A framework for engineering quantum likelihood functions for expectation estimation

Dax Enshan Koh[*1], Guoming Wang[†1], Peter D. Johnson[‡1], and Yudong Cao[§1]

[1] *Zapata Computing, Inc.*

Abstract

We develop a framework for characterizing and analyzing engineered likelihood functions (ELFs), which play an important role in the task of estimating the expectation values of quantum observables. These ELFs are obtained by choosing tunable parameters in a parametrized quantum circuit that minimize the expected posterior variance of an estimated parameter. We derive analytical expressions for the likelihood functions arising from certain classes of quantum circuits and use these expressions to pick optimal ELF tunable parameters. Finally, we show applications of ELFs in the Bayesian inference framework.

1 Introduction

Likelihood functions play a fundamental role in various quantum algorithms ranging from quantum channel parameter estimation [1], quantum metrology [2], to quantum phase estimation [3–5] and amplitude estimation [6]. The ability to realize likelihood functions that are otherwise infeasible with only classical resources allows one to take advantage of quantum mechanics to significantly accelerate sampling and measurement processes. The problems encountered in many of these settings [1, 2, 6] can be likened to the problem of finding the bias $\frac{1}{2}(1+q)$ of a coin, where $q \in [-1, 1]$ is an unknown parameter. Instead of the likelihood function $p(d|q) = \frac{1}{2}[1 + (-1)^d q]$, $d \in \{0, 1\}$, that arises in attempting to classically estimate q by direct coin flipping, quantum techniques allow for realizing likelihood functions of the form $p(d|q) = \frac{1}{2}[1 + (-1)^d f(q)]$ where $f(q)$ is a function that depends on the specifics of the quantum scheme. This opens up fundamentally new opportunities for accelerating parameter estimation[1].

Here we focus on the task of estimating the amplitude of overlap Π between two quantum states that differ by a unitary transformation U. This task encompasses a broad variety of quantum algorithms and techniques, including the SWAP test [7] and variants for estimating general state overlaps, expectation estimation [8, 9] in the special case where U is also a Hermitian operator, and phase estimation [3, 4, 10, 11] in the special case where $U = \mathrm{e}^{-\mathrm{i}Ht}$ for some Hermitian operator H and both states are identical and an eigenvector[2] of U. Though some of these works may not explicitly refer to the output distribution $p(d|\Pi)$ of the measurement outcome d as a likelihood function, it can be argued that the quantum advantage of these schemes derive from the ability to realize likelihood functions that are beyond classical possibilities.

In this work, we are motivated by likelihood functions as a way of shedding light on an origin of quantum advantage that has perhaps been previously underappreciated. Our specific approach is to investigate the

*dax.koh@zapatacomputing.com
†guoming.wang@zapatacomputing.com
‡peter@zapatacomputing.com
§yudong@zapatacomputing.com

[1] It is also possible to classically generate likelihood functions that have a non-linear dependence on q by combining multiple flips. However, one can demonstrate that the class of functions f possible with quantum schemes are clearly hard to produce classically.

[2] There are also cases where the eigenvector requirement can be dropped for phase estimation but these do not fit the overlap estimation paradigm in a straightforward way.

inferential power of not only the likelihood functions that have commonly arisen in the literature (which we call "Chebyshev likelihood functions" (CLF) in this paper), but also a broad class of tunable likelihood functions that can be generated by quantum mechanical means (namely the "quantum-generated likelihood functions"). Our results show that the quantum-generated likelihood functions allow for more effective information gain than CLFs during inference, leading to further quantum speedup than previous CLF-based schemes. In addition, we develop an alternative scheme for amplitude estimation based on quantum-generated likelihood functions that does not involve ancilla qubits. A similar scheme has also been considered recently [6] using CLFs.

This work complements the results presented in [12], where we introduced efficient schemes based on engineering likelihood functions to estimate the expectation values of quantum observables. These engineered likelihood functions (ELFs) are obtained by choosing tunable parameters in a parameterized quantum circuit that minimize the expected posterior variance of an estimated parameter. The algorithm in [12] utilizes a multi-round Bayesian updating scheme to reduce the expected posterior variance of the parameter of interest. In this paper, we present a framework for characterizing and analyzing ELFs that forms the basis for the approximate optimization methods used in [12]. While [12] considers the more realistic case of noisy ELFs, this paper concentrates on the idealized case where the ELFs are noiseless. It turns out that studying this idealized case already suffices for gaining useful insights on the behavior and properties of these ELFs.

The rest of the paper is organized as follows. In Section 2, we introduce the concept of quantum-generated likelihood functions through the lens of amplitude estimation. We describe two schemes, called the ancilla-based and ancilla-free schemes, and derive analytical expressions for their likelihood functions as cosine polynomials and trigono-multivariate polynomial functions. In Section 3, we define our objective function, namely the expected posterior variance, and derive simplified expressions for it as we make various assumptions. In Section 4, we combine the results of Section 2 and 3 to show how the tunable parameters in the likelihood functions can be optimized to reduce the expected posterior variance. We also present some numerical results based on solving this optimization problem that demonstrate the performance of various ELFs. For a glossary of some of the definitions and notation used in this paper, see Appendix A.

2 Quantum-generated likelihood functions

2.1 Formulation of problem

For positive integers $n \in \mathbb{Z}^+$, let A be an n-qubit unitary operator and P be an n-qubit Hermitian unitary operator, i.e.

$$A^\dagger A = I, \quad P = P^\dagger, \quad P^2 = I. \tag{1}$$

Write $|A\rangle = A\,|0^n\rangle$, where $|0^n\rangle$ is the all-zero computational basis state on n qubits. The computational task that we consider in this paper is the estimation of the (arccosine of the) expectation value

$$\theta = \arccos(\langle A|\, P\, |A\rangle) \in [0, \pi]. \tag{2}$$

For simplicity, we assume that $\theta \notin \{0, \pi\}$, i.e.

$$|\, \langle A|\, P\, |A\rangle\, | \neq 1. \tag{3}$$

Consider the subspace $\mathscr{S} = \mathrm{span}\{|A\rangle\,, P\,|A\rangle\}$. Note that the assumption in Eq. (3) implies that S is a two-dimensional subspace. This allows us to define $\left|A^\perp\right\rangle$ as the unique state in S that is orthogonal to $|A\rangle$, i.e.

$$\left|A^\perp\right\rangle = \frac{P\,|A\rangle - \langle A|\, P\, |A\rangle\, |A\rangle}{\sqrt{1 - \langle A|\, P\, |A\rangle^2}}. \tag{4}$$

By construction, $\mathscr{B} = \{|A\rangle\,, \left|A^\perp\right\rangle\}$ forms an orthonormal basis for $\mathscr{S}$. Henceforth, we shall label $|A\rangle$ and $|A^\perp\rangle$ as $|\bar{0}\rangle = \begin{pmatrix}1\\0\end{pmatrix}$ and $|\bar{1}\rangle = \begin{pmatrix}0\\1\end{pmatrix}$ respectively and refer to $\mathscr{S}$ as the *logical space*, which is to be contrasted with the *physical space* describing the n-qubit system. We shall extend the bar notation to certain operators on $\mathscr{S}$. For example, we will denote the Pauli matrices with respect to the basis $\mathscr{B}$ by

$$\bar{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \bar{X} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \bar{Y} = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \quad \bar{Z} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{5}$$

In the basis $\mathscr{B}$, P may be written as

$$P = \cos(\theta)\bar{Z} + \sin(\theta)\bar{X} = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}. \tag{6}$$

To indicate explicit dependence of P on θ, we will write $P = P(\theta)$.

Next, we define the following two types of unitary operators:

$$\begin{aligned} U(\theta;\alpha) &= \exp\left(-\mathrm{i}\alpha P(\theta)\right) \\ &= \cos(\alpha)\, I - \mathrm{i}\sin(\alpha)\, P(\theta) \\ &= \cos(\alpha)\, \bar{I} - \mathrm{i}\sin(\alpha)\left(\cos(\theta)\, \bar{Z} + \sin(\theta)\, \bar{X}\right), \end{aligned} \tag{7}$$

and

$$\begin{aligned} V(\beta) &= \exp\left(-\mathrm{i}\beta(2|\bar{0}\rangle\langle\bar{0}| - I)\right) \\ &= \exp\left(-\mathrm{i}\beta\bar{Z}\right) \\ &= \cos(\beta)\, \bar{I} - \mathrm{i}\sin(\beta)\, \bar{Z}, \end{aligned} \tag{8}$$

where $\alpha, \beta \in \mathbb{R}$.

In constructing our estimation algorithm, we assume that we are able to perform the following primitive operations. First, we assume that we are able to prepare computational basis states $|0^n\rangle$ and apply A to them to obtain $|\bar{0}\rangle$. Next, we assume that we are allowed to apply the following operations on the physical qubits: $U(\theta;\alpha)$ for any angle $\alpha \in \mathbb{R}$ and $V(\beta)$ for any $\beta \in \mathbb{R}$, as well as controlled versions of these operations, namely c-$U(\theta;\alpha) = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes U(\theta;\alpha)$ and c-$V(\beta) = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes V(\beta)$. Furthermore, we assume that we can apply the single-qubit Hadamard gate $H = \frac{1}{\sqrt{2}}\left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right)$ to the physical qubits. Finally, we assume that we are allowed to perform computational basis measurements as well as measurements of the operator P (here, a measurement of P corresponds to the projection-valued measure $\{\frac{I+P}{2}, \frac{I-P}{2}\}$).

2.2 Schemes for expectation estimation

We now describe two different schemes for expectation estimation that we will focus on in this paper. Each of these schemes, which will use only the primitive operations that were listed in Section 2.1, depends on some tunable parameters, which we collect in a single vector $\vec{x} = x_1 x_2 \ldots x_{2L} \in \mathbb{R}^{2L}$. Here, the variable $L \in \mathbb{Z}^+$ represents half the number of tunable parameters in each scheme. The vector $\vec{x}$ is tunable in the sense that we will later (specifically, in Section 4) tune it to optimize some objective function. For this section, however, it would suffice to treat $\vec{x}$ as fixed.

The first scheme, called the *ancilla-free* (AF) scheme because it does not make use of any ancilla qubits, proceeds as follows:

Algorithm 1 Ancilla-free (AF) scheme

1: Start with the n-qubit state $|\bar{0}\rangle$
2: **for** $i = 1, \ldots, L$ **do**
3: Apply the operator $U(\theta; x_{2i-1})$
4: Apply the operator $V(x_{2i})$
5: Perform a measurement corresponding to the Hermitian operator P, i.e. perform the projective measurement $(\frac{I+P}{2}, \frac{I-P}{2})$ with outcomes $(0, 1)$

The second scheme, called the *ancilla-based* (AB) scheme because it involves the use of a single ancilla register, proceeds as follows:

Algorithm 2 Ancilla-based (AB) scheme

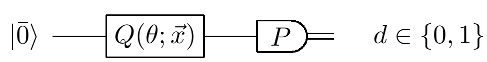
Start with the input $|\bar{0}\rangle$ in the n-qubit data register and $|0\rangle$ in the single-qubit ancilla register
Apply the Hadamard gate H to the ancilla register
for $i = 1, \ldots, L$ **do**
    Apply the operator c-$U(\theta; x_{2i-1}) = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes U(\theta; x_{2i-1})$, with the control being the ancilla register and the target being the data register
    Apply the operator c-$V(x_{2i}) = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes V(x_{2i})$, with the control being the ancilla register and the target being the data register
Apply the Hadamard gate H to the ancilla register
Perform a computational-basis measurement on the ancilla register to obtain outcomes $(0, 1)$

Circuit diagrams illustrating these schemes are given in Figure 2.1, where we have introduced the operator

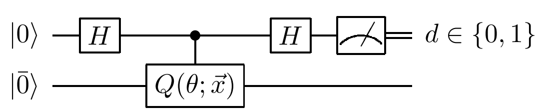
$$Q(\theta; \vec{x}) = V(x_{2L})U(\theta; x_{2L-1})V(x_{2L-2})U(\theta; x_{2L-3}) \ldots V(x_4)U(\theta; x_3)V(x_2)U(\theta; x_1), \tag{9}$$

to describe the overall unitary operator resulting from $2L-1$ alternations of the rotations $U(\theta; \cdot)$ and $V(\cdot)$. For each of these schemes, we denote the single-bit measurement outcome by $d \in \{0, 1\}$.

