Buffon Noodles

Ed Waymire

1. INTRODUCTION. During the 1992 spring term, physicist Corinne Manogue gave a colloquium talk to our mathematics department in which she surveyed various mathematical notions which arise in a string theorist's approach to quantum field theory. All this talk of strings and probability by Corinne made an impression on my colleague Robby Robson, who had an interesting hallway question the following day. Robby was wondering if I knew what would happen if Buffon had tossed a noodle in place of a needle (actually Buffon tossed baguettes)? Here the image of noodles is that of the long thin stringy type which always seem an entangled mess. So this story will feature entangled noodles.

2. BUFFON NEEDLE. Let's recall the Buffon needle problem. Snell ([7]) has a very nice treatment together with original references and a wonderful dose of historical perspective blended with modern computer simulations; that's where I learned about the baguette. One tosses a needle of length L onto a plane surface (floor) marked by parallel lines of width $D > L$ units apart. One asks for the probability of the event C that the needle intersects its closest line. The answer is a (linear) function

$$p_0(x) := P(C) = \frac{2L}{\pi D} = \gamma_0 x \tag{1}$$

of the ratio $x = L/D$ with slope $\gamma_0 = 2/\pi$ ($\approx .636$). A particularly interesting aspect of this example is that one obtains a statistical method of estimating numerical values of π from the crossing frequency of repeated needle tossings.

One needs a model of the needle tossing experiment to compute the probability in (1). For this, let Y denote the vertical distance from the midpoint of the needle to the nearest parallel and let Θ denote the acute angle made by the needle and the nearest parallel. Then one assumes that (Y, Θ) is uniformly distributed over the range of values $[0, D/2] \times [0, \pi/2]$. Now the problem is a geometry problem. It is interesting for the ensuing discussion that the crossing probability (1) does not change if one assumes that an *endpoint* of the needle is randomly (uniformly) selected on a fixed vertical line and, independently, the angle made by the needle with the vertical line is obtained by a random rotation from 0 to 2π.

3. BROWNIAN NOODLE. While others (eg.,see [1], [5]) have considered "noodle" extensions defined by bending needles into various specific convex shapes to be tossed, we shall allow the noodles to become randomly *tangled*. To begin, let us first consider what happens in the case of a two-dimensional standard Brownian motion $\{B(t) = (B_1(t), B_2(t))\}$ over a time interval of unit length and started at random along a vertical line of length D, FIGURE 1a. Although the total length of the Brownian noodle is ∞, one may introduce L as a scale parameter by taking

$$S(t) := \frac{\sqrt{2}L}{2}B(t), \qquad 0 \leq t \leq 1. \tag{2}$$

The answer in this case is

$$p_1(x) := P(C) = \frac{8}{\pi^2}\sum_{m=1}^{\infty}\frac{1}{(2m-1)^2}\left(1 - e^{-(2m-1)^2(\pi^2/4)(L/D)^2}\right). \tag{3}$$

Now the parameter L is a scale parameter and the gap is D. The crossing probability in (3) depends only on the ratio L/D. The extra factor $\sqrt{2}/2$ is being introduced here for the convenience of making comparisons with computations later. For convenience now let $L' := \sqrt{2}L/2$.

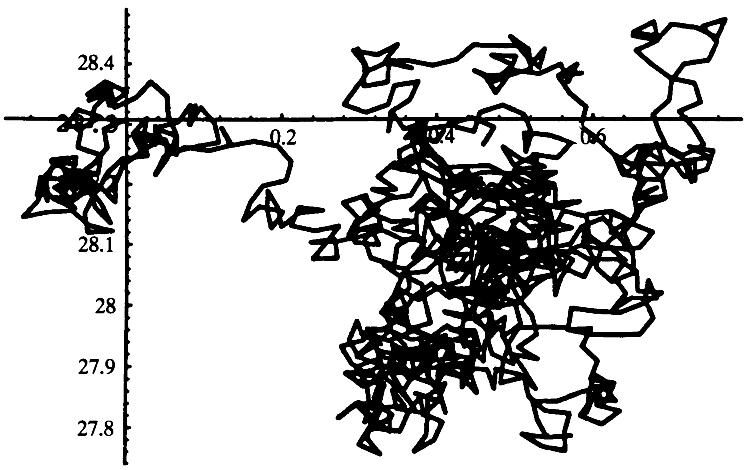

Figure 1a. Sample Realization of Brownian Noodle Toss.

