CitizenGW-Compute: Turning Computers into Gravitational Wave Sensor (P-I)

Saikat Mohanta @CosmiKodes,
Independent Researcher
Open Science Project: CitizenGW-Compute
Email: saikatmohantabkp@gmail.com

Abstract—Gravitational waves (GWs) have been observed by LIGO and Virgo in the frequency range 10 Hz - 10 kHz, confirming one of the key predictions of prof. Einstien's famous General Relativity. While these interferometers belongs to state-of-the-art technological marvel, auxiliary methods that complement them in different frequency ranges or through alternative detection strategies could open new avenues in GW astronomy. This paper proposes a novel approach: leveraging the timing stability of modern computing systems (CPUs, GPUs, and memory caches) as indirect probes of GW-induced perturbations. Our method exploits deterministic artificial neural network (ANN) workloads and ultra-stable computational benchmarks to detect deviations in instruction throughput during known GW events. The project is designed as a distributed, open-source, citizen science initiative, where geographically separated users collectively log timing anomalies, cross-correlate with official GW alerts, and analyze results using machine learning. We outline the theoretical motivation, architecture, expected challenges, and roadmap. We present the weak-field GW physics background, derive a mathematical model with simplified assumptions relating GW strain to fractional frequency and throughput deviations, express the computational workload cost in Big-O notation, model thermal and electronic jitter as noise processes, and develop statistical detection metrics (cross-correlation, matched filtering, false-alarm control). The goal is to provide a mathematical basis to guide prototype design and to estimate the scale required for plausible detection or meaningful upper limits.

Index Terms—Gravitational waves, citizen science, distributed computing, artificial neural networks, weak field approximation, timing stability, timing stability, matched filtering, distributed detection, ANN benchmark, thermal jitter.

I. INTRODUCTION

The direct detection of gravitational waves (GWs) by LIGO and Virgo marked the beginning of gravitational wave astronomy. These instruments are sensitive to spacetime strains on the order of 10^{-21}, a scale far beyond everyday measurement technologies. However, computing systems—whose performance depends critically on nanosecond-level clock precision—may serve as unconventional probes if timing anomalies can be correlated across multiple independent systems.

This proposal investigates whether predictable computational workloads can serve as timing baselines for auxiliary GW monitoring. Unlike interferometric detectors, which are optimized for strain sensitivity in the 10 Hz–10 kHz band, our system targets lower-frequency ranges (0.1 Hz–10 Hz), where subtle timing perturbations may leave statistical imprints in computation throughput.

II. WEAK-FIELD GRAVITATIONAL-WAVE EFFECTS ON CLOCKS AND COMPUTATION

In the transverse-traceless (TT) gauge, a weak gravitational wave propagating in the z direction induces a perturbation of the Minkowski metric:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}(t, z), \qquad |h_{\mu\nu}| \ll 1. \tag{1}$$

For a monochromatic plane wave with a single polarization component we may write

$$h(t) = h_0 \sin(2\pi f_{\mathrm{gw}} t + \phi_0), \tag{2}$$

where h_0 is the strain amplitude and f_{gw} its frequency.

A local clock (or oscillator) measures proper time τ defined by $d\tau^2 = -g_{\mu\nu} dx^\mu dx^\nu$. To leading order in h the GW produces a fractional modulation in observed clock rate. For a clock at rest in the chosen coordinates the fractional frequency shift $\Delta f / f$ induced by the GW is (to first order):

$$\frac{\Delta f(t)}{f} \approx h(t). \tag{3}$$

This result follows from the interpretation of $h(t)$ as a perturbation of the time-time portion or via the phase-modulation viewpoint: an oscillator with unperturbed phase $\phi_0(t) = 2\pi f t$ acquires an extra phase

$$\Delta\phi(t) = 2\pi f \int^t h(t') \, dt' \tag{4}$$

whose time derivative yields $\Delta f / f \sim h(t)$.

If a computational workload performs R arithmetic operations per second that scale linearly with clock frequency (a simplifying assumption), then the induced fractional change in the measured rate of operations is

$$\frac{\Delta R(t)}{R} \approx h(t). \tag{5}$$

Therefore the absolute change in operations/sec is

$$\Delta R(t) \approx R \, h(t). \tag{6}$$

III. COMPUTATIONAL WORKLOAD: COST AND SCALING

Consider a cascaded feed-forward fully connected ANN with L layers where each layer is an $n \times n$ dense transformation (for simplicity). A single forward pass requires $\mathcal{O}(Ln^2)$ scalar multiply-add operations for matrix-vector products. If

we instead use matrix–matrix operations (batched inputs or heavier kernels), per-iteration cost can be $\mathcal{O}(Ln^3)$.

