On the Complexity of the Conditional Independence Implication Problem with Bounded Cardinalities

Michał Makowski
Faculty of Mathematics, Informatics and Mechanics, University of Warsaw, Poland

Abstract

We show that the conditional independence (CI) implication problem with bounded cardinalities, which asks whether a given CI implication holds for all discrete random variables with given cardinalities, is co-NEXPTIME-hard. The problem remains co-NEXPTIME-hard if all variables are binary. The reduction goes from a variant of the tiling problem and is based on a prior construction used by Cheuk Ting Li to show the undecidability of a related problem where the cardinality of some variables remains unbounded. The CI implication problem with bounded cardinalities is known to be in EXPSPACE, as its negation can be stated as an existential first-order logic formula over the reals of size exponential with regard to the size of the input.

2012 ACM Subject Classification Mathematics of computing → Information theory; Theory of computation → Problems, reductions and completeness

Keywords and phrases Conditional independence implication, exponential time, tiling problem

Digital Object Identifier 10.4230/LIPIcs.MFCS.2024.73

Related Version *Full Version*: `https://arxiv.org/abs/2408.02550` [10]

Acknowledgements The author would like to thank Damian Niwiński for his guidance and the anonymous reviewers for their valuable comments.

1 Introduction

The implication problem for conditional independence statements is one of the major decision problems arising in multivariate statistical modeling and other applications [2]. The problem asks whether a list of conditional independence statements implies another such statement (see the exact formulation below). The problem is a special case of the conditional entropic inequality problem, as the statement X *and* Y *are independent, given* Z (sometimes denoted $X \perp Y \mid Z$) is equivalent to the equation $I(X;Y|Z) = 0$ in information theory. Here the random variables in consideration are of the form $(X_{i_1}, \ldots, X_{i_\ell})$, abbreviated by X_Z with $Z = \{i_1, \ldots, i_\ell\}$, selected from a fixed n-tuple of variables $X_1, \ldots, X_n$ considered with a joint distribution. As with the general problem, this can be considered over continuous, infinite discrete or finite discrete random variables. Furthermore, the CI implication problem can be refined by imposing certain requirements on the sets A, B, C in $X_A \perp X_B \mid X_C$, e.g., they must be pairwise disjoint for *disjoint CI*, and for *saturated CI* they must additionally satisfy $A \cup B \cup C = \{1, \ldots, n\}$. We will focus on discrete random variables with a finite domain and without constraints on the sets, addressing disjoint CI in the full version.

If the domain size is bounded, this problem is decidable since the conditional independence can be expressed as an arithmetic formula in terms of elementary events' probabilities. Considering all possible domain sizes yields a semi-algorithm for finding a counter-example to the implication, showing that the unbounded problem is co-recursively enumerable (as noted by Khamis *et al* [5]). The decidability of the general CI implication problem was unknown for a long time, with only special cases resolved. Finally, Cheuk Ting Li published two papers [8, 9] proving the problem to be undecidable. This was also shown independently

by Kühne and Yashfe [6]. Still unknown is the complexity of the bounded problem – the method of constructing an existential formula of Tarski's arithmetic yields an upper bound of EXPSPACE (cf. [1]), while the other published algorithms appear to be mostly heuristics. Hannula *et al* [4] conjectured that the problem can be actually easier, especially in the case where all variables are binary, as the arithmetic formula in question is of a very special form.

In this paper we show that the problem is in general co-NEXPTIME-hard, and the hardness result continues to hold if all variables are binary. Our reduction is an adaptation of the construction presented by Li [8] to show that the problem is undecidable if the cardinalities of some variables are bounded. While there is still a gap between the lower and upper bound, this shows that the complexity of the CI implication problem is harder than it might have been expected.

2 Problem statement

Denote by $\mathrm{card}(X)$ the cardinality of a random variable X. Formally, the following problem will be considered:

Bounded CI Implication

Input: Integers m, n, given in unary. A list of $m+1$ triples (A_i, B_i, C_i) of subsets of $\{1, \ldots, n\}$. A list of n integers K_j, given in binary.

Question: Determine whether the implication

$$\bigwedge_{i \in \{1,\ldots,m\}} (I(X_{A_i}; X_{B_i} | X_{C_i}) = 0) \Rightarrow I(X_{A_{m+1}}; X_{B_{m+1}} | X_{C_{m+1}}) = 0$$

holds for all jointly distributed random variables $(X_1, \ldots, X_n)$ with $\mathrm{card}(X_j) \leq K_j$ for all $j \in \{1, \ldots, n\}$.

We define Constant-bounded CI Implication as a variant of the above problem in which all K_i are fixed to be equal to 2 rather than given as input. We show the following:

▶ **Theorem 1.** *Bounded CI Implication and Constant-bounded CI Implication are co-NEXPTIME-hard. This also holds in the disjoint CI case, i. e. when for each i the sets A_i, B_i, C_i are pairwise disjoint.*

We focus on the first part of the theorem – the proof of the second part differs little from that given by Li [8] and is given in the full version.

In order to state the tiling-based problems utilized in the reduction, we introduce some definitions based on those in [7]. We define a *tiling system* as a triple $\mathcal{D} = (D, H, V)$, where D is a finite set of tiles and $H, V \subseteq D^2$ are the horizontal and vertical constraints, accordingly, which give the pairs of tiles that may be neighbors. This is a generalization of Wang tiles, where a set of colors C is given and tiles (formally quadruples from the set C^4) are represented by squares with colored edges with the requirement that only edges of the same color may touch. As stated in [12], Wang tiles correspond exactly to those tiling systems for which the implication $(a\,R\,b \wedge a\,R\,c \wedge d\,R\,b) \Rightarrow d\,R\,c$ holds for all $a, b, c, d \in D$ and both $R \in \{H, V\}$.

We define a $k \times l$ *tiling* by $\mathcal{D}$ as a function $f : \{0, \ldots, k-1\} \times \{0, \ldots, l-1\} \to D$ such that:

- $(f(m, n), f(m+1, n)) \in H$ for all $m < k-1, n < l$,
- $(f(m, n), f(m, n+1)) \in V$ for all $m < k, n < l-1$.

A *periodic tiling* is one that also has $(f(k-1, n), f(0, n)) \in H$ and $(f(m, l-1), f(m, 0)) \in V$ for all $m < k, n < l$. For a (non-periodic) $k \times l$ tiling f, the *starting tile* and *final tile* are the values of $f(0, 0)$ and $f(k-1, l-1)$, respectively.

We will show a polynomial-time many-one reduction from the following problem, which is known to be NEXPTIME-complete [7, exercise 7.2.2.]:

Binary Bounded Tiling
Input: A tiling system $\mathcal{D}$, a starting tile $d_0 \in D$, an integer k given in binary.
Question: Determine whether there exists a $k \times k$ tiling f by $\mathcal{D}$ such that $f(0,0) = d_0$.

The reduction consists of two parts, the first being a purely tiling-based reduction from the above to the following intermediate problem:

Periodic Bounded Tiling
Input: A tiling system $\mathcal{D}$, a designated tile t, integers m, n given in binary.
Question: Determine whether there exists a periodic tiling by $\mathcal{D}$ of size at most $m \times n$ which uses tile t.

The second part is a reduction from Periodic Bounded Tiling to the complement of Bounded CI Implication, based on a construction by Li [8].

3 First part of the reduction

We first show a polynomial-time many-one reduction from Binary Bounded Tiling to Periodic Bounded Tiling. This means that given a tiling system $\mathcal{D}$, starting tile d_0 and integer k, we will construct in polynomial time a tiling system $\mathcal{D}''$, designated tile t and integers m, n such that Binary Bounded Tiling gives a positive answer for input $(\mathcal{D}, d_0, k)$ iff Periodic Bounded Tiling gives a positive answer for input $(\mathcal{D}'', t, m, n)$. This consists of two steps:

1. Modify $\mathcal{D}$ into system $\mathcal{D}'$ such that valid tilings by $\mathcal{D}$ of size $k \times k$ correspond to valid tilings by $\mathcal{D}'$ with certain corner constraints.
2. Modify $\mathcal{D}'$ into system $\mathcal{D}''$ and tile t such that valid tilings by $\mathcal{D}'$ with the above corner constraints correspond to periodic tilings by $\mathcal{D}''$ of size $(k+1) \times (k+1)$ that use tile t.

