Stochastic Computing on DNA Strands through Hydroxyl Radical Nicking

Arnav Solanki, Tonglin Chen, Marc Riedel

Department of Electrical and Computer Engineering
University of Minnesota, Twin-Cities

Introduction

We present a novel scheme for storing information and performing computation on randomly nicked DNA. Previous research has shown that hydroxyl radicals can be used to nick the DNA backbone randomly, independent of the sequence [7]. Fine-grained control can be exerted 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 based on the exposed bases scaled to the strand's length. This procedure does not require DNA sequence synthesis. We use *toehold-mediated DNA strand displacement*, a powerful tool for performing computation on DNA [9], [6]. 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 [3],[4]. With random nicking, we can exploit the theory of *stochastic computing* to transform stored fractional values similar to bit-streams [5]. We demonstrate the basic operations of data storage and computation, the NOT and AND operation in this paradigm.

Creating Fractional Strands

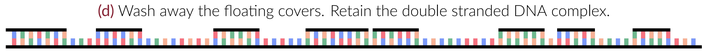
Figure 1 shows how a fractional strand can be obtained from a starting nicking rate x and denaturing threshold k (we assume a naive model of DNA denaturing). The fractional strand is a DNA complex where the number of exposed bases (we term these as gap regions) to the total strand length is the fraction stored.

(a) Start with a template double strand of DNA - top strand is the nicking strand, bottom strand is the base strand.

(b) Randomly nick the DNA backbone on the nicking strand at a nicking rate of x. The nicking rate is the probability of the backbone being nicked next to a random nucleotide. The nicking strand segment between any two nicks is called a cover.

(c) Denature the double strand at a controlled temperature - here we assume that all covers with a length less than $k = 5$ nucleotides are denatured. k is the denaturing threshold.

(d) Wash away the floating covers. Retain the double stranded DNA complex.



(e) Seal all extraneous nicks with Ligase. The resulting DNA complex is a fractional strand that encodes the fraction y.

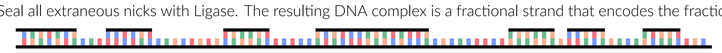


Figure 1. The procedure for creating the fractional strand Y that stores the fraction y, given the parameters x, the nicking rate, and k, the denaturing threshold.

The transformation from nicking rate x and denaturing threshold k to the fraction stored y is given by the equation -

$$y = 1 - (1 + kx)(1 - x)^k. \tag{1}$$

This equation was observed as a polynomial fit for y in x for our simulations shown in Figure 2

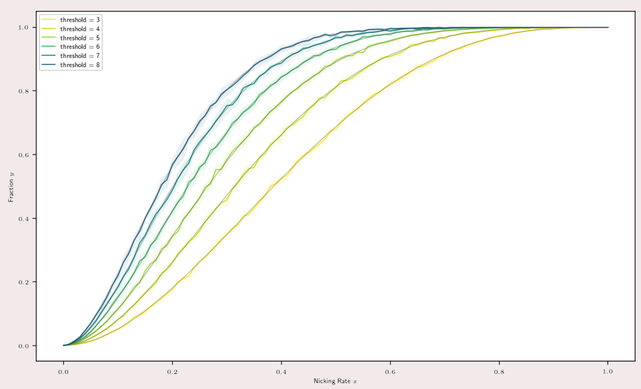


Figure 2. Plots for how the fraction y varies with the nicking rate x. We simulated 10 strands of 1000 sequence length with denaturing thresholds in [3,8]. The black lines within each threshold's colored plot are our predicted polynomial fits.

NOT Operation

The NOT operation allows the formation of a fractional strand that stores the 1-minus operation on an input strand. Figure 3 shows how the NOT of Strand A results in a strand storing the value $1 - a$.

(a) Start with strand A.

(b) Fill in the gap regions of the strand with RNA by using RNA Polymerase to create a DNA-RNA complex.

(c) Degrade all DNA in the complex with DNAse to free the RNA covers.

(d) Reverse transcribe the RNA covers into DNA covers. This could be implemented in two steps by using Reverse Transcriptase to obtain cDNA and then DNA polymerase to obtain the DNA covers.

(e) Hybridize the DNA covers with the base strand of the DNA template strand A was obtained from.

(f) The resulting fractional strand C stores the NOT value of strand A.

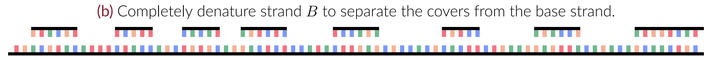
Figure 3. The procedure for performing a NOT operation on a fractional strand A that stores the fraction a. This NOT operation is synonymous to a 1-minus or complement operation on a. The result of this operation is a fractional strand C that stores $c = 1 - a$.

AND Operation

The AND operation creates a fractional strand that stores the product of two fractions. Figure 4 shows how the AND of Strands A and B results in a strand storing the value $a \times b$.

(a) Start with strand B that stores the fraction b.

