

SPARSE KNESER GRAPHS ARE HAMILTONIAN

TORSTEN MÜTZE, JERRI NUMMENPALO, AND BARTOSZ WALCZAK

ABSTRACT. For integers $k \geq 1$ and $n \geq 2k + 1$, the *Kneser graph* $K(n, k)$ is the graph whose vertices are the k -element subsets of $\{1, \dots, n\}$ and whose edges connect pairs of subsets that are disjoint. The Kneser graphs of the form $K(2k + 1, k)$ are also known as the *odd graphs*. We settle an old problem due to Meredith, Lloyd, and Biggs from the 1970s, proving that for every $k \geq 3$, the odd graph $K(2k + 1, k)$ has a Hamilton cycle. This and a known conditional result due to Johnson imply that all Kneser graphs of the form $K(2k + 2^a, k)$ with $k \geq 3$ and $a \geq 0$ have a Hamilton cycle. We also prove that $K(2k + 1, k)$ has at least $2^{2^{k-6}}$ distinct Hamilton cycles for $k \geq 6$. Our proofs are based on a reduction of the Hamiltonicity problem in the odd graph to the problem of finding a spanning tree in a suitably defined hypergraph on Dyck words.

1. INTRODUCTION

The question whether a given graph has a Hamilton cycle is one of the oldest and most fundamental problems in graph theory and computer science, shown to be NP-complete in Karp’s seminal paper [Kar72]. The problem originates from the 19th-century “Hamilton puzzle”, which involves finding a Hamilton cycle along the edges of a dodecahedron. Efficient methods of generating Hamilton cycles in highly symmetric graphs (in particular, so-called Gray codes) are particularly important from the point of view of practical applications [Knu11, Sav97]. Still, for various natural and extensively studied families of graphs, it is conjectured that a Hamilton cycle always exists, but finding one is a notoriously hard problem; see for instance [MN20, SW20]. In this paper, we focus on a well-known instance of this phenomenon—the so-called Kneser graphs.

1.1. Kneser graphs. For any two integers $k \geq 1$ and $n \geq 2k + 1$, the *Kneser graph* $K(n, k)$ has the k -element subsets of $[n] := \{1, \dots, n\}$ as vertices and the pairs of those subsets that are disjoint as edges. These graphs were introduced by Lovász in his celebrated proof of Kneser’s conjecture [Lov78]. The proof uses topological methods to show that the chromatic number of $K(n, k)$ is equal to $n - 2k + 2$. Lovász’s result initiated an exciting line of research [Bár78, Gre02, Mat04, Zie02] and gave rise to the nowadays flourishing fields of topological combinatorics and computational topology, see e.g. [ČKM⁺13, ČKM⁺14]. Apart from the above, Kneser graphs have many other interesting properties. For instance, the maximum size of an independent set in $K(n, k)$ is equal to $\binom{n-1}{k-1}$, by the famous Erdős-Ko-Rado theorem [EKR61].

Torsten Mütze: Department of Computer Science, University of Warwick, Coventry CV47AL, United Kingdom

Jerri Nummenpalo: Department of Computer Science, ETH Zürich, 8092 Zürich, Switzerland

Bartosz Walczak: Department of Theoretical Computer Science, Faculty of Mathematics and Computer Science, Jagiellonian University, Kraków, Poland

E-mail addresses: torsten.mutze@warwick.ac.uk, njerri@inf.ethz.ch, walczak@tcs.uj.edu.pl.

An extended abstract of this paper appeared in *STOC 2018: Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, pages 912–919. ACM, New York, 2018.

Torsten Mütze is also affiliated with Charles University, Faculty of Mathematics and Physics, and was supported by Czech Science Foundation grant GA 19-08554S, and by German Science Foundation grant 413902284.

Bartosz Walczak was partially supported by National Science Centre of Poland grant 2015/17/D/ST1/00585.

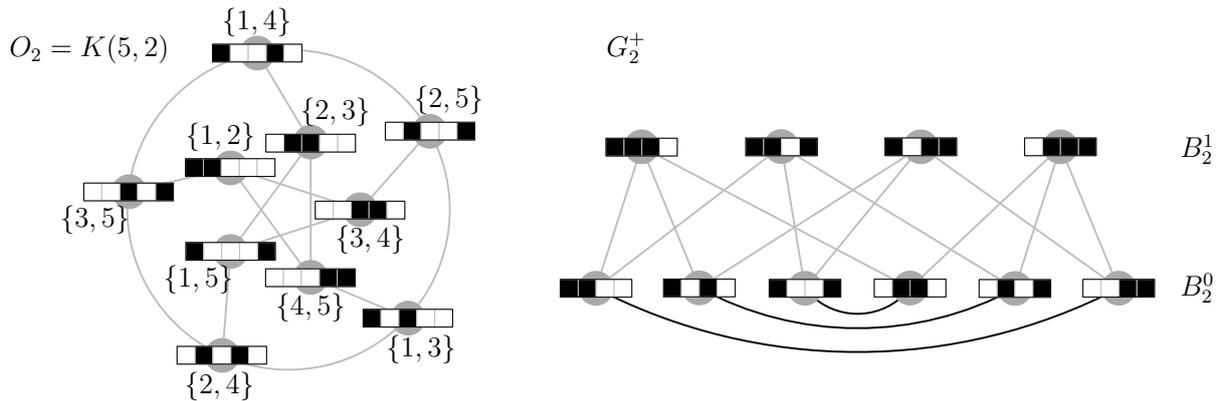


FIGURE 1. The Petersen graph $O_2 = K(5, 2)$ (left) and the graph G_2^+ (right) that is isomorphic to it. The isomorphism is defined in the proof of Lemma 4. The vertices of the Petersen graph are shown as 2-element subsets of $[5] = \{1, 2, 3, 4, 5\}$, together with the corresponding characteristic bitstrings. Black squares represent 1-bits and white squares represent 0-bits.

1.2. Hamilton cycles in Kneser graphs. As indicated before, it has long been conjectured that Kneser graphs have Hamilton cycles. Apart from one obvious exception, namely the Petersen graph $K(5, 2)$ shown in Figure 1, no other negative instances are apparent. Observe that Kneser graphs are vertex-transitive, that is, they look the same from the point of view of any vertex. This makes them an excellent test case for a famous and vastly more general conjecture due to Lovász [Lov70], which asserts that any connected and vertex-transitive graph has a Hamilton cycle, apart from the Petersen graph and four other exceptional instances.

We proceed by giving an account of the long history of finding Hamilton cycles in Kneser graphs. The degree of every vertex in $K(n, k)$ is $\binom{n-k}{k}$, so for fixed k , increasing n also increases the vertex degrees, which intuitively makes the task of finding a Hamilton cycle easier. The density is also witnessed by cliques of size $c \geq 3$, which are present for $n \geq ck$ and absent for $n < ck$. The sparsest case, for which finding a Hamilton cycle is intuitively hardest, is when $n = 2k + 1$. The corresponding graphs $O_k := K(2k + 1, k)$, for $k \geq 1$, are known as *odd graphs*. They include the Petersen graph $O_2 = K(5, 2)$. The odd graphs O_2 and O_3 are illustrated in Figures 1 and 2, respectively. Note that all vertices in the odd graph O_k have degree $k + 1$, which is only logarithmic in the number of vertices. The conjecture that the odd graph O_k has a Hamilton cycle for every $k \geq 3$ originated in the 1970s, in papers by Meredith and Lloyd [ML72, ML73] and by Biggs [Big79]. A stronger version of the conjecture asserts that O_k even has $\lfloor (k + 1)/2 \rfloor$ edge-disjoint Hamilton cycles. Already Balaban [Bal72] exhibited a Hamilton cycle for the cases $k = 3$ and $k = 4$, and Meredith and Lloyd described one for $k = 5$ and $k = 6$. Later, Mather [Mat76] also solved the case $k = 7$. With the help of computers, Shields and Savage [SS04] found Hamilton cycles in O_k for all values of k up to 13. They also found Hamilton cycles in $K(n, k)$ for all $n \leq 27$ (except for the Petersen graph).

There is a long line of research devoted to proving that sufficiently dense Kneser graphs have a Hamilton cycle. Heinrich and Wallis [HW78] showed that $K(n, k)$ has a Hamilton cycle if $n \geq 2k + k/(\sqrt[k]{2} - 1) = (1 + o(1))k^2/\ln 2$. This was improved by B. Chen and Lih [CL87], whose results imply that $K(n, k)$ has a Hamilton cycle if $n \geq (1 + o(1))k^2/\log k$, see [CI96]. In another breakthrough, Y. Chen [Che00] showed that $K(n, k)$ is Hamiltonian when $n \geq 3k$. A particularly nice and clean proof for the cases where $n = ck$, $c \in \{3, 4, \dots\}$, was obtained

by Y. Chen and Füredi [CF02]. Their proof uses Baranyai's well-known partition theorem for complete hypergraphs [Bar75] to partition the vertices of $K(ck, k)$ into cliques of size c . This proof method was extended by Bellmann and Schülke to any $n \geq 4k$ [BS19]. The asymptotically best result currently known, again due to Y. Chen [Che03], is that $K(n, k)$ has a Hamilton cycle if $n \geq (3k + 1 + \sqrt{5k^2 - 2k + 1})/2 = (1 + o(1))2.618 \dots k$.

Another line of attack towards proving Hamiltonicity is to find long cycles in $K(n, k)$. To this end, Johnson [Joh04] showed that there exists a constant $c > 0$ such that the odd graph O_k has a cycle that visits at least a $1 - c/\sqrt{k}$ proportion of all vertices, which is almost all vertices as k tends to infinity. This was generalized and improved in [MS17], where it was shown that $K(n, k)$ has a cycle visiting a $2k/n$ proportion of all vertices. The last result implies that O_k has a cycle visiting a $1 - 1/(2k + 1)$ proportion of the vertices (e.g., the Petersen graph O_2 has a cycle that visits 8 of its 10 vertices).

A different relaxation of proving Hamiltonicity is to construct a cycle factor, that is, a collection of vertex-disjoint cycles that together cover all vertices of the graph. From this point of view, a Hamilton cycle is a cycle factor consisting of a single cycle. In this direction, Johnson and Kierstead [JK04] showed that the edges of O_k can be partitioned into cycle factors for odd k and into cycle factors and one matching for even k . A different cycle factor in O_k , which turns out to be crucial for our present result, was constructed in [MSW18]. It is shown in Figure 2 for the case $k = 3$.

1.3. Bipartite Kneser graphs. Bipartite Kneser graphs form another family of vertex-transitive graphs closely related to Kneser graphs. The *bipartite Kneser graph* $H(n, k)$ has all k -element and all $(n - k)$ -element subsets of $[n]$ as vertices and all pairs of these subsets such that one is contained in the other as edges. It has been a long-standing problem to show that $H(n, k)$ has a Hamilton cycle. A detailed account of the historic developments is given in [MS17]. Also here, the sparsest case $H(2k + 1, k)$ resisted all attacks for more than three decades, and the question whether $H(2k + 1, k)$ has a Hamilton cycle became known as the *middle levels conjecture*. This conjecture has been recently solved affirmatively in [Müt16] (see also [GMN18]), and the general case, the Hamiltonicity of $H(n, k)$, has been settled subsequently in [MS17]. Note that proving Hamiltonicity for the Kneser graph $K(n, k)$ is arguably harder than for the bipartite Kneser graph $H(n, k)$. In particular, proving that the odd graphs $O_k = K(2k + 1, k)$ are Hamiltonian is harder than the middle levels conjecture. Specifically, from a Hamilton cycle (x_1, \dots, x_N) in $K(n, k)$, where $N = \binom{n}{k}$, we can easily construct a Hamilton cycle or a Hamilton path in $H(n, k)$, as follows. Consider the sequences $C_1 := (x_1, \bar{x}_2, x_3, \bar{x}_4, \dots)$ and $C_2 := (\bar{x}_1, x_2, \bar{x}_3, x_4, \dots)$, where $\bar{x}_i := [n] \setminus x_i$. If N is odd, then C_1 and C_2 together form a Hamilton cycle in $H(n, k)$. If N is even, then C_1 and C_2 are two cycles in $H(n, k)$ that can be joined to form a Hamilton path. In fact, the arguments given in this paper easily give a Hamilton cycle in $H(2k + 1, k)$ for all $k \geq 1$, providing an alternative proof of the middle levels conjecture; see Section 6.

1.4. Our results. We prove that the odd graphs $O_k = K(2k + 1, k)$ with $k \geq 3$ contain Hamilton cycles. That is, we resolve the sparsest case of the conjecture on the Hamiltonicity of Kneser graphs in the affirmative.

Theorem 1. *For any integer $k \geq 3$, the odd graph $O_k = K(2k + 1, k)$ has a Hamilton cycle.*

Using the conditional results proved by Johnson [Joh11], Theorem 1 immediately yields the following more general statement.

Theorem 2. *For any integers $k \geq 3$ and $a \geq 0$, the Kneser graph $K(2k + 2^a, k)$ has a Hamilton cycle.*

We also establish the following counting version of Theorem 1.

Theorem 3. *For any integer $k \geq 6$, the odd graph $O_k = K(2k+1, k)$ has at least $2^{2^{k-6}}$ distinct Hamilton cycles.*

The double-exponential growth of the number of Hamilton cycles guaranteed by Theorem 3 is essentially best possible: since O_k has $\binom{2k+1}{k}$ vertices, the number of Hamilton cycles in O_k is at most $(\binom{2k+1}{k})! = 2^{2^{\Theta(k)}}$. Note also that applying automorphisms of O_k to a single Hamilton cycle yields at most $(2k+1)! = 2^{\Theta(k \log k)}$ distinct Hamilton cycles, substantially fewer than guaranteed by Theorem 3. In other words, Theorem 3 is not an immediate consequence of Theorem 1.

