AIMS Mathematics, 11(6): 16366–16394.
DOI:10.3934/math.2026672
Received: 30 December 2025
Revised: 10 May 2026
Accepted: 22 May 2026
Published: 08 June 2026

https://www.aimspress.com/journal/Math

Research article

A rigorous and self-contained proof of the Grover-Rudolph state preparation algorithm

Antonio Falcó[1,*], **Daniela Falcó-Pomares**[2] **and Hermann G. Matthies**[3]

[1] Departamento de Matemáticas, Física y Ciencias Tecnológicas, Universidad Cardenal Herrera-CEU, CEU Universities San Bartolomé 55, 46115 Alfara del Patriarca (Valencia), Spain
[2] Grupo de Investigación Bisite *Universidad de Salamanca* Calle Espejo s/n, 37007 Salamanca, Spain
[3] Institute of Scientific Computing, Technische Universität Braunschweig, Universitätsplatz 2, 38106 Braunschweig, Germany

* **Correspondence:** Email: afalco@uchceu.es.

Abstract: We give a rigorous and self-contained analysis of the Grover-Rudolph quantum state-preparation algorithm, which encodes a probability distribution $\{p_k\}$ as an n-qubit amplitude state $\sum_k \sqrt{p_k}\,|k\rangle$ via a hierarchy of controlled R_y rotations determined by a dyadic refinement of the target. We formalize the dyadic probability tree, derive the trigonometric factorization of conditional masses, and prove by induction that the circuit prepares exactly the desired measurement law. We further prove that perturbing each rotation angle by at most η changes the output distribution by at most $\min(1, n\eta)$ in total variation, and combine this with a Hoeffding concentration bound to obtain an explicit design rule: $b \geq \log_2(2n\pi/\varepsilon)$ bits and $S \geq 2^{n+1}\log(2/\delta)/\varepsilon^2$ shots suffice to achieve accuracy ε with confidence $1-\delta$. As a circuit-theoretic complement, we provide an ancilla-free transpilation of each stage into $\{\mathrm{R}_y(\cdot), X, \mathrm{CNOT}\}$ via Gray-code ladders and a Walsh-Hadamard angle transform.

Keywords: quantum state preparation; amplitude encoding; Grover-Rudolph algorithm; dyadic partitions; uniformly controlled rotations; Gray code; ancilla-free transpilation; controlled rotations
Mathematics Subject Classification: 81P65, 81P68

1. Introduction

Preparing quantum states whose amplitudes encode classical data is a basic primitive in quantum algorithm design. In particular, *amplitude encoding* of probability distributions underlies quantum Monte Carlo-type routines and amplitude-estimation methods, where expectation values or integrals are converted into success probabilities and then estimated quadratically faster than classically in idealized settings (see, e.g., the amplitude-estimation framework of Brassard-Høyer-Mosca-Tapp [2]).

A particularly influential and widely reused state-preparation procedure is due to Grover and Rudolph [4]. Given a probability distribution over 2^n outcomes (or, more generally, a nonnegative function on $[0, 1]$ sampled on a dyadic grid), their construction prepares an n-qubit state $\sum_{z\in\mathbb{Z}_2^n} \sqrt{p_z}\,|z\rangle$ by recursively refining a dyadic partition and applying, at each level, a family of controlled one-qubit rotations whose angles are defined from conditional masses on a binary tree. The algorithm is conceptually elegant and has become a standard reference point for state preparation from classical distributions.

Despite its popularity, the Grover-Rudolph procedure is often presented at a high level, with correctness justified informally or with implicit assumptions about the dyadic refinement and the controlled-gate semantics. The first objective of this paper is therefore *mathematical*: we provide a rigorous and self-contained treatment of the dyadic probability tree, the induced angle map, and the resulting amplitude identities, culminating in a fully explicit proof that the circuit prepares the desired target state. In particular, we isolate the key telescoping/trigonometric factorizations and prove them by transparent induction, so that the overall correctness argument can be read independently of external compilation folklore.

Grover and Rudolph's original note [4] introduced the recursive dyadic construction that we formalize here. Beyond this specific algorithm, state-preparation and multiplexing problems have been extensively studied in the circuit-synthesis literature. A prominent line of work concerns *uniformly controlled one-qubit gates* (also called *multiplexed* or *uniformly controlled rotations*), for which Gray-code techniques yield systematic decompositions into elementary gates without ancillas. In particular, the Gray-code ladder idea appears as a standard tool for implementing uniformly controlled rotations and related multiqubit structures (see, e.g., the uniformly controlled gate framework in Bergholm et al. [1] and the companion discussion of uniformly controlled one-qubit gates in [9]).

Since the original Grover-Rudolph proposal, the state-preparation problem has continued to attract significant attention, with contributions addressing improved gate complexity for general and structured states. Zhang, Li, and Yuan [11] showed that any n-qubit state can be prepared with $\Theta(n)$-depth circuits at the cost of exponentially many ancilla qubits, and achieved $\Theta(\log(nd))$-depth for sparse states with d nonzero amplitudes. Marin-Sanchez, Gonzalez-Conde, and Sanz [8] developed quantum circuits that efficiently encode smooth functions through a dyadic-tree approach closely related to the Grover-Rudolph construction, making the connection between classical function approximation and amplitude encoding more explicit. Gonzalez-Conde, Watts, Rodriguez-Grasa, and Sanz [3] extended this line by proposing efficient amplitude encoding for polynomial functions via block-encoding techniques. Mao, Tian, and Sun [7] established a nearly optimal circuit size of $O(nd/\log(n+m)+n)$ for sparse state preparation using m ancillary qubits, improving upon prior work by a factor of $\log d$. More recently, Luo, Li, and Li [6] established an explicit depth-ancilla trade-off for sparse state preparation, achieving circuit depth $O(\frac{nd\log m}{m\log(m/n)} + \log nd)$ with $m \geq 6n$ ancilla qubits. Ramacciotti, Lefterovici, and Rotundo [10] analyzed the Grover-Rudolph algorithm directly in the sparse regime, proving that the gate complexity is linear in the number of nonzero amplitudes and linear in the number of qubits after a simple modification.

The present paper differs from these works in three complementary ways. First, our primary objective is *mathematical rigor*: we provide a fully self-contained correctness proof of the original Grover-Rudolph construction in the general dense, amplitude-encoded setting, making explicit the dyadic refinement structure, the angle assignment, and the inductive amplitude argument that is often

left implicit in the literature. Second, we prove a quantitative stability estimate for imperfectly implemented Grover-Rudolph angles, showing how deterministic angle errors propagate to the final output distribution in total variation distance. Third, we integrate the circuit-synthesis perspective into the same rigorous framework: each Grover-Rudolph stage is identified as a uniformly controlled R_y rotation, and an explicit ancilla-free Gray-code transpilation is derived through a Walsh-Hadamard angle transform. Although Gray-code decompositions are standard in quantum compilation, our contribution is to connect them directly with the dyadic Grover-Rudolph hierarchy and to provide complete correctness proofs in this specific setting.

The main contributions can be summarized as follows:

- A rigorous dyadic-tree formalization of the Grover-Rudolph angle construction, including explicit refinement identities and a clean inductive correctness proof of the prepared amplitudes.
- A quantitative stability theorem for imperfectly implemented Grover-Rudolph angles, showing that if each angle is perturbed by at most η, then the resulting output distribution changes by at most $\min(1, n\eta)$ in total variation distance. This separates deterministic angle-synthesis errors from finite-shot statistical fluctuations.
- A stage-wise circuit formulation showing that each Grover-Rudolph stage is a uniformly controlled one-qubit rotation, together with an explicit ancilla-free transpilation into $\{R_y(\cdot), X, \mathrm{CNOT}\}$ using a Gray-code ladder and a Walsh-Hadamard angle transform, with full correctness proofs and implementable pseudo-code.
- A systematic sensitivity study of the total-variation error as a function of angle-quantization precision ($b \in \{8, 16, 32\}$) and number of measurement shots ($S \in \{256, 1024, 4096\}$) for $n \in \{2, 3, 4\}$ qubits, showing that shot noise dominates at practically relevant shot counts and that $b = 8$-bit angle precision already reaches the shot-noise floor in the tested instances. These empirical findings are underpinned by a combined theoretical bound (Corollary 5.1) that unifies the stability theorem with the Hoeffding inequality and yields an explicit design rule: $b \geq \log_2(2n\pi/\varepsilon)$ bits and $S \geq 2^{n+1}\log(2/\delta)/\varepsilon^2$ shots suffice to guarantee $\mathrm{TV}(p, \hat{p}) \leq \varepsilon$ with confidence $1 - \delta$.

The paper is organized as follows: Section 2 fixes notation and collects the required preliminaries, including the dyadic partition formalism for probability masses, the embedding conventions for one-qubit gates, and the selective pattern-controlled gate construction used throughout; it also states the Grover-Rudolph state-preparation circuit and the main theorem. Section 3 contains the rigorous and self-contained correctness proof, centered on the trigonometric factorization of dyadic masses and the inductive amplitude formula that propagates across Grover-Rudolph stages. Section 4 proves the stability estimate for imperfect rotation angles and derives the finite-precision consequence for angle synthesis, thereby separating deterministic synthesis error from finite-shot sampling error. In Section 5, we provide a worked example together with an idealized noise-free circuit-level verification of the construction, as well as a systematic sensitivity analysis of the total-variation error as a function of angle-quantization precision and number of measurement shots. Finally, Section 6 develops the active-register viewpoint and gives an explicit ancilla-free transpilation of each Grover-Rudolph stage into the elementary gate dictionary $\{R_y(\cdot), X, \mathrm{CNOT}(\cdot \to \cdot)\}$ via Gray-code ladder decompositions, including implementable pseudo-code and circuit diagrams.

2. Preliminary results and statement of the main theorem

Throughout the paper, we write $\mathbb{Z}_2 := \{0, 1\}$ and, for $n \geq 1$,

$$\mathbb{Z}_{2^n} := \{0, 1, \ldots, 2^n - 1\}, \qquad N := 2^n.$$

For a bitstring $\mathbf{z} = z_1 \cdots z_m \in \mathbb{Z}_2^m$, we use the shorthand $|\mathbf{z}\rangle := |z_1\rangle \otimes \cdots \otimes |z_m\rangle$ for the corresponding computational basis vector.

2.1. Binary representations and dyadic partitions

For $k \in \mathbb{Z}_{2^n}$, let $b_n(k) = z_1 z_2 \cdots z_n \in \mathbb{Z}_2^n$ denote the (length-n) binary representation of k, i.e.,

$$k = b_n^{-1}(z_1 z_2 \cdots z_n) = z_1 2^0 + z_2 2^1 + \cdots + z_n 2^{n-1}.$$

For each level $1 \leq \ell \leq n$ and $0 \leq k \leq 2^\ell$, define the dyadic grid points

$$z^{(\ell)}_{b_\ell(k)} := \frac{k}{2^\ell}, \quad \text{with the convention } z^{(\ell)}_{b_\ell(2^\ell)} := 1.$$

We then define the dyadic intervals

$$I^{(\ell)}_{b_\ell(k)} := \Big[z^{(\ell)}_{b_\ell(k)},\, z^{(\ell)}_{b_\ell(k+1)}\Big] = \Big[\frac{k}{2^\ell},\, \frac{k+1}{2^\ell}\Big], \qquad k \in \mathbb{Z}_{2^\ell}.$$

At the finest level $\ell = n$, we write

$$I_k := I^{(n)}_{b_n(k)} = \Big[\frac{k}{2^n}, \frac{k+1}{2^n}\Big], \qquad k \in \mathbb{Z}_{2^n}.$$

2.2. Quantum registers and one-qubit embeddings

We consider the n-qubit Hilbert space

$$\mathcal{H}_n := (\mathbb{C}^2)^{\otimes n} \cong \mathbb{C}^{2^n},$$

endowed with the computational basis $\{\, |b_n(k)\rangle : k \in \mathbb{Z}_{2^n} \,\}$. Given a one-qubit unitary $U \in \mathrm{U}(2)$ and an index $1 \leq j \leq n$, we define the standard embedding (acting on the j-th qubit)

$$w^{(n)}_j(U) := \mathsf{I}_2^{\otimes(j-1)} \otimes U \otimes \mathsf{I}_2^{\otimes(n-j)} \in \mathrm{U}(2^n).$$

2.3. Selective and controlled one-qubit operations

Fix $1 \leq \ell \leq n$ and $U \in \mathrm{U}(2)$. For a pattern $\mathbf{z} = z_1 \cdots z_{\ell-1} z_{\ell+1} \cdots z_n \in \mathbb{Z}_2^{n-1}$, we define the selective (rank-one supported) operator

$$\mathsf{C}^{(\ell)}_{\mathbf{z}}(U) := |z_1 \cdots z_{\ell-1}\rangle\langle z_1 \cdots z_{\ell-1}| \otimes U \otimes |z_{\ell+1} \cdots z_n\rangle\langle z_{\ell+1} \cdots z_n|.$$

This operator acts as $w^{(n)}_\ell(U)$ on basis states whose $(n-1)$ remaining bits match $\mathbf{z}$, and annihilates all other basis states. From $\mathsf{C}^{(\ell)}_{\mathbf{z}}(U)$, we build the corresponding controlled unitary (identity outside the selected branch)

$$\mathsf{CC}^{(\ell)}_{\mathbf{z}}(U) := \mathsf{C}^{(\ell)}_{\mathbf{z}}(U) + \sum_{\mathbf{z}' \in \mathbb{Z}_2^{n-1} \setminus \{\mathbf{z}\}} \mathsf{C}^{(\ell)}_{\mathbf{z}'}(\mathsf{I}_2) \;\in\; \mathrm{U}(2^n).$$

In particular, $\mathsf{CC}^{(\ell)}_{\mathbf{z}}(U)$ applies U to qubit ℓ if and only if the other qubits match the control pattern $\mathbf{z}$, and otherwise acts as the identity.

