# Trace estimation via asynchronous stochastic rounding

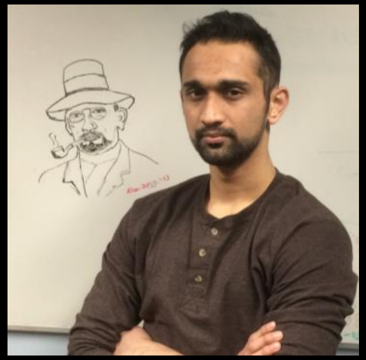
10th International Congress on Industrial and Applied Mathematics
ICIAM 2023 TOKYO

## ICIAM 2023

Minisymposium on Stochastic Rounding for Reduced-Precision Arithmetic in Scientific Computing



Vasileios Kalantzis


Shashanka Ubaru


Georgios Kollias


Lior Horesh


Chai Wah Wu

# Introduction

- We consider the problem of computing the trace of an implicitly-defined $n$ by $n$ matrix $A = f(G)$.
- This problem is important in many applications, where the function $f$ can be matrix exponential, matrix log, fractional powers or entropy.
- Computational difficulties arise when the order $n$ is large and/or all the information of $G$ is not available in the computer's main memory.
- We consider approaches to estimate the trace of a large matrix $A$ using a restricted amount of information from $A$.

# Inaccurate or approximate computing

- Due to resource (time, power, size, etc.) limitation, it might not be possible to compute the full trace of $A$.
- In this case, we are limited to a restricted access to information about $A$ in order to estimate $trace(A)$.
- We consider two such restrictions which can be considered in the same framework.
  1. Asynchronous randomized trace estimates
  2. Trace estimation via stochastic rounding

# Matrix trace estimation and graph analytics

- Matrix trace estimation is ubiquitous in graph analytics.
- In chemical graph theory, the Estrada index is a topological index of protein folding. The index was first defined by Ernesto Estrada as a measure of the degree of folding of a protein, which is represented as a path-graph weighted by the dihedral or torsional angles of the protein backbone. This index of degree of folding has found multiple applications in the study of protein functions and protein-ligand interactions. The Estrada index is equal to

$$\texttt{trace}(e^A).$$

- Computing the transitivity ratio of a (sub-)graph leads to e-commerce opportunities, e.g., high ratio implies similarity between nodes, thus creating marketing opportunities in of e-commerce platforms (for example, suggest to user $i$ what you suggested to users $j$ and $k$ if they form a triangle). The number of triangles can be determined by computing

$$\texttt{trace}(A^3)/6.$$

# Generalized Adversarial Networks

- The Fréchet inception distance (FID) is a metric used to assess the quality of images created by a generative model, like a generative adversarial network (GAN).
- Unlike the earlier inception score (IS), which evaluates only the distribution of generated images, the FID compares the distribution of generated images with the distribution of a set of real images ("ground truth").
- For multivariate variables this is equivalent to computing

$$\mathtt{FID} = \|\mu_X - \mu_Y\| + \mathtt{trace}(\Sigma_X + \Sigma_Y - 2\sqrt{\Sigma_X \Sigma_Y}).$$

- We need to compute the trace of covariance matrices.

# Hutchinson's trace estimator

- The standard approach to compute the trace of an implicitly-defined matrix $A$ is to apply Monte Carlo trace estimation.
- Let $x$ be a random vector with zero mean and unit variance. We have

$$\mathbb{E}[x^T A x] = \sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}\mathbb{E}[x_i x_j] = \sum_{i=1}^{n}\left[A_{ii}\mathbb{E}[x_i x_i] + \sum_{j\neq i}^{n} A_{ij}\mathbb{E}[x_i x_j]\right] = \sum_{i=1}^{n} A_{ii} = \texttt{trace}(A).$$

- Thus, the following trace estimator, known as Hutchinson's trace estimator, is an unbiased estimator of $\texttt{trace}(A)$:

$$\text{Hutchinson's trace estimator}: \ \frac{1}{m}\sum_{k=1}^{m} x_k^T A x_k,$$

where $x_k$ is an $n$-length Rademacher vector (i.e., each entry is equal to $\pm 1$ with equal probability).
- The convergence of Hutchinson's trace estimator is governed by $O(1/\sqrt{m})$.

# Asynchronous randomized trace estimates

- Asynchronous computations arise naturally in distributed-memory implementations for the iterative computation of fixed points so as to reduce idle time between different processing elements via reducing synchronization points.

- While asynchronous iterations generally lead to slower convergence, the ever-increasing gap between the time required to share a floating-point number between different processing elements and the time needed to perform a single floating-point operation by one of the processing elements, has led to a revived interest in the analysis and application of asynchronous algorithms in numerical linear algebra.

