\mathcal{C} =$ $\{(1,1),(1,2),\ldots,(1,L),(2,1),(2,2),\ldots,(2,L),\ldots,(L,L)\}$ where L is the length of the square classroom.

$\mathcal{L} =$ $f(c) \leftarrow [L_c = \{(i+\delta i, j+\delta j)\ \forall\ (\delta i \wedge \delta j), \delta i, \delta j \in \{-1,0,1\}\}$ as a mapping for outer-totalistic Moore neighborhood of radius $r = 1$ with a fixed boundary condition.

$\mathcal{N} =$ $\{00000000, 00000001, \ldots, 11111111\}$ such that the representation of the neighborhood state $N_c \in \mathcal{N}$ is equivalent to $N_c = \{s_{i+\delta i, j+\delta j} \forall \delta i, \delta j \in \{-1,0,1\}\}$.

$\mathcal{R} =$ the probabilistic rule defined by Equation 2.2.

The numerical procedure is outlined in Figure 2.1. Each simulation for the traditional model starts with all students unlearned. The simulation is considered finished once all the students have learned.

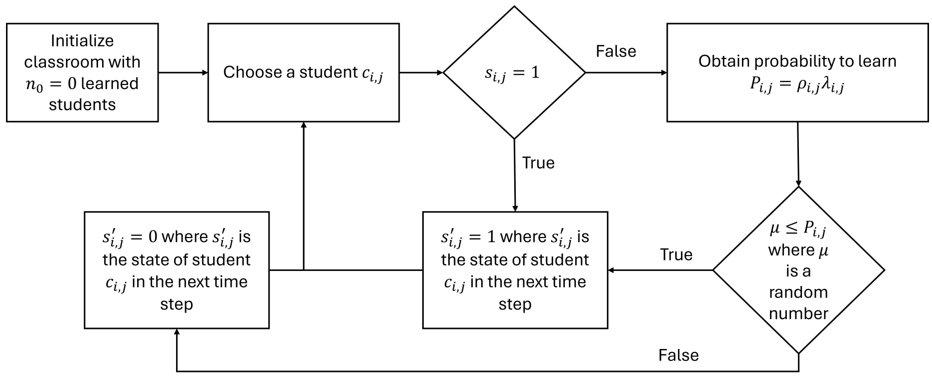
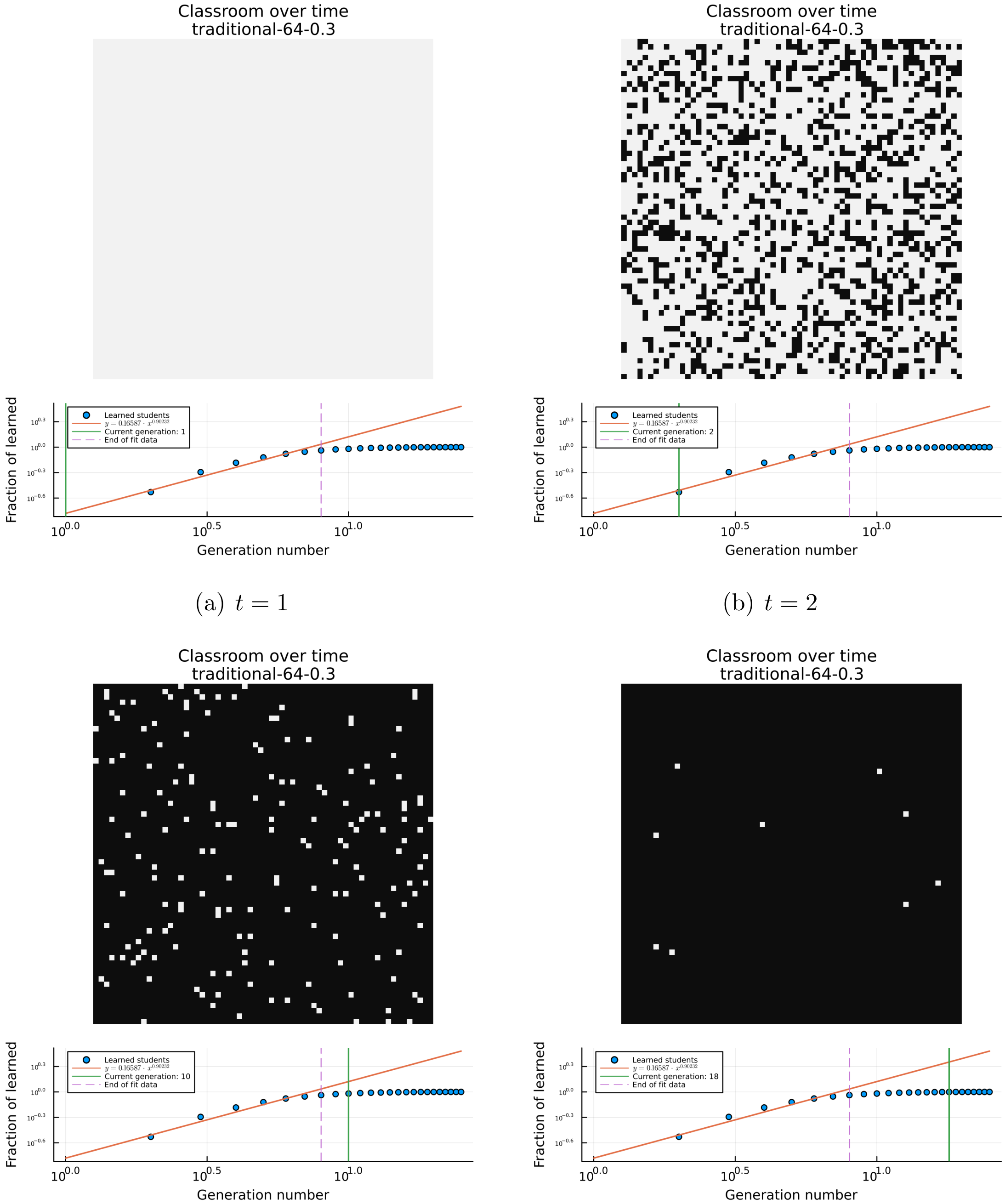


Figure 2.1: Numerical process for simulation of 2D BPCA for traditional setups.

Figure 2.2 shows an example of a classroom's evolution over time in the traditional set up. The sample classroom is set with $L = 64$, $\lambda_0 = 1$, and $\rho_0 = 0.5$.

2.2.2 Peer instruction (PI)

To model the learning process in a PI set up, we set the probability for the student to learn in each time P_{ij} to be dependent on three factors. First (1), their learning rate $\lambda_{i,j}$. Secondly (2), the positional learning factor $\rho_{i+\delta i, j+\delta j}$ which describes how likely it is to learn from the neighbor $c_{i+\delta i, j+\delta j}$ based solely on their relative position with respect to $c_{i,j}$. Lastly (3), the aptitude level of the neighbor $s_{i+\delta i, j+\delta j}$ which dictates



(a) $t = 1$

(b) $t = 2$

(c) $t = 10$

(d) $t = 18$

Figure 2.2: Sample classroom evolution for $L = 64$, $\lambda_0 = 1$, and $\rho_0 = 0.3$ with traditional instruction. Black squares represent learned students while white squares represent unlearned students. The accompanying graphs show the fraction of learned students as a function of time step. The blue circle represents the data points while the orange line shows the power law fit. The broken pink line shows the where we truncate the data for fitting the power law and the green line shows the current time step in the plot.

whether student $c_{i,j}$ can learn from them. The probability for a student to learn in each time step is then determined by the following equation:

$$P_{ij} = 1 - \prod_{\forall \delta i, \delta j} [1 - (\lambda_{ij})(\rho_{i+\delta i, j+\delta j})(s_{i+\delta i, j+\delta j})] \tag{2.3}$$

where

$P_{i,j} \in [0,1]$ is the probability of student $c_{i,j}$ to learn in each time step,

$\lambda_{i,j} = 1$ is the learning rate of student $c_{i,j}$

$\rho_{i+\delta i, j+\delta j} \in [0,1]$ is the probability of $c_{i,j}$ to learn from their neighbors in seats $\{c_{i+\delta i, j+\delta j} \forall \delta i, \delta j \in \{-1,0,1\}\}$ solely based from their relative positions with each other, and

$s_{i+\delta i, j+\delta j} = \{\text{unlearned, learned}\} = \{0,1\}$ are the neighbors' aptitude level.

The derivation of Equation 2.3 is shown in Appendix A.1.

The numerical procedure is outlined in Figure 2.3. Each simulation starts the classroom with four learned students $n_0 = 4$ placed in different seats in the classroom. The simulation is considered finished once all the students have learned.

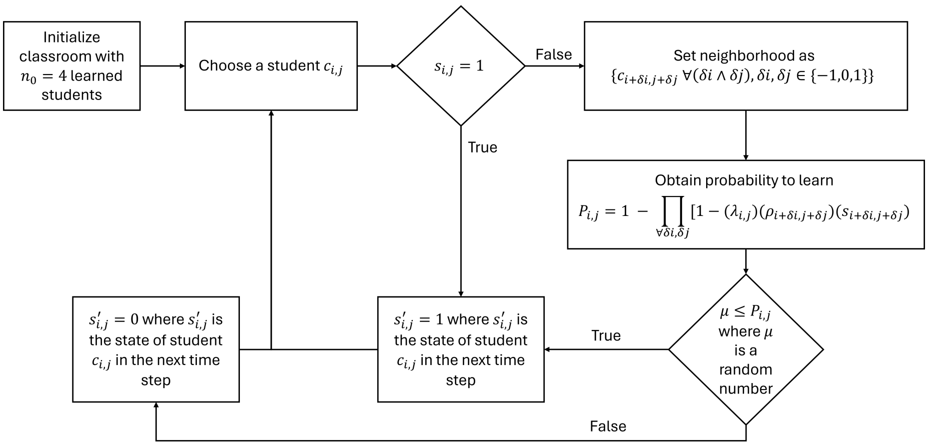


Figure 2.3: Numerical process for simulation of 2D BPCA for PI setups.

The seating arrangement (SA) were chosen from a previous study that shows that the SA can affect the learning process [12]. These SA's are namely: inner corner, outer corner, center, and random. The different SA's are shown in Figure 2.4.

Figure 2.5 shows an example of classroom's evolution over time. The sample classroom is set with $L = 64$, $\lambda_0 = 1$, and $\rho_0 = 0.5$ with the inner corner SA.

