

On Approximating Univariate NP-Hard Integrals

Ohad Asor* and Avishy Carmi†

1 Jan 2016

Abstract

Approximating $I_{\#PART} = \int_0^1 \prod_{k=1}^n \cos(x_k \pi t) dt$ to within an accuracy of 2^{-n} where the input integers $\{x_k\}_{k=1}^n$ are given in binary radix, is equivalent to counting the number of equal-sum partitions of the integers $\{x_k\}$ and is thus a $\#P$ problem. Similarly, integrating this function from zero to infinity and deciding whether the result is either zero or infinity is an NP-Complete problem. Efficient numerical integration methods such as the double exponential formula and the sinc approximation have been around since the mid 70's. Noting the hardness of approximating $I_{\#PART}$ we argue that the proven rates of convergence of such methods cannot possibly be correct since they give rise to an anomalous result as $P=\#P$.

1 Background

The Partition Counting Problem ($\#PART$) is the following: given n positive integers $\{x_k\}_{k=1}^n$, in how many ways is it possible to divide them into two equal-sum subsets. If the input $\{x_k\}$ is given in binary rather unary radix, then solving this problem in worst-case subexponential time wrt the input's length would prove $P=\#P$. Merely deciding whether an equal-sum partition exists is an NP-Complete problem. Assuming the exponential time hypothesis, $\#PART$ in its strong form (as we define more accurately below) cannot be solved in subexponential time.

The treatment in [1] and subsequently in other places e.g. [11, 12, 13] express the number of equal-sum partitions by the integral

$$2^n I_{\#PART} = 2^{n-1} \int_0^1 \prod_{k=1}^n \cos(x_k \pi t) dt$$

The decision version (NP-Complete) of the partition problem is deciding whether $I_{\#PART} = 0$ (or equivalently $I_{\#PART} < 2^{-n}$). It can also be expressed by taking this integral from 0 to ∞ and deciding whether it equals zero or infinity, as it must be either one.

The double-exponential (DE) tanh-sinh quadrature is a numerical integration technique whose its and its flavors convergence rate proven to be at least exponential with the number of evaluation points [2, 3, 7, 8, 9, 10]. It is currently considered as the fastest high-precision quadrature technique and a large amount of experimental results are available¹.

The DE analysis has led to numerous other results, such as the near optimality of sinc approximation discussed on this paper, that results with exponential convergence bounds as well.

*ohadasor@gmail.com

†Center for Quantum Information Science and Technology & Faculty of Engineering Sciences, Ben-Gurion University of the Negev, Beer-Sheva 8410501, Israel. avishycarmi@gmail.com

¹It is interesting to see Problem 8 in [14] which is somehow related to our discussion.

Those results imply that the answer to #PART problem (and even #SAT) can be expressed as a sum of a subexponential number of summands. Noting the hardness of approximating $I_{\#PART}$ we argue here that, unless $P=\#P$, the DE and Sinc convergence rate as stated in [2, 3, 7, 8, 9, 10] cannot be correct.

2 The Partition Problem

Given $n \in \mathbb{N}$ and $\{x_k\}_{k=1}^n \subset \mathbb{Z}$, we seek $\sigma \in \{-1, 1\}^n$ such that $\langle \sigma, \mathbf{x} \rangle = 0$, where $\langle \sigma, \mathbf{x} \rangle = \sum_{k=1}^n \sigma_k x_k$ denotes the inner product. Deciding whether such σ exists is an NP-Complete problem, while counting how many such σ 's exists, is in #P. That is true only when the inputs $\{x_k\}$ are given in binary radix² and we denote by d_k the number of binary digits of x_k . The partition problem is known to be Weak-NP since it has a polynomial-time algorithm if the input is supplied in unary radix, which is exponentially large than the binary one. To get a feeling about typical dimensions of hard problems, the reduction of c -clause and k -variables 3SAT into the partition problem ends up with $\mathcal{O}(c+k)$ integers to partition, each having up to $\mathcal{O}(c+k)$ digits [8]. The exponential time hypothesis therefore implies that it is impossible to solve the partition problem in runtime complexity of $\mathcal{O}(\text{poly}(\sum_{k=1}^n d_k))$. We therefore keep in mind considering the strong partition problem setting that we have three dependencies on n :

1. The integrand consists of product of n cosines.
2. Every cosine is of a natural number with up to n digits multiplied by the variable of integration.
3. The integral has to be approximated up to precision of n binary digits.

