DISCRETE BERNOULLI CONVOLUTIONS: AN ALGORITHMIC APPROACH TOWARD BOUND IMPROVEMENT

NEIL J. CALKIN, JULIA DAVIS, MICHELLE DELCOURT, ZEBEDIAH ENGBERG,
JOBBY JACOB, AND KEVIN JAMES

ABSTRACT. In this paper we consider a discrete version of the Bernoulli convolution problem traditionally studied via functional analysis. We develop an algorithm which bounds the Bernoulli sequences, and we give a significant improvement on the best known bound. [1]

1. INTRODUCTION

The classic Bernoulli convolution problem in analysis has an elegant discrete analogue which we now describe. Consider the two maps $\mathrm{dup}_n, \mathrm{shf}_n : \mathbb{R}^n \longrightarrow \mathbb{R}^{3n}$ defined by

$$(1.1) \quad \mathrm{dup}_n : (a_1, a_2, ..., a_{n-1}, a_n) \longmapsto (a_1, a_1, a_2, a_2, ..., a_{n-1}, a_{n-1}, a_n, a_n, \overbrace{0, ..., 0}^{n \text{ times}})$$

$$(1.2) \quad \mathrm{shf}_n : (a_1, a_2, ..., a_{n-1}, a_n) \longmapsto (\overbrace{0, ..., 0}^{n \text{ times}}, a_1, a_1, a_2, a_2, ..., a_{n-1}, a_{n-1}, a_n, a_n).$$

The names "dup" and "shf" reference the duplication and shifting of the coordinates. Consider the finite sequences of increasing length recursively given by $B_0 = (1)$ and $B_{n+1} = \mathrm{dup}_n(B_n) + \mathrm{shf}_n(B_n)$. We call B_n the *Bernoulli sequence on level n*. Likewise, we refer to the map $(\mathrm{dup}_n + \mathrm{shf}_n) : \mathbb{R}^n \longrightarrow \mathbb{R}^{3n}$ seen in (1.1) and (1.2) as the process of *duplicate, shift, add* or DSA for short.

Figure 1 shows the first few levels of the Bernoulli sequence.

In this paper, we are interested in the rate at which the maximum $m_n := \max(B_n)$ is growing with n. It is easy to see that the mean $\mu(B_n) = (4/3)^n$ [2]. The major question to be asked is if m_n also grows like $(4/3)^n$. We develop an algorithm which allows one to bound m_n. Upon adjusting the parameter of our algorithm, we achieve the best known bound $m_n = O((1.33997599)^n)$.

This problem is equivalent to an unsolved problem in the theory of Bernoulli convolutions. A Bernoulli convolution is obtained as an infinite convolution of Bernoulli measures. The Bernoulli measure, denoted by $b(X)$, is the measure corresponding to the discrete probability density function on the real line with value $1/2$ at 1 and -1. The Bernoulli convolution with parameter q for $0 < q < 1$ is the measure

$$\mu_q(X) = b(X) * b(X/q) * b(X/q^2) * ....$$

```
1 ➝ 1 1
        1 1
    ─────────
      1 2 1
```

```
1 2 1 ➝ 1 1 2 2 1 1
              1 1 2 2 1 1
        ───────────────────
        1 1 2 3 2 3 2 1 1
```

```
1 1 2 3 2 3 2 1 1 ➝ 1 1 1 1 2 2 3 3 2 2 3 3 2 2 1 1 1 1
                          1 1 1 1 2 2 3 3 2 2 3 3 2 2 1 1 1 1
                    ─────────────────────────────────────────────
                    1 1 1 1 2 2 3 3 2 3 4 4 3 4 3 4 4 3 2 3 3 2 2 1 1 1 1
```

FIGURE 1.   This shows the process of DSA at three low levels.
The figure demonstrates computation of the Bernoulli sequence on
level $n+1$ from the Bernoulli sequence on level n for $n = 0, 1$ and 2
using the definition.

A different perspective on Bernoulli convolutions is obtained through a functional
equation. For $0 < q < 1$, consider the functional equation

(1.3) $$F(t) = \frac{1}{2}F\left(\frac{t-1}{q}\right) + \frac{1}{2}F\left(\frac{t+1}{q}\right)$$

for t on the interval $I_q := [-1/(1-q), 1/(1-q)]$. $F_q(t)$ is the distribution function
of μ_q, that is $F_q(t) = \mu_q((-\infty, t])$. Jessen and Winter, in [3], showed that $F_q(t)$
is either absolutely continuous or purely singular. The major question regarding
the solutions of (1.3) is to determine the values of q that make $F_q(t)$ absolutely
continuous. The only value of q for which it is shown that $F_q(t)$ is absolutely
continuous is $q = 1/2$. In this paper, we investigate this question for $q = 2/3$. $F_q(t)$
is absolutely continuous at $q = 2/3$ if and only if $m_n = O((4/3)^n)$. For a detailed
introduction and motivation on the discrete analogue of Bernoulli convolutions, see
[2].