2.4. Rotation primitives and amplitude encoding objective

We will use the real rotation gate $\mathsf{R}(\alpha) \in \mathrm{U}(2)$ defined by

$$\mathsf{R}(\alpha) := \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}, \qquad \alpha \in \mathbb{R}.$$

Given a nonnegative density $\varrho : [0,1] \to [0,\infty)$ with $\int_0^1 \varrho(x)\,dx = 1$, we define the target probabilities

$$p_k := \int_{I_k} \varrho(x)\,dx, \qquad k \in \mathbb{Z}_{2^n}.$$

The state-preparation (amplitude-encoding) objective is to construct a circuit $U \in \mathrm{U}(2^n)$ such that

$$U|0\rangle^{\otimes n} = \sum_{k\in\mathbb{Z}_{2^n}} \sqrt{p_k}\,|b_n(k)\rangle,$$

up to a global phase.

2.5. Elementary identities for dyadic refinement

The binary representation b_m and the dyadic grid points and intervals $z^{(\ell)}_{b_\ell(k)}$, $I^{(\ell)}_{b_\ell(k)}$ are as defined in Section 2.1; here, we extend the index m to an arbitrary positive integer. The following two lemmas collect the shift and refinement identities used throughout the proof of the main theorem.

Lemma 2.1 (Binary shift identities)**.** *Let* $1 \le \ell \le n-1$.

(i) *For* $k \in \mathbb{Z}_{2^\ell}$*:*

$$b_{\ell+1}(2k) = 0\,b_\ell(k), \qquad b_{\ell+1}(2k+1) = 1\,b_\ell(k).$$

(ii) *For* $k \in \mathbb{Z}_{2^\ell}$ *where* $k \le 2^\ell - 2$ *:*

$$b_{\ell+1}(2k+2) = 0\,b_\ell(k+1).$$

Proof. Write $k = \sum_{j=1}^{\ell} z_j 2^{j-1}$, i.e., $b_\ell(k) = z_1 \cdots z_\ell$. Then $2k = \sum_{j=1}^{\ell} z_j 2^j$, hence the length-$(\ell+1)$ binary expansion of $2k$ is $0z_1 \cdots z_\ell = 0\,b_\ell(k)$, proving the first identity in (i). Since $2k+1 = 1 + \sum_{j=1}^{\ell} z_j 2^j$, the leading bit is 1, and the remaining bits reproduce $b_\ell(k)$, giving $b_{\ell+1}(2k+1) = 1\,b_\ell(k)$. For (ii), the hypothesis $k \le 2^\ell - 2$ ensures $k+1 \le 2^\ell - 1$, so $b_\ell(k+1)$ is well-defined; the identity $b_{\ell+1}(2k+2) = 0 b_\ell(k+1)$ then follows by applying the first identity in (i) with k replaced by $k+1$. □

Corollary 2.1 (Dyadic refinement)**.** *Let* $1 \le \ell \le n-1$ *and* $k \in \mathbb{Z}_{2^\ell}$*. Then*

$$I^{(\ell)}_{b_\ell(k)} = I^{(\ell+1)}_{b_{\ell+1}(2k)} \cup I^{(\ell+1)}_{b_{\ell+1}(2k+1)} = I^{(\ell+1)}_{0b_\ell(k)} \cup I^{(\ell+1)}_{1b_\ell(k)},$$

and the union is disjoint up to the common endpoint. Consequently, for any integrable ϱ*,*

$$\int_{I^{(\ell)}_{b_\ell(k)}} \varrho(x)\,dx = \int_{I^{(\ell+1)}_{0b_\ell(k)}} \varrho(x)\,dx + \int_{I^{(\ell+1)}_{1b_\ell(k)}} \varrho(x)\,dx.$$

Lemma 2.2 (Suffix extraction via scaling)**.** *Let* $k \in \mathbb{Z}_{2^n}$ *with* $b_n(k) = z_1 z_2 \cdots z_n$*. For* $1 \le s \le n-1$*, we have*

$$\left\lfloor 2^{-s}k \right\rfloor = \sum_{j=s+1}^{n} z_j\, 2^{j-s-1}, \qquad \textit{and hence} \qquad b_{n-s}\Big(\lfloor 2^{-s}k \rfloor\Big) = z_{s+1} \cdots z_n.$$

Proof. Using $k = \sum_{j=1}^{n} z_j 2^{j-1}$,

$$2^{-s}k = \sum_{j=1}^{s} z_j 2^{j-s-1} + \sum_{j=s+1}^{n} z_j 2^{j-s-1}.$$

The first sum belongs to $[0, 1)$, so taking the integer part yields the second sum, which is exactly the claimed binary value. □

2.6. Measurement probabilities and statement of the main theorem

We briefly recall how probabilities are extracted from a quantum state, since this is the criterion used to certify the correctness of the Grover-Rudolph construction.

A (mixed) quantum state is represented by a density matrix

$$\rho \in \mathbb{M}_N(\mathbb{C}), \qquad \rho \succeq 0, \qquad \operatorname{tr}(\rho) = 1,$$

and pure states correspond to rank-one projectors $\rho = |\psi\rangle\langle\psi|$ with $\||\psi\|| = 1$. Given a unitary $U \in \mathrm{U}(N)$ and an input state ρ_0, the output state is

$$\rho = U\rho_0 U^{\star}.$$

For the computational-basis measurement, define the orthogonal projectors

$$\Pi_k := |b_n(k)\rangle\langle b_n(k)|, \qquad k \in \mathbb{Z}_{2^n}, \qquad \sum_{k\in\mathbb{Z}_{2^n}} \Pi_k = I_N.$$

The associated outcome is a discrete random variable A taking values in $\mathbb{Z}_{2^n}$, and the Born rule yields the probability law

$$\mathbb{P}_\rho(\mathsf{A} = k) = \operatorname{tr}(\rho\,\Pi_k), \qquad k \in \mathbb{Z}_{2^n}. \tag{2.1}$$

In particular, if $\rho = |\psi\rangle\langle\psi|$ is pure, then

$$\mathbb{P}_\rho(\mathsf{A} = k) = |\langle b_n(k)|\psi\rangle|^2.$$

Theorem 2.1 (Grover-Rudolph state preparation)**.** *Let $n \geq 1$ and set $N := 2^n$. Let $\varrho : [0, 1] \to [0, \infty)$ satisfy $\int_0^1 \varrho(x)\,dx = 1$ and define $\{p_k\}_{k\in\mathbb{Z}_{2^n}}$ by $p_k = \int_{I_k} \varrho(x)\,dx$. Then there exists a quantum circuit $U \in \mathrm{U}(N)$, constructed as a product of multi-controlled single-qubit rotations, such that for $\rho_0 = |0\rangle^{\otimes n}\langle 0|^{\otimes n}$ and $\rho := U\rho_0 U^{\star}$, one has*

$$\mathbb{P}_\rho(\mathsf{A} = k) = \operatorname{tr}(\rho\,\Pi_k) = p_k, \qquad \forall\, k \in \mathbb{Z}_{2^n}.$$

Moreover, the circuit can be chosen with $N - 1$ elementary controlled-rotation gates.

3. Proof of the Grover-Rudolph theorem

Throughout this section, we set $N = 2^n$ and we work in $\mathbb{M}_N(\mathbb{C})$ with the computational basis $\{|b_n(k)\rangle : k \in \mathbb{Z}_{2^n}\}$, initial state $\rho_0 = |b_n(0)\rangle\langle b_n(0)| = |0\rangle^{\otimes n}\langle 0|^{\otimes n}$, and the measurement random variable A associated with the projectors $|b_n(k)\rangle\langle b_n(k)|$. For a bit $z \in \mathbb{Z}_2$ and an angle $x \in \mathbb{R}$, we use

$$\mathsf{T}_z(x) := (\cos x)^{1-z}(\sin x)^z.$$

3.1. Dyadic integrals and trigonometric factorization

Let $\varrho : [0,1] \to [0,\infty)$ be a density, i.e., $\int_0^1 \varrho(x)\,dx = 1$. For $m \geq 1$ and a binary word $\mathbf{w} \in \mathbb{Z}_2^m$, let $k(\mathbf{w}) \in \mathbb{Z}_{2^m}$ be the unique integer such that $b_m(k(\mathbf{w})) = \mathbf{w}$ (equivalently, if $\mathbf{w} = z_1 \cdots z_m$, then $k(\mathbf{w}) = \sum_{r=1}^m z_r 2^{r-1}$). We associate $\mathbf{w}$ to the dyadic interval of length 2^{-m}

$$I_{\mathbf{w}} := \Big[\frac{k(\mathbf{w})}{2^m},\ \frac{k(\mathbf{w})+1}{2^m}\Big] \subset [0,1],$$

and define its probability mass by

$$p_{\mathbf{w}} := \int_{I_w} \varrho(x)\,dx.$$

In particular, if $k \in \mathbb{Z}_{2^n}$ and $b_n(k) = \mathbf{w} = z_1 \cdots z_n$, then $k(\mathbf{w}) = k$ and thus

$$p_{z_1\cdots z_n} = p_{\mathbf{w}} = \int_{I_w} \varrho(x)\,dx = \int_{k/2^n}^{(k+1)/2^n} \varrho(x)\,dx.$$

By Corollary 2.1, every dyadic interval indexed by a word $\mathbf{w} \in \mathbb{Z}_2^m$ with $m < n$ splits into its two children $0\mathbf{w}$ and $1\mathbf{w}$. Therefore, for the associated probability masses, we have the refinement identity

$$p_{\mathbf{w}} = \int_{I_w^{(m)}} \varrho(x)\,dx = \int_{I_{0w}^{(m+1)}} \varrho(x)\,dx + \int_{I_{1w}^{(m+1)}} \varrho(x)\,dx = p_{0\mathbf{w}} + p_{1\mathbf{w}}.$$

Angle assignment. For each word $\mathbf{w} \in \mathbb{Z}_2^m$, $0 \leq m \leq n-1$, with $p_{\mathbf{w}} > 0$, we define $\theta_{\mathbf{w}} \in [0, \pi/2]$ by

$$\cos^2\theta_{\mathbf{w}} = \frac{p_{0\mathbf{w}}}{p_{\mathbf{w}}}, \qquad \sin^2\theta_{\mathbf{w}} = \frac{p_{1\mathbf{w}}}{p_{\mathbf{w}}}.$$

This is well defined because $p_{\mathbf{w}} = p_{0\mathbf{w}} + p_{1\mathbf{w}}$. If $p_{\mathbf{w}} = 0$, then $p_{0\mathbf{w}} = p_{1\mathbf{w}} = 0$, and the choice of $\theta_{\mathbf{w}}$ is immaterial for the prepared state; we set $\theta_{\mathbf{w}} := 0$ for definiteness.

For the empty word $\emptyset$, we write $\theta := \theta_\emptyset$. This is the stage-1 angle corresponding to the top-level dyadic split, namely

$$\theta = \theta_\emptyset := \arccos\sqrt{\frac{p_0}{p_\emptyset}} = \arccos\sqrt{\int_0^{1/2} \varrho(x)\,dx},$$

where $p_\emptyset = \int_0^1 \varrho(x)\,dx = 1$.

Proposition 3.1 (Trigonometric factorization)**.** *For each $k \in \mathbb{Z}_{2^n}$ with $b_n(k) = z_1 \cdots z_n$ one has*

$$p_{z_1\cdots z_n} = \mathsf{T}^2_{z_1}(\theta_{z_2\cdots z_n})\,\mathsf{T}^2_{z_2}(\theta_{z_3\cdots z_n})\cdots\mathsf{T}^2_{z_{n-1}}(\theta_{z_n})\,\mathsf{T}^2_{z_n}(\theta). \tag{3.1}$$

Proof. Fix $k \in \mathbb{Z}_{2^n}$ and write $b_n(k) = z_1 \cdots z_n$. For $r \in \{1,\dots,n\}$, let $\mathbf{w}_r := z_r z_{r+1} \cdots z_n$ denote the length-$(n-r+1)$ suffix, and set $\mathbf{w}_{n+1} := \emptyset$. We prove by backward induction on r that

$$p_{\mathbf{w}_r} = \mathsf{T}^2_{z_r}(\theta_{\mathbf{w}_{r+1}})\,\mathsf{T}^2_{z_{r+1}}(\theta_{\mathbf{w}_{r+2}})\cdots\mathsf{T}^2_{z_{n-1}}(\theta_{\mathbf{w}_n})\,\mathsf{T}^2_{z_n}(\theta), \tag{3.2}$$

where $\theta = \theta_\emptyset$ and, for any word $\mathbf{w}$, the angle $\theta_{\mathbf{w}}$ is defined by the convention used in the manuscript (in particular, if $p_{\mathbf{w}} = 0$, we set $\theta_{\mathbf{w}} := 0$).

Base case ($r = n$). We must show $p_{z_n} = \mathsf{T}^2_{z_n}(\theta)$. This holds by the definition of $\theta = \theta_\emptyset$:

$$p_0 = p_\emptyset \cos^2\theta = \cos^2\theta = \mathsf{T}^2_0(\theta), \qquad p_1 = p_\emptyset \sin^2\theta = \sin^2\theta = \mathsf{T}^2_1(\theta),$$

since $p_\emptyset = \int_0^1 \varrho(x)\,dx = 1$.

Inductive step. Assume (3.2) holds for $r+1$, and set $\mathbf{w} := \mathbf{w}_{r+1} = z_{r+1}\cdots z_n$. We prove (3.2) for r. There are two cases.

