Packing Density of Combinatorial Settlement Planning Models

Mate Puljiz ⬤, Stjepan Šebek ⬤, and Josip Žubrinić

Abstract. We consider a combinatorial settlement model on a rectangular grid where each house must be exposed to sunlight from east, south, or west. We are interested in maximal configurations, where no additional houses can be added. Once the settlement is completely built, it seems natural to consider the building density of the obtained maximal configuration. In this article we consider two different random models which produce maximal configurations and, using simulations, we plot an estimate of the distribution of the building density (actually, the occupancy—the total number of houses built) and we conjecture that the means of these distributions converge to a certain limit as the grid dimensions grow to infinity.

1. INTRODUCTION. Consider the following problem: a rectangular $m \times n$ tract of land, whose sides are oriented north–south and east–west as in Figure 1, consists of mn square lots of size 1×1. Each 1×1 square lot can be either empty, or occupied by a single house. A house is said to be *blocked from sunlight* if the three lots immediately to its east, west and south are all occupied (it is assumed that sunlight always comes from the south[1]). The tracts of land are not adjacent to any other buildings, i.e., along the boundary of the rectangular $m \times n$ grid, there are no obstructions to sunlight. We refer to the models of such rectangular tracts of land, with certain lots occupied, as configurations. Of interest are maximal configurations, where no house is blocked from the sunlight, and any further addition of a house to the configuration on any empty lot would result in either that house being blocked from the sunlight, or it would cut off sunlight from some previously built house, or both.
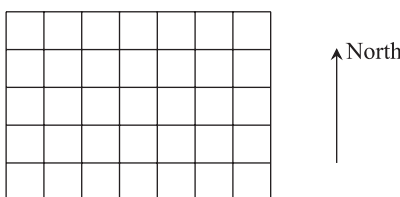

Figure 1. An example of a tract of land ($m = 5, n = 7$).

We can encode any fixed configuration as a 0-1, $m \times n$ matrix C, with $C_{i,j} = 1$ if and only if a house is built on the lot (i, j) (i-th row and j-th column, counted from the top left corner). We can, equivalently, think of C as a subset of $[m] \times [n] = \{(i, j) : 1 \le i \le m, 1 \le j \le n\}$, where, again, $(i, j) \in C$ if and only if a house is built on the lot (i, j).

It is natural to define the *building density* of a configuration C as $\dfrac{|C|}{mn}$, where

$$|C| = \sum_{i=1}^{m} \sum_{j=1}^{n} C_{i,j}$$

[1]Our Southern Hemisphere friends are welcome to turn the page upside down when inspecting the figures in our paper.

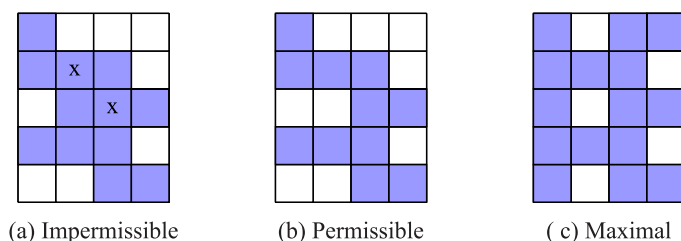(a) Impermissible   (b) Permissible   (c) Maximal

Figure 2. Examples of impermissible, permissible, and maximal configuration on a 5×4 tract of land.

is the total number of occupied lots in the configuration C, i.e., the cardinality of C when C is interpreted as a subset of $[m] \times [n]$. We also call $|C|$ the *occupancy* of C.

A configuration C is said to be *permissible* if no house in it is blocked from the sunlight, otherwise it is called *impermissible*.

A configuration C is said to be *maximal* if it is permissible and no other permissible configuration strictly contains it, i.e., no further houses can be added to it, whilst ensuring that all the houses still get some sunlight. See Figure 2 for examples of impermissible, permissible, and maximal configurations on a 5×4 tract of land. Shaded squares represent houses and unshaded squares represent empty lots on the tract of land. The houses that are blocked from the sunlight are marked with the letter "x".

We were introduced to this problem by Juraj Božić who came up with it during his studies at the Faculty of Architecture, University of Zagreb. His main goal was to design a model for settlement planning where the impact of the architect would be as small as possible and people would have a lot of freedom in the process of building the settlement. This minimal intervention from the side of the architect is given through the condition that houses are not allowed to be blocked from the sunlight and that the tracts of land on which the settlements are built are of rectangular shapes.

The model studied here also appeared in [**9**]. In that paper, the authors study the set of all maximal configurations and derive the bounds on the possible building densities that can be attained by these maximal configurations. The key results from that paper are discussed later in the text.

In the present paper, we aim to study two stochastic versions of this model and infer the distribution of some related statistics. In particular, we are interested in the distribution of building density of the random maximal configuration. The first of the two models we introduce in Section 2 samples a maximal configuration uniformly at random from the set of all maximal configurations on a grid of fixed size. The other, perhaps more natural, sampling is related to the so-called random sequential adsorption (RSA). On a grid of fixed size, plots are sampled sequentially and the houses are built if they do not violate the model rules. This process is stopped when a maximal configuration is reached. It turns out that these two samplings from the set of maximal configurations produce different probability distributions over maximal configurations and the resulting distributions of building density are also different, see Claim 1. In Section 3 we sample from the latter stochastic model by simulating random sequential building of houses and then plot the estimated probability distribution of the building density for some small grids. Finally, in Section 4 we discuss RSA and related models. The notion of saturation coverage in RSA (also known elsewhere as packing fraction, packing density, or jamming limit) motivates us to study the asymptotics of the mean building density when the dimensions of the tract of land increase to infinity. Using simulations we estimate these packing densities for grids up to size $100{,}000 \times 100{,}000$, and conjecture that the limits exist. Their actual values, however, remain unknown to us.

2. MODELING OF A PROBLEM.

Let $\mathcal{M}_{m,n}$ denote the set of all maximal configurations on an $m \times n$ tract of land. We sample at random one maximal configuration C from $\mathcal{M}_{m,n}$. The main questions that we are interested in are:

- What is the distribution of the random variable that measures the building density of the sampled configuration C (i.e., what is the distribution of $(mn)^{-1} \cdot |C|$)?
- How does the expected value of this random variable behave when we let m or n or both (m and n) go to infinity?

