SCALING UNIVERSALITY: A FRAMEWORK FOR FUNCTION REPRESENTATIONS

Anonymous authors
Paper under double-blind review

ABSTRACT

A fundamental question in function approximation theory is whether physical operator scale bounds permanently restrict representational capacity or merely reparameterize it. Standard universal approximation theorems evaluate function classes under fixed input and output conventions, leaving a theoretical gap for scale-bounded computational substrates even when target functions remain fully recoverable via output rescaling. We resolve this gap by introducing scaling universality, an algebraic framework that incorporates input dynamic ranges and output scale bounds directly into the representation criterion. The central theoretical insight is the formulation of an output scale threshold (τ): rather than requiring representation at a fixed output scale, scaling universality requires representations to hold across all output scales exceeding τ. This threshold requirement elevates output scaling from an ad-hoc heuristic to a closed algebraic framework, guaranteeing structural closure and representation existence for both exact and approximate function classes under linear precomposition, output scaling, addition, and composition across operational depth. To demonstrate the framework in the norm-bounded regime of quantum computing, we instantiate scaling universality via Generalized Quantum Signal Processing (GQSP), proving exact scaling universality for real polynomials on $[-1, 1]$ via real-valued GQSP primitives and deriving the analytical normalization threshold governing this space. Extending to multivariate domains, we combine GQSP primitives with Linear Combinations of Unitaries (LCU) and a scale-compatible recursive superposition architecture to prove approximate scaling universality for continuous functions $C(\mathbb{R}^n)$ on $[-1, 1]^n$. This work establishes the representational completeness and algebraic closure of scale-bounded primitives, providing a rigorous theoretical foundation for functional representation decoupled from downstream resource trade-offs such as sampling complexity, compilation depth, and parameter optimization.

1 INTRODUCTION

A fundamental question in function approximation theory is whether strict operator bounds permanently restrict representation capacity or merely reparameterize it. In many physical computational substrates—from digital architectures with finite numerical dynamic ranges (Jacob et al., 2018) to unitarily constrained quantum systems (Motlagh & Wiebe, 2024)—the underlying mathematical operations enforce strict norm bounds on admissible input domains and output amplitudes. Standard universal approximation theorems (Cybenko, 1989; Hornik et al., 1989) evaluate expressivity under fixed scale conventions, relying implicitly on unbounded output-layer combinations. Consequently, a bounded functional class may appear strictly limited under a fixed output scale convention, even when target functions remain fully recoverable through appropriate input and output rescaling.

In practice, scaling is often treated via localized normalization—such as fixed-point shifts (Higham, 2002), gain attenuation (Manolakis & Proakis, 1996), or sub-normalization in quantum block-encodings (Martyn et al., 2021). While these procedures map individual target functions onto bounded primitives, they treat scaling as a static operational step rather than a structural property of the function class itself. Lacking a foundational framework, it has remained open whether scale-constrained representation classes satisfy formal structural closure under recursive composition while preserving domain and range compatibility across operational bounds.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

To resolve this gap, we introduce scaling universality—a mathematical framework that incorporates input dynamic ranges and output scale bounds directly into the representation criterion. The central theoretical insight is the formulation of an output scale threshold (τ). Rather than requiring representation at a single fixed output scale, scaling universality requires representation for every output scale exceeding τ. This threshold requirement elevates scaling from a static heuristic to a closed algebraic framework: we establish structural closure properties for both exact and approximate scaling universality under linear precomposition, output scaling, addition, and composition, subject to corresponding domain and range compatibility conditions. These results provide explicit algebraic rules to track threshold evolution and guarantee that deep architectures built from scale-bounded primitives retain representation capacity across recursive compositions.

To demonstrate the framework on a strictly norm-bounded setting, we instantiate scaling universality in quantum function processing. Using real-valued Generalized Quantum Signal Processing (GQSP) representations (Motlagh & Wiebe, 2024) as a primitive block, we establish exact scaling universality for real polynomials on $[-1, 1]$ and derive the analytical normalization threshold governing this space. Combining this primitive with Linear Combinations of Unitaries (LCU) and a scale-compatible recursive superposition architecture, we prove approximate scaling universality for multivariate continuous functions $C(\mathbb{R}^n)$ on $[-1, 1]^n$.

By establishing scaling universality as an algebraic guarantee, this work decouples foundational representational capability from runtime execution conventions. Crucially, our framework isolates functional existence from downstream operational resource allocation: while output rescaling bridges norm-bounded primitives and target magnitudes, it isolates expressivity bounds from backend resource costs such as measurement shot noise, compilation depth, or parameter optimization. The resulting theory provides a rigorous mathematical basis for functional representation under bounded operational primitives.

2 Scaling Universality

2.1 Scaling Universality Framework

We introduce scaling universality to distinguish representation-dependent constraints from intrinsic representational expressivity. Input scaling accommodates differences in input domains, while output scaling accounts for normalization and magnitude constraints. Allowing the input and output scales to vary separately for each target function, however, would not generally yield a common output scale that can be propagated through addition and composition. We therefore require the correspondence for arbitrary input scales and all sufficiently large output scales, so that a common sufficiently large scale can be chosen when combining representations through these operations.

Definition 2.1 (Common domain of a function class)**.** Let $\mathcal{F}$ be a class of functions. The common domain of $\mathcal{F}$ is defined by

$$\mathrm{dom}(\mathcal{F}) := \bigcap_{f \in \mathcal{F}} \mathrm{dom}(f). \tag{1}$$

Definition 2.2 (Input-scaling admissibility)**.** Let $\mathcal{F}$ be a function class and let D be a set. We say that $\mathcal{F}$ is input-scaling admissible on D if, for every $\alpha \in \mathbb{R}$,

$$\alpha D \subseteq \mathrm{dom}(\mathcal{F}). \tag{2}$$

Definition 2.3 (Exact scaling universality)**.** Let $\mathcal{F}$ and $\mathcal{G}$ be function classes, and let D be a set such that $\mathcal{F}$ is input-scaling admissible on D. We say that $\mathcal{G}$ is exactly scaling-universal for $\mathcal{F}$ on D if, for every $f \in \mathcal{F}$ and every $\alpha \in \mathbb{R}$, there exists $\tau \geq 0$ such that, for every $|\beta| > \tau$, there exists $g \in \mathcal{G}$ such that $D \subseteq \mathrm{dom}(g)$ and

$$\beta g(x) = f(\alpha x) \tag{3}$$

for all $x \in D$.

Remark. While positive output scales are standard in practice, we allow signed output scales to maintain algebraic symmetry under sign inversion. We emphasize that β serves as a classical mathematical reconstruction factor—such as a readout multiplier or statistical sampling denominator—used to recover the target function f, rather than an unconstrained physical signal amplifier.

Definition 2.4 (Uniform norm)**.** Let D be a set and let h be a function satisfying $D \subseteq \mathrm{dom}(h)$. The uniform norm of h on D is defined by

$$\|h\|_D := \sup_{x \in D} |h(x)|. \tag{4}$$

Definition 2.5 (Approximate scaling universality)**.** Let $\mathcal{F}$ and $\mathcal{G}$ be function classes, and let D be a set such that $\mathcal{F}$ is input-scaling admissible on D. We say that $\mathcal{G}$ is approximately scaling-universal for $\mathcal{F}$ on D if, for every $f \in \mathcal{F}$, every $\alpha \in \mathbb{R}$, and every $\varepsilon > 0$, there exists $\tau \geq 0$ such that, for every $|\beta| > \tau$, there exists $g \in \mathcal{G}$ such that $D \subseteq \mathrm{dom}(g)$ and

$$\|\beta g - f(\alpha \cdot)\|_D < \varepsilon. \tag{5}$$

Proposition 2.6 (Inclusion monotonicity)**.** *Let $\mathcal{F}_1$, $\mathcal{F}_2$, $\mathcal{G}_1$, and $\mathcal{G}_2$ be function classes such that $\mathcal{F}_2 \subseteq \mathcal{F}_1$ and $\mathcal{G}_1 \subseteq \mathcal{G}_2$, and let D be a set. If $\mathcal{G}_1$ is exactly scaling-universal for $\mathcal{F}_1$ on D, then $\mathcal{G}_2$ is exactly scaling-universal for $\mathcal{F}_2$ on D. If $\mathcal{G}_1$ is approximately scaling-universal for $\mathcal{F}_1$ on D, then $\mathcal{G}_2$ is approximately scaling-universal for $\mathcal{F}_2$ on D.*

2.2 Elementary Operations

The main architectural role of scaling universality is its stability under the elementary operations used to construct function representations. We establish closure under four core structural operations—linear precomposition, scaling, addition, and composition—which together form the inductive building blocks for complex recursive architectures.

Definition 2.7 (Image of a function class)**.** Let $\mathcal{F}$ be a class of functions and let D be a set. The image of $\mathcal{F}$ on D is defined by

$$\mathcal{F}(D) := \bigcup_{f \in \mathcal{F}} f(D \cap \mathrm{dom}(f)). \tag{6}$$

Definition 2.8 (Input-scaling invariance)**.** A function class $\mathcal{F}$ is called input-scaling-invariant if, for every $f \in \mathcal{F}$ and every $\alpha \in \mathbb{R}$, the function $f(\alpha \cdot)$ belongs to $\mathcal{F}$.

Lemma 2.9 (Linear precomposition closure)**.** *Let $\mathcal{F}$ and $\mathcal{G}$ be function classes, let π be a linear map, and let D_1 and D_2 be sets such that $\pi(D_1) \subseteq D_2$. If $\mathcal{G}$ is exactly scaling-universal for $\mathcal{F}$ on D_2, then $\mathcal{G} \circ \pi$ is exactly scaling-universal for $\mathcal{F} \circ \pi$ on D_1. If $\mathcal{G}$ is approximately scaling-universal for $\mathcal{F}$ on D_2, then $\mathcal{G} \circ \pi$ is approximately scaling-universal for $\mathcal{F} \circ \pi$ on D_1.*

Lemma 2.10 (Scaling closure)**.** *Let $\mathcal{F}$ and $\mathcal{G}$ be function classes, let D be a set, and let $a, b \in \mathbb{R} \setminus \{0\}$. If $\mathcal{G}$ is exactly scaling-universal for $\mathcal{F}$ on D, then $b\mathcal{G}$ is exactly scaling-universal for $a\mathcal{F}$ on D. If $\mathcal{G}$ is approximately scaling-universal for $\mathcal{F}$ on D, then $b\mathcal{G}$ is approximately scaling-universal for $a\mathcal{F}$ on D.*

Lemma 2.11 (Additivity closure)**.** *Let $\mathcal{F}_1, \mathcal{G}_1, \mathcal{F}_2, \mathcal{G}_2$ be function classes, and let D be a set. If $\mathcal{G}_1$ is exactly scaling-universal for $\mathcal{F}_1$ on D, and $\mathcal{G}_2$ is exactly scaling-universal for $\mathcal{F}_2$ on D, then $\mathcal{G}_1 + \mathcal{G}_2$ is exactly scaling-universal for $\mathcal{F}_1 + \mathcal{F}_2$ on D. If $\mathcal{G}_1$ is approximately scaling-universal for $\mathcal{F}_1$ on D, and $\mathcal{G}_2$ is approximately scaling-universal for $\mathcal{F}_2$ on D, then $\mathcal{G}_1 + \mathcal{G}_2$ is approximately scaling-universal for $\mathcal{F}_1 + \mathcal{F}_2$ on D.*

Lemma 2.12 (Exact composition closure)**.** *Let $\mathcal{F}_1, \mathcal{G}_1, \mathcal{F}_2, \mathcal{G}_2$ be function classes, and let D_1 and D_2 be sets such that $\mathcal{G}_1(D_1) \subseteq D_2$. If $\mathcal{F}_2$ is input-scaling-invariant, $\mathcal{G}_1$ is exactly scaling-universal for $\mathcal{F}_1$ on D_1, and $\mathcal{G}_2$ is exactly scaling-universal for $\mathcal{F}_2$ on D_2, then $\mathcal{G}_2 \circ \mathcal{G}_1$ is exactly scaling-universal for $\mathcal{F}_2 \circ \mathcal{F}_1$ on D_1.*

Lemma 2.13 (Approximate composition closure)**.** *Let $\mathcal{F}_1, \mathcal{G}_1, \mathcal{F}_2, \mathcal{G}_2$ be function classes, and let D_1 be a compact set and D_2 be a set such that $\mathcal{G}_1(D_1) \subseteq D_2$ and $\mathcal{F}_1(D_1) \subseteq \mathrm{int}(\mathrm{dom}(\mathcal{F}_2))$. If $\mathcal{F}_1$ and $\mathcal{F}_2$ are input-scaling-invariant, every function in $\mathcal{F}_1$ and $\mathcal{F}_2$ is continuous on its domain, $\mathcal{G}_1$ is approximately scaling-universal for $\mathcal{F}_1$ on D_1, and $\mathcal{G}_2$ is approximately scaling-universal for $\mathcal{F}_2$ on D_2, then $\mathcal{G}_2 \circ \mathcal{G}_1$ is approximately scaling-universal for $\mathcal{F}_2 \circ \mathcal{F}_1$ on D_1.*

Remark. The input-scaling-invariance assumption on $\mathcal{F}_2$ is stronger than strictly necessary in Lemma 2.13. We impose input-scaling invariance on both $\mathcal{F}_1$ and $\mathcal{F}_2$ because it gives a symmetric and simpler statement and is satisfied by the function classes considered in this work.