1.5. Gray code algorithms. Hamilton cycles in Kneser graphs and bipartite Kneser graphs are closely related to Gray codes. A *combinatorial Gray code* is the algorithmic problem of generating all objects in a combinatorial class, such as bitstrings, permutations, combinations, partitions, trees, or triangulations, etc., in some well-defined order. Gray codes have found widespread use in areas such as circuit testing, signal encoding, data compression, graphics, and image processing etc.—see the survey [Sav97] and the references therein. The ultimate goal for Gray code algorithms is to generate each new object from the previous one in constant time, which entails that consecutive objects may differ only by a constant amount. A Gray code thus corresponds to a Hamilton cycle in a graph whose vertices are the combinatorial objects and whose edges connect objects that differ only by such an elementary transformation. More than half of the most recent volume of Knuth’s seminal series *The Art of Computer Programming* [Knu11] is devoted to this fundamental subject. The two hardest Gray code problems mentioned in Knuth’s book (Problem 71 in Section 7.2.1.2 and Problem 56 in Section 7.2.1.3), including the middle levels conjecture, have been solved in the meantime, and efficient algorithms to generate these Gray code have been developed in [MN20] and [SW20]. Recall from Section 1.3 that Hamiltonicity of the odd graphs is arguably harder than the middle levels conjecture.

Our proof of Theorem 1 is constructive and translates straightforwardly into an algorithm to compute a Hamilton cycle in the odd graph O_k in polynomial time (polynomial in the size of the graph, which is exponential in k). We can identify each k -element subset of $[2k+1]$ with a bitstring of length $2k+1$, where the i th bit is set to 1 if the element i is contained in the set and it is set to 0 otherwise; see Figure 1. A Hamilton cycle in the odd graph thus corresponds to a Gray code listing of all bitstrings of length $2k+1$ with exactly k many 1-bits, such that consecutive bitstrings differ in all but one position. It remains open whether our proof can be translated into a constant-time algorithm to generate this Gray code, that is, an algorithm that in each step computes the bit that is not flipped in constant time, using only $\mathcal{O}(k)$ memory space and polynomial initialization time. To avoid costly complementation operations, such an algorithm could maintain two bitstrings, one the complement of the other, along with a flag indicating which of the two bitstrings is the current one; then, in each step, only a single bit in both bitstrings and the flag would need to be flipped.

1.6. Proof idea. We construct a Hamilton cycle in the odd graph O_k as follows; see Figure 2. We start with the cycle factor \mathcal{C}_k in the odd graph O_k described in [MSW18]. It has the property that all of its cycles have the same length $2k+1$ and the number of cycles is the k th Catalan number. Furthermore, the cycles in \mathcal{C}_k can be identified with so-called *Dyck words* of length $2k$, that is, bitstrings of length $2k$ with the property that every prefix has at least as many 1-bits as 0-bits. It is well known that the number of such Dyck words is equal to the k th Catalan number [Sta99].

Given the cycle factor \mathcal{C}_k , we modify it locally to join its cycles into a single Hamilton cycle in O_k . Each such modification involves ℓ cycles C_1, \dots, C_ℓ from the factor \mathcal{C}_k and a 2ℓ -cycle C'

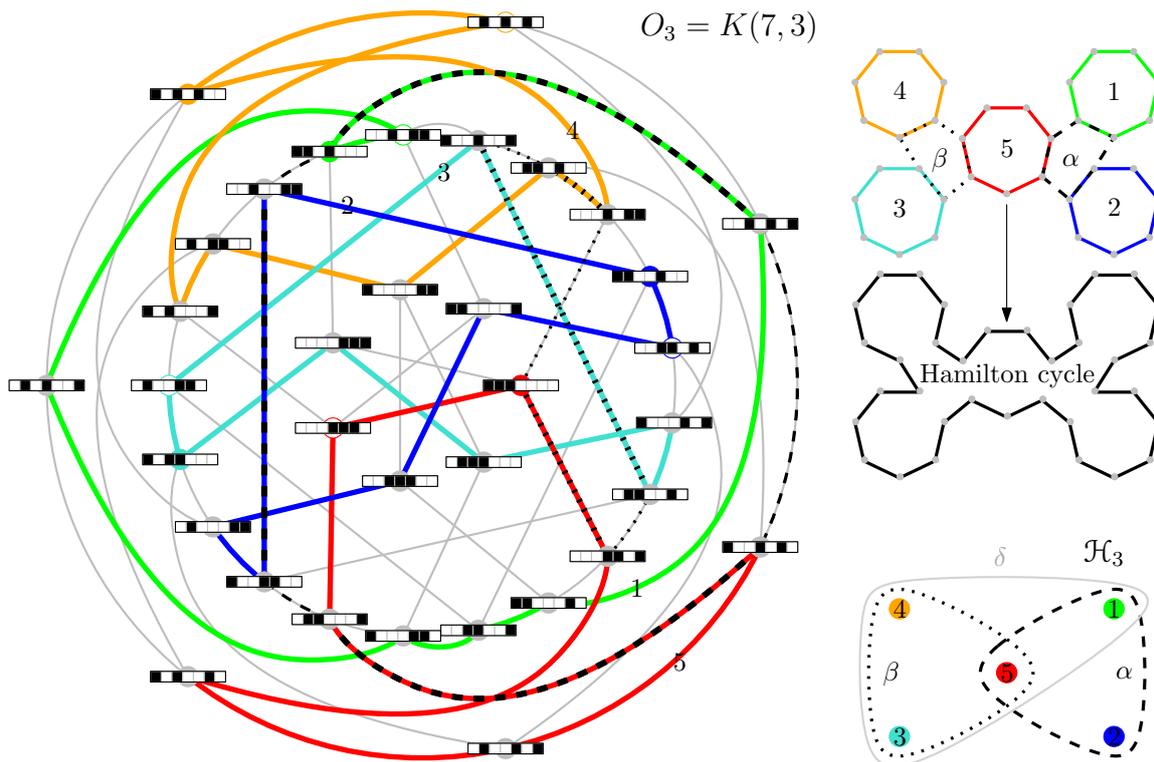


FIGURE 2. Illustration of our Hamiltonicity proof for the odd graph $O_3 = K(7, 3)$. The vertices are represented as bitstrings, where 1-bits are drawn as black squares and 0-bits as white squares. The bold cycles labeled $1, \dots, 5$ constitute the cycle factor \mathcal{C}_3 . Two flipping cycles of length 6 are highlighted by dashed and dotted lines. The five cycles from \mathcal{C}_3 correspond to the vertices of the hypergraph \mathcal{H}_3 , and the two flipping cycles correspond to the hyperedges α and β . There is another flipping cycle of length 8 in the graph O_3 , corresponding to the hyperedge δ in \mathcal{H}_3 , but this cycle is not shown in the figure. As $\{\alpha, \beta\}$ is a spanning tree in \mathcal{H}_3 , taking the symmetric difference of the edge sets of the cycles in \mathcal{C}_3 with the edge sets of the two corresponding flipping cycles indicated in the figure yields a Hamilton cycle in the graph O_3 . This construction step is shown schematically in the top right part of the figure.

that shares exactly one edge with each of C_1, \dots, C_ℓ . Specifically, C' shares every second of its edges with one of the ℓ cycles, and every other edge of C' goes between two different cycles. Consequently, taking the symmetric difference of the edge set of C' with the edge sets of C_1, \dots, C_ℓ yields a single cycle on the vertex set of all C_1, \dots, C_ℓ . We call a cycle C' with this property a *flipping cycle*. In Figure 2, two flipping 6-cycles are highlighted with dashed and dotted lines. We perform this operation simultaneously with an appropriate set of mutually edge-disjoint flipping cycles so as to join all cycles in \mathcal{C}_k into a single cycle. Although the joining operation can work with flipping 2ℓ -cycles for any $\ell \geq 2$, we will use only 6-cycles ($\ell = 3$) and 8-cycles ($\ell = 4$). We cannot use flipping 4-cycles ($\ell = 2$), because the odd graph O_k has no 4-cycles at all.

This approach can be formalized as follows. We construct a hypergraph \mathcal{H}_k whose vertices are the Dyck words of length $2k$ representing the cycles of the factor \mathcal{C}_k . Each ℓ -edge (3-edge or 4-edge) of \mathcal{H}_k represents a flipping 2ℓ -cycle (6-cycle or 8-cycle, respectively) that can be used to join ℓ cycles from \mathcal{C}_k as described before. In the example illustrated in Figure 2, the

hypergraph \mathcal{H}_3 consists of three hyperedges labeled α , β , and δ of cardinalities 3, 3, and 4, respectively. Here is the key insight about the hypergraph \mathcal{H}_k : in order to prove that the odd graph O_k has a Hamilton cycle, it suffices to prove that the hypergraph \mathcal{H}_k has a *spanning tree*, that is, a connected and acyclic set of hyperedges covering all vertices. In such a spanning tree, any two hyperedges intersect in at most one element. For instance, the hypergraph \mathcal{H}_3 in Figure 2 has a spanning tree $\{\alpha, \beta\}$. The hypergraph \mathcal{H}_k that we construct has the property that the flipping cycles represented by the hyperedges in any spanning tree are mutually edge-disjoint. Consequently, every spanning tree in \mathcal{H}_k corresponds to a collection of flipping cycles such that the symmetric difference of their edge sets and the edges of the cycles in \mathcal{C}_k results in a Hamilton cycle in the odd graph O_k .

The proof of Theorem 3 exploits the degrees of freedom that are inherent in the construction above to provide double-exponentially many distinct spanning trees in \mathcal{H}_k , which give rise to double-exponentially many distinct Hamilton cycles in O_k . This general approach of reducing a Hamilton cycle problem to a spanning tree problem in a suitably defined auxiliary (hyper)graph has also been exploited in several other papers; see e.g. [GMN18, GSWW18, Hol17, HRW12, Joh09, Joh11, Müt16, SW20].

1.7. Outline of the paper. In Section 2, we introduce notation and terminology that will be used throughout this paper, and we recall the construction of the cycle factor \mathcal{C}_k given in [MSW18]. In Section 3, we describe how the cycles in \mathcal{C}_k are joined to form a Hamilton cycle in O_k , and we present the proofs of Theorems 1–3. The proofs of some technical lemmas are deferred to Sections 4 and 5. In Section 6, we give an alternative proof of the middle levels conjecture. We conclude with some open problems in Section 7.

2. PRELIMINARIES

2.1. Bitstrings and Dyck paths. A *bitstring* is a finite sequence of digits 0 and 1 called the *bits* of the bitstring. The empty bitstring is denoted by ϵ . The concatenation of two bitstrings x and y is denoted by xy . For every bitstring x , we define $x^0 := \epsilon$ and $x^n := x^{n-1}x$ for $n \geq 1$. The length of a bitstring x is denoted by $|x|$. The *complement* of a bitstring x , denoted by \bar{x} , is the bitstring obtained from x by *flipping* every bit, that is, by replacing every 1-bit by a 0-bit and vice versa.

The *weight* of a bitstring x is the number of 1-bits in x . We let B_k^0 and B_k^1 denote the sets of bitstrings of length $2k$ with weights k and $k+1$, respectively, and we let $B_k := B_k^0 \cup B_k^1$. It follows that $|B_k^0| = \binom{2k}{k}$, $|B_k^1| = \binom{2k}{k-1}$, and $|B_k| = \binom{2k+1}{k}$. We let D_k denote the set of bitstrings of length $2k$ with weight k and with the property that in every prefix, the number of 1-bits is at least the number of 0-bits. It is a well known fact that $|D_k| = \frac{1}{k+1} \binom{2k}{k} = \frac{1}{2k+1} \binom{2k+1}{k}$, which is the k th Catalan number. We also define $D := \bigcup_{k=0}^{\infty} D_k$, and we call every bitstring in D a *Dyck word*.

It is sometimes convenient to represent a Dyck word $x \in D_k$ by a *Dyck path* of length $2k$ in the integer lattice \mathbb{Z}^2 . Every 1-bit in the Dyck word x is represented by an *up-step*, which changes the current coordinates by $(+1, +1)$, and every 0-bit is represented by a *down-step*, which changes the current coordinates by $(+1, -1)$; see Figure 3. The prefix property from the definition of D_k corresponds to the property that the lattice path never goes below the abscissa.

For a Dyck word $x = b_1 b_2 \cdots b_{2k} \in D_k$, where $b_1, \dots, b_{2k} \in \{0, 1\}$, we let $\overleftarrow{x} := \overline{b_{2k} b_{2k-1} \cdots b_1}$. That is, \overleftarrow{x} is the complement of the reverse of x , which is itself a Dyck word in D_k . For example, if $x = 110010$, then $\overleftarrow{x} = 101100$. We call the operation $x \mapsto \overleftarrow{x}$ *mirroring*. In terms of Dyck path representation, it corresponds to taking the mirror image with respect to the vertical line $x = k$.