Ancilla-free circuit:



Ancilla-based circuit:



where

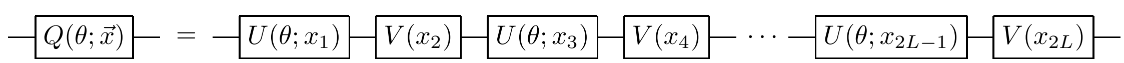


Figure 2.1: Quantum circuits for both the ancilla-free and ancilla-based schemes. The output of each of these circuits is a single bit, denoted by $d \in \{0, 1\}$.

Each of these schemes is associated with a likelihood function, which will play a central role in this paper. Specifically, the likelihood function associated with the scheme $\mathcal{A}$, where $\mathcal{A} = \text{AF, AB}$ (which stand for ancilla-free and ancilla-based respectively), and tunable parameters $\vec{x}$ is defined to be the likelihood of the random variable θ (which encodes information about the expectation value $\langle A| P |A\rangle$) given the measurement outcome d. In other words, this likelihood function, denoted $\mathcal{L}^{\mathcal{A}}(\theta; d, \vec{x}) = \Pr(d|\theta; \vec{x})$, is the probability of obtaining the outcome d in scheme $\mathcal{A}$ given the unknown parameter θ and tunable parameters $\vec{x}$. Since d takes only the values 0 and 1, the likelihood function can be written as

$$\mathcal{L}^{\mathcal{A}}(\theta; d, \vec{x}) = \frac{1}{2}\left[1 + (-1)^d \Lambda^{\mathcal{A}}(\theta; \vec{x})\right] \tag{10}$$

for some function $\Lambda^{\mathcal{A}}(\theta; \vec{x})$. We shall call the function $\Lambda^{\mathcal{A}}$ the *bias* associated with scheme $\mathcal{A}$.

In the next proposition, we derive expressions for the biases $\Lambda^{\mathcal{A}}(\theta; \vec{x})$, which will be useful for the rest of this paper. For this purpose, we first extend the definition in Eq. (9) slightly so that the second argument of Q can be a vector of arbitrary nonzero length. Let $\alpha \in \mathbb{Z}^+$ be a positive integer. For $\theta \in \mathbb{R}$ and an arbitrary vector $\vec{z} = (z_1, \ldots, z_\alpha) \in \mathbb{R}^\alpha$, define

$$Q(\theta; \vec{z}) := R(z_\alpha) \ldots V(z_4)U(\theta; z_3)V(z_2)U(\theta; z_1), \tag{11}$$

where

$$R(\,\cdot\,) = \begin{cases} U(\theta;\,\cdot\,), & \alpha \text{ odd}, \\ V(\,\cdot\,), & \alpha \text{ even}. \end{cases} \tag{12}$$

Next, define

$$Q_{00}[\vec{z}](\theta) := \langle\bar{0}|\, Q(\theta;\vec{z})\,|\bar{0}\rangle \tag{13}$$

to be the expectation value of $Q(\theta;\vec{z})$ in the state $|\bar{0}\rangle$. These allow us to state the following proposition.

Proposition 1. *Let $\theta \in \mathbb{R}$ and $\vec{x} \in \mathbb{R}^{2L}$. The ancilla-free and ancilla-based biases are given by*

$$\Lambda^{\mathrm{AF}}(\theta;\vec{x}) = \mathrm{i}Q_{00}\left[\vec{x}, \tfrac{\pi}{2}, -\vec{x}^{R}\right](\theta) \tag{14}$$

$$\Lambda^{\mathrm{AB}}(\theta;\vec{x}) = \operatorname{Re} Q_{00}[\vec{x}](\theta), \tag{15}$$

where $\vec{x}^R$ denotes the reverse of $\vec{x}$.

Proof. In the ancilla-free case, the state just before the measurement is $Q(\theta;\vec{x})\,|\bar{0}\rangle$, whose density operator can be written as

$$\varrho = Q(\theta;\vec{x})|\bar{0}\rangle\langle\bar{0}|Q(\theta;\vec{x})^\dagger = Q(\theta;\vec{x})\left(\frac{\bar{I}+\bar{Z}}{2}\right)Q(\theta;\vec{x})^\dagger = \frac{1}{2}(1 + Q(\theta;\vec{x})\bar{Z}Q(\theta;\vec{x})^\dagger). \tag{16}$$

Hence, the probability of obtaining the outcome $d \in \{0,1\}$ when ϱ is measured is

$$\begin{aligned}\mathcal{L}^{\mathrm{AF}}(\theta;d,\vec{x}) &= \mathrm{tr}\left[\frac{1}{2}\left(1 + (-1)^d P(\theta)\right)\varrho\right] \\ &= \frac{1}{2}\left[1 + (-1)^d \mathrm{tr}(P(\theta)\varrho)\right].\end{aligned} \tag{17}$$

Comparing this with Eq. (10), we get the following expression for the ancilla-free bias

$$\begin{aligned}\Lambda^{\mathrm{AF}}(\theta;\vec{x}) &= \mathrm{tr}(P(\theta)\varrho) \\ &= \mathrm{tr}\left[P(\theta)\frac{1}{2}\left(I + Q(\theta;\vec{x})\bar{Z}Q(\theta;\vec{x})^\dagger\right)\right] \\ &= \frac{1}{2}\mathrm{tr}(Q(\theta;\vec{x})^\dagger P(\theta) Q(\theta;\vec{x})\bar{Z}) \\ &= \frac{1}{2}\mathrm{tr}\left[Q(\theta;\vec{x})^\dagger P(\theta)Q(\theta;\vec{x})(2|\bar{0}\rangle\langle\bar{0}| - I)\right] \\ &= \langle\bar{0}|\,Q(\theta;\vec{x})^\dagger P(\theta) Q(\theta;\vec{x})\,|\bar{0}\rangle \\ &= \langle\bar{0}|\,U(\theta,-x_1)V(-x_2)\dots U(\theta,-x_{2L-1})V(-x_{2L})\ \mathrm{i}U(\theta,\pi/2) \\ &\qquad \times V(x_{2L})U(\theta;x_{2L-1})\dots V(x_2)U(\theta;x_1)\,|\bar{0}\rangle \\ &= \mathrm{i}Q_{00}\left[x_1, x_2, \dots, x_{2L-1}, x_{2L}, \tfrac{\pi}{2}, -x_{2L}, -x_{2L-1}, \dots, -x_2, -x_1\right](\theta) \\ &= \mathrm{i}Q_{00}\left[\vec{x}, \tfrac{\pi}{2}, -\vec{x}^R\right](\theta),\end{aligned} \tag{18}$$

where the third line follows from the fact that $P(\theta)$ is traceless, the fifth line follows from

$$\mathrm{tr}(Q(\theta;\vec{x})^\dagger P(\theta)Q(\theta;\vec{x})) = \mathrm{tr}(Q(\theta;\vec{x})Q(\theta;\vec{x})^\dagger P(\theta)) = \mathrm{tr}(P(\theta)) = 0 \tag{19}$$

and the sixth line follows from the fact that $P(\theta) = \mathrm{i}U(\theta,\pi/2)$.

In the ancilla-based case, the state just before measurement is

$$\begin{aligned}\varsigma &= (H\otimes I)\text{c-}Q(\theta;\vec{x})(H\otimes I)\,|0\rangle\,|\bar{0}\rangle \\ &= \frac{1}{2}\,|0\rangle\,(I + Q(\theta;\vec{x}))\,|\bar{0}\rangle + \frac{1}{2}\,|1\rangle\,(I - Q(\theta;\vec{x}))\,|\bar{0}\rangle\,,\end{aligned} \tag{20}$$

where the second line follows from a straightforward calculation. Hence, the probability of obtaining the outcome $d \in \{0,1\}$ when ς is measured is

$$\begin{aligned}\mathcal{L}^{\mathrm{AB}}(\theta; d, \vec{x}) &= \left\| \frac{1}{2} \left(I + (-1)^d Q(\theta; \vec{x}) \right) |\bar{0}\rangle \right\|^2 \\ &= \frac{1}{2} \left[1 + (-1)^d \operatorname{Re} \langle \bar{0} | Q(\theta; \vec{x}) | \bar{0} \rangle \right]. \end{aligned} \tag{21}$$

Comparing this with Eq. (10), we obtain

$$\Lambda^{\mathrm{AB}}(\theta; \vec{x}) = \operatorname{Re} \langle \bar{0} | Q(\theta; \vec{x}) | \bar{0} \rangle = \operatorname{Re} Q_{00}[\vec{x}](\theta). \tag{22}$$

□

2.3 Series expansions of the ancilla-free and ancilla-based biases

In this subsection, we will derive series expansion formulas[3] for both the ancilla-free and ancilla-based biases, which will be useful in Section 4. In order to state and prove the theorems, we first present some necessary preliminary definitions and lemmas in Sections 2.3.1 and 2.3.2.

2.3.1 Mathematical preliminaries I: Expansion formulas

Let $n, k \in \mathbb{N}$ and $u, v \in \mathbb{F}_2$. The central object that we introduce here is the set Θ^n_{ukv}, which is (informally) defined as follows:

Θ^n_{ukv} = set of strings $\vec{x} = x_1 x_2 \dots x_n \in \mathbb{F}_2^n$ for which the string $r^{x_n} \dots p^{x_4} q^{x_3} p^{x_2} q^{x_1}$ can be converted to $p^u (qp)^k q^v$ by repeatedly applying the replacement rules $pp \to \varepsilon$ and $qq \to \varepsilon$, where $r = p$ if n is even and $r = q$ if n is odd. Here, $p^x = \varepsilon$ if $x = 0$ and $p^x = p$ if $x = 1$.

For example, the string $101101011 \in \Theta^9_{010}$ since

$$q^1 p^1 q^0 p^1 q^0 p^1 q^1 p^0 q^1 = qpppqq \to q\cancel{pp}p\cancel{qq} = qp = p^0 (qp)^1 q^0. \tag{23}$$

For a formal definition of Θ^n_{ukv}, see Appendix B. By convention, we take $\Theta^n_{ukv} = \emptyset$ whenever $k \notin \mathbb{N}$. The following theorem, which we prove in Appendix B, characterizes the set of k values for which Θ^n_{ukv} is nonempty.

Lemma 2. *(also Theorem 17) Let $n \in \mathbb{N}$, $u, v \in \mathbb{F}_2$ and $k \in \mathbb{N}$. Then,*

$$\Theta^n_{ukv} \neq \emptyset \iff k \leq \left\lfloor \frac{n-1}{2} \right\rfloor - u \mathbb{1}_{n \in 2\mathbb{Z}+1}. \tag{24}$$

We now introduce our key lemma, which we prove in Appendix B:

Lemma 3. *(also Lemma 19) Let $n \in \mathbb{Z}^+$ and $P^2 = Q^2 = I$. Let $\{a^y_x : x \in \mathbb{F}_2,\ y \in [n]\} \subset \mathbb{C}$. Then,*

$$\begin{aligned} &(a^n_0 + a^n_1 R) \dots \left(a^4_0 + a^4_1 P\right) \left(a^3_0 + a^3_1 Q\right) \left(a^2_0 + a^2_1 P\right) \left(a^1_0 + a^1_1 Q\right) \\ &= \sum_{k=0}^{\infty} \sum_{u,v \in \mathbb{F}_2} \left(\sum_{\vec{x} \in \Theta^n_{ukv}} a^1_{x_1} a^2_{x_2} \dots a^n_{x_n} \right) P^u (QP)^k Q^v, \end{aligned} \tag{25}$$

where $R = P$ if n is even and $R = Q$ if n is odd.

