Quantum algorithms for approximate function loading

Gabriel Marin-Sanchez,1 Javier Gonzalez-Conde,1,2, † and Mikel Sanz1,2,3,4, †

1Department of Physical Chemistry, University of the Basque Country UPV/EHU, Apartado 644, 48080 Bilbao, Spain
2EHU Quantum Center, University of the Basque Country UPV/EHU, Apartado 644, 48080 Bilbao, Spain
3IKERBASQUE, Basque Foundation for Science, Plaza Euskadi 5, 48009, Bilbao, Spain
4Basque Center for Applied Mathematics (BCAM), Alameda de Mazarredo, 14, 48009 Bilbao, Spain
(Dated: September 25, 2023)

Loading classical data into quantum computers represents an essential stage in many relevant quantum algorithms, especially in the field of quantum machine learning. Therefore, the inefficiency of this loading process means a major bottleneck for the application of these algorithms. Here, we introduce two approximate quantum-state preparation methods for the NISQ era inspired by the Grover-Rudolph algorithm, which partially solve the problem of loading real functions. Indeed, by allowing for an infidelity $\epsilon$ and under certain smoothness conditions, we prove that the complexity of the implementation of the Grover-Rudolph algorithm without ancillary qubits, first introduced by Möttönen et al, results into $O(2^{\log n})$, with $n$ the number of qubits and $k_0(\epsilon)$ asymptotically independent of $n$. This leads to a dramatic reduction in the number of required two-qubit gates. Aroused by this result, we also propose a variational algorithm capable of loading functions beyond the aforementioned smoothness conditions. Our variational Ansatz is explicitly tailored to the landscape of the function, leading to a quasi-optimized number of hyperparameters. This allows us to achieve high fidelity in the loaded state with high speed convergence for the studied examples.

I. INTRODUCTION

Quantum computing has triggered a great interest in the last decades due to its theoretical capability to outperform classical information processing. Even though noise and decoherence are major drawbacks for the computational capacity of current quantum computers, quantum advantage has been experimentally achieved [1–3]. Unfortunately, these accomplishments lack any industrial or scientific relevance, so the search of a useful application still remains. In this sense, the realistic experimental implementation of many promising quantum algorithms in several fields like solving systems of linear equations [4, 5], performing data fitting [6], computing scattering cross sections [7, 8], pricing financial derivatives [9–11] or initial conditions in differential equations [12–14], is constrained by the assumption that data can be efficiently loaded into a quantum device. In this context, the efficient loading of classical data into quantum computers is a particularly important problem, and represents a major bottleneck of the practical application of quantum computation in the NISQ-Era, especially with the emergence of the quantum machine learning field [15–20].

There exist different quantum embedding techniques transforming classical data into quantum information [18, 21]. In particular, we can distinguish two main embedding protocols depending on how the information is encoded. On the one hand, the basis embedding, in which each bit value “0” or “1” is mapped to a computational qubit state $|0\rangle$ or $|1\rangle$, respectively [22]. In this way, the embedded quantum state corresponds to a uniform superposition of the bit-wise translations of binary strings. On the other hand, the amplitude-embedding technique encodes the normalized vector of classical data, which is now not necessarily binary, into the amplitudes of a quantum state [23–25, 27–35, 55]. In particular, these feature maps have been proposed to load discretized real valued functions [24, 25, 27, 55] with relevant applications in loading initial conditions for solving partial derivatives equations [11–14], computing Monte-Carlo integrations [9, 10, 36] and quantum field theory [37, 38]. However, a practical implementation of these approaches generally incurs into an overhead of resources, which can be reflect into either an exponential number of entangling gates [27–29, 39, 55, 56], or the employment of a huge number of ancillary qubits [31–33].

A rather different approach sustained by the Solovay–Kitaev theorem [41, 42] is based on the application of quantum generative models to efficiently accomplish an approximate amplitude encoding of discretized real valued functions [43, 44]. Nonetheless, in these generic ansätze, increasing the number of hyperparameters does not necessarily reflect in improving the expressability of the Ansatz to capture the function details [45]. Additionally, these variational methods usually suffer from training problems such as local minima and barren plateaus [46].

In this article, we present two approximate quantum algorithms to load real functions into quantum computers for the NISQ era. Our first protocol, inspired by the Grover-Rudolph algorithm and its implementation proposed by Möttönen et al [27, 55], implements the algorithm without ancillary qubits with complexity $O(2^{\log n \epsilon})$ for functions whose second logarithm derivative is upper bounded by a constant, with $n$ the number of qubits, $\epsilon$ the infidelity respect to the exact state, and $k_0(\epsilon)$ asymptotically independent of $n$. This leads to a dramatic reduction in the number of required two-qubit gates. Inspired by this result, we also introduce and benchmark a variational quantum circuit with applications in a broader family of functions. Our proposed Ansatz is adapted to the structure of the function, which intuitively correlates hyperparameters and expressibility. Moreover, by taking the angles provided by the Grover-Rudolph protocol, we can define a suitable ini-

* Corresponding author: javier.gonzalez@ehu.eus
† Corresponding author: mikel.sanz@ehu.eus
tial training angle set, which considerably improves the training process, avoiding barren plateaus and local minima. Finally we have numerically proven resilience of our algorithm against several noise sources.

II. THE GROVER-RUDOLPH ALGORITHM

The method of Grover and Rudolph, originally proposed in Ref. [55], describes a constructive protocol to load into an $n$-qubit quantum state the discretized version $\{f\}$ of certain integrable density function $f: \{x_{\min}, x_{\max}\} \subset \mathbb{R} \rightarrow \mathbb{R}$ as

$$|\Psi(f)\rangle_n = \sum_{i=0}^{2^n-1} \sqrt{f_i} |i\rangle.$$  \hfill (1)

A. Grover and Rudolph Algorithm without ancillas

The first proposal in the literature that provided an explicit circuit implementation of the Grover-Rudolph idea without using ancillary qubits was proposed by M. M"ott"onen et al [27]. Without using ancillas, this protocol provides a constructive algorithm which applies a sequence of operation blocks, $F_k^{(k-1)}(y, \theta^{(k-1)})$, to the initial state $|0\rangle^{\otimes n}$. For $k \in \{1, \ldots, n\}$, each $k$-qubit block, $F_k^{(k-1)}(y, \theta^{(k-1)}) = \sum_{l=0}^{2^{k-1}-1} |l\rangle \otimes R_y(\theta_l^{(k-1)})$, corresponds to a uniformly-controlled $y$-axis rotation, where the $l$-th component of the angle vector $\theta^{(k-1)}$ is calculated as

$$\theta_l^{(k-1)}(l) = 2 \arccos \left( \frac{\int_{x_{\min}+\delta_l}^{x_{\max}+\delta_l} f(x) dx}{\int_{x_{\min}+\delta_l}^{x_{\max}+\delta_l} f(x) dx} \right).$$ \hfill (2)

Here, $\delta_l = \frac{x_{\max} - x_{\min}}{2^n}$ and $l \in \{0, \ldots, 2^{k-1} - 1\}$ is the index corresponding to the $(l+1)$-th subinterval of the $2^{k-1}$ partition of the interval $[x_{\min}, x_{\max}]$. Each multi-controlled gate comprising $F_k^{(k-1)}(y, \theta^{(k-1)})$ is denoted by $\wedge_{k-1} \{ R_y(\theta_l^{(k-1)}) \}$ and bisects the $(l+1)$-th partition interval of the function by using the conditional probability of being in right or left side of the interval, as depicted in Fig. 1.

In the complete algorithm, the total number of angles needed scales as $\sum_{m=1}^{n} 2^{m-1} = 2^n - 1$, which is exponential in the number of qubits and requires an exponential number of multi-controlled C-NOT gates. Therefore, without using ancillary qubits, the required number of two qubit gates scales exponentially as the number of angles required do, and therefore the circuit complexity of implementing Grover-Rudolph algorithm without ancillas is $O(2^n)$.

We can conclude that this protocol without ancillas is theoretically capable of loading a discretized density function at the cost of an exponential overhead of resources [39, 56] to prepare the state in Eq. (1) using blocks $F_k^{(k-1)}(y, \theta^{(k-1)})$.

B. Grover and Rudolph Algorithm with ancillas

According to the original Grover-Rudolph algorithm [55], for each step $k$, we can efficiently prepare $|\Psi_k\rangle = \sum_{l=0}^{2^n-1} \sqrt{f_l} |l\rangle$ by first loading the rotation angles, Eq. 2, with bit precision $m$ into a bit string of $m$ ancillary registers and then performing $k$ controlled rotations of angle $2\pi/2^l$, $j = 1, \ldots, m$ controlled by the ancillary registers. Indeed, the encoding of the rotation angles can be efficiently achieved if we have access to the oracle:

$$|\Psi_k\rangle = \sum_{l=0}^{2^n-1} \sqrt{f_l} |l\rangle |0\rangle^{\otimes m} \rightarrow |\Psi_k\rangle = \sum_{l=0}^{2^n-1} \sqrt{f_l} |l\rangle |\theta^{(k-1)}_l(I)\rangle^{m \text{ bit precision}}$$ \hfill (3)

An example case when this operation can be efficiently performed is when the function $\theta^{(k-1)}_l(l)$ can be well approximated by a polynomial and then, implemented by employing a polynomial amount of classical half adder operations, i.e. NAND gates. Lastly, the NAND gates are efficiently mapped into Toffoli quantum gates by making use of at most 3 qubits per classical bit [47]. As this map is input-dependent, it enables us to compute it simultaneously for all $\theta^{(k-1)}_l(l)$ when the input state is a quantum superposition. However, this oracle is not explicitly provided in the original manuscript and the efficient circuit to which the authors refer is not presented, remaining as an oracle. Additionally, the original manuscript of Grover and Rudolph does not provide the analytical bounds of these approximations ($m$ bit precision, oracle implementation), as well as implicitly makes use of additional ancillary qubits which incurs into an important cost for the NISQ era. Last but not least, some recent works have risen criticism about the feasibility of this original proposal [14, 48, 49]. Moreover, to our best knowledge, there is no explicit efficient implementation of this oracle in terms of gates without employing ancillary qubits.

![Fig. 1. Effect of the uniformly-controlled y-axis rotation $F_k^{(k-1)}(y, \theta^{(k-1)})$ for $k = 3$ and $n$ qubits. Each multi-controlled gate comprising $F_k^{(k-1)}(y, \theta^{(k-1)})$ bisects the $(l+1)$-th partition interval of the function by using the conditional probability of being in the right or left side of the interval.](image-url)
two contiguous angles is bounded by the second derivative of the logarithm of \( f \) in the following way:

\[
|\hat{\theta}^{(k-1)}_{l+1} - \hat{\theta}^{(k-1)}_{l}| \leq \frac{\delta_k^2}{4} \max_{y \in [\delta_0 \pi, \delta_k \pi]} \left| (\partial_y^2 \log f(y)) \right|, \tag{5}
\]

where \( \delta_k = \frac{\pi}{2^{k+1}}, l \in \{0, \ldots, 2^{k-1} - 2\} \).

