Probabilistic Rounding Error Analysis for Sums

Eric Hallman

North Carolina State University

Spring 2020

Introduction

The problem: for $x_1, \ldots, x_n \in \mathbb{R}$, compute

$$s_n := \sum_{i=1}^{n} x_i$$

Compute the sum using recursive summation:

$$\begin{aligned} \hat{s}_1 &:= x_1, \\ \hat{s}_i &:= \mathrm{fl}\big(\hat{s}_{i-1} + x_i\big), \quad i = 2 : n. \end{aligned}$$

The goal: bound the error $|\hat{s}_n - s_n|$.

Introduction

Classical model for floating-point arithmetic:

Model (Classical)

For any floating point numbers a and b,

$$\mathrm{fl}(a \odot b) = (a \odot b)(1+\delta), \quad |\delta| \le u, \quad \odot \in \{+, -, \times, /, \surd\},$$

where u is the unit roundoff.

Outline

1. Background
   - Deterministic Error Bounds
   - Probabilistic Error Analysis
2. Intermediate Sums
3. New Bounds
4. Experiments/Conclusions

Deterministic Bound

Computing the partial sums:

$$
\begin{aligned}
\hat{s}_1 &= x_1, \\
\hat{s}_2 &= (\hat{s}_1 + x_2)(1+\delta_2) \\
&= x_1(1+\delta_2) + x_2(1+\delta_2), \\
\hat{s}_3 &= (\hat{s}_2 + x_3)(1+\delta_3) \\
&= x_1(1+\delta_2)(1+\delta_3) + x_2(1+\delta_2)(1+\delta_3) + x_3(1+\delta_3), \\
&\ \vdots \\
\hat{s}_n &= \sum_{i=1}^{n} x_i \left(\prod_{j=\max\{2,i\}}^{n} (1+\delta_j) \right)
\end{aligned}
$$

Deterministic Bound

Lemma (Deterministic Error Bound)

If $|\delta_i| \leq u$ *for* $i = 1 : n$*, and* $nu < 1$*, then*

$$\prod_{i=1}^{n}(1 + \delta_i) = 1 + \theta_n,$$

where

$$|\theta_n| \leq \frac{nu}{1 - nu} =: \gamma_n.$$

Consequently,

$$|\hat{s}_n - s_n| \leq \gamma_n \sum_{i=1}^{n} |x_i|.$$

- $\gamma_n \approx nu$ as long as $nu \ll 1$.

Deterministic Bound

Higham (2002):

> *Whenever we write γ_n there is an implicit assumption that nu < 1, which is true in virtually any circumstance that might arise with IEEE single or double precision arithmetic.*

Low-Precision Arithmetic

Normalized non-zero floating point numbers:

$$x = (-1)^s (1.f)_2 \cdot 2^e$$

Precision	Sign	Exp	Float	u
Double (f64)	1	11	52	$1.11 \cdot 10^{-16}$
Single (f32)	1	8	23	$5.96 \cdot 10^{-8}$
Half (f16)	1	5	10	$4.88 \cdot 10^{-4}$
Quarter (??)	1	3	4	$3.13 \cdot 10^{-2}$

- Problem sizes getting larger
- Half precision increasingly common
- When $nu > 1$, bounds using γ_n become useless

Outline

Probabilistic Model

Solution: abandon worst-case bounds, try probabilistic analysis instead

Model (Probabilistic)

For any floating point numbers a and b,

$$\mathrm{fl}(a \odot b) = (a \odot b)(1+\delta), \quad |\delta| \le u, \quad \odot \in \{+, -, \times, /, \surd\},$$

where u is the unit roundoff. The quantities δ for each computation are independent random variables with mean zero.

This model is wrong. But is it *useful*?

Probabilistic Model

Central Limit Theorem: if $e = \sum_{i=1}^{n} \delta_i t_i$ and $|\delta_i| \leq u$, then

$$|e| \leq \lambda u \left(\sum_{i=1}^{n} t_i^2 \right)^{1/2}$$

with high probability for large n.