[3] For the series expansion formulas, see Theorems 10 and 11, which this subsection will culminate in.

By virtue of Lemma 2, the sum (25) has only a finite number of nonzero terms. Keeping only the nonzero terms in the expansion gives

$$\begin{aligned}&\left(a_0^n + a_1^n R\right)\dots\left(a_0^4 + a_1^4 P\right)\left(a_0^3 + a_1^3 Q\right)\left(a_0^2 + a_1^2 P\right)\left(a_0^1 + a_1^1 Q\right)\\ &= \sum_{u,v\in\mathbb{F}_2} \sum_{k=0}^{\left\lfloor \frac{n-1}{2}\right\rfloor - u\mathbb{1}_{n\in 2\mathbb{Z}+1}} \left(\sum_{\vec{x}\in\Theta_{ukv}^n} a_{x_1}^1 a_{x_2}^2 \dots a_{x_n}^n\right) P^u (QP)^k Q^v. \end{aligned} \tag{26}$$

The next important object that we introduce is Ξ, which is defined as follows: for $\alpha \in \mathbb{Z}^+$ and $l \in \mathbb{N}$, let

$$\begin{aligned}\Xi_l^\alpha &= \bigcup_{u,v\in\mathbb{F}_2} \Theta_{u,l-v,v}^\alpha \\ &= \Theta_{0l0}^\alpha \cup \Theta_{1l0}^\alpha \cup \Theta_{0,l-1,1}^\alpha \cup \Theta_{1,l-1,1}^\alpha. \end{aligned} \tag{27}$$

It is straightforward to check that

$$\Theta_{ukv}^\alpha \subseteq \Xi_{k+v}^\alpha. \tag{28}$$

Next, we use Lemma 2 to characterize the set of k values for which Ξ_l^n is nonempty.

Lemma 4. *Let $\alpha \in \mathbb{Z}^+$ and $l \in \mathbb{N}$. Then,*

$$\Xi_l^\alpha \neq \emptyset \iff l \leq \lceil \alpha/2 \rceil. \tag{29}$$

Proof. We first prove the forward direction. Assume that $\Xi_l^\alpha \neq \emptyset$. Then at least one of the following holds: (i) $\Theta_{0l0}^\alpha \neq \emptyset$, (ii) $\Theta_{1l0}^\alpha \neq \emptyset$, (iii) $\Theta_{0,l-1,1}^\alpha \neq \emptyset$, (iv) $\Theta_{1,l-1,1}^\alpha \neq \emptyset$. Making use of Lemma 2, we find that

1. If $\Theta_{0l0}^\alpha \neq \emptyset$, then $l \leq \left\lfloor \frac{\alpha-1}{2} \right\rfloor \leq \left\lfloor \frac{\alpha+1}{2} \right\rfloor = \lceil \alpha/2 \rceil$.
2. If $\Theta_{1l0}^\alpha \neq \emptyset$, then $l \leq \left\lfloor \frac{\alpha-1}{2} \right\rfloor - \mathbb{1}_{n\in 2\mathbb{Z}+1} \leq \left\lfloor \frac{\alpha-1}{2} \right\rfloor \leq \left\lfloor \frac{\alpha+1}{2} \right\rfloor = \lceil \alpha/2 \rceil$.
3. If $\Theta_{0,l-1,1}^\alpha \neq \emptyset$, then $l-1 \leq \left\lfloor \frac{\alpha-1}{2} \right\rfloor \implies l \leq 1 + \left\lfloor \frac{\alpha-1}{2} \right\rfloor = \left\lfloor \frac{\alpha+1}{2} \right\rfloor = \lceil \alpha/2 \rceil$.
4. If $\Theta_{1,l-1,1}^\alpha \neq \emptyset$, then $l-1 \leq \left\lfloor \frac{\alpha-1}{2} \right\rfloor - \mathbb{1}_{n\in 2\mathbb{Z}+1} \leq \left\lfloor \frac{\alpha-1}{2} \right\rfloor \implies l \leq 1 + \left\lfloor \frac{\alpha-1}{2} \right\rfloor = \left\lfloor \frac{\alpha+1}{2} \right\rfloor = \lceil \alpha/2 \rceil$.

In all these cases, $l \leq \lceil \alpha/2 \rceil$.

Next, we prove the reverse direction. Assume that $l \leq \lceil \alpha/2 \rceil = \left\lfloor \frac{\alpha+1}{2} \right\rfloor = 1 + \left\lfloor \frac{\alpha-1}{2} \right\rfloor$. which implies that $l - 1 \leq \left\lfloor \frac{\alpha-1}{2} \right\rfloor$. By Lemma 2, $\Theta_{0,l-1,1}^\alpha \neq \emptyset$, which implies that $\Xi_l^\alpha \neq \emptyset$. □

In Appendix C, we strengthen both Lemmas 2 and 4 by finding the cardinalities of the sets Ξ_l^α and Θ_{ukv}^n. These cardinalities will be useful for bounding the space complexity of computing various sums related to Eq. (25).

2.3.2 Mathematical preliminaries II: Trigono-multilinear and trigono-multiquadratic functions

For $k \in \mathbb{Z}^+$, let $\vec{x} = (x_1, \dots, x_k) \in \mathbb{R}^k$ and $\vec{y} = y_1 \dots y_k \in \{0,1\}^k$. Define

$$\zeta_{\vec{y}}(\vec{x}) := \prod_{a:y_a=0} \cos\left(x_a\right) \prod_{b:y_b=1} \sin\left(x_b\right). \tag{30}$$

For example,

$$\zeta_{00101}(x_1, x_2, x_3, x_4, x_5) = \cos(x_1)\cos(x_2)\sin(x_3)\cos(x_4)\sin(x_5).$$

When $k = 1$, each $\zeta_y(\cdot)$ is a trigonometric function: $\zeta_0(x) = \cos(x)$ and $\zeta_1(x) = \sin(x)$, i.e.

$$\zeta_y(x) = (\sin x)^y (\cos x)^{1-y}. \tag{31}$$

A nice property that $\zeta_{\vec{y}}(\vec{x})$ satisfies is the following multiplicative property:

$$\zeta_{y_1 y_2 \ldots y_n}(x_1, x_2, \ldots, x_k) = \zeta_{y_1}(x_1)\zeta_{y_2}(x_2)\ldots\zeta_{y_k}(x_k). \tag{32}$$

The functions $\zeta_{\vec{y}}(\vec{x})$ can be used to define the notions of trigono-multilinearity and trigono-multiquadraticity.

Definition 5. Let $k \in \mathbb{Z}^+$. A k-ary function $f : \mathbb{R}^k \to \mathbb{C}$ is *trigono-multilinear* if for all $\vec{y} \in \{0,1\}^k$, there exists $\xi_{\vec{y}} \in \mathbb{C}$ such that for all $\vec{x} \in \mathbb{R}^k$,

$$f(\vec{x}) = \sum_{\vec{y} \in \{0,1\}^k} \xi_{\vec{y}} \zeta_{\vec{y}}(\vec{x}). \tag{33}$$

Definition 6. Let $k \in \mathbb{Z}^+$. A k-ary function $f : \mathbb{R}^k \to \mathbb{C}$ is *trigono-multiquadratic* if for all $\vec{y}, \vec{z} \in \{0,1\}^k$, there exists $\xi_{\vec{y}\vec{z}} \in \mathbb{C}$ such that for all $\vec{x} \in \mathbb{R}^k$,

$$f(\vec{x}) = \sum_{\vec{y}, \vec{z} \in \{0,1\}^k} \xi_{\vec{y}\vec{z}} \zeta_{\vec{y}\vec{z}}(\vec{x}, \vec{x}), \tag{34}$$

where $\vec{y}\vec{z} = y_1 \ldots y_k z_1 \ldots z_k \in \{0,1\}^{2k}$ is the string formed from concatenating $\vec{y}$ and $\vec{z}$.

Equivalently, the trigono-multilinear functions are those that can be written as

$$f(x_1, \ldots, x_k) = \sum_{y_1 \ldots y_k \in \{0,1\}^k} \xi_{y_1 \ldots y_k} \prod_{a : y_a = 0} \cos(x_a) \prod_{b : y_b = 1} \sin(x_b) \tag{35}$$

and the trigono-multiquadratic functions are those that can be written as

$$f(x_1, \ldots, x_k) = \sum_{y_1 \ldots y_k \in \{0,1,2\}^k} \eta_{y_1 \ldots y_k} \prod_{a : y_a = 0} \cos^2(x_a) \prod_{b : y_b = 1} \sin^2(x_b) \prod_{c : y_c = 2} \cos(x_c) \sin(x_c), \tag{36}$$

where $\xi_{y_1 \ldots y_k}, \eta_{y_1 \ldots y_k} \in \mathbb{C}$.

When $k = 1$, the above expressions simplify as follows. A unary trigono-multilinear function is of the form

$$f(x) = \xi_0 \cos(x) + \xi_1 \sin(x) \tag{37}$$

and a unary trigono-multiquadratic function is of the form

$$f(x) = \eta_0 \cos^2(x) + \eta_1 \sin^2(x) + \eta_2 \sin(x)\cos(x), \tag{38}$$

where $\xi_0, \xi_1, \eta_0, \eta_1, \eta_2 \in \mathbb{C}$.

For a larger example, consider $k = 4$. A 4-ary trigono-multilinear function takes the form

$$\begin{aligned} f(x_1, x_2, x_3, x_4) = {} & \xi_{0000} \cos(x_1)\cos(x_2)\cos(x_3)\cos(x_4) \\ & + \xi_{0001} \cos(x_1)\cos(x_2)\cos(x_3)\sin(x_4) \\ & + \ldots \\ & + \xi_{1110} \sin(x_1)\sin(x_2)\sin(x_3)\cos(x_4) \\ & + \xi_{1111} \sin(x_1)\sin(x_2)\sin(x_3)\sin(x_4) \end{aligned} \tag{39}$$

and a 4-ary trigono-multiquadratic function takes the form

$$\begin{aligned} f(x_1,x_2,x_3,x_4) = {} & \eta_{0000}\cos^2(x_1)\cos^2(x_2)\cos^2(x_3)\cos^2(x_4) \\ & + \eta_{0001}\cos^2(x_1)\cos^2(x_2)\cos^2(x_3)\sin^2(x_4) \\ & + \eta_{0002}\cos^2(x_1)\cos^2(x_2)\cos^2(x_3)\cos(x_4)\sin(x_4) \\ & + \dots \\ & + \eta_{2220}\cos(x_1)\sin(x_1)\cos(x_2)\sin(x_2)\cos(x_3)\sin(x_3)\cos^2(x_4) \\ & + \eta_{2221}\cos(x_1)\sin(x_1)\cos(x_2)\sin(x_2)\cos(x_3)\sin(x_3)\sin^2(x_4) \\ & + \eta_{2222}\cos(x_1)\sin(x_1)\cos(x_2)\sin(x_2)\cos(x_3)\sin(x_3)\cos(x_4)\sin(x_4)\,. \end{aligned} \tag{40}$$

Trigono-multilinear and trigono-multiquadratic functions have various nice properties that are useful and of independent interest. We explore some of these properties in Appendix D.

2.3.3 Applying the expansion formulas to the quantum-generated biases

We will now apply the results of Sections 2.3.1 and 2.3.2 to the quantum-generated biases $\Lambda^{\mathcal{A}}(\theta;\vec{x})$. Specifically, we will use the expansion formula of Eq. (25) to expand the functions $Q(\theta,\vec{z})$, $Q_{00}[\vec{z}](\theta)$ and the biases $\Lambda^{\mathcal{A}}(\theta;\vec{x})$ and show that $\Lambda^{\mathrm{AF}}(\theta;\vec{x})$ and $\Lambda^{\mathrm{AB}}(\theta;\vec{x})$ are trigono-multiquadratic and trigono-multilinear functions (of θ) respectively. For an example, we refer the reader to Appendix E, where we work out explicit expressions for the expansion formulas in the case when $L=1$.