This is a fairly typical reduction between tilings, similar to problems considered for instance in [3].

Limiting tiling size

For the first step, we create a tiling system $\mathcal{C}$ (of polynomial size with regard to the length of k), along with starting tile c_0 and final tile c_1, implementing a binary counter that counts down from an appropriately chosen $k' \leq k$ (close to k) and whose position shifts by 1 with each decrement. The tile c_1 occurs when the counter reaches 0. Similar constructions have been shown [11], our example is given in Figure 1. Thus, any tiling by $\mathcal{C}$ with c_0 in the top-right corner and c_1 in the bottom-left must be of size exactly $k \times k$.

Consider a "layering" of $\mathcal{D}$ and $\mathcal{C}$ into system $\mathcal{D} \times \mathcal{C} = (D \times C, H_{D\times C}, V_{D\times C})$, where the relations are defined as $R_{D\times C} = \{((d,c),(d',c')) : (d,d') \in R_D \wedge (c,c') \in R_C\}$ for both $R \in \{H, V\}$. This way, any tiling by $\mathcal{D} \times \mathcal{C}$ corresponds to a pair of tilings by $\mathcal{D}$ and $\mathcal{C}$. We let $\mathcal{D}' = \mathcal{D} \times \mathcal{C}$ and define the corner constraints mentioned in point 2 by restricting the possible starting and final tiles to $S = \{(d_0, c_0)\}, F = \{(d, c_1) : d \in D\}$ respectively. This yields tilings by $\mathcal{D}'$ which consist of a tiling by $\mathcal{D}$ with starting tile d_0 and a tiling by $\mathcal{C}$ with starting tile c_0 and final tile c_1.

Periodic tilings

The second step realizes the above constraint while also converting between periodic and non-periodic tilings. It consists of adding 5 new tiles to $\mathcal{D}'$ – ,  ,  ,  ,  – yielding $\mathcal{D}''$. The added constraints are shown in Figure 2 and an example tiling in Figure 3. The distinguished tile t is  – the idea is that its usage forces a "border" of  and  tiles. Within each of these borders (there can be multiple if  is used more than once) there is a valid tiling by the system $\mathcal{D}'$ additionally satisfying the starting and final constraints S, F.

Final conversion

Combining the above steps, a tiling by $\mathcal{D}$ of size $k \times k$ is represented by a periodic $(k+1)\times(k+1)$ tiling by $\mathcal{D}''$ and thus we let $m = n = k + 1$. The designated tile is set to  , completing the reduction from Binary Bounded Tiling to Periodic Bounded Tiling.

Variant with only powers of two

Consider the following variant of the tiling problem:

Power-of-two Periodic Bounded Tiling
Input: Integers m, n given in unary, a tiling system $\mathcal{D}$, a designated tile t.
Question: Determine whether there exists a periodic tiling by $\mathcal{D}$ of size at most $2^m \times 2^n$ which uses tile t.

Note that this is the same as taking the input in binary while restricting it only to powers of two. This variant is also NEXPTIME-hard because in the above reduction, the only possible sizes of tiling by the constructed tiling system are multiples of $k + 1$, both in width and height. Letting $m = n = \lceil \log_2(k + 1) \rceil$, we have $k + 1 \leq 2^m < 2(k + 1)$ and the same for 2^n. Therefore, the possible tilings are the same for size bound $(k + 1) \times (k + 1)$ and $2^{\lceil \log_2(k+1) \rceil} \times 2^{\lceil \log_2(k+1) \rceil}$.

4 Second part of the reduction

Given a tiling system $\mathcal{D}$, a designated tile t and integers m, n given in binary, we will construct (in polynomial time) a CI implication with bounded cardinalities which does not hold iff Periodic Bounded Tiling has a solution for input $(\mathcal{D}, t, m, n)$. This is based on the construction of Li [8]. Proofs which differ from the original only by specifying cardinality bounds are deferred to the full version. The main changes to Li's construction are as follows:

- provide bounds for the random variables used in the construction – this is done as the implication is being constructed;
- reduce the size of the implication from exponential to polynomial with regard to the input – only one part (the predicate COL) needs to be replaced by a polynomial-size equivalent;
- modify the representation of tiles to better suit bounded size and non-Wang tiles;
- add the requirement of the usage of a given tile – this is done by modifying the consequent of the implication.

4.1 Preliminaries

We denote by $\mathrm{Unif}(S)$ a uniform distribution over set S and by $\mathrm{Bern}(p)$ a Bernoulli distribution with parameter p. We use the shorthand X^k to represent a tuple of random variables $(X_1, \ldots, X_k)$, which is in itself also a random variable.

						$_1$	1_0	0_1	1
					$_1$	1_0	0_0	0	
				$_0$	0_1	1_1^*	1^*		
			$_0$	0_1	1_0	0			
		$_0$	0_0	0_1	1^*				
	$_0$	0_0	0_0	0					
$\star$	1_1^*	1_1^*	1^*						

Figure 1 A tiling implementation of a binary counter which shifts its position on every decrement, with the starting and final tile in the top-right and bottom-left corners respectively. This example counts down from 5 (101 in binary). Every tile except for the final $\star$ is of the form a_b^c, where a is the bit value (possibly blank), b is the value of the bit directly to the right (or blank if there is none), and c is optionally $*$ if a borrow operation is required. The shaded tiles of the top row function in the same manner, but they are "memorized" within the tiling system such that the placement of the top-right tile forces the top row to write out the binary initial value. The tiles in the lower rows are chosen deterministically based on their right and top neighbor. Finally, the $\star$ tile only occurs when the tile above is blank and the one to the right requires a borrow, which indicates that the counter has just gone below zero. In order to be unable to further count down, we disallow any tiles being below or to the left of tile $\star$. The size of the tiling is $(k'+b+2)\times(k'+2)$, where b is the number of bits and k' is the initial value; however, this could be modified such that the final tiling has size $(k'+b+2)\times(k'+b+2)$ by padding with b dummy rows at the top. For sufficiently large k, we can always efficiently find b, k' such that $k'+b+2=k$.

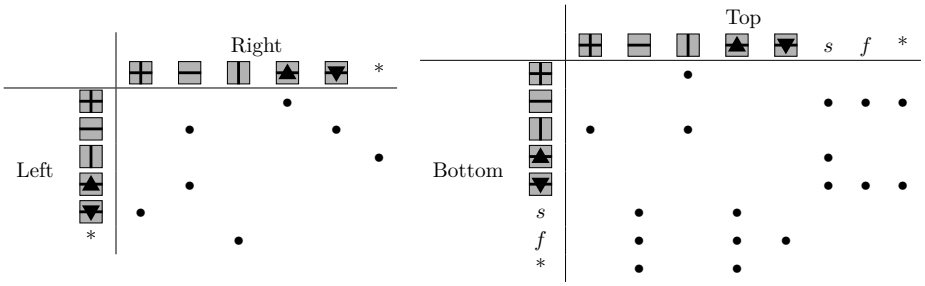

Figure 2 Modified adjacency relation for the system $\mathcal{D}''$ – only adjacencies marked by $\bullet$ are permitted, as well as all adjacencies from the original tiling system. The asterisk denotes any tiles from the system $\mathcal{D}'$, while s, f represent any tile from the initial and final subset of tiles, respectively (S and F defined above).

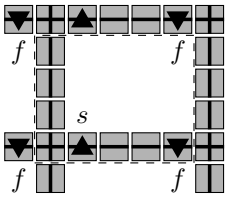

Figure 3 An example "border" created by the above tiling, with tiles from the original set not shown and s, f representing tiles from S, F respectively. The dashed rectangle represents the actual rectangle being tiled, while the tiles outside are periodic copies added to better illustrate the construction.

