# Hexagonal Run-Length Zero Capacity Region—Part I: Analytical Proofs

Spencer Congero<sup>®</sup>, *Student Member, IEEE*, and Kenneth Zeger<sup>®</sup>

Abstract—The zero capacity region for hexagonal (d, k) runlength constraints is known for many, but not all, d and k. The pairs (d, k) for which it has been unproven whether the capacity is zero or positive consist of: (i) k = d + 2 when  $d \ge 2$ ; (ii) k = d + 3 when  $d \ge 1$ ; (iii) k = d + 4 when either d = 4 or d is odd and  $d \ge 3$ ; and (iv) k = d + 5when d = 4. Here, we prove that the capacity is zero for all of case (i), and for case (ii) whenever  $d \ge 7$ . The method used in this paper is to reduce an infinite search space of valid labelings to a finite set of configurations that we exhaustively examine using backtracking. In Part II of this two-part series, we use automated procedures to prove that the capacity is zero in case (i) when 2 < d < 9, in case (ii) when 3 < d < 11, and in case (iii) when  $d \in \{4, 5, 7, 9\}$ , and that the capacity is positive in case (ii) when  $d \in \{1, 2\}$ , in case (iii) when d = 3, and in case (iv). Thus, the only remaining unknown cases are now when k = d + 4, for any odd d > 11.

*Index Terms*— Channel capacity, run length coding, hexagonal constraint.

## I. INTRODUCTION

ONE-DIMENSIONAL run-length constraint imposes both lower and upper bounds on the number of zeros that occur between consecutive ones in a binary string. Specifically, if d and k are nonnegative integers, or  $\infty$ , then a binary string is said to satisfy a (d, k) constraint if every consecutive pair of ones in the string has at least d zeros between them and the string never has more than k zeros in a row. It is known that if k > d, then the number of (one-dimensional) N-bit binary strings that satisfy the (d, k) constraint grows exponentially in N (e.g., [14]) and that the logarithm (base two) of that number, divided by N, approaches a positive limit as N grows to infinity. This limit is known as the "capacity" of the constraint.

The concepts of (d, k) constraints and capacities have been generalized to two dimensions, where the one-dimensional (d, k) constraint is imposed both vertically and horizontally. Sometimes these two-dimensional constraints are referred to as "rectangular constraints." To determine the capacity of a rectangular constraint, one counts the number of binary labelings of an  $N \times N$  square that satisfy the constraint, takes its

Digital Object Identifier 10.1109/TIT.2021.3120090

logarithm, and then divides by the area  $N^2$  of the square. It is known that this quantity approaches a limit  $C_{\text{rect}}(d, k)$  (called the "capacity" again) as N grows to infinity (e.g., [18]).<sup>1</sup>

The zero capacity region for a particular type of constraint is the set of all pairs (d, k) for which the (d, k) capacity equals zero. If a particular constraint has zero capacity, then the number of valid labelings of a region does not grow exponentially fast in terms of the volume (e.g., length for 1 dimension, area for 2 dimensions, etc.) of the region.

During 1998-2013, various studies of the two-dimensional rectangular capacity were performed for the particular case  $C_{\text{rect}}(1,\infty) \approx 0.587891162$  by Calkin and Wilf [5], Weeks and Blahut [34], Baxter [3], Marcus and Pavlov [23], [24], [28], and in [26]. This rectangular  $(1,\infty)$  constraint is sometimes referred to as the "hard square model" by physicists [2], and its capacity is known to equal the rectangular capacity  $C_{\text{rect}}(0,1)$ .

For two-dimensional rectangular (d, k) constraints, the zero capacity region was completely characterized in 1999 in [18], where it was shown that the capacity satisfies  $C_{\text{rect}}(d, k) > 0$  if and only if  $k \ge d + 2$ , when  $d \ge 1$  (i.e.,  $C_{\text{rect}}(d, k) = 0$  when k = d + 1). It is also known that  $C_{\text{rect}}(0, k) > 0$  and  $C_{\text{rect}}(k, \infty) > 0$  for all  $k \ge 1$ . Bounds on the two-dimensional rectangular (d, k) capacity were given in [18], by Sharov and Roth in [31], and were later improved and generalized to higher dimensions by Schwartz and Vardy in [30]. Schwartz and Bruck [29] introduced an interesting rigorous method for obtaining the capacity of general two-dimensional constrained systems, although it is not presently known how to effectively apply it to the hexagonal (d, k) case.

In 2016, Elishco *et al.* [10] introduced the notion of "semiconstrained systems," in which certain prescribed patterns are forbidden to appear more often than particular designated frequencies. These systems generalize (d, k) constrained systems, since (d, k) constraints require that the forbidden patterns (i.e., patterns violating the *d* or *k* constraints) must never occur.<sup>2</sup> Bounds and asymptotics for the capacity of semiconstrained systems were obtained in [10], and also, in 2018, by the same authors in [11], and for the multidimensional case in [12].

Other two-dimensional constraints have been studied in the literature as well. In 1961, Kasteleyn [17] counted the asymptotic number of arrangements of  $1 \times 2$  tiles that cover

Manuscript received September 21, 2020; revised August 22, 2021; accepted October 10, 2021. Date of publication October 14, 2021; date of current version December 23, 2021.

The authors are with the Department of Electrical and Computer Engineering, University of California at San Diego, La Jolla, CA 92093 USA (e-mail: scongero@ucsd.edu; ken@zeger.us).

Communicated by M. Schwartz, Associate Editor for Coding Techniques. Color versions of one or more figures in this article are available at https://doi.org/10.1109/TIT.2021.3120090.

<sup>&</sup>lt;sup>1</sup>Or, equivalently, one may count the number of  $N \times M$  rectangles satisfying the constraint, take its logarithm, divide by the area NM, and then let both N and M tend to infinity in any manner.

<sup>&</sup>lt;sup>2</sup>The (d, k) constrained systems were called "fully constrained" systems in [10].

<sup>0018-9448 © 2021</sup> IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.



Fig. 1. Converting a hexagonal lattice into a square lattice using northeast diagonals.

a square lattice. In 2006, Forchhammer and Laursen [13] estimated the capacity of a two-dimensional binary code forbidding "isolated bits," i.e., a code where each 0 and 1 cannot be surrounded entirely by bits of the opposite parity. In 2010, Louidor and Marcus [21] determined the capacity of two different two-dimensional constrained systems, namely "charge constrained" and "odd constrained" systems.

Also, in 1961, Wang [32] considered finite sets of certain equal-sized squares, each of whose sides are labeled by one of a given set of colors. These squares later became known as "Wang tiles." Such Wang tiles are used to tile the plane under the constraint that adjacent tiles (horizontally and vertically) share a common color where they meet. Durand et al. [9] in 2014, and Chen et al. [7] in 2016, counted the number of such Wang tilings and computed a quantity they called the "entropy," or alternatively the "spatial entropy," which is analogous to the capacity calculation described above. The authors in [7] used the phrase "spatial chaos" to describe when the spatial entropy is positive, and gave conditions on when the spatial entropy is zero. In [9], a specific aperiodic tile set (i.e., a tile set such that every tiling of the plane by tiles from this set is aperiodic) was shown to have positive spatial entropy. However, there is no known direct connection between twodimensional (d, k) constraints and Wang tilings.

We now focus on the family of two-dimensional constraints studied in this paper. A "hexagonal" (d, k) constraint is a different type of two-dimensional run-length constraint, that imposes one-dimensional (d, k) constraints on a hexagonal lattice. Each hexagon in such a lattice has six neighbors, and thus three axes run through it. The one-dimensional constraint must be satisfied along each of the three axes for each hexagon in the lattice. An equivalent way to view the hexagonal constraint on a rectangular lattice is to impose the (d, k) constraint both horizontally and vertically, and also along one of the two diagonal directions (we will use the northeast-southwest direction, but refer to it as the "northeast diagonal") [2, p. 409] (see Figure 1). The same diagonal constraint direction is chosen for all squares in the lattice.

The hexagonal (d, k) capacity  $C_{\text{hex}}(d, k)$  is known to be positive for certain pairs (d, k). In fact, if  $C_{\text{hex}}(d, k) > 0$ , then it immediately follows that  $C_{\text{hex}}(d', k') > 0$  whenever either d' < d or k' > k (or both), since the constraints weaken in either instance. Positive lower bounds on the hexagonal (d, k) capacity were previously proven for d = 0, and for all values of  $d \ge 5$  for sufficiently large k (for example, k = d + 5suffices), and now also for  $1 \le d \le 4$  with our results in Part II. In what follows, we will summarize, for each d > 0, the smallest known k such that  $C_{\text{hex}}(d, k) > 0$ .

The only exactly known non-zero capacity of a hexagonal (d, k) constraint is for the case  $(1, \infty)$ , which is known in the physics literature as the "hard hexagon model." As with the rectangular constraint, it is easy to show that the hexagonal (0, 1) and  $(1, \infty)$  capacities are the same, by reversing the roles of 0s and 1s. The problem of counting the number of patterns in a bounded area that satisfy the hexagonal  $(1, \infty)$  constraint was considered in the context of Ising models in physics, as early as in 1944 by Onsager [27], and in 1950 by Wannier [33]. An equivalent problem is to find the number of configurations of non-attacking kings on a chessboard with regular hexagonal cells.

In 1978, Metcalf and Yang [25] conjectured that the capacity of the hexagonal  $(1, \infty)$  constraint was  $\log_2 e^{1/3} \approx 0.48090$ , but this was disproven in 1980 by Baxter and Tsang [4], who obtained a slightly more accurate estimate.

Baxter [1], [2], later in 1980, and then Joyce [15], [16] in 1988, performed numerous intricate calculations, which when combined determine the exact capacity<sup>3</sup> of the hexagonal  $(1,\infty)$  constraint (the approximate value is  $C_{\text{hex}}(1,\infty) = C_{\text{hex}}(0,1) \approx 0.4807676$ , which is fairly close to the incorrect conjecture of Metcalf and Yang). As a result, one deduces that  $C_{\text{hex}}(0,k) > 0$  for all  $k \ge 1$ .

In 2001, using the technique of finding two distinct tileable squares, Censor and Etzion [6] proved that  $C_{\text{hex}}(d, d+4) > 0$ 

 $^{3}$ The exact value is remarkably given by Baxter and Joyce as the logarithm, base two, of the product

$$4^{-1}3^{5/4}11^{-5/12}c^{-2}$$

$$\cdot \left(1 - \sqrt{1 - c} + \sqrt{2 + c + 2\sqrt{1 + c + c^2}}\right)^2$$

$$\cdot \left(-1 - \sqrt{1 - c} + \sqrt{2 + c + 2\sqrt{1 + c + c^2}}\right)^2$$

$$\cdot \left(\sqrt{1 - a} + \sqrt{2 + a + 2\sqrt{1 + a + a^2}}\right)^{-1/2}$$

where  $a = -\frac{124}{363} \cdot 11^{1/3}$ ,  $b = \frac{2501}{11979} \cdot 33^{1/2}$ , and  $c = (\frac{1}{4} + \frac{3}{8}a((b+1)^{1/3} - (b-1)^{1/3}))^{1/3}$ .

for all even  $d \ge 6$ . An immediate consequence is that  $C_{\text{hex}}(d, d+5) > 0$  for all odd  $d \ge 5$ , since the hexagonal (d, d+5) constraint is weaker than the (d+1, d+5) constraint. In Part II of our papers, we present a tiling algorithm that automatically generates distinct tileable square labelings that demonstrate positive hexagonal (d, k) capacities for certain pairs (d, k). In particular, we prove that the capacities  $C_{\text{hex}}(1, 4)$ ,  $C_{\text{hex}}(2, 5)$ ,  $C_{\text{hex}}(3, 7)$ , and  $C_{\text{hex}}(4, 9)$  are all positive.

Also, we note that the positive hexagonal (d, k) capacities obtained in [6] were for the case of k = d + 4 when d is even and  $d \ge 6$ , but the proof technique does not apply to odd  $d \ge 5$ . In Part II, in contrast to even  $d \ge 6$ , we show that some of the open cases with k = d + 4 when d is odd have zero capacity.

