Maximizing Minimum Cycle Bases Intersection

Ylène Aboulfath[1], Dimitri Watel[2,3], Marc-Antoine Weisser[4], Thierry Mautor[1], and Dominique Barth[1]

[1] DAVID, Université Versailles Saint-Quentin-En-Yvelines, France
ylene.aboulfath@uvsq.fr, dominique.barth@uvsq.fr, thierry.mautor@uvsq.fr
[2] SAMOVAR, Evry, France
[3] ENSIIE, Evry, France dimitri.watel@ensiie.fr
[4] LISN, CentraleSupélec, France marc-antoine.weisser@centralesupelec.fr

Abstract. We address a specific case of the matroid intersection problem: given a set of graphs sharing the same set of vertices, select a minimum cycle basis for each graph to maximize the size of their intersection. We provide a comprehensive complexity analysis of this problem, which finds applications in chemoinformatics. We establish a complete partition of subcases based on intrinsic parameters: the number of graphs, the maximum degree of the graphs, and the size of the longest cycle in the minimum cycle bases. Additionally, we present results concerning the approximability and parameterized complexity of the problem.

Keywords: Minimum cycle basis, Matroids intersection problem, Complexity

1 Introduction

In chemoinformatics and bioinformatics, a molecular dynamics trajectory represents evolution of the 3D positions of atoms constituting a molecule, thus forming a sequence of molecular graphs derived from these positions at discrete time intervals. These graphs share the same vertex set (atoms) but vary in their edges, particularly those representing hydrogen bonds, which may appear or disappear over time, unlike covalent bonds, which are persistent. A research objective is to characterize the evolution of molecular structure during the trajectory [3,11].

In many studies, the structure of a molecule is intricately linked to the interactions among elementary cycles within its associated graph. Usually, only short cycles intervenne in the characterization of the molecule, therefore, the structure is commonly represented by a minimum cycle basis of the graph [5,6,8,9,14]. Cycle bases are a concise representation of cycles within an undirected graph, and finding a minimum cycle basis (*i.e.* minimizing the total weight of cycles in the base) can be done in polynomial time [2,7]. Thus, given a sequence of molecular graphs modeling the trajectory, to evaluate the conservation of the molecular structure during the trajectory [1], we seek to obtain a minimum cycle basis for each graph such that they have the most cycles in common.

In this context, we define the problem, referred to as MAX-MCBI, as follows: given a set of k graphs $\{G_1, G_2, \ldots, G_k\}$ with the same vertex set, find for each graph a minimum cycle basis such that the size of their overall intersection is maximum. Note that MAX-MCBI is a special case of the matroid intersection problem (MI) wherein, given k matroids with the same ground set C, we search for one independent set in each matroid such that the size of their intersection is maximum. This is primarily because the set of cycles in an undirected graph forms a vector space. MI is NP-Complete when $k = 3$ [15] but polynomial in $|C|$ when $k = 2$ [4] and $\frac{1}{k}$-approximable [10]. Transferring this positive results to MAX-MCBI can be achieved once we address the challenge posed by the potentially exponential number of cycles in a graph.

Our contributions. In this paper, we exploit the distinctive features of our specific instance of MI to establish a NP-Hard/Polynomial partition of subcases based on intrinsic parameters. Additionally, we investigate the parameterized complexity and approximability of the problem MAX-MCBI and its decision version MCBI. In the decision version, given k graphs and a non-negative integer K, the objective is to determine whether there exists a minimum cycle basis for each graph such that the size of their intersection is greater than K or not.

We distinguish four intrinsic parameters: the number of graphs k; the maximum size γ of the cycles in a minimum cycle basis of any graph G_i; the maximum degree Δ in any graph G_i; and the decision integer K. The first parameter k directly arises from the complexity of MI which is polynomial for $k = 2$ and $NP-complete$ otherwise. Parameters γ and Δ arise from our application. Those are classical parameters studied in molecular contexts. Finally, K is a classical parameter in parameterized complexity.

The results are summarized in Table 1. Note that the case $\Delta = 2$ is trivially polynomial as each graph is a set of disjoint cycles. Considering the parameterized complexity and the approximability, few questions remain open.

Section 2 and 3 are dedicated to formal definitions and the proof that MAX-MCBI is indeed a subproblem of MI. As a result we get the $\frac{1}{k}$- approximation algorithm, the polynomial case when $k = 2$ and the belonging of MCBI to XP with respect to K. In Section 4, we prove the hardness results. Finally Section 5 is dedicated to the cases where $\gamma = 3$ or where $\gamma = 4$ and $\Delta = 3$.

2 Formal Definitions

We adopt a standard definition of a cycle in a graph G as any subgraph in which each vertex has an even degree[5] [7]. The sum of two cycles, denoted as $c_1 \oplus c_2$, is the subgraph containing the set of edges present in one and only one of the two cycles. In this paper, when D represents a set of cycles, we denote the sum of cycles as $\bigoplus D = \bigoplus_{d \in D} d$. Note also that if $\bigoplus D = c$ then $\bigoplus D \oplus c = 0$. This general definition of cycles with the sum operation $\oplus$ defines a vector space in the field $\mathbb{Z}/2\mathbb{Z}$. A cycle basis of a graph G is a linearly independent set B of cycles that spans the cycle space of G. The terms *span* and *linearly independent* refer to the classical linear algebra definitions.

[5] It's important to note that a cycle can then be composed of elementary cycles, which may seem counter-intuitive at first.

Table 1: Summary of the results of this paper. The hardness results hold true only when the parameters in the *blank cells* are not fixed and remain valid even when the parameters in the *non-blank cells* are fixed to the specified values (or higher). The polynomial results hold true only when the parameters in the *non-blank cells* are fixed to the specified value (or lower) and remain valid regardless of the values of the parameters in the *blank cells*. Additionally, parameters in the parameterized results are indicated with a cross.

k	Δ	γ	K	MCBI	MAX-MCBI	
3	4 3	4 5		NP-Complete	NP-Hard	
	4 3	4 5	-	-	$\frac{1}{k}$ Inapprox.	Theorem 3
	4 3	4 5	$\times$	W[1]-Hard	-	
	2			P	P	Section 1
		3		P	P	Theorem 4
	3	4		P	P	Theorem 5
2				P	P	
			-	-	$\frac{1}{k}$ Approx.	Theorem 2
			$\times$	XP		

– B is linearly independent if, for all $B' \subseteq B$, $\bigoplus B' \neq 0$.
– B spans a cycle d if there exists $B' \subseteq B$ such that $\bigoplus B' = d$. If B is a basis, then the subset B' is unique for each cycle d.

We define $\lambda_B : B \times \mathcal{C} \to \{0, 1\}$ as the function such that if B' is the subset of B with $\bigoplus B' = d$, then $c \in B'$ if and only if $\lambda_B(c, d) = 1$.

