An efficient algorithm to compute the minimum free energy of interacting nucleic acid strands

Ahmed Shalaby Damien Woods

Hamilton Institute, Department of Computer Science, Maynooth University, Ireland

We give the first polynomial time algorithm for predicting optimal multistranded structures, answering the open problem from Dirks, Bois, Schaeffer, Winfree and Pierce [SIAM Review 2007].

Practitioners need algorithms to predict the most favoured DNA/RNA secondary structures, called minimum free energy (MFE) structures, or to compute a partition function that allows assigning a probability to any structure. MFE prediction is NP-hard in the presence of pseudoknots—base pairings that violate a restricted planarity condition. However, for single-stranded unpseudoknotted structures, there are polynomial time dynamic programming algorithms. For multiple strands, the problem is significantly more complicated: Codon, Hajiaghayi and Thachuk [DNA27, 2021] proved it NP-hard for N bases and $\mathcal{O}(N)$ strands. In 2007, Dirks et al gave a polynomial time partition function algorithm for multiple ($\mathcal{O}(1)$) strands, now widely-used in software packages like NUPACK and ViennaRNA, however their technique did not generalise to MFE which they left open.

We give an $\mathcal{O}(N^4)$ time algorithm for unpseudoknotted multiple ($\mathcal{O}(1)$) strand MFE prediction [Shalaby, Woods, ICALP 2025], answering the open problem from Dirks et al. The challenge lies in considering the rotational symmetry of secondary structures, a global feature not immediately amenable to local subproblem decomposition used in dynamic programming. Our proof has two main technical contributions: First, a characterisation of symmetric secondary structures implying only quadratically many need to be considered when computing the rotational symmetry penalty. Second, that bound is leveraged by a backtracking algorithm to efficiently find the MFE in an exponential space of contenders.

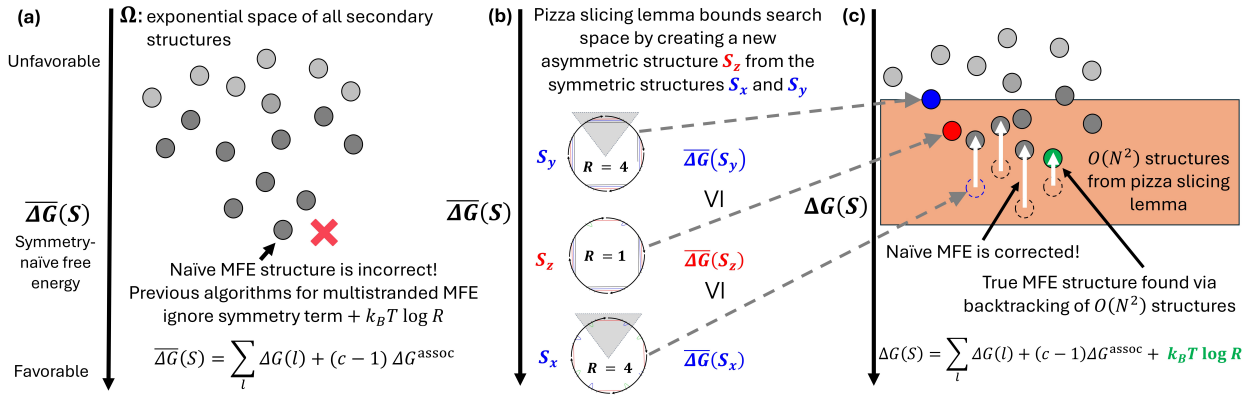

Figure 1: Overview of our result. (a) Current state-of-the-art MFE prediction algorithms for $\mathcal{O}(1)$ interacting strands are not guaranteed to predict the optimal MFE structure since they ignore the rotational symmetry penalty [Dirks et al, 2007; NUPACK; ViennaRNA]. (b) Slicing and swapping strategy for constructing new asymmetric structure S_z by combining two symmetric structures S_x, S_y that have the same symmetric backbone cut. (c) The slicing and swapping strategy guarantees efficient and correct MFE prediction within a backtracking subspace of $O(N^2)$ structures.