We next summarize the pairs (d, k) for which it was previously known that  $C_{\text{hex}}(d,k) = 0$ . It suffices, for each d, to give the largest k that makes  $C_{\text{hex}}(d, k) = 0$ . Since each rectangular (d, k) constraint is weaker than the corresponding hexagonal (d, k) constraint, it immediately follows that  $C_{\text{hex}}(d,k) \leq C_{\text{rect}}(d,k)$  for all d and k. Thus, in particular,  $C_{\text{hex}}(d,k) = 0$  at least whenever  $C_{\text{rect}}(d,k) = 0$ , namely when  $k = d + 1 \ge 2$ . A stronger result was stated in [19], namely that  $C_{\text{hex}}(d, d+2) = 0$  for all  $d \ge 1$ , but no proof has been published. We prove this result in Section IV in Theorem 4.1 for  $d \ge 3$ , and in Theorem 4.3 for  $d \in \{1, 2\}$ . In our Part II, it is also implied by the Forbidden String Algorithm for the cases d = 3 and d = 5, and by the Constant Position Algorithm when  $1 \le d \le 9$ . Even though for  $d \ge 7$  the k = d + 2 case is implied by the stronger result we prove for the k = d+3 case, the proof of Theorem 4.1 provides a relatively less complex introduction to the technique used in the stronger case. We note that the proofs we provide here of  $C_{\text{hex}}(d, d+2) = 0$  when  $2 \le d \le 6$  were neither in the previous literature, nor implied by our k = d + 3 results in this paper.

In [20], it was stated that  $C_{hex}(d, d + 3) = 0$  when  $d \in \{3, 4, 5, 7, 9, 11\}$ . In [6], Censor and Etzion considered an octagonal (d, k) constraint, which assumes the hexagonal (d, k) constraint plus an additional constraint along the northwest diagonal, and proved that the octagonal (d, k)capacity is zero whenever k = d + 3 and d > 0. However, they did not give any results about whether the hexagonal (d, k) capacity is zero when k = d + 3, but did pose it as an open question, which partially motivated our present paper. In summary, there have been an infinite number of cases for k = d+3, prior to our present paper, where it was unknown if the hexagonal capacity is zero. We answer this open question in completion here.

Specifically, whether  $C_{\text{hex}}(d, k)$  is positive or zero has been unproven<sup>4</sup> for the following cases:

- (i) k = d + 2 when  $d \ge 2$
- (ii) k = d + 3 when  $d \ge 1$
- (iii) k = d + 4 when either d = 4 or d is odd and  $d \ge 3$
- (iv) k = d + 5 when d = 4.

Among these cases, we prove here (in Theorem 4.1 and Theorem 4.3) that the hexagonal capacity equals zero in all of case (i), and (in Theorem 2.1) in case (ii) for all  $d \ge 7$ . In Part II, we prove that the capacity is zero in case (i) when  $2 \le d \le 9$ , in case (ii) when  $3 \le d \le 11$ , and in case (iii) when  $d \in \{4, 5, 7, 9\}$ , and that the capacity is positive in case (ii) when  $d \in \{1, 2\}$ , in case (iii) when d = 3, and in case (iv).

Table I summarizes the present knowledge of the zero capacity region when d is less than 19 and k is less than 25, including the results we present in Parts I and II of these papers. The results from Part I are shown surrounded by squares and the results from Part II are shown surrounded by circles. We note that four of the results turn out to be produced by both the methods in Part I and Part II, and we denote them in the table being surrounded by both a circle and a square. Proofs of the results in Part I or Part II have not previously appeared in the literature. We note that although we provide here the first published proofs of the cases where k = d + 2, those satisfying  $d \ge 7$  are not listed as new results in the table, since they directly follow from our stronger (but more complex) k = d + 3 proof,

The four cases shown in our Part II are denoted by "+" signs inside circles. For any fixed d, the leftmost "+" in row d of Table I represents the smallest k for which it is known that the hexagonal (d, k) capacity is positive. Every "+" in the table represents a positive lower bound, rather than an exact capacity, except for the (0, 1) case. Exact values appear difficult to obtain.

In contrast to proving that a capacity is positive, demonstrating that a capacity is zero requires new techniques, which can be very complex. One technique was used in [18] to prove that the rectangular (d, d + 1) capacity is zero for all  $d \ge 1$ . The technique showed that, asymptotically, the values of the bits stored in a linear amount of space of an  $N \times N$  square determine the values of the bits in the remaining quadratic amount space in the square. In other words, the number of different valid labelings of such squares is  $2^{O(N)}$ , which implies the constraint has zero capacity. In contrast, for a constraint to have positive capacity, there would need to be  $2^{\Omega(N^2)}$  different valid labelings of an  $N \times N$  square. The same general goal, although with a significantly different approach, will be used in the present paper to show that certain hexagonal constraints have zero capacity.

Specifically, our approach in Part I to proving a particular hexagonal (d, k) capacity is zero is to show that for large enough squares of side length N, with a fixed labeling of a thin outer "frame" of width k + 1, at most one valid labeling of the square's interior is possible. This is accomplished by means of assuming, to the contrary, that there exist at least two valid square labelings for a given frame labeling, and then drawing (rather laborious) logical inferences which lead to a contradiction. A series of assumptions is made using a manual backtracking method, which ultimately leads to the desired contradiction. This approach becomes very complex, depending on the level of pushing and popping on the stack of assumptions. Then, since an  $N \times N$  square's frame contains O(N) bit locations, the total number of distinct labelings of

<sup>&</sup>lt;sup>4</sup>Some of these cases were stated in [19] and [20] and are included in Part II for archival purposes.

# TABLE I

Summary of the Known Hexagonal (d, k) Zero Capacity Region for Small d and k. Zero and Positive Capacities Are Denoted by "0" and "+," Respectively. The Zeros in Squares Denote Our Contributions in the Present Paper (Part I), While the Circled Symbols Are From Our Part II [8], and Those With Both Squares and Circles Occurred in Both Parts I and II. The Question Marks Denote Remaining Unsolved Cases

$d \backslash k$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
0	0	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
1		0	0	0	(+)	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
2			0	0	$\bigcirc$	(+)	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
3				0	0	$\bigcirc$	$\bigcirc$	(+)	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
4					0	0	$\bigcirc$	$\bigcirc$	$\bigcirc$	(+)	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
5						0	0	$\bigcirc$	$\bigcirc$	$\bigcirc$	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
6							0	0	$\bigcirc$	$\bigcirc$	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
7								0	0	0	$\bigcirc$	$\bigcirc$	+	+	+	+	+	+	+	+	+	+	+	+	+
8									0	0	0	$\bigcirc$	+	+	+	+	+	+	+	+	+	+	+	+	+
9										0	0	0	$\bigcirc$	$\bigcirc$	+	+	+	+	+	+	+	+	+	+	+
10											0	0	0	$\bigcirc$	+	+	+	+	+	+	+	+	+	+	+
11												0	0	0	$\bigcirc$	?	+	+	+	+	+	+	+	+	+
12													0	0	0	0	+	+	+	+	+	+	+	+	+
13														0	0	0	0	?	+	+	+	+	+	+	+
14															0	0	0	0	+	+	+	+	+	+	+
15																0	0	0	0	?	+	+	+	+	+
16																	0	0	0	0	+	+	+	+	+
17																		0	0	0	0	?	+	+	+
18																			0	0	0	0	+	+	+

the square is  $2^{O(N)}$ , instead of the required  $2^{\Omega(N^2)}$  for positive capacity, which proves the capacity is zero.

## **II. PRELIMINARIES**

A square is an  $N \times N$  two-dimensional array, for some positive integer N. A labeling of a subset of a square assigns a 0 or 1 to each element of the subset. We will refer to horizontal, vertical, and northeast diagonal lines in a square as rows, columns, and diagonals, respectively, or more generally as files. In an  $N \times N$  square, all rows and columns have length N, whereas diagonals can have lengths ranging from 1 (at two of the corners) to N. For any positive integer  $\delta$ , the frame of width  $\delta$  of a square S is the union of the first  $\delta$ and last  $\delta$  rows and the first  $\delta$  and last  $\delta$  columns of S (see Figure 2).

A labeling l on a square S is said to *satisfy the hexagonal* (d, k) constraint (or is *valid*) if in every file there are at least d zeros between any two ones, and any run of 0s has length at most k.

The *capacity* of the hexagonal (d, k) constraint is defined as

$$C_{\text{hex}}(d,k) = \lim_{N \to \infty} \frac{\log_2 |L|}{N^2},$$

where L is the set of all labelings of an  $N \times N$  square that satisfy the hexagonal (d, k) constraint. The capacity is known



Fig. 2. Frame of width  $\delta$  in an  $N \times N$  square.

to exist for all d, k (e.g., [18]). If  $C_{\text{hex}}(d, k) > 0$ , then the number of valid labelings is lower bounded as  $|L| = 2^{\Omega(N^2)}$ .

The main result (proven in Section V) of this paper is that  $C_{\text{hex}}(d, d+3) = 0$  whenever  $d \ge 3$ , as stated in the following theorem.

Theorem 2.1: The capacity of the hexagonal (d, k) constraint is zero whenever  $d \ge 3$  and k = d + 3.

It has been known [20] that  $C_{hex}(3,6) = C_{hex}(4,7) = C_{hex}(5,8) = 0$ , and it is shown in our Part II that  $C_{hex}(6,9) = 0$ . We therefore restrict attention to  $d \ge 7$  in the proof of Theorem 2.1.

An overview of the proof of this result is given in Section III-A, and the actual proof is given in Section V. In Section IV, as a preview to Section V, an illustration of the proof technique is given for the simpler case of the (d, d+2) constraint.

Lemma 2.2: For any nonnegative integers d, k, and  $\delta$ , if the capacity of the hexagonal (d, k) constraint is positive, then there exists a sufficiently large square on which some pair of distinct labelings satisfying the hexagonal (d, k) constraint agree on the square's frame of width  $\delta$ .

**Proof:** Let L be the set of valid labelings of an  $N \times N$  square and suppose the capacity is positive; then  $|L| = 2^{\Omega(N^2)}$ . Since the frame's area is  $4\delta(N - \delta)$ , the number of valid labelings of the frame is at most  $2^{4\delta(N-\delta)}$ , which grows more slowly than the number of valid labelings of the entire  $N \times N$  square. Thus for large enough N, there must exist two distinct labelings of an  $N \times N$  square that agree on the frame.

Let r and b be labelings of a square S. The *disagreement* set of r and b is

$$D_0 = \{(i, j) \in S : r(i, j) \neq b(i, j)\}.$$

For each point (i, j) in the disagreement set of labelings r and b, we say that (i, j) is colored red if r(i, j) = 1 and is colored black if b(i, j) = 1.

For any subset D of a square and any file f, if  $X \in f \cap D$ , then we say X is a point of f in D. If X and Y are points of a file f, then X is said to be before Y if X is to the left (respectively, below, or southwest) of Y, if f is a row (respectively, a column, or diagonal). If D is a subset of the disagreement set, then the *first* element of f in D is the minimum element for this ordering among points in  $f \cap D$ .

For any subset of the disagreement set, we define a coordinate system whose origin (0,0) is the first element in the bottommost row, and assume, without loss of generality, that the origin is colored black. A *disagreement diagram* is an illustration showing the color of each point in a subset of the disagreement set. The files in a disagreement diagram are numbered as shown in Figure 3. The leftmost point in the bottommost nonempty row of a disagreement subset will be referred to as the *lowest-left point* of the subset.

#### **III. DESCRIPTION OF PROOF TECHNIQUE**

# A. Overview

For the remainder of the paper, we set k = d+3. Our goal is to prove that the hexagonal (d, k) capacity equals zero. We will assume, to the contrary, that  $C_{\text{hex}}(d, k) > 0$ , and attempt to derive a contradiction.

Let N be a positive integer and let r and b be any two distinct valid labelings of an  $N \times N$  square, such that the labelings agree on the square's frame of width (k + 1). Such labelings are guaranteed to exist for sufficiently large N by Lemma 2.2. The value of N could conceivably be arbitrarily large, and we know of no analytical upper bound on the size of the square being labeled.



Fig. 3. Illustration of file indexing in disagreement diagrams. The first disagreement point in the bottommost row is colored black and defined to be located at the origin (0, 0), and each file is indexed relative to this origin point. In particular, rows and columns increase in the north and east directions, respectively, and diagonals increase in the southeast direction (i.e., the diagonal containing a point (i, j) is indexed as i-j). For example, the red square in the figure is in row 0, column 3, and diagonal 3.

