Extending perfect matchings to Gray codes with prescribed ends

Petr Gregor* Tomáš Novotný

Department of Theoretical Computer Science and Mathematical Logic
Charles University, Malostranské nám. 25, 118 00 Prague
Czech Republic

gregor@ktiml.mff.cuni.cz

Riste Škrekovski

Faculty of Mathematics and Physics, University of Ljubljana &
Faculty of Information Studies, Novo Mesto &
FAMNIT, University of Primorska, Koper
Slovenia

Submitted: Mar 27, 2017; Accepted: Jun 6, 2018; Published: Jun 22, 2018

Abstract

A *binary (cyclic) Gray code* is a (cyclic) ordering of all binary strings of the same length such that any two consecutive strings differ in a single bit. This corresponds to a Hamiltonian path (cycle) in the hypercube. Fink showed that every perfect matching in the n-dimensional hypercube Q_n can be extended to a Hamiltonian cycle, confirming a conjecture of Kreweras. In this paper, we study the "path version" of this problem. Namely, we characterize when a perfect matching in Q_n extends to a Hamiltonian path between two prescribed vertices of opposite parity. Furthermore, we characterize when a perfect matching in Q_n with two faulty vertices extends to a Hamiltonian cycle. In both cases we show that for all dimensions $n \geqslant 5$ the only forbidden configurations are so-called half-layers, which are certain natural obstacles. These results thus extend Kreweras' conjecture with an additional edge, or with two faulty vertices. The proof for the case $n = 5$ is computer-assisted.

Mathematics Subject Classifications: 05C38, 05C45, 05A15

1 Introduction

A *binary* (*cyclic*) *n-bit Gray code* is a (cyclic) ordering of all 2^n binary strings of length n such that any two consecutive strings differ in a single bit. Alternatively, it can be viewed

*The corresponding author.

as a Hamiltonian path (cycle) in the n-dimensional hypercube Q_n. It is named after Frank Gray who in 1953 patented a scheme based on this code to convert analog signals to digital [10] but its description can be traced back much earlier in history [11, 12].

Gray codes turned out to be useful and flexible at the same time [6]. Applications have been found in such diverse areas as data compression, graphics and image processing, information retrieval, signal encoding or processor allocation in hypercubic networks [17]. Gray codes satisfying certain additional properties have been subject of extensive research with recent breakthroughs, e.g. confirmation of the Middle Levels Conjecture [14].

Ruskey and Savage [16] asked if any matching in a hypercube Q_n can be extended to a Hamiltonian cycle. This question is still open in general, although for small dimensions ($n \leqslant 5$) [20], for small matchings (bounded by a $O(n^2)$ function) [3, 7, 19], or for an extension into 2-factors [9, 18] it has been answered positively. Another special case when the matching is perfect was independently conjectured by Kreweras [13]. This has been affirmatively answered by Fink [8] who proved a strengthened version of Kreweras' conjecture for the complete graph $K(Q_n)$ on the vertices of the hypercube Q_n.

Theorem 1 (Fink [8])**.** *Let P be a perfect matching of $K(Q_n)$. There exists a perfect matching R of Q_n such that $P \cup R$ induces a Hamiltonian cycle of $K(Q_n)$.*

In this paper, we study the "path version" of this problem. Namely, in the following theorem we characterize when there is a Hamiltonian path in the hypercube Q_n between two given vertices x, y of opposite parity containing a given perfect matching. This can be also viewed as an extension of Kreweras' conjecture with one additional edge xy. Similarly as Fink employed a strengthened version in Theorem 1 for the graph $K(Q_n)$, we need a strengthened version for the complete bipartite graph $B(Q_n)$ on the vertices of Q_n (with the same bipartition).

Theorem 2. *Let P be a perfect matching of $B(Q_n)$, $n \geqslant 5$, and let x, y be vertices of opposite parity with $xy \notin P$. There exists a matching $R \subseteq E(Q_n)$ such that $P \cup R$ forms a Hamiltonian xy-path of $B(Q_n)$ if and only if $(P \cup x^P y^P) \setminus \{xx^P, yy^P\}$ does not contain a half-layer, where x^P and y^P are the vertices such that $xx^P, yy^P \in P$.*

A *half-layer* in Q_n is a set of edges with the same direction and the same parity, see the definitions below. Our result shows that these easy to recognize configurations are the only obstacles. For convenience in the induction proof we actually prove the following equivalent and shorter form of Theorem 2.

Theorem 3. *Let x, y be vertices of opposite parity in Q_n, $n \geqslant 5$, and let P be a perfect matching of $B(Q_n - \{x, y\})$. There exists a matching $R \subseteq E(Q_n)$ such that $P \cup R$ forms a Hamiltonian cycle of $B(Q_n - \{x, y\})$ if and only if P does not contain a half-layer.*

To see that Theorem 3 implies Theorem 2, assume that P is a perfect matching of $B(Q_n)$, $n \geqslant 5$, and x, y are vertices of opposite parity with $xy \notin P$. Let $P' = (P \cup x^P y^P) \setminus \{xx^P, yy^P\}$ where x^P and y^P are the vertices such that $xx^P, yy^P \in P$. Observe that for any matching $R \subseteq E(Q_n)$ the set of edges $P \cup R$ forms a Hamiltonian path between x and y in $B(Q_n)$ if and only if $P' \cup R$ is a Hamiltonian cycle of $B(Q_n - \{x, y\})$. By Theorem 3 there exists such matching R if and only if P' does not contain a half-layer.

On the other hand, to see that Theorem 2 implies Theorem 3, assume that x, y are vertices of opposite parity in Q_n, $n \geqslant 5$, and P is a perfect matching of $B(Q_n - \{x, y\})$. Choose any edge $x^P y^P \in P$ and let $P' = (P \setminus \{x^P y^P\}) \cup \{xx^P, yy^P\}$ assuming that x^P and y^P have opposite parity to x and y, respectively. Observe that for any matching $R \subseteq E(Q_n)$ the set of edges $P \cup R$ forms a Hamiltonian cycle of $B(Q_n - \{x, y\})$ if and only if $P' \cup R$ is a Hamiltonian xy-path of $B(Q_n)$. By Theorem 2 there exists such matching R if and only if P does not contain a half-layer.

Theorem 3 can be also viewed as another extension of Kreweras' conjecture for hypercubes with two faulty vertices x, y of opposite parity. In [1] the authors showed that Kreweras' conjecture also holds for sparse spanning regular subgraphs of hypercubes. As for other related results, in a more general setting when the prescribed edges can be incident (i.e. not necessarily a matching) it is known [4, 5] that any $2n-3$ (resp. $2n-4$) edges satisfying certain necessary conditions can be prescribed for a Hamiltonian cycle (resp. for a path between given vertices), and these bounds are tight. In a sense complementary results characterize when there is a Hamiltonian cycle (resp. path) in Q_n that avoids a given matching [2]. In particular, a Hamiltonian cycle exists in $Q_n - M$ for $n \geqslant 4$ if and only if the forbidden matching M does not contain a half-layer [2]. Thus half-layers are the only obstacles in the complementary problem as well.

It should be noted that our proof is computer-assisted; the case $n = 5$, which serves us as a base of induction, was verified on computer by an exhaustive checking of all non-isomorphic configurations.

The paper is organized as follows. In Section 2 we introduce definitions and notation. In Section 3 we study half-layers and quad-layers that play a key role as obstacles in our results. In Section 4 we prove the induction step of Theorem 3. In Section 5 we conclude with discussion of a possible generalizations. In Appendix we describe our algorithm that verifies the base of induction for $n = 5$. The source code of the verifying algorithm in C++ is available on [15].