Consequently, if \( |\partial_y^2 \log f(y)| \leq \eta \), then for each block comprising a uniformly-controlled rotation of \( k \) qubits, the difference between any two angles is bounded by

\[
|\hat{\theta}^{(k-1)}_{l} - \hat{\theta}^{(k-1)}_{l'}| \leq \frac{\delta_k^2}{8} \eta \quad \forall l, l' \in \{0, \ldots, 2^k - 1\}. \tag{6}
\]

This result allows us to cluster angles of each block in which Eq. (6) is fulfilled. Thus, we define a cluster representative angle, \( \tilde{\theta}^{(k-1)} \), as

\[
|\hat{\theta}^{(k-1)} - \tilde{\theta}^{(k-1)}| \leq \frac{\delta_k^2}{8} \eta :\eta_k \quad \forall l \in \{0, \ldots, 2^k - 1\}. \tag{7}
\]

\( \forall k \geq k_0, \) with \( k_0 + 1 \) the index of the first block in which Eq. (6) is fulfilled (successive blocks also verify it). Note that \( |\partial_y^2 \log f(y)| \leq \eta \) is a sufficient condition for the bound in Eq. (5), however, if in a singularity point it grew slower than \( \frac{1}{\eta^2} \), the difference between consecutive angles still vanishes as we will analyze later.

We now analyze how clustering the angles according to Eq. (7) affects the final error of the process, measured by means of the fidelity with respect to the exact discretized state \( |f(x)\rangle \).

Let us analyze each part of the algorithm. First, we provide a sufficient condition on the target density function to guarantee an upper bound over the difference between two contiguous angles.

**Lemma 1.** Let \( f \) be a continuous function such that \( f : [0, 1] \to \mathbb{R}^+ \) and consider a block comprising a uniformly-controlled rotation of \( k \) qubits. Then, the difference between

\[
\text{Definition 1.} \text{ Let } f : [0, 1] \to \mathbb{R}^+ \text{ be a positive function in } L^2([0, 1]). \text{ We define the } n\text{-qubit normalized representative state of } f(x) \text{ as the } n\text{-qubit state } |f(x)\rangle_n = \sum_{\delta_n}^n f(\delta_n)|\delta_n\rangle, \text{ with } \delta_n = \frac{1}{2^n} \text{ and } \sum_{\delta_n}^n f^2(\delta_n) = 1.
\]

According to this definition, we encode the discretized function into the amplitude of a quantum state, in contrast to the Grover-Rudolph algorithm, which does it in the probability, Eq. (1). We also consider our function defined in the interval [0, 1] as a standardization required.

**Theorem 1.** Let \( f : [0, 1] \to \mathbb{R}^+ \) be a positive integrable function in \( L^2([0, 1]) \) and \( 0 \leq \eta \leq 8\pi \) a constant such that \( \eta = \sup_{x \in [0, 1]} |\partial_x^2 \log f^2(x)| \). Then, it is possible to approximate the \( n\)-qubit representative state of \( f(x), |f(x)\rangle_n \), by a quantum state \( |\Psi(f)\rangle_n \) such that the fidelity \( ||\Psi(f)||f(x)\rangle_n^2 \geq 1 - \epsilon \) with at most \( 2^{k_0} - 1 \) two-qubit gates, with

\[
k_0(\epsilon) = \max \left\{ \left\lceil -\frac{1}{2} \log_2(4\pi^2 - \frac{96}{\eta^2} \log(1 - \epsilon)) \right\rceil, 2 \right\}, \tag{4}
\]

and the circuit to perform it is provided in Fig. 2(i).

Let us analyze each part of the algorithm. First, we provide a sufficient condition on the target density function to guarantee an upper bound over the difference between two contiguous angles.

**Lemma 1.** Let \( f \) be a continuous function such that \( f : [0, 1] \to \mathbb{R}^+ \) and consider a block comprising a uniformly-controlled rotation of \( k \) qubits. Then, the difference between

\[
\text{FIG. 2. (i) Quantum circuit performing the protocol presented in Th. 4, based on the Grover-Rudolph method for a system of } n \text{ qubits. We cluster the angles of the blocks for } k \geq k_0 + 1 > 2, \text{ leading to a drastic reduction in the number of gates needed. (ii) Values of } k_0 \text{ for } \eta \in [0, 8\pi], \epsilon \in [0.0001, 0.01] \text{ and } n \to \infty, \text{ following Eq. (10), where the dotted lines correspond to the contour lines and denote the change of values. We can appreciate that } k_0 \leq 10 \text{ for most of the cases.}
\]
is at most $\eta_k$, i.e. $|\tilde{\theta}^{(k)} - \theta^{(k)}| \leq \eta_k$ for $l = 0, \ldots, 2^{k-1} - 1$ and $k = 1, \ldots, n$. Then, the following lemma can be proven:

**Lemma 2.** Consider a system of $n$ qubits and an error $\eta_k$ between any angle of the k-th block and its representative such that $\eta_k \leq \pi$, with $k = 1, \ldots, n$. Then, the fidelity between the final states with and without clustering, $F = |\langle \phi|\tilde{U}_n|\phi\rangle|^2$, satisfies

$$F \geq \prod_{k=1}^{n} \cos^2 (\eta_k/2). \tag{8}$$

Assuming that we cluster angles of the blocks for $k \geq k_0 + 1 > 2$, then

$$F \geq \prod_{k=k_0+1}^{n} \cos^2 \left( \frac{\eta}{8^k} \right) \geq e^{-\frac{\eta}{2} (\log 8^{-1})^2} \,: = \, F_0, \tag{9}$$

since $\cos(x) \geq e^{-x^2}$ for $x \leq \pi/2$. Therefore, if an infidelity $\epsilon = 1 - F_0$, is allowed and the angles of all blocks comprised of more than $k_0(\epsilon)$ qubits are clustered, with $k_0$ given by Eq. (4), then the fidelity satisfies $F \geq 1 - \epsilon$. In the asymptotic limit of $n \to \infty$, we have that $k_0(\epsilon)$ tends to

$$k_0(\epsilon) \to \max \left\{ \left| -\frac{1}{2} \log_2 (\frac{96}{\epsilon^2} \log(1-\epsilon)) \right|, 2 \right\}, \tag{10}$$

which is independent of the system size, $n$. In Fig. (2)(ii), we have depicted the values of $k_0$ at this limit, for $\eta \in [0, 8\pi]$ and $\epsilon \in [0.0001, 0.01]$.

We finally study the implementation cost of the proposed protocol. We only take into account the latter type of gates and ignore single-qubit operations [50]. Using the result of [39, 56], which illustrates how a uniformly-controlled rotation of $k$ qubits can be implemented with $2^k-1$ CNOTs, the complexity of the circuit described in Th. 4 is $O(2^{k(\epsilon)}).$

**A. Normal Distribution**

We apply the algorithm given by Th. 4 to a normal distribution with a mean value $\mu = 0.5$ and for different values of the variance $\sigma$. We numerically benchmark the fidelity attained by our first protocol using the value of $k_0$ resulting of Eq.(4) when we assume a maximum infidelity of $\epsilon = 0.05$ and a system of 8 qubits. Notice that for this distribution, $\eta = 2/\sigma^2$. In Fig. 3(i) and Table I, we have depicted the results from the simulations of $|f(x)\rangle_8$ and $|\Psi(f)\rangle_8$, for different values of $\sigma$. We appreciate that the condition of $F \geq 0.95 = 1 - \epsilon$ is not only satisfied in all cases, but the fidelities obtained are considerably better. Furthermore, the significant reduction in the number of two-qubit gates required to achieve these results is noteworthy. In the worst case, for $\sigma = 0.3$, the quantity of gates needed represents the 12.16% of the original set, while the fidelity of the experiment reaches 0.99841.

**B. Generalization: Singular Points**

In this section we study the generalization of the Th. 4 when the functions are allowed to have singularities on the
Consider a function \( f : [0, 1] \to \mathbb{R}^+ \) with
\[
|\partial^2_x \log f(x)| \xrightarrow{x \to 0} \infty, \quad (11)
\]
An example is the beta density function, defined as
\[
f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)}, \quad (12)
\]
with \( \alpha, \beta > 0 \) and \( B(\alpha, \beta) \) the beta function. Then, the second derivative of its logarithm is
\[
\partial^2_x \log f(x) = \frac{1 - \alpha}{x^2} + \frac{1 - \beta}{(1-x)^2}. \quad (13)
\]
Thus, for \( \alpha \neq 1 \), we have a singularity at \( x = 0 \). Also, we see that if \( \beta \neq 1 \), there is a singularity at \( x = 1 \).

In this situation, \( |\partial^2_x \log f(x)| \) can not be bounded by a finite factor \( \eta \) in the whole interval and, hence, Th. 4 can not be applied. However, under certain circumstances, this issue can be solved. Recall that, in Lem. 1, we obtained that the difference between two consecutive angles in a block of \( k \) qubits satisfies
\[
|\partial^k_{l+1} - \partial^k_l| \leq \frac{\delta^2}{4} \left| \partial^2_x \log f(x) \right|, \quad (14)
\]
with \( \delta_k = \frac{1}{\eta^2}, l \in \{0, 1, \ldots, 2^k - 2\} \), and \( x = l\delta_k \). Since the singularity is found in \( x = 0 \), there exists a \( k_{\text{max}} \) from which the maximum of \( |\partial^2_x \log f(x)| \) in \([2^{-k_{\text{max}}+1}, 1]\) is found in \( x = 2^{-k_{\text{max}}+1} \). Then,
\[
|\partial^k_{l+1} - \partial^k_l| \leq \frac{1}{4} \left| \partial^2_x \log f(x)|_{x=2^{-k_{\text{max}}}} \right|. \quad (15)
\]
Next, if for the limit \( k \to \infty \) the term \( |\partial^2_x \log f(x)|_{x=2^{-k+1}} \to 0 \), this value is decreasing and there exists a \( k^* \geq k_{\text{max}} \) from which an \( \eta = |\partial^2_x \log f(x)|_{x=2^{-k+1}} \) can be set so the clustering conditions \( \eta \leq \pi \Rightarrow \eta \leq 2^{-k+1}8\pi \) are satisfied. This \( k^* \) acts as the \( k_0 \) of Th. 4 and represents the last block without clustering. Then, for \( k > k^* \), the angles corresponding to the interval \([2^{-k+k}, 1]\) can be clustered, following our protocol. The case for a singularity in \( x = 1 \) is analogical and in Fig. 4 the resulting gates corresponding to the process of clustering the inner angles for a block of 3 qubits are depicted. Given that \( R_x(\tilde{\theta})R_x(-\tilde{\theta}) = 1 \), we can complete the identity of the clustered block by subtracting \( \tilde{\theta} \) from the rest of the angles.

