STOCHASTIC COMPUTATION ON DNA STRANDS THROUGH HYDROXYL NICKING

Tonglin Chen, Arnav Solanki, Marc Riedel

University of Minnesota
Department of Electrical and Computer Engineering

1 Introduction

This abstract demonstrates a novel scheme for storing information and performing computation on randomly nicked DNA. Previous research has shown that hydroxyl radicals can be used to cleave DNA – we will use the term “‘nick” DNA – randomly along its backbone [1]. One can exert fine-grained control on the rate of nicking. We exploit this process to store *fractional values*: the value stored in DNA is a fraction between 0 and 1, relative to a maximum rate of nicking. We use *toehold-mediated DNA strand displacement*, a powerful tool for performing computation on DNA [2], [3]. We also use DNA enzymes such as ligase, which repairs nicks on the DNA backbone, and flap endonuclease, which snips off overhanging single-strand flaps [4],[5]. With random nicking, we can exploit the theory of *stochastic computing* to transform stored fractional values [6]. We demonstrate the basic operations of data storage and computation in this paradigm.

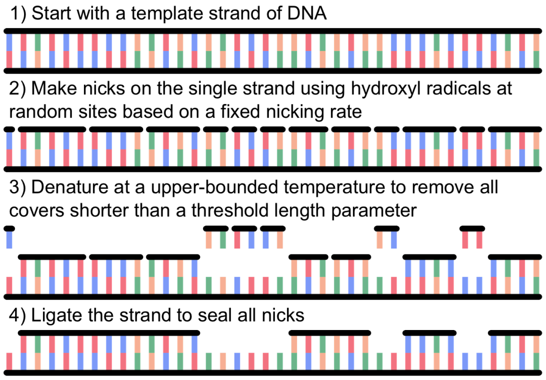

Fig. 1: Procedure for encoding a fractional value on a template strand of DNA.

2 Encoding Data on DNA

We propose nicking DNA strands using hydroxyl radicals. Nicking produces double-stranded DNA complexes with random cuts in the phosphate backbone. After denaturing, bases opposite adjacent nicks are exposed. Call the *fractional value* that is stored the probability that a given base is exposed. Equivalently, by the law of large numbers, it is the ratio of the total number of bases exposed to the total number of bases. We will call contiguous exposed regions in this strand *gaps* and their counterparts *covers*. For example, in Figure 1 (4), the complex has 11 bases exposed out of a total of 38 bases, so the value stored by the strand is $11/38 \approx 0.2894$. There are 4 gaps and 4 covers.

3 Nicking Rate Transformation

This process of data storage assumes that random nicking can be performed selectively on the phosphate backbone of only one of the two strands in a double-stranded complex. When the complex is lightly denatured, bases between adjacent nicks fall off. We call the minimum distance between adjacent nicks for this to occur the *threshold* for denaturing. The strand can the be treated with ligase to repair any extraneous nicks after this stage. Call the *nicking rate*, the rate at which nicks are produced, relatively to a maximum rate. We have determined that the following function predicts the fractional value of a strand f based on the nicking rate x and threshold for denaturing k:

$$f = 1 - (1 + kx)(1 - x)^k$$

This result is based on simulations as well as analytical reasoning.

4 Logical `AND` operation

Based on the theory of stochastic computing, a logical `AND` operation produces a result that is the *product* of two fractional values [6]. Figure 2 depicts how the operation can be performed: the probability of a certain base being exposed in the final strand C is the probability of it being simultaneously exposed in strand A and in strand B; this is the product of the fractional values of both strands, assuming these were created randomly and independently. Note that with $f_A = 11/38$ and $f_B = 13/38$,

---

This work was funded by DARPA Grant #W911NF-18-2-0032. Poster could be found at `http://mriedel.ece.umn.edu/wiki/images/2/29/FNANO_poster2020.pdf`

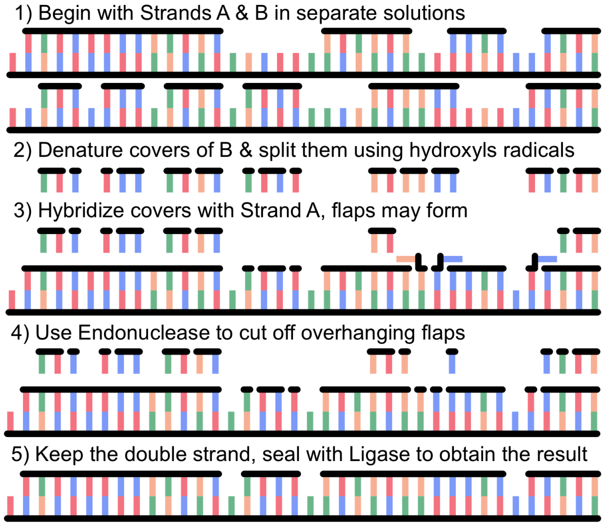

Fig. 2: Logical `AND` operation on two fractional values.

we have $f_C = 4/38 \approx 0.1052$. This approximates $f_A \times f_B \approx 0.0990$.