2 Preliminaries

For a positive integer n we denote by $[n]$ the set $\{1, 2, \ldots, n\}$. As usual, the vertex and the edge sets of a graph G are denoted by $V(G)$ and $E(G)$, respectively. For sets of so-called *faulty* elements $V \subseteq V(G)$ and $E \subseteq E(G)$, let $G - V$ denote the subgraph of G induced by $V(G) \setminus V$, and let $G - E$ denote the graph with vertices $V(G)$ and edges $E(G) \setminus E$. For a graph G we denote by $K(G)$ the complete graph on vertices $V(G)$, and if G is bipartite with a unique bipartition, we denote by $B(G)$ the complete bipartite graph on vertices $V(G)$ with the same bipartition. Note that $E(G) \subseteq E(B(G)) \subseteq E(K(G))$.

The *n-dimensional hypercube* Q_n is a (bipartite) graph with all binary vectors of length n as vertices and with edges joining every two vertices that differ in exactly one coordinate, i.e.

$$V(Q_n) = \{0, 1\}^n \quad \text{and} \quad E(Q_n) = \{uv \mid |\Delta(u, v)| = 1\},$$

where $\Delta(u,v) = \{i \in [n] \mid u_i \neq v_i\}$. Thus the distance of vertices u and v is $d(u,v) = |\Delta(u,v)|$. The distance of two edges uv and xy is the minimal distance between a vertex of uv and a vertex of xy.

The *weight* $w(v)$ of a vertex v is the number of 1's in v, i.e. $w(v) = d(v, \mathbf{0})$ where $\mathbf{0} = (0 \cdots 0)$. Furthermore, for an edge uv we define its weight $w(uv) = \min\{w(u), w(v)\}$, i.e. it is the weight of its vertex closer to $\mathbf{0}$. The *parity* of a vertex or an edge is the parity of its weight. Note that vertices of each parity form bipartite sets of Q_n. We denote (on illustrations) the set of even (odd) vertices by W (resp. B). Consequently, any two vertices u and v have the same parity if and only if $d(u,v)$ is even.

For any $i \in [n]$, let Q^i_L and Q^i_R be the $(n-1)$-dimensional subcubes that are induced by all vertices with fixed 0 and 1 (respectively) in the i-th coordinate. Clearly, each $(n-1)$-dimensional subcube is Q^i_L or Q^i_R for a unique i. The set of all edges between Q^i_L and Q^i_R is called a *layer (of direction i)*. Thus, $E(Q_n)$ can be partitioned into n layers, one layer for each $i \in [n]$. For a vertex x of Q_n and $i \in [n]$ let x^i denote the neighbor of x in direction i.

3 Half-layers and quad-layers

In this section we study half-layers and quad-layers which are natural obstacles that need to be avoided in our inductive construction.

Notice that an edge xy of direction $i \in [n]$ in Q_n such that $x \in V(Q^i_L)$ and $y \in V(Q^i_R)$ is *even* (resp. *odd*) if the vertex x is even (resp. odd). The set of all edges between Q^i_L and Q^i_R of the same parity is called a *half-layer* of Q_n. A half-layer is *odd* (resp. *even*) if its edges are odd (resp. even). Every layer comprises two disjoint half-layers, one odd and one even. See Figure 1 for an illustration. Clearly, any two edges in a half-layer have distance at least 2. Furthermore, observe that a half-layer of Q_n does not belong to any $(n-1)$-dimensional subcube if $n \geqslant 3$.

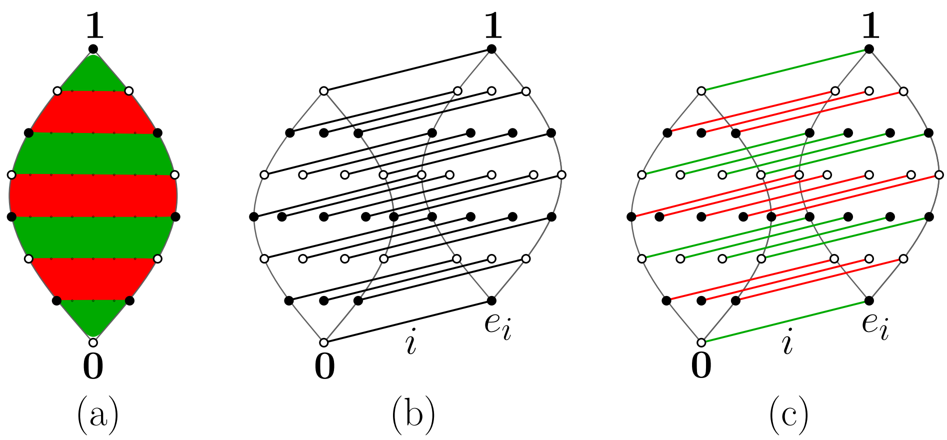


Figure 1: A lattice representation of the hypercube: (a) even (green) and odd (red) edges, (b) a layer in the direction i, (c) even (green) and odd (red) half-layers in the direction i.

The following lemma shows that two half-layers of different directions are incident in a strong sense.

Lemma 4. *Let H_i and H_j be half-layers of Q_n of different directions i and j, respectively. Then, every edge of H_i is incident with some edge of H_j.*

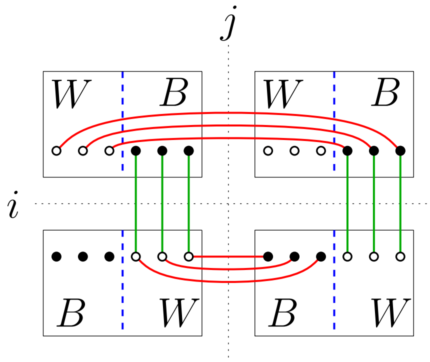


Figure 2: An illustration for Lemma 1: every edge of the green half-layer is incident with some edge of the red half-layer.

Proof. Let xx^i be an edge from H_i. Notice that the edges xx^j and $x^i(x^i)^j$ are both of direction j and both are incident with the edge xx^i. Moreover, they are of different parity, so precisely one of them is in H_j. See Figure 2 for an illustration. $\square$

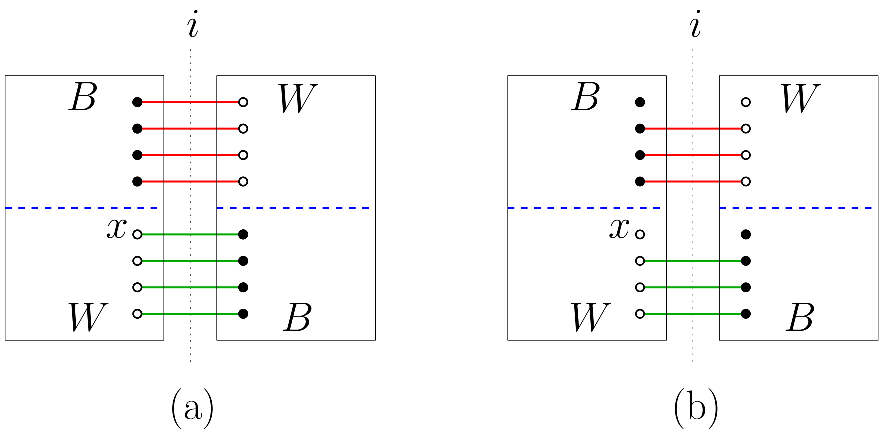
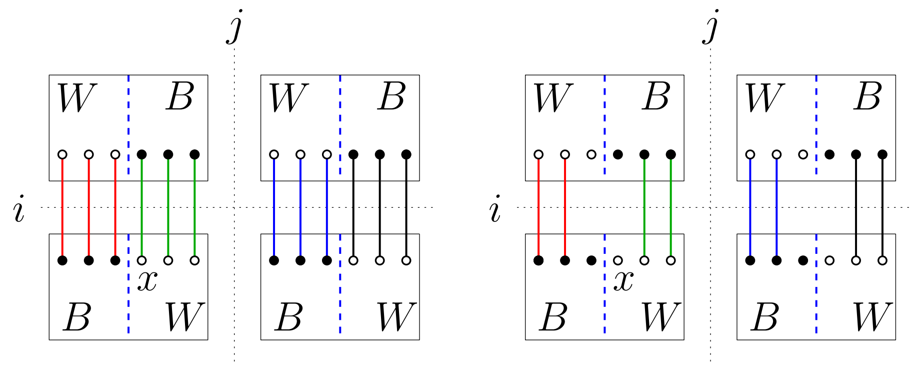