The weight of a cycle is defined as its number of edges, denoted for a cycle c by $\omega(c)$. The weight of a cycle basis B is given by $\sum_{c \in B} \omega(c)$. Therefore, a minimum cycle basis is a cycle basis that minimizes its weight. The set of minimum cycle bases of a graph G is denoted by $\mathcal{MCB}(G)$. Polynomial time algorithms for finding a minimum cycle basis have been proposed [2,7].

Problem 1 (MAX-MCBI). Given a set of k graphs $G_1, G_2, \ldots, G_k$, find a subset of cycles B of $\bigcap_{i=1}^{k} G_i$ such that, for all $i \in [\![1; k]\!]$, there exists $B_i \in \mathcal{MCB}(G_i)$ with $B \subseteq B_i$, and maximizing $|B|$.

The decision problem MCBI associated with MAX-MCBI is, given an integer K, to determine if there exists a solution with $|B| \geq K$.

The rest of this section proves that the search for B can be performed within a polynomial-size subset of cycles. We begin by presenting common lemmas on minimum cycle bases. Their proofs are provided in the Appendix A.

Lemma 1. *Given B a cycle basis of a graph G with two cycles $c_1 \in B$ and $c_2 \notin B$, if $\lambda_B(c_1, c_2) = 1$ then $(B \setminus \{c_1\}) \cup \{c_2\}$ is a cycle basis of G.*

Lemma 2. *$B \in \mathcal{MCB}(G)$ if and only if B is a cycle basis and for c_1, c_2 with $c_1 \in B$ and $c_2 \notin B$ such that $\lambda_B(c_1, c_2) = 1$, we have $\omega(c_1) \leq \omega(c_2)$.*

Lemma 3. *If $B_1, B_2 \in \mathcal{MCB}(G)$, for every $c_1 \in B_1 \backslash B_2$, there exists $c_2 \in B_2 \backslash B_1$ such that $(B_1 \backslash \{c_1\}) \cup \{c_2\} \in \mathcal{MCB}(G)$.*

Given G, let $\mathcal{M}(G) = (C, I)$ be the couple where C is the set of cycles of G and I are the subsets D of C such that there exists $B \in \mathcal{MCB}(G)$ with $D \subseteq B$.

Lemma 4. *$\mathcal{M}(G)$ is a matroid.*

Proof. Lemma 3 proves the basis exchange axiom. As $\mathcal{MCB}(G)$ is not empty (it possibly contains the empty set of G is a tree), then $\mathcal{M}(G)$ satisfies the basis axioms and is then a matroid. □

Note that MAX-MCBI can be simply rewritten as the search for a maximum-size set of cycles that are independent in each matroid $\mathcal{M}(G_i)$ using the matroid terminology for *independent*. In order to avoid confusion with the *linear independency*, whenever referring to linear algebra independence, we will explicitly use that terminology. This proves that MAX-MCBI is indeed a subproblem of the matroid intersection problem (MI). However, we cannot use algorithms dedicated to MI to prove any polynomial complexity result as, currently, the ground sets of our matroids have exponential size. We address this in the next section; however, we introduce a polynomial time independency oracle for $\mathcal{M}(G)$.

Lemma 5. *Given a subset D of cycles of G, we can check in polynomial time if D is independent in $\mathcal{M}(G)$.*

Proof. This can be achieved by running a modified version of the Horton algorithm [7]. Given a graph $G = (V, E)$, the Horton algorithm generates a minimum cycle basis by enumerating a list L of $O(|V||E|)$ cycles, then sorting L from the smallest cycles to the largest and finally using a greedy polynomial-time procedure to build a minimum cycle basis. To adapt this algorithm, we introduce a modification. Before the sorting step, we replace L with $D \cup L$ and during the sorting process, in case of a tie, cycles from D are given priority. The set D is independent if and only if all cycles of D are in the resulting basis. □

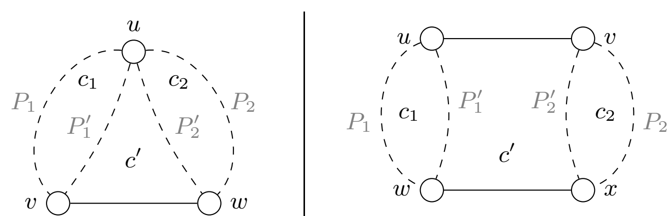
3 Case with $k = 2$, Approximability and Parameterized Complexity

With Theorem 1, we prove we may only focus on a polynomial subset of cycles.

Theorem 1. *Given an instance $\{G_1, G_2, \ldots, G_k\}$ of MAX-MCBI and let L be the list returned by the function* CANDIDATESLIST$(G_1, G_2, \ldots, G_k)$ *described in Algorithm 1, there exists an optimal solution B^* such that $B^* \subseteq L$.*

Algorithm 1 Building the list of candidate cycles containing an optimal solution of MAX-MCBI

1: **function** CandidatesList($G_1, G_2, \dots, G_k$)
2: $\quad L \leftarrow \emptyset$
3: $\quad G = (V, E) \leftarrow \bigcap_{i=1}^{k} G_i$
4: $\quad$ **for** $u \in V$, $(v, w) \in E$ **do** add to L the cycle consisting in the edge (v, w), a shortest path from u to v and a shortest path from u to w in G, if such a cycle is elementary.
5: $\quad$ **for** $(u, v) \in E$, $(w, x) \in E$ **do** add to L two cycles consisting in (u, v), (w, x), one with the shortest paths from u to w and from v to x in G, and a second with the shortest paths from u to x and from v to w in G, if such a cycle is elementary.
6: $\quad$ **return** L



(a) c an odd cycle $\quad$ (b) c an even cycle

Fig. 1: Example of cycles such that $c_1 \oplus c' \oplus c_2 = c$.

Proof. Let B^* be any optimal solution and B_i a minimum cycle basis of G_i containing B^* for every $i \in [\![1; k]\!]$.

Let us suppose that there exists a cycle $c \in B^* \backslash L$. Hereinafter, we prove that c can be replaced by a cycle $c' \in L \backslash B^*$ such that $B^* \backslash \{c\} \cup \{c'\}$ is still optimal.

Assuming c is an odd cycle containing an edge (v, w). Let P_1 and P_2 be the two paths included in c connecting respectively v and w to the same node u such that $|P_1| = |P_2|$. At line 4 of Algorithm 1, when the loop enumerates u and (v, w), we obtain a cycle c' that is added to L if it is elementary. As depicted by Figure 1a, there exist two (possibly non-elementary) cycles c_1 and c_2 such that $c = c_1 \oplus c' \oplus c_2$. As P'_1 and P'_2 are shortest paths and as $|P_1| = |P_2| = (|c| - 1)/2$, we have $\omega(c_1) < \omega(c), \omega(c') \leq \omega(c)$ and $\omega(c_2) < \omega(c)$. If c' is not elementary, then c is the sum of strictly smaller cycles, this contradicts Lemma 2. Consequently, c' is elementary and is added to L. For the same reason, $\omega(c') = \omega(c)$.

