



Graph-Theoretic Partitioning of RNAs and Classification of Pseudoknots

Louis Petingi^{1,2}  and Tamar Schlick^{3,4}

¹ Computer Science Department,
College of Staten Island (CUNY), New York City, USA
louis.petingi@csi.cuny.edu

² City University of New York Graduate Center, New York City, USA

³ Department of Chemistry,
Courant Institute of Mathematical Sciences, NYU, New York City, USA
schlick@nyu.edu

⁴ NYU-East China Normal University Center for Computational Chemistry,
New York University Shanghai, Shanghai, China

Abstract. Dual graphs have been applied to model RNA secondary structures with pseudoknots, or intertwined base pairs. In a previous work, a linear-time algorithm was introduced to partition dual graphs into maximal topological components called blocks and determine whether each block contains a pseudoknot or not. This characterization allowed us to efficiently isolate smaller RNA fragments and classify them as pseudoknotted or pseudoknot-free regions, while keeping these substructures intact. In this paper we extend the partitioning algorithm by classifying a pseudoknot as either recursive or non-recursive. A pseudoknot is recursive if it contains independent regions or fragments. Each of these regions can be also identified by the modified algorithm, continuing with our current research in the development of a library of building blocks for RNA design by fragment assembly. Partitioning and classification of RNAs using dual graphs provide a systematic way for study of RNA structure and prediction.

Keywords: Graph theory · RNA secondary structures · Partitioning · Bi-connectivity · Pseudoknots

1 Introduction

Let $G = (V, E)$ be undirected graph composed of by a finite set of *vertices* V and a set E of unordered pair of vertices called *edges*, where each edge represents a relation between two vertices.

Our RNA analysis is based on dual graphs, introduced in 2003 by Gan et al. [7], to model RNA secondary structures (2D). The 2D elements of RNA molecules consist of double-stranded (stem) regions defined by base pairing such as Adenine-Uracil, Guanine-Cytosine, Guanine-Uracil, and single stranded loops; stems and loops are mapped to the vertices and edges of the corresponding dual

graph, respectively (later we present an alternative definition of dual graphs). Dual graphs can represent complex RNA structures called pseudoknots (*PKs*); these structures involve an intertwining of two-base-paired regions of the RNA and are common elements in many biologically relevant RNAs.

In [19, 20] a linear-time partitioning algorithm was introduced based on the dual graph representation of RNA 2D. This algorithm partitions a dual graph into connected components called *blocks* and then determines whether each block contains a pseudoknot or is a regular region. Thus our procedure provides a systematic approach to partition an RNA 2D, into smaller classified regions, while providing a topological perspective for the analysis of RNAs.

Pseudoknots can be classified into two main groups: *recursive* and *non-recursive* pseudoknot [9, 23]. The former is distinguished from the latter because it contains an internal pseudoknotted or regular region that does not intertwine with external stems within the PK; in this work, the original algorithm is extended to classify PKs into these two main categories. In addition, as a recursive PK comprises independent regions or fragments, our modified algorithm can also identify each of these regions, to be later cataloged and applied in the analysis of RNAs with pseudoknots.

In the next section, we present background material and definitions relevant to this paper, and we review the partitioning algorithm introduced in [19, 20], as well as its applications, as for example the development of a library of building blocks for RNA design by fragment assembly [13]. Following this line of research, in Sect. 3 it is shown how the partitioning approach can be extended so if a block contains a pseudoknot, then it can be classified as either recursive or non-recursive; in the case the PK is recursive, the algorithm can also identify each independent region. We summarize the findings and outline new directions in Sect. 4. An Appendix section includes computational tests performed by the modified algorithm, on some RNA's motifs.

2 Background

2.1 Biological and Topological Definitions

In 2003, Gan et al. [7] introduced *tree* and *dual* graph-theoretic representations of RNA 2D motifs in a framework called RAG (RNA-As-Graphs) [5, 8, 12, 16]. A pseudoknot is an intertwining of two-based-paired regions (stems) of an RNA (see for example Fig. 1).