Clearly, to be able to answer these questions, we need to specify precisely how we are sampling at random from the set $\mathcal{M}_{m,n}$. There are (at least) two natural ways of doing this:

Model (i): Sampling uniformly at random

In this model, we assign equal probabilities to all the maximal configurations in $\mathcal{M}_{m,n}$ and we sample one of them uniformly at random (see e.g., Figures 4–6 for illustrations of the set $\mathcal{M}_{m,n}$). Notice that we need to know all the maximal configurations to be able to analyze this model and this turns out to be quite involved. As previously mentioned, we are interested in the random variable measuring the building density of such a uniformly sampled maximal configuration. Denoting this random variable with $X_{m,n}^u$, we have

$$\mathbb{P}\left(X_{m,n}^u = \frac{k}{mn} \right) = \frac{|\{C \in \mathcal{M}_{m,n} : |C| = k\}|}{|\mathcal{M}_{m,n}|}, \qquad k \in \mathbb{N}.$$

Model (ii): Sequential building of houses

In this model, we build houses one-by-one. Let C_r denote the set of occupied lots after the r-th step of the algorithm. Let B_r denote the set of lots where we tried to build a house (in the first r steps), but were not able to, because this house, or some other already existent house, would then become blocked from the sunlight.

We start with $C_0 = B_0 = \emptyset$. In the first step, we sample uniformly at random one element (i, j) from the set $[m] \times [n]$, we build a house on the corresponding lot, we set $C_1 = \{(i, j)\}$, and still keep $B_1 = \emptyset$. In the r-th step, we again sample uniformly at random one element (i, j), but this time from the set $([m] \times [n]) \setminus (C_{r-1} \cup B_{r-1})$. If, after building a house on the sampled lot, our configuration stays permissible, we set $C_r = C_{r-1} \cup \{(i, j)\}$ and $B_r = B_{r-1}$. Otherwise, we set $C_r = C_{r-1}$ and $B_r = B_{r-1} \cup \{(i, j)\}$. In the end (after mn steps), we have $C_{mn} = C$, for some maximal configuration $C \in \mathcal{M}_{m,n}$, and $B_{mn} = C^c$.

Notice that sampling random lots in this fashion results in one random permutation of the elements of the set $[m] \times [n]$. Therefore we can assign a unique maximal configuration to any permutation of the set $[m] \times [n]$. Clearly, many different permutations will give us the same maximal configuration, see Figure 3. Let S_X be the set of all permutations of the elements of the set X (recall that $|S_X| = |X|!$). Denote by $G : S_{[m] \times [n]} \to \mathcal{M}_{m,n}$ the function that maps a permutation in $S_{[m] \times [n]}$ to the corresponding maximal configuration in $\mathcal{M}_{m,n}$, as explained above. In this model of the sequential building of houses, we do not sample maximal configurations uniformly at random, instead we sample permutations from $S_{[m] \times [n]}$ uniformly at random and then consider the corresponding maximal configurations. Hence denoting the random variable that measures the building density of a maximal configuration (that we got by random sampling a permutation $\sigma \in S_{[m] \times [n]}$) by $X_{m,n}^s$, we have

$$\mathbb{P}\left(X_{m,n}^s = \frac{k}{mn} \right) = \frac{|\{\sigma \in S_{[m] \times [n]} : |G(\sigma)| = k\}|}{(mn)!}, \qquad k \in \mathbb{N}.$$
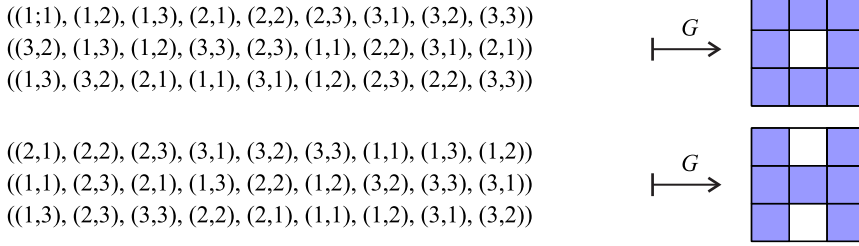
$((1;1), (1,2), (1,3), (2,1), (2,2), (2,3), (3,1), (3,2), (3,3))$
$((3,2), (1,3), (1,2), (3,3), (2,3), (1,1), (2,2), (3,1), (2,1))$
$((1,3), (3,2), (2,1), (1,1), (3,1), (1,2), (2,3), (2,2), (3,3))$

$((2,1), (2,2), (2,3), (3,1), (3,2), (3,3), (1,1), (1,3), (1,2))$
$((1,1), (2,3), (2,1), (1,3), (2,2), (1,2), (3,2), (3,3), (3,1))$
$((1,3), (2,3), (3,3), (2,2), (2,1), (1,1), (1,2), (3,1), (3,2))$

Figure 3. Acting of the function G on some permutations in $S_{[3]\times[3]}$.

(1) 129,600   (2) 25,920   (3) 25,920   (4) 25,920   (5) 25,920

(6) 25,920   (7) 25,920   (8) 25,920   (9) 25,920   (10) 25,920
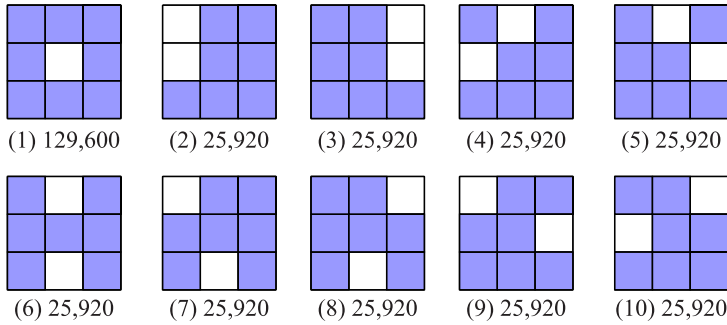
Figure 4. All the elements of $\mathcal{M}_{3,3}$ together with the number of permutations in $S_{[3]\times[3]}$ that are mapped to each by the function G.