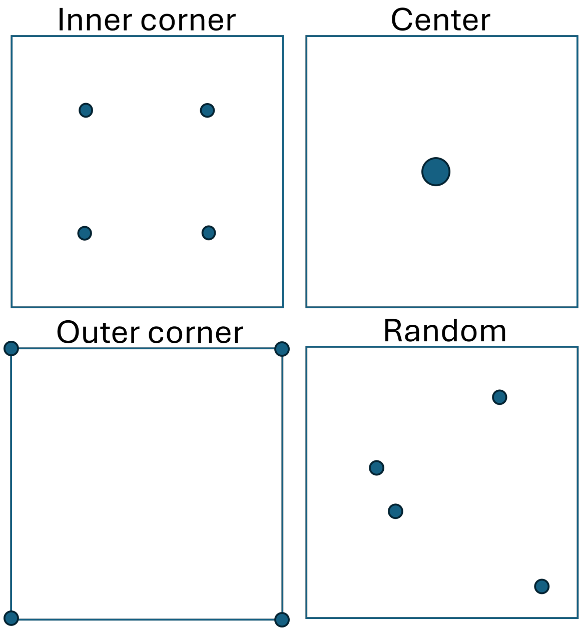
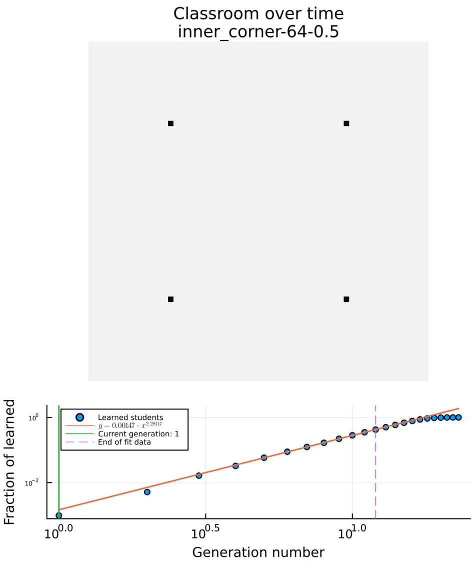
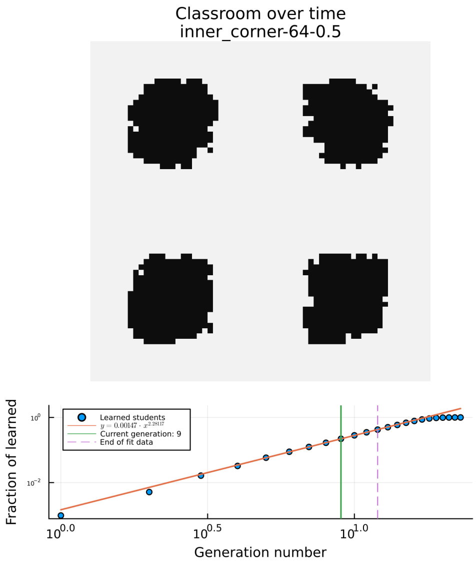


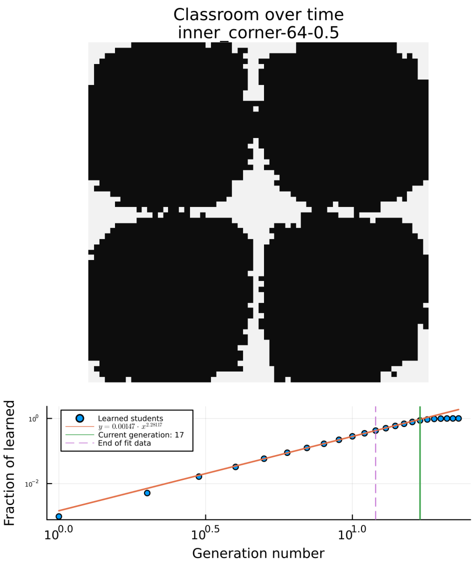
Figure 2.4: Peer instruction seating arrangements. Circles denote high aptitude students. The inner corner SA places high aptitude students halfway between the center and the corner of the classroom. The outer corner SA places high aptitude students at the corner of the classroom. The center SA places high aptitude students in the center of the classroom. The random SA places high aptitude students randomly throughout the classroom.



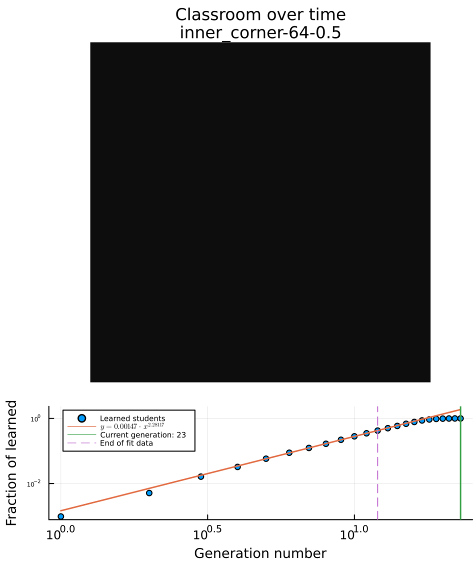
(a) $t = 1$



(b) $t = 9$



(c) $t = 17$



(d) $t = 23$

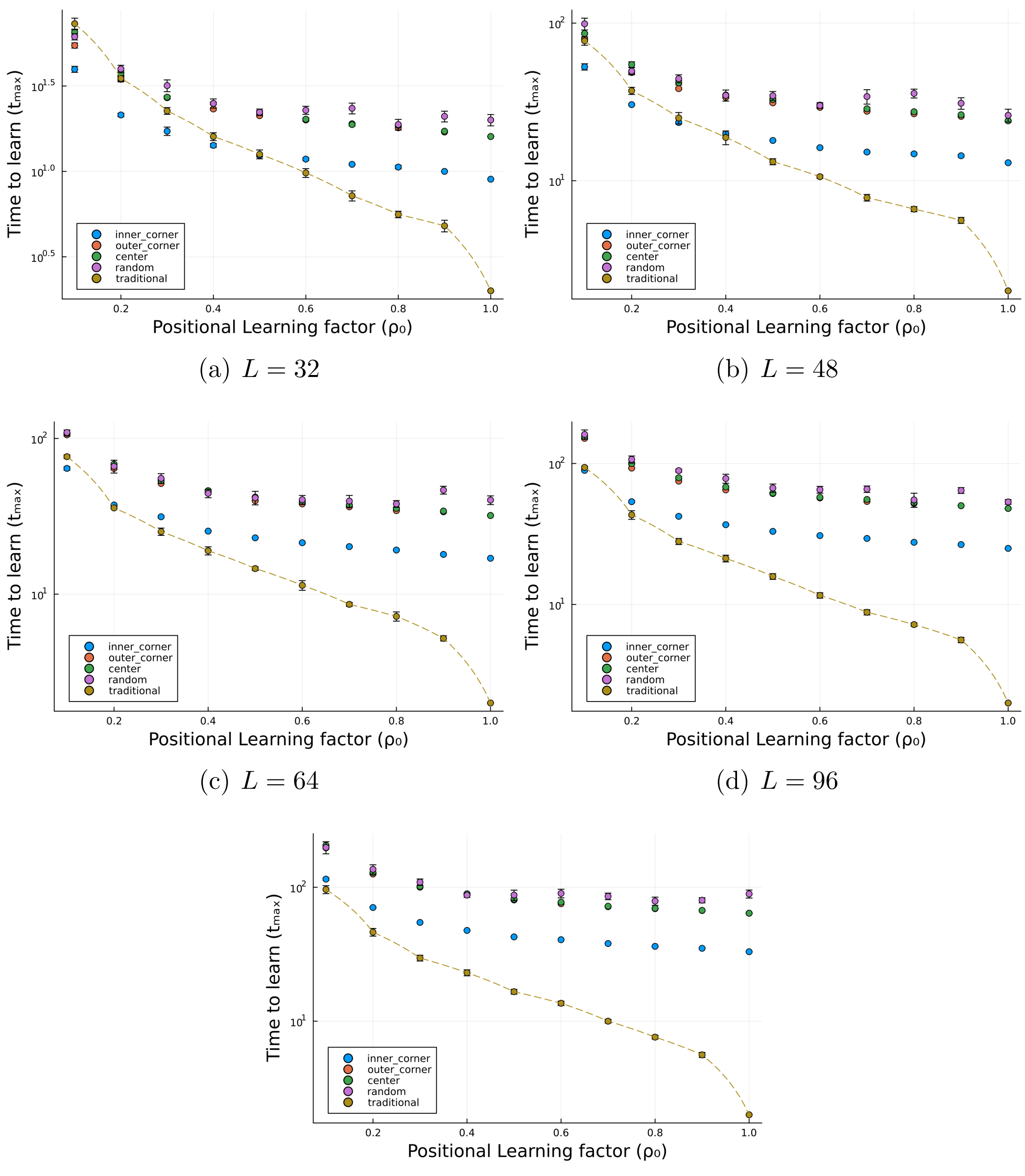
Figure 2.5: Sample classroom evolution for $L = 64$, $\lambda_0 = 1$, and $\rho_0 = 0.5$ with the inner corner SA. Black squares represent learned students while white squares represent unlearned students. The accompanying graphs show the fraction of learned students as a function of time step. The blue circle represents the data points while the orange line shows the power law fit. The broken pink line shows the where we truncate the data for fitting the power law and the green line shows the current time step in the plot.

2.3 Results: Homogenous PI vs Traditional

In this chapter, we only consider the case where all students have the same learning rate $\lambda_{i,j} = \lambda_0 = 1$ and an isotropic positional learning factor $\rho_{i,j} = \rho_0 \ \forall \ (\delta i \wedge \delta j), \delta i, \delta j \in \{-1, 0, 1\}$. From the simulations, we compared both the average number of time steps $\langle t_{max} \rangle$ it takes for all the students in the classroom to learn and the average learning rate $\langle m \rangle$ across different configurations over 5 independent runs. The class learning rate m for each trial was obtained by using a Levenberg-Marquardt algorithm to fit a power law $(y = ax^m)$ to the fraction of learned students as a function of time step. We only considered the first 50% of the data for the PI model and the first 25% of the data for the traditional model. This truncation was done so that we only fit the part of the data before the finite size effect starts to affect the simulation.

2.3.1 Time to learn t_{max} vs positional learning factor ρ_0

The data in Figure 2.6 shows that the time to learn t_{max} decreases with increasing positional learning factor ρ_0 for all classroom sizes for both traditional instruction and PI. Among PI SAs, the inner corner SA has the lowest t_{max} for all classroom sizes. The outer corner and center SAs performed similarly, while the random SA has the highest t_{max}. The traditional model performs situationally better than even the inner corner PI model, something we will investigate in Section 2.3.3.



(a) $L = 32$

(b) $L = 48$

(c) $L = 64$

(d) $L = 96$

(e) $L = 128$

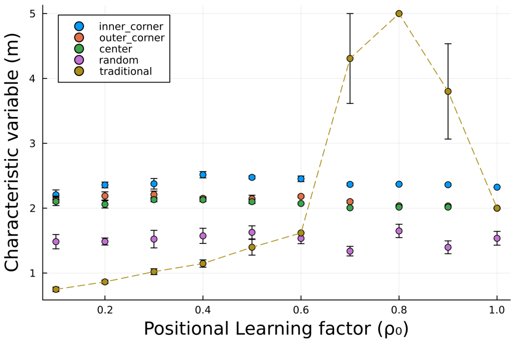
Figure 2.6: Time to learn t_{max} as a function of positional learning factor ρ_0 for different classroom sizes $L \in \{32, 48, 64, 96, 128\}$. Lower time to learn t_{max} indicate better performance.

2.3.2 Class learning rate m vs positional learning factor ρ_0

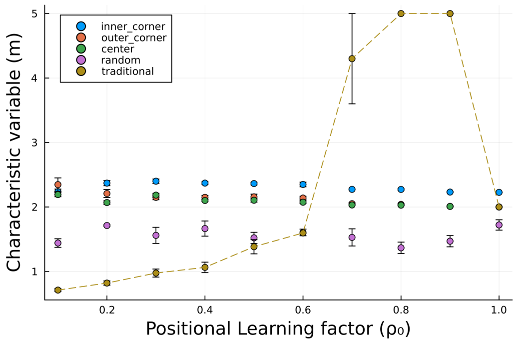
The data in Figure 2.7 shows that the class learning rate m does not generally change for different positional learning factors ρ_0 when compared against similar SAs for PI. In traditional instruction, however, there is an increase in class learning rate m with increasing ρ_0 for all classroom sizes. Similar to the findings in Section 2.3.1, the inner corner SA has the highest learning rate m for all classroom sizes. The outer corner and center SA's performed similarly, while the random SA has the lowest class learning rate m. We still find that the traditional model performs situationally better than even the inner corner PI model.