For an in depth report on Bernoulli convolutions, refer to *Sixty years of Bernoulli
convolutions* [4]. A good introduction on the computational aspects of the problem
is found in Chapter 5 of [1].

2. Bounding Bernoulli

2.1. **Preliminaries.** Since B_n has length 3^n, we index it by the first 3^n nonnegative
integers. It is advantageous to normalize the indexing in such a way that each index
is in the interval $[0, 1]$. We simply take the image of $k \in \{0, 1, 2, ..., 3^n - 1\}$ under
the map $k \mapsto k/3^n$ to normalize the index. For a subset $S \subset [0, 1]$, we define

$$\Gamma_n(S) = \max_{x \in \bar{S}} g_n(x)$$

where $\bar{S} = S \cap \{0, 1/3^n, 2/3^n, ..., (3^n - 1)/3^n\}$ and $g_n(x)$ denotes the value corre-
sponding to x in the nth level Bernoulli sequence where now $x \in [0, 1]$. In other
words,

(2.1) $$g_n\left(\frac{k}{3^n}\right) = b_k \qquad \text{for } k = 0, 1, ...3^n - 1.$$

Consider the two maps predup, preshf : $[0, 1] \cup \{\emptyset\} \longrightarrow [0, 1] \cup \{\emptyset\}$ defined by

$$\text{predup} : x \longmapsto \begin{cases} (3/2)x & \text{if } x \in [0, 2/3] \\ \emptyset & \text{otherwise} \end{cases}$$

$$\text{preshf} : x \longmapsto \begin{cases} (3/2)(x - 1/3) & \text{if } x \in [1/3, 1] \\ \emptyset & \text{otherwise.} \end{cases}$$

Furthermore, define $\text{predup}(\emptyset) = \emptyset$ and $\text{preshf}(\emptyset) = \emptyset$.

For a given entry $g_n(x)$ on the nth level, $g_n(x)$ can be written as a sum of two entries on level $n - 1$, provided x is in the middle third. If x is in the first third or the last third then $g_n(x)$ only comes from one entry on the previous level. The maps predup and preshf give the preimage of the index x on the duplicated copy of the $(n-1)^{th}$ level and the shifted copy of the $(n-1)^{th}$ level, respectively. These maps reflect the behavior of the indices in the duplication process and shifting. Indeed, in this new setting the process of *duplicate, shift, add* translates to the equation

(2.2) $$g_{n+1}(x) = g_n(\text{predup}(x)) + g_n(\text{preshf}(x))$$

where $g_n(\emptyset) = 0$.

2.2. **Description of algorithm.** We wish to give a reasonable bound on the growth of $\Gamma_n := \Gamma_n([0, 1])$ in terms of n. Note that Γ_n is equal to m_n, which is defined in Section 1. The following procedure gives a method for computing a positive real number θ so that $\Gamma_n = O(\theta^n)$. The exact number θ will depend on several factors discussed in the remarks in Section 2.4.

(1) Write the interval $[0, 1]$ as a union

(2.3) $$[0, 1] = \bigcup_{j=1}^{w} D_j$$

where each D_j is a closed interval.

(2) Fix an index j and let $D = D_j$. Consider the image of D under predup and preshf. We "pull D back" one level to obtain the two intervals $\text{predup}(D)$ and $\text{preshf}(D)$. Note that one of $\text{predup}(D)$ or $\text{preshf}(D)$ may be empty, but this poses no problem. Pulling D back two levels, we are left with the four intervals

$$(\text{predup}(\text{predup}(D))), \quad (\text{preshf}(\text{predup}(D))),$$

$$(\text{predup}(\text{preshf}(D))), \quad (\text{preshf}(\text{preshf}(D))).$$

Continuing in this manner, we are left with 2^r intervals after considering the image of D under all pullbacks to level $n - r$. Call these intervals $E_1, E_2, ..., E_{2^r}$.

