Gravitational wave matched filtering by quantum Monte Carlo integration and quantum amplitude amplification

Koichi Miyamoto†, Gonzalo Morrás‡, Takahiro S. Yamamoto‡, Sachiko Kuroyanagi†‡, Savvas Nesseris†
†Center for Quantum Information and Quantum Biology, Osaka University, Toyonaka, Osaka, Japan
‡Instituto de Física Teórica UAM-CSIC, Universidad Autonóma de Madrid, Cantoblanco, Madrid, Spain

Abstract—The speedup of heavy numerical tasks by quantum computing is now actively investigated in various fields including data analysis in physics and astronomy. In this paper, we propose a new quantum algorithm for matched filtering in gravitational wave (GW) data analysis based on the previous work by Gao et al., Phys. Rev. Research 4, 023006 (2022) [arXiv:2109.01535]. Our approach uses the quantum algorithm for Monte Carlo integration for the signal-to-noise ratio (SNR) calculation instead of the fast Fourier transform used in Gao et al. and searches signal templates with high SNR by quantum amplitude amplification. In this way, we achieve an exponential reduction of the qubit number compared with Gao et al.’s algorithm, keeping a quadratic speedup over classical GW matched filtering with respect to the template number.

Index Terms—Quantum algorithms, quantum applications, gravitational wave, matched filtering, quantum Monte Carlo integration, quantum amplitude amplification.

I. INTRODUCTION

In this paper, we study an application of some quantum algorithms to matched filtering in gravitational wave (GW) detection experiments. Matched filtering is a commonly used technique to search a signal buried in noisy data and is widely used in GW data analysis. Given a target signal waveform, called a template, we take an inner product between a template and data to cancel out the noise contribution and extract a signal. In fact, the LIGO detectors discovered the first GW in 2015 using matched filtering [1].

A challenging point in GW matched filtering is the large number of templates. The functional form of a GW signal is predicted by the general relativity based theory for astronomical events such as black hole collisions, but it has some parameters such as the masses of the black holes. In order to get a high signal-to-noise ratio (SNR), we must perform matched filtering using a template with appropriate parameters. Therefore, usually, we set sufficiently many points in the parameter space and repeat matched filtering for them one by one. This is an extremely time-consuming task and expected to be sped up by quantum computing.

Recently, Ref. [2] presented a quantum algorithm for GW matched filtering. With the SNR calculation implemented as a quantum circuit, Grover’s search algorithm [3] finds a template with SNR higher than a threshold $\rho_{th}$ with $O(M/\sqrt{r(\rho_{th})})$ complexity, where $r(\rho_{th})$ is the fraction of templates that yield SNR $\rho \geq \rho_{th}$ and $M$ is the number of points in the time-series data of the detector output, or, equivalently, the number of frequency bins of Fourier transformed data. This is in fact a quadratic speedup over the classical method, which has $O(M/r(\rho_{th}))$ complexity, with respect to the template number.

However, this quantum algorithm has the following subtlety. It uses Fast Fourier Transform (FFT) [4] for the SNR calculation, which is also used in the usual classical way. FFT simultaneously calculates SNR for $M$ possible values of a parameter called time of coalescence, with other parameter fixed, in $O(M \log M)$ time, whereas naively such a computation takes $O(M^2)$ time without FFT. However, in order to store the intermediate and final calculation results, FFT requires $O(M)$ qubits, which is a somewhat large number since $M$ is typically of order $O(10^6)$ [5]. This might cause an issue on feasibility, since fault-tolerant quantum computers will have a limitation on the number of logical qubits available even in the future, because of the large physical qubit overhead for error correction [6].