Figure 3: A schematic representation of half-layers (the blue line represents the bipartition): (a) even (green) and odd (red) half-layers of direction i, the red half-layer is x-dangerous, but the green is not; (b) the red almost half-layer is x-dangerous, but the green is not.

Since we proceed in the proof of Theorem 3 inductively by splitting Q_n into two $(n-1)$-dimensional subcubes, we need to consider so-called quad-layers. A *quad-layer* *(of direction i)* is a half-layer (of direction i) in some $(n-1)$-dimensional subcube. If j is the fixed direction of this $(n-1)$-dimensional subcube (i.e. it is one of Q_L^j and Q_R^j), then it is called a *j-separate quad-layer (of direction i)*. Note that such j is unique for every quad-layer if $n \geqslant 4$. (For $n = 3$ a quad-layer of direction i contains a single edge and is j-separate for both $j \in [n] \setminus \{i\}$.) Observe that every half-layer can be partitioned into two quad-layers in $n-1$ ways. See Figures 3(a) and 4(a) for an illustration.



(a) (b)

Figure 4: A schematic representation of quad-layers: (a) red, green, blue, and black are j-separate quad-layers of direction i, only the red quad-layer is x-dangerous; (b) red, green, blue, and black are (j-separate) almost quad-layers, only the red almost quad-layer is x-dangerous.

The following lemma shows that half-layers and quad-layers may occur in a matching of $K(Q_n)$ (and thus also of $B(Q_n)$) only in a limited number of directions. It is used only in the proof of Lemma 7.

Lemma 5. *Let P be a matching of $K(Q_n)$. Then,*

(i) P contains half-layers in at most one direction;

(ii) if $n \geqslant 4$, then P contains quad-layers in at most two directions.

Proof. The first claim follows immediately by Lemma 4. For the second claim, suppose that we have three quad-layers L_1, L_2, L_3 in P of distinct directions d_1, d_2, d_3, respectively. Let L_i be an s_i-separate quad-layer for $i = 1, 2, 3$.

If $s_1 = s_2 = s_3$, then at least two of the quad-layers belong to the same $(n-1)$-dimensional subcube $Q_L^{s_1}$ or $Q_R^{s_1}$, say L_1 and L_2 belong to $Q_L^{s_1}$. But then L_1 and L_2 are half-layers of $Q_L^{s_1}$, and we obtain a contradiction by Lemma 4. So at least one of s_1, s_2, s_3 is distinct from the other two. By renumbering the quad-layers we may assume in the sequel that it is s_1, so $s_1 \neq s_2$ and $s_1 \neq s_3$.

As $s_1 \neq s_2$, we argue in the following way. Without loss of generality, we may assume that L_1 is in $Q_L^{s_1}$ and L_2 is in $Q_L^{s_2}$. Now we claim that *the $(n-2)$-cube $Q_L^{s_1} \cap Q_L^{s_2}$ does not contain any edge of $L_1 \cup L_2$*. Otherwise, suppose that the edge $xy \in L_1$ is contained in

$Q_L^{s_1} \cap Q_L^{s_2}$. As $d_1 \neq d_2$, the direction of xy is distinct from d_2. So, it follows that either x or y is incident with an edge of L_2. This is a contradiction that establishes our claim. As $n \geqslant 4$, it follows that the $(n-2)$-cube $Q_L^{s_2} - V(Q_L^{s_1}) = Q_R^{s_1} \cap Q_L^{s_2}$ is too small to contain all edges of L_2. So we conclude that the edges of L_2 have one vertex in $Q_L^{s_1} \cap Q_L^{s_2}$ and the other in $Q_R^{s_1} \cap Q_L^{s_2}$. This implies that $d_2 = s_1$. Similarly $d_1 = s_2$.

In an analogous way, from $s_1 \neq s_3$ we can conclude that $d_1 = s_3$ and $d_3 = s_1$. However, now we obtain that $d_2 = s_1 = d_3$, which is a contradiction with the assumption that directions d_1, d_2, d_3 are distinct. □

Notice that a quad-layer in Q_3 consists of a single edge and Q_3 contains a matching consisting of three edges of distinct directions, thus the bound $n \geqslant 4$ in the second claim of the above lemma cannot be decreased.

In the proof of Theorem 3, namely in Case 2(b)(i), we need to add a single given edge to the prescribed matching in one subcube without introducing a half-layer in the matching. This motivates the following definitions.

Let L be a layer and uv be an edge of L. We say that $L \setminus \{uv\}$ is an *almost layer*. Similarly, we define an *almost half-layer* and an *almost quad-layer*. So, an almost layer, an almost half-layer, and an almost quad-layer miss one edge to become a layer, a half-layer, and a quad-layer, respectively. Observe that any almost half-layer of Q_n does not belong to any $(n-1)$-dimensional subcube if $n \geqslant 4$. See Figures 3(b) and 4(b) for an illustration.

For almost half and quad-layers we have the following lemma, analogous to Lemma 5. It is also used only in the proof of Lemma 7.

Lemma 6. *Let P be a matching of $K(Q_n)$ with $n \geqslant 4$. Then,*

(i) P contains almost half-layers in at most one direction;

(ii) if $n \geqslant 6$, then P contains almost quad-layers in at most two directions.

Proof. Let L_1, L_2 be two almost half-layers of Q_n in different directions. Since $n \geqslant 4$, the set L_1 contains at least 3 edges. By Lemma 4, at most one of them is not incident with L_2. Thus L_1, L_2 cannot be both in the matching P, which proves the first claim. For the second claim, we proceed similarly as in the proof of Lemma 5. Suppose that we have three almost quad-layers L_1, L_2, L_3 in P of distinct directions d_1, d_2, d_3, respectively. Let L_i be an s_i-separate almost quad-layer for $i = 1, 2, 3$.

If $s_1 = s_2 = s_3$, then at least two of the almost quad-layers belong to the same $(n-1)$-dimensional subcube $Q_L^{s_1}$ or $Q_R^{s_1}$, say L_1 and L_2 belong to $Q_L^{s_1}$. But then L_1 and L_2 are almost half-layers of $Q_L^{s_1}$, and we obtain a contradiction with the claim (i) for $P \cap Q_L^{s_1}$. So at least one of s_1, s_2, s_3 is distinct from the other two. By renumbering the almost quad-layers we may assume in the sequel that it is s_1, so $s_1 \neq s_2$ and $s_1 \neq s_3$.