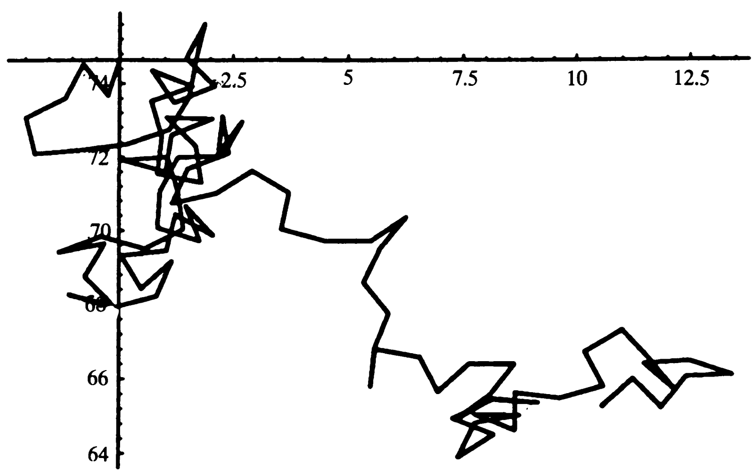

Figure 1b. Sample Realization of 100 Strand Noodle Toss.

To obtain the solution (3) one first observes that it is enough to consider the one-dimensional Brownian motion $\{B_1(t)\}$ on the interval $[0, D/L']$ with uniform initial distribution on $[0, D/L']$. We will need the joint distribution of (M_1, m_1) where

$$M_1 = \max_{0 \le t \le 1} B_1(t), \qquad m_1 = \min_{0 \le t \le 1} B_1(t). \tag{4}$$

For then the answer is simply $1 - P(0 < m_1 < M_1 < D/L')$. There are various ways to compute the joint distribution of the maximum and minimum in (4). One approach is to first compute the transition probabilities of the Brownian motion on $[0, D/L']$ with absorbing boundaries at 0, D/L' and starting at $0 < x < D/L'$. Since these transition probabilities satisfy the heat equation on the interior of the interval with Dirichlet boundary conditions one may compute the following eigenfunction expansion given in ([**2**], p. 412–413):

$$p(t; x, y) = 2\frac{L'}{D} \sum_{m=1}^{\infty} \exp\left\{-\frac{m^2\pi^2 t}{2}\left(\frac{L'}{D}\right)^2\right\} \sin\left(\frac{m\pi L' x}{D}\right) \sin\left(\frac{m\pi L' y}{D}\right),$$
$$\left(0 < x, y < \frac{D}{L'}\right). \tag{5}$$

Now the desired probability (3) is the complementary probability to that obtained by performing the indicated integrations in

$$P\left(0 < m_1 < M_1 < \frac{D}{L'}\right) = \frac{L'}{D}\int_0^{D/L'}\int_0^{D/L'} p(1; x, y)\, dy\, dx. \tag{6}$$

Here one also uses the familiar fact that the sum of reciprocal squares is $\pi^2/6$, so that restricting the sum to reciprocals of squares of odd numbers yields

$$\frac{\pi^2}{6} - \frac{1}{4}\frac{\pi^2}{6} = \frac{\pi^2}{8}.$$

FIGURE 2a contains a plot of the two probabilities (1) and (3) as a function of the ratio $x = L/D$. The corresponding result of a computer simulation is given in FIGURE 2b.