In light of this, we propose an alternative quantum algorithm for GW matched filtering, in which the SNR calculation with FFT is replaced with the quantum algorithm for Monte Carlo integration (QMCI) [7]. Based on quantum amplitude estimation [8], QMCI estimates the expectation of a random variable, and can be applied to calculations of integrals and sums, with the oracle to compute the integrand or the summand given. Therefore, it can be used to calculate the SNR, which is in practice represented as a sum of contributions from many Fourier modes. We then search the high SNR templates by quantum amplitude amplification [8], an extension of Grover’s algorithm. In this approach, the required qubit number scales on $M$ as $O(poly(\log M))$, which means an exponential reduction from the FFT approach.

Note that this is not just a straightforward application of another quantum algorithm to a part of an existing method, since the use of QMCI causes the following issue. Unlike FFT, which calculates SNR deterministically, QMCI inevitably

$^{1}$In the big-O notation, we use a symbol $\tilde{O}(\cdot)$, which hides logarithmic factors in $O(\cdot)$.
accompanies errors, and thus comparing the SNR calculated by QMCI with a single SNR threshold \( \rho_\text{th} \) leads to a false alarm that the detector output yields SNR larger than \( \rho_\text{th} \) for some templates despite there being no such event. As a solution to this, we propose to set two thresholds \( \rho_\text{hard} \) and \( \rho_\text{soft} \) that have the following meanings: we should never miss events with SNR \( \rho \geq \rho_\text{hard} \); and we do not want to be falsely alarmed by events with \( \rho < \rho_\text{soft} \). Then, with QMCI accuracy set according to the diﬀerence between \( \rho_\text{hard} \) and \( \rho_\text{soft} \), the proposed algorithm says “there is a signal” for events with SNR \( \rho \geq \rho_\text{hard} \) with high probability, “there is no signal” for events with SNR \( \rho < \rho_\text{soft} \) with certainty, and either of these messages for events with SNR \( \rho \in [\rho_\text{soft}, \rho_\text{hard}) \). The query complexity in this algorithm is of order \( \tilde{O}(M/\sqrt{\rho_\text{hard}}) \), which still indicates a quadratic speedup.

As the full version of this paper, see [9].

II. Gravitational wave matched filtering

Suppose that we are given the detector output \( s(t) \) as a function of time \( t \), which is a sum of the signal \( h(t) \) and the noise \( n(t) \):

\[
s(t) = h(t) + n(t). \tag{1}
\]

We assume that the noise is Gaussian, which means that

\[
\mathbb{E}_n[\hat{n}(f)\hat{n}^*(f')] = \frac{1}{2}S_n(|f|)\delta(f - f'), \tag{2}
\]

where \( * \) denotes the complex conjugate, \( \hat{q}(f) := \int_{-\infty}^{\infty} dt e^{2\pi i ft}q(t) \) is the Fourier transform of the function \( q(t) \) in time domain, \( \mathbb{E}_n[\cdot] \) denotes an expectation with respect to randomness of the noise, \( S_n \) is the single-sided power spectrum density and \( \delta(\cdot) \) is the Dirac delta function. We define the inner product of two functions \( q(t) \) and \( q'(t) \) in time domain as

\[
(q|q') := 4\Re \left( \int_{-\infty}^{\infty} df \frac{\hat{q}^*(f)\hat{q}(f)}{S_n(f)} \right). \tag{3}
\]

The matched filtering search is performed by taking an inner product of \( s(t) \) and a function \( Q(t) \) that yields a large inner product with the targeted signal. \( Q(t) \) is often called a template and normalized as \( \langle Q|Q \rangle = 1 \). The template bank, the collection of templates, is prepared based on theoretically predicted waveform of signals. The SNR is then defined as

\[
\rho = \frac{\langle Q(s) \rangle}{\sqrt{\mathbb{E}_n[|Q(n)|^2]}} = 4\Re \left( \int_{0}^{\infty} df \frac{\hat{Q}^*(f)\hat{S}(f)}{S_n(f)} \right). \tag{4}
\]

Hereafter, we write each Fourier transformed template as \( \hat{Q}(f)e^{-2\pi i f_0} \), where \( f_0 \) is a parameter called the time of coalescence and the dependency on other parameters (intrinsic parameters) is put into \( \hat{Q} \). Here, we assume that there are \( N_\text{temp} \) candidates of the intrinsic parameter set in the template bank and label the functions \( \hat{Q}_m \) by \( m \in [N_\text{temp}] \).