The *entropy* of a finite discrete random variable X with domain $\mathcal{X}$ is defined as

$$H(X) = -\sum_{x\in\mathcal{X}} \mathbf{P}(X = x) \log \mathbf{P}(X = x).$$

The *entropy of X conditioned on Y*, with X, Y finite discrete random variables with domains $\mathcal{X}, \mathcal{Y}$ respectively, is defined as

$$H(X|Y) = -\sum_{x\in\mathcal{X}, y\in\mathcal{Y}} \mathbf{P}(X = x \wedge Y = y) \log \frac{\mathbf{P}(X = x \wedge Y = y)}{\mathbf{P}(X = x)}.$$

Finally, the *conditional mutual information of X and Y given Z* can be defined as

$$I(X;Y|Z) = H(X|Z) - H(X|Y,Z).$$

Similarly, the *mutual information* of X and Y is defined as

$$I(X;Y) = H(X) - H(X|Y).$$

Note that the conditional independence (CI) statement $I(X;Y|Z) = 0$ is equivalent to the fact that X and Y are independent given Z, and thus can be expressed without the usage of logarithms as

$$\mathbf{P}(X = x \wedge Y = y \wedge Z = z)\mathbf{P}(Z = z) = \mathbf{P}(X = x \wedge Z = z)\mathbf{P}(Y = y \wedge Z = z)$$

for all $x \in \mathcal{X}, y \in \mathcal{Y}, z \in \mathcal{Z}$. Further, the functional dependence statement $H(X|Y) = 0$, which states that for any $y \in \mathcal{Y}$, there exists exactly one $x \in \mathcal{X}$ such that $\mathbf{P}(X = x \wedge Y = y) \neq 0$ can be expressed equivalently as a (non-disjoint) CI statement $I(X;X|Y) = 0$.

We will follow Li's construction [8], providing cardinality bounds and modifications where necessary. While the final goal is a CI implication, we will mostly construct *affine existential information predicates* (AEIP) [8], converting to a CI implication at the end. We will only consider a special form of AEIP which consists of an existentially quantified conjunction of CI statements. This family of predicates is closed under conjunction, in particular we can use a predicate within the definition of another predicate (implicitly renaming variables in the case of a naming conflict).

Since our goal is to construct a bounded CI implication, every quantified variable will be given a *cardinality bound* – the maximum allowed size of its domain. We denote the existential quantification of a variable X with $\mathrm{card}(X) \leq k$ by the shorter notation $\exists X \leq k$, similarly for tuple variables $\exists X^2 \leq k$ represents the existence of variables X_1, X_2 with $\mathrm{card}(X_1), \mathrm{card}(X_2) \leq k$. Whenever a predicate takes arguments, their cardinalities are already bounded since they have been quantified. We may need to refer to these bounds when quantifying new variables, denoting by K_X the bound already given to variable X. Finally, whenever $\leq$ is replaced by $\leq_i$, this indicates an "implicit" bound, that is one which does not change the meaning of the predicate because it is already satisfied by any such quantified variable even without the explicit bound. An example of this is that whenever $H(X|Y) = 0$, X is functionally dependent on Y and so $X \leq_i K_Y$.

The first defined predicate is TRIPLE:

$$\begin{aligned}\mathrm{TRIPLE}(Y_1, Y_2, Y_3) : {} & H(Y_1|Y_2, Y_3) = H(Y_2|Y_1, Y_3) = H(Y_3|Y_1, Y_2) = 0 \\ & \wedge I(Y_1; Y_2) = I(Y_1; Y_3) = I(Y_2; Y_3) = 0.\end{aligned}$$

By definition, predicate TRIPLE of three variables Y_1, Y_2, Y_3 is satisfied iff Y_1, Y_2, Y_3 are pairwise independent and each variable is functionally dependent on the other two. Functional dependency is a special case of conditional independence, since $H(X|Y) = I(X;X|Y)$. The following is shown in [13]:

▶ **Lemma 1.** *If* $\mathrm{TRIPLE}(X, Y, Z)$ *is satisfied, then* X, Y, Z *are all uniformly distributed and have the same cardinality.*

This is used in the next predicate, UNIF:

$$\mathrm{UNIF}(X) : \exists U_1 \leq_i K_X, U_2 \leq_i K_X : \mathrm{TRIPLE}(X, U_1, U_2).$$

By definition, predicate UNIF of one variable X is satisfied iff there exist discrete random variables U_1, U_2 jointly distributed with X such that $\mathrm{TRIPLE}(X, U_1, U_2)$ holds, which in turn is equivalent to X being uniformly distributed over its support. Lemma 1 immediately shows that the implicit cardinality bounds for U_1, U_2 are correct.

For any constant k, predicate $\mathrm{UNIF}_k(X)$ is defined to imply X being uniformly distributed over a domain of size k:

$$\mathrm{UNIF}_k(X) : \mathrm{UNIF}(X) \wedge \alpha_k \leq H(X) \leq \alpha_{k+1},$$

where $\alpha_k \in \mathbb{Q}$ is some rational with $\log(k-1) < \alpha_k < \log k$ (because the entropy of a uniform variable is the logarithm of the cardinality of its support). While this is still a valid AEIP, it is not a conjunction of CI statements. However, in the bounded cardinality setting UNIF_k can be restated in this form [8, 10]. Note that this predicate imposes an exact domain size constraint, while the cardinality bounds given as input provide only an upper bound. Finally, note that the predicates UNIF, UNIF_k are satisfiable – for any $k > 0$, there exists a variable X which satisfies $\mathrm{UNIF}_k(X)$ and so $\mathrm{UNIF}(X)$.

Define the *characteristic bipartite graph* of random variables X_1, X_2 with (disjoint) supports $\mathcal{X}_1, \mathcal{X}_2$ as the undirected graph with set of vertices $V = \mathcal{X}_1 \cup \mathcal{X}_2$ and set of edges $E = \{(x_1, x_2) : x_1 \in \mathcal{X}_1, x_2 \in \mathcal{X}_2, \mathbf{P}(X_1 = x_1 \wedge X_2 = x_2) > 0\}$.

Li constructs the predicate

$$\begin{aligned}\mathrm{CYCS}(X_1, X_2) : \exists U \leq_i 2 : \mathrm{UNIF}(X_1) \wedge \mathrm{UNIF}(X_2) \wedge \mathrm{UNIF}_2(U)\\ \wedge I(X_1; U) = I(X_2; U) = 0\\ \wedge H(X_1|X_2, U) = H(X_2|X_1, U) = 0\\ \wedge H(U|X_1, X_2) = 0\end{aligned}$$

and shows the following (without cardinality bounds):

▶ **Lemma 2.** $\mathrm{CYCS}(X_1, X_2)$ *is satisfied iff* X_1, X_2 *are uniform and the characteristic bipartite graph of* X_1, X_2 *consists only of vertex-disjoint simple cycles.*

Furthermore, this predicate is satisfiable – for any finite collection of even-length cycles, we can clearly find X_1, X_2 such that their characteristic bipartite graph consists exactly of this collection of cycles.

4.2 Overview of the construction

We now give an overview of the following steps of the construction. Section 4.3 defines the predicate $\mathrm{TORI}'(X^2, Y^2, Z)$, which enforces that the characteristic bipartite graph of X_1 and X_2 is a collection of cycles, similarly for Y_1 and Y_2. Finally, we require that Z be distributed uniformly over two values and that the three variables X^2, Y^2, Z be independent. The distribution of (X^2, Y^2) then represents a collection of tori, with each quadruple of values of (X_1, X_2, Y_1, Y_2) representing a vertex in some torus. The addition of variable Z effectively creates a corresponding copy of this collection.

With the goal of creating meaningful labels to be applied to the vertices of the aforementioned graph, which are represented by k-tuple of binary variables W^k, Section 4.4 defines the predicates $\mathrm{SW}'(W^k, V^k, \bar{V}^k, F)$ and $\mathrm{COL}'(W^k, V^k, \bar{V}^k, F)$. The former ensures that $V_i = (1 - W_i)F, \bar{V}_i = W_i F$ (up to relabeling) with the side-effect of requiring each $W_i \sim \mathrm{Bern}(\frac{1}{2})$. The latter predicate restricts W^k such that the only values that are possible have either $W_k = 1$ and exactly one $W_i = 0$, or $W_k = 0$ and exactly one $W_i = 1$, for some $i \in \{1, \ldots, k-1\}$.