(b) Completely denature strand B to separate the covers from the base strand.



(c) Wash away the base strand, retain the floating covers.

(d) Hybridize the covers from B with strand A (same as in the NOT operation in Figure 3)

(e) After hybridization, some covers may remain floating. Flaps may form - these flaps will exist in a dynamic equilibrium such that flaps from either overlapping cover may be hanging free.

(f) Cut off the overhanging flaps with Flap Endonuclease to separate the extraneous covers from the DNA complex

(g) Wash away the floating covers.

(h) Seal all extraneous nicks with Ligase. The resulting fractional strand D stores the AND of strands A and B.

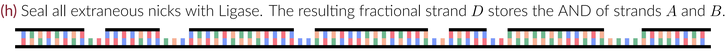


Figure 4. The procedure for performing an AND operation on fractional strands A and B, that store the fractions a and b respectively. This AND operation is synonymous to a intersection of sets. The result of this operation is a fractional strand D that stores $d = a \times b$.

Application via Cascading Operations

Research in stochastic computing has shown that complex polynomial functions can be performed by composing the AND and NOT operations [5] in cascades. For instance, a 4^{th} order e^{-x} and a 6^{th} order $\cos(x)$ can computed with Taylor series expansions with only 11 and 12 such operations respectively:

$$e^{-x} \approx 1 - x + \frac{x^2}{2} - \frac{x^3}{3!} + \frac{x^4}{4!} \approx \left(1 - x\left(1 - \frac{1}{2}x\left(1 - \frac{1}{3}x\left(1 - \frac{1}{4}x\right)\right)\right)\right), \tag{2}$$

$$\cos(x) \approx 1 - \frac{x^2}{2} + \frac{x^4}{4!} - \frac{x^6}{6!} \approx \left(1 - \frac{1}{2}x^2\left(1 - \frac{1}{4}x^2\left(1 - \frac{1}{6}x^2\right)\right)\right). \tag{3}$$

We have designed and verified computation of such functions through simulation, showing promising results performing within 15% error rate. We are interested in implementing our methods physically in laboratories and welcome further collaboration.

Discussion

An important note for implementing fractional strands as stochastic datastructures is the notion of independence - the AND and NOT operation must operate without any bias or dependence to allow for large scale cascading. This means effects such as hybridization biases due to the template strand's sequence, cross-hybridization across multiple fractional strands in a single solution and incomplete denaturing must be reduced to minimize noise in stochastic operations. For computing functions involving high degrees of a single fraction variable (for example equations 2 and 3), it is essential to generate independent copies of all inputs, even repeated values, to prevent cascading noise.

The formation of fractional strands can be done on native DNA strands, however the performance of such strands in stochastic computing is currently untested. Our method does not require any DNA synthesis of a custom sequence, thereby capable of lowering costs for DNA data storage. Fractions stored on strands could be measured with the use of sequencing technologies like Oxford Nanopore [2], or spectral analysis methods to test for DNA Hyperchromity [1].

Further work can be done in this domain to improve our model - Our denaturing model in Figure 1 is a naive approach to DNA denaturing - we plan on incorporating sequence specific melting point calculations in our next simulations to test performance. A read-copy operation, a mechanism to measure a strand's fraction and create an independent copy of it, would also be a lucrative step to implement to allow for mulitple fan-out from cascaded operations.

Project Details

This research was funded by DARPA Grant #W911NF-18-2-0032. We thank David Soloveichik, Olgica Milenkovic, Andrew Ellington, Charles Schroeder, Thomas Tullius, Owen Hoffend, Krisna Van Dyke and Adam Sychla for helpful discussions.

References

[1] David Freifelder and Peter F Davison.
Hyperchromicity and strand separation in bacterial dna.
Biophysical journal, 2(3):249, 1962.

[2] Miten Jain, Hugh E Olsen, Benedict Paten, and Mark Akeson.
The oxford nanopore minion: delivery of nanopore sequencing to the genomics community.
Genome biology, 17(1):239, 2016.

[3] I Robert Lehnman.
Dna ligase: structure, mechanism, and function.
Science, 186(4166):790–797, 1974.

[4] Yuan Liu, Hui-I Kao, and Robert A Bambara.
Flap endonuclease 1: a central component of dna metabolism.
Annual review of biochemistry, 73(1):589–615, 2004.

[5] 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, 60(1):93–105, 2011.

[6] David Soloveichik, Georg Seelig, and Erik Winfree.
Dna as a universal substrate for chemical kinetics.
Proceedings of the National Academy of Sciences, 107(12):5393–5398, 2010.

[7] 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, 83(15):5469–5473, Aug 1986.

[8] Weikang Qian and M. D. Riedel.
The synthesis of robust polynomial arithmetic with stochastic logic.
In *2008 45th ACM/IEEE Design Automation Conference*, pages 648–653, June 2008.

[9] Bernard Yurke.
A dna-fuelled molecular machine made of dna.
Nature, 406(6796: 605), 2000.