For all $i \in [\![1; k]\!]$, let B_i be a minimum cycle basis of G_i containing B^*. By Lemma 2, every d such that $\lambda_{B_i}(d, c_1) = 1$ has a weight $\omega(d) \leq \omega(c_1)$. Similarly for c_2. Thus those cycles cannot be c' and c. As $c' = \bigoplus\{d : \lambda_{B_i}(d, c_1)\} \oplus \bigoplus\{d : \lambda_{B_i}(d, c_2)\} \oplus c$, then, $c' \notin B_i$, otherwise B_i is not linearly independent.

Consequently, $\omega(c) = \omega(c')$ and $\lambda_{B_i}(c, c') = 1$ for all basis B_i. We can then replace c by c' in B_i by Lemma 1. The same property occurs for even cycles (see Figure 1b). This operation can be repeated until $B^* \subseteq L$. As the size of B^* is unchanged, we get another optimal solution. $\square$

MAX-MCBI can now be seen as a subproblem of the matroid intersection problem, MI. We deduce the following theorem.

Theorem 2. *MCBI and* MAX*-MCBI are polynomial when* $k = 2$*,* MAX*-MCBI is* $\frac{1}{k}$*-approximable and, finally MCBI is XP with respect to* K*.*

Proof. MAX-MCBI consists in solving MI in $\{\mathcal{M}(G_i), i \in [\![1; k]\!]\}$. By Theorem 1, we can restrict each matroid to L, the cycles output by Algorithm 1. Let $\mathcal{M}(G_i)_{|L}$ be the resulting restricted matroid. By Lemma 5 and Theorem 1 $\mathcal{M}(G_i)_{|L}$ is a matroid with a polynomial size ground set and a polynomial time independence oracle. In that case MI is polynomial when $k = 2$ [4] and is $\frac{1}{k}$-approximable [10]. Finally if K is fixed, we can simply enumerate every subset of L of size K and check if that subset is independent in all the matroids. □

4 Hardness of MCBI

This section gives the hardness proofs for MAX-MCBI and MCBI. The latter is in NP as by Lemma 5 we can check independence in polynomial time.

Theorem 3. *MCBI is NP-Complete even if* $k = 3$*. In addition, MCBI is W[1]-Hard with respect to* K*. Moreover, unless P = NP, for every* $\varepsilon > 0$*, there is no polynomial approximation algorithm with ratio* $\frac{1}{k^{1-\varepsilon}}$ *for* MAX*-MCBI. All those results remain true even if* $\Delta = 3$ *and* $\gamma = 5$ *or if* $\Delta = 4$ *and* $\gamma = 4$*.*

Proof. We provide a reduction from the *Maximum Independent Set* in a graph that consists, given a graph $H = (V, E)$ and an integer K', in the search for an independent set of H of size at least K', that is a subset in which no pair of nodes is linked by an edge in E. In order to avoid confusions with the word *independent*, we then use the *Stable set* terminology instead of Independent set.

The reduction does not depend on the value of γ except for a useful procedure described hereinafter and in Figure 2. Let $c_1 = (u_1, u_2, \ldots, u_l)$ and $c_2 = (v_1, v_2, \ldots, v_l)$ be two disjoint cycles where $l \in \{4, 5\}$. The procedure $CONN(c_1, c_2)$ connects the nodes of the two cycles. We add the edges (u_i, v_i) for $i \in [\![1; l]\!]$ and, if $l = 4$, we add the edges $(u_1, v_2), (u_2, v_3), (u_3, v_4)$ and (u_4, v_1). Note that, by Lemma 2, after using the procedure, no more minimum cycle basis containing c_1 and c_2 at the same time.

We now describe a general reduction with more than 3 graphs and then show how to reduce k. Let then $(H = (V, E), K')$ be an instance of the maximum Stable set problem. We create an instance of MCBI as follows. We set l either to 4 or 5. We set $K = K'$. For each edge $e \in E$, we build a graph G_e. For each node $v \in V$, we add a cycle c_v of size l to all the graphs. All the cycles are disjoint. Finally, if $e = (u, v)$, then, in the graph G_e, we connect c_u and c_v using the procedure $CONN(c_u, c_v)$. Every other cycle of G_e is not connected to the rest of the graph and is its own connected component.

Note that $l = \gamma$ and if $\gamma = 4$, then $\Delta = 4$ and if $\gamma = 5$ then $\Delta = 3$. Note also that only the cycles $\{c_v | v \in V\}$ belong to the intersection of the graphs $\{G_e | e \in E\}$. A feasible solution contains only those cycles.

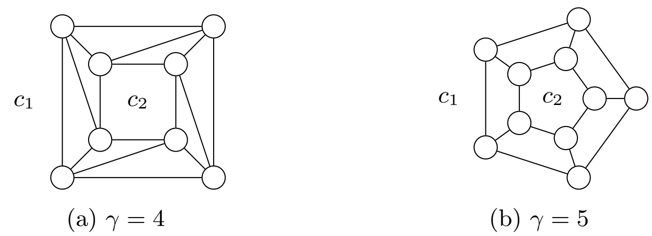


Fig. 2: Illustration of the procedure $CONN(c_1, c_2)$

Given an edge $e = (u, w) \in E$, the graph G_e contains exactly two minimum cycles bases. If $\gamma = 4$ (resp. 5), let D be the set of triangles (resp. squares) connecting c_u and c_w. Then $\mathcal{MCB}(G_e) = \{D \cup \{c_v | v \in V \backslash \{w\}\}, D \cup \{c_v | v \in V \backslash \{u\}\}\}$. As a consequence, for every subset of nodes $V' \subset V$, the set $\{c_v | v \in V'\}$ is independent in G_e if and only if $u \notin V'$ or $w \notin V'$. Thus, there exists a feasible set of cycles of size K if and only if there exists a stable set of size K in H. As the Maximum Stable Set is NP-Complete and W[1]-Hard with respect to K', this reduction proves the NP-Completeness of MCBI and the W[1]-Hardness with respect to K.

Now note that, if $e = (u, v)$ and $f = (u', v')$ are two non-incident edges, then instead of G_e and G_f we can add to the instance a graph $G_{(e,f)}$ containing the union of G_e and G_f. As the two edges are not incident, we have that a feasible set of cycles cannot contain c_u and c_v at the same time and the same for $c_{u'}$ and $c_{v'}$. Note that this transformation does not change the values of γ and Δ. We can extend this idea to any matching M of H. By the Vizing Theorem [12], using a polynomial greedy algorithm, edges of H can be covered with $|V|$ disjoint matchings. This reduces the number k of graphs from $|E|$ to $|V|$. By [16], unless P = NP, for all $\varepsilon > 0$, there is no polynomial approximation with ratio $\frac{1}{|V|^{1-\varepsilon}}$ for the Maximum Stable Set problem. This proves the inapproximability result.