2.2. Graphs G_k and G_k^+ . We use standard graph-theoretic terminology, where the edges of every graph that we consider are unordered pairs of vertices of the form $\{u, v\}$. We define G_k

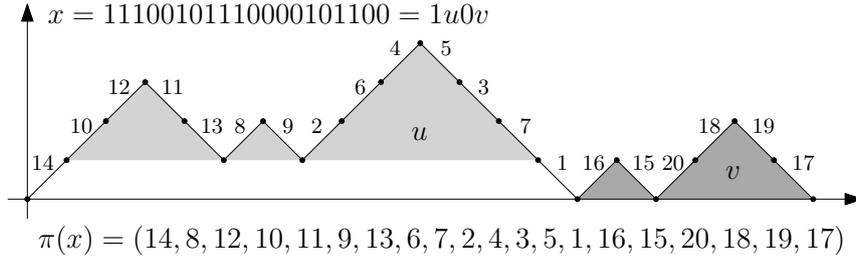


FIGURE 3. Dyck path representation of a Dyck word $x \in D_{10}$ and the permutation $\pi(x)$. The numbers on top of the Dyck path edges represent the order in which its up-steps and down-steps occur in the sequence $\pi(x)$, i.e., they are equal to the inverse permutation $(\pi(x))^{-1}$ when read from left to right.

as the graph with vertex set B_k and with edges that connect pairs of bitstrings that differ by exactly one bit. In other words, G_k is the subgraph of the $2k$ -dimensional hypercube induced by the bitstrings with weights k and $k + 1$. For $k \geq 1$, we also define G_k^+ as the graph obtained from G_k by adding all edges of the form $\{x, \bar{x}\}$ where $x \in B_k^0$. This construction is illustrated on the right hand side of Figure 1, where the edges $\{x, \bar{x}\}$ are highlighted in black. Observe that while the graph G_k is bipartite, the graph G_k^+ is not.

Lemma 4. *For every $k \geq 1$, the graph G_k^+ is isomorphic to the odd graph O_k .*

Proof. A natural isomorphism between G_k^+ and O_k is obtained by mapping every $x \in B_k^0$ to $x0$ and every $x \in B_k^1$ to $\bar{x}1$ and by interpreting the resulting bitstrings of length $2k + 1$ and weight k as characteristic vectors of k -element subsets of $[2k + 1]$. It is straightforward to verify that this mapping preserves edges and non-edges. \square

To prove Theorems 1 and 3, we will use Lemma 4 and construct Hamilton cycles in G_k^+ for all $k \geq 3$.

2.3. Cycle factor \mathcal{C}_k in G_k^+ . A *cycle factor* in a graph is a collection of vertex-disjoint cycles that together cover all vertices of the graph. The cycle factor \mathcal{C}_k in G_k^+ , which we will define shortly, was introduced and analyzed in [MSW18]. The cycles in \mathcal{C}_k correspond to Dyck words in D_k as follows. For every Dyck word $x \in D_k$, we define a permutation $\pi(x)$ of the set $[2k]$. Then, we define a path $P(x)$ in G_k whose subsequent vertices are obtained by starting from x and flipping the bits one by one at positions determined by the sequence $\pi(x) = (a_1, \dots, a_{2k})$, ending at \bar{x} . Finally, we add the edge $\{x, \bar{x}\}$ to $P(x)$, obtaining a cycle $C(x)$ in G_k^+ that becomes a member of \mathcal{C}_k .

We let (a_1, \dots, a_n) denote the sequence of integers a_1, \dots, a_n . We generalize this notation allowing a_i to be itself an integer sequence—in that case, if $a_i = (b_1, \dots, b_m)$, then (a_1, \dots, a_n) is shorthand for $(a_1, \dots, a_{i-1}, b_1, \dots, b_m, a_{i+1}, \dots, a_n)$. The empty integer sequence is denoted by $()$. For an integer sequence $\pi = (a_1, \dots, a_n)$ and an integer a , we define

$$a + \pi := (a + a_1, \dots, a + a_n), \quad a - \pi := (a - a_1, \dots, a - a_n).$$

It is clear that every non-empty Dyck word $x \in D$ has a unique decomposition of the form $x = 1u0v$, where $u, v \in D$; see Figure 3. Using this fact, for every Dyck word $x \in D$, we define an integer sequence $\pi(x)$ of length $|x|$ as follows, by induction on $|x|$:

$$\begin{aligned} \pi(\epsilon) &:= (), \\ \pi(1u0v) &:= (|u| + 2, (|u| + 2) - \pi(\bar{u}), 1, (|u| + 2) + \pi(v)) \quad \text{for any } u, v \in D. \end{aligned} \tag{1}$$

The sequence $\pi(x)$ for $x \in D_k$ satisfies the following properties:

- (i) $\pi(x)$ is a permutation of the set $[2k]$;
- (ii) if $\pi(x) = (a_1, \dots, a_{2k})$, then the bit of x at position a_i is 0 for i odd and 1 for i even.

To see why, we apply induction on $|x|$. The base case $x = \epsilon$ clearly satisfies both (i) and (ii). If $x \neq \epsilon$, then the sequence $\pi(x) = \pi(1u0v)$ is, by definition, a concatenation of four sequences that are, by induction, permutations of the sets $\{|u| + 2\}$, $\{(|u| + 2) - |u|, \dots, (|u| + 2) - 1\}$, $\{1\}$, and $\{(|u| + 2) + 1, \dots, (|u| + 2) + |v|\}$. These sets form a partition of $[2k]$, which proves (i). To prove (ii), we distinguish on which of the four aforementioned sets a_i belongs to. Suppose $a_i \in \{2, \dots, |u| + 1\}$. It follows from (i) and (1) that a_i is the $(i - 1)$ th entry of the sequence $(|u| + 2) - \pi(\overleftarrow{u})$. Let $\pi(\overleftarrow{u}) = (b_1, \dots, b_{|u|})$. By the induction hypothesis, the bit of \overleftarrow{u} at position b_{i-1} is 0 if and only if i is even. The bit of x at position a_i is the bit of u at position $a_i - 1 = (|u| + 1) - b_{i-1}$, which is the complement of the bit of \overleftarrow{u} at position b_{i-1} . Therefore, the bit of x at position a_i is 1 if and only if i is even, as claimed in (ii). We leave the analysis of the remaining cases to the reader.

In terms of Dyck path representation, we can interpret $\pi(x)$ as the alternating order of down-steps and up-steps of the Dyck path x ; see Figure 3. The first term of $\pi(x)$ represents the first down-step that touches the abscissa—it goes from $(|u| + 1, 1)$ to $(|u| + 2, 0)$. The next part of $\pi(x)$ represents the up-steps and down-steps of the part u of the Dyck path between $(1, 1)$ to $(|u| + 1, 1)$ in the order obtained recursively on the mirror image of u . The next term of $\pi(x)$ represents the first up-step, which goes from $(0, 0)$ to $(1, 1)$. The final part of $\pi(x)$ represents the down-steps and up-steps of the part v of the Dyck path between $(|u| + 2, 0)$ to $(|x|, 0)$ ordered recursively.

Now, let $x \in D_k$ and $\pi(x) = (a_1, \dots, a_{2k})$. Using the properties (i) and (ii) above, we define a path $P(x) = (x_0, x_1, \dots, x_{2k})$ in the graph G_k so that $x_0 = x$ and x_i is obtained from x_{i-1} by flipping the bit at position a_i for every $i \in [2k]$, whence it follows that $x_{2k} = \bar{x}$. We call $\pi(x)$ the *bit-flip sequence* for $P(x)$. We define the set of paths \mathcal{P}_k by

$$\mathcal{P}_k := \{P(x) \mid x \in D_k\}.$$

The set of paths \mathcal{P}_3 with the corresponding bit-flip sequences is illustrated in Figure 4.

The following lemma is a consequence of the results of [MSW18]. For the reader's convenience, we provide a short direct proof of it in Section 4.

Lemma 5 ([MSW18]). *For every $k \geq 1$, the paths in \mathcal{P}_k are mutually vertex-disjoint, and together they cover all vertices of G_k .*

For every Dyck word $x \in D_k$ and every bit position $i \in [2k]$, we let $e(x, i)$ denote the edge of the path $P(x)$ along which the i th bit is flipped. That is, if $\pi(x) = (a_1, \dots, a_{2k})$, then the path $P(x)$ contains edges $e(x, a_1), \dots, e(x, a_{2k})$ in this order along the path from x to \bar{x} . For example, for x_1 as in Figure 4, we have $e(x_1, 3) = \{101101, 100101\}$ and $e(x_1, 1) = \{100111, 000111\}$.

For every Dyck word $x \in D_k$ with $k \geq 1$, the first vertex x and the last vertex \bar{x} of $P(x)$ are adjacent in G_k^+ . We let $C(x)$ denote the cycle in G_k^+ obtained by adding the edge $\{x, \bar{x}\}$ to the path $P(x)$. We define

$$\mathcal{C}_k := \{C(x) \mid x \in D_k\}.$$

It follows from Lemma 5 that the set of cycles \mathcal{C}_k is a cycle factor in G_k^+ . Figure 2 illustrates the cycles in \mathcal{C}_3 , which are obtained by closing the paths in \mathcal{P}_3 illustrated in Figure 4 and applying the isomorphism between G_k^+ and O_k described in the proof of Lemma 4.

3. CONSTRUCTION OF A HAMILTON CYCLE

We describe how to modify the cycle factor \mathcal{C}_k to join its cycles to a single Hamilton cycle. As indicated in Section 1.6, the modification operation consists in taking the symmetric difference

$x_1 = 111000$		$x_2 = 110100$		$x_3 = 110010$		$x_4 = 101100$		$x_5 = 101010$	
$P(x_1)$	$\pi(x_1)$	$P(x_2)$	$\pi(x_2)$	$P(x_3)$	$\pi(x_3)$	$P(x_4)$	$\pi(x_4)$	$P(x_5)$	$\pi(x_5)$
$\overline{111000}$	6	$\overline{110100}$	6	110010	4	101100	2	101010	2
$\overline{111001}$	2	$\overline{110101}$	4	$\overline{110110}$	2	111100	1	$\overline{111010}$	1
101001	4	$\overline{110001}$	5	$\overline{100110}$	3	011100	6	$\overline{011010}$	4
101101	3	$\overline{110011}$	2	101110	1	011101	4	011110	3
$\overline{100101}$	5	100011	3	001110	6	$\overline{011001}$	5	010110	6
$\overline{100111}$	1	101011	1	001111	5	$\overline{011011}$	3	010111	5
000111		001011		001101		$\overline{010011}$		010101	

FIGURE 4. The set of paths $\mathcal{P}_3 = \{P(x_1), \dots, P(x_5)\}$ in the graph G_3 together with the bit-flip sequences $\pi(x_1), \dots, \pi(x_5)$ that generate them. The edges on the three flipping cycles that witness the flippable tuples $\alpha(\epsilon)$, β , and δ defined in (4) are indicated by dashed, dotted, and solid frames, respectively.

with a carefully chosen set of cycles of length 6 or 8. The key ingredient of our argument is Lemma 6 below, which reduces the Hamiltonicity problem to a spanning tree problem in a suitably defined hypergraph. To make these ideas formal, we introduce a few definitions.

A *flipping cycle* on D_k is a cycle in G_k of length 2ℓ that has exactly ℓ edges in common with ℓ distinct paths in the set \mathcal{P}_k (one common edge with each path), and along the cycle these ℓ common edges alternate with ℓ edges that go between pairs of distinct paths. Recall that each path $P(x) \in \mathcal{P}_k$ in G_k extends to the cycle $C(x)$ in the graph G_k^+ by adding the edge $\{x, \bar{x}\}$, so a flipping cycle has exactly ℓ edges in common with ℓ distinct cycles in the set \mathcal{C}_k .

A *marked Dyck word* is a non-empty Dyck word in which exactly one bit has been *marked*. More formally, a *marked Dyck word* is a pair (x, m) with $x \in D_k$ and $m \in [2k]$ for some $k \geq 1$, where m is the position of the marked bit in x . We simplify notation of marked Dyck words by underlining the marked bit. For instance, $1011\underline{0}0$ denotes the marked Dyck word $(101100, 5)$. We define prepending to, appending to, and mirroring a marked Dyck word (x, m) in a natural way, as follows:

$$\begin{aligned} u(x, m)v &:= (uxv, |u| + m) \quad \text{for any bitstrings } u \text{ and } v \text{ such that } uv \in D, \\ \overleftarrow{(x, m)} &:= (\overline{x}, |x| + 1 - m). \end{aligned} \tag{2}$$

For instance, if $(x, m) = 1011\underline{0}0$, then $1(x, m)010 = 11011\underline{0}0010$ and $\overleftarrow{(x, m)} = 1\underline{1}0010$. In terms of Dyck path representation, (x, m) is a Dyck path where the m th step is marked; see Figure 5. Under the operations of prepending, appending, and mirroring, the marked step remains at the same relative position.

A *marked ℓ -tuple* on a set of Dyck words $X \subseteq D_k$ is an unordered ℓ -tuple of marked Dyck words of the form $\tau = \{(x_1, m_1), \dots, (x_\ell, m_\ell)\}$, where x_1, \dots, x_ℓ are distinct Dyck words in X , $m_1, \dots, m_\ell \in [2k]$, and $\ell \geq 3$. The set $\{x_1, \dots, x_\ell\} \subseteq X$ is called the *support* of such a marked ℓ -tuple τ and it is denoted by $\text{supp } \tau$. The index m_i is called the *mark* of x_i in τ . A marked ℓ -tuple $\tau = \{(x_1, m_1), \dots, (x_\ell, m_\ell)\}$ on X is called a *flippable ℓ -tuple* on X if there is a flipping 2ℓ -cycle in G_k that contains exactly the edges $e(x_1, m_1), \dots, e(x_\ell, m_\ell)$ of the paths $P(x_1), \dots, P(x_\ell)$, respectively. We say that such a flipping cycle *witnesses* the flippable tuple τ .