# Probabilistic framework

- Let $\mathcal{T}$ denote a random subset of $T \in \mathbb{N}$ integers (without replacement) from the set $\{1, 2, \ldots, N\}$. We define the asynchronous MV $y = A\,|_{\mathcal{T}}\, x$ between the matrix $A \in \mathbb{R}^{N \times N}$ and a vector $x \in \mathbb{R}^N$ as a function of $\mathcal{T}$ such that:

$$e_i^T y = \begin{cases} [Ax]_i & \text{if } i \in \mathcal{T} \\ 0 & \text{if } i \notin \mathcal{T}. \end{cases}$$

- In other words, the operator $|_{\mathcal{T}}$ is equivalent to the regular MV $Ax$ with the exception that the $i$th row of $A$ is now replaced by an $N$-length zero row vector for any $i \notin \mathcal{T}$.
- Given $T$, the random subset of $\mathcal{T}$ picks any $T \equiv |\mathcal{T}|$ integers of $\{1, 2, \ldots, N\}$ with equal probability, i.e., each one of the $\binom{N}{T}$ possible row sets of $A$ is picked with probability $\binom{N}{T}^{-1}$.

# Probabilistic framework

- Let $k = 1, 2, \ldots, m,\ m \in \mathbb{N}$, and denote by $\mathcal{T}_k$ a random subset of $|\mathcal{T}_k| \in \mathbb{N}$ integers (without replacement) from $1$ to $N$. The deterministic integer $|\mathcal{T}_k|$ is an instance of the integer-valued random variable $T \in \{1, 2 \ldots, N\}$. Then, for any $N$-length instances $x_1, x_2, \ldots, x_m$, of a random vector $x$, we define the asynchronous randomized trace estimator

$$\Gamma_m = \frac{1}{m} \sum_{k=1}^{m} x_k^T (A \mid_{\mathcal{T}_k} x_k) = \frac{1}{m} \sum_{k=1}^{m} \sum_{i \in \mathcal{T}_k} [x_k]_i^T [A \mid_{\mathcal{T}_k} x_k]_i .$$

- The second equality of $\Gamma_m$ follows by recalling that the $i$th entry of the product $A|_{\mathcal{T}_k} x_k$ is nonzero if and only if $i \in \mathcal{T}_k$.

The vectors $x_1, \ldots, x_m$ are instances of a random vector sampled from a zero mean distribution and i.i.d. components with variance 1, i.e. $E[xx^T] = I$.

# Probabilistic framework

- Consider now the diagonal random matrix formed by the summation of $T$ canonical outer products

$$D_{\mathcal{T}} = \sum_{i \in \mathcal{T}} e_i e_i^T,$$

where both the cardinality $T$ and the row subset $\mathcal{T}$ are random variables.

- When $T \equiv N$, as in the synchronous case, the matrix $D_{\mathcal{T}}$ is equal to the $N \times N$ identity matrix. The asynchronous randomized trace estimator can be then written equivalently as

$$\Gamma_m = \frac{1}{m} \sum_{k=1}^{m} x_k^T D_{\mathcal{T}_k} A x_k = \frac{1}{m} \sum_{k=1}^{m} x_k^T Q(\mathcal{T}_k) x_k,$$

- Here, $Q(\mathcal{T}_k) = D_{\mathcal{T}_k} A$ and $D_{\mathcal{T}_k} A x = A|_{\mathcal{T}_k} x$.

# Probabilistic framework

- Let $Q$ denote a random matrix and $x$ denote an independent random vector of the same length as $Q$ such that $\mathbb{E}[x] = 0$ and $\mathbb{E}[xx^T] = I$. Then,

$$\mathbb{E}[x^T Q x] = \mathtt{Tr}(\mathbb{E}[Q]).$$

- If the sample space of the random matrix $Q$ is formed by all possible matrices $Q(\mathcal{T}) = D_{\mathcal{T}} A$ such that, for a given sample integer value of a uniform $T$ in the interval $[1, N]$, the random subset of $\mathcal{T}$ picks any $T \equiv |\mathcal{T}|$ integers of $\{1, 2, \ldots, N\}$ with equal probability, then $\Gamma_m$ is an unbiased estimator of $\mathtt{Tr}(\mathbb{E}[Q])$.
- The main question now becomes whether we can exploit $\Gamma_m$ to approximate the trace of the implicit deterministic matrix $A$.
- Let $\mu_T = \mathbb{E}[T]$ denote the expectation of the random variable $T$. Then,

$$\mathbb{E}[Q] = \frac{\mu_T}{N} A, \qquad \text{and} \qquad \mathbb{E}\left[\Gamma_m\right] = \frac{\mu_T}{N} \mathtt{Tr}(A),$$

i.e., the randomized estimator $\dfrac{N}{\mu_T}\Gamma_m$ is an unbiased estimator of $\mathtt{Tr}(A)$.