A priori, it is not clear whether these two models are the same or not. It turns out that in general they are not the same. We illustrate this fact in the following claim.

Claim 1. $X_{3,3}^u \overset{(d)}{\neq} X_{3,3}^s.$[2]

Proof. Using the programming language R, we calculated the value of the function G for all $(3 \cdot 3)! = 362,880$ permutations in $S_{[3]\times[3]}$. In this way we obtained all the maximal configurations in $\mathcal{M}_{3,3}$ (these are illustrated in Figure 4) and, furthermore, we saw how many permutations are mapped to each of these maximal configurations. Even though there are 10 maximal configurations in $\mathcal{M}_{3,3}$, it turns out that for each of the 9 maximal configurations with 7 occupied lots, there are only 25,920 permutations that the function G maps to them, and there are 129,600 permutations that G maps to the maximal configuration with 8 occupied sites. Therefore

$$X_{3,3}^u \sim \begin{pmatrix} \frac{7}{9} & \frac{8}{9} \\ \frac{9}{10} & \frac{1}{10} \end{pmatrix}, \qquad X_{3,3}^s \sim \begin{pmatrix} \frac{7}{9} & \frac{8}{9} \\ \frac{9}{14} & \frac{5}{14} \end{pmatrix}.$$

Above we used the standard notation for discrete distributions where the numbers in the top row denote the possible values the random variable can attain with positive probability, and below are the corresponding probabilities. ∎

Using the same code in R, we found all the elements of sets $\mathcal{M}_{3,4}$ and $\mathcal{M}_{4,3}$, together with the number of permutations in $S_{[3]\times[4]}$ and $S_{[4]\times[3]}$, respectively, that are mapped to each of them (see Figures 5 and 6). Inspecting the results, we see that for

---

[2]We write $X \overset{(d)}{=} Y$ if two random variables X and Y are equal in distribution. Since we are dealing with discrete random variables, this is the same as requiring $\mathbb{P}(X = z) = \mathbb{P}(Y = z)$ for all $z \in \mathbb{R}$.
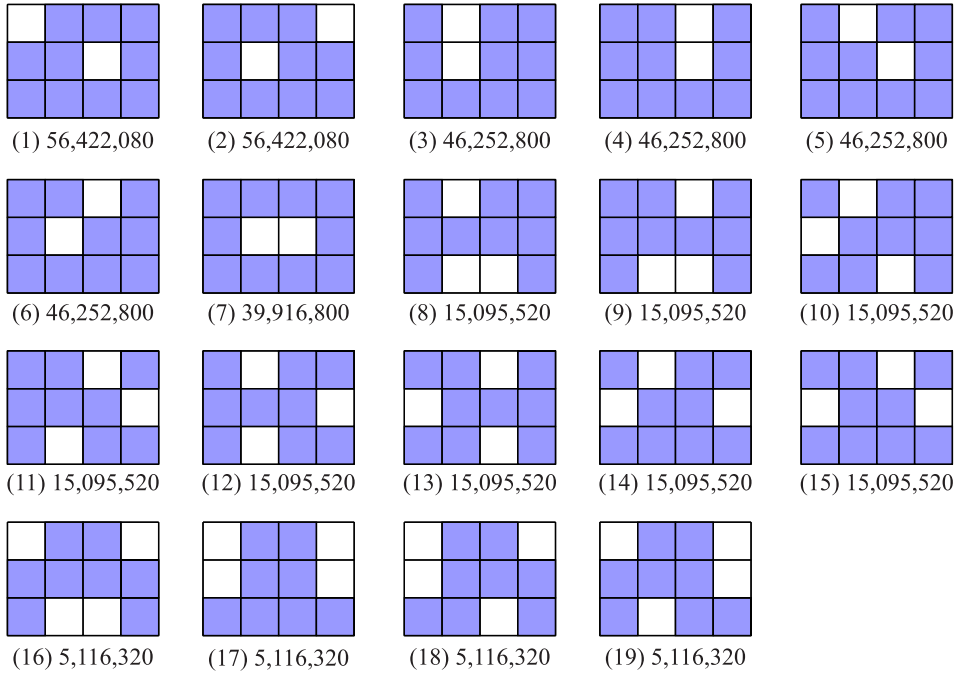
Figure 5. All the elements of $\mathcal{M}_{3,4}$ together with the number of permutations in $S_{[3]\times[4]}$ that are mapped to each by the function G.

these dimensions of the tract of land, we also have

$$X_{3,4}^u \overset{(d)}{\neq} X_{3,4}^s \qquad \text{and} \qquad X_{4,3}^u \overset{(d)}{\neq} X_{4,3}^s.$$

Furthermore, we can see that

$$X_{3,4}^u \overset{(d)}{\neq} X_{4,3}^u \qquad \text{and} \qquad X_{3,4}^s \overset{(d)}{\neq} X_{4,3}^s.$$

Not only do these variables not have the same distribution, but they also do not even have the same support. Therefore orientation of the tract of land is important. The reason for that is the difference in the role of north and south. Notice that east and west are symmetric since the houses can get sunlight from both of those sides, but north and south play very different roles since houses cannot get sunlight from the north.

In Figures 4–6 one can see how the set $\mathcal{M}_{m,n}$ looks for several different values of m and n. More precisely, we showed that $|\mathcal{M}_{3,3}| = 10$, $|\mathcal{M}_{3,4}| = 19$ and $|\mathcal{M}_{4,3}| = 25$. A natural question that arises is the following.