To get an intuition for these definitions, consider the Dyck words x_1 , x_2 , and x_3 in the first three columns in Figure 4. Then $\tau = \{1110\underline{0}0, 1101\underline{0}0, 1\underline{1}0010\}$ is a marked triple on D_3 with support $\text{supp } \tau = \{x_1, x_2, x_3\}$. In fact, it is a flippable triple on D_3 witnessed by a flipping

6-cycle $W = (100101, 100111, 100110, 110110, 110100, 110101)$ that contains the edges $e(x_1, 5)$ of $P(x_1)$, $e(x_2, 6)$ of $P(x_2)$, and $e(x_3, 2)$ of $P(x_3)$, indicated in Figure 4 by dashed frames. By taking the symmetric difference with W , the cycles $C(x_1)$, $C(x_2)$, and $C(x_3)$ become joined into a single cycle. This observation motivates the definitions that follow.

Let $X \subseteq D_k$ (where $k \geq 2$), let \mathcal{X} be a set of flippable tuples on X , and let $\mathcal{H} = (X, \mathcal{X})$. We call such a structure \mathcal{H} a *flippability hypergraph* on X , and we apply a few standard hypergraph-theoretic terms to \mathcal{H} (as follows), although the reader should realize that the members of \mathcal{X} convey the marks as extra information in addition to the standard hypergraph structure. Thus, the *subhypergraph* of \mathcal{H} induced by a non-empty set $U \subseteq X$ is defined as

$$\mathcal{H}[U] := (U, \{\tau \in \mathcal{X} \mid \text{supp } \tau \subseteq U\}).$$

A *spanning tree* of \mathcal{H} is a subset of \mathcal{X} defined as follows, by induction on $|X|$. If $|X| = 1$, then the only spanning tree of \mathcal{H} is the empty set. If $|X| \geq 2$, then a set $\mathcal{T} \subseteq \mathcal{X}$ is a spanning tree of \mathcal{H} if and only if there are a flippable ℓ -tuple $\tau \in \mathcal{T}$, a partition of X into non-empty subsets X_1, \dots, X_ℓ , and spanning trees $\mathcal{T}_1, \dots, \mathcal{T}_\ell$ of $\mathcal{H}[X_1], \dots, \mathcal{H}[X_\ell]$ (respectively) such that $\mathcal{T} = \{\tau\} \cup \mathcal{T}_1 \cup \dots \cup \mathcal{T}_\ell$ and $|\text{supp } \tau \cap X_i| = 1$ for each $i \in [\ell]$. For instance, a one-element set $\mathcal{T} = \{\tau\} \subseteq \mathcal{X}$ is a spanning tree of \mathcal{H} if and only if $\text{supp } \tau = X$, and a two-element set $\mathcal{T} = \{\tau_1, \tau_2\} \subseteq \mathcal{X}$ is a spanning tree of \mathcal{H} if and only if $\text{supp } \tau_1 \cup \text{supp } \tau_2 = X$ and $|\text{supp } \tau_1 \cap \text{supp } \tau_2| = 1$. In general, straightforward induction shows that the supports of any two flippable tuples in a spanning tree have at most one element of X in common. A *conflict-free set* in \mathcal{H} is a subset \mathcal{U} of \mathcal{X} such that the following two conditions are satisfied:

- (i) the supports of any two flippable tuples in \mathcal{U} have at most one element of X in common;
- (ii) for any two distinct flippable tuples $\tau_1, \tau_2 \in \mathcal{U}$ whose supports have a common element $x \in X$, the mark of x in τ_1 is different from the mark of x in τ_2 .

As it is mentioned above, every spanning tree satisfies condition (i), and thus a *conflict-free spanning tree* is a spanning tree that additionally satisfies condition (ii).

The following lemma is the cornerstone behind our proofs of Theorems 1 and 3. It reduces the problem of finding a Hamilton cycle in the graph G_k^+ (which is isomorphic to the odd graph O_k) to the problem of finding a conflict-free spanning tree in a flippability hypergraph on D_k .

Lemma 6. *Let \mathcal{H} be a flippability hypergraph on D_k , where $k \geq 3$. If \mathcal{H} has a conflict-free spanning tree, then the graph G_k^+ has a Hamilton cycle. Moreover, distinct conflict-free spanning trees of \mathcal{H} give rise to distinct Hamilton cycles in G_k^+ .*

Proof. For every flippable tuple τ on D_k , fix a flipping cycle $W(\tau)$ in G_k^+ that witnesses τ . For a non-empty set $X \subseteq D_k$, let $G_k^+[X]$ denote the subgraph of G_k^+ induced by the set of all vertices of the cycles $C(x)$ with $x \in X$. For a non-empty set $X \subseteq D_k$ and a conflict-free set \mathcal{U} in $\mathcal{H}[X]$, let $S(X, \mathcal{U})$ denote the symmetric difference of the edge sets of the cycles $C(x)$ with $x \in X$ and the cycles $W(\tau)$ with $\tau \in \mathcal{U}$. We prove the following statement, which immediately yields the lemma:

Claim. Let X be a non-empty subset of D_k . For every conflict-free spanning tree \mathcal{T} in $\mathcal{H}[X]$, the set $S(X, \mathcal{T})$ forms a Hamilton cycle in $G_k^+[X]$. Moreover, if \mathcal{T} is a conflict-free spanning tree in $\mathcal{H}[X]$ and \mathcal{U} is a conflict-free set in $\mathcal{H}[X]$ such that $S(X, \mathcal{T}) = S(X, \mathcal{U})$, then $\mathcal{T} = \mathcal{U}$.

The proof of the claim goes by induction on $|X|$. If $|X| = 1$, then the empty set is the unique conflict-free spanning tree in $\mathcal{H}[X]$, and $S(X, \emptyset) = C(x)$ for the unique $x \in X$. For the rest of the proof, suppose that $|X| \geq 2$ and that the claim holds for all subsets of D_k smaller than X .

Let \mathcal{T} be a conflict-free spanning tree in $\mathcal{H}[X]$. By the definition of a spanning tree, there are a flippable ℓ -tuple $\tau \in \mathcal{T}$, a partition of X into non-empty subsets X_1, \dots, X_ℓ , and spanning

trees $\mathcal{T}_1, \dots, \mathcal{T}_\ell$ of $\mathcal{H}[X_1], \dots, \mathcal{H}[X_\ell]$ (respectively) such that $\mathcal{T} = \{\tau\} \cup \mathcal{T}_1 \cup \dots \cup \mathcal{T}_\ell$ and $|\text{supp } \tau \cap X_i| = 1$ for each $i \in [\ell]$. Since $\mathcal{T}_1, \dots, \mathcal{T}_\ell \subset \mathcal{T}$, the spanning trees $\mathcal{T}_1, \dots, \mathcal{T}_\ell$ are conflict-free. Therefore, by the induction hypothesis, the sets $S(X_1, \mathcal{T}_1), \dots, S(X_\ell, \mathcal{T}_\ell)$ form Hamiltonian cycles in $G_k^+[X_1], \dots, G_k^+[X_\ell]$. Suppose $\tau = \{(x_1, m_1), \dots, (x_\ell, m_\ell)\}$, where $x_i \in X_i$ and $m_i \in [2k]$ for $i \in [\ell]$. The unique common edge of $W(\tau)$ with $G_k^+[X_i]$ is the edge $e(x_i, m_i)$, which belongs to $S(X_i, \mathcal{T}_i)$, as \mathcal{T} is conflict-free, for $i \in [\ell]$. The set $S(X, \mathcal{T})$ is the symmetric difference of $S(X_1, \mathcal{T}_1), \dots, S(X_\ell, \mathcal{T}_\ell)$ and $W(\tau)$, which is therefore a single cycle—a Hamiltonian cycle in $G_k^+[X]$.

Now, suppose that $\mathcal{H}[X]$ has another conflict-free set \mathcal{U} such that $S(X, \mathcal{T}) = S(X, \mathcal{U}) =: S$. For each $i \in [\ell]$, since the only edges in S that connect X_i with $X \setminus X_i$ are those that precede and follow $e(x_i, m_i)$ on $W(\tau)$, these two edges along with $e(x_i, m_i)$ belong to the same cycle witnessing some flippable tuple in \mathcal{U} . It follows that the entire flipping cycle $W(\tau)$ witnesses one of the flippable tuples in \mathcal{U} , which implies $\tau \in \mathcal{U}$. The symmetric difference of S and $W(\tau)$ is the disjoint union of $S(X_1, \mathcal{T}_1), \dots, S(X_\ell, \mathcal{T}_\ell)$. This implies that every flippable tuple in $\mathcal{U} \setminus \{\tau\}$ is a flippable tuple on one of X_1, \dots, X_ℓ . Therefore, we have $\mathcal{U} = \{\tau\} \cup \mathcal{U}_1 \cup \dots \cup \mathcal{U}_\ell$, where \mathcal{U}_i is a conflict-free set in $\mathcal{H}[X_i]$ such that $S(X_i, \mathcal{T}_i) = S(X_i, \mathcal{U}_i)$ for each $i \in [\ell]$. This and the induction hypothesis yield $\mathcal{T}_i = \mathcal{U}_i$ for each $i \in [\ell]$ and therefore $\mathcal{T} = \mathcal{U}$. \square

To apply Lemma 6, we need to define a flippability hypergraph on D_k that admits a conflict-free spanning tree. In other words, we need to construct a sufficiently large set of flippable tuples. Our construction works inductively and is based on the next lemma, which allows us to generate more flippable tuples from existing ones by prepending and appending certain bitstrings to them. We introduce the following auxiliary notation for every flippable tuple $\tau = \{(x_1, m_1), \dots, (x_\ell, m_\ell)\}$ on D_k :

$$\begin{aligned} u\tau v &:= \{u(x_1, m_1)v, \dots, u(x_\ell, m_\ell)v\} = \{(ux_1v, |u| + m_1), \dots, (ux_\ell v, |u| + m_\ell)\} \\ &\quad \text{for any bitstrings } u \text{ and } v \text{ such that } uv \in D, \quad (3) \\ \overleftarrow{\tau} &:= \{\overleftarrow{(x_1, m_1)}, \dots, \overleftarrow{(x_\ell, m_\ell)}\} = \{(\overleftarrow{x_1}, 2k + 1 - m_1), \dots, (\overleftarrow{x_\ell}, 2k + 1 - m_\ell)\}. \end{aligned}$$

Lemma 7. *If τ is a flippable tuple, then*

- (i) $u\tau v$ is a flippable tuple for any bitstrings u and v such that $uv \in D$ and $|u|$ is even,
- (ii) $u\overleftarrow{\tau}v$ is a flippable tuple for any bitstrings u and v such that $uv \in D$ and $|u|$ is odd.

The proofs of Lemma 7 and of all subsequent lemmas stated in this section are deferred to Sections 4 and 5.

We now specify the base case for our inductive construction of flippable tuples. We found the following tuples on D_3 and D_4 with the help of a computer. In fact, the computer search gave many more flippable tuples, and we carefully selected a subset that can be used to create a conflict-free spanning tree on D_k . The flippable tuples in this basic set Φ of flippable tuples are called *patterns*. We let $\Phi := \{\alpha(w) \mid w \in D\} \cup \{\beta, \gamma, \delta\}$, where

$$\begin{aligned} \alpha(w) &:= \{1w11000, 1w10100, 1w10010\}, & \gamma &:= \{11001100, 11011000, 11101000\}, \\ \beta &:= \{111000, 101100, 101010\}, & \delta &:= \{111000, 110100, 101100, 101010\}. \end{aligned} \quad (4)$$

The Dyck path representation of these tuples is shown in Figure 5. The next lemma asserts that these definitions indeed yield flippable tuples.

Lemma 8. *Every pattern in Φ defined by (4) is a flippable tuple.*

Figure 4 shows three flipping cycles that witness the patterns $\alpha(\epsilon)$, β , and δ .

We use Lemmas 7 and 8 to construct a set Ψ of flippable tuples. Namely, we define

$$\Psi := \{u\varphi v \mid \varphi \in \Phi, uv \in D, \text{ and } |u| \text{ is even}\} \cup \{u\overleftarrow{\varphi}v \mid \varphi \in \Phi, uv \in D, \text{ and } |u| \text{ is odd}\}. \quad (5)$$

By Lemmas 7 and 8, every marked tuple in Ψ is flippable. Observe that the set of flippable tuples Ψ is already closed with respect to the operation described in Lemma 7, that is,

$$\Psi = \{u\tau v \mid \tau \in \Psi, uv \in D, \text{ and } |u| \text{ is even}\} \cup \{u\overleftarrow{\tau}v \mid \tau \in \Psi, uv \in D, \text{ and } |u| \text{ is odd}\}. \quad (6)$$

Next, for each $k \geq 2$, we define a set Ψ_k by extracting only the flippable tuples on D_k from Ψ :

$$\Psi_2 := \emptyset, \quad \Psi_k := \{\tau \in \Psi \mid \tau \text{ is a flippable tuple on } D_k\} \quad \text{for } k \geq 3. \quad (7)$$

Finally, we define a flippability hypergraph $\mathcal{H}_k := (D_k, \Psi_k)$ for $k \geq 2$.