Case 1: $p_\mathbf{w} = 0$. By Corollary 2.1 (dyadic refinement), $p_{0\mathbf{w}} + p_{1\mathbf{w}} = p_\mathbf{w} = 0$, and since all probabilities are nonnegative, it follows that $p_{0\mathbf{w}} = p_{1\mathbf{w}} = 0$. In particular, $p_{\mathbf{w}_r} = p_{z_r\mathbf{w}} = 0$. On the other hand, our convention $\theta_\mathbf{w} = 0$ implies that $\mathsf{T}^2_{z_r}(\theta_\mathbf{w}) \in \{0,1\}$, and the right-hand side of (3.2) contains the factor $\mathsf{T}^2_{z_r}(\theta_\mathbf{w})$ multiplying the remaining terms. But the remaining product equals $p_\mathbf{w}$ by the induction hypothesis applied at level $r+1$ (since its left-hand side is $p_\mathbf{w}$), and therefore, the whole right-hand side equals $\mathsf{T}^2_{z_r}(\theta_\mathbf{w})\,p_\mathbf{w} = 0$. Hence, (3.2) holds.

Case 2: $p_\mathbf{w} > 0$. By definition of $\theta_\mathbf{w}$, we have

$$p_{0\mathbf{w}} = p_\mathbf{w}\cos^2\theta_\mathbf{w} = p_\mathbf{w}\,\mathsf{T}^2_0(\theta_\mathbf{w}), \qquad p_{1\mathbf{w}} = p_\mathbf{w}\sin^2\theta_\mathbf{w} = p_\mathbf{w}\,\mathsf{T}^2_1(\theta_\mathbf{w}).$$

Therefore, for $z_r \in \{0,1\}$,

$$p_{\mathbf{w}_r} = p_{z_r\mathbf{w}} = p_\mathbf{w}\,\mathsf{T}^2_{z_r}(\theta_\mathbf{w}).$$

Applying the induction hypothesis at level $r+1$ to $p_\mathbf{w} = p_{\mathbf{w}_{r+1}}$ yields

$$p_{\mathbf{w}_r} = \mathsf{T}^2_{z_r}(\theta_{\mathbf{w}_{r+1}})\,\mathsf{T}^2_{z_{r+1}}(\theta_{\mathbf{w}_{r+2}})\cdots\mathsf{T}^2_{z_{n-1}}(\theta_{\mathbf{w}_n})\,\mathsf{T}^2_{z_n}(\theta),$$

which is exactly (3.2).

This completes the induction. Taking $r = 1$ proves (3.1). □

3.2. Controlled rotations and state preparation

For $\alpha \in \mathbb{R}$, let

$$\mathsf{R}(\alpha) := \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \in \mathrm{U}(2).$$

For $j \in \{2,\dots,n\}$ and a binary word $\mathbf{w} \in \mathbb{Z}_2^{j-1}$, we denote by

$$\mathsf{CC}^{(1)}_\mathbf{w}(\mathsf{R}(\theta_\mathbf{w})) \in \mathrm{U}(2^j),$$

the multi-controlled one-qubit rotation acting on the *first* qubit conditional on the last $(j-1)$ qubits being equal to $\mathbf{w}$, and acting as the identity on all other basis states. Equivalently, $\mathsf{CC}^{(1)}_\mathbf{w}(\mathsf{R}(\theta_\mathbf{w}))$ applies $\mathsf{R}(\theta_\mathbf{w})$ to qubit 1 if and only if the control register (qubits $2,\dots,j$ in the j-qubit active subregister or, in the full n-qubit register, the last $(j-1)$ wires) matches the pattern $\mathbf{w}$.

Define the stage unitaries by

$$\mathsf{U}_1 := \mathsf{I}_{2^{n-1}} \otimes \mathsf{R}(\theta), \qquad \mathsf{U}_j := \prod_{\mathbf{w}\in\mathbb{Z}_2^{j-1}} \Big(\mathsf{I}_{2^{n-j}} \otimes \mathsf{CC}^{(1)}_\mathbf{w}(\mathsf{R}(\theta_\mathbf{w}))\Big), \quad 2 \le j \le n.$$

In particular, only the gates $\mathsf{CC}^{(1)}_\mathbf{w}(\mathsf{R}(\theta_\mathbf{w}))$ (with a fixed target qubit) are required for the inductive construction, while U_1 is an uncontrolled rotation on the last qubit.

Finally, we define the full circuit

$$\mathsf{U} := \mathsf{U}_n \mathsf{U}_{n-1} \cdots \mathsf{U}_1 \in \mathrm{U}(2^n).$$

By construction, U_j is a product of 2^{j-1} elementary controlled-rotation gates, hence the total number of elementary gates is

$$\sum_{j=1}^{n} 2^{j-1} = 2^n - 1 = N - 1,$$

i.e., U has circuit length $N - 1$.

Lemma 3.1 (Inductive amplitude form)**.** *Let* $\left|\psi_j\right\rangle := \mathsf{U}_j \mathsf{U}_{j-1} \cdots \mathsf{U}_1 |0\rangle^{\otimes n}$. *Then, for every* $1 \le j \le n$,

$$\left|\psi_j\right\rangle = |0\rangle^{\otimes(n-j)} \otimes \sum_{\mathbf{w}=z_{n-j+1}\cdots z_n \in \mathbb{Z}_2^j} \left(\mathsf{T}_{z_{n-j+1}}(\theta_{z_{n-j+2}\cdots z_n}) \cdots \mathsf{T}_{z_{n-1}}(\theta_{z_n}) \mathsf{T}_{z_n}(\theta)\right) |\mathbf{w}\rangle .$$

In particular,

$$|\psi_n\rangle = \sum_{z_1\cdots z_n \in \mathbb{Z}_2^n} \left(\mathsf{T}_{z_1}(\theta_{z_2\cdots z_n}) \cdots \mathsf{T}_{z_n}(\theta)\right) |z_1 \cdots z_n\rangle .$$

Proof. We proceed by induction on j.

Base case ($j = 1$)**.** By definition, $\mathsf{U}_1 = \mathsf{I}_{2^{n-1}} \otimes \mathsf{R}(\theta)$ acts only on the last qubit. Hence,

$$|\psi_1\rangle = \mathsf{U}_1 |0\rangle^{\otimes n} = |0\rangle^{\otimes(n-1)} \otimes \mathsf{R}(\theta) |0\rangle .$$

Since $\mathsf{R}(\theta) |0\rangle = \cos\theta\, |0\rangle + \sin\theta\, |1\rangle$, we obtain

$$|\psi_1\rangle = |0\rangle^{\otimes(n-1)} \otimes (\cos\theta\, |0\rangle + \sin\theta\, |1\rangle) = |0\rangle^{\otimes(n-1)} \otimes \sum_{z_n \in \mathbb{Z}_2} \mathsf{T}_{z_n}(\theta) |z_n\rangle ,$$

which is the claimed formula for $j = 1$.

Inductive step. Assume the statement holds for some $j - 1$ with $2 \le j \le n$, i.e.,

$$\left|\psi_{j-1}\right\rangle = |0\rangle^{\otimes(n-(j-1))} \otimes \sum_{\mathbf{w}=z_{n-j+2}\cdots z_n \in \mathbb{Z}_2^{j-1}} A_{\mathbf{w}} |\mathbf{w}\rangle , \tag{3.3}$$

where, for each $\mathbf{w} = z_{n-j+2} \cdots z_n$, the amplitude $A_{\mathbf{w}}$ is given by the product

$$A_{\mathbf{w}} = \mathsf{T}_{z_{n-j+2}}(\theta_{z_{n-j+3}\cdots z_n}) \cdots \mathsf{T}_{z_{n-1}}(\theta_{z_n})\, \mathsf{T}_{z_n}(\theta). \tag{3.4}$$

At this stage, the basis vectors $|\mathbf{w}\rangle$ label the computational states of the last $(j - 1)$ qubits. Their identification with suffixes of the global binary index follows from the fact that truncating the first $(j - 1)$ bits of an integer corresponds to extracting its suffix. More precisely, if $k \in \mathbb{Z}_{2^n}$ has binary expansion $b_n(k) = z_1 \cdots z_n$, then

$$b_{n-(j-1)}(\lfloor 2^{-(j-1)} k \rfloor) = z_{n-j+2} \cdots z_n,$$

as shown in Lemma 2.2.

It will be convenient to isolate the $(n-j+1)$-th qubit (the next qubit to be rotated at stage j). Write the tensor product in (3.3) as

$$|0\rangle^{\otimes(n-(j-1))} = |0\rangle^{\otimes(n-j)} \otimes |0\rangle,$$

so that

$$\left|\psi_{j-1}\right\rangle = |0\rangle^{\otimes(n-j)} \otimes \sum_{\mathbf{w}\in\mathbb{Z}_2^{j-1}} A_{\mathbf{w}} \left(|0\rangle \otimes |\mathbf{w}\rangle\right), \tag{3.5}$$

where the factor $|0\rangle$ corresponds to qubit $(n-j+1)$ and $|\mathbf{w}\rangle$ to the last $(j-1)$ qubits.

Action of U_j *on basis branches.* By definition,

$$\mathsf{U}_j = \prod_{\mathbf{w}\in\mathbb{Z}_2^{j-1}} \left(\mathsf{I}_{2^{n-j}} \otimes \mathsf{CC}_{\mathbf{w}}^{(1)}(\mathsf{R}(\theta_{\mathbf{w}}))\right).$$

Inside the active j-qubit register (qubits $(n-j+1),\ldots,n$), the gate $\mathsf{CC}_{\mathbf{w}}^{(1)}(\mathsf{R}(\theta_{\mathbf{w}})) \in \mathrm{U}(2^j)$ acts as follows: it applies $\mathsf{R}(\theta_{\mathbf{w}})$ to the *first* qubit of that register (i.e., qubit $(n-j+1)$) if and only if the remaining $(j-1)$ qubits are in the computational basis state $|\mathbf{w}\rangle$, and acts as the identity on all other computational basis states. Consequently, for each $\mathbf{w} \in \mathbb{Z}_2^{j-1}$,

$$\mathsf{CC}_{\mathbf{w}}^{(1)}(\mathsf{R}(\theta_{\mathbf{w}}))\left(|0\rangle \otimes |\mathbf{w}\rangle\right) = \left(\mathsf{R}(\theta_{\mathbf{w}})\,|0\rangle\right) \otimes |\mathbf{w}\rangle, \tag{3.6}$$

and for $\mathbf{w}' \neq \mathbf{w}$,

$$\mathsf{CC}_{\mathbf{w}}^{(1)}(\mathsf{R}(\theta_{\mathbf{w}}))\left(|0\rangle \otimes |\mathbf{w}'\rangle\right) = |0\rangle \otimes |\mathbf{w}'\rangle.$$

Since the projectors onto distinct control branches are orthogonal, these controlled rotations act on disjoint subspaces and therefore commute; hence, the product over $\mathbf{w}$ applies the correct rotation in each branch independently.

Splitting of amplitudes. Applying U_j to (3.5) and using (3.6) gives

$$\left|\psi_j\right\rangle = \mathsf{U}_j\left|\psi_{j-1}\right\rangle = |0\rangle^{\otimes(n-j)} \otimes \sum_{\mathbf{w}\in\mathbb{Z}_2^{j-1}} A_{\mathbf{w}} \left(\mathsf{R}(\theta_{\mathbf{w}})\,|0\rangle\right) \otimes |\mathbf{w}\rangle.$$

Finally, $\mathsf{R}(\theta_{\mathbf{w}})\,|0\rangle = \cos\theta_{\mathbf{w}}\ |0\rangle + \sin\theta_{\mathbf{w}}\ |1\rangle = \sum_{z\in\mathbb{Z}_2} \mathsf{T}_z(\theta_{\mathbf{w}})\,|z\rangle$, so

$$\begin{aligned}\left|\psi_j\right\rangle &= |0\rangle^{\otimes(n-j)} \otimes \sum_{\mathbf{w}\in\mathbb{Z}_2^{j-1}} A_{\mathbf{w}} \sum_{z_{n-j+1}\in\mathbb{Z}_2} \mathsf{T}_{z_{n-j+1}}(\theta_{\mathbf{w}})\ \left|z_{n-j+1}\right\rangle \otimes |\mathbf{w}\rangle \\ &= |0\rangle^{\otimes(n-j)} \otimes \sum_{z_{n-j+1}\in\mathbb{Z}_2} \sum_{\mathbf{w}\in\mathbb{Z}_2^{j-1}} \left(A_{\mathbf{w}}\,\mathsf{T}_{z_{n-j+1}}(\theta_{\mathbf{w}})\right) \left|z_{n-j+1}\mathbf{w}\right\rangle.\end{aligned}$$

Renaming the concatenated word $z_{n-j+1}\mathbf{w}$ as

$$\mathbf{w}' = z_{n-j+1}z_{n-j+2}\cdots z_n \in \mathbb{Z}_2^j,$$

and substituting $A_{\mathbf{w}}$ from (3.4), we obtain exactly the claimed product form

$$\left|\psi_j\right\rangle = |0\rangle^{\otimes(n-j)} \otimes \sum_{\mathbf{w}'=z_{n-j+1}\cdots z_n\in\mathbb{Z}_2^j} \left(\mathsf{T}_{z_{n-j+1}}(\theta_{z_{n-j+2}\cdots z_n})\cdots\mathsf{T}_{z_{n-1}}(\theta_{z_n})\mathsf{T}_{z_n}(\theta)\right)|\mathbf{w}'\rangle.$$

This completes the induction. The final statement for $|\psi_n\rangle$ is the special case $j = n$. □

3.3. Completion of the proof of Theorem 2.1

Proof of Theorem 2.1. Let $\rho := \mathsf{U}\rho_0\mathsf{U}^\star = |\psi_n\rangle\langle\psi_n|$, with $|\psi_n\rangle$ as in Lemma 3.1. For $k \in \mathbb{Z}_{2^n}$ with $b_n(k) = z_1 \cdots z_n$, we compute

$$\mathbb{P}_\rho(\mathsf{A} = k) = \mathrm{tr}(\rho\ |b_n(k)\rangle\langle b_n(k)|) = |\langle b_n(k)|\psi_n\rangle|^2 = \mathsf{T}^2_{z_1}(\theta_{z_2\cdots z_n}) \cdots \mathsf{T}^2_{z_n}(\theta).$$

By Proposition 3.1, the right-hand side equals $p_{z_1\cdots z_n} = \int_{k/2^n}^{(k+1)/2^n} \varrho(x)\,dx$; hence,

$$\mathbb{P}_\rho(\mathsf{A} = k) = \int_{k/2^n}^{(k+1)/2^n} \varrho(x)\,dx, \qquad \forall\, k \in \mathbb{Z}_{2^n}.$$

Finally, the circuit length of U is $N - 1$ by the gate count above. This proves the theorem. □

4. Stability under imperfect rotation angles

The results above establish the correctness of the Grover-Rudolph construction when all rotation angles are implemented exactly. In practice, however, the angles may be affected by finite-precision rounding, synthesis error, or calibration error. In this section, we quantify how such deterministic angular perturbations propagate to the final output distribution.

