

On the energy landscape of symmetric quantum signal processing

Jiasu Wang¹, Yulong Dong¹, and Lin Lin^{1,2,3}

¹Department of Mathematics, University of California, Berkeley, CA 94720, USA.

²Challenge Institute for Quantum Computation, University of California, Berkeley, CA 94720, USA

³Applied Mathematics and Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

Symmetric quantum signal processing provides a parameterized representation of a real polynomial, which can be translated into an efficient quantum circuit for performing a wide range of computational tasks on quantum computers. For a given polynomial f , the parameters (called phase factors) can be obtained by solving an optimization problem. However, the cost function is non-convex, and has a very complex energy landscape with numerous global and local minima. It is therefore surprising that the solution can be robustly obtained in practice, starting from a fixed initial guess Φ^0 that contains no information of the input polynomial. To investigate this phenomenon, we first explicitly characterize all the global minima of the cost function. We then prove that one particular global minimum (called the maximal solution) belongs to a neighborhood of Φ^0 , on which the cost function is strongly convex under the condition $\|f\|_\infty = \mathcal{O}(d^{-1})$ with $d = \deg(f)$. Our result provides a partial explanation of the aforementioned success of optimization algorithms.

Keywords—Quantum signal processing, quantum algorithm, nonlinear optimization.

Contents

1	Introduction	2
1.1	Main results	4
1.2	Background and related works	6
1.3	Discussion and open questions	7
2	Preliminaries	8
2.1	Chebyshev polynomials	8
2.2	Notation	9
2.3	Quantum signal processing	10
2.4	Quantum signal processing as an optimization problem	11
3	Symmetric quantum signal processing	11
3.1	Sign of the leading coefficient of $Q(x)$	11
3.2	Existence of symmetric phase factors	12
3.3	Proof of Theorem 1	14
3.4	Optimization based method for finding symmetric phase factors	14
4	Global minimizer	15
4.1	Characterizing complementary polynomials	15
4.2	Maximal solution	19
4.3	Relation to other constructive methods	21
5	Local convergence	21
5.1	Sketch of the proof	22
5.2	Recursion relation in terms of Chebyshev coefficients	22
5.3	Bounding Chebyshev coefficients of the maximal solution	25
5.4	Proof of Theorem 5	27
6	Local strong convexity	29
6.1	Hessian matrix at the maximal solution is positive definite	29
6.2	Lower bounding the distance to the initial point	33
6.3	Local strong convexity of the optimization problem	35
6.4	Projected gradient descend algorithm for symmetric phase-factor evaluation	36
7	Additional numerical results	38
7.1	Hessian matrix without symmetry constraint is singular at global optima	38
7.2	Existence of local minima	38
7.3	Numerical behavior near the maximal solution	41
A	Proof of Lemma 10	45
B	Proof of Lemma 11	45
C	Proof of Lemma 13	46

1 Introduction

Given a target polynomial $f \in \mathbb{R}[x]$ satisfying (1) $\deg(f) = d$, (2) the parity of f is $d \bmod 2$, (3) $\|f\|_\infty := \max_{x \in [-1,1]} |f(x)| < 1$, the problem of quantum signal processing (QSP) [14]

is to find a set of parameters (called phase factors) $\Phi := (\phi_0, \dots, \phi_d) \in [-\pi, \pi)^{d+1}$ so that

$$f(x) = g(x, \Phi) := \text{Re}[\langle 0|U(x, \Phi)|0\rangle], \quad x \in [-1, 1], \quad (1)$$

with

$$U(x, \Phi) := e^{i\phi_0 Z} e^{i \arccos(x) X} e^{i\phi_1 Z} e^{i \arccos(x) X} \dots e^{i\phi_{d-1} Z} e^{i \arccos(x) X} e^{i\phi_d Z}. \quad (2)$$

Here

$$X := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Z := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (3)$$

are Pauli X and Z matrices, respectively. $\langle 0|U(x, \Phi)|0\rangle$ stands for the upper-left entry of the matrix $U(x, \Phi) \in \text{SU}(2)$, and $g(x, \Phi)$ defines a mapping from $[-1, 1]$ to $\mathbb{R}$ for fixed Φ . When the phase factors are restricted to be symmetric, *i.e.*,

$$\Phi = (\phi_0, \phi_1, \phi_2, \dots, \phi_2, \phi_1, \phi_0) \in [-\pi, \pi)^{d+1}, \quad (4)$$

this is referred to as the symmetric quantum signal processing. The simplest example is $\Phi = (0, \dots, 0)$. This gives $U(x, \Phi) = e^{i d \arccos(x) X}$ and $g(x, \Phi) = \cos(d \arccos(x)) = T_d(x)$, where T_d is the Chebyshev polynomial of the first kind of degree d .

Due to the parity constraint, the number of degrees of freedom in the target polynomial $f(x)$ is $\tilde{d} := \lceil \frac{d+1}{2} \rceil$. Hence $f(x)$ is entirely determined by its values on $\tilde{d}$ distinct points. Throughout the paper, we choose these points to be $x_k = \cos\left(\frac{2k-1}{4\tilde{d}}\pi\right)$, $k = 1, \dots, \tilde{d}$, *i.e.*, positive nodes of the Chebyshev polynomial $T_{2\tilde{d}}(x)$. For any target polynomial $f(x)$ defined above, the solution to Eq. (1) exists [14, 6], and Ref. [5] suggests that the solution can be restricted to be symmetric. The existence of the solution implies that the problem of symmetric quantum signal processing can be equivalently solved via the following optimization problem

$$\Phi^* = \underset{\substack{\Phi \in [-\pi, \pi)^{d+1}, \\ \text{symmetric.}}}{\text{argmin}} F(\Phi), \quad F(\Phi) := \frac{1}{\tilde{d}} \sum_{k=1}^{\tilde{d}} |g(x_k, \Phi) - f(x_k)|^2, \quad (5)$$

i.e., any solution Φ^* to Eq. (1) achieves the global minimum of the cost function with $F(\Phi^*) = 0$, and vice versa.

However, the energy landscape of the cost function $F(\Phi)$ is very complex, and has numerous global as well as local minima (see Section 7). This is already the case with the symmetry constraint. (Without the symmetry constraint, the number of variables is larger than the number of equations and there should be an infinite number of global minima.) Starting from a random initial guess, an optimization algorithm can easily be trapped at a local minima already when d is small¹. It is therefore surprising that starting from a special symmetric initial guess

$$\Phi^0 = (\pi/4, 0, 0, \dots, 0, 0, \pi/4), \quad (6)$$

which is independent of the target function, at least one global minimum can be robustly identified using standard unconstrained optimization algorithms even when d is as large as 10,000 [5], and the optimization method is free from being trapped by any local minima. Direct calculation shows that $g(x, \Phi^0) = 0$, and therefore Φ^0 does not contain any *a priori* information of the target polynomial $f(x)$!

¹An early attempt (not using optimization-based approach) showed that even finding Φ with $d > 30$ can be very costly [4, Appendix H.3].

1.1 Main results

In this work, we provide a theory to characterize the energy landscape of the cost function $F(\Phi)$, and to explain this unusual behavior of numerical optimization. Under the assumptions on the maxnorm $\|f\|_\infty$ as stated in Corollary 7, this leads to the first provable algorithm for solving the QSP problem without referring to extended precision arithmetic operations [8].

Let the domain of the symmetric phase factors be

$$D_d = \begin{cases} [-\frac{\pi}{2}, \frac{\pi}{2}]^{\frac{d}{2}} \times [-\pi, \pi) \times [-\frac{\pi}{2}, \frac{\pi}{2}]^{\frac{d}{2}}, & d \text{ is even,} \\ [-\frac{\pi}{2}, \frac{\pi}{2}]^{d+1}, & d \text{ is odd.} \end{cases} \quad (7)$$

Our first main result is the existence and uniqueness of symmetric phase factors for a class of polynomial matrices in $SU(2)$.

Theorem 1 (Existence and uniqueness of symmetric phase factors). *Consider any $P \in \mathbb{C}[x]$ and $Q \in \mathbb{R}[x]$ satisfying the following conditions*

- (1) $\deg(P) = d$ and $\deg(Q) = d - 1$.
- (2) P has parity $(d \bmod 2)$ and Q has parity $(d - 1 \bmod 2)$.
- (3) (Normalization condition) $\forall x \in [-1, 1] : |P(x)|^2 + (1 - x^2)|Q(x)|^2 = 1$.
- (4) If d is odd, then the leading coefficient of Q is positive.

There exists a unique set of symmetric phase factors $\Phi := (\phi_0, \phi_1, \dots, \phi_1, \phi_0) \in D_d$ such that

$$U(x, \Phi) = \begin{pmatrix} P(x) & iQ(x)\sqrt{1-x^2} \\ iQ(x)\sqrt{1-x^2} & P^*(x) \end{pmatrix}. \quad (8)$$

This result is a refinement of [6, Theorem 4], which only shows the existence of the phase factors without symmetry constraints. Theorem 1 states that in the presence of the symmetric constraint, the phase factors are also unique. Besides, under some constraint, the converse of Theorem 1 also holds (see Remark 12). It follows that there is a bijection between the global minimizers of Eq. (5) and all possible pairs of $(P(x), Q(x))$ satisfying the assumption in Theorem 1 and $\text{Re}[P](x) = f(x)$. The conditions (1), (2) for the target polynomial f are compatible with the first two requirements in Theorem 1. The condition (3) on the maxnorm, i.e., $\|f\|_\infty < 1$ is compatible with the normalization condition, which is itself a natural condition due to the unitarity of $U(x, \Phi)$.

To simplify the discussion, we introduce the following definition of admissible pair of polynomials associated with a target polynomial. Here $P_{\text{Re}} := \text{Re}[P]$, $P_{\text{Im}} := \text{Im}[P]$.

Definition 2 (Admissible pair of polynomials). *Let $f \in \mathbb{R}[x]$ be a target polynomial satisfying (1) $\deg(f) = d$, (2) the parity of f is $(d \bmod 2)$, (3) $\|f\|_\infty < 1$. Then (P, Q) is an admissible pair of polynomials associated with f if the conditions (1)-(3) in Theorem 1 are satisfied, together with*

- (4) The leading coefficient of Q is positive,
- (5) $P_{\text{Re}}(x) = f(x)$.

The polynomials $P_{\text{Im}}, Q \in \mathbb{R}[x]$ are also called complementary polynomials to f .

Comparing Theorem 1 and Definition 2, the main modification is that the leading coefficient of the complementary polynomial Q is always restricted to be positive. This allows us to unify the discussion of odd and even values of d . The bijection relation can then be concisely formulated as follows.

Corollary 3 (Bijection between global minima and admissible pairs). *If d is odd, there is a bijection between the global minima of Eq. (5) and all admissible pairs (P, Q) .*

If d is even, there is a bijection between the global minima of Eq. (5) and all pairs of polynomials $(P, \pm Q)$, where (P, Q) is an admissible pair.

The proof of Theorem 1 is constructive. So given an admissible pair, we have an algorithm to evaluate the symmetric phase factor. Theorem 4 generalizes the result of [8, Lemma 4] by explicitly constructing *all* admissible pairs. Together with Theorem 1, we have a complete description of all global minima of the cost function in Eq. (5).

Theorem 4 (Construction of admissible pairs). *Given a target polynomial $f(x)$, all admissible pairs (P, Q) must take the following form,*

$$\begin{aligned} P_{\text{Im}}(x) &= \sqrt{\alpha} \frac{e(x + i\sqrt{1-x^2}) + e(x - i\sqrt{1-x^2})}{2}, \\ Q(x) &= \sqrt{\alpha} \frac{e(x + i\sqrt{1-x^2}) - e(x - i\sqrt{1-x^2})}{2i\sqrt{1-x^2}}. \end{aligned} \quad (9)$$

Here, $e(z) := z^{-d} \prod_{i=1}^{2d} (z - r_i)$, where each r_i ($i = 1, \dots, 2d$) is a root of the function

$$\mathfrak{F}(z) := 1 - \left[f\left(\frac{z + z^{-1}}{2}\right) \right]^2, \quad (10)$$

and $\alpha \in \mathbb{C}$ satisfies $\mathfrak{F}(z) = \alpha e(z)e(z^{-1})$.

Note that the set $\{r_i\}_{i=1}^{2d}$ does not include all roots of the *Laurent polynomial* $\mathfrak{F}(z)$, which has $4d$ roots in total. By properly choosing the roots (see a more detailed description in Remark 17), we can construct all the admissible pairs. The number of global minima is finite, which is a consequence of the compactness of the domain D_d . Theorem 4 shows that the number of global minima can grow combinatorially with respect to d .

Unfortunately, the procedure described above for finding phase factors is numerically unstable. It requires extended precision arithmetic operations and is therefore very expensive when d is large (see detailed discussions in Section 1.2). This is noticeably different from solving the optimization problem in Eq. (5), which is numerically stable and can be readily performed using standard double precision arithmetic operations.

Among the myriad of global minima, Theorem 4 allows us to identify a special global minimum, which is obtained by choosing $\{r_i\}_{i=1}^{2d}$ to be the roots of $\mathfrak{F}(z)$ within the unit disc. The unique symmetric phase factor associated with this admissible pair is referred to as the *maximal solution* (the reason for the naming is technical and is explained in Section 4.2).

The maximal solution enjoys many desirable properties. For any target polynomial f with $\|f\|_\infty \leq \frac{1}{2}$, the maximal solution lies in the neighborhood of Φ^0 , which is defined in Eq. (6). To characterize this property, we first introduce the definition of reduced phase factors. Given any symmetric phase factors Φ of length $d+1$, reduced phase factors is referred to the left half part of Φ , namely $(\phi_0, \phi_1, \dots, \phi_{\tilde{d}-1})$, and is denoted as $\tilde{\Phi}$. Recall

that $\tilde{d} := \left\lceil \frac{d+1}{2} \right\rceil$. Due to symmetry, Φ is fully determined by $\tilde{\Phi}$. Hence, we may adopt Euclidean distance between $\tilde{\Phi}^*$ and $\tilde{\Phi}^0$ instead of that between Φ^* and Φ^0 . Furthermore, throughout this paper, we choose $\tilde{\Phi}$ as the free variables. With some abuse of notation, we identify the cost function $F(\Phi)$ with $F(\tilde{\Phi})$.

Theorem 5 (Distance between the maximal solution and Φ^0). *Let Φ^* be the maximal solution for the target function $f(x)$. Denote $\tilde{\Phi}^*$ and $\tilde{\Phi}^0$ as the corresponding reduced phase factors of Φ^* and Φ^0 respectively. If $\|f(x)\|_\infty \leq \frac{1}{2}$, then*

$$\|\tilde{\Phi}^* - \tilde{\Phi}^0\|_2 \leq \frac{\pi}{\sqrt{3}} \|f(x)\|_\infty. \quad (11)$$

In particular, if the target polynomial is $f = 0$, then the maximal solution is the initial guess Φ^0 . We then prove that when $\|f\|_\infty$ is sufficiently small, the maximal solution belongs to a neighborhood of Φ^0 , on which the cost function $F(\tilde{\Phi})$ is strongly convex, i.e., the Hessian matrix denoted by $\text{Hess}(\tilde{\Phi})$ is positive definite.

Theorem 6 (Local strong convexity). *If the target polynomial satisfies $\|f\|_\infty \leq \frac{\sqrt{3}}{20\pi d}$, for any symmetric phase factors Φ of length $d+1$ satisfying $\|\tilde{\Phi} - \tilde{\Phi}^0\|_2 \leq \frac{1}{20d}$, the following estimate holds:*

$$\frac{1}{4} \leq \lambda_{\min}(\text{Hess}(\tilde{\Phi})) \leq \lambda_{\max}(\text{Hess}(\tilde{\Phi})) \leq \frac{25}{4}. \quad (12)$$

The local strong convexity result in Theorem 6 immediately implies that when $\|f\|_\infty$ is sufficiently small, standard optimization algorithms, such as the projected gradient method [1, 11], can converge in the neighborhood of Φ^0 , without being trapped by any local minima.

Corollary 7 (Convergence of projected gradient method). *If the target polynomial satisfies $\|f\|_\infty \leq \frac{\sqrt{3}}{20\pi d}$, starting from Φ^0 , the projected gradient method with step size $t = \frac{1}{L}$ converges exponentially to the maximal solution Φ^* , i.e., at the ℓ -th iteration*

$$\|\tilde{\Phi}^\ell - \tilde{\Phi}^*\|_2^2 \leq e^{-\frac{\sigma}{L}\ell} \|\tilde{\Phi}^0 - \tilde{\Phi}^*\|_2^2. \quad (13)$$

Here $\sigma = \frac{1}{4}$ and $L = \frac{25}{4}$.

Therefore when $\|f\|_\infty \leq \frac{\sqrt{3}}{20\pi d}$, in order to reach precision ϵ , the projected gradient method can be terminated after $\mathcal{O}(\log(1/\epsilon))$ steps independent of the details of f . Since each iteration is numerically stable, the maximal solution can be readily obtained using standard double precision arithmetic operations in practice.

1.2 Background and related works

The mapping in Eqs. (1) and (2) may seem a very peculiar way for encoding a polynomial (and it is). However, such an $\text{SU}(2)$ representation can be directly translated into a quantum circuit, which is so far the most concise way for performing eigenvalue transformation $f(A)$ (when A is an Hermitian matrix) [14], and singular value transformations $f^{\text{SV}}(A)$ (when A is a general matrix) on quantum computers [6]. Many tasks in quantum computation can be formulated using such transformations. When f is not a polynomial, it can be approximated by a polynomial with bounded precision. This strategy can be used to unify a large class of quantum algorithms including Grover's search and quantum phase estimation [17, 16], and to perform a wide range of computational tasks, such as solving

linear system of equations [17, 13], eigenvalue problems [12], Hamiltonian simulation [14, 6] etc.

After the initial efforts [14, 4], significant progress has been made by using direct methods to obtain phase factors [6, 8]. These methods are based on finding roots of high degree polynomials to high precision and are not numerically stable. Specifically, these algorithms require $\mathcal{O}(d \log(d/\epsilon))$ bits of precision, where d is the degree of $f(x)$ and ϵ is the target accuracy [8]. It is worth mentioning that the extended precision needed in these algorithms is not an artifact of the proof technique. For instance, for $d \approx 500$, the number of bits needed to represent each floating point number can be as large as $1000 \sim 2000$ [5], which is much larger than the 64 bits provided by standard double precision floating point format.

There have been two recent improvements of the factorization based method, based on the capitalization method [3], and the Prony method [19], respectively. Although the two methods differ significantly, empirical results indicate that both methods are numerically stable, and are applicable to polynomials of large degrees. The algorithm in [3] is a variation of the algorithm in [6], which introduces a small perturbation to the high order Chebyshev coefficients to enhance the numerical stability. The Prony method in [19] avoids the root finding of high degree polynomials by directly constructing the admissible pairs for a given target polynomial, and uses randomness to enhance the numerical stability. Ref. [5] proposes a very different optimization based algorithm and suggests the use of the symmetric phase factors. This work provides a detailed discussion of the analytic structure of quantum signal processing with symmetric phase factors, and demonstrates the effectiveness of the optimization based approach under the condition that $\|f\|_\infty$ is sufficiently small. This provides a partial solution to the open problem of finding phase factors using numerically stable algorithms, and justifies the usage of standard double precision arithmetic operations in practice.

1.3 Discussion and open questions

Without the symmetry constraint, the optimization problem in Eq. (5) has d degrees of freedom. This is larger than the number of degrees of freedom of target function f , which is only $\tilde{d}$ due to the parity constraint. As a result, the Hessian matrix is always singular, preventing us from proving the exponential convergence of optimization methods [18].

Our analytical result is the first result trying to explain the unexpected success of optimization algorithms in the problem of finding phase factors. In Theorem 6, we prove the local strong convexity of the cost function when $\|f\|_\infty = \mathcal{O}(d^{-1})$. Therefore for a given target polynomial, in order to apply Corollary 7, we need to first multiply the target polynomial by a scaling factor $c = \mathcal{O}(d^{-1})$ and find the phase factors for $cf(x)$ instead. Such a scaling factor can be undesirable and may lead to suboptimal query complexities when implementing certain quantum algorithms. On the other hand, numerical results suggest that optimization algorithms can robustly converge to a global minimum when $\|f\|_\infty$ is smaller than a constant (less than 1) that is independent of d . Note that the estimate in Theorem 5 is already independent of d . Hence the gap is mainly due to Theorem 6, where the eigenvalue estimate of the Hessian matrix is obtained via the estimate of each matrix entry, which introduces the d -dependence. It is a natural open question to prove the strong convexity of the cost function near Φ^0 where $\|f\|_\infty$ can be independent of d .

The proof of Corollary 7 requires the projection of $\tilde{\Phi}^\ell$ to the domain $\{\tilde{\Phi} : \|\tilde{\Phi} - \tilde{\Phi}^0\|_2 \leq \frac{1}{20d}\}$ at each iteration. This is because the steepest descent method may potentially overshoot and deviate away from the locally strong convex region. Numerical observation

shows that unconstrained optimizers, such as quasi-Newton methods can robustly achieve the global minimum, which is beyond our current theoretical analysis.

Finally, since Φ^0 is only one of the solutions when $f = 0$, we may ask what if we run the projected gradient method starting from other solutions when $f = 0$, and to obtain optimal phase factors for more general target polynomials. These solutions correspond to non-maximal solutions. Numerical observations show that the convergence rate towards such non-maximal solutions can be significantly slower than that towards the maximal solution (see Figs. 5 and 6). The analysis of such behavior requires the generalization of estimates of in Theorems 5 and 6 to non-maximal solutions. These will be our future works.

Acknowledgments:

This work was partially supported by the NSF Quantum Leap Challenge Institute (QLCI) program through grant number OMA-2016245 (J.W., Y.D.), by Department of Energy under Grant No. DE-SC0017867 and No. DE-AC02-05CH11231 (L.L.). L.L. is a Simons Investigator.

2 Preliminaries

2.1 Chebyshev polynomials

We first collect some basic facts of Chebyshev polynomials, which is used to formulate the optimization problem in Eq. (5) and will be extensively used in Sections 5 and 6. Given $x \in [-1, 1]$, the Chebyshev polynomial of the first kind is $T_n(x) = \cos(n \arccos(x))$, and that of the second kind is given by $U_n(x) \sin(\arccos(x)) = \sin((n+1) \arccos(x))$.