Question 2. *What is the cardinality of the set $\mathcal{M}_{m,n}$ (for $m, n \in \mathbb{N}$)?*

Maximal configurations (for a particular $m, n \in \mathbb{N}$) with the highest building density possible are called *efficient*, while those with the lowest building density possible are called *inefficient*. Another interesting thing that can be observed from the previous examples is that among all the maximal configurations, the ones that have the most permutations mapped to them are some of the efficient ones. Furthermore, maximal configurations that have the least permutations mapped to them are the inefficient ones. However, there are more maximal configurations that are not efficient than maximal configurations that are efficient. Therefore it can happen that more permutations
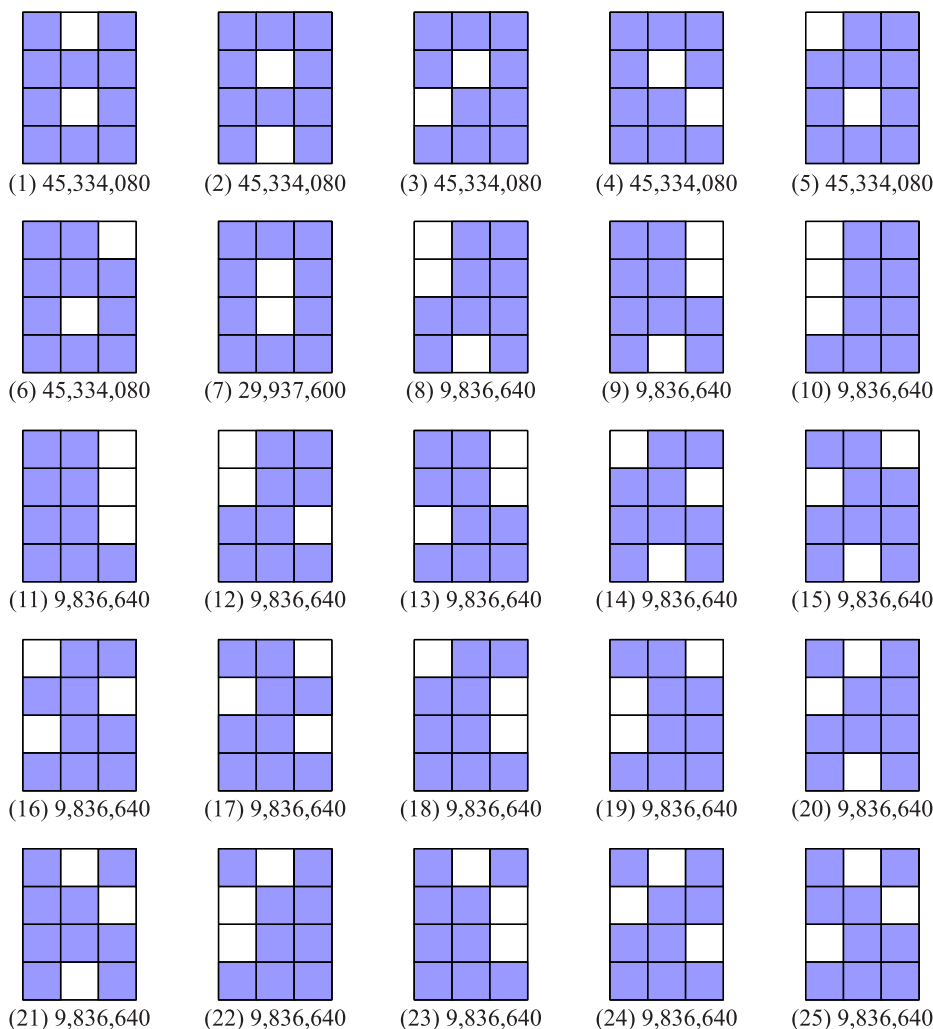
Figure 6. All the elements of $\mathcal{M}_{4,3}$ together with the number of permutations in $S_{[4] \times [3]}$ that are mapped to each by the function G.

(in total) are mapped to maximal configurations that are not efficient than to maximal configurations that are efficient. Namely, for $m = n = 3$, the efficient configuration (with building density $\frac{8}{9}$) is 5 times more likely to occur that any of the other configurations (notice that in this specific example, all the other maximal configurations are inefficient). However, since there are 9 inefficient configurations (all of them having building density $\frac{7}{9}$), more permutations are mapped to maximal configurations with building density $\frac{7}{9}$ then to the efficient configurations. In other words, when considering particular configurations, the most likely configuration to be obtained—assuming the random sequential building model—is an efficient one; but when considering their occupancies, the most likely occupancy to be obtained is strictly smaller than the occupancy of the efficient configurations.

At first, the discussion above may seem connected to Claim 1. However, the mere fact that different maximal configurations occur with different probabilities under the two models is not sufficient to conclude that the resulting building densities follow different distributions. The numbers could possibly conspire to produce the same building density distribution, although that would be quite surprising.

© THE MATHEMATICAL ASSOCIATION OF AMERICA
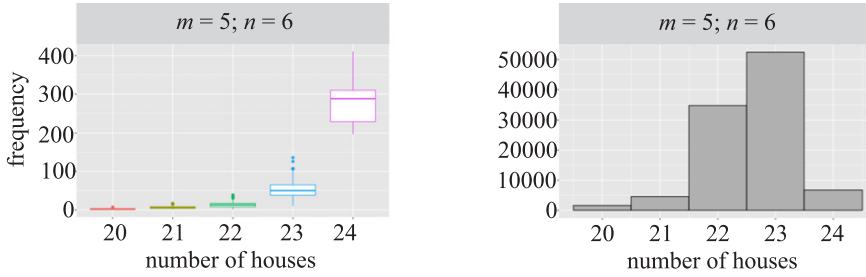
(a) Each data point represents a different configuration, plotted above its occupancy number. The frequency is the number of sampled permutations which were mapped to that particular configuration.

(b) The histogram of the occupancies of configurations obtained from those same 100,000 sampled permutations.

Figure 7. Comparison of the number of permutations (out of 100,000 sampled) which the function G maps to a particular maximal configuration and the number of permutations (out of those same 100,000) which the function G maps to any of the maximal configurations with a particular occupancy. The two charts display the same data. Note that there are more distinct configurations with occupancy 23 than those with occupancy 24. As a result, the occupancy 23 in histogram outweighs the occupancy 24, even though each particular configuration with occupancy 23 occurs less frequently than any particular configuration with occupancy 24.