First, the expansion formula when applied to $Q(\theta;\vec{z})$ gives the following expression.

Theorem 7. *Let $\vec{z}\in\mathbb{R}^{\alpha}$. Then, written in the basis $\{|\bar{0}\rangle, |\bar{1}\rangle\}$,*

$$Q(\theta;\vec{z}) = \sum_{k=0}^{\infty}\sum_{u,v\in\mathbb{F}_2}\sum_{\vec{y}\in\Theta^{\alpha}_{ukv}}(-\mathrm{i})^{\mathrm{wt}(\vec{y})}\zeta_{\vec{y}}(\vec{z})\begin{pmatrix}\cos[(k+v)\theta] & -(-1)^v\sin[(k+v)\theta] \\ (-1)^u\sin[(k+v)\theta] & (-1)^{u+v}\cos[(k+v)\theta]\end{pmatrix}. \tag{41}$$

Proof. Substituting Eqs. (7) and (8) into Eq. (11) allows us to express $Q(\theta;\vec{z})$ in the following form:

$$\begin{aligned} Q(\theta;\vec{z}) &= (\cos z_\alpha - \mathrm{i}\sin z_\alpha W)\dots(\cos z_4 - \mathrm{i}\sin z_4\bar{Z})(\cos z_3 - \mathrm{i}\sin z_3 P) \\ &\quad\times(\cos z_2 - \mathrm{i}\sin z_2\bar{Z})(\cos z_1 - \mathrm{i}\sin z_1 P) \\ &= (a_0^\alpha + a_1^\alpha W)\dots\left(a_0^4 + a_1^4\bar{Z}\right)\left(a_0^3 + a_1^3 P\right)\left(a_0^2 + a_1^2\bar{Z}\right)\left(a_0^1 + a_1^1 P\right) \end{aligned} \tag{42}$$

where P denotes $P(\theta)$,

$$W = \begin{cases} P(\theta), & \alpha \text{ odd} \\ \bar{Z}, & \alpha \text{ even}\end{cases} \tag{43}$$

and for $j\in\{0,1\}$ and $k\in\{1,2,\dots,\alpha\}$,

$$\begin{aligned} a_j^k &= \begin{cases}\cos z_k, & j=0 \\ -\mathrm{i}\sin z_k, & j=1\end{cases} \\ &= (-\mathrm{i}\sin z_k)^j(\cos z_k)^{1-j} \\ &= (-\mathrm{i})^j\zeta_j(z_k), \end{aligned} \tag{44}$$

where $\zeta_j(z_k)$ was defined in Eq. (30). Since $\bar{Z}^2 = P^2 = I$, we can expand Eq. (42) according to the Lemma 3.

This gives

$$Q(\theta;\vec{z}) = \sum_{k=0}^{\infty}\sum_{u,v\in\mathbb{F}_2}\Bigg(\sum_{\vec{y}\in\Theta^{n}_{ukv}}\underbrace{a^1_{y_1}a^2_{y_2}\dots a^\alpha_{y_\alpha}}_{①}\Bigg)\underbrace{\bar{Z}^u(P\bar{Z})^k P^v}_{②}. \tag{45}$$

Now,

$$\textcircled{1} = \prod_{j=1}^{\alpha} a_{y_j}^k = \prod_{j=1}^{\alpha}(-\mathrm{i})^{y_j}\zeta_{y_j}(z_k) = (-\mathrm{i})^{\mathrm{wt}(\vec{y})}\prod_{j=1}^{\alpha}\zeta_{y_j}(z_k) = (-\mathrm{i})^{\mathrm{wt}(\vec{y})}\zeta_{\vec{y}}(\vec{z}). \tag{46}$$

And,

$$\begin{aligned}
\textcircled{2} &= \bar{Z}^u(P\bar{Z})^k P^v \\
&= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}^u \begin{pmatrix} \cos k\theta & -\sin k\theta \\ \sin k\theta & \cos k\theta \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}^v \\
&= \begin{pmatrix} 1 & 0 \\ 0 & (-1)^u \end{pmatrix} \begin{pmatrix} \cos k\theta & -\sin k\theta \\ \sin k\theta & \cos k\theta \end{pmatrix} \begin{pmatrix} \cos v\theta & \sin v\theta \\ \sin v\theta & (-1)^v\cos v\theta \end{pmatrix} \\
&= \begin{pmatrix} \cos[(k+v)\theta] & -(-1)^v\sin[(k+v)\theta] \\ (-1)^u\sin[(k+v)\theta] & (-1)^{u+v}\cos[(k+v)\theta] \end{pmatrix}.
\end{aligned} \tag{47}$$

Substituting Eqs. (46) and (47) into Eq. (45) gives Eq. (41).

□

Theorem 7 can be used to expand Q_{00} as follows.

Theorem 8. *Let $\vec{z} \in \mathbb{R}^\alpha$. Then,*

$$Q_{00}[\vec{z}](\theta) = \sum_{l=0}^{\lceil\alpha/2\rceil}\left\{\sum_{y\in\Xi_l^\alpha}(-\mathrm{i})^{\mathrm{wt}(\vec{y})}\zeta_{\vec{y}}(\vec{z})\right\}\cos(l\theta) \tag{48}$$

$$= \sum_{\vec{y}\in\{0,1\}^\alpha}\left\{(-\mathrm{i})^{\mathrm{wt}(\vec{y})}\cos(l_{\vec{y}}\theta)\right\}\zeta_{\vec{y}}(\vec{z}), \tag{49}$$

where l_y is the unique l for which $\vec{y} \in \Xi_l^\alpha$.

In other words,

$$Q_{00}[\vec{z}](\theta) = \sum_{l=0}^{\lceil\alpha/2\rceil} a_l(\vec{z})\cos(l\theta) = \sum_{\vec{y}\in\{0,1\}^\alpha}\xi_{\vec{y}}(\theta)\zeta_{\vec{y}}(\vec{z}) \tag{50}$$

is a

(i) cosine polynomial in θ of degree α with Fourier coefficients

$$a_l(\vec{z}) = \sum_{\vec{y}\in\Xi_l^\alpha}(-\mathrm{i})^{\mathrm{wt}(\vec{y})}\zeta_{\vec{y}}(\vec{z}). \tag{51}$$

(ii) trigono-multilinear function in $\vec{z}$ of arity 2α with coefficients

$$\xi_{\vec{y}}(\theta) = (-\mathrm{i})^{\mathrm{wt}(\vec{y})}\cos(l_{\vec{y}}\theta). \tag{52}$$

Proof.

(i) Using Eq. (41), we find that

$$\begin{aligned}
Q_{00}[\vec{z}](\theta) &= \langle\bar{0}|Q(\theta;\vec{z})|\bar{0}\rangle \\
&= \sum_{k=0}^{\infty}\sum_{u,v\in\mathbb{F}_2}\sum_{y\in\Theta_{ukv}^\alpha}(-\mathrm{i})^{\mathrm{wt}(\vec{y})}\zeta_{\vec{y}}(\vec{z})\,\langle\bar{0}|\begin{pmatrix} \cos[(k+v)\theta] & -(-1)^v\sin[(k+v)\theta] \\ (-1)^u\sin[(k+v)\theta] & (-1)^{u+v}\cos[(k+v)\theta] \end{pmatrix}|\bar{0}\rangle \\
&= \sum_{k=0}^{\infty}\sum_{u,v\in\mathbb{F}_2}\sum_{\vec{y}\in\Theta_{ukv}^\alpha}(-\mathrm{i})^{\mathrm{wt}(\vec{y})}\zeta_{\vec{y}}(\vec{z})\cos[(k+v)\theta].
\end{aligned} \tag{53}$$

Denoting $\nu_{\vec{y}}(\vec{z}) = (-\mathrm{i})^{\mathrm{wt}(\vec{y})}\zeta_{\vec{y}}(\vec{z})$, the above expression simplifies to

$$\begin{aligned}
Q_{00}[\vec{z}](\theta) &= \sum_{k=0}^{\infty}\sum_{u,v\in\mathbb{F}_2}\left(\sum_{\vec{y}\in\Theta^{\alpha}_{ukv}}\nu_{\vec{y}}(\vec{z})\right)\cos[(k+v)\theta] \\
&= \sum_{u\in\mathbb{F}_2}\left[\sum_{k=0}^{\infty}\sum_{\vec{y}\in\Theta^{\alpha}_{ul0}}\nu_{\vec{y}}(\vec{z})\cos(k\theta) + \sum_{k=0}^{\infty}\sum_{\vec{y}\in\Theta^{\alpha}_{uk1}}\nu_{\vec{y}}(\vec{z})\cos[(k+1)\theta]\right] \\
&= \sum_{u\in\mathbb{F}_2}\left[\sum_{l=0}^{\infty}\sum_{\vec{y}\in\Theta^{\alpha}_{ul0}}\nu_{\vec{y}}(\vec{z})\cos(k\theta) + \sum_{l=1}^{\infty}\sum_{y\in\Theta^{\alpha}_{u,l-1,1}}\nu_{\vec{y}}(\vec{z})\cos(l\theta)\right] \\
&= \sum_{u\in\mathbb{F}_2}\left(\sum_{l=0}^{\infty}\sum_{\vec{y}\in\Theta^{\alpha}_{ul0}} + \sum_{l=1}^{\infty}\sum_{y\in\Theta^{\alpha}_{u,l-1,1}}\right)\nu_{\vec{y}}(\vec{z})\cos(l\theta) \\
&= \sum_{l=0}^{\infty}\left\{\sum_{u\in\mathbb{F}_2}\left(\sum_{\vec{y}\in\Theta^{\alpha}_{ul0}} + \sum_{\vec{y}\in\Theta^{\alpha}_{u,l-1,1}}\right)\nu_{\vec{y}}(\vec{z})\right\}\cos(l\theta), \quad \text{since } \Theta^{\alpha}_{u\alpha 0} = \Theta^{\alpha}_{u,-1,0} = \emptyset \\
&= \sum_{l=0}^{\infty}\left\{\sum_{\vec{y}\in\Xi^{\alpha}_{l}}\nu_{\vec{y}}(\vec{z})\right\}\cos(l\theta) \\
&= \sum_{l=0}^{\lceil\alpha/2\rceil}\left\{\sum_{\vec{y}\in\Xi^{\alpha}_{l}}(-\mathrm{i})^{\mathrm{wt}(\vec{y})}\zeta_{\vec{y}}(\vec{z})\right\}\cos(l\theta), \qquad (54)
\end{aligned}$$

where the last line follows from the fact that $\Xi^{\alpha}_{l} \neq \emptyset \iff l \leqslant \lceil\alpha/2\rceil$ (by Lemma 4).

(ii) Let $l_{\vec{y}}$ be the unique element in $\{l \in \mathbb{N} : \vec{y} \in \Xi^{\alpha}_{l}\}$. Then,

$$\begin{aligned}
Q_{00}[\vec{z}](\theta) &= \sum_{l=0}^{\lceil\alpha/2\rceil}\left\{\sum_{\vec{y}\in\Xi^{\alpha}_{l}}(-\mathrm{i})^{\mathrm{wt}(\vec{y})}\zeta_{\vec{y}}(\vec{z})\cos(l_{\vec{y}}\theta)\right\} \\
&= \sum_{\vec{y}\in\{0,1\}^{\alpha}}\left\{(-\mathrm{i})^{\mathrm{wt}(\vec{y})}\cos(l_{\vec{y}}\theta)\right\}\zeta_{\vec{y}}(\vec{z}), \qquad (55)
\end{aligned}$$

where the last line follows from the fact that

$$\bigcup_{l=0}^{\lceil\alpha/2\rceil}\Xi^{\alpha}_{l} = \{0,1\}^{\alpha}.$$

□

Next, we will use Theorem 8 to expand the biases $\Lambda^{\mathcal{A}}(\theta;\vec{x})$. But before we do so, we first prove the following lemma:

Lemma 9 (Closure under reversal)**.** *Let $\vec{a},\vec{c} \in \{0,1\}^{2L}$. Then,*

$$\vec{c}1\vec{a}^{R} \in \Xi^{4L+1}_{l} \iff \vec{a}1\vec{c}^{R} \in \Xi^{4L+1}_{l}. \qquad (56)$$

Proof. We shall prove the forward direction. By symmetry, the proof of the backward direction is obtained by switching the roles of a and c in the proof.