Throughout this section, we keep the rotation convention used in the previous sections, namely

$$\mathsf{R}(\alpha) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}, \qquad \mathsf{R}(\alpha)\,|0\rangle = \cos\alpha\ |0\rangle + \sin\alpha\ |1\rangle\,.$$

Thus the Grover-Rudolph angle $\theta_{\mathbf{w}}$ directly determines the binary split

$$\rho_{\mathbf{w}}(0) = \cos^2\theta_{\mathbf{w}}, \qquad \rho_{\mathbf{w}}(1) = \sin^2\theta_{\mathbf{w}}.$$

Let

$$\Theta = \{\theta_{\mathbf{w}} : \mathbf{w} \in \mathbb{Z}_2^m,\ 0 \le m \le n-1\}$$

be the exact family of Grover-Rudolph angles. We consider a perturbed family

$$\widetilde{\Theta} = \{\widetilde{\theta}_{\mathbf{w}} : \mathbf{w} \in \mathbb{Z}_2^m,\ 0 \le m \le n-1\}, \qquad \widetilde{\theta}_{\mathbf{w}} = \theta_{\mathbf{w}} + \varepsilon_{\mathbf{w}}.$$

Throughout this section, we adopt the convention $\mathbb{Z}_2^0 := \{\emptyset\}$, so that the index set $\bigcup_{m=0}^{n-1} \mathbb{Z}_2^m$ includes the empty word $\emptyset$ at level $m = 0$; the corresponding angle $\theta_\emptyset$ (resp. $\widetilde{\theta}_\emptyset$, $\varepsilon_\emptyset$) is the top-level rotation angle (resp. its perturbation, its error). With this convention, all sums and maxima over $\mathbf{w} \in \mathbb{Z}_2^m$, $0 \le m \le n-1$, are well defined. The perturbed circuit is obtained from the exact Grover-Rudolph circuit by replacing each occurrence of $\mathsf{R}(\theta_{\mathbf{w}})$ by $\mathsf{R}(\widetilde{\theta}_{\mathbf{w}})$, while keeping the same dyadic tree and the same control structure.

For a binary word $\mathbf{z} = z_1 \cdots z_n$, we use the suffix notation

$$\mathbf{z}_{r+1:n} := z_{r+1} \cdots z_n, \qquad 1 \le r \le n-1,$$

and we extend the convention to $r = n$ by setting $\mathbf{z}_{n+1:n} := \emptyset$, so that $\theta_{\mathbf{z}_{n+1:n}} = \theta_\emptyset$ is the top-level angle. With these conventions, the factorization proved in Proposition 3.1 reads

$$p_{\mathbf{z}} = \prod_{r=1}^{n} \mathsf{T}^2_{z_r}(\theta_{\mathbf{z}_{r+1:n}}),$$

where the suffix $\mathbf{z}_{r+1:n}$ has length $n - r \in \{0, 1, \ldots, n-1\}$, covering all control words $\mathbf{w} \in \mathbb{Z}_2^m$, $0 \le m \le n-1$. The perturbed output distribution is therefore

$$\widetilde{p}_{\mathbf{z}} = \prod_{r=1}^{n} \mathsf{T}^2_{z_r}(\widetilde{\theta}_{\mathbf{z}_{r+1:n}}).$$

We measure the distance between probability distributions on $\mathbb{Z}_2^n$ using total variation distance,

$$\mathrm{TV}(p, q) := \frac{1}{2} \sum_{\mathbf{z} \in \mathbb{Z}_2^n} |p_{\mathbf{z}} - q_{\mathbf{z}}|.$$

Theorem 4.1 (Stability under angular perturbations)**.** *Let p be the probability distribution prepared by the exact Grover-Rudolph circuit, and let $\widetilde{p}$ be the probability distribution obtained by replacing each angle $\theta_{\mathbf{w}}$ by*

$$\widetilde{\theta}_{\mathbf{w}} = \theta_{\mathbf{w}} + \varepsilon_{\mathbf{w}}.$$

Then

$$\mathrm{TV}(p, \widetilde{p}) \le \min\left(1, \sum_{m=0}^{n-1} \max_{\mathbf{w} \in \mathbb{Z}_2^m} |\varepsilon_{\mathbf{w}}|\right).$$

In particular, if

$$|\varepsilon_{\mathbf{w}}| \le \eta \qquad \textit{for all } \mathbf{w} \in \mathbb{Z}_2^m,\ 0 \le m \le n-1,$$

then

$$\mathrm{TV}(p, \widetilde{p}) \le \min(1, n\eta).$$

Proof. For each word $\mathbf{w} \in \mathbb{Z}_2^m$, $0 \le m \le n-1$, define the exact binary transition probability

$$\rho_{\mathbf{w}}(0) = \cos^2 \theta_{\mathbf{w}}, \qquad \rho_{\mathbf{w}}(1) = \sin^2 \theta_{\mathbf{w}},$$

and the perturbed binary transition probability

$$\widetilde{\rho}_{\mathbf{w}}(0) = \cos^2 \widetilde{\theta}_{\mathbf{w}}, \qquad \widetilde{\rho}_{\mathbf{w}}(1) = \sin^2 \widetilde{\theta}_{\mathbf{w}}.$$

Thus, for $\mathbf{z} = z_1 \cdots z_n$, using the convention $\mathbf{z}_{n+1:n} := \emptyset$,

$$p_{\mathbf{z}} = \prod_{r=1}^{n} \rho_{\mathbf{z}_{r+1:n}}(z_r), \qquad \widetilde{p}_{\mathbf{z}} = \prod_{r=1}^{n} \widetilde{\rho}_{\mathbf{z}_{r+1:n}}(z_r).$$

For $s = 0, \ldots, n$, define a hybrid distribution $p^{(s)}$ by using the perturbed transitions for the last s steps of the suffix recursion and the exact transitions for the remaining steps:

$$p^{(s)}_{\mathbf{z}} = \prod_{r=1}^{n-s} \rho_{\mathbf{z}_{r+1:n}}(z_r) \prod_{r=n-s+1}^{n} \widetilde{\rho}_{\mathbf{z}_{r+1:n}}(z_r).$$

Then, $p^{(0)} = p$ and $p^{(n)} = \widetilde{p}$. Each $p^{(s)}$ is a probability distribution on $\mathbb{Z}_2^n$: nonnegativity is clear since every factor $\rho_{\mathbf{z}_{r+1:n}}(z_r)$ and $\widetilde{\rho}_{\mathbf{z}_{r+1:n}}(z_r)$ is nonnegative, and the condition $\sum_{\mathbf{z}\in\mathbb{Z}_2^n} p_{\mathbf{z}}^{(s)} = 1$ follows by summing the product bit by bit: for each r, the identity $\sum_{z_r\in\{0,1\}} \rho_{\mathbf{z}_{r+1:n}}(z_r) = 1$ (respectively $\sum_{z_r\in\{0,1\}} \widetilde{\rho}_{\mathbf{z}_{r+1:n}}(z_r) = 1$) collapses the corresponding factor to 1, so the full sum telescopes to 1. Hence, since TV is a metric on the simplex of probability distributions, the triangle inequality gives

$$\mathrm{TV}(p, \widetilde{p}) \le \sum_{s=0}^{n-1} \mathrm{TV}(p^{(s)}, p^{(s+1)}).$$

The two hybrid distributions $p^{(s)}$ and $p^{(s+1)}$ differ only in the transition associated with words of length s. More precisely, conditioning on the suffix $\mathbf{w} \in \mathbb{Z}_2^s$, the only change is the replacement of the binary distribution $\rho_{\mathbf{w}}$ by $\widetilde{\rho}_{\mathbf{w}}$. Therefore,

$$\mathrm{TV}(p^{(s)}, p^{(s+1)}) \le \max_{\mathbf{w}\in\mathbb{Z}_2^s} \mathrm{TV}(\rho_{\mathbf{w}}, \widetilde{\rho}_{\mathbf{w}}).$$

For a binary distribution of the form

$$\rho_{\mathbf{w}} = (\cos^2\theta_{\mathbf{w}}, \sin^2\theta_{\mathbf{w}}),$$

we have

$$\mathrm{TV}(\rho_{\mathbf{w}}, \widetilde{\rho}_{\mathbf{w}}) = \left|\cos^2\theta_{\mathbf{w}} - \cos^2\widetilde{\theta}_{\mathbf{w}}\right|.$$

Using

$$\cos^2 a - \cos^2 b = -\sin(a+b)\sin(a-b),$$

we obtain

$$\left|\cos^2\theta_{\mathbf{w}} - \cos^2\widetilde{\theta}_{\mathbf{w}}\right| \le |\theta_{\mathbf{w}} - \widetilde{\theta}_{\mathbf{w}}| = |\varepsilon_{\mathbf{w}}|.$$

Consequently,

$$\mathrm{TV}(p^{(s)}, p^{(s+1)}) \le \max_{\mathbf{w}\in\mathbb{Z}_2^s} |\varepsilon_{\mathbf{w}}|.$$

Summing over $s = 0, \ldots, n-1$ gives

$$\mathrm{TV}(p, \widetilde{p}) \le \min\left(1, \sum_{s=0}^{n-1} \max_{\mathbf{w}\in\mathbb{Z}_2^s} |\varepsilon_{\mathbf{w}}|\right).$$

The uniform estimate follows immediately. □

Corollary 4.1 (Finite-precision angle synthesis)**.** *Assume that each Grover-Rudolph angle is approximated by an angle $\widetilde{\theta}_{\mathbf{w}}$ satisfying*

$$|\widetilde{\theta}_{\mathbf{w}} - \theta_{\mathbf{w}}| \le \eta.$$

Then, the output distribution $\widetilde{p}$ of the perturbed circuit satisfies

$$\mathrm{TV}(p, \widetilde{p}) \le \min(1, n\eta).$$

In particular, if the angles are rounded to a uniform grid of mesh Δ, so that

$$|\widetilde{\theta}_{\mathbf{w}} - \theta_{\mathbf{w}}| \leq \frac{\Delta}{2},$$

then

$$\mathrm{TV}(p, \widetilde{p}) \leq \min\left(1, \frac{n\Delta}{2}\right).$$

Remark 4.1 (Physical R_y-angles)**.** *The previous estimates are stated in terms of the Grover-Rudolph rotation convention*

$$\mathsf{R}(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Many quantum-computing libraries use instead the standard convention

$$\mathrm{R}_y(\varphi) = \begin{pmatrix} \cos(\varphi/2) & -\sin(\varphi/2) \\ \sin(\varphi/2) & \cos(\varphi/2) \end{pmatrix}.$$

Thus,

$$\mathsf{R}(\theta) = \mathrm{R}_y(2\theta).$$

If the implemented physical angle is

$$\widetilde{\varphi}_{\mathbf{w}} = \varphi_{\mathbf{w}} + \delta_{\mathbf{w}}, \qquad \varphi_{\mathbf{w}} = 2\theta_{\mathbf{w}},$$

then the corresponding Grover-Rudolph angular perturbation is

$$\varepsilon_{\mathbf{w}} = \frac{\delta_{\mathbf{w}}}{2}.$$

Therefore, Theorem 4.1 gives

$$\mathrm{TV}(p, \widetilde{p}) \leq \min\left(1, \frac{1}{2}\sum_{m=0}^{n-1} \max_{\mathbf{w}\in\mathbb{Z}_2^m} |\delta_{\mathbf{w}}|\right),$$

and, under the uniform physical-angle bound $|\delta_{\mathbf{w}}| \leq \delta$*,*

$$\mathrm{TV}(p, \widetilde{p}) \leq \min\left(1, \frac{n\delta}{2}\right).$$

Remark 4.2 (Interpretation of the bound)**.** *The estimate in Theorem 4.1 exploits the hierarchical structure of the Grover-Rudolph construction. A direct operator-norm perturbation bound for a product of* $2^n - 1$ *rotations would lead to a worst-case estimate proportional to the total number of rotations. By contrast, the present argument works at the level of the conditional probability tree and shows that the untruncated perturbation sum grows at most linearly with the depth* n *of the tree; the* $\min(1, \cdot)$ *truncation then reflects the fact that* $\mathrm{TV} \leq 1$ *always holds, so the bound is non-trivial only when* $n\eta < 1$*.*

5. Numerical validation and sensitivity analysis

This section provides the numerical validation of the Grover-Rudolph construction. Section 5.1 illustrates Theorem 2.1 on a three-qubit example, comparing the target probabilities obtained from the density ϱ with the empirical frequencies produced by a circuit simulation. Section 5.2 then carries out a systematic sensitivity study that quantifies the total-variation error as a joint function of angle-quantization precision and measurement shot count, for $n \in \{2, 3, 4\}$ qubits, thereby validating the theoretical bounds established in Section 4.