As $s_1 \neq s_2$, we argue in the following way. Without loss of generality, we may assume that L_1 is in $Q_L^{s_1}$ and L_2 is in $Q_L^{s_2}$. Now we claim that *the $(n-2)$-cube $Q_L^{s_1} \cap Q_L^{s_2}$ contains at most one edge of L_1 and at most one edge of L_2*. Otherwise, suppose that two edges $x_1y_1, x_2y_2 \in L_1$ are contained in $Q_L^{s_1} \cap Q_L^{s_2}$. As L_2 is an almost half-layer in $Q_L^{s_2}$ of direction

$d_2 \neq d_1$, at least one of the vertices x_1, x_2, y_1, y_2 is incident with L_2. This is a contradiction that establishes our claim.

As $n \geqslant 6$, it follows that the $(n-2)$-cube $Q_L^{s_2} - V(Q_L^{s_1}) = Q_R^{s_1} \cap Q_L^{s_2}$ is too small to contain all the edges of L_2 that are not in $Q_L^{s_1} \cap Q_L^{s_2}$. Indeed, L_2 has $2^{n-3} - 1$ edges, at most one of them is in $Q_L^{s_1} \cap Q_L^{s_2}$, and $Q_R^{s_1} \cap Q_L^{s_2}$ can contain at most 2^{n-4} edges with mutual distance at least 2, but $2^{n-3} - 2 > 2^{n-4}$ for $n \geqslant 6$. So we conclude that the edges of L_2 have one vertex in $Q_L^{s_1} \cap Q_L^{s_2}$ and the other in $Q_R^{s_1} \cap Q_L^{s_2}$. This implies that $d_2 = s_1$. Similarly $d_1 = s_2$.

In an analogous way, from $s_1 \neq s_3$ we can conclude that $d_1 = s_3$ and $d_3 = s_1$. However, now we obtain that $d_2 = s_1 = d_3$, which is a contradiction with the assumption that directions d_1, d_2, d_3 are distinct. □

In the proof of Theorem 3, quad-layers may lead to obstacles only if they are in the $(n-1)$-dimensional subcube containing one of the removed vertices. This motivates the following additional definitions.

Let x be a vertex of Q_n. If x is not incident with any edge of a half-layer L of Q_n, then we say that L is *x-dangerous*. A quad-layer L is *x-dangerous*, if it is an x-dangerous half-layer in some $(n-1)$-dimensional subcube which contains x. Furthermore, we define an *x-dangerous almost half-layer* as an x-dangerous half-layer without one edge. Similarly, an *x-dangerous almost quad-layer* is an x-dangerous quad-layer without one edge. See Figures 3 and 4 for an illustration. The following statements are simple observations:

(o1) *Every x-dangerous half-layer covers $n-1$ neighbors of x. Namely, if the half-layer is of direction i, then the only uncovered neighbor of x is x^i.*

(o2) *Every x-dangerous quad-layer covers $n-2$ neighbors of x. Namely, if the quad-layer is j-separated and of direction i, then the only uncovered neighbors of x are x^i and x^j.*

(o3) *Every x-dangerous almost quad-layer covers at least $n-3$ neighbors of x.*

In the following lemma we restate the above two lemmas for hypercubes with two faulty vertices of opposite parity. Additionally, we include similar statements also for x-dangerous half and quad-layers. This lemma is used in the proof of Theorem 3.

Lemma 7. *Let P be a matching of $K(Q_n - \{x, y\})$, where x, y are vertices of opposite parity. Then,*

(i) *P contains half-layers in at most one direction;*

(ii) *P contains quad-layers in at most two directions for $n \geqslant 4$;*

(iii) *P contains x-dangerous quad-layers in at most one direction for $n \geqslant 4$;*

(iv) *P contains almost half-layers in at most one direction for $n \geqslant 4$;*

(v) *P contains almost quad-layers in at most two directions for $n \geqslant 6$;*

(vi) P contains x-dangerous almost quad-layers in at most one direction for $n \geqslant 6$.

Proof. The claims *(i)*, *(ii)* and *(iv)*, *(v)* follow directly from Lemmas 5 and 6, respectively. Now we consider the claim *(iii)*. By *(ii)*, x-dangerous quad-layers could be in at most two directions, say i and j. Suppose that L_i and L_j are x-dangerous quad-layers of directions i and j, respectively. Since x is not covered by $L_i \cup L_j$ and $(x^i)^j$ is not covered by both L_i, L_j, it follows that at least one of x^i, x^j is not covered by $L_i \cup L_j$. Since each of L_i, L_j covers $n-2$ neighbors of x by $(o2)$ and $n \geqslant 4$, some neighbor must be covered by both, a contradiction.

The claim *(vi)* is proven in the same way as the claim *(iii)*. The only difference is in counting. Each of x-dangerous almost quad-layers L_i, L_j covers at least $n-3$ neighbors of x by $(o3)$ and $n \geqslant 6$, which again leads to a contradiction. □

4 Proof of Theorem 3

Proof. First we prove the necessity. Assume that P contains a half-layer L of direction i. We may assume that $x \in V(Q_L^i)$ and that the half-layer L is even, i.e. it covers the set A of even vertices in Q_L^i and x is odd. Let B denote the set of odd vertices in Q_L^i distinct from x. Assume by way of contradiction that there is a matching $R \subseteq E(Q_n - \{x, y\})$ such that $P \cup R$ forms a Hamiltonian cycle of $B(Q_n - \{x, y\})$. However, this is impossible, as R must match each element of A to a distinct element of B and $|A| > |B|$.

For the other direction (sufficiency) assume that P does not contain a half-layer. We proceed by induction on n. For $n = 5$ we verified the sufficiency part of Theorem 2 by computer, see the Appendix for a description of the verifying algorithm. This implies that also the sufficiency part of Theorem 3 holds for $n = 5$.

Now we assume that the statement holds for $n-1$ and we prove it for $n \geqslant 6$. The main idea is to cut the cube into two subcubes $Q_L := Q_L^d$ and $Q_R := Q_R^d$ through a carefully selected direction d in order to apply induction. In the case when both x, y belong to the same subcube, say Q_L, we first apply induction in Q_L and use Theorem 1 in Q_R and finally combine both matchings in order to obtain the required matching R. In the case when x and y are in different parts, say x in Q_L and y in Q_R, we choose some $y' \in V(Q_L)$ and $x' \in V(Q_R)$ in order to apply induction in both parts, and afterwards we again combine the obtained matchings.

Now we specify how to determine d. If P contains an x-dangerous or y-dangerous (almost) quad-layer, then we use the direction of such (almost) quad-layer for d. Note that by Lemma 7*(vi)* we may have two choices for d, one for an x-dangerous (almost) quad-layer and one for an y-dangerous (almost) quad-layer. If P does not contain an x-dangerous or y-dangerous (almost) quad-layer, we choose for d a direction with the maximal number of edges of P between Q_L^d and Q_R^d. As $n \geqslant 6$, an almost quad-layer has at least 7 edges. A direction with the maximal number of edges of P between the subcubes contains at least $(2^{n-1} - 1)/n \geqslant 6$ edges. Thus, there are at least 6 edges of P between Q_L^d and Q_R^d. Let P_L and P_R denote the set of edges of P in Q_L and Q_R, respectively, and let A denote the set of edges of P between the subcubes; thus, $P = P_L \cup P_R \cup A$ and