(3) For each $i = 1, 2, ..., 2^r$, consider the interval E_i. If E_i is empty, we disregard E_i. Also, if E_i only contains 0 or 1, we disregard it. The reason for this is that $\Gamma_n(\{0\}) = 1$ and $\Gamma_n(\{1\}) = 0$. Hence intervals will contribute nothing—this is explained in the proof. Otherwise there exists a $y \in E_i$ so that $|y - 1/2| \le |z - 1/2|$ for all other $z \in E_i$. In other words, we select the element from E_i that is closest to $1/2$. If $y > 1/2$, replace y with $1 - y \in (0, 1/2]$ in the following step.

(4) Consider the sequence of rationals $a_0, a_1, a_2, \ldots$ where

$$a_k = \frac{1}{2}\left(\frac{2}{3}\right)^k.$$

We see that $a_k \to 0$ as $k \to \infty$, hence there exists $k = k_i$ so that $a_{k+1} \leq y < a_k$.

(5) For this index i, define the monomial $\mathrm{mon}_i(X) := X^{n-r-k_i} \in \mathbb{R}[X]$. If $E_i = \emptyset, \{0\},$ or $\{1\}$, then set $\mathrm{mon}_i(X) = 0$. Recall that $D = D_j$ and consider

(2.4)
$$f_j(X) = X^n - \sum_{i=1}^{2^r} \mathrm{mon}_i(X).$$

Note that f_j is not necessarily a polynomial. If n is small, then the powers $n - r - k_i$ of X may turn out to be negative. However, we assume that n is large enough so that f_j is indeed a polynomial. Increasing the value of n only increases the multiplicity of the root at $X = 0$. For each $j = 1, 2, \ldots, w$, let θ_j be the greatest real root of $f_j(X)$.

(6) Let $\theta = \max_j \theta_j$.

After computing θ, we have $\Gamma_n = O(\theta^n)$.

2.3. **Verification of algorithm.** After first reading the above algorithm, it is not clear how or even why it works. We now prove that this algorithm does indeed provide a bound for Γ_n.

Theorem 2.1. *Suppose θ is found using the above method. Then $\Gamma_N = O(\theta^N)$ for all positive integers N.*

Proof. We proceed by induction on N. Whenever N is small, Γ_N is a bounded integer. Hence $\Gamma_N < C\theta^N$ for a large enough constant C. This establishes the base case; moreover it establishes the bound for any N less than some fixed positive integer. Before we begin the inductive step, we revisit our algorithm. We will see that the above algorithm is essentially the computation involved in the inductive step.

Using the aforementioned notation, we see that in step 1 we have

$$\Gamma_n = \max_{j \in \{1, \ldots, w\}} \Gamma_n(D_j).$$

Now using the properties of the functions predup and preshf as in (2.2), in step 2 we obtain the estimate:

$$\Gamma_n(D) \leq \sum_{i=1}^{2^r} \Gamma_{n-r}(E_i).$$

Consider the sequence $g_h(x)$ on some level h. There is very little we know regarding the location of the maximum. However, one very trivial statement we can make is that the maximum occurs on $[1/3, 2/3]$; in the process of *duplicate, shift, add* the first third of $g_h(x)$ is simply the first half of $g_{h-1}(x)$. By extending this inductively, if $x \in [0, (1/2)(2/3)^k]$, then $g_h(x)$ can be realized as a term of the $(h-k)^{th}$ level. A similar argument applies for $x \in [1 - (1/2)(2/3)^k, 1]$. For such x, we obtain that

$$g_h(x) \leq \Gamma_{h-k}.$$

In step 3, we are determining the $y \in E_i$ that is closest to $1/2$. In step 4, we compute the largest value of k so that every $x \in E_i$ satisfies either $x \leq (1/2)(2/3)^k$ or $x \geq 1 - (1/2)(2/3)^k$. This gives the estimate

$$\Gamma_{n-r}(E_i) \leq \Gamma_{n-r-k_i}.$$

Putting these two inequalities together, we obtain

$$\Gamma_n(D) \leq \sum_{i=1}^{2^r} \Gamma_{n-r-k_i}.$$

We now tackle the inductive step in the proof. Assume that $\Gamma_n < C\theta^n$ for all $n < N$. We wish to show that $\Gamma_N < C\theta^N$.