5.1. A numerical example

Consider the piecewise-linear probability density (Figure 1)

$$\varrho(x) = \begin{cases} 4x, & 0 \le x \le \frac{1}{2}, \\ 4 - 4x, & \frac{1}{2} \le x \le 1, \end{cases} \qquad \text{so that} \qquad \int_0^1 \varrho(x)\,dx = 1.$$

We set $n = 3$ (hence $N = 2^3 = 8$) and define the target probabilities

$$p_k := \int_{k/2^3}^{(k+1)/2^3} \varrho(x)\,dx, \qquad k \in \mathbb{Z}_{2^3}.$$

By Theorem 2.1, there exists a circuit $U \in \mathrm{U}(8)$, built from multi-controlled one-qubit rotations, such that for $\rho_0 = |0\rangle^{\otimes 3} \langle 0|^{\otimes 3}$ and $\rho = U\rho_0 U^\star$ one has

$$\mathbb{P}_\rho(\mathsf{A} = k) = p_k, \qquad \forall k \in \mathbb{Z}_{2^3},$$

where A denotes the computational-basis measurement random variable. Moreover, the construction uses $2^3 - 1 = 7$ controlled rotations, hence it is fully determined by 7 angles.

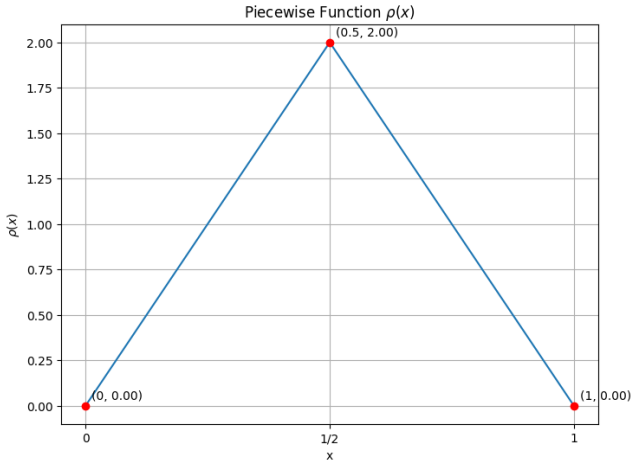

Figure 1. The probability density function $\varrho(x)$.

Step 1: computing U_1. The first stage acts on the last qubit and is given by

$$\mathsf{U}_1 := \mathsf{I}_{2^2} \otimes \mathsf{R}(\theta), \qquad \theta = \arccos \sqrt{\frac{\int_0^{1/2} \varrho(x)\,dx}{\int_0^1 \varrho(x)\,dx}} = \arccos \sqrt{\int_0^{1/2} \varrho(x)\,dx} = \frac{\pi}{4}.$$

Figure 2 illustrates the integrals involved in the evaluation of θ.

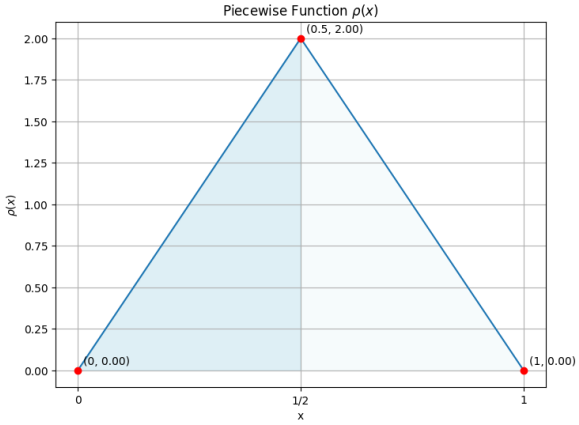

Figure 2. The integral $\int_0^1 \varrho(x)\,dx$ (blue) and $\int_0^{1/2} \varrho(x)\,dx$ (dark blue).

Step 2: computing U_2. At the second stage, we apply two controlled rotations (one per branch), namely

$$\mathsf{U}_2 := \prod_{z_2 \in \mathbb{Z}_2} \Big(\mathsf{I}_2 \otimes \mathsf{CC}_{z_2}^{(1)} \mathsf{R}(\theta_{z_2}) \Big),$$

with parameters

$$\theta_0 = \arccos \sqrt{\frac{\int_0^{1/4} \varrho(x)\,dx}{\int_0^{1/2} \varrho(x)\,dx}} = \frac{\pi}{3}, \qquad \theta_1 = \arccos \sqrt{\frac{\int_{1/2}^{3/4} \varrho(x)\,dx}{\int_{1/2}^{1} \varrho(x)\,dx}} = \frac{\pi}{6}.$$

Figure 3 shows the corresponding decomposition of the integrals.

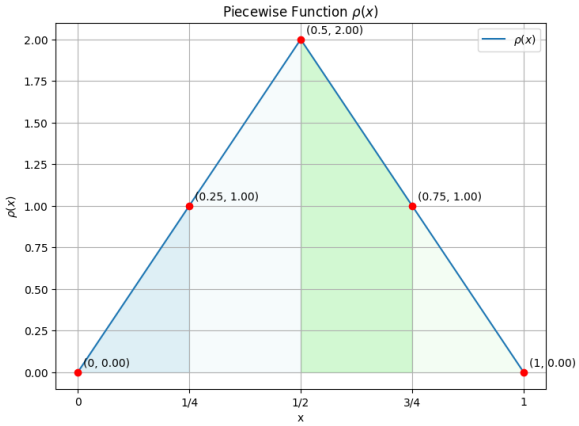

Figure 3. The integrals $\int_0^{1/2} \varrho(x)\,dx$ (blue) and $\int_{1/2}^1 \varrho(x)\,dx$ (green), and their dyadic refinement.

Step 3: computing U_3. Finally, we apply four controlled rotations corresponding to the four leaves of the depth-2 binary tree:

$$\mathsf{U}_3 := \prod_{z_1 z_2 \in \mathbb{Z}_2^2} \mathsf{CC}^{(1)}_{z_1 z_2}\, \mathsf{R}(\theta_{z_1 z_2}),$$

where the angles are

$$\theta_{00} = \frac{\pi}{3}, \qquad \theta_{11} = \frac{\pi}{6}, \qquad \theta_{01} = \arccos\frac{\sqrt{21}}{6}, \qquad \theta_{10} = \arccos\frac{\sqrt{15}}{6}.$$

Figure 4 illustrates the integral regions used to determine these parameters. The complete circuit is, therefore,

$$\mathsf{U} := \mathsf{U}_3\mathsf{U}_2\mathsf{U}_1 = \left(\prod_{z_1 z_2 \in \mathbb{Z}_2^2} \mathsf{CC}^{(1)}_{z_1 z_2}\, \mathsf{R}(\theta_{z_1 z_2})\right)\left(\prod_{z_2 \in \mathbb{Z}_2} \left(\mathsf{I}_2 \otimes \mathsf{CC}^{(1)}_{z_2}\, \mathsf{R}(\theta_{z_2})\right)\right)(\mathsf{I}_{2^2} \otimes \mathsf{R}(\theta)).$$

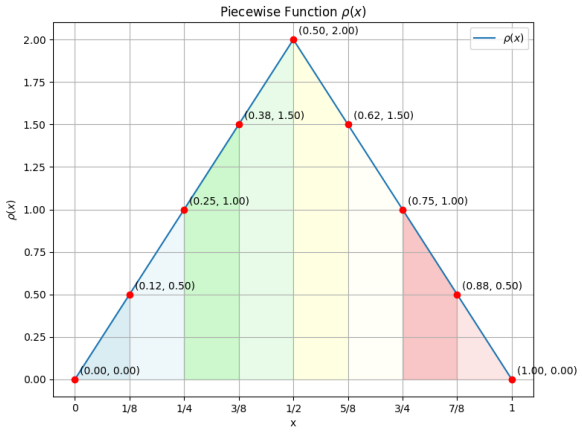

Figure 4. Integral regions used to compute θ_{00}, θ_{01}, θ_{10}, and θ_{11}.

Quantum circuit implementation and simulation. Figure 5 displays the corresponding circuit. We simulate this circuit using Qiskit's Aer module `qasm_simulator`. Running 2048 shots on the simulated 3-qubit device yields the empirical distribution shown in Figure 6.

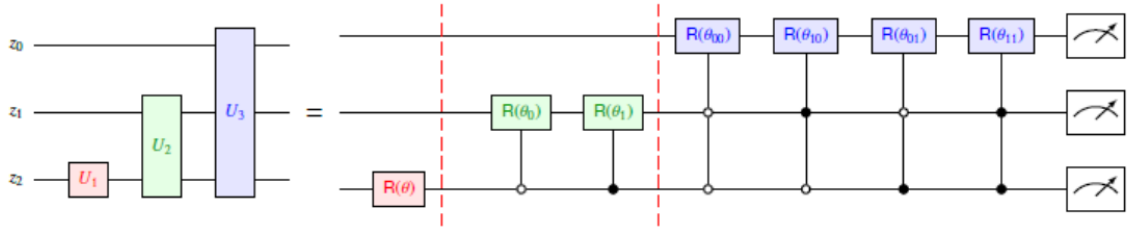

Figure 5. Quantum circuit U implementing Theorem 2.1 for $n = 3$.

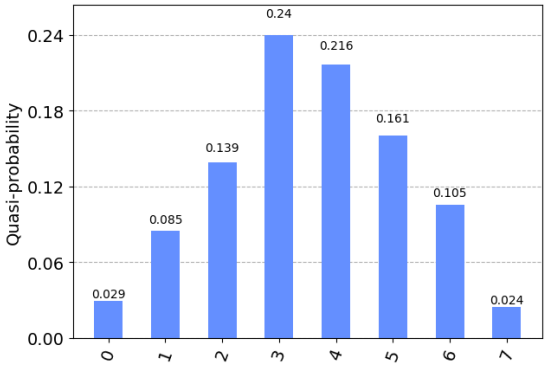

Figure 6. Empirical outcome distribution obtained from 2048 shots of the simulated circuit.

Discussion. This example illustrates concretely how the Grover-Rudolph recursion translates dyadic integrals of ϱ into a hierarchy of controlled rotation angles. Comparing the theoretical probabilities $\{p_k\}$ with the empirical frequencies produced by the simulator, we observe an overall good agreement. In particular, the maximum deviation occurs on the interval $[3/8, 1/2]$, where the estimated probability is 0.240, whereas the exact value is 0.21875, yielding an absolute error of approximately 0.02125. The minimum deviation is observed on $[1/2, 5/8]$, with an absolute error of about 0.00275. These discrepancies are consistent with finite-shot sampling fluctuations and with the finite-precision representation of rotation angles in the simulation.

A natural next step is to implement the same circuit on a real quantum device, where hardware noise, gate calibration errors, and decoherence effects would provide a more realistic assessment of the method's practical performance. Nevertheless, even in this idealized setting, the example provides a transparent verification of the construction and supports the theoretical results established in the previous sections.

5.2. Sensitivity to angle precision and shot noise

The Grover-Rudolph circuit depends on $N - 1$ real-valued rotation angles. In any practical implementation, two independent sources of error degrade the output distribution: (i) *angle quantization*, arising from finite-precision representation of the $\theta_{\mathbf{w}}$ values, and (ii) *shot noise*, arising from finite sampling of the prepared quantum state. We study both effects on the triangle density ϱ for $n \in \{2, 3, 4\}$ qubits.

Experimental setup. Let $\{p_k\}_{k\in\mathbb{Z}_{2^n}}$ denote the exact target probabilities and let $\{p_k^{(b)}\}$ denote the output probabilities obtained when each angle $\theta_{\mathbf{w}}$ is approximated by a finite-precision value. As recalled in Remark 4.1, the Grover-Rudolph angles $\theta_{\mathbf{w}} \in [0, \pi/2]$ defined by $\cos^2\theta_{\mathbf{w}} = p_{0\mathbf{w}}/p_{\mathbf{w}}$ are passed to Qiskit as physical R_y angles $\phi_{\mathbf{w}} := 2\theta_{\mathbf{w}} \in [0, \pi]$. We quantize the physical angles $\phi_{\mathbf{w}}$ to the nearest multiple of $\pi/2^{b-1}$ (a b-bit uniform grid on $[0, \pi]$), which is equivalent to quantizing $\theta_{\mathbf{w}}$ to the nearest multiple of $\pi/2^b$ on $[0, \pi/2]$. We measure accuracy by the *total variation distance* $\mathrm{TV}(p, q)$ defined in Section 4, which equals half the ℓ^1 error and is bounded in $[0, 1]$. All simulations use Qiskit's `AerSimulator`—the same backend employed in Section 5.1—with the `statevector` method for angle-precision experiments (noise-free exact output) and the `automatic` method (`qasm_simulator`)

for shot-noise experiments. Shot-noise experiments average 10 independent runs with fixed random seeds for reproducibility.

Effect of angle precision (Table 1). Table 1 reports $\mathrm{TV}(p, p^{(b)})$ computed via statevector simulation (no shot noise) for $b \in \{8, 16, 32\}$ bits and $n \in \{2, 3, 4\}$.

Table 1. TV distance due to angle quantization alone (`AerSimulator` statevector method, no shot noise). Each entry is $\mathrm{TV}(p, p^{(b)})$, where $p^{(b)}$ uses a b-bit uniform grid on the physical R_y angle $\phi_{\mathbf{w}} = 2\theta_{\mathbf{w}} \in [0, \pi]$, equivalently a grid of step $\pi/2^b$ on $\theta_{\mathbf{w}} \in [0, \pi/2]$.

n	$b = 8$	$b = 16$	$b = 32$
2	3.55×10^{-3}	1.38×10^{-5}	2.11×10^{-10}
3	3.56×10^{-3}	1.66×10^{-5}	1.51×10^{-10}
4	3.07×10^{-3}	8.80×10^{-6}	1.91×10^{-10}

The TV error decreases rapidly with b: going from 8 to 16 bits reduces the error by roughly three orders of magnitude, and 32-bit precision yields errors below 10^{-9}. Notably, the error is nearly *independent of* n: rounding each of the $2^n - 1$ physical angles independently does not cause an accumulation of errors in TV distance for fixed b, because each stage operates on an orthogonal subspace and errors do not compound multiplicatively across stages in total variation.