The Maximum Stable Set remains NP-Complete even if H has degree at most 3 and if any path linking two nodes with degree 3 contains at least 3 edges [13]. Such a graph is always 3-edge-colorable. This proves the NP-Completeness of MCBI when $k = 3$. $\square$

5 Cases $\gamma = 3$, and $\gamma = 4, \Delta = 3$

We describe two useful lemmas to prove that we can focus on the cycles size by size independently. In this section, we call L the list returned by Algorithm 1, and we denote by $T(G)$ and $S(G)$ the triangles and squares of a graph G.

Lemma 6. *If $B_1, B_2 \in \mathcal{MCB}(G)$, for every $l \in \mathbb{N}$, $\{c \in B_1 | \omega(c) \neq l\} \cup \{c \in B_2 | \omega(c) = l\} \in \mathcal{MCB}(G)$.*

Proof. Using Lemma 3, for every $c \in B_1 \backslash B_2$, there is some $d \in B_2 \backslash B_1$ such that $B_1 \backslash \{c\} \cup \{d\} \in \mathcal{MCB}(G)$. Note that $\omega(c) = \omega(d)$. We can then swap the cycles until all the cycles of B with size l are in B_2. After the exchanges, $|\{c \in B|\omega(c) = l\}| = |\{c \in B_2|\omega(c) = l\}|$, otherwise, the basis exchange axiom would be false for B and B_2. Then $B = \{c \in B_1|\omega(c) \neq l\} \cup \{c \in B_2|\omega(c) = l\}$. □

Using Lemma 6 we see that all the subset of cycles of same size in two bases are interchangeable. This means that, from two feasible solutions B_3 and B_4 respectively maximizing the number of triangles and squares, we can build a feasible solution maximizing both of them. We define with $\mathcal{M}(G, l)$ the matroid $\mathcal{M}(G)$ retricted to the cycles of size l in L. Maximizing the triangles (respectively squares) consists in finding a maximum set of cycles that are independent in $\mathcal{M}(G, 3)$ (resp. $\mathcal{M}(G, 4)$). The following lemma gives a characterization of the independent sets. Due to lack of space, the proof may be found in Appendix A.

Lemma 7. *B is independent in $\mathcal{M}(G, l)$ if and only if, for all $D \subseteq B$, $\bigoplus D \notin span(cycle\ c|\omega(c) \leq l - 1)$.*

Theorem 4. *Given an instance of* MAX*-MCBI, one can find a feasible solution maximizing the number of triangles in polynomial time. As a consequence, MCBI and* MAX*-MCBI are polynomial when $\gamma = 3$.*

Proof. Because we work with simple graphs, there are no cycles of size 2. By Lemma 7, any set of linearly independent triangles of L is independent in $\mathcal{M}(G_i, 3)$ for all i. We can then start with an empty solution B and add each cycle $c \in T(G) \cap L$ to B if $c \notin span(B)$. □

Now we consider the case for $\gamma = 4$ and $\Delta = 3$. An instance where $\gamma = 4$ has triangles and squares in L. Contrary to the previous case, we can have two squares that are linearly independent but not together independent in $\mathcal{M}(G_i, 4)$ for all i. For instance one of the squares may contain a diagonal in G_i or the sum of the squares may belong to the span of the triangles like in Figure 2a.

Interestingly, the algorithm we use is almost the same as for $\gamma = 3$. It is described in the proof of Theorem 5.

We consider that L do not contain any square that is spanned by triangles of $T(G)$ as such a square cannot belong to any independent set of $\mathcal{M}(G_i, 4)$. Such cycles can be removed from L in polynomial time.

In the following lemmas, we use the matroid terminology of *circuit*. By Lemma 7, a circuit of $\mathcal{M}(G, 4)$ is a subset $C \subseteq S(G)$ such that $\bigoplus C \in span(T(G))$ but, for all subsets C' of C, we have $\bigoplus C' \notin span(T(G))$. We now caracterize the circuits of the graphs when $\Delta = 3$.

Lemma 8. *Let $G \in \{G_i, i \leq k\}$ with $\Delta = 3$ and $s_1 = (a_1, b_1, c_1, d_1) \in L$, $s_2 = (a_2, b_2, c_2, d_2) \in L$, with $s_1 \neq s_2$ and $t_3 = (a_3, b_3, c_3) \in T(G)$. Figure 3 gives the possible intersections of s_1 and s_2, and of s_1 and t_3, up to an isomorphism.*

Proof. Recall that we removed from L the squares generated by triangles of $T(G)$: they do not contain a chord. For the squares intersection, if s_1 and s_2 have only one common node, that node has degree 4 and 4 common nodes imply

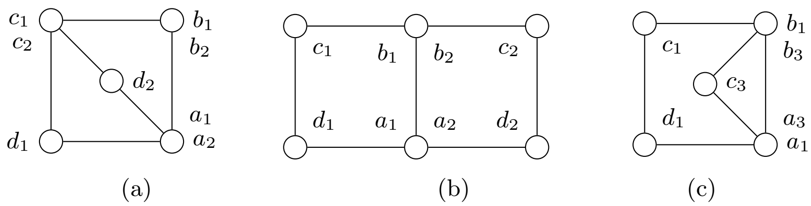


Fig. 3: Possible cases of intersection of squares and of triangles.

two diagonals. Thus they have three common nodes (Figure 3a) or two common nodes without diagonal (Figure 3b). Similarly, if s_1 and t_3 have 1 or 3 common nodes, there is a contradiction. Without diagonal, this leads to Figure 3c. $\square$

Lemma 9. *Let $G \in \{G_i, i \leq k\}$ with $\Delta = 3$ and $C \subseteq L$ be a circuit of $\mathcal{M}(G_i, 4)$. Then $\bigcup C$ is a connected subgraph of G.*

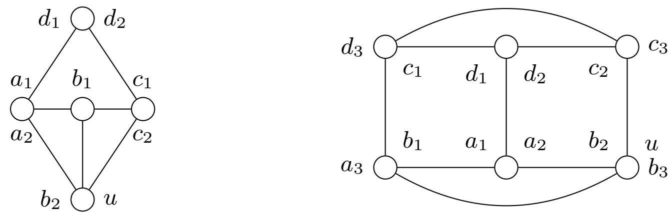
Proof. As C is a circuit of $\mathcal{M}(G_i, 4)$, by Lemma 7, there exists $T \subseteq T(G)$ such that $\bigoplus T = \bigoplus C$. We assume that T is minimal, meaning that for $T' \subsetneq T$, $\bigoplus T \neq 0$. We now show that $\bigcup C \cup \bigcup T$ is a connected graph. Indeed, otherwise, let assume there exists a connected component G' of $\bigcup C \cup \bigcup T$, and let C' and T' be respectively the proper subets of C and T in G'. No cycle of $C \backslash C'$ and $T \backslash T'$ intersects the edges of $\bigcup C' \cup \bigcup T'$. As $\bigoplus T = \bigoplus C$ then $\bigoplus T' = \bigoplus C'$. Consequently either C' is empty in which case $\bigoplus T' = 0$ or C' is a dependent proper subset of C. The first case contradicts the minimality of T and the second one contradicts the fact that C is a circuit. The only left possibility is that $T' = C' = \emptyset$ meaning that $\bigcup C \cup \bigcup T$ is connected.