# Algorithm

**0.** Receive $m$, set $\Gamma = \widehat{\Gamma} = 0$

**1.** Do $k = 1, \ldots, m$

- Sample $x$ from the Rademacher distribution
- Update $\widehat{\Gamma} = \widehat{\Gamma} + x_k^T D_{\mathcal{T}_k} A x_k$
- Set $\Gamma = \widehat{\Gamma}/k$

**2.** End

**3.** Return $\Gamma_m = \Gamma$

- For each $k$, the random subset $\mathcal{T}_k$ picks any $|\mathcal{T}_k|$ integers of $\{1, 2, \ldots, N\}$ with equal probability.
- Each cardinality $|\mathcal{T}_k|$ is an instance of an integer $T$ and takes values between $1$ and $N$.

# Variance of asynchronous trace estimator

**Theorem 1.** *Let $\sigma_T^2$ denote the variance of the random variable $T$, and define the scalars*

$$K_1 = \frac{3(N\mu_T - \sigma_T^2 - \mu_T^2)}{N(N-1)},\ K_2 = \frac{2\left(\sigma_T^2 + \mu_T^2 - \mu_T\right)}{N(N-1)},\ K_3 = \frac{\left(\sigma_T^2 + \frac{1}{N}\mu_T^2 - \mu_T\right)}{N(N-1)} \text{ and } K_4 = K_1 - 2\frac{\mu_T}{N}.$$

*The variance of a single sample of the asynchronous randomized trace estimator* $\mathrm{Var}(x^T Qx)$ *is then equal to*

$$K_1 \mathtt{Tr}\left(\mathrm{diag}(A)^2\right) + K_2 \mathtt{Tr}\left(A^2\right) + K_3 \mathtt{Tr}\left(A\right)^2,$$

*when $x \in \mathcal{N}(0, I)$, and equal to*

$$K_4 \mathtt{Tr}\left(\mathrm{diag}(A)^2\right) + K_2 \mathtt{Tr}\left(A^2\right) + K_3 \mathtt{Tr}\left(A\right)^2,$$

*when $x$ is a Rademacher random vector.*

# Variations of asynchronous trace estimator

Notice that when $T \equiv N$, the randomized trace estimation becomes synchronous, and we have $\sigma_T^2 = 0$ and $\mu_T = N$. Plugging these values in Theorem 1 gives us $K_1 = K_3 = 0$, $K_2 = 2$, $K_4 = -2$, and $\mathrm{Var}(x^T Qx) = 2\|A\|_F^2$ when $x \in \mathcal{N}(0, I)$, and $\mathrm{Var}(x^T Qx) = 2(\|A\|_F^2 - \sum_{i=1}^N A_{ii}^2)$ when $x$ is a Rademacher vector. These variances are identical to those of the randomized trace estimator in the synchronous case [2]. In the general asynchronous case, $T$ can be less than $N$, and one can distinguish three important cases for $T$:

1. $T$ is a fixed integer (deterministic) in the range $1 \leq T \leq N$,
2. $T$ takes on integer values in $[1, N]$ with equal probability,
3. $\mathcal{T}$ is obtained by choosing each element in $[1, N]$ with probability $p$. Note that for a fixed $T$, each subset $\mathcal{T}$ such that $T \equiv |\mathcal{T}|$ occurs with the same probability.

# Stochastic rounding

We can consider the asynchronous setting as a case of approximate and inaccurate computing where only a random approximation $Q$ of the matrix $A$ is used each time and requiring that the expectation of the random variable is *proportional* to $A$ (as expressed by $E[Q] = (\mu_T/N)A$).

Another important method of random approximation is *stochastic rounding*, where a real number is approximated by neighboring quantization levels with probability proportional to the distance to the opposite quantization level.

More precisely, if $q_1 \leq x \leq q_2$ lies between quantization levels $q_1$ and $q_2$, the stochastic rounding of $x$ is defined as $\mathrm{sr}(x) = q_1$ with probability $\frac{q_2 - x}{q_2 - q_1}$ and $\mathrm{sr}(x) = q_2$ otherwise[1]. It is easy to see that $\mathbb{E}[\mathrm{sr}(x)] = x$ and $\mathrm{Var}(\mathrm{sr}(x)) = x(q_1 + q_2 - x) - q_1 q_2$. Let us denote $r(x) = x - q_1$ and $\Delta(x) = q_2 - q_1$ in which case we can write $\mathrm{Var}(\mathrm{sr}(x)) = r(x)(\Delta(x) - r(x))$, and $\mathbb{E}[\mathrm{sr}(x)^2] = q_2 r(x) + x q_1$.