2.3 SUPERPOSITION ARCHITECTURE

We next extend scaling universality to recursively constructed superposition architectures.

Definition 2.14 (Superposition architecture)**.** A superposition architecture $\mathfrak{A}$ with input maps $\{\pi_j\}_{j=1}^n$ is recursively generated by

$$\mathfrak{A} ::= \square \circ \pi_j \mid c\mathfrak{C} \mid \mathfrak{C}_1 + \mathfrak{C}_2 \mid \square \circ \mathfrak{C}, \tag{7}$$

where $j \in \{1, \dots, n\}$, $c \in \mathbb{R} \setminus \{0\}$, and $\mathfrak{C}$, $\mathfrak{C}_1$ and $\mathfrak{C}_2$ are superposition architectures with the same input maps as $\mathfrak{A}$. For a function class $\mathcal{F}$, let $\mathfrak{A}[\mathcal{F}]$ denote the function class obtained by instantiating each occurrence of $\square$ with a function from $\mathcal{F}$.

Definition 2.15 (Superposition architecture equivalence)**.** Let $\mathfrak{A}$ and $\mathfrak{B}$ be superposition architectures with the same input maps. We say that $\mathfrak{A}$ and $\mathfrak{B}$ are equivalent, denoted by $\mathfrak{A} \sim \mathfrak{B}$, if they have the same recursive structure, with corresponding scaling nodes differing only by nonzero scalar factors.

Definition 2.16 (Pre-composition subarchitectures)**.** Let $\mathfrak{A}$ be a superposition architecture. Define

$$\mathrm{Pre}(\mathfrak{A}) := \{\mathfrak{P} \mid \square \circ \mathfrak{P} \in \mathrm{Sub}(\mathfrak{A})\}, \tag{8}$$

where $\mathrm{Sub}(\mathfrak{A})$ denotes the set of subarchitectures of $\mathfrak{A}$.

Theorem 2.17 (Exact superposition closure)**.** *Let $\mathcal{F}$ and $\mathcal{G}$ be function classes, let $\mathfrak{A}$ and $\mathfrak{B}$ be superposition architectures with linear input maps $\{\pi_j\}_{j=1}^n$ such that $\mathfrak{A} \sim \mathfrak{B}$, and let D_1 and D_2 be sets such that*

$$\bigcup_{\mathfrak{P} \in \mathrm{Pre}(\mathfrak{A})} \mathfrak{P}[\mathcal{G}](D_1) \subseteq D_2. \tag{9}$$

If $\mathcal{F}$ is input-scaling-invariant and $\mathcal{G}$ is exactly scaling-universal for $\mathcal{F}$ on D_2, then $\mathfrak{B}[\mathcal{G}]$ is exactly scaling-universal for $\mathfrak{A}[\mathcal{F}]$ on D_1.

Proof. We proceed by structural induction on the common recursive structure of $\mathfrak{A}$ and $\mathfrak{B}$. The base case follows from Lemma 2.9. The scaling, addition, and composition cases follow respectively from Lemma 2.10, Lemma 2.11, and Lemma 2.12. □

Theorem 2.18 (Approximate superposition closure)**.** *Let $\mathcal{F}$ and $\mathcal{G}$ be function classes, let $\mathfrak{A}$ and $\mathfrak{B}$ be superposition architectures with linear input maps $\{\pi_j\}_{j=1}^n$ such that $\mathfrak{A} \sim \mathfrak{B}$, and let D_1 and D_2 be compact sets such that*

$$\bigcup_{\mathfrak{P} \in \mathrm{Pre}(\mathfrak{A})} \mathfrak{P}[\mathcal{G}](D_1) \subseteq D_2, \tag{10}$$

and

$$\bigcup_{\mathfrak{P} \in \mathrm{Pre}(\mathfrak{A})} \mathfrak{P}[\mathcal{F}](D_1) \subseteq \mathrm{int}(\mathrm{dom}(\mathcal{F})). \tag{11}$$

If $\mathcal{F}$ is input-scaling-invariant, every function in $\mathcal{F}$ is continuous on its domain, and $\mathcal{G}$ is approximately scaling-universal for $\mathcal{F}$ on D_2, then $\mathfrak{B}[\mathcal{G}]$ is approximately scaling-universal for $\mathfrak{A}[\mathcal{F}]$ on D_1.

Proof. We proceed by structural induction on the common recursive structure of $\mathfrak{A}$ and $\mathfrak{B}$. The base case follows from Lemma 2.9. The scaling and addition cases follow from Lemma 2.10 and Lemma 2.11, respectively. For a composition node, the two range conditions in the hypothesis ensure that the realizable and target ranges of the pre-composition subarchitecture remain in the domains required by the outer primitive. Since $\mathcal{F}$ is input-scaling-invariant and its functions are continuous, the hypotheses of Lemma 2.13 are therefore satisfied, and the composition case follows. Hence the result holds for the entire recursive architecture. □

Together, Theorem 2.17 and Theorem 2.18 establish the architecture-level principle underlying our framework: once scaling universality is established for the primitive function classes, it propagates whenever the corresponding elementary operations satisfy the closure conditions.

3 QUANTUM POLYNOMIAL TRANSFORMATIONS

To show that scale-constrained quantum primitives satisfy our framework, we instantiate scaling universality using real-valued GQSP representations targeting polynomials over the interval $[-1, 1]$.

3.1 ENCODING

We use projected unitary encodings to represent contractions as matrix elements of unitary operators.

Definition 3.1 (Projected unitary encoding)**.** Let $A \in \mathbb{C}^{N_L \times N_R}$ satisfy $\|A\|_2 \leq 1$. Given a unitary operator $U \in \mathbb{C}^{M \times M}$ and isometries $\Pi_L \in \mathbb{C}^{M \times N_L}$ and $\Pi_R \in \mathbb{C}^{M \times N_R}$, the tuple (U, Π_L, Π_R) is referred to as a projected unitary encoding of A if

$$\Pi_L^\dagger U \Pi_R = A. \tag{12}$$

Definition 3.2 (Hermitian projected unitary encoding)**.** Let $A \in \mathbb{C}^{N \times N}$ be Hermitian. Given a unitary operator $U \in \mathbb{C}^{M \times M}$ and an isometry $\Pi \in \mathbb{C}^{M \times N}$, the tuple (U, Π) is referred to as a Hermitian projected unitary encoding of A if U is a Hermitian unitary operator and (U, Π, Π) forms a projected unitary encoding of A.

For any $x \in [-1, 1]$, consider the Hermitian unitary

$$U(x) := \begin{bmatrix} x & \sqrt{1-x^2} \\ \sqrt{1-x^2} & -x \end{bmatrix}. \tag{13}$$

It satisfies

$$\langle 0|U(x)|0\rangle = x. \tag{14}$$

Therefore, $(U(x), |0\rangle)$ forms a Hermitian projected unitary encoding of the scalar x.

The real part of the encoded scalar can, in principle, be estimated by the Hadamard test (Cleve et al., 1998).

3.2 POLYNOMIAL CONSTRUCTION

Given a Hermitian projected unitary encoding $(U, |\psi\rangle)$ of a scalar x, we construct a new projected unitary encoding whose encoded value is a polynomial transformation of x.

Definition 3.3 (Generalized single-qubit rotation operator)**.** For real parameters $\theta, \phi, \lambda \in \mathbb{R}$, the generalized single-qubit rotation operator is defined as

$$R(\theta, \phi, \lambda) := \begin{bmatrix} e^{i(\lambda+\phi)} \cos\theta & e^{i\phi} \sin\theta \\ e^{i\lambda} \sin\theta & -\cos\theta \end{bmatrix}. \tag{15}$$

Definition 3.4 (GQSP operator)**.** Let U be a unitary operator, $d \in \mathbb{N}$, and let $\boldsymbol{\theta} := (\theta_0, \ldots, \theta_d) \in \mathbb{R}^{d+1}$, $\boldsymbol{\phi} := (\phi_0, \ldots, \phi_d) \in \mathbb{R}^{d+1}$, and $\lambda \in \mathbb{R}$. The GQSP operator is defined as

$$Q(U; \boldsymbol{\theta}, \boldsymbol{\phi}, \lambda) := \left(\prod_{j=1}^{d} (R(\theta_j, \phi_j, 0) \otimes I)(|0\rangle\langle 0| \otimes U + |1\rangle\langle 1| \otimes I) \right) (R(\theta_0, \phi_0, \lambda) \otimes I). \tag{16}$$

Definition 3.5 (Reflection operator)**.** The reflection operator R_ψ about the state $|\psi\rangle$ is defined as

$$R_\psi := 2|\psi\rangle\langle\psi| - I. \tag{17}$$

Lemma 3.6 (Hermitian involution)**.** *Let $(U, |\psi\rangle)$ be a Hermitian projected unitary encoding of the real scalar $x \in (-1, 1)$. Define*

$$|\psi^\perp\rangle := \frac{U|\psi\rangle - x|\psi\rangle}{\sqrt{1-x^2}}. \tag{18}$$

Then, $|\psi^\perp\rangle$ is a unit vector orthogonal to $|\psi\rangle$, and the two-dimensional subspace $\mathcal{S} = \mathrm{span}\{|\psi\rangle, |\psi^\perp\rangle\}$ is invariant under U. Moreover, with respect to the basis $\{|\psi\rangle, |\psi^\perp\rangle\}$, the restriction of U to $\mathcal{S}$ is

$$U|_{\mathcal{S}} = \begin{bmatrix} x & \sqrt{1-x^2} \\ \sqrt{1-x^2} & -x \end{bmatrix}. \tag{19}$$

Lemma 3.7 (Qubitized walk)**.** *Let $(U, |\psi\rangle)$ be a Hermitian projected unitary encoding of a real scalar $x \in (-1, 1)$. Define the qubitized walk operator*

$$W := R_\psi U. \tag{20}$$

Let

$$|\psi^\perp\rangle := \frac{U|\psi\rangle - x|\psi\rangle}{\sqrt{1-x^2}} \tag{21}$$

and

$$|\Psi_\pm\rangle := \frac{|\psi\rangle \pm i|\psi^\perp\rangle}{\sqrt{2}}. \tag{22}$$

Then, $|\Psi_+\rangle$ and $|\Psi_-\rangle$ are orthonormal eigenvectors of W with corresponding eigenvalues $e^{\pm i \arccos x}$.

Definition 3.8 (Admissible GQSP polynomial class)**.** Let $d \geq 0$. Define the admissible GQSP polynomial class of degree at most d by

$$\mathcal{A}_d := \left\{ P \;\middle|\; \deg(P) \leq d,\, \forall \theta \in \mathbb{R},\, |P(e^{i\theta})| \leq 1 \right\}. \tag{23}$$

Lemma 3.9 (Spectral representation)**.** *Let $(U, |\psi\rangle)$ be a Hermitian projected unitary encoding of the real scalar $x \in [-1, 1]$, and let*

$$W := R_\psi U. \tag{24}$$

Then, for every $P \in \mathcal{A}_d$,

$$\langle\psi| P(W) |\psi\rangle = \frac{P(e^{i \arccos x}) + P(e^{-i \arccos x})}{2}. \tag{25}$$

Definition 3.10 (Realizable polynomial class)**.** Let $d \geq 0$. Define the realizable polynomial class of degree at most d by

$$\mathcal{R}_d := \left\{ q \;\middle|\; \exists P \in \mathcal{A}_d,\, \forall x \in [-1, 1],\, q(x) = \frac{P(e^{i \arccos x}) + P(e^{-i \arccos x})}{2} \right\}. \tag{26}$$

Theorem 3.11 (Polynomial realizability characterization)**.** *Let $(U(x), |\psi\rangle)$ be a family of Hermitian projected unitary encodings of the real scalar $x \in [-1, 1]$, and define the corresponding qubitized walk operator*

$$W(x) := R_\psi U(x). \tag{27}$$

A polynomial p belongs to $\mathcal{R}_d$ if and only if there exist parameters $\boldsymbol{\theta}, \boldsymbol{\phi} \in \mathbb{R}^{d+1}$ and $\lambda \in \mathbb{R}$ such that $(Q(W(x); \boldsymbol{\theta}, \boldsymbol{\phi}, \lambda), |0\rangle \otimes |\psi\rangle, |0\rangle \otimes |\psi\rangle)$ is a projected unitary encoding of $p(x)$ for all $x \in [-1, 1]$.