Our goal will be to show that, in fact, r and b cannot be distinct valid labelings, due to conflicts that would arise if they were. Once established, this fact implies that a valid labeling of an  $N \times N$  square is completely determined by the values of the labeling on the square's frame of width (k+1). Thus, the number of possible valid labelings of an  $N \times N$ square is limited to the number of possible valid labelings of the square's frame, which is of order  $2^{O(N)}$ . This quantity is too small to induce a positive capacity, since  $2^{\Omega(N^2)}$  is needed.

In order to achieve our goal, we analyze the disagreement set of the two hypothetically different valid labelings of the square, and deduce that no such disagreement set can actually exist. As a first step toward this result, we examine the leftmost point in the bottommost row of the disagreement set, arbitrarily color it black, place it at the origin in a coordinate system, and proceed to make a series of additional assumptions about the location of a different point in the disagreement set colored red. We establish (Lemmas 2.2 and 2.3) that such a red point indeed exists, and, furthermore, that any such red point must lie within at most three neighboring positions of the origin in either direction along some row, column, or diagonal. We choose one of these three files and consider each of the six possible positions for the red point in that file, and determine whether the (d, d+3) constraint can be preserved when this point is added to the disagreement set. In other words, six "assumptions" are made for the chosen file.

### B. A Search Tree for Invalidating Hypothetical Labelings

We construct a *search tree* to enable us to prove that any two distinct labelings that agree on a frame of a square cannot both be valid under the hexagonal constraint. The search tree is constructed as follows.

Each node in the tree is a collection of points in a grid, which represents a *potential* disagreement subset of some disagreement set  $D_0$  (depending on the particular hypothetical labelings r and b). If this potential disagreement subset equals an actual disagreement subset of a particular  $D_0$  when the lowest-left points of the two sets are aligned, then the potential disagreement subset will be identified with the actual disagreement subset.

The root of T is constrained to contain exactly one point. For the disagreement set  $D_0$  of r and b on the  $N \times N$  square, the root is, by default, the origin of  $D_0$ , i.e., the root represents the first point of row 0 in  $D_0$ . On the other hand, we may choose the root's single point to be a point in  $D_0$  other than the origin of  $D_0$ . In either case, the lowest-left point of every node in the search tree T is aligned with and identified as the single point in the root node. (It will turn out that any nonorigin point at the root will always be chosen as a "pseudoorigin.") When the root's single point is not the origin, it will be explicitly noted for clarity.

Each node in the tree is a potential disagreement subset, which represents a possible sequence of assumptions. For a given disagreement subset, each of the six assumptions described in Section III-A will be an edge in the search tree. Such an edge is given by a starting disagreement subset, together with another disagreement subset that is obtained from the starting one by making the described assumption.

The tree provides a mechanism for invalidating any possible pair of hypothetical labelings r and b, by traversing a particular path for each such r and b. Such a path originates at the tree's root, and the root node itself is calibrated to a specific location within the disagreement set  $D_0$  of the given r and b.

For any given  $D_0$  and choice of the root of T, some of the nodes of T are disagreement subsets, and other are not. We will refer to the nodes of T as disagreement subsets, even though only some of them may be actual disagreement subsets, depending on which particular  $D_0$  is used to search the tree. Note that the topology of the tree T is fixed, and does not vary with the choice of  $D_0$ . However, any particular choice of  $D_0$  and the root of T induces a specific path through the tree T.

Each node in the tree T represents the assumptions made on the edges in the unique path from the root of the tree to that node. In particular, each non-root node of the tree corresponds to the set of assumptions of its parent node, together with the one added assumption corresponding to the edge from its parent node to itself.

In some cases, a particular assumption leads to a contradiction, and so that assumption can be eliminated. This elimination corresponds to a node in the tree having no out-edges, i.e., it is a leaf node. On the other hand, if no contradiction to the (d, k) constraint is observed, then the process is repeated at that node by choosing a particular file and then examining the six possible assumptions that can be made as out-edges of that node. This process is repeated at all non-leaf nodes until hopefully all paths terminate, which would result in a finite rooted tree, thus establishing the overall contradiction desired. Fortunately, this occurred and we present the discovered tree in this paper as our main result.

In this way, we build a search tree to represent the various sequences of assumptions made about the contents of the disagreement set, and use this tree to establish that the original positive capacity assumption was false. In other words, for any particular hypothetical pair of distinct labelings that agree on the frame of a square, we can show that at least one of the labelings is not valid under the hexagonal constraint by following a unique path (determined by the disagreement set induced by the pair of labelings) through the search tree and arriving at a contradiction at a leaf node.



Fig. 4. Search tree  $T_0$ .

Specifically, the discovered tree contains 50 internal nodes and 151 leaf nodes. Of the 151 leaf nodes, 110 of them can be eliminated quickly by immediate conflicts, i.e., the set of assumptions associated with each node contradicts the hexagonal (d, k) constraint, under the original positive capacity assumption. This leaves only 41 for more careful analysis. All but 9 of these 41 can be readily classified according to three types of relatively easy disagreement patterns. The remaining 9 are particular non-standard, more complicated, "special conflicts" that must be handled separately, but do indeed cause contradictions as well. We note that this method was implemented by hand, not with a computer.

In order to speed up the general tree building technique described above, we observed that at eight leaf nodes and one internal node, it was possible to reduce the complexity of the tree growing process using a concept of "pseudo-origins," which is described in Section V.

We also note that the detailed construction of the tree used in the proof does not depend on any particular choice of the square size N, but does depend in many places on the fact that k = d + 3.

# C. Constructing the Search Tree

The search tree is denoted by T, and its nodes are disagreement subsets. The set of disagreement points corresponding to any node in the tree is a proper subset of the set of disagreement points corresponding to each of its children in the tree. We depict the disagreement subsets using red and black squares, where a red square in a certain position means that position is labeled 1 by r, and a black square in a certain position means that position is labeled 1 by b. Note that no position in a node of T can contain both a red and black square, since these positions are in the disagreement set of r and b by assumption.

By Lemma 2.3, there are at most six possible locations for the first red disagreement point of row 0, but by symmetry we need only consider the three locations to the right of (0,0). Therefore, the root node has branches leading to these three possible positions of the first red disagreement point of row 0 (see Figure 4). We use these three disagreement subsets as root nodes for three subtrees, labeled  $T_1$ ,  $T_2$ , and  $T_3$ (see Figures 5–7).

Let D be a subset of the disagreement set  $D_0$  of the two distinct labelings r and b, and let f be a file. We say f is D-minimal if f intersects D and the first point x of f in D is



Fig. 5. Search tree  $T_1$ . Internal nodes are labeled as either R (rows), C (columns), or D (diagonals) to indicate the disagreement subset files used to make further assumptions regarding the possible (d, k) validity of the labelings r and b.

also the first point of f in the disagreement set  $D_0$  with the same color as x (see Figure 8).

If a conflict cannot immediately be found in a node, then children are added by an exhaustive procedure. First, we find a file (if it exists) that is minimal with respect to the current disagreement subset, and that contains only one point of the disagreement subset. Then, we add children corresponding to each possible location in that file for the first disagreement point of the other color, which is guaranteed to exist by Lemma 2.2. We denote these different positions by their distance offset  $\Delta$  from the first disagreement point in the given file. By Lemma 2.3, we have  $-3 \leq \Delta \leq 3$ . If such a file does not exist (which happens just once in the search tree), then the node is treated as a special conflict (see Special Conflict 1 in Lemma 2.6).

If a conflict can indeed be found in a node (either by commonly occurring configurations or by a special argument), then this node is made a leaf node of the search tree.

This process is continued until each branch terminates in a leaf node. If all leaves of a search tree violate the constraint, then the disagreement subset shown in the root node of the tree cannot occur, where the lowest-left point of the disagreement subset is the lowest-left point (or, possibly, another type of disagreement point we will call a "pseudo-origin") of the full disagreement set of the two valid labelings.

Frequently, certain  $\Delta$  values can be immediately eliminated due to a conflict with other points in the disagreement subset, such as when two positions labeled 1 are positioned closer than distance d apart. Such invalid disagreement subsets are not shown as explicit nodes in the search tree, but are instead displayed as a collection of eliminated delta values grouped in a box branching from a node.

An illustration of how internal tree nodes are handled is given in Section IV. We also provide a guide to reading the tree diagrams in Figure 9.

In Subsection A of Appendix B, conflicts in the leaf nodes are discussed in detail, along with various useful lemmas.

IV. 
$$C_{\text{HEX}}(d, d+2) = 0$$
 Whenever  $d \ge 1$ 

We illustrate our general proof technique with a relatively simple and already-known (but unpublished) case of zero hexagonal (d, k) capacity (e.g., see [19]), namely when k = d + 2. Traversing the steps of this simplified proof facilitates comprehension of the more complex general result, since they share many common ideas.



Fig. 6. Search tree  $T_2$ . Internal nodes are labeled as either R (rows), C (columns), or D (diagonals) to indicate the disagreement subset files used to make further assumptions regarding the possible (d, k) validity of the labelings r and b.

The following theorem covers all  $d \ge 3$ . The remaining cases of d = 1 and d = 2 are proven later in this section in Theorem 4.3. We note that Theorem 4.1 also follows from the stronger, but more complex, result that we prove in Theorem 2.1, for the cases when  $d \ge 7$ .

# A. $C_{hex}(d, d+2) = 0$ Whenever $d \ge 3$

Theorem 4.1: The capacity of the hexagonal (d, k) constraint is zero whenever  $d \ge 3$  and k = d + 2.

**Proof:** Suppose, to the contrary, that  $C_{\text{hex}}(d, d+2) > 0$ . Then, by Lemma 2.2, for sufficiently large N, there exist two distinct labelings, r and b, of an  $N \times N$  square that agree on the square's frame of width k+1. Let  $D_0$  be the disagreement set of r and b. The points of the square are assigned integer coordinates with the lowest-left point of  $D_0$  denoted by p and located at the origin (0,0). Without loss of generality, suppose the disagreement point p is colored black, i.e., b(0,0) = 1and r(0,0) = 0. We make finite sequences of assumptions about the contents of  $D_0$  that exhaust all possible scenarios, using a depth first search on a tree (see Figure 10) that we build for this purpose. We show that every path through this tree leads to a contradiction, implying that the original assumption of two different valid labelings was false. Thus, there cannot exist any nonempty disagreement set  $D_0$ , so, in fact, any two labelings that agree on the frame also agree on the rest of the square. This fact limits the number of possible valid labelings of squares to a growth rate which is too small to sustain a positive hexagonal (d, d + 2)capacity.

The same proof used in Lemma 2.3 also shows that for the (d, d+2) constraint, since p is the first point of the lowest row of  $D_0$ , there must be a red point  $q \in D_0$  within  $\pm 2$  positions of p in row 0 (instead of  $\pm 3$  as in the (d, d+3) constraint). However, since p is the lowest-left disagreement point, q cannot be to the left of p, so q must be located either at (1, 0) or (2, 0).



Fig. 7. Search tree  $T_3$ . Internal nodes are labeled as either R (rows), C (columns), or D (diagonals) to indicate the disagreement subset files used to make further assumptions regarding the possible (d, k) validity of the labelings r and b.



(a) For the disagreement subset  $D_1$  consisting of the three bottommost disagreement points, column 0 is  $D_1$ -minimal. On the other hand, for the disagreement subset  $D_2$  consisting of the three topmost disagreement points, column 2 is not  $D_2$ -minimal.



(b) For the disagreement subset  $D_1$  consisting of the two leftmost disagreement points, row 0 is  $D_1$ -minimal. On the other hand, for the disagreement subset  $D_2$  consisting of the four rightmost disagreement points, row 2 is not  $D_2$ -minimal.

Fig. 8. Examples of files that either are, or are not, D-minimal. In both figures, the labelings agree at any points below or to the left of the solid black border.

These two possible arrangements of the black point p and the red point q are displayed in disagreement diagrams 1.1 and 2.1, respectively.<sup>5</sup>

We next show that both arrangements of p and q shown in disagreement diagrams 1.1 and 2.1 lead to at least one of the labelings r or b violating the hexagonal (d, d + 2) constraint. To this end, we consider every possible arrangement of the points of  $D_0$  that can arise from the assumptions made in

<sup>5</sup>All disagreement diagrams for Theorem 4.1 are found in Figure 11.