# Stochastic rounded asynchronous randomized trace estimator

**Definition 3.** *Let $\mathcal{T}$ denote a random subset of $T \in \mathbb{N}$ integers (without replacement) from the set $\{1, 2, \ldots, N\}$. We define the stochastically rounded asynchronous matrix-vector product (SRAMVP) $y = A\,|_{\mathcal{T}}\, x$ between $A \in \mathbb{R}^{N\times N}$ and a vector $x \in \mathbb{R}^N$ as a function of $\mathcal{T}$ such that:*

$$e_i^T y = \begin{cases} \mathrm{sr}([Ax]_i) & \text{if } i \in \mathcal{T} \\ 0 & \text{if } i \notin \mathcal{T}. \end{cases}$$

*In other words, the operator $|_{\mathcal{T}}$ is equivalent to the regular matrix-vector multiplication $Ax$ with the difference that the entries are replaced with a stochastic rounding representation and the $i$th row of $A$ is replaced by an $N$-length zero row vector unless $i \in \mathcal{T}$. We assume that the stochastic rounding is independent from the random subset $\mathcal{T}$.*

# Stochastic rounded asynchronous randomized trace estimator

Let $\tilde{A}$ be the random matrix where each entry $\tilde{A}_{ij} = \mathrm{sr}(A_{ij})$ independently.

**Theorem 3.** *The variance of the stochastically rounded asynchronous randomized trace estimator* $\mathrm{Var}(x^T Qx)$ *is equal to*

$$K_1 \mathtt{Tr}\left(\mathbb{E}[\mathrm{diag}(\tilde{A})^2]\right) + K_2 \mathtt{Tr}\left(\mathbb{E}[\tilde{A}^2]\right) + K_3 \mathbb{E}\left[\mathtt{Tr}\left(\tilde{A}\right)^2\right],$$

*when* $x \in \mathcal{N}(0, I)$, *and equal to*

$$K_4 \mathtt{Tr}\left(\mathbb{E}[\mathrm{diag}(\tilde{A})^2]\right) + K_2 \mathtt{Tr}\left(\mathbb{E}[\tilde{A}^2]\right) + K_3 \mathbb{E}\left[\mathtt{Tr}\left(\tilde{A}\right)^2\right],$$

*when* $x$ *is a Rademacher random vector, where* $K_1,\ K_2,\ K_3$, *and* $K_4$, *are defined in Theorem 1.*

# Stochastic rounded asynchronous randomized trace estimator

As for the random vectors $x$, note that by symmetry the Rademacher vectors can be considered a stochastic rounding of Gaussian vectors with two quantization levels when the stochastic rounding is independent from the Gaussian random variable. More generally, we replace $x$ with $\mathrm{sr}(x)$ and obtain

$$\tilde{\Gamma}_m = \frac{1}{m}\sum_{k=1}^{m} \mathrm{sr}(x_k)^T Q(\mathcal{T}_k)\mathrm{sr}(x_k). \tag{2}$$

Assuming the quantization levels are symmetric around 0, then for $x$ symmetric around 0 (e.g., Gaussian) we have $\mathbb{E}\left[\mathrm{sr}(x)\mathrm{sr}(x)^T\right] \propto I$ and Eq. (2) after scaling is an unbiased estimator of $\mathtt{Tr}(A)$.

# Experimental results

- A sample of sparse matrices from the SuiteSparse Matrix Collection

| **Id** | **Matrix name** | $N$ | $\mathtt{nnz}(A)$ | $\mathtt{Tr}(A)$ |
|---|---|---|---|---|
| 1 | Pajek/yeast | 2361 | 13828 | 536 |
| 2 | SNAP/ca-HepTh | 9877 | 51971 | 25 |
| 3 | Botonakis/thermomech_TC | 102158 | 711558 | 585.871 |
| 4 | SNAP/web-Stanford | 281903 | 2312497 | 0 |
| 5 | LAW/cnr-2000 | 325557 | 3216152 | 87442 |