For bigger tracts of land, it becomes quite computationally demanding to evaluate the function G for all permutations in $S_{[m]\times[n]}$, but we can sample some permutations from $S_{[m]\times[n]}$ uniformly at random and see how the function G acts on them—thus essentially sampling configurations from the sequential building model. We did this for a 5×6 tract of land. We sampled 100,000 random permutations from the set $S_{[5]\times[6]}$ and we acted with the function G on them. The results can be seen in Figure 7. From the box plot it is clear that again many more permutations were mapped to a particular efficient configuration than to the maximal configurations that are not efficient. More precisely, the box plot in Figure 7 shows us that most of the efficient configurations (with occupancy 24) that were obtained from sampled random permutations of the set $S_{[5]\times[6]}$ have between 200 and 300 permutations mapped to them and some even have 400 permutations mapped to them (out of 100,000 randomly sampled permutations of the set $S_{[5]\times[6]}$). On the other hand, most of the maximal configurations with occupancy 23 (one less than the efficient ones) have around 50 different permutations mapped to them. This number decreases even more when we go to the maximal configurations with fewer than 23 occupied lots. However, as the histogram shows, there are obviously many more different maximal configurations with 23 or even 22 occupied lots (than there are efficient ones) which results in building densities 23/30 and 22/30 being more probable than the building density 24/30. It is easy to check that the building density of the efficient configurations on a 5×6 tract of land is precisely 24/30.

Conjecture 3. *Any maximal configuration that has the highest number of permutations mapped to it (by the function G) is an efficient configuration. Moreover, any maximal configuration that has the fewest permutations mapped to it is an inefficient configuration.*

Notice that the claim of Conjecture 3 is intuitive, but not trivial. Namely, every empty lot in the maximal configuration implies some constraints on the permutation that resulted in this particular configuration. More precisely, some specific lots had to appear in the permutation before the lot that remained empty. The more empty lots we want to obtain in the final configuration, the more constraints on a permutation we impose and this should result in fewer permutations mapping to maximal configura-
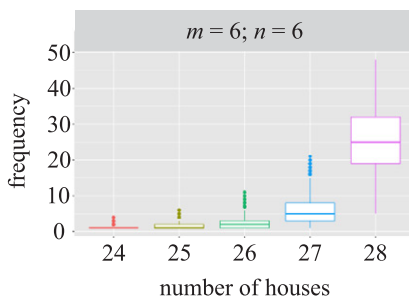
Figure 8. We sampled 100,000 permutations from the set $S_{[6]\times[6]}$ and acted upon those samples with function G. Each data point represents a different configuration, plotted above its occupancy number. The frequency is the number of sampled permutations which were mapped to that particular configuration. Note that the results suggest that one could have maximal configurations that are not efficient with more permutations mapped to them than to some of the efficient configurations.

tions with more empty lots. On the other hand, the situation is not completely trivial since in the case when the empty lots are grouped together, they imply fewer constraints on a permutation (compared to when they are further apart) and we can expect that the more empty lots we have, the more grouped together they can be. That is why, in general, one expects to see efficient configurations that have fewer permutations mapped to them than to some maximal configurations that are not efficient. Simulations on the 6×6 grid strongly support this claim (see Figure 8). Notice that this is not in a contradiction with Conjecture 3.

3. DISTRIBUTION OF RANDOM VARIABLES $X_{m,n}^u$ AND $X_{m,n}^s$. Ideally, one would like to obtain analytically the distributions of random variables $X_{m,n}^u$ and $X_{m,n}^s$ for arbitrary m and n. For small tracts of land, one can obtain the exact probability distributions of variables $X_{m,n}^u$ and $X_{m,n}^s$ computationally, by exhaustive search. However, this soon becomes infeasible due to the computational complexity. As we have not been able to make much progress on this problem, we leave it open:

Question 4. *What is the distribution of random variables $X_{m,n}^u$ and $X_{m,n}^s$ (for $m, n \in \mathbb{N}$)?*

If this question proves too hard, one would at least like to know the support of these random variables. In [**9**, Lemma 2.1] it was shown that the occupancy of any maximal configuration on the $m \times n$ grid lies between $\frac{1}{2}mn$ and $\frac{3}{4}mn + \frac{m-1}{2} + \frac{n}{4}$ when $m, n \geq 2$. From this it follows that for large grids, as both $m \to \infty$ and $n \to \infty$, the building density of maximal configurations must be between $\frac{1}{2}$ and $\frac{3}{4}$. It is also interesting to note that these limit building densities are in, in fact, attained on the infinite grid by periodic configurations shown in Figure 9.
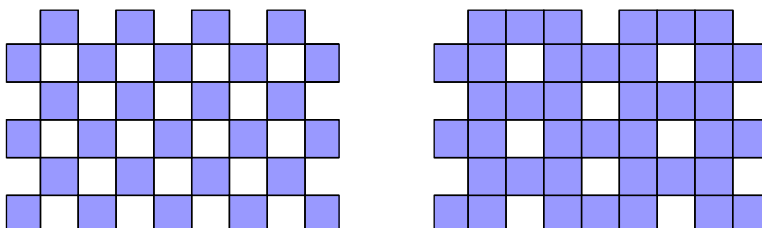

Figure 9. Periodic patterns on the infinite grid attaining densities $\frac{1}{2}$ and $\frac{3}{4}$.

The lowest possible value of the building density on the $m \times n$ tract of land is actually known explicitly. By reformulating the results from [9, Definition 3.8 and Theorem 4.5] one can see that for any $m, n \geq 2$,

$$X_{m,n}^{u/s} \geq \begin{cases} \frac{1}{2} + \frac{2}{mn}, & \text{if } n \equiv 0 \pmod 4, \\ \frac{1}{2} + \frac{1}{n}, & \text{if } n \equiv 2 \pmod 4, \\ \frac{1}{2} + \frac{1}{2n} + \frac{1}{mn}, & \text{if } n \equiv 1 \pmod 2, \end{cases} \tag{1}$$

and this lower bound is attained with positive probability. The tightest upper bound from [9] for the support of the random variables $X_{m,n}^u$ and $X_{m,n}^s$ is not explicit but expressed through a recurrence relation (see [9, Theorem 4.12]). These results imply that the bounds of $\frac{1}{2}$ and $\frac{3}{4}$ on the building density of maximal configurations when both m and n tend to infinity are in fact sharp.