Now we assume that $\bigcup C$ is disconnected. In $\bigcup C \cup \bigcup T$, any two connected components of $\bigcup C$ are connected by a chain P included in $\bigcup T$. Let $s_1 = (a_1, b_1, c_1, d_1)$ and $s_2 = (a_2, b_2, c_2, d_2)$ be two squares of C respectively in the first and second component linked by P. Let $w_1, w_2, \ldots, w_q$ be the nodes of P, with $w_1 \in s_1$ and $w_q \in s_2$. Without loss of generality, we state that $w_1 = a_1$ and $w_q = a_2$. Note that we cannot have $q = 1$ as the squares do not intersect.

Each edge (w_i, w_{i+1}) belongs to a triangle t_i. As $w_1 = a_1$ has 3 neighbors, b_1, d_1 and w_2, then t_1 is either (w_1, w_2, b_1) or (w_1, w_2, d_1). We assume, wlog, that it is (w_1, w_2, b_1). As a result, $q \neq 2$, otherwise, $w_2 = a_2$ has four neighbors (a_1, b_1, b_2 and d_2). Thus $q \geq 3$. As w_2 has already 3 neighbors, w_3, a_1 and b_1, the nodes of t_2 are w_2, w_3 and either a_1 or b_1. This means that w_3 is connected to that node, and then the degree of a_1 or b_1 is 4. We then have a contradiction. $\square$

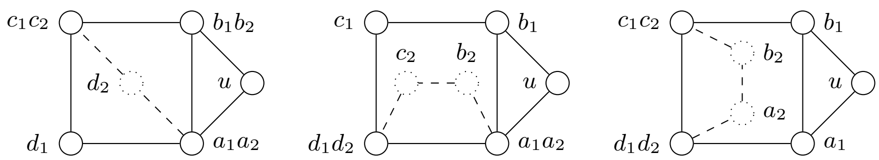
Lemma 10. *Let $G \in \{G_i, i \leq k\}$ with $\Delta = 3$. Then Figure 4 gives the possible circuits C of G such that $\bigoplus C \neq 0$.*

Proof. As $\bigoplus C \neq 0$, then, there exists $T \subseteq T(G)$ with $T \neq \emptyset$ and $\bigoplus C = \bigoplus T$. There exists at least one triangle t that intersects a square s_1 of C. By Lemma 8, that intersection is the graph depicted by Figure 3c. We set (a_1, b_1, c_1, d_1) and



(a) Circuit with two squares. (b) Circuit with three squares.

Fig. 4: Possible circuits containing squares in a graph G with degree at most 3. Each square contains four node (a_i, b_i, c_i, d_i).



(a) (b) (c)

Fig. 5: Illustrations for the proof of Lemma 10

$t = (a_1, b_1, u)$. As we removed from L the squares generated by triangles, $|C| > 1$. By Lemma 9, there exists another square s_2 that intersects s_1. Let $s_2 = (a_2, b_2, c_2, d_2)$. In the following, we rename the nodes of s_2 so that if a_1 (resp. b_1, c_1, d_1) is in $s_1 \cap s_2$, then $a_1 = a_2$ (resp. $b_1 = b_2$, $c_1 = c_2$, $d_1 = d_2$).

- If s_1 and s_2 have three common nodes as in Figure 3a, then:
  - $s_1 \cap s_2 = \{a_1 = a_2, b_1 = b_2, c_1 = c_2\}$ and $d_1 \neq d_2$ (see Figure 5a). Then u, b_1, d_1 and d_2 are neighbors of a_1, then two of those nodes are equal. Otherwise the degree of a_1 is 4. By hypothesis, $d_1 \neq d_2$. We cannot have $d_1 = b_1$ as s_1 would not be a square, similarly $b_1 \neq u$. And $d_1 \neq u$ otherwise s_1 would contain a diagonal. The only possibilities are $d_2 = u$ or $d_2 = b_1$. In the first case s_2 contains a diagonal (b_2, d_2). In the second case, $d_2 = b_2$. Thus there is a contradiction.
  - Or $s_1 \cap s_2 = \{a_1 = a_2, c_1 = c_2, d_1 = d_2\}$ and $b_1 \neq b_2$. Similarly we can deduce that $b_2 = u$. We obtain the graph of Figure 4a.
  - The two other intersections are symmetrical cases.
- If s_1 and s_2 have two common nodes as in Figure 3b, then:
  - $s_1 \cap s_2 = \{a_1 = a_2, b_1 = b_2\}$ and $c_1 \neq c_2, d_1 \neq d_2$. Then $c_2 = u$ and $d_2 = u$, otherwise the degree of a_1 or b_1 is 4. But c_2 cannot equal d_2.
  - Or $s_1 \cap s_2 = \{a_1 = a_2, d_1 = d_2\}$ and $b_1 \neq b_2, c_1 \neq c_2$ (see Figure 5b). Then $b_2 = u$ otherwise the degree of a_1 is 4, and $c_2 \neq u$.
    The edge $(b_2 = u, c_2)$ necessarily belongs either to a triangle or to another square of C, otherwise, we cannot have $\bigoplus T + \bigoplus C = 0$. Assuming it belongs to a triangle t'. As u has 3 neighbors, a_1, b_1 and c_2, the third

node of t' is necessarily a_1 or b_1. This implies that a_1 or b_1 are linked to c_2. This is a contradiction as that nodes has then 4 neighbors.

Consequently, the edge $(b_2 = u, c_2)$ belongs to a third square $s_3 = (a_3, b_3 = b_2 = u, c_3 = c_2, d_3)$. Again, as u has already 3 neighbors, we have that $a_3 \in \{a_1 = a_2, b_1\}$. If $a_3 = a_1 = a_2$ then d_3 must be an existing neighbor of a_1, that is b_1 or $d_1 = d_2$. It cannot be d_2 otherwise $s_2 = s_3$. Then $b_1 = d_2$, b_1 has four neighbors and there is a contradiction. If, on the other hand, $a_3 = b_1$ then d_3 must be an existing neighbor of b_1, that is a_1 or c_1. In the first case, a_1 has four neighbors. In the second case, we obtain the graph of Figure 4b.