Fig. 9. Illustration explaining aspects of the search tree diagrams given in Figures 4–7.

the cases shown in diagrams 1.1 and 2.1. A point of the disagreement subset is selected that is the first point of a D-minimal file, and for which there is not yet a point of the other color within  $\pm 2$  positions. Analogous to the (d, d + 3) case, such a point of the other color is guaranteed to exist in one of these locations by Lemma 2.3.

We then assume each of these four possibilities, one at a time, and show that each leads to a contradiction. For some of these assumptions a contradiction to r and b being different valid labelings appears immediately, while other assumptions are more complicated. In such cases, we add the assumption to the disagreement subset and repeat the process for this new, augmented disagreement subset. This corresponds to adding a new node to the tree with an edge from the previous disagreement subset node.

Each disagreement subset in the tree (in Figure 10) is listed in what follows, with the first few being described in detail. Specifically, the coordinates of the first disagreement point of its file are given, as well as the name of the file (i.e., the row, column, or diagonal), and the coordinates of all points within  $\pm 2$  positions of the first point within that file. These points are labeled -2, -1, 1, and 2 in Figure 12 for cases 1.1, 1.2, and 1.3. For each of these points, an explanation is given of the contradiction that arises from it being included in the disagreement set, or else a new arrangement is considered for further exploration.

• 1.1: point (0,0); diagonal 0;  $\Delta$  locations (-2,-2), (-1,-1), (1,1), (2,2).

This case corresponds to disagreement diagram 1.1, and we consider the four listed neighbors of (0,0) along the diagonal stemming from (0,0) within  $\pm 2$  of (0,0) (see Figure 12). Since (0,0) is colored black, any disagreement point at one of the listed  $\Delta$  locations must be red. However, there cannot be a disagreement point at (-2, -2) or (-1, -1), since that would contradict the point at (0,0) being the lowest-left disagreement point. Also, there cannot be a red disagreement point at (1,1), for then the red point at (1,0) would vertically violate the d constraint in r (since  $d \geq 3$ ). There is no immediate contradiction to coloring a disagreement point at (2,2)red, so we address this possibility in case 1.2.

• 1.2: point (0,0); column 0;  $\Delta$  locations (0,-2), (0,-1), (0,1), (0,2).

Since (0,0) is colored black, any disagreement point at one of the listed  $\Delta$  locations is colored red.

However, there cannot be a disagreement point at (0, -2) or (0, -1), since (0, 0) is the lowest-left disagreement point. Also, there cannot be a red disagreement point at (0, 2), for then the red point at (2, 2) would horizontally violate the d constraint in r (since  $d \ge 3$ ). There is no immediate contradiction to coloring a disagreement point at (0, 1) red, so we address this possibility in case 1.3.

• 1.3: point (1,0); column 1;  $\Delta$  locations (1,-2), (1,-1), (1,1), (1,2).

The chosen file is the column immediately to the right of the point (0,0). Since (1,0) is colored red, any disagreement point at one of the listed  $\Delta$  locations is colored black. However, there cannot be a disagreement point at (1,-2) or (1,-1), since (0,0) is the lowestleft disagreement point. Also, there cannot be a black disagreement point at (1,1), for then the black point at (0,0) would diagonally violate the *d* constraint in *b* (since  $d \geq 3$ ). There is no immediate contradiction to coloring a disagreement point at (1,2) black, so we address this possibility in case 1.4.

• 1.4: Conflict 2.

The four hollow circles indicate the squares of  $D_0$  that are used to obtain a contradiction in this case. Whereas these squares of  $D_0$  are under the hexagonal (d, d + 2)constraint, their arrangement is analogous to an example of a Conflict 2 arrangement shown in Figure 15(d) for the hexagonal (d, d + 3) constraint.

A slight modification of Lemma 2.5 yields a conflict in this case with the hexagonal (d, d + 2) constraint, i.e., at least one of the labelings r and b must not be valid. Since the arrangement of disagreement points in 1.4 does not allow both r and b to be valid labelings, the arrangement of disagreement points in 1.1 also does not allow both r and b to be valid labelings.

- 2.1: point (0,0); diagonal 0;  $\Delta$  locations (-2,-2), (-1,-1), (1,1), (2,2).
- 2.2: point (0,0); column 0;  $\Delta$  locations (0,-2), (0,-1), (0,1), (0,2).
- 2.3: point (2,0); column 2;  $\Delta$  locations (2,-2), (2,-1), (2,1), (2,2).
- 2.4: point (0,2); diagonal -2;  $\Delta$  locations (-2,0), (-1,1), (1,3), (2,4).
- 2.5: Conflict 1 (see Figure 14(a) for the analogous version of this conflict for the k = d + 3 case).

Since the two arrangements of disagreement points shown in 1.1 and 2.1 do not allow r and b to be valid labelings, the proof is complete.

B.  $C_{hex}(d, d+2) = 0$  When  $d \in \{1, 2\}$ 

If m and b are real numbers and l is a labeling of  $Z^2$ , then the set  $\{(x, y) \in Z^2 : y = mx + b\}$  is called a *line of* 1s of slope m and intercept b if l(x, y) = 1 for all (x, y) in the set.

In the proof of the following lemma, we say that a line of 1s *intersects* a string  $10^{z}1$  on the left (respectively, right) if the string is horizontal and its leftmost (respectively, rightmost) 1 is on the line of 1s.



Fig. 10. Search tree for the proof of Theorem 4.1 with the hexagonal (d, d+2) constraint. Each tree node corresponds to one of the diagrams, 1.1 - 1.4 or 2.1 - 2.5, in Figure 11, and the colored squares in those diagrams represent assumed values of the labelings r and b, under the assumption that the labelings r and b disagree. The rectangles above and to the right of internal nodes list the  $\Delta$  values corresponding to disagreement subsets that immediately result in a contradiction (i.e., these subsets are leaves of the tree). Every path from the root to a leaf results in a contradiction, thus disproving the assumption that  $C_{\text{hex}}(d, d+2) > 0$ .



Fig. 11. The figures show disagreement subsets corresponding to nodes of the search tree in Figure 10 for Theorem 4.1. The leftmost black point in the lowest row of each diagram has coordinates (0, 0).



Fig. 12. Diagrams 1.1, 1.2, and 1.3 from Figure 11 with labeled  $\Delta$  locations.

We thank Zsolt Kukorelly for some of the ideas in the following lemma.

Lemma 4.2: If a labeling satisfies the hexagonal (2,4) constraint and has a line of 1s with slope -1, 1/2, or 2 that intersects the string  $10^{3}1$  or  $10^{4}1$ , then the labeling consists entirely of parallel lines of 1s.

*Proof:* We consider each of the three slopes m separately. In each case where a line of 1s is assumed to intersect a string on the left or right, the 1 in the string that lies on the line of 1s will be assumed, without loss of generality, to lie at the origin. Let l denote the binary labeling of points in  $Z^2$ .

• Suppose m = -1.

A line of 1s cannot intersect  $10^{3}1$  on the left or right, for otherwise 101 would occur diagonally.

If the line of 1s intersects  $10^41$  on the left (respectively, right), then a line of 1s is forced, with slope -1 and intercept b = 5 (respectively, b = -5).

To see this, note that l(0,0) = l(5,0) = l(1,-1) = 1 and l(2,-1) = l(3,-1) = l(4,-1) = l(5,-1) = 0, which

implies l(6, -1) = 1 to prevent  $0^5$  horizontally, so by induction half of the line of 1s,  $\{(x, y) \in Z^2 : x+y=5\}$ , is forced downward (i.e. when  $y \leq 0$ ). Also, l(4, 0) =l(4, -1) = l(4, -2) = l(4, -3) = 0, so l(4, 1) = 1 to prevent  $0^4$  vertically. This implies that the half line of 1s extends upward to give an entire line of 1s. A symmetric argument gives the result when the line of 1s intersects  $10^41$  on the right.

• Suppose m = 1/2.

A line of 1s cannot intersect  $10^31$  on the left or right, for otherwise 101 would occur vertically. To see this, note that l(0,0) = l(4,0) = l(2,-2) = 1 implies that 101 occurs diagonally from (2,-2) to (4,0). And similarly on the other side of the line of 1s.

If the line of 1s intersects  $10^41$  on the left (respectively, right), then a line of 1s is forced, with slope 1/2 and intercept b = -(5/2) (respectively, b = (5/2)).

To see this, note that l(3,1) = l(4,1) = l(5,1) = l(6,1) = 0 implies that l(7,1) = 0 to prevent  $0^5$ 



Fig. 13. Special Conflicts. The lowest-left point of disagreement subset D is at position (0, 0) and colored black. By the assumption that all critical files for D are D-minimal, both assumed-to-be-valid labelings agree in all colored circles. A square of side length d = 7 helps visualize the conflicts, but any larger value of d results in the same conflicts for the same reasons given.

horizontally, so by induction half of the line of 1s,  $\{(x,y) \in Z^2 : y = (x/2) - (5/2)\}$ , is forced upward (i.e. when  $y \ge 0$ ). Also, since l(3,2) = l(3,1) = l(3,0) = l(3,-2) = 0, we must have l(3,-1) = 1 to prevent  $0^5$  vertically. This implies that the half line of 1s extends downward to give an entire line of 1s. A symmetric argument gives the result when the line of 1s intersects  $10^{4}1$  on the right.

• Suppose m = 2.

If the line of 1s intersects  $10^{3}1$  on the left (respectively, right), then two lines of 1s are forced, whose slopes are 2. One of them has intercept b = -8 (respectively, b = 8), and the other has intercept either b = -3 or b = -5 (respectively, b = 3 or b = 5).

To see this, note that l(2,2) = l(3,2) = l(4,2) = l(6,2) = 0 implies l(5,2) = 1, and



Fig. 14. Examples of Conflicts 1 and 2 for the hexagonal (d, d+3) constraint.

l(0,-2) = l(1,-2) = l(2,-2) = l(4,-2) = 0 implies l(3,-2) = 1, in both cases to prevent  $0^5$  horizontally. This forces a line of 1s, namely,  $\{(x,y) \in Z^2 : y = 2x - 8\}$ . If additionally l(2,1) = 1, then it is easy to see that the line of 1s,  $\{(x,y) \in Z^2 : y = 2x - 3\}$ , is forced, but alternatively if l(2,1) = 0, then the line of 1s,  $\{(x,y) \in Z^2 : y = 2x - 3\}$ , is forced, but alternatively if l(2,1) = 0, then the line of 1s,  $\{(x,y) \in Z^2 : y = 2x - 5\}$ , is forced. A symmetric argument gives the result when the line of 1s intersects  $10^{4}1$  on the right.

If the line of 1s intersects  $10^41$  on the left (respectively, right), then two lines of 1s are forced, with slopes 2, and intercepts b = -5 and b = -10 (respectively, b = 5 and b = 10).

To see this, note that l(5,3) = 0 to avoid  $0^5$  diagonally from (1,1) to (5,5), so that l(6,4) = 0 to avoid  $0^5$ horizontally from (2,2) to (6,2), and then l(4,1) = 0 to avoid  $0^5$  diagonally from (2,0) to (6,4). But l(1,-2) =l(3,0) = l(4,1) = l(5,2) = 0 implies l(2,-1) = 1to prevent  $0^5$  diagonally. It then follows that l(1,2) =l(6,2) = l(3,1) = l(-1,-2) = l(4,-2) = 1, and by induction we get the following two lines of 1s:  $\{(x,y) \in$  $Z^2 : y = 2x - 5\}$  and  $\{(x,y) \in Z^2 : y = 2x - 10\}$ .

Each of these three cases shows that starting with a line of 1s forces new lines of 1s of the same slope to its left and to its right, provided the original line of 1s intersected either  $10^{3}1$  or  $10^{4}1$ . If, instead, the line of 1s intersected  $10^{2}1$  on the left (respectively, right), then it would force a line of 1s through the rightmost (respectively, leftmost) 1 in  $10^{2}1$ . This is because if any other 1 in the line of 1s was the leftmost bit in  $10^{3}1$  or  $10^{4}1$ , then as previously shown it would force a line of 1s through the rightmost 1 in that string, contradicting the assumed string  $10^{2}1$  that intersects the line of 1s.

All of the locations between the original line of 1s and each of these new lines is labeled by 0. By induction, if this process is continued, one concludes that the labeling of the entire plane consists only of parallel lines of 1s.

Theorem 4.3: The capacity of the hexagonal (d, k) constraint is zero when  $d \in \{1, 2\}$  and k = d + 2.