A useful fact is that both T_n and U_n form a sequence of orthogonal polynomials. The Chebyshev polynomial of the first kind T_n are orthogonal with respect to the weight $\frac{1}{\sqrt{1-x^2}}$ over $[-1, 1]$, and that of the second kind U_n are orthogonal with respect to the weight $\sqrt{1-x^2}$ over $[-1, 1]$,

$$\begin{aligned} \frac{1}{\pi} \int_{-1}^1 T_n(x) T_m(x) \frac{1}{\sqrt{1-x^2}} dx &= \frac{1}{2} \delta_{nm} (1 + \delta_{n0}), \\ \frac{1}{\pi} \int_{-1}^1 U_n(x) U_m(x) \sqrt{1-x^2} dx &= \frac{1}{2} \delta_{nm}. \end{aligned}$$

The induced norms

$$\|f\|_T^2 := \frac{1}{\pi} \int_{-1}^1 f^2(x) \frac{1}{\sqrt{1-x^2}} dx, \quad \|f\|_U^2 := \frac{1}{\pi} \int_{-1}^1 f^2(x) \sqrt{1-x^2} dx$$

are referred to as the Chebyshev norm of the first kind and the second kind respectively. Furthermore, any function h with $\|h\|_T < \infty$ can be uniquely expressed as a series of Chebyshev polynomials of the first kind,

$$h(x) = \sum_{n=0}^{\infty} c_n T_n(x), \quad c_n = \frac{2}{\pi(1 + \delta_{n0})} \int_{-1}^1 h(x) T_n(x) \frac{1}{\sqrt{1-x^2}} dx. \quad (14)$$

Similarly, any function h with $\|h\|_U < \infty$ can be uniquely expressed as a series of Chebyshev polynomials of the second kind,

$$h(x) = \sum_{n=0}^{\infty} c_n U_n(x), \quad c_n = \frac{2}{\pi} \int_{-1}^1 h(x) U_n(x) \sqrt{1-x^2} dx. \quad (15)$$

Apart from the weighted orthogonality, Chebyshev polynomials of the first kind T_n satisfies the discrete orthogonality condition:

$$\sum_{j=1}^{2\tilde{d}} T_n(x_j) T_m(x_j) = \begin{cases} 0 & n \neq m \\ 2\tilde{d} & n = m = 0 \\ \tilde{d} & n = m \neq 0 \end{cases}, \quad (16)$$

where $x_j = \cos\left(\pi \frac{2j-1}{4\tilde{d}}\right)$ are the nodes of $T_{2\tilde{d}}$ and $\max(n, m) \leq 2\tilde{d}$. This set of points is also referred to as Chebyshev nodes. Furthermore, if n and m have the same parity which implies $T_n(-x)T_m(-x) = T_n(x)T_m(x)$, we have

$$\sum_{j=1}^{\tilde{d}} T_n(x_j) T_m(x_j) = \frac{1}{2} \sum_{j=1}^{2\tilde{d}} T_n(x_j) T_m(x_j) = \begin{cases} 0 & n \neq m \\ \tilde{d} & n = m = 0 \\ \frac{\tilde{d}}{2} & n = m \neq 0 \end{cases}. \quad (17)$$

2.2 Notation

Given a positive integer n , $\mathbb{K}_n[x]$ denotes the set of all polynomials of degree at most n with variable x , and coefficients taken in the field $\mathbb{K}$. The Laurent polynomial is a linear combination of positive as well as negative powers of the variable x with coefficients in the field $\mathbb{K}$, and $\mathbb{K}[x, x^{-1}]$ denotes the set of all Laurent polynomials. Specifically, we are only interested in the cases where $\mathbb{K}$ is either $\mathbb{R}$ or $\mathbb{C}$. Unless otherwise noted, the term “polynomial” refers to real polynomial or complex polynomial, which only have nonnegative power of the variable x .

We assume that the length of the set of phase factors is $d+1$ unless otherwise noted. We also assume that a pair of polynomials $(P, Q) \in \mathbb{C}_d[x] \times \mathbb{C}_{d-1}[x]$ satisfies the conditions in Theorem 1. Given a polynomial $P \in \mathbb{C}[x]$, we define $P_{\text{Re}}(x) := \text{Re}[P(x)]$ and $P_{\text{Im}}(x) := \text{Im}[P(x)]$. The real polynomial $P_{\text{Re}}(x)$ is usually denoted by $f(x)$ when it refers to the target of the optimization.

The QSP problem involves some delicate relations of polynomial coefficients. Throughout the paper, we will interchangeably expand a polynomial using the monomial basis and the Chebyshev basis, *i.e.*,

$$f(x) = \sum_{j=0}^d \mathbf{f}_j x^j = \sum_{j=0}^d f_j T_j(x), \quad \mathbf{f}_j, f_j \in \mathbb{R}. \quad (18)$$

Similarly,

$$\begin{aligned} P(x) &= \sum_{j=0}^d \mathbf{p}_j x^j = \sum_{j=0}^d p_j T_j(x), \quad \mathbf{p}_j, p_j \in \mathbb{C}, \\ P_{\text{Im}}(x) &= \sum_{j=0}^d \mathbf{s}_j x^j = \sum_{j=0}^d s_j T_j(x), \quad \mathbf{s}_j, s_j \in \mathbb{R}, \\ Q(x) &= \sum_{j=0}^{d-1} \mathbf{q}_j x^j = \sum_{j=0}^{d-1} q_j U_j(x) \quad \mathbf{q}_j, q_j \in \mathbb{C}. \end{aligned} \quad (19)$$

For convenience, we set the values of $\mathbf{f}_j, f_j, \mathbf{p}_j, p_j, \mathbf{s}_j, s_j, \mathbf{q}_j, q_j$ with a negative integer index to zero.

The coefficients in the monomial and Chebyshev basis satisfy the relation

$$\mathbf{f}_j = 2^{j-1}f_j, \quad \mathbf{p}_j = 2^{j-1}p_j, \quad \mathbf{s}_d = 2^{j-1}s_j, \quad \mathbf{q}_j = 2^j q_j, \quad \forall j \geq 0. \quad (20)$$

The set $[n] := \{0, 1, \dots, n-1\}$ is referred to as the index set generated by a positive integer n . The row and column indices of a n -by- n matrix run from 0 to $n-1$, namely in the index set $[n]$. For a matrix $A \in \mathbb{C}^{m \times n}$, the transpose, Hermitian conjugate and complex conjugate are denoted by $A^\top$, $A^\dagger$, A^* , respectively. The same notations are also used for the operations on a vector. We use the convention in quantum computing to write the basis vectors in $\mathbb{C}^2$. It is equivalent to the following identification

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \langle 0| = |0\rangle^\dagger = (1, 0), \quad \text{and} \quad \langle 1| = |1\rangle^\dagger = (0, 1).$$

Given a 2-by-2 matrix $A = (a_{ij})$, we have $\langle i|A|j\rangle = a_{ij}$ for any $i, j \in \{0, 1\}$.

For a matrix $A \in \mathbb{C}^{m \times n}$, we denote the operator norm $\|\cdot\|_2$ induced by the vector ℓ_2 norm as

$$\|A\|_2 := \sup_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2}, \quad (21)$$

and denote the Frobenius norm $\|\cdot\|_F$ as

$$\|A\|_F := \sqrt{\sum_{i \in [m], j \in [n]} |a_{ij}|^2}. \quad (22)$$

We use $\sigma_k(A)$ to denote its $(k+1)$ -th largest singular value, *i.e.*,

$$0 \leq \sigma_{\min(m,n)-1}(A) \leq \dots \leq \sigma_1(A) \leq \sigma_0(A).$$

We define $\sigma_{\min}(A) := \sigma_{\min(m,n)-1}(A)$ and $\sigma_{\max}(A) := \sigma_0(A)$. Here, the index k starts from 0 which agrees with the convention of matrix index.

For a Hermitian matrix $A \in \mathbb{C}^{n \times n}$, we use the notation $\lambda_k(A)$ to designate the $(k+1)$ -th largest eigenvalue, *i.e.*,

$$\lambda_{n-1}(A) \leq \dots \leq \lambda_1(A) \leq \lambda_0(A).$$

We define $\lambda_{\min}(A) := \lambda_{n-1}(A)$ and $\lambda_{\max}(A) := \lambda_0(A)$.

2.3 Quantum signal processing

In the absence of the symmetry constraint, the existence of phase factors for a unitary representation of a pair of polynomials (P, Q) is proved in [6].

Theorem 8 (Quantum signal processing [6, Theorem 4]). *Let $d \in \mathbb{N}$. There exists a set of phase factors $\Phi := (\phi_0, \dots, \phi_d) \in \mathbb{R}^{d+1}$ such that*

$$U(x, \Phi) = e^{i\phi_0 Z} \prod_{j=1}^d \left(W(x) e^{i\phi_j Z} \right) = \begin{pmatrix} P(x) & iQ(x)\sqrt{1-x^2} \\ iQ^*(x)\sqrt{1-x^2} & P^*(x) \end{pmatrix}, \quad (23)$$

where

$$W(x) = e^{i \arccos(x) X} = \begin{pmatrix} x & i\sqrt{1-x^2} \\ i\sqrt{1-x^2} & x \end{pmatrix},$$

if and only if $P, Q \in \mathbb{C}[x]$ satisfy that

- (1) $\deg(P) \leq d, \deg(Q) \leq d-1$,
- (2) $P(x)$ has parity $(d \bmod 2)$ and $Q(x)$ has parity $(d-1 \bmod 2)$,
- (3) $|P(x)|^2 + (1-x^2)|Q(x)|^2 = 1, \forall x \in [-1, 1]$.

When the real component of P is of interest only, the conditions for the existence of phase factors can be further simplified.

Corollary 9 (Real quantum signal processing [6, Corollary 5]). *Let $f \in \mathbb{R}[x]$ be a degree- d polynomial for some $d \geq 1$ such that*

- $f(x)$ has parity $(d \bmod 2)$,
- $|f(x)| < 1, \forall x \in [-1, 1]$.

Then there exists some $P, Q \in \mathbb{C}[x]$ satisfying properties (1)–(3) of Theorem 8 such that $f(x) = \text{Re}[P(x)]$.

2.4 Quantum signal processing as an optimization problem

The existence result in Corollary 9 enables the following optimization based strategy to find phase factors. Given the target polynomial $f(x) \in \mathbb{R}[x]$, we may minimize the $\mathcal{L}_2$ distance between $g(x, \Phi)$ and target function $f(x)$, i.e.,

$$L(\Phi) = \left(\int_{-1}^1 [g(x, \Phi) - f(x)]^2 dx \right)^{\frac{1}{2}}. \quad (24)$$

The desired phase factors Φ^* satisfy $L(\Phi^*) = 0$.

Since both $g(x, \Phi)$ and $f(x)$ are polynomials of degree d with definite parity, the number of degrees of freedom is $\tilde{d} := \lceil \frac{d+1}{2} \rceil$. These polynomials are determined by their values at $x_k = \cos\left(\frac{2k-1}{4d}\pi\right)$, $k = 1, \dots, \tilde{d}$, which are positive Chebyshev nodes of $T_{2\tilde{d}}(x)$. This leads to the following optimization problem,

$$\Phi^* = \underset{\Phi \in [-\pi, \pi]^{d+1}}{\text{argmin}} F(\Phi), \quad F(\Phi) := \frac{1}{\tilde{d}} \sum_{k=1}^{\tilde{d}} |g(x_k, \Phi) - f(x_k)|^2. \quad (25)$$

3 Symmetric quantum signal processing

In this section, we present some basic properties of symmetric quantum signal processing. We first explain a constraint of the sign of the leading coefficient of $Q(x)$ when d is odd. Then we introduce a constructive method for finding phase factors given suitable polynomials (P, Q) . In the end, we prove Theorem 1, which establishes a bijection between the global minimizers of Eq. (5) and all suitable complementary polynomials (P_{Im}, Q) .

3.1 Sign of the leading coefficient of $Q(x)$

We first derive an explicit expression of the coefficients of T_d and U_{d-1} in the Chebyshev expansion of P and Q respectively.

Lemma 10. *Let $\Phi \in \mathbb{R}^{d+1}$ be a set of phase factors, $P(x)$ and $Q(x)$ be defined by Eq. (23). Considering the Chebyshev expansion of P and Q in Eq. (19), one has*

$$p_d = e^{i(\phi_0 + \phi_d)} \prod_{j=1}^{d-1} \cos(\phi_j), \quad q_{d-1} = e^{i(\phi_0 - \phi_d)} \prod_{j=1}^{d-1} \cos(\phi_j). \quad (26)$$

Proof. See Appendix A. $\square$

After imposing the symmetry constraint on phase factors, we find that there is an additional constraint on the leading coefficient of Q when d is odd. It originates from the observation that when Φ is symmetric, one has $\phi_0 = \phi_d$ and then the second equality in Eq. (26) becomes

$$q_{d-1} = \prod_{j=1}^{d-1} \cos(\phi_j) = \prod_{j=1}^{\tilde{d}-1} \cos^2(\phi_j) \geq 0. \quad (27)$$

Note that $\deg(Q)$ may be less than $d-1$. The more precise statement is given in Lemma 11.

Lemma 11. *Let $\Phi \in \mathbb{R}^{d+1}$ be a set of phase factors, $P(x)$ and $Q(x)$ be defined by Eq. (23). If Φ is symmetric and d is odd, then the leading coefficient of Q in the Chebyshev basis has the same sign as $(-1)^{\frac{d-1-\deg(Q)}{2}}$.*

Proof. See Appendix B. $\square$

As a remark, the statement in Lemma 11 also holds with the Chebyshev basis replaced by the monomial basis, which is an consequence of directly applying the relation in Eq. (20).

Remark 12. *By directly applying Lemma 11 and Theorem 8, one gets that the converse of Theorem 1 also holds, given that $\deg(P) = d$ and $\deg(Q) = d-1$.*

3.2 Existence of symmetric phase factors

We present a useful lemma, which evaluates the outmost symmetric phase factor, and reduces the polynomial degree by 2. It also connects the leading coefficients in the monomial basis before and after the reduction. The result of Lemma 13 will be used throughout the paper. To clarify, we only consider the expansion of polynomial in the monomial basis in rest of this section.

Lemma 13. *Suppose that $P(x) \in \mathbb{C}[x]$, $Q(x) \in \mathbb{R}[x]$ satisfy the conditions in Theorem 1 and $d \geq 2$. Choose*

$$e^{2i\phi_0} = \frac{\mathbf{q}_{d-1}}{\mathbf{p}_d^*} \quad (28)$$

and define

$$\begin{aligned} & \begin{pmatrix} P^{(1)}(x) & iQ^{(1)}(x)\sqrt{1-x^2} \\ iQ^{(1)}(x)\sqrt{1-x^2} & (P^{(1)})^*(x) \end{pmatrix} \\ & := W^{-1} e^{-i\phi_0 Z} \begin{pmatrix} P(x) & iQ(x)\sqrt{1-x^2} \\ iQ(x)\sqrt{1-x^2} & P^*(x) \end{pmatrix} e^{-i\phi_0 Z} W^{-1}. \end{aligned} \quad (29)$$

Consider the expansion of P , Q , $P^{(1)}$ and $Q^{(1)}$ in the monomial basis,

$$\begin{aligned} P(x) &= \sum_{i=0}^d \mathbf{p}_i x^i, & Q(x) &= \sum_{j=0}^{d-1} \mathbf{q}_j x^j. \\ P^{(1)}(x) &= \sum_{i=0}^{d+2} \mathbf{p}_i^{(1)} x^i, & Q^{(1)}(x) &= \sum_{j=0}^{d+1} \mathbf{q}_j^{(1)} x^j. \end{aligned} \quad (30)$$

Then $Q^{(1)}(x) \in \mathbb{R}[x]$ and $P^{(1)}, Q^{(1)}$ satisfy the conditions (1)-(3) in Theorem 1 with d replaced by $d-2$.

Furthermore, if $d = 2$, then

$$P^{(1)} = -\mathbf{p}_0^* e^{2i\phi_0} = -\left(\frac{\mathbf{p}_0}{\mathbf{q}_2}\right)^* \mathbf{q}_1, \quad Q^{(1)} = 0. \quad (31)$$

If $d \geq 3$, then

$$\begin{aligned} \mathbf{p}_{d-2}^{(1)} &= \frac{\mathbf{q}_{d-1}}{4} - \frac{1}{\mathbf{q}_{d-1}} \left| \text{Im}[\mathbf{p}_{d-2} e^{-2i\phi_0}] \right|^2 + i \text{Im}[\mathbf{p}_{d-2} e^{-2i\phi_0}], \\ \mathbf{q}_{d-3}^{(1)} &= \frac{\mathbf{q}_{d-1}}{4} + \frac{1}{\mathbf{q}_{d-1}} \left| \text{Im}[\mathbf{p}_{d-2} e^{-2i\phi_0}] \right|^2. \end{aligned} \quad (32)$$

And $P^{(1)}, Q^{(1)}$ also satisfy the condition (4) in Theorem 1.

Proof. See Appendix C. $\square$

For convenience we denote $(P^{(0)}, Q^{(0)}) = (P, Q)$. Inspired by Lemma 13, we choose ϕ_ℓ to satisfy

$$e^{2i\phi_\ell} = \frac{\mathbf{p}_{d-2\ell}^{(\ell)}}{\mathbf{q}_{d-1-2\ell}^{(\ell)}} \quad (33)$$

and repeat the construction

$$\begin{aligned} &\begin{pmatrix} P^{(\ell+1)}(x) & iQ^{(\ell+1)}(x)\sqrt{1-x^2} \\ iQ^{(\ell+1)}(x)\sqrt{1-x^2} & (P^{(\ell+1)})^*(x) \end{pmatrix} \\ &:= W^{-1} e^{-i\phi_\ell Z} \begin{pmatrix} P^{(\ell)}(x) & iQ^{(\ell)}(x)\sqrt{1-x^2} \\ iQ^{(\ell)}(x)\sqrt{1-x^2} & (P^{(\ell)})^*(x) \end{pmatrix} e^{-i\phi_\ell Z} W^{-1}. \end{aligned} \quad (34)$$

as long as $\deg(P^{(\ell)}) \geq 2$. Notice that

$$\deg(P^{(\ell)}) = \deg(P^{(\ell-1)}) - 2 = \deg(P^{(0)}) - 2\ell.$$

Thus, the iteration stops when $\ell = \tilde{d} - 1$. Moreover, $\deg(P^{(\tilde{d}-1)}) = 0$ if d is even, and $\deg(P^{(\tilde{d}-1)}) = 1$ if d is odd. For convenience, we introduce the notation of $\hat{d}$,

$$\hat{d} := \begin{cases} \tilde{d} - 1 & \text{if } d \text{ is odd,} \\ \tilde{d} - 2 & \text{if } d \text{ is even.} \end{cases} \quad (35)$$

Lemma 14. Suppose that $P(x), Q(x)$ satisfy the conditions in Theorem 1. Then there exists a set of symmetric phase factor $\Phi := (\phi_0, \phi_1, \dots, \phi_1, \phi_0) \in D_d$ such that Eq. (23) holds, where ϕ_ℓ satisfies

$$e^{2i\phi_\ell} = \frac{\mathbf{p}_{d-2\ell}^{(\ell)}}{\mathbf{q}_{d-1-2\ell}^{(\ell)}}, \quad \forall 0 \leq \ell \leq \hat{d}. \quad (36)$$

In particular, if d is even, $\phi_{\tilde{d}-1}$ satisfies

$$e^{i\phi_{\tilde{d}-1}} = \mathbf{p}_0^{(\tilde{d}-1)}. \quad (37)$$

Here $\mathbf{p}_{d-2\ell}^{(\ell)}$ and $\mathbf{q}_{d-1-2\ell}^{(\ell)}$ are the leading coefficient of the expansion of $P^{(\ell)}$ and $Q^{(\ell)}$ in the monomial basis respectively. We use the notation $(P^{(0)}, Q^{(0)}) = (P, Q)$ and define $(P^{(\ell)}, Q^{(\ell)})$ by Eq. (34) for $1 \leq \ell \leq \tilde{d} - 1$.

Proof. By repeating the construction in Eq. (34), we obtain a sequence of polynomials $(P^{(\ell)}, Q^{(\ell)})$ where $0 \leq \ell \leq \tilde{d} - 1$. Eq. (36) determines the value of ϕ_ℓ over D_d for $0 \leq \ell \leq \tilde{d} - 2$.

When d is odd, $\deg(P^{(\tilde{d}-1)}) = 1$ and $\deg(Q^{(\tilde{d}-1)}) = 0$. Lemma 13 implies that the leading coefficient of $Q^{(\tilde{d}-1)}$ is positive as well. By the normalization condition, $P^{(\tilde{d}-1)}$ must be of the form cx where $|c| = 1$ and $Q^{(\tilde{d}-1)}$ must be 1. Choose $\phi_{\tilde{d}-1} \in [-\frac{\pi}{2}, \frac{\pi}{2})$ according to Eq. (36) and one can check

$$\begin{pmatrix} P^{(\tilde{d}-1)}(x) & iQ^{(\tilde{d}-1)}(x)\sqrt{1-x^2} \\ iQ^{(\tilde{d}-1)}(x)\sqrt{1-x^2} & (P^{(\tilde{d}-1)})^*(x) \end{pmatrix} = e^{i\phi_{\tilde{d}-1}Z} W e^{i\phi_{\tilde{d}-1}Z}.$$

In this way, $\Phi = (\phi_0, \dots, \phi_{\tilde{d}-1}, \phi_{\tilde{d}-1}, \dots, \phi_0)$ satisfies the requirements.

When d is even, we know that $P^{(\tilde{d}-1)}$ is a constant of absolute value 1 and $Q^{(\tilde{d}-1)} = 0$, *i.e.*,

$$\begin{pmatrix} P^{(\tilde{d}-1)}(x) & iQ^{(\tilde{d}-1)}(x)\sqrt{1-x^2} \\ iQ^{(\tilde{d}-1)}(x)\sqrt{1-x^2} & (P^{(\tilde{d}-1)})^*(x) \end{pmatrix} = e^{i\phi_{\tilde{d}-1}Z},$$

where $\phi_{\tilde{d}-1}$ satisfies Eq. (37). Furthermore, we may choose $\phi_{\tilde{d}-1} \in [-\pi, \pi)$ and $\Phi = (\phi_0, \dots, \phi_{\tilde{d}-2}, \phi_{\tilde{d}-1}, \phi_{\tilde{d}-2}, \dots, \phi_0)$ satisfies the requirements. $\square$

3.3 Proof of Theorem 1

Now we are ready to prove Theorem 1.