$$
\begin{aligned}
\Gamma_N &= \max_j \Gamma_N(D_j) \\
&\leq \max_j \sum_{i=1}^{2^r} \Gamma_{N-r-k_i} \\
&< \max_j \sum_{i=1}^{2^r} C\theta^{N-r-k_i} \\
&= C \max_j \left(\theta^N - f_j(\theta) \right)
\end{aligned}
$$

where the last equality comes from (2.4). Since $f_j(X)$ is a monic polynomial, $f_j(X) \to \infty$ as $X \to \infty$. In particular, $f_j(X) \geq 0$ for all $X \geq \theta_j$, where θ_j the largest real root of $f_j(X)$. Since $\theta \geq \theta_j$ for all j, it follows that $f_j(\theta) \geq 0$ for all j. Hence

$$C \max_j \left(\theta^N - f_j(\theta) \right) \leq C \max_j \theta^N = C\theta^N.$$

$\square$

2.4. **Remarks on algorithm and its implementation.** We can improve the efficiency of this algorithm as follows. In step 1 of the outline, we only need to consider breaking up the interval $[1/3, 1/2]$ in (2.3) as a union of closed intervals because the maximum is guaranteed to occur on this interval. Also, steps 3 and 4 could be omitted from the algorithm at the expense of obtaining a weaker bound. In this case, we set $\mathrm{mon}_i(X) = X^{n-r}$ if $E_i \neq \emptyset$ and $\mathrm{mon}_i(X) = 0$ if $E_i = \emptyset$.

We implemented the algorithm in Python. The code can easily be broken up and run at many computing nodes; indeed this is how we implemented step 1. We have found that the majority of the running time is spent in steps 3 and 4. Specifically, for a given interval D, steps 3 and 4 must be repeated at most 2^r times. Since some of the intervals E_i become empty under repeated applications of predup and preshf, the number of nonempty E_i will not be quite this high. However, the number of nonempty intervals still grows exponentially.

There are two independent parameters we control when running this algorithm. We are able to choose exactly how we break up the interval $[1/3, 1/2]$ in (2.3) and we choose the pullback number r defined in step 2. There is a computational trade off when setting these two values—it is hard to make both large at once. After testing various ways of breaking up $[1/3, 1/2]$ in (2.3) and various pullback numbers r, we

find that a small increase in r will yield a better bound than a finer partition in (2.3).

2.5. **Data.** We ran many jobs using the high-throughput computing facilities of Condor at Clemson University. Before implementing this algorithm, the best known bound on m_n was $O((\sqrt{2})^n)$ [2]. We succeeded in significantly improving the bound. The best bound we were able to achieve is $m_n = O(\theta^n)$ where $\theta = 1.33997599527$. Specifically θ is the largest real root of the polynomial

$$X^{33} - 752X^8 - 520X^7 - 319X^6 - 231X^5 - 141X^4 - 101X^3 - 54X^2 - 50X - 83.$$

This polynomial was obtained using $r = 25$ and $w = 300$.

We conclude this paper with the following conjecture.

Conjecture 2.1.1. $m_n = O((4/3)^n)$.

References

[1] David H. Bailey, Jonathan M. Borwein, Neil J. Calkin, Roland Girgensohn, D. Russell Luke, and Victor H. Moll, *Experimental mathematics in action*, A K Peters Ltd., Wellesley, MA, 2007. MR MR2320374 (2007m:00003)

[2] Neil Calkin, Julia Davis, Michelle Delcourt, Zebediah Engberg, Jobby Jacob, and Kevin James, *Taking the convoluted out of bernoulli convolutions: A discrete approach*, Submitted.

[3] Børge Jessen and Aurel Wintner, *Distribution functions and the Riemann zeta function*, Trans. Amer. Math. Soc. **38** (1935), no. 1, 48–88. MR MR1501802

[4] Yuval Peres, Wilhelm Schlag, and Boris Solomyak, *Sixty years of Bernoulli convolutions*, Fractal geometry and stochastics, II (Greifswald/Koserow, 1998), Progr. Probab., vol. 46, Birkhäuser, Basel, 2000, pp. 39–65. MR MR1785620 (2001m:42020)

Neil Calkin, Department of Mathematical Sciences, Clemson University, Clemson, SC 29634
E-mail address: calkin@ces.clemson.edu

Julia Davis, Department of Mathematics, Grove City College, Grove City, PA 16127
E-mail address: davisjl1@gcc.edu

Michelle Delcourt, School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332
E-mail address: mdelcourt3@gatech.edu

Zebediah Engberg, Department of Mathematics, Dartmouth College, Hanover, NH 03755
E-mail address: zeb@dartmouth.edu

Jobby Jacob, School of Mathematical Sciences, Rochester Institute of Technology, Rochester,   NY 14623
E-mail address: jxjsma@rit.edu

Kevin James, Department of Mathematical Sciences, Clemson University, Clemson, SC 29634
E-mail address: kevja@clemson.edu