The partition algorithm is based on topological properties of graphs, suggests an alternative way to look at the problem of detection and classification of PKs and of general RNAs. As base pairing in PKs is not well-nested, making the presence of PKs in RNA sequences more difficult to predict by the more classical dynamic programming [3] and context-free grammars standard methods [2].

Following (Kravchenko [17]), we define our biological variables as follows.

Definition 1. *General terms:*

- a. *RNA primary structure: a sequence of linearly ordered bases x_1, x_2, \dots, x_r , where $x_i \in \{A, U, C, G\}$.*
- b. *canonical base pair: a base pair $(x_i, x_j) \in \{(A, U), (U, A), (C, G), (G, C), (G, U), (U, G)\}$.*
- c. *RNA secondary structure without pseudoknot - or regular structure, encapsulated in the region (i_0, \dots, k_0) : an RNA 2D structure in which no two base pairs $(x_i, x_j), (x_l, x_m)$, satisfy $i_0 \leq i < l < j < m \leq k_0$ (i.e., no two base pairs intertwined).*
- d. *a base pair stem: a tuple $(x_i, x_{i+1}, \dots, x_{i+r}, x_{i+(r+1)}, \dots, x_{j-1}, x_j)$ in which $(x_i, x_j), (x_{i+1}, x_{j-1}), \dots, (x_{i+r}, x_{i+(r+1)})$ form base pairs.*
- e. *loop region: a tuple (x_1, x_2, \dots, x_r) in which $\forall_{i \leq j \leq r} (x_i, x_j)$ does not form a base pair.*
- f. *a pseudoknot encapsulated in the region (i_0, \dots, k_0) : if $\exists l, m, (i_0 < l < m < k_0)$ such that (x_{i_0}, x_m) and (x_l, x_{k_0}) are base pairs (i.e., at least two base pairs intertwined).*

A graphical representation is a natural way to describe an RNA 2D structure (see Fig. 1(a), (b)), in which the x -axis is labeled according to the primary linearly ordered sequence of bases (Definition 1a), and a stem (Definition 1d) is represented by arcs connecting base pairs. A region on the x -axis between the end-points of the arcs representing stems is called a *segment*.

A dual graph can be defined from the graphical representation of an RNA 2D structure as follows (Fig. 1).

Definition 2. *The dual graph is defined by mapping stems and the segments between stems (x -axis), of the graphical representation of an RNA 2D structure, to the vertices and edges of the dual graph, respectively.*

In the next section we present our partitioning approach as of a dual graph G , into subgraphs $G' \subseteq G$, called blocks.

2.2 Graph Partitioning Algorithm

The graph-theoretic partitioning algorithm is based on identifying *articulation points* of the dual graph representation of an RNA 2D. An articulation point is a vertex of a graph whose deletion disconnects a graph or an isolated vertex remains.

We need to define the following.

Definition 3. *Connectivity*

- a. *A vertex v is an articulation point or cut-vertex if $G - v$ results in a disconnected graph (i.e., at least two connected components remain) or an isolated vertex remains.*
- b. *A connected component is non-separable if it does not have an articulation point (or cut-vertex). Please note that single edges or isolated points are non-separable.*

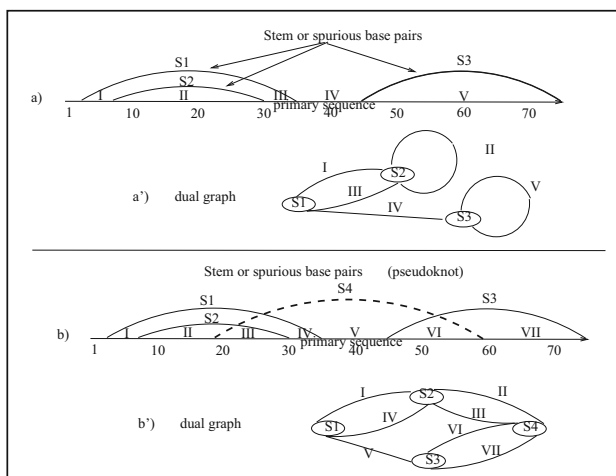


Fig. 1. Graphical and dual graph representations of an RNA 2D structure. (a) graphical representation of a pseudoknot-free RNA primary sequence and embedded stems or base pairs; (a') corresponding dual graph representation. (b) graphical representation of a pseudoknotted RNA 2D structure; (b') corresponding dual graph. This figure was originally depicted in [20].