The sudden decrease in class learning rate m for the traditional model at $\rho_0 = 1.0$ is due to the improper truncation and fitting of the data. In cases of traditional instruction where $\rho_0 = 1.0$, all the students will transition from unlearned to learned in one time step, so fitting the power law to the first 25% of the data would yield a class learning rate m that poorly represents the simulations' dynamics.

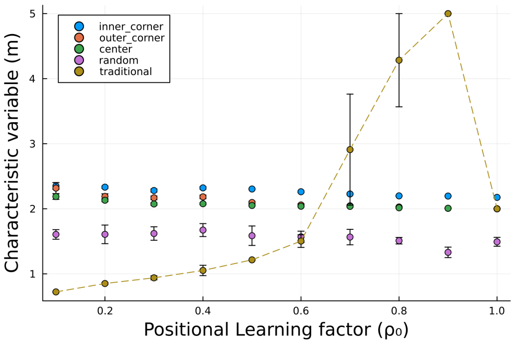
The results in this section show that traditional instruction performs better than PI in terms of class learning rate m for all classroom sizes at low ρ_0. This is in contrast to the results in Section 2.3.1 where PI performed better only in small classrooms with low positional learning factor ρ_0. This discrepancy is likely due to the different metrics used to evaluate the performance of the models. The class learning rate m is a measure of how quickly the students learn, while the time to learn t_{max} is a measure of how long it takes for all the students to learn. In cases when the classroom is large and positional learning factor ρ_0 is small, using traditional instruction enables the teacher to teach all the students in a shorter amount of time despite PI being able to teach the students more effectively.



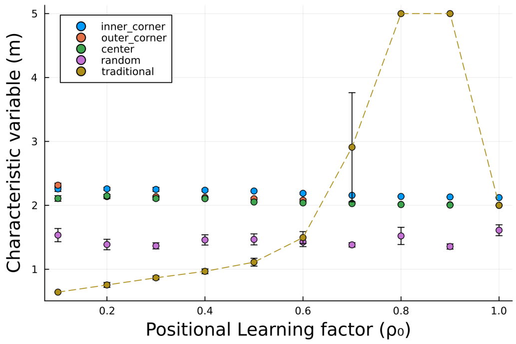
(a) $L = 32$



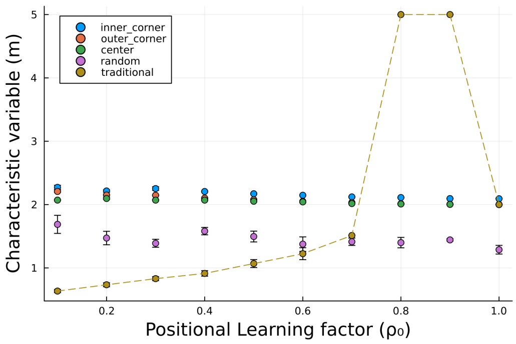
(b) $L = 48$



(c) $L = 64$



(d) $L = 96$



(e) $L = 128$

Figure 2.7: Class learning rate m as a function of positional learning factor ρ_0 for different classroom sizes $L \in \{32, 48, 64, 96, 128\}$. Higher class learning rate m values indicate better performance.

2.3.3 Time to learn t_{max} vs class size N

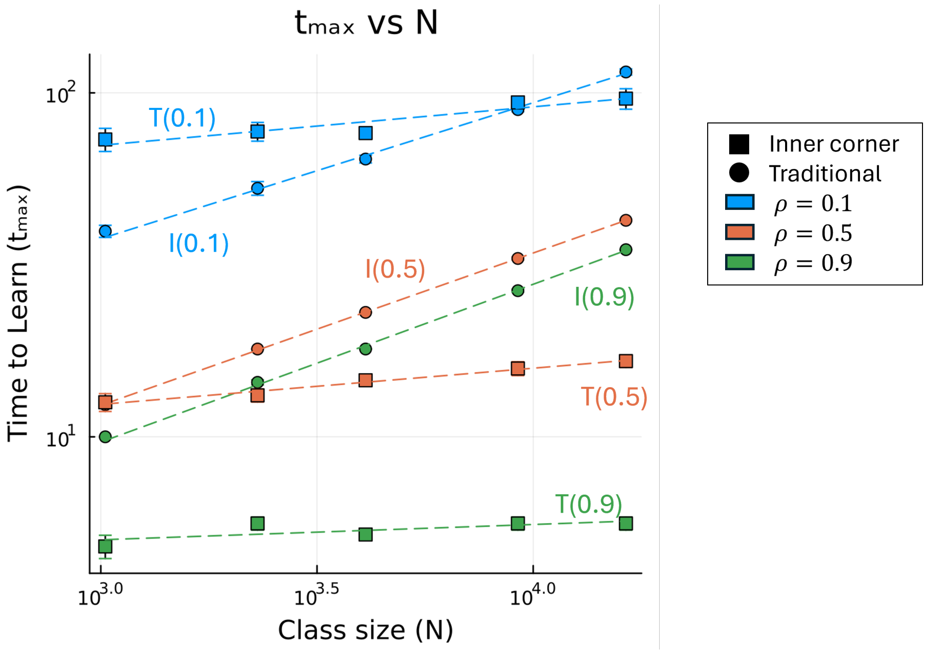


Figure 2.8: Time to learn t_{max} as a function of class size N for PI and traditional models. Circles and squares denote data points for PI (inner corner SA) and traditional instruction, respectively. The dash lines are the corresponding power law fit of the form $t_{max} = a \cdot N^b$ using $\rho \in \{0.1, 0.5, 0.9\}$. The fitted parameters are shown in the table with an average of $\langle b \rangle = 0.4532 \pm 0.019$ for the inner corner and $\langle b \rangle = 0.087 \pm 0.021$ for the traditional model. Lower time to learn t_{max} indicate better performance.

	Inner corner		**Traditional**	
	a	b	a	b
$\lambda = 0.1$	2.4515	0.3955	32.4884	0.1119
$\lambda = 0.5$	0.5795	0.4428	6.0079	0.1052
$\lambda = 0.9$	0.4055	0.4589	3.6914	0.0445

Table 2.1: Power law fit parameters a and b for t_{max} vs N, where $t_{max} = a \cdot N^b$

When we analyze the dependence of the time to learn t_{max} on the class size N, we find that the trends for both models follow the power law $t_{max} = a \cdot N^b$. In Figure 2.8 we observe transition points where the traditional instruction becomes more efficient than PI. These points happen at progressively lower class sizes as the value of ρ_0 increases. We also find two major groups of power laws based on their b values, as summarized in Table 2.1. The first group has an average b value of $\langle b \rangle = 0.4532 \pm 0.019$ for the inner corner SA, while the second group has an average

b value of $\langle b \rangle = 0.087 \pm 0.021$ for traditional instruction. The inner corner SA has higher b values than traditional instruction, indicating that traditional instruction is less affected by class size N and so is more scalable than PI.

2.4 Conclusions

In this chapter, we have shown that between different seating arrangements for PI, the inner corner SA is the most efficient in terms of time to learn t_{max} and class learning rate m. The outer corner and center SAs performed similarly worse than the inner corner SA, while the random SA performed the worst. This is different from what existing literature has shown, where the outer corner SA performed the best [12]. This is likely because of the simplifications made in our model. Our model does not incorporate factors such as the similarity effect mentioned in previous studies [9, 12]. This effect is the phenomenon wherein students of similar aptitude levels tend to learn from each other better when seated together regardless of their actual aptitude. Our model also does not consider anisotropic positional learning factors $\rho_{i,j}$, where the probability of learning from a neighbor varies with the neighbors' relative positions. Besides being binary, which introduces granularity, our model also does not consider that not all students are equally receptive to peer instruction or have the same learning rate. Introducing these factors, some of which we will explore in the next chapters, may make our model reflect reality better and provide us with better understanding of the learning dynamics in the classroom.

The nature of our model lends itself to be heavily influenced by geometric factors. The most impactful factor in this case is then the distance d_{max} of any student to a high aptitude or learned student as shown in figure 2.9. The inner corner SA minimizes d_{max}, which is why it is the most efficient. The outer corner and center SAs have equal d_{max} higher than the inner corner SA, which explains why they performed very similarly albeit worse than the inner corner SA.

Our analysis on the dependence of t_{max} on N explains why the traditional model performed better than the PI model in most cases. With lower b-values, the traditional model is less affected by class size than the PI models are. Because of this, the traditional model performed better in cases with larger classrooms.

Regarding the performance of PI methods with traditional teaching methods, we found that classrooms with higher ρ_0 values performed better in traditional instruc-

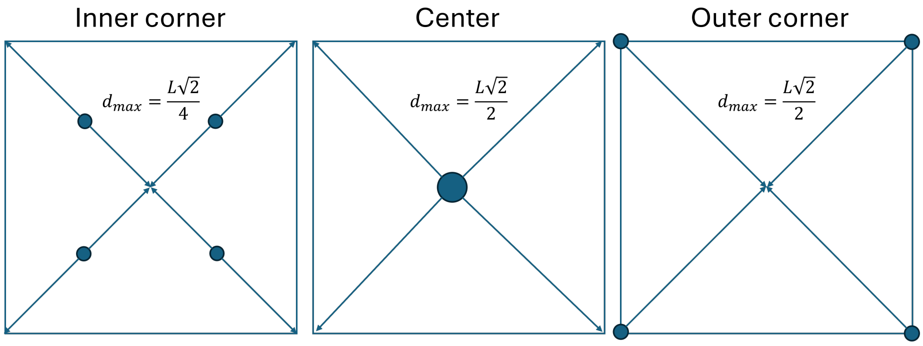


Figure 2.9: Distance d_{max} of any student to a high aptitude or learned student for different seating arrangements.

tion, while classrooms with lower ρ_0 values performed better in PI. A previous similar study [12] found that classes with lower aptitude levels were the ones who benefited the most from the PI methods. Although we cannot directly conclude this from our model, we were able to show that students with a lower probability of learning or students who learn slower, can benefit as much from PI as in traditional setups. This reinforces other studies' findings [7] that show that PI methods can be effective even regardless of the students' actual aptitude levels.

Chapter 3

Heterogeneous Learning Rates

To introduce learning rate heterogeneity in the classroom, we revisit equation 2.3.

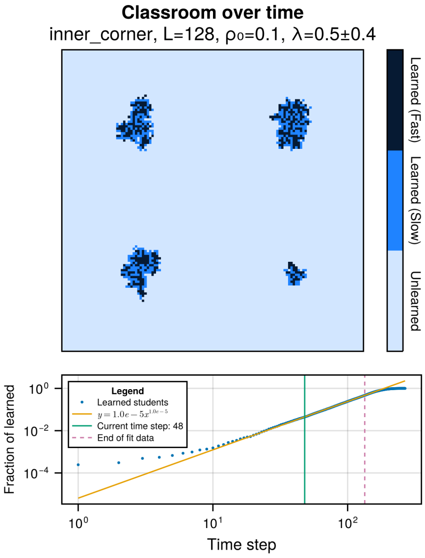
$$P_{ij} = 1 - \prod_{\forall \delta i, \delta j} \left[1 - (\lambda_{ij})(\rho_{i+\delta i, j+\delta j})(s_{i+\delta i, j+\delta j})\right] \qquad \text{(2.3 revisted)}$$

We can adjust the parameter λ_{ij} to introduce heterogeneity in each student's learning rate. We set a student's learning rate as $\lambda_{ij} = \lambda_0 \pm \delta\lambda$ where $\delta\lambda \in \{0.0, 0.1, 0.2, 0.3, 0.4\}$ and $\lambda_0 = 0.5$. Each student has an equal chance of having a learning rate that is either faster $(\lambda_0 + \delta\lambda)$ or slower $(\lambda_0 - \delta\lambda)$ than the average learning rate λ_0.

