Addendum: topology and prediction of RNA pseudoknots

Christian M. Reidys[1,*], Fenix W. D. Huang[1], Jørgen E. Andersen[2], Robert C. Penner[2,3,4], Peter F. Stadler[5] and Markus E. Nebel[6]

[1]Department of Mathematics and Computer Science, University of Southern Denmark, [2]Department of Mathematics and Computer Science, Center for Quantum Geometry of Moduli Spaces, Aarhus University, Århus C, Odense, Denmark, [3]Department of Mathematics, [4]Department of Physics, California Institute of Technology, Pasadena, CA, USA, [5]Department of Computer Science, University of Leipzig, Härtelstrasse 16-18, Leipzig and [6]Department of Computer Science, University of Kaiserslautern, Kaiserslautern, Germany

Associate Editor: Anna Tramontano

It has come to our attention that several concepts and results underlying the `gfold` software presented in our article 'Topology and prediction of RNA pseudoknots' (Reidys *et al.*, 2011) are also present in earlier work by Bon *et al.* (2008); Orland and Zee (2002); Pillsbury *et al.* (2005b); Vernizzi *et al.* (2005) and (Pillsbury *et al.*, 2005a). Here, we briefly examine these works in relation to the results of our paper.

The classification and expansion of pseudoknotted RNA structures in terms of the topological genus of an associated fatgraph or double line graph were first proposed by Orland and Zee (2002) and Bon *et al.* (2008), although fatgraphs were applied to RNA secondary structures already by Penner and Waterman (1993) and Penner (2004). The enumerative results initiated by Orland and Zee (2002) are based on matrix models, while our generating functions are derived via representation theory (Zagier, 1995). Enumeration results on RNA structures according to genus were already obtained by Vernizzi *et al.* (2005), again using the formal framework of the matrix model. Genus as well as other topological invariants of fatgraphs were introduced and studied as descriptors of proteins in Penner *et al.* (2010).

Pillsbury *et al.* (2005a) report recursion relations of time complexity $O(N^6)$ to generate RNA structures of genus one in the context of an RNA folding algorithm that is substantially different from our algorithm `gfold`. Aside from not incorporating loop-based energy models, `gfold` is not restricted to genus one RNA structures. The four basic irreducible shadows of genus one in Theorem 2.3 of our paper appeared first in Bon *et al.* (2008); Pillsbury *et al.* (2005b). The shadows of Reidys *et al.* (2011) are derived from (i) the notion of irreducibility formulated by Kleitman (1970) and (ii) the work on pseudoknot shapes by Jin and Reidys (2009); Reidys and Wang (2010). Irreducibility is equivalent to the concept of primitivity introduced by Bon *et al.* (2008), inspired by the work of Dyson (1949).

The equation to compute the genus of a fatgraph is classical going back to Euler (1752) and was first applied in the context representing RNA structures by Orland and Zee (2002) and Bon *et al.* (2008).

Additivity of genus under topological sums is elementary (Massey, 1967) and for reducible and nested RNA structures first discussed by Bon *et al.* (2008). Our Equations (2.1), (2.2) and (2.4) are thus textbook knowledge. Lemma 2.1 is also well known and was used e.g. by Penner and Waterman (1993) and Bon *et al.* (2008).

REFERENCES

Bon,M. *et al.* (2008) Topological classification of RNA structures. *J. Mol. Biol.*, **379**, 900–911.

Dyson,F.J. (1949) The S matrix in quantum electrodynamics. *Phys. Rev.*, **75**, 1736–1755.

Euler,L. (1752) Elementa doctrinae solidorum. *Novi Comm. Acad. Sci. Imp. Petropol.*, **4**, 109–140.

Jin,E. and Reidys,C. (2009) Combinatorial design of pseudoknot RNA. *Adv. Appl. Math.*, **42**, 135–151.

Kleitman,D. (1970) Proportions of irreducible diagrams. *Studies Appl. Math.*, **49**, 297–299.

Massey,W.S. (1967) *Algebraic Topology: An Introduction*. Springer, New York.

Orland,H. and Zee,A. (2002) RNA folding and large *n* matrix theory. *Nuclear Physics B*, **620**, 456–476.

Penner,R.C. (2004) Cell decomposition and compactification of Riemann's moduli space in decorated Teichmüller theory. In Tongring,N. and Penner,R.C. (eds) *Woods Hole Mathematics-Perspectives in Math and Physics*, World Scientific, Singapore, arXiv: math.GT/0306190, pp. 263–301.

Penner,R.C. and Waterman,M.S. (1993) Spaces of RNA secondary structures. *Adv. Math.*, **101**, 31–49.

Penner,R.C. *et al.* (2010) Fatgraph models of proteins. *Comm. Pure Appl. Math.*, **63**, 1249–1297.

Pillsbury,M. *et al.* (2005a) An algorithm for RNA pseudoknots. arXiv: condmat/0310505v2.

Pillsbury,M. *et al.* (2005b) Steepest descent calculation of RNA pseudoknots. *Phys. Rev. E*, **72**, 011911.

Reidys,C.M. and Wang,R. (2010) Shapes of RNA pseudoknot structures. *J. Comput. Biol.*, **17**, 1575–1590.

Reidys,C.M. *et al.* (2011) Topology and prediction of RNA pseudoknots. *Bioinformatics*, **27**, 1076–1085.

Vernizzi,G. *et al.* (2005) Enumeration of RNA structures by matrix models. *Phys. Rev. Lett.*, **94**, 168103.

Zagier,D. (1995) On the distribution of the number of cycles of elements in symmetric groups. *Nieuw Arch. Wisk. IV*, **13**, 489–495.

*To whom correspondence should be addressed.