Effect of shot noise (Table 2). Table 2 reports the mean TV distance between p and the empirical frequency distribution obtained from $S \in \{256, 1024, 4096\}$ shots with exact angles, simulated via the `AerSimulator` `qasm_simulator` method.

Table 2. TV distance due to shot noise alone (`AerSimulator` `qasm_simulator`, exact angles, mean over 10 independent runs).

n	$S = 256$	$S = 1024$	$S = 4096$
2	0.03477	0.01855	0.01311
3	0.06250	0.02559	0.01733
4	0.09609	0.04482	0.02412

The TV distance scales approximately as $O(\sqrt{2^n/S})$, consistent with the standard Hoeffding bound stated in Remark 5.1 below. For fixed S, the error grows with n because there are more outcomes to estimate; for fixed n, the error decreases at the classical $O(1/\sqrt{S})$ rate.

Remark 5.1 (Hoeffding bound for the TV distance)**.** *Let $\hat{p}_k = n_k/S$ be the empirical frequency of outcome k obtained from S independent measurements of the prepared state, so that $(n_0, \ldots, n_{N-1})$ follows a multinomial distribution with parameters S and $(p_0, \ldots, p_{N-1})$. By the Hoeffding inequality [5] applied to each coordinate $\hat{p}_k - p_k \in [-1, 1]$, together with the union bound over the $N = 2^n$ outcomes, one obtains that with probability at least $1 - \delta$,*

$$\mathrm{TV}(p, \hat{p}) = \frac{1}{2} \sum_{k=0}^{N-1} |\hat{p}_k - p_k| \;\le\; \sqrt{\frac{2^n \log(2/\delta)}{2S}}.$$

In particular, achieving $\mathrm{TV}(p, \hat{p}) \le \varepsilon$ *with confidence* $1 - \delta$ *requires*

$$S \;\ge\; \frac{2^n \log(2/\delta)}{2\varepsilon^2},$$

i.e., $S = \Omega(2^n/\varepsilon^2)$ *shots. This should be contrasted with the deterministic angle-perturbation estimate of Corollary 4.1: for b-bit angle precision on* $\theta_{\mathbf{w}} \in [0, \pi/2]$ *(equivalently, b-bit precision on the physical* R_y *angle* $\phi_{\mathbf{w}} = 2\theta_{\mathbf{w}} \in [0, \pi]$*), the mesh is* $\Delta = \pi/2^b$*, and the angle-induced TV error satisfies* $\mathrm{TV}(p, \tilde{p}) \le n\pi/2^{b+1}$*, which is independent of* S *and decreases* exponentially *in b. The numerical results in Tables 1 and 2 confirm that the shot-noise term* $O(\sqrt{2^n/S})$ *dominates over the angle-precision term* $O(n/2^b)$ *for all practically relevant values of* $b \ge 8$ *and* $S \le 4096$*.*

Table 3 shows the combined TV distance for $b \in \{8, 16, 32\}$ bits with varying shots; the results for $b = 16$ and $b = 32$ are numerically indistinguishable at all shot counts, while $b = 8$ shows a small but visible excess at low shot counts.

Table 3. TV distance for combined angle quantization (b bits) and shot noise (`AerSimulator` `qasm_simulator`, mean over 10 runs). Results for $b = 16$ are numerically indistinguishable from $b = 32$; the small excess visible for $b = 8$ at low shot counts reflects the angle-quantization contribution $\mathrm{TV}(p, \tilde{p}) \lesssim 4 \times 10^{-3}$.

		$S = 256$	$S = 1024$	$S = 4096$
$b = 8$	$n = 2$	0.03594	0.01768	0.01343
	$n = 3$	0.06211	0.02627	0.01755
	$n = 4$	0.09648	0.04521	0.02456
$b = 16$	$n = 2$	0.03477	0.01855	0.01311
	$n = 3$	0.06250	0.02559	0.01733
	$n = 4$	0.09609	0.04482	0.02412
$b = 32$	$n = 2$	0.03477	0.01855	0.01313
	$n = 3$	0.06250	0.02559	0.01733
	$n = 4$	0.09609	0.04482	0.02412

The combined results confirm that *shot noise dominates* at practically relevant shot counts: for $S \ge 256$, the TV distance with $b = 8$ bits differs from the corresponding value with $b = 32$ bits by at most 0.00117 in absolute terms (e.g., 0.03594 vs 0.03477 for $n = 2$, $S = 256$), while the shot-noise contribution alone ranges from 0.01311 to 0.09648 across the $n \in \{2, 3, 4\}$, $S \in \{256, 4096\}$ combinations. In other words, for the triangle density with $n \le 4$ qubits, even 8-bit angle quantization introduces an angle-precision error ($\lesssim 4 \times 10^{-3}$ in TV) that is at least one order of magnitude smaller than the shot-noise contribution at $S = 256$. This suggests that in practice, moderate angle precision ($b = 8$ to 16 bits) is entirely sufficient, and the dominant cost of increasing n is the number of shots $S = \Omega(2^n/\varepsilon^2)$ needed to achieve a given TV accuracy ε.

The following corollary combines Theorem 4.1 with the Hoeffding bound of Remark 5.1 to give a single closed-form guarantee on the total TV error from both sources simultaneously.

Corollary 5.1 (Combined angle-quantization and shot-noise bound)**.** *Let* $n \ge 1$*,* $\delta \in (0, 1)$*, and* $\varepsilon \in (0, 1]$*. Suppose each Grover-Rudolph angle* $\theta_{\mathbf{w}} \in [0, \pi/2]$ *is implemented as a physical* $\mathrm{R}_y(2\widetilde{\theta}_{\mathbf{w}})$ *gate,*

with $|\widetilde{\theta}_{\mathbf{w}} - \theta_{\mathbf{w}}| \le \eta$ *for some* $\eta \ge 0$*. Let* $\widetilde{p}$ *denote the output distribution of the perturbed circuit, and let* $\hat{p}$ *be the empirical frequency distribution obtained from* S *independent measurements of the perturbed circuit. Then, with probability at least* $1 - \delta$*,*

$$\mathrm{TV}(p, \hat{p}) \;\le\; \min\left(1,\; n\eta + \sqrt{\frac{2^n \log(2/\delta)}{2S}}\right). \tag{5.1}$$

In particular, if the angles are quantized to a uniform b*-bit grid on* $[0, \pi/2]$ *(grid step* $\pi/2^b$*, maximum error* $\eta = \pi/2^{b+1}$*), then with probability at least* $1 - \delta$*,*

$$\mathrm{TV}(p, \hat{p}) \;\le\; \min\left(1,\; \frac{n\pi}{2^{b+1}} + \sqrt{\frac{2^n \log(2/\delta)}{2S}}\right). \tag{5.2}$$

Consequently, for $\varepsilon \in (0, 1]$ *and* $\delta \in (0, 1)$*, to achieve* $\mathrm{TV}(p, \hat{p}) \le \varepsilon$ *with confidence* $1 - \delta$*, it suffices to choose* b *and* S *such that*

$$\frac{n\pi}{2^{b+1}} \le \frac{\varepsilon}{2} \qquad \textit{and} \qquad S \;\ge\; \frac{2^n \log(2/\delta)}{2(\varepsilon/2)^2} = \frac{2^{n+1} \log(2/\delta)}{\varepsilon^2}. \tag{5.3}$$

The first condition requires $b \ge \log_2(2n\pi/\varepsilon)$*, i.e., the precision* b *only needs to grow* logarithmically *in* n *and* $1/\varepsilon$*. The second condition shows that the required shot count grows* linearly *in* $N = 2^n$ *(the number of outcomes) for fixed* ε *and* δ*.*

Proof. By the triangle inequality for total variation,

$$\mathrm{TV}(p, \hat{p}) \le \mathrm{TV}(p, \widetilde{p}) + \mathrm{TV}(\widetilde{p}, \hat{p}).$$

Theorem 4.1 gives $\mathrm{TV}(p, \widetilde{p}) \le n\eta$. The empirical distribution $\hat{p}$ is drawn from S independent samples of the perturbed circuit whose output distribution is $\widetilde{p}$; applying Remark 5.1 with $\widetilde{p}$ in place of p gives $\mathrm{TV}(\widetilde{p}, \hat{p}) \le \sqrt{2^n \log(2/\delta)/(2S)}$ with probability at least $1 - \delta$. Combining the two bounds yields (5.1). Inequality (5.2) follows by substituting $\eta = \pi/2^{b+1}$. The design rule (5.3) is obtained by splitting the error budget equally between the two terms and inverting each inequality. □

Remark 5.2 (Separation of deterministic and stochastic errors)**.** *Corollary 5.1 makes explicit the qualitatively different scaling of the two error sources. The angle-precision term* $n\eta$ *is* deterministic*, independent of* S*, and decreases* exponentially *in the bit depth* b*. The shot-noise term* $\sqrt{2^n \log(2/\delta)/(2S)}$ *is* stochastic*, independent of* b*, and decreases only as* $1/\sqrt{S}$*. As a consequence, once* b *is large enough that* $n\pi/2^{b+1}$ *is negligible relative to* ε*, increasing* b *further provides no benefit: the bottleneck is entirely the shot budget. The numerical experiments of Tables 1–3 confirm this picture for* $n \in \{2, 3, 4\}$*: already at* $b = 8$ *bits, the angle-precision error (*$\lesssim 4 \times 10^{-3}$*) is dominated by the shot-noise contribution at any practical shot count* $S \le 4096$*.*

6. Ancilla-free transpilation of the Grover-Rudolph circuit in a $\{\mathrm{R}_y(\cdot), X, \mathrm{CNOT}(\cdot \to \cdot)\}$ gate dictionary

In practical implementations, the circuit U in Theorem 2.1 must be expressed over a fixed elementary gate dictionary. The choice of dictionary depends on the target hardware; here we focus on

a gate set that is widely supported by current quantum processors and quantum-computing simulators. Throughout this section, we work with

$$\mathcal{G} := \{\mathrm{R}_y(\cdot),\ X,\ \mathrm{CNOT}(\cdot \to \cdot)\},$$

where $\mathrm{CNOT}(c \to t)$ denotes the standard two-qubit controlled-NOT (i.e., a controlled-X on the target t conditioned on the control c being in state $|1\rangle$).

6.1. Active-register viewpoint and target gates

At Grover-Rudolph stage $j \in \{2, \dots, n\}$, the construction acts nontrivially only on the last j qubits of the n-qubit register. We therefore introduce the *active j-qubit register*

$$(q_1, \dots, q_j) := (n - j + 1,\ n - j + 2,\ \dots,\ n), \qquad \text{i.e.,} \qquad q_r := n - j + r,\ \ r = 1, \dots, j,$$

so that any $G \in \mathrm{U}(2^j)$ acting on the active register is embedded into $\mathcal{H}_n = (\mathbb{C}^2)^{\otimes n}$ as $\mathsf{I}_{2^{n-j}} \otimes G \in \mathrm{U}(2^n)$.

For $\mathbf{w} \in \mathbb{Z}_2^{j-1}$ we write

$$\mathsf{CC}^{(1)}_{\mathbf{w}}(\mathsf{R}(\theta_{\mathbf{w}})) \in \mathrm{U}(2^j),$$

for the pattern-controlled one-qubit rotation acting on the *target* q_1 conditioned on the control register $(q_2, \dots, q_j)$ being equal to $\mathbf{w}$, and acting as the identity on all other computational basis states.

6.1.1. Dictionary gates

For a qubit q, we denote by X_q the Pauli-X gate on q, and by $\mathrm{R}_y(\alpha)_q$ the y-rotation

$$\mathrm{R}_y(\alpha) = \exp\Big(-\frac{i\alpha}{2} Y\Big) = \begin{pmatrix} \cos(\alpha/2) & -\sin(\alpha/2) \\ \sin(\alpha/2) & \cos(\alpha/2) \end{pmatrix}.$$

For qubits c (control) and t (target), the standard $\mathrm{CNOT}(c \to t)$ is defined by

$$\mathrm{CNOT}(c \to t)\,|x\rangle_c\,|y\rangle_t = |x\rangle_c\,|y \oplus x\rangle_t\,, \qquad x, y \in \mathbb{Z}_2,$$

and acts as the identity on all other qubits.

6.1.2. CNOT in terms of fully pattern-controlled X

If one wishes to express CNOT using the manuscript's fully pattern-controlled primitives $\mathsf{CC}^{(1)}_v(X) \in \mathrm{U}(2^j)$ (with $v \in \mathbb{Z}_2^{j-1}$ specifying the entire control word on $(q_2, \dots, q_j)$), then for $r \in \{1, \dots, j-1\}$ one has

$$\mathrm{CNOT}(q_{r+1} \to q_1) = \prod_{u \in \mathbb{Z}_2^{j-2}} \mathsf{CC}^{(1)}_{u_1 \cdots u_{r-1}\, 1\, u_r \cdots u_{j-2}}(X). \tag{6.1}$$

The factors commute because they act on pairwise orthogonal control subspaces, and exactly one factor is active on any branch with the r-th control bit equal to 1.

*6.2. Stage-as-*UCRY *compilation and Gray-code ladder*

6.2.1. Rotation primitive

Recall that

$$\mathsf{R}(\alpha) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} = \mathrm{R}_y(2\alpha),$$

so at stage j, it suffices to compile controlled R_y rotations with angles $\phi_{\mathbf{w}} := 2\theta_{\mathbf{w}}$.