- λ modest in size, controls probability
- Deterministic bound $u \sum_{i=1}^{n} |t_i|$ can be factor of $\sqrt{n}$ larger

Probabilistic Model

Wilkinson (1961):

> *In general, the statistical distribution of the rounding errors will reduce considerably the function of n occurring in the relative errors. We might expect in each case that this function should be replaced by something which is* **no bigger than its square root** *and is usually appreciably smaller.*

Ultimate goal: replace nu in error bounds with something that grows like $\sqrt{n}u$

Probabilistic Bound

Theorem (Hoeffding's Inequality)

Let $X_1, \ldots, X_n$ be independent random variables satisfying

$$|X_i| \leq c_i, \quad i = 1 : n.$$

The sum $S = \sum_{i=1}^n X_i$ satisfies

$$Pr(|S - \mathbb{E}[S]| \geq \xi) \leq 2 \exp\left(-\frac{\xi^2}{2\sum_{i=1}^n c_i^2}\right).$$

- Assumptions: random variables are independent and **bounded**.
- Holds for all n, not just as $n \to \infty$

Probabilistic Bound

Lemma (Higham/Mary 2018)

Assume the probabilistic model for roundoff errors. If $|\delta_i| \le u$ for $i = 1 : n$, and $nu < 1$, then

$$\prod_{i=1}^{n}(1+\delta_i) = 1 + \tilde{\theta}_n,$$

where

$$|\tilde{\theta}_n| \le \tilde{\gamma}_n(\lambda) := \exp\left(\lambda\sqrt{n}u + \frac{nu^2}{1-u}\right) - 1 \approx \lambda\sqrt{n}u$$

with failure probability at most

$$Q(\lambda) = 2\exp\left(-\frac{\lambda^2(1-u)^2}{2}\right).$$

Probabilistic Bound

Deterministic bound:

$$|\hat{s}_n - s_n| \leq \gamma_n \sum_{i=1}^{n} |x_i| \approx nu \sum_{i=1}^{n} |x_i|.$$

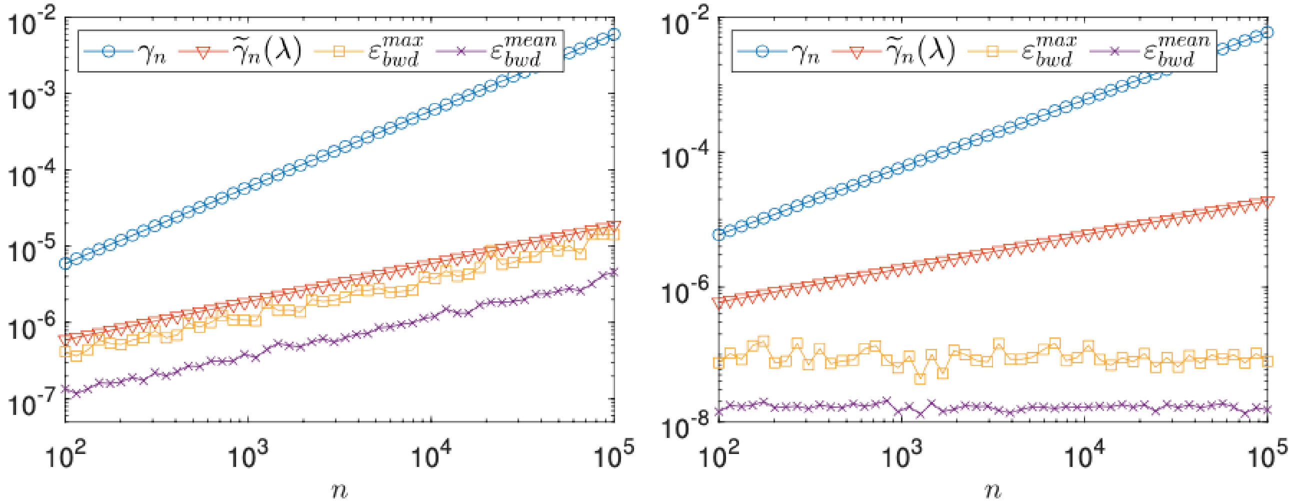
Probabilistic bound: WFP at most $Q(\lambda)$,

$$|\hat{s}_n - s_n| \leq \tilde{\gamma}_n(\lambda) \sum_{i=1}^{n} |x_i| \approx \lambda\sqrt{n}u \sum_{i=1}^{n} |x_i|.$$

- Probabilistic approximation holds while $\lambda\sqrt{n}u \ll 1$

Probabilistic Bound

How does the Higham/Mary bound perform in practice?