Proof of Theorem 1. The existence is given by Lemma 14. We only need to show the uniqueness. If there exists another set of symmetric phase factors $\Psi := (\psi_0, \psi_1, \dots, \psi_1, \psi_0) \in D_d$ satisfying the requirement, we may consider

$$\begin{aligned} & \begin{pmatrix} P_\Psi^{(1)}(x) & iQ_\Psi^{(1)}(x)\sqrt{1-x^2} \\ iQ_\Psi^{(1)}(x)\sqrt{1-x^2} & (P_\Psi^{(1)})^*(x) \end{pmatrix} \\ &:= W^{-1} e^{-i\psi_0 Z} \begin{pmatrix} P(x) & iQ(x)\sqrt{1-x^2} \\ iQ(x)\sqrt{1-x^2} & P^*(x) \end{pmatrix} e^{-i\psi_0 Z} W^{-1}. \end{aligned}$$

Since $P_\Psi^{(1)}(x)$ can be obtained by a set of symmetric phase factors of length $d-2$, the coefficient of x^{d+2} of $P_\Psi^{(1)}$ is zero, *i.e.*,

$$\mathbf{p}_{d+2}^{(1)} = -2\mathbf{q}_{d-1} + \mathbf{p}_d^* e^{2i\psi_0} + \mathbf{p}_d e^{-2i\psi_0} = 0,$$

which is obtained by applying Eq. (103). Together with $|\mathbf{p}_d| = |\mathbf{q}_{d-1}|$, it leads to $e^{2i\psi_0} = \frac{\mathbf{q}_{d-1}}{\mathbf{p}_d^*}$ and this equation has only one solution over $[-\frac{\pi}{2}, \frac{\pi}{2})$, hence $\psi_0 = \phi_0$. Then $\Phi = \Psi$ follows inductively. $\square$

3.4 Optimization based method for finding symmetric phase factors

Ref. [5] suggests that the optimization problem Eq. (5) should be solved by imposing the symmetry constraint on the phase factors. This allows us to use the reduced phase factors $\tilde{\Phi} = (\phi_0, \phi_1, \dots, \phi_{\tilde{d}-1})$, as optimization variables. For any scalar valued function of the set

of reduced phase factors, we use subscripts to represent the partial derivative with respect to the corresponding phase factor, for example, $F_i(\tilde{\Phi}) := \partial_{\phi_i} F(\tilde{\Phi})$. We also use subscripts to represent the corresponding entry of any matrix. These notations and conventions can simplify the discussion in Sections 5 and 6.

The Hessian matrix of the problem in Eq. (5) can be written element-wise as

$$\begin{aligned} F_{ij}(\tilde{\Phi}) &= \frac{2}{\tilde{d}} \sum_{k=1}^{\tilde{d}} g_i(x_k, \tilde{\Phi}) g_j(x_k, \tilde{\Phi}) + \frac{2}{\tilde{d}} \sum_{k=1}^{\tilde{d}} [g(x_k, \tilde{\Phi}) - f(x_k)] g_{ij}(x_k, \tilde{\Phi}) \\ &= \frac{2}{\tilde{d}} (A^\top A)_{ij} + R_{ij}, \end{aligned} \quad (38)$$

where $A(\tilde{\Phi}) := [g_j(x_i, \tilde{\Phi})] \in \mathbb{R}^{\tilde{d} \times \tilde{d}}$ is the Jacobian matrix of $\tilde{\Phi} \mapsto (g(x_1, \tilde{\Phi}), \dots, g(x_{\tilde{d}}, \tilde{\Phi}))^\top$, and R is a real $\tilde{d} \times \tilde{d}$ matrix closely related to the residual with entry (note that R_{ij} does not mean the partial derivative here)

$$R_{ij}(\tilde{\Phi}) := \frac{2}{\tilde{d}} \sum_{k=1}^{\tilde{d}} [g(x_k, \tilde{\Phi}) - f(x_k)] g_{ij}(x_k, \tilde{\Phi}).$$

Due to the existence of phase factors stated in Theorem 1, at the optimal point $\tilde{\Phi}^*$, the second term vanishes, namely $R(\tilde{\Phi}^*) = 0$, and $\text{Hess}(\tilde{\Phi}^*) = \frac{2}{\tilde{d}} (A^*)^\top A^*$, where $A^* := A(\tilde{\Phi}^*)$ and $\tilde{\Phi}^*$ are the reduced phase factors corresponding to symmetric phase factors Φ^* .

4 Global minimizer

According to the last section, the problem of finding all global minima of the optimization problem is converted to that of finding all admissible pairs of (P, Q) associated with the target polynomial f . In this section, we provide a characterization of *all* possible complementary polynomials (P_{Im}, Q) .

4.1 Characterizing complementary polynomials

Before constructing all possible pairs of $(P(x), Q(x))$ satisfying Definition 2, we first introduce the following lemma, which views the problem from the perspective of Laurent polynomials.

Lemma 15. *Given any real polynomial $f(x)$ of degree d and with $\|f\|_\infty < 1$, $\mathfrak{F}(z) = 1 - \left(f\left(\frac{z+z^{-1}}{2}\right)\right)^2$ has a decomposition*

$$\mathfrak{F}(z) = \alpha \prod_{i=1}^{2d} (z - r_i)(z^{-1} - r_i), \quad (39)$$

such that $\alpha, r_i \in \mathbb{C} \setminus \{0\}$.

The decomposition is unique in the following sense: if $\tilde{\alpha} \in \mathbb{C}$ and $\tilde{r}_i \in \mathbb{C}$, $i = 1, \dots, n$ such that

$$\mathfrak{F}(z) = \tilde{\alpha} \prod_{i=1}^n (z - \tilde{r}_i)(z^{-1} - \tilde{r}_i), \quad (40)$$

then

$$n = 2d, \quad \alpha \prod_{i=1}^{2d} r_i = \tilde{\alpha} \prod_{i=1}^{2d} \tilde{r}_i, \quad \frac{r_i + r_i^{-1}}{2} = \frac{\tilde{r}_i + \tilde{r}_i^{-1}}{2}, \quad (41)$$

up to permutation of the roots.

Proof. According to the fundamental theorem of algebra, $1 - f(x)^2$ has a unique decomposition

$$1 - f(x)^2 = c \prod_{i=1}^{2d} (s_i - x), \quad s_i \in \mathbb{C}.$$

Since $\|f\|_\infty < 1$, $s_i \neq 0$. The equation $s = \frac{z+z^{-1}}{2}$ always has 2 nonzero solutions, say z_1, z_2 . In particular, $z_1 = z_2^{-1}$. For any s_i , pick one solution r_i , use equality

$$r_i(s_i - x) = \frac{1}{2}(z - r_i)(z^{-1} - r_i),$$

and one gets

$$\mathfrak{F}(z) = c \prod_{i=1}^{2d} \frac{1}{2r_i} (z - r_i)(z^{-1} - r_i) = \alpha \prod_{i=1}^{2d} (z - r_i)(z^{-1} - r_i),$$

where $\alpha = c \prod_{i=1}^{2d} \frac{1}{2r_i}$. It provides a decomposition of $\mathfrak{F}(z)$ satisfying the requirement.

On the other hand, given any decomposition

$$\mathfrak{F}(z) = \tilde{\alpha} \prod_{i=1}^n (z - \tilde{r}_i)(z^{-1} - \tilde{r}_i),$$

then one has

$$\mathfrak{F}(z) = \tilde{\alpha} \prod_{i=1}^n (2\tilde{r}_i) \left(\frac{\tilde{r}_i + \tilde{r}_i^{-1}}{2} - \frac{z + z^{-1}}{2} \right).$$

It provides a decomposition of $1 - f(x)^2$, *i.e.*,

$$1 - f(x)^2 = \tilde{\alpha} \left(\prod_{i=1}^n 2\tilde{r}_i \right) \prod_{i=1}^n \left(\frac{\tilde{r}_i + \tilde{r}_i^{-1}}{2} - x \right).$$

By the uniqueness decomposition of $1 - f(x)^2$ in $\mathbb{C}[x]$, we obtain Eq. (41) up to permutation of the roots. $\square$

In particular, we are interested in the following decomposition of $\mathfrak{F}(z)$, which is referred to as an *admissible decomposition*,

- $\mathfrak{F}(z) = \alpha \prod_{i=1}^{2d} (z - r_i)(z^{-1} - r_i)$, $\alpha > 0, r_i \in \mathbb{C} \setminus \{0\}$.
- Multiset $\mathcal{D} := \{r_i : 1 \leq i \leq 2d\}$ is closed under complex conjugation and additive inverse.

Firstly, $\mathcal{D}$ contains half of the roots of $\mathfrak{F}(z)$ with multiplicity (hence $\mathcal{D}$ is a multiset rather than set). Secondly, note that $\alpha > 0$ naturally comes from that $\alpha = \frac{1-f(1)^2}{\prod_{r \in \mathcal{D}} (1-r)^2} > 0$ for any $\mathcal{D}$ closed under complex conjugation. Another observation is that the roots of $\mathfrak{F}(z)$ cannot be any number of absolute value 1. Otherwise, $1 - f\left(\frac{r_i + r_i^{-1}}{2}\right)^2 = 0$, *i.e.*, $\|f\|_\infty = 1$, which contradicts the constraint on the maxnorm of f .

To be more precise, we rephrase Theorem 4 as follows and present its proof.

Theorem 16. *The pair of polynomials $(P(x), Q(x))$ satisfy Definition 2 if only if there exists an admissible decomposition of $\mathfrak{F}(z)$,*

$$\mathfrak{F}(z) = \alpha \prod_{i=1}^{2d} (z - r_i)(z^{-1} - r_i), \quad (42)$$

such that

$$\begin{aligned} P_{\text{Im}}(x) &= \sqrt{\alpha} \frac{e(x + i\sqrt{1-x^2}) + e(x - i\sqrt{1-x^2})}{2}, \\ Q(x) &= \sqrt{\alpha} \frac{e(x + i\sqrt{1-x^2}) - e(x - i\sqrt{1-x^2})}{2i\sqrt{1-x^2}}, \end{aligned} \quad (43)$$

where $e(z) := z^{-d} \prod_{i=1}^{2d} (z - r_i)$.

Proof. $\implies$: Denote $p_{\text{Im}}(z) := P_{\text{Im}}\left(\frac{z+z^{-1}}{2}\right)$ and $q(z) := Q\left(\frac{z+z^{-1}}{2}\right)$. Then the normalization condition gives

$$\begin{aligned} \mathfrak{F}(z) &= p_{\text{Im}}(z)^2 - q(z)^2 \left(\frac{z - z^{-1}}{2}\right)^2 \\ &= \left(p_{\text{Im}}(z) + \frac{z - z^{-1}}{2} q(z)\right) \left(p_{\text{Im}}(z) - \frac{z - z^{-1}}{2} q(z)\right). \end{aligned}$$

Define $H(z) := p_{\text{Im}}(z) + \frac{z - z^{-1}}{2} q(z)$. Since $p_{\text{Im}}(z) = p_{\text{Im}}(z^{-1})$ and $q(z) = q(z^{-1})$, one gets

$$\mathfrak{F}(z) = H(z)H(z^{-1}).$$

Notice that $z^d H(z)$ is a real polynomial of degree no more than $2d$. Thus $z^d H(z)$ can be represented as $c \prod_{r \in \mathcal{D}} (z - r)$, where $c \in \mathbb{C}$ and the multiset $\mathcal{D}$ contains all the roots of $z^d H(z)$ with multiplicity. Since $H(z)$ is real, $\mathcal{D}$ is closed under complex conjugation and $c \in \mathbb{R}$.

With $H(z) = cz^{-d} \prod_{r \in \mathcal{D}} (z - r)$, one gets

$$\mathfrak{F}(z) = H(z)H(z^{-1}) = c^2 \prod_{r \in \mathcal{D}} (z - r)(z^{-1} - r).$$

By Lemma 15, the cardinality of $\mathcal{D}$ is $2d$ and thus $z^d H(z)$ is of degree $2d$. Moreover, $r \in \mathbb{C} \setminus \{0\}$, $\forall r \in \mathcal{D}$.

Besides, we can check that the parity of $z^d H(z)$ is even by examining the parity of $p_{\text{Im}}(z)$ and $q(z)$. Thus $\mathcal{D}$ is closed under additive inverse and Eq. (43) holds as well.

$\Leftarrow$: Denote $p_{\text{Im}}(z) := \sqrt{\alpha} \frac{e(z) + e(z^{-1})}{2}$ and $q(z) := \sqrt{\alpha} \frac{e(z) - e(z^{-1})}{z - z^{-1}}$. Then

$$\begin{aligned} \mathfrak{F}(z) &= \alpha \prod_{i=1}^{2d} (z - r_i)(z^{-1} - r_i) = \alpha e(z)e(z^{-1}) \\ &= \left(\sqrt{\alpha} \frac{e(z) + e(z^{-1})}{2}\right)^2 - \left(\sqrt{\alpha} \frac{e(z) - e(z^{-1})}{2}\right)^2 \\ &= \left(\sqrt{\alpha} \frac{e(z) + e(z^{-1})}{2}\right)^2 + \left(1 - \left(\frac{z + z^{-1}}{2}\right)^2\right) \left(\sqrt{\alpha} \frac{e(z) - e(z^{-1})}{z - z^{-1}}\right)^2 \\ &= p_{\text{Im}}(z)^2 + \left(1 - \left(\frac{z + z^{-1}}{2}\right)^2\right) q(z)^2. \end{aligned}$$

Since $\mathcal{D}$ is closed under complex conjugation, $\prod_{i=1}^{2d}(z - r_i)$ is a real polynomial and $e(z) = z^d + \sum_{i=1}^{d-1} e_i z^i + \left(\prod_{i=1}^{2d} r_i\right) z^{-d}$ for some $e_i \in \mathbb{R}$. Notice that for $k > 0$

$$\begin{aligned} \left(x + i\sqrt{1-x^2}\right)^k + \left(x + i\sqrt{1-x^2}\right)^{-k} &= 2T_k(x), \\ \left(x + i\sqrt{1-x^2}\right)^k - \left(x + i\sqrt{1-x^2}\right)^{-k} &= 2iU_{k-1}(x)\sqrt{1-x^2}. \end{aligned}$$

We can verify that $p_{\text{Im}}(x + i\sqrt{1-x^2})$ and $q(x + i\sqrt{1-x^2})$ are real polynomials of x . Hence

$$P_{\text{Im}}(x) := p_{\text{Im}}(x + i\sqrt{1-x^2}), \quad Q(x) := q(x + i\sqrt{1-x^2}).$$

are well defined and $(P(x) = f(x) + iP_{\text{Im}}(x), Q(x))$ satisfies the normalization condition,

$$1 - f(x)^2 = P_{\text{Im}}(x)^2 + (1 - x^2)Q(x)^2.$$

Since $\mathcal{D}$ is closed under additive inverse,

$$e(-z) = (-z)^{-d} \prod_{i=1}^{2d} (-z - r_i) = (-1)^d z^{-d} \prod_{i=1}^{2d} (z - r_i) = (-1)^d e(z).$$

It leads to that $P_{\text{Im}}(x)$ has parity $(d \bmod 2)$ and $Q(x)$ has parity $(d - 1 \bmod 2)$.

Another important observation is that $\prod_{r \in \mathcal{D}} r < 0$. It comes from

$$\mathfrak{F}(z) = \alpha \prod_{i=1}^{2d} (z - r_i)(z^{-1} - r_i) = \alpha \left(\prod_{i=1}^{2d} r_i \right) z^{-2d} \prod_{i=1}^{2d} (z - r_i)(z - r_i^{-1}) \quad (44)$$

and that the coefficient of z^{2d} in $\mathfrak{F}(z)$ is negative. According to Eq. (44), one has that the coefficient of z^{2d} in $\mathfrak{F}(z)$ is $\alpha \left(\prod_{i=1}^{2d} r_i \right)$. Therefore,

$$\alpha \left(\prod_{i=1}^{2d} r_i \right) < 0 \implies \left(\prod_{i=1}^{2d} r_i \right) < 0. \quad (45)$$

Since

$$\begin{aligned} \frac{e(z) - e(z^{-1})}{z - z^{-1}} &= \frac{1}{z - z^{-1}} \left[\left(1 - \prod_{i=1}^{2d} r_i \right) (z^d - z^{-d}) + \sum_{i=1}^{d-1} e_i (z^i - z^{-i}) \right] \\ &= \left(1 - \prod_{i=1}^{2d} r_i \right) (z^{d-1} + z^{-d+1}) + l.o.. \end{aligned}$$

Here, we use *l.o.* to denote any linear combination of z^j with coefficients in $\mathbb{C}$ and $|j| < d-1$. By comparing with the coefficient of z^{d-1} in the monomial expansion for $Q(x)$, $\mathbf{q}_{d-1}$, one gets

$$\sqrt{\alpha} \left(1 - \prod_{i=1}^{2d} r_i \right) = 2^{1-d} \mathbf{q}_{d-1}. \quad (46)$$

In this way, $\deg(Q) = d-1$, $\deg(P) = d$ and the leading coefficient of $Q(x)$ is positive. $\square$

Remark 17 (Constructing complementary polynomials). *Theorem 16 indicates that there exists a bijection between the admissible decomposition of $\mathfrak{F}(z)$ and the admissible pair (P, Q) satisfying Definition 2. Furthermore, it provides a constructive way to find all admissible complementary polynomials:*

1. Choose a multiset $\mathcal{D}$ such that

- (a) the joint union of $\mathcal{D}$ and its reciprocal contains all roots of $\mathfrak{F}(z)$ with multiplicity.
- (b) $\mathcal{D}$ is closed under complex conjugation and additive inverse.

2. Construct P_{Im} and Q according to Eq. (43).

4.2 Maximal solution

Among all possible choices of $\mathcal{D}$, we are interested in choosing $\mathcal{D}$ to contain all roots of $\mathfrak{F}(z)$ within the unit disc and outside the unit disc. As will be shown below, the associated complementary polynomials satisfy many desirable properties. Let (P_{Im}, Q) , $(\bar{P}_{\text{Im}}, \bar{Q})$ denote the corresponding pair of complementary polynomials respectively. Moreover, throughout the paper, unless otherwise noted, we use $(\bar{P}_{\text{Im}}, \bar{Q})$ to denote the complementary polynomials constructed according to Eq. (43), where $\mathcal{D}$ contains all roots of $\mathfrak{F}(z)$ outside the unit disc.

Lemma 18 (Maximal magnitude of leading coefficients). *The leading coefficients of the expansion of Q and $\bar{Q}$ in the monomial basis have the largest magnitude among all possible complementary polynomials.*

Proof. Given any multiset $\mathcal{D}$ satisfying the condition mentioned in Remark 17, note that $\mathfrak{F}(z)$ can be represented as

$$\mathfrak{F}(z) = \beta z^{-2d} \prod_{r \in \mathcal{D}} (z - r)(z - r^{-1})$$

for some $\beta \in \mathbb{C}$. By checking the leading coefficient, $\beta = -\frac{\mathbf{f}_d^2}{4^d}$ is a quantity fully determined by f .

Recall the proof of Theorem 16, we know that $\alpha \prod_{r \in \mathcal{D}} r = \beta$ by Eq. (44). Together with Eq. (46), we obtain

$$\mathbf{q}_{d-1}^2 = 4^{d-1} |\beta| \left(K + \frac{1}{K} + 2 \right), \quad (47)$$

where $K := -\prod_{r \in \mathcal{D}} r > 0$ by Eq. (45).

When $\mathcal{D}$ contains all the roots within the unit disc or outside the unit disc, $\frac{1}{K} + K$ is the largest among all the possibilities, that is, the leading coefficient of $Q(x)$ and $\bar{Q}(x)$ has the largest magnitude. $\square$

The following lemma shows that (P_{Im}, Q) and $(\bar{P}_{\text{Im}}, \bar{Q})$ are directly related to each other.

Lemma 19. *The polynomials (P_{Im}, Q) and $(\bar{P}_{\text{Im}}, \bar{Q})$ are related to each other:*

$$P_{\text{Im}}(x) = -\bar{P}_{\text{Im}}(x), \quad Q(x) = \bar{Q}(x). \quad (48)$$

Proof. Denote $\mathcal{D} := \{r \text{ is a root of } \mathfrak{F}(z) : |r| < 1\}$, and

$$e(z) := z^{-d} \prod_{r \in \mathcal{D}} (z - r), \quad \bar{e}(z) := z^{-d} \prod_{r \in \mathcal{D}} (z - r^{-1}).$$

Then for some $\alpha, \bar{\alpha} > 0$,

$$\mathfrak{F}(z) = \alpha e(z) e(z^{-1}) = \bar{\alpha} \bar{e}(z) \bar{e}(z^{-1}). \quad (49)$$

Notice that

$$\begin{aligned}\bar{e}(z) &= z^{-d} \prod_{r \in \mathcal{D}} (z - r^{-1}) = z^{-d} \prod_{r \in \mathcal{D}} (r^{-1}(rz - 1)) = z^d \prod_{r \in \mathcal{D}} (r^{-1}(r - z^{-1})) \\ &= \left(\prod_{r \in \mathcal{D}} r^{-1} \right) z^{-d} \prod_{r \in \mathcal{D}} (z^{-1} - r) = -\frac{1}{K} e(z^{-1}).\end{aligned}\tag{50}$$

where $K = -\prod_{r \in \mathcal{D}} r$. Similarly, $\bar{e}(z^{-1}) = -\frac{1}{K} e(z)$. Thus, $\bar{\alpha} = K^2 \alpha$. By Theorem 16, we have

$$\begin{aligned}\bar{P}_{\text{Im}}(x) &= \sqrt{\bar{\alpha}} \frac{\bar{e}(x + i\sqrt{1-x^2}) + \bar{e}(x - i\sqrt{1-x^2})}{2}, \\ &= -\sqrt{\alpha K^2} \frac{1}{K} \frac{e(x - i\sqrt{1-x^2}) + e(x + i\sqrt{1-x^2})}{2} = -P_{\text{Im}}(x),\end{aligned}\tag{51}$$

and

$$\begin{aligned}\bar{Q}(x) &= \sqrt{\bar{\alpha}} \frac{\bar{e}(x + i\sqrt{1-x^2}) - \bar{e}(x - i\sqrt{1-x^2})}{2i\sqrt{1-x^2}} \\ &= -\sqrt{\alpha K^2} \frac{1}{K} \frac{e(x - i\sqrt{1-x^2}) - e(x + i\sqrt{1-x^2})}{2i\sqrt{1-x^2}} = Q(x).\end{aligned}\tag{52}$$

This proves the lemma. $\square$

Due to this relation, we may only focus on the admissible pair $(P(x), Q(x))$ obtained by choosing $\mathcal{D}$ to contain all the roots of $\mathfrak{F}(z)$ within the unit disc. Due to the property in Lemma 18, the corresponding phase factors in D_d will be referred to as the maximal solution.

Remark 20 (Maximal solution at $f = 0$). *Lemma 19 agrees with the observation that two particular admissible pairs associated with $f = 0$ are $(iT_d(x), U_{d-1}(x))$ and $(-iT_d(x), U_{d-1}(x))$, which are obtained by*

$$\Phi^0 := \left(\frac{\pi}{4}, 0, \dots, 0, \frac{\pi}{4} \right) \quad \bar{\Phi}^0 := \left(-\frac{\pi}{4}, 0, \dots, 0, -\frac{\pi}{4} \right)\tag{53}$$

respectively. In the following discussion, we will show that the maximal solution will converge to Φ^0 as $\|f\|_\infty \rightarrow 0$. The same argument is suitable for $(\bar{P}_{\text{Im}}, \bar{Q})$ with Φ^0 replaced by $\bar{\Phi}^0$.