Note that 2 and 3 can sometimes be relaxed as in Corollary 2 below.

Lemma 1. *Let $\{x_k\}_{k=1}^n \subset \mathbb{Z}$ be integers given in binary radix. Let also $\psi(t) = \prod_{k=1}^n \cos(\pi x_k t)$. Then evaluating $I_{\#PART} = \int_0^1 \psi(t) dt$ up to accuracy of n binary digits is in #P.*

Proof. This lemma can be proved in many interesting ways, many of them seem to go back to the classical monograph by Kac [1]. Slightly different proofs of this lemma may be found in [12, 11]. Our derivation is based on the formula

$$\prod_{k=1}^n \cos(z_k) = 2^{-n} \sum_{\sigma \in \{-1, 1\}^n} \cos \langle \sigma, \mathbf{z} \rangle \quad (1)$$

for every $\mathbf{z} \in \mathbb{C}^n$, which follows from a repeated application of the identity

$$4 \cos(z_1) \cos(z_2) = \cos(z_1 + z_2) + \cos(z_1 - z_2) + \cos(-z_1 + z_2) + \cos(-z_1 - z_2) \quad (2)$$

Using this and the cosine being an even function, the integral reads

$$\begin{aligned} I_{\#PART} &= 2^{-n-1} \sum_{\sigma \in \{-1, 1\}^n} \int_{-1}^1 \cos(\pi t \langle \sigma, \mathbf{x} \rangle) dt = 2^{-n} \sum_{\sigma \in \{-1, 1\}^n} \frac{\sin \pi \langle \sigma, \mathbf{z} \rangle}{\pi \langle \sigma, \mathbf{z} \rangle} \\ &= 2^{-n} \sum_{\sigma \in \{-1, 1\}^n} \begin{cases} 1 & \text{if } \langle \sigma, \mathbf{z} \rangle = 0 \\ 0 & \text{if } \langle \sigma, \mathbf{z} \rangle \neq 0 \end{cases} \quad (3) \end{aligned}$$

²Or other encodings with similar order of magnitude.

Thus, $I_{\# \text{PART}}$ is precisely the fraction of zero partitions for $\{x_k\}_{k=1}^n$ divided by 2^n . This also shows why an accuracy of 2^{-n} is required. \square

From Theorem 1, the periodicity of the integrand, and the partition problem being agnostic to multiplication of the input by a constant, it directly follows that:

Corollary 2. *Let $\{x_k\}_{k=1}^n \subset (0, 1)$ be rationals. Then they have an equal sum partition if and only if $\int_0^\infty \cos(x_k t) dt = 0$. Moreover, if and only if $\int_0^\infty \cos(x_k t) dt \neq \infty$.*

3 State Of The Art Univariate Approximations

Following (2.b.10) and (2.b.15) in [2], Takahasi&Mori show that there exists positive $\tau \approx \frac{\pi}{2}$ such that for all f analytic at $(-1, 1)$, putting

$$I = \int_{-1}^1 f(t) dt = \int_{-\infty}^{\infty} \frac{\cosh t}{\cosh^2\left(\frac{\pi}{2} \sinh t\right)} f\left(\tanh \frac{\pi}{2} \sinh t\right) dt \quad (4)$$

$$\hat{I} = \frac{\pi}{2} h \sum_{m=-N}^N \frac{\cosh(mh)}{\cosh^2\left(\frac{\pi}{2} \sinh(mh)\right)} f\left(\tanh \frac{\pi}{2} \sinh(mh)\right) \quad (5)$$

then

$$|I - \hat{I}| = \mathcal{O}\left(e^{-\frac{2\pi\tau}{h}}\right) + \mathcal{O}\left(e^{-\frac{2\pi\tau N}{\ln(4\tau N)}}\right) \quad (6)$$

showing that the integral $I_{\# \text{PART}}$ can be expressed as a sum of subexponential number of summands, since the desired accuracy is 2^{-n} even at the Strong-#P setting of the problem. Since Takahasi&Mori result a lot of research was done on this field up till today. Contemporary methods suggest combining the DE method with the Sinc approximation. On [7],[8], Sugihara showed the near-optimality of the sinc approximation

$$f(x) \approx \sum_{m=-n}^n f(mh) \frac{\sin\left[\frac{\pi}{h}(x - mh)\right]}{\frac{\pi}{h}(x - mh)} \quad (7)$$

providing existing and new lower and upper bounds.