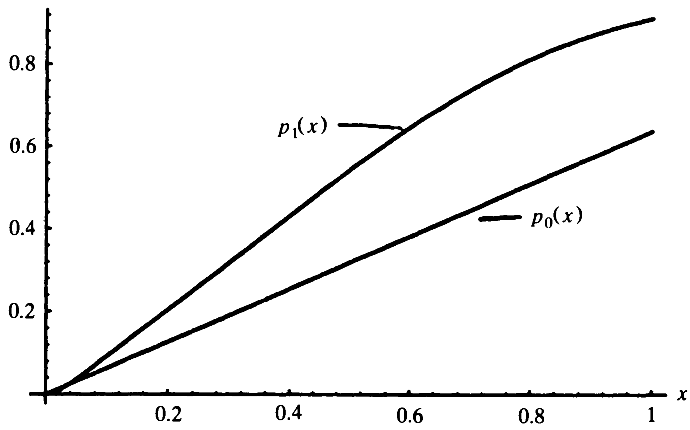

Figure 2a. Theoretical Brownian Noodle Crossing Probabilities $p_1(x)$.

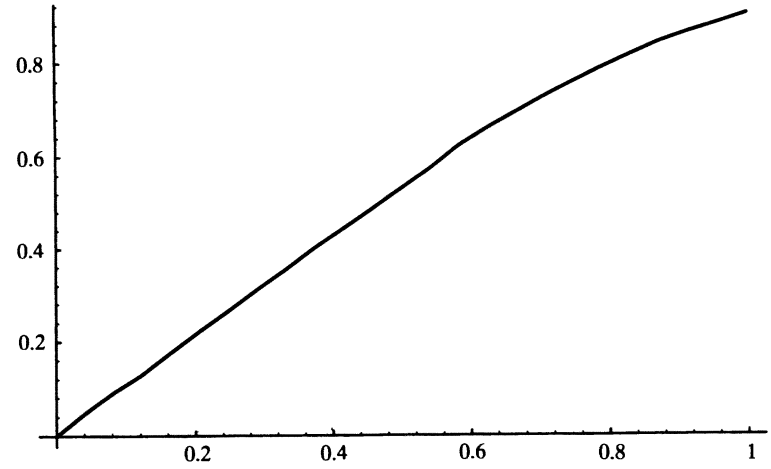

Figure 2b. 5000 Toss Simulated Crossing Frequencies for 100 Strand Brownian Noodle.

4. "REAL" NOODLES. Next let us consider a real string of n needles of lengths $L/\sqrt{n}$ each, and strung together in independent random directions. The total length is then $n\ (L/\sqrt{n}) = L\sqrt{n}$. In order to preserve the ratio as $x = L/D$, the gap between parallel lines is chosen as $D\sqrt{n}$. Since we already know the answer for $n = 1$ by (1), let's do the problem for $n \gg 1$. In this case we first note that the *functional central limit theorem* provides an approximation to the distribution of the string of needles by the Brownian noodle. In particular, observe that the lengths of the projections of the needle are i.i.d. with, up to the scale factor $L/\sqrt{n}$, the arcsine distribution $P(Y \in dy) = (1/\pi\sqrt{1-y^2})\,dy$. This distribution has mean 0 and variance 1/2. Thus the Brownian motion approximation has diffusion coefficient $L^2/2$. This explains the scaling factor $(\sqrt{2}/2)L$ introduced earlier.

5. FIRST ROUND APPROXIMATIONS. In the limit the approximation involves a noodle $\{S(t): 0 \le t \le 1\}$ of length ∞ in a gap of width ∞, but in the ratio $x = L/D$. This is how string theorists discussing quantum field theory can sound! One might "expect" the noodle crossing frequency to be estimated by $p_1(x/\sqrt{n})$ as this corresponds to formally replacing D by $\sqrt{n}\,D$ in (3). This curve is given in FIGURE 3a with $n = 100$. FIGURE 3b provides the crossing frequencies $\bar{p}_5(x)$ obtained from a simulation of 5000 tosses with $n = 100$ for comparison. However, since the effect of fixing the gap to length ratio in the above computation is a second rescaling of the string of noodles by $1/\sqrt{n}$, the central limit theorem approximation by the Brownian noodle is not justified. On the other hand, the formula (3) *is* applicable to strings composed of a large number n of needles of (fixed) lengths x each tossed onto a surface with gap width $\sqrt{n}$. This fact is illustrated by the simulation in FIGURE 2b of the corresponding crossing frequencies $\bar{p}_1(x)$ for a Brownian noodle with $n = 100$, $N = 5000$; i.e. these are the crossing frequencies of $-\sqrt{n}/2$ or $\sqrt{n}/2$ by a string of n randomly (uniformly) oriented needles of length x each with endpoint randomly distributed over $[-\sqrt{n}/2, \sqrt{n}/2]$.