Remark 21. *There is a useful observation that the coefficient of x^d for P_{Im} is positive. By Eq. (43), we have*

$$\begin{aligned}P_{\text{Im}}(x) &= \sqrt{\alpha} \frac{e(x + i\sqrt{1-x^2}) + e(x - i\sqrt{1-x^2})}{2} \\ &= \sqrt{\alpha} \frac{1 + \prod_{r \in \mathcal{D}} r}{2} \left[(x + i\sqrt{1-x^2})^d + (x - i\sqrt{1-x^2})^d \right] + l.o. \\ &= \sqrt{\alpha} \left(1 + \prod_{r \in \mathcal{D}} r \right) T_d(x) + l.o..\end{aligned}\tag{54}$$

Here $l.o.$ represents any polynomial of degree $< d$. To achieve the second equality, we use the expansion of $e(z) = z^{-d} \prod_{r \in \mathcal{D}} (z - r)$ in the monomial basis and discover that only

terms z^d and $(\prod_{r \in \mathcal{D}} r) z^{-d}$ contribute to the first term on the second line, if we replace z, z^{-1} with $x + i\sqrt{1-x^2}, x - i\sqrt{1-x^2}$ respectively.

The specific choice of $\mathcal{D}$ yields $(1 + \prod_{r \in \mathcal{D}} r) > 0$. Hence the leading coefficient of P_{Im} is positive.

4.3 Relation to other constructive methods

There are two other methods for constructing complementary polynomials, which will be referred to as the GSLW method [6] and the Haah method [8], respectively. The two methods are equivalent. The GSLW method solves the problem without using Laurent polynomials. The Haah method first points out that the decomposition can be viewed more naturally in terms of Laurent polynomials.

The Haah method considers

$$1 - P_{\text{Re}} \left(\frac{z + z^{-1}}{2} \right)^2 + \left(\frac{z - z^{-1}}{2} \right)^2 Q_{\text{Im}} \left(\frac{z + z^{-1}}{2} \right)^2 \quad (55)$$

and the corresponding decomposition:

$$1 - P_{\text{Re}} \left(\frac{z + z^{-1}}{2} \right)^2 + \left(\frac{z - z^{-1}}{2} \right)^2 Q_{\text{Im}} \left(\frac{z + z^{-1}}{2} \right)^2 = \alpha \prod_{r \in \mathcal{D}} (z - r)(z^{-1} - r) \quad (56)$$

where $\mathcal{D}$ contains all the roots within the unit disc.

In our construction, we may omit the term Q_{Im} due to the symmetry constraint. However, one can check that our method is also suitable for the case that $Q_{\text{Im}} \neq 0$ by simply replacing $\mathfrak{F}(z)$ by Eq. (55). By the same method as that in Theorem 16, it can be shown that any admissible complementary polynomial $(P_{\text{Re}}, Q_{\text{Re}})$ must take the form of Eq. (43), where Q is replaced by Q_{Re} . Hence our method can be generalized to find all admissible complementary polynomials, by going through all possible $e(z) = z^{-d} \prod_{r \in \mathcal{D}} (z - r)$. In particular, the Haah method considers the special case that $\mathcal{D}$ contains all the roots of $\mathfrak{F}(z)$ within the unit disc, and therefore does not generate all admissible complementary polynomials.

We would like to clarify that Ref [8] states that we can construct P_{Im} and Q as long as the joint union of a conjugate-closed multiset $\mathcal{D}$ and its reciprocal $\{\frac{1}{r} : r \in \mathcal{D}\}$ is the list of all the roots of Eq. (55), if there is no parity requirement for P_{Im} and Q . After imposing the parity condition, according to Theorem 16, another necessary condition for suitable complementary polynomials is that $\mathcal{D}$ is closed under additive inverse. Without it, the parity condition may not hold. This can be verified by the proof of Theorem 16.

5 Local convergence

We prove Theorem 5 in this section. Throughout this section, polynomials are expanded in the Chebyshev basis without otherwise noted. This is more convenient than the expansion in the monomial basis.

According to the definition of the Chebyshev norms in Section 2.1 and the convention of the coefficients in Section 2.2, we specifically list the following relations which exploit the weighted orthogonality and will be frequently used in the proof,

$$\|f\|_T^2 = f_0^2 + \frac{1}{2} \sum_{k=1}^d f_k^2, \quad \|P_{\text{Im}}\|_T^2 = s_0^2 + \frac{1}{2} \sum_{k=1}^d s_k^2, \quad \|Q\|_U^2 = \frac{1}{2} \sum_{k=0}^{d-1} q_k^2. \quad (57)$$

By integrating the point-wise normalization condition $f(x)^2 + P_{\text{Im}}(x)^2 + (1-x^2)Q(x)^2 = 1$ with the weight function $\frac{1}{\sqrt{1-x^2}}$, we obtain the following Parseval's equality:

$$f_0^2 + s_0^2 + \frac{1}{2} \sum_{k=1}^d (f_k^2 + s_k^2) + \frac{1}{2} \sum_{k=0}^{d-1} q_k^2 = 1. \quad (58)$$

5.1 Sketch of the proof

We first give a sketch of the proof as a road map. For simplicity, we assume d is odd in the sketch, while the proof works for any d .

1. We first show that each element of $\tilde{\Phi}^* - \tilde{\Phi}^0$ is bounded by $\frac{\pi}{4}$ by properly restricting $\|f\|_\infty$.
2. We further upper bound the distance $\|\tilde{\Phi}^* - \tilde{\Phi}^0\|_2$ by $|\sin(2\phi_0 - \frac{\pi}{2})|^2 + \sum_{\ell=1}^{\tilde{d}-1} |\sin(2\phi_\ell)|^2$. It follows the inequality $|x| \leq \frac{\pi}{2} |\sin(x)|$ which holds for any $|x| \leq \frac{\pi}{2}$. The bound in the first step ensures that the inequality can be invoked.
3. We upper bound the phase factors by the leading Chebyshev coefficients in the recursion sequence of polynomials. The upper bound turns out to be

$$\left| \sin\left(2\phi_0 - \frac{\pi}{2}\right) \right|^2 \leq \frac{8}{3} \|f\|_\infty^2, \\ |\sin(2\phi_\ell)|^2 \leq \frac{8}{3} \left(\left(q_{d-1-2\ell}^{(\ell)} \right)^2 - \left(q_{d+1-2\ell}^{(\ell-1)} \right)^2 \right), \quad 1 \leq \ell \leq \tilde{d}-1.$$

4. Combining the previous results, we finally get $\|\tilde{\Phi}^* - \tilde{\Phi}^0\|_2^2 \leq \frac{\pi^2}{6} (\|f\|_\infty^2 + 1 - q_{d-1}^2)$. The upper bound can be further simplified to be $\frac{\pi^2}{3} \|f\|_\infty^2$ by exploiting the estimate to q_{d-1} (see Lemma 23). Then the statement in Theorem 5 follows.

5.2 Recursion relation in terms of Chebyshev coefficients

In this subsection, we will convert the recursion relation in terms of monomial coefficients to that in terms of Chebyshev coefficients. Furthermore, we will derive some useful properties which will be used as intermediate steps in the proof of Theorem 5.

Let $(P, Q) \in \mathbb{C}_d[x] \times \mathbb{R}_{d-1}[x]$ be a pair of polynomials satisfying the conditions in Theorem 1. The symmetric reduction procedures yield a sequence of polynomials $(P^{(\ell)}, Q^{(\ell)}) \in \mathbb{C}_{d-2\ell}[x] \times \mathbb{R}_{d-2\ell-1}[x]$ satisfying the conditions and an associated sequence of phase factors ϕ_ℓ so that $e^{2i\phi_\ell} = \frac{\mathbf{p}_{d-2\ell}^{(\ell)}}{\mathbf{q}_{d-1-2\ell}^{(\ell)}}$. In particular, $(P^{(0)}, Q^{(0)}) = (P, Q)$. An important recursion relation generating the coefficients of new polynomials in the monomial basis is given in Eq. (32). For convenience, we restate it as follows: if $d \geq 3$, then

$$\mathbf{p}_{d-2}^{(1)} = \frac{\mathbf{q}_{d-1}}{4} - \frac{1}{\mathbf{q}_{d-1}} \left| \text{Im}[\mathbf{p}_{d-2} e^{-2i\phi_0}] \right|^2 + i \text{Im}[\mathbf{p}_{d-2} e^{-2i\phi_0}], \\ \mathbf{q}_{d-3}^{(1)} = \frac{\mathbf{q}_{d-1}}{4} + \frac{1}{\mathbf{q}_{d-1}} \left| \text{Im}[\mathbf{p}_{d-2} e^{-2i\phi_0}] \right|^2.$$

Here, recall that the boldface symbols refer to the coefficients in the monomial basis. We will first rewrite this relation using the Chebyshev basis. As a remark, it suffices

to consider the case when $\ell = 0$, and the general case can be recovered by substituting $(P, Q, P^{(1)}, Q^{(1)}, \phi_0) \leftarrow (P^{(\ell)}, Q^{(\ell)}, P^{(\ell+1)}, Q^{(\ell+1)}, \phi_\ell)$ in the arguments as long as $\deg(P^{(\ell)}) \geq 3$.

For simplicity, let us denote the Chebyshev polynomial in the monomial basis as $T_d(x) = 2^{d-1}x^d + t_d x^{d-2} + l.o.$ where we use *l.o.* to denote any polynomial of degree no more than $d-3$ and t_d is a real number whose exact value is irrelevant in the proof.

Expanding the polynomials in monomial basis and Chebyshev basis, we have

$$\begin{aligned} P(x)e^{-2i\phi_0} &= \sum_{k=0}^d \mathbf{p}_k e^{-2i\phi_0} x^k = e^{-2i\phi_0} \mathbf{p}_d x^d + e^{-2i\phi_0} \mathbf{p}_{d-2} x^{d-2} + l.o. \\ &= \sum_{k=0}^d p_k e^{-2i\phi_0} T_k(x) = e^{-2i\phi_0} 2^{d-1} p_d x^d + \left(e^{-2i\phi_0} p_d t_d + e^{-2i\phi_0} p_{d-2} 2^{d-3} \right) x^{d-2} + l.o.. \end{aligned}$$

Then, by comparing the coefficients of terms x^d and x^{d-2} , we have $\mathbf{p}_d = 2^{d-1} p_d$ and

$$e^{-2i\phi_0} \mathbf{p}_{d-2} = e^{-2i\phi_0} p_d t_d + e^{-2i\phi_0} p_{d-2} 2^{d-3} = q_{d-1} t_d + e^{-2i\phi_0} p_{d-2} 2^{d-3}.$$

In the last equality, we use

$$e^{2i\phi_0} = \frac{\mathbf{p}_d}{\mathbf{q}_{d-1}} = \frac{2^{d-1} p_d}{2^{d-1} q_{d-1}} = \frac{p_d}{q_{d-1}} \Rightarrow e^{-2i\phi_0} p_d = q_{d-1} \in \mathbb{R}.$$

An important observation is

$$\text{Im}[\mathbf{p}_{d-2} e^{-2i\phi_0}] = 2^{d-3} \text{Im}[p_{d-2} e^{-2i\phi_0}], \quad (59)$$

which follows $q_{d-1} t_d \in \mathbb{R}$.

Plugging the previous equations in terms of Chebyshev coefficients in Eq. (32), we have the following recursion relation

$$\begin{aligned} p_{d-2}^{(1)} &= q_{d-1} - \frac{1}{4q_{d-1}} \left| \text{Im}[p_{d-2} e^{-2i\phi_0}] \right|^2 + i \text{Im}[p_{d-2} e^{-2i\phi_0}], \\ q_{d-3}^{(1)} &= q_{d-1} + \frac{1}{4q_{d-1}} \left| \text{Im}[p_{d-2} e^{-2i\phi_0}] \right|^2. \end{aligned} \quad (60)$$

The second equality implies the following sign preserving property and the monotonicity

$$\text{sgn}(q_{d-1}) = \text{sgn}(q_{d-3}^{(1)}), \text{ and } |q_{d-1}| \leq |q_{d-3}^{(1)}|.$$

Regarding the next phase factor ϕ_1 , its defined equation is converted similarly to Chebyshev basis

$$e^{2i\phi_1} = \frac{\mathbf{p}_{d-2}^{(1)}}{\mathbf{q}_{d-3}^{(1)}} = \frac{p_{d-2}^{(1)}}{q_{d-3}^{(1)}}.$$

Furthermore, we can use the derived recursion relation in Eq. (60) to obtain

$$\sin(2\phi_1) = \text{Im} \left[\frac{p_{d-2}^{(1)}}{q_{d-3}^{(1)}} \right] = \frac{\text{Im}[p_{d-2} e^{-2i\phi_0}]}{q_{d-3}^{(1)}}.$$

The derived results can be generalized to $(P^{(\ell)}, Q^{(\ell)})$ where $1 \leq \ell \leq \hat{d}$ ($\hat{d}$ is defined by Eq. (35)). We summarize these results as the following theorem.

Theorem 22. Given $(P(x), Q(x))$ satisfying the assumption in Theorem 1 and $\Phi = (\phi_0, \phi_1, \dots, \phi_1, \phi_0)$ being a set of symmetric phase factors in $\mathcal{D}_d$ such that Eq. (8) is satisfied, the following statements hold.

1. For any $1 \leq \ell \leq \hat{d}$, the leading coefficients of $P^{(\ell)}$ and $Q^{(\ell)}$ satisfy the recursion relation

$$\begin{aligned} p_{d-2\ell}^{(\ell)} &= q_{d+1-2\ell}^{(\ell-1)} - \frac{1}{4q_{d+1-2\ell}^{(\ell-1)}} \left| \text{Im}[p_{d-2\ell}^{(\ell-1)} e^{-2i\phi_{\ell-1}}] \right|^2 + i \text{Im}[p_{d-2\ell}^{(\ell-1)} e^{-2i\phi_{\ell-1}}], \\ q_{d-1-2\ell}^{(\ell)} &= q_{d+1-2\ell}^{(\ell-1)} + \frac{1}{4q_{d+1-2\ell}^{(\ell-1)}} \left| \text{Im}[p_{d-2\ell}^{(\ell-1)} e^{-2i\phi_{\ell-1}}] \right|^2. \end{aligned} \quad (61)$$

2. For any $0 \leq \ell \leq \hat{d}$,

$$e^{2i\phi_\ell} = \frac{p_{d-2\ell}^{(\ell)}}{q_{d-1-2\ell}^{(\ell)}}, \quad (62)$$

and if $1 \leq \ell \leq \hat{d}$,

$$\sin(2\phi_\ell) = \frac{\text{Im}[p_{d-2\ell}^{(\ell-1)} e^{-2i\phi_{\ell-1}}]}{q_{d-1-2\ell}^{(\ell)}}. \quad (63)$$

3. If d is even, $e^{i\phi_{\tilde{d}-1}} = p_0^{(\tilde{d}-1)}$ and

$$\cos(\phi_{\tilde{d}-1}) = q_1^{(\tilde{d}-2)}, \quad \sin(\phi_{\tilde{d}-1}) = -ip_0^{(\tilde{d}-2)} e^{-2i\phi_{\tilde{d}-2}}. \quad (64)$$

4. (The monotonicity of the leading coefficient of $Q^{(\ell)}$ in the Chebyshev basis)

If $q_{d-1}^{(0)} > 0$, then

$$0 < q_{d-1}^{(0)} \leq q_{d-3}^{(1)} \leq \dots \leq q_{d-1-2\ell}^{(\ell)} \leq \dots \leq q_{d-1-2\hat{d}}^{(\hat{d})} \leq 1. \quad (65)$$

Proof. The statements 1, 2 and 4 follow the previous discussion. Hence, we only need to verify Eq. (64).

Note that when $\ell = \tilde{d} - 2$, we have $\deg(P^{(\tilde{d}-2)}) = 2$ and

$$\begin{aligned} \langle 0 | e^{i\phi_{\tilde{d}-2}Z} W(x) e^{i\phi_{\tilde{d}-1}} W(x) e^{i\phi_{\tilde{d}-2}Z} | 0 \rangle &= e^{2i\phi_{\tilde{d}-2}} \left(2 \cos(\phi_{\tilde{d}-1}) x^2 - e^{-i\phi_{\tilde{d}-1}} \right) \\ &= P^{(\tilde{d}-2)}(x) = p_2^{(\tilde{d}-2)} T_2(x) + p_0^{(\tilde{d}-2)} T_0(x) = 2p_2^{(\tilde{d}-2)} x^2 - p_2^{(\tilde{d}-2)} + p_0^{(\tilde{d}-2)}. \end{aligned}$$

The first equality is obtained by direct computation. Comparing the constant terms, we have

$$\begin{aligned} e^{i\phi_{\tilde{d}-1}} &= \left(p_2^{(\tilde{d}-2)} e^{-2i\phi_{\tilde{d}-2}} \right)^* - \left(p_0^{(\tilde{d}-2)} e^{-2i\phi_{\tilde{d}-2}} \right)^* \\ &= q_1^{(\tilde{d}-2)} - \left(p_0^{(\tilde{d}-2)} e^{-2i\phi_{\tilde{d}-2}} \right)^*. \end{aligned}$$

Taking the modulus on both sides, we get

$$\begin{aligned} 1 &= \left(q_1^{(\tilde{d}-2)} - \text{Re}[p_0^{(\tilde{d}-2)} e^{-2i\phi_{\tilde{d}-2}}] \right)^2 + \left(\text{Im}[p_0^{(\tilde{d}-2)} e^{-2i\phi_{\tilde{d}-2}}] \right)^2 \\ &= \left(q_1^{(\tilde{d}-2)} \right)^2 + \left| p_0^{(\tilde{d}-2)} \right|^2 - 2q_1^{(\tilde{d}-2)} \text{Re}[p_0^{(\tilde{d}-2)} e^{-2i\phi_{\tilde{d}-2}}]. \end{aligned} \quad (66)$$

Applying Parseval's equality derived from $|P^{(\tilde{d}-2)}(x)|^2 + (1-x^2)|Q^{(\tilde{d}-2)}(x)|^2 = 1$, we get

$$\left|p_0^{(\tilde{d}-2)}\right|^2 + \frac{1}{2}\left|p_2^{(\tilde{d}-2)}\right|^2 + \frac{1}{2}\left(q_1^{(\tilde{d}-2)}\right)^2 = 1.$$

Note that $\left|p_2^{(\tilde{d}-2)}\right| = \left|p_2^{(\tilde{d}-2)}e^{-2i\phi_{\tilde{d}-2}}\right| = \left|q_1^{(\tilde{d}-2)}\right|$, we have

$$\left|p_0^{(\tilde{d}-2)}\right|^2 + \left(q_1^{(\tilde{d}-2)}\right)^2 = 1. \quad (67)$$

Plugging the last equality into Eq. (66), we obtain $\text{Re}[p_0^{(\tilde{d}-2)}e^{-2i\phi_{\tilde{d}-2}}] = 0$ since $q_1^{(\tilde{d}-2)} \neq 0$. Therefore, $p_0^{(\tilde{d}-2)}e^{-2i\phi_{\tilde{d}-2}}$ is pure imaginary which yields Eq. (64). $\square$

5.3 Bounding Chebyshev coefficients of the maximal solution

In this subsection, we assume that (P, Q) is the admissible pair corresponding to the maximal solution. We will derive bounds on the leading coefficients of Q and $P_{\text{Im}}(x)$ in the Chebyshev basis, *i.e.*, q_{d-1} and s_d respectively.

Lemma 23. *The leading coefficient of Q in the Chebyshev basis satisfies the following estimate: $1 - \|f\|_\infty^2 \leq q_{d-1}^2 \leq 1 - \|f\|_T^2 + \frac{1}{2}f_d^2$.*

Proof. Expanding the polynomials in the monomial basis, the normalization condition reduces to

$$f(x)^2 + P_{\text{Im}}(x)^2 + (1-x^2)Q(x)^2 = (\mathbf{f}_d^2 + \mathbf{s}_d^2 - \mathbf{q}_{d-1}^2)x^{2d} + l.o. = 1,$$

where *l.o.* denotes any polynomial of degree less than $2d$. It implies that $\mathbf{f}_d^2 + \mathbf{s}_d^2 = \mathbf{q}_{d-1}^2$ which is further $f_d^2 + s_d^2 = q_{d-1}^2$. Hence, we have

$$\begin{aligned} q_{d-1}^2 &= \frac{1}{2}(f_d^2 + s_d^2 + q_{d-1}^2) \\ &= 1 - \left(f_0^2 + s_0^2 + \frac{1}{2}\sum_{k=1}^{d-1}(f_k^2 + s_k^2) + \frac{1}{2}\sum_{k=0}^{d-2}q_k^2\right) \\ &\leq 1 - \left(f_0^2 + \frac{1}{2}\sum_{k=1}^{d-1}f_k^2\right) = 1 - \|f\|_T^2 + \frac{1}{2}f_d^2. \end{aligned}$$

In the second line, we exploit Eq. (58), the Parseval's equality derived from the normalization condition.

To lower bound q_{d-1} , we first remark that Eq. (47) can be rewritten as

$$q_{d-1} = |\beta| \left(K + \frac{1}{K} + 2\right), \quad (68)$$

where $K := -\prod_{r \in \mathcal{D}} r > 0$, and $\mathcal{D}$ contains all roots of $\mathfrak{F}(z)$ within the unit disc for the maximal solution.