For our prototype we assume per-core execution of I iterations of L layers on a vector of dimension n. The total floating-point operations (approximate multiply-adds) per benchmark run is:

$$\text{FLOPs} \sim I \cdot L \cdot 2n^2. \tag{7}$$

If a core executes R FLOPs/s (baseline), then the time for one run is $T = \text{FLOPs}/R$. The choice of (I, L, n) trades off per-run duration versus resolution: longer runs integrate more samples (reducing statistical uncertainty) but reduce temporal resolution for tracking low-frequency GWs.

IV. NOISE MODEL: THERMAL AND ELECTRONIC JITTER

Real hardware measurements are affected by noise sources that must be modeled to discriminate a GW-induced signal. We write the observed (normalized) fractional throughput deviation for node k as

$$x_k(t) = s(t) + n_k(t), \tag{8}$$

where $s(t) \equiv h(t)$ is the common GW-induced fractional signal (under the linear model) and $n_k(t)$ denotes node-specific noise.

We model $n_k(t)$ as a stochastic process with two dominant components:

$$n_k(t) = n_{\text{th},k}(t) + n_{\text{oss},k}(t), \tag{9}$$

where n_{th} is thermal/electronic jitter (approx. wide-sense stationary) and n_{oss} is low-frequency environmental/OS noise (power-line, scheduling spikes). A simple working model is Gaussian, zero-mean noise with one-sided power spectral density (PSD)

$$\mathcal{S}_n(f) = S_0 + \frac{S_1}{1 + (f/f_c)^\alpha}, \tag{10}$$

where S_0 models high-frequency white noise floor, and the second term models $1/f^\alpha$-like low-frequency excess with corner frequency f_c.

For a measurement integrated over time T (single run), the variance of the mean fractional deviation is approximately

$$\sigma_n^2(T) \approx \frac{1}{T^2} \int_{-\infty}^{\infty} \mathcal{S}_n(f) |\tilde{W}_T(f)|^2 df, \tag{11}$$

where $\tilde{W}_T(f)$ is the Fourier transform of the integration window (e.g., sinc for rectangular windows). For long T (low-pass), the integrated noise decreases roughly as $\sigma_n \propto 1/\sqrt{T}$ for white-noise dominated regimes.

V. DETECTION STATISTICS AND CROSS-CORRELATION

With N geographically distributed nodes (assumed independent noise realizations), we can form an averaged estimator of the common signal:

$$\bar{x}(t) = \frac{1}{N} \sum_{k=1}^{N} x_k(t) = s(t) + \bar{n}(t), \tag{12}$$

with $\text{Var}[\bar{n}(t)] = \sigma_n^2/N$ for identically distributed independent noise. Thus the signal-to-noise ratio (SNR) for a simple time-domain average over a measurement of duration T is

$$\text{SNR}_{\text{avg}} = \frac{\langle s \rangle_T}{\sigma_n/\sqrt{N}} \approx \frac{h_0}{\sigma_n/\sqrt{N}}, \tag{13}$$

where $\langle s \rangle_T$ denotes the characteristic amplitude of $s(t)$ over T (for narrow-band signals $\sim h_0$).

For a sinusoidal GW $s(t) = h_0 \sin(2\pi f_{\text{gw}} t)$, a frequency-domain matched filter provides an optimal linear detector given known waveform shape. The matched filter SNR (single node) is

$$\rho^2 = 4 \int_0^\infty \frac{|\tilde{s}(f)|^2}{\mathcal{S}_n(f)} df. \tag{14}$$

For N independent nodes coherently combined (assuming alignment and synchronization), SNR scales as $\sqrt{N}$.

A. Required ensemble size for a target SNR

Rearranging (13) yields an approximate required number of nodes to reach a target SNR_t:

$$N \gtrsim \left(\frac{\text{SNR}_t \, \sigma_n}{h_0} \right)^2. \tag{15}$$

This scaling is critical: because h_0 for astrophysical GW events can be extremely small (e.g., $h_0 \sim 10^{-21}$ at Earth for LIGO-scale events), achieving useful SNR requires either astronomically small noise σ_n or impractically large N unless one can increase the proportional response (e.g., amplify R or use specialized external clocks).