Lemma 9. *For every $k \geq 3$, the set Ψ_k of flippable tuples defined by (7) has the property that for any $\tau_1, \tau_2 \in \Psi_k$, if $\text{supp } \tau_1 \cap \text{supp } \tau_2 = \{x\}$ where $x \in D_k$, then the mark of x in τ_1 is different from the mark of x in τ_2 . In particular, every spanning tree of the hypergraph \mathcal{H}_k is conflict-free.*

In view of Lemmas 6 and 9, it remains to prove that the hypergraph \mathcal{H}_k has a spanning tree (many distinct spanning trees) to complete the proofs of Theorems 1 and 3.

Lemma 10. *For every $k \geq 3$, the hypergraph \mathcal{H}_k has a spanning tree.*

Lemma 11. *For every $k \geq 6$, the hypergraph \mathcal{H}_k has at least $2^{2^{k-6}}$ distinct spanning trees.*

Proof of Theorem 1. Combine Lemma 6, Lemma 9, and Lemma 10. □

Proof of Theorem 2. Combine Theorem 1 and [Joh11, Theorem 1]. □

Proof of Theorem 3. Combine Lemma 6, Lemma 9, and Lemma 11. □

4. PROOFS OF LEMMAS 5, 7, 8, AND 9

Proof of Lemma 5. We extend the mirroring notation to arbitrary bitstrings: for a bitstring $w = b_1 b_2 \cdots b_{2k}$, where $b_1, \dots, b_{2k} \in \{0, 1\}$, we define $\overleftarrow{w} := \overline{b_{2k} b_{2k-1} \cdots b_1}$.

We have $P(\epsilon) = (\epsilon)$ by definition. When $k \geq 1$ and $x \in D_k$, the path $P(x) = (x_0, x_1, \dots, x_{2k})$ can be described recursively as follows. Let $x = 1u0v$ be the unique decomposition of x with $u, v \in D$ that is used in (1). Let $\ell = \frac{1}{2}|u| + 1$, so that $u \in D_{\ell-1}$ and $v \in D_{k-\ell}$. Then we have

$$\begin{aligned} x_i &= 1u0v && \text{for } i = 0, \\ x_i &= 1u_{i-1}1v && \text{for } i \in \{1, \dots, 2\ell - 1\}, \quad \text{where } P(\overleftarrow{u}) = (\overleftarrow{u}_0, \overleftarrow{u}_1, \dots, \overleftarrow{u}_{2\ell-2}), \\ x_i &= 0\overline{u}1v_{i-2\ell} && \text{for } i \in \{2\ell, \dots, 2k\}, \quad \text{where } P(v) = (v_0, v_1, \dots, v_{2k-2\ell}). \end{aligned} \quad (8)$$

Conversely, for every bitstring $y \in B_k$ with $k \geq 1$, exactly one of the following three cases holds:

- (i) y has a unique decomposition $y = 1u0v$ with $u \in D_{\ell-1}$ and $v \in D_{k-\ell}$ where $\ell \in [k]$;
- (ii) y has a unique decomposition $y = 1w1v$ with $\overleftarrow{w} \in B_{\ell-1}$ and $v \in D_{k-\ell}$ where $\ell \in [k]$;
- (iii) y has a unique decomposition $y = 0\overline{u}1w$ with $u \in D_{\ell-1}$ and $w \in B_{k-\ell}$ where $\ell \in [k]$.

We prove that for every bitstring $y \in B_k$ with $k \geq 0$, there is exactly one pair (x, i) such that $x \in D_k$, $i \in \{0, \dots, 2k\}$, and $y = x_i$ given that $P(x) = (x_0, x_1, \dots, x_{2k})$. The proof goes by induction on k . The base case $k = 0$ is trivial. For the induction step, let $y \in B_k$ with $k \geq 1$, and suppose the statement holds for all bitstrings in $B_0 \cup \cdots \cup B_{k-1}$.

We consider the three cases of the decomposition of y described in (i)–(iii). In case (i), it follows from (8) that $(1u0v, 0)$ is the unique pair (x, i) with the required properties. In case (ii), we apply the induction hypothesis to \overleftarrow{w} to get a unique pair (u, j) such that $u \in D_{\ell-1}$, $j \in \{0, \dots, 2\ell - 2\}$, and $\overleftarrow{w} = \overleftarrow{u}_j$ (that is, $w = u_j$) given that $P(\overleftarrow{u}) = (\overleftarrow{u}_0, \overleftarrow{u}_1, \dots, \overleftarrow{u}_{2\ell-2})$; it follows from (8) that $(1u0v, j + 1)$ is the unique pair (x, i) with the required properties. Finally, in case (iii), we apply the induction hypothesis to w to get a unique pair (v, j) such that $v \in D_{k-\ell}$,

$j \in \{0, \dots, 2k - 2\ell\}$, and $w = v_j$ given that $P(v) = (v_0, v_1, \dots, v_{2k-2\ell})$; it follows from (8) that $(1u0v, 2\ell + j)$ is the unique pair (x, i) with the required properties. \square

Proof of Lemma 7. We will prove the following three special cases of the statements claimed in the lemma: if τ is a flippable tuple, then

- (iii) $u\tau$ is a flippable tuple for every $u \in D$,
- (iv) τv is a flippable tuple for every $v \in D$,
- (v) $1\overleftarrow{\tau}0$ is a flippable tuple.

Statements (i) and (ii) then follow by straightforward induction, because the operations described therein can be obtained by repeated application of the operations described in (iii)–(v).

We will need the following simple observation.

Claim. For any $x, y \in D$, we have $\pi(xy) = (\pi(x), |x| + \pi(y))$.

The claim is proved by induction on $|x|$. If $x = \epsilon$, then the claim holds trivially. Otherwise, let $x = 1u0v$ be the unique decomposition of x such that $u, v \in D$. By (1) and by the induction hypothesis applied to vy , we have

$$\begin{aligned} \pi(1u0vy) &= (|u| + 2, (|u| + 2) - \pi(\overleftarrow{u}), 1, (|u| + 2) + \pi(vy)) \\ &= (|u| + 2, (|u| + 2) - \pi(\overleftarrow{u}), 1, (|u| + 2) + \pi(v), (|u| + 2 + |v|) + \pi(y)) \\ &= (\pi(1u0v), (|u| + 2 + |v|) + \pi(y)), \end{aligned}$$

which proves the claim.

Let $\tau = \{(x_1, m_1), \dots, (x_\ell, m_\ell)\}$ be a flippable tuple in D_k , where $k \geq 1$, and let $C = (y_1, \dots, y_{2\ell})$ be a flipping cycle of length 2ℓ that witnesses τ .

For the proof of (iii), let $u \in D$. The claim implies that for each $i \in [\ell]$, the final part of the path $P(ux_i)$ looks as follows:

$$P(ux_i) = (\dots, \overline{u}z_0, \overline{u}z_1, \dots, \overline{u}z_{2k}), \quad \text{where } P(x_i) = (z_0, z_1, \dots, z_{2k}).$$

Therefore, $(\overline{u}y_1, \dots, \overline{u}y_{2\ell})$ is a flipping cycle that witnesses $u\tau$. This proves (iii).

For the proof of (iv), let $v \in D$. The claim implies that for each $i \in [\ell]$, the initial part of the path $P(x_iv)$ looks as follows:

$$P(x_iv) = (z_0v, z_1v, \dots, z_{2k}v, \dots), \quad \text{where } P(x_i) = (z_0, z_1, \dots, z_{2k}).$$

Therefore, $(y_1v, \dots, y_{2\ell}v)$ is a flipping cycle that witnesses τv . This proves (iv).

Finally, we prove (v). For each $i \in [\ell]$, by (1), we have $\pi(1x_i0) = (2k + 2, (2k + 2) - \pi(\overleftarrow{x_i}), 1)$, which implies that

$$P(1x_i0) = (1x_i0, 1z_01v, 1z_11v, \dots, 1z_{2k}1v, 0\overleftarrow{x_i}1), \quad \text{where } P(\overleftarrow{x_i}) = (\overleftarrow{z_0}, \overleftarrow{z_1}, \dots, \overleftarrow{z_{2k}}).$$

Therefore, $(1\overleftarrow{y_1}1, \dots, 1\overleftarrow{y_{2\ell}}1)$ is a flipping cycle that witnesses $1\overleftarrow{\tau}0$. This proves (v). \square

Proof of Lemma 8. Consider the sequences

$$\begin{aligned} C_{\alpha(w)} &:= (1w00101, 1w00111, 1w00110, 1w10110, 1w10100, 1w10101) \quad \text{for } w \in D, \\ C_\beta &:= (111000, 111001, 011001, 011011, 011010, 111010), \\ C_\gamma &:= (11011100, 10011100, 10011101, 10011001, 11011001, 11011000), \\ C_\delta &:= (111000, 111001, 110001, 110011, 010011, 011011, 011010, 111010). \end{aligned} \tag{9}$$

It is easy to verify that each of these sequences is a cycle in G_k for the appropriate value of k .

We claim that for each pattern $\varphi \in \Phi$, the cycle C_φ from (9) is a flipping cycle that witnesses φ . For $\varphi \in \{\beta, \delta\}$ this can be verified directly from Figure 4, as follows. The bitstrings on the

cycles C_β and C_δ are indicated in the figure by dotted and solid frames, respectively. Both cycles have exactly one edge in common with each of the paths that start at the respective vertices in the tuples β and δ . Furthermore, the bits flipped along the indicated edges are precisely those that are marked in β and δ .

In the same way, the claim can be verified for $\varphi = \alpha(w)$ when $w = \epsilon$. The bitstrings on the cycle $C_{\alpha(\epsilon)}$ are indicated in Figure 4 by dashed frames. For general $w \in D$, by the definition (1), the initial parts of the paths in $\mathcal{P}_{3+|w|/2}$ that start at the members of $\alpha(w)$ look as follows:

$$\begin{aligned} P(1w11000) &= (1w11000, 1w11001, 1w01001, 1w01101, 1w00101, 1w00111, \dots), \\ P(1w10100) &= (1w10100, 1w10101, \dots), \\ P(1w10010) &= (1w10010, 1w10110, 1w00110, \dots). \end{aligned} \tag{10}$$

For $w = \epsilon$, these paths are exactly the same as $P(x_1)$, $P(x_2)$, and $P(x_3)$ in Figure 4. The 6-cycle $C_{\alpha(w)}$ defined by (9) intersects every path from (10) exactly at the last edge explicitly shown in (10). Furthermore, the bits flipped along the intersection edges are exactly the marked bits of the members of $\alpha(w)$ as defined by (4).

Finally, we consider the case $\varphi = \gamma$. The initial parts of the paths in \mathcal{P}_4 that start at the members of γ look as follows:

$$\begin{aligned} P(11001100) &= (11001100, 11011100, 10011100, \dots), \\ P(11011000) &= (11011000, 11011001, \dots), \\ P(11101000) &= (11101000, 11101001, 10101001, 10111001, 10011001, 10011101, \dots). \end{aligned} \tag{11}$$

The cycle C_γ defined by (9) intersects every path from (11) exactly at the last edge explicitly shown in (11), and the bits flipped along the intersection edges are exactly the marked bits of the members of γ as defined by (4). \square

Proof of Lemma 9. Let $\overleftarrow{\Phi} := \{\overleftarrow{\varphi} \mid \varphi \in \Phi\}$. Let τ_1 and τ_2 be flippable tuples in Ψ_k such that $\text{supp } \tau_1 \cap \text{supp } \tau_2 = \{x\}$, where $x \in D_k$. Let $i \in \{1, 2\}$. By (5), we have $\tau_i = u_i \varphi_i v_i$, where u_i and v_i are some bitstrings such that $u_i v_i \in D$ and

$$\varphi_i \in \Phi \quad \text{if } |u_i| \text{ is even,} \quad \varphi_i \in \overleftarrow{\Phi} \quad \text{if } |u_i| \text{ is odd.} \tag{12}$$

Since $x \in \text{supp } \tau_i$, we have $x = u_i x_i v_i$ for some Dyck word $x_i \in \text{supp } \varphi_i$. Let m_i be the mark of x_i in φ_i . It follows that $|u_i| + m_i$ is the mark of x in τ_i .

Suppose for the sake of contradiction that the mark of x in τ_1 is the same as the mark of x in τ_2 , that is, $|u_1| + m_1 = |u_2| + m_2 =: m$. Let $p := \max(1 - m_1, 1 - m_2)$ and $q := \min(|x_1| - m_1, |x_2| - m_2)$. Thus we have $p \leq 0 \leq q$. The fact that $u_1 x_1 v_1 = u_2 x_2 v_2$ and $|u_1| + m_1 = |u_2| + m_2$ implies that for each $i \in \{p, p + 1, \dots, q\}$, the $(m_1 + i)$ th bit of x_1 is equal to the $(m_2 + i)$ th bit of x_2 . We claim that this is possible only when $(x_1, m_1) = (x_2, m_2)$. The proof of this claim involves consideration of all possible cases of φ_1 and φ_2 satisfying (12) and all possible cases of marked Dyck words $(x_1, m_1) \in \varphi_1$ and $(x_2, m_2) \in \varphi_2$. To avoid tedious case distinctions, we propose a visual argument using the Dyck path representation of x . Figure 5 presents the Dyck paths of the members of the patterns in Φ in which the steps representing the marked bits have been marked white or black according to the following *marking rule*: steps at odd positions are marked white, and steps at even positions are marked black.