(a) Random uniform $[0, 1]$ vectors. (b) Random uniform $[-1, 1]$ vectors.

Fig. 4.1: Backward error and its bounds for the computation in single precision of the inner product $y = a^T b$, for vectors a and b with random uniform entries. Here, $N_{test} = 100$ and $\lambda = 1$.

- Much stronger than deterministic bound
- Slightly pessimistic on random $[0, 1]$ data
- Way off the mark on random $[-1, 1]$ data

Running Error Bound

Idea (Higham 2002): rewrite the computed sums as

$$\hat{s}_i = \frac{\hat{s}_{i-1} + x_i}{1 + \delta_i}, \quad |\delta_i| \le u.$$

Result: the deterministic *running error bound*

$$|\hat{s}_n - s_n| = \left|\sum_{i=1}^{n} \delta_i \hat{s}_i\right| \le u \sum_{i=1}^{n} |\hat{s}_i|$$

This can be much smaller than the *a priori* bounds!

Running Error Bound

Higham 2002:

> *In designing or choosing a summation method to achieve high accuracy, the aim should be to minimize the absolute values of the intermediate sums [$\hat{s}_i$].*

- Our goal: find the probabilistic version of the running error bound
- Problem: Quantities $\hat{s}_i$ depend on δ terms
  - Difficult to apply concentration bounds directly to $\sum_{i=1}^{n} \delta_i \hat{s}_i$

Intermediate Sums

Our strategy: rewrite error as

$$\hat{s}_n - s_n = \sum_{i=2}^{n} \left(s_i \delta_i \prod_{j=i+1}^{n} (1 + \delta_j) \right).$$

Deterministic bound:

$$|\hat{s}_n - s_n| \le u(1 + \gamma_n) \sum_{i=2}^{n} |s_i|.$$

Two small wrinkles in getting the probabilistic version:

- δ_i independent $\not\Rightarrow \delta_i \prod_{j=i+1}^{n}(1+\delta_j)$ independent
- The term $(1 + \gamma_n)$ is a problem when $nu > 1$

Martingales

Use a martingale!

Definition (Martingale)

A squence of random variables $X_1, X_2, \ldots$ is a martingale with respect to $\delta_1, \delta_2, \ldots$ if for $i \geq 1$

1. X_i is a function of $\delta_1, \ldots, \delta_{i-1}$,
2. $\mathbb{E}[|X_i|] < \infty$
3. $\mathbb{E}[X_{i+1}|\delta_1, \ldots, \delta_{i-1}] = X_i$.

- Examples: unbiased random walk, gambler playing a fair game
- The increments $(X_{i+1} - X_i)$ do **not** need to be independent!

Martingales

We can relax the requirements of Hoeffding's Inequality.

Theorem (Azuma's Inequality)

Suppose a martingale $\{X_1, \ldots, X_n\}$ *satisfies*

$$|X_i - X_{i-1}| \leq c_i, \quad i = 2:n.$$

Then

$$Pr(|X_n - X_1| \geq \xi) \leq 2\exp\left(-\frac{\xi^2}{2\sum_{i=2}^{n} c_i^2}\right).$$

Martingales

To construct the martingale, work backwards:

$$\begin{aligned} X_1 &= s_n, \\ X_2 &= X_1 + s_n\delta_n, \\ X_3 &= X_2 + s_{n-1}\delta_{n-1}(1+\delta_n), \\ &\vdots \\ X_n &= X_{n-1} + s_2\delta_2\prod_{j=3}^{n}(1+\delta_j) \end{aligned}$$

Apply Azuma's inequality with

$$\begin{aligned} c_i &= |s_{n-i+2}|u(1+\gamma_n), \\ \xi &= \lambda u(1+\gamma_n)\left(\sum_{i=2}^{n} s_i^2\right)^{1/2} \end{aligned}$$

Martingales

Deterministic bound:

$$|\hat{s}_n - s_n| \leq u(1+\gamma_n) \sum_{i=2}^{n} |s_i|$$

Probabilistic bound:

$$|\hat{s}_n - s_n| \leq \lambda u(1+\gamma_n) \left(\sum_{i=2}^{n} s_i^2 \right)^{1/2}$$

with failure probability at most $2\exp\left(-\frac{\lambda^2}{2}\right)$

- Not good enough: $(1+\gamma_n)$ blows up when $nu > 1$!