*Proof:* When d = 1 and k = 3, the string 101 is forbidden horizontally, for otherwise the string  $0^4$  would occur horizontally above it. Thus  $C_{\text{hex}}(1,3) = C_{\text{hex}}(2,3) \leq C_{\text{rect}}(2,3) = 0$ . Under the (2,4) constraint, the only possible zero runs are

 $0^2$ ,  $0^3$ , and  $0^4$ . If a valid labeling does not have any runs  $0^3$  or  $0^4$ , then all

zero runs are  $0^2$ , and there are only three possibilities for the labeling of each row, and each such labeling determines the labeling everywhere else.

Next suppose there is at least one run  $0^3$  or  $0^4$  in any hexagonal (2, 4) labeling. We will next show that every valid (2, 4) labeling has at least one line of 1s with slope either -1, 1/2, or 2.

We will first consider the case when  $10^31$  appears somewhere and then the case when  $10^41$  appears somewhere. Without loss of generality, we will assume such strings are horizontal and start at the origin.

- Assume l(0,0) = l(4,0) = 1.
- Then l(3,2) = 0 to prevent  $0^6$  horizontally from (0,1) to (5,1). Since l(0,2) = l(2,2) = l(3,2) = l(4,2) = l(6,2) = 0, we must have l(1,2) = l(5,2) = 1, to prevent  $0^5$  horizontally. The original two assumptions are thus still true if shifted by (1,2), and, by symmetry about row 0, if shifted by (-1,-2) as well. Then, by induction, we deduce that the following line of 1s is forced:  $\{(x,y) \in Z^2 : y = 2x\}$ .
- Assume l(0,0) = l(5,0) = l(4,2) = 1.
- Then l(3,1) = l(4,1) = l(5,1) = l(6,1) = 0, so l(2,1) = l(7,1) = 1, to prevent  $0^5$  horizontally. Also, l(2,-1) = l(3,0) = l(4,1) = l(5,2) = 0, so l(6,3) = l(1,-2) = 1 to prevent  $0^5$  diagonally. The original three assumptions are thus still true if shifted by (2,1), i.e., l(2,1) = l(7,1) = l(6,3) = 1. Furthermore, since l(0,0) = l(5,0) = l(1,-2) = 1, the original assumptions are true if rotated about row 0. Then, by this symmetry and induction, we deduce that the following line of 1s is forced:  $\{(x,y) \in Z^2 : y = x/2\}$ .
- Assume l(0,0) = l(5,0) = l(3,2) = 1.
- This implies l(0, 1) = l(1, 1) = l(2, 1) = l(3, 1) = 0so l(-1, 1) = l(4, 1) = 1. Since l(2, 2) = l(2, 1) = l(2, 0) = l(2, -1) = 0 we must have l(2, 3) = l(2, -2) = 1. The original three assumptions are thus still true if shifted by (-1, 1), i.e., l(-1, 1) = l(4, 1) = l(2, 3) = 1. Furthermore, since l(0, 0) = l(5, 0) = l(2, -2) = 1, the original assumptions are true if rotated about row 0. Thus, by this symmetry and induction, we deduce that the following line of 1s is forced:  $\{(x, y) \in Z^2 : y = -x\}$ .
- Assume l(0,0) = l(5,0) = 1 and l(4,2) = l(3,2) = 0. Then l(2,2) = l(3,2) = l(4,2) = l(5,2) = 0, which implies l(1,2) = l(6,2) = 1. Also l(2,1) = 0 to prevent  $0^5$  vertically from (4,-1) to (4,3), and l(4,1) = 0to prevent  $0^5$  vertically from (2,-1) to (2,3). Thus, l(3,1) = 1 to prevent  $0^5$  horizontally from (0,1) to (4,1). This forces l(4,3) = 1 to prevent  $0^5$  vertically from (4,-1) to (4,3), so l(5,4) = l(4,4) = 0. The original four assumptions are thus still true if shifted by (1,2). Also, since l(2,0) = l(2,1) = l(2,2) = l(2,3) = 0, we get l(2,-1) = 1, which implies l(1,-2) = l(2,-2) = 0.

Therefore, since l(0,0) = l(5,0) = 1, the original assumptions are true if rotated about row 0. Thus, by this symmetry and induction, we deduce that the following line of 1s is forced:  $\{(x, y) \in Z^2 : y = 2x\}$ .

We have now shown that all possible strings  $10^{3}1$  and  $10^{4}1$  force a line of 1s with slope -1, 1/2, or 2. So, by Lemma 4.2 the entire labeling consists of parallel lines of 1s. This means that the labeling of any row of a rectangle induces one of at most three possible labelings of the entire rectangle, corresponding to the three possible slopes of the parallel lines of 1s. Then, since there are at most  $3 \cdot 2^N$  valid  $N \times N$  square labelings that contain at least one of  $10^{3}1$  or  $10^{4}1$ , and there are only three valid  $N \times N$  square labelings containing only  $10^{2}1$ , we have  $C_{\text{hex}}(2,4) \leq \lim_{N \to \infty} \frac{1}{N^2} \log_2(3 \cdot 2^N + 3) = 0$ .

V. Main Result:  $C_{\text{hex}}(d, d+3) = 0$  Whenever  $d \geq 3$ 

In this section we establish that  $C_{\text{hex}}(d, d+3) = 0$  whenever  $d \ge 3$  (Theorem 2.1). Of these infinite cases, as noted in Section I, it has been previously shown [20] that  $C_{\text{hex}}(d, d+3) = 0$  when  $d \in \{3, 4, 5, 7, 9, 11\}$  and also when d = 6 in our Part II. In what follows we prove the result for all  $d \ge 7$ , which suffices to complete the proof.

The proof of Theorem 2.1 relies directly on Lemmas 2.2, 2.8, and 2.11, the latter two of which are derived in this section from Lemmas 2.6, 2.9, and 2.10.

Since the lowest-left point of the disagreement set  $D_0$  of the two distinct labelings r and b is defined to lie at the origin (0,0), the labelings r and b agree at certain positions, such as:

(i) at any point in row 0 to the left of (0,0).

(ii) at any point (x, y) satisfying  $x \le 4$  and  $y \le -1$ .

Similarly, we call a point  $p = (x, y) \in D_0$  a *pseudo-origin* if the labelings r and b agree in the following positions:

(i) at any point in row y to the left of p.

(ii) at any point (x', y') satisfying  $x' \le x + 4$  and  $y' \le y - 1$ .

In the construction of the search tree T, certain files of disagreement subsets are particularly useful in the analysis. For example, each edge of T corresponds to a specific file used for expansion, and this file is marked above the node in the tree diagrams illustrated in Figures 4–7. Such files in this category are used to examine six possible assumptions for the position of a particular disagreement point, corresponding to the six locations described in Lemma 2.3. Another important category of files used in the construction of T consists of those used to achieve conflicts at the leaf nodes. These are enumerated in Table II. We call these two categories of files critical, as described in the following definition.

Definition 5.1: A file f is critical for a disagreement subset D if f corresponds to an edge from the node D in the tree T or is used to demonstrate a conflict in the leaf D of the tree T.

It is noted that if a file is critical for disagreement subset D, then D contains at least one point in that file.

In the exhaustive search of disagreement subsets using the search tree T, we often find conflicts with the constraint

TABLE II The Critical Files Used to Demonstrate Conflicts at Leaf Nodes in the Search Tree  ${\cal T}$ 

Leaf node	critical files							
1.5	Columns 0,1,2; Diagonals -1,0,1							
1.7	Columns 0,1,4; Diagonals -1,0,1							
1.10	Columns 0,1,2; Diagonals -1,0,1							
1.14	Columns 0,1,2; Diagonals -1,0,1							
1.15	Diagonals -1,0							
1.16	Diagonals 0,1							
1.17	Diagonals 0,1							
1.21	Column 2							
1.22	Diagonals 0,1							
1.23	Columns 0,1							
1.29	Columns 0,1,2; Diagonals -3,-1							
1.30	Diagonals -3,-1,0,1							
1.31	Column 0; Diagonal 1							
1.32	Columns 0,1							
2.6	Diagonals 0,2							
2.7	Diagonals -1,0							
2.8	Diagonals -1,0							
2.12	Diagonal -2							
2.13	Row 2							
2.16	Diagonals 1,2							
2.17	Diagonals 1,2							
2.20	Columns 2,3							
2.22	Row 2							
2.23	Columns 0,2; Diagonals 0,2							
2.24	Diagonals 0,2							
2.30	Diagonal -3							
2.32	Diagonal 2							
2.33	Columns 0,1							
2.34	Columns 0,1							
3.4	Rows 0,1							
3.6	Row 1							
3.7	Diagonal -1							
3.10	Rows 0,1							
3.11	Rows 0,2							
3.14	Rows 0,1							
3.16	Columns 0,1							
3.17	Columns 0,1							
3.21	Rows 0,2							
3.22	Diagonals 2,3							
3.24	Column 4							
3.25	Column 0: Diagonal 3							

relatively close to the lowest-left point of a disagreement subset. Specifically, the indices of rows, columns, and diagonals that are critical for some disagreement subset in the constructed search tree T are fortunately at most 2, 4, and 4, respectively. This upper bound for the critical columns and diagonals motivates our definition of pseudo-origin, and leads to the following lemma.

Lemma 5.2: Let r and b be distinct labelings of a square that satisfy the hexagonal (d, d+3) constraint with  $d \ge 7$  and agree on the square's frame of width k+1. If x is a pseudoorigin and D is a potential disagreement set in the search



Fig. 15. Diagrams (a)–(f) show Conflict 2 arrangements. The same arrangements with the colors red and black switched also cause conflicts Diagrams (g)–(l) show arrangements that do not immediately cause conflicts. All arrangements apply to the hexagonal (d, d + 3) constraint.

tree T with root x, then row 0 and any critical columns and diagonals for D are D-minimal.

*Proof:* First note that if the two distinct labelings r and b agree in all positions of a file up to a point p, then any disagreement point within the first 8 positions after p in the file is the first disagreement point of its color in the file. This is because two points of the same color must be separated by at least  $d \ge 7$  positions, and so two such points cannot be contained within 8 consecutive positions.

Since x is a pseudo-origin, the two labelings agree in the region containing all points to the left of x, and all points below the row containing x in columns or diagonals with index at most 4. Then by inspecting every potential disagreement subset D in the search tree T, it can be verified that for row 0 or any column or diagonal with index at most 4 that intersects D, the first disagreement points of both colors of such a file f in D occur within the first 7 positions after the last agreement point of f. Therefore, by the first paragraph, these first disagreement points of each such file in D are the first disagreement points of each such file in D. Thus any such file (including, in particular, row 0) is D-minimal, and since any critical column or diagonal has index at most 4, the lemma is proved.

In the proof of Lemma 2.6, use is made of various diagrams in Figure 13. Before stating the lemma, we explain in detail how these diagrams are used.

In each special conflict diagram in Figure 13, the red and black squares indicate disagreement points as previously discussed. In contrast, the circles represent certain important agreement points of the labelings r and b. The circles in each diagram are contained in some critical column or diagonal of the disagreement subset in the same diagram, and these critical files are listed in each diagram's caption. Specifically, the circles in a given column or diagonal denote the possible locations of the last agreement point labeled 1 in that file before any point in the disagreement subset. Additionally, among the circles corresponding to critical columns, the values of r and b do not change within a given color of circle, and the same holds true for the circles corresponding to critical diagonals.

Each critical file contains exactly two disagreement points, one red and one black. The lowest-left black square is always a pseudo-origin, so the two labelings r and b agree at all points below its row (i.e., row 0) when the column index is 4 or less. The colored circles in the diagrams lie in these agreement regions and are associated with critical files.

In any critical file, to satisfy the k constraint, a point labeled 1 must occur within the (k + 1) positions before the second of the two shown disagreement points in that file. But since  $d \geq 7$  (by assumption) and the largest row index of any disagreement point in a critical file is 6, the d constraint implies this point labeled 1 must occur in the agreement region of the two labelings. Then by the d constraint applied to the first of the two shown disagreement points in the critical file, this agreement point labeled 1 must occur at least d positions away from this first disagreement point. Since k = d + 3, this leaves a window of at most 3 positions for such an agreement point labeled 1, and since  $d \ge 7$  by assumption, exactly one of the positions in this window must be labeled 1. The circles in each critical file occupy the positions in this window. However, the various configurations of circles do not allow the unique circle labeled 1 in each critical file to be arbitrarily chosen while satisfying the hexagonal (d, k) constraint.