Due to the additional rescaling of the gap width, there will be no loss in generality to consider the crossing probability of strings composed of n needles of length $x \le 1$ each and a gap of width n. A quick and clean (rigorous) upper bound

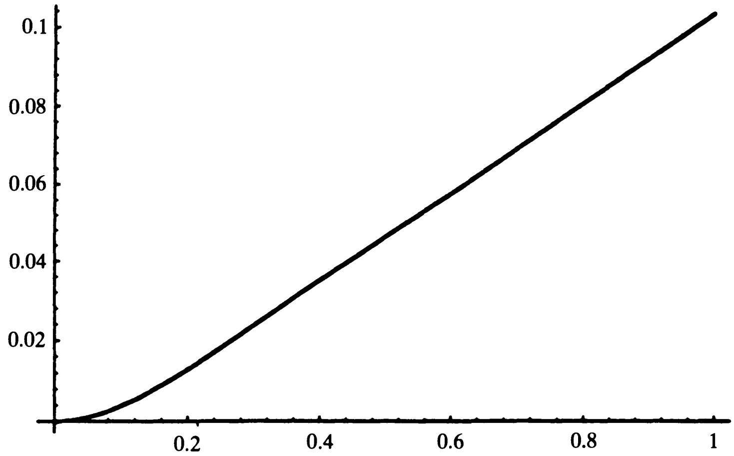

Figure 3a. Theoretical Brownian Noodle Crossing Probabilities $p_1(x/\sqrt{100})$.

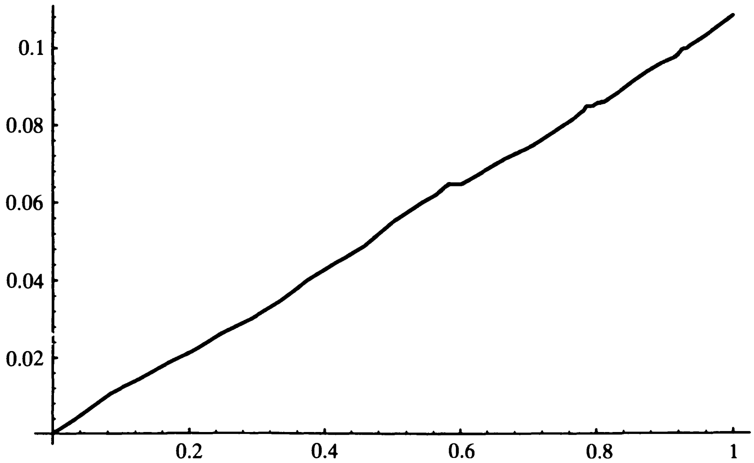

Figure 3b. 5000 Toss Simulated Crossing Frequencies for 100 Strand Noodle.

on the crossing probability can be obtained with *Doob's maximal inequality* ([**2**], pp. 49–52). In particular one obtains, using translation invariance of the problem to pose it with one end started uniformly in $[-n/2,\ n/2]$, that

$$P(C) \le \frac{4}{n}x^2 + \frac{2}{3}. \tag{7}$$

For large n this bounds the probability by the essentially constant value $2/3+$. Recall that the maximum crossing probability for the needle (1) is $2/\pi \approx 2/3-$. To make this estimate, let $\{S_n\}$ denote the random walk starting at 0 with displacement distribution $P(Y \in dy) = (1/\pi\sqrt{1-y^2})\,dy$, and let $Z^{(n)} \equiv nU$ be uniformly distributed on $[-n/2, n/2]$ independently of the random walk. The details of the distribution of Y do not play a role beyond first and second moments. The essential observation is that by symmetry $\{xS_k + Z^{(n)}\colon k = 0, 1, \dots\}$ is a *martingale*. Thus, Doob's inequality gives (7), after conditioning on $Z^{(n)}$ and using symmetry, as follows:

$$
\begin{aligned}
P(C) &= EP\left(\max_{k \le n}\left(\left\{xS_k + Z^{(n)}\right) \ge \frac{n}{2}\right\} \cup \left\{\min_{k \le n}\left(xS_k + Z^{(n)}\right) \le -\frac{n}{2}\right\} |Z^{(n)}\right) \\
&\le EP\left(\max_{k \le n}\left(xS_k + Z^{(n)}\right) \ge \frac{n}{2} |Z^{(n)}\right) + EP\left(\min_{k \le n}\left(xS_k + Z^{(n)}\right) \le -\frac{n}{2} |Z^{(n)}\right) \\
&= 2EP\left(\max_{k \le n}\left(xS_k + Z^{(n)}\right) \ge \frac{n}{2} |Z^{(n)}\right) \le 2\frac{4}{n^2} E\left(xS_n + Z^{(n)}\right)^2 \\
&= \frac{8}{n^2}\left(x^2 ES_n^2 + E\left(Z^{(n)}\right)^2\right) = \frac{8}{n^2}\left(\frac{x^2 n}{2} + \frac{n^2}{12}\right) = \frac{4}{n}x^2 + \frac{2}{3}. \qquad (8)
\end{aligned}
$$

A bound as universal as (7) will be much too large for many cases of interest. More precise estimates of the crossing probability upper bound are given in the next section.

For a lower bound, let $D(\beta)$ denote the event that the initial end of the string falls within βx, $\beta \le 1$, units from the boundary. Then $D(\beta)$ has probability $2\beta x/n$. Also, from the geometry $P(C|\mathrm{D}(\beta)) \ge (\mathrm{Cos}^{-1}(\beta)/\pi)$ since, for any n, crossing is implied by the occurence of the angle of the first needle between $\pm\mathrm{Cos}^{-1}(\beta)$. Thus, for any $0 \le \beta \le 1$, $P(C) \ge 2\beta x/n(\mathrm{Cos}^{-1}(\beta)/\pi)$ so that

$$
P(C) \ge \lambda(n)x, \quad \text{where } \lambda(n) = \frac{2}{n\pi} \max_{0 \le \beta \le 1} \beta\mathrm{Cos}^{-1}(\beta) \approx \frac{.357}{n}. \qquad (9)
$$

6. LARGE DEVIATIONS. As noted above, the functional central limit theorem approximation involved in replacing the string of needles by a Brownian motion for large n provides probabilities of events involving $O(1)$ fluctuations in the shape of the string of needles; the individual needle lengths already being $O(1/\sqrt{n})$. However, the crossing probabilities of $\sqrt{n}D$ involve $O(\sqrt{n})$ deviations from this picture which are not obtained by this approximation. These fluctuations have probabilities which also depend on the details of the angular distribution of the needles in the string. The precise probability computation desired here is a *large deviation problem* of the type which arises in the actuarial mathematics of insurance risk. In fact, the classic work of Cramer on large deviations was motivated by actuarial problems. Posed this way, the noodle problem also has an interesting twist in that it involves a "two-sided ruin event."

To see how the distribution will depend on the angle distribution, consider the degenerate case in which the angles made by the needles comprising the string with the horizontal are each $\pi/2$ (nonrandom). Then the noodle is a needle with fixed orientation and the only randomness is in the distribution of the endpoint; see FIGURE 4a. For this case one readily finds that $p_3(x) = P(C) = x$. (This prqbability is unchanged if one randomizes the choice of endpoint). To see how large deviation estimates go in a simple but less trivial situation for which explicit computations are still possible, consider the case in which the needles comprising the string make angles $\pm\pi/4$ with the horizontal with equal probabilities; see FIGURE 4b for a sample outcome of a simulation of a single toss with $n = 100$. Since the variance in step size is one in this case, the rescaling by $\sqrt{2}/2$ is unnecessary for the comparison with the (nonrigorous) central limit theorem prediction; i.e. replace x by $x/2$ in FIGURE 3a to correct for the $\sqrt{2}/2$ before comparing with the observed FIGURE 5. In this case the lengths of the projections of the needle are i.i.d. with, up to the scale factor L, a Bernoulli distribution