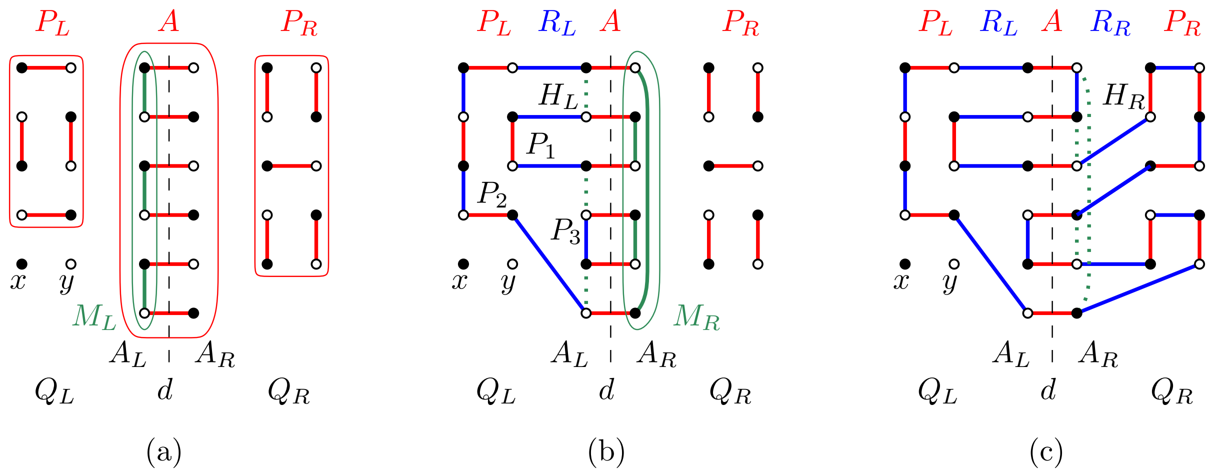


Figure 5: Three steps in Case 1: (a) extending P_L by a matching M_L on vertices of A_L; (b) applying induction in $B(Q_L - \{x, y\})$ and the matching M_R obtained on vertices of A_R; (c) applying Theorem 1 in Q_R.

$|A| \geqslant 6$. Let A_L and A_R denote the set of vertices from Q_L and Q_R, respectively, incident with the edges of A.

By Lemma 7(vi) and the choice of d, we obtain that x or y has no (almost) half-layer from P inside its subcube. So, without loss of generality, we can assume that x has this property; otherwise we swap x and y. Moreover, we can assume that x is in Q_L; otherwise we swap Q_L and Q_R. Thus P_L does not contain an (almost) half-layer of Q_L. Now, we distinguish two cases regarding whether y belongs to Q_L or Q_R.

Case 1: *y is in Q_L.* See Figure 5 for an illustration. Since P_L contains no (almost) half-layer of Q_L and $|A_L| \geqslant 6$, we can easily extend P_L on the vertices of A_L by some new edges M_L of $B(Q_L - \{x, y\})$ to a perfect matching $P_L^* = P_L \cup M_L$ of $B(Q_L - \{x, y\})$ so that P_L^* contains no half-layer. Indeed, take an arbitrary matching M_L of $B(Q_L - \{x, y\})$ on A_L and if $P_L \cup M_L$ contains a half-layer B, choose an edge $u_B u_W \in M_L \cap B$ and any other edge $v_B v_W \in M_L$ and swap their endvertices; that is, take $M_L := (M_L \setminus \{u_B u_W, v_B v_W\}) \cup \{u_B v_W, v_B v_W\}$. Then the modified M_L with P_L contains no half-layer by Lemma 4. By applying induction in Q_L for P_L^* we obtain a set of edges $R_L \subseteq E(Q_L)$ such that $P_L^* \cup R_L$ forms a Hamiltonian cycle H_L of $B(Q_L - \{x, y\})$.

Notice that $H_L - M_L$ is a union of some vertex-disjoint paths $P_1, P_2, \ldots, P_k$, each with endvertices of opposite parity. Let us denote the endvertices of P_j by $u_j, v_j \in A_L$, and furthermore let us denote by $u_j^A, v_j^A \in A_R$ their neighbors through the edges of A, i.e. $u_j u_j^A, v_j v_j^A \in A$. Then $M_R = \{u_j^A v_j^A \mid j = 1, \ldots, k\}$ is a matching in $B(Q_R)$ that extends the matching P_R on the vertices of A_R to a perfect matching $P_R^* = P_R \cup M_R$ of $B(Q_R)$.

Now, by applying Theorem 1 in Q_R for P_R^* we obtain a set of edges $R_R \subseteq E(Q_R)$ such that $P_R^* \cup R_R$ forms a Hamiltonian cycle H_R of $B(Q_R)$. It remains to observe that $(H_L \setminus M_L) \cup (H_R \setminus M_R) \cup A$ forms a Hamiltonian cycle of $B(Q_n - \{x, y\})$. Therefore the required matching is $R = R_L \cup R_R$.

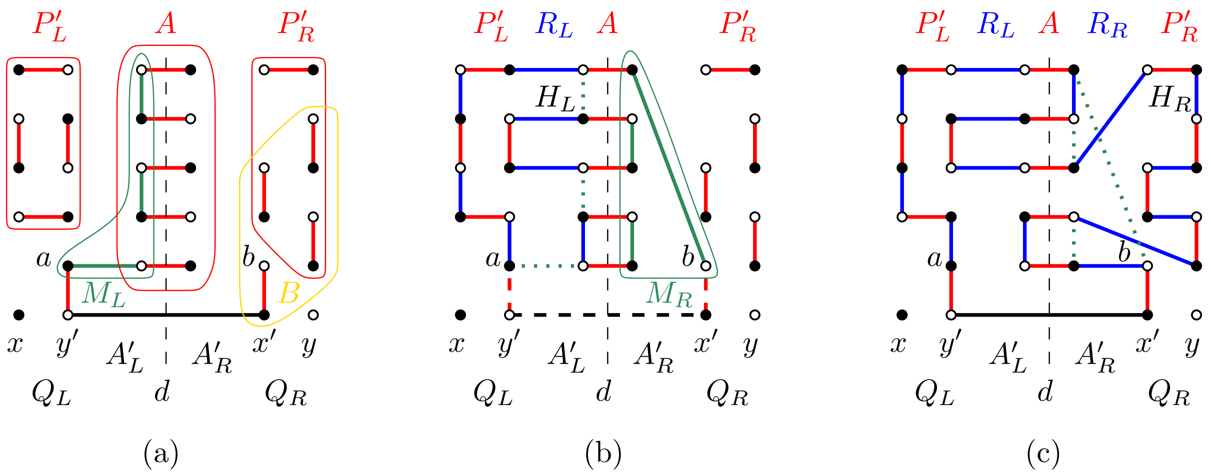


Figure 6: Three steps in Case 2(a): (a) extending P'_L by a matching M_L on vertices of A'_L; (b) applying induction in $B(Q_L - \{x, y'\})$ and the matching M_R obtained on vertices of A'_R; (c) applying induction in $B(Q_R - \{x', y\})$.

Case 2: *y is in Q_R.* We distinguish two subcases:

(a) *P_R contains a half-layer B of Q_R.* See Figure 6 for an illustration. Notice that B is y-dangerous. First we choose a vertex x' in Q_R of opposite parity to y that is covered by some edge $x'b$ of B, and we put $P'_R = P_R \setminus \{x'b\}$, $A'_R = A_R \cup \{b\}$. Let $y' = (x')^d$ denote the neighbor of x' in Q_L, $P'_L = P_L \setminus \{y'a\}$, and $A'_L = A_L \cup \{a\}$, where $a \in V(Q_L)$ such that $y'a \in P_L$.