# Experimental results

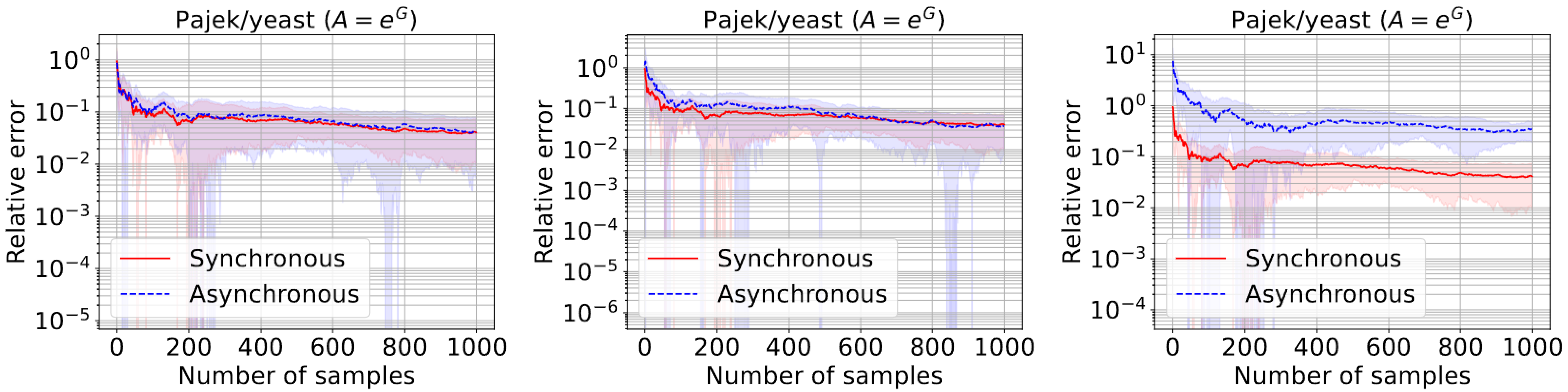



Figure 1: $A = e^G$. *Left to right: fixed* $T = \lceil Np \rceil$, *uniform* $T$, *fixed* $p$. *We use* $p = 0.6$.