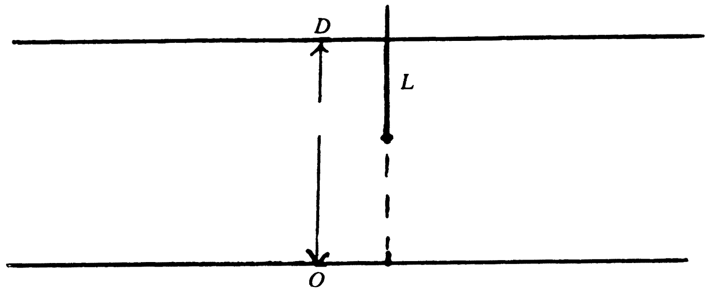

Figure 4a. Sample 0 Degree Single Strand Noodle Realization.

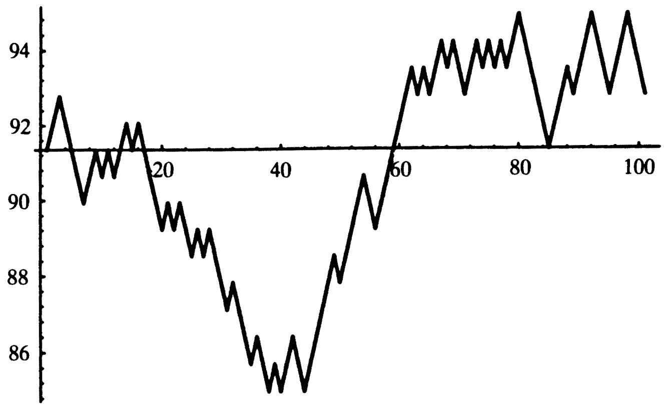

Figure 4b. Sample 100 Strand 45 Degree Noodle Realization.

$P(Y_1 = +1) = P(Y_1 = -1) = 1/2$. Let $S_n := (Y_1 + Y_2 + \cdots + Y_n)$, $S_0 = 0$. More generally, assume $\phi(s) = Ee^{sY_1}$ finite and note for i.i.d. $\{Y_n\}$, $(\phi(s))^n = Ee^{sS_n} \geq e^{nsb}P(S_n \geqslant nb)$. That is, for arbitrary s

$$P(S_n \geq nb) \leq e^{n\{\log(\phi(s)) - sb\}}. \tag{10}$$

In particular,

$$P(S_n > nb) \leq e^{n\chi^*(b)}, \quad \text{where} \quad \chi^*(b) = \inf_s(\log Ee^{sY} - sb). \tag{11}$$

The reverse (asymptotic) inequality for $b > EY_1$ yielding

$$\frac{1}{n}\log P(S_n > nb) \to \chi^*(b) \quad \text{as } n \to \infty, \tag{12}$$

is the often cited *Cramer-Chernoff theorem* on large deviations of S_n/n from EY_1. That is to say, $P(S_n > nb) \sim e^{n\chi^*(b)}$ as $n \to \infty$; the rate $-\chi^*(b)$ is a *large deviation rate*. The reader may consult the developments given in ([3], pp. 147–148) and ([4], pp. 57–63) for some basics of large deviation theory.

Using the reflection principle for simple symmetric random walk one may also check that ([2], p. 10, 70), $P(\max_{k \leq n} S_k \geq b) = 2P(S_n \geq b) - P(S_n = b)$. In particular one obtains from this,

$$P\Big(\max_{k \leq n} S_k \geq b\Big) \leq 2P(S_n \geq b). \tag{13}$$

These results will serve as our basic tools. Notice that for large n it is no more than a venal sin to think of the inequality (13) as equality; i.e. the resulting probability bound may be expected to be close to the actual probability.