Consider the case when $l > 0$. Let

$$\begin{aligned} \vec{c}1\vec{a}^R &\in \Xi_l^{4L+1} = \Theta_{0l0}^{4L+1} \cup \Theta_{1l0}^{4L+1} \cup \Theta_{0,l-1,1}^{4L+1} \cup \Theta_{1,l-1,1}^{4L+1} && (57)\\ \implies q^{a_1}p^{a_2}\dots q^{a_{2L+1}}p^{a_{2L}}qp^{c_{2L}}q^{c_{2L-1}}\dots p^{c_2}q^{c_1} &\sim (qp)^l \text{ or } p(qp)^l \text{ or } (qp)^{l-1}q \text{ or } p(qp)^{l-1}q, && (58) \end{aligned}$$

where $\sim$ is the equivalence relation defined in Eq. (206).

Since string reversal preserves $\sim$ (see Eq. (207)),

$$(q^{a_1}p^{a_2}\dots q^{a_{2L+1}}p^{a_{2L}}qp^{c_{2L}}q^{c_{2L-1}}\dots p^{c_2}q^{c_1})^R \sim ((qp)^l)^R \text{ or } (p(qp)^l)^R \text{ or } ((qp)^{l-1}q)^R \text{ or } (p(qp)^{l-1}q)^R. \quad (59)$$

Now, $p(qp)^l$ and $(qp)^{l-1}q$ are palindromes and mapped to themselves under $(\cdot)^R$; and $(qp)^l$ and $p(qp)^{l-1}q$ are mapped to each other under $(\cdot)^R$. Hence, Eq. (59) becomes

$$\begin{aligned} q^{c_1}p^{c_2}\dots q^{c_{2L+1}}p^{c_{2L}}qp^{a_{2L}}q^{a_{2L-1}}\dots p^{a_2}q^{a_1} &\sim (qp)^l \text{ or } p(qp)^l \text{ or } (qp)^{l-1}q \text{ or } p(qp)^{l-1}q\\ \implies \vec{a}1\vec{c}^R &\in \Theta_{0l0}^{4L+1} \cup \Theta_{1l0}^{4L+1} \cup \Theta_{0,l-1,1}^{4L+1} \cup \Theta_{1,l-1,1}^{4L+1} = \Xi_l^{4L+1}. && (60) \end{aligned}$$

The proof for the case $l = 0$ proceeds similarly, with the last two clauses of each 'or' statement above deleted. □

This lemma may be used to expand the ancilla-free bias.

Theorem 10. *Let $\vec{x} \in \mathbb{R}^{2L}$. Then,*

$$\begin{aligned} \Lambda^{\mathrm{AF}}(\theta;\vec{x}) &= \sum_{l=0}^{2L+1}\left[\left(\sum_{\substack{\vec{a}1\vec{c}^R\in\Xi_l^{4L+1}\\ \mathrm{wt}(\vec{c})-\mathrm{wt}(\vec{a})\equiv_4 0}} - \sum_{\substack{\vec{a}1\vec{c}^R\in\Xi_l^{4L+1}\\ \mathrm{wt}(\vec{c})-\mathrm{wt}(\vec{a})\equiv_4 2}}\right)\zeta_{\vec{a}\vec{c}}(\vec{x},\vec{x})\right]\cos(l\theta) && (61)\\ &= \sum_{\vec{a},\vec{c}\in\{0,1\}^{2L}}\left[\nu_{\mathrm{wt}(\vec{c})-\mathrm{wt}(\vec{a})}\cos(l_{\vec{a}\vec{c}}\theta)\right]\zeta_{\vec{a}\vec{c}}(\vec{x},\vec{x}), && (62) \end{aligned}$$

where $l_{\vec{a}\vec{c}}$ is the unique l for which $\vec{a}1\vec{c}^R \in \Xi_l^{4L+1}$ and

$$\nu_s = \mathrm{Re}(\mathrm{i}^s) = \begin{cases} 1 & s \equiv 0 \bmod 4,\\ 0 & s \equiv 1 \text{ or } 3 \bmod 4,\\ -1 & s \equiv 2 \bmod 4. \end{cases} \quad (63)$$

In other words,

$$\Lambda^{\mathrm{AF}}(\theta;\vec{x}) = \sum_{l=0}^{2L+1}\mu_l^{\mathrm{AF}}(\vec{x})\cos(l\theta) = \sum_{\vec{a},\vec{c}\in\{0,1\}^{2L}}\xi_{\vec{a}\vec{c}}^{\mathrm{AF}}(\theta)\zeta_{\vec{a}\vec{c}}(\vec{x},\vec{x}) \quad (64)$$

is a

(i) cosine polynomial in θ of degree $2L+1$ with Fourier coefficients

$$\mu_l^{\mathrm{AF}}(\vec{x}) = \left(\sum_{\substack{\vec{a}1\vec{c}^R\in\Xi_l^{4L+1}\\ \mathrm{wt}(\vec{c})-\mathrm{wt}(\vec{a})\equiv_4 0}} - \sum_{\substack{\vec{a}1\vec{c}^R\in\Xi_l^{4L+1}\\ \mathrm{wt}(\vec{c})-\mathrm{wt}(\vec{a})\equiv_4 2}}\right)\zeta_{\vec{a}\vec{c}}(\vec{x},\vec{x}). \quad (65)$$

(ii) trigono-multiquadratic function in $\vec{x}$ of arity $2L$ with coefficients

$$\xi_{\vec{a}\vec{c}}^{\mathrm{AF}}(\theta) = \nu_{\mathrm{wt}(\vec{c})-\mathrm{wt}(\vec{a})}\cos(l_{\vec{a}\vec{c}}\theta). \quad (66)$$

In the above, $s \equiv_n t$ means $s \equiv t \pmod n$.

Proof.

(i) By Eq. (14),

$$\Lambda^{\mathrm{AF}}(\theta;\vec{x}) = \mathrm{i}Q_{00}\left[\vec{x},\tfrac{\pi}{2},-\vec{x}^R\right](\theta) \tag{67}$$

$$= \sum_{l=0}^{\lceil(4L+1)/2\rceil}\left\{\mathrm{i}\sum_{\vec{y}\in\Xi_l^{4L+1}}(-\mathrm{i})^{\mathrm{wt}(\vec{y})}\zeta_{\vec{y}}\left(\vec{x},\tfrac{\pi}{2},-\vec{x}^R\right)\right\}\cos(l\theta) \tag{68}$$

$$= \sum_{l=0}^{2L+1}\mu_l^{\mathrm{AF}}(\vec{x})\cos(l\theta), \tag{69}$$

where

$$\mu_l^{\mathrm{AF}}(\vec{x}) = \mathrm{i}\sum_{\vec{y}\in\Xi_l^{4L+1}}(-\mathrm{i})^{\mathrm{wt}(\vec{y})}\underbrace{\zeta_{\vec{y}}\left(\vec{x},\tfrac{\pi}{2},-\vec{x}^R\right)}_{\circledast}. \tag{70}$$

Writing $\vec{y}=\vec{a}b\vec{c}^R = a_1\dots a_{2L}bc_{2L}\dots c_1$, where $\vec{a},\vec{c}\in\{0,1\}^{2L}$ and $b\in\{0,1\}$,

$$\begin{aligned}\circledast &= \zeta_{a_1,\dots,a_{2L},b,c_{2L},\dots,c_1}\left(x_1,\dots,x_{2L},\tfrac{\pi}{2},-x_{2L},\dots,x_1\right)\\ &= \zeta_{a_1}(x_1)\dots\zeta_{a_{2L}}(x_{2L})\zeta_b(\pi/2)\zeta_{c_{2L}}(-x_{2L})\dots\zeta_{c_1}(-x_1).\end{aligned} \tag{71}$$

Since $\zeta_b(\pi/2)=b$ and $\zeta_{c_i}(-x)=(-1)^{c_i}\zeta_{c_i}(x)$ for all i,

$$\begin{aligned}\circledast &= (-1)^{c_1+\dots c_{2L}}b\zeta_{a_1}(x_1)\dots\zeta_{a_{2L}}(x_{2L})\zeta_{c_1}(x_1)\dots\zeta_{c_{2L}}(x_{2L})\\ &= (-1)^{\mathrm{wt}(\vec{c})}b\zeta_{\vec{a}}(\vec{x})\zeta_{\vec{c}}(\vec{x}).\end{aligned} \tag{72}$$

Substituting this back into Eq. (70) gives

$$\begin{aligned}\mu_l^{\mathrm{AF}}(\vec{x}) &= \mathrm{i}\sum_{\vec{a}b\vec{c}^R\in\Xi_l^{4L+1}}(-\mathrm{i})^{\mathrm{wt}(\vec{a})+\mathrm{wt}(b)+\mathrm{wt}(\vec{c})}(-1)^{\mathrm{wt}(\vec{c})}b\zeta_{\vec{a}}(\vec{x})\zeta_{\vec{c}}(\vec{x})\\ &= \sum_{\vec{a}1\vec{c}^R\in\Xi_l^{4L+1}}(-\mathrm{i})^{\mathrm{wt}(\vec{a})}\mathrm{i}^{\mathrm{wt}(\vec{c})}\zeta_{\vec{a}}(\vec{x})\zeta_{\vec{c}}(\vec{x})\\ &= \left(\underbrace{\sum_{\substack{\vec{a}1\vec{c}^R\in\Xi_l^{4L+1}\\ a=c}}}_{①}+\underbrace{\sum_{\substack{\vec{a}1\vec{c}^R\in\Xi_l^{4L+1}\\ \vec{a}<\vec{c}}}}_{②}+\underbrace{\sum_{\substack{\vec{a}1\vec{c}^R\in\Xi_l^{4L+1}\\ \vec{c}<\vec{a}}}}_{③}\right)\underbrace{(-1)^{\mathrm{wt}(\vec{a})}\mathrm{i}^{\mathrm{wt}(\vec{a})+\mathrm{wt}(\vec{c})}\zeta_{\vec{a}}(\vec{x})\zeta_{\vec{c}}(\vec{x})}_{(**)},\end{aligned} \tag{73}$$

where $<$ denotes any lexicographical ordering of strings.