In this situation, if we consider that singularities exist in \( x = 0, 1 \), the total number of two-qubit gates required to load \( f \) into a quantum state is
\[
\#TQGs = \sum_{k=1}^{k^*} 2^{k-1} + 2 \sum_{k=k^*+1}^{n} (80k - 398) = 2^{k^*} - 1 + 2(n-k^*)(20k^* + 2n - 179), \quad (16)
\]
see the Appendix A for further details in gates decomposition.

Notice that this result is only valid for \( k^* > 6 \), but since we are interested in the asymptotical behaviour of the protocol, it is not an issue. All in all, we obtain that the complexity of this process is exponentially dependent on \( k^* \) with an extra polynomial term.

On the other hand, if a singularity is found in \((0, 1)\), we end up with multiple clusters of angles in each block. The reason behind this is that the clusters are formed with contiguous angles. Therefore, the reduction of gates is not significant, and the protocol can not be performed efficiently (polynomial). However, if the representatives of the disjointed clusters are equal, then we can create a single cluster, so the reduction is doable. In this sense, we have numerically observed that far from the singularities, all angles converge to a value of \( \pi/2 \), but we have not found an analytical proof yet.

Let us see an example of a function that meets the previous description and analyze the outcome of the protocol. Consider the function
\[
f(x) = \frac{1}{N} e^{x^{1/2}}, \quad (17)
\]
in \([0, 1]\), where \( N \) is the normalization factor, and a system of \( n = 10 \) qubits. The second derivative of its logarithm is
\[
\log f(x) = x^{1/2} - \log N \Rightarrow \partial_x \log f(x) = \frac{3}{2} \sqrt{x} \quad (18)
\]
\[
\Rightarrow \partial^2_x \log f(x) = \frac{3}{4 \sqrt{x}}. \quad (19)
\]
Hence, \( \partial^2_x \log f(x) \) has a singularity at \( x = 0 \). First of all, since \( \partial^2_x \log f(x) \) is a monotonic decreasing function, its maximum in the interval \([2^{-k+k+1}, 1]\) is found in \( x = 2^{-k+k+1} \), for any \( k = 1, \ldots, n \). Then, we can set \( k_{\text{max}} = 1 \). Next, we need to compute the limit of \( |\partial^2_x \log f(x)|_{x=2^{-k+k+1}} \) to \( 2^{-2k+2} \):
\[
|\partial^2_x \log f(x)|_{x=2^{-k+k+1}} = 2^{-2k+2} = \frac{3}{4 \sqrt{x}}|_{x=2^{-k+k+1}} = \frac{3}{4 \sqrt{x}} e^{x^{1/2}} = \frac{3}{4 \sqrt{x}} e^{x^{1/2}} \xrightarrow{k \to \infty} 0. \quad (20)
\]
in NISQ quantum processors, we observe a trade-off, see Fig. 6, between the clusteringization error and the experimental error, quite similar to the trade-off observed in digital quantum simulations between the number of Trotter steps and the experimental error in Refs. [52, 53]. This balance is crucial for algorithms in noisy quantum processors. Reproducing a similar reasoning as the one in the aforementioned reference and references thereof, we first propose an approximated model of how the experimental error combined with the clustering error of our algorithm scales as a function of the number of non-parallel two-qubit gates. Then, we perform some numerical simulations introducing multiple realistic noises in our algorithm to support our theoretical predictions.

In order to establish a theoretical framework to understand the behavior of our system when clusterization and experimental errors are considered, we make the assumption that the main source of experimental noise comes from the two-qubit gates, while ignoring the noise arising from single-qubit rotations. Additionally, we consider that the application of each of these gates onto a quantum state, denoted as \( \rho \), is modeled in the following form

\[
\rho \to (1 - \xi)U^\dagger \rho U + \xi U^\dagger \rho U \tag{23}
\]

Where \( U \) corresponds to the desired dynamics while the second term introduces a certain Taylor expansion of the deviation from the exact evolution, with \( \xi \ll 1 \). After \( m \) non-parallel gates, ignoring quadratic terms in \( \xi \), the infidelity of the state has approximately evolved according to

\[
1 - \exp F \sim m \xi \|U\| \tag{24}
\]

as the leading term. Note that the argument could be extended by replacing \( U \to T(\cdot) \), an arbitrary quantum channel, but the calculation is more complicated to produce a similar argument. In a rough approximation, for a certain value of the clustering index \( k_0 \), the total fidelity of the protocol run on a

![Graph](image)

**FIG. 5.** Simulation of \(|f(x)\rangle_{10}\) and \(|\Psi(f)\rangle_{10}\) for the function defined in Eq. (17), \( n = 10, \epsilon = 0.01 \) and the clustering process performed for \( k = 3, \ldots, n \).

Therefore, the difference in the angles is bounded and we can select a \( k^* \geq 1 \) for which the conditions of Th. 4 are satisfied in \([2^{-k^*+1}, 1]\). The first condition we need to check is the inequality given by

\[
\frac{\delta_k}{8} \eta \leq \pi, \tag{21}
\]

with \( \eta = |\partial^2 \log f(x)|_{x=2^{-k^*+1}} = \frac{32}{3} \frac{1}{2^{k^*+1}} \). Then,

\[
\frac{3}{32} 2^{(k^*-1)} 2^{-k+1} = \frac{3}{32} 2^{1/2(1-k^*)}. \tag{22}
\]

Since this term is decreasing, its maximum is found when \( k = 1 \), with a value of \( \frac{3}{32} \pi \). Thus, this condition is never met for any \( k = 1, \ldots, n \). As we can select \( k^* = 1 \). Additionally, if we compute the value of \( k_0 \) with \( \epsilon = 0.01 \) and \( \eta = |\partial^2 \log f(x)|_{x=2^{-k^*+1}} = 0.75 \) using Eq. (4), we obtain \( k_0 = 2 \). Then, the last block to remain unclustered must be the maximum between \( k_0 \) and \( k^* \), which in this case is 2.

Now, in Fig. 5, we have depicted the result of the experiment of the considered function for a system of \( n = 10 \) qubits and the clustering starting with the 3-qubit block, following the protocol described in this section. With a fidelity of 0.99975, larger than the one required, we have that the final state \(|\Psi(f)\rangle_{10}\) successfully captures the features of \( f \) with a reduction of the complexity of two-qubit gates from \( O(2^n) \) to \( O(2^3) \).

### C. Analysis of resilience to noise in NISQ era

In this subsection, we present a theoretical and numerical analysis of how experimental errors affect different clustering levels in our first protocol and their impact on the final fidelity. The crucial point here is that when digital accuracy increases, the number of gates requested grows exponentially. As the introduction of these gates implies a growing experimental error

![Graph](image)

**FIG. 6.** Total fidelity, \( F_{\text{total}} \), in terms of the clustering infidelity assumed in the protocol according to the expression \( \epsilon(k_0) \sim 1 - e^{-F/24(4^0.4 - 4^{-k})} \), where \( \eta = 22.22, k_0 = 1, \ldots, 6 \) and \( n = 6 \). As we can appreciate, the expected fidelity of our protocol is above the fidelity resulting of implementing the protocol without clustering in noisy devices, which dramatically tends to 0. Actually, the fidelity reaches a maximum for \( k_0 = 4 \), which corresponds to a clustering error \( \epsilon_0 = 0.0726 \).
Decoherence time (us)  Jakarta
0.01 2.457E-04 8.328E-03

CX gate time (ns)  114.84

Dephasing time (us)  0.7 0.8 0.9 1.0 0.0 0.1 0.2 0.3 0.4 0.5 0.6

FIG. 7. (i) and (ii) Total fidelity, $F_{\text{total}}$, in terms of the clustering infidelity assumed and $k_0$ in the protocol according to the expression $\epsilon(k_0) = e^{-\eta/244^{4-0.4-n}}$, where $\eta = 22.22$, $k_0 = 1, ... , 6$, and $n = 6$ for the ideal and noisy cases. (iii) $l_2$ normalized error in terms of the clustering in the same conditions. The maximal fidelity, respectively the minimal error measure with the $l_2$ norm, is achieved for a clusterization level $k_0 = 2$, which corresponds to $\epsilon_0 = 0.7222$

A NISQ device has a contribution coming from this experimental noise together with the digital infidelity due to the clustering procedure. We can approximate this by

$$F_{\text{total}} \sim e^{-\frac{\sigma^2}{244^{4-0.4-n}}} - 2\xi \|\bar{U}\| (1 + 2^{k_0})$$

which in terms of the clustering error, $\epsilon$, by using that $1 - \epsilon \sim e^{-\frac{\sigma^2}{244^{4-0.4-n}}}$, and hence,

$$2^{k_0} - 1 \sim \left(4^n - \frac{96}{\eta^2} \log(1 - \epsilon)\right)^{\frac{1}{2}} - 1,$$

the function. This condition holds

$$F_{\text{total}} \sim 1 - \epsilon - 2\xi \|\bar{U}\| \left(\left(4^n - \frac{96}{\eta^2} \log(1 - \epsilon)\right)^{\frac{1}{2}} - 1\right).$$

From the equation above, we can find the value of clustering error $\epsilon_0$ for which $F_{\text{total}}$ achieves its maximum by deriving the function. This condition holds

$$4 (1 - \epsilon_0)^2 (4^n - \mu \log(1 - \epsilon_0))^3 = \alpha^2 \mu^2$$

with $\alpha = 2\xi \|\bar{U}\|$ and $\mu = \frac{96}{\eta^2}$, which is a transcendent equality, so it cannot be analytically solved. However, for the range of parameters in which we are interested and for sufficiently small, we can approximate $\epsilon_0 \approx \frac{\alpha^2 \mu^2}{\sqrt{\alpha^2 \mu^2}}$.

In order to provide a numerical example for this expression, we analyze a normal distribution with $\sigma = 0.3$ encoded in $n = 6$ qubits. For the experimental error, we consider $\alpha = 2\xi \|\bar{U}\| = 0.0087$. We depict the results of this analysis in Fig. 1. If we compare results from this analysis, we can see that the maximum fidelity is achieved for a value of $\epsilon_0 = 0.0726$ while the predicted value reads $\epsilon_0 = 0.1131$.