Probabilistic Azuma

Second idea: why use the bound

$$\left| \delta_i \prod_{j=i+1}^{n} (1+\delta_j) \right| \le u(1+\gamma_n),$$

when the left-hand side is **probably** much smaller?

Probabilistic Azuma

Relax Azuma's inequality by allowing the bounds to fail with small probability.

Theorem (Azuma's Inequality, Probabilistic Version)

Suppose a martingale $\{X_1, \ldots, X_n\}$ *satisfies*

$$|X_i - X_{i-1}| \leq c_i, \quad i = 2 : n$$

with total failure probability at most η. *Then*

$$Pr(|X_n - X_1| \geq \xi) \leq 2 \exp\left(-\frac{\xi^2}{2\sum_{i=2}^{n} c_i^2}\right) + \eta.$$

Using $\eta = Q(\lambda)$, we can replace γ_n with $\tilde{\gamma}_n(\lambda)$ in our bounds at minimal cost!

New Bound

Deterministic bound:

$$|\hat{s}_n - s_n| \le u(1+\gamma_n)\sum_{i=2}^{n}|s_i|$$

Probabilistic bound:

$$|\hat{s}_n - s_n| \le \lambda u(1+\tilde{\gamma}_n(\lambda))\left(\sum_{i=2}^{n} s_i^2\right)^{1/2}$$

with failure probability at most $2Q(\lambda)$.