Given a polynomial fully factored into $g(z) = \gamma \prod_{i=1}^n (z - a_i) \in \mathbb{C}[z]$, its *Mahler measure* [15] is defined as

$$M(g) := |\gamma| \prod_{i=1}^n \max\{|a_i|, 1\}.$$

This quantity is equal to the geometric mean of $|g(z)|$ on the unit circle by using Jensen's formula in [10]

$$M(g) = \exp \left(\frac{1}{2\pi} \int_0^{2\pi} \ln |g(e^{i\theta})| d\theta \right). \quad (69)$$

Remarkably, note that $|\beta| \frac{1}{K}$ is the Mahler measure of a polynomial $h(z)$ defined as

$$h(z) := \beta \prod_{r \in \mathcal{D}} (z - r)(z - r^{-1}) = z^{2d} \mathfrak{F}(z). \quad (70)$$

Therefore, one has

$$\begin{aligned} |\beta| \frac{1}{K} &= M(h) = \exp \left(\frac{1}{2\pi} \int_0^{2\pi} \ln |h(e^{i\theta})| d\theta \right) \\ &= \exp \left(\frac{1}{2\pi} \int_0^{2\pi} \ln (1 - f(\cos \theta)^2) d\theta \right) \\ &\geq \exp \left(\frac{1}{2\pi} \int_0^{2\pi} \ln (1 - \|f\|_\infty^2) d\theta \right) \\ &= 1 - \|f\|_\infty^2. \end{aligned} \quad (71)$$

To conclude,

$$q_{d-1}^2 = |\beta| \left(K + \frac{1}{K} + 2 \right) \geq |\beta| \frac{1}{K} \geq 1 - \|f\|_\infty^2, \quad (72)$$

which completes the proof. $\square$

Remark 24 (Convergence of the maximal solution towards Φ^0). *The derived bounds on q_{d-1} consequentially yields the bounds on s_d . Using the equality $f_d^2 + s_d^2 = q_{d-1}^2$, we get the following estimate*

$$1 - \|f\|_\infty^2 - f_d^2 \leq s_d^2 \leq 1 - \|f\|_T^2 - \frac{1}{2} f_d^2. \quad (73)$$

Combining the derived results, the remaining Chebyshev coefficients of P_{Im} and Q can be upper bounded as follows

$$\begin{aligned} s_0^2 + \frac{1}{2} \sum_{k=1}^{d-1} s_k^2 + \frac{1}{2} \sum_{k=0}^{d-2} q_k^2 \\ (\text{Parseval's equality}) &= 1 - f_0^2 - \frac{1}{2} \sum_{k=1}^d f_k^2 - \frac{1}{2} s_d^2 - \frac{1}{2} q_{d-1}^2 \\ (f_d^2 + s_d^2 = q_{d-1}^2) &= 1 - q_{d-1}^2 - f_0^2 - \frac{1}{2} \sum_{k=1}^{d-1} f_k^2 \\ (\text{Lemma 23}) &\leq \|f\|_\infty^2 - f_0^2 - \frac{1}{2} \sum_{k=1}^{d-1} f_k^2 \leq \|f\|_\infty^2. \end{aligned} \quad (74)$$

As a consequence of the upper bound, we have $s_i, q_j \rightarrow 0 \ \forall i \in [d], j \in [d-1]$ as $\|f\|_\infty \rightarrow 0$. Remark 21 implies that $s_d > 0$. Together with $f_d^2 + s_d^2 = q_{d-1}^2$, it follows that $s_d = q_{d-1}$ in that limit since $f_d \rightarrow 0$, and the normalization condition yields $s_d = q_{d-1} = 1$ when the leading coefficient of Q is fixed to be positive.

To conclude, as $\|f\|_\infty \rightarrow 0$, the maximal solution exhibits a limit $(P, Q) \rightarrow (iT_d, U_{d-1})$ which agrees with the polynomials generated from the set of phase factors $\Phi^0 = (\frac{\pi}{4}, 0, \dots, 0, \frac{\pi}{4})$.

5.4 Proof of Theorem 5

In this subsection, we will prove Theorem 5 based on the previous results. We start from a technical lemma strengthening the results in Theorem 22 under additional assumptions.

Lemma 25. *When $\|f\|_\infty \leq \frac{1}{2}$, it holds that*

1. $\frac{\sqrt{3}}{2} \leq \sqrt{1 - \|f\|_\infty^2} \leq q_{d-1}^{(0)} \leq q_{d-3}^{(1)} \leq \dots \leq q_{d-1-2\ell}^{(\ell)} \leq \dots \leq q_{d-1-2\hat{d}}^{(\hat{d})} \leq 1.$
2. $|\phi_0 - \frac{\pi}{4}| < \frac{\pi}{4}$ and $|\phi_\ell| < \frac{\pi}{4}, \quad \forall 1 \leq \ell \leq \tilde{d} - 1.$

Proof. The first statement directly follows Lemma 23.

To show the second statement, we first notice that according to Theorem 22,

$$\cos\left(2\phi_0 - \frac{\pi}{2}\right) = \sin(2\phi_0) = \frac{\text{Im}[p_d^{(0)}]}{q_{d-1}^{(0)}}.$$

According to the Eq. (54) in Remark 21, we have

$$\text{Im}[p_d^{(0)}] = \sqrt{\alpha} \left(1 + \prod_{r \in \mathcal{D}} r\right) > 0, \quad (75)$$

where $\mathcal{D}$ contains the roots of $\mathfrak{F}(z)$ in the unit disc. Therefore, $|\phi_0 - \frac{\pi}{4}| < \frac{\pi}{4}$ follows $\cos(2\phi_0 - \frac{\pi}{2}) > 0$.

When $1 \leq \ell \leq \hat{d}$, according to Eqs. (61) and (62), we have

$$\cos(2\phi_\ell) = \frac{1}{q_{d-1-2\ell}^{(\ell)}} \left(q_{d+1-2\ell}^{(\ell-1)} - \frac{1}{4q_{d+1-2\ell}^{(\ell-1)}} \left| \text{Im}[p_{d-2\ell}^{(\ell-1)} e^{-2i\phi_{\ell-1}}] \right|^2 \right). \quad (76)$$

Using Parseval's equality, the normalization condition $|P^{(\ell-1)}|^2 + (1-x^2)|Q^{(\ell-1)}|^2 = 1$ yields

$$\begin{aligned} 1 &= |p_0^{(\ell-1)}|^2 + \frac{1}{2} \sum_{j=1}^{d-2\ell+2} |p_j^{(\ell-1)}|^2 + \frac{1}{2} \sum_{j=0}^{d-2\ell+1} (q_j^{(\ell-1)})^2 \\ &\geq \frac{1}{2} |p_{d-2\ell}^{(\ell-1)}|^2 + \frac{1}{2} |p_{d-2\ell+2}^{(\ell-1)}|^2 + \frac{1}{2} (q_{d-2\ell+1}^{(\ell-1)})^2. \end{aligned}$$

Furthermore, the cancellation of the leading order in the normalization condition implies $|p_{d-2\ell+2}^{(\ell-1)}|^2 = (q_{d-2\ell+1}^{(\ell-1)})^2$. Thus

$$\frac{1}{2} |p_{d-2\ell}^{(\ell-1)}|^2 + (q_{d-2\ell+1}^{(\ell-1)})^2 \leq 1.$$

Using the first statement frequently, we have

$$\begin{aligned} \left| \text{Im}[p_{d-2\ell}^{(\ell-1)} e^{-2i\phi_{\ell-1}}] \right|^2 &\leq |p_{d-2\ell}^{(\ell-1)}|^2 \leq 2 \left(1 - (q_{d-2\ell+1}^{(\ell-1)})^2 \right) \\ &\leq 2 \left(1 - (q_{d-1}^{(0)})^2 \right) \leq 2 \|f\|_\infty^2 \leq 4 \|f\|_\infty^2 < 4 \left(1 - \|f\|_\infty^2 \right) \leq 4 (q_{d+1-2\ell}^{(\ell-1)})^2. \end{aligned} \quad (77)$$

Plugging Eq. (77) into Eq. (76), we have $\cos(2\phi_\ell) > 0$. Hence, $|\phi_\ell| < \frac{\pi}{4}$.

If d is even, following Theorem 22 statement 3, we have

$$\cos(\phi_{\tilde{d}-1}) = q_1^{(\tilde{d}-2)} = q_1^{(\hat{d})} \geq \frac{\sqrt{3}}{2}.$$

Here, $\hat{d} = \tilde{d} - 2$ when d is even and the first statement is invoked. It implies that $|\phi_{\tilde{d}-1}| < \frac{\pi}{4}$. $\square$

The second statement of the previous lemma ensures that the inequality $\frac{\pi}{2} |\sin(x)| \geq |x|$ for any $|x| \leq \frac{\pi}{2}$ can be invoked. Now we are ready to prove Theorem 5.

Proof of Theorem 5. We first estimate $|\sin(2\phi_0 - \frac{\pi}{2})|^2$ as follows

$$\left| \sin(2\phi_0 - \frac{\pi}{2}) \right|^2 = \left| \frac{\operatorname{Re}[p_d]}{q_{d-1}} \right|^2 = \frac{f_d^2}{q_{d-1}^2} \leq \frac{4}{3} f_d^2 \leq \frac{8}{3} \|f\|_\infty^2, \quad (78)$$

where we use the inequality $f_d^2 \leq 2 \|f\|_T^2 \leq 2 \|f\|_\infty^2$.

For any $1 \leq \ell \leq \hat{d}$, by Eq. (63), we have

$$\begin{aligned} |\sin(2\phi_\ell)|^2 &= \left| \frac{\operatorname{Im}[p_{d-2\ell}^{(\ell-1)} e^{-2i\phi_{\ell-1}}]}{q_{d-1-2\ell}^{(\ell)}} \right|^2 \\ &\stackrel{(\text{Lemma 25 1.})}{\leq} \frac{4}{3} \left| \operatorname{Im}[p_{d-2\ell}^{(\ell-1)} e^{-2i\phi_{\ell-1}}] \right|^2 \\ &\stackrel{(\text{Theorem 22 1.})}{=} \frac{16}{3} \left(q_{d-1-2\ell}^{(\ell)} - q_{d+1-2\ell}^{(\ell-1)} \right) q_{d+1-2\ell}^{(\ell-1)} \\ &\stackrel{(\text{Theorem 22 4.})}{\leq} \frac{8}{3} \left(q_{d-1-2\ell}^{(\ell)} - q_{d+1-2\ell}^{(\ell-1)} \right) \left(q_{d-1-2\ell}^{(\ell)} + q_{d+1-2\ell}^{(\ell-1)} \right) \\ &= \frac{8}{3} \left(\left(q_{d-1-2\ell}^{(\ell)} \right)^2 - \left(q_{d+1-2\ell}^{(\ell-1)} \right)^2 \right). \end{aligned}$$

Here, the monotonicity in Theorem 22 implies that $q_{d-1-2\ell}^{(\ell)} - q_{d+1-2\ell}^{(\ell-1)} \geq 0$.

If d is odd, one can derive an upper bound by telescoping and Lemma 25

$$\begin{aligned} \sum_{\ell=1}^{\hat{d}} |\sin(2\phi_\ell)|^2 &\leq \frac{8}{3} \sum_{\ell=1}^{\hat{d}} \left(\left(q_{d-1-2\ell}^{(\ell)} \right)^2 - \left(q_{d+1-2\ell}^{(\ell-1)} \right)^2 \right) \\ &\leq \frac{8}{3} \left(1 - \left(q_{d-1}^{(0)} \right)^2 \right) \leq \frac{8}{3} \|f\|_\infty^2. \end{aligned}$$

Therefore, we have

$$\begin{aligned} \|\tilde{\Phi}^* - \tilde{\Phi}^0\|_2^2 &= \left(\phi_0 - \frac{\pi}{4} \right)^2 + \sum_{\ell=1}^{\hat{d}} \phi_\ell^2 \\ &\leq \frac{1}{4} \frac{\pi^2}{4} \left| \sin(2\phi_0 - \frac{\pi}{2}) \right|^2 + \frac{1}{4} \frac{\pi^2}{4} \sum_{\ell=1}^{\hat{d}} |\sin(2\phi_\ell)|^2 \\ &\leq \frac{1}{4} \frac{\pi^2}{4} \left(\frac{8}{3} \|f\|_\infty^2 + \frac{8}{3} \|f\|_\infty^2 \right) = \frac{1}{3} \pi^2 \|f\|_\infty^2. \end{aligned}$$

If d is even, we have to treat $\phi_{\tilde{d}-1}$ separately. Using Theorem 22 statement 3, we have

$$\sin^2(\phi_{\tilde{d}-1}) = 1 - \cos^2(\phi_{\tilde{d}-1}) = 1 - \left(q_1^{(\tilde{d}-2)} \right)^2.$$

Recall that $|\phi_{\tilde{d}-1}| < \frac{\pi}{4}$, we apply a finer estimate $|x| \leq \frac{\pi}{2\sqrt{2}} |\sin(x)|$ for any $|x| \leq \frac{\pi}{4}$ and get

$$|\phi_{\tilde{d}-1}|^2 \leq \frac{\pi^2}{8} \sin^2(\phi_{\tilde{d}-1}). \quad (79)$$

Similarly,

$$\sum_{\ell=1}^{\hat{d}} |\sin(2\phi_\ell)|^2 \leq \frac{8}{3} \sum_{\ell=1}^{\hat{d}} \left((q_{d-1-2\ell}^{(\ell)})^2 - (q_{d+1-2\ell}^{(\ell-1)})^2 \right) = \frac{8}{3} (q_1^{(\tilde{d}-2)})^2 - \frac{8}{3} (q_{d-1}^{(0)})^2.$$

Furthermore, combining the derived results, one has

$$\begin{aligned} \|\tilde{\Phi}^* - \tilde{\Phi}^0\|_2^2 &= \left(\phi_0 - \frac{\pi}{4} \right)^2 + \sum_{\ell=1}^{\hat{d}} \phi_\ell^2 + \phi_{\tilde{d}-1}^2 \\ &\leq \frac{1}{4} \frac{\pi^2}{4} \left| \sin(2\phi_0 - \frac{\pi}{2}) \right|^2 + \frac{1}{4} \frac{\pi^2}{4} \sum_{\ell=1}^{\hat{d}} |\sin(2\phi_\ell)|^2 + \frac{\pi^2}{8} \left| \sin(\phi_{\tilde{d}-1}) \right|^2 \\ &\leq \frac{\pi^2}{4} \left(\frac{2}{3} \|f\|_\infty^2 + \frac{2}{3} (q_1^{(\tilde{d}-2)})^2 - \frac{2}{3} (q_{d-1}^{(0)})^2 + \frac{1}{2} - \frac{1}{2} (q_1^{(\tilde{d}-2)})^2 \right) \\ (\text{Theorem 22 4}) &\leq \frac{\pi^2}{4} \left(\frac{2}{3} \|f\|_\infty^2 + \frac{2}{3} - \frac{2}{3} (q_{d-1}^{(0)})^2 \right) \\ (\text{Lemma 25 1}) &\leq \frac{\pi^2}{4} \left(\frac{2}{3} \|f\|_\infty^2 + \frac{2}{3} \|f\|_\infty^2 \right) = \frac{1}{3} \pi^2 \|f\|_\infty^2. \end{aligned}$$

In conclusion, we have

$$\|\tilde{\Phi}^* - \tilde{\Phi}^0\|_2 \leq \frac{\pi}{\sqrt{3}} \|f\|_\infty.$$

□

6 Local strong convexity

6.1 Hessian matrix at the maximal solution is positive definite

We first show that the Hessian matrix at the initial point $\tilde{\Phi}^0$ is positive definite.

Theorem 26. *At the initial point $\tilde{\Phi}^0 := (\frac{\pi}{4}, 0, \dots, 0) \in \mathbb{R}^{\tilde{d}}$, the Hessian matrix of the optimization problem Eq. (5) is*

$$\text{Hess}(\tilde{\Phi}^0) = \frac{2}{\tilde{d}} A(\tilde{\Phi}^0)^\top A(\tilde{\Phi}^0) = \begin{cases} 4I & , \text{if } d \text{ is odd,} \\ \text{diag}\{4, \dots, 4, 2\} & , \text{if } d \text{ is even.} \end{cases} \quad (80)$$

Proof. According to [5, Lemma 3],

$$\text{Re} \left[\langle 0 | U(x, \tilde{\Phi}) | 0 \rangle \right] = -\text{Im} \left[\langle 0 | e^{-i\frac{\pi}{4}Z} U(x, \tilde{\Phi}) e^{-i\frac{\pi}{4}Z} | 0 \rangle \right],$$

which suggests us to extract $\frac{\pi}{4}$ from the phase factors at the edges and consider $e^{-i\frac{\pi}{4}Z} U(x, \tilde{\Phi}^0) e^{-i\frac{\pi}{4}Z} = W(x)^d$. The integer power of the matrix $W(x) := e^{i \arccos(x) X}$ exactly generates the Chebyshev polynomial of the first kind as its upper-left matrix element. Furthermore, its upper-left is real. The claim follows the property of the exponential map of Pauli matrices $W(x)^d = e^{id \arccos(x) X} = \cos(d \arccos(x)) + i \sin(d \arccos(x)) X$. Then, noting that $\langle 0 | X | 0 \rangle = 0$, it yields $\langle 0 | W(x)^d | 0 \rangle = \cos(d \arccos(x)) = T_d(x)$.

To proceed, let us consider the monomial $M^{d_1, d_2, d_3}(x) := W(x)^{d_1} (iZ) W(x)^{d_2} (iZ) W(x)^{d_3}$ for any $d_1, d_2, d_3 \geq 0$. We will show that the upper-left element $\langle 0 | M^{d_1, d_2, d_3}(x) | 0 \rangle$ is real. Following the anticommutation relation $ZXZ = -X$, a useful equality can be derived $ZW(x)Z = e^{i \arccos(x)ZXZ} = e^{-i \arccos(x)X} = W(x)^{-1}$. Consequentially, by moving Z gates to the same site for cancellation, it can be shown that $\langle 0 | M^{d_1, d_2, d_3}(x) | 0 \rangle = -T_{|d_1 + d_3 - d_2|}(x)$ is real.

Note that taking second-order derivative on the unitary defined in Equation (5) at Φ^0 results in an integer-coefficient superposition of monomials of the defined form. To illustrate, given an odd integer $d = 2\tilde{d} - 1$ and $0 \leq i < j < \tilde{d}$, one can verify that

$$\frac{\partial^2}{\partial \phi_i \partial \phi_j} U(x, \tilde{\Phi}) \Big|_{\tilde{\Phi}^0} = M^{i, j-i, d-j}(x) + M^{i, d-j-i, j}(x) + M^{d-j, j-i, i}(x) + M^{j, d-j-i, i}(x).$$

Then, by taking the imaginary component, the second-order derivative g_{ij} is vanishing at $\tilde{\Phi}^0$. Thus, the second term in the right-hand side of Equation (38) vanishes since $g_{ij}(x, \tilde{\Phi}^0) = 0$.

When $d = 2(\tilde{d} - 1)$ is even, $g_i(x, \tilde{\Phi}^0) = -2\text{Im} \left[\langle 0 | iW(x)^i ZW(x)^{d-i} | 0 \rangle \right] = -2T_{d-2i}(x)$ for $i = 0, \dots, \tilde{d} - 2$ and $g_{\tilde{d}-1}(x, \tilde{\Phi}^0) = -1$. According to the discrete orthogonality of Chebyshev nodes, we have

$$\sum_{k=1}^{\tilde{d}} g_i(x_k, \tilde{\Phi}^0) g_j(x_k, \tilde{\Phi}^0) = \tilde{d} \left(2\delta_{ij} - \delta_{i, \tilde{d}-1} \delta_{j, \tilde{d}-1} \right).$$

When $d = 2\tilde{d} - 1$ is odd, we have similarly $\sum_{k=1}^{\tilde{d}} g_i(x_k, \tilde{\Phi}^0) g_j(x_k, \tilde{\Phi}^0) = 2\tilde{d}\delta_{ij}$.

Following Eq. (38), the Hessian matrix at $\tilde{\Phi}^0$ takes the form in the theorem which completes the proof. $\square$

To show that the Hessian matrix is bounded from below, we need the following lemma regarding the perturbation analysis to the singular values of a matrix.

Lemma 27 ([7, Corollary 8.6.2]). *Given $A, A+E \in \mathbb{R}^{m \times n}$ with $m \geq n$, we have*

$$|\sigma_k(A+E) - \sigma_k(A)| \leq \sigma_0(E) = \|E\|_2, \quad \forall k \in [n].$$

Then we bound the perturbation to the Jacobian matrix via that of the phase factors.

Lemma 28. *Given any symmetric phase factors Φ of length $d+1$, let $E := A(\tilde{\Phi} + \epsilon) - A(\tilde{\Phi})$. Then $\|E\|_2 \leq 4\tilde{d}^{\frac{3}{2}} \|\epsilon\|_2$.*

Proof. Recall that in the proof of Theorem 26, we have shown that taking the derivative w.r.t. ϕ_i yields terms by adding an iZ multiplication left to $e^{i\phi_i Z}$. Note that $iZ = e^{i\frac{\pi}{2}Z}$, it shifts the i -th site by $\frac{\pi}{2}$ which essentially breaks the symmetry constraint. That means we have to use asymmetric set of phase factors to characterize the derivative. To capture this, we define the evaluation map of asymmetric phase factors as follows. Given an asymmetric set of phase factors $\Phi_{\text{as}} := (\phi_0, \phi_1, \dots, \phi_d)$, we define

$$U^{(\text{as})}(x, \Phi_{\text{as}}) := e^{i\phi_0 Z} e^{i \arccos(x) X} e^{i\phi_1 Z} e^{i \arccos(x) X} \dots e^{i\phi_{d-1} Z} e^{i \arccos(x) X} e^{i\phi_d Z}.$$

We use a superscript (as) to distinguish it with $U(x, \tilde{\Phi})$ which takes the reduced phase factors $\tilde{\Phi}$ of a symmetric set as argument.

When d is odd, we define an asymmetric set of phase factors as

$$\Phi_{\text{as}}^{(i)} := \left(\phi_0, \dots, \phi_{i-1}, \phi_i + \frac{\pi}{2}, \phi_{i+1}, \dots, \phi_{\tilde{d}-1}, \phi_{\tilde{d}-1}, \dots, \phi_{i+1}, \phi_i, \phi_{i-1}, \dots, \phi_0 \right)$$

which unrolls the full set of phase factors and add a $\frac{\pi}{2}$ shift on the i -th site. As a remark, the shifted site i can be greater than $\tilde{d}$ which adds a phase shift on the right half.

Taking derivative yields that

$$\begin{aligned} g_i(x, \tilde{\Phi}) &= \text{Re} \left[\langle 0 | \frac{\partial}{\partial \phi_i} U(x, \tilde{\Phi}) | 0 \rangle \right] = \text{Re}[\langle 0 | U^{(\text{as})}(x, \Phi_{\text{as}}^{(i)}) | 0 \rangle] + \text{Re}[\langle 0 | U^{(\text{as})}(x, \Phi_{\text{as}}^{(d-i)}) | 0 \rangle] \\ &= \text{Re}[\langle 0 | U^{(\text{as})}(x, \Phi_{\text{as}}^{(i)}) | 0 \rangle] + \text{Re}[\langle 0 | U^{(\text{as})}(x, \Phi_{\text{as}}^{(i)})^\top | 0 \rangle] = 2\text{Re}[\langle 0 | U^{(\text{as})}(x, \Phi_{\text{as}}^{(i)}) | 0 \rangle]. \end{aligned}$$

Due to the unitarity, we have

$$|g_i(x, \tilde{\Phi})| \leq 2, \quad \forall i \in [\tilde{d}].$$

To further take the second derivative, let us first introduce a double-shifted asymmetric set of phase factors. When $0 \leq i < j < \tilde{d}$,

$$\Phi_{\text{as}}^{(i,j)} := \left(\phi_0, \dots, \phi_{i-1}, \phi_i + \frac{\pi}{2}, \phi_{i+1}, \dots, \phi_{j-1}, \phi_j + \frac{\pi}{2}, \phi_{j+1}, \dots, \phi_{\tilde{d}-1}, \phi_{\tilde{d}-1}, \dots, \phi_0 \right).$$

When $0 \leq i = j < \tilde{d}$,

$$\Phi_{\text{as}}^{(i,i)} := \left(\phi_0, \dots, \phi_{i-1}, \phi_i + \pi, \phi_{i+1}, \dots, \phi_{\tilde{d}-1}, \phi_{\tilde{d}-1}, \dots, \phi_{i+1}, \phi_i, \phi_{i-1}, \dots, \phi_0 \right).$$

Similarly, when i or $j \geq \tilde{d}$, the phase shift is added on the right half accordingly.