Consider the Dyck path x in which the m th step has been marked white or black according to this rule. Let $i \in \{1, 2\}$. The common part of such a marked Dyck path x with the infinite vertical strip $[|u_i|, |u_i x_i|] \times \mathbb{R}$ is a translated copy of the Dyck path x_i in which the m_i th step has been marked according to the marking rule above. That is, it has the form of one of the

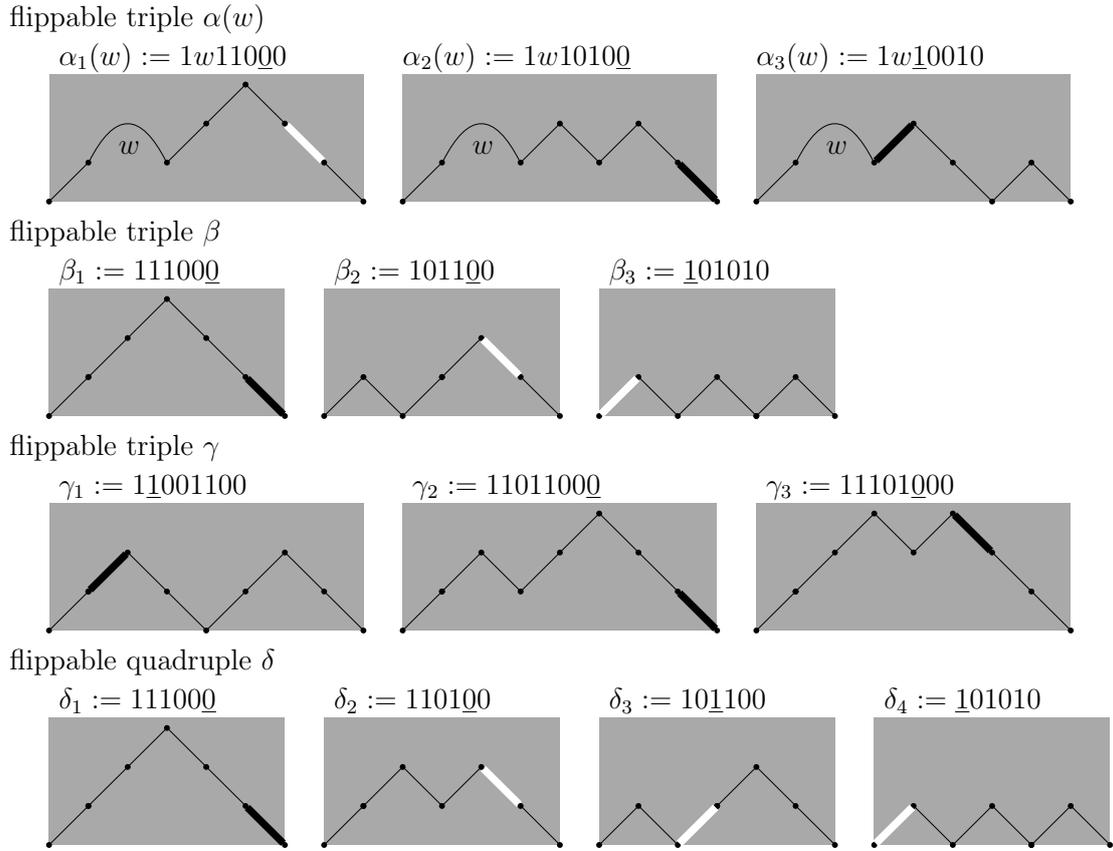


FIGURE 5. Dyck path representation of the patterns Φ defined by (4). The steps of the Dyck paths that represent the marked bits are highlighted white at odd positions and black at even positions.

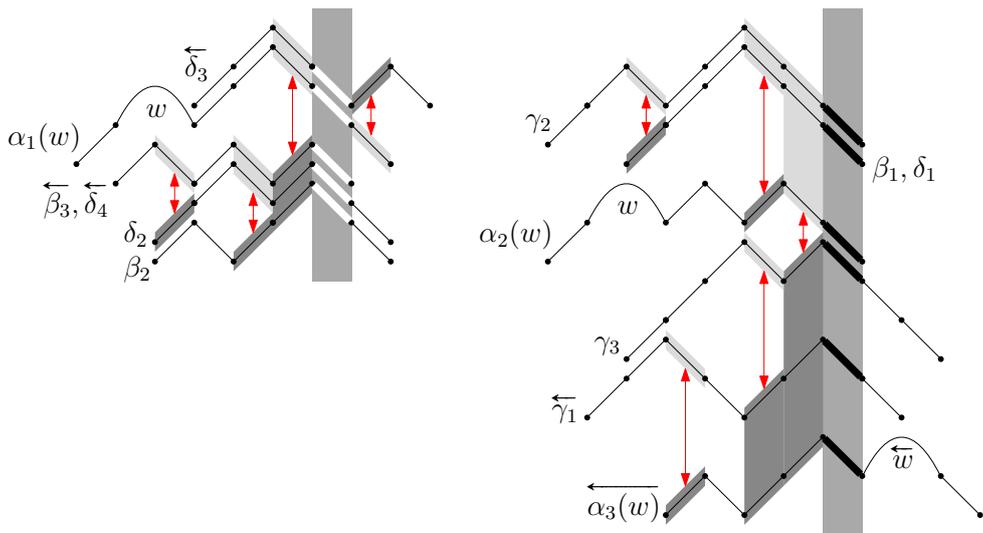


FIGURE 6. Illustration for the proof of Lemma 9. The double arrows indicate mismatches between the Dyck paths aligned at the marked down-steps.

marked Dyck paths in Figure 5 if $|u_i|$ is even or the form of the mirror image of one of the marked Dyck paths in Figure 5 if $|u_i|$ is odd. Note that the mirroring does *not* change the mark colors—even though the parity of the relative position of the mark from the left within the Dyck path changes, this is compensated by the opposite parity of $|u_i|$. The two translated marked Dyck paths x_1 and x_2 must coincide on the common part of the two vertical strips.

Figure 6 presents every Dyck path from Figure 5 or its mirror image, so that the marked step is a down-step. It shows them aligned horizontally with respect to the marked down-steps, separately for each color. It can be checked in the figure that no two of these marked Dyck paths can coincide on the common part of the two vertical strips considered in the argument above, unless they are the same marked Dyck path. Specifically, a mismatch between any two distinct Dyck paths is indicated by a double arrow in the figure. The situation when the marked step is an up-step is analogous, by symmetry.

We have argued that $(x_1, m_1) = (x_2, m_2)$. This is possible only when $\{\varphi_1, \varphi_2\} = \{\beta, \delta\}$, $\{\varphi_1, \varphi_2\} = \{\overleftarrow{\beta}, \overleftarrow{\delta}\}$, or otherwise $\varphi_1 = \varphi_2$. In any case, we have $|\text{supp } \varphi_1 \cap \text{supp } \varphi_2| \geq 3$. The assumption that $|u_1| + m_1 = |u_2| + m_2$ implies $|u_1| = |u_2|$, which implies $u_1 = u_2$ and $v_1 = v_2$. Therefore, for each $y \in \text{supp } \varphi_1 \cap \text{supp } \varphi_2$, we have $u_1 y v_1 = u_2 y v_2 \in \text{supp } u_1 \varphi_1 v_1 \cap \text{supp } u_2 \varphi_2 v_2 = \text{supp } \tau_1 \cap \text{supp } \tau_2$. It follows that $|\text{supp } \tau_1 \cap \text{supp } \tau_2| \geq 3$, which contradicts the assumption that $\text{supp } \tau_1 \cap \text{supp } \tau_2 = \{x\}$.

The second statement of Lemma 9 is an immediate consequence of the first statement and the definition of a conflict-free spanning tree. \square

5. PROOFS OF LEMMAS 10 AND 11

Before proceeding with the proofs, we generalize the notation (2) and (3). For every $k \geq 2$ and every set $X \subseteq D_k$, we define

$$\begin{aligned} uXv &:= \{uxv \mid x \in X\} \quad \text{for any bitstrings } u \text{ and } v \text{ such that } uv \in D, \\ \overleftarrow{X} &:= \{\overleftarrow{x} \mid x \in X\}. \end{aligned}$$

Similarly, for every $k \geq 2$ and every set \mathcal{X} of flippable tuples on D_k , we define

$$\begin{aligned} u\mathcal{X}v &:= \{u\tau v \mid \tau \in \mathcal{X}\} \quad \text{for any bitstrings } u \text{ and } v \text{ such that } uv \in D, \\ \overleftarrow{\mathcal{X}} &:= \{\overleftarrow{\tau} \mid \tau \in \mathcal{X}\}. \end{aligned}$$

As a direct consequence of the definitions above and (6), if $X \subseteq D_k$ and \mathcal{X} is a spanning tree of $\mathcal{H}_k[X]$, then the following holds for any bitstrings u and v such that $uv \in D$:

- (i) if $|u|$ is even, then $u\mathcal{X}v$ is a spanning tree of $\mathcal{H}_{k+|uv|/2}[uXv]$;
- (ii) if $|u|$ is odd, then $u\overleftarrow{\mathcal{X}}v$ is a spanning tree of $\mathcal{H}_{k+|uv|/2}[u\overleftarrow{X}v]$.

We will use this property extensively in the proofs below.

Proof of Lemma 10. For the reader's convenience, this proof is illustrated in Figure 7.

For $k \geq 2$, we partition the set of Dyck words D_k into two sets E_k and F_k as follows:

$$\begin{aligned} E_2 &:= \{1010\}, & E_3 &:= D_3 \setminus \{110010\}, & E_k &:= 10D_{k-1} & \text{for } k \geq 4, \\ F_2 &:= \{1100\}, & F_3 &:= \{110010\}, & F_k &:= D_k \setminus 10D_{k-1} & \text{for } k \geq 4. \end{aligned} \tag{13}$$

In particular, we have the following, for $k \geq 2$:

$$1010(10)^{k-2} \in E_k, \quad 1100(10)^{k-2} \in F_k. \tag{14}$$

We prove the following more general statement, which directly implies the lemma:

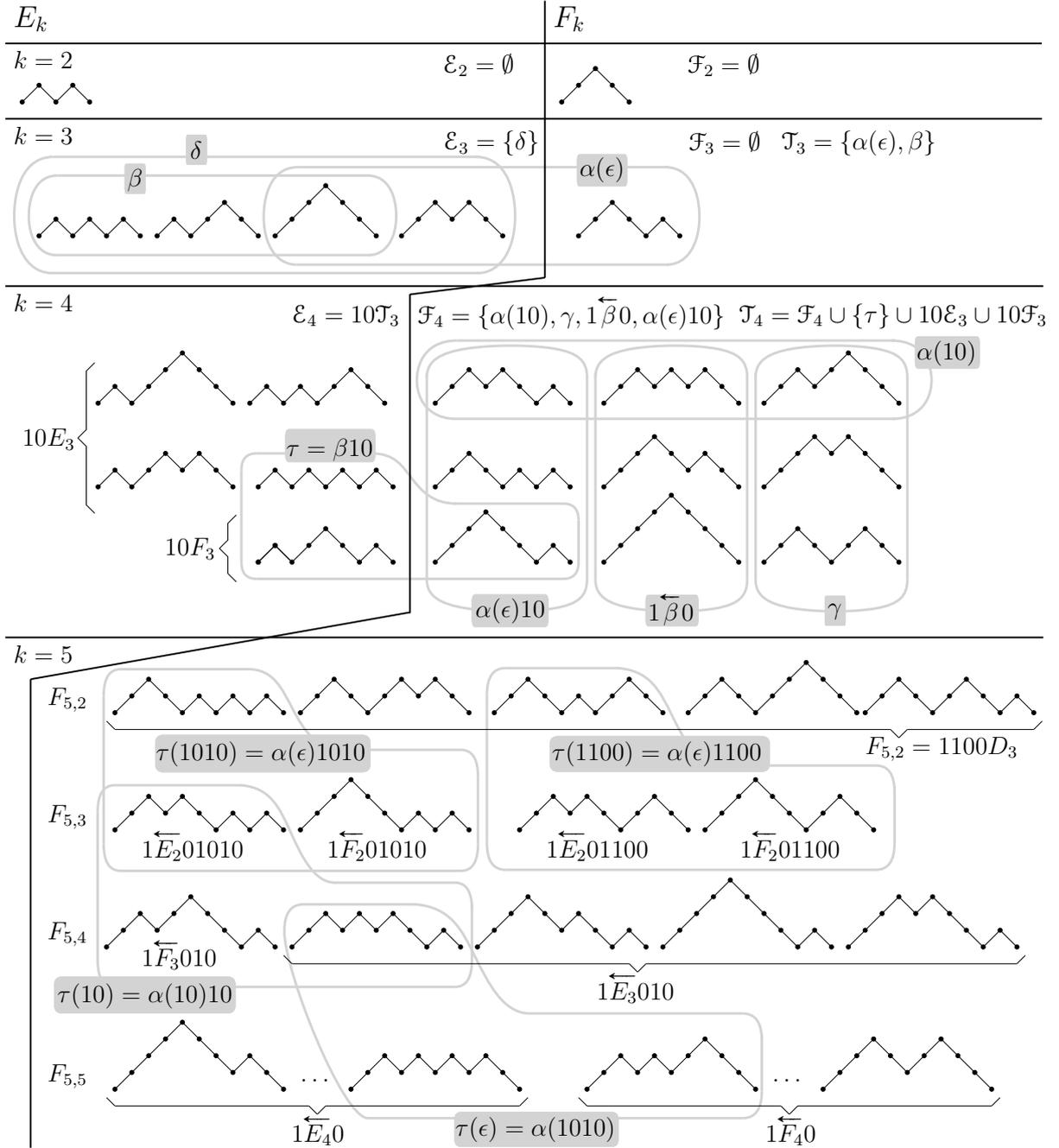


FIGURE 7. Illustration of the proof of Lemma 10. The inductive construction of \mathcal{T}_k is shown for $k = 4$, and the inductive construction of \mathcal{F}_k is shown for $k = 5$.