Since we were not able to give an analytical answer to Question 4, we turned to simulations. From now on, we focus only on the random variable $X_{m,n}^s$ as this is a more realistic model for the problem that we had in mind and we also do not need to know all the maximal configurations to be able to sample this random variable. Using simulations, we obtained an approximate distribution of the random variable $X_{m,n}^s$ by sampling a large number of permutations from $S_{[m]\times[n]}$, acting with function G on those permutations, and then counting the number of occupied lots in the resulting maximal configurations. The results can be seen in Figure 10. We simulated the distribution of random variables $X_{5,5}^s$, $X_{5,15}^s$, $X_{15,5}^s$, and $X_{15,15}^s$ by sampling 50,000 random permutations from $S_{[m]\times[n]}$ for each combination of parameters m and n. There are several reasons for choosing these particular values of m and n. The first reason is that for all values of m and n smaller than 17 the authors in [9] calculated precise values, not only of the lowest, but also of the highest building density possible so we already know the support of the random variables we simulated. The lowest building densities possible can easily be read from relation (1) and the highest possible building densities were obtained by posing the integer programming formulation of the problem and then solving it using IBM ILOG CPLEX (see [2]). For the random variables we simulated, the
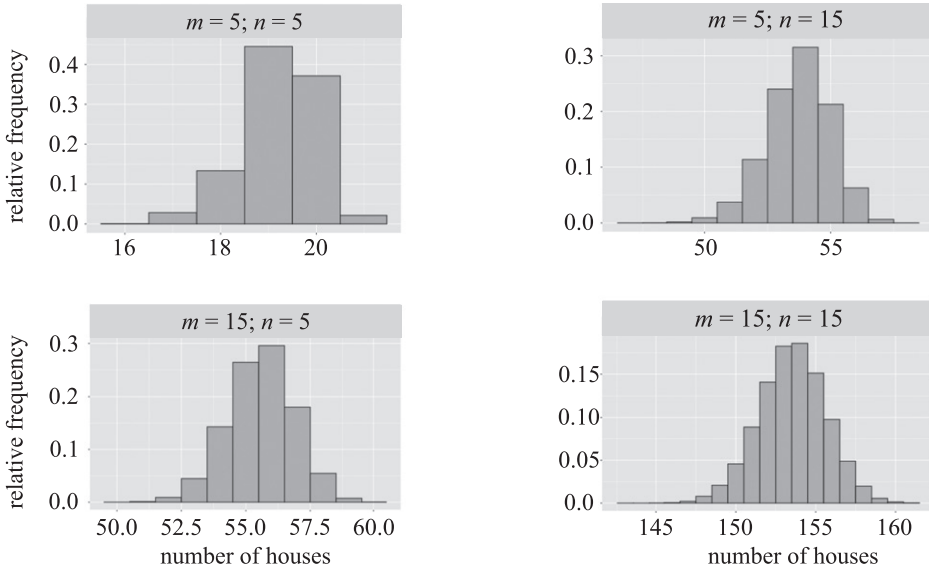

Figure 10. Histograms that approximate the distribution of the random variable $(mn) \cdot X_{m,n}^s$, i.e., the total number of houses (occupancy) in a maximal configuration.

sharp bounds are as follows

$$\frac{16}{25} \le X_{5,5}^s \le \frac{21}{25}, \quad \frac{41}{75} \le X_{5,15}^s \le \frac{58}{75}, \quad \frac{46}{75} \le X_{15,5}^s \le \frac{61}{75}, \quad \frac{121}{225} \le X_{15,15}^s \le \frac{173}{225}.$$

The other reason for our choices of m and n is that in the next section we will be interested in what happens with the $\mathbb{E}[X_{m,n}^s]$ when at least one of the indices m or n goes to infinity. Box plots in Figure 11 already suggest that increasing m or n causes the expectation to drop and the size of the drop depends on whether we increase only m, only n, or both m and n. This is exactly what we will show in the next section, with the aid of simulations.
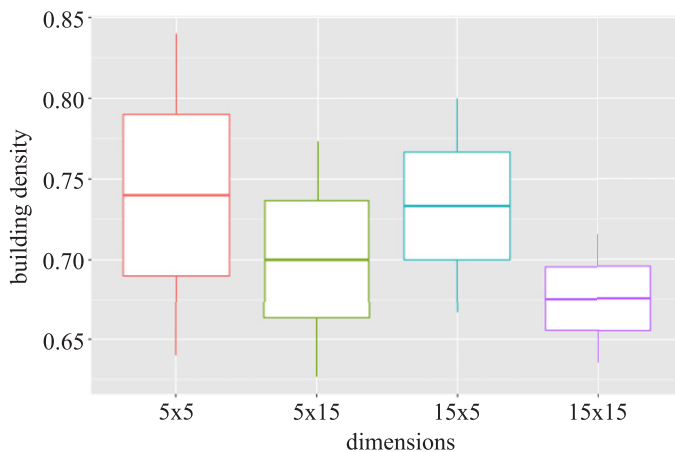

Figure 11. Comparison of distributions of random variables $X_{m,n}^s$ for different values of m and n.

4. PACKING DENSITY. Sequential building of houses has a lot of similarities with the so-called random sequential adsorption (RSA). RSA refers to a process where particles are randomly introduced in a system and, if they do not overlap with any previously adsorbed particle, they adsorb and remain fixed for the rest of the process. In our model, we assume that houses are built one-by-one. Once a house is built, we can say that it is adsorbed and that it remains fixed for the rest of the process. However, even though it will never happen that an occupied lot is chosen (hence there are no overlaps like in a standard RSA model), it can happen that a house cannot be built on the chosen lot (i.e., it is not adsorbed) if that house would block some other house from the sunlight or if that house itself would be blocked from the sunlight. RSA can be carried out in computer simulations, in a mathematical analysis, or in experiments. The same holds for our model. Our main goal for future work is to find a way to mathematically analyze our model. To gain a better understanding of what to expect, we first carried out computer simulations of this model. The standard RSA was first studied by one-dimensional models. The most famous examples are the attachment of pendant groups in a polymer chain by Paul Flory [4], and the car-parking problem by Alfréd Rényi [10] and Ewan Stafford Page [7]. Page studied a model analogous to the one considered by Rényi, but on a discrete interval (so called *discrete Rényi packing problem* or *unfriendly seating problem*). Other early works include those of Benjamin Widom [13]. In two and higher dimensions many systems have been studied by computer simulations, including disks, randomly oriented squares and rectangles, aligned squares and rectangles, k-mers (particles occupying k adjacent sites), and various other shapes.