- The case $s_1 \cap s_2 = \{b_1, c_1\}$ is symmetrical.
- Or $s_1 \cap s_2 = \{c_1 = c_2, d_1 = d_2\}$ and $a_1 \neq a_2, b_1 \neq b_2$ (see Figure 5c). The node a_2 cannot equal a_1 or b_1. It is then either u or an additional node. The same property occurs for b_2. If $a_2 = u$ then u has four neighbors a_1, b_1, $d_1 = d_2$ and b_2. As b_2 cannot equal any of the three first nodes, there is a contradiction. Similarly, there is a contradiction if $b_2 = u$.

  The edge (a_1, d_1) necessarily belongs to either a triangle or another square of C, otherwise, we cannot have $\bigoplus T + \bigoplus C = 0$. It cannot belong to a triangle as the two nodes have three neighbors and no common neighbor. There is then a third square $s_3 = (a_3 = a_1, b_3, c_3, d_3 = d_1)$ containing that edge. Then b_3 is a neighbor of a_1 and c_3 is a neighbor of d_3. As a consequence $b_3 = u$ and $c_3 = a_2$. Thus u and a_2 are neighbors. Similarly, the edge (b_1, c_1) belongs to a fourth square and u and b_2 are neighbors, and then u has four neighbors. This is a contradiction.

Note finally that no square can be added to extend the circuits of Figure 4 as, by definition, C is a minimal dependent set of cycles. □

Theorem 5. *MCBI and* MAX*-MCBI are polynomial when* $\gamma = 4$ *and* $\Delta = 3$.

Proof. We use the following algorithm. First, we compute the squares of the list L with Algorithm 1. We then remove from L any square that is generated by $T(G_i)$ for some $i \in [\![1;k]\!]$. Then, we initialize an empty solution B and, for each cycle $c \in L$, add c to B if $c \notin \bigcup_{i=1}^{k} span(B \cup T(G_i))$. We finally return B.

Let B^* be an optimal solution and B be the solution resulting from the algorithm. Let $s_{(j)}$ be the j-th added cycle of L and let also $B_{(j)}$ be the set B at the beginning of the j-th iteration of the algorithm (before adding $s_{(j)}$). Let α be the first index where $s_{(\alpha)} \in B^* \backslash B$ or $s_{(\alpha)} \in B \backslash B^*$. We assume that, among all the optimal solutions, B^* is the solution maximizing the index α.

If $s_{(\alpha)} \in B^* \backslash B$, then at the α-th iteration of the algorithm, $s_{(\alpha)}$ is not added to B, meaning that $s_{(\alpha)} \in span(B \cup T(G_i))$ for some i, that is $\{s_{(i)}\} \cup B_{(i)}$ is not independent in $\mathcal{M}(G_i, 4)$. Note that, by definition of α, $B_{(\alpha)} \subseteq B^*$. And because $s_{(\alpha)} \in B^*$ then $\{s_{(\alpha)}\} \cup B_{(\alpha)} \subseteq B^*$, which is a contradiction.

If, on the other hand, $s_{(\alpha)} \in B \backslash B^*$, then $B^* \cup \{s_{(\alpha)}\}$ contains a circuit C in $\mathcal{M}(G_{i_C}, 4)$, with $s_{(\alpha)} \in C$, for some $i_C \in [\![1;k]\!]$. We first make the assumption that C is dependent in all the graphs. Let $s \in C \cap B^*$.

– Either the set $B(s) = B^*\backslash\{s\} \cup \{s_{(\alpha)}\}$ is feasible.
– Or for some $i \in [\![1;k]\!]$, $B^*\backslash\{s\} \cup \{s_{(\alpha)}\}$ is dependent in $\mathcal{M}(G_i, 4)$. In that last case, there exists a circuit C' of $\mathcal{M}(G_i, 4)$ with $C' \subset B(s)$ and $s_{(\alpha)} \in C'$. As $s \in C$ and $s \notin C'$ then $C \neq C'$; and thus $(C \cup C')\backslash(C \cap C')$ is dependent in $\mathcal{M}(G_i, 4)$. However $(C \cup C')\backslash(C \cap C') \subset B^*$ and this is a contradiction.

Therefore $B(s)$ is feasible for every $s \in C \cap B^*$. Note that $B(s)$ is also optimal. Let $j_1 < j_2 < \cdots < j_{|C|}$ be the $|C|$ indices of the squares of C in L (one of them is α). By definition of α, for every $j \leq \alpha$, $s_{(j)} \in B$. As B is independent, $C \not\subseteq B$: there exists a square $s_{(j)}$, with $j \in [\![\alpha + 1; j_{|C|}]\!]$ such that $s_{(j)} \notin B$. The set $B(s_{(j)})$ is optimal and then contradicts the maximization of α by B^*.

Consequently, C is not dependent in all the graphs. Let T be the set of triangles of G_{i_C} such that $\bigoplus C = \bigoplus T$. There exists a graph G_i such that $\bigcup T \not\subseteq G_i$. Indeed, if we assume the contrary, we get that C is dependent in all graphs. As a consequence, first $T \neq \emptyset$ and at least one edge in $\bigcup T$ is not part of any square in C. This implies that C is not the circuit depicted by Figure 4b.

By Lemma 10, C is the circuit depicted by Figure 4a. It contains $s_{(\alpha)}$ and another square s_ϵ. Let $s_{(\alpha)} = (a_1, b_1, c_1, d_1)$ and $s_\epsilon = (a_2 = a_1, b_2 \neq b_1, c_2 = c_1, d_2 = d_1)$. We now prove that, in each graph $G \in \{G_1, G_2, \ldots, G_k\}$, if C is not a circuit in $\mathcal{M}(G, 4)$ then $= B^* \cup \{s_{(\alpha)}\}$ is independent in that matroid. Indeed, we have that $(b_1, b_2) \notin G$. Also, if $B^* \cup \{s_{(\alpha)}\}$ is not independent, then there is another circuit $C' \subseteq B^* \cup \{s_{(\alpha)}\}$ in $\mathcal{M}(G, 4)$ containing $s_{(\alpha)}$. This circuit C' is again depicted by Figure 4a, similar to C. Let $s_{(\lambda)} = (a_3, b_3, c_3, d_3)$ be the second square of C'. The four possible cases are considered.