Proof Sketch. If $p \in \mathcal{R}_d$, the GQSP theorem (Motlagh & Wiebe, 2024) guarantees phase parameters $\boldsymbol{\theta}, \boldsymbol{\phi}, \lambda$ that construct an admissible polynomial $P(W(x))$ in the upper-left block. Projecting with state $|\psi\rangle$ applies Lemma 3.9 to yield $p(x)$. Conversely, taking the projected unitary encoding and applying GQSP in reverse constructs the corresponding admissible polynomial $P \in \mathcal{A}_d$, placing $p \in \mathcal{R}_d$. □

Theorem 3.11 gives an exact characterization of the real polynomial class realizable by the quantum construction considered here. The construction serves as a proof of quantum realizability and is not intended to be resource-optimal.

3.3 Polynomial Normalization

GQSP imposes a unit-circle normalization constraint on the associated complex polynomial. We characterize how this constraint translates to real polynomial representations on $[-1, 1]$ and derive the resulting exact normalization threshold.

Definition 3.12 (Real polynomial space)**.** Let $d \geq 0$. Define the real polynomial space of degree at most d by

$$\mathcal{P}_d := \{ p : \mathbb{R} \to \mathbb{R} \mid \deg(p) \leq d \}. \tag{28}$$

For arbitrary finite degree, define

$$\mathcal{P}_\infty := \bigcup_{d=0}^{\infty} \mathcal{P}_d. \tag{29}$$

Since GQSP encodings yield complex polynomials $q \in \mathcal{R}_d$ in general, we restrict our attention to $\mathrm{Re}(\mathcal{R}_d)$, which corresponds directly to real-valued observable readouts (e.g., via the Hadamard test).

Definition 3.13 (Real realizable polynomial class)**.** Let $d \geq 0$. Define the real realizable polynomial class of degree at most d by

$$\mathrm{Re}(\mathcal{R}_d) := \{r \mid \exists q \in \mathcal{R}_d,\, \forall x \in [-1,1],\, r(x) = \mathrm{Re}(q(x))\}\,. \tag{30}$$

For arbitrary finite degree, define

$$\mathrm{Re}(\mathcal{R}_\infty) := \bigcup_{d=0}^{\infty} \mathrm{Re}(\mathcal{R}_d). \tag{31}$$

Proposition 3.14 (Real-part closure)**.** *Let $d \geq 0$. Then,*

$$\mathrm{Re}(\mathcal{R}_d) \subseteq \mathcal{R}_d. \tag{32}$$

Definition 3.15 (Associated polynomial)**.** Let $d \geq 0$ and $p \in \mathcal{P}_d$ have Chebyshev expansion

$$p(x) = \sum_{j=0}^{d} a_j T_j(x). \tag{33}$$

The associated polynomial of p is defined by

$$P(z) := \sum_{j=0}^{d} a_j z^j. \tag{34}$$

Lemma 3.16 (Associated polynomial characterization)**.** *Let $d \geq 0$, $p \in \mathcal{P}_d$, and P be a polynomial of degree at most d. Then, P is the associated polynomial of p if and only if*

$$\frac{P(e^{i \arccos x}) + P(e^{-i \arccos x})}{2} = p(x) \tag{35}$$

for all $x \in [-1,1]$.

Definition 3.17 (Normalization threshold)**.** Let $d \geq 0$ and $p \in \mathcal{P}_d$ have associated polynomial P. The normalization threshold of p is defined by

$$\Gamma(p) := \max_{|z|=1} |P(z)|. \tag{36}$$

Proposition 3.18 (Exact normalization threshold)**.** *Let $d \geq 0$, $p \in \mathcal{P}_d$, and $\beta \in \mathbb{R} \setminus \{0\}$. Then,*

$$\frac{p}{\beta} \in \mathrm{Re}(\mathcal{R}_d) \tag{37}$$

if and only if

$$|\beta| \geq \Gamma(p). \tag{38}$$

3.4 Polynomial Scaling Universality

The normalization result above can be combined with arbitrary input scaling to compare the realizable polynomial class with the full real polynomial space under scaling universality. In particular, these representation-dependent normalization constraints do not restrict the finite-degree real polynomial function class under scaling universality.

Theorem 3.19 (Finite-degree polynomial scaling universality)**.** *Let $d \geq 0$. Then, $\mathrm{Re}(\mathcal{R}_d)$ is exact scaling-universal for $\mathcal{P}_d$ on $[-1,1]$.*

Proof Sketch. Given $p \in \mathcal{P}_d$ and $\alpha \in \mathbb{R}$, the rescaled function $\widetilde{p}(x) := p(\alpha x)$ preserves degree bounds, ensuring $\widetilde{p} \in \mathcal{P}_d$. Applying the normalization threshold from Proposition 3.18 allows us to scale $\widetilde{p}$ into $\mathrm{Re}(\mathcal{R}_d)$ for any $|\beta| > \Gamma(\widetilde{p})$, satisfying exact scaling universality. □

Corollary 3.20 (Arbitrary-degree polynomial scaling universality)**.** $\mathrm{Re}(\mathcal{R}_\infty)$ *is exact scaling-universal for $\mathcal{P}_\infty$ on $[-1,1]$.*

Theorem 3.21 (Continuous-function scaling universality)**.** $\mathrm{Re}(\mathcal{R}_\infty)$ *is approximately scaling-universal for* $C(\mathbb{R})$ *on* $[-1, 1]$.

Proof Sketch. Given $f \in C(\mathbb{R})$, $\alpha \in \mathbb{R}$, and $\varepsilon > 0$, we use the Weierstrass theorem (Rudin, 1976) to approximate f on $[-|\alpha|, |\alpha|]$ with a polynomial $p \in \mathcal{P}_\infty$. Applying exact scaling universality for polynomials (Corollary 3.20) yields $g \in \mathrm{Re}(\mathcal{R}_\infty)$ such that $\beta g(x) = p(\alpha x)$ for large $|\beta|$, bounded by $\|\beta g - f(\alpha\cdot)\|_{[-1,1]} < \varepsilon$. □

4 Multivariate Quantum Superposition Architectures

We now extend single-variable real polynomial representations to multivariate continuous functions.

4.1 Superposition Construction

To bridge our abstract algebraic closure results with physical implementations, we provide explicit quantum circuit constructions for multivariate superposition architectures. In our framework, superposition nodes rely on scalar addition up to a non-zero scaling factor, which can be implemented naturally via an equal-weight Linear Combination of Unitaries (LCU) construction (Childs & Wiebe, 2012).

Let $\{U_i\}_{i=0}^{2^m-1}$ be a collection of unitary operators sharing a common state $|\psi\rangle$, such that each tuple $(U_i, |\psi\rangle, |\psi\rangle)$ constitutes a projected unitary encoding of a scalar x_i. Define

$$V_L := (H^{\otimes m} \otimes I) \left(\sum_{i=0}^{2^m-1} |i\rangle\langle i| \otimes U_i \right) (H^{\otimes m} \otimes I). \tag{39}$$

Then, the tuple $(V_L, |0\rangle^{\otimes m} \otimes |\psi\rangle, |0\rangle^{\otimes m} \otimes |\psi\rangle)$ is a projected unitary encoding of

$$y := \frac{1}{2^m} \sum_{i=0}^{2^m-1} x_i, \tag{40}$$

so the LCU realizes the scaled addition node with the nonzero factor 2^{-m} imposed by normalization.

Since the resulting unitary need not be Hermitian, we convert it to a Hermitian projected unitary encoding. Suppose that $(U, |\psi\rangle, |\psi\rangle)$ is a projected unitary encoding of $x \in \mathbb{C}$. Define

$$V_H := (H \otimes I)(|0\rangle\langle 0| \otimes U + |1\rangle\langle 1| \otimes U^\dagger)(XH \otimes I). \tag{41}$$

Then, $(V_H, |0\rangle \otimes |\psi\rangle)$ is a Hermitian projected unitary encoding of $\mathrm{Re}(x)$.

Applied recursively, these constructions realize the elementary operations required by the superposition architectures, with necessary scaling provided by the LCU addition nodes. These circuits establish the quantum realizability of the corresponding recursive representations. Note that this architecture does not require a single, fully coherent quantum circuit; intermediate states may be measured, classically processed, and re-encoded into subsequent blocks. Implementation overheads—such as measurement shot counts, classical latency, and state preparation—are beyond the scope of this representational analysis.

4.2 Superposition Scaling Universality

The finite-degree polynomial result can now be propagated from the primitive representation to compatible superposition architectures.

Proposition 4.1 (Superposition polynomial scaling universality)**.** *Let* $d \geq 0$*, and let* $\mathfrak{A}$ *and* $\mathfrak{B}$ *be superposition architectures with linear input maps* $\{\pi_j\}_{j=1}^n$ *satisfying* $\mathfrak{A} \sim \mathfrak{B}$ *and*

$$\bigcup_{\mathfrak{P} \in \mathrm{Pre}(\mathfrak{A})} \mathfrak{P}[\mathrm{Re}(\mathcal{R}_d)]([-1, 1]^n) \subseteq [-1, 1]. \tag{42}$$

Then, $\mathfrak{B}[\mathrm{Re}(\mathcal{R}_d)]$ *is exactly scaling-universal for* $\mathfrak{A}[\mathcal{P}_d]$ *on* $[-1, 1]^n$.

Proof. By Theorem 3.19, $\mathrm{Re}(\mathcal{R}_d)$ is exactly scaling-universal for $\mathcal{P}_d$ on $[-1,1]$. Since $\mathcal{P}_d$ is input-scaling-invariant, Theorem 2.17 implies that $\mathfrak{B}[\mathrm{Re}(\mathcal{R}_d)]$ is exactly scaling-universal for $\mathfrak{A}[\mathcal{P}_d]$ on $[-1,1]^n$. □

Theorem 4.2 (Multivariate bounded continuous-function scaling universality)**.** *Let $n \geq 2$. There exists an architecture $\mathfrak{A}_n$ such that $\mathfrak{A}_n[\mathrm{Re}(\mathcal{R}_\infty)]$ is approximately scaling-universal for $C_b(\mathbb{R}^n)$ on $[-1,1]^n$.*

Proof. By the Laczkovich superposition theorem (Laczkovich, 2021), there exists an architecture $\mathfrak{A}_n$ with coordinate projections as input maps such that

$$C_b(\mathbb{R}^n) \subseteq \mathfrak{A}_n[C(\mathbb{R})]. \tag{43}$$

By Theorem 3.21 and Theorem 2.18, $\mathfrak{A}_n[\mathrm{Re}(\mathcal{R}_\infty)]$ is approximately scaling-universal for $\mathfrak{A}_n[C(\mathbb{R})]$ on $[-1,1]^n$. By Proposition 2.6, $\mathfrak{A}_n[\mathrm{Re}(\mathcal{R}_\infty)]$ is approximately scaling-universal for $C_b(\mathbb{R}^n)$ on $[-1,1]^n$. □

Remark. Alternative superposition representations—such as those by Kolmogorov (1961), Sprecher (1965), or Doss (1977)—do not generally satisfy the hypotheses of Theorem 2.18. In our construction, primitive functions in $\mathrm{Re}(\mathcal{R}_\infty)$ map $[-1,1]$ into $[-1,1]$. Out-of-domain outputs from addition nodes are rescaled back into $[-1,1]$ via internal scaling factors prior to evaluation, ensuring domain compatibility.

Corollary 4.3 (Multivariate continuous-function scaling universality)**.** *Let $n \geq 2$. There exists a superposition architecture $\mathfrak{A}_n$ such that $\mathfrak{A}_n[\mathrm{Re}(\mathcal{R}_\infty)]$ is approximately scaling-universal for $C(\mathbb{R}^n)$ on $[-1,1]^n$.*

5 Discussion

The primary utility of the scaling threshold τ lies in its structural propagation across recursive architectures. By requiring functional preservation for all output scales $|\beta| > \tau$, scaling universality provides a hardware-agnostic abstraction layer guaranteeing structural closure under composition without violating domain or operational norm bounds. Crucially, this framework isolates representational capability from operational execution and optimization. In line with foundational approximation theory (Cybenko, 1989; Hornik et al., 1989), we establish existence guarantees and threshold tracking rules independently of parameter training, gradient landscapes, or backend compilation overheads (Low & Chuang, 2017).