First, we calculate

$$\begin{aligned}③(**) &= \sum_{\substack{\vec{a}1\vec{c}^R\in\Xi_l^{4L+1}\\ \vec{c}<\vec{a}}}(-1)^{\mathrm{wt}(\vec{a})}\mathrm{i}^{\mathrm{wt}(\vec{a})+\mathrm{wt}(\vec{c})}\zeta_{\vec{a}}(\vec{x})\zeta_{\vec{c}}(\vec{x})\\ &= \sum_{\substack{\vec{c}1\vec{a}^R\in\Xi_l^{4L+1}\\ \vec{a}<\vec{c}}}(-1)^{\mathrm{wt}(\vec{c})}\mathrm{i}^{\mathrm{wt}(\vec{a})+\mathrm{wt}(\vec{c})}\zeta_{\vec{a}}(\vec{x})\zeta_{\vec{c}}(\vec{x}) \qquad \text{by interchanging } \vec{a} \text{ and } \vec{c}\\ &= \sum_{\substack{\vec{a}1\vec{c}^R\in\Xi_l^{4L+1}\\ \vec{a}<\vec{c}}}(-1)^{\mathrm{wt}(\vec{c})}\mathrm{i}^{\mathrm{wt}(\vec{a})+\mathrm{wt}(\vec{c})}\zeta_{\vec{a}}(\vec{x})\zeta_{\vec{c}}(\vec{x}) \qquad \text{by Lemma 9.}\end{aligned} \tag{74}$$

Hence,

$$\left(②+③\right)(**) = \sum_{\substack{\vec{a}1\vec{c}^R\in\Xi_l^{4L+1}\\ \vec{a}<\vec{c}}} \left[(-1)^{\mathrm{wt}(\vec{c})+\mathrm{wt}(\vec{a})}\right] \mathrm{i}^{\mathrm{wt}(\vec{a})+\mathrm{wt}(\vec{c})}\zeta_{\vec{a}}(\vec{x})\zeta_{\vec{c}}(\vec{x})$$
$$= \sum_{\substack{\vec{a}1\vec{c}^R\in\Xi_l^{4L+1}\\ \vec{a}<\vec{c}}} \nu_{\mathrm{wt}(\vec{c})-\mathrm{wt}(\vec{a})}\zeta_{\vec{a}\vec{c}}(\vec{x},\vec{x}), \tag{75}$$

where we used the fact that for all $\alpha,\beta\in\mathbb{Z}$,

$$[(-1)^\alpha+(-1)^\beta]\mathrm{i}^{\alpha+\beta} = 2\nu_{\beta-\alpha}, \tag{76}$$

where ν_a is given by Eq. (63).

Therefore,

$$\left(②+③\right)(**) = 2\left(\sum_{\substack{\vec{a}1\vec{c}^R\in\Xi_l^{4L+1}\\ \vec{a}<\vec{c}\\ \mathrm{wt}(\vec{c})-\mathrm{wt}(\vec{a})\equiv_4 0}} + \sum_{\substack{\vec{a}1\vec{c}^R\in\Xi_l^{4L+1}\\ \vec{a}<\vec{c}\\ \mathrm{wt}(\vec{c})-\mathrm{wt}(\vec{a})\equiv_4 1 \text{ or } 3}} + \sum_{\substack{\vec{a}1\vec{c}^R\in\Xi_l^{4L+1}\\ \vec{a}<\vec{c}\\ \mathrm{wt}(\vec{c})-\mathrm{wt}(\vec{a})\equiv_4 2}}\right)\nu_{\mathrm{wt}(\vec{c})-\mathrm{wt}(\vec{a})}\zeta_{\vec{a}\vec{c}}(\vec{x},\vec{x})$$
$$= 2\left(\sum_{\substack{\vec{a}1\vec{c}^R\in\Xi_l^{4L+1}\\ \vec{a}<\vec{c}\\ \mathrm{wt}(\vec{c})-\mathrm{wt}(\vec{a})\equiv_4 0}} - \sum_{\substack{\vec{a}1\vec{c}^R\in\Xi_l^{4L+1}\\ \vec{a}<\vec{c}\\ \mathrm{wt}(\vec{c})-\mathrm{wt}(\vec{a})\equiv_4 2}}\right)\zeta_{\vec{a}\vec{c}}(\vec{x},\vec{x}). \tag{77}$$

Note that replacing $\vec{a}<\vec{c}$ with $\vec{c}<\vec{a}$ does not change the value of $\left(②+③\right)(**)$.

Indeed,

$$\left(②+③\right)(**) = 2\left(\sum_{\substack{\vec{a}1\vec{c}^R\in\Xi_l^{4L+1}\\ \vec{a}<\vec{c}\\ \mathrm{wt}(\vec{c})-\mathrm{wt}(\vec{a})\equiv_4 0}} - \sum_{\substack{\vec{a}1\vec{c}^R\in\Xi_l^{4L+1}\\ \vec{a}<\vec{c}\\ \mathrm{wt}(\vec{c})-\mathrm{wt}(\vec{a})\equiv_4 2}}\right)\zeta_{\vec{a}\vec{c}}(\vec{x},\vec{x})$$
$$= 2\left(\sum_{\substack{\vec{c}1\vec{a}^R\in\Xi_l^{4L+1}\\ \vec{a}<\vec{c}\\ \mathrm{wt}(\vec{c})-\mathrm{wt}(\vec{a})\equiv_4 0}} - \sum_{\substack{\vec{c}1\vec{a}^R\in\Xi_l^{4L+1}\\ \vec{a}<\vec{c}\\ \mathrm{wt}(\vec{c})-\mathrm{wt}(\vec{a})\equiv_4 2}}\right)\zeta_{\vec{a}\vec{c}}(\vec{x},\vec{x}), \quad \text{by Lemma 9}$$
$$= 2\left(\sum_{\substack{\vec{a}1\vec{c}^R\in\Xi_l^{4L+1}\\ \vec{c}<\vec{a}\\ \mathrm{wt}(\vec{a})-\mathrm{wt}(\vec{c})\equiv_4 0}} - \sum_{\substack{\vec{a}1\vec{c}^R\in\Xi_l^{4L+1}\\ \vec{c}<\vec{a}\\ \mathrm{wt}(\vec{a})-\mathrm{wt}(\vec{c})\equiv_4 2}}\right)\zeta_{\vec{c}\vec{a}}(\vec{x},\vec{x}), \quad \text{by switching the labels } a \text{ and } c$$
$$= 2\left(\sum_{\substack{\vec{a}1\vec{c}^R\in\Xi_l^{4L+1}\\ \vec{c}<\vec{a}\\ \mathrm{wt}(\vec{c})-\mathrm{wt}(\vec{a})\equiv_4 0}} - \sum_{\substack{\vec{a}1\vec{c}^R\in\Xi_l^{4L+1}\\ \vec{c}<\vec{a}\\ \mathrm{wt}(\vec{c})-\mathrm{wt}(\vec{a})\equiv_4 2}}\right)\zeta_{\vec{a}\vec{c}}(\vec{x},\vec{x}), \tag{78}$$

where the last line follows from the facts that (i) $t\equiv_4 0 \Leftrightarrow -t\equiv_4 0$, (ii) $t\equiv_4 2\Leftrightarrow -t\equiv_4 2$, and (iii) $\zeta_{\vec{a}\vec{c}}(\vec{x},\vec{x}) = \zeta_{\vec{c}\vec{a}}(\vec{x},\vec{x})$.

By taking the average of Eqs. (77) and (78), we obtain

$$\left(②+③\right)(**) = \left(\sum_{\substack{\vec{a}1\vec{c}^R\in\Xi_l^{4L+1}\\ \vec{c}\neq\vec{a}\\ \mathrm{wt}(\vec{c})-\mathrm{wt}(\vec{a})\equiv_4 0}} - \sum_{\substack{\vec{a}1\vec{c}^R\in\Xi_l^{4L+1}\\ \vec{c}\neq\vec{a}\\ \mathrm{wt}(\vec{c})-\mathrm{wt}(\vec{a})\equiv_4 2}}\right)\zeta_{\vec{a}\vec{c}}(\vec{x},\vec{x}). \tag{79}$$

Next, we calculate

$$①(**) = \sum_{\vec{a}1\vec{a}^R \in \Xi_l^{4L+1}} \zeta_{\vec{a}}(\vec{x})^2, \tag{80}$$

which can be expressed as

$$①(**) = \left(\sum_{\substack{\vec{a}1\vec{c}^R \in \Xi_l^{4L+1} \\ \vec{c}=\vec{a} \\ \mathrm{wt}(\vec{c})-\mathrm{wt}(\vec{a}) \equiv_4 0}} - \sum_{\substack{\vec{a}1\vec{c}^R \in \Xi_l^{4L+1} \\ \vec{c}=\vec{a} \\ \mathrm{wt}(\vec{c})-\mathrm{wt}(\vec{a}) \equiv_4 2}} \right) \zeta_{\vec{a}\vec{c}}(\vec{x}, \vec{x}). \tag{81}$$

Substituting Eq. (79) and (81) into Eq. (73) gives

$$\mu_l^{\mathrm{AF}}(\vec{x}) = \left(\sum_{\substack{\vec{a}1\vec{c}^R \in \Xi_l^{4L+1} \\ \mathrm{wt}(\vec{c})-\mathrm{wt}(\vec{a}) \equiv_4 0}} - \sum_{\substack{\vec{a}1\vec{c}^R \in \Xi_l^{4L+1} \\ \mathrm{wt}(\vec{c})-\mathrm{wt}(\vec{a}) \equiv_4 2}} \right) \zeta_{\vec{a}\vec{c}}(\vec{x}, \vec{x}). \tag{82}$$

(ii) By arranging the terms in the sum differently, we obtain

$$\begin{aligned}
\Lambda^{\mathrm{AF}}(\theta; \vec{x}) &= \sum_{l=0}^{2L+1} \left[\left(\sum_{\substack{\vec{a}1\vec{c}^R \in \Xi_l^{4L+1} \\ \mathrm{wt}(\vec{c})-\mathrm{wt}(\vec{a}) \equiv_4 0}} - \sum_{\substack{\vec{a}1\vec{c}^R \in \Xi_l^{4L+1} \\ \mathrm{wt}(\vec{c})-\mathrm{wt}(\vec{a}) \equiv_4 2}} \right) \zeta_{\vec{a}\vec{c}}(\vec{x}, \vec{x}) \right] \cos(l\theta) \\
&= \sum_{l=0}^{2L+1} \sum_{\vec{a}1\vec{c}^R \in \Xi_l^{4L+1}} \nu_{\mathrm{wt}(\vec{c})-\mathrm{wt}(\vec{a})} \zeta_{\vec{a}\vec{c}}(\vec{x}, \vec{x}) \cos(l_{\vec{a}\vec{c}}\theta) \\
&= \sum_{\vec{a},\vec{c} \in \{0,1\}^{2L}} \left[\nu_{\mathrm{wt}(\vec{c})-\mathrm{wt}(\vec{a})} \cos(l_{\vec{a}\vec{c}}\theta) \right] \zeta_{\vec{a}\vec{c}}(\vec{x}, \vec{x}),
\end{aligned} \tag{83}$$

which completes the proof of the theorem.

□

Note that the Fourier coefficients (65) can also be expressed as

$$\mu_l^{\mathrm{AF}}(\vec{x}) = \sum_{a \in \Gamma_l^{2L}} \zeta_{\vec{a}}(\vec{x})^2 + 2 \left(\sum_{\vec{b} \in \Omega_{l,0}^{4L}} - \sum_{\vec{b} \in \Omega_{l,2}^{4L}} \right) \zeta_{\vec{b}}(\vec{x}, \vec{x}), \tag{84}$$

where

$$\Omega_{l,K}^{4L} = \{\vec{a}\vec{c} \in \{0,1\}^{2L} \times \{0,1\}^{2L} : \vec{a}1\vec{c}^R \in \Xi_l^{4L+1},\ a < c,\ \mathrm{wt}(c) - \mathrm{wt}(a) \equiv K \bmod 4\}, \tag{85}$$

$$\Gamma_l^{2L} = \{\vec{a} \in \{0,1\}^{2L} : \vec{a}1\vec{a}^R \in \Xi_l^{4L+1}\}. \tag{86}$$

Next, we expand the ancilla-based bias.

Theorem 11. *Let $\vec{x} \in \mathbb{R}^{2L}$. Then,*

$$\Lambda^{\mathrm{AB}}(\theta; \vec{x}) = \sum_{l=0}^{L} \left[\left(\sum_{\substack{\vec{y} \in \Xi_l^{2L} \\ \mathrm{wt}(\vec{y}) \equiv_4 0}} - \sum_{\substack{\vec{y} \in \Xi_l^{2L} \\ \mathrm{wt}(\vec{y}) \equiv_4 2}} \right) \zeta_{\vec{y}}(\vec{x}) \right] \cos(l\theta) \tag{87}$$

$$= \sum_{\vec{y} \in \{0,1\}^{2L}} \left[\nu_{\mathrm{wt}(\vec{y})} \cos(l_{\vec{y}}\theta) \right] \zeta_{\vec{y}}(\vec{x}) \tag{88}$$

where $l_{\vec{y}}$ is the unique l for which $\vec{y} \in \Xi_l^{2L}$ and ν_s is given by Eq. (63).