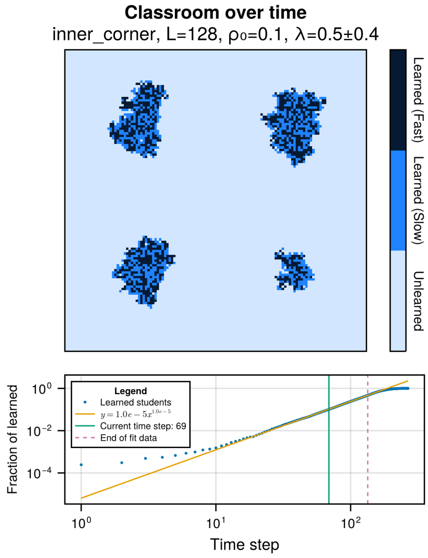
3.1 Effects on classroom evolution

3.1.1 Peer Instruction (PI)

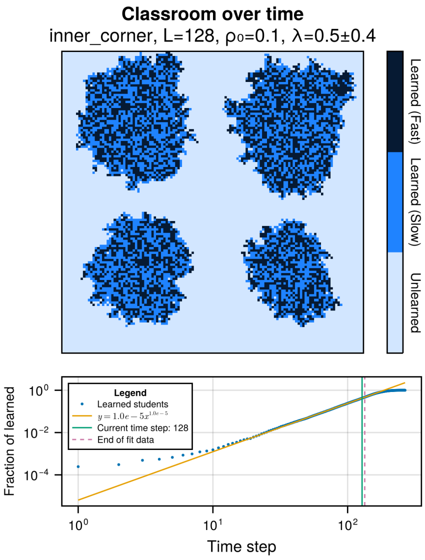
When introducing heterogeneity to PI setups, we see that the general trend of the classroom evolution is still the same as the homogenous case. As shown in the figures in this section, we notice values of ρ_0 is still the most important factor in determining class performance. However, irregularities in the shapes of the "wave of learning" become more pronounced for lower values of ρ_0 and high values of $\delta\lambda$ as shown in Figure 3.1. These irregularities are also more pronounced at the start of the classroom and fades over time. For high values of $\delta\lambda$, low values of ρ_0 leads to the irregularity in the "wave of learning" to persist longer than those with higher values of ρ_0. The irregularities in the shape of the "wave of learning" and the deviation from the power law are less pronounced for higher values of ρ_0, regardless of the value of $\delta\lambda$, as shown in Figures 3.2 and 3.3. In these cases, the class learning rate is slower than the power law fit at the start of the simulation.



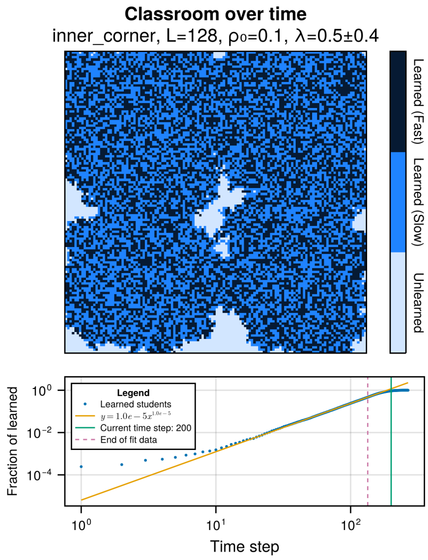
(a) $t = 48$



(b) $t = 69$

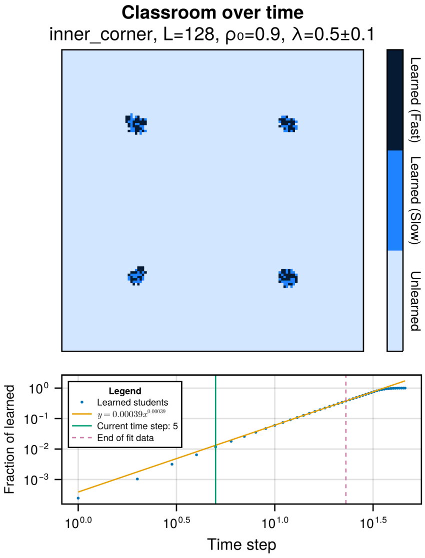


(c) $t = 128$

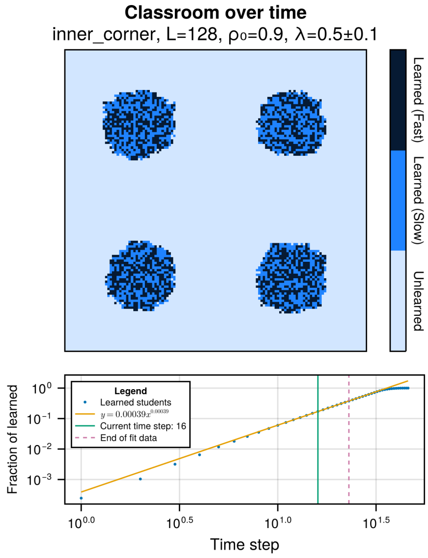


(d) $t = 200$

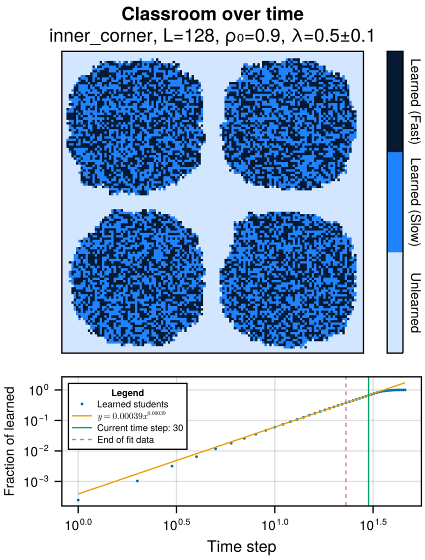
Figure 3.1: Sample classroom evolutions for PI with the inner corner SA with $L = 128$ at different times t for positional learning coefficient $\rho_0 = 0.1$, $\delta\lambda = 0.4$. Dark blue squares represent learned students with learning rate $\lambda = \lambda_0 + \delta\lambda$, blue squares represent learned students with learning rate $\lambda = \lambda_0 - \delta\lambda$, and light blue squares represent unlearned students. The accompanying graph shows the fraction of learned students as a function time step. The blue dots represent data points. The yellow line shows the power law fit. The pink dashed vertical line shows where we truncate the data for fitting the power law. The green vertical line shows the current time step in the simulation.



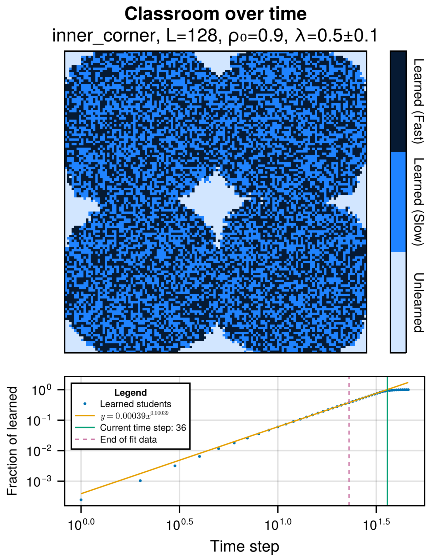
(a) $t = 5$



(b) $t = 16$

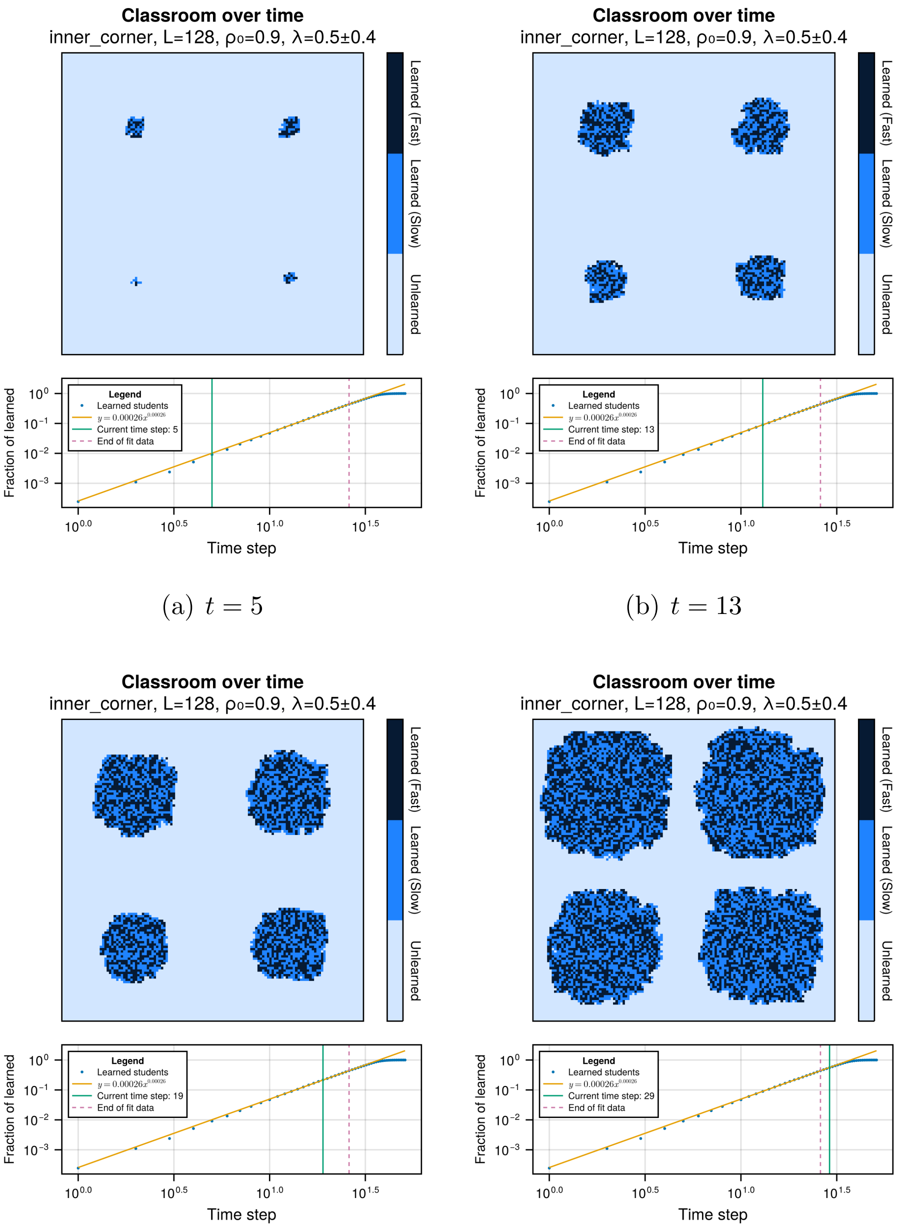


(c) $t = 30$



(d) $t = 36$

Figure 3.2: Sample classroom evolutions for PI with the inner corner SA with $L = 128$ at different times t for positional learning coefficient $\rho_0 = 0.9$, $\delta\lambda = 0.1$. Dark blue squares represent learned students with learning rate $\lambda = \lambda_0 + \delta\lambda$, blue squares represent learned students with learning rate $\lambda = \lambda_0 - \delta\lambda$, and light blue squares represent unlearned students. The accompanying graph shows the fraction of learned students as a function time step. The blue dots represent data points. The yellow line shows the power law fit. The pink dashed vertical line shows where we truncate the data for fitting the power law. The green vertical line shows the current time step in the simulation.