By decoupling functional expressivity from physical execution, our threshold tracking rules provide an algebraic foundation to quantify downstream operational trade-offs—such as signal-to-noise ratios in analog processing (Sarpeshkar, 1998) or measurement sampling overheads under quantum sub-normalization (Martyn et al., 2021; Gilyén et al., 2019). This foundational layer opens clear avenues for downstream compilation, including platform-specific parameter synthesis for high-degree multivariate systems (Chao et al., 2020; Dong et al., 2021) and the extension of algebraic closure rules to stochastic physical noise and dynamic range drift.

6 Conclusion

This work introduces scaling universality, bridging universal approximation theory and scale-bounded computational substrates. By formulating an output scale threshold τ, we prove algebraic closure for exact and approximate representations under linear precomposition, output scaling, addition, and composition across operational depth. Instantiating our framework via real-valued GQSP primitives and an LCU-based superposition architecture guarantees exact scaling universality for real polynomials on $[-1,1]$ and approximate scaling universality for continuous functions on $[-1,1]^n$. By decoupling functional expressivity from downstream sampling complexity, compilation depth, and parameter optimization, we establish a rigorous algebraic foundation for scale-bounded function representation, opening new avenues for scale-conscious quantum circuit synthesis and architecture design.

AI USE STATEMENT

We used an AI language model to aid and polish the writing of the manuscript, including improving clarity, concision, grammar, organization, and presentation. We have reviewed all AI-assisted work.

REFERENCES

Rui Chao Chao, Dawei Ding, Andrew Girth, Richard Kueng, Guang Hao Low, and Martin Suchara. Finding angles for quantum signal processing with polynomial root finding. *arXiv preprint arXiv:2003.02831*, 2020.

Andrew M Childs and Nathan Wiebe. Hamiltonian simulation using linear combinations of unitary operations. *arXiv preprint arXiv:1202.5822*, 2012.

Richard Cleve, Artur Ekert, Chiara Macchiavello, and Michele Mosca. Quantum algorithms revisited. *Proceedings: Mathematical, Physical and Engineering Sciences*, pp. 339–354, 1998.

George Cybenko. Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*, 2(4):303–314, 1989.

Yulong Dong, Xiangling Meng, K Birgitta Whaley, and Lin Lin. Efficient phase-factor determination in quantum signal processing. *Physical Review A*, 103(4):042419, 2021.

Raouf Doss. A superposition theorem for unbounded continuous functions. *Transactions of the American mathematical society*, 233:197–203, 1977.

András Gilyén, Yuan Su, Guang Hao Low, and Nathan Wiebe. Quantum singular value transformation and its algorithmic applications. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing (STOC)*, pp. 193–204, 2019.

Nicholas J Higham. *Accuracy and stability of numerical algorithms*. SIAM, 2002.

Kurt Hornik, Maxwell Stinchcombe, and Halbert White. Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5):359–366, 1989.

Benoit Jacob, Skirmantas Kligys, Bo Chen, Menglong Meng, Mark Tang, Andrew Howard, Dmitry Allen, and Dmitry Kalenichenko. Quantization and training of neural networks for efficient integer-arithmetic-only inference. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2704–2713, 2018.

Andrey Nikolaevich Kolmogorov. *On the representation of continuous functions of several variables by superpositions of continuous functions of a smaller number of variables*. American Mathematical Society, 1961.

Miklós Laczkovich. A superposition theorem of kolmogorov type for bounded continuous functions. *Journal of Approximation Theory*, 269:105609, 2021.

Guang Hao Low and Isaac L Chuang. Optimal hamiltonian simulation by quantum signal processing. *Physical review letters*, 118(1):010501, 2017.

Dimitris G Manolakis and John G Proakis. *Digital Signal Processing: Principles, Algorithms, and Applications*. Prentice Hall Englewood Cliffs, NJ, 1996.

John M Martyn, Trevor M Rossi, Andrew K Tan, and Isaac L Chuang. Grand unification of quantum algorithms. *PRX Quantum*, 2(4):040203, 2021.

Danial Motlagh and Nathan Wiebe. Generalized quantum signal processing. *PRX Quantum*, 5(2): 020368, 2024.

James R Munkres. *Topology*. Prentice Hall, 2nd edition, 2000.

Walter Rudin. *Principles of Mathematical Analysis*. McGraw-Hill, New York, 3rd edition, 1976.

Rahul Sarpeshkar. Analog versus digital: Computational efficiency, precision, and dynamic range. *IEEE Micro*, 18(6):40–51, 1998.

David A Sprecher. On the structure of continuous functions of several variables. *Transactions of the American Mathematical Society*, 115:340–355, 1965.

A Overview

This paper introduces a unified theoretical framework for scaling universality and applies it to analyze scale-constrained quantum representation architectures. Our primary contributions are summarized as follows:

1. **Scaling-universality framework.** We formalize the notion of scaling universality to quantify representation capacity when computational primitives operate under explicit output-scale constraints (τ). We prove that scaling universality is closed under linear precomposition, output scaling, addition, and composition, establishing a modular framework for scale-bounded architecture analysis.
2. **Real polynomial quantum representations.** We adapt Generalized Quantum Signal Processing (GQSP)—traditionally formulated for complex polynomials on the unit circle—to construct quantum encodings for real-valued polynomials over $[-1, 1]$. This construction defines the realizable polynomial class $\mathcal{R}_d$ and its real counterpart $\mathrm{Re}(\mathcal{R}_d)$.
3. **Finite-degree exact scaling-universality.** We prove that for any finite degree d, the class $\mathrm{Re}(\mathcal{R}_d)$ is exactly scaling-universal for the space of real polynomials $\mathcal{P}_d$ on $[-1, 1]$.
4. **Univariate continuous-function scaling-universality.** We extend this result to the infinite-degree limit, demonstrating that $\mathrm{Re}(\mathcal{R}_\infty)$ is approximately scaling-universal for univariate continuous functions $C(\mathbb{R})$ on $[-1, 1]$.
5. **Multivariate continuous-function scaling-universality.** By embedding GQSP primitives into an LCU-based superposition architecture $\mathfrak{A}_n$, we prove that $\mathfrak{A}_n[\mathrm{Re}(\mathcal{R}_\infty)]$ is approximately scaling-universal for multivariate continuous functions $C(\mathbb{R}^n)$ on $[-1, 1]^n$.
6. **Quantum realizability and scope.** We establish that scale-constrained quantum operators meet the structural criteria for universal expressivity. Our constructions provide fundamental lower bounds on representational capacity, intentionally decoupled from compilation depth, shot noise, and hardware resource optimization.

B Proofs of Main Results

This section provides the full proofs of the theoretical results stated in the main text.

Proof of Proposition 2.6. We prove the exact and approximate cases separately.

Exact case. Let $f \in \mathcal{F}_2$ and $\alpha \in \mathbb{R}$ be arbitrary. Since $\mathcal{F}_2 \subseteq \mathcal{F}_1$, we have $f \in \mathcal{F}_1$. Since $\mathcal{G}_1$ is exactly scaling-universal for $\mathcal{F}_1$ on D, there exists $\tau \geq 0$ such that, for every $|\beta| > \tau$, there exists $g \in \mathcal{G}_1$ satisfying

$$\beta g(x) = f(\alpha x) \tag{44}$$

for all $x \in D$. Since $\mathcal{G}_1 \subseteq \mathcal{G}_2$, the same g belongs to $\mathcal{G}_2$. Thus, $\mathcal{G}_2$ is exactly scaling-universal for $\mathcal{F}_2$ on D.

Approximate case. Let $f \in \mathcal{F}_2$, $\alpha \in \mathbb{R}$, and $\varepsilon > 0$ be arbitrary. Since $\mathcal{F}_2 \subseteq \mathcal{F}_1$, we have $f \in \mathcal{F}_1$. Since $\mathcal{G}_1$ is approximately scaling-universal for $\mathcal{F}_1$ on D, there exists $\tau \geq 0$ such that, for every $|\beta| > \tau$, there exists $g \in \mathcal{G}_1$ satisfying

$$\|\beta g - f(\alpha \cdot)\|_D < \varepsilon. \tag{45}$$

Since $\mathcal{G}_1 \subseteq \mathcal{G}_2$, the same g belongs to $\mathcal{G}_2$. Thus, $\mathcal{G}_2$ is approximately scaling-universal for $\mathcal{F}_2$ on D. □

Proof of Lemma 2.9. We prove the exact and approximate cases separately.

Exact case. Let $f \in \mathcal{F}$ and $\alpha \in \mathbb{R}$. Since $\mathcal{G}$ is exactly scaling-universal for $\mathcal{F}$ on D_2, there exists $\tau \geq 0$ such that, for every $|\beta| > \tau$, there exists $g \in \mathcal{G}$ satisfying

$$\beta g(y) = f(\alpha y) \tag{46}$$

for all $y \in D_2$. Therefore, for every $|\beta| > \tau$, we have

$$\begin{aligned} \beta(g \circ \pi)(x) &= \beta g(\pi(x)) \\ &= f(\alpha\pi(x)) \\ &= f(\pi(\alpha x)) \\ &= (f \circ \pi)(\alpha x), \end{aligned} \tag{47}$$

for all $x \in D_1$. Thus, $\mathcal{G} \circ \pi$ is exactly scaling-universal for $\mathcal{F} \circ \pi$ on D_1.

Approximate case. Let $f \in \mathcal{F}$, $\alpha \in \mathbb{R}$, and $\varepsilon > 0$. Since $\mathcal{G}$ is approximately scaling-universal for $\mathcal{F}$ on D_2, there exists $\tau \geq 0$ such that, for every $|\beta| > \tau$, there exists $g \in \mathcal{G}$ satisfying

$$\|\beta g - f(\alpha \cdot)\|_{D_2} < \varepsilon. \tag{48}$$

Therefore, for every $|\beta| > \tau$, we have

$$\begin{aligned} \|\beta(g \circ \pi) - (f \circ \pi)(\alpha \cdot)\|_{D_1} &= \|(\beta g) \circ \pi - f \circ (\pi(\alpha \cdot))\|_{D_1} \\ &= \|(\beta g) \circ \pi - f \circ (\alpha\pi)\|_{D_1} \\ &= \|(\beta g) \circ \pi - f(\alpha \cdot) \circ \pi\|_{D_1} \\ &\leq \|\beta g - f(\alpha \cdot)\|_{D_2} \\ &< \varepsilon, \end{aligned} \tag{49}$$

Thus, $\mathcal{G} \circ \pi$ is approximately scaling-universal for $\mathcal{F} \circ \pi$ on D_1. □

Proof of Lemma 2.10. We prove the exact and approximate cases separately.

Exact case. Let $f \in \mathcal{F}$ and $\alpha \in \mathbb{R}$ be arbitrary. Since $\mathcal{G}$ is exactly scaling-universal for $\mathcal{F}$ on D, there exists $\tau \geq 0$ such that, for every $|\beta| > \tau$, there exists $g \in \mathcal{G}$ satisfying

$$\beta g(x) = f(\alpha x) \tag{50}$$

for all $x \in D$.

Define

$$\widetilde{\tau} := \frac{\tau|a|}{|b|}. \tag{51}$$

For any $|\widetilde{\beta}| > \widetilde{\tau}$, we have

$$\left|\frac{b}{a}\widetilde{\beta}\right| > \tau, \tag{52}$$

so there exists $g \in \mathcal{G}$ satisfying

$$\frac{b}{a}\widetilde{\beta} g = f(\alpha \cdot). \tag{53}$$

Define

$$\widetilde{f} := af \in a\mathcal{F} \tag{54}$$

and

$$\widetilde{g} := bg \in b\mathcal{G}. \tag{55}$$

It follows that

$$\widetilde{\beta}\widetilde{g}(x) = \widetilde{f}(\alpha x) \tag{56}$$

for all $x \in D$. Thus, $b\mathcal{G}$ is exactly scaling-universal for $a\mathcal{F}$ on D.