# Experimental results

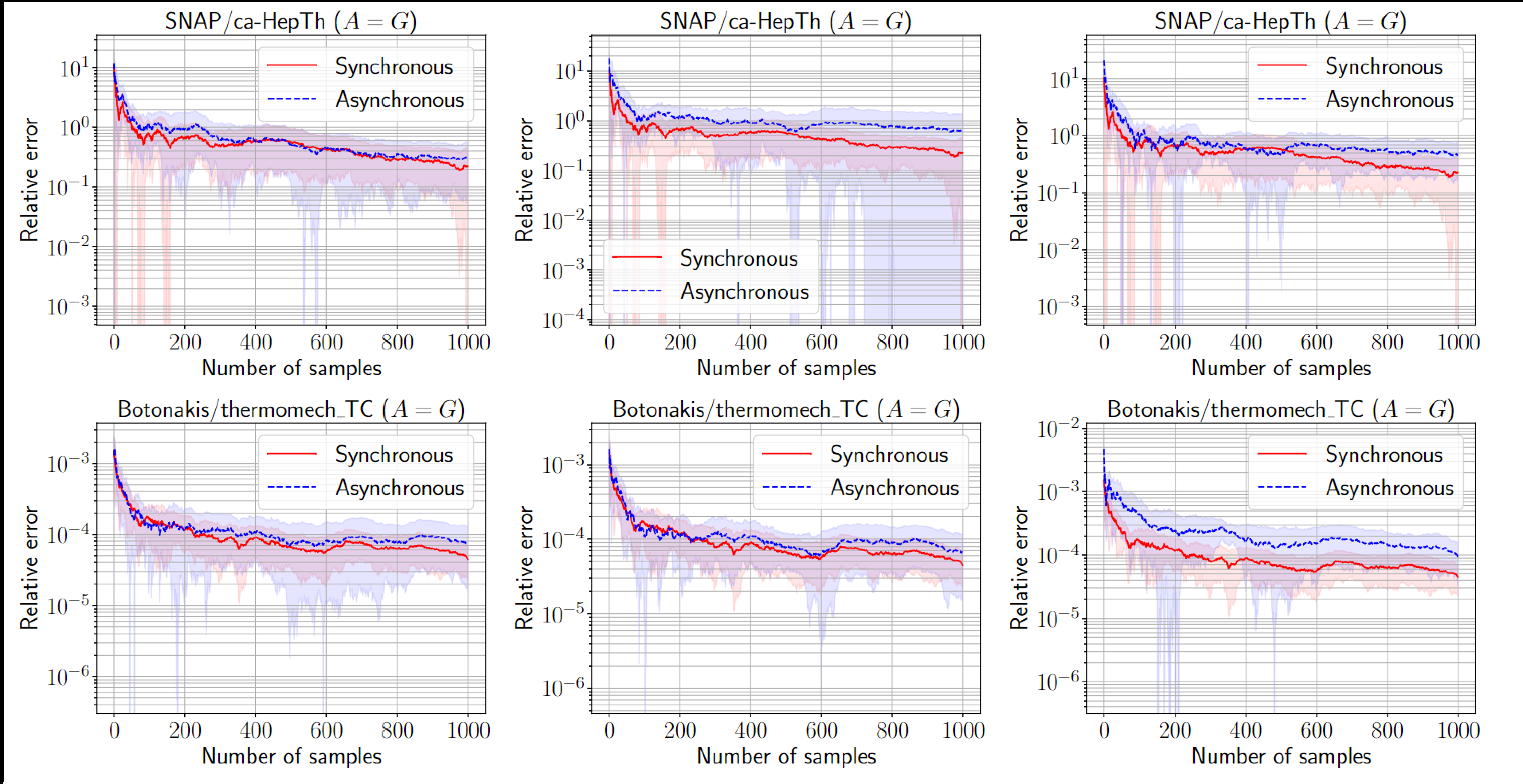



Figure 2: *Left to right: fixed $T = \lceil Np \rceil$, uniform $T$, fixed $p$. We use $p = 0.6$.*

# Experimental results

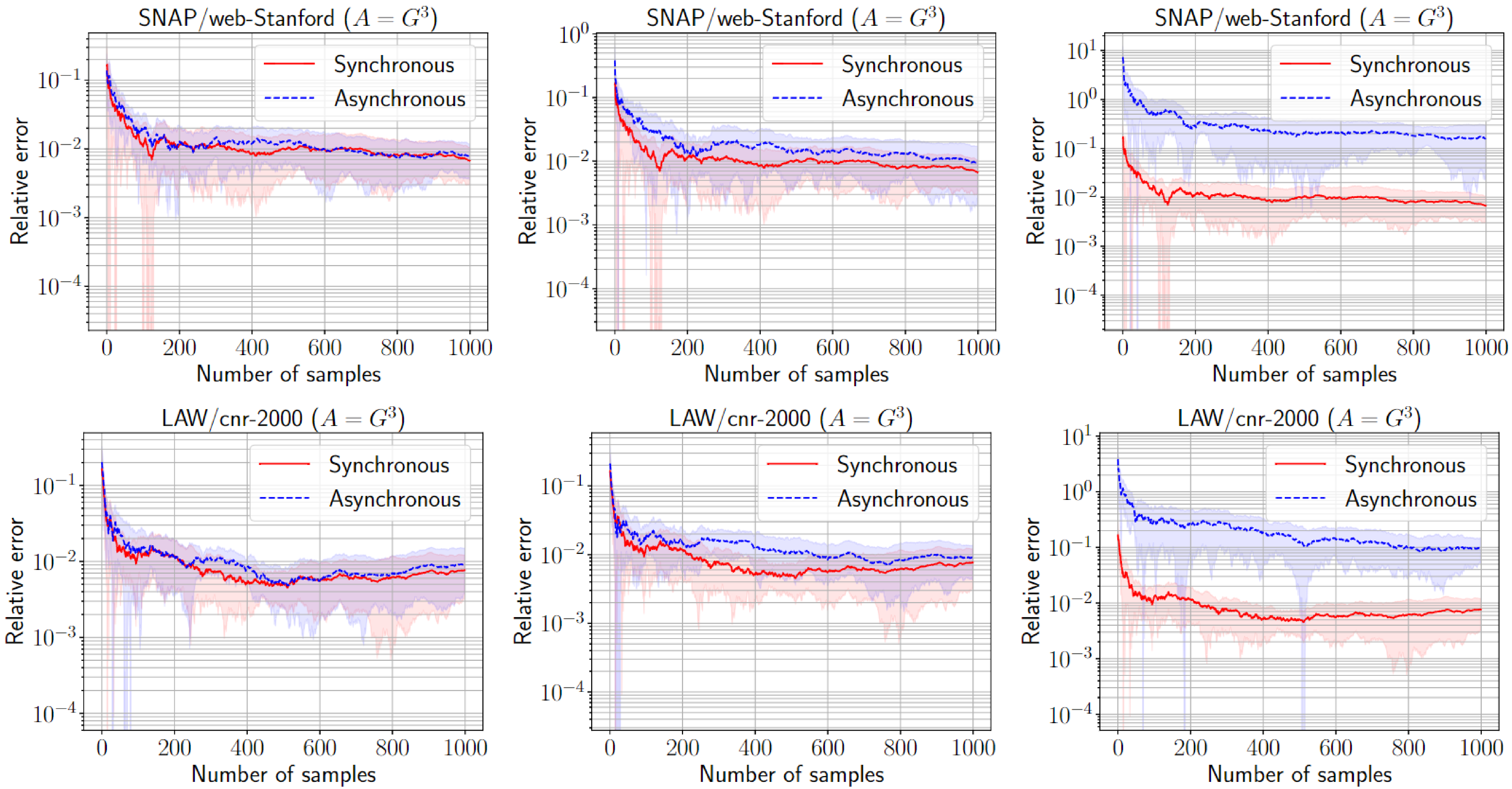



Figure 3: $A = G^3$. *Left to right: fixed* $T = \lceil Np \rceil$, *uniform* $T$, *fixed* $p$. *We use* $p = 0.6$.