Section 4.5 combines the prior predicates in predicate $\mathrm{CTORI}'(X^2, Y^2, Z, W^k, V^k, \bar{V}^k, F)$ in such a way that each vertex is assigned exactly one label, i. e. W^k depends functionally on (X^2, Y^2, Z). This is extended in predicate $\mathrm{OTORI}'(X^2, Y^2, Z, W^k, V^k, \bar{V}^k, F)$, which assigns four possible *groups* to vertex labels and enforces a structure as shown in Figure 4.

Finally, the predicate $\mathrm{TTORI}'(X^2, Y^2, Z, W^k, V^k, \bar{V}^k, F)$ defined in Section 4.7 restricts the possible labels of vertices connected by edges so as to enforce the given vertical and horizontal constraints of the given tile system.

4.3 Grid

A collection of tori is constructed by Li using the following predicate (where X^2 represents a pair of variables (X_1, X_2), similarly for Y^2):

$$\mathrm{TORI}(X^2, Y^2) : \mathrm{CYCS}(X^2) \wedge \mathrm{CYCS}(Y^2) \wedge I(X^2; Y^2) = 0,$$

When the above holds, fixing any three of the variables X_1, X_2, Y_1, Y_2 leaves two possible values of the fourth. Similarly, fixing one variable from X_1, X_2 and one from Y_1, Y_2 gives the remaining two variables a distribution over four values. This is visualized by a graph similar in idea to the bipartite characteristic graph: its vertices are quadruples of values (x_1, x_2, y_1, y_2) which satisfy $\mathbf{P}(X_1 = x_1 \wedge X_2 = x_2 \wedge Y_1 = y_1 \wedge Y_2 = y_2) > 0$, with edges connecting any two quadruples which differ in exactly one out of these four values. When arranged in a grid with possible pairs (X_1, X_2) on one axis and (Y_1, Y_2) on the other, as in Figure 4, the torus structure becomes apparent. Again, this predicate is satisfiable in the sense that any collection of tori which is a product of two collections of even-length cycles has a representation by random variables X^2, Y^2.

Our construction departs slightly from the construction of Li, adding another coordinate Z, corresponding to taking two copies of the collection of tori, with the edges and faces described above preserved when Z is fixed. Additionally, fixing X^2 and Y^2 but not Z "connects" two corresponding vertices in the two copies. The predicate for this is as follows:

$$\begin{aligned}\mathrm{TORI}'(X^2, Y^2, Z) : \mathrm{CYCS}(X^2) \wedge \mathrm{CYCS}(Y^2) \wedge \mathrm{UNIF}_2(Z) \\ \wedge I(X^2, Y^2; Z) = 0 \wedge I(X^2, Z; Y^2) = 0 \wedge I(Y^2, Z; X^2) = 0.\end{aligned}$$

4.4 Vertex labels

The basis for constructing labels which will later on be assigned to vertices is the following predicate defined by Li:

$$\begin{aligned}\mathrm{FLIP}(F, G_1, G_2) : \exists U \leq_i 4, Z^2 \leq_i 3 : \mathrm{UNIF}_4(U) \wedge \mathrm{UNIF}_2(F) \\ \wedge H(F, G_1, G_2 | U) = I(G_1; G_2 | F) = 0 \\ \wedge \mathrm{UNIF}_3(Z_1) \wedge I(Z_1; G_1) = H(U | G_1, Z_1) = 0 \\ \wedge \mathrm{UNIF}_3(Z_2) \wedge I(Z_2; G_2) = H(U | G_2, Z_2) = 0.\end{aligned}$$

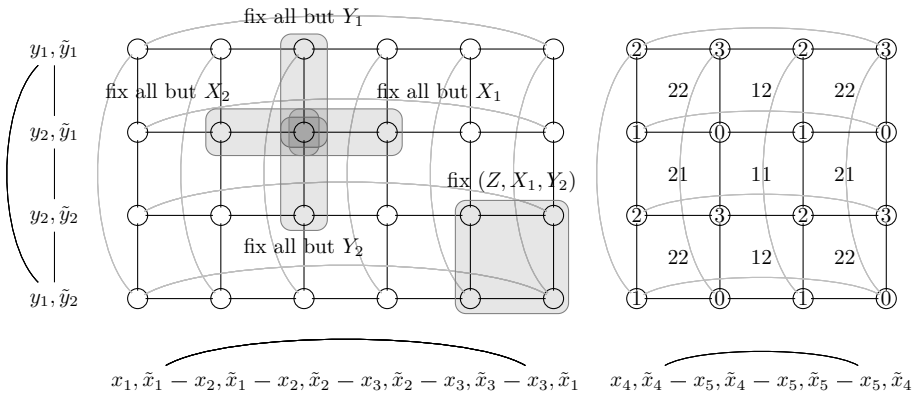

Figure 4 Visualization of the tori which are a product of the cycles created by (X_1, X_2) and (Y_1, Y_2), with each vertex corresponding to a quadruple of the values of (X_1, X_2, Y_1, Y_2). The axes show these cycles – a quadruple's (X_1, X_2) (resp. (Y_1, Y_2)) values are determined by projecting onto the horizontal (resp. vertical) axis. Additionally, the left torus shows highlighted edges which arise when all but one variable (of X_1, X_2, Y_1, Y_2, Z) are fixed as well as an example face which arises when Z and two other variables are fixed. The right torus has each face labeled with its type and each vertex labeled with its group – these are used in Section 4.6 in order to restrict allowed labelings of the vertices. The second "corresponding" torus and the Z axis are omitted for clarity.

Recalling that $\mathrm{Unif}(S)$ denotes a uniform distribution over set S, the following is shown:

▶ **Lemma 3.** $\mathrm{FLIP}(F, G_1, G_2)$ *is satisfied iff, up to relabeling,* (F, G_1, G_2) *has the distribution* $\mathit{Unif}(\{(0,0,0),(0,1,0),(1,0,0),(1,0,1)\})$.

For any $k \geq 4$, Li defines the predicate SW which allows us to represent strings of k bits. Here, W^k represents a tuple of variables $(W_1, \ldots, W_k)$, same for $V^k, \bar{V}^k$: Intuitively, the value of W_i determines whether F is copied into V_i or $\bar{V}_i$, hence the name SW for *switch*.

$$\begin{aligned}\mathrm{SW}(W^k, V^k, \bar{V}^k, F) : \exists G \leq_i 2 : I(W^k; F, G) = 0 \\ \wedge \bigwedge_{i \in \{1,\ldots,k\}} \big(\mathrm{UNIF}_2(W_i) \wedge H(V_i, \bar{V}_i | W_i, F) = I(V_i; \bar{V}_i | W_i) = 0 \\ \wedge \mathrm{FLIP}(F, G, V_i) \wedge \mathrm{FLIP}(F, G, \bar{V}_i)\big)\end{aligned}$$

▶ **Lemma 4.** *If* $\mathrm{SW}(W^k, V^k, \bar{V}^k, F)$ *is satisfied, then we have (without loss of generality)* $V_i = (1 - W_i)F, \bar{V}_i = W_i F$ *for all* i.

The predicate SW is satisfiable: we let (F, G) take each of the values $(0,0)$, $(0,1)$ with probability $\frac{1}{4}$, in which case we let $V_i = \bar{V}_i = 0$, and the value $(1,0)$ with probability $\frac{1}{2}$, in which case $V_i = 1 - W_i, \bar{V}_i = W_i$. As long as each $W_i \sim \mathrm{Bern}(\frac{1}{2})$ (recall that $\mathrm{Bern}(p)$ denotes a Bernoulli distribution with parameter p), we can satisfy the predicate for any distribution of W^k. This predicate additionally has the following property:

▶ **Lemma 5.** *If* $\mathrm{SW}(W^k, V^k, \bar{V}^k, F)$ *is satisfied, then for any* $S, \bar{S} \subseteq \{1, \ldots, k\}$ *we have*

$$H(F | V_S, \bar{V}_{\bar{S}}, W^k) = \mathbf{P}(\mathit{sat}(W^k, S, \bar{S}) = 1),$$

where

$$\mathit{sat}(w^k, S, \bar{S}) = \Big(\prod_{i \in S} w_i\Big)\Big(\prod_{i \in \bar{S}} (1 - w_i)\Big),$$

i.e. if $w_i = 1$ *for all* $i \in S$ *and* $w_i = 0$ *for* $i \in \bar{S}$ *then* $\mathit{sat}(w^k, S, \bar{S})$ *equals 1 and otherwise it equals 0.*