Once the theoretical framework has been established, let us carry out some numerical simulations analyzing the robustness (in terms of fidelity) against different noises. The main objective is to support the aforementioned analytical findings. The guidelines of the numerical experiment is described as follows. The circuit is transpiled to a native set of gates given by $\text{CNOT}, \text{Id}, R_z(\theta), X$ and $S_x$. The noise quantum channels considered are

- Bit-Flip (Pbf)
- Amplitude-Damping (T1)
- Dephasing (T2)
- Gate errors (rD, CNOT error)
- Measurement error (pmeas)

For a realistic scenario, we have taken the value of these errors from the IBM Jakarta quantum processor, which are summarized in Table II. Considering this set up, we have studied again the normal distribution with $\sigma = 0.3$ encoded in $n = 6$ qubits, focusing on the fidelity and $l_2$ norm (normalized with the system size such that it converges to the $L_2$ norm in the continuous limit) with respect to the exact discretized state. The numerical results are depicted in Fig. 7, which also shows a trade-off between the clustering and the experimental error. More explicitly, the maximal fidelity, respectively the minimal error measure with the norm $L_2$, is achieved for a clusterization level $k_0 = 2$, which reproduces the structure showing a trade off between errors predicted by our theoretical model. However the maximum is reached for a smaller value of $k_0$, compared to the predictions, as the theoretical model is a first order simplification which becomes not that accurate when the

| Parameter | Description | Value |
|-----------|-------------|-------|
| SQG time | Single qubit gate time (ns) | 35 |
| CX time | CX gate time (ns) | 540 |
| rD | Deviation ratio for the single qubit gates | 2.457E-04 |
| Pbf | Bit-flip error during the rz gate | 2.457E-04 |
| CNOTerror | Deviation ratio for CNOT gate | 8.328E-03 |
| pmeas | Readout error | 2.23E-01 |
| Pth | Thermal population of the ground state | 0.01 |
| T1 | Decoherence time (us) | 114.84 |
| T2 | Dephasing time (us) | 38.65 |

TABLE II. Noise parameters description and their value. We have estimated the numerical values from the calibration data provided for the IBM device ‘ibm jakarta’.
FIG. 8. Illustration of proposed variational circuit for a general function. (a) Target function. The proposed landscape generalizes the possible cases of special points that lead to singularities in the second logarithmic derivative. The special points are classified in the legend of the graph, where \( z \) denotes the zeros, \( s \) the singularities, and the index denotes whether the slope in the contiguous points has the same sign, index \( i = 1 \), or not, index \( i = 2 \). This is translated to a minimum amount of 7 necessary parameters required to capture the local behavior of the function in these points, one per index \( i = 1 \) and 2 per index \( i = 2 \). We add an additional parameter which represents the clusterization. (b) Variational circuit. The blocks which have a number of angles lower or equal to the necessary parameters will remain unclustered. From the first block which has more angles, we proceed to reduce the number of angles to the number of parameters by clustering the ones that do not correspond to the special points. This leads to a significant reduction in the number of gates. (c) Zoom in of the \( F_{4,5}(y, \theta) \) block clusterization. As we only need 8 parameters, the 16 angles of this block are reduced to the half. We keep one local rotation for the clustering parameter and 7 multi-controlled gates corresponding to the intervals where the singularity is located.

Consequently, although our protocol introduces a controllable error, it significantly reduces the depth required for the computation as well, resulting in a more balanced and reliable outcome. In the presence of experimental noise, our algorithm achieves a good balance between fidelity (experimental + clustering errors) and feasibility (realistic depth), which is crucial for this stage of quantum computing, where computers are characterized by high error rates and limited coherence times.

IV. VARIATIONAL QUANTUM CIRCUIT

A. Ansatz

Based on the previous protocol, we propose a variational Ansatz for loading functions beyond the conditions required by Th. 1. We consider \((z + s + 1)p(k)\) hyperparameters for each \(k\)-qubit block satisfying \(k > k_0\), where \(z\) and \(s\) are respectively the number of zeros and singular points in the function, \(p(k)\) is polynomial number in \(k\) denoting the number of hyperparameters allowed per singular point/zero for the \(k\)-th block, and \(k_0 = \max\{ k \mid z + s + 1 \geq 2^k \}\). Assuming \(p(k) \geq 1\), the minimum number of hyperparameters needed to capture the singular behavior of the function is \(z + s + 1\). Therefore, for each \(k\)-qubit block \(f^{(k-1)}(g, \theta)\) comprising more than \(z + s + 1\)
parameters, we cluster the angles which do not correspond to the position of zeros or singularities. In this form, this proposal establishes an intuitive correlation between the number of hyperparameters \( p(k) \) and the expressability of the circuit to capture the details of the target functions in the most relevant points. By using the decomposition of the multi-controlled rotations \([39, 56]\), the number of two-qubit gates which must be included in the variational circuit sums up to

\[
\#\text{TQGs} = \sum_{k=1}^{k_0} 2^{k-1} + (s+z) \sum_{k=k_0+1}^{n} p(k)(80k - 398)
\]

\[= 2^{k_0} - 1 + r_{k_0+1}(n, s + z),\]  

(28)

with \( m \) the number of zeros/singularities and \( r_{k_0+1}(n, s + z) \) growing polynomially with the system size \( n \). A remarkable advantage of this Ansatz is the training procedure, since Grover-Rudolph algorithm provides a suitable set of initial training angles which considerably enhances the convergence of the protocol with respect to a random initialization. This fact, together with a scaling in the number of hyperparameters \( p(k) \) substantially slower than the system size, allow us to avoid the training procedure to get stuck into both local minima and barren plateaus. In Fig. 8, we illustrate how to select the hyperparameters of the variational circuit for a general function containing multiple zeros and singular points.

### B. Training Process

We proceed to illustrate the training method for the variational circuit proposed in this article. This process consists of iterative steps in which a loss function that measures how far the outcome is from the desired state is recursively minimized to obtain the optimal parameters.

Consider a function \( f \) and \( n \) qubits. The desired quantum state is

\[|\Psi(f)\rangle_n = \sum_{l=0}^{2^n-1} f_l |l\rangle,\]  

(29)

where the \( f_l \) terms are discrete approximations to the objective function, as presented in Def. 1. Now, given a set of angles \( \theta \), our Ansatz \( U(\theta) \) returns the state

\[|\Psi(\theta)\rangle_n = U(\theta) |0\rangle,\]  

(30)

where \( |0\rangle \equiv |0\rangle^\otimes n \), and the components of the obtained state are products of sines and cosines.

Let us now introduce the mean squared error loss function, defined as

\[L(n, f, \theta) = \frac{1}{2^n} \sum_{l=0}^{2^n-1} (f_l - \Psi_l(\theta))^2\]

\[= \frac{1}{2^n} \sum_{l=0}^{2^n-1} \left( f_l^2 + \Psi_l(\theta)^2 - 2 f_l \Psi_l(\theta) \right),\]  

(31)

where \( \Psi_l(\theta) := \langle l | \Psi(\theta) \rangle_n \).

This process aims to find the optimal parameters \( \theta \) for which \( |\Psi(\theta)\rangle_n \) approximates \( |\Psi(f)\rangle_n \), which is equivalent to minimizing the loss function. Here, we use the gradient descent method \([51]\) to do so.

Then, given any angle \( \theta_{l}^{(k-1)} \), with \( k \in \{1, \ldots, n\} \) and \( l \in \{0, 2^{k-1} - 1\} \), its value gets updated after each training step in the following way:

\[\theta_{l}^{(k-1)} = \theta_{l}^{(k-1)} - \gamma \frac{\partial L(n, f, \theta)}{\partial \theta_{l}^{(k-1)}},\]  

(32)

where \( \gamma \) is the learning rate. Let us now compute the expression of the derivative:

\[
\frac{\partial L(n, f, \theta)}{\partial \theta_{l}^{(k-1)}} = \frac{1}{2^n} \sum_{l=0}^{2^n-1} 2 (\Psi_l(\theta) - f_l) \frac{\partial \Psi_l(\theta)}{\partial \theta_{l}^{(k-1)}}.\]  

(33)

Since \( \Psi_l(\theta) \) is a product of sines and cosines, its partial derivative with respect to a given angle is

\[
\frac{\partial \Psi_l(\theta)}{\partial \theta_{l}^{(k-1)}} = \begin{cases} 
- \frac{1}{2} \sin(\theta_l^{(k-1)/2}) \Psi_l(\theta) & \text{if } j 2^{n-k-1} \leq l < (j+1) 2^{n-k}, \\
\frac{1}{2} \cos(\theta_l^{(k-1)/2}) \Psi_l(\theta) & \text{if } (j+1) 2^{n-k} \leq l < (j+1)2^{n-k+1}, \\
0 & \text{otherwise.}
\end{cases}
\]  

(34)

Finally, different stopping criteria exist to put to an end the training process. A simple example is to fix a number of training steps. Another criteria, which is the one we consider for our numerical simulations, is to set a tolerance for the loss function. Therefore, when the difference between the cost function of two consecutive steps is less than the given tolerance, the process is considered satisfactory.

As a summary, in Algorithm 1 we have depicted the pseudocode corresponding to the training process.
π distribution between the values of 0 and initialization with Grover-Rudolph method angles. These randomly initialized parameters are generated by drawing samples from a uniform Scholes distribution. We compare the training results (numbers of step and final infidelity) of 40 randomly initialized parameters versus the performance, while most of random initializations get stuck around a fidelity value (vertical lines). Training parameters: fidelities are close to 1.

(ii) Results of training the variational ansatz for a different density functions.

Algorithm 1 Algorithm for the training process.

Require: n qubits, normalized function \( f \), objective state \( |\Psi(f)\rangle \),
circuit \( U(\theta) \) with hyperparameters, stopping criteria and learning rate \( \gamma \).

1: Initialize parameters \( \theta \)
2: while stopping criteria is not met do
3: \( |\Psi(\theta)\rangle \leftarrow U(\theta)|0\rangle \)
4: Compute loss function \( L(n, f, \theta) \)
5: \( \theta \leftarrow \theta - \gamma \frac{\partial L(n, f, \theta)}{\partial \theta} \)
6: end while

C. Examples

Let us now illustrate the behavior the variational circuit different density functions.

1. Sine Function

As a first example, we have tested the variational circuit with the normalized sine function \( \sin x \) in the domain \([0, \frac{\pi}{2}]\). This example contains zeros in \( x = 0 \) and \( x = \pi \) and clearly does not satisfy the conditions in Th. 4 or a possible generalization, since it is not even positive. In order to train the parametrized quantum circuit, we use the mean squared error loss function and the gradient descent as optimization method, with a learning rate \( \gamma = 1.5 \). In Fig. 3(ii), we illustrate the arising loaded states of our trained circuits for different values of \( p(k) = 1, 2, 3 \) and \( k \) (the last case means the introduction of \( k \) hyperparameters per zero of the function) and a system comprising \( n = 5 \) qubits, with \( k_0 = 2 \). For the studied cases, our trained Ansatz is able to load the sine function with a fidelity larger than 0.97. Additionally, in Fig. 9, we show the resulting fidelity of the different combinations of the number of qubits and \( p(k) \), for \( n \in \{5, \ldots, 10\} \) and \( k_0 = 3 \).