– $b_1 = b_3, c_1 = c_3, d_1 = d_3$ and $(a_1, a_3) \in G$. Note that $a_3 \neq b_2$ as the edge $(b_1 = b_3, b_2)$ is not in G. However in that case a_1 has four neighbors.
– The case $a_1 = a_3, b_1 = b_3, d_1 = d_3$ and $(c_1, c_3) \in G$ is symmetrical.
– $a_1 = a_3, c_1 = c_3, d_1 = d_3$ and $(b_1, b_3) \in G$. Note that $b_3 \neq b_2$ as $(b_1, b_2) \notin G$. However in that case a_1 has four neighbors.
– $a_1 = a_3, b_1 = b_3, c_1 = c_3$ and $(d_1, d_3) \in G$. If $d_3 = b_2$, then $(d_3 = b_2, d_1 = d_2)$ is a diagonal of $s_{(\varepsilon)}$. However, if $d_3 \neq b_2$ then a_1 has four neighbors.

The existence of C' is a contradiction, $B^* \cup \{s_{(\alpha)}\}$ is independent in $\mathcal{M}(G, 4)$.

As a consequence, $B^*\backslash\{s_{(\varepsilon)}\} \cup \{s_{(\alpha)}\}$ is feasible and optimal. However, as B is independent and contains $s_{(\alpha)}$, then $s_{(\varepsilon)} \notin B$, this means that $\varepsilon > \alpha$. The optimality of $B^*\backslash\{s_{(\varepsilon)}\} \cup \{s_{(\alpha)}\}$ contradicts the fact that B^* maximises the value of α. As a conclusion no such square $s_{(\alpha)}$ exists and B is optimal. $\square$

6 Concluding remarks

This paper introduces the problem of maximizing the intersection of minimum cycle bases of graphs and studies the complexity based on four natural parameters. Several questions remain open regarding the minimum value of k for which inapproximability and W[1]-hardness hold. Furthermore, from a chemical perspective, one could argue that the difference between the sets of edges of the graphs (the edit distance) may be small relative to k, Δ, and γ. This observation may lead to the discovery of new tractable algorithms for the problem.

References

1. Aboulfath, Y., Bougueroua, S., Cimas, A. Barth, D., Gaigeot, M.P.: Time-resolved graphs of polymorphic cycles for h-bonded network identification in flexible biomolecules. Journal of Chemical Theory and Computation (2024)
2. Amaldi, E., Iuliano, C., Jurkiewicz, T., Mehlhorn, K., Rizzi, R.: Breaking the o(m2n) barrier for minimum cycle bases. In: Algorithms - ESA 2009. pp. 301–312. European Symposium on Algorithms, Springer Berlin Heidelberg (2009)
3. Bougueroua, S., Spezia, R., Pezzotti, S., Vial, S., Quessette, F., Barth, D., Gaigeot, M.P.: Graph theory for automatic structural recognition in molecular dynamics simulations. The Journal of Chemical Physics **149**(18), 184102 (nov 2018). https://doi.org/10.1063/1.5045818
4. Edmonds, J.: Submodular Functions, Matroids, and Certain Polyhedra, pp. 11–26. Springer Berlin Heidelberg, Berlin, Heidelberg (2003). https://doi.org/10.1007/3-540-36478-1_2, https://doi.org/10.1007/3-540-36478-1_2
5. Gaüzère, B., Brun, L., Villemin, D.: Relevant cycle hypergraph representation for molecules. In: Kropatsch, W.G., Artner, N.M., Haxhimusa, Y., Jiang, X. (eds.) Graph-Based Representations in Pattern Recognition. pp. 111–120. Springer Berlin Heidelberg, Berlin, Heidelberg (2013)
6. Gleiss, P.M., Stadler, P.F., Wagner, A., Fell, D.A.: Relevant cycles in chemical reaction networks. Advances in complex systems **4**(02n03), 207–226 (jun 2001). https://doi.org/10.1142/s0219525901000140
7. Horton, J.D.: A polynomial-time algorithm to find the shortest cycle basis of a graph. SIAM Journal on Computing **16**(2), 358–366 (apr 1987)
8. Ilemo, S.N., Barth, D., David, O., Quessette, F., Weisser, M.A., Watel, D.: Improving graphs of cycles approach to structural similarity of molecules. PLOS ONE **14**(12), e0226680 (dec 2019). https://doi.org/10.1371/journal.pone.0226680
9. Kavitha, T., Liebchen, C., Mehlhorn, K., Michail, D., Rizzi, R., Ueckerdt, T., Zweig, K.A.: Cycle bases in graphs characterization, algorithms, complexity, and applications. Computer Science Review **3**(4), 199–243 (nov 2009). https://doi.org/10.1016/j.cosrev.2009.08.001
10. Korte, B., Hausmann, D.: An analysis of the greedy heuristic for independence systems. In: Annals of Discrete Mathematics, vol. 2, pp. 65–74. Elsevier (1978)
11. Likhachev, I.V., Balabaev, N., Galzitskaya., O.: Available instruments for analyzing molecular dynamics trajectories. The Open Biochemistry Journal **10**(1), 1–11 (march 2016)
12. Misra, J., Gries, D.: A constructive proof of vizing's theorem. Information Processing Letters **41**(3), 131–133 (mar 1992). https://doi.org/10.1016/0020-0190(92)90041-s
13. Murphy, O.J.: Computing independent sets in graphs with large girth. Discrete Applied Mathematics **35**(2), 167–170 (1992)
14. Vismara, P.: Union of all the minimum cycle bases of a graph. The Electronic Journal of Combinatorics **4**(1) (jan 1997). https://doi.org/10.37236/1294
15. Welsh, D.J.A.: Matroid theory, p. 131. Courier Dover Publications (1976)
16. Zuckerman, D.: Linear degree extractors and the inapproximability of max clique and chromatic number. In: Proceedings of the Thirty-Eighth Annual ACM Symposium on Theory of Computing. p. 681–690. STOC '06, Association for Computing Machinery, New York, NY, USA (2006). https://doi.org/10.1145/1132516.1132612, https://doi.org/10.1145/1132516.1132612

A Proofs of Lemmas 1, 2, 3 and 7

Lemma 1. *Given B a cycle basis of a graph G with two cycles $c_1 \in B$ and $c_2 \notin B$, if $\lambda_B(c_1, c_2) = 1$ then $(B \setminus \{c_1\}) \cup \{c_2\}$ is a cycle basis of G.*

Proof. Let $B' = (B \backslash \{c_1\}) \cup c_2$. Let d be any cycle of G. As $\lambda_B(c_1, c_2) = 1$, then there exists $D \subset B \backslash \{c_1\}$ such that $c_2 = c_1 \oplus \bigoplus D$. Thus, $c_1 = c_2 \oplus \bigoplus D$. As a consequence, c_1 is generated by B'.

Then any cycle that is generated by B is also generated by B'. As $|B'| = |B|$, it is a cycle basis. □

We prove two intermediates lemma to prove Lemmas 2 and 3.