With the defined notation, we are able to compute the second derivative as

$$\begin{aligned} g_{ij}(x, \tilde{\Phi}) &:= \frac{\partial}{\partial \phi_j} g_i(x, \tilde{\Phi}) = 2\text{Re}[\langle 0 | \frac{\partial}{\partial \phi_j} U^{(\text{as})}(x, \Phi_{\text{as}}^{(i)}) | 0 \rangle] \\ &= 2\text{Re}[\langle 0 | U^{(\text{as})}(x, \Phi_{\text{as}}^{(i,j)}) | 0 \rangle] + 2\text{Re}[\langle 0 | U^{(\text{as})}(x, \Phi_{\text{as}}^{(i,d-j)}) | 0 \rangle]. \end{aligned}$$

Due to the unitarity, we have

$$|g_{ij}(x, \tilde{\Phi})| \leq 4, \quad \forall i, j \in [\tilde{d}].$$

Therefore, for any given odd integer d and symmetric $\Phi \in \mathbb{R}^{d+1}$, we have

$$\|\nabla g_i(x, \tilde{\Phi})\|_2 = \sqrt{\sum_{j \in [\tilde{d}]} |g_{ij}(x, \tilde{\Phi})|^2} \leq 4\sqrt{\tilde{d}}, \quad \forall i \in [\tilde{d}]. \quad (81)$$

When d is even, the full set of phase factors is $(\phi_0, \dots, \phi_{\tilde{d}-2}, \phi_{\tilde{d}-1}, \phi_{\tilde{d}-2}, \dots, \phi_0)$. The only difference in the unrolled structure comparing to that in the odd case is that there is a single unpaired phase factor $\phi_{\tilde{d}-1}$ at the middle. Therefore, when the derivative only involves $\phi_0, \dots, \phi_{\tilde{d}-2}$, the analysis is exactly the same as that when d is odd.

When $i = \tilde{d} - 1$, taking derivative *w.r.t.* $\phi_{\tilde{d}-1}$ is equivalently shifting this site by $\frac{\pi}{2}$, namely,

$$g_{\tilde{d}-1}(x, \tilde{\Phi}) = \text{Re}[\langle 0 | U(x, \tilde{\Psi}) | 0 \rangle] = g(x, \tilde{\Psi}),$$

where

$$\tilde{\Psi} := \Phi + \frac{\pi}{2} \mathbf{e}_{\tilde{d}-1} = (\phi_0, \dots, \phi_{\tilde{d}-2}, \phi_{\tilde{d}-1} + \frac{\pi}{2}, \phi_{\tilde{d}-2}, \dots, \phi_0).$$

Note that Ψ is a set of symmetric phase factors (and therefore we use a different notation). Then, taking the second derivative on $g(x, \tilde{\Phi})$ is equivalent to taking the first derivative on $g(x, \tilde{\Psi})$ whose analysis is the same as that in the previous case. Therefore,

$$g_{i, \tilde{d}-1}(x, \tilde{\Phi}) = g_i(x, \tilde{\Psi}) = 2\text{Re} \left[\langle 0 | U^{(\text{as})}(x, \Psi_{\text{as}}^{(i)}) | 0 \rangle \right], \quad \forall i \in [\tilde{d}-1],$$

and

$$g_{\tilde{d}-1, \tilde{d}-1}(x, \tilde{\Phi}) = g_{\tilde{d}-1}(x, \tilde{\Psi}) = -g(x, \tilde{\Phi}).$$

These results lead to the inequality

$$|g_{i, \tilde{d}-1}(x, \tilde{\Phi})| \leq 2, \quad \forall i \in [\tilde{d}].$$

Thus, the inequality in Eq. (81) also holds when d is even.

To conclude, for any given integer d and symmetric $\Phi \in \mathbb{R}^{d+1}$, we obtain a uniform upper bound of the gradient of $g_j(x, \tilde{\Phi})$

$$\|\nabla g_j(x, \tilde{\Phi})\|_2 \leq 4\sqrt{\tilde{d}}, \quad \forall j \in [\tilde{d}], \quad \forall x \in [-1, 1]. \quad (82)$$

Following the mean value theorem, there exists $t \in (0, 1)$ so that for any $j \in [\tilde{d}]$ and $x \in [-1, 1]$

$$|g_j(x, \tilde{\Phi} + \epsilon) - g_j(x, \tilde{\Phi})| \leq \|\nabla g_j(x, \tilde{\Phi} + t\epsilon)\|_2 \|\epsilon\|_2 \leq 4\sqrt{\tilde{d}} \|\epsilon\|_2.$$

Furthermore

$$\|A(\tilde{\Phi} + \epsilon) - A(\tilde{\Phi})\|_F = \sqrt{\sum_{j=0}^{\tilde{d}-1} \sum_{k=1}^{\tilde{d}} (g_j(x_k, \tilde{\Phi} + \epsilon) - g_j(x_k, \tilde{\Phi}))^2} \leq 4\tilde{d}^{\frac{3}{2}} \|\epsilon\|_2.$$

Then, the theorem follows

$$\|A(\tilde{\Phi} + \epsilon) - A(\tilde{\Phi})\|_2 \leq \|A(\tilde{\Phi} + \epsilon) - A(\tilde{\Phi})\|_F \leq 4\tilde{d}^{\frac{3}{2}} \|\epsilon\|_2.$$

□

Then the perturbation analysis to the singular value of the Jacobian matrix is obtained by directly applying Lemma 27.

Corollary 29. *If $\|\epsilon\|_2 \leq \frac{1}{20\tilde{d}}$, then*

$$\frac{4}{5}\sqrt{\tilde{d}} \leq \sigma_{\min}(A(\tilde{\Phi}^0 + \epsilon)) \leq \sigma_{\max}(A(\tilde{\Phi}^0 + \epsilon)) \leq \left(\frac{1}{5} + \sqrt{2}\right)\sqrt{\tilde{d}}.$$

Proof. Applying Lemma 27, we have

$$|\sigma_k(A(\tilde{\Phi}^0 + \epsilon)) - \sigma_k(A(\tilde{\Phi}^0))| \leq 4\tilde{d}^{\frac{3}{2}} \|\epsilon\|_2. \quad (83)$$

As a consequence of Theorem 26, $\sigma_{\max}(A(\tilde{\Phi}^0)) = \sqrt{2\tilde{d}}$ and

$$\sigma_{\min}(A(\tilde{\Phi}^0)) = \begin{cases} \sqrt{\tilde{d}} & , \text{ when } d \text{ is even,} \\ \sqrt{2\tilde{d}} & , \text{ when } d \text{ is odd.} \end{cases}$$

When $\|\epsilon\|_2 \leq \frac{1}{20\tilde{d}}$, we have

$$\sigma_{\min}(A(\tilde{\Phi}^0 + \epsilon)) \geq \sigma_{\min}(A(\tilde{\Phi}^0)) - \|E\|_2 \geq \sqrt{\tilde{d}} - 4\tilde{d}^{\frac{3}{2}} \|\epsilon\|_2 \geq \frac{4}{5}\sqrt{\tilde{d}}. \quad (84)$$

and

$$\sigma_{\max}(A(\tilde{\Phi}^0 + \epsilon)) \leq \sigma_{\max}(A(\tilde{\Phi}^0)) + \|E\|_2 \leq \sqrt{2\tilde{d}} + 4\tilde{d}^{\frac{3}{2}} \|\epsilon\|_2 \leq \left(\frac{1}{5} + \sqrt{2}\right) \sqrt{\tilde{d}}. \quad (85)$$

□

Recall that the Hessian matrix at the maximal solution satisfies $\text{Hess}(\tilde{\Phi}^*) = \frac{2}{\tilde{d}}A(\tilde{\Phi}^*)^\top A(\tilde{\Phi}^*)$. The positive definiteness of this matrix follows by combining the previous results with Theorem 5.

Theorem 30. *Let Φ^* be the optimal phase factors corresponding to the maximal solution. If $\|f\|_\infty \leq \frac{\sqrt{3}}{20\pi\tilde{d}}$, then $\text{Hess}(\tilde{\Phi}^*)$ is positive definite and*

$$\lambda_{\min}(\text{Hess}(\tilde{\Phi}^*)) \geq \frac{32}{25}. \quad (86)$$

Proof. Theorem 5 implies that $\|\tilde{\Phi}^* - \tilde{\Phi}^0\|_2 \leq \frac{1}{20\tilde{d}}$. As a consequence of Corollary 29, we have $\sigma_{\min}(A(\tilde{\Phi}^*)) \geq \frac{4}{5}\sqrt{\tilde{d}}$. Therefore

$$\lambda_{\min}(\text{Hess}(\tilde{\Phi}^*)) = \lambda_{\min}\left(\frac{2}{\tilde{d}}A(\tilde{\Phi}^*)^\top A(\tilde{\Phi}^*)\right) = \frac{2}{\tilde{d}}\sigma_{\min}^2(A(\tilde{\Phi}^*)) \geq \frac{32}{25},$$

which completes the proof. □

6.2 Lower bounding the distance to the initial point

In the last subsection, we show that the Hessian matrix is positive definite at the maximal solution. To show the positive definiteness of the Hessian matrix in a neighborhood of the optima, we need the following lower bound as an intermediate step.

Theorem 31. *Given any symmetric phase factor $\tilde{\Phi}$ satisfying $\|\tilde{\Phi} - \tilde{\Phi}^0\|_2 \leq \frac{1}{\sqrt{2}}$, the following inequality holds*

$$\frac{1}{\sqrt{3}} \|g(x, \tilde{\Phi})\|_T \leq \|\tilde{\Phi} - \tilde{\Phi}^0\|_2. \quad (87)$$

Proof. We denote the expansion of the polynomial $g(x, \tilde{\Phi})$ in the Chebyshev basis as

$$g(x, \tilde{\Phi}) = \sum_{k=0}^d g_k(\tilde{\Phi}) T_k(x).$$

For simplicity, we drop the $\tilde{\Phi}$ -dependence in the Chebyshev coefficient g_k (in particular, g_k is not the derivative of $g(x, \tilde{\Phi})$ with respect to ϕ_k as in the previous section).

Then we apply Eq. (26) to estimate q_{d-1} , the coefficient of $Q(x)$. Notice that $\cos(\phi_\ell)$ is positive for any $1 \leq \ell \leq \tilde{d} - 1$, which follows $|\phi_\ell| \leq \|\tilde{\Phi} - \tilde{\Phi}^0\|_2 \leq \frac{1}{\sqrt{2}} < \frac{\pi}{4}$. We use Eq. (26) and the inequality $\cos(\phi_{\tilde{d}-1}) \geq \cos(\phi_{d-1})^2$ when d is even, and obtain

$$q_{d-1}^{(0)} \geq \prod_{\ell=1}^{\tilde{d}-1} \cos^2(\phi_\ell) = \prod_{\ell=1}^{\tilde{d}-1} (1 - \sin^2(\phi_\ell)) \geq 1 - \sum_{\ell=1}^{\tilde{d}-1} \phi_\ell^2 \geq 1 - \|\tilde{\Phi} - \tilde{\Phi}^0\|_2^2 \geq \frac{1}{2}. \quad (88)$$

Hence $\deg(Q) = d - 1$ and the leading coefficient of Q is positive. Now we are able to apply Theorem 22 and Lemma 23 and have

$$\sqrt{1 - \|g(x, \tilde{\Phi})\|_\infty^2} \leq q_{d-1}^{(0)} \leq q_{d-3}^{(1)} \leq \cdots \leq q_{d-1-2\ell}^{(\ell)} \leq \cdots \leq q_{d-1-2\hat{d}}^{(\hat{d})} \leq 1.$$

According to the reduction procedure, the leading Chebyshev coefficients and the phase factor are connected as follows

$$\left| \sin(2\phi_0 - \frac{\pi}{2}) \right|^2 = \left| \frac{\text{Re}[p_d]}{q_{d-1}} \right|^2 = \frac{g_d^2}{q_{d-1}^2} \geq g_d^2, \quad (89)$$

where the last inequality follows $|q_{d-1}| \leq 1$.

Following Eq. (36), for any $1 \leq \ell \leq \hat{d}$, we have

$$\begin{aligned} |\sin(2\phi_\ell)|^2 &= \left| \frac{\text{Im}[p_{d-2\ell}^{(\ell-1)} e^{-2i\phi_{\ell-1}}]}{q_{d-1-2\ell}^{(\ell)}} \right|^2 \geq \left| \text{Im}[p_{d-2\ell}^{(\ell-1)} e^{-2i\phi_{\ell-1}}] \right|^2 \\ &= 4q_{d+1-2\ell}^{(\ell-1)} \left(q_{d-1-2\ell}^{(\ell)} - q_{d+1-2\ell}^{(\ell-1)} \right) \geq 4q_{d-1}^{(0)} \left(q_{d-1-2\ell}^{(\ell)} - q_{d+1-2\ell}^{(\ell-1)} \right), \end{aligned}$$

where the last inequality follows the monotonicity in Theorem 22.

Now we are going to show

$$\|\tilde{\Phi} - \tilde{\Phi}^0\|_2^2 \geq \frac{1}{4}g_d^2 + q_{d-1}^{(0)} \left(1 - q_{d-1}^{(0)} \right), \quad (90)$$

where g_d denotes the leading Chebyshev coefficient of $g(x, \tilde{\Phi})$.

When d is odd, we have

$$\sum_{\ell=1}^{\hat{d}} |\sin(2\phi_\ell)|^2 \geq 4q_{d-1}^{(0)} \sum_{\ell=1}^{\hat{d}} \left(q_{d-1-2\ell}^{(\ell)} - q_{d+1-2\ell}^{(\ell-1)} \right) = 4q_{d-1}^{(0)} \left(q_0^{(\hat{d})} - q_{d-1}^{(0)} \right) = 4q_{d-1}^{(0)} \left(1 - q_{d-1}^{(0)} \right).$$

Here, we have used the fact that $q_0^{(\hat{d})} = 1$ obtained from the proof of Lemma 14. Hence

$$\begin{aligned} \|\tilde{\Phi} - \tilde{\Phi}^0\|_2^2 &= \left(\phi_0 - \frac{\pi}{4} \right)^2 + \sum_{\ell=1}^{\hat{d}} \phi_\ell^2 \\ &\geq \frac{1}{4} \left| \sin(2\phi_0 - \frac{\pi}{2}) \right|^2 + \frac{1}{4} \sum_{\ell=1}^{\hat{d}} |\sin(2\phi_\ell)|^2 \\ &\geq \frac{1}{4}g_d^2 + q_{d-1}^{(0)} \left(1 - q_{d-1}^{(0)} \right). \end{aligned}$$

When d is even, the dangling unpaired phase factor $\phi_{\tilde{d}-1}$ has to be analyzed individually

$$\left| \sin(\phi_{\tilde{d}-1}) \right|^2 = 1 - \left(q_1^{(\tilde{d}-2)} \right)^2.$$

Furthermore, we have

$$\sum_{\ell=1}^{\hat{d}} |\sin(2\phi_\ell)|^2 \geq 4q_{d-1}^{(0)} \sum_{\ell=1}^{\hat{d}} \left(q_{d-1-2\ell}^{(\ell)} - q_{d+1-2\ell}^{(\ell-1)} \right) = 4q_{d-1}^{(0)} \left(q_1^{(\tilde{d}-2)} - q_{d-1}^{(0)} \right).$$

Unlike the odd case, the coefficient $q_1^{(\tilde{d}-2)}$ is not necessary equal to 1 when d is even. Then

$$\begin{aligned}
\|\tilde{\Phi} - \tilde{\Phi}^0\|_2^2 &= \left(\phi_0 - \frac{\pi}{4}\right)^2 + \sum_{\ell=1}^{\hat{d}} \phi_\ell^2 + \phi_{\tilde{d}-1}^2 \\
&\geq \frac{1}{4} \left| \sin(2\phi_0 - \frac{\pi}{2}) \right|^2 + \frac{1}{4} \sum_{\ell=1}^{\hat{d}} |\sin(2\phi_\ell)|^2 + \left| \sin(\phi_{\tilde{d}-1}) \right|^2 \\
&\geq \frac{1}{4} g_d^2 + q_{d-1}^{(0)} \left(q_1^{(\tilde{d}-2)} - q_{d-1}^{(0)} \right) + 1 - \left(q_1^{(\tilde{d}-2)} \right)^2 \\
&= \frac{1}{4} g_d^2 + q_{d-1}^{(0)} \left(1 - q_{d-1}^{(0)} \right) + \left(1 - q_{d-1}^{(0)} \right) \left(1 - q_1^{(\tilde{d}-2)} \right) + q_1^{(\tilde{d}-2)} \left(1 - q_1^{(\tilde{d}-2)} \right) \\
&\geq \frac{1}{4} g_d^2 + q_{d-1}^{(0)} \left(1 - q_{d-1}^{(0)} \right).
\end{aligned}$$

According to Eq. (88) and Lemma 23, we have $\frac{1}{2} \leq q_{d-1}^{(0)} \leq \sqrt{1 - \|g(x, \tilde{\Phi})\|_T^2 + \frac{1}{2} g_d^2}$. It implies that

$$\begin{aligned}
\|\tilde{\Phi} - \tilde{\Phi}^0\|_2^2 &\geq \frac{1}{4} g_d^2 + \sqrt{1 - \|g(x, \tilde{\Phi})\|_T^2 + \frac{1}{2} g_d^2} - 1 + \|g(x, \tilde{\Phi})\|_T^2 - \frac{1}{2} g_d^2 \\
&\geq \|g(x, \tilde{\Phi})\|_T^2 - \frac{1}{4} g_d^2 - \frac{\|g(x, \tilde{\Phi})\|_T^2 - \frac{1}{2} g_d^2}{\sqrt{1 - \|g(x, \tilde{\Phi})\|_T^2 + \frac{1}{2} g_d^2} + 1} \\
&\geq \|g(x, \tilde{\Phi})\|_T^2 - \frac{1}{4} g_d^2 - \frac{2 \|g(x, \tilde{\Phi})\|_T^2 - g_d^2}{3} \\
&\geq \frac{1}{3} \|g(x, \tilde{\Phi})\|_T^2.
\end{aligned}$$

□

6.3 Local strong convexity of the optimization problem

Now we are going to present the proof of Theorem 6, where we use Corollary 29 and the following lemma regarding the perturbative analysis to the eigenvalues of a matrix.

Lemma 32 ([7, Theorem 8.1.4] **Wielandt–Hoffman**). *If $A, A+E \in \mathbb{R}^{n \times n}$ are symmetric matrices, then*

$$\sum_{k \in [n]} (\lambda_k(A+E) - \lambda_k(A))^2 \leq \|E\|_F^2.$$

Although Weyl's inequality may give a sharper estimate about the perturbation of eigenvalues, we adopt Lemma 32 in our analysis. That is because the former exploits the operator norm of the perturbation of matrix, which is hard to estimate in our problem.

Proof of Theorem 6. Recall that

$$F(\tilde{\Phi}) = \frac{2}{\tilde{d}} \left(A^\top(\tilde{\Phi}) A(\tilde{\Phi}) \right) + R(\tilde{\Phi}), \tag{91}$$

where $R_{ij} := \frac{2}{\tilde{d}} \sum_{k=1}^{\tilde{d}} \left(g(x_k, \tilde{\Phi}) - f(x_k) \right) g_{ij}(x_k, \tilde{\Phi})$. The residual term is estimated as follows

$$\begin{aligned}
\|R\|_F^2 &= \sum_{i=0}^{\tilde{d}-1} \sum_{j=0}^{\tilde{d}-1} \left(\frac{2}{\tilde{d}} \sum_{k=1}^{\tilde{d}} \left(g(x_k, \tilde{\Phi}) - f(x_k) \right) g_{ij}(x_k, \tilde{\Phi}) \right)^2 \\
&\leq \frac{4}{\tilde{d}^2} \sum_{i=0}^{\tilde{d}-1} \sum_{j=0}^{\tilde{d}-1} \left(\sum_{k=1}^{\tilde{d}} \left(g(x_k, \tilde{\Phi}) - f(x_k) \right)^2 \right) \left(\sum_{k=1}^{\tilde{d}} g_{ij}(x_k, \tilde{\Phi})^2 \right) \\
&\leq 8 \left(\sum_{k=1}^{\tilde{d}} g(x_k, \tilde{\Phi})^2 + \sum_{k=1}^{\tilde{d}} f(x_k)^2 \right) \max_{i,j} \left(\sum_{k=1}^{\tilde{d}} g_{ij}(x_k, \tilde{\Phi})^2 \right) \\
&= 8\tilde{d}^2 \left(\|g(x, \tilde{\Phi})\|_T^2 + \|f(x)\|_T^2 \right) \max_{i,j} \|g_{ij}(x, \tilde{\Phi})\|_T^2.
\end{aligned} \tag{92}$$

Here, the last line uses the discrete orthogonality of Chebyshev polynomial.

Then, we have the following inequality

$$\|R\|_F^2 \leq 128\tilde{d}^2 \left(3 \|\tilde{\Phi} - \tilde{\Phi}^0\|_2^2 + \|f\|_\infty^2 \right) \leq \frac{24}{25} \left(1 + \frac{1}{\pi^2} \right), \tag{93}$$

where the first inequality uses Theorem 31, the fact $\|f\|_T \leq \|f\|_\infty$, and its consequence $\|g_{ij}(x, \tilde{\Phi})\|_T \leq \|g_{ij}(x, \tilde{\Phi})\|_\infty \leq 4$ derived in the proof of Lemma 28.