By logical deduction in each Special Conflict diagram, one can verify that, of the possible labelings of all of the circles in the critical columns, either 0, 1, or 2 labelings satisfy the d constraint. If zero such labelings satisfy the d constraint, we depict the circles as hollow with no coloring. If exactly one labeling satisfies the d constraint, we color circles orange to represent a label of 1, and green to represent a label of 0. If two labelings satisfy the d constraint, which happens only in Special Conflict 1, it suffices to color the circles labeled 1 orange in one labeling and blue in the other labeling. In both cases, we color circles green when the point is labeled 0 in every labeling. We then repeat this colorization process for the circles in critical diagonals.

As a result, Special Conflict 1 contains two possible labelings for both critical columns and diagonals, and Special Conflict 6 contains zero possible labelings of critical diagonals. The other 7 Special Conflicts contain a unique labeling for the circles in critical columns, and a unique labeling for the circles in critical diagonals.

Following this coloration, violations of the d constraint or kconstraint could occur due to the possible labelings of the circles in the critical columns and diagonals. We describe such a violation in each Special Conflict, and deduce a contradiction.

To illustrate the coloration process, consider the diagram corresponding to Special Conflict 2. The diagram's caption indicates that the four critical files are columns 0,1, and 4, together with diagonal 1. Considering the columns first, the single circle in column 4 must be labeled 1, which then implies the bottom circle in column 1 must be labeled 1 to avoid violating the d constraint along a diagonal, and this in turn implies the top circle in column 0 must also be labeled 1. In the only critical diagonal, the single circle must be labeled 1. This fixes the coloring of the circles based on the labelings. Then, a conflict can be seen in the row immediately below the black  $d \times d$  square outline, since the two orange circles in that row have only (d-1) positions labeled 0 between them.

The figures assume d = 7 for these special conflicts, as seen in the  $7 \times 7$  square outline, but, in fact, any value of  $d \ge 7$  causes a conflict for the same reasons (to be given in Subsection B of Appendix B) by considering a larger square outline with side length d. Specifically, the  $d \times d$ square outline shown in each image can be used to apply the arguments for any  $d \ge 7$  by examining the positions of the circles relative to each other.

Some technical lemmas used in the proof of the main result are given in Subsection B of Appendix B.

The main result then readily follows from Lemmas 2.2, 2.8, and 2.11.

Proof of Theorem 2.1: In Part II of this two-part series [8], we prove that  $C_{\text{hex}}(d, d+3) = 0$  when  $3 \le d \le 11$ , which overlaps with the cases  $7 \le d \le 11$  shown here.

Let  $d \ge 7$  and k = d + 3 and suppose to the contrary that  $C_{\text{hex}}(d,k) > 0$ . Then by Lemma 2.2, there exist two distinct valid labelings r and b of a sufficiently large  $N \times N$  square that agree on the square's frame of width k + 1.

By Lemma 2.11, there exists a pseudo-origin x in the disagreement set of r and b with the row minimality property (see Definition 2.7 in Appendix B). Then by Lemma 2.8, there is a conflict with the constraint in either labeling r or b, so it cannot be true that both labelings are valid, a contradiction. Thus,  $C_{\text{hex}}(d, d+3) = 0$  for all  $d \ge 7$ .

#### APPENDIX A

The figures below (labeled 1.1–1.32, 2.1–2.34, 3.1–3.25) show disagreement subsets corresponding to nodes of the search tree for the main result in Section V.

## APPENDIX B

## A. Conflicts in Leaf Nodes

In this subsection we include lemmas that are used in the proof of the main theorem. In particular, we focus on establishing conflicts at leaf nodes that help complete our proof by contradiction.

Lemma 2.1: Let r and b be distinct labelings of a square that satisfy the hexagonal (d, k) constraint and agree on the

square's frame of width  $\delta > k + 1$ . Then every file of length at least 2(k+1) has at least two positions where both r and b are labeled 1. In particular, at least one of these positions comes before every disagreement point in the file, and another of these positions comes after every disagreement point in the file.

*Proof:* Let f be a file of length at least 2(k+1). In both r and b, one of the first (k+1) positions of f must be labeled 1 or else the labelings would violate the k constraint. But since r and b agree on the frame, which has width (k+1), r and b must agree at such a position labeled 1 within the first (k+1)positions of f. Similarly for the last (k+1) positions.

The following lemma shows that if a file contains a disagreement point labeled 1 by one labeling, then the file also contains a disagreement point labeled 1 by the other labeling.

Lemma 2.2: Let r and b be distinct labelings of a square that satisfy the hexagonal (d, d+3) constraint with  $d \ge 3$  and agree on the square's frame of width k+1. For each file f of the square, if f contains a disagreement point of a particular color, then f also contains a disagreement point of the other color.

*Proof:* Suppose f is a file that intersects the disagreement set such that x is the first disagreement point of f. Without loss of generality, suppose x is black. Suppose there does not exist a red disagreement point in f. Then for any point  $z \in f$ , b(z) = 0 implies r(z) = 0.

Since b(x) = 1, the d constraint implies there are at least d positions on both sides of x in f labeled 0 by b. Therefore, there are also at least d positions on either side of x in flabeled 0 by r. Then since r(x) = 0, r must have a run of at least 2d + 1 consecutive 0s in f. But  $d \ge 3$  implies 2d+1 > d+3 = k, and so the run of 2d+1 consecutive 0s in r violates the k constraint, a contradiction.

As previously mentioned, the quantity  $\Delta$  denotes the distance offset from the first disagreement point in a given file to the point of the other color (guaranteed to exist by Lemma 2.2) in the same file. The following lemma establishes that if the file is D-minimal, then  $-3 \le \Delta \le 3$ .

Lemma 2.3: Let r and b be distinct labelings of a square that satisfy the hexagonal (d, d+3) constraint with d > 5and agree on the square's frame of width k + 1. Let D be a subset of the disagreement set  $D_0$ , and let f be a file that is Dminimal. Let  $x \in D$  be the first point of f in D. Then the first point of f in  $D_0$  of opposite color to x is located in one of six possible locations, namely within  $\pm 3$  positions from x in f.

*Proof:* Without loss of generality, let x be colored black. Since f is D-minimal, x is the first black disagreement point of f in  $D_0$ . By Lemma 2.2, there exists a red disagreement point in f; let y be the first such point. Without loss of generality, we can assume y is after x in f (if not, we can switch the roles of red and black).

Consider the last position z before x where r(z) = 1and b(z) = 1. There must exist at least one such z by Lemma 2.1. By the d constraint in labeling b, there must be d positions labeled 0 between z and x, and by the k constraint in labeling r, there cannot be more than d+3 zeros between z and y. Therefore, x and y cannot be separated by more than 3 positions.

145



Given a disagreement subset  $D \subseteq D_0$ , the following lemma shows that for a file f that is not D-minimal, there exist at least two points (of different colors) of f in  $D_0$  before the first point of f in D.

Lemma 2.4: Let r and b be distinct labelings of a square that satisfy the hexagonal (d, d+3) constraint with  $d \ge 7$  and agree on the square's frame of width k + 1. Let  $D_0$  be the disagreement set of r and b, and let  $D \subseteq D_0$ . Suppose there exists a file f intersecting D that is not D-minimal, and let z be the first point of f in D. Then there exist at least two points (of different colors) of f in  $D_0$  before z.

**Proof:** Without loss of generality, suppose z is colored black. Since f is not D-minimal, z is not the first black point of f in  $D_0$ . Therefore, there exists another point x before z that is the first black disagreement point of f. Furthermore, since  $d \ge 7$ , x is at least 7 positions before z.



By Lemmas 2.2 and 2.3, there exists a red disagreement point y that is at most 3 positions away from x, which guarantees y occurs before z. Thus, x and y are two points of different colors of f in  $D_0$  before z.

The proof of Lemma 2.4 in fact applies to the stronger case where  $d \ge 3$ , but we need only  $d \ge 7$  for our analysis.

A disagreement subset D may contain an arrangement of points that causes at least one of r or b to violate the hexagonal (d, k) constraint, provided that certain files containing these points are D-minimal. We call such arrangements *conflicts*. In particular, the following three types of conflicts arise often in our proofs, and we will refer to them as Conflict 1, Conflict 2, and Conflict 3. Examples of these conflicts are shown in Figure 14.

• Conflict 1.

In this arrangement, the first disagreement points of two parallel files in a disagreement subset D are arranged as shown, for example, in Figure 14a. In addition, these files must be D-minimal and separated by fewer than d files. Conflict 2

• <u>Conflict 2</u>.

In this arrangement, the first disagreement points of two adjacent D-minimal files in a disagreement subset D are arranged as shown, for example, in Figure 14b. Figure 15 provides a full catalog of possible Conflict 2 arrangements, as well as arrangements of disagreement points that may resemble Conflict 2, but are not.

• Conflict 3.

In this arrangement, the first disagreement points of opposite color of a D-minimal file in a disagreement subset D are separated by 3 positions.

The following lemma shows that the arrangements of points in Conflict 1, Conflict 2, and Conflict 3 all do indeed cause at least one of the labelings r and b to violate the hexagonal (d, k) constraint.

Lemma 2.5: Let D be a subset of the disagreement set of two distinct valid labelings r and b of an  $N \times N$  square that agree on the square's frame of width k + 1. Then the arrangements of points in Conflict 1, Conflict 2, and Conflict 3 each cause at least one of r and b to violate the hexagonal (d, k) constraint.

*Proof:* Each of the three types of conflicts previously defined are examined to establish the lemma.

- Conflict 1.

In Figure 14a, let the bottom two points be in row 0 and the upper two points be in row i, and suppose  $i \leq d$ . By Lemma 2.1, in each of rows 0 and i there exists a point before the displayed points where both labelings equal 1. Let the last positions where both labelings equal 1 before the displayed points in rows 0 and i be denoted  $p_0$  and  $p_i$ , respectively.

Since rows 0 and *i* are *D*-minimal and the displayed points are the first disagreements points of their rows in *D* (as required by Conflict 1), the displayed points are the first disagreement points of their rows in the disagreement set of the two labelings. Therefore, the only possible column that can contain  $p_0$  and  $p_i$  is the column that is separated by exactly *d* columns from the column containing the leftmost disagreement point in each row. However, since row 0 and row *i* are separated by i-1 < d rows,  $p_0$  and  $p_i$  are separated by fewer than *d* rows. Therefore this arrangement causes a conflict with the *d* constraint.

Similar arguments show the lemma in cases where the disagreement points are in columns or diagonals instead of rows.

- Conflict 2.

In Figure 14b, let the bottom two points be in row 0 and the upper two points be in row 1. By Lemma 2.1, in each of rows 0 and 1 there exists a point before the displayed

points where both labelings equal 1. Let the last positions where both labelings equal 1 before the displayed points in rows 0 and 1 be denoted  $p_0$  and  $p_1$ , respectively.

Since rows 0 and i are D-minimal and the displayed points are the first disagreements points of their rows in D (as required by Conflict 2), the displayed points are the first disagreement points of their rows in the disagreement set of the two labelings. Therefore, the only possible column that can contain  $p_0$  is the column that is separated by exactly d columns from the column containing the leftmost disagreement point in row 0. Also, the only possible columns that can contain  $p_1$  are the columns that are separated by exactly d or exactly (d + 1) columns from the column containing the leftmost disagreement point in row 1. However, both of these positions for  $p_1$ cause a conflict with the d constraint, since  $p_0$  would be either vertically or diagonally adjacent to  $p_1$ . Therefore this arrangement causes a conflict with the d constraint. Similar arguments show the lemma in cases displayed in diagrams (a)-(f) in Figure 15.

- Conflict 3.

This arrangement of points causes a conflict since the first disagreement points of opposite color of a D-minimal file in a disagreement subset D can be separated by at most 2 positions, by Lemma 2.3.

# Example 1: Configuration 3.10 in Appendix A.

The four points of r and b that are involved in the conflict are labeled with white dots. They are arranged according to Conflict 2 described in Lemma 2.5, so the arrangement shown in this configuration causes at least one of r and b to violate the hexagonal (d, k) constraint.

Example 2: Configuration 3.25 in Appendix A.

This configuration does not contain a commonly occurring conflict, and so we treat it as a special conflict in Lemma 2.6.