Second, we extend P'_L on A'_L by some edges M_L to a perfect matching P^*_L of $B(Q_L - \{x, y'\})$ so that P^*_L contains no half-layer of Q_L. This can be done since P_L does not contain a half-layer in Q_L and $|A_L| \geqslant 6$, recall the choice of M_L in Case 1. By applying induction in Q_L for P^*_L we obtain a set of edges $R_L \subseteq E(Q_L)$ such that $P^*_L \cup R_L$ forms a Hamiltonian cycle H_L of $B(Q_L - \{x, y'\})$.

Next, we take the matching M_R of $B(Q_R - \{x', y\})$ on vertices of A'_R exactly as in Case 1 but with respect to the set $A' = A \cup \{ab\}$ instead of A. Thus M_R extends the matching P'_R on A'_R to a perfect matching $P^*_R = P'_R \cup M_R$ of $B(Q_R - \{x', y\})$. By Lemma 7(iv), P^*_R does not contain a half-layer of Q_R since P'_R already contains an almost half-layer $B \setminus \{x'b\}$ and the edge $x'b$ cannot occur in M_R (it is not an edge of $B(Q_R - \{x', y\})$).

Thus, we may apply induction again in Q_R for P^*_R to obtain a set of edges $R_R \subseteq E(Q_R)$ such that $P^*_R \cup R_R$ forms a Hamiltonian cycle H_R of $B(Q_R - \{x', y\})$. It remains to observe that $(H_L \setminus M_L) \cup (H_R \setminus M_R) \cup A \cup \{ay', y'x', x'b\}$ forms a Hamiltonian cycle of $B(Q_n - \{x, y\})$. Therefore the required matching is $R = R_L \cup R_R \cup \{x'y'\}$.

(b) *P_R contains no half-layer of Q_R.* In this case, to be able to apply induction in Q_R for the matching P^*_R we will carefully choose the edges of the matching M_L so that it is guaranteed that P^*_R has no half-layer in Q_R. First, since P contains no half-layer of Q_n, there is a vertex x' in Q_R of opposite parity to y such that $x'y' \notin P$ where $y' = (x')^d$ is the neighbor of x' in Q_L. Similarly as in the previous subcase, we define $P'_R = P_R \setminus \{x'b\}$, $P'_L = P_L \setminus \{y'a\}$, $A'_R = A_R \cup \{b\}$, $A'_L = A_L \cup \{a\}$ where a, b are the vertices such that $x'b, y'a \in P$.

We say that a direction $j \in [n] \setminus \{d\}$ is *dangerous* if P'_R can be extended on A'_R to a perfect matching P^*_R of $B(Q_R - \{x', y\})$ containing a half-layer of Q_R in direction j. Clearly, such half-layers are y-dangerous.

(i) *At most one direction is dangerous.* If no direction is dangerous then we may proceed exactly as in Subcase (a). Now assume that j is a dangerous direction. Our aim is to first choose one edge for the matching M_L so that it is guaranteed that later the matching M_R does not introduce the half-layer in direction j into $P^*_R = P'_R \cup M_R$.

Let $u^A v^A$ be any edge of the y-dangerous half-layer in direction j of Q_R such that $u^A, v^A \in A'_R$, and let $u, v \in A'_L$ be the neighbors of u^A, v^A through the edges of A, i.e. $uu^A, vv^A \in A$. Note that $P'_L \cup \{uv\}$ does not contain a half-layer since there was no almost half-layer in P_L, and recall that $|A_L| \geqslant 6$. Thus we may extend P'_L on A'_L by some edges M_L including the edge uv to a perfect matching P^*_L of $B(Q_L - \{x, y'\})$ so that P^*_L contains no half-layer of Q_L. The choice of M_L can be done as in the previous cases except that the edge uv has to be kept in M_L. The rest is the same as in the subcase (a). Note that the edge $u^A v^A$ cannot appear in M_R, since M_L contains the edge uv and hence u, v cannot be the endvertices of the same path in $H_L - M_L$. Therefore it is guaranteed that $P^*_R = P'_R \cup M_R$ contains no half-layer as required.

(ii) *There are $k \geqslant 2$ dangerous directions.* We proceed similarly as in Subcase (i). Our aim is to first choose k edges for the matching M_L, one for each dangerous direction, so that we again "forbid" all possible half-layers later in P^*_R. We claim that

$$|P'_L| \leqslant 2^{n-k-2} - 1. \tag{1}$$

Notice that $|P'_L| = |P'_R|$. By Lemma 4, no edge of a y-dangerous half-layer in a dangerous direction is in P'_R. Furthermore, y-dangerous half-layers of Q_R in k dangerous directions cover (together) all but 2^{n-k-1} vertices of Q_R as they have the same parity. Since the matching P'_R can pair only those 2^{n-k-1} vertices and x', y are not paired, we obtain that $|P'_R| \leqslant 2^{n-k-2} - 1$. This establishes the claim (1).

Now we pick the edges to "forbid" dangerous half-layers. Let S be the following set of n independent edges between all the neighbors of y and distinct vertices at distance 2 from y:

$$S = \{y^i (y^i)^{i+1 \bmod n} \mid i \in [n]\}.$$

Clearly, every direction appears exactly once in S. Furthermore, if $uv \in S$ has a dangerous direction, then uv belongs to a y-dangerous half-layer and thus

$u, v \in A'_R$. For each such edge uv we choose the edge $u^A v^A$ for M_L where $u^A, v^A \in A'_L$ such that $uu^A, vv^A \in A$.

If we add these k edges chosen for M_L into P'_L, can P'_L contain a half-layer of Q_L? By claim (1), we would have at most $2^{n-k-2}-1+k$ edges, which is strictly less than 2^{n-3} as $k \geqslant 2$, the size of a half-layer in Q_L, so the answer is negative. Therefore, P'_L can be extended on A'_L by some edges M_L including the above already chosen edges to a perfect matching $P^*_L = P'_L \cup M_L$ of $B(Q_L - \{x, y'\})$ so that P^*_L contains no half-layer of Q_L. Similarly as before, the choice of other edges of M_L can be done arbitrarily and if $P'_L \cup M_L$ contains a half-layer, it can be corrected by switching the endvertices of two suitable edges. The rest is the same as before. □

5 Conclusions

In this section we discuss possible extensions. In both Theorems 2 and 3 we assume that $n \geqslant 5$. As for smaller dimensions, it is easy to check that they both hold also for $n = 3$. However, for $n = 4$ there are other particular exceptional configurations, see their list for Theorem 3 on Figures 7 and 8.

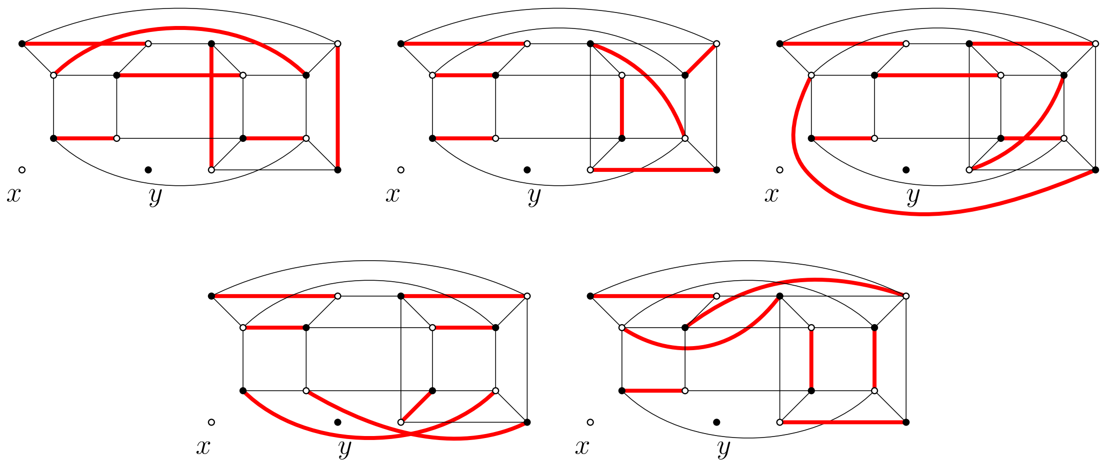


Figure 7: All (up to isomorphism) configurations of (thick red) perfect matchings of $B(Q_4 - \{x, y\})$ for $d(x, y) = 1$ that do not contain a half-layer but cannot be extended by (thin black) edges of $Q_4 - \{x, y\}$ to a Hamiltonian cycle.