- c. A block is a maximal (edge-wise) non-separable graph.
- d. An edge-set X is an edge-disconnecting set if the removal of X from G results in a disconnected graph. The edge-connectivity of a graph $\lambda(G)$ is the size of a minimum edge-disconnecting of G .
- e. The degree of a vertex v of G is the number of $d_G(v)$ is the number of edges incident at v .

Articulation points allow us to identify blocks (see Fig. 2); since a block is a maximally non-separable component, a pseudoknot cannot be then contained in two different blocks. Thus identification of these block components allows us to isolate pseudoknots (as well as pseudoknot-free blocks), without breaking their structural properties.

An algorithm for identifying (bi-connected) block components in a graph was introduced by Hopcroft and Tarjan [11], and runs in linear computational time.

A *hairpin* loop occurs when two regions of the same strand, usually complementary in nucleotide sequence when read in opposite directions, base-pair to form a double helix that ends in an unpaired loop. A self-loop in the dual graph, i.e., an edge having the same vertex as the end-points, represents a hairpin, and as it does not connect two different vertices (i.e., stems), it is formally deleted from the dual graph.

From Definition 1c, an RNA 2D structure is a regular-region (pseudoknot-free) and encapsulated in a region (i_0, \dots, k_0) , if no two base pairs $(x_i, x_j), (x_l, x_m)$, satisfy $i < l < j < m, i_0 \leq i, j, l, m \leq m_0$, otherwise the region is a pseudoknot; this definition yields the following main result.

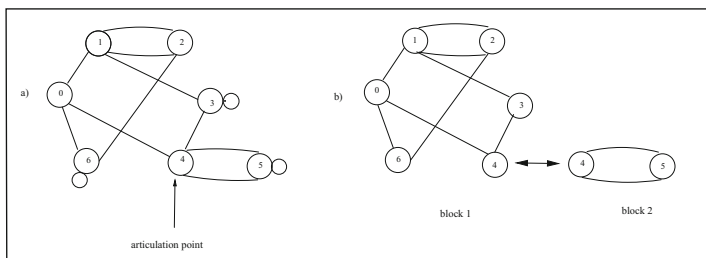


Fig. 2. Identification of (a) articulation points and (b) partitioning of a dual graph.

Corollary 4. [19, 20] *Given a dual graph representation of RNA 2D structure, a block represents a pseudoknot if and only if the block has a vertex of degree (Definition 3e) at least 3.*

The partitioning algorithm performs the following steps.

1. *Partition the dual graph into blocks by application of Hopcroft and Tarjan's algorithm.*
2. *Analyze each block to determine whether contains a vertex of degree at least 3. If that is the case then the block contains a pseudoknot, according to Corollary 4. If not then the block represents a pseudoknot-free structure.*

Consider as an example the dual graph shown in Fig. 2. This graph is decomposed into 2 blocks. According to Corollary 4, block 1 is a pseudoknot as it has a vertex of degree at least 3, while block 2, a cycle, corresponds to a regular region.

Our partitioning algorithm was applied recently [13] to analyze the modular units of RNAs for a representative database of experimentally determined RNA structures and to develop a library of building blocks for RNA design by fragment assembly, as done recently for tree graphs, along with supporting chemical mapping experiments [14]. Among the 22 frequently occurring motifs we found for known RNAs up to 9 vertices, 15 contain pseudoknots [13]. Thus, further classification of the pseudoknotted RNAs could help in cataloging and applications to RNA design. Another application of the partitioning algorithm to small and large units of ribosomal RNAs of various prokaryotic and eukaryotic organisms helped identify common subgraphs and ancestry relationships [13].