## B. Lemmas for Main Result in Section V

Lemma 2.6: Let r and b be distinct labelings of a square that satisfy the hexagonal (d, d + 3) constraint with  $d \ge 7$ and agree on the square's frame of width k + 1. Then all disagreement subsets shown as Special Conflicts in Figure 13 cause at least one of r or b to violate the hexagonal (d, d + 3)constraint.

*Proof:* In each Special Conflict diagram in Figure 13, let the row directly below the square outline be row i (rows decrease moving downward).

*Special Conflict 1:* Under the given assumptions, there are two possible valid labelings of the circles in the critical columns, and also two possible valid labelings of the circles in the critical diagonals. The green circles denote positions labeled 0 in all possible labelings,

For the circles in critical columns, either orange circles are labeled 0 and blue circles are labeled 1, or vice versa. The same property holds for critical diagonals. However, the labeling associated with orange column circles does not have to agree with the labeling associated with orange diagonal circles (similarly for blue). To demonstrate that there is a conflict caused by this disagreement subset, we show that all four pairings of these orange and blue arrangements of positions labeled by 1s generate a conflict.

- (*Diagonal Orange* = 1, *Column Orange* = 1) The 1 in row (*i* + 1) has a run of (*d* + 5) positions labeled 0 to the right. This violates the *k* = *d* + 3 constraint.
- (Diagonal Orange = 1, Column Blue = 1) The two 1s in row i are separated by (d-1) positions. This causes a conflict with the d constraint.
- (Diagonal Blue = 1, Column Orange = 1) The two 1s in row (i 1) are separated by (d + 4) positions labeled 0. This causes a conflict with the k = d + 3 constraint.
- (Diagonal Blue = 1, Column Blue = 1) The 1 in row (i + 1) has a run of (d + 5) positions labeled 0 to the left. This violates the k = d + 3 constraint.

Special Conflict 2: The two 1s in row i (i.e., the two orange circles immediately below the square outline) are separated by a distance of (d - 1). This causes a conflict with the d constraint.

Special Conflict 3: The 1 in row (i+1) has a run of (d+5) positions labeled 0 to the right. This violates the k = d+3 constraint.

Special Conflict 4: The 1 in row (i-2) has a run of (d+4) positions labeled 0 to the right. This violates the k = d+3 constraint.

Special Conflict 5: The 1 in row (i+2) has a run of (d+4) positions labeled 0 to the right. This violates the k = d+3 constraint.

Special Conflict 6: Diagonals -1 and -3 each contain a single position that must be labeled 1, denoted by hollow circles. If one were to try labeling these positions with 1s, then three of the hollow circles below (in diagonals 0 and 1) would be labeled with 0s, by the *d* constraint. But then diagonals 0 and 1 would each contain a single remaining position that must be labeled 1, and these positions are vertically adjacent. Labeling both of these positions by a 1 would therefore cause a conflict with the *d* constraint.

Special Conflict 7: The two 1s in row i are separated by a distance of (d-1). This causes a conflict with the d constraint.

Special Conflict 8: The two 1s in row (i-1) are separated by a distance of (d-1). This causes a conflict with the dconstraint.

Special Conflict 9: The two 1s in row i are separated by a distance of (d-3). This causes a conflict with the d constraint.

Definition 2.7: We say that a pseudo-origin  $x \in D_0$  has the row minimality property if for any  $D \subseteq D_0$  in the search tree T with root x, every critical file for D is D-minimal.

Lemma 2.8: Let r and b be distinct labelings of a square that satisfy the hexagonal (d, d+3) constraint with  $d \ge 7$  and agree on the square's frame of width k + 1. Suppose there exists a pseudo-origin x in the disagreement set of r and b with the row minimality property. Then at least one of r and b conflicts with the constraint.

*Proof:* We traverse the unique path (determined by the disagreement set) of the search tree T with root x. Since x has the row minimality property, every critical file for each

disagreement subset in the search tree T with root x is minimal for that disagreement subset.

At each non-leaf node, we choose one particular file and consider the six possible cases (i.e.,  $\Delta = \pm 1, \pm 2, \pm 3$ ) required by Lemma 2.3. Note that the file chosen for each node is displayed above the node in the tree diagrams in Figures 4–7, and the disagreement subsets corresponding to the nodes are shown in Appendix A. Certain values of  $\Delta$  corresponding to relatively easy conflicts are shown in boxes above the nodes, but for the remaining values of  $\Delta$ , out-edges are shown leading to other nodes in the tree.

At six particular leaf nodes (namely, 1.21, 2.12, 2.30, 2.32, 3.7, 3.24), we choose a file and show that all six possible values of  $\Delta$  lead to conflicts. At the remaining leaf nodes, we establish a conflict in the given disagreement subset by using the disagreement points in more than one file. Thus, in any case, all leaf nodes in the tree T lead to contradictions with the hexagonal (d, d + 3) constraint, and therefore the original assumption of positive capacity cannot be true.

The conflicts established at the leaf nodes rely on the fact that all files in Table II for a disagreement subset D are D-minimal, since the pseudo-origin x has the row minimality property. The examination of the six possible  $\Delta$  values in a given file at tree nodes also relies on pseudo-origin x having the row minimality property, since D-minimality of the file is required in Lemma 2.3.

Section III-C describes the steps used to exhaustively build these search trees in more detail, and Subsection A of Appendix B describes the verification of contradictions at the tree leaves, as proven in Lemma 2.5 and Lemma 2.6. In particular, Lemma 2.6 verifies the contradictions to the hexagonal (d, k)constraint for the 7 special conflicts of tree  $T_1$ , the one special conflict of  $T_2$ , and the one special conflict of  $T_3$ .

The hypothesis of Lemma 2.8 assumes the existence of a pseudo-origin  $x \in D_0$  with the row minimality property. The following lemmas establish that such a point x indeed exists. They rely on the previously stated fact that the values of the indices of the rows, columns, and diagonals that are critical for some disagreement subset in the search tree T are upper bounded by 2, 4, and 4, respectively.

Lemma 2.9: Let r and b be distinct labelings of an  $N \times N$  square with disagreement set  $D_0$  that satisfy the hexagonal (d, d + 3) constraint, with  $d \ge 7$ , and which agree on the square's frame of width k + 1. Let  $c_m$  be the column index of the leftmost point of  $D_0$  in row m. Let  $D \subseteq D_0$  be a disagreement subset in the search tree T whose root is a pseudo-origin. Suppose row j is critical for D, but is not D-minimal, and let (i, j) be the leftmost point of row j in D. Then the leftmost point of row j = 1.

*Proof:* By Lemma 2.4 there exist both red and black points in row j in  $D_0$  to the left of the point (i, j). Let y be the leftmost point of row j in  $D_0$ , and let z be the leftmost point of the other color between y and (i, j). If y and (i, j) are colored the same, then the column index of y satisfies  $c_j \leq i - d - 1$ by the d constraint. If y and (i, j) are colored differently, then the column index of z is less than or equal to i - d - 1 by the d constraint. Therefore, in either case, since  $c_i$  is less than the column index of z, we have  $c_j \leq i - d - 1 \leq i - 8$ .

Since y is the leftmost disagreement point of row j in  $D_0$ , the two labelings agree at all points to the left of y in row j. Whenever  $0 \le m \le j-1$ , the integer  $c_m$  is the column index of the leftmost point of row m in  $D_0$ , and so the two labelings agree at all points in each row m to the left of column  $c_m$ . Therefore, since  $c_i + 4 \leq i - 4 < c_m$  whenever  $0 \leq m \leq m$ j-1, the two labelings agree at points with row index m, for  $0 \le m \le j-1$ , and with column or diagonal index less than or equal to  $c_i + 4$ . The two labelings agree in these columns and diagonals at all points below row 0 as well, since the lowest-left point of D is a pseudo-origin. Therefore, y is a pseudo-origin.

The following lemma shows that for any  $x \in D_0$  that is a pseudo-origin without the row minimality property, there exists another pseudo-origin that is above and to the left of x in  $D_0$ . This property will be exploited in an inductive argument in Lemma 2.11.

Lemma 2.10: Let r and b be distinct labelings of an  $N \times$ N square with disagreement set  $D_0$  that satisfy the hexagonal (d, d+3) constraint, with  $d \ge 7$ , and which agree on the square's frame of width k + 1. Let  $x \in D_0$  be a pseudoorigin without the row minimality property. Then there exists a pseudo-origin in  $D_0$  that is above and to the left of x.

*Proof:* For any row m, let  $c_m$  be the column index of the leftmost point of row m in  $D_0$ . By Lemma 5.2, since x is a pseudo-origin, row 0 and any critical columns and diagonals in any  $D \subseteq D_0$  in the search tree T with root x are D-minimal. Therefore, since x does not have the row minimality property, there exists  $D \subseteq D_0$  in the search tree T with root x but for which there exists a critical row that is not D-minimal. The nodes in the search tree where rows are critical (as opposed to columns or diagonals) are cases 2.13, 2.22, 3.4, 3.6, 3.10, 3.11, 3.14, 3.21, and in the expansion from 2.28 (in Appendix A).

In each of the subsets  $D \subseteq D_0$  in the following itemized cases, row 0 is D-minimal because the lowest-left point of Dis a pseudo-origin, and so the labelings r and b agree to the left of that point. Therefore, by Lemma 2.9, to show a row i > 0 has a pseudo-origin, it suffices to check that  $i - 4 < c_m$ , whenever  $0 \le m \le j-1$ , where i is the column index of the leftmost point of row j in D.

Figure 16 can be used for visualization in the following cases.

# • Configurations 2.13 and 2.22

Let D be the disagreement subset in one of configurations 2.13 or 2.22. The critical row for D is row 2, so suppose row 2 is not *D*-minimal. There are two cases to consider: row 1 is *D*-minimal, or row 1 is not *D*-minimal.

- Case 1

If row 1 is D-minimal, then the point in row 1 (i.e., at position (1,1) in D is one of the leftmost two points of row 1 in  $D_0$ . So either this point is the leftmost disagreement point of row 1, or the leftmost disagreement point of row 1 is at most 3 positions to the left of this point. In either situation (and in either choice of configuration),  $c_1 \ge -2$ . The column index *i* of the leftmost point of row 2 in D is 0, and  $c_0 = 0$ .

So  $i - 4 = -4 < -2 \leq c_1$  and  $i - 4 < c_0$ , and so the leftmost point of row 2 in  $D_0$  is a pseudo-origin by Lemma 2.9.

# - Case 2

Alternatively, suppose row 1 is not *D*-minimal. The column index i of the leftmost point of row 1 in D (in either choice of configuration) is at most 2, and  $c_0 = 0$ . So  $i - 4 \le -2 < 0 = c_0$ , and so the leftmost point of row 1 in  $D_0$  is a pseudo-origin by Lemma 2.9.

# • Expansion from configuration 2.28

When we add children to the search tree from 2.28 by expanding on row 2, we are assuming that row 2 is D-minimal. So suppose row 2 is not D-minimal. There are two cases to consider: row 1 is D-minimal, or row 1 is not *D*-minimal.

- Case 1

If row 1 is *D*-minimal, then the shown points in row 1 constitute the leftmost two points of row 1 in  $D_0$ . Therefore,  $c_1 = 1$ . The column index *i* of the leftmost point of row 2 in D is 1, and  $c_0 = 0$ . So  $i - 4 = -3 < 1 = c_1$  and thus  $i - 4 < c_0$ , and so the leftmost point of row 2 in  $D_0$  is a pseudo-origin by Lemma 2.9.

- Case 2

Alternatively, suppose row 1 is not D-minimal. The column index i of the leftmost point of row 1 in Dis 1, and  $c_0 = 0$ . So  $i - 4 = -3 < 0 = c_0$ , and so the leftmost point of row 1 in  $D_0$  is a pseudo-origin by Lemma 2.9.

# Configurations 3.4, 3.10, and 3.14

Let D be the disagreement subset in one of configurations 3.4, 3.10, or 3.14. The critical rows for D are row 0 and row 1. Row 0 is D-minimal, so suppose row 1 is not D-minimal. The column index i of the leftmost point of row 1 in D is at most 1 in any of the three configurations, and  $c_0 = 0$ . So  $i - 4 \le -3 < 0 = c_0$ , and so the leftmost point of row 1 in  $D_0$  is a pseudo-origin by Lemma 2.9.