(a) $t = 5$

(b) $t = 13$

(c) $t = 19$

(d) $t = 29$

Figure 3.3: Sample classroom evolutions for PI with the inner corner SA with $L = 128$ at different times t for positional learning rate $\rho_0 = 0.9$, $\delta\lambda = 0.4$. Dark blue squares represent learned students with learning rate $\lambda = \lambda_0 + \delta\lambda$, blue squares represent learned students with learning rate $\lambda = \lambda_0 - \delta\lambda$, and light blue squares represent unlearned students. The accompanying graph shows the fraction of learned students as a function time step. The blue dots represent data points. The yellow line shows the power law fit. The pink dashed vertical line shows where we truncate the data for fitting the power law. The green vertical line shows the current time step in the simulation.

3.1.2 Traditional Instruction

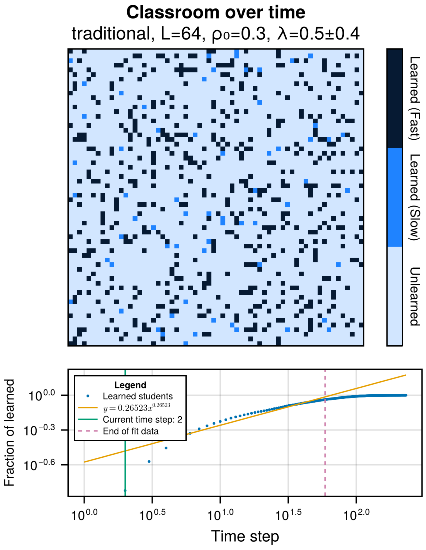
For traditional instruction, even though the value of positional learning coefficient ρ_0 is still an important factor in determining class performance, the effect of $\delta\lambda$ is more pronounced compared to PI. The trend we see in traditional instruction that is absent or less evident in PI, is that majority of the students that learn earlier in the simulations are fast learning students (Figure 3.4(a)). After this period, most of the simulation time is spent waiting for the slow learning students to learn. This behavior is more pronounced for higher values of $\delta\lambda$ and lower values of ρ_0 as shown in Figure 3.4. This could explain the deviation from the power law fit at earlier times t for the traditional model considering that the deviation shows that the class learning rate is higher than the fitted power law.

3.1.3 Comparing temporal learning dynamics of different classroom configurations

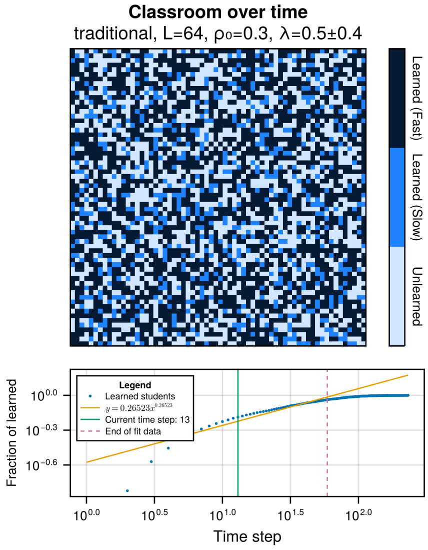
When varying the class size N, as in Figure 3.5(a), we see that in traditional instruction, the dynamics remains generally unchanged, with only the time to learn t_{max} increasing as the class size N increases. For PI, an increase in class size N adds a time delay to when the learning starts to speed up. Despite the added time delay, the general shape of the learning curve remains the same. An increase in class size N also does not affect the time to learn t_{max} for PI as much as it does for the traditional model.

When varying the positional learning factor ρ_0, we see that it varies the shape of the learning curve for traditional instruction, especially for lower values of ρ_0 as shown in Figure 3.5(b). For traditional instruction, the value of ρ_0 also greatly changes the initial number of students that learn in the second time step. The same figure shows that for PI, the value of ρ_0 has a similar effect to class size N, adding only a delay before learning accelerates. Notably, the effect of ρ_0 is nonlinear for either instruction models. We can see in the same figure that those with extremely low values of ρ, like that of $\rho_0 = 0.1$ perform much worse with the learning curves being further from $\rho_0 = 0.5$ than $\rho_0 = 0.9$.

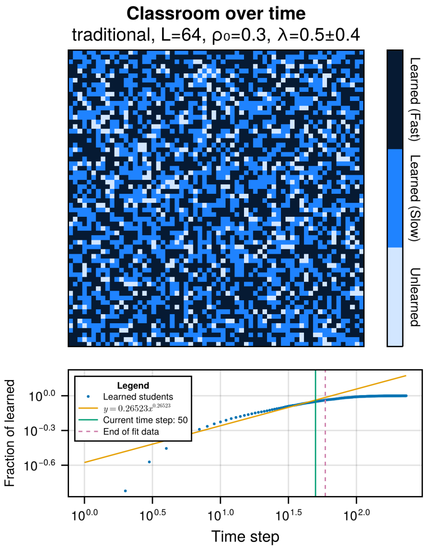
Figure 3.5(c) shows that an increase in heterogeneity $\delta\lambda$ changes how fast the slope of the learning curve changes for the traditional model without changing the initial number of learned students. As for PI, only very high $\delta\lambda$ values noticeably



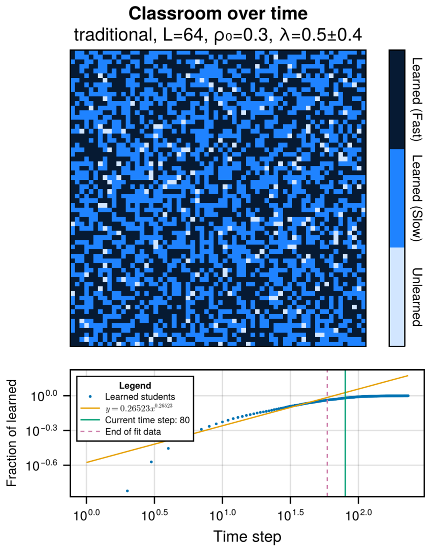
(a) $t = 2$



(b) $t = 13$



(c) $t = 50$



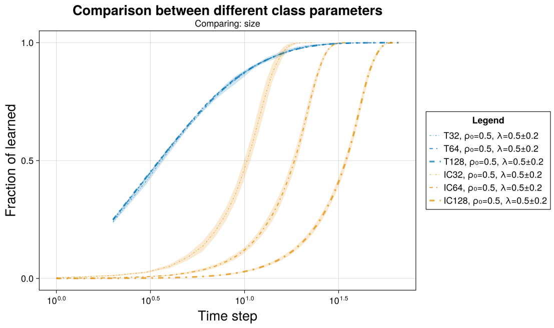
(d) $t = 80$

Figure 3.4: Sample classroom evolutions for PI with the inner corner SA with $L = 128$ at different times t for positional learning coefficient $\rho_0 = 0.9$, $\delta\lambda = 0.4$. Dark blue squares represent learned students with learning rate $\lambda = \lambda_0 + \delta\lambda$, blue squares represent learned students with learning rate $\lambda = \lambda_0 - \delta\lambda$, and light blue squares represent unlearned students. The accompanying graph shows the fraction of learned students as a function time step. The blue dots represent data points. The yellow line shows the power law fit. The pink dashed vertical line shows where we truncate the data for fitting the power law. The green vertical line shows the current time step in the simulation.

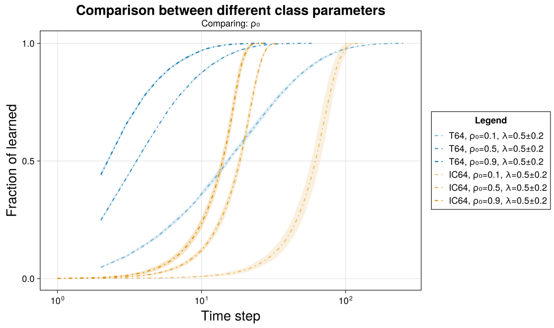
impact the learning curve. The effect of heterogeneity for PI is also similar to the effects of class size N and positional learning factor ρ_0 where it only adds a time delay, shifting the learning curve to the right.

As shown in Figure 3.5(d), when comparing between the different SA's for PI and traditional instruction, at least for $L = 64, \rho_0 = 0.5, \lambda = 0.5 \pm 0.2$, traditional instruction generally performs better than PI, especially in the early time steps. Among the different SA's for PI, the inner corner SA still performs the best while the center and outer corner SAs perform similarly, the same conclusion we got from the homogenous case. Contrary to the results of the homogenous classes, the random SA performed better than the center and outer corner SAs.

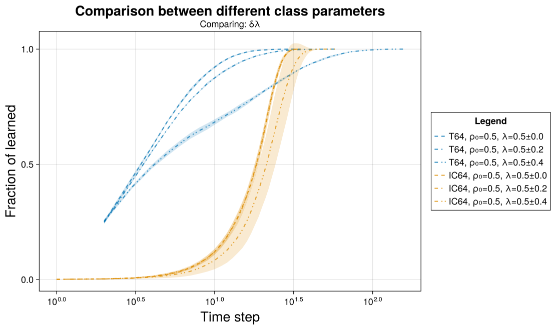
One observation that is consistent with all the comparisons shown in Figure 3.5 is that even though using traditional instruction has more students learning in the early time steps, it also spends more time waiting for the last few students to learn. In contrast, PI setups take some time for learning to accelerate, but once the majority learn, the remaining students quickly follow regardless of their individual learning rate. This makes PI have a shorter time to learn t_{max}, despite being outperformed by the traditional model in early time steps. This phenomenon is further investigated in the next sections where we look closer at the different factors that can affect which set up is better for different classroom parameters.



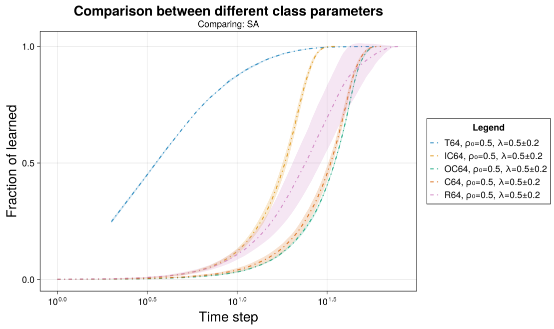
(a) Size comparison



(b) ρ_0 comparison



(c) $\delta\lambda$ comparison



(d) SA comparison

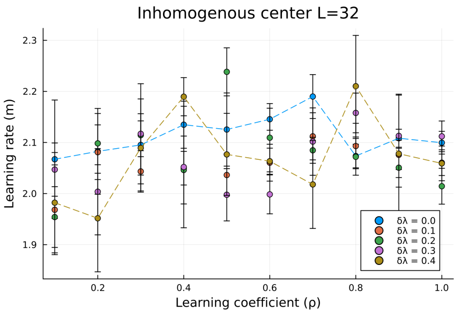
Figure 3.5: Comparison of time to learn t_{max} and fraction of learned students for different representative classroom configurations. Each SA corresponds to a different color, blue for traditional, yellow for inner corner, green for outer corner, orange for center, and pink for random. Different values ρ_0 corresponds to varying alpha or transparency, where lower ρ_0 values are more transparent. Different values of $\delta\lambda$ corresponds to different line styles, where $\delta\lambda = 0.0$ are represented by dashed lines, $\delta\lambda = 0.2$ are represented by lines with alternating dots and dashes, $\delta\lambda = 0.4$ are represented by two dots followed by a dash. Different classroom sizes L corresponds to different line widths, where bigger classroom sizes correspond to thicker lines. The bands around each line show the standard deviation of the data over 5 trials. Higher fraction of learned indicates better performance.