Approximate case. Let $f \in \mathcal{F}$, $\alpha \in \mathbb{R}$, and $\varepsilon > 0$ be arbitrary. Since $\mathcal{G}$ is approximately scaling-universal for $\mathcal{F}$ on D, there exists $\tau \geq 0$ such that, for every $|\beta| > \tau$, there exists $g \in \mathcal{G}$ satisfying

$$\|\beta g - f(\alpha \cdot)\|_D < \frac{\varepsilon}{|a|}. \tag{57}$$

Define

$$\widetilde{\tau} := \frac{\tau|a|}{|b|}. \tag{58}$$

For any $|\widetilde{\beta}| > \widetilde{\tau}$, we have

$$\left|\frac{b}{a}\widetilde{\beta}\right| > \tau, \tag{59}$$

so there exists $g \in \mathcal{G}$ satisfying

$$\left\|\frac{b}{a}\widetilde{\beta}g - f(\alpha\cdot)\right\|_D < \frac{\varepsilon}{|a|}. \tag{60}$$

Define

$$\widetilde{f} := af \in a\mathcal{F} \tag{61}$$

and

$$\widetilde{g} := bg \in b\mathcal{G}. \tag{62}$$

It follows that

$$\begin{aligned}\|\widetilde{\beta}\widetilde{g} - \widetilde{f}(\alpha\cdot)\|_D &= \|b\widetilde{\beta}g - af(\alpha\cdot)\|_D \\ &= |a|\left\|\frac{b}{a}\widetilde{\beta}g - f(\alpha\cdot)\right\|_D \\ &< \varepsilon.\end{aligned} \tag{63}$$

Thus, $b\mathcal{G}$ is approximately scaling-universal for $a\mathcal{F}$ on D. □

Proof of Lemma 2.11. We prove the exact and approximate cases separately.

Exact case. Let $f \in \mathcal{F}_1 + \mathcal{F}_2$ and $\alpha \in \mathbb{R}$ be arbitrary. Then, there exist $f_1 \in \mathcal{F}_1$ and $f_2 \in \mathcal{F}_2$ such that

$$f = f_1 + f_2. \tag{64}$$

Since $\mathcal{G}_1$ is exactly scaling-universal for $\mathcal{F}_1$ on D, there exists $\tau_1 \geq 0$ such that, for every $|\beta| > \tau_1$, there exists $g_1 \in \mathcal{G}_1$ satisfying

$$\beta g_1(x) = f_1(\alpha x). \tag{65}$$

for all $x \in D$. Similarly, since $\mathcal{G}_2$ is exactly scaling-universal for $\mathcal{F}_2$ on D, there exists $\tau_2 \geq 0$ such that, for every $|\beta| > \tau_2$, there exists $g_2 \in \mathcal{G}_2$ satisfying

$$\beta g_2(x) = f_2(\alpha x) \tag{66}$$

for all $x \in D$.

Let

$$\tau := \max\{\tau_1, \tau_2\}. \tag{67}$$

Then, for every $|\beta| > \tau$, we have

$$\begin{aligned}\beta(g_1 + g_2)(x) &= \beta g_1(x) + \beta g_2(x) \\ &= f_1(\alpha x) + f_2(\alpha x) \\ &= f(\alpha x)\end{aligned} \tag{68}$$

for all $x \in D$. Since $g_1 + g_2 \in \mathcal{G}_1 + \mathcal{G}_2$, it follows that $\mathcal{G}_1 + \mathcal{G}_2$ is exactly scaling-universal for $\mathcal{F}_1 + \mathcal{F}_2$ on D.

Approximate case. Let $f \in \mathcal{F}_1 + \mathcal{F}_2$, $\alpha \in \mathbb{R}$, and $\varepsilon > 0$ be arbitrary. Then, there exist $f_1 \in \mathcal{F}_1$ and $f_2 \in \mathcal{F}_2$ such that

$$f = f_1 + f_2. \tag{69}$$

Since $\mathcal{G}_1$ is approximately scaling-universal for $\mathcal{F}_1$ on D, there exists $\tau_1 \geq 0$ such that, for every $|\beta| > \tau_1$, there exists $g_1 \in \mathcal{G}_1$ satisfying

$$\|\beta g_1 - f_1(\alpha\cdot)\|_D < \frac{\varepsilon}{2}. \tag{70}$$

Similarly, since $\mathcal{G}_2$ is approximately scaling-universal for $\mathcal{F}_2$ on D, there exists $\tau_2 \geq 0$ such that, for every $|\beta| > \tau_2$, there exists $g_2 \in \mathcal{G}_2$ satisfying

$$\|\beta g_2 - f_2(\alpha\cdot)\|_D < \frac{\varepsilon}{2}. \tag{71}$$

Let

$$\tau := \max\{\tau_1, \tau_2\}. \tag{72}$$

Then, for every $|\beta| > \tau$, the triangle inequality gives

$$\begin{aligned} \|\beta(g_1 + g_2) - f(\alpha\cdot)\|_D &= \|\beta g_1 - f_1(\alpha\cdot) + \beta g_2 - f_2(\alpha\cdot)\|_D \\ &\le \|\beta g_1 - f_1(\alpha\cdot)\|_D + \|\beta g_2 - f_2(\alpha\cdot)\|_D \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= \varepsilon. \end{aligned} \tag{73}$$

Since $g_1 + g_2 \in \mathcal{G}_1 + \mathcal{G}_2$, it follows that $\mathcal{G}_1 + \mathcal{G}_2$ is approximately scaling-universal for $\mathcal{F}_1 + \mathcal{F}_2$ on D. □

Proof of Lemma 2.12. Let $f \in \mathcal{F}_2 \circ \mathcal{F}_1$ and $\alpha \in \mathbb{R}$ be arbitrary. Then, there exist $f_2 \in \mathcal{F}_2$ and $f_1 \in \mathcal{F}_1$ such that

$$f = f_2 \circ f_1. \tag{74}$$

Since $\mathcal{G}_1$ is exactly scaling-universal for $\mathcal{F}_1$ on D_1, there exists $\tau_1 \ge 0$ such that, for every $|\beta_1| > \tau_1$, there exists $g_1 \in \mathcal{G}_1$ satisfying

$$\beta_1 g_1(y) = f_1(\alpha y) \tag{75}$$

for all $y \in D_1$. Choose any β_1 with $|\beta_1| > \tau_1$ and let g_1 be a corresponding function. Since $\mathcal{F}_2$ is input-scaling-invariant,

$$f_2(\beta_1\cdot) \in \mathcal{F}_2. \tag{76}$$

Since $\mathcal{G}_2$ is exactly scaling-universal for $\mathcal{F}_2$ on D_2, there exists $\tau_2 \ge 0$ such that, for every $|\beta_2| > \tau_2$, there exists $g_2 \in \mathcal{G}_2$ satisfying

$$\beta_2 g_2(x) = f_2(\beta_1 x) \tag{77}$$

for all $x \in D_2$.

Therefore, for every $|\beta_2| > \tau_2$, we have

$$\begin{aligned} \beta_2(g_2 \circ g_1)(x) &= \beta_2 g_2(g_1(x)) \\ &= f_2(\beta_1 g_1(x)) \\ &= f_2(f_1(\alpha x)) \\ &= f(\alpha x) \end{aligned} \tag{78}$$

for all $x \in D_1$. Since $g_2 \circ g_1 \in \mathcal{G}_2 \circ \mathcal{G}_1$, it follows that $\mathcal{G}_2 \circ \mathcal{G}_1$ is exactly scaling-universal for $\mathcal{F}_2 \circ \mathcal{F}_1$ on D_1. □

Proof of Lemma 2.13. Let $f \in \mathcal{F}_2 \circ \mathcal{F}_1$, $\alpha \in \mathbb{R}$, and $\varepsilon > 0$ be arbitrary. Then, there exist $f_2 \in \mathcal{F}_2$ and $f_1 \in \mathcal{F}_1$ such that

$$f = f_2 \circ f_1. \tag{79}$$

Since $\mathcal{F}_1$ is input-scaling-invariant, $f_1(\alpha\cdot) \in \mathcal{F}_1$. Hence, by the continuity of functions in $\mathcal{F}_1$, $f_1(\alpha\cdot)$ is continuous on D_1 and

$$f_1(\alpha D_1) \subseteq \mathcal{F}_1(D_1) \subseteq \mathrm{int}(\mathrm{dom}(\mathcal{F}_2)). \tag{80}$$

Since D_1 is compact, $f_1(\alpha D_1)$ is compact. Therefore, there exists a compact set K such that

$$f_1(\alpha D_1) \subseteq \mathrm{int}(K) \subseteq K \subseteq \mathrm{dom}(\mathcal{F}_2). \tag{81}$$

Hence, there exists $\rho > 0$ such that

$$\|\widetilde{f}_1 - f_1(\alpha\cdot)\|_{D_1} < \rho \tag{82}$$

implies

$$\widetilde{f}_1(D_1) \subseteq K. \tag{83}$$

By the continuity of functions in $\mathcal{F}_2$, f_2 is continuous on K and therefore uniformly continuous on K. Thus, there exists $\delta > 0$ such that, whenever $\widetilde{f}_1(D_1) \subseteq K$ and

$$\|\widetilde{f}_1 - f_1(\alpha\cdot)\|_{D_1} < \delta, \tag{84}$$

we have

$$\|f_2 \circ \widetilde{f}_1 - f_2 \circ f_1(\alpha\cdot)\|_{D_1} < \frac{\varepsilon}{2}. \tag{85}$$

Set

$$\varepsilon_1 := \min\{\delta, \rho\}. \tag{86}$$

Since $\mathcal{G}_1$ is approximately scaling-universal for $\mathcal{F}_1$ on D_1, there exists $\tau_1 \geq 0$ such that, for every $|\beta_1| > \tau_1$, there exists $g_1 \in \mathcal{G}_1$ satisfying

$$\|\beta_1 g_1 - f_1(\alpha\cdot)\|_{D_1} < \varepsilon_1. \tag{87}$$

Choose any β_1 with $|\beta_1| > \tau_1$ and let g_1 be a corresponding function. Since $g_1 \in \mathcal{G}_1$ and $\mathcal{G}_1(D_1) \subseteq D_2$, we have

$$g_1(D_1) \subseteq D_2. \tag{88}$$

Moreover,

$$\beta_1 g_1(D_1) \subseteq K, \tag{89}$$

and

$$\|f_2 \circ (\beta_1 g_1) - f_2 \circ f_1(\alpha\cdot)\|_{D_1} < \frac{\varepsilon}{2}. \tag{90}$$

Since $\mathcal{G}_2$ is approximately scaling-universal for $\mathcal{F}_2$ on D_2, there exists $\tau_2 \geq 0$ such that, for every $|\beta_2| > \tau_2$, there exists $g_2 \in \mathcal{G}_2$ satisfying

$$\|\beta_2 g_2 - f_2(\beta_1\cdot)\|_{D_2} < \frac{\varepsilon}{2}. \tag{91}$$

Since $g_1(D_1) \subseteq D_2$,

$$\|(\beta_2 g_2) \circ g_1 - f_2(\beta_1\cdot) \circ g_1\|_{D_1} \leq \|\beta_2 g_2 - f_2(\beta_1\cdot)\|_{D_2} < \frac{\varepsilon}{2}. \tag{92}$$

Therefore, for every $|\beta_2| > \tau_2$, the triangle inequality gives

$$\begin{aligned}
\|\beta_2(g_2 \circ g_1) - f(\alpha\cdot)\|_{D_1} &= \|(\beta_2 g_2) \circ g_1 - f_2 \circ f_1(\alpha\cdot)\|_{D_1} \\
&\leq \|(\beta_2 g_2) \circ g_1 - f_2 \circ (\beta_1 g_1)\|_{D_1} + \|f_2 \circ (\beta_1 g_1) - f_2 \circ f_1(\alpha\cdot)\|_{D_1} \\
&= \|(\beta_2 g_2) \circ g_1 - f_2(\beta_1\cdot) \circ g_1\|_{D_1} + \|f_2(\beta_1\cdot) \circ g_1 - f_2 \circ f_1(\alpha\cdot)\|_{D_1} \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\
&= \varepsilon.
\end{aligned} \tag{93}$$

Since $g_2 \circ g_1 \in \mathcal{G}_2 \circ \mathcal{G}_1$, it follows that $\mathcal{G}_2 \circ \mathcal{G}_1$ is approximately scaling-universal for $\mathcal{F}_2 \circ \mathcal{F}_1$ on D_1. □

Proof of Lemma 3.6. We first verify that $|\psi^{\perp}\rangle$ is normalized:

$$\begin{aligned}
\left\||\psi^{\perp}\rangle\right\|^2 &= \frac{\|(U|\psi\rangle - x|\psi\rangle)\|^2}{1-x^2} \\
&= \frac{(\langle\psi|U - x\langle\psi|)(U|\psi\rangle - x|\psi\rangle)}{1-x^2} \\
&= \frac{\langle\psi|U^2|\psi\rangle - 2x\langle\psi|U|\psi\rangle + x^2\langle\psi|\psi\rangle}{1-x^2} \\
&= \frac{1 - 2x^2 + x^2}{1-x^2} \\
&= 1.
\end{aligned} \tag{94}$$

Moreover, $|\psi^\perp\rangle$ is orthogonal to $|\psi\rangle$, since

$$\langle\psi|\psi^\perp\rangle = \frac{\langle\psi| U |\psi\rangle - x \langle\psi|\psi\rangle}{\sqrt{1-x^2}} = \frac{x - x}{\sqrt{1-x^2}} = 0. \tag{95}$$

Thus, $\{|\psi\rangle, |\psi^\perp\rangle\}$ is an orthonormal basis for $\mathcal{S}$.