This immediately yields the following:

▶ **Lemma 6.** *The equality* $H(F|V_{\{i\in\{1,\dots,k\}\,:\,w_i=1\}}, \bar{V}_{\{i\in\{1,\dots,k\}\,:\,w_i=0\}}, W^k) = \mathbf{P}(W^k = w^k)$ *holds for any* $w^k \in \{0,1\}^k$. ⌟

Using these properties, we can disallow certain values of W^k from occurring using a CI statement. This is used by Li to limit the possible values of W^k to the set $T_k \subseteq \{0,1\}^k$, which consists of $2(k-1)$ *labels* (referred to as *colors* by Li), each one with a *value* and *sign*. The set consists of strings in which exactly one bit differs from the last bit, that is for any $w^k \in T_k$ we have $w_j \neq w_k$ for exactly one $j \neq k$. The sign of the label is determined by w_k – negative when $w_k = 1$, positive when $w_k = 0$ – and j is the value of the label. For example, the elements of $T_4 = \{0111, 1011, 1101, 1000, 0100, 0010\}$ correspond in order to labels $\{-1, -2, -3, +1, +2, +3\}$.

Li's predicate for enforcing this is simple:

$$\begin{aligned}\mathrm{COL}(W^k, V^k, \bar{V}^k, F): \ & \mathrm{SW}(W^k, V^k, \bar{V}^k, F) \\ & \wedge \bigwedge_{w^k \in \{0,1\}^k \setminus T_k} \left(H(F|V_{\{i:w_i=1\}}, \bar{V}_{\{i:w_i=0\}}, W^k) = 0\right),\end{aligned}$$

This predicate simply disallows the occurrence of any $w^k \notin T_k$, hence we have

▶ **Lemma 7.** *If* $\mathrm{COL}(W^k, V^k, \bar{V}^k, F)$ *is satisfied, then* $\mathbf{P}(W^k \notin T_k) = 0$. ⌟

While sufficient for showing undecidability, this predicate cannot be used in a polynomial-time reduction due to its size being exponential with regard to k. However, an equivalent polynomial-size predicate can be easily constructed: with

$$\bigwedge_{\substack{i,j\in\{1,\dots,k-1\},\\ i<j}} (H(F|V_{\{i,j\}}, \bar{V}_{\{k\}}, W^k) = 0),$$

we disallow all w^k such that $w_k = 0$ and $w_i = w_j = 1$ for some $1 \leq i < j < k$. Similarly,

$$\bigwedge_{\substack{i,j\in\{1,\dots,k-1\},\\ i<j}} (H(F|V_{\{k\}}, \bar{V}_{\{i,j\}}, W^k) = 0)$$

disallows all w^k such that $w_k = 1$ and $w_i = w_j = 0$ for some $1 \leq i < j < k$. The only remaining w^k have either $w_k = 0$ and at most one 1 in $\{1, \dots, k-1\}$ or have $w_k = 1$ and at most one 0 in $\{1, \dots, k-1\}$. The strings 0^k and 1^k are disallowed by $H(F|V_{\{1,\dots,k\}}, \bar{V}_{\varnothing}, W^k) = 0$ and $H(F|V_{\varnothing}, \bar{V}_{\{1,\dots,k\}}, W^k) = 0$. Combined, these yield the polynomial-size predicate

$$\begin{aligned}\mathrm{COL}'(W^k, V^k, \bar{V}^k, F): & \mathrm{SW}(W^k, V^k, \bar{V}^k, F) \\ & \wedge \bigwedge_{\substack{i,j\in\{1,\dots,k-1\},\\ i<j}} (H(F|V_{\{i,j\}}, \bar{V}_{\{k\}}, W^k) = 0) \\ & \wedge \bigwedge_{\substack{i,j\in\{1,\dots,k-1\},\\ i<j}} (H(F|V_{\{k\}}, \bar{V}_{\{i,j\}}, W^k) = 0) \\ & \wedge H(F|V_{\{1,\dots,k\}}, \bar{V}_{\varnothing}, W^k) = 0 \\ & \wedge H(F|V_{\varnothing}, \bar{V}_{\{1,\dots,k\}}, W^k) = 0\end{aligned}$$

equivalent to the exponential-size COL.

4.5 Edge constraints

The next predicate is the following, with COL used in place of COL′ in Li's original construction:

$$\begin{aligned}\mathrm{COLD}(X, W^k, V^k, \bar{V}^k, F) : \mathrm{COL}'(W^k, V^k, \bar{V}^k, F)\\ \wedge H(W^k|X) = I(V^k, \bar{V}^k, F; X|W^k) = 0.\end{aligned}$$

This predicate will represent the labeling of vertices, with X representing the coordinate and W^k (which depends functionally on X) representing its label. For any $x \in \mathcal{X}$, denote by $w^k(x)$ the unique value of w^k which satisfies $\mathbf{P}(W^k = w^k|X = x) > 0$.

Suppose that $\mathrm{COLD}(X, W^k, V^k, \bar{V}^k, F)$ is satisfied and let E be any random variable which *splits X into sets of (a constant) size l* – we say this is the case when $H(E|X) = 0$ and $X|E = e$ is uniform over l values for all $e \in \mathcal{E}$. This is not verified by a predicate; rather, we will only choose E which have this property by definition. Fixing subsets of indices $S, \bar{S} \subseteq \{1, \ldots, k\}$, we define for any $e \in \mathcal{E}$ the value a_e:

$$a_e = |\{x : \mathbf{P}(X = x|E = e) > 0 \wedge \mathrm{sat}(w^k(x), S, \bar{S}) = 1\}|$$

In order to impose restrictions on the possible values of a_e, Li defines the following predicates:

$$\begin{aligned}\mathrm{SAT}_{\neq 1/2,S,\bar{S}}(E, W^k, V^k, \bar{V}^k, F) : \exists U \leq_i 2 : \mathrm{UNIF}_2(U) \wedge I(U; E, V_S, \bar{V}_{\bar{S}}) = 0\\ \wedge H(F|V_S, \bar{V}_{\bar{S}}, E, U) = 0,\end{aligned}$$

$$\begin{aligned}\mathrm{SAT}_{\leq 1/2,S,\bar{S}}(E, W^k, V^k, \bar{V}^k, F) : \exists U \leq_i 3 : \mathrm{UNIF}_3(U) \wedge I(U; E, V_S, \bar{V}_{\bar{S}}) = 0\\ \wedge H(F|V_S, \bar{V}_{\bar{S}}, E, U) = 0,\end{aligned}$$

$$\begin{aligned}\mathrm{SAT}_{\leq 3/4,S,\bar{S}}(E, W^k, V^k, \bar{V}^k, F) : \exists U \leq_i 105 : \mathrm{UNIF}_{105}(U) \wedge I(U; E, V_S, \bar{V}_{\bar{S}}) = 0\\ \wedge H(F|V_S, \bar{V}_{\bar{S}}, E, U) = 0.\end{aligned}$$

The predicates satisfy the following properties.