3.2 Class learning rate m vs positional learning factor ρ_0

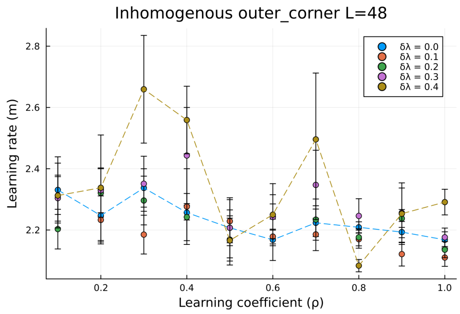
Figure 3.6 shows that class learning rate becomes inconsistent with the positional learning factor ρ_0 in PI models when introducing heterogeneity. The inconsistency may be caused by sampling problem where the first 50% of the data may not be just capturing the initial dynamics of the classroom, which should be the only basis of the class learning rate m. With heterogeneity, PI models no longer show trends in learning rate m as a function of positional learning factor ρ_0.

The traditional instruction model shows a similar trend in learning rate m as with time to learn t_{max} where homogenous classrooms perform better than heterogeneous classrooms. This might be explained by the spatio-temporal dynamics of traditional learning models where the fast learning students learn first, and the slow learning students learn last, as discussed in Section 3.1 and shown in Figure 3.4.

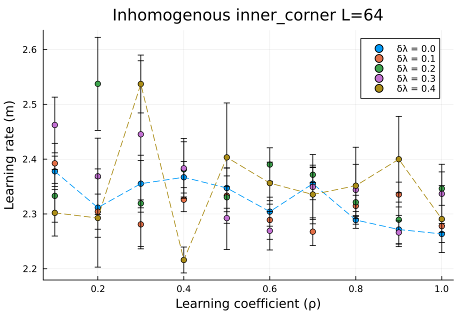
Moving forward, we will only focus on the time to learn t_{max} as a metric for performance.



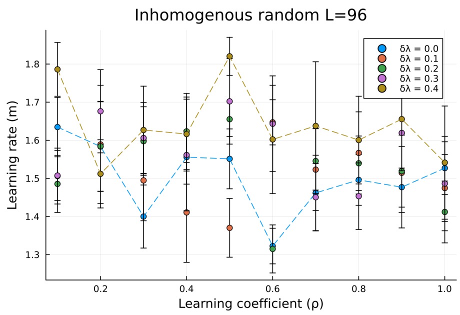
(a) Center $L = 32$



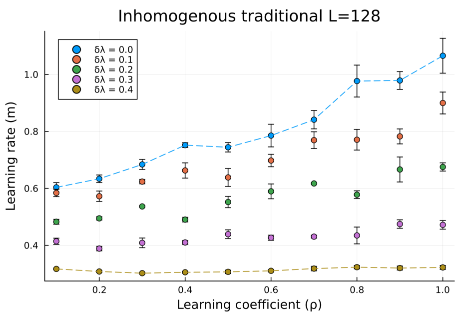
(b) Outer corner $L = 48$



(c) Inner corner $L = 64$



(d) Random $L = 96$



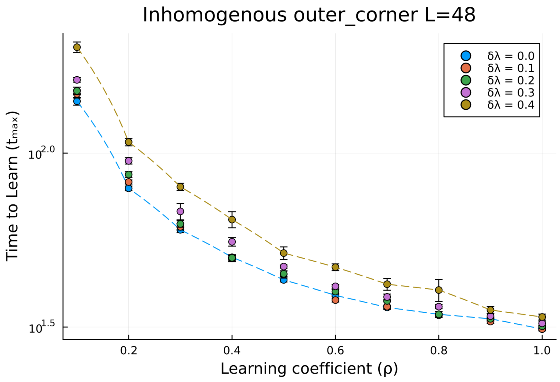
(e) Traditional $L = 128$

Figure 3.6: Representative plots for class learning rate m as a function of positional learning coefficient ρ_0. Each subplot represents a different SA and classroom size L. In each plot, the circles represent the data points, color represents a different value of $\delta\lambda \in \{0.0, 0.1, 0.2, 0.3, 0.4\}$, dashed lines connect the data points for $\delta\lambda \in \{0.0, 0.4\}$. Higher class learning rate m values indicate better performance.

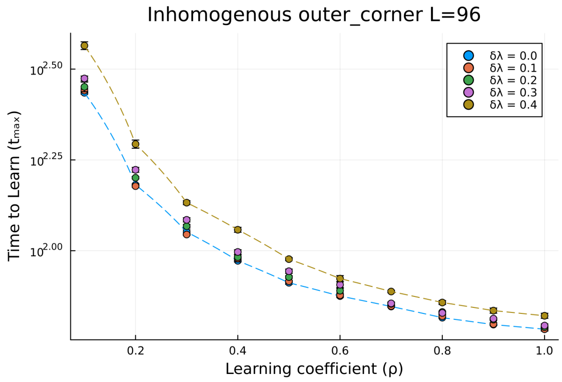
3.3 Time to learn t_{max} vs positional learning factor ρ_0

Figure 3.7 shows the time to learn t_{max} is directly affected by the level of heterogeneity $\delta\lambda$. This means that the heterogeneity $\delta\lambda$ has a significant effect on the time to learn t_{max}, with lower values of $\delta\lambda$ leading to lower time to learn t_{max} for all values of ρ_0. We also notice that the traditional model is more affected by classroom heterogeneity compared to PI models.

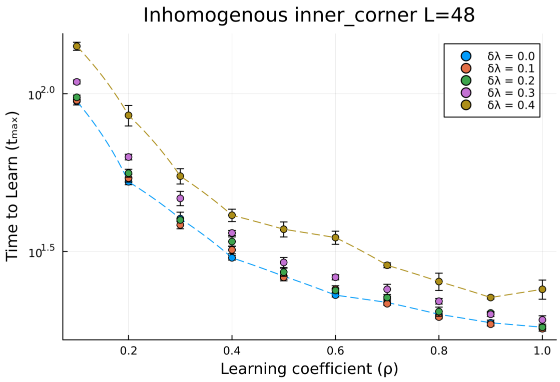
When aggregating the different results for each class size L, as shown in Figure 3.8, it affirms our previous findings that PI models are better for smaller classes and lower values of ρ_0. This stays consistent when comparing the performance of a similar classroom size L and heterogeneity $\delta\lambda$ using the traditional model. However, the comparison is not as clear when comparing between different levels of heterogeneity $\delta\lambda$. When $L \geq 64$, the PI models can perform better than the traditional model depending on the heterogeneity $\delta\lambda$, even with the same positional learning factor ρ_0.



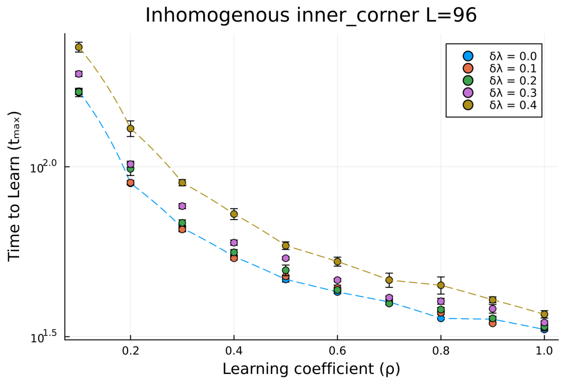
(a) Outer corner $L = 48$



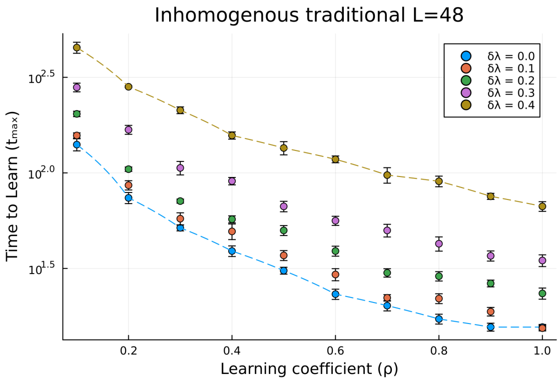
(b) Outer corner $L = 96$



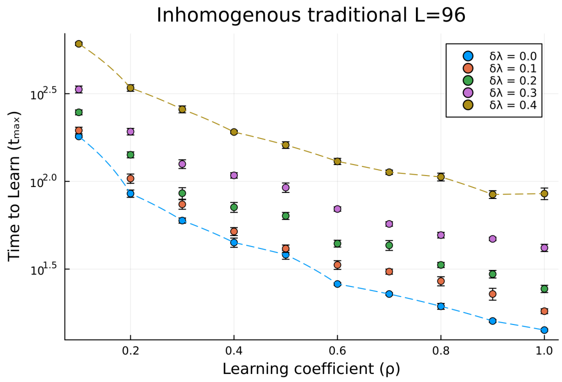
(c) Inner corner $L = 48$



(d) Inner corner $L = 96$

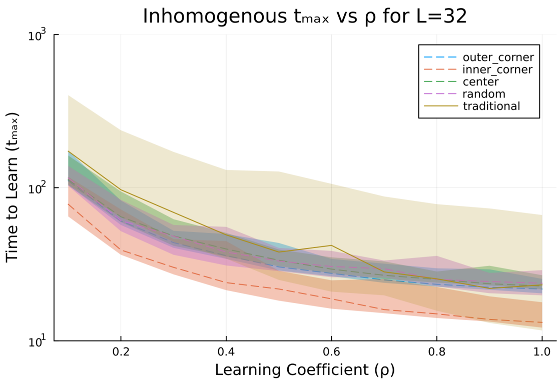


(e) Traditional $L = 48$

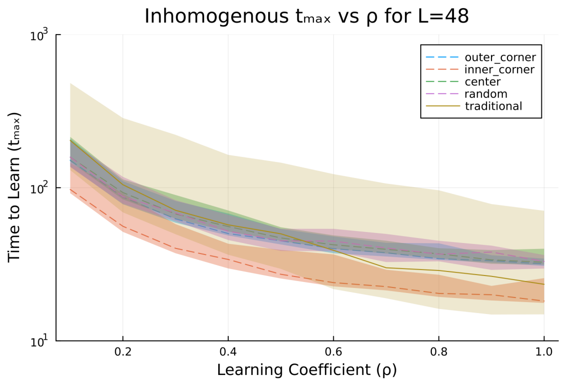


(f) Traditional $L = 96$

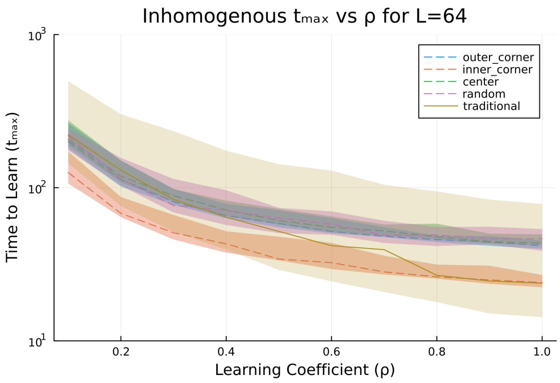
Figure 3.7: Time to learn t_{max} as a function of positional learning factor ρ_0 for different representative classroom configurations and sizes with varying heterogeneity $\delta\lambda \in \{0, 0.1, 0.2, 0.3, 0.4\}$. Lower time to learn t_{max} indicates better performance.