First let us compute the so-called *Legendre transform* $\chi^*(b)$ of the function $\chi(s) = \log Ee^{sY}$ required in (11). One may easily check by calculating limits as $s \to \pm\infty$ that $\chi^*(b) = -\infty$ if $|b| > 1$. For $|b| < 1$, equating the derivative with respect to s of $\{\log((e^s + e^{-s})/2) - sb\}$ to 0 and solving for s gives $e^{2s} = ((1 + b)/(1 - b))$ and therefore

$$\chi^*(b) = -\frac{1+b}{2}\log(1+b) - \frac{1-b}{2}\log(1-b). \tag{14}$$

As before let $Z^{(n)} \equiv nU$ be uniformly distributed over $[-n/2, n/2]$ and independent of $\{S_n\}$. Then one has along the lines of (8) that

$$P(C) \le 2P\left(\max_{k\le n} S_k \ge n\left\{\frac{1}{2x} + \frac{U}{x}\right\}\right) \le 4P\left(S_n \ge n\left\{\frac{1}{2x} + \frac{U}{x}\right\}\right), \tag{15}$$

first using the symmetry of the distribution of both $\{S_n\}$ and U, and then the reflection principle. By conditioning on U, using (15), the argument for (11) conditionally, and then (14), one has after taking expected values (with respect to U),

$$\begin{aligned} P(C) &\le 4EP\left(S_n \ge n\left(\frac{1}{2x} + \frac{U}{x}\right)|U\right) \le 4Ee^{n\chi^*(1/2x+U/x)} = 4x\int_0^{1/x} e^{n\chi^*(y)}\,dy \\ &= 4x\int_0^1 \exp\left\{-\frac{n}{2}[(1+y)\log(1+y) + (1-y)\log(1-y)]\right\}dy := \gamma_4(n)x. \end{aligned} \tag{16}$$

Note that the (concave) function

$$\chi(y) = -(1/2)[(1+y)\log(1+y) + (1-y)\log(1-y)]$$

and its first derivative are zero at $y = 0$. The second derivative at 0 is -1. As a result one may bound by a Taylor's approximation $-(1/2)y^2$ to obtain by a simple change of variable and the fact that

$$\int_0^{\sqrt{n}} e^{-(1/2)z^2}\,dz \sim \sqrt{2\pi}/2,$$

$$\begin{aligned} \gamma_4(n) &= 4\int_0^1 \exp\left\{-\frac{n}{2}[(1+y)\log(1+y) + (1-y)\log(1-y)]\right\}dy \\ &\le 2\sqrt{2\pi}\,n^{-1/2} \approx 5.01n^{-1/2}. \end{aligned} \tag{17}$$

Observe from this that the crossing probabilities are theoretically predicted to be smaller than the universal bound (7) for 45 Degree strings consisting of n strands. FIGURE 5 depicts the crossing frequencies $\bar{p}_4(x)$ for a simulation of $N = 1000$ tosses with $n = 100$. A sample realization of a toss was plotted in FIGURE 4b. Finally, observe that the argument used to arrive at the lower bound (9) can be simply adapted to this case to give

$$\frac{x}{n} \le P(C) \le \gamma_4(n)x \sim \frac{2\sqrt{2\pi}}{\sqrt{n}}x. \tag{18}$$

To extend these estimates to the case of uniform angles on $[0, 2\pi)$ first requires a general inequality of the form (15). Such an extension is possible for symmetric

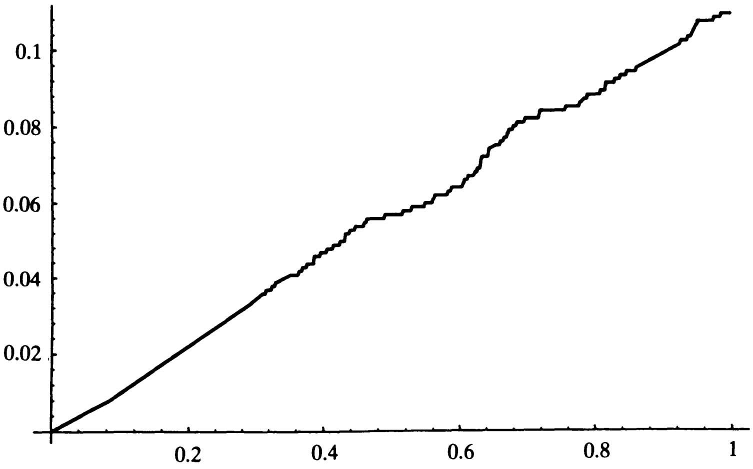

Figure 5. 5000 Toss Simulation Crossing Frequencies for 100 Strand 45 Deg Noodle.