In other words,

$$\Lambda^{\mathrm{AB}}(\theta;\vec{x}) = \sum_{l=0}^{L} \mu_l^{\mathrm{AB}}(\vec{x})\cos(l\theta) = \sum_{\vec{y}\in\{0,1\}^{2L}} \xi_{\vec{y}}^{\mathrm{AB}}(\theta)\zeta_{\vec{y}}(\vec{x}) \tag{89}$$

is a

(i) cosine polynomial in θ of degree L with Fourier coefficients

$$\mu_l^{\mathrm{AB}}(\vec{x}) = \left(\sum_{\substack{\vec{y}\in\Xi_l^{2L} \\ \mathrm{wt}(\vec{y})\equiv_4 0}} - \sum_{\substack{\vec{y}\in\Xi_l^{2L} \\ \mathrm{wt}(\vec{y})\equiv_4 2}} \right) \zeta_{\vec{y}}(\vec{x}). \tag{90}$$

(ii) trigono-multilinear function in $\vec{x}$ of arity $2L$ with coefficients

$$\xi_{\vec{y}}^{\mathrm{AB}}(\theta) = \nu_{\mathrm{wt}(\vec{y})}\cos(l_{\vec{y}}\theta). \tag{91}$$

Proof.

(i) From Eq. (13), the ancilla-based bias may be written as

$$\Lambda^{\mathrm{AB}}(\theta;\vec{x}) = \operatorname{Re} Q_{00}[\vec{x}](\theta) = \sum_{l=0}^{L} \left\{ \sum_{y\in\Xi_l^{2L}} \operatorname{Re}(-\mathrm{i})^{\mathrm{wt}(\vec{y})}\zeta_{\vec{y}}(\vec{x}) \right\} \cos(l\theta), \tag{92}$$

which follows from Eq. (48).

Therefore, the Fourier coefficients (90) can be simplified as

$$\mu_l^{\mathrm{AB}}(\vec{x}) = \operatorname{Re}(-\mathrm{i})^{\mathrm{wt}(\vec{y})}\zeta_{\vec{y}}(\vec{z}) = \sum_{\vec{y}\in\Xi_l^{2L}} \nu_{\mathrm{wt}(\vec{y})}\zeta_{\vec{y}}(\vec{x}) \tag{93}$$

$$= \left(\sum_{\substack{\vec{y}\in\Xi_l^{2L} \\ \mathrm{wt}(\vec{y})\equiv_4 0}} + \sum_{\substack{\vec{y}\in\Xi_l^{2L} \\ \mathrm{wt}(\vec{y})\equiv_4 1 \text{ or } 3}} + \sum_{\substack{\vec{y}\in\Xi_l^{2L} \\ \mathrm{wt}(\vec{y})\equiv_4 2}} \right) \nu_{\mathrm{wt}(\vec{y})}\zeta_{\vec{y}}(\vec{x}) = \left(\sum_{\substack{\vec{y}\in\Xi_l^{2L} \\ \mathrm{wt}(\vec{y})\equiv_4 0}} - \sum_{\substack{\vec{y}\in\Xi_l^{2L} \\ \mathrm{wt}(\vec{y})\equiv_4 2}} \right) \zeta_{\vec{y}}(\vec{x}). \tag{94}$$

(ii) From Eqs. (93) and (92),

$$\begin{aligned}\Lambda^{\mathrm{AB}}(\theta;\vec{x}) &= \sum_{l=0}^{L} \sum_{\vec{y}\in\Xi_l^{2L}} \nu_{\mathrm{wt}(\vec{y})}\zeta_{\vec{y}}(\vec{x})\cos(l\theta) = \sum_{l=0}^{L} \sum_{\vec{y}\in\Xi_l^{2L}} \left\{ \nu_{\mathrm{wt}(\vec{y})}\zeta_{\vec{y}}(\vec{x})\cos(l_{\vec{y}}\theta) \right\} \\ &= \sum_{\vec{y}\in\{0,1\}^{2L}} \left[\nu_{\mathrm{wt}(\vec{y})}\cos(l_{\vec{y}}\theta) \right] \zeta_{\vec{y}}(\vec{x}). \end{aligned} \tag{95}$$

□

In summary, Theorems 10 and 11 give the expansion formulas for the ancilla-free and ancilla-based biases respectively, which we will use in Section 4. At least two special cases of these expansion formulas yield nice simple expressions and are of interest: (i) when the angles $\vec{x}$ are chosen to be $(\pi/2)^{2L}$ and (ii) when $L = 1$. We study Case (i) in Section 2.4 and as mentioned above, Case (ii) in Appendix E. Finally, in Appendix F, as an application of the expansion formulas, we derive expressions for the leading terms in the cosine expansions (65) and (90).

2.4 Special case: Chebyshev likelihood functions

Here we will consider the special case when the tunable parameters $\vec{x} \in \mathbb{R}^{2L}$ in Eq. (10) are chosen to be $\vec{x} = (\pi/2)^{2L}$ and show how the results in Sections 2–2.3 simplify in this case. In particular, we shall show that the biases of the likelihood function in both the ancilla-free and ancilla-bassed schemes can be written in terms of Chebyshev polynomials in $\langle\bar{0}| P(\theta) |\bar{0}\rangle$. Due to this, we refer to the likelihood functions (10) with parameters $\vec{x} = (\pi/2)^{2L}$ as *Chebyshev likelihood functions* (CLFs).

Setting $\alpha = \beta = \pi/2$ in Eqs. (7) and (8) gives

$$\begin{aligned} U(\theta; \tfrac{\pi}{2}) &= -\mathrm{i}P(\theta), \\ V(\tfrac{\pi}{2}) &= -\mathrm{i}\bar{Z}. \end{aligned} \tag{96}$$

Hence, Eq. (9) evaluates to

$$\begin{aligned} Q\left(\theta; (\tfrac{\pi}{2})^{2L}\right) &= \left(V(\tfrac{\pi}{2})U(\theta; \tfrac{\pi}{2})\right)^L \\ &= (-1)^L(\bar{Z}P(\theta))^L \\ &= (-1)^L \begin{pmatrix} \cos L\theta & \sin L\theta \\ -\sin L\theta & \cos L\theta \end{pmatrix}, \end{aligned} \tag{97}$$

where the matrix above is written in the basis $\{|\bar{0}\rangle, |\bar{1}\rangle\}$.

By making use of Proposition 1, for example, the ancilla-free and the ancilla-based biases may be expressed as

$$\begin{aligned} \Lambda^{\mathrm{AF}}\left(\theta; (\tfrac{\pi}{2})^{2L}\right) &= \cos[(2L+1)\theta], \\ \Lambda^{\mathrm{AB}}\left(\theta; (\tfrac{\pi}{2})^{2L}\right) &= (-1)^L \cos(L\theta). \end{aligned} \tag{98}$$

Hence, by Eq. (10) the likelihood functions corresponding to these biases are

$$\begin{aligned} \mathcal{L}^{\mathrm{AF}}(\theta; d, \vec{x}) &= \frac{1}{2}\left[1 + (-1)^d \cos[(2L+1)\theta]\right], \\ \mathcal{L}^{\mathrm{AB}}(\theta; d, \vec{x}) &= \frac{1}{2}\left[1 + (-1)^{d+L} \cos(L\theta)\right], \end{aligned} \tag{99}$$

By rewriting θ in terms of the expectation value $\langle\bar{0}| P(\theta) |\bar{0}\rangle$ using Eq. (2), the above biases can be expressed in terms of Chebyshev polynomials:

$$\begin{aligned} \Lambda^{\mathrm{AF}}\left(\theta; (\tfrac{\pi}{2})^{2L}\right) &= \cos[(2L+1)\arccos\langle\bar{0}| P(\theta) |\bar{0}\rangle] = \mathcal{T}_{2L+1}(\langle\bar{0}| P(\theta) |\bar{0}\rangle), \\ \Lambda^{\mathrm{AB}}\left(\theta; (\tfrac{\pi}{2})^{2L}\right) &= (-1)^L \cos(L\arccos\langle\bar{0}| P(\theta) |\bar{0}\rangle) = (-1)^L\mathcal{T}_L(\langle\bar{0}| P(\theta) |\bar{0}\rangle), \end{aligned} \tag{100}$$

where $\mathcal{T}_m(x) = \cos(m \arccos x)$, for $|x| \leq 1$, is the nth Chebyshev polynomial of the first kind. As explained above, it is for this reason that the likelihood functions in (99) are called Chebyshev likelihood functions.

Note that the biases (98) can trivially be expanded as Fourier series:

$$\begin{aligned} \Lambda^{\mathrm{AF}}\left(\theta; (\tfrac{\pi}{2})^{2L}\right) &= \sum_{l=0}^{2L+1} \delta_{l,2L+1} \cos(l\theta), \\ \Lambda^{\mathrm{AB}}\left(\theta; (\tfrac{\pi}{2})^{2L}\right) &= \sum_{l=0}^{L} (-1)^L \delta_{l,L} \cos(l\theta). \end{aligned} \tag{101}$$

Hence, the Fourier coefficients given by Eqs. (65) and (90) can be read off to be

$$\begin{aligned} \mu_l^{\mathrm{AF}}\left((\tfrac{\pi}{2})^{2L}\right) &= \delta_{l,2L+1}, \\ \mu_l^{\mathrm{AB}}\left((\tfrac{\pi}{2})^{2L}\right) &= (-1)^L \delta_{l,L}. \end{aligned} \tag{102}$$

3 The expected posterior variance: a Bayesian perspective

In this section, we develop tools for understanding the expected posterior variance, which will be useful in Section 4 when we apply them to the quantum-generated likelihood functions that we introduced in Section 2. The treatment in this section, which can be read independently of Section 2, is fairly general and applicable to a broad range of likelihood functions.

3.1 General form

Suppose that we have a (continuous) prior distribution $p(\theta) = p_\uptheta(\theta)$ that reflects our current state of knowledge about the value of some parameter $\theta \in \Theta$, where $\Theta \subseteq \mathbb{R}$. To update our knowledge about θ, we may design an experiment that yields measurement outcomes $d \in \Omega$, where the set of possible measurement outcomes $\Omega \subseteq \mathbb{N}$ is assumed to be a discrete set. The probability of obtaining an outcome d given the unknown parameter θ is captured by the (continuous) likelihood function $\mathcal{L}(\theta; d) = \mathbb{P}_{\mathrm{d}|\uptheta}(d|\theta)$. Our knowledge of θ after performing Bayesian updating is described by the (continuous) posterior distribution

$$p_{\uptheta|\mathrm{d}}(\theta|d) = \frac{\mathcal{L}(\theta; d)p(\theta)}{\mathbb{P}_\mathrm{d}(d)} \tag{103}$$

where

$$\mathbb{P}_\mathrm{d}(d) = \int \mathrm{d}\theta \ \mathcal{L}(\theta; d)p(\theta) \tag{104}$$

is the marginal distribution of d. For Eq. (103) to be well-defined, the outcome d must have a nonzero probability of occurring, i.e. $d \in \Omega' := \{d \in \Omega : \mathbb{P}_\mathrm{d}(d) \neq 0\}$.