Claim. There are a spanning tree \mathcal{T}_k of \mathcal{H}_k for $k \geq 3$, a spanning tree \mathcal{E}_k of $\mathcal{H}_k[E_k]$ for $k \geq 2$, and a spanning tree \mathcal{F}_k of $\mathcal{H}_k[F_k]$ for $k \geq 2$.

We prove the claim by induction on k .

For $k = 2$, we let $\mathcal{E}_2 := \emptyset$ and $\mathcal{F}_2 := \emptyset$, which trivially satisfy the conditions for a spanning tree, as $|E_2| = |F_2| = 1$.

For $k = 3$, we let $\mathcal{E}_3 := \{\delta\}$ and $\mathcal{F}_3 := \emptyset$, which satisfy the conditions for a spanning tree, as $\text{supp } \delta = E_3$ and $|F_3| = 1$. We also let $\mathcal{T}_3 := \{\alpha(\epsilon), \beta\}$, which satisfies the conditions for a spanning tree, because $\text{supp } \alpha(\epsilon) \cup \text{supp } \beta = D_3$ and $|\text{supp } \alpha(\epsilon) \cap \text{supp } \beta| = 1$.

For $k = 4$, we let $\mathcal{F}_4 := \{\tau_1, \tau_2, \tau_3, \tau_4\}$, where

$$\begin{aligned}\tau_1 &:= \{11011000, 11010100, 11010010\} = \alpha(10), \\ \tau_2 &:= \{11001100, 11011000, 11101000\} = \gamma, \\ \tau_3 &:= \{11110000, 11100100, 11010100\} = 1\overleftarrow{\beta}0, \\ \tau_4 &:= \{11100010, 11010010, 11001010\} = \alpha(\epsilon)10.\end{aligned}$$

They belong to Ψ_k by construction. Moreover, we have $\text{supp } \tau_1 \cup \text{supp } \tau_2 \cup \text{supp } \tau_3 \cup \text{supp } \tau_4 = F_4$, $|\text{supp } \tau_1 \cap \text{supp } \tau_i| = 1$ for every $i \in \{2, 3, 4\}$, and $\text{supp } \tau_i \cap \text{supp } \tau_j = \emptyset$ for any distinct $i, j \in \{2, 3, 4\}$. This shows that \mathcal{F}_4 is indeed a spanning tree of $\mathcal{H}_4[F_4]$.

Finally, we proceed by induction on k to construct \mathcal{T}_k and \mathcal{E}_k for $k \geq 4$, and \mathcal{F}_k for $k \geq 5$. The construction of \mathcal{E}_k makes use of \mathcal{T}_{k-1} . The construction of \mathcal{F}_k (for $k \geq 5$) makes use of $\mathcal{E}_2, \dots, \mathcal{E}_{k-1}, \mathcal{F}_2, \dots, \mathcal{F}_{k-1}$, and \mathcal{T}_{k-2} (this is why the case $k = 4$ needs to be considered separately, as \mathcal{H}_2 has no spanning tree). Finally, the construction of \mathcal{T}_k makes use of $\mathcal{F}_k, \mathcal{E}_{k-1}$, and \mathcal{F}_{k-1} .

Constructing \mathcal{E}_k from \mathcal{T}_{k-1} for $k \geq 4$ is straightforward: since $E_k = 10D_{k-1}$, it suffices to take $\mathcal{E}_k := 10\mathcal{T}_{k-1}$, which is a spanning tree of $\mathcal{H}_k[E_k]$.

Now, we show how to construct \mathcal{F}_k from $\mathcal{E}_2, \dots, \mathcal{E}_{k-1}, \mathcal{F}_2, \dots, \mathcal{F}_{k-1}$, and \mathcal{T}_{k-2} for $k \geq 5$. For $2 \leq j \leq k$, let $F_{k,j} := \bigcup_{i=2}^j \{1\overleftarrow{u}0v \mid u \in D_{i-1} \text{ and } v \in D_{k-i}\}$. Since $F_k = F_{k,k}$, the following statement, which we prove by auxiliary induction on j , directly implies the existence of a spanning tree of $\mathcal{H}_k[F_k]$:

Claim. For $2 \leq j \leq k$, there is a spanning tree $\mathcal{F}_{k,j}$ of $\mathcal{H}_k[F_{k,j}]$.

For $j = 2$, we have $F_{k,2} = \{1100v \mid v \in D_{k-2}\} = 1100D_{k-2}$, so we let $\mathcal{F}_{k,2} := 1100\mathcal{T}_{k-2}$. Now, suppose $3 \leq j \leq k$. The fact that E_{j-1} and F_{j-1} form a partition of D_{j-1} yields the following partition of the set $F_{k,j}$:

$$\begin{aligned}F_{k,j} &= F_{k,j-1} \cup \{1\overleftarrow{u}0v \mid u \in D_{j-1} \text{ and } v \in D_{k-j}\} \\ &= F_{k,j-1} \cup \bigcup_{v \in D_{k-j}} 1\overleftarrow{D_{j-1}}0v \\ &= F_{k,j-1} \cup \bigcup_{v \in D_{k-j}} (1\overleftarrow{E_{j-1}}0v \cup 1\overleftarrow{F_{j-1}}0v).\end{aligned}\tag{15}$$

For every $v \in D_{k-j}$, consider the following marked triple on $F_{k,j}$:

$$\tau(v) := \{1(10)^{j-3}11000v, 1(10)^{j-3}10100v, 1(10)^{j-3}10010v\} = \alpha((10)^{j-3})v.\tag{16}$$

It belongs to Ψ_k by construction. We have

$$1(10)^{j-3}11000v \in 1\overleftarrow{F_{j-1}}0v, \quad 1(10)^{j-3}10100v \in 1\overleftarrow{E_{j-1}}0v, \quad 1(10)^{j-3}10010v \in F_{k,j-1},$$

where the first two memberships follow from (14). We take the spanning trees of the subhypergraphs of \mathcal{H}_k induced by the sets of the partition of $F_{k,j}$ given by (15) and connect them into a single spanning tree of $\mathcal{H}_k[F_{k,j}]$ using the triples $\tau(v)$ for all $v \in D_{k-j}$. That is, we let

$$\mathcal{F}_{k,j} := \mathcal{F}_{k,j-1} \cup \bigcup_{v \in D_{k-j}} (\{\tau(v)\} \cup 1\overleftarrow{E_{j-1}}0v \cup 1\overleftarrow{F_{j-1}}0v),$$

which is a spanning tree of $\mathcal{H}_k[F_{k,j}]$.

Finally, we show how to construct \mathcal{T}_k from \mathcal{F}_k , \mathcal{E}_{k-1} and \mathcal{F}_{k-1} . Consider the following marked triple on D_k :

$$\tau := \{111000(10)^{k-3}, 101100(10)^{k-3}, \underline{1}01010(10)^{k-3}\} = \beta(10)^{k-3}.$$

It belongs to Ψ_k by construction. We have

$$111000(10)^{k-3} \in F_k, \quad 101100(10)^{k-3} \in 10E_{k-1}, \quad 101010(10)^{k-3} \in 10F_{k-1},$$

where the first membership is by the definition of F_k for $k \geq 4$ and the other two follow from (14). The sets F_k , $10E_{k-1}$, and $10F_{k-1}$ form a partition of D_k . We take the spanning trees of the subhypergraphs induced by these partition sets and connect them into a single spanning tree of \mathcal{H}_k using the triple τ . That is, we let $\mathcal{T}_k := \mathcal{F}_k \cup \{\tau\} \cup 10\mathcal{E}_{k-1} \cup 10\mathcal{F}_{k-1}$, which is a spanning tree of \mathcal{H}_k . \square

Proof of Lemma 11. The proof proceeds along the same lines as the proof of Lemma 10, so we only highlight the differences.

Apart from the partition of D_4 into two sets E_4 and F_4 defined by (13), we will use another one—a partition into sets E'_4 and F'_4 defined as follows:

$$E'_4 := D_4 \setminus \{11001100\}, \quad F'_4 := \{11001100\}.$$

It has the following property analogous to (14):

$$10101100 \in E'_4, \quad 11001100 \in F'_4. \quad (17)$$

We define spanning trees $\mathcal{E}'_4 := \{\tau'_1, \tau'_2, \tau'_3, \tau'_4, \tau'_5\}$ of $\mathcal{H}_4[E'_4]$ and $\mathcal{F}'_4 := \emptyset$ of $\mathcal{H}_4[F'_4]$, where

$$\begin{aligned} \tau'_1 &:= \{1\underline{1}110000, 11\underline{1}01000, 1110\underline{0}100, 110101\underline{0}0\} = 1\overleftarrow{\delta}0, \\ \tau'_2 &:= \{11011\underline{0}00, 110101\underline{0}0, 1101\underline{0}010\} = \alpha(10), \\ \tau'_3 &:= \{11100\underline{0}10, 110100\underline{1}0, \underline{1}001010\} = \alpha(\epsilon)10, \\ \tau'_4 &:= \{11100\underline{0}10, 1011\underline{0}010, \underline{1}0101010\} = \beta10, \\ \tau'_5 &:= \{10111\underline{0}00, 101101\underline{0}0, 1010\underline{1}100, 10\underline{1}01010\} = 10\delta. \end{aligned}$$

We have $\text{supp } \tau'_1 \cup \dots \cup \text{supp } \tau'_5 = E'_4$, $|\text{supp } \tau'_{i-1} \cap \text{supp } \tau'_i| = 1$ for $i \in \{2, 3, 4, 5\}$, and $\text{supp } \tau'_i \cap \text{supp } \tau'_j = \emptyset$ whenever $|i - j| \geq 2$, which shows that \mathcal{E}'_4 is indeed a spanning tree of $\mathcal{H}_4[E'_4]$, and \mathcal{F}'_4 is a spanning tree of $\mathcal{H}_4[F'_4]$ because $|F'_4| = 1$.

To obtain many spanning trees of \mathcal{H}_k for $k \geq 6$, we proceed as in the proof of Lemma 10 except that we introduce variants to the construction of the spanning tree $\mathcal{F}_{k,5}$ of $\mathcal{H}_k[F_{k,5}]$. Consider an arbitrary partition of D_{k-5} into two sets X and Y . The fact that E_4 and F_4 as well as E'_4 and F'_4 form partitions of D_4 yields the following partition of $F_{k,5}$ analogous to (15):

$$\begin{aligned} F_{k,5} &= F_{k,4} \cup \{1\overleftarrow{u}0v \mid u \in D_4 \text{ and } v \in D_{k-5}\} \\ &= F_{k,4} \cup \bigcup_{v \in D_{k-5}} 1\overleftarrow{D}_4 0v \\ &= F_{k,4} \cup \bigcup_{v \in X} (1\overleftarrow{E}_4 0v \cup 1\overleftarrow{F}_4 0v) \cup \bigcup_{v \in Y} (1\overleftarrow{E}'_4 0v \cup 1\overleftarrow{F}'_4 0v). \end{aligned} \quad (18)$$

Consider the following triples on $F_{k,5}$, where the first one is a special case of (16):

$$\begin{aligned} \tau(v) &:= \{110101100\underline{0}v, 110101010\underline{0}v, 1101010010v\} = \alpha(1010)v \quad \text{for } v \in X, \\ \tau'(v) &:= \{111001100\underline{0}v, 111001010\underline{0}v, 1110010010v\} = \alpha(1100)v \quad \text{for } v \in Y. \end{aligned}$$

They belong to Ψ_k by construction. We have

$$\begin{aligned} 1101011000v &\in 1\overleftarrow{F}_4 0v, & 1101010100v &\in 1\overleftarrow{E}_4 0v, & 1101010010v &\in F_{k,4}, & \text{for } v \in X, \\ 1110011000v &\in 1\overleftarrow{F}'_4 0v, & 1110010100v &\in 1\overleftarrow{E}'_4 0v, & 1110010010v &\in F_{k,4}, & \text{for } v \in Y, \end{aligned}$$

where the first two memberships follow from (14) and (17), respectively. We take the spanning trees of the subhypergraphs of \mathcal{H}_k induced by the sets of the partition of $F_{k,5}$ given by (18) and connect them into a single spanning tree of $\mathcal{H}_k[F_{k,5}]$ using the triples $\tau(v)$ for all $v \in X$ and the triples $\tau'(v)$ for all $v \in Y$. That is, we let

$$\mathcal{F}_{k,5} := \mathcal{F}_{k,4} \cup \bigcup_{v \in X} (\{\tau(v)\} \cup 1\overleftarrow{E}_4 0v \cup 1\overleftarrow{F}_4 0v) \cup \bigcup_{v \in Y} (\{\tau'(v)\} \cup 1\overleftarrow{E}'_4 0v \cup 1\overleftarrow{F}'_4 0v),$$

which is a spanning tree of $\mathcal{H}_k[F_{k,5}]$. Then, we continue with the constructions of \mathcal{F}_k and \mathcal{T}_k exactly as in the proof of Lemma 10.