random variables (*Levy's inequality*), ([**5**], p. 14; [**2**], Exc. #4, p. 70). The corresponding bound in place of (16) for this case is, following the first two inequalities of (16),

$$P(C) \leq 4Ee^{n\chi^*(1/2x+U/x)} = 4x\int_0^{1/x} e^{n\chi^*(y)}\,dy = 4x\int_0^1 e^{n\chi^*(y)}\,dy, \tag{19}$$

where now for $|y| > 1, \chi^*(y) = -\infty$, but for $|y| < 1$

$$\chi^*(y) = \inf_s \left\{ \log\left(\frac{1}{\pi}\int_{-1}^{1} \frac{e^{st}}{\sqrt{1-t^2}}\,dt \right) - sy \right\}. \tag{20}$$

The crossing frequencies $\bar{p}_5(x)$ obtained from a simulation of 5000 tosses with $n = 100$ strands were provided earlier in FIGURE 3b. To obtain the asymptotic value of $\gamma_5(n)$ one may proceed as in the above case to compute the values of $\chi^*(y)$ and its first two derivatives at $y = 0$. This is most easily accomplished by the following Legendre transform *duality equations*: Let

$$\chi(s) = \frac{1}{\pi}\int_{-1}^{1} \frac{e^{st}}{\sqrt{1-t^2}}\,dt = \frac{1}{\pi}\int_{-(\pi/2)}^{\pi/2} e^{s\,\mathrm{Sin}\,\theta}\,d\theta. \tag{21}$$

Then

$$\chi^*(y) = \chi(s(y)) - ys(y), \qquad \frac{d\chi(s)}{ds} = y, \qquad \frac{d\chi^*(y)}{dy} = -s. \tag{22}$$

One may check from this that $\chi^*(y)$ and its first derivative are 0 at $y = 0$ and the second derivative is $-1/2$. It follows as above that

$$\gamma_5(n) = 4\int_0^1 e^{n\chi^*(y)}\,dy \leq 4\sqrt{\pi}\,n^{-1/2} \approx 7.09n^{-1/2}. \tag{23}$$

A lower bound was given at (9). In this case the crossing probability prediction falls below the bound (7).

What we have presented here is the stuff with which probability bounds and estimates can be made. The computer simulations provide an order of magnitude feel for crossings, but no more than this for n large. The reader desiring a more precise recipe for "n-noodle pi" may want to give more careful consideration to the small n cases, beginning with $n = 2$. We only hope to have whet the appetite for more probability and noodles.

REFERENCES

1. Adhikari, A. and J. Pitman (1989): The Shortest Plane Arc of Width 1, *American Math. Monthly*, 309–327.
2. Bhattacharya, R. and E. Waymire (1990): *Stochastic Processes with Applications*, John Wiley and Sons, NY.
3. Billingsley, P. (1986): *Probability and Measure*, John Wiley and Sons, NY.
4. Durrett, R. (1991): *Probability: Theory and Examples*, Wadsworth and Brooks/Cole, Pacific Grove, CA.
5. Kahane, J-P. (1985): *Some Random Series of Functions*, 2nd ed, Cambridge University Press, NY.
6. Ramaley, J. F. (1969): Buffon's Noodle Problem, *American Math. Monthly*, no. 8, 916–918.
7. Snell, J. L. (1988): *Introduction to Probability*, Random House/Birkhauser, Boston.

Department of Mathematics
Oregon State University
Corvallis, OR 97331
waymire@math.orst.edu

PICTURE PUZZLE

(from the collection of Paul Halmos)

What can these men talk about?
(see page 570)