In the next section we extend our algorithm to classify PKs as either recursive or non-recursive; the algorithm can also identify each recursive region.

3 Classification of Pseudoknots as Either Recursive or Non-recursive and Identification of Each Recursive Region

The RNA 2D dual graph and graphical representations depicted in this section are based upon New York University's *RAG*-database [12], and *R-Chie* visualization software [18], respectively.

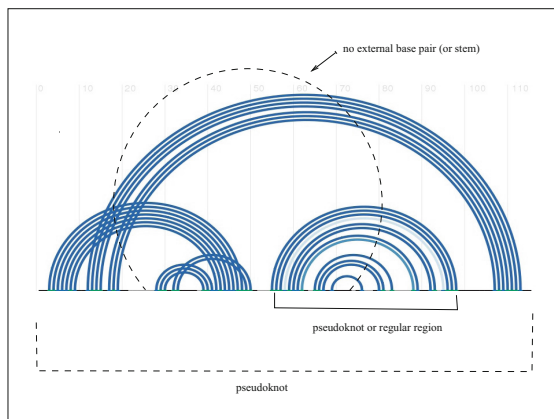


Fig. 3. Recursive pseudoknot.

The definition of a recursive pseudoknot follows the one of Wong et al. [23]. Let $A = x_1x_2\dots x_m$ be a sequence of linearly ordered bases, and M be the 2D of A . M is represented as a set of base pair positions, i.e., $M = \{(i, j) | 1 \leq i < j \leq m, (x_i, x_j) \text{ is a base pair}\}$. Let $M_{x,y} \subseteq M$ be the set of base pairs within the subsequence $x_1x_2\dots x_m$, $1 \leq x < y \leq m$, i.e., $M_{x,y} = \{(i, j) \in M | x \leq i < j \leq y\}$, with $M = M_{1,m}$.

Definition 5. $M_{x,y}$ is a recursive pseudoknot if $M_{x,y}$ is a pseudoknot (see Definition 1f), and $\exists a_1, b_1, \dots, a_s, b_s, (x < a_1 < b_1 < \dots < a_s < b_s < y)$ that satisfy the followings.

Each M_{a_i, b_i} is called a recursive region.

- M_{a_i, b_i} , for $1 \leq i \leq s$, is a recursive pseudoknot.
- For each M_{a_i, b_i} , $1 \leq i \leq s$, there does not exist a base pair $(i, j) \in M$ that $i \in [a_i, b_i]$ but $j \notin [a_i, b_i]$, or $i \notin [a_i, b_i]$ but $j \in [a_i, b_i]$.
- $M_{x,y} \cup \bigcup_{i=1}^s M_{a_i, b_i}$ is either a regular structure or a pseudoknot.

A recursive pseudoknot is a pseudoknot $M_{x,y}$ that contains a pseudoknotted or regular region $M_{a,b}$, and there does not exist a base pair (c, d) , such that x_d is a base of $M_{a,b}$ and x_c is a base of M external to $M_{a,b}$ (see Fig. 3). Here we are assuming that $M_{a,b}$ is contained in $M_{x,y}$, that is, $x < a < b < y$.

Wong et al. definitions [23] also incorporated the concepts of *standard* and *non-standard* pseudoknots; however it is not within the scope of this work to study them from the dual graph representation perspective.

A graph is *Eulerian* if there exist a trail with no repetition of edges from a vertex v_0 of G , ending at vertex v_k , covering all the edges of the topology; if $v_0 = v_k$ then the graph is an *Eulerian cycle* (see [10], p. 64). Dual graph representations of general RNA 2D structures, and specifically of PKs, can be easily shown to be Eulerian graphs from Definition 2. By starting from the origin

on the x -axis of the graphical representation and traversing to the right, a unique trail in its dual graph can be described, where all edges are covered.