Lemma 11. *If $B \in \mathcal{MCB}(G)$ then for c_1, c_2 with $c_1 \in B$ and $c_2 \notin B$ such that $\lambda_B(c_1, c_2) = 1$, we have $\omega(c_1) \leq \omega(c_2)$.*

Proof. Let $B \in \mathcal{MCB}(G)$. Then by definition B is a cycle basis. Let $c_1 \in B$ and $c_2 \notin B$ with $\lambda_B(c_1, c_2) = 1$. By Lemma 1, $(B \backslash \{c_1\}) \cup \{c_2\}$ is a cycle basis. Then, if $\omega(c_2) < \omega(c_1)$, then the weight of B' is lower than the one of B, this contradict the optimality of B. □

Lemma 12. *Let B_1, B_2 be two cycle bases of G such that*

- *for each $B \in \{B_1, B_2\}$, for every c_1, c_2 where $c_1 \in B$ and $c_2 \notin B$ such that $\lambda_B(c_1, c_2) = 1$, we have $\omega(c_1) \leq \omega(c_2)$*

then

- *for every $c_2 \in B_2 \backslash B_1$, there exists $c_1 \in B_1 \backslash B_2$ such that $(B_2 \backslash \{c_2\}) \cup \{c_1\}$ is a cycle basis with same weight as B_2.*

Proof. The result may be proved with Lemma 1 if we demonstrate there exists $c_1 \in B_1$ with $\omega(c_1) = \omega(c_2)$ and $\lambda_{B_2}(c_2, c_1) = 1$.

Let $D_1 = \{d \in B_1 \cap B_2 | \lambda_{B_1}(d, c_2) = 1\}$, $D_2 = \{d \in B_1 \backslash B_2 | \lambda_{B_1}(d, c_2) = 1 \text{ and } \omega(d) < \omega(c_2)\}$ and $D_3 = \{d \in B_1 \backslash B_2 | \lambda_{B_1}(d, c_2) = 1 \text{ and } \omega(d) = \omega(c_2)\}$. By the hypothesis on B_1, no cycle of B_1 with weight greater than c_2 generates c_2. Then $c_2 = \bigoplus D_1 \oplus \bigoplus D_2 \oplus \bigoplus D_3$.

By the hypothesis on B_2, for all $d \in D_2$ and $e \in B_2$ such that $\lambda_{B_2}(e, d) = 1$, $\omega(e) \leq \omega(d) < \omega(c_2)$ thus $e \neq c_2$: for all $d \in D_2$, $\lambda_{B_2}(c_2, d) = 0$. Note also that $c_2 \notin D_1$ as $D_1 \subseteq B_1$. Thus, if we have for all $d \in D_3$, $\lambda_{B_2}(c_2, d) = 0$ then

$$c_2 = \bigoplus D_1 \oplus \bigoplus_{d \in D_2 \cup D_3} \bigoplus_{\substack{e \in B_2 \backslash c_2 \\ \lambda_{B_2}(e,d)=1}} e$$

As c_2 does not belong to the right side of the equation, B_2 is not linearly independent. This is a contradiction. There is then $d \in D_3$ such that $\lambda_{B_2}(c_2, d) = 1$. By Lemma 1, $(B_2 \backslash \{c_2\}) \cup \{c_1\}$ is a cycle basis with same weight as B_2. As $d \in B_1 \backslash B_2$, the lemma is proved. □

Lemma 2. *$B \in \mathcal{MCB}(G)$ if and only if B is a cycle basis and for c_1, c_2 with $c_1 \in B$ and $c_2 \notin B$ such that $\lambda_B(c_1, c_2) = 1$, we have $\omega(c_1) \leq \omega(c_2)$.*

Proof. The forward direction is proved by Lemma 11.

Now given a basis B_1 satisfying the second property and B_2 be a minimum cycle basis. By Lemma 11, then we can apply Lemma 12 with B_1 and B_2. Given $c_2 \in B_2 \backslash B_1$, there exists $c_1 \in B_1 \backslash B_2$ such that $(B_2 \backslash \{c_2\}) \cup \{c_1\} \in \mathcal{MCB}(G)$. We can then successively exchange all the cycles of B_2 with cycle of B_1 until the two coincide. As the new basis is still a minimum cycle basis, this demonstrates that B_1 is a minimum cycle basis. □

Lemma 3. *If $B_1, B_2 \in \mathcal{MCB}(G)$, for every $c_1 \in B_1 \backslash B_2$, there exists $c_2 \in B_2 \backslash B_1$ such that $(B_1 \backslash \{c_1\}) \cup \{c_2\} \in \mathcal{MCB}(G)$.*

Proof. This is a direct consequence of Lemmas 11 and 12. □

Lemma 7. *B is independent in $\mathcal{M}(G, l)$ if and only if, for all $D \subseteq B$,*
$\bigoplus D \notin span(cycle\ c | \omega(c) \leq l - 1)$.

Proof. We denote by S_{l-1} the set $span(\text{cycle } c | \omega(c) \leq l - 1)$.

We first prove the necessary condition. Let B be a subset of cycles of size l such that B is independent in $\mathcal{M}(G, l)$. Then B is independent in $\mathcal{M}(G)$: there exists a minimum cycle basis B' in G containing B.

We assume there exists $D \subseteq B$ such that $\bigoplus D \in S_{l-1}$. Then there exists E containing cycles of G with weight at most $l - 1$ such that $\bigoplus D = \bigoplus E$. By Lemma 2, for all $e \in E$ and $f \in B'$ such that $\lambda_{B'}(f, e) = 1$, we have $\omega(f) \leq \omega(e) = l - 1$. This means that $f \notin D$. Consequently

$$\bigoplus D = \bigoplus_{e \in E} \bigoplus_{\substack{f \in B' \backslash D \\ \lambda_{B'}(f,e)=1}} f$$

As no cycle of D belongs to the right side of the equation, B' is linearly dependent, which is a contradiction. This proves that no such set D exists.

Now, we prove the sufficient condition, assuming that for all $D \subset B$, $\bigoplus D \notin S_{l-1}$. Then first B is linearly independent (otherwise for some D, we would have $\bigoplus D = 0 \in S_{l-1}$). Let $B^* \in \mathcal{MCB}(G)$. We now prove that if $B \not\subseteq B^*$ then for all $d \in B \backslash B^*$, there exists $c \in B^*$ such that $B^* \backslash \{c\} \cup \{d\} \in \mathcal{MCB}(G)$.

Let $d \in B \backslash B^*$. If, for all $c \in B^*$ such that $\lambda_{B^*}(c, d) = 1$, we have $\omega(c) \neq l$, then by Lemma 2, $\omega(c) \leq l - 1$. Consequently, $d \in S_{l-1}$, this contradicts the hypothesis on B. Given c with $\lambda_{B^*}(c, d) = 1$ and $\omega(c) = l$, Lemma 1 proves $B^* \backslash \{c\} \cup \{d\} \in \mathcal{MCB}(G)$. By doing such exchanges we eventually get a minimum cycle basis containing B.

□