VI. PRACTICAL NUMERIC ESTIMATE (ILLUSTRATIVE)

Take a baseline per-node throughput $R = 10^{11}$ ops/s (high-end GPU/aggregate), and a GW strain amplitude $h_0 = 10^{-21}$. The expected absolute change in operations/sec (from (5)) is

$$\Delta R_{\text{amp}} = R \, h_0 = 10^{11} \times 10^{-21} = 10^{-10} \text{ ops/s.} \tag{16}$$

This is essentially zero compared to typical measurement quantization and noise. Realistic measurement noise in throughput due to thermal/OS effects can easily be $\sigma_R \gtrsim 10^3 - 10^6$ ops/s for single-node short integrations. The corresponding fractional noise is $\sigma_n = \sigma_R/R \gtrsim 10^{-8} - 10^{-5}$.

Using (15), for $\text{SNR}_t = 5$ and $\sigma_n = 10^{-6}$ (optimistic), required number of nodes is

$$N \gtrsim \left(\frac{5 \times 10^{-6}}{10^{-21}} \right)^2 \sim 2.5 \times 10^{31}, \tag{17}$$

an astronomically large number, demonstrating that direct detection with commodity hardware under the linear, per-clock proportional model is effectively impossible for strains of order 10^{-21}.

Nevertheless, this calculation is instructive: the experiment can still be valuable to place upper limits, study systematic noise sources, and explore non-linear coupling mechanisms or resonant amplification that might increase the effective response above h.

VII. MATHEMATICAL STRATEGIES TO INCREASE CONFIDENCE AND REDUCE FALSE TRIGGERS

Several mathematical and algorithmic techniques can improve detection confidence or reduce false positives:

1) **Matched filtering / coherent stacking:** If the expected GW waveform (or narrowband sinusoid) is known, matched filtering across time and across nodes maximizes SNR under Gaussian noise.
2) **Cross-correlation with interferometer triggers:** Use official LIGO/Virgo event times as a priori windows to restrict analysis and reduce trials factors.
3) **Time-frequency analysis:** Compute spectrograms and look for common excess power in the target 0.1–10 Hz band across many nodes; apply clustering algorithms to identify coherent features.
4) **Noise whitening and adaptive filtering:** Model each node's noise PSD $\mathcal{S}_{n,k}(f)$ and whiten the time series before stacking.
5) **Outlier rejection and robust statistics:** Use median-based estimators, trimmed means, or M-estimators to reduce sensitivity to transient OS spikes.
6) **False alarm control via permutation tests:** Estimate null distributions by time-shifting node data (off-source windows) and computing empirical false-alarm probabilities when searching for coincidences.

VIII. THERMAL JITTER MODELING AND MITIGATION

Thermal noise in oscillators and variations in power consumption produce timing jitter. A simplified oscillator phase noise model yields a single-sideband phase noise $\mathcal{L}(f)$ which can be mapped to fractional frequency noise PSD $\mathcal{S}_y(f)$. For practical mitigation:

- Measure each node's noise PSD during long idle windows to characterize $\mathcal{S}_{n,k}(f)$.
- Use temperature sensors and correlate throughput deviations with thermal variations; include these covariates in regression models to remove correlated noise.
- Employ Allan deviation estimates to identify stability on required timescales.

IX. STATISTICAL CORRELATION AND DETECTION CONFIDENCE

Given a candidate coincidence (many nodes showing correlated excess during an interferometer event window), compute a detection statistic Λ (e.g., normalized cross-correlation energy or matched-filter SNR). Estimate the background distribution of Λ using off-source time shifts or bootstrap resampling. Convert Λ to a p-value and assign a false alarm probability (FAP). Controlling FAP across multiple trials requires adjustments (Bonferroni, or better, controlling the false discovery rate).

X. CONCLUSIONS AND PRACTICAL IMPLICATIONS

The linear model predicts an impractically small direct coupling between astrophysical GW strains ($h \sim 10^{-21}$) and commodity-computer throughput. However, the mathematical framework above provides:

- Clear scaling laws (Eqn. 15) showing why large ensembles are required under linear coupling assumptions.
- A toolkit of statistical and signal-processing techniques to maximize sensitivity, reject false triggers, and place quantitative upper limits.
- A roadmap for experimental extensions: higher-precision local clocks, resonant mechanical transducers, or FPGA-based timestamping which could increase the proportional response.

REFERENCES

[1] B.P. Abbott et al., "Observation of Gravitational Waves from a Binary Black Hole Merger," *Phys. Rev. Lett.*, vol. 116, p. 061102, 2016.
[2] M. Maggiore, *Gravitational Waves, Vol. 1*, Oxford Univ. Press, 2008.
[3] D.W. Allan, "Statistics of atomic frequency standards," *Proc. IEEE*, 54(2):221–230, 1966.