2. Black–Scholes Distribution

As a second example, let us now apply the proposed variational circuit to the Black–Scholes distribution, which is given by

\[
BS(x) = \begin{cases} 
K - e^{-\gamma} s & \text{if } -\log (Ks) \leq x < 0, \\
K - e^{\gamma} s & \text{if } 0 < x \leq \log (Ks), 
\end{cases}
\]

(35)

with \( s = Kc \), where \( K \) and \( c \) are the parameters. The zeros are found in \( x = \pm \log (Ks) \).

We benchmark the performance of our Ansatz for loading the Black-Scholes distribution, Eq.(35), into a 5-qubit system.

| \( p(k) \) | Fidelity | \#Angles | %Angles |
|---|---|---|---|
| 1 | 0.999303 | 33 | 0.806 |
| 2 | 0.99838 | 52 | 1.269 |
| 3 | 0.99890 | 70 | 1.709 |
| \( k \) | 0.99913 | 142 | 3.468 |
We choose the parameters of the distribution $K = 45$ and $c = 3$, and for the training parameters, we consider, $k_0 = 2$, a learning rate of $y = 1.5$ and a tolerance of $10^{-9}$. In Fig. 10 (i), we depict the results of the from the numerical simulations of the discretized state $|Ψ(\text{BS}(x))\rangle_5$ and the trained Ansatz $|Ψ(θ)\rangle_5$, for different values of $p(k)$ and $n = 5$ qubits. We can appreciate that in all cases the circuit is able to capture the main features of the target state, consequently all fidelities are close to 1. We highlight that for $p(k) = 3$ the resulting state and the objective one are almost equal, $ε = 5e−5$.

We also provide an analysis for larger systems. In Table III, we depict the numerical results for $|Ψ(\text{BS}(x))\rangle_{12}$ and $|Ψ(θ)\rangle_{12}$ in a system of 12-qubit system, with $k_0 = 2$, and different values of $p(k)$. We are able to obtain large fidelities with only a few percentage of the original angles, which implies that with only a small portion of the initial parameter space, the circuit can approximately simulate the target state efficiently.

Finally, in Fig 10 (ii) and Table IV we analyze the performance of initializing the value of the parameters with the angles provided by the Grover-Rudolph method for different sizes of the system. We present the training results of the Ansatz initialized with several random angles sets and with the angles provided by the Grover-Rudolph method. We analyze the performance in terms of infidelity and number of steps of the training. We can observe how this initialization reduces drastically the number of steps, as well as enables us to achieve the largest fidelities.

V. CONCLUSIONS

In this article, we have considered the problem of loading real valued functions into a quantum computer, which is a major bottleneck for solving partial derivatives equations [11–14], computing Monte-Carlo integrations [9, 10, 36] and quantum field theory [37, 38] and quantum machine learning [15, 17–20]. Firstly, inspired by the Grover-Rudolph algorithm [55] without ancillas, we have analytically proven that the complexity for implementing our method on a $n$-qubit system scales as $O(2^{(ε)})$, with $ε$ the infidelity with respect to the exact state and $k_0(ε)$ asymptotically independent of $n$. This reduction of two-qubit gates leads to a significant speedup, which allows us to implement quantum protocols involving data embeddings in large qubit systems. Additionally, we have generalized this method for functions containing a certain type of singularities, obtaining promising results for density functions with singularities that satisfy the expanded theorem conditions. Furthermore, we have proposed a variational Ansatz inspired in our previous protocol. We have observed that it can efficiently and accurately load functions with zeros and singularities. Our proposed Ansatz is tailored to the landscape of the function, providing an intuitive correlation between hyperparameters and expressability. Moreover, our previous protocol allows us to define a suitable initial training angle set, which considerably improves the training process, avoiding barren plateaus and local minima. As a future work, tensor networks could be used to prove the quasi-optimality in the minimal number of hyperparameters in the variational Ansatz.

ACKNOWLEDGMENTS

We thank J. J. García-Ripoll for the useful discussions regarding the quasi optimality of the variational Ansatz, T. Nguyen for the useful discussions regarding the Grover Rudolph algorithm, and M. Garcia-de-Andoin for the discussion on the noise impact. The authors acknowledge financial support from OpenSuperQ+100 (Grant No. 101113946) of the EU Flagship on Quantum Technologies, as well as from the EU FET-Open project EPIQUS (Grant No. 899368), also from Project Grant No. PID2021-125823N-100 595 and Spanish Ramón y Cajal Grant No. RYC-2020-030503-I funded by MCIN/AEI/10.13039/50110011033 and by “ERDF a way of making Europe” and “ERDF Invest in your Future,” Basque Government through Grant No. IT1470-22 and the IKUR Strategy under the collaboration agreement between Ikerbasque Foundation and BCAM on behalf of the Department of Education of the Basque Government, as well as from and UPV/EHU Ph.D. Grant No. PIF20/276.

APPENDIX A: GATES DECOMPOSITION AND COMPLEXITY

A.1 Uniformly-Controlled Rotation

Let us consider $n−1$ control qubits, a target $n$-th qubit, an axis of rotation $u$, and a vector of $2^{n−1}$ angles $θ$. Then, a uniformly controlled rotation $P_{n−1}(u, θ)$, depicted in Fig. 11, is a sequence of multi-controlled rotations $Λ_{n−1}(R_u(θ))$ compris-
ing the $2^{n-1}$ combinations of control bits.

\[
\begin{array}{c|c|c|c|c|c|c|c|c}
\hline
n & R_0(\theta) & R_1(\theta) & \cdots & R_{n-1}(\theta) \\
\hline
1 & \cdots & \cdots & \cdots & \cdots \\
\end{array}
\]

FIG. 11. Circuit corresponding to the uniformly controlled rotation gate $F_n^{-1}(u, \theta)$, with $n-1$ control qubits, a target at the $n$th qubit, an axis of rotation $u$, and a set of angles $\theta$.

Analytically this gate can be expressed as

\[
F_n^{-1}(u, \theta) = \prod_{i=0}^{2^{n-1}-1} (|i\rangle \langle i| \otimes R_u(\theta)) + (|1\rangle_{n-1} \langle 1| \otimes 1),
\]

and since $\sum_{i=0}^{2^{n-1}-1} |i\rangle \langle i| = 1_{\text{sys}}$ and $\langle i|j\rangle = \delta_{ij}$, we can simplify it to

\[
F_n^{-1}(u, \theta) = \sum_{i=0}^{2^{n-1}-1} |i\rangle \langle i| \otimes R_u(\theta).
\]

### A.2 Cost of Multi-Controlled Rotations

Let us now analyze the cost of implementing a multi-controlled rotation in terms of single and two-qubit gates.

**Theorem 1.** Let us consider an $n$-qubit system. Then, the multi-controlled rotation $\wedge_{n-1}(R_{j}(\theta))$ can be decomposed employing two one-qubit controlled rotations $\wedge_{1}(R_{j}(\pm\theta/2))$ and two $\wedge_{n-2}(X)$, as illustrated in Fig. 12.

**Proof.** First, let us see that $R_{j}(\theta) = XR_{j}(\theta/2)XR_{j}(\theta/2)$:

\[
XR_{j}(\theta/2)X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} \cos(\frac{\theta}{4}) & \sin(\frac{\theta}{4}) \\ -
\sin(\frac{\theta}{4}) & \cos(\frac{\theta}{4}) \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \cos(\frac{\theta}{4}) & \sin(\frac{\theta}{4}) \\ -
\sin(\frac{\theta}{4}) & \cos(\frac{\theta}{4}) \end{pmatrix} = R_{j}(\theta/2)
\]

Then,

\[
XR_{j}(\theta/2)XR_{j}(\theta/2) = R_{j}(\theta/2)R_{j}(\theta/2) = R_{j}(\theta).
\]

Now, let us check that the circuits are equivalent. On the left hand side, we have:

\[
\wedge_{n-1} \left( R_{j}(\theta) \right) = |1\rangle^{\otimes(n-1)} \langle 1|^{\otimes(n-1)} \otimes R_{j}(\theta)
\]

\[
+ (1^{\otimes(n-1)} - |1\rangle^{\otimes(n-1)} \langle 1|^{\otimes(n-1)}) \otimes 1.
\]

On the right-hand side:

\[
\wedge_{n-2}(X) \wedge_1 (R_{j}(\theta/2)) \wedge_{n-2}(X) \wedge_1 (R_{j}(\theta/2))
\]

\[
= \left[ |1\rangle^{\otimes(n-2)} \langle 1|^{\otimes(n-2)} \otimes 1 \otimes X
\]

\[
+ (1^{\otimes(n-1)} - |1\rangle^{\otimes(n-1)} \langle 1|^{\otimes(n-1)}) \otimes 1 \otimes X
\]

\[
\right].
\]

where we have ignored the identities on the left hand side of the equation for the sake of simplicity. If we expand the previous expression, we obtain the following:

\[
\left[ |1\rangle \langle 1| \otimes XR_{j}(\theta/2)
\right.

\[
+ |1\rangle \langle 1| \otimes XR_{j}(\theta/2)
\]

\[
+ |1\rangle \langle 1| \otimes XR_{j}(\theta/2)
\]

\[
+ |1\rangle \langle 1| \otimes XR_{j}(\theta/2)
\]

where we have defined the projector

\[
F(\alpha) \equiv \left[ 1^{\otimes(n-2)} \otimes (|1\rangle \langle 1| \otimes R_{j}(\theta/2) + |0\rangle \langle 0| \otimes 1)
\right.

\[
+ 1^{\otimes(n-2)} \otimes (|1\rangle \langle 1| \otimes R_{j}(\theta/2) + |0\rangle \langle 0| \otimes 1)
\]

\[
+ 1^{\otimes(n-2)} \otimes (|1\rangle \langle 1| \otimes R_{j}(\theta/2) + |0\rangle \langle 0| \otimes 1)
\]

\[
+ 1^{\otimes(n-2)} \otimes (|1\rangle \langle 1| \otimes R_{j}(\theta/2) + |0\rangle \langle 0| \otimes 1)
\]

\[
\left. + 1^{\otimes(n-2)} \otimes (|1\rangle \langle 1| \otimes R_{j}(\theta/2) + |0\rangle \langle 0| \otimes 1) \right] = 1.
\]