We next show that $\mathcal{S}$ is invariant under U. By the definition of $|\psi^\perp\rangle$,

$$U |\psi\rangle = x |\psi\rangle + \sqrt{1-x^2} |\psi^\perp\rangle . \tag{96}$$

Hence, $U |\psi\rangle \in \mathcal{S}$.

It remains to determine the action of U on $|\psi^\perp\rangle$. Using $U^2 = I$ and the preceding expression for $U |\psi\rangle$, we obtain

$$\begin{aligned} U |\psi^\perp\rangle &= \frac{U^2 |\psi\rangle - xU |\psi\rangle}{\sqrt{1-x^2}} \\ &= \frac{|\psi\rangle - x(x |\psi\rangle + \sqrt{1-x^2} |\psi^\perp\rangle)}{\sqrt{1-x^2}} \\ &= \sqrt{1-x^2} |\psi\rangle - x |\psi^\perp\rangle . \end{aligned} \tag{97}$$

Therefore, $U |\psi^\perp\rangle \in \mathcal{S}$ as well, and consequently $\mathcal{S}$ is invariant under U.

Finally, with respect to the orthonormal basis $\{|\psi\rangle, |\psi^\perp\rangle\}$, the restriction of U to $\mathcal{S}$ is

$$U\big|_{\mathcal{S}} := \begin{bmatrix} \langle\psi| U |\psi\rangle & \langle\psi| U |\psi^\perp\rangle \\ \langle\psi^\perp| U |\psi\rangle & \langle\psi^\perp| U |\psi^\perp\rangle \end{bmatrix} = \begin{bmatrix} x & \sqrt{1-x^2} \\ \sqrt{1-x^2} & -x \end{bmatrix} \tag{98}$$

□

Proof of Lemma 3.7. By Lemma 3.6, the vectors $|\psi\rangle$ and $|\psi^\perp\rangle$ are orthonormal. Therefore,

$$\langle\Psi_+|\Psi_+\rangle = \frac{(\langle\psi| - i \langle\psi^\perp|)(|\psi\rangle + i |\psi^\perp\rangle)}{2} = \frac{\langle\psi|\psi\rangle + \langle\psi^\perp|\psi^\perp\rangle}{2} = 1, \tag{99}$$

$$\langle\Psi_-|\Psi_-\rangle = \frac{(\langle\psi| + i \langle\psi^\perp|)(|\psi\rangle - i |\psi^\perp\rangle)}{2} = \frac{\langle\psi|\psi\rangle + \langle\psi^\perp|\psi^\perp\rangle}{2} = 1, \tag{100}$$

$$\langle\Psi_+|\Psi_-\rangle = \frac{(\langle\psi| - i \langle\psi^\perp|)(|\psi\rangle - i |\psi^\perp\rangle)}{2} = \frac{\langle\psi|\psi\rangle - \langle\psi^\perp|\psi^\perp\rangle}{2} = 0, \tag{101}$$

Hence, $|\Psi_+\rangle$ and $|\Psi_-\rangle$ are orthonormal.

Moreover, Lemma 3.6 gives

$$\begin{aligned} W |\Psi_\pm\rangle &= R_\psi U |\Psi_\pm\rangle \\ &= \frac{R_\psi(U |\psi\rangle \pm iU |\psi^\perp\rangle)}{\sqrt{2}} \\ &= \frac{R_\psi((x |\psi\rangle + \sqrt{1-x^2} |\psi^\perp\rangle) \pm i(\sqrt{1-x^2} |\psi\rangle - x |\psi^\perp\rangle))}{\sqrt{2}} \\ &= \frac{R_\psi((x \pm i\sqrt{1-x^2}) |\psi\rangle + (\sqrt{1-x^2} \mp ix) |\psi^\perp\rangle)}{\sqrt{2}}. \end{aligned} \tag{102}$$

Since $R_\psi |\psi\rangle = |\psi\rangle$ and $R_\psi |\psi^\perp\rangle = - |\psi^\perp\rangle$, this becomes

$$\begin{aligned} W |\Psi_\pm\rangle &= \frac{(x \pm i\sqrt{1-x^2}) |\psi\rangle - (\sqrt{1-x^2} \mp ix) |\psi^\perp\rangle}{\sqrt{2}} \\ &= \frac{(x \pm i\sqrt{1-x^2}) |\psi\rangle \pm i(x \pm i\sqrt{1-x^2}) |\psi^\perp\rangle}{\sqrt{2}} \\ &= \frac{(x \pm i\sqrt{1-x^2})(|\psi\rangle \pm i |\psi^\perp\rangle)}{\sqrt{2}} \\ &= (x \pm i\sqrt{1-x^2}) |\Psi_\pm\rangle . \end{aligned} \tag{103}$$

Since $x \in (-1, 1)$,

$$x \pm i\sqrt{1-x^2} = e^{\pm i \arccos x}. \tag{104}$$

Hence, $|\Psi_\pm\rangle$ are eigenvectors of W with eigenvalues $e^{\pm i \arccos x}$. □

Proof of Lemma 3.9. Suppose first that $x \in (-1, 1)$. By Lemma 3.7, W has orthonormal eigenvectors

$$|\Psi_\pm\rangle := \frac{\left(|\psi\rangle \pm i\,|\psi^\perp\rangle\right)}{\sqrt{2}} \tag{105}$$

with corresponding eigenvalues $e^{\pm i \arccos x}$. Since

$$|\psi\rangle = \frac{|\Psi_+\rangle + |\Psi_-\rangle}{\sqrt{2}}, \tag{106}$$

using the spectral decomposition of W and polynomial functional calculus gives

$$\begin{aligned} \langle\psi|\, P(W)\, |\psi\rangle &= \frac{(\langle\Psi_+| + \langle\Psi_-|)P(W)(|\Psi_+\rangle + |\Psi_-\rangle)}{2} \\ &= \frac{(\langle\Psi_+| + \langle\Psi_-|)(P(e^{i\arccos x})\,|\Psi_+\rangle + P(e^{-i\arccos x})\,|\Psi_-\rangle)}{2} \\ &= \frac{P(e^{i\arccos x}) + P(e^{-i\arccos x})}{2}. \end{aligned} \tag{107}$$

It remains to consider the endpoints. If $x = \pm 1$, then

$$U\,|\psi\rangle = x\,|\psi\rangle\,, \tag{108}$$

and hence

$$W\,|\psi\rangle = xR_\psi\,|\psi\rangle = x\,|\psi\rangle\,. \tag{109}$$

It follows that

$$\langle\psi|\, P(W)\, |\psi\rangle = P(x). \tag{110}$$

Since $e^{i\arccos x} = e^{-i\arccos x} = x$ for $x = \pm 1$,

$$P(x) = \frac{P(e^{i\arccos x}) + P(e^{-i\arccos x})}{2}. \tag{111}$$

Thus the claimed identity holds for all $x \in [-1, 1]$. □

Proof of Theorem 3.11. We prove both directions.

Forward direction. Suppose that $p \in \mathcal{R}_d$. Then, there exists an admissible polynomial $P \in \mathcal{A}_d$ such that

$$p(x) = \frac{P(e^{i\arccos x}) + P(e^{-i\arccos x})}{2} \tag{112}$$

for all $x \in [-1, 1]$. By the GQSP theorem (Motlagh & Wiebe, 2024), there exist parameters $\boldsymbol{\theta}, \boldsymbol{\phi} \in \mathbb{R}^{d+1}$ and $\lambda \in \mathbb{R}$ such that

$$P(W(x)) = (\langle 0| \otimes I)Q(W(x); \boldsymbol{\theta}, \boldsymbol{\phi}, \lambda)(|0\rangle \otimes I) \tag{113}$$

for all $x \in [-1, 1]$. Then, by Lemma 3.9,

$$(\langle 0| \otimes \langle\psi|)Q(W(x); \boldsymbol{\theta}, \boldsymbol{\phi}, \lambda)(|0\rangle \otimes |\psi\rangle) = \frac{P(e^{i\arccos x}) + P(e^{-i\arccos x})}{2} = p(x) \tag{114}$$

for all $x \in [-1, 1]$. Thus, $(Q(W(x); \boldsymbol{\theta}, \boldsymbol{\phi}, \lambda), |0\rangle \otimes |\psi\rangle\,, |0\rangle \otimes |\psi\rangle)$ is a projected unitary encoding of $p(x)$ for all $x \in [-1, 1]$.

Converse. Suppose that there exist parameters $\boldsymbol{\theta}, \boldsymbol{\phi} \in \mathbb{R}^{d+1}$ and $\lambda \in \mathbb{R}$ such that $(Q(W(x); \boldsymbol{\theta}, \boldsymbol{\phi}, \lambda), |0\rangle \otimes |\psi\rangle\,, |0\rangle \otimes |\psi\rangle)$ is a projected unitary encoding of $p(x)$ for all $x \in [-1, 1]$. Then,

$$p(x) = (\langle 0| \otimes \langle\psi|)Q(W(x); \boldsymbol{\theta}, \boldsymbol{\phi}, \lambda)(|0\rangle \otimes |\psi\rangle) \tag{115}$$

for all $x \in [-1, 1]$. By the GQSP theorem (Motlagh & Wiebe, 2024), the parameters $\boldsymbol{\theta}, \boldsymbol{\phi}, \lambda$ determine an admissible polynomial $P \in \mathcal{A}_d$ satisfying

$$(\langle 0| \otimes I)Q(W(x); \boldsymbol{\theta}, \boldsymbol{\phi}, \lambda)(|0\rangle \otimes I) = P(W(x)) \tag{116}$$

for all $x \in [-1, 1]$. Hence, by Lemma 3.9,

$$\frac{P(e^{i \arccos x}) + P(e^{-i \arccos x})}{2} = (\langle 0| \otimes \langle \psi|)Q(W(x); \boldsymbol{\theta}, \boldsymbol{\phi}, \lambda)(|0\rangle \otimes |\psi\rangle) = p(x). \tag{117}$$

Thus, $p \in \mathcal{R}_d$. □

Proof of Proposition 3.14. Let $r \in \mathrm{Re}(\mathcal{R}_d)$. Then, there exists $q \in \mathcal{R}_d$ such that

$$r(x) = \mathrm{Re}(q(x)) \tag{118}$$

for all $x \in [-1, 1]$. Since $q \in \mathcal{R}_d$, there exists $P \in \mathcal{A}_d$ such that

$$q(x) = \frac{P(e^{i \arccos x}) + P(e^{-i \arccos x})}{2}. \tag{119}$$

for all $x \in [-1, 1]$.

Define

$$\widetilde{P}(z) := \frac{P(z) + \overline{P(\overline{z})}}{2}. \tag{120}$$

Then,

$$|\widetilde{P}(z)| \leq \frac{|P(z)| + |P(\overline{z})|}{2} \leq 1 \tag{121}$$

for all $|z| = 1$, so $\widetilde{P} \in \mathcal{A}_d$. Moreover,

$$\frac{\widetilde{P}(e^{i \arccos x}) + \widetilde{P}(e^{-i \arccos x})}{2} = \mathrm{Re}\left(\frac{P(e^{i \arccos x}) + P(e^{-i \arccos x})}{2}\right) = r(x). \tag{122}$$

for all $x \in [-1, 1]$. Thus, $r \in \mathcal{R}_d$. □

Proof of Lemma 3.16. Let

$$p(x) = \sum_{j=0}^{d} a_j T_j(x) \tag{123}$$

be the Chebyshev expansion of p. We prove both directions.

Forward direction. Suppose P is the associated polynomial of p. Then,

$$P(z) = \sum_{j=0}^{d} a_j z^j. \tag{124}$$

Using $T_j(\cos\theta) = \cos(j\theta)$,

$$\begin{aligned} \frac{P(e^{i \arccos x}) + P(e^{-i \arccos x})}{2} &= \sum_{j=0}^{d} a_j \cos(j \arccos x) \\ &= \sum_{j=0}^{d} a_j T_j(x) \\ &= p(x). \end{aligned} \tag{125}$$

Converse. Suppose

$$\frac{P(e^{i \arccos x}) + P(e^{-i \arccos x})}{2} = p(x) \tag{126}$$

for all $x \in [-1, 1]$. Write

$$P(z) = \sum_{j=0}^{d} b_j z^j. \tag{127}$$

Then,

$$\begin{aligned} p(x) &= \frac{P(e^{i \arccos x}) + P(e^{-i \arccos x})}{2} \\ &= \sum_{j=0}^{d} b_j \cos(j \arccos x) \\ &= \sum_{j=0}^{d} b_j T_j(x). \end{aligned} \tag{128}$$

Uniqueness of the Chebyshev expansion gives $b_j = a_j$ for every j, so P is the associated polynomial of p. □

Proof of Proposition 3.18. We prove both directions.