Lemma 6. [19, 20] *The dual graph representations of RNA 2D structures and of PKs are Eulerian by following the primary sequence of bases.*

As depicted in Fig. 1(b), the alternating sequence of stems and segments $\{S_1, I, S_2, II, S_4, III, S_2, IV, S_1, V, S_3, VI, S_4, VII, S_3\}$ of the graphical representation (b) forms an Eulerian trail in its dual graph (b').

A pseudoknotted block can be classified as recursive by just calculating the edge-connectivity (see Definition 3d) of the block. As an example consider the *Hepatitis Delta Virus Ribozyme* (see Fig. 4), necessary for viral replication. The stem labeled 4 in the graphical representation (or vertex labeled 4 in the dual graph) is attached to the pseudoknot by the segments a and b in its graphical representation, or edges labeled a and b in the dual graph representation. It is clear that if the PK is recursive then the edge-connectivity of the pseudoknotted block must be 2. However it is not obvious that the converse is necessarily true, that is, if the pseudoknotted block has edge-connectivity 2 then it is recursive. The following Lemma settles this question.

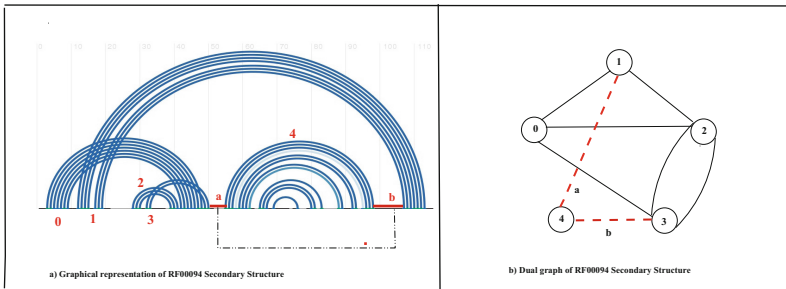


Fig. 4. Hepatitis Delta Virus Ribozyme secondary structure. (a) Graphical representation. (b) Dual graph representation.

Lemma 7. *The dual graph representation of a pseudoknotted block is recursive if and only if the block has edge-connectivity 2.*

Proof. If the block $M_{x,y}$ is a recursive pseudoknot then it contains an internal region $M_{a,b}$ with $x < a < b < y$ according the aforementioned definition. As there does not exist a base pair (c, d) in which x_d is a base in the internal region and x_c is a base of the pseudoknot outside this internal region, then $M_{a,b}$ must be adjacent to the remaining of the PK in the graphical representation by two segments, or equivalently, by two edges in the dual graph of the pseudoknot (see Fig. 4).

Conversely suppose that the dual graph of a PK, $G = (V, E)$, has edge-connectivity 2 and let $E' = \{e_1, e_2\}$ represent a minimum size disconnecting set. As G is connected then deletion of E' from G will result in exactly two connected components, $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ (see Fig. 5a). From Lemma 6 one knows that there exist an Eulerian path in G following the primary sequence in linear order, starting from the origin on the x -axis of the graphical representation and traversing to the right. Let the Eulerian path $P = P_1.e_1.P_2.e_2.P_3$ where P_1 starts at the initial base x_1 (w.l.o.g. we assume that x_1 is in G_1) (see Fig. 5b). It is the case that P_2 must cover all the edges of G_2 (following an ordered sequence of bases) because when P reaches e_2 to continue with P_3 , P can not go back to G_2 again as e_1 and e_2 were already used by the Eulerian path. Therefore G_2 is a region composed of all edges (and vertices) corresponding to an ordered sequence of bases, that is, G_2 is a well-defined region within the pseudoknotted block G . \square

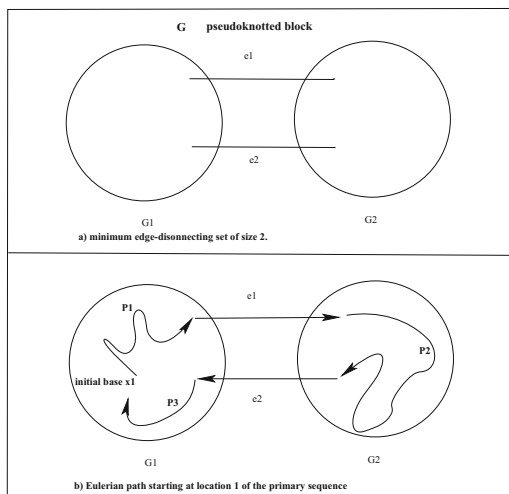


Fig. 5. Minimum disconnecting set of size 2. (a) Two connected subgraphs. (b) Eulerian path covering all the edges of the dual graph and following the primary sequence.