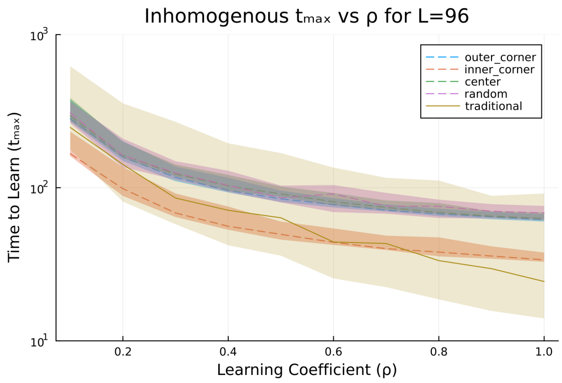
(a) $L = 32$



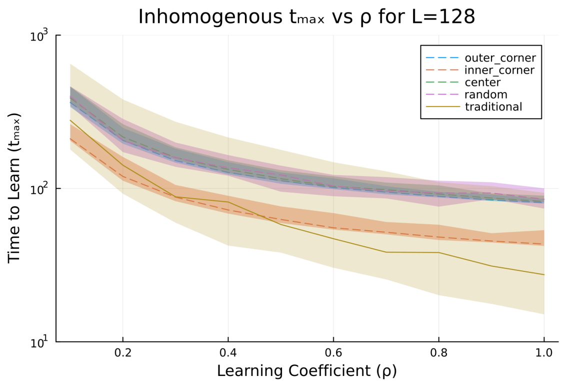
(b) $L = 48$



(c) $L = 64$



(d) $L = 96$



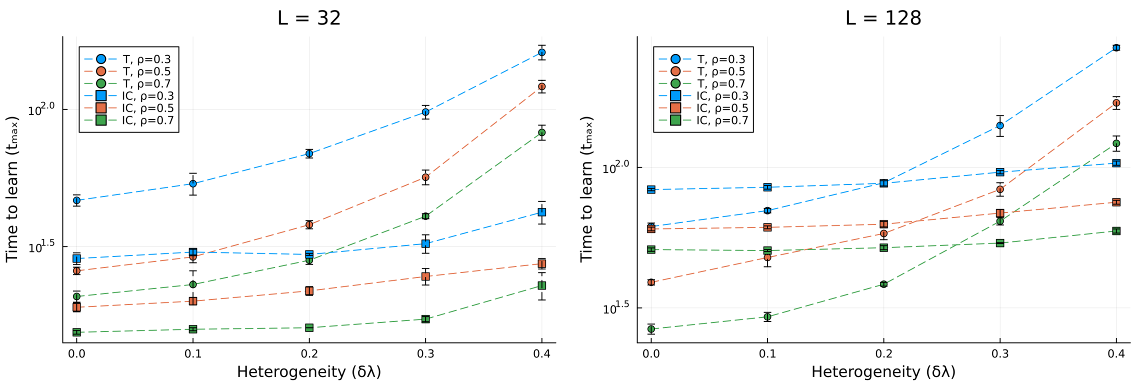
(e) $L = 128$

Figure 3.8: Time to learn t_{max} as a function of positional learning factor ρ_0 for the heterogeneous model of all seating arrangements and classroom sizes L. Each ribbon series represent the range of values for t_{max} for a given seating arrangements with heterogeneity $\delta\lambda = \{0.0, 0.1, 0.2, 0.3, 0.4\}$. Lower time to learn t_{max} indicates better performance.

3.4 Time to learn t_{max} vs heterogeneity $\delta\lambda$

Figure 3.9 shows that as heterogeneity $\delta\lambda$ increases, PI methods perform better than traditional methods even when PI methods have lower positional learning factors ρ_0. For our representative ρ_0 values, PI performs better than traditional methods regardless of the ρ_0 value, even in a large classroom $L = 128$. However, in larger classrooms, the advantage in performance of PI over traditional methods is less pronounced. Furthermore, as class size increases, traditional methods tend to stay advantageous than PI methods for increasing values of heterogeneity $\delta\lambda$.

For example, at $L = 32$ shown in figure 3.9(a), PI methods are better than traditional methods for all values of $\delta\lambda$ when comparing between equal ρ_0 values. However, at $L = 128$ shown in figure 3.9(b), the traditional model performs better than the PI model for $0 \leq \delta\lambda \leq 0.2$ when comparing between equal ρ_0 values.

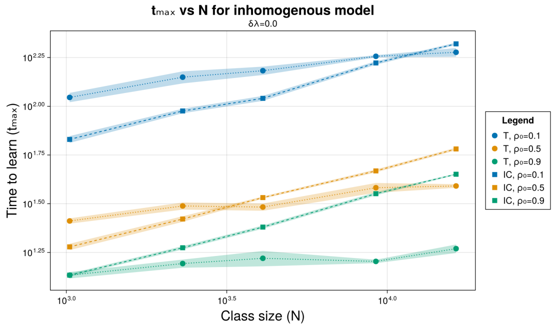


(a) $L = 32$ (b) $L = 128$

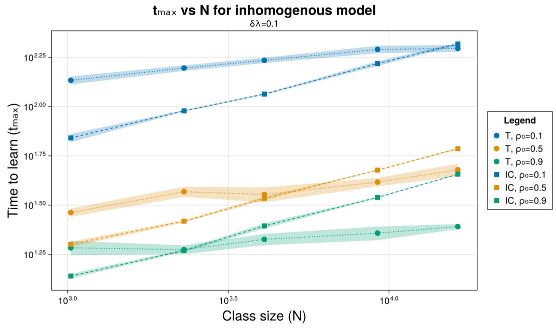
Figure 3.9: Time to learn t_{max} as a function of heterogeneity $\delta\lambda$ for the heterogeneous models of the PI (inner corner SA) and traditional models with varying positional learning factor $\rho_0 \in \{0.3, 0.5, 0.7\}$ and classroom sizes $L \in \{32, 128\}$. Each color represents a different value of ρ_0, while the circle and square symbols represent the traditional and PI models respectively. Lower time to learn t_{max} indicates better performance.

3.5 Time to learn t_{max} vs class size N

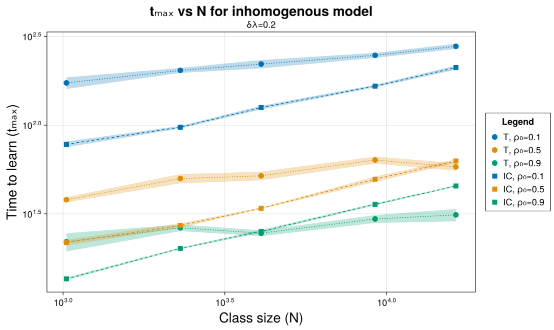
When investigating the size dependence of time to learn t_{max} shown in Figure 3.10, we see a similar trend to what we found in section 2.3.3 where PI models are better for smaller classes and traditional models are better for larger classes. Heterogeneity $\delta\lambda$ generally has a negative effect on the time to learn t_{max} for both PI and traditional models. However, as shown in sections 3.3 and 3.4, the traditional model is more affected by heterogeneity $\delta\lambda$ compared to PI models. Because of the sensitivity of traditional instruction to heterogeneity $\delta\lambda$, at low values of ρ_0, even for large classes, PI models can perform better than traditional models for higher values of heterogeneity $\delta\lambda$. This is shown in Figures 3.10(d) and 3.10(e) where the PI model performs better than the traditional model for all class sizes (N) and the same value of ρ_0.



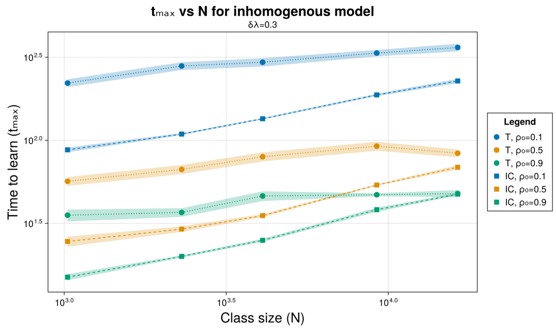
(a) $\delta\lambda = 0.0$



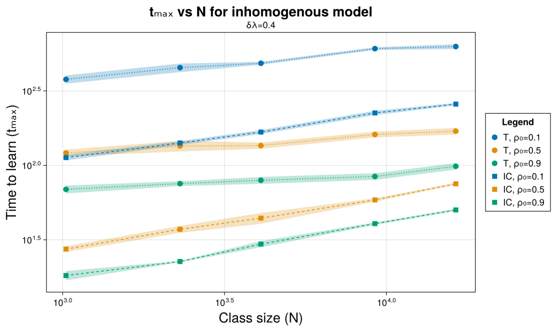
(b) $\delta\lambda = 0.1$



(c) $\delta\lambda = 0.2$



(d) $\delta\lambda = 0.3$



(e) $\delta\lambda = 0.4$

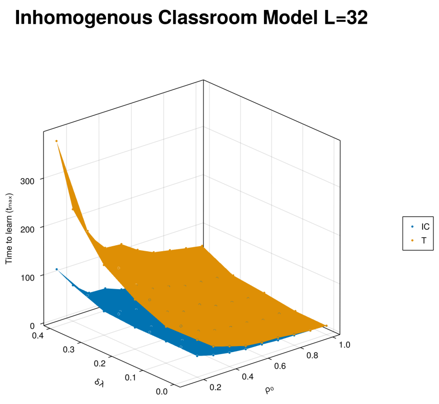
Figure 3.10: Time to learn t_{max} as a function of class size N for the heterogeneous models of the PI (inner corner SA) and traditional models with varying positional learning factor $\rho_0 \in \{0.1, 0.5, 0.9\}$ and heterogeneity $\delta\lambda \in \{0.0, 0.1, 0.2, 0.3, 0.4\}$. Each color represents a different value of ρ_0, while the circle and square symbols represent the traditional and PI models respectively. Each band represents the standard deviation of time to learn t_{max} over 5 trials. Lower time to learn t_{max} indicates better performance.

3.6 An overview of the different parameters that affect time to learn t_{max}

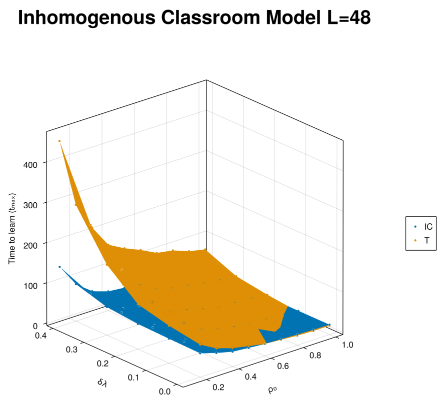
Figure 3.11 gives us a better visualization of the relationships between the positional learning factor ρ_0, heterogeneity $\delta\lambda$, and time to learn t_{max} for the PI and traditional models. In addition to our previous findings in Sections 3.3 and 3.4, we see that PI is the safe choice between the two instruction modes regardless of the class conditions. When traditional instruction is better than PI, there is not much of a difference in the time to learn t_{max}. However, when PI performs better than traditional instruction, there are cases where the time to learn t_{max} is significantly lower for PI compared to traditional instruction.

We also see that as $N = L^2$ increases, the trends for the time to learn t_{max} remain consistent, with the biggest difference being the shifting of the blue surface for PI upwards.

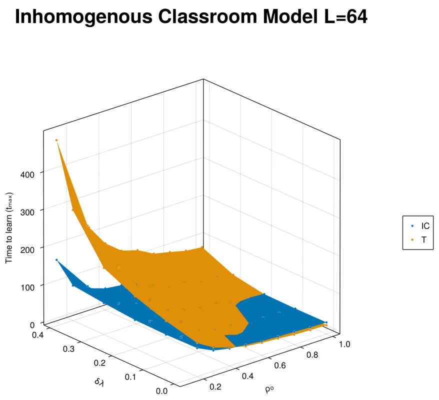
It should be noted that while the z-axis of the plots shown in Figure 3.11 do not have the same limits, the grid lines remain consistent having the same interval of 100 time steps between each grid line.