Then, applying Lemma 32, one has

$$\begin{aligned}
\lambda_{\min}(\text{Hess}(\tilde{\Phi})) &\geq \lambda_{\min} \left(\frac{2}{\tilde{d}} A(\tilde{\Phi})^\top A(\tilde{\Phi}) \right) - \|R\|_F \geq \frac{32}{25} - \sqrt{\frac{24}{25} \left(1 + \frac{1}{\pi^2} \right)} \geq \frac{1}{4}, \\
\lambda_{\max}(\text{Hess}(\tilde{\Phi})) &\leq \lambda_{\max} \left(\frac{2}{\tilde{d}} A(\tilde{\Phi})^\top A(\tilde{\Phi}) \right) + \|R\|_F \leq 2 \left(\frac{1}{5} + \sqrt{2} \right)^2 + \sqrt{\frac{24}{25} \left(1 + \frac{1}{\pi^2} \right)} \leq \frac{25}{4}.
\end{aligned} \tag{94}$$

□

Remark 33. According to Theorem 6, we know that if $\|f\|_\infty \leq \frac{\sqrt{3}}{20\pi\tilde{d}}$, the optimization problem Eq. (5) is strongly convex in the following region

$$\mathcal{S} := \left\{ \tilde{\Phi} \in \mathbb{R}^{\tilde{d}} : \|\tilde{\Phi} - \tilde{\Phi}^0\|_2 \leq \frac{1}{20\tilde{d}} \right\}. \tag{95}$$

Furthermore, the maximal solution is included in that region $\tilde{\Phi}^* \in \mathcal{S}$. For convenience, the lower bound of the Hessian matrix is referred to as $\sigma = \frac{1}{4}$, while the upper bound is referred to as $L = \frac{25}{4}$.

6.4 Projected gradient descend algorithm for symmetric phase-factor evaluation

We list a practical algorithm for symmetric phase-factor evaluation in Algorithm 1 which is based on the projected gradient descent method [2]. In Corollary 7, we give the convergence analysis of this algorithm and here is the proof.

Proof of Corollary 7. According to Theorem 6, the cost function $F(\tilde{\Phi})$ is σ -strongly convex and L -smooth on ball $\mathcal{S} := \{\tilde{\Phi} : \|\tilde{\Phi} - \tilde{\Phi}^0\| \leq \frac{1}{20\tilde{d}}\}$, where $\sigma = \frac{25}{4}$ and $L = \frac{1}{4}$. Using

Algorithm 1 Projected gradient descend algorithm for optimizing symmetric phase factors (with performance guarantee)

Input: A target real polynomial f of degree d so that it has definite parity and $\|f\|_\infty \leq \frac{\sqrt{3}}{20\pi d}$, an initial guess $\tilde{\Phi}^0 = (\frac{\pi}{4}, 0, \dots, 0) \in \mathcal{S}$ and the error tolerance ϵ .

Output: A set of reduced phase factors $\tilde{\Phi}$ so that $F(\tilde{\Phi}) \leq \epsilon$ implying $\|g(x, \tilde{\Phi}) - f(x)\|_\infty \leq d\sqrt{\epsilon}$ [5, Theorem 4].

Evaluate the objective function $F(\tilde{\Phi})$ at $\tilde{\Phi}^0$ according to Eq. (5).

while $F(\tilde{\Phi}) > \epsilon$ **do**

 Evaluate the gradient $\nabla F(\tilde{\Phi})$ of the objective.

 Update $\tilde{\Phi}_i$ by gradient descent $\tilde{\Phi} \leftarrow \tilde{\Phi} - \frac{1}{L} \nabla F(\tilde{\Phi})$.

if $\|\tilde{\Phi} - \tilde{\Phi}^0\|_2 > \frac{1}{20d}$ **then**

 Project the point onto the ball $\tilde{\Phi} \leftarrow \tilde{\Phi}^0 - \frac{\tilde{\Phi} - \tilde{\Phi}^0}{20d\|\tilde{\Phi} - \tilde{\Phi}^0\|_2}$.

end if

end while

return $\tilde{\Phi}$

Theorem 5, we know that $\tilde{\Phi}^* \in \mathcal{S}$, since $\|\tilde{\Phi}^* - \tilde{\Phi}^0\| \leq \frac{\pi}{\sqrt{3}} \|f\|_\infty \leq \frac{1}{20d}$. Applying [2, Theorem 3.10], the projected gradient method with step size $\frac{1}{L}$ converges exponentially, that is,

$$\|\tilde{\Phi}^\ell - \tilde{\Phi}^*\|_2 \leq e^{-\frac{\sigma}{L}\ell} \|\tilde{\Phi}^0 - \tilde{\Phi}^*\|_2,$$

where $\tilde{\Phi}^\ell$ is the set of reduced phase factors in the ℓ -th iteration step. In particular, $\tilde{\Phi}^\ell \in \mathcal{S}$ for all ℓ follows the construction of the algorithm. $\square$

Let $\tilde{\Phi}^\ell$ be the set of reduced phase factors in the ℓ -th iteration step. Note that $F(\tilde{\Phi}^*) = 0$, $\nabla F(\tilde{\Phi}^*) = 0$ and $\tilde{\Phi}^* \in \mathcal{S}$. Following Taylor's theorem, there exists $s \in (0, 1)$ so that

$$\begin{aligned} F(\tilde{\Phi}^\ell) &= F(\tilde{\Phi}^*) + \nabla F(\tilde{\Phi}^*)^\top (\tilde{\Phi}^\ell - \tilde{\Phi}^*) + \frac{1}{2} (\tilde{\Phi}^\ell - \tilde{\Phi}^*)^\top \text{Hess}(s\tilde{\Phi}^\ell + (1-s)\tilde{\Phi}^*) (\tilde{\Phi}^\ell - \tilde{\Phi}^*) \\ &= \frac{1}{2} (\tilde{\Phi}^\ell - \tilde{\Phi}^*)^\top \text{Hess}(s\tilde{\Phi}^\ell + (1-s)\tilde{\Phi}^*) (\tilde{\Phi}^\ell - \tilde{\Phi}^*). \end{aligned}$$

Because $\tilde{\Phi}^\ell, \tilde{\Phi}^* \in \mathcal{S}$ and $\mathcal{S}$ is convex, the intermediate point also lies in the convex set $s\tilde{\Phi}^\ell + (1-s)\tilde{\Phi}^* \in \mathcal{S}$. Then, the Hessian is upper bounded at the intermediate point by Theorem 6. Thus

$$F(\tilde{\Phi}^\ell) \leq \frac{1}{2} \lambda_{\max} \left(\text{Hess}(s\tilde{\Phi}^\ell + (1-s)\tilde{\Phi}^*) \right) \|\tilde{\Phi}^\ell - \tilde{\Phi}^*\|_2^2 \leq \frac{L}{2} \|\tilde{\Phi}^\ell - \tilde{\Phi}^*\|_2^2.$$

Furthermore, using Corollary 7, we have

$$F(\tilde{\Phi}^\ell) \leq \frac{L}{2} \|\tilde{\Phi}^\ell - \tilde{\Phi}^*\|_2^2 \leq \frac{L}{2} e^{-\frac{\sigma}{L}\ell} \|\tilde{\Phi}^0 - \tilde{\Phi}^*\|_2^2.$$

To obtain $F(\tilde{\Phi}^\ell) \leq \epsilon$, it suffices to take $\ell = \mathcal{O} \left(\log \left(\frac{1}{\epsilon} \right) \right)$ iterations.

Remarkably, Algorithm 1 does not require a large number of bits. To see the claim, we consider the model of finite precision arithmetic (see standard axioms in [9]). The parameter u is referred to the unit roundoff. In each iteration step, the rounding error when computing the function F and its gradient is $\mathcal{O}(d^2 u)$. Furthermore, the projection is

needed to be executed when the proposed new Φ lies outside $\mathcal{S}$. Under this circumstance, the possible error amplification by taking square roots and division does not contribute significantly and the total rounding error accumulated in each iteration step remains $\mathcal{O}(d^2u)$. To conclude, the numerical error of the algorithm is upper bounded by invoking triangle inequality. Then, there exists some constant C so that

$$Cd^2u \log \frac{1}{\epsilon} \leq \epsilon \Rightarrow \frac{1}{u} := 2^{n_{\text{bits}}} \geq \frac{1}{C} \frac{d^2 \log(1/\epsilon)}{\epsilon} \Rightarrow n_{\text{bits}} = \mathcal{O} \left(\log \left(\frac{d \log(1/\epsilon)}{\epsilon} \right) \right).$$

Therefore, the number of bits required by Algorithm 1 is $\mathcal{O} \left(\log \left(\frac{d \log(1/\epsilon)}{\epsilon} \right) \right)$. Therefore the algorithm is numerically stable, and it is sufficient to use the standard double precision arithmetic operations.

7 Additional numerical results

In this section, we present some additional numerical results demonstrating some of the features of the energy landscape of the optimization problem for symmetric phase factors in Eq. (5).

7.1 Hessian matrix without symmetry constraint is singular at global optima

When the symmetry constraint is not imposed, the Hessian matrix can be derived from Eq. (38). Note that the Jacobian matrix $A \in \mathbb{R}^{\tilde{d} \times d}$ does not have full column rank because $d > \tilde{d}$. Then, following that $\ker(A^\top A) = \ker(A)$, we have $\text{rank}(A^\top A) = \text{rank}(A) \leq \tilde{d} < d$. Note that at the global optima, the Hessian matrix is precisely $\text{Hess}(\Phi^*) = (A^*)^\top A^*$ due to the vanishing residual term $R = 0$. Here, A^* refers to the Jacobian matrix at the optima. Thus, the Hessian matrix at optima is singular. Because the residual term R is small near the optima by continuity, the Hessian matrix is almost singular in a neighborhood of the optima. This behavior is demonstrated by the numerical result displayed in Fig. 1 (a). On the other hand, when the symmetry constraint is imposed, Fig. 1 (b) shows that the Hessian matrix is positive definite in a neighborhood of the maximal solution and it agrees our theoretical results about the local strong convexity of the Hessian matrix.

7.2 Existence of local minima

To visualize the full energy landscape of the optimization problem, we first consider the simplest scenario so that only two phase factors suffice to solve the optimization problem exactly. We choose two target polynomials $f(x) = x^2 - \frac{1}{2}$ and $f(x) = \frac{1}{\sqrt{3}}x^3 - \frac{2}{\sqrt{3}}x$. The energy landscape on the irreducible domain is displayed in Fig. 2. The global minima derived from the proposed method are annotated in the domain. In this special case, there are no local minima. However, this is a rare exception rather than the norm. Existence of local minima can be observed by slightly increasing the degree of the target polynomial $f(x) = 0.2103T_4(x) + 0.1231T_2(x) + 0.1666$. The local minimizer is numerically searched by randomly initiating the stochastic gradient descent algorithm to solve Eq. (5). The value of the objective function at Φ^{loc} is 0.0218. The Hessian matrix has eigenvalues 0.1359, 4.5815, 7.7510, which indicates that Φ^{loc} is indeed a local minimum. To visualize the landscape, we plot a two-dimensional section of the optimization landscape spanned by the affine plane spanned by the eigenvectors of the least two eigenvalues of the Hessian matrix (Fig. 2).

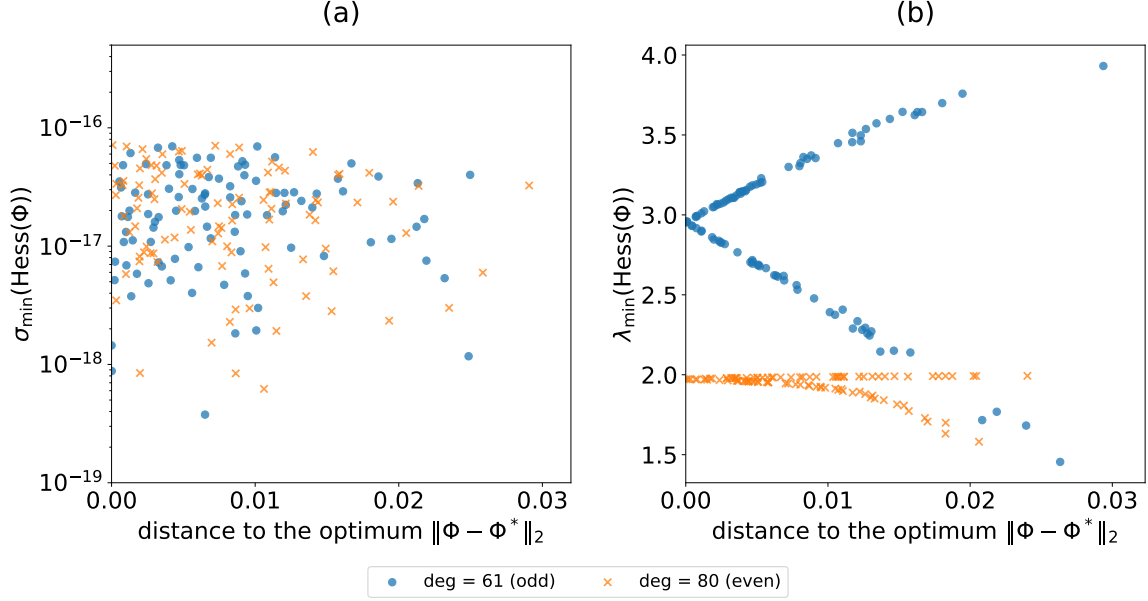


Figure 1: The smallest singular value ($\sigma_{\min}$) or eigenvalue ($\lambda_{\min}$) of the Hessian matrix evaluated at 100 randomly sampled Φ near the optimizer Φ^* of a given polynomial. Blue dots correspond to an odd polynomial of degree 61 and orange crosses correspond to an even polynomial of degree 80. (a) The symmetry constraint is not imposed. The plotted data concentrates around the machine precision $\sim 10^{-17}$, which implies a large basin near the optimizer on which the Hessian matrix is singular if the symmetry constraint is not imposed. (b) The symmetry constraint is imposed. The numerical results show that the Hessian matrix is well conditioned near the optimizer, and it is positive definite which agrees our result of local strong convexity.

Next we present an example to demonstrate the complexity of the full landscape of the optimization problem. We choose $f(x) = \frac{1}{\alpha}(T_4(x) + 2T_2(x) + T_0(x))$ as the target polynomial. The scale factor $\alpha = 440$ is set to be large enough and satisfies the requirements of Theorem 6. We first found a global minimum around $\tilde{\Phi}^0 = (\frac{\pi}{4}, 0, 0)$, by taking $\tilde{\Phi}^0$ as the initial guess for Algorithm 1. We also numerically searched two local minimizers by randomly initiating the stochastic gradient descent algorithm to solve Eq. (5). We use $\tilde{\Phi}^*$, $\tilde{\Phi}^1$ and $\tilde{\Phi}^2$ to denote the global minimum and two local minima respectively. Their values are as follows:

$$\begin{aligned}\tilde{\Phi}^* &= (0.7843, -0.0023, -0.0023), \\ \tilde{\Phi}^1 &= (1.2317, 1.5685, 2.2467), \\ \tilde{\Phi}^2 &= (0.4370, 1.5731, 0.6991).\end{aligned}$$

The values of the objective function at $\tilde{\Phi}^1$ and $\tilde{\Phi}^2$ are both around $2.5769\text{e-}06$. The Hessian matrix at both points are positive semidefinite, which indicates that they are indeed local minima. To visualize the landscape, we plot a two-dimensional section of the optimization landscape spanned by the affine plane spanned by $\tilde{\Phi}^1 - \tilde{\Phi}^*$ and $\tilde{\Phi}^2 - \tilde{\Phi}^*$ (Fig. 4). It agrees with our conclusion that the energy landscape of the optimization problem remains complex and has many local minima as well as global minima. This confirms that even when $\|f\|_\infty$ is sufficiently small, the global energy landscape is still very complex, and the local energy landscape described by Theorem 6 is nontrivial.

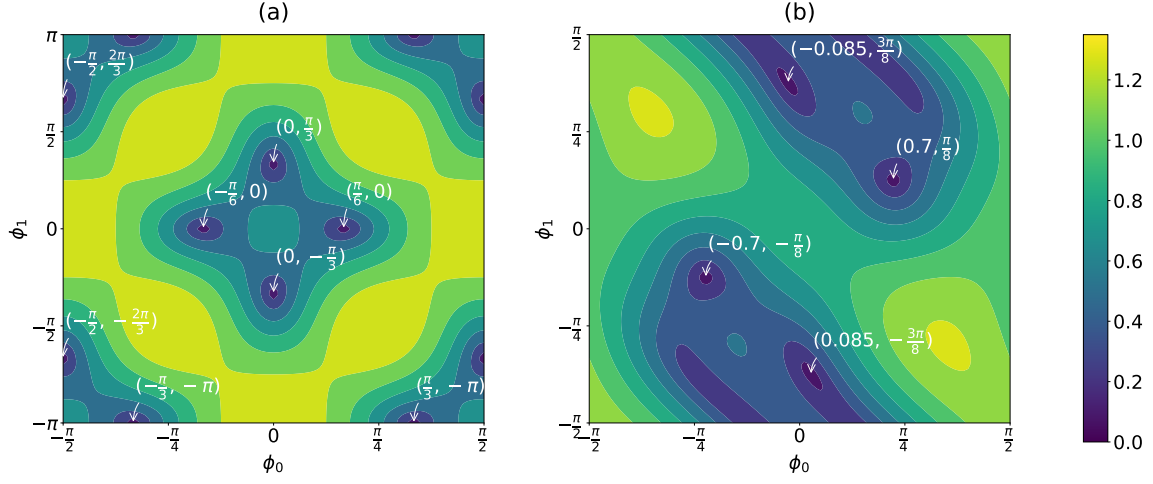


Figure 2: The optimization landscape of the modified cost function $F(\tilde{\Phi})^{1/3}$ on the irreducible domain D_d . Here, the cube root is taken to signify the structure of the landscape near optima. The annotated values are all optima exactly computed from the proposed method. (a) The target function is set to $f(x) = x^2 - \frac{1}{2}$ which yields 8 inequivalent optima. (b) The target function is set to $f(x) = \frac{1}{\sqrt{3}}x^3 - \frac{2}{\sqrt{3}}x$ which yields 4 inequivalent optima.

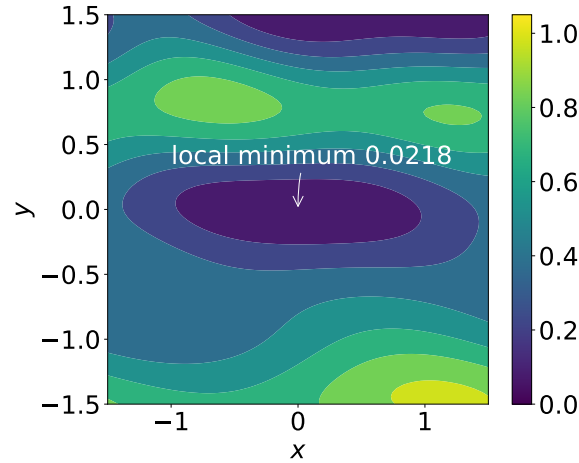


Figure 3: A two-dimensional section of the optimization landscape $F(\tilde{\Phi}^{\text{loc}} + xv_1 + yv_2)$, where Φ^{loc} is a local minimum of a given degree-5 target polynomial, and v_1, v_2 are unit eigenvectors corresponding to the least two eigenvalues of the Hessian matrix evaluated at Φ^{loc} . The objective value of the local minimum is annotated on the figure.

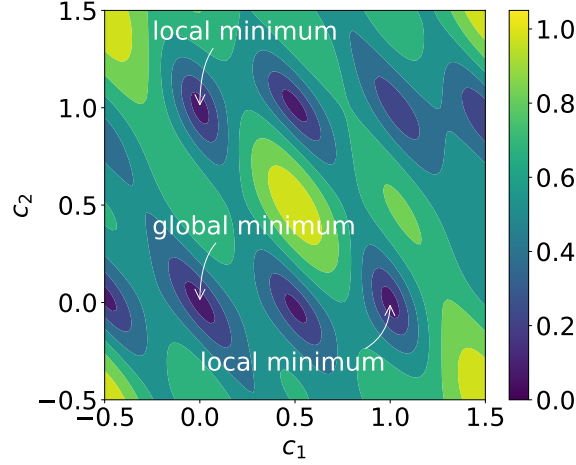


Figure 4: A two-dimensional section of the optimization landscape $F(\tilde{\Phi}^* + c_1 v_1 + c_2 v_2)$. Here Φ^* is a global minimum of the target polynomial $f(x) = \frac{1}{\alpha}(T_4(x) + 2T_2(x) + T_0(x))$ with $\alpha = 440$, and $v_1 = \tilde{\Phi}^1 - \tilde{\Phi}^*$, $v_2 = \tilde{\Phi}^2 - \tilde{\Phi}^*$, where $\tilde{\Phi}^1$ and $\tilde{\Phi}^2$ are two local minima found by numerical schemes. The objective values of the global minimum and local minima are annotated on the figure.

7.3 Numerical behavior near the maximal solution

Here we demonstrate that among all global minima, the maximal solution is particularly desirable from the perspective of numerical optimization. Given an arbitrarily chosen polynomial $f(x) = \frac{1}{4}T_6(x) + \frac{5}{4}T_4(x) + \frac{1}{8}T_2(x) - T_0(x)$, we consider a sequence of scaled polynomials $f^{(k)}(x) := 10^{-k}f(x)$, so that $\lim_{k \rightarrow \infty} f^{(k)}(x) = 0$. All exact solutions to the optimization problem associated to $f^{(k)}(x)$ can be constructed explicitly via Remark 17. This procedure yields several sequences of phase factors, which converge to different limiting points as $k \rightarrow \infty$. For the specific degree-6 polynomial above, there are four distinct limits which is referred to as Φ_0^{class} ,

- class 1: $\Phi_0^1 = (\frac{\pi}{4}, 0, 0, 0, 0, \frac{\pi}{4})$,
- class 2: $\Phi_0^2 = (\frac{\pi}{4}, 0, \frac{\pi}{4}, -\frac{\pi}{2}, \frac{\pi}{4}, 0, \frac{\pi}{4})$,
- class 3: $\Phi_0^3 = (\frac{\pi}{4}, 0, -\frac{\pi}{4}, \frac{\pi}{2}, -\frac{\pi}{4}, 0, \frac{\pi}{4})$,
- class 4: $\Phi_0^4 = (\frac{\pi}{4}, \frac{\pi}{4}, 0, -\frac{\pi}{2}, 0, \frac{\pi}{4}, \frac{\pi}{4})$.

Specifically, the first class is associated with the maximal solution. The sequence of the set of phase factors corresponding to the scaled polynomial $f^{(k)}(x)$ and a given class of construction is referred to as $\Phi_*^{\text{class}}(k)$. Fig. 5 shows that the class associated with the maximal solution distinguishes from other classes, in the sense that the convergence rate of $\Phi_*^1(k)$ towards Φ_0^1 is much faster. This suggests that the convergence basin near Φ_0^1 is much flatter, which justifies the choice of using Φ_0^1 as the initial guess of the optimization.