# **Configuration 3.6**

Let D be the disagreement subset in configuration 3.6. The only critical row for D is row 1, so suppose row 1 is not D-minimal. The column index i of the leftmost point of row 1 in D is 0, and  $c_0 = 0$ . So  $i - 4 = -4 < 0 = c_0$ , and so the leftmost point of row 1 in  $D_0$  is a pseudoorigin by Lemma 2.9.

# • Configuration 3.11

Let D be the disagreement subset in configuration 3.11. The critical rows for D are rows 0 and 2. Row 0 is D-minimal, so suppose row 2 is not D-minimal. There are two cases to consider: row 1 is D-minimal, or row 1 is not D-minimal.

- Case 1.

If row 1 is D-minimal, then the point in row 1 in D is one of the leftmost two points in row 1 in  $D_0$ . So either this point is the leftmost disagreement point of row 1, or the leftmost disagreement point of row 1 is at most 3 positions to the left of this point in row 1.





Fig. 16. Images depicting the cases in Lemma 2.10. The point labeled Y is the rightmost point of its row that could be a pseudo-origin. The green area indicates the region where the labelings agree, which shows that the point labeled Y satisfies the requirements of being a pseudo-origin (as long as the labelings agree at any point to the left of the point labeled Y). The value of d used in the figures is 7, but any larger value of d would push the point labeled Y even farther to the left.

In either situation,  $c_1 \ge -2$ . The column index *i* of the leftmost point of row 2 in *D* is 0, and  $c_0 = 0$ . So  $i - 4 = -4 < -2 \le c_1$  and  $i - 4 < c_0$ , and so the leftmost point of row 2 in  $D_0$  is a pseudo-origin by Lemma 2.9.

- Case 2.

Alternatively, suppose row 1 is not *D*-minimal. The column index *i* of the leftmost point of row 1 in *D* is 1, and  $c_0 = 0$ . So  $i - 4 = -3 < 0 = c_0$ , and so the leftmost point of row 1 in  $D_0$  is a pseudo-origin by Lemma 2.9.

# • Configuration 3.21

Let D be the disagreement subset in configuration 3.21. The critical rows for D are rows 0 and 2. Row 0 is D-minimal, so suppose row 2 is not D-minimal. There are two cases to consider: row 1 is *D*-minimal, or row 1 is not *D*-minimal.

## - Case 1.

If row 1 is *D*-minimal, then the point in row 1 in *D* is one of the leftmost two points in row 1 in  $D_0$ . So either this point is the leftmost disagreement point of row 1, or the leftmost disagreement point of row 1 is at most 3 positions to the left of this point. In either situation,  $c_1 \ge 0$ . The column index *i* of the leftmost point of row 2 in *D* is 2, and  $c_0 = 0$ . So  $i - 4 = -2 < c_0 \le c_1$ , and so the leftmost point of row 2 in  $D_0$  is a pseudo-origin by Lemma 2.9. - *Case 2*.

Alternatively, suppose row 1 is not D-minimal. The column index i of the leftmost point of row 1 in D

is 3, and  $c_0 = 0$ . So  $i - 4 = -1 < 0 = c_0$ , and so the leftmost point of row 1 in  $D_0$  is a pseudo-origin by Lemma 2.9.

Lemma 2.11: Let r and b be distinct labelings of an  $N \times N$  square with disagreement set  $D_0$  that satisfy the hexagonal (d, d + 3) constraint, with  $d \ge 7$ , and which agree on the square's frame of width k + 1. Then there exists a pseudo-origin in  $D_0$  with the row minimality property.

*Proof:* Let x be the lowest-left point of  $D_0$ , and without loss of generality suppose its color is black and that it is located at position (0,0). Clearly x is a pseudo-origin since x is the lowest-left point of  $D_0$ .

Suppose x does not have the row minimality property. Then by Lemma 2.10, there exists another pseudo-origin that is above and to the left of x in  $D_0$ . If this process is repeated, then either a pseudo-origin with the row minimality property will be found in a finite number of steps (since  $D_0$  has height less than N), or else the highest row containing a pseudoorigin of  $D_0$  will be reached. The pseudo-origin with greatest row index must have the row minimality property, or else Lemma 2.10 would show the existence of another pseudoorigin in a higher row, a contradiction.

## REFERENCES

- [1] R. J. Baxter, "Hard hexagons: Exact solution," J. Phys. A, Math. Gen., vol. 13, pp. 1023–1030, Mar. 1980.
- [2] R. J. Baxter, Exactly Solved Models in Statistical Mechanics. New York, NY, USA: Academic, 1982.
- [3] R. J. Baxter, "Planar lattice gases with nearest-neighbor exclusion," Ann. Combinatorics, vol. 191, pp. 191–203, Jun. 1999.
- [4] R. J. Baxter and S. K. Tsang, "Entropy of hard hexagons," J. Phys. A, Math. Gen., vol. 13, no. 3, pp. 1023–1030, Mar. 1980.
  [5] N. J. Calkin and H. S. Wilf, "The number of independent sets in
- [5] N. J. Calkin and H. S. Wilf, "The number of independent sets in a grid graph," *SIAM J. Discrete Math.*, vol. 11, no. 1, pp. 54–60, Feb. 1998.
- [6] K. Censor and T. Etzion, "The positive capacity region of twodimensional run-length-constrained channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 5128–5140, Nov. 2006.
- [7] J.-Y. Chen, Y.-J. Chen, W.-G. Hu, and S.-S. Lin, "Spatial chaos of Wang tiles with two symbols," *J. Math. Phys.*, vol. 57, no. 2, Feb. 2016, Art. no. 022704.
- [8] S. Congero and K. Zeger, "Hexagonal run-length zero capacity region—Part II: Automated proofs," *IEEE Trans. Inf. Theory*, vol. 68, no. 1, pp. 153–177, Jan. 2021.
- [9] B. Durand, G. Gamard, and A. Grandjean, "Aperiodic tilings and entropy," in *Proc. Int. Conf. Develop. Lang. Theory.* Cham, Switzerland: Springer, 2014, pp. 166–177.
- [10] O. Elishco, T. Meyerovitch, and M. Schwartz, "Semiconstrained systems," *IEEE Trans. Inf. Theory*, vol. 62, no. 4, pp. 1688–1702, Apr. 2016.
- [11] O. Elishco, T. Meyerovitch, and M. Schwartz, "On encoding semiconstrained systems," *IEEE Trans. Inf. Theory*, vol. 64, no. 4, pp. 2474–2484, Apr. 2018.
- [12] O. Elishco, T. Meyerovitch, and M. Schwartz, "On independence and capacity of multidimensional semiconstrained systems," *IEEE Trans. Inf. Theory*, vol. 64, no. 10, pp. 6461–6483, Oct. 2018.
- [13] S. Forchhammer and T. V. Laursen, "A model for the two-dimensional no isolated bits constraint," in *Proc. IEEE Int. Symp. Inf. Theory*, Seattle, WA, USA, Jul. 2006, pp. 1189–1193.
- [14] K. S. Immink, Codes for Mass Data Storage Systems, 2nd ed. Eindhoven, The Netherlands: Shannon Foundation, 2004.
- [15] G. S. Joyce, "On the hard-hexagon model and the theory of modular functions," *Phil. Trans. Roy. Soc. London A*, vol. 325, pp. 643–702, Aug. 1988.
- [16] G. S. Joyce, "Exact results for the activity and isothermal compressibility of the hard-hexagon model," *J. Phys. A, Math. Gen.*, vol. 21, no. 20, pp. L983–L988, Oct. 1988.

- [17] P. W. Kasteleyn, "The statistics of dimers on a lattice. I. The number of dimer arrangements on a quadratic lattice," *Physica*, vol. 27, pp. 1209–1225, Dec. 1961.
- [18] A. Kato and K. Zeger, "On the capacity of two-dimensional runlength constrained channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1527–1540, Jul. 1999.
- [19] Zs. Kukorelly and K. Zeger, "The capacity of some hexagonal (d,k)constraints," in *Proc. IEEE Int. Symp. Inf. Theory*, Washington, DC, USA, Jun. 2001, p. 64.
- [20] Zs. Kukorelly and K. Zeger, "Automated theorem proving for hexagonal run length constrained capacity computation," in *Proc. IEEE Int. Symp. Inf. Theory*, Seattle, WA, USA, Jul. 2006, pp. 1199–1203.
- [21] E. Louidor and B. H. Marcus, "Improved lower bounds on capacities of symmetric 2D constraints using Rayleigh quotients," *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1624–1639, Apr. 2010.
- [22] B. Marcus, "Capacity of higher-dimensional constrained systems," in *Coding Theory and Applications*, vol. 3, R. Pinto, P. R. Malonek, and P. Vettori, Eds. Cham, Switzerland: Springer, 2015, pp. 3–21.
- [23] B. Marcus and R. Pavlov, "Approximating entropy for a class of Z<sup>2</sup> Markov random fields and pressure for a class of functions on Z<sup>2</sup> shifts of finite type," *Ergodic Theory Dyn. Syst.*, vol. 33, no. 1, pp. 186–220, 2013.
- [24] B. Marcus and R. Pavlov, "Computing bounds for entropy of stationary Z<sup>d</sup> Markov random fields," *SIAM J. Discrete Math.*, vol. 27, no. 3, pp. 1544–1558, Jan. 2013.
- [25] B. D. Metcalf and C. P. Yang, "Degeneracy of antiferromagnetic ising lattices at critical magnetic field and zero temperature," *Phys. Rev. B, Condens. Matter*, vol. 18, no. 5, pp. 2304–2307, Sep. 1978.
- [26] Zs. Nagy and K. Zeger, "Capacity bounds for the three-dimensional (0,1) run length limited channel," *IEEE Trans. Inf. Theory*, vol. 46, no. 3, pp. 1030–1033, May 2000.
- [27] L. Onsager, "Crystal statistics. I. A two-dimensional model with an order-disorder transition," *Phys. Rev.*, vol. 65, nos. 3–4, pp. 117–149, Feb. 1944.
- [28] R. Pavlov, "Approximating the hard square entropy constant with probabilistic methods," Ann. Probab., vol. 40, no. 6, pp. 2362–2399, Nov. 2012.
- [29] M. Schwartz and J. Bruck, "On the capacity of the precision-resolution system," *IEEE Trans. Inf. Theory*, vol. 56, no. 3, pp. 1028–1037, Mar. 2010.
- [30] M. Schwartz and A. Vardy, "New bounds on the capacity of multidimensional run-length constraints," *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4373–4382, Jul. 2011.
- [31] A. Sharov and R. M. Roth, "Two-dimensional constrained coding based on tiling," *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1800–1807, Apr. 2010.
- [32] H. Wang, "Proving theorems by pattern recognition—II," Bell Syst. Tech. J., vol. 40, no. 1, pp. 1–41, Jan. 1961.
- [33] G. H. Wannier, "Antiferromagnetism. The triangular ising net," *Phys. Rev.*, vol. 79, no. 2, pp. 357–364, Jul. 1950.
- [34] W. Weeks and R. E. Blahut, "The capacity and coding gain of certain checkerboard codes," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 1193–1203, May 1998.

**Spencer Congero** (Student Member, IEEE) was born in Hartford, CT, USA, in 1994. He received the bachelor's degree in electrical engineering from the University of Southern California in 2016. He is currently pursuing the Ph.D. degree with the University of California, San Diego.

Kenneth Zeger was born in Boston, in 1963. He received the S.B. and S.M. degrees in electrical engineering and computer science from the Massachusetts Institute of Technology in 1984, and the M.A. degree in mathematics and the Ph.D. degree in electrical engineering from the University of California, Santa Barbara, in 1989 and 1990, respectively. He was an Assistant Professor of electrical engineering with the University of Hawai'i from 1990 to 1992. He was with the Department of Electrical and Computer Engineering and the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, as an Assistant Professor, from 1992 to 1995, and an Associate Professor from 1995 to 1996. He was with the Department of Electrical and Computer Engineering, University of California at San Diego, as an Associate Professor, from 1996 to 1998, where he has been a Professor since 1998. He served as a member of the Board of Governors of the IEEE Information Theory Society for the periods 1998-2000, 2005-2007, and 2008-2010. He received NSF Presidential Young Investigator Award in 1991. He served as an Associate Editor At-Large for the IEEE TRANSACTIONS ON INFORMATION THEORY for the period 1995-1998.