5 Logical `NOT` operation

Based on the theory of stochastic computing, given a fractional value x, a logical `NOT` operation produces a value $1 - x$. Figure 3 shows a `NOT` operation: the denaturing step separates out the covers for the output strand that exactly fill the gaps in the input strand.

6 Conclusion

In this abstract, we presented a novel scheme for storing data in DNA using random nicking with hydroxl groups. We also proposed methods for implementing basic computation, namely products with `AND` operations and $1-x$ computations with `NOT` operations. The methods are fundamental yet powerful. Research in stochastic computing has shown that complex functions can be performed by composing these operations [6]. For instance, a Taylor series expansion of e^{-x} can computed with only 7 such operations:

$$e^{-x} \approx 1 - x + \frac{x^2}{2!} - \frac{x^3}{3!} = 1 - x(1 - \frac{x}{2}(1 - \frac{x}{3}))$$

We have designed and verified computation of such functions through simulation. We are collaborating with the Soloveichik and Milenkovic groups at the Univ. of Texas and the Univ. of Illinois, respectively, to demonstrate the computation experimentally on DNA.

7 References

[1] T. D. Tullius and B. A. Dombroski, “Hydroxyl radical “footprinting”: high-resolution information about dna-protein contacts and application to lambda repressor and cro protein,” *Proceedings of the National Academy of Sciences*, vol. 83, no. 15, pp. 5469–5473, Aug 1986.

[2] Bernard Yurke, “A dna-fuelled molecular machine made of dna,” *Nature*, vol. 406, no. 6796: 605, 2000.

[3] David Soloveichik, Georg Seelig, and Erik Winfree, “Dna as a universal substrate for chemical kinetics,” *Proceedings of the National Academy of Sciences*, vol. 107, no. 12, pp. 5393–5398, 2010.

[4] I Robert Lehnman, “Dna ligase: structure, mechanism, and function,” *Science*, vol. 186, no. 4166, pp. 790–797, 1974.

[5] Yuan Liu, Hui-I Kao, and Robert A Bambara, “Flap endonuclease 1: a central component of dna metabolism,” *Annual review of biochemistry*, vol. 73, no. 1, pp. 589–615, 2004.

[6] W. Qian, X. Li, Marc Riedel, K. Bazargan, and D. J. Lilja, “An architecture for fault-tolerant computation with stochastic logic,” *IEEE Transactions on Computers*, vol. 60, no. 1, pp. 93–105, 2011.

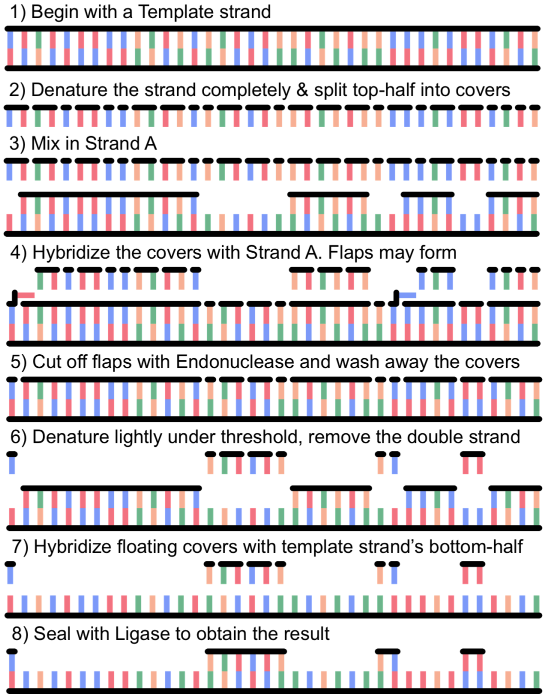

Fig. 3: Logical `NOT` operation on a single fractional value.