FIG. 12. Circuit showing the decomposition of $\wedge_{n-1}(R_{j}(\theta))$ employing two controlled rotations $\wedge_{1}(R_{j}(\pm\theta/2))$ and two $\wedge_{n-2}(X)$ gates.
It is straightforward to see that both terms \( |1\cdots1\rangle \langle 1\cdots1| \otimes XR_1(\alpha) \cdot F(\alpha) \) and \( |1\cdots10\rangle \langle 1\cdots10| \otimes X \cdot F(\alpha) \) are equal to 0, due to the projector \( F(\alpha) \). Then, we have

\[
\wedge_{n-2}(X) \wedge_1 (R_\alpha(-\theta/2)) \wedge_{n-2}(X) \wedge_1 (R_\alpha(\theta/2)) \\
= |1\cdots1\rangle \langle 1\cdots1| \otimes R_\alpha(\theta) + |1\cdots10\rangle \langle 1\cdots10| \otimes \mathbb{I} \quad \text{(48)}
+ F(-\theta/2)F(\theta/2), \quad \text{(49)}
\]

where we have used that \( XR_\alpha(-\theta/2)XR_\alpha(\theta/2) = R_\alpha(\theta) \) and \( XX = \mathbb{I} \). Let us now compute the last term:

\[
F(-\theta/2)F(\theta/2) \\
= \mathbb{I}^{\otimes(n-2)} \otimes |1\rangle \langle 1| \otimes R_\alpha(-\theta/2)R_\alpha(\theta/2) \quad \text{(50)}
- |1\cdots1\rangle \langle 1\cdots1| \otimes R_\alpha(-\theta/2)R_\alpha(\theta/2) \quad \text{(51)}
+ \mathbb{I}^{\otimes(n-2)} \otimes |1\rangle \langle 1| \otimes -|1\cdots10\rangle \langle 1\cdots10| \otimes \mathbb{I} \\
- |1\cdots10\rangle \langle 1\cdots10| \otimes \mathbb{I} + |1\cdots1\rangle \langle 1\cdots1| \otimes \mathbb{I} \\
- |1\cdots1\rangle \langle 1\cdots1| \otimes R_\alpha(-\theta/2)R_\alpha(\theta/2) \quad \text{(52)}
+ |1\cdots1\rangle \langle 1\cdots1| \otimes R_\alpha(-\theta/2)R_\alpha(\theta/2)
\]

where we have used that \( R_\alpha(-\theta/2)R_\alpha(\theta/2) = \mathbb{I} \) and

\[
\mathbb{I}^{\otimes(n-2)} \otimes |1\rangle \langle 1| \otimes \mathbb{I} + \mathbb{I}^{\otimes(n-2)} \otimes |0\rangle \langle 0| \otimes \mathbb{I} = \mathbb{I}^{\otimes(n-1)} \otimes \mathbb{I}. \quad \text{(53)}
\]

Finally,

\[
\wedge_{n-2}(X) \wedge_1 (R_\alpha(-\theta/2)) \wedge_{n-2}(X) \wedge_1 (R_\alpha(\theta/2)) \\
= |1\cdots1\rangle \langle 1\cdots1| \otimes R_\alpha(\theta) + |1\cdots10\rangle \langle 1\cdots10| \otimes \mathbb{I} \\
+ \left( \mathbb{I}^{\otimes(n-1)} - |1\cdots1\rangle \langle 1\cdots1| - |1\cdots10\rangle \langle 1\cdots10| \right) \otimes \mathbb{I} \\
= \wedge_{n-1}(R_\alpha(\theta)), \quad \text{(54)}
\]

as we wanted to prove.

\[\blacksquare\]

**Corollary 1.** The order of complexity of implementing the \( \wedge_{n-1}(R_\alpha(\theta)) \) gate is linear in \( n \), for \( n > 6 \). In fact, \( \wedge_{n-1}(R_\alpha(\theta)) \) can be decomposed with \( 80n - 398 \) two-qubit gates.

**Proof.** Following Th. 1, we know that the multi-controlled rotation \( \wedge_{n-1}(R_\alpha(\theta)) \) can be decomposed into two one-qubit controlled rotations, \( \wedge_1(R_\alpha(\theta/2)) \) and \( \wedge_1(R_\alpha(-\theta/2)) \), and two \( \wedge_{n-2}(X) \) gates. Therefore, we need to study the number of two-qubit gates necessary to implement the last two gates \( \wedge_{n-2}(X) \).

Corollary 7.4 from [54] states that, for \( n > 6 \), \( \wedge_{n-2}(X) \) can be realized with \( 8(n-5)\wedge_2(X) \). Then, we reduce the problem to obtain the cost of the Toffoli gate, \( \wedge_2(X) \).

Lemma 6.1 from the previous reference [54] states that any one-qubit unitary, \( \wedge_2(U) \) can be implemented with 5 two-qubit gates comprising two CNOTs, two \( \wedge_1(V) \), and one \( \wedge_1(V^\dagger) \), with \( V^2 = U \). In our case, we have \( U = X \). Then, we can choose \( V = \frac{1}{\sqrt{2}}(I + iX) \). We provide the exact decomposition of the Toffoli gate using the previous description in Fig. 13.

Finally, if \( n > 6 \), then the number of gates required for decomposing \( \wedge_{n-1}(R_\alpha(\theta)) \), considering only two-qubit operators, is

\[
\#\text{gates}(\wedge_{n-1}(R_\alpha(\theta))) = 2 + 2 \cdot (8(n - 5) \cdot 5) = 80n - 398. \quad \text{(55)}
\]

\[\blacksquare\]

**APPENDIX B: THEOREM PROOFS**

**B.1 Relation Between the Difference in the Angles and the Target Function**

**Theorem 2.** Let \( f : [0, 1] \rightarrow \mathbb{R} \) be a continuous function such that \( f \) is uniformly controlled rotation of \( k \) qubits. Then, the difference between two contiguous angles is bounded by the second derivative of the logarithm of \( f \) in the following way:

\[
|\theta^{(k-1)}_{l+1} - \theta^{(k-1)}_l| \leq \frac{\delta_k^2}{4} \max_{x \in [\delta_k, (k+1)\delta_k]} \left| \frac{\partial^2 \log f(x)}{\partial x^2} \right|, \quad \text{(56)}
\]

where \( \delta_k = \frac{1}{2^{k-1}}, l \in \{1, \ldots, 2^{k-1} - 1\} \).

**Proof.** The discrete expression of the angles for a block comprised of uniformly controlled rotation with \( k \) qubits, first introduced in [55], is given by

\[
\theta^{(k-1)}_l = \arccos \left( \frac{\int_{\delta_k}^{(l+1)\delta_k} f(x)dx}{\int_{\delta_k}^{(l+1)\delta_k} f(x)dx} \right), \quad \text{(57)}
\]

with \( \delta_k = \frac{1}{2^{k-1}} \) and \( l \in \{1, \ldots, 2^{k-1} - 1\} \). We can define the continuous extension of the previous function as

\[
\theta^{(k-1)}(y) := \arccos \left( \frac{\int_{\delta_k}^{y} f(x)dx}{\int_{\delta_k}^{y} f(x)dx} \right), \quad \text{(58)}
\]

where the original expression can be recovered replacing \( y = l\delta_k \).

Now, given a function \( g \) and a displacement \( \mu \), we define the numerical derivative of \( g \) with respect to \( \mu \) as

\[
\partial_y^{(\mu)} g(y) := \frac{g(y + \mu) - g(y)}{\mu}. \quad \text{(59)}
\]
Notice that for $\mu \to 0$, $\partial_x^\mu g(y) \to \partial_x g(y)$, the usual derivative. Then, for two consecutive angles, we have
\[
\left|\theta^{(k-1)}_{l+1} - \theta^{(k-1)}_l\right| = \left|\partial_y \left(\partial_y^{(k-1)}(y)\right)\right|_{y=y_0}\] (60).

Now, we require a connection between the difference in the angles and the partial derivative of Eq.(58). The Mean Value Theorem guarantees that the numerical derivative, which corresponds to the slope of the straight line connecting $\theta^{(k-1)}_l$ and $\theta^{(k-1)}_{l+1}$, is constrained by the maximum absolute value of the exact derivative in that interval. Then, we have
\[
\left|\theta^{(k-1)}_{l+1} - \theta^{(k-1)}_l\right| \leq \delta_k \max_{y' \in [l_0(l+1)\delta_k]} \left|\partial_y \theta^{(k-1)}(y')\right|_{y'=y'}\] (61).

Let us now develop the term of the exact derivative of the function $\theta^{(k-1)}(y)$:
\[
\partial_y \theta^{(k-1)}(y) = \frac{(f(y + \delta_k) - f(y)) \int_{y^+\delta_k}^y f(x)dx - (f(y + \delta_k/2) - f(y)) \int_{y^+\delta_k/2}^y f(x)dx}{\int_{y^+\delta_k}^y f(x)dx \int_{y^+\delta_k/2}^y f(x)dx}\] (62).

By introducing $(f(y + \delta_k/2) - f(y))$ in the first term and reorganizing the expression, we have:
\[
\partial_y \theta^{(k-1)}(y) = \frac{(f(y + \delta_k) - f(y + \delta_k/2)) \int_{y^+\delta_k}^y f(x)dx - (f(y + \delta_k/2) - f(y)) \int_{y^+\delta_k/2}^y f(x)dx}{\int_{y^+\delta_k}^y f(x)dx \int_{y^+\delta_k/2}^y f(x)dx}\]
\[
\leq \frac{\delta_k}{4} \max_{y' \in [l_0,l_1+1\delta_k]} \left|\partial_y \log f(y')\right|_{y'=y'}\] (63).

Finally, by plugging this result in Eq. (61), we obtain
\[
\left|\theta^{(k-1)}_{l+1} - \theta^{(k-1)}_l\right| \leq \frac{\delta_k}{4} \max_{y' \in [l_0,l_1+1\delta_k]} \left|\partial_y \log f(y')\right|_{y'=y'}\] (70).

**Corollary 2.** Let $f$ be a continuous function such that $f : [0, 1] \to \mathbb{R}^n$ and consider a block of uniformly controlled rotations of $k$ qubits. Then, if $\exists \eta > 0$ such that $|\partial_y f(y)| \leq \eta \forall y \in [0, 1]$,
\[
\left|\theta^{(k-1)}_{l+1} - \theta^{(k-1)}_l\right| \leq \frac{\delta_k^2}{4}\eta\] (72).
Corollary 3. For a non-standardized function \( f : [\xi_{\min}, \xi_{\max}] \to \mathbb{R}^+, \) the result in Corollary 2 holds with the modification in the bound

\[
|\theta_{\ell+1}^{(k)} - \theta_{\ell}^{(k)}| \leq \frac{\delta_k \cdot \eta}{4L^2},
\]

(73)

where \( L = \xi_{\max} - \xi_{\min} \) and \( \delta_k = \frac{L}{2^{k+1}}. \)

Proof. The change of variables that map the \( x' \in [0, 1] \) with \( x \in [\xi_{\max} - \xi_{\min}] \) is given by

\[
x = \xi_{\min} + x'L.
\]

(74)

Now, by using the chain rule, we obtain:

\[
\frac{\partial f}{\partial y'} = \frac{\partial f}{\partial y} \frac{\partial y'}{\partial y'} = \frac{\partial f}{\partial y} L \implies |\frac{\partial^2_f \log(f(y))}{\partial y^2}| \leq \frac{\eta}{L^2}.
\]

(75)

\[\blacksquare\]

Corollary 4. Let \( \eta \) be such that \( |\delta^L \log(f(y))| \leq \eta \forall y \in [0, 1]. \) Then, the difference between any two angles is

\[
|\theta_{l}^{(k)} - \theta_{l'}^{(k)}| \leq \frac{\delta_k}{4} \eta,
\]

(76)

\( \forall l, l' \in \{1, \ldots, 2^{k-1}\}. \)

Proof. The worst scenario is when \( l = 1 \) and \( l' = 2^{k-1}. \) In this case, by using the triangular inequality, we have

\[
|\theta_{1}^{(k-1)} - \theta_{2^{k-1}}^{(k-1)}| \leq |\theta_{1}^{(k-1)} - \theta_{2}^{(k-1)}| + \cdots + |\theta_{2^{k-1}-1}^{(k-1)} - \theta_{2^{k-1}}^{(k-1)}| \leq \frac{\delta_k}{4} \eta,
\]

(77)

since there are less than \( \delta_k \) elements in the sum.