Remark 3. Besides the results mentioned on this section, it is interesting to see that our specific integrand has an intimate relation to the sinc function, even beyond what can be seen in formula (3). Recall that sinc is the reproducing kernel of the Hilbert space of so-called Bandlimited functions³ and our integrand is a bandlimited function indeed, as can be seen when written as a sum of 2^n cosines. c.f. also Paley-Weiner theorem for a wider treatment of bandlimited functions.

Quoting some definitions and notations from [7]:

Definition 4. Let $1 < p \leq \infty$ and $d > 0$:

1. The $2d$ -wide strip \mathcal{D}_d is defined as $\mathcal{D}_d = \{z \in \mathbb{C} \mid |\text{Im}z| < d\}$. Note that $\mathbb{R} \subset \mathcal{D}_d$.
2. $B(\mathcal{D}_d)$ is the set of all f analytic in \mathcal{D}_d satisfying $\lim_{x \rightarrow \pm\infty} \int_{-d}^d |f(x + iy)| dy = 0$ and $\lim_{y \rightarrow d^-} \int_{-\infty}^{\infty} |f(x + iy)| + |f(x - iy)| dx < \infty$.

³Functions with Fourier transform having finite-measure support.

3. If f is analytic in \mathcal{D}_y for some y and setting q such that $p^{-1} + q^{-1} = 1$, we denote

$$\mathcal{N}_p^*(f, y) = \quad (8)$$

$$\begin{cases} \sqrt[p]{\int_{-\infty}^{\infty} |f(x + iy)|^p |\cosh(x + iy)|^{\frac{2p}{q}} + |f(x - iy)|^p |\cosh(x - iy)|^{\frac{2p}{q}} dx} & 1 < p < \infty \\ \sup_{z \in \mathcal{D}_y} |f(z) \cosh^2 z| & p = \infty \end{cases}$$

and define $\mathbf{H}_p^*(\mathcal{D}_{\frac{\pi}{4}})$ to be the set of functions f such that f is analytic in $\mathcal{D}_{\frac{\pi}{4}}$, $\mathcal{N}_p^*(f, y) < \infty$ for $0 \leq y \leq \frac{\pi}{4}$, and requiring the following norm $\|f\|_p^* \equiv \lim_{y \rightarrow \frac{\pi}{4}-0} \mathcal{N}_p^*(f, y)$ to be finite.

4. Let $N = 2n + 1$ and denote

$$E_{N,h}^{\text{sinc}}[\mathbf{H}_p^*(\mathcal{D}_{\frac{\pi}{4}})] = \sup_{\|f\|_p^* \leq 1} \left\{ \sup_{x \in \mathbb{R}} \left| f(x) - \sum_{j=-n}^n f(jh) \frac{\sin\left[\frac{\pi}{h}(x - jh)\right]}{\frac{\pi}{h}(x - jh)} \right| \right\} \quad (9)$$

while $E_{N,h}^{\text{min}}$ goes over all the following N -point approximations that might consider also f 's derivatives:

$$E_{N,h}^{\text{min}}[\mathbf{H}_p^*(\mathcal{D}_{\frac{\pi}{4}})] = \inf_{1 \leq \ell \leq N} \inf_{\{m_k\}_{k=1}^{\ell}, \sum_k m_k = N} \inf_{a_j \in \mathcal{D}_{\frac{\pi}{4}}, i \neq j \implies a_i \neq a_j} \inf_{\phi_{jk}} \left\{ \sup_{\|f\|_p^* \leq 1} \left\{ \sup_{x \in \mathbb{R}} \left| f(x) - \sum_{j=1}^{\ell} \sum_{k=0}^{m_j-1} f^{(k)}(a_j) \phi_{jk}(x) \right| \right\} \right\} \quad (10)$$

where ϕ_{jk} are required to be analytic on $\mathcal{D}_{\frac{\pi}{4}}$. The a_j 's represent the node points, ϕ_{jk} 's the basis functions, and m_j 's the maximal order of derivatives.