- Now works well when $nu > 1$ and $\lambda\sqrt{n}u \ll 1$

Numerical Experiments

Half precison, uniform [-1,1] data

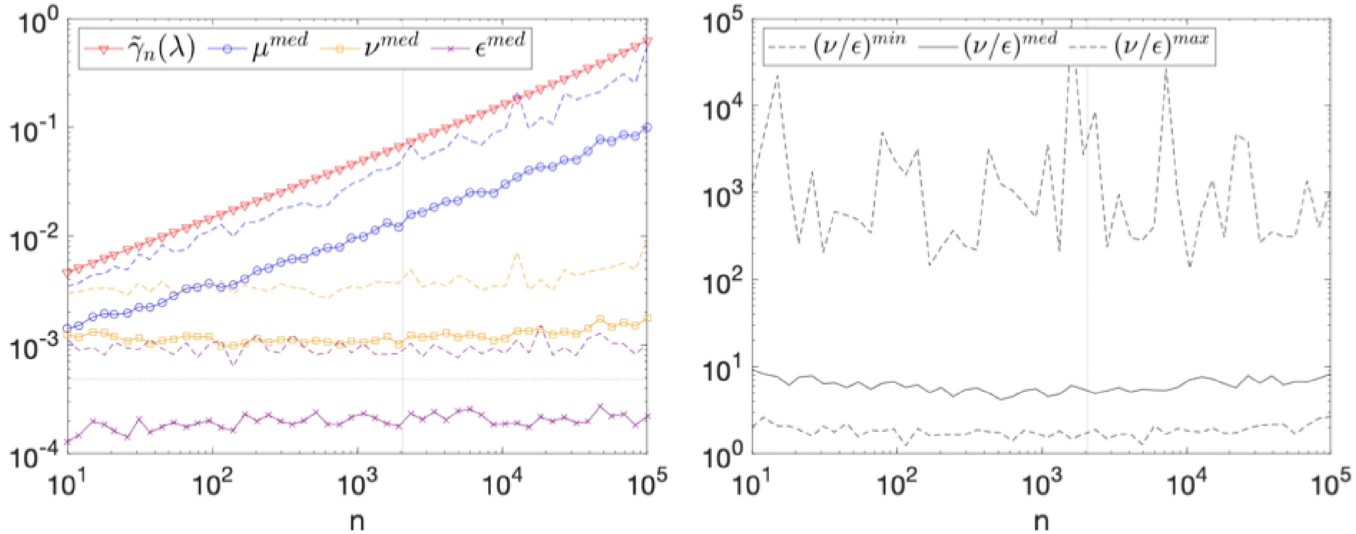


FIG. 6.2. *Backward error bounds for the half precision computation of* $\mathbf{a}^T\mathbf{b}$, *where* $\mathbf{a}$ *and* $\mathbf{b}$ *are random uniform* $[-1,1]$ *vectors,* $N_{test} = 100$, *and* $\lambda = 3$. *Left: The horizontal line represents the unit roundoff* $u = 2^{-11}$, *and the vertical line represents the value of* n *for which* $nu > 1$.

- $\tilde{\gamma}_n$: the Higham/Mary probabilistic bound
- μ: the deterministic running bound
- ν: our new probabilistic bound
- ϵ: the true error

Numerical Experiments

Quarter precison, uniform [-1,1] data

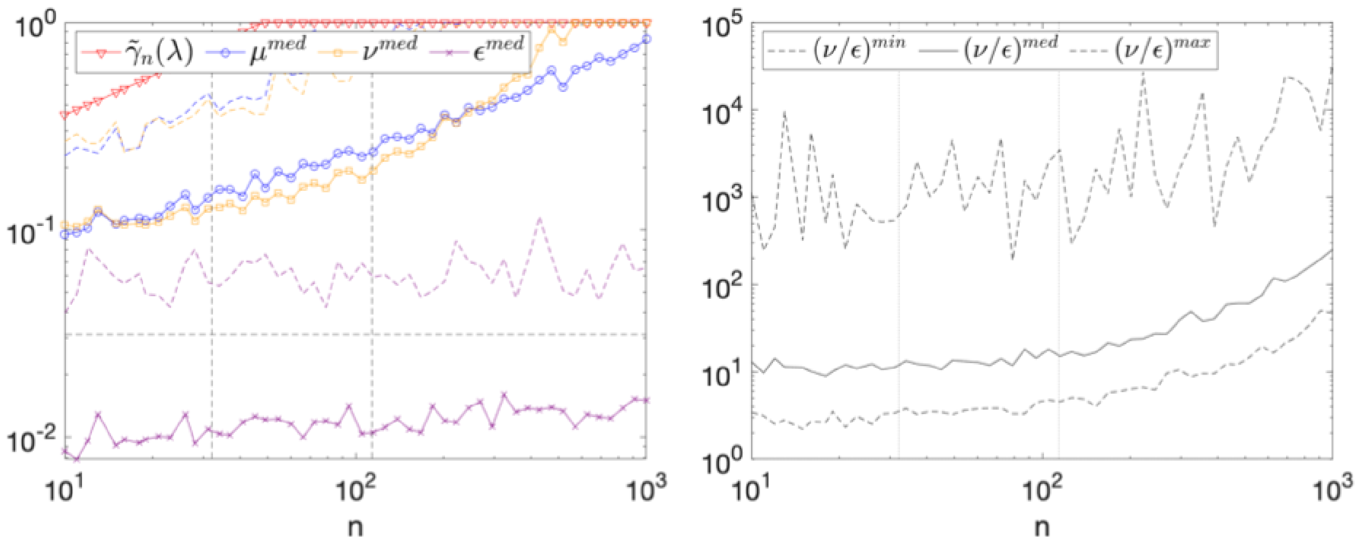


FIG. 6.3. *Backward error bounds for the quarter precision computation of* $\mathbf{a}^T\mathbf{b}$*, where* $\mathbf{a}$ *and* $\mathbf{b}$ *are random uniform* $[-1,1]$ *vectors,* $N_{test} = 100$*, and* $\lambda = 3$*. Error bounds are plotted as 1 when their value exceeds 1. Left: the horizontal line represents the unit roundoff* $u = 2^{-5}$*. Vertical lines represent the values of* n *for which* $nu > 1$ *and* $\lambda\sqrt{nu} > 1$*.*