6.2.2. Uniformly controlled viewpoint

Define the uniformly controlled y-rotation on the active register by

$$\mathrm{UCRY}(\phi) := \prod_{\mathbf{w}\in\mathbb{Z}_2^{j-1}} \mathsf{CC}_{\mathbf{w}}^{(1)}(\mathrm{R}_y(\phi_{\mathbf{w}})) \ \in \ \mathrm{U}(2^j), \qquad \phi = \{\phi_{\mathbf{w}}\}_{\mathbf{w}\in\mathbb{Z}_2^{j-1}}.$$

With $\phi_{\mathbf{w}} := 2\theta_{\mathbf{w}}$, the active stage unitary is precisely

$$U_j^{(\mathrm{act})} = \mathrm{UCRY}(\{2\theta_{\mathbf{w}}\}_{\mathbf{w}\in\mathbb{Z}_2^{j-1}}), \qquad \text{and the full-}n\text{ embedding is } \ \mathsf{I}_{2^{n-j}} \otimes U_j^{(\mathrm{act})} \in \mathrm{U}(2^n).$$

6.2.3. Gray-code conventions

Let $(\mathbf{g}_k)_{k=0}^{2^m-1}$ be the standard (binary-reflected) Gray code on $\mathbb{Z}_2^m$. For $k \in \mathbb{Z}_{2^m}$, define

$$\mathbf{g}_k := b_m(k) \oplus b_m\left(\left\lfloor \frac{k}{2} \right\rfloor\right) \in \mathbb{Z}_2^m, \qquad \gamma(k) := b_m^{-1}(\mathbf{g}_k) \in \mathbb{Z}_{2^m},$$

where $\oplus$ denotes componentwise XOR on $\mathbb{Z}_2^m$, and b_m^{-1} is the inverse map that sends a word $z_1 \cdots z_m$ to the integer $\sum_{r=1}^m z_r 2^{r-1}$. Then, consecutive words differ in exactly one bit. For $k \geq 1$, set

$$d_k := \gamma(k) \oplus \gamma(k-1) \in \mathbb{Z}_{2^m},$$

where $\oplus$ now denotes bitwise XOR on integers; in particular, d_k is a power of two. Let $j(k) \in \{1, \ldots, m\}$ be the unique index such that $d_k = 2^{\,j(k)-1}$. Because b_m^{-1} uses the LSB-first encoding ($b_m^{-1}(z_1 \cdots z_m) = \sum_r z_r 2^{r-1}$), a flip in bit position r of the word $\mathbf{g}_{k-1} \to \mathbf{g}_k$ contributes $\pm 2^{r-1}$ to $\gamma(k) - \gamma(k-1)$, so $d_k = 2^{r-1}$ and thus $j(k) = r$. Equivalently, $j(k)$ is the unique bit position $r \in \{1, \ldots, m\}$ that flips from $\mathbf{g}_{k-1}$ to $\mathbf{g}_k$, where positions are counted from z_1 (LSB) to z_m (MSB).

Example 6.1 (Binary-reflected Gray code for $m = 3$)**.** *We illustrate the binary-reflected Gray code using the paper's convention*

$$b_m : \mathbb{Z}_{2^m} \to \mathbb{Z}_2^m, \qquad b_m(k) = z_1 z_2 \cdots z_m \iff k = \sum_{r=1}^m z_r \, 2^{r-1},$$

(i.e., z_1 is the least significant bit). For $m = 3$ (so $2^m = 8$) and $k = 0, \ldots, 7$, define the Gray word

$$\mathbf{g}_k := b_3(k) \oplus b_3\left(\left\lfloor \frac{k}{2} \right\rfloor\right) \in \mathbb{Z}_2^3, \qquad \gamma(k) := b_3^{-1}(\mathbf{g}_k) \in \{0, \ldots, 7\},$$

where $\oplus$ denotes componentwise XOR on $\mathbb{Z}_2^3$, and b_3^{-1} is the inverse map.

Table 4 lists $b_3(k)$, $b_3(\lfloor k/2 \rfloor)$, $\mathbf{g}_k$, the integer $\gamma(k) = b_3^{-1}(\mathbf{g}_k)$ computed with the LSB-first convention $b_3^{-1}(z_1 z_2 z_3) = z_1 + 2z_2 + 4z_3$, and the flip mask $d_k = \gamma(k) \oplus \gamma(k-1)$ (integer XOR, undefined for $k = 0$).

Thus, in the paper's bit ordering, the Gray-code sequence is

$$\mathbf{g}_0 = 000, \ \mathbf{g}_1 = 100, \ \mathbf{g}_2 = 110, \ \mathbf{g}_3 = 010, \ \mathbf{g}_4 = 011, \ \mathbf{g}_5 = 111, \ \mathbf{g}_6 = 101, \ \mathbf{g}_7 = 001,$$

and consecutive words differ in exactly one bit. For $k \geq 1$, define the (integer) flip mask

$$d_k := \gamma(k) \oplus \gamma(k-1) \in \{1, 2, 4\},$$

where $\oplus$ now denotes bitwise XOR on integers. Since $\gamma(k)$ is computed with the LSB-first encoding ($b_3^{-1}(z_1z_2z_3) = z_1 + 2z_2 + 4z_3$), the bit z_r contributes 2^{r-1} to γ, so a flip in position r produces $d_k = 2^{r-1}$. In this example, one obtains

$$(d_1, \ldots, d_7) = (1, 2, 1, 4, 1, 2, 1),$$

so each d_k is a power of two. The flip index $j(k) \in \{1, 2, 3\}$ is the unique integer such that $d_k = 2^{j(k)-1}$, hence

$$(j(1), \ldots, j(7)) = (1, 2, 1, 3, 1, 2, 1).$$

Equivalently, $j(k)$ records the unique bit position (in the word $z_1z_2z_3$) that flips from $\mathbf{g}_{k-1}$ to $\mathbf{g}_k$. For instance,

$$\mathbf{g}_1 = 100 \longrightarrow \mathbf{g}_2 = 110$$

flips z_2 (the second bit), which corresponds to $d_2 = \gamma(2) \oplus \gamma(1) = 3 \oplus 1 = 2 = 2^{2-1}$ and thus $j(2) = 2$.

Table 4. Binary-reflected Gray code for $m = 3$: binary words $b_3(k)$, $b_3(\lfloor k/2 \rfloor)$, Gray word $\mathbf{g}_k$, integer $\gamma(k) = b_3^{-1}(\mathbf{g}_k)$ with LSB-first convention, and flip mask $d_k = \gamma(k) \oplus \gamma(k-1)$ (undefined for $k = 0$).

k	$b_3(k)$	$b_3(\lfloor k/2 \rfloor)$	$\mathbf{g}_k$	$\gamma(k)$	d_k
0	000	000	000	0	–
1	100	000	100	1	1
2	010	100	110	3	2
3	110	100	010	2	1
4	001	010	011	6	4
5	101	010	111	7	1
6	011	110	101	5	2
7	111	110	001	4	1

Proposition 6.1 (Gray-code ladder for a uniformly controlled R_y gate in $\mathrm{U}(2^j)$)**.** *Let $j \geq 2$ and consider the active register $(q_1, \ldots, q_j)$, where q_1 is the target and $C = (q_2, \ldots, q_j)$ are the $m := j - 1$ control qubits. Given an angle list $\phi = \{\phi_{\mathbf{w}}\}_{\mathbf{w} \in \mathbb{Z}_2^m}$, define the uniformly controlled rotation*

$$\mathrm{UCRY}(\phi) := \prod_{\mathbf{w} \in \mathbb{Z}_2^m} \mathsf{CC}_{\mathbf{w}}^{(1)}(\mathrm{R}_y(\phi_{\mathbf{w}})) \ \in \ \mathrm{U}(2^j), \tag{6.2}$$

i.e., $\mathrm{UCRY}(\phi)$ *applies $\mathrm{R}_y(\phi_{\mathbf{w}})$ to q_1 conditioned on the control register being $|\mathbf{w}\rangle$.*

Let $(\mathbf{g}_k)_{k=0}^{2^m-1}$ be the standard (binary-reflected) Gray code on $\mathbb{Z}_2^m$ and let $j(k) \in \{1, \ldots, m\}$ denote the unique bit that flips from $\mathbf{g}_{k-1}$ to $\mathbf{g}_k$. Define a new angle list $\boldsymbol{\alpha} = \{\alpha_{\mathbf{v}}\}_{\mathbf{v} \in \mathbb{Z}_2^m}$ by the Walsh-Hadamard transform

$$\alpha_{\mathbf{v}} := \frac{1}{2^m} \sum_{\mathbf{w} \in \mathbb{Z}_2^m} (-1)^{\langle \mathbf{v}, \mathbf{w} \rangle} \, \phi_{\mathbf{w}}, \qquad \mathbf{v} \in \mathbb{Z}_2^m, \tag{6.3}$$

where $\langle \mathbf{v}, \mathbf{w} \rangle := \sum_{r=1}^{m} v_r w_r \pmod 2$. Then $\mathrm{UCRY}(\phi)$*, admits the ancilla-free Gray-code ladder decomposition (cf. [1])*

$$\mathrm{UCRY}(\phi) = \mathrm{R}_y(\alpha_{\mathbf{g}_0})_{q_1} \prod_{k=1}^{2^m-1} \Big[\mathrm{CNOT}(q_{1+j(k)} \to q_1) \, \mathrm{R}_y(\alpha_{\mathbf{g}_k})_{q_1} \Big]. \tag{6.4}$$

In particular, the circuit uses exactly 2^m single-qubit R_y gates and $2^m - 1$ CNOT gates and requires no ancillas.

Proof. **Gray-walk viewpoint.** Fix a control word $\mathbf{u} = u_1 \cdots u_m \in \mathbb{Z}_2^m$ and restrict the Gray-code ladder circuit (6.4) to the branch $\mathbb{C}^2 \otimes |\mathbf{u}\rangle$, where $\mathbb{C}^2$ refers to the target qubit q_1 and $|\mathbf{u}\rangle$ to the control register $(q_2, \ldots, q_j)$. On this branch, each CNOT$(q_{1+j(k)} \to q_1)$ applies X to the target if and only if the corresponding control bit equals 1, hence it acts as $X^{u_{j(k)}}$ on q_1. Therefore, the branch-restricted unitary reads

$$U_{\mathbf{u}} = \mathrm{R}_y(\alpha_{\mathbf{g}_0}) \prod_{k=1}^{2^m-1} \Big(X^{u_{j(k)}} \, \mathrm{R}_y(\alpha_{\mathbf{g}_k}) \Big),$$

where we omit the target subscript q_1 for readability.

Let $S_k(\mathbf{u}) := \sum_{\ell=1}^{k} u_{j(\ell)} \pmod 2$ be the parity of the number of X toggles applied up to step k on this branch. Using the elementary conjugation identity

$$X \, \mathrm{R}_y(\alpha) \, X = \mathrm{R}_y(-\alpha) \qquad (\alpha \in \mathbb{R}),$$

we can rewrite each factor $\mathrm{R}_y(\alpha_{\mathbf{g}_k})$ as conjugated by the cumulative toggle $X^{S_{k-1}(\mathbf{u})}$, so that the effective angle at step k acquires a sign

$$\mathrm{R}_y(\alpha_{\mathbf{g}_k}) \;\longmapsto\; \mathrm{R}_y((-1)^{S_{k-1}(\mathbf{u})} \, \alpha_{\mathbf{g}_k}).$$

For the binary-reflected Gray code, the parity $S_{k-1}(\mathbf{u})$ equals the $\mathbb{Z}_2$ inner product $\langle \mathbf{u}, \mathbf{g}_k \rangle := \sum_{r=1}^{m} u_r (\mathbf{g}_k)_r \pmod 2$, hence the branch action simplifies to

$$U_{\mathbf{u}} = \mathrm{R}_y\Big(\sum_{k=0}^{2^m-1} (-1)^{\langle \mathbf{u}, \mathbf{g}_k \rangle} \, \alpha_{\mathbf{g}_k} \Big) \, X^{S_{2^m-1}(\mathbf{u})}.$$

Finally, along the Gray-code ladder, each control index $r \in \{1, \ldots, m\}$ appears exactly 2^{m-1} times among the flips $j(1), \ldots, j(2^m - 1)$. Therefore,

$$S_{2^m-1}(\mathbf{u}) = \sum_{r=1}^{m} u_r \, 2^{m-1} \equiv 0 \pmod 2 \qquad (m \geq 2),$$

so the net trailing toggle is $X^{S_{2^m-1}(\mathbf{u})} = \mathrm{I}$ and no residual Pauli factor remains. Consequently, on each branch $\mathbf{u}$, the ladder realizes a *pure* R_y rotation with angle $\sum_k (-1)^{\langle \mathbf{u}, \mathbf{g}_k \rangle} \alpha_{\mathbf{g}_k}$, which is precisely the uniformly controlled action once $\boldsymbol{\alpha}$ is chosen as the Walsh-Hadamard transform of $\boldsymbol{\phi}$ in (6.3).