In reality, the detector output is a sequence of discrete points in time. Denoting the number of the points and the time interval by \( M \) and \( \Delta t \) respectively, the SNR \( \rho_{m,j} \) for the \( m \)th intrinsic parameter set and \( l_0 = j\Delta t \) \((j \in [M])\) now becomes

\[
\rho_{m,j} = \frac{4}{M\Delta t} \Re \left( \sum_{l=1}^{N-1} \frac{\hat{Q}_m(f_k)\hat{S}(f_k)e^{2\pi i j f_k}}{S_n(f_k)} \right). \tag{5}
\]

where \( f_k := k/M\Delta t \) and \( M \) is assumed to be even.

Then, the problem boils down to the following: Determine whether there exists any \((m, j) \in [N_\text{temp}] \times [M] \) such that \( \rho_{m,j} \) in Eq. (5) exceeds some given value or not. If there are such integer pairs, find one of them.

III. Quantum algorithm for gravitational wave matched filtering

Instead of FFT in [2], the proposed quantum algorithm for GW matched filtering calculates the SNR by QMCI. Although FFT uses \( O(M) \) qubits to store intermediate and final calculation results, QMCI calculates the summation in Eq. (5) in the parallelism by the quantum superposition and requires only \( O(\text{poly}(\log M)) \) qubits.

The query complexity, that is, the number of the queries to the oracle to compute the summation in the proposed algorithm is evaluated as follows. We use the version of QMCI for the integrand with the bounded variance [7]. Given the bound \( \sigma^2 \) and the required accuracy \( \epsilon \), the query complexity of QMCI is \( \tilde{O}(\sigma/\epsilon) \). With respect to the search by QAA, the complexity increases by a \( O(\sqrt{M}) \) factor compared with the Grover search in Gao et al.’s FFT-based method. This is because in our method QMCI calculates \( \rho_{m,j} \) for one pair of \((m, j) \) and the pair with high \( \rho_{m,j} \) is searched in \([N_\text{temp}][M] \), whereas in the FFT-based method \( \rho_{m} := \max_{j}[M] \rho_{m,j} \) with fixed \( m \) is calculated and \( m \) with the high \( \rho_{m} \) is searched in \([N_\text{temp}] \). In total, the query complexity in our method is \( \tilde{O}(\sigma \sqrt{M}/\epsilon \sqrt{\rho_\text{th}}) \). Since we set \( \epsilon = O(1) \) as described below and we have \( \sigma = O(\sqrt{M}) \) (see [9]), the query complexity is resulting \( \tilde{O}(M/\sqrt{\rho_\text{th}}) \), which is similar to the FFT-based method.

We should also note that the nature of QMCI causes the following issue. The output of QMCI inevitably accompanies an error, and thus, even if the SNR of a given template calculated by QMCI exceeds the threshold \( \rho_\text{th} \), its true SNR might be below \( \rho_\text{th} \). We may think that we can evade such a misjudge by setting the accuracy \( \epsilon \) in QMCI extremely small, but it comes with large complexity. Therefore, we need to reasonably set the accuracy: following the nature of the problem under consideration, we should derive the error tolerance for SNR and set the QMCI accuracy matching it.

For this, we propose the following way. We consider GW matched filtering as a system that alarms us when the detector output seems to contain a signal. Besides, we consider the two levels of SNR threshold denoted by \( \rho_\text{hard} \) and \( \rho_\text{soft} \), which have the following meaning.