- Single/Half precision: we typically overestimate the error by a factor of 10
- Quarter precison: similar performance to the deterministic running bound

Numerical Experiments

Half precison, uniform [-1,1] data

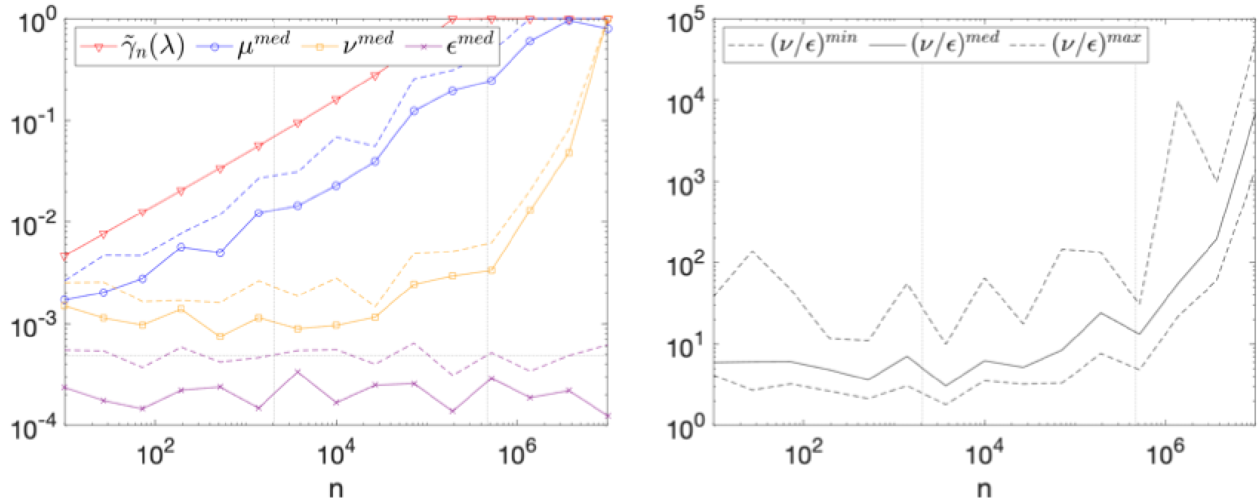


FIG. 6.4. *Results in half precision arithmetic with $N_{test} = 10$ and $\lambda = 3$. Error bounds are plotted as 1 when their value exceeds 1. Left: the horizontal line represents the unit roundoff $u = 2^{-11}$. Vertical lines represent the values of n for which $nu > 1$ and $\lambda\sqrt{nu} > 1$.*

- Our estimates break down when $\lambda\sqrt{nu} > 1$
- Still fails to capture behavior in practice for large n

Conclusions

Summary of bounds:

	Deterministic	Probabilistic
Data	$\gamma_n \sum_{i=1}^n \lvert x_i \rvert$	$\tilde{\gamma}_n \sum_{i=1}^n \lvert x_i \rvert$
Intermediate Sums	$u(1+\gamma_n) \sum_{i=2}^n \lvert s_i \rvert$	$\lambda u(1+\tilde{\gamma}_n(\lambda)) \left(\sum_{i=2}^n s_i^2\right)^{1/2}$
Running Bound	$u \sum_{i=1}^n \lvert \hat{s}_i \rvert$	???

- (How) can we drop the $(1+\tilde{\gamma}_n(\lambda))$ term?
- (How) can we develop an effective running bound?

Conclusions

Thanks for listening!

For Further Reading I

- Nicholas J. Higham
  Accuracy and Stability of Numerical Algorithms, 2ed. 2002.
- Cleve Moler
  "Half Precision" 16-bit Floating Point Arithmetic
  https://blogs.mathworks.com/cleve/2017/05/08/half-precision-16-bit-floating-point-arithmetic/
- Nicholas J. Higham and Theo Mary
  A New Approach to Probabilistic Rounding Error Analysis
  http://eprints.maths.manchester.ac.uk/2673/1/paper.pdf
- Fan Chung and Linyuan Lu
  Concentration Inequalities and Martingale Inequalities—a Survey
  http://people.math.sc.edu/lu/papers/concen.pdf