To further test the performance of numerical optimization associated with different initial guesses, we use the limit point Φ_0^{class} as the initial guess and find the optimal phase factors $\Phi_*^{\text{class}}(k)$ generating the scaled polynomial $\frac{1}{2}f(x)$ by running QSPPACK [5]. The trajectory of the optimization process is displayed in Fig. 6. It shows that the optimization starting from Φ_0^1 also has the fastest convergence rate.

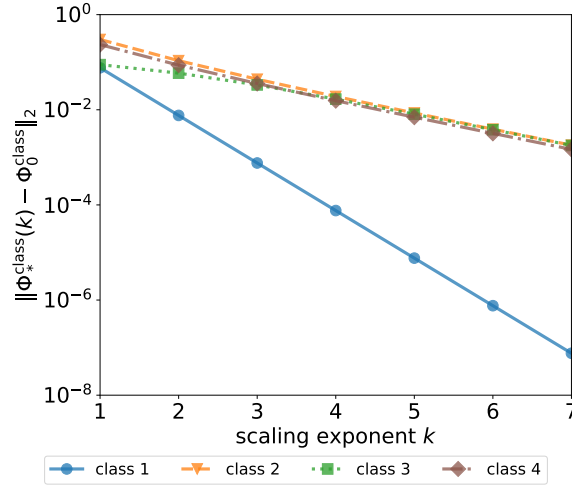


Figure 5: The convergence to the limit point of different class as the target polynomial being scaled down. Given a polynomial $f(x)$, the optimal set of phase factors $\Phi_*^{\text{class}}(k)$ parameterizes the scaled polynomial $f^{(k)}(x) := 10^{-k}f(x)$. Here, Φ_0^{class} is the limit point of the corresponding class.

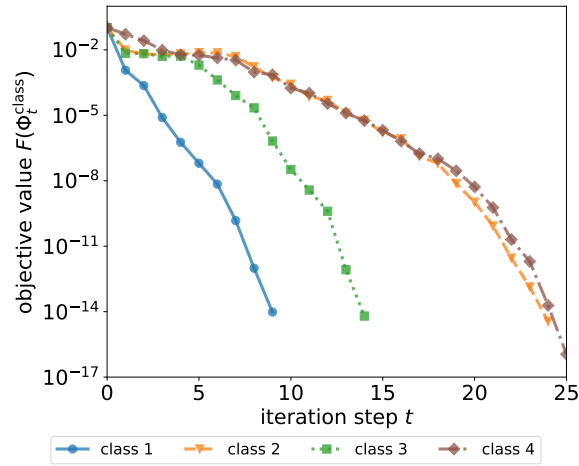


Figure 6: The objective value in each iteration step by running QSPACK for target function $\frac{1}{2}f(x)$. Initiating the optimization from the limit point of a class Φ_0^{class} , Φ_t^{class} is the set of phase factors in the t -th optimization step. Different symbols correspond to the optimization trajectories starting from distinct classes of initial guess.

References

- [1] D. P. Bertsekas. On the Goldstein-Levitin-Polyak gradient projection method. *IEEE Transactions on automatic control*, 21(2):174–184, 1976. doi:10.1109/TAC.1976.1101194.
- [2] S. Bubeck. Convex optimization: Algorithms and complexity. *Foundations and Trends in Machine Learning*, 8(3-4):231–357, 2015. doi:10.1561/22000000050.
- [3] R. Chao, D. Ding, A. Gilyen, C. Huang, and M. Szegedy. Finding angles for quantum signal processing with machine precision, 2020. arXiv:2003.02831.
- [4] A. M. Childs, D. Maslov, Y. Nam, N. J. Ross, and Y. Su. Toward the first quantum simulation with quantum speedup. *Proc. Nat. Acad. Sci.*, 115(38):9456–9461, 2018. doi:10.1073/pnas.1801723115.
- [5] Y. Dong, X. Meng, K. B. Whaley, and L. Lin. Efficient phase factor evaluation in quantum signal processing. *Phys. Rev. A*, 103:042419, 2021. doi:10.1103/PhysRevA.103.042419.
- [6] A. Gilyén, Y. Su, G. H. Low, and N. Wiebe. Quantum singular value transformation and beyond: exponential improvements for quantum matrix arithmetics. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, pages 193–204. ACM, 2019. doi:10.1145/3313276.3316366.
- [7] G. H. Golub and C. F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, third edition, 1996.
- [8] J. Haah. Product decomposition of periodic functions in quantum signal processing. *Quantum*, 3:190, 2019. doi:10.22331/q-2019-10-07-190.
- [9] N. J. Higham. *Accuracy and Stability of Numerical Algorithms*. Society for Industrial and Applied Mathematics, second edition, 2002. doi:10.1137/1.9780898718027.
- [10] J. L. W. V. Jensen. Sur un nouvel et important théorème de la théorie des fonctions. *Acta Mathematica*, 22:359 – 364, 1900. doi:10.1007/BF02417878.
- [11] C. T. Kelley. *Iterative methods for optimization*, volume 18. SIAM, 1999. doi:10.1137/1.9781611970920.
- [12] L. Lin and Y. Tong. Near-optimal ground state preparation. *Quantum*, 4:372, 2020. doi:10.22331/q-2020-12-14-372.
- [13] L. Lin and Y. Tong. Optimal quantum eigenstate filtering with application to solving quantum linear systems. *Quantum*, 4:361, 2020. doi:10.22331/q-2020-11-11-361.
- [14] G. H. Low and I. L. Chuang. Optimal hamiltonian simulation by quantum signal processing. *Physical review letters*, 118(1):010501, 2017. doi:10.1103/PhysRevLett.118.010501.
- [15] K. Mahler. On some inequalities for polynomials in several variables. *Journal of The London Mathematical Society-second Series*, pages 341–344, 1962. doi:10.1112/JLMS/S1-37.1.341.
- [16] J. M. Martyn, Z. M. Rossi, A. K. Tan, and I. L. Chuang. A grand unification of quantum algorithms. *American Physical Society (APS)*, 2(4), 2021. doi:10.1103/PRXQuantum.2.040203.
- [17] M. A. Nielsen and I. Chuang. *Quantum computation and quantum information*. Cambridge Univ. Pr., 2000. doi:10.1017/CB09780511976667.

- [18] J. Nocedal and S. J. Wright. *Numerical optimization*. Springer Verlag, 1999. doi:
[10.1007/b98874](https://doi.org/10.1007/b98874).
- [19] L. Ying. Stable factorization for phase factors of quantum signal processing. *Quantum*,
6:842, 2022. doi:[10.22331/q-2022-10-20-842](https://doi.org/10.22331/q-2022-10-20-842).

A Proof of Lemma 10

Proof. One can check that

$$\begin{aligned} P(x) &= e^{i(\phi_0+\phi_d)} \langle 0 | W(x) \prod_{j=1}^{d-1} [e^{i\phi_j Z} W(x)] | 0 \rangle, \\ i\sqrt{1-x^2}Q(x) &= e^{i(\phi_0-\phi_d)} \langle 0 | W(x) \prod_{j=1}^{d-1} [e^{i\phi_j Z} W(x)] | 1 \rangle. \end{aligned} \quad (96)$$

Then, we expand the product of matrices from left to right by using $e^{i\phi Z} = \cos(\phi) + i\sin(\phi)Z$ and $W(x)Z = ZW(x)^{-1}$. We want to show that the term containing $\sin(\phi_j)$ does not contribute to p_d and q_{d-1} . We use *l.o.* to denote any polynomial or polynomial-valued matrix of degree $< d$. Here, we relax the definition of polynomial and classify the entry of the matrix of our interest as polynomial by considering $\sqrt{1-x^2}$ as another variable. Note that

$$\begin{aligned} &W(x) \prod_{j=1}^{d-1} [e^{i\phi_j Z} W(x)] \\ &= \cos(\phi_1) W(x)^2 \prod_{j=2}^{d-1} [e^{i\phi_j Z} W(x)] + i\sin(\phi_1) W(x)ZW(x) \prod_{j=2}^{d-1} [e^{i\phi_j Z} W(x)] \\ &= \cos(\phi_1) W(x)^2 \prod_{j=2}^{d-1} [e^{i\phi_j Z} W(x)] + i\sin(\phi_1) Z \prod_{j=2}^{d-1} [e^{i\phi_j Z} W(x)] \\ &= \cos(\phi_1) W(x)^2 \prod_{j=2}^{d-1} [e^{i\phi_j Z} W(x)] + l.o., \end{aligned} \quad (97)$$

where the product of the second term is indeed a matrix whose elements are polynomials of degree $\leq d-2$.

Using the fact that

$$[W(x)]^d = \begin{pmatrix} T_d(x) & i\sqrt{1-x^2}U_{d-1}(x) \\ i\sqrt{1-x^2}U_{d-1}(x) & T_d(x) \end{pmatrix}, \quad (98)$$

we may prove inductively,

$$\begin{aligned} P(x) &= e^{i(\phi_0+\phi_d)} \prod_{j=1}^{d-1} \cos(\phi_j) T_d(x) + l.o. \\ i\sqrt{1-x^2}Q(x) &= i\sqrt{1-x^2}e^{i(\phi_0-\phi_d)} \prod_{j=1}^{d-1} \cos(\phi_j) U_{d-1}(x) + l.o., \end{aligned} \quad (99)$$

which proves the lemma. □

B Proof of Lemma 11

Proof. Now that d is odd and Φ is symmetric, Lemma 10 indicates

$$q_{d-1} = \prod_{j=1}^{\tilde{d}-1} \cos^2(\phi_j) \geq 0.$$

If $\phi_j \neq \frac{\pi}{2} + k\pi$ for all $1 \leq j \leq \tilde{d}-1$, where k is some integer, then the leading Chebyshev coefficient of Q is q_{d-1} , which is positive. Otherwise, there exists positive integer j_0 and integer k such that $\phi_{j_0} = \phi_{d-j_0} = \frac{\pi}{2} + k\pi$. And either of the following equalities holds

$$e^{i\phi_{j_0}Z} = e^{i\phi_{d-j_0}Z} = e^{i\frac{\pi}{2}Z}, \quad e^{i\phi_{j_0}Z} = e^{i\phi_{d-j_0}Z} = e^{-i\frac{\pi}{2}Z}.$$

Without losing generality, we only consider the cases that $\phi_{j_0} = \phi_{d-j_0} = \pm\frac{\pi}{2}$.

If $j_0 < \tilde{d}-1$, using the identity

$$W(x)e^{\pm i\frac{\pi}{2}Z}W(x) = e^{\pm i\frac{\pi}{2}Z},$$

the product of matrices

$$e^{i\phi_{j_0-1}Z}W(x)e^{i\phi_{j_0}Z}W(x)e^{i\phi_{j_0+1}Z}$$

can be replaced by

$$e^{i(\phi_{j_0-1}+\phi_{j_0}+\phi_{j_0+1})Z}.$$

We can perform a similar reduction for

$$e^{i\phi_{d-j_0-1}Z}W(x)e^{i\phi_{d-j_0}Z}W(x)e^{i\phi_{d-j_0+1}Z}.$$

Therefore, $U(x, \Phi)$ can be constructed by a new set of symmetric phase factors whose full length is $d-3$ and $\deg(Q) \leq d-5$. Apply Lemma 10 again and one has $q_{d-5} \geq 0$.

If $j_0 = \tilde{d}-1$, the product of matrices

$$e^{i\phi_{d-2}^{\sim}Z}W(x)e^{i\phi_{d-1}^{\sim}Z}W(x)e^{i\phi_d^{\sim}Z}W(x)e^{i\phi_{d+1}^{\sim}Z}$$

can be replaced by

$$-e^{i\phi_{d-2}^{\sim}Z}W(x)e^{i\phi_{d+1}^{\sim}Z}.$$

Therefore, $U(x, \Phi)$ can be constructed by a new set of symmetric phase factors whose full length is $d-1$ and $\deg(Q) \leq d-3$. Apply Lemma 10 again and one has $q_{d-3} \leq 0$.

Repeat the procedure above until the length of the new set of symmetric phase factors reaches $\deg(Q)+2$, and it follows that the leading Chebyshev coefficient of Q has the same sign as $(-1)^{\frac{d-1-\deg(Q)}{2}}$. $\square$

C Proof of Lemma 13

First, we consider the expansion of $P(x)$ and $Q(x)$ in the monomial basis,

$$P(x) = \sum_{i=0}^d \mathbf{p}_i x^i, \quad Q(x) = \sum_{j=0}^{d-1} \mathbf{q}_j x^j.$$

and derive some useful equalities from the normalization condition, which comes from the unitarity of $U(x, \Phi)$.

Lemma 34. *For any $P \in \mathbb{C}[x]$, $Q \in \mathbb{R}[x]$ satisfying the condition (1)-(3) in Theorem 1, one can define $e^{2i\phi_0} := \frac{\mathbf{q}_{d-1}}{\mathbf{p}_d^*}$ and derive two useful equalities*

$$2\text{Re}[\mathbf{p}_{d-2}e^{-2i\phi_0}] + \mathbf{q}_{d-1} - 2\mathbf{q}_{d-3} = 0, \quad \forall d \geq 2, \quad (100)$$

and

$$2\text{Re}[\mathbf{p}_{d-4}e^{-2i\phi_0}] + \mathbf{q}_{d-3} - 2\mathbf{q}_{d-5} = -\frac{\mathbf{q}_{d-1}}{4} - \frac{1}{\mathbf{q}_{d-1}} \left| \text{Im}[\mathbf{p}_{d-2}e^{-2i\phi_0}] \right|^2, \quad \forall d \geq 3. \quad (101)$$

Here we use the convention that $\mathbf{q}_k = \mathbf{p}_k = 0$ if $k < 0$.

Proof. The normalization condition implies that $|\mathbf{p}_d| = |\mathbf{q}_{d-1}|$, so the complex argument ϕ_0 is well defined. Since $Q(x)$ is real, $\mathbf{p}_d e^{-2i\phi_0} = \mathbf{p}_d^* e^{2i\phi_0} = \mathbf{q}_{d-1} \in \mathbb{R}$ and the normalization condition becomes

$$e^{-2i\phi_0} P(x) \cdot e^{2i\phi_0} P^*(x) + (1 - x^2) Q(x)^2 = 1. \quad (102)$$

For $d \geq 2$, the coefficient of x^{2d-2} in Eq. (102) is

$$\begin{aligned} 0 &= \mathbf{p}_d e^{-2i\phi_0} \mathbf{p}_{d-2}^* e^{2i\phi_0} + \mathbf{p}_d^* e^{2i\phi_0} \mathbf{p}_{d-2} e^{-2i\phi_0} + \mathbf{q}_{d-1}^2 - 2\mathbf{q}_{d-1} \mathbf{q}_{d-3} \\ &= \mathbf{q}_{d-1} \left(\mathbf{p}_{d-2}^* e^{2i\phi_0} + \mathbf{p}_{d-2} e^{-2i\phi_0} + \mathbf{q}_{d-1} - 2\mathbf{q}_{d-3} \right) \\ \implies 2\text{Re}[\mathbf{p}_{d-2} e^{-2i\phi_0}] + \mathbf{q}_{d-1} - 2\mathbf{q}_{d-3} &= 0. \end{aligned}$$

For $d \geq 3$, the coefficients of x^{2d-4} in Eq. (102) is

$$\begin{aligned} 0 &= \mathbf{p}_d e^{-2i\phi_0} \mathbf{p}_{d-4}^* e^{2i\phi_0} + \mathbf{p}_d^* e^{2i\phi_0} \mathbf{p}_{d-4} e^{-2i\phi_0} + \mathbf{p}_{d-2}^* e^{2i\phi_0} \mathbf{p}_{d-2} e^{-2i\phi_0} \\ &\quad + 2\mathbf{q}_{d-1} \mathbf{q}_{d-3} - \mathbf{q}_{d-3}^2 - 2\mathbf{q}_{d-1} \mathbf{q}_{d-5} \\ &= \mathbf{q}_{d-1} \left(2\text{Re}[\mathbf{p}_{d-4} e^{-2i\phi_0}] + \mathbf{q}_{d-3} - 2\mathbf{q}_{d-5} \right) - \mathbf{q}_{d-3}^3 + \mathbf{q}_{d-1} \mathbf{q}_{d-3} + \left| \mathbf{p}_{d-2} e^{-2i\phi_0} \right|^2 \\ &= \mathbf{q}_{d-1} \left(2\text{Re}[\mathbf{p}_{d-4} e^{-2i\phi_0}] + \mathbf{q}_{d-3} - 2\mathbf{q}_{d-5} \right) - \mathbf{q}_{d-3}^3 + \mathbf{q}_{d-1} \mathbf{q}_{d-3} \\ &\quad + \left| \text{Re}[\mathbf{p}_{d-2} e^{-2i\phi_0}] \right|^2 + \left| \text{Im}[\mathbf{p}_{d-2} e^{-2i\phi_0}] \right|^2 \\ &= \mathbf{q}_{d-1} \left(2\text{Re}[\mathbf{p}_{d-2} e^{-2i\phi_0}] + \mathbf{q}_{d-3} - 2\mathbf{q}_{d-5} \right) - \mathbf{q}_{d-3}^3 + \mathbf{q}_{d-1} \mathbf{q}_{d-3} \\ &\quad + \left(\frac{2\mathbf{q}_{d-3} - \mathbf{q}_{d-1}}{2} \right)^2 + \left| \text{Im}[\mathbf{p}_{d-2} e^{-2i\phi_0}] \right|^2 \\ &= \mathbf{q}_{d-1} \left(2\text{Re}[\mathbf{p}_{d-2} e^{-2i\phi_0}] + \mathbf{q}_{d-3} - 2\mathbf{q}_{d-5} \right) + \frac{\mathbf{q}_{d-1}^2}{4} + \left| \text{Im}[\mathbf{p}_{d-2} e^{-2i\phi_0}] \right|^2 \\ \implies 2\text{Re}[\mathbf{p}_{d-2} e^{-2i\phi_0}] + \mathbf{q}_{d-3} - 2\mathbf{q}_{d-5} &= -\frac{\mathbf{q}_{d-1}}{4} - \frac{1}{\mathbf{q}_{d-1}} \left| \text{Im}[\mathbf{p}_{d-2} e^{-2i\phi_0}] \right|^2, \end{aligned}$$

where we exploit Eq. (100) in the fourth equality. $\square$

Proof of Lemma 13. Direct computation shows that

$$P^{(1)} = 2(1 - x^2)xQ - (1 - x^2)P^* e^{2i\phi_0} + x^2 P e^{-2i\phi_0}, \quad (103)$$

and

$$Q^{(1)} = (2x^2 - 1)Q - xP^* e^{2i\phi_0} - xP e^{-2i\phi_0} = (2x^2 - 1)Q - 2x\text{Re}[P^* e^{2i\phi_0}]. \quad (104)$$

Hence, $Q(x) \in \mathbb{R}[x]$. Normalization condition is preserved due to unitarity. We can also verify that the parity of $P^{(1)}$ is the same as P and the parity of $Q^{(1)}$ is the same as Q .

Now we apply the results of Lemma 34 and first examine the coefficients of $P^{(1)}$. The coefficient of x^{d+2} is

$$\mathbf{p}_{d+2}^{(1)} = -2\mathbf{q}_{d-1} + \mathbf{p}_d^* e^{2i\phi_0} + \mathbf{p}_d e^{-2i\phi_0} = 0.$$

If $d \geq 2$, the coefficient of x^d is

$$\begin{aligned} \mathbf{p}_d^{(1)} &= 2\mathbf{q}_{d-1} - 2\mathbf{q}_{d-3} + \mathbf{p}_{d-2}^* e^{2i\phi_0} + \mathbf{p}_{d-2} e^{-2i\phi_0} - \mathbf{p}_d^* e^{2i\phi_0} \\ &= \left(2\text{Re}[\mathbf{p}_{d-2} e^{-2i\phi_0}] + \mathbf{q}_{d-1} - 2\mathbf{q}_{d-3} \right) + \left(\mathbf{q}_{d-1} - \mathbf{p}_d^* e^{2i\phi_0} \right) = 0. \end{aligned}$$

If $d = 2$, $P^{(1)} = -\mathbf{p}_0^* e^{2i\phi_0} = -\left(\frac{\mathbf{p}_0}{\mathbf{p}_2}\right)^* \mathbf{q}_1$. And if $d \geq 3$, the coefficient of x^{d-2} is

$$\begin{aligned} \mathbf{p}_{d-2}^{(1)} &= 2\mathbf{q}_{d-3} - 2\mathbf{q}_{d-5} - \mathbf{p}_{d-2}^* e^{2i\phi_0} + \mathbf{p}_{d-4}^* e^{2i\phi_0} + \mathbf{p}_{d-4} e^{-2i\phi_0} \\ &= \mathbf{q}_{d-3} - \mathbf{p}_{d-2}^* e^{2i\phi_0} - \frac{\mathbf{q}_{d-1}}{4} - \frac{1}{\mathbf{q}_{d-1}} \left| \text{Im}[\mathbf{p}_{d-2} e^{-2i\phi_0}] \right|^2 \\ &= \frac{\mathbf{q}_{d-1}}{4} - \frac{1}{\mathbf{q}_{d-1}} \left| \text{Im}[\mathbf{p}_{d-2} e^{-2i\phi_0}] \right|^2 + i \text{Im}[\mathbf{p}_{d-2} e^{-2i\phi_0}] \end{aligned}$$

where we use Eq. (100) in the second equality and Eq. (101) in the third equality.

Then we also apply the results of Lemma 34 to examine the coefficients of $Q^{(1)}$. The coefficient of x^{d+1} is

$$\mathbf{q}_{d+1}^{(1)} = 2\mathbf{q}_{d-1} - \mathbf{p}_d e^{-2i\phi_0} - \mathbf{p}_d^* e^{2i\phi_0} = 0.$$

If $d \geq 2$, the coefficient of x^{d-1} is

$$\mathbf{q}_{d-1}^{(1)} = 2\mathbf{q}_{d-3} - \mathbf{q}_{d-1} - \mathbf{p}_{d-2} e^{-2i\phi_0} - \mathbf{p}_{d-2}^* e^{2i\phi_0} = 0.$$

If $d = 2$, $Q^{(1)} = 0$. And if $d \geq 3$, the coefficient of x^{d-3} is

$$\begin{aligned} \mathbf{q}_{d-3}^{(1)} &= 2\mathbf{q}_{d-5} - \mathbf{q}_{d-3} - \mathbf{p}_{d-4}^* e^{2i\phi_0} - \mathbf{p}_{d-4} e^{-2i\phi_0} \\ &= \frac{\mathbf{q}_{d-1}}{4} + \frac{1}{\mathbf{q}_{d-1}} \left| \text{Im}[\mathbf{p}_{d-2} e^{-2i\phi_0}] \right|^2, \end{aligned}$$

where we exploit Eq. (101). Furthermore, we get that $\mathbf{q}_{d-3}^{(1)}$ is positive if d is odd. □