▶ **Lemma 8.** *If* $\mathrm{COLD}(X, W^k, V^k, \bar{V}^k, F)$ *is satisfied and E splits X into sets of size 2, then* $\mathrm{SAT}_{\neq 1/2,S,\bar{S}}(E, W^k, V^k, \bar{V}^k, F)$ *is satisfied iff $a_e \neq 1$ for all $e \in \mathcal{E}$.*

▶ **Lemma 9.** *If* $\mathrm{COLD}(X, W^k, V^k, \bar{V}^k, F)$ *is satisfied and E splits X into sets of size 2, then* $\mathrm{SAT}_{\leq 1/2,S,\bar{S}}(E, W^k, V^k, \bar{V}^k, F)$ *is satisfied iff $a_e \leq 1$ for all $e \in \mathcal{E}$.*

▶ **Lemma 10.** *If* $\mathrm{COLD}(X, W^k, V^k, \bar{V}^k, F)$ *is satisfied and E splits X into sets of size 4, then* $\mathrm{SAT}_{\leq 3/4,S,\bar{S}}(E, W^k, V^k, \bar{V}^k, F)$ *is satisfied iff $a_e \leq 3$ for all $e \in \mathcal{E}$.*

The next defined predicate is the following:

$$\begin{aligned}\mathrm{CTORI}'(X^2, Y^2, Z, W^k, V^k, \bar{V}^k, F) : \mathrm{TORI}'(X^2, Y^2, Z)\\ \wedge \mathrm{COLD}((X^2, Y^2, Z), W^k, V^k, \bar{V}^k, F),\end{aligned}$$

which simply implies applying labels (without any constraints) to the vertices of the tori. In the original predicate, which uses TORI instead of TORI′, there is no coordinate Z. Clearly, this predicate is satisfiable in the sense that any collection of pairs of tori of even size which is labeled in a manner that satisfies the requirement $W_i \sim \mathrm{Bern}(\frac{1}{2})$ has a corresponding representation by $X^2, Y^2, Z, W^k, V^k, \bar{V}^k, F$. In particular, this W_i requirement is satisfied by any labeling in which any pair of corresponding vertices (those which differ only in the Z

coordinate) has labels of the same value but opposite sign. This is because negating the sign of a label corresponds to negating all of its bits. For a set of labels $\mathcal{L} = \{0, \ldots, l-1\}$, we now label the vertices with labels from the set $\{0, \ldots, 4l-1\}$ and so we set $k = 4l+1$. For any $i \in \{0, \ldots, l\}$, $j \in \{0, \ldots, 3\}$, we identify all four labels $4i+j$ with the original label i, referring to any vertex whose label is $4i+j$ as a *group j vertex*. The group of a vertex is used to orient it relative to its neighbors – this is achieved by the following predicate:

$$
\begin{aligned}
&\mathrm{OTORI}'(X^2, Y^2, Z, W^k, V^k, \bar{V}^k, F) : \\
&\ \mathrm{CTORI}'(X^2, Y^2, Z, W^k, V^k, \bar{V}^k, F) \\
&\wedge\ \mathrm{SAT}_{\neq 1/2,\{k\},\varnothing}((X_1, X_2, Y_1, Z), W^k, V^k, \bar{V}^k, F) \\
&\wedge\ \mathrm{SAT}_{\neq 1/2,\{k\},\varnothing}((X_1, X_2, Y_2, Z), W^k, V^k, \bar{V}^k, F) \\
&\wedge\ \mathrm{SAT}_{\neq 1/2,\{k\},\varnothing}((X_1, Y_1, Y_2, Z), W^k, V^k, \bar{V}^k, F) \\
&\wedge\ \mathrm{SAT}_{\neq 1/2,\{k\},\varnothing}((X_2, Y_1, Y_2, Z), W^k, V^k, \bar{V}^k, F) \\
&\wedge \bigwedge_{j_1, j_2 \in J_1} \big(\mathrm{SAT}_{\leq 1/2,\varnothing,\{1,\ldots,k\}\setminus\{j_1,j_2\}}((X_1, X_2, Y_1, Z), W^k, V^k, \bar{V}^k, F) \\
&\qquad\qquad \wedge \mathrm{SAT}_{\leq 1/2,\varnothing,\{1,\ldots,k\}\setminus\{j_1,j_2\}}((X_1, X_2, Y_2, Z), W^k, V^k, \bar{V}^k, F) \\
&\qquad\qquad \wedge \mathrm{SAT}_{\leq 1/2,\{1,\ldots,k\}\setminus\{j_1,j_2\},\varnothing}((X_1, X_2, Y_1, Z), W^k, V^k, \bar{V}^k, F) \\
&\qquad\qquad \wedge \mathrm{SAT}_{\leq 1/2,\{1,\ldots,k\}\setminus\{j_1,j_2\},\varnothing}((X_1, X_2, Y_2, Z), W^k, V^k, \bar{V}^k, F)\big) \\
&\wedge \bigwedge_{j_1, j_2 \in J_2} \big(\mathrm{SAT}_{\leq 1/2,\varnothing,\{1,\ldots,k\}\setminus\{j_1,j_2\}}((X_1, Y_1, Y_2, Z), W^k, V^k, \bar{V}^k, F) \\
&\qquad\qquad \wedge\ \mathrm{SAT}_{\leq 1/2,\varnothing,\{1,\ldots,k\}\setminus\{j_1,j_2\}}((X_2, Y_1, Y_2, Z), W^k, V^k, \bar{V}^k, F) \\
&\qquad\qquad \wedge\ \mathrm{SAT}_{\leq 1/2,\{1,\ldots,k\}\setminus\{j_1,j_2\},\varnothing}((X_1, Y_1, Y_2, Z), W^k, V^k, \bar{V}^k, F) \\
&\qquad\qquad \wedge\ \mathrm{SAT}_{\leq 1/2,\{1,\ldots,k\}\setminus\{j_1,j_2\},\varnothing}((X_2, Y_1, Y_2, Z), W^k, V^k, \bar{V}^k, F)\big) \\
&\wedge\ \mathrm{SAT}_{\leq 1/2,\{k\},\varnothing}((X_1, X_2, Y_1, Y_2), W^k, V^k, \bar{V}^k, F) \\
&\wedge\ \mathrm{SAT}_{\leq 1/2,\varnothing,\{k\}}((X_1, X_2, Y_1, Y_2), W^k, V^k, \bar{V}^k, F),
\end{aligned}
$$

where

$$
\begin{aligned}
J_1 &= \{(j_1, j_2) \in \{1, \ldots, k-1\} : \{j_1 \bmod 4, j_2 \bmod 4\} \notin \{\{0,1\},\{2,3\}\}\}, \\
J_2 &= \{(j_1, j_2) \in \{1, \ldots, k-1\} : \{j_1 \bmod 4, j_2 \bmod 4\} \notin \{\{1,2\},\{0,3\}\}\}.
\end{aligned}
$$

The only difference between OTORI' and Li's original OTORI is the added variable Z and the final two $\mathrm{SAT}_{\leq 1/2}$ predicates. We have the following fact:

▶ **Lemma 11.** *If* $\mathrm{OTORI}'(X^2, Y^2, Z, W^k, V^k, \bar{V}^k, F)$ *is satisfied, then the following statements hold:*

1. *within each torus, all vertices' labels have the same sign;*
2. *any two vertices differing only in the Z coordinate have opposite sign;*
3. *any pair of vertices connected by a vertical edge either has groups 1 and 0 or 2 and 3;*
4. *any pair of vertices connected by a horizontal edge either has groups 1 and 2 or 3 and 0.*

Proof. Recall that fixing the variable Z along with any three of the variables X_1, X_2, Y_1, Y_2 leaves two possible values for the remaining variable. These correspond to two vertices of an edge, as illustrated in Figure 4. Therefore, the first four $\mathrm{SAT}_{\neq 1/2,\{k\},\varnothing}$ predicates state that for any edge (u, v), the value of w_k for u and v cannot differ, which implies exactly Point 1 above. Point 2 follows directly from the last two $\mathrm{SAT}_{\leq 1/2}$ predicates – of the two vertices which differ only in the Z coordinate, at most one can have $w_k = 0$ and at most one can have $w_k = 1$.

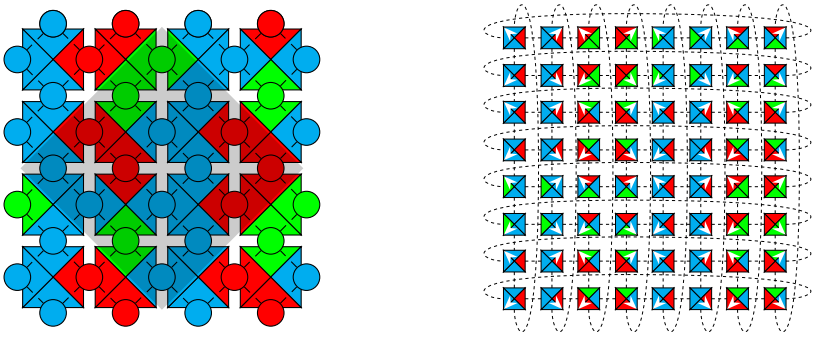

Figure 5 Left: Li's representation of a 4×4 torus coloring and the 4×4 tiling that it yields. The conversion from 16 vertices to 32 gives the tiling an additional diagonal periodicity. Combined with the fact that this does not work for non-square $n \times m$ tilings (without arranging them into a $\mathrm{lcm}(n,m) \times \mathrm{lcm}(n,m)$ square), this proves problematic for restricting the size of a periodic tiling. Right: the corresponding 8×8 torus labeling in our representation. Each tile has 4 labels, one for each corner of the tile (indicated by the arrow in this case). The tiles do not need to be Wang tiles.