It is worth mentioning that our induction step would allow us to prove a stronger version of Theorems 2 and 3 with the graph $B(Q_n)$ resp. $B(Q_n - \{x, y\})$ being replaced by $K(Q_n)$ resp. $K(Q_n - \{x, y\})$. Let us state it formally only for Theorem 3.

Conjecture 8. Let x, y be vertices of opposite parity in Q_n, $n \geqslant 5$, and let P be a perfect matching of $K(Q_n - \{x, y\})$. There exists a matching $R \subseteq E(Q_n)$ such that $P \cup R$ forms a Hamiltonian cycle of $K(Q_n - \{x, y\})$ if and only if P does not contain a half-layer.

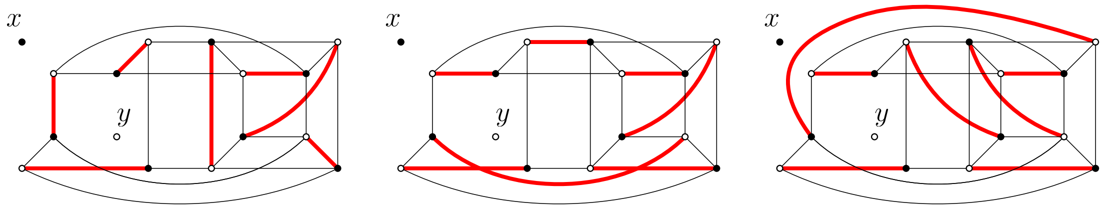


Figure 8: All (up to isomorphism) configurations of (thick red) perfect matchings of $B(Q_4 - \{x, y\})$ for $d(x, y) = 3$ that do not contain a half-layer but cannot be extended by (thin black) edges of $Q_4 - \{x, y\}$ to a Hamiltonian cycle.

The proof of the necessity in Conjecture 8 is the same as in our proof. However, for the sufficiency in Conjecture 8 we were not able to verify the case $n = 5$ even with the help of computer (the number of non-isomorphic configurations is unmanageable in this case.) Our proof in Section 4 thus only provides the following implication.

Theorem 9. *If Conjecture 8 holds for $n = 5$ then it holds for all $n \geqslant 5$.*

For the sake of completeness, we also note that for $n = 4$ there are two additional exceptional configurations for Conjecture 8, see Figure 9.

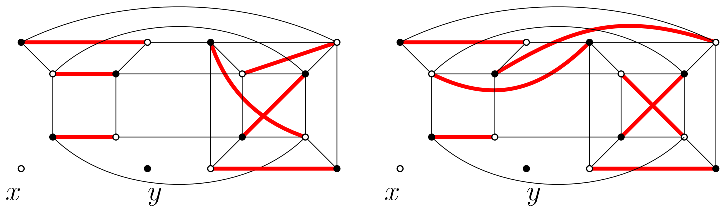


Figure 9: Two additional (up to isomorphism) configurations of (thick red) perfect matchings of $K(Q_4 - \{x, y\})$ that do not contain a half-layer but cannot be extended by (thin black) edges of $Q_4 - \{x, y\}$ to a Hamiltonian cycle.

Acknowledgements

The authors would like to thank Jernej Azarija for an earlier attempt to verify the case $n = 5$, Tomáš Dvořák for discussions regarding the verifying algorithm, and the referees for their helpful remarks. This research was supported by the Czech Science Foundation grant GA14-10799S and Slovenian research agency ARRS, program no. P1–0383.

References

[1] A. Alahmadi, R.E.L. Aldred, A. Alkenani, R. Hijazi, P. Solé, C. Thomassen, *Extending a perfect matching to a Hamiltonian cycle*, Discrete Math. Theor. Comput. Sci. 17 (2015), 241–254.

[2] D. Dimitrov, T. Dvořák, P. Gregor, R. Škrekovski, *Gray codes avoiding matchings*, Discrete Math. Theor. Comput. Sci. 11 (2009), 123–147.

[3] T. Dvořák, *Matchings of quadratic size extend to long cycles in hypercubes*, Discrete Math. Theor. Comput. Sci. 18 (2016), #12.

[4] T. Dvořák, *Hamiltonian cycles with prescribed edges in hypercubes*, SIAM J. Discrete Math. 19 (2005), 135–144.

[5] T. Dvořák, P. Gregor, *Hamiltonian paths with prescribed edges in hypercubes*, Discrete Math. 307 (2007), 1982–1998.

[6] T. Dvořák, P. Gregor, *Hamiltonian fault-tolerance of hypercubes*, in Proc. European Conference on Combinatorics, Graph Theory and Applications (EuroComb 2007), Electron. Notes in Discrete Math. 29 (2007), 471–477.

[7] T. Dvořák, J. Fink, *Gray codes extending quadratic matchings*, J. Graph Theory (2018). doi:10.1002/jgt.22371.

[8] J. Fink, *Perfect matchings extend to Hamilton cycles in hypercubes*, J. Combin. Theory Ser. B 97 (2007), 1074–1076.

[9] J. Fink, *Matchings extend into 2-factors in hypercubes*, Combinatorica (2018). doi:10.1007/s00493-017-3731-8.

[10] F. Gray, *Pulse Code Communication,* U.S. Patent 2,632,058, filed 13 November 1947, issued 17 March 1953.

[11] L. A. Gross, *Theorie du Baguenaudier*, Aimé Vingtrinier, Lyon, 1872.

[12] D. Knuth, *The Art of Computer Programming, Volume 4A*, Addison-Wesley, 2011.

[13] G. Kreweras, *Matchings and Hamiltonian cycles on hypercubes*, Bull. Inst. Combin. Appl. 16 (1996), 87–91.

[14] T. Mütze, *Proof of the middle levels conjecture*, Proc. Lond. Math. Soc. 112 (2016), 677–713.

[15] T. Novotný, https://github.com/novotnyt94/Hypothesis-checker, 2017.

[16] F. Ruskey, C. D. Savage, *Hamilton cycles that extend transposition matchings in Cayley graphs of S_n*, SIAM J. Discrete Math. 6 (1993), 152–166.

[17] C. Savage, *A survey of combinatorial Gray codes*, SIAM Rev. 39 (1997), 605–629.

[18] J. Vandenbussche, D. B. West, *Matching extendibility in hypercubes*, SIAM J. Discrete Math. 23 (2009), 1539–1547.

[19] F. Wang, H. Zhang, *Prescribed matchings extend to Hamiltonian cycles in hypercubes with faulty edges*, Discrete Math. 321 (2014), 35–44.

[20] E. Zulkoski, C. Bright, A. Heinle, I. Kotsireas, K. Czarnecki, and V. Ganesh, *Combining SAT Solvers with Computer Algebra Systems to Verify Combinatorial Conjectures*, J. Autom. Reasoning 58 (2017), 313–339.