The edge-connectivity of a graph $G = (V, E)$ can be determined in polynomial time in order $(|V||E|^2)$ using the max-flow min-cut theorem of network flows by Edmond and Karp [4], or Ford and Fulkerson [6].

As it is shown in the proof of by Lemma 7, we can also delete each pair of edges and determine if the graph is disconnected using Depth-First-Search [10] in time $(|E|^3)$; this variation allows us to find every internal recursive region of the recursive pseudoknot if such pair of edges exist. For example if edges a and b of Hepatitis Delta Virus Ribozyme 2D dual graph representation (Fig. 4) are

deleted, then vertex labeled 4 corresponding to the stem labeled 4 of the graphical representation will be isolated. Thus the disconnecting set $\{a, b\}$ uniquely identifies a recursive region (i.e., stem labeled 4).

Corollary 4 and Lemma 7 yield the following partitioning and classification of pseudoknots algorithm.

Algorithm for Partitioning and Classification of PKs

- i. Input dual graph $G = (V, E)$ as the Adjacency Matrix, of a RNA 2D.
- ii. Output partitioning of the RNA 2D into recursive PK, non-recursive PK, and regular regions.
 - 1. Partition the dual graph into blocks by application of Hopcroft and Tarjan’s algorithm;
 - 2. Analyze each block to determine whether each contains a vertex of degree at least 3;
 - 3. **IF** the block has a vertex of degree ≥ 3 then the block is a **pseudoknot**;
 - Apply max-flow min-cut theorem to determine edge-connectivity;
 - *if* edge-connectivity = 2 then the block is a **recursive pseudoknot**;
else the **pseudoknot is not recursive**;
 - 4. **ELSE** the **block is a regular region**;

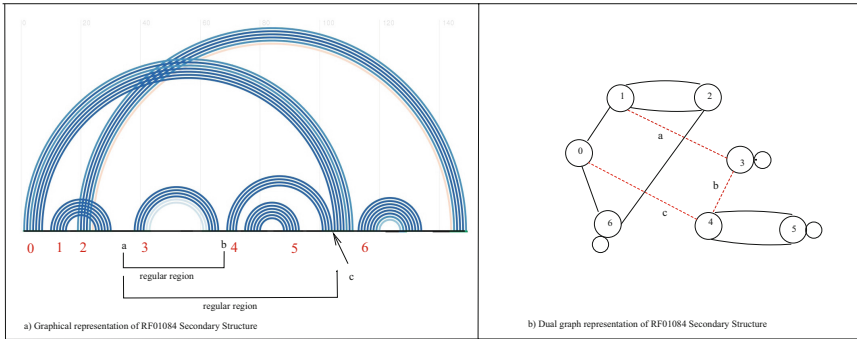


Fig. 6. A *tRNA-like-structure*. (a) Graphical representation. (b) Dual graph representation.

As another example consider a *tRNA-like-structure* [1], linked to regulation of plant virus replication (see Fig. 6). Note that the vertex labeled 4 is an articulation point, therefore this dual graph will be partitioned into two blocks, one is a pseudoknot, because it contains a vertex of degree 3 or greater (self-loops are deleted in the analysis), while the other block is a regular region. The pseudoknotted block can be classified as recursive because it has edge-connectivity 2. In addition every disconnecting set of size 2 represents a recursive internal region of the PK.

As an example of a non-recursive pseudoknot consider the *Translational repression of the Escherichia coli alpha operon mRNA* [22], illustrated in Fig. 7. The dual graph representation of this motif 2D has edge-connectivity 3, thus it is not a recursive PK.