For the next two points, note that for $j_1, j_2 \neq k$, we have $\mathrm{sat}(w^k, \{1,\dots,k\} \setminus \{j_1, j_2\}, \varnothing) = 1$ iff w^k represents one of the labels $\{-j_1, -j_2\}$ – for positive labels, we have $w_k = 0$ and for negative labels other than $-j_1, -j_2$, we have $w_i = 0$ for some $i \notin \{j_1, j_2\}$. Analogously, $\mathrm{sat}(w^k, \varnothing, \{1,\dots,k\} \setminus \{j_1, j_2\}) = 1$ iff w^k represents one of $\{+j_1, +j_2\}$. Thus, the four $\mathrm{SAT}_{\leq 1/2, \varnothing, \{1,\dots,k\} \setminus \{j_1,j_2\}}$ predicates within the conjunction over $j_1, j_2 \in J_1$ imply exactly Point 3, with the conjunction over $j_1, j_2 \in J_2$ implying Point 4 analogously. Clearly, OTORI′ is satisfiable – an example torus (with coordinate Z omitted) is shown in Figure 4. ◀

4.6 Tiles

Li denotes the set of four possible vertices when (Z, X_i, Y_j) is fixed as a *type ij face* for any $i, j \in \{1, 2\}$ – e. g. fixing (Z, X_1, Y_2) yields a type 12 face. The relation between face types and vertex groups is illustrated in Figure 4.

Li's construction utilizes the fact that the four corner-neighbors of any type 11 face are type 22 faces and vice versa. Because the original construction makes use of a Wang tiling system, these connecting corner vertices can represent the edge colors of the touching tiles. An example is shown in Figure 5. The diagonal nature of this tiling proves problematic when we wish to restrict its size, thus we simply use faces (of type 11) to directly represent tiles, which also allows us to use the general form of tiling systems. Figure 5 illustrates a torus labeling in our representation. For given tiling system $\mathcal{D} = (D, H, V)$, we define the following sets:

$$\mathcal{D}_{11} = \{(4t+1, 4t+2, 4t+3, 4t) : t \in D\},$$
$$\mathcal{D}_{12} = \{(4v, 4v+3, 4u+2, 4u+1) : (u,v) \in V\},$$
$$\mathcal{D}_{21} = \{(4u+2, 4u+1, 4v, 4v+3) : (u,v) \in H\},$$

and for each $i \in \{11, 12, 21\}$,

$$I_i = \{j_1, \dots, j_4 \in \{1,\dots,k-1\} : j_i \bmod 4 \text{ distinct}, \{j_1,\dots,j_4\} \notin \mathcal{D}_i\}.$$

The final predicate to enforce that the coloring represents a valid tiling (of size at most $m \times n$) is defined as follows:

$$\begin{aligned}\mathrm{TTORI}'_{\mathcal{D}} \ : \exists X^2 \le m, Y^2 \le n, Z \le_i 2, W^k \le_i 2, V^k \le_i 2, \bar{V}^k \le_i 2, F \le_i 2 : \\ \mathrm{OTORI}'(X^2, Y^2, Z, W^k, V^k, \bar{V}^k, F) \\ \wedge \bigwedge_{j_1,\dots,j_4 \in I_{11}} \big(\mathrm{SAT}_{\le 3/4,\varnothing,\{1,\dots,k\}\setminus\{j_1,\dots,j_4\}}((X_1,Y_1,Z),W^k,V^k,\bar{V}^k,F) \\ \wedge \mathrm{SAT}_{\le 3/4,\{1,\dots,k\}\setminus\{j_1,\dots,j_4\},\varnothing}((X_1,Y_1,Z),W^k,V^k,\bar{V}^k,F)\big) \\ \wedge \bigwedge_{j_1,\dots,j_4 \in I_{12}} \big(\mathrm{SAT}_{\le 3/4,\varnothing,\{1,\dots,k\}\setminus\{j_1,\dots,j_4\}}((X_1,Y_2,Z),W^k,V^k,\bar{V}^k,F) \\ \wedge \mathrm{SAT}_{\le 3/4,\{1,\dots,k\}\setminus\{j_1,\dots,j_4\},\varnothing}((X_1,Y_2,Z),W^k,V^k,\bar{V}^k,F)\big) \\ \wedge \bigwedge_{j_1,\dots,j_4 \in I_{21}} \big(\mathrm{SAT}_{\le 3/4,\varnothing,\{1,\dots,k\}\setminus\{j_1,\dots,j_4\}}((X_2,Y_1,Z),W^k,V^k,\bar{V}^k,F) \\ \wedge \mathrm{SAT}_{\le 3/4,\{1,\dots,k\}\setminus\{j_1,\dots,j_4\},\varnothing}((X_2,Y_1,Z),W^k,V^k,\bar{V}^k,F)\big).\end{aligned}$$

▶ **Lemma 12.** $\mathrm{TTORI}'_{\mathcal{D}}$ *is satisfied iff* $\mathcal{D}$ *admits a periodic tiling of size at most* $m \times n$.

Proof. All implicit bounds follow from previously defined predicates used within $\mathrm{TTORI}'_{\mathcal{D}}$: TORI' for Z, SW for W^k, and FLIP for V^k, $\bar{V}^k$, F. The bounds of m, n for X^2, Y^2 imply that the tori can take any even size up to $2m \times 2n$.

For the "only if" direction, note that the three conjunctions, similarly as in the predicate OTORI, forbid faces of type 11, 12, 21 from being not of the form in $\mathcal{D}_{11}, \mathcal{D}_{12}, \mathcal{D}_{21}$, respectively. Clearly, the set $\mathcal{D}_{11}$ consists of exactly those tiles (represented as type 11 faces) which are in D. Similarly, $\mathcal{D}_{12}, \mathcal{D}_{21}$ consist of all those "glue" type 12 and 21 faces which are allowed by H, V respectively. Therefore, in a distribution satisfying $\mathrm{TTORI}'_{\mathcal{D}}$, the type 11 faces represent tiles and neighboring tiles respect the constraints H and V. Therefore, each torus corresponds to a periodic tiling by $\mathcal{D}$. Since the tori are of size at most $2m \times 2n$, the tiling is of size at most $m \times n$.

For the "if" direction, for any given tiling by $\mathcal{D}$ of size at most $m \times n$, we create a pair of corresponding tori, labeled such that one represents the positive version of this tiling and the other the negative version. The satisfiability arguments show that this can be done for any given tiling. ◀

4.7 Final construction

Once fully expanded, the $\mathrm{TTORI}'_{\mathcal{D}}$ predicate is of the form (omitting the bounds for clarity)

$$\exists B, X^2, Y^2, W^k, V^k, \bar{V}^k, F, \dots : \Big(B \sim \mathrm{Bern}(1/2) \wedge \textstyle\bigwedge_i (I(A_i; B_i | C_i) = 0)\Big),$$

where A_i, B_i, C_i are tuples of the quantified variables. Its negation can be equivalently rewritten:

$$\begin{aligned}&\neg\exists B, \dots : \big(B \sim \mathrm{Bern}(1/2) \wedge \textstyle\bigwedge_i (I(A_i; B_i | C_i) = 0)\big) \\ \Leftrightarrow\ &\forall B, \dots : \big(B \not\sim \mathrm{Bern}(1/2) \vee \neg \textstyle\bigwedge_i (I(A_i; B_i | C_i) = 0)\big) \\ \Leftrightarrow\ &\forall B, \dots : \big(\big(\textstyle\bigwedge_i (I(A_i; B_i | C_i) = 0)\big) \Rightarrow B \not\sim \mathrm{Bern}(1/2)\big) \\ \Leftrightarrow\ &\forall B, \dots : \big(\big(\mathrm{UNIF}(B) \wedge |B| \le 2 \wedge \textstyle\bigwedge_i (I(A_i; B_i | C_i) = 0)\big) \Rightarrow H(B) = 0\big).\end{aligned}$$

The last equivalence holds because a variable B with $|B| \le 2$ and $\mathrm{UNIF}(B)$ can either be uniform over one value and have entropy 0 or uniform over two values and have entropy 1. This final form is a valid CI implication with (partial) cardinality bounds. In the bounded case, the established cardinality bounds are preserved.