Remark 5. The definition of $E_{N,h}^{\text{min}}$ gains more interest at the scope $I_{\# \text{PART}}$ since as we take higher derivatives of the integrand, we result with binomially large formulas such that their length would eventually sum to exactly 2^n . In fact, if we could efficiently calculate the Taylor coefficients of our product-of-cosines integral, then we could solve the $\# \text{PART}$ at that same efficiency.

Before stating the main theorem in [7] we first state Theorem 1.1 there quoting results from around the 1970's. They give exponential convergence bounds to the sinc approximation and showing that it is close to the best possible approximation at following sense:

Theorem 6. *Let $1 < p < \infty$, then there exists constants C_p, C_p' such that for all $f \in \mathbf{H}_p^*(\mathcal{D}_{\frac{\pi}{4}})$*

$$C_p N^{-\frac{1}{2p}} e^{-\pi \sqrt{\frac{N}{2q}}} \leq E_{N,h}^{\text{min}}[\mathbf{H}_p^*(\mathcal{D}_{\frac{\pi}{4}})] \leq E_{N,h}^{\text{sinc}}[\mathbf{H}_p^*(\mathcal{D}_{\frac{\pi}{4}})] \leq C_p' \sqrt{N} e^{-\frac{\pi}{2} \sqrt{\frac{N}{q}}} \quad (11)$$

where $p^{-1} + q^{-1} = 1$, $N = 2n + 1$ and $h = \frac{\pi}{2} \sqrt{\frac{q}{2n}}$. If $p = \infty$ then there exists constants $\nu, C_{\infty}, C_{\infty}'$ satisfying

$$C_{\infty} N^{\nu} e^{-\pi \sqrt{N}} \leq E_{N,h}^{\text{min}}[\mathbf{H}_{\infty}^*(\mathcal{D}_{\frac{\pi}{4}})] \leq E_{N,h}^{\text{sinc}}[\mathbf{H}_{\infty}^*(\mathcal{D}_{\frac{\pi}{4}})] \leq C_{\infty}' \sqrt{N} e^{-\frac{\pi}{2} \sqrt{N}} \quad (12)$$

with $h = \frac{\pi}{2} \sqrt{\frac{1}{2n}}$.

It should be noted that the space $\mathbf{H}_p^*(\mathcal{D}_{\frac{\pi}{4}})$ contains essentially analytic functions that decay exponentially fast at infinity and its imaginary neighbourhood. We can see how the exponential convergence of the DE family is implicitly considered in Theorem 7, while showing that

sinc approximation has the same order of magnitude. Nevertheless, our integrand $I_{\# \text{PART}}$ does not decay and we need to change the variable to produce exponentially or double-exponentially decaying integrand (e.g. the classical DE transforms) before we apply the sinc method and enjoy the convergence guarantee. The transform translates the sum in (10) into the form $\sum_{m=-n}^n f(\tau(mh)) \frac{\sin[\frac{\pi}{h}(\tau^{-1}(x)-mh)]}{\frac{\pi}{h}(\tau^{-1}(x)-mh)}$. Full derivation of this procedure is available on [7], section 5, remark 2.

Sugihara [7] generalized Theorem 9 for a wider class of asymptotic decays (while for $I_{\# \text{PART}}$ change of variable would still be required), yet concludes that sinc is near-optimal even then. Define $\mathbf{H}^\infty(\mathcal{D}_d, \omega)$ to be the space of all $f : \mathcal{D}_d \rightarrow \mathbb{C}$ such that f is analytic in \mathcal{D}_d and $\|f\| \equiv \sup_{z \in \mathcal{D}_d} \left| \frac{f(z)}{\omega(z)} \right| < \infty$. Then such f decays like ω at infinity in \mathcal{D}_d , by definition. We then define $E_{N,h}^{\text{sinc}}[\mathbf{H}^\infty(\mathcal{D}_{\frac{\pi}{4}})]$, $E_{N,h}^{\min}[\mathbf{H}^\infty(\mathcal{D}_{\frac{\pi}{4}})]$ as in (10) and (11), only requiring f to be in $\mathbf{H}^\infty(\mathcal{D}_{\frac{\pi}{4}})$. Summarizing theorems 3.1, 3.2, 3.3 in [13]:

Let $\omega(z) \in B(\mathcal{D}_d)$ be nonzero in all \mathcal{D}_d . Then

1. It is impossible for ω to satisfy $w(x) = \mathcal{O}(e^{-\beta e^{\gamma|x|}})$ as $|x| \rightarrow \infty$, for $\beta > 0$ and $\gamma > \frac{\pi}{2d}$.
2. If the decay rate of ω on the real axis satisfies

$$\alpha_1 e^{-(\beta|x|)^\rho} \leq |\omega(x)| \leq \alpha_2 e^{-(\beta|x|)^\rho} \quad (13)$$

for $\alpha_1, \alpha_2, \beta > 0, \rho \geq 1, x \in \mathbb{R}$, then

$$E_{N,h}^{\text{sinc}}[\mathbf{H}^\infty(\mathcal{D}_{\frac{\pi}{4}})] \leq C_{d,\omega} N^{\frac{1}{1+\rho}} e^{-(\frac{1}{2}\pi d \beta N)^{\frac{\rho}{1+\rho}}} \quad (14)$$

where $h = (\pi d)^{\frac{1}{1+\rho}} (\beta n)^{\frac{-\rho}{1+\rho}}$ and $C_{d,\omega}$ is a constant depending on d, ω , and

$$E_{N,h}^{\min}[\mathbf{H}^\infty(\mathcal{D}_{\frac{\pi}{4}})] \geq C'_{d,\omega} e^{-\left(\frac{2(\pi d \beta N)^\rho}{1+\rho}\right)^{\frac{1}{1+\rho}}} \quad (15)$$

where $C'_{d,\omega}$ is another constant depending on d, ω .

3. If the decay rate of ω on the real axis satisfies

$$\alpha_1 e^{-\beta_1 e^{\gamma|x|}} \leq |\omega(x)| \leq \alpha_2 e^{-\beta_2 e^{\gamma|x|}} \quad (16)$$

for $\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma > 0, x \in \mathbb{R}$, then

$$E_{N,h}^{\text{sinc}}[\mathbf{H}^\infty(\mathcal{D}_{\frac{\pi}{4}})] \leq C_{d,\omega} e^{-\frac{\pi d \gamma N}{2 \ln \frac{\pi d \gamma N}{2 \beta_2}}} \quad (17)$$

where $h = \frac{1}{\gamma n} \log \frac{\pi d \gamma n}{\beta_2}$ and $C_{d,\omega}$ is a constant depending on d, ω , and

$$E_{N,h}^{\min}[\mathbf{H}^\infty(\mathcal{D}_{\frac{\pi}{4}})] \geq C'_{d,\omega} e^{-\frac{\pi d \gamma N}{2 \ln \frac{\pi d \gamma N}{2 \beta_1}}} \quad (18)$$

where $C'_{d,\omega}$ is another constant depending on d, ω .

See also [8] for more discussion and contemporary results on both DE and Sinc bounds. On [9, 10] Okayama calculates improved bounds with explicit constants for both the DE formula and the Sinc approximation. We do not bring the explicit expressions here but the reader can verify by [9, 10] that the constants are indeed independent of the integrand and does not seem to be unpractically large.

4 Local Complexities

After showing that the solution to the $\#SAT^4$ problem is claimed to be able to be represented as a sum of a subexponential number of terms wrt the problem's size, we have to show that the complexity of computing each term is low enough. If we want to approximate $I_{\#PART}$ up to accuracy of 2^{-n} , then by Kahan summation ([13, 14]) we can compute the integrand only up to that accuracy without loss of accuracy during summation, paying only in doing about four times more operations per addition. So we have to see how to efficiently calculate a product of n cosines up to the n 'th binary digit.

Denote by $M(n)$ the complexity of multiplying two n -digit numbers up to accuracy of $2n$. Then multiplying three numbers can be done by multiplying the first two in no more than $M(n)$ and taking only the first n digits of the result. Afterwards we're left again with two n -digit numbers to multiply, ending with total complexity of no more than $2M(n)$. Continuing this way, the complexity of multiplying n numbers up to precision of n digits takes no more than $\mathcal{O}((n-1)M(n))$. Note that the multiplicands need not be more accurate than n digits, since higher digits won't impact lower digits in the result as long as we multiply numbers in $(0, 1)$.