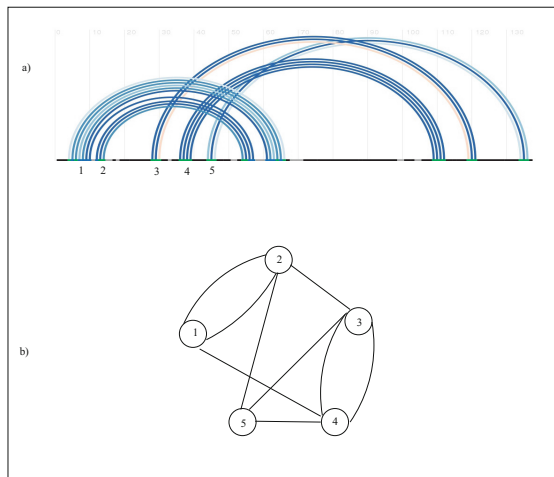


Fig. 7. *Translational repression of the Escherichia coli alpha operon mRNA.* (a) Graphical representation; (b) Dual graph representation.

The Appendix illustrates the output generated by the modified algorithm when is run on some of the aforementioned motifs. The algorithm is written in C++ and is archived for public use [21].

4 Conclusions and Ongoing Work

We have extended our partitioning algorithm of the dual graph representation of RNA 2D structures into maximal non-separable components called blocks, to classify pseudoknots as either recursive or non-recursive. In [19,20] it was shown that an RNA 2D contains a pseudoknot if and only if the dual graph representation has a block in which one of the vertices is of degree 3 or larger. This paper showed that a pseudoknotted block is recursive if and only if the block has edge-connectivity 2. Moreover each disconnecting set of size 2 represents an internal recursive region of the pseudoknot allowing further classification of modular units for RNA design. These results also offer an alternative and simple way to visualize and classify PKs based on graph theoretical properties, allowing a systematic analysis of RNAs.

With our recent extension of our graph growing algorithm to generate dual graph libraries of possible RNA motifs, thousands more potential graphs were analyzed and classified [15]. The modified algorithm will be useful for these motifs for further studies of RNA structure and design.

Acknowledgments. We would like to thank first the referees for improving the content of the paper. NIGMS support from award R35GM122562 to T. S. is gratefully acknowledged. The work of L. P. was supported by PSC-CUNY Award # 61249-00-49 of the City University of New York.

Appendix

Let (a, b) represents an edge of a dual graph with end-vertices a and b .

We are next illustrating the output of the partitioning algorithm tested on the tRNA-like-structure dual graph (see Fig. 6).

```

----- Motif :RF01084 -----
===== New Block =====
(4,5) - (4,5) -
--- this block represents a regular-region ---
===== New Block =====
(4,0) - (3,4) - (1,3) - (6,0) - (2,6) - (1,2) - (1,2) - (0,1) -
removed edges (2,6) and (6,0), these two edges are a disconnecting set:
The block is a recursive PK.
----- Summary information for Motif :RF01084 -----
----- Total number of blocks: 2
----- number of non-recursive PK blocks: 0
----- number of recursive PK blocks: 1
----- number of regular blocks : 1
-----