© THE MATHEMATICAL ASSOCIATION OF AMERICA

An important result related to these models is the maximum surface coverage, called the saturation coverage, the packing fraction, or the packing density; also known as the jamming limit in relation to parking problems. This is exactly the question that was in the center of our attention. More precisely, we were interested in values

$$\lim_{m\to\infty} \mathbb{E}[X_{m,n}^s] \ \text{ (for fixed } n), \quad \lim_{n\to\infty} \mathbb{E}[X_{m,n}^s] \ \text{ (for fixed } m), \quad \lim_{\substack{n\to\infty \\ m\to\infty}} \mathbb{E}[X_{m,n}^s]. \qquad (2)$$

The precise constants in the limit are called the packing densities. There are some famous packing densities calculated explicitly and many more estimated using simulations. The most common notion of convergence considered in these kinds of problems is convergence of expected values. However, in some instances it was possible to prove convergence in probability for these random variables in addition to convergence of expectations (see e.g., [7] for discrete Rényi car-parking problem). Note that here, due to boundedness of random variables, the convergence in probability is strictly stronger than the convergence of expected values. The techniques used in the references cited above seem to exploit the one-dimensionality of the problem and it is not clear how one could extend them to work for 2D grids.

For saturation coverage of k-mers on one-dimensional lattice systems, see [6]. Analogous results on two-dimensional lattice systems can be found in [12] and [11]. Saturation coverage of $k \times k$ squares on a two-dimensional square lattice is studied in [8] and [1]. The model which has some resemblance to our model is one of the models studied in [5]. In this model, dimers of random orientations are deposited on a square lattice. A dimer can occupy a pair of nearest-neighbor sites only if both of the chosen sites are vacant. Interpreting each dimer as a house and an empty lot (still keeping in mind that the sun cannot come from the north), our model looks relatively similar, but the big difference is that in our model empty lots can overlap and that is why it doesn't belong to standard RSA models.

Another crucial difference of our model and various other instances of RSA is that in our model the orientation of the grid plays a role. The north side which never receives sunlight is prominent in our model and breaks the 4-fold symmetry that RSA models usually possess. One might argue that this is the most interesting feature of our model. For a systematic overview of all of the mentioned results and many more, see [3,14].

To get a better insight into the behavior of $\mathbb{E}[X_{m,n}^s]$, when we let m or n go to infinity, we ran several simulations, the results of which are shown in Figures 12 and 13. In Figure 12(a) we fixed $m = 5$ and varied n from 10 to 100 with step size 10
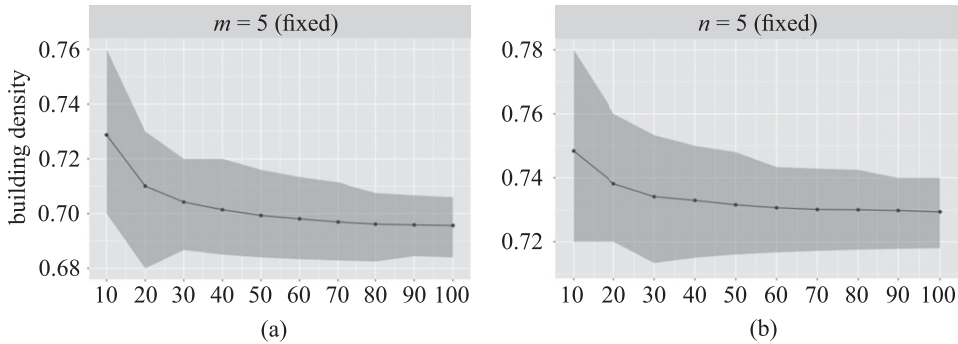

Figure 12. Behavior of $\mathbb{E}[X_{m,n}^s]$ when we let just one of the parameters, m or n, go to infinity and fix the other one. For each combination of m and n, we ran 2000 simulations and calculated the mean building density and furthermore the 5th and 95th percentile of the obtained building densities. Means are shown with the black dots that are connected with lines, and the area from 5th to 95th percentile is shaded gray.
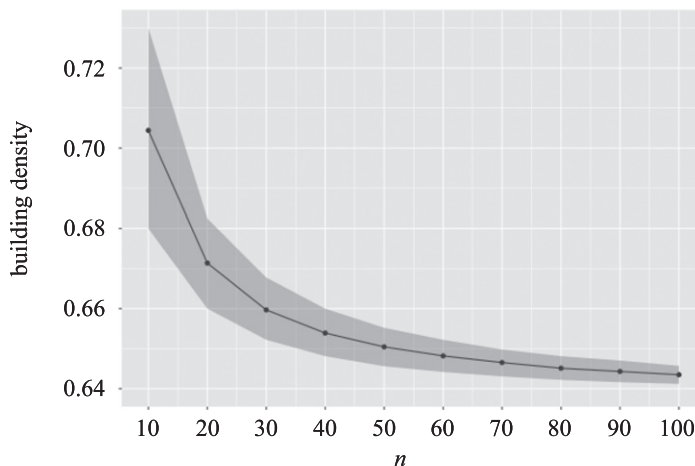
Figure 13. Behavior of $\mathbb{E}[X_{m,n}^s]$ when we let both of the parameters, m and n, go to infinity at the same rate.

Table 1. Packing densities.

m \ n	5	10	50	100	500	1,000	10,000	100,000
5	0.769	0.728	0.699	0.697	0.693	0.692	0.692	0.692
10	0.749	0.703	0.672	0.668	0.665	0.665	0.664	0.664
50	0.730	0.684	0.651	0.646	0.643	0.642	0.642	0.642
100	0.729	0.682	0.648	0.643	0.640	0.640	0.639	0.639
500	0.728	0.680	0.646	0.641	0.638	0.637	0.637	0.637
1,000	0.728	0.680	0.645	0.641	0.638	0.637	0.637	0.637
10,000	0.727	0.680	0.645	0.641	0.637	0.637	0.636	0.636
100,000	0.727	0.680	0.645	0.641	0.637	0.637	0.636	

and in Figure 12(b) we switched the roles of m and n. Finally, we ran the simulations where we varied both m and n at the same rate from 10 to 100 with step size 10, see Figure 13. Based on the results of these simulations we pose the following conjecture:

Conjecture 5. *The double sequence $\{\mathbb{E}[X_{m,n}^s]\}_{m,n}$ is decreasing in m and in n.*

Since we know that this expectation is strictly positive, from Conjecture 5 it would immediately follow that all the limits from (2) exist. We give Monte Carlo estimation of these limits in Table 1. For each combination of parameters m and n, we ran 100 simulations and then calculated the mean.