Enforcing the usage of a designated tile

We extend Li's above construction to enforce that a given tile t be used in the tiling. Recall from Lemma 6 that for any $w^k \in \{0,1\}^k$,

$$H(F|V_{\{i:w_i=1\}}, \bar{V}_{\{i:w_i=0\}}, W^k) = \mathbf{P}(W^k = w^k).$$

Furthermore, variable F is constrained by the predicate $\mathrm{UNIF}_2(F)$. It can be equivalently restated as $\mathrm{UNIF}_=(F,B)$, where $\mathrm{UNIF}_=$ is defined by Li as follows:

$$\mathrm{UNIF}_=(Y,Z) : \exists U^3 \leq_i K_Y : \mathrm{TRIPLE}(Y,U_1,U_2) \wedge \mathrm{TRIPLE}(Z,U_1,U_3).$$

Clearly, $\mathrm{UNIF}_=(F,B)$ holds iff F and B are both uniform with the same cardinality and hence $\mathrm{UNIF}_2(F)$ can be replaced by $\mathrm{UNIF}_=(F,B)$. Therefore, if the (conditional) entropy of F is nonzero, then the entropy of B must also be nonzero and so we must have $B \sim \mathrm{Bern}(\frac{1}{2})$.

Given a designated tile t, let $w^k \in T_k$ be the value of W^k corresponding to label t (without loss of generality, of vertex group 0 and positive sign) and $S = \{i : w_i = 1\}, \bar{S} = \{i : w_i = 0\}$. The modified implication is as follows:

$$\forall B, \ldots : \big(\big(\mathrm{UNIF}(B) \wedge |B| \leq 2 \wedge \bigwedge\nolimits_i (I(A_i; B_i|C_i) = 0)\big) \Rightarrow H(F|V_S, \bar{V}_{\bar{S}}, W^k) = 0\big).$$

The counterexamples of this implication are exactly those labelings which use the tile t. Altogether, this chapter has shown the following theorem:

▶ **Theorem 2.** *For any given tiling system $\mathcal{D}$ along with tile t and natural numbers m, n, there exists a bounded CI implication which holds iff $\mathcal{D}$ does not admit a periodic tiling of size at most $m \times n$ which makes use of tile t. Moreover, the implication can be computed in time polynomial with regard to the size of the tiling system, while the bounds can be computed in time polynomial w.r.t. to the size of m, n.*

Theorem 2 gives exactly a polynomial-time many-one reduction from Periodic Bounded Tiling to the complement of Bounded CI Implication, in particular because m, n and K are encoded in the same manner. In the case of Constant-bounded CI Implication, the above argument does not work since we have constant bounds larger than 2 as well as bounds whose value depends on the input. However, any variable X with cardinality bound 2^j can be replaced by the tuple $(X_1, \ldots, X_j)$, where X_i has cardinality bound 2 for each $i \in \{1, \ldots, j\}$. These are clearly equivalent, since each X_i can correspond to the i-th bit of X. More generally, a variable with cardinality bound l can be replaced by $(X_1, \ldots, X_{\lceil \log l \rceil})$, with each X_i's cardinality bounded by 2 and the additional requirement $\mathrm{UNIF}'_k((X_1, \ldots, X_{\lceil \log l \rceil}))$. Here $\mathrm{UNIF}'_k(Y)$ is a modification of the UNIF_k predicate such that it enforces Y being uniform with $|Y| \leq k$. The construction of both of these follows closely that of Li and is given in detail in the full version. Note that the resulting predicate can be large, but this is only important when the bound is not constant – in our case these are only the two bounds $X^2 \leq m, Y^2 \leq n$. To avoid this issue, we reduce from Power-of-two Periodic Bounded Tiling – then X_i, Y_i (for $i \in \{1,2\}$) have bounds $2^m, 2^n$ respectively, while the remaining variables have constant bounds. Replacing X_1, X_2, Y_1, Y_2 by tuples of binary variables as shown above, we obtain a CI implication of size $N \cdot O(m+n)$, where N is the size of the original implication. The number of random variables grows similarly. The newly created bounds are all constant, and the reduction takes time polynomial with regard to the input size. The values of the remaining constant bounds are known and therefore each such variable can be converted in constant time. Together, these two results yield Theorem 1.

References

1 John F. Canny. Some algebraic and geometric computations in PSPACE. In Janos Simon, editor, *Proceedings of the 20th Annual ACM Symposium on Theory of Computing, May 2-4, 1988, Chicago, Illinois, USA*, pages 460–467. ACM, 1988. doi:10.1145/62212.62257.

2 Dan Geiger and Judea Pearl. Logical and algorithmic properties of conditional independence and graphical models. *The Annals of Statistics*, 21(4):2001–2021, 1993.

3 Daniel Gottesman and Sandy Irani. The quantum and classical complexity of translationally invariant tiling and hamiltonian problems. *Proceedings - Annual IEEE Symposium on Foundations of Computer Science, FOCS*, May 2009. doi:10.1109/FOCS.2009.22.

4 Miika Hannula, Åsa Hirvonen, Juha Kontinen, Vadim Kulikov, and Jonni Virtema. Facets of distribution identities in probabilistic team semantics. In Francesco Calimeri, Nicola Leone, and Marco Manna, editors, *Logics in Artificial Intelligence - 16th European Conference, JELIA 2019, Rende, Italy, May 7-11, 2019, Proceedings*, volume 11468 of *Lecture Notes in Computer Science*, pages 304–320. Springer, 2019. doi:10.1007/978-3-030-19570-0_20.

5 Mahmoud Abo Khamis, Phokion G. Kolaitis, Hung Q. Ngo, and Dan Suciu. Decision problems in information theory. In Artur Czumaj, Anuj Dawar, and Emanuela Merelli, editors, *47th International Colloquium on Automata, Languages, and Programming, ICALP 2020, July 8-11, 2020, Saarbrücken, Germany (Virtual Conference)*, volume 168 of *LIPIcs*, pages 106:1–106:20. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2020. doi:10.4230/LIPIcs.ICALP.2020.106.

6 Lukas Kühne and Geva Yashfe. On entropic and almost multilinear representability of matroids. *CoRR*, abs/2206.03465, 2022. arXiv:2206.03465.

7 Harry R. Lewis and Christos H. Papadimitriou. *Elements of the theory of computation, 2nd Edition*. Prentice Hall, 1998.

8 Cheuk Ting Li. The undecidability of conditional affine information inequalities and conditional independence implication with a binary constraint. *IEEE Transactions on Information Theory*, 68(12):7685–7701, 2022. doi:10.1109/TIT.2022.3190800.

9 Cheuk Ting Li. Undecidability of network coding, conditional information inequalities, and conditional independence implication. *IEEE Transactions on Information Theory*, 2023. doi:10.1109/TIT.2023.3247570.

10 Michał Makowski. On the complexity of the conditional independence implication problem with bounded cardinalities, 2024. Full version of this paper. arXiv:2408.02550.

11 Paul W. K. Rothemund and Erik Winfree. The program-size complexity of self-assembled squares (extended abstract). In *Proceedings of the Thirty-Second Annual ACM Symposium on Theory of Computing*, STOC '00, pages 459–468, New York, NY, USA, 2000. Association for Computing Machinery. doi:10.1145/335305.335358.

12 Peter van Emde Boas. The convenience of tilings. In A. Sorbi, editor, *Complexity, Logic, and recursion Theory*, volume 187 of *Lecture Notes in Pure and Applied Logic*, pages 331–363. Marcel Dekker, Inc., 1997.

13 Z. Zhang and R.W. Yeung. A non-Shannon-type conditional inequality of information quantities. *IEEE Transactions on Information Theory*, 43(6):1982–1986, 1997. doi:10.1109/18.641561.