Forward direction. Suppose that

$$r := \frac{p}{\beta} \in \mathrm{Re}(\mathcal{R}_d). \tag{129}$$

By Proposition 3.14, $r \in \mathcal{R}_d$, so there exists $P \in \mathcal{A}_d$ such that

$$\frac{P(e^{i \arccos x}) + P(e^{-i \arccos x})}{2} = r(x) \tag{130}$$

for all $x \in [-1, 1]$. By Lemma 3.16, P is the associated polynomial of r.

Let $\widetilde{P}$ be the associated polynomial of p. Since the associated polynomial is linear in the Chebyshev coefficients,

$$P = \frac{\widetilde{P}}{\beta}. \tag{131}$$

Since $P \in \mathcal{A}_d$,

$$|P(z)| \leq 1 \tag{132}$$

for all $|z| = 1$. Therefore,

$$|\beta| \geq \max_{|z|=1} |\widetilde{P}(z)| = \Gamma(p). \tag{133}$$

Converse. Suppose that

$$|\beta| \geq \Gamma(p). \tag{134}$$

Let $\widetilde{P}$ be the associated polynomial of p, and define

$$P := \frac{\widetilde{P}}{\beta}. \tag{135}$$

Then,

$$|P(z)| = \frac{|\widetilde{P}(z)|}{|\beta|} \leq \frac{\Gamma(p)}{|\beta|} \leq 1 \tag{136}$$

for all $|z| = 1$, so $P \in \mathcal{A}_d$.

Let

$$r := \frac{p}{\beta}. \tag{137}$$

Since the associated polynomial is linear in the Chebyshev coefficients, P is the associated polynomial of r. By Lemma 3.16,

$$\frac{P(e^{i \arccos x}) + P(e^{-i \arccos x})}{2} = r(x). \tag{138}$$

for all $x \in [-1, 1]$. Thus,

$$r = \frac{p}{\beta} \in \mathcal{R}_d. \tag{139}$$

Since r is real-valued,

$$r \in \mathrm{Re}(\mathcal{R}_d). \tag{140}$$

□

Proof of Theorem 3.19. Let $p \in \mathcal{P}_d$ and $\alpha \in \mathbb{R}$ be arbitrary, and define the rescaled polynomial

$$\widetilde{p}(x) := p(\alpha x). \tag{141}$$

Since

$$\deg(\widetilde{p}) \le \deg(p) \le d, \tag{142}$$

we have $\widetilde{p} \in \mathcal{P}_d$. Set

$$\tau := \Gamma(\widetilde{p}). \tag{143}$$

For every $|\beta| > \tau$, Proposition 3.18 gives

$$\frac{\widetilde{p}}{\beta} \in \mathrm{Re}(\mathcal{R}_d). \tag{144}$$

Moreover,

$$\beta \frac{\widetilde{p}(x)}{\beta} = p(\alpha x) \tag{145}$$

for all $x \in [-1, 1]$. Thus, $\mathrm{Re}(\mathcal{R}_d)$ is exactly scaling-universal for $\mathcal{P}_d$ on $[-1, 1]$. □

Proof of Corollary 3.20. Let $p \in \mathcal{P}_\infty$ and $\alpha \in \mathbb{R}$ be arbitrary. Since

$$\mathcal{P}_\infty = \bigcup_{d=0}^{\infty} \mathcal{P}_d, \tag{146}$$

there exists $d \ge 0$ such that $p \in \mathcal{P}_d$.

By Theorem 3.19, $\mathrm{Re}(\mathcal{R}_d)$ is exact scaling-universal for $\mathcal{P}_d$ on $[-1, 1]$. Hence, there exists $\tau \ge 0$, for every $|\beta| > \tau$ there exists $r \in \mathrm{Re}(\mathcal{R}_d)$ satisfying

$$\beta r(x) = p(\alpha x) \tag{147}$$

for all $x \in [-1, 1]$.

Since

$$\mathrm{Re}(\mathcal{R}_d) \subseteq \mathrm{Re}(\mathcal{R}_\infty), \tag{148}$$

we also have $r \in \mathrm{Re}(\mathcal{R}_\infty)$. Therefore, for the arbitrary $p \in \mathcal{P}_\infty$ and $\alpha \in \mathbb{R}$, there exists $\tau \ge 0$, for every $|\beta| > \tau$, there exists $r \in \mathrm{Re}(\mathcal{R}_\infty)$ satisfying

$$\beta r(x) = p(\alpha x) \tag{149}$$

for all $x \in [-1, 1]$. Thus, $\mathrm{Re}(\mathcal{R}_\infty)$ is exact scaling-universal for $\mathcal{P}_\infty$ on $[-1, 1]$. □

Proof of Theorem 3.21. Let $f \in C(\mathbb{R})$, $\alpha \in \mathbb{R}$, and $\varepsilon > 0$ be arbitrary. By the Weierstrass approximation theorem (Rudin, 1976), there exists a polynomial $p \in \mathcal{P}_\infty$ such that

$$\|p - f\|_{[-|\alpha|, |\alpha|]} < \varepsilon. \tag{150}$$

By Corollary 3.20, $\mathrm{Re}(\mathcal{R}_\infty)$ is exact scaling-universal for $\mathcal{P}_\infty$ on $[-1, 1]$. Hence, there exists $\tau \ge 0$ such that, for every $|\beta| > \tau$, there exists $g \in \mathrm{Re}(\mathcal{R}_\infty)$ satisfying

$$\beta g(x) = p(\alpha x) \tag{151}$$

for all $x \in [-1, 1]$. Hence,

$$\begin{aligned} \|\beta g - f(\alpha \cdot)\|_{[-1,1]} &= \|p(\alpha \cdot) - f(\alpha \cdot)\|_{[-1,1]} \\ &= \|p - f\|_{[-|\alpha|, |\alpha|]} \\ &< \varepsilon. \end{aligned} \tag{152}$$

Therefore, $\mathrm{Re}(\mathcal{R}_\infty)$ is approximately scaling-universal for $C(\mathbb{R})$ on $[-1, 1]$. □

Proof of Theorem 4.3. Let $f \in C(\mathbb{R}^n)$, $\alpha \in \mathbb{R}$, and $\varepsilon > 0$ be arbitrary. Since $[-|\alpha|, |\alpha|]^n$ is compact, the restriction of f to this set is bounded. By the Tietze extension theorem (Munkres, 2000), there exists $\widetilde{f} \in C_b(\mathbb{R}^n)$ such that

$$\widetilde{f}(x) = f(x) \tag{153}$$

for all $x \in [-|\alpha|, |\alpha|]^n$. Thus,

$$\widetilde{f}(\alpha x) = f(\alpha x) \tag{154}$$

for all $x \in [-1, 1]^n$.

By Theorem 4.2, there exists $\tau \geq 0$ such that, for every $|\beta| > \tau$, there exists $g \in \mathfrak{A}_n(\mathrm{Re}(\mathcal{R}_\infty))$ satisfying

$$\|\beta g - \widetilde{f}(\alpha \cdot)\|_{[-1,1]^n} < \varepsilon. \tag{155}$$

Therefore,

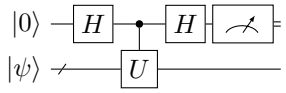
$$\|\beta g - f(\alpha \cdot)\|_{[-1,1]^n} < \varepsilon. \tag{156}$$

Thus, $\mathfrak{A}_n(\mathrm{Re}(\mathcal{R}_\infty))$ is approximately scaling-universal for $C(\mathbb{R}^n)$ on $[-1, 1]^n$. □

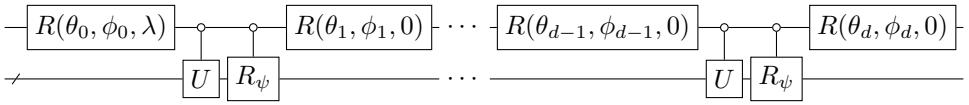
C Quantum Circuit Implementation Details

This section presents the quantum circuits corresponding to the constructions introduced in the main text.

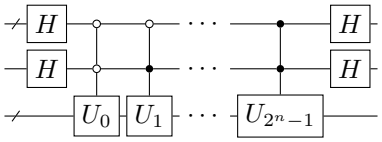
The Hadamard test used to extract the real part of a projected unitary encoding is given by



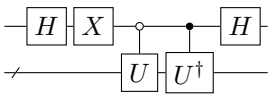
The quantum circuit implementing the polynomial transformation is



The normalized LCU construction is implemented by



Finally, the Hermitianization of a projected unitary encoding is implemented by



D Additional Algebraic Properties of Scaling Universality

This section presents additional properties of scaling universality that are conceptually meaningful but not required for the main results.

Proposition D.1 (Exact-to-approximate scaling universality)**.** *Let $\mathcal{F}$ and $\mathcal{G}$ be function classes, and let D be a set. If $\mathcal{G}$ is exactly scaling-universal for $\mathcal{F}$ on D, then $\mathcal{G}$ is approximately scaling-universal for $\mathcal{F}$ on D.*

Proof. Let $f \in \mathcal{F}$, $\alpha \in \mathbb{R}$, and $\varepsilon > 0$ be arbitrary. Since $\mathcal{G}$ is exactly scaling-universal for $\mathcal{F}$ on D, there exists $\tau \geq 0$ such that, for every $|\beta| > \tau$, there exists $g \in \mathcal{G}$ satisfying

$$\beta g(x) = f(\alpha x) \tag{157}$$

for all $x \in D$. Hence,

$$\|\beta g - f(\alpha \cdot)\|_D = 0 < \varepsilon. \tag{158}$$

Thus, $\mathcal{G}$ is approximately scaling-universal for $\mathcal{F}$ on D. □

Proposition D.2 (Restriction monotonicity)**.** *Let $\mathcal{F}$ and $\mathcal{G}$ be function classes, and let D_1 and D_2 be sets $D_2 \subseteq D_1$. If $\mathcal{G}$ is exactly scaling-universal for $\mathcal{F}$ on D_1, then $\mathcal{G}$ is exactly scaling-universal for $\mathcal{F}$ on D_2. If $\mathcal{G}$ is approximately scaling-universal for $\mathcal{F}$ on D_1, then $\mathcal{G}$ is approximately scaling-universal for $\mathcal{F}$ on D_2.*

Proof. We prove the exact and approximate cases separately.

Exact case. Let $f \in \mathcal{F}$ and $\alpha \in \mathbb{R}$ be arbitrary. Since $\mathcal{G}$ is exactly scaling-universal for $\mathcal{F}$ on D_1, there exists $\tau \geq 0$ such that, for every $|\beta| > \tau$, there exists $g \in \mathcal{G}$ satisfying

$$\beta g(x) = f(\alpha x) \tag{159}$$

for all $x \in D_1$. Since $D_2 \subseteq D_1$, the same equality holds for all $x \in D_2$. Thus, $\mathcal{G}$ is exactly scaling-universal for $\mathcal{F}$ on D_2.

Approximate case. Let $f \in \mathcal{F}$, $\alpha \in \mathbb{R}$, and $\varepsilon > 0$ be arbitrary. Since $\mathcal{G}$ is approximately scaling-universal for $\mathcal{F}$ on D_1, there exists $\tau \geq 0$ such that, for every $|\beta| > \tau$, there exists $g \in \mathcal{G}$ satisfying

$$\|\beta g - f(\alpha \cdot)\|_{D_1} < \varepsilon. \tag{160}$$

Since $D_2 \subseteq D_1$,

$$\|\beta g - f(\alpha \cdot)\|_{D_2} \leq \|\beta g - f(\alpha \cdot)\|_{D_1} < \varepsilon. \tag{161}$$

Thus, $\mathcal{G}$ is approximately scaling-universal for $\mathcal{F}$ on D_2. □

Proposition D.3 (Transitivity)**.** *Let $\mathcal{F}$, $\mathcal{G}$, and $\mathcal{H}$ be function classes, and let D be a set. If $\mathcal{G}$ is exactly scaling-universal for $\mathcal{F}$ on D and $\mathcal{H}$ is exactly scaling-universal for $\mathcal{G}$ on D, then $\mathcal{H}$ is exactly scaling-universal for $\mathcal{F}$ on D. If $\mathcal{G}$ is approximately scaling-universal for $\mathcal{F}$ on D and $\mathcal{H}$ is approximately scaling-universal for $\mathcal{G}$ on D, then $\mathcal{H}$ is approximately scaling-universal for $\mathcal{F}$ on D.*

Proof. We prove the exact and approximate cases separately.