Evaluation of the effective branch angle (Walsh-Hadamard inversion). From the Gray-walk analysis above (and the fact that the trailing toggle cancels for $m \geq 2$), the restriction of the ladder circuit (6.4) to the control branch $\mathbb{C}^2 \otimes |\mathbf{u}\rangle$ is a pure rotation

$$U_{\mathbf{u}} = \mathrm{R}_y(\widetilde{\phi}(\mathbf{u})), \qquad \widetilde{\phi}(\mathbf{u}) := \sum_{k=0}^{2^m-1} (-1)^{\langle \mathbf{u}, \mathbf{g}_k \rangle} \, \alpha_{\mathbf{g}_k},$$

(for $m = 1$, the ladder reduces to the standard two-gate controlled-R_y synthesis, and the same conclusion holds). Since the Gray code $(\mathbf{g}_k)_{k=0}^{2^m-1}$ enumerates $\mathbb{Z}_2^m$ exactly once, we can re-index the sum by $\mathbf{v} \in \mathbb{Z}_2^m$ and write

$$\widetilde{\phi}(\mathbf{u}) = \sum_{\mathbf{v} \in \mathbb{Z}_2^m} (-1)^{\langle \mathbf{u}, \mathbf{v} \rangle} \, \alpha_{\mathbf{v}}.$$

Using the definition (6.3),

$$\alpha_{\mathbf{v}} = \frac{1}{2^m} \sum_{\mathbf{w}\in\mathbb{Z}_2^m} (-1)^{\langle \mathbf{v},\mathbf{w}\rangle}\, \phi_{\mathbf{w}},$$

we obtain, by exchanging sums,

$$\widetilde{\phi}(\mathbf{u}) = \frac{1}{2^m} \sum_{\mathbf{w}\in\mathbb{Z}_2^m} \phi_{\mathbf{w}} \sum_{\mathbf{v}\in\mathbb{Z}_2^m} (-1)^{\langle \mathbf{u}+\mathbf{w},\mathbf{v}\rangle}.$$

By orthogonality of characters of the group $(\mathbb{Z}_2^m, +)$,

$$\sum_{\mathbf{v}\in\mathbb{Z}_2^m} (-1)^{\langle \mathbf{a},\mathbf{v}\rangle} = \begin{cases} 2^m, & \mathbf{a} = \mathbf{0}, \\ 0, & \mathbf{a} \neq \mathbf{0}, \end{cases}$$

hence, only the term $\mathbf{w} = \mathbf{u}$ survives, and $\widetilde{\phi}(\mathbf{u}) = \phi_{\mathbf{u}}$. Therefore, $U_{\mathbf{u}} = \mathrm{R}_y(\phi_{\mathbf{u}})$ on every branch $\mathbf{u}$, which is exactly the defining action of $\mathrm{UCRY}(\phi)$ in (6.2). □

6.2.4. Conclusion: transpiling the Grover-Rudolph stage unitaries

At Grover-Rudolph stage $j \in \{2, \ldots, n\}$, the circuit applies, on the active j-qubit register, the product of pattern-controlled rotations

$$U_j^{(\mathrm{act})} = \prod_{\mathbf{w}\in\mathbb{Z}_2^{j-1}} \mathsf{CC}_{\mathbf{w}}^{(1)}(\mathsf{R}(\theta_{\mathbf{w}})) = \prod_{\mathbf{w}\in\mathbb{Z}_2^{j-1}} \mathsf{CC}_{\mathbf{w}}^{(1)}(\mathrm{R}_y(2\theta_{\mathbf{w}})).$$

Equivalently, $U_j^{(\mathrm{act})}$ is a uniformly controlled R_y gate with angle list $\phi = \{\phi_{\mathbf{w}}\}_{\mathbf{w}\in\mathbb{Z}_2^{j-1}}$ given by $\phi_{\mathbf{w}} := 2\theta_{\mathbf{w}}$. Proposition 6.1 then yields an explicit ancilla-free Gray-code ladder realization of $U_j^{(\mathrm{act})}$ over the gate dictionary $\mathcal{G} = \{X, \mathrm{R}_y(\cdot), \mathrm{CNOT}(\cdot \to \cdot)\}$, using exactly 2^{j-1} R_y gates and $2^{j-1} - 1$ CNOT gates. When embedded into the full n-qubit space, this acts as $\mathsf{I}_{2^{n-j}} \otimes U_j^{(\mathrm{act})} \in \mathrm{U}(2^n)$ (see Figure 7 for the ladder pattern).

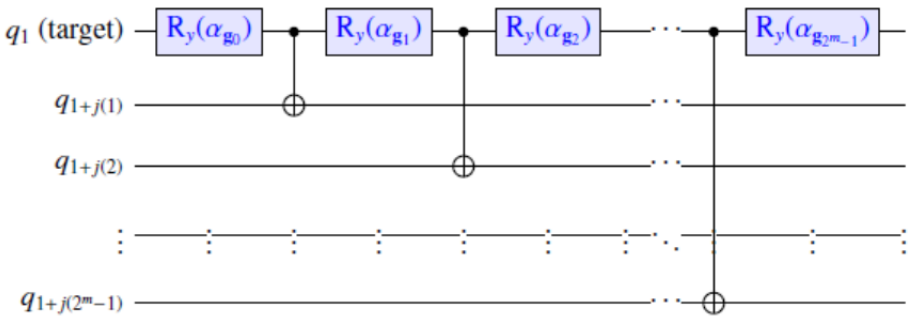

Figure 7. Gray-code ladder realization of $\mathsf{CC}_{\mathbf{1}}^{(1)}(\mathrm{R}_y(\phi))$ on the active register $(q_1, \ldots, q_j)$, where $m = j - 1$. At step k, the CNOT uses control $q_{1+j(k)}$ and target q_1, followed by $\mathrm{R}_y(\alpha_{\mathbf{g}_k})$ on q_1.

Remark 6.1 (practical compilation guideline)**.** *In implementations, it is typically advantageous to transpile each Grover-Rudolph stage as a single uniformly controlled* R_y *gate (multiplexor) rather than transpiling the* 2^{j-1} *pattern-controlled factors individually; this avoids redundant synthesis and directly yields the Gray-code ladder structure of Proposition 6.1 (cf. [1]).*

6.3. *Pseudo-code: ancilla-free Gray-code ladder compilation of a Grover-Rudolph stage as* UCRY

For completeness, Algorithm 1 gives an explicit pseudo-code implementation of the ancilla-free Gray-code ladder used to compile each Grover-Rudolph stage as a uniformly controlled R_y gate.

Algorithm 1: Ancilla-free Gray-code ladder compilation of UCRY(ϕ) $\in$ U(2^j).

Input: Active register $(q_1, \ldots, q_j)$ with target q_1 and controls $(q_2, \ldots, q_j)$;
angle list $\phi = \{\phi_{\mathbf{w}}\}_{\mathbf{w}\in\mathbb{Z}_2^m}$ with $m \leftarrow j-1$. (In Grover-Rudolph stage j: $\phi_{\mathbf{w}} \leftarrow 2\theta_{\mathbf{w}}$.)
Output: A gate list over $\{R_y(\cdot), \mathrm{CNOT}(\cdot \to \cdot)\}$ implementing UCRY(ϕ).
$m \leftarrow j-1$;

(A) Compute Walsh-Hadamard angles α:;
foreach $\mathbf{v} \in \mathbb{Z}_2^m$ **do**
  $\alpha_{\mathbf{v}} \leftarrow \dfrac{1}{2^m} \displaystyle\sum_{\mathbf{w}\in\mathbb{Z}_2^m} (-1)^{\langle \mathbf{v},\mathbf{w}\rangle}\, \phi_{\mathbf{w}}$;

// In practice, compute α with a fast Walsh-Hadamard transform in $O(m2^m)$.

(B) Gray-code ladder order on the control space:;
for $k \leftarrow 0$ **to** $2^m - 1$ **do**
  $\mathbf{g}_k \leftarrow b_m(k) \oplus b_m(\lfloor k/2 \rfloor) \in \mathbb{Z}_2^m$;
  $\gamma_k \leftarrow b_m^{-1}(\mathbf{g}_k) \in \mathbb{Z}_{2^m}$;
  if $k = 0$ **then**
    Apply $R_y(\alpha_{\mathbf{g}_0})$ on q_1;
  else
    $d \leftarrow \gamma_k \oplus \gamma_{k-1}$; // integer XOR; d is a power of two
    choose the unique $s \in \{1, \ldots, m\}$ such that $d = 2^{s-1}$;
    Apply $\mathrm{CNOT}(q_{1+s} \to q_1)$;
    Apply $R_y(\alpha_{\mathbf{g}_k})$ on q_1;

Note. For the Grover-Rudolph construction, set $\phi = 2\theta_{\mathbf{w}}$, since $\mathsf{R}(\theta_{\mathbf{w}}) = R_y(2\theta_{\mathbf{w}})$.

7. Conclusions

We have presented a rigorous and self-contained proof of correctness of the Grover-Rudolph state preparation algorithm. The analysis makes explicit the dyadic refinement structure underlying the method, derives a trigonometric factorization of the target probabilities in terms of conditional masses, and proves by induction that the resulting hierarchy of controlled rotations produces exactly the desired measurement law in the computational basis.

Beyond exact correctness, we established a quantitative stability theorem for imperfectly implemented Grover-Rudolph angles (Theorem 4.1). The estimate shows that, if every Grover-Rudolph angle is perturbed by at most η, then the resulting output distribution differs from the target distribution by at most $\min(1, n\eta)$ in total variation distance. This linear-in-depth bound exploits the conditional binary-tree structure of the construction. Combining this estimate with the Hoeffding concentration inequality, Corollary 5.1 gives a unified closed-form guarantee: for $\varepsilon \in (0, 1]$, with probability at

least $1-\delta$,

$$\mathrm{TV}(p,\hat{p}) \;\le\; \min\left(1,\; \frac{n\pi}{2^{b+1}} + \sqrt{\frac{2^n \log(2/\delta)}{2S}}\right),$$

where b is the angle bit depth, and S, the number of measurement shots. Inverting this bound yields an explicit design rule: $b \ge \log_2(2n\pi/\varepsilon)$ bits and $S \ge 2^{n+1}\log(2/\delta)/\varepsilon^2$ shots suffice to achieve $\mathrm{TV}(p,\hat{p}) \le \varepsilon$ with confidence $1-\delta$. A key practical consequence (Remark 5.2) is that the required bit depth b grows only *logarithmically* in n and $1/\varepsilon$, whereas the shot budget S must grow *linearly* in $N = 2^n$; the two error sources therefore have qualitatively different scaling, and once b is large enough, increasing it further provides no benefit.

The numerical experiments of Section 5.2 confirm these theoretical predictions for the triangle density with $n \in \{2,3,4\}$ qubits: the angle-precision error at $b = 8$ bits is at most 4×10^{-3} in TV, already dominated by shot noise at all tested shot counts $S \in \{256, 1024, 4096\}$, consistently with the $O(\sqrt{2^n/S})$ rate of the Hoeffding bound.

Finally, as a practical complement, we described an ancilla-free compilation viewpoint in which each Grover-Rudolph stage is treated as a uniformly controlled one-qubit rotation, or multiplexor, admitting a Gray-code ladder decomposition. This provides a direct route to gate-set transpilation on devices with a native dictionary such as $\{\mathrm{R}_y(\cdot), X, \mathrm{CNOT}\}$, while remaining logically separate from the proof of the Grover-Rudolph theorem.

Possible directions for future work include: (i) sharper *distribution-dependent* stability estimates that exploit cancellations in the dyadic tree or non-uniform angular error profiles, beyond the worst-case uniform bound of Theorem 4.1; (ii) hardware-aware cost models for compiling the stage multiplexors under connectivity constraints; and (iii) extensions to structured families of distributions, such as log-concave or sparse distributions, in the spirit of [10], where additional synthesis savings may be achievable.

Author contributions

Antonio Falcó, Daniela Falcó-Pomares and Hermann G. Matthies: Conceptualization, Methodology, Formal analysis, Investigation, Writing—original draft, Writing—review & editing. All authors contributed equally to this work.

Use of Generative-AI tools declaration

The authors used generative-AI tools for language editing and/or formatting assistance. All scientific content, derivations, and conclusions were produced and verified by the authors, who take full responsibility for the manuscript.

Acknowledgments

This work was supported by the Generalitat Valenciana under grant COMCUANTICA/007 (QUANTWin), by the Agreement between the Directorate-General for Innovation of the Ministry of Innovation, Industry, Trade and Tourism of the Generalitat Valenciana and the Universidad CEU Cardenal Herrera, and by Universidad CEU Cardenal Herrera under grants INDI25/17 and GIR25/14.

The funders had no role in the study design; in the collection, analysis, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Conflict of interest

The authors declare that they have no conflict of interest.

References

1. V. Bergholm, J. J. Vartiainen, M. Möttönen, M. M. Salomaa, Quantum circuits with uniformly controlled one-qubit gates, *Phys. Rev. A*, **71** (2005), 052330. https://doi.org/10.1103/PhysRevA.71.052330
2. G. Brassard, P. Høyer, M. Mosca, A. Tapp, Quantum amplitude amplification and estimation, *Contemp. Math.*, **305** (2002), 53–74. https://doi.org/10.1090/conm/305/05215
3. J. González-Conde, T. W. Watts, P. Rodriguez-Grasa, M. Sanz, Efficient quantum amplitude encoding of polynomial functions, *Quantum*, **8** (2024), 1297. https://doi.org/10.22331/q-2024-03-21-1297
4. L. K. Grover, T. Rudolph, Creating superpositions that correspond to efficiently integrable probability distributions, arXiv: quant-ph/0208112. https://doi.org/10.48550/arXiv.quant-ph/0208112
5. W. Hoeffding, Probability inequalities for sums of bounded random variables, *J. Amer. Stat. Assoc.*, **58** (1963), 13–30. https://doi.org/10.1080/01621459.1963.10500830
6. J. Luo, G. Li, L. Li, Space-time tradeoff for sparse quantum state preparation, arXiv: 2506.16964. https://doi.org/10.48550/arXiv.2506.16964
7. R. Mao, G. Tian, X. Sun, Toward optimal circuit size for sparse quantum state preparation, *Phys. Rev. A*, **110** (2024), 032439. https://doi.org/10.1103/PhysRevA.110.032439
8. G. Marín-Sánchez, J. González-Conde, M. Sanz, Quantum algorithms for approximate function loading, *Phys. Rev. Res.*, **5** (2023), 033114. https://doi.org/10.1103/PhysRevResearch.5.033114
9. M. Möttönen, J. J. Vartiainen, V. Bergholm, M. M. Salomaa, Transformation of quantum states using uniformly controlled rotations, *Quant. Inf. Comp.*, **5** (2005), 467–473.
10. D. Ramacciotti, A. I. Lefterovici, A. F. Rotundo, A simple quantum algorithm to efficiently prepare sparse states, *Phys. Rev. A*, **110** (2024), 032609. https://doi.org/10.1103/PhysRevA.110.032609
11. X. Zhang, T. Li, X. Yuan, Quantum state preparation with optimal circuit depth: implementations and applications, *Phys. Rev. Lett.*, **129** (2022), 230504. https://doi.org/10.1103/PhysRevLett.129.230504