Appendix A

In this appendix we describe the algorithm that verified the case $n = 5$ of Theorem 2, which serves us as the base of induction in our proof. The source code of the algorithm in C++ is available on [15].

We start with terminology required for the description of the algorithm. We say that two matchings in $B(Q_n)$ are *isomorphic* if there exists an automorphism of Q_n mapping one to the other. Furthermore, we say that an edge $uv \in B(Q_n)$ has *length* l if the distance of vertices $u, v \in V(Q_n)$ is equal to l. An edge is *short* if it lies in Q_n (i.e. its length is 1). Note that in this appendix we represent vertices of Q_n by corresponding integers $\{0, \dots, 2^n - 1\}$ via the standard binary encoding.

The algorithm is composed of two major parts: a generation of the set $\mathcal{P}$ of all non-isomorphic perfect matchings of $B(Q_n)$ and a search for Hamiltonian paths between x, y that extend P with some short edges, for every matching $P \in \mathcal{P}$ and every two vertices x, y of opposite parity with $xy \notin P$.

A.1 Generation of matchings

The generation is based on backtracking with cutting-off branches with already visited states or violating the condition on the number of short edges. First, we make following observations showing that pruning made in backtracking does not miss any perfect matching, up to isomorphism:

Observation 10. *If a perfect matching P contains at least one edge of length l then there exist an isomorphic perfect matching with the edge $\{0, 2^l - 1\}$.*

Observation 11. *When two perfect matchings differ in numbers of their short edges, they cannot be isomorphic. Therefore, generating non-isomorphic perfect matchings separately for all possible numbers (0 to 2^{n-1}) of short edges generates all of them exactly once.*

Observation 12. *When all possible perfect matchings $\mathcal{P}'$ obtainable from extending some non-perfect matching M into perfect were generated, then each perfect matching obtainable from extending any matching isomorphic to M is also isomorphic to some perfect matching in $\mathcal{P}'$.*

The backtracking is based on adding edges one-by-one into a partial matching to obtain all possible configurations. More precisely, in each level of recursion we find the first uncovered vertex, which is then connected to all uncovered vertices in the other partite sets and for each possibility (if it is not cut) the recursion continues with another step until all edges are selected. The resultant perfect matching is tested for isomorphism with previously generated perfect matchings and if it is new, it is added into the set of generated perfect matchings.

By Observation 10, we can set the first edge to be $\{0, 1\}$ if at least one of the edges is required to be short. If not (i.e. the required number of short edges is set to 0), then this case must be handled separately.

The diameter of Q_n is n, so each edge has length at most n. Therefore, if there exists an edge of length 3, we may set the first edge to be $\{0, 7\}$. For $n = 3$ and $n = 4$ there are no other possibilities. However, for $n = 5$ it is possible to have all edges of length 5. But then all edges are uniquely determined, so we may add that perfect matching separately.

Observation 11 allows to divide the program into $2^{n-1}+1$ independent cases depending on the number of short edges so we can run it for all of these cases in parallel.

A.1.1 Pruning non-perspective branches

Let us suppose that we want to add an edge uv into a partial matching. There are two simple cases when this branch can be skipped:

1. Adding the selected edge into the matching violates the requirement for the number of short edges.
2. Adding the selected edge creates a partial matching isomorphic to some matching generated earlier.

The first condition is clear, the second comes from Observation 12. In the program, the first condition is checked in two cases. If $uv \in E(Q_n)$, then we increase the number of used short edges by one. If this number extends the requirement, this edge cannot be used. Otherwise $uv \notin E(Q_n)$; then we compare the number of short edges still needed to be added and the number of available short edges on uncovered vertices. If there are not enough short edges left, we cannot fulfil the requirement.

The second condition is checked only partially, only when the number of used edges is not greater than 10. This relaxation must be done because of fast increasing memory requirements to store all visited states. Moreover, searching for the correct isomorphism is a relatively expensive operation, thus more frequent filtering can even be slower.

A.1.2 Isomorphism testing

The following characterization of the hypercube automorphisms is a well-known fact.

Observation 13. *Every automorphism of hypercube is composed of a unique transposition (i.e. switching certain coordinates) and a unique permutation of coordinates.*

For example, for $n = 5$ there are exactly $2^5 \cdot 5!$ different automorphisms of Q_n. Therefore, searching for an isomorphic matching among all generated matchings by trying all of the automorphisms would be too slow.

Definition 14. For a matching M, let the *signature* of M be the 2^n-dimensional vector $(a_0, a_1, \ldots, a_{2^n-1})$, where for every $i \in \{0, 1, \ldots, 2^n - 1\}$,

$$a_i = \begin{cases} \text{the index of the neighbor of } i \text{ in } M & \text{if } i \text{ is covered by } M, \\ 2^n & \text{otherwise.} \end{cases}$$

Definition 15. For the set $\mathcal{I}(M)$ of all matchings isomorphic to M we say that M_1 is *lexicographically smaller* than M_2, (where $M_1, M_2 \in \mathcal{I}(M)$) if the signature of M_1 is lexicographically smaller than the signature of M_2. Matching M_0 isomorphic to M is *lexicographically minimal* (denoted by $\operatorname{lexmin}(M)$), if it is lexicographically smaller than all $M' \in \mathcal{I}(M) \setminus \{M_0\}$.

Observation 16. *For every two matchings A, B, $\operatorname{lexmin}(A)$ and $\operatorname{lexmin}(B)$ exist, they are unique, and $\operatorname{lexmin}(A) = \operatorname{lexmin}(B)$ if and only if A and B are isomorphic.*

We use the Observation 16 to reduce the complexity of search. When we generate a possibly new matching M, we count $\operatorname{lexmin}(M)$ and then we try to find in the list of all so far generated matchings. If it is already there, we know that some matching isomorphic to M was generated earlier, otherwise M is new and then we insert $\operatorname{lexmin}(M)$ to the list.

We can also use Observation 13 to estimate the total number of non-isomorphic perfect matchings in $B(Q_n)$ for $n = 5$. There are exactly 16! different perfect matchings in $B(Q_n)$ and for every perfect matching P it holds $|\mathcal{I}(P)| \leqslant 2^5 \cdot 5!$. Therefore

$$|\{\mathcal{I}(P) : P \text{ is a perfect matching of } B(Q_n)\}| \geqslant \frac{16!}{2^5 \cdot 5!} = 5,448,643,200.$$

In fact, there are exactly 5,450,821,743 non-isomorphic perfect matchings, which is relatively larger only by factor 0.0004 compared to the estimate.

A.2 Testing matchings

All generated matchings are tested independently for existence of required paths, thus let us suppose we are testing a perfect matching P. For every non-connected pair of vertices of opposite parity, the Hamiltonian path between them that contains P is being searched using DFS with the most-constrained-first heuristics.

Suppose that we want to find a Hamiltonian path H between vertices x and y. We know that $P \subset H$. At first, we count for every vertex the directions in which a short edge may be added to this partial path without creating a loop. Then we select a vertex with the least possibilities and try adding incident edges one-by-one. After adding an edge, we update the list of possibilities and repeat the procedure until H is found or a vertex with no possibilities appears; then we try to use a different one on the previous level.

If we return to the beginning without finding the path, then the path does not exist. In that case, we need to check that the matching with these endvertices x, y does not satisfy the necessary condition in Theorem 2. If it is so (all cases for dimensions $n = 3$ and $n = 5$), no action has to be done. Otherwise (in the cases for dimension $n = 4$ corresponding to Figures 7 and 8), the violation is reported.