# Experimental results

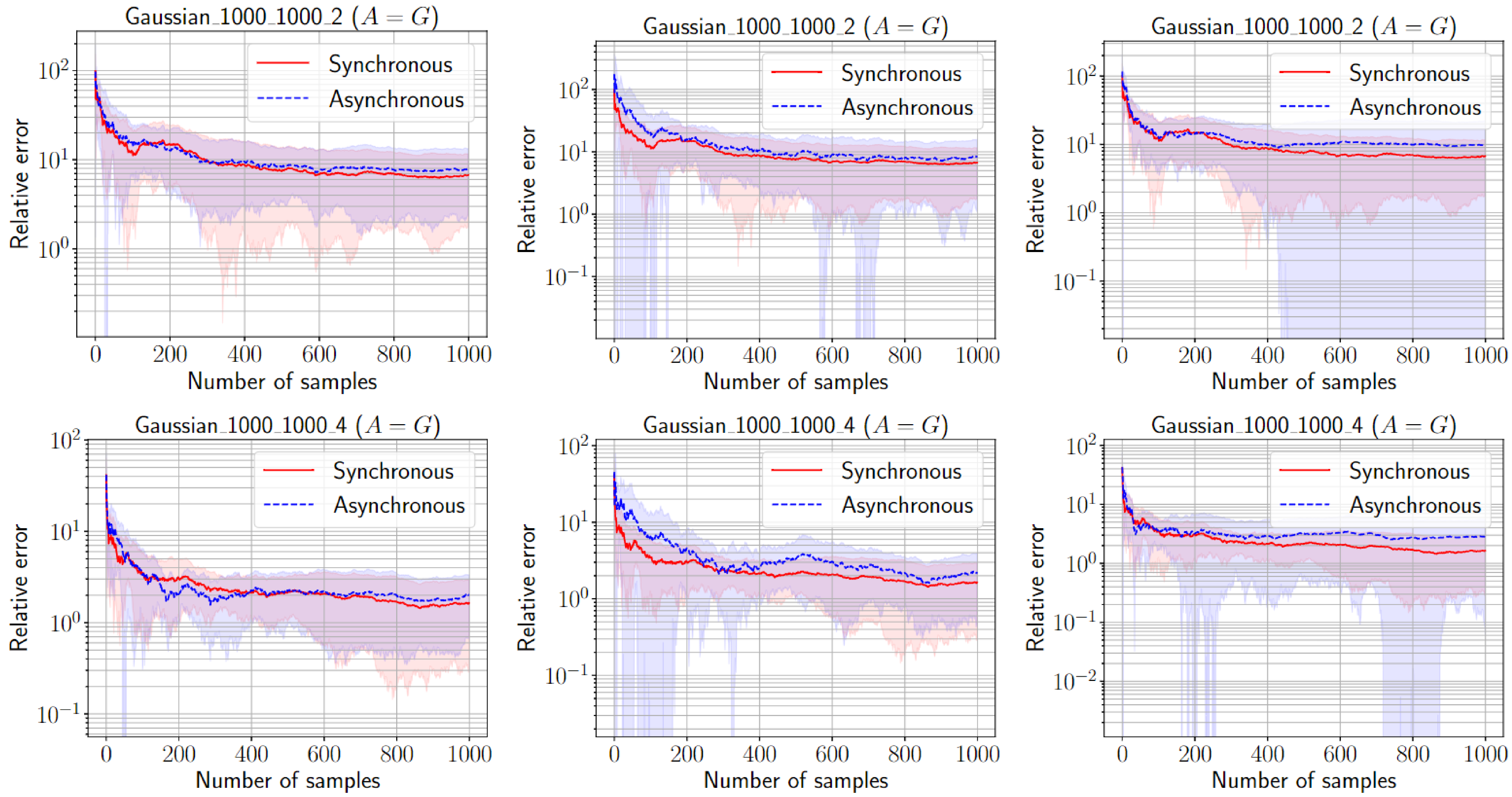



Figure 4: *[Rademacher samples]* Stochastic rounding. Matrix of size $N = 1000$ with entries sampled from standard normal distribution scaled by 1000. Left to right: fixed $T$, uniform $T$, fixed $p$; $p = 0.6$, $T = \lceil Np \rceil$. Top to bottom: Different numbers of quantization levels: $2, 4$