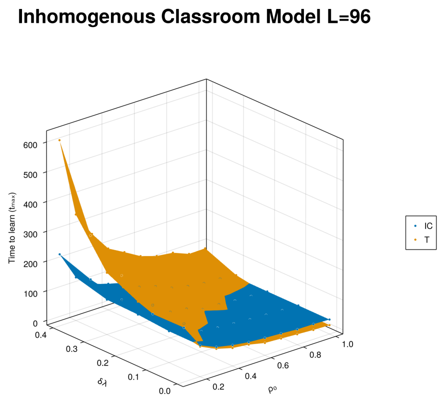
(a) $L = 32$



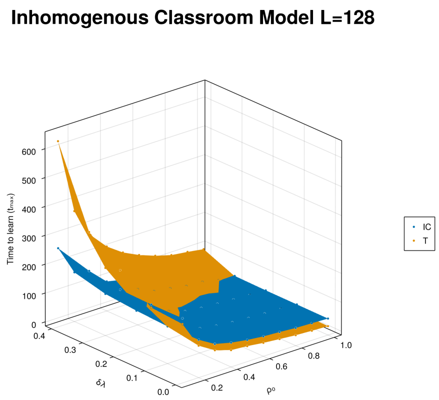
(b) $L = 48$



(c) $L = 64$



(d) $L = 96$



(e) $L = 128$

Figure 3.11: Time to learn t_{max} as a function of positional learning factor ρ_0 and heterogeneity $\delta\lambda$ for the heterogeneous models of the PI (inner corner SA) and traditional models with varying positional learning factor. The blue surface represents PI and the orange surface represents traditional instruction. Lower time to learn t_{max} indicates better performance.

3.7 Conclusions

To better model the real world, we introduced heterogeneity to the students' learning rate. Adding heterogeneity does not affect the general trend of our instruction models. For the case of a heterogeneous system, we still find traditional instruction fares better than PI for larger classes with high positional learning factor rho_0.

Besides not affecting the general trend, we found that traditional instuction is more sensitive to heterogeneity. As discussed in section 3.1, the difference in sensitivity is because of the dynamics that were seen in the classes' evolution. This is related to the spatio-temporal dynamics of traditional learning models where the fast learning students learn first and the slow learning students learn last. The time to learn t_{max} for traditional instruction is then dependent on waiting for the slow students.

In contrast to the spatio-temporal dynamics of traditional instruction, the advantage of PI is that it is not as heavily dependent on the students themselves as traditional instruction. PI is more dependent on the geometry of the classroom and the number of learning sources available to the students. For the case of PI, the "wavefront of learning" (see: section 3.1) provides students with more than one source of learning, unlike in traditional instruction. This phenomenon offsets the increase in the time to learn t_{max} that is caused by waiting for the slow learning students. PI models therefore have a shorter time to learn t_{max}, even though the traditional model initially performs better.

Given the foregoing, initially performing traditional instruction and then proceeding with PI later would maximize each method's strengths. The traditional instruction phase would yield high learning rates at the start of the simulation, by allowing the fast learning students to learn first. The learned students would then be able to help the slow learning students learn in the PI phase, shortening the time to learn t_{max}.

As a secondary strategy, if one has to choose to implement only one instruction method, PI is the safe option. Despite not always being the most optimal, the performance advantage of traditional over PI is not as pronounced as the performance advantage of PI over traditional.

Chapter 4

Conclusion and Recommendations

We proposed a probabilistic cellular automata model as a new way of investigating classroom learning dynamics. We are able to investigate the effects of different factors governing two very different methods of instruction - traditional instruction and peer instruction (PI).

We found traditional instruction to be more scalable and less dependent on class size than PI. On the one hand, the teacher is the only source of information in traditional instruction and is also able to teach all the students at the same time. On the other hand, PI is more efficient in smaller classes because students have more sources of information (i.e. their seatmates). This offsets the simultaneous teaching advantage of traditional instruction.

We investigated different seating arrangements (SA) for PI and found that the inner corner SA is the most efficient setup from all SAs considered. This optimal SA differs from a previous study where the outer corner SA is found to be the most efficient [12]. The current model used in this work does not incorporate the effect of aptitude similarity [9] between students in PI and isotropy in the learning direction.

To improve our model in reflecting real world factors, we also introduced heterogeneity in students' learning rates. In doing this, it further emphasized one of the advantages of PI over traditional instruction - that PI is less affected by heterogeneity in students' learning rates. In traditional instruction, the students with fast learning rates learn in the first few time steps, while the rest of the time it takes for the class to learn is waiting for the slow learners to learn. This causes a two-stage learning process in traditional instruction. The effect of heterogeneity in PI is less pronounced, only affecting the shape of the "wave of learning" but generally leaving the relevant trends and metrics unaffected. These dynamics lead to PI performing better than traditional instruction in cases where class heterogeneity is high. Additionally, we found that classes with slower learners, due to either the assigned positional learning factor ρ_0 or the learning rate λ, generally do better in PI than in traditional instruction.

Most previous studies focused on students' pre- and post-test scores to assess the advantages of PI over traditional instruction. That was not something we could do in this study because of the binary-state nature of our model. However, we have confirmed results of experimental studies done before — that PI can perform either similarly or better than traditional instruction [6–9]. We also found similar results regarding class heterogeneity — that both methods of instruction are more effective with homogenous classrooms [12].

It is important to note that our model does not fully model the process of PI in real-world scenarios since it only models part of the PI process where students get to discuss the concept test questions with their seatmates. We recommend that future researchers who want to delve deeper into this topic consider the effect of aptitude similarity in PI as described by Smith et al. (2009) [9]. Future works could also consider transitioning to a continuous-state model to better reflect real world situations. Considering that our study suggests that the optimal method of instruction would be to have a short segment of traditional instruction followed by PI, which is similar to the actual process of PI in the classroom (i.e. reading and short lecture, concept test, sharing), future researchers who plan to model the classroom learning similarly should investigate the time delay between the two methods of instruction.

Appendix A

Appendix

A.1 Derivation of Equation 2.3

For any event e, the desired outcome occurs with probability p or not with probability q where

$$p + q = 1. \tag{A.1a}$$

For n events, each being event e:

$$\prod_{\forall e}(p_e + q_e) = 1. \tag{A.1b}$$

Expanding equation A.1b, we get

$$\prod_{\forall e} p_e + \ldots + \prod_{\forall e} q_e = 1. \tag{A.1c}$$

where the sum of the first $n - 1$ terms is the probability of the desired outcome occurring at least once over n events and the last term is the of the probability of the desired outcome not occurring at all. Thus, we can rewrite the probability of the desired outcome occurring at least once as

$$P = 1 - \prod_{\forall e} q_e. \tag{A.1d}$$

substituting equation A.1a into equation A.1d, we get

$$P = 1 - \prod_{\forall e}(1 - p_e). \tag{A.1e}$$

A.2 Julia packages used

This study used the following `Julia` packages and their hash codes:

1. `Alert = "28312eec-4d86-447d-83ad-bc2b262de792"`
2. `CSV = "336ed68f-0bac-5ca0-87d4-7b16caf5d00b"`
3. `CairoMakie = "13f3f980-e62b-5c42-98c6-ff1f3baf88f0"` [21]
4. `ColorSchemes = "35d6a980-a343-548e-a6ea-1d62b119f2f4"`
5. `DataFrames = "a93c6f00-e57d-5684-b7b6-d8193f3e46c0"` [22]
6. `GLMakie = "e9467ef8-e4e7-5192-8a1a-b1aee30e663a"` [21]
7. `LaTeXStrings = "b964fa9f-0449-5b57-a5c2-d3ea65f4040f"`
8. `LsqFit = "2fda8390-95c7-5789-9bda-21331edee243"`
9. `Measurements = "eff96d63-e80a-5855-80a2-b1b0885c5ab7"` [23]
10. `Plots = "91a5bcdd-55d7-5caf-9e0b-520d859cae80"` [24]
11. `ProgressMeter = "92933f4c-e287-5a05-a399-4b506db050ca"`

A.3 GitHub repository

The GitHub repository containing the codes and datasets used for the study can be found here: https://github.com/ioakimsy/BS-Thesis

Bibliography

[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*, vol. 67, pp. 359–360, 04 1999.

[3] R. T. Johnson and D. W. Johnson, "Active learning: Cooperation in the classroom," *The annual report of educational psychology in Japan*, vol. 47, pp. 29–30, 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*, vol. 40, no. 4, pp. 206–209, 2002.

[6] C. H. Crouch and E. Mazur, "Peer instruction: Ten years of experience and results," *American journal of physics*, vol. 69, no. 9, pp. 970–977, 2001.

[7] N. Lasry, E. Mazur, and J. Watkins, "Peer instruction: From harvard to the two-year college," *American journal of Physics*, vol. 76, no. 11, pp. 1066–1069, 2008.

[8] 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*, vol. 62, no. 7, pp. 627–633, 1994.

[9] 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*, vol. 323, no. 5910, pp. 122–124, 2009.

[10] V. P. Coletta and J. A. Phillips, "Interpreting fci scores: Normalized gain, pre-instruction scores, and scientific reasoning ability," *American journal of physics*, vol. 73, no. 12, pp. 1172–1182, 2005.

[11] S. Tobias, "They're not dumb. they're different.: A new "tier of talent" for science," *Change: The Magazine of Higher Learning*, vol. 22, no. 4, pp. 11–30, 1990.

[12] R. Roxas, S. Carreon-Monterola, and C. Monterola, "Seating arrangement, group composition and competition-driven interaction: Effects on students' performance in physics," in *AIP Conference Proceedings*, vol. 1263, pp. 155–158, American Institute of Physics, 2010.

[13] 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*, vol. 66, pp. 64–74, 01 1998.

[14] H. Nitta, "A mathematical model of peer instruction and its applications," *Upgrading Physics Education to Meet the Needs of Society*, pp. 87–98, 2019.

[15] C. M. Bordogna and E. V. Albano, "Theoretical description of teaching-learning processes: A multidisciplinary approach," *Physical Review Letters*, vol. 87, no. 11, p. 118701, 2001.

[16] C. M. Bordogna and E. V. Albano, "Simulation of social processes: application to social learning," *Physica A: Statistical Mechanics and its Applications*, vol. 329, no. 1-2, pp. 281–286, 2003.

[17] 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*, vol. 4, no. 1, p. 010109, 2008.

[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*, vol. 2009, no. 07, p. P07040, 2009.

[19] R. X. Ramos, “Dynamics of a nonlinear cellular automata model of biological neurons and its applications to young and aged neurons,” Master’s thesis, University of the Philippines Diliman, 2023.

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

[21] S. Danisch and J. Krumbiegel, “Makie.jl: Flexible high-performance data visualization for Julia,” *Journal of Open Source Software*, vol. 6, no. 65, p. 3349, 2021.

[22] M. Bouchet-Valat and B. Kamiński, “Dataframes.jl: Flexible and fast tabular data in julia,” *Journal of Statistical Software*, vol. 107, no. 4, pp. 1–32, 2023.

[23] M. Giordano, “Uncertainty propagation with functionally correlated quantities,” *ArXiv e-prints*, Oct. 2016.

[24] S. Christ, D. Schwabeneder, C. Rackauckas, M. K. Borregaard, and T. Breloff, “Plots.jl – a user extendable plotting api for the julia programming language,” 2023.