Clearly, distinct choices of the partition of D_{k-5} into two sets X and Y in the procedure above give rise to distinct spanning trees $\mathcal{F}_{k,5}$ of $\mathcal{H}_k[F_{k,5}]$, which consequently give rise to distinct spanning trees \mathcal{F}_k of $\mathcal{H}_k[F_k]$ and \mathcal{T}_k of \mathcal{H}_k . Since $|D_{k-5}| = \frac{1}{k-4} \binom{2k-10}{k-5} \geq 2^{k-6}$ for $k \geq 6$, there are at least $2^{2^{k-6}}$ distinct partitions of D_{k-5} into sets X and Y , which give rise to at least $2^{2^{k-6}}$ distinct spanning trees of \mathcal{H}_k . \square

6. ALTERNATIVE PROOF OF THE MIDDLE LEVELS CONJECTURE

As we explained in Section 1.3, Theorem 1 implies that the bipartite Kneser graph $H(2k+1, k)$ has a Hamilton path for every $k \geq 1$. We proceed to prove that it even has a Hamilton cycle, yielding an alternative proof of the middle levels conjecture, first proved in [Müt16]. Plugging in Theorem 3 instead of Theorem 1, we obtain an alternative proof of the fact that $H(2k+1, k)$ contains double-exponentially many distinct Hamilton cycles, also first proved in [Müt16].

Theorem 12. *For every integer $k \geq 1$, the bipartite Kneser graph $H(2k+1, k)$ has a Hamilton cycle. For every integer $k \geq 6$, the bipartite Kneser graph $H(2k+1, k)$ has at least $2^{2^{k-6}}$ distinct Hamilton cycles.*

Proof. For any graph G whose vertices are bitstrings, we let Gx denote the graph obtained by appending a bitstring x to all vertices of G , and we let \overline{G} denote the graph obtained from G by taking the complement of each vertex. The graph $H(2k+1, k)$ is isomorphic to the graph M_k obtained as the disjoint union of $G_k 0$ and $\overline{G}_k 1$ plus the *matching edges* $\{x0, x1\}$ with $x \in B_k^0$. The isomorphism is given by interpreting all bitstrings of length $2k+1$ as characteristic vectors of k -element and $(k+1)$ -element subsets of $[2k+1]$.

For $k=1$ and $k=2$, the theorem can be verified directly. For $k \geq 3$, Theorem 1 yields a Hamilton cycle in the graph G_k^+ . Moreover, it follows from the proof of Theorem 1 that this Hamilton cycle is obtained as the symmetric difference of the cycle factor \mathcal{C}_k in G_k^+ with some set of flipping cycles. The cycle factor \mathcal{C}_k is obtained from a collection of paths \mathcal{P}_k in G_k by adding the edges $\{x, \overline{x}\}$ with $x \in D_k$, where the path $P(x) \in \mathcal{P}_k$ connects x and \overline{x} . The flipping cycles are cycles in G_k . Therefore, the resulting Hamilton cycle C in G_k^+ contains some edges of G_k plus the edges $\{x, \overline{x}\}$ with $x \in D_k$. We remove the latter edges from C , thus obtaining a collection of paths \mathcal{Q} in G_k . Then, to obtain a Hamilton cycle in M_k , we take the paths in $\mathcal{Q}0$ and $\overline{\mathcal{P}}_k 1$ and add the matching edges $\{x0, x1\}$ and $\{\overline{x}0, \overline{x}1\}$ with $x \in D_k$. This is indeed a Hamilton cycle in M_k , obtained from the Hamilton cycle $C0$ in $G_k^+ 0$ by removing every edge of the form $\{x0, \overline{x}0\}$ (with $x \in D_k$) and replacing it by the path going from $\overline{x}0$ to $\overline{x}1$, then along $\overline{P(x)}1$ from $\overline{x}1$ to $x1$, and then to $x0$.

For $k \geq 6$, Theorem 3 yields at least $2^{2^{k-6}}$ distinct Hamilton cycles in G_k^+ with the properties discussed above, which give rise to at least $2^{2^{k-6}}$ distinct Hamilton cycles in $H(2k+1, k)$ by the construction described above. \square

7. OPEN PROBLEMS

As mentioned in Section 1.5, our construction of a Hamilton cycle in the odd graph O_k translates straightforwardly into an algorithm to compute this cycle in polynomial time, that is, polynomial in the size of the graph. It remains open whether this can be improved to an algorithm whose running time is polynomial in k , ideally, constant for each generated vertex. This might even yield a simpler algorithmic solution for the middle levels conjecture (recall Section 6 and [MN20]).

As Hamiltonicity of the odd graphs O_k is now settled, one may turn to Biggs' more general conjecture that O_k has $\lfloor (k+1)/2 \rfloor$ edge-disjoint Hamilton cycles for $k \geq 3$ [Big79]. Clearly, edge-disjointness is a much stronger requirement than distinctness guaranteed by Theorem 3. A starting point for such a construction may be the factorizations of O_k into cycles described by Johnson and Kierstead [JK04]. Our construction of a cycle factor is unrelated to these factorizations, and we do not even see how to extend it to a construction of two edge-disjoint cycle factors.

Another intriguing open problem is whether all Kneser graphs $K(n, k)$ except the Petersen graph $O_2 = K(5, 2)$ are Hamiltonian. Despite serious attempts, we have not been able to generalize the methods presented in this paper, nor the inductive decomposition technique of Johnson [Joh04], to the general case. In light of Theorem 2, the sparsest open case is $n = 2k + 3$.

A strengthening of the concept of containing a Hamilton cycle is to contain the r th power of a Hamilton cycle. To this end, Katona [Kat05] conjectured that the vertices of $K(n, k)$ can be ordered so that any $r + 1$ consecutive vertices for $r := \lfloor n/k \rfloor - 2$ are disjoint sets, which he proved for $k = 2$ using Walecki's theorem. It seems plausible that Katona's conjecture holds even for $r := \lceil n/k \rceil - 2$. Theorem 1 confirms this for the case $n = 2k + 1$ (where $r = 1$).

REFERENCES

- [Bal72] A. T. Balaban. Chemical graphs. XIII. Combinatorial patterns. *Rev. Roumain Math. Pures Appl.*, 17(1):3–16, 1972.
- [Bar75] Zs. Baranyai. On the factorization of the complete uniform hypergraph. In *Infinite and finite sets (Colloq., Keszthely, 1973; dedicated to P. Erdős on his 60th birthday)*, Vol. I, pages 91–108. Colloq. Math. Soc. János Bolyai, Vol. 10. North-Holland, Amsterdam, 1975.
- [Bár78] I. Bárány. A short proof of Kneser's conjecture. *J. Combin. Theory Ser. A*, 25(3):325–326, 1978.
- [Big79] N. Biggs. Some odd graph theory. In *Second International Conference on Combinatorial Mathematics (New York, 1978)*, volume 319 of *Ann. New York Acad. Sci.*, pages 71–81. New York Acad. Sci., New York, 1979.
- [BS19] J. Bellmann and B. Schülke. Short proof that Kneser graphs are Hamiltonian for $n \geq 4k$. *arXiv:1912.08186*, Dec 2019.
- [CF02] Y. Chen and Z. Füredi. Hamiltonian Kneser graphs. *Combinatorica*, 22(1):147–149, 2002.
- [Che00] Y. Chen. Kneser graphs are Hamiltonian for $n \geq 3k$. *J. Combin. Theory Ser. B*, 80(1):69–79, 2000.
- [Che03] Y. Chen. Triangle-free Hamiltonian Kneser graphs. *J. Combin. Theory Ser. B*, 89(1):1–16, 2003.
- [CI96] W. E. Clark and M. E. H. Ismail. Binomial and Q -binomial coefficient inequalities related to the Hamiltonicity of the Kneser graphs and their Q -analogues. *J. Combin. Theory Ser. A*, 76(1):83–98, 1996.
- [ČKM⁺13] M. Čadek, M. Krčál, J. Matoušek, L. Vokřínek, and U. Wagner. Extending continuous maps: polynomiality and undecidability. In *STOC '13: Proceedings of the 45th Annual ACM Symposium on Theory of Computing*, pages 595–604. ACM, New York, 2013.

- [ČKM⁺14] M. Čadek, M. Krčál, J. Matoušek, F. Sergeraert, L. Vokřínek, and U. Wagner. Computing all maps into a sphere. *J. ACM*, 61(3):Art. 17, 44 pages, 2014.
- [CL87] B. Chen and K. Lih. Hamiltonian uniform subset graphs. *J. Combin. Theory Ser. B*, 42(3):257–263, 1987.
- [EKR61] P. Erdős, C. Ko, and R. Rado. Intersection theorems for systems of finite sets. *Quart. J. Math. Oxford Ser. (2)*, 12:313–320, 1961.
- [GMN18] P. Gregor, T. Mütze, and J. Nummenpalo. A short proof of the middle levels theorem. *Discrete Analysis*, 2018:8:12 pp., 2018.
- [Gre02] J. E. Greene. A new short proof of Kneser’s conjecture. *Amer. Math. Monthly*, 109(10):918–920, 2002.
- [GSWW18] D. Gabric, J. Sawada, A. Williams, and D. Wong. A framework for constructing de Bruijn sequences via simple successor rules. *Discrete Math.*, 341(11):2977–2987, 2018.
- [Hol17] A. E. Holroyd. Perfect snake-in-the-box codes for rank modulation. *IEEE Trans. Inform. Theory*, 63(1):104–110, 2017.
- [HRW12] A. E. Holroyd, F. Ruskey, and A. Williams. Shorthand universal cycles for permutations. *Algorithmica*, 64(2):215–245, 2012.
- [HW78] K. Heinrich and W. D. Wallis. Hamiltonian cycles in certain graphs. *J. Austral. Math. Soc. Ser. A*, 26(1):89–98, 1978.
- [JK04] J. R. Johnson and H. A. Kierstead. Explicit 2-factorisations of the odd graph. *Order*, 21(1):19–27, 2004.
- [Joh04] J. R. Johnson. Long cycles in the middle two layers of the discrete cube. *J. Combin. Theory Ser. A*, 105(2):255–271, 2004.
- [Joh09] J. R. Johnson. Universal cycles for permutations. *Discrete Math.*, 309(17):5264–5270, 2009.
- [Joh11] J. R. Johnson. An inductive construction for Hamilton cycles in Kneser graphs. *Electron. J. Combin.*, 18(1):Paper 189, 12 pages, 2011.
- [Kar72] R. M. Karp. Reducibility among combinatorial problems. In *Complexity of computer computations (Proc. Sympos., IBM Thomas J. Watson Res. Center, Yorktown Heights, N.Y., 1972)*, pages 85–103. Plenum, New York, 1972.
- [Kat05] Gy. O. H. Katona. Constructions via Hamiltonian theorems. *Discrete Math.*, 303(1-3):87–103, 2005.
- [Knu11] D. E. Knuth. *The Art of Computer Programming. Vol. 4A. Combinatorial Algorithms. Part 1*. Addison-Wesley, Upper Saddle River, NJ, 2011.
- [Lov70] L. Lovász. Problem 11. In *Combinatorial Structures and Their Applications (Proc. Calgary Internat. Conf., Calgary, Alberta, 1969)*. Gordon and Breach, New York, 1970.
- [Lov78] L. Lovász. Kneser’s conjecture, chromatic number, and homotopy. *J. Combin. Theory Ser. A*, 25(3):319–324, 1978.
- [Mat76] M. Mather. The Rugby footballers of Croam. *J. Combin. Theory Ser. B*, 20(1):62–63, 1976.
- [Mat04] J. Matoušek. A combinatorial proof of Kneser’s conjecture. *Combinatorica*, 24(1):163–170, 2004.
- [ML72] G. H. J. Meredith and E. K. Lloyd. The Hamiltonian graphs O_4 to O_7 . In *Combinatorics (Proc. Conf. Combinatorial Math., Math. Inst., Oxford, 1972)*, pages 229–236. Inst. Math. Appl., Southend-on-Sea, 1972.
- [ML73] G. H. J. Meredith and E. K. Lloyd. The footballers of Croam. *J. Combin. Theory Ser. B*, 15:161–166, 1973.
- [MN20] T. Mütze and J. Nummenpalo. A constant-time algorithm for middle levels gray codes. *Algorithmica*, 82(5):1239–1258, 2020.
- [MS17] T. Mütze and P. Su. Bipartite Kneser graphs are Hamiltonian. *Combinatorica*, 37(6):1207–1219, 2017.
- [MSW18] T. Mütze, C. Standke, and V. Wiechert. A minimum-change version of the Chung-Feller theorem for Dyck paths. *European J. Combin.*, 69:260–275, 2018.
- [Müt16] T. Mütze. Proof of the middle levels conjecture. *Proc. London Math. Soc.*, 112(4):677–713, 2016.
- [Sav97] C. D. Savage. A survey of combinatorial Gray codes. *SIAM Rev.*, 39(4):605–629, 1997.
- [SS04] I. Shields and C. D. Savage. A note on Hamilton cycles in Kneser graphs. *Bull. Inst. Combin. Appl.*, 40:13–22, 2004.
- [Sta99] R. P. Stanley. *Enumerative combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999.
- [SW20] J. Sawada and A. Williams. Solving the sigma-tau problem. *ACM Trans. Algorithms*, 16(1):11:1–11:17, 2020.
- [Zie02] G. M. Ziegler. Generalized Kneser coloring theorems with combinatorial proofs. *Invent. Math.*, 147(3):671–691, 2002.