We're now left with how to compute transcendental functions arising in the integrand according to our derivation (sinc, cos, exp) up to accuracy of n digits. Observe that $e^{2^{-n}} = \sum_{k=0}^{\infty} \frac{1}{k!} 2^{-nk}$ prescribes the digits of the exponent's result nicely right away up to a single division, as can be seen on the cosine's or sinc's Taylor series as well. It can also be achieved directly from the input's digits, by writing $x = \sum_k d_k 2^{-k} \implies \prod_k \left(e^{2^{-k}}\right)^{d_k}$ where in binary we have $d_k \in \{0, 1\}$, suggesting $e^{2^{-n}}$ to be precomputed. We also note that the formulas for $\cos(a+b)$, $\sin(a+b)$ can be applied to calculate the trigonometric functions of n -digit binary number in linear amount of arithmetic operations, by simply following its 1 digits and taking $b = 2^{-k}$ for all k up to n . Therefore computing the transcendental functions in concern is $\mathcal{O}(nM(n))$ per one input, so we end up with complexity of maximum $\mathcal{O}(n^2 M^2(n)) \approx \mathcal{O}(n^5)$ per computing the integrand once up to the desired accuracy.

5 Conclusion

We have pointed out that the DE numerical integration convergence rate together with contemporary associated approximation methods should be reexamined for they currently suggest the existence of a polynomial time solution to a $\#P$ problem, by expressing the solution in a subexponential number of terms, each of them having polynomial complexity.

6 Acknowledgments

The authors thank HunterMinerCrafter for many valuable discussions. We also thank Prof. Brendan McKay for his comments on MathOverflow.

References

- [1] Kac, "Statistical Independence in Probability, Analysis and Number Theory". Carus Mathematical Monographs, No. 12, Wiley, New York (1959) 1, 2

⁴We used the setting of $\#PART$ as obtained after using the current known reduction from $\#SAT$ to $\#PART$, as mentioned in section 2.

- [2] Takahasi, Mori, “Double Exponential Formulas for Numerical Integration” Publications of the Research Institute for Mathematical Sciences, Vol. 9 (3), pp. 721 – 741. (1974). 1, 3
- [3] Trefethen, Weideman, “The Exponentially Convergent Trapezoidal Rule”. SIAM Review, Vol. 56 (3) pp. 385–458. (2014) 1
- [4] Kahan, “Further remarks on reducing truncation errors”. Communications of the ACM, 8 (1): 40. (1965). 1, 4
- [5] Higham, “The accuracy of floating point summation”. SIAM Journal on Scientific Computing, 14 (4): 783–799. (1993). 1, 4
- [6] Sipser, “Introduction to the Theory of Computation”. International Thomson Publishing (1996). 1, 2, 3, 3
- [7] Sugihara, “Near Optimality of the Sinc Approximation”, Mathematics of Computation, Volume 72, number 242 pages 767-786 (2002) 1, 3, 3, 3, 3
- [8] Okayama, Tanaka, Matsuo, Sugihara, ”The Double-Exponential Transformation is Not Always Better than the Tanh transformation – Theoretical Convergence Analysis”, Mathematical engineering technical reports; METR 2011-43 1, 2, 3, 3
- [9] Okayama, “Error estimates with explicit constants for Sinc quadrature and Sinc indefinite integration over infinite intervals, Reliable Computing, Vol. 19 (2013), pp. 45–65 1, 3
- [10] Okayama, ”Explicit error bound for the tanh rule and the DE formula for integrals with logarithmic singularity” JSIAM Letters Vol.6 (2014) pp.9–11 2014 1, 3
- [11] Johnson, Garey, “Computers and Intractability”. W. H. Freeman and Company (1979). 1, 2
- [12] Moore, Mertens, “The Nature of Computation”. Oxford University Press, Inc., New York, NY, USA. (2010). 1, 2
- [13] Krantz, “Handbook of Logic and Proof Techniques for Computer Science”. Springer-Verlag New York, NY, USA. (2002). 1, 4
- [14] Bailey, Borwein, Kapoor, Weisstein, ”Ten Problems in Experimental Mathematics” (2006) 1, 4