```

References

1. Barends, S., Rudinger-Thirion, J., Florentz, C., Giegé, R., Pleij, C., Kraal, B.: tRNA-Like structure regulates translation of brome mosaic virus RNA. *J. Virol.* **78**(8), 4003–4010 (2004)
2. Condon, A., Davy, B., Rastegari, B., Zhao, S., Tarrant, F.: Classifying RNA pseudoknotted structures. *Theor. Comput. Sci.* **320**(1), 35–50 (2004)
3. Dirks, R.-M., Pierce, N.-A.: A partition function algorithm for nucleic acid secondary structure including pseudoknots. *J. Comput. Chem.* **24**(13), 1664–1677 (2003)
4. Edmonds, J., Karp, R.-M.: Theoretical improvements in algorithmic efficiency for network flow problems. *J. ACM* **19**(2), 248–264 (1972)
5. Fera, D., et al.: RAG: RNA-As-Graphs web resource. *BMC Bioinform.* **5**, 88 (2004)
6. Ford, L.R., Fulkerson, D.R.: Maximal flow through a network. *Can. J. Math.* **8**(1956), 399–404 (1956)
7. Gan, H.-H., Pasquali, S., Schlick, T.: Exploring the repertoire of RNA secondary motifs using graph theory; implications for RNA design. *Nucleic Acids Res.* **31**(11), 2926–2943 (2003)
8. Gan, H.-H., et al.: RAG: RNA-As-Graphs database-concepts, analysis, and features. *Bioinformatics* **20**(8), 1285–1291 (2004)
9. Han, B., Dost, B., Bafna, V., Zhang, S.: Structural alignment of pseudoknotted RNA. *J. Comput. Biol.* **15**(5), 489–504 (2008)

10. Harary, F.: Graph Theory. Addison-Wesley, Boston (1969)
11. Hopcroft, J., Tarjan, R.: Efficient algorithms for graph manipulation. *Communi. ACM* **16**(6), 372–378 (1973)
12. Izzo, J.-A., Kim, N., Elmetwaly, S., Schlick, T.: RAG: an update to the RNA-As-Graphs resource. *BMC Bioinform.* **12**, 219 (2011)
13. Jain, S., Bayrak, C.-S., Petingi, L., Schlick, T.: Dual graph partitioning highlights a small group of pseudoknot-containing RNA submotifs. *Genes* **9**(8), 371 (2018)
14. Jain, S., Ramos, S.-B., Laederach, A., Schlick, T.: A pipeline for computational design of novel RNA-like topologies. *NAR* **46**(14), 7040–7051 (2018)
15. Jain, S., Saju, S., Petingi, L., Schlick, T.: An extended dual graph library and partitioning algorithm applicable to pseudoknotted RNA structures. *Methods* (2019). <https://doi.org/10.1016/j.ymeth.2019.03.022>. Accessed 1 Apr 2019
16. Kim, N., Petingi, L., Schlick, T.: Network theory tools for RNA modeling. *WSEAS Trans. Math.* **12**(9), 941–955 (2013)
17. Kravchenko, A.: Predicting RNA secondary structures including pseudoknots. University of Oxford internal report (2009)
18. Lai, D., Proctor, J.-R., Zhu, J.-Y., Meyer, I.-M.: R-CHIE: a web server and R package for visualizing RNA secondary structures. *Nucleic Acids Res.* **40**(12), e95 (2012). e-RNA: <https://www.e-rna.org/r-chie/rfam.cgi>. Accessed 4 Jan 2019
19. Petingi, L., Schlick, T.: Partitioning RNAs into pseudonotted and pseudoknot-free regions modeled as dual graphs. *q-bio.QM*, [arXiv:1601.04259](https://arxiv.org/abs/1601.04259) (2016). <http://arxiv.org/abs/1601.04259>. Accessed 4 Oct 2018
20. Petingi, L., Schlick, T.: Partitioning and classification of RNA secondary structures into pseudonotted and pseudoknot-free regions using a graph-theoretical approach. *IAENG Int. J. Comput. Sci.* **44**(2), 241–246 (2017)
21. Petingi, L.: Dual graph partitioning code. <https://github.com/Louis-Petingi/Partition-Algorithm-3>. Accessed 3 Jan 2019
22. Schlax, P.-J., Xavier, K.-A., Gluick, T.-C., Draper, D.-E.: Translational repression of the Escherichia coli alpha operon mRNA: importance of an mRNA conformational switch and a ternary entrapment complex. *J. Biol. Chem.* **276**(42), 38494–38501 (2001)
23. Wong, T.-K., Lam, T.-W., Sung, W.-K., Cheung, B.-W., Yiu, S.-M.: Structural alignment of RNA with complex pseudoknot structure. *J. Comput. Biol.* **18**(1), 97–108 (2011)