# Experimental results

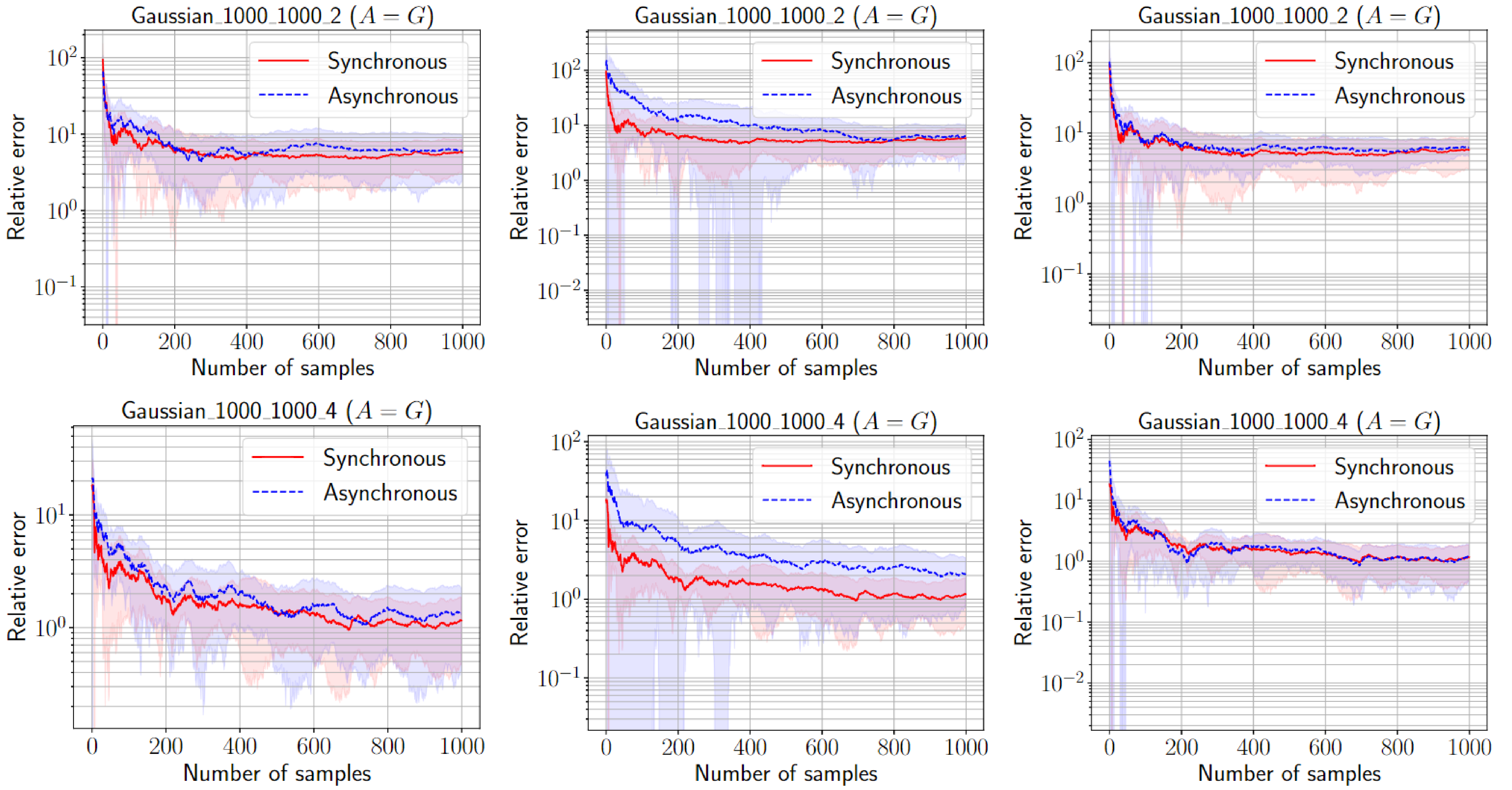



Figure 5: *[Gaussian samples]* Stochastic rounding. Matrix of size $N = 1000$ with entries sampled from standard normal distribution scaled by 1000. Left to right: fixed $T$, uniform $T$, fixed $p$; $p = 0.6$, $T = \lceil Np \rceil$. Top to bottom: Different numbers of quantization levels: 2, 4