In this situation, we can define the representative angle \( \tilde{\theta}^{(k-1)} \) for the clustering process as the one corresponding to the middle part of the interval, satisfying

\[
|\tilde{\theta}^{(k-1)} - \theta_{l}^{(k-1)}| \leq \frac{\delta_k}{8} \eta \quad \forall l \in \{0, 1, \ldots, 2^{k-1} - 1\}.
\]

(78)

B.2 Error Bound of the Algorithm

1. Relation Between the Difference in the Angles and the Fidelity

Let us consider a system with \( n \) qubits, thus the unitary gate to prepare the quantum state representing the target density function can be written in the terms of the blocks as

\[
U_n = U_{n-1}(\theta^{(n-1)}) \cdots U_0(\theta^{(0)}),
\]

(79)

where we define

\[
U_{k-1}(\theta^{(k-1)}) \coloneqq F_k^{-1}(y, \theta^{(k-1)}) \otimes I^\otimes(n-k).
\]

(80)

Let \( \tilde{U}_n \) denote the operation \( U_n \) given a representative with an error \( \eta_k \) between the angles for each block, i.e. \( |\theta_l^{(k-1)} - \tilde{\theta}_l^{(k-1)}| \leq \eta_k \) for \( l = 0, \ldots, 2^{k-1} - 1 \) and \( k = 1, \ldots, n. \)

Theorem 3. Consider a system of \( n \) qubits and an error \( \eta_k \) between any angle of the \( k \)-th block and its representative such that \( \eta_k \leq \pi, \) with \( k = 1, \ldots, n. \) Then, the fidelity between the final states with and without clustering, \( F = |\langle 0^{\otimes n} U_n \tilde{U}_n 0^{\otimes n} \rangle|^2, \) satisfies

\[
F \geq \prod_{k=1}^{n} \cos^2(\eta_k/2).
\]

(81)

Proof. We use induction to prove the inequality. Hence, we start with the elemental case of \( n = 1 \) and later proceed assuming it is satisfied for \( n - 1 \) and check if it holds for \( n. \)

- \( n = 1: \)

\[
\langle 0 | U_1^{\dagger} U_1 | 0 \rangle = \langle 0 | R^{\dagger}_y(\theta(0)) R_0(\tilde{\theta}(0)) | 0 \rangle = \langle 0 | R^{\dagger}_y(\theta(0) - \tilde{\theta}(0)) | 0 \rangle = \cos \frac{(\theta(0) - \tilde{\theta}(0))}{2}
\]

(82)

Notice that the angles, given by

\[
\delta^{(n-1)} = 2 \arccos \left( \frac{\int_{\xi_{\min}}^{\xi_{\max} + \delta^{(n)}} f(x)dx}{\int_{\xi_{\min} + \delta^{(n)}}^{\xi_{\max} + \delta^{(n)}} f(x)dx} \right),
\]

(83)

take values between \( 2 \arccos(1) = 0 \) and \( 2 \arccos(0) = \pi. \) Then, since the cosine is a decreasing function in the interval \([0, \pi/2] \) and \( \eta \leq \pi, \) we have that

\[
|\theta(0) - \tilde{\theta}(0)| \leq \eta_1 \implies \cos \frac{(\theta(0) - \tilde{\theta}(0))}{2} \geq \cos \frac{(\eta_1/2)}{2}
\]

(84)

- \( n > 1: \) Assume the condition holds for \( n - 1 \). We aim to recover the expression of \( |\langle 0^{\otimes n-1} U_{n-1}^{\dagger} U_{n-1} 0^{\otimes n-1} \rangle|^2 \) so we can use the induction hypothesis. Therefore, as we did for \( n = 1, \) we sandwich the operator

\[
\left( F_n^{-1}(y, \theta^{(n-1)}) \right)^\dagger F_n^{-1}(y, \tilde{\theta}^{(n-1)})
\]

(85)

with the state \( |0\rangle \) at each side. From now on during this proof, to simplify the notation, we denote the gate \( F_k^{-1}(y, \tilde{\theta}^{(k-1)}) \) as \( U_{k-1}(\tilde{\theta}^{(k-1)}) \). Then, for \( n: \)

\[
\langle 0^{\otimes n} U_n^{\dagger} \tilde{U}_n 0^{\otimes n} \rangle
\]

\[
= \langle 0^{\otimes n} U_n^{\dagger} (\theta^{(0)}) \cdots U_{n-1}^{\dagger} (\theta^{(n-1)}) U_{n-2} (\tilde{\theta}^{(n-2)}) \cdots U_0 (\tilde{\theta}^{(0)}) | 0 \rangle^{\otimes n}
\]

\[
= \langle 0^{\otimes n-1} \otimes (U_0(\tilde{\theta}^{(0)}) \otimes I^\otimes(n-2)) \rangle \cdot \left( U_{n-1}^{\dagger} (\tilde{\theta}^{(n-1)}) \right) \left( U_{n-2} (\tilde{\theta}^{(n-2)}) \otimes I^\otimes(2) \right) \cdots \left( U_0 (\tilde{\theta}^{(0)}) \otimes I^\otimes(2) \right) | 0 \rangle^{\otimes n-1} \otimes | 0 \rangle
\]

\[
= \langle 0^{\otimes n-1} (U_0(\tilde{\theta}^{(0)}) \otimes I^\otimes(n-2)) \rangle \cdots \left( U_{n-1}^{\dagger} (\tilde{\theta}^{(n-1)}) \right) \left( U_{n-2} (\tilde{\theta}^{(n-2)}) \otimes I^\otimes(2) \right) \cdots \left( U_0 (\tilde{\theta}^{(0)}) \otimes I^\otimes(2) \right) | 0 \rangle^{\otimes n-1} \otimes | 0 \rangle
\]

(86)

Notice that now we can substitute the rotation terms in the following way:

\[
\langle 0 | R^{\dagger}_y (\tilde{\theta}^{(n-1)} - \theta^{(n-1)}) | 0 \rangle
\]
\[
\theta(k) \equiv \frac{\varnothing}{k} \eta. \quad \text{Therefore, we obtain the following condition for the value of } \eta \text{ for Th. 3 to be satisfied:}
\]
\[
\frac{\delta k}{8} \eta \leq \pi \implies \eta \leq 2^{k-1} \cdot 8\pi.
\]

In particular, if this holds for the smallest \( k \), with value 1, it will be satisfied for the rest of the blocks. Then, we can write the previous condition simply as \( \eta \leq 8\pi \).

2. Expression of \( k_0 \)

Assume only clusterization for \( k > k_0 \):

\[
F \geq \prod_{k=k_0+1}^{n} \cos^2 \left( \frac{\eta}{4 \cdot 2^{k}} \right) \geq \prod_{k=k_0+1}^{n} e^{-2 \cdot 2^{k}} = e^{-2^{k_0} \sum_{k=k_0}^{n} 4^k},
\]

since \( \cos(x) \geq e^{-x^2} \) for \( x \lesssim \pi/2 \). Also, the sum is a geometric series with value

\[
\sum_{k=k_0+1}^{n} 4^{-k} = \frac{1}{3} \left( 4^{-k_0} - 4^{-n} \right).
\]

Then,

\[
F \geq e^{-\frac{1}{4} \sum_{k=k_0}^{n} 4^{k_0} \cdot 4^{-n}} = F_{k_0}.
\]

Let us finally find an expression for \( k_0 \) given minimum fidelity \( F_{k_0} = 1 - \epsilon \):

\[
-2 \frac{\eta^2}{16} \left( 4^{-k_0} - 4^{-n} \right) = \log F_{k_0}
\]

\[
4^{-k_0} - 4^{-n} = -\frac{3}{2 \eta^2} \log F_{k_0}
\]

\[
-4^{-k_0} = 4^{-n} = \frac{3}{2 \eta^2} \log F_{k_0}
\]

\[
k_0 \log 4 = \log \left( 4^{-n} - \frac{3}{2 \eta^2} \log F_{k_0} \right)
\]

\[
k_0 = -\frac{\log \left( 4^{-n} - \frac{3}{2 \eta^2} \log F_{k_0} \right)}{\log 4}.
\]

Note that in the limit for both \( F \to 1 \) and \( \eta \to \infty \), we have that

\[
k_0 \to -\log(4^{-n})/\log(4) = n,
\]

as expected.

Also, since \( k_0 \) must be an integer, we take the ceiling of its value. In addition, we consider the minimum of this parameter to be 2, so the condition of \( \eta \leq 8\pi \) is satisfied. Hence, the final expression for \( k_0 \) is

\[
k_0 = \max \left\{ \left[ -\frac{1}{2} \log_2(4^{-n} - \frac{96}{\eta^2} \log(1 - \epsilon)) \right], 2 \right\}.
\]
where we have defined $\epsilon \equiv 1 - F_0$.

Once we have a fixed number of uniformly controlled rotation blocks, $k_0$, for which there is no clustering, we count the number of necessary gates. The gate $F_k^{-1}(\tilde{g}, \theta)$ can be implemented with $2^{k-1}$ CNOTs and $2^{k-1}$ single-qubit gates [56]. Thus, the total number of gates is:

\[
\text{#CNOTs} = \sum_{k=1}^{k_0} 2^{k-1} = 2^{k_0} - 1, \tag{104}
\]

\[
\text{#SQGs} = \sum_{k=1}^{k_0} 2^{k-1} + \sum_{k=k_0+1}^{n} 1 = 2^{k_0} - 1 + n - k_0. \tag{105}
\]

**B.3 Final Protocol**

**Definition 1.** Let $f: [x_{\min}, x_{\max}] \to \mathbb{R}^+$ be a positive function in $L^2([x_{\min}, x_{\max}])$. We define the n-qubit normalized representative state of $f(x)$ as the n-qubit state $|f(x)\rangle_\eta = \sum_{j=0}^{2^n-1} f(x_{\min} + j\delta_n) |j\rangle$, with $\delta_n = \frac{x_{\max} - x_{\min}}{2^n}$ and $\sum_{j=0}^{2^n-1} f^2(x_{\min} + j\delta_n) = 1$.