- If some templates have SNR \( \rho \geq \rho_\text{hard} \) for a given detector output, we want to be alarmed with certainty.
- We never want to be falsely alarmed when all templates have SNR \( \rho < \rho_\text{soft} \).

\[\text{For } n \in \mathbb{N}, \text{ we define } [n]_0 := [0, 1, ..., n - 1] \]

789
\( \rho_{\text{true}} \)

judged as mismatched
judged as matched or mismatched
judged as matched

\( \epsilon = \frac{\rho_{\text{hard}} - \rho_{\text{soft}}}{2} \); accuracy in QMCI

\( \rho_{\text{soft}} \geq \rho_{\text{hard}} \) or \( \rho_{\text{true}} < \rho_{\text{hard}} \) with high probability.

- When no template has SNR \( \rho \geq \rho_{\text{hard}} \) but some have \( \rho \in (\rho_{\text{soft}}, \rho_{\text{hard}}) \), it is not needed but fine to be alarmed.

In this situation, we set the QMCI accuracy to \( (\rho_{\text{hard}} - \rho_{\text{soft}})/2 \) and judge a template as matched if its SNR calculated by QMCI exceeds \( (\rho_{\text{hard}} + \rho_{\text{soft}})/2 \) and mismatched otherwise. In this strategy, with high probability, a template with a true SNR \( \rho \geq \rho_{\text{hard}} \) is judged as matched and that with a true SNR \( \rho < \rho_{\text{hard}} \) is judged as mismatched. A template with a true SNR \( \rho_{\text{soft}} \leq \rho < \rho_{\text{hard}} \) is an intermediate case where the data may contain a signal near the threshold and can be judged as either matched or mismatched due to the QMCI error. We present an illustration of this strategy in Figure 1. A reasonable setting of them is \( \rho_{\text{soft}} \) equal to some typical value (e.g. 8) and \( \rho_{\text{hard}} = \rho_{\text{soft}} + 1 \), where ‘1’ corresponds to the variance of the SNR due to the randomness of the noise.

Lastly, we present the theorem on the query complexity and the number of qubits used in the proposed method.

**Theorem 1.** (Informal) Suppose that we have accesses to oracles \( O_{\text{Re}} \) and \( O_{\text{Im}} \) such that, for every \( (m,k) \in [N_{\text{temp}}]_0 \times [M]_0 \),

\[
O_{\text{Re}}(m,k) |0\rangle = \begin{cases} |m, k\rangle |0\rangle & \text{if } k = 0 \\ |m, k\rangle \begin{pmatrix} 2Q_\text{m}(f_k) |0\rangle \langle 0| \\ 2Q_\text{m}(f_k) |1\rangle \langle 1| \end{pmatrix} & \text{otherwise} \end{cases}
\]

\[
O_{\text{Im}}(m,k) |0\rangle = \begin{cases} |m, k\rangle |0\rangle & \text{if } k = 0 \\ |m, k\rangle \begin{pmatrix} 2Q_\text{m}(f_k) |0\rangle \langle 0| \\ 2Q_\text{m}(f_k) |1\rangle \langle 1| \end{pmatrix} & \text{otherwise} \end{cases}
\]

Let \( \rho_{\text{soft}} \) and \( \rho_{\text{hard}} \) be real numbers such that \( 0 < \rho_{\text{soft}} < \rho_{\text{hard}} \) and \( \delta \in (0,1) \). Then, under some additional assumptions, there is a quantum algorithm that uses

\[
O\left( \log M + \log \left( \frac{\sqrt{M}}{\rho_{\text{hard}} - \rho_{\text{soft}}} \right) \right) \log \left( \frac{\sqrt{M}}{\rho_{\text{hard}} - \rho_{\text{soft}}} \right) \times \log \log \left( \frac{\sqrt{M}}{\rho_{\text{hard}} - \rho_{\text{soft}}} \right) \log \left( \frac{N_{\text{temp}} M}{\delta} \right)
\]

\[O_{\text{Re}}(m,k) |0\rangle = \begin{cases} |m, k\rangle |0\rangle & \text{if } k = 0 \\ |m, k\rangle \begin{pmatrix} 2Q_\text{m}(f_k) |0\rangle \langle 0| \\ 2Q_\text{m}(f_k) |1\rangle \langle 1| \end{pmatrix} & \text{otherwise} \end{cases}
\]

\( \rho_{\text{true}} < \rho_{\text{hard}} \) and \( \rho_{\text{true}} \geq \rho_{\text{hard}} \) with high probability.