Already for a $100{,}000 \times 100{,}000$ tract of land we did not have enough computational power to run the simulation. Regardless of that, the pattern is pretty clear. Values in each particular row and each particular column are decreasing and then stabilizing at some point. Even though the literature suggests that analytically obtaining the precise packing density in any of the studied cases (letting m go to infinity, letting n go to infinity or letting both m and n go to infinity) is very involved, one can still ask the following question

Question 6. *What are the exact values of the limits*

$$\lim_{m \to \infty} \mathbb{E}[X_{m,n}^s], \qquad \lim_{n \to \infty} \mathbb{E}[X_{m,n}^s] \qquad \text{and} \qquad \lim_{\substack{m \to \infty \\ n \to \infty}} \mathbb{E}[X_{m,n}^s]?$$

Even if Question 6 turns out to be too demanding, one might pursue a simpler task of obtaining better approximations for packing density. A wide range of techniques for approximating packing densities in random sequential adsorption models

have already been developed. Therefore possible directions for future research include using simulations designed in a better and more appropriate way, or adopting some of the approaches from the RSA literature.

ORCID
Mate Puljiz ⬥ https://orcid.org/0000-0003-0912-8345
Stjepan Šebek ⬥ https://orcid.org/0000-0002-1802-1542

REFERENCES

[1] Brosilow, B. J., Ziff, R. M., Vigil, R. D. (1991). Random sequential adsorption of parallel squares. *Phys. Rev. A*. 43: 631–638. doi.org/10.1103/PhysRevA.43.631.

[2] Cplex, I. I. (2009). Version 12. 1: User's manual for CPLEX. *IBM Corp.* 46(53): 157.

[3] Evans, J. W. (1993). Random and cooperative sequential adsorption. *Rev. Mod. Phys.* 65(4): 1281–1329.

[4] Flory, P. J. (1939). Intramolecular reaction between neighboring substituents of vinyl polymers. *J. Am. Chem. Soc.* 61(6): 1518–1521.

[5] Gan, C. K., Wang, J. S. (1998). Extended series expansions for random sequential adsorption. *J. Chem. Phys.* 108: 3010–3012. doi.org/10.1063/1.475687.

[6] Krapivsky, P. L., Redner, S., Ben-Naim, E. (2010). *A kinetic view of statistical physics*. Cambridge: Cambridge Univ. Press. doi.org/10.1017/CBO9780511780516.

[7] Page, E. S. (1959). The distribution of vacancies on a line. *J. Roy. Statist. Soc. Ser. B.* 21: 364–374.

[8] Privman, V., Wang, J. S., Nielaba, P. (1991). Continuum limit in random sequential adsorption. *Phys. Rev. B.* 43: 3366–3372. doi.org/10.1103/PhysRevB.43.3366.

[9] Puljiz, M., Šebek, S., Žubrinić, J. (2021). Combinatorial settlement planing. *To appear in Contrib. Discrete Math.* https://arxiv.org/abs/2107.07555

[10] Rényi, A. (1958). On a one-dimensional problem concerning random space filling. *Magyar Tud. Akad. Mat. Kutató Int. Közl.* 3(1–2): 109–127.

[11] Slutskii, M. G., Barash, L. Y., Tarasevich, Y. Y. (2018). Percolation and jamming of random sequential adsorption samples of large linear *k*-mers on a square lattice. *Phys. Rev. E.* 98. doi.org/10.1103/PhysRevE.91.012109.

[12] Tarasevich, Y. Y., Laptev, V. V., Vygornitskii, N. V. (2015). Impact of defects on percolation in random sequential adsorption of linear *k*-mers on square lattices. *Phys. Rev. E.* 91. doi.org/10.1103/PhysRevE.91.012109.

[13] Widom, B. (1966). Random sequential addition of hard spheres to a volume. *J. Chem. Phys.* 44: 3888–3894. doi.org/10.1063/1.1726548.

[14] Wikipedia. Random sequential adsorption. https://en.wikipedia.org/wiki/Random_sequential_adsorption

MATE PULJIZ was born in Croatia in 1988. He received his master's degree from the University of Zagreb, Croatia, in 2012 and his Ph.D. in Pure Mathematics from the University of Birmingham, United Kingdom, in 2017. He is currently an Assistant Professor with the Department of Applied Mathematics, Faculty of Electrical Engineering and Computing, Zagreb, Croatia. His research interests include abstract dynamical systems, particularly the topology of their orbits, particle models, and fractional calculus.
Department of Applied Mathematics, Faculty of Electrical Engineering and Computing, University of Zagreb, Zagreb, Croatia
mate.puljiz@fer.hr

STJEPAN ŠEBEK was born in Croatia in 1990. He received his master's degree in 2014 and his Ph.D. in Mathematics in 2019, both from the University of Zagreb, Croatia. He is currently an Assistant Professor with the Department of Applied Mathematics, Faculty of Electrical Engineering and Computing, Zagreb, Croatia. His research interests include geometry and potential theory of random walks.
Department of Applied Mathematics, Faculty of Electrical Engineering and Computing, University of Zagreb, Zagreb, Croatia
stjepan.sebek@fer.hr

JOSIP ŽUBRINIĆ was born in Croatia in 1993. He received his master's degree in 2016 and his Ph.D. in Mathematics in 2022, both from the University of Zagreb, Croatia. He is currently a Postdoc with the Department of Applied Mathematics, Faculty of Electrical Engineering and Computing, Zagreb, Croatia. His research interests include homogenization and dimension reduction in the theory of elasticity.
Department of Applied Mathematics, Faculty of Electrical Engineering and Computing, University of Zagreb, Zagreb, Croatia
josip.zubrinic@fer.hr