**Theorem 4.** Let $f : [0, 1] \to \mathbb{R}^+$ be a positive integrable function in $L^2([0, 1])$ and $0 \leq \eta \leq 8 \pi$ a constant such that $\eta = \sup_{x \in [0, 1]} |\tilde{f}_x^2 \log f^2(x)|$. Then, it is possible to approximate the n-qubit representative state of $f(x)$, $|f(x)\rangle_{\eta}$, by a quantum state $|\Psi(f)\rangle_{\eta}$ such that the fidelity $\langle \Psi(f) | f(x) \rangle_{\eta}^2 \geq 1 - \epsilon$ with at most $2^{k_0(\epsilon)} - 1$ two-qubit gates, with

\[
k_0(\epsilon) = \max \left\{ \left\lfloor -\frac{1}{2} \log_2 \left(4^{-\eta} - \frac{96}{\eta^2} \log(1 - \epsilon)\right) \right\rfloor, 2 \right\}; \tag{106}
\]

and the circuit to perform it is given in Fig. 14.

**Proof.** Using Th. 2 and Cor. 4, we can define the representative angle for each block as in Eq. (78). Then, by applying Th. 3 with the development between Eqs. (94) and (101), it is clear that the fidelity will be larger than or equal to $1 - \epsilon$. Finally, Eq. (104) states that the number of two-qubit gates necessary to realize this protocol is $2^{k_0} - 1$, as we wanted to prove.  

[1] F. Arute, K. Arya, R. Babbush et al., Quantum supremacy using a programmable superconducting processor. Nature 574, 505 (2019).
[2] F. Zhong, Y. Deng, J. Qin et al., Strong quantum computational advantage using a superconducting quantum processor. ArXiv:2106.14734 (2021)
[3] F. Zhong, H. Wang, Y. Deng et al., Quantum computational advantage using photons. Science 370, 1460 (2020).
[4] A. W. Harrow, A. Hassidim and S. Lloyd, Quantum algorithm for solving linear systems of equations. Phys. Rev. Lett. vol. 15, no. 103, pp. 150502 (2009).
[5] A. M. Childs, R. Kothari and R. D. Somma, Quantum algorithm for systems of linear equations with exponentially improved dependence on precision. SIAM Journal on Computing 46, 1920–1950 (2017).
[6] N. Wiebe, D. Braun, and S. Lloyd, Quantum Algorithm for Data Fitting. Phys. Rev. Lett. 109, 050505 (2012).
[7] B. D. Clader, B. C. Jacobs, and C. R. Sprouse, Preconditioned Quantum Linear System Algorithm. Rev. Lett. 110, 250504 (2013).
[8] A. Scherer, B. Valiron, S.-C. Mau, S. Alexander, E. van den Berg, and T. E. Chapuran, Concrete resource analysis of the quantum linear-system algorithm used to compute the electromagnetic scattering cross section of a 2D target.. Quantum Information Processing 16 (2017).
[9] P. Rebentrost, B. Gupta, T. R. Bromley Quantum computational finance: Monte Carlo pricing of financial derivatives. Phys. Rev. A 98, 022321 (2018).
[10] N. Stamatopoulos, D. J. Egger, Y. Sun, C. Zoufal, R. Iten, N. Shen, and S. Woerner, Option Pricing using Quantum Computers. Quantum 4, 291 (2020).
[11] J. Gonzalez-Conde, A. Rodriguez-Rozas, E. Solano, and M. Sanz, Pricing Financial Derivatives with Exponential Quantum Speedup. arXiv:2101.04023 (2021).
[12] J. Liu, H. Ø. Kolden, H. K. Krovi, N. F. Loureiro, K. Trivisa, A. M. Childs, Efficient quantum algorithm for dissipative nonlinear differential equations. Proceedings of the National Academy of Sciences 118, e2026805118 (2021)
[13] B. Zanger, C. B. Mendl, M. Schulz, and M. Schreiber, Quantum Algorithms for Solving Ordinary Differential Equations via Classical Integration Methods. Quantum 5, 502 (2021).
[14] J. J. Garc ía-Ripoll, Quantum-inspired algorithms for multivariate analysis: from interpolation to partial differential equations. Int. J. Mod. Phys. C, 22, Quantum 5, 431 (2021).
[15] S. Lloyd, M. Mohseni and P. Rebentrost, Quantum algorithms for supervised and unsupervised machine learning. ArXiv:1307.0411v2 (2013).
[16] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, S. Lloyd, Quantum Machine Learning. Nature 549, 195-202 (2017).
[17] H. Huang, M. Broughton, M. Mohseni, R. Babbush, S. Boixo, H. Neven, and J. R. McClean, Power of data in quantum machine learning. Nature Communications volume 12, 2631 (2021)
[18] M. Schuld and N. Killoran, Quantum machine learning in feature Hilbert spaces. Phys. Rev. Lett. 122, 040504 (2019).
Key questions for the quantum machine learner to Supervised Learning with Quan-

Quantum Algorithms for Digitization of Scalar Fields for

Fast Black-Box Quantum State Preparation

Quantum circuts for general multi-qubit gates. Phys. Rev. Lett. 93, 130502 (2004).

Approximate amplitude encoding in shallow parameterized quantum circuits and its application to financial market indicator. arXiv:2103.13211 (2021).

Quantum generative adversarial networks for learning and loading random distributions. Npj Quantum Information 5, 1–9 (2019).

Supervised learning with quantum enhanced feature spaces. arXiv:2103.01958 (2018).

Implementtion of Conditional Phase Gates Based on Tunable ZZ Interactions. Npj Quantum Information 6, 625-644 (2021).

Variational Quantum Algorithms. Physical Review Letters 125 (2020).

Digital quantum simulation of spin systems in superconducting circuits. 2014. Phys. Rev. Lett. 112:200501

Quantum circuits for general multi-qubit gates. Phys. Rev. Lett. 93, 130502 (2004).

Elementary gates for quantum computation. Phys. Rev. A 71, 052335 (2019).

Creating superpositions that divide-and-conquer algorithm for quantum state preparation. Scientific Reports 11, 6329 (2021).

A divide-and-conquer algorithm for quantum state preparation. Scientific Reports 11, 6329 (2021).

Quantum speedup of Monte Carlo methods. Proc. Roy. Soc. Ser. A. 471, 2181, 20150301, (2015).

Approximate amplitude encoding in shallow parameterized quantum circuits and its application to financial market indicator. arXiv:2103.13211 (2021).

Quantum generative adversarial networks for learning and loading random distributions. Npj Quantum Information 5, 1–9 (2019).

Supervised learning with quantum enhanced feature spaces. arXiv:2103.01958 (2018).

Implemention of Conditional Phase Gates Based on Tunable ZZ Interactions. Npj Quantum Information 6, 625-644 (2021).

Variational Quantum Algorithms. Physical Review Letters 125 (2020).

Digital quantum simulation of spin systems in superconducting circuits. 2014. Phys. Rev. Lett. 112:200501

Quantum circuits for general multi-qubit gates. Phys. Rev. Lett. 93, 130502 (2004).

Elementary gates for quantum computation. Phys. Rev. A 71, 052335 (2019).

Creating superpositions that divide-and-conquer algorithm for quantum state preparation. Scientific Reports 11, 6329 (2021).

A divide-and-conquer algorithm for quantum state preparation. Scientific Reports 11, 6329 (2021).

Quantum speedup of Monte Carlo methods. Proc. Roy. Soc. Ser. A. 471, 2181, 20150301, (2015).

Approximate amplitude encoding in shallow parameterized quantum circuits and its application to financial market indicator. arXiv:2103.13211 (2021).

Quantum generative adversarial networks for learning and loading random distributions. Npj Quantum Information 5, 1–9 (2019).

Supervised learning with quantum enhanced feature spaces. arXiv:2103.01958 (2018).

Implemention of Conditional Phase Gates Based on Tunable ZZ Interactions. Npj Quantum Information 6, 625-644 (2021).

Variational Quantum Algorithms. Physical Review Letters 125 (2020).

Digital quantum simulation of spin systems in superconducting circuits. 2014. Phys. Rev. Lett. 112:200501

Quantum circuits for general multi-qubit gates. Phys. Rev. Lett. 93, 130502 (2004).

Elementary gates for quantum computation. Phys. Rev. A 71, 052335 (2019).

Creating superpositions that divide-and-conquer algorithm for quantum state preparation. Scientific Reports 11, 6329 (2021).

A divide-and-conquer algorithm for quantum state preparation. Scientific Reports 11, 6329 (2021).

Quantum speedup of Monte Carlo methods. Proc. Roy. Soc. Ser. A. 471, 2181, 20150301, (2015).

Approximate amplitude encoding in shallow parameterized quantum circuits and its application to financial market indicator. arXiv:2103.13211 (2021).

Quantum generative adversarial networks for learning and loading random distributions. Npj Quantum Information 5, 1–9 (2019).

Supervised learning with quantum enhanced feature spaces. arXiv:2103.01958 (2018).

Implemention of Conditional Phase Gates Based on Tunable ZZ Interactions. Npj Quantum Information 6, 625-644 (2021).

Variational Quantum Algorithms. Physical Review Letters 125 (2020).

Digital quantum simulation of spin systems in superconducting circuits. 2014. Phys. Rev. Lett. 112:200501

Quantum circuits for general multi-qubit gates. Phys. Rev. Lett. 93, 130502 (2004).

Elementary gates for quantum computation. Phys. Rev. A 71, 052335 (2019).

Creating superpositions that divide-and-conquer algorithm for quantum state preparation. Scientific Reports 11, 6329 (2021).

A divide-and-conquer algorithm for quantum state preparation. Scientific Reports 11, 6329 (2021).

Quantum speedup of Monte Carlo methods. Proc. Roy. Soc. Ser. A. 471, 2181, 20150301, (2015).

Approximate amplitude encoding in shallow parameterized quantum circuits and its application to financial market indicator. arXiv:2103.13211 (2021).

Quantum generative adversarial networks for learning and loading random distributions. Npj Quantum Information 5, 1–9 (2019).

Supervised learning with quantum enhanced feature spaces. arXiv:2103.01958 (2018).

Implemention of Conditional Phase Gates Based on Tunable ZZ Interactions. Npj Quantum Information 6, 625-644 (2021).

Variational Quantum Algorithms. Physical Review Letters 125 (2020).

Digital quantum simulation of spin systems in superconducting circuits. 2014. Phys. Rev. Lett. 112:200501

Quantum circuits for general multi-qubit gates. Phys. Rev. Lett. 93, 130502 (2004).

Elementary gates for quantum computation. Phys. Rev. A 71, 052335 (2019).

Creating superpositions that divide-and-conquer algorithm for quantum state preparation. Scientific Reports 11, 6329 (2021).

A divide-and-conquer algorithm for quantum state preparation. Scientific Reports 11, 6329 (2021).