\( (B) \) a message “there is no signal”.

- If \( r(\rho_{\text{hard}}) > 0 \), the algorithm outputs \( (A) \) with probability at least \( 1 - \delta \). In the algorithm, the number of queries to \( O_{\text{Re}} \) and \( O_{\text{Im}} \) is of order

\[
\mathcal{O}\left( \frac{\gamma M}{\rho_{\text{hard}} - \rho_{\text{soft}}} \sqrt{\rho_{\text{hard}} - \rho_{\text{soft}}} \right)
\]

- If \( r(\rho_{\text{hard}}) = 0 \) and \( r(\rho_{\text{soft}}) > 0 \), the algorithm outputs either \( (A) \) or \( (B) \). In the algorithm, the number of queries to \( O_{\text{Re}} \) and \( O_{\text{Im}} \) is of order as in Eq. (9).

\[
\text{Here}^3 \gamma := \max_{(m,k) \in [N_{\text{temp}}]_0 \times [M]_0} \frac{|Q_{\text{m}}(f_k)|}{\sqrt{\delta} \sqrt{M} \Delta t}.
\]

Note that, if \( r(\rho_{\text{soft}}) = 0 \), the algorithm outputs \( (B) \) with certainty, since, suggested some \( (m,j) \in [N_{\text{temp}}]_0 \times [M]_0 \) by the algorithm, we check \( \rho_{m,j} \) by classical calculation. For the detail of the algorithm, see [9].

**Acknowledgements**

G.M., S.K., and S.N. acknowledge support from the research project PGC2018-094773-B-C32, and the Spanish Research Agency (Agencia Estatal de Investigación) through the Grant IFT Centro de Excelencia Severo Ochoa No CEX2020-001007-S, funded by MCIN/AEI/10.13039/501100011033. S.K. is supported by the Spanish Attraction of Talento contract no. 2019-T1/TIC-13177 granted by Comunidad de Madrid, the I+D grant PID2020-118159GA-C42 of the Spanish Ministry of Science and Innovation and the i-LINK 2021 grant LINKA20416 of CSIC. T.Y. and S.K. are supported by Japan Society for the Promotion of Science (JSPS) KAKENHI Grant no. JP20H01899 and JP20H05853. K.M. is supported by MEXT Quantum Leap Flagship Program (MEXT Q-LEAP) Grant no. JPMXS0120319794 and JSPS KAKENHI Grant no. JP22K11924.

**References**

[1] B. P. Abbott et al. Phys. Rev. Lett., 116(13):131103, 2016.
[2] S. Gao et al. Phys. Rev. Res., 4(2):023006, 2022.
[3] L. K. Grover. In Proceedings of the twenty-eighth annual ACM symposium on Theory of computing, pages 212–219, 1996.
[4] J. W. Cooley, P. A. W. Lewis, and P. D. Welch. Proceedings of the IEEE, 55(10):1675, 1967.
[5] B. Allen et al. Phys. Rev. D, 85(12):122006, 2012.
[6] A. G. Fowler et al. Phys. Rev. A, 86(3):032324, 2012.
[7] A. Montanaro. Proc. R. Soc. A, 471(2181):20150301, 2015.
[8] G. Brassard et al. Contemporary Mathematics, 305:53, 2002.
[9] K. Miyamoto et al. arXiv preprint arXiv:2205.05966, 2022.

\( ^3 \) We expect that \( \gamma = O(1) \). See [9].