Exact case. Let $f \in \mathcal{F}$ and $\alpha \in \mathbb{R}$ be arbitrary. Since $\mathcal{G}$ is exactly scaling-universal for $\mathcal{F}$ on D, there exists $\tau_1 \geq 0$ such that, for every $|\beta_1| > \tau_1$, there exists $g \in \mathcal{G}$ satisfying

$$\beta_1 g(x) = f(\alpha x) \tag{162}$$

for all $x \in D$.

Fix any β_1 satisfying $|\beta_1| > \tau_1$, and let $g \in \mathcal{G}$ be the corresponding function. Since $\mathcal{H}$ is exactly scaling-universal for $\mathcal{G}$ on D, there exists $\tau_2 \geq 0$ such that, for every $|\beta_2| > \tau_2$, there exists $h \in \mathcal{H}$ satisfying

$$\beta_2 h(x) = g(x) \tag{163}$$

for all $x \in D$.

Define

$$\tau := |\beta_1| \tau_2. \tag{164}$$

For any $|\beta| > \tau$, let

$$\beta_2 := \frac{\beta}{\beta_1}. \tag{165}$$

Then, $|\beta_2| > \tau_2$, so there exists $h \in \mathcal{H}$ satisfying

$$\beta_2 h(x) = g(x) \tag{166}$$

for all $x \in D$. Therefore,

$$\beta h(x) = \beta_1 \beta_2 h(x) = \beta_1 g(x) = f(\alpha x) \tag{167}$$

for all $x \in D$. Thus, $\mathcal{H}$ is exactly scaling-universal for $\mathcal{F}$ on D.

Approximate case. Let $f \in \mathcal{F}$, $\alpha \in \mathbb{R}$, and $\varepsilon > 0$ be arbitrary. Since $\mathcal{G}$ is approximately scaling-universal for $\mathcal{F}$ on D, there exists $\tau_1 \geq 0$ such that, for every $|\beta_1| > \tau_1$, there exists $g \in \mathcal{G}$ satisfying

$$\|\beta_1 g - f(\alpha\cdot)\|_D < \frac{\varepsilon}{2}. \tag{168}$$

Fix any β_1 satisfying $|\beta_1| > \tau_1$, and let $g \in \mathcal{G}$ be the corresponding function. Since $\mathcal{H}$ is approximately scaling-universal for $\mathcal{G}$ on D, there exists $\tau_2 \geq 0$ such that, for every $|\beta_2| > \tau_2$, there exists $h \in \mathcal{H}$ satisfying

$$\|\beta_2 h - g\|_D < \frac{\varepsilon}{2|\beta_1|}. \tag{169}$$

Define

$$\tau := |\beta_1|\tau_2. \tag{170}$$

For any $|\beta| > \tau$, let

$$\beta_2 := \frac{\beta}{\beta_1}. \tag{171}$$

Then, $|\beta_2| > \tau_2$, so there exists $h \in \mathcal{H}$ satisfying

$$\|\beta_2 h - g\|_D < \frac{\varepsilon}{2|\beta_1|}. \tag{172}$$

Consequently,

$$\begin{aligned} \|\beta h - f(\alpha\cdot)\|_D &= \|\beta_1\beta_2 h - f(\alpha\cdot)\|_D \\ &\leq |\beta_1|\|\beta_2 h - g\|_D + \|\beta_1 g - f(\alpha\cdot)\|_D \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= \varepsilon. \end{aligned} \tag{173}$$

Thus, $\mathcal{H}$ is approximately scaling-universal for $\mathcal{F}$ on D. □

E Comparison of Universality Frameworks

E.1 Standard Universality

We begin with standard universality, which compares functions under a fixed input domain and without input or output scaling.

Definition E.1 (Exact standard universality)**.** Let $\mathcal{F}$ and $\mathcal{G}$ be function classes, and let D be a set. We say that $\mathcal{G}$ is exactly universal for $\mathcal{F}$ on D if, for every $f \in \mathcal{F}$, there exists $g \in \mathcal{G}$ such that

$$g(x) = f(x) \tag{174}$$

for all $x \in D$.

Definition E.2 (Approximate standard universality)**.** Let $\mathcal{F}$ and $\mathcal{G}$ be function classes, and let D be a set. We say that $\mathcal{G}$ is approximately universal for $\mathcal{F}$ on D if, for every $f \in \mathcal{F}$ and every $\varepsilon > 0$, there exists $g \in \mathcal{G}$ such that

$$\|g - f\|_D < \varepsilon. \tag{175}$$

However, it does not directly capture correspondences between function classes that differ only through input or output scaling. For example, consider the function classes

$$\mathcal{F} := \{f : \mathbb{R} \to \mathbb{R} \mid \exists A, \omega \in \mathbb{R},\ f(x) = A\sin(\omega x)\} \tag{176}$$

and

$$\mathcal{G} := \{g : (-1, 1) \to \mathbb{R} \mid \exists A \in [-1, 1], \omega \in \mathbb{R},\ g(x) = A\sin(\omega x)\} \tag{177}$$

The two classes have the same underlying functional form, but differ in their input domain and output magnitude. Standard universality therefore does not capture their correspondence directly: functions in $\mathcal{G}$ are defined only on $(-1, 1)$ and satisfy $|g(x)| \leq 1$, whereas functions in $\mathcal{F}$ have domain $\mathbb{R}$ and may have arbitrary amplitude. Input and output scaling resolve these differences without changing the underlying functional structure. This motivates a notion of universality that explicitly accounts for input and output scaling.

E.2 Existential scaling Universality

Existential scaling universality removes this restriction by allowing the input and output scales to be chosen separately for each target function.

Definition E.3 (Exact existential scaling universality)**.** Let $\mathcal{F}$ and $\mathcal{G}$ be function classes, and let D be a set. We say that $\mathcal{G}$ is exactly existential-scaling-universal for $\mathcal{F}$ on D if, for every $f \in \mathcal{F}$, there exist $\alpha \in \mathbb{R}$, $\beta \in \mathbb{R} \setminus \{0\}$, and $g \in \mathcal{G}$ such that $\alpha D \subseteq \mathrm{dom}(f)$, $D \subseteq \mathrm{dom}(g)$, and

$$\beta g(x) = f(\alpha x) \tag{178}$$

for all $x \in D$.

Definition E.4 (Approximate existential scaling universality)**.** Let $\mathcal{F}$ and $\mathcal{G}$ be function classes, and let D be a set. We say that $\mathcal{G}$ is approximately existential-scaling-universal for $\mathcal{F}$ on D if, for every $f \in \mathcal{F}$ and every $\varepsilon > 0$, there exist $\alpha \in \mathbb{R}$, $\beta \in \mathbb{R} \setminus 0$, and $g \in \mathcal{G}$ such that $\alpha D \subseteq \mathrm{dom}(f)$, $D \subseteq \mathrm{dom}(g)$, and

$$\|\beta g - f(\alpha \cdot)\|_D < \varepsilon. \tag{179}$$

However, existential scaling universality does not in general provide the scale compatibility required by addition and composition. To see this, suppose that f_1 and f_2 are represented by g_1 and g_2 through

$$\beta_1 g_1(x) = f_1(\alpha_1 x) \tag{180}$$

and

$$\beta_2 g_2(x) = f_2(\alpha_2 x). \tag{181}$$

For addition, a representation of $f_1 + f_2$ using a single input scale α and output scale β would require

$$\beta(g_1(x) + g_2(x)) = f_1(\alpha x) + f_2(\alpha x). \tag{182}$$

Existential scaling universality does not guarantee that the independently selected scales $\alpha_1, \alpha_2, \beta_1, \beta_2$ can be replaced by common input and output scales α and β.

Composition imposes a similar compatibility condition: the output of the inner representation becomes the input to the outer representation, so the inner output scale and outer input scale must be compatible. Existential scaling universality selects these scales independently and therefore does not guarantee this compatibility.

Therefore, existential scaling universality does not generally propagate through addition or composition, and hence does not automatically propagate through recursively constructed architectures.

E.3 Uniform Existential Scaling Universality

A stronger variant of existential scaling universality requires the input and output scales to be chosen independently of the target function, while still allowing a single pair of scales to be chosen existentially.

Definition E.5 (Exact uniform existential scaling universality)**.** Let $\mathcal{F}$ and $\mathcal{G}$ be function classes, and let D be a set. We say that $\mathcal{G}$ is exactly uniform-existential-scaling-universal for $\mathcal{F}$ on D if there exist $\alpha \in \mathbb{R}$ and $\beta \in \mathbb{R} \setminus \{0\}$ such that $\alpha D \subseteq \mathrm{dom}(\mathcal{F})$ and, for every $f \in \mathcal{F}$, there exists $g \in \mathcal{G}$ such that $D \subseteq \mathrm{dom}(g)$ and

$$\beta g(x) = f(\alpha x) \tag{183}$$

for all $x \in D$.

Definition E.6 (Approximate uniform existential scaling universality)**.** Let $\mathcal{F}$ and $\mathcal{G}$ be function classes, and let D be a set. We say that $\mathcal{G}$ is approximately uniform-existential-scaling-universal for $\mathcal{F}$ on D if there exist $\alpha \in \mathbb{R}$ and $\beta \in \mathbb{R} \setminus \{0\}$ such that $\alpha D \subseteq \mathrm{dom}(\mathcal{F})$ and, for every $f \in \mathcal{F}$ and every $\varepsilon > 0$, there exists $g \in \mathcal{G}$ such that $D \subseteq \mathrm{dom}(g)$ and

$$\|\beta g - f(\alpha \cdot)\|_D < \varepsilon. \tag{184}$$

Uniform existential scaling universality removes the function-dependence of the scales in existential scaling universality, but it still provides only one existentially chosen pair (α, β). Therefore, like existential scaling universality, it does not generally propagate through addition or composition. This motivates requiring the scaling correspondence to remain valid under prescribed scales, rather than allowing the scales to be selected separately for each target function.

E.4 Unrestricted Scaling Universality

Unrestricted scaling universality requires the correspondence to hold for every input scaling and every nonzero output scaling.

Definition E.7 (Exact unrestricted scaling universality)**.** Let $\mathcal{F}$ and $\mathcal{G}$ be function classes, and let D be a set such that $\mathcal{F}$ is input-scaling admissible on D. We say that $\mathcal{G}$ is exactly unrestricted-scaling-universal for $\mathcal{F}$ on D if, for every $f \in \mathcal{F}$, every $\alpha \in \mathbb{R}$, and every $\beta \in \mathbb{R} \setminus \{0\}$, there exists $g \in \mathcal{G}$ such that $D \subseteq \mathrm{dom}(g)$ and

$$\beta g(x) = f(\alpha x) \tag{185}$$

for all $x \in D$.

Definition E.8 (Approximate unrestricted scaling universality)**.** Let $\mathcal{F}$ and $\mathcal{G}$ be function classes, and let D be a set such that $\mathcal{F}$ is input-scaling admissible on D. We say that $\mathcal{G}$ is approximately unrestricted-scaling-universal for $\mathcal{F}$ on D if, for every $f \in \mathcal{F}$, every $\alpha \in \mathbb{R}$, every $\beta \in \mathbb{R} \setminus \{0\}$, and every $\varepsilon > 0$, there exists $g \in \mathcal{G}$ such that $D \subseteq \mathrm{dom}(g)$ and

$$\|\beta g - f(\alpha \cdot)\|_D < \varepsilon. \tag{186}$$

Unrestricted scaling universality imposes a stronger output-scaling requirement than is compatible with bounded representations. In particular, suppose that $\mathcal{F}$ contains a function f with $\|f\|_D > 0$ and that $\mathcal{G}$ is uniformly bounded on D, so that

$$M := \sup_{g \in \mathcal{G}} \|g\|_D < \infty. \tag{187}$$

If $M = 0$, the claim is immediate. Otherwise, choosing

$$|\beta| < \frac{\|f\|_D}{2M} \tag{188}$$

gives, for every $g \in \mathcal{G}$,

$$\|\beta g - f\|_D \geq \|f\|_D - |\beta| \|g\|_D > \frac{1}{2}\|f\|_D. \tag{189}$$

Thus, unrestricted scaling universality cannot hold. This motivates restricting the output-scaling quantifier to sufficiently large $|\beta|$ in Definition 2.3 and Definition 2.5.