Note that different experiments could give rise to different likelihood functions, and hence, different amounts of information gain about θ. Our goal is to engineer likelihood functions that allow us to minimize the *expected posterior variance*, defined as

$$\mathbb{E}_\mathrm{d}\mathrm{Var}_{\uptheta|\mathrm{d}}(\theta|d) = \sum_{d\in\Omega'} \mathbb{P}_\mathrm{d}(d)\mathrm{Var}_{\uptheta|\mathrm{d}}(\theta|d). \tag{105}$$

In the above formula, $\mathrm{Var}_{\uptheta|\mathrm{d}}(\theta|d)$ is the variance of the posterior distribution upon obtaining outcome d, i.e.

$$\mathrm{Var}_{\uptheta|\mathrm{d}}(\theta|d) = \mathbb{E}_{\uptheta|\mathrm{d}}[\theta^2|d] - \left(\mathbb{E}_{\uptheta|\mathrm{d}}[\theta|d]\right)^2, \tag{106}$$

where

$$\mathbb{E}_{\uptheta|\mathrm{d}}[\theta|d] = \int \mathrm{d}\theta \ \theta \ p_{\uptheta|\mathrm{d}}(\theta|d) \tag{107}$$

$$\mathbb{E}_{\uptheta|\mathrm{d}}[\theta^2|d] = \int \mathrm{d}\theta \ \theta^2 \ p_{\uptheta|\mathrm{d}}(\theta|d) \tag{108}$$

are the first and second moments of $p_{\uptheta|\mathrm{d}}(\cdot|d)$ respectively.

The expression for the expected posterior variance as given by Eqs. (105)–(108) involves the computation of several integrals, which may be computationally intensive in general. Hence, our goal in this section is to provide various simplifications of the formula (105) as we specialize to specific experiments that are relevant to this work. To this end, our first step is to show that the expected posterior variance can be written in terms of the prior mean

$$\mu = \int \mathrm{d}\theta \ \theta p(\theta) \tag{109}$$

and prior variance

$$\sigma^2 = \int \mathrm{d}\theta \ (\theta - \mu)^2 p(\theta) \tag{110}$$

as follows:

Theorem 12. *The expected posterior variance is given by*

$$\mathbb{E}_{\mathrm{d}} \operatorname{Var}_{\theta|\mathrm{d}}(\theta|d) = \sigma^2 + \mu^2 - \sum_{d\in\Omega'} \frac{I_1(d)^2}{I_0(d)} \tag{111}$$

where

$$I_k(d) = \int \mathrm{d}\theta\ \theta^k \mathcal{L}(\theta; d) p(\theta) \tag{112}$$

is the kth moment of the function $\mathcal{L}(\,\cdot\,; d)p(\cdot)$.

Proof.

$$\begin{aligned}
\mathbb{E}_{\mathrm{d}} \operatorname{Var}_{\theta|\mathrm{d}}(\theta|d) &= \sum_{d\in\Omega'} \mathbb{P}_{\mathrm{d}}(d) \operatorname{Var}_{\theta|\mathrm{d}}(\theta|d) \\
&= \sum_{d\in\Omega'} \mathbb{P}_{\mathrm{d}}(d) \left\{ \int \mathrm{d}\theta\ \theta^2 p_{\theta|\mathrm{d}}(\theta|d) - \left[\int \mathrm{d}\theta\ \theta p_{\theta|\mathrm{d}}(\theta|d) \right]^2 \right\} \\
&= \sum_{d\in\Omega'} \mathbb{P}_{\mathrm{d}}(d) \int \mathrm{d}\theta\ \theta^2 \frac{\mathcal{L}(\theta; d)p(\theta)}{\mathbb{P}_{\mathrm{d}}(d)} - \sum_{d\in\Omega'} \mathbb{P}_{\mathrm{d}}(d) \left(\int \mathrm{d}\theta\ \theta \frac{\mathcal{L}(\theta; d)p(\theta)}{\mathbb{P}_{\mathrm{d}}(d)} \right)^2 \\
&= \sum_{d\in\Omega'} \int \mathrm{d}\theta\ \theta^2 \mathcal{L}(\theta; d)p(\theta) - \sum_{d\in\Omega'} \frac{\left[\int \mathrm{d}\theta\ \theta \mathcal{L}(\theta; d)p(\theta)\right]^2}{\int \mathrm{d}\theta\ \mathcal{L}(\theta; d)p(\theta)} \\
&= \int \mathrm{d}\theta\ \theta^2 \underbrace{\sum_{d\in\Omega'} \mathcal{L}(\theta; d)}_{=1} p(\theta) - \sum_{d\in\Omega'} \frac{I_1(d)^2}{I_0(d)} \\
&= \sigma^2 + \mu^2 - \sum_{d\in\Omega'} \frac{I_1(d)^2}{I_0(d)}.
\end{aligned}$$

where we used the fact that

$$\sigma^2 = \left(\int \mathrm{d}\theta\ \theta^2 p(\theta) \right) - \left(\int \mathrm{d}\theta\ \theta p(\theta) \right)^2 = \left(\int \mathrm{d}\theta\ \theta^2 p(\theta) \right) - \mu^2.$$

in the last line.

□

Note that Eq. (111) can also be expressed in the following form:

$$\mathbb{E}_{\mathrm{d}} \operatorname{Var}_{\theta|\mathrm{d}}(\theta|d) = \sigma^2(1 - \sigma^2 \mathcal{V}) \tag{113}$$

where

$$\mathcal{V} = \frac{1}{\sigma^4} \left[\sum_{d\in\Omega'} \frac{I_1(d)^2}{I_0(d)} - \mu^2 \right]. \tag{114}$$

We shall call the symbol $\mathcal{V}$ defined by Eq. (113) the *variance reduction factor*. The motivation for introducing $\mathcal{V}$ is that working with $\mathcal{V}$ turns out to be more convenient than working directly with the expected posterior variance when we specialize to specific domains Ω, prior distributions $p(\theta)$ and likelihood functions $\mathcal{L}(\theta; d)$. We refer the reader to Figure 3.1 for a summary of the specializations that we will consider in Sections 3.2–3.4.1 as well as the results we obtained.

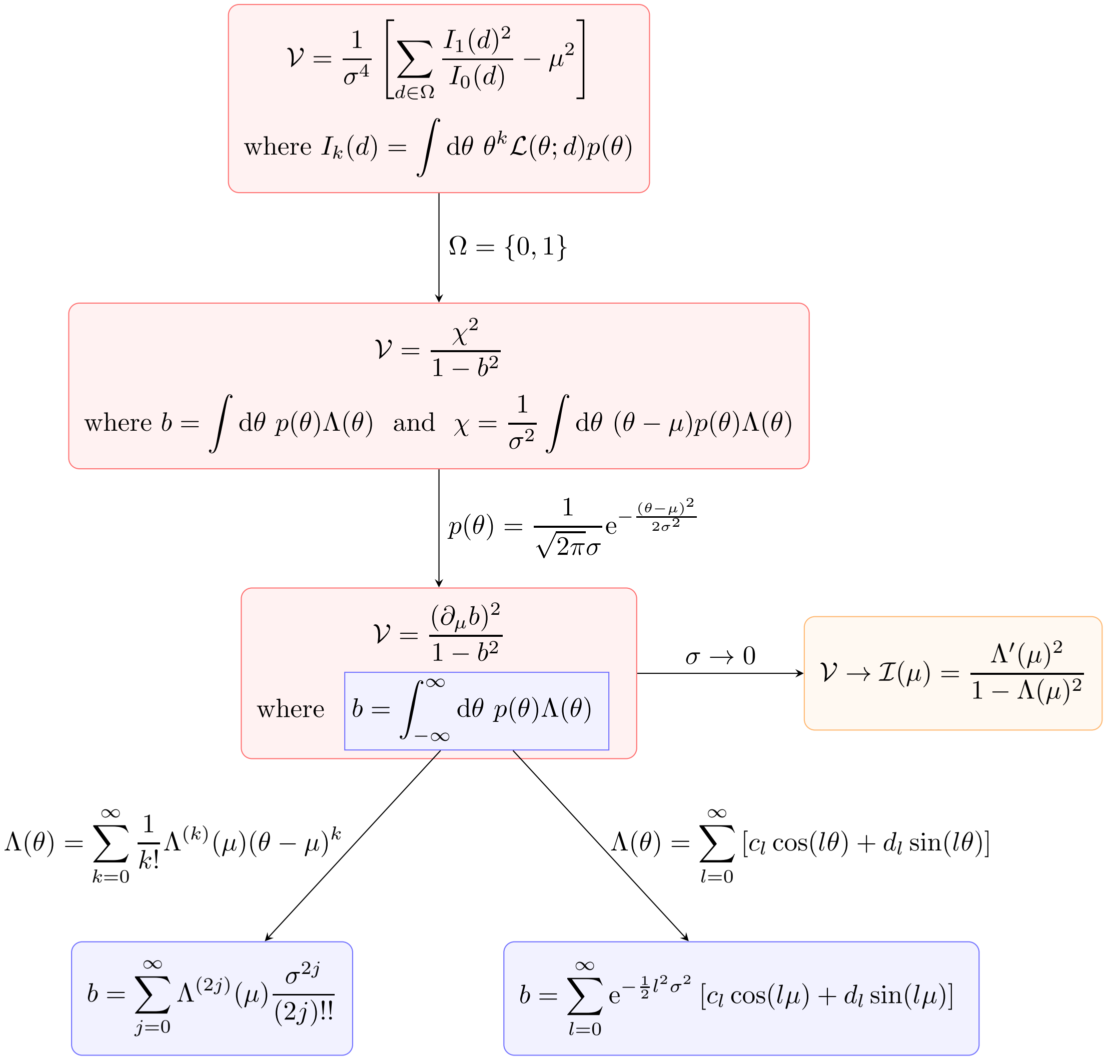


Figure 3.1: Flowchart showing expressions for the variance reduction factor $\mathcal{V}$ under various assumptions. At the top of the flowchart is the most general expression for $\mathcal{V}$. As we specialize to specific Ω, $p(\theta)$ and $\Lambda(\theta)$ by moving down the flowchart, the expression for $\mathcal{V}$ simplifies. Here, (i) Ω represents the set of measurement outcome values, (ii) $p(\theta)$ is the prior distribution, (iii) $\mathcal{L}(\theta;d) = \mathbb{P}_{\mathrm{d}|\theta}(d|\theta)$ is the likelihood function, (iv) $\mu = \int \mathrm{d}\theta\ \theta p(\theta)$ is the prior mean, and (v) $\sigma^2 = \int \mathrm{d}\theta\ (\theta-\mu)^2 p(\theta)$ is the prior variance. If $\Omega = \{0,1\}$, let $\Lambda(\theta) = 2\mathcal{L}(\theta;0) - 1$ be the bias. In this flowchart, we assume that $b \neq 1$, $|\Lambda(\mu)| \neq 1$, and $I_0(d) \neq 0$ for all $d \in \Omega$.

3.2 Two-outcome likelihood functions

In the rest of this paper, we specialize to the case when $\Omega = \{0,1\}$. In this case, we show that the expression for the expected posterior variance simplifies as follows.

Theorem 13. *Assuming that $\Omega = \{0,1\}$, the expected posterior variance may be written as*

$$\mathbb{E}_{\mathrm{d}}\,\mathrm{Var}_{\theta|\mathrm{d}}(\theta|d) = \begin{cases}\sigma^2 & I_0 \in \{0,1\}\\ \sigma^2 - \frac{(I_1-\mu I_0)^2}{I_0(1-I_0)} & I_0 \notin \{0,1\}\end{cases} \tag{115}$$

where for $k \in \{0,1\}$,

$$I_k := I_k(0) = \int \mathrm{d}\theta\ \theta^k \mathcal{L}(\theta;0)p(\theta) \tag{116}$$

is the kth moment of the function $\mathcal{L}(\,\cdot\,;0)p(\cdot)$.

Proof. Consider

$$\sum_{d\in\{0,1\}} I_k(d) = \int \mathrm{d}\theta\ \theta^k \underbrace{\sum_{d\in\{0,1\}} \mathcal{L}(\theta;d)}_{=1} p(\theta) \tag{117}$$

$$= \int \mathrm{d}\theta\ \theta^k p(\theta). \tag{118}$$

Equivalently,

$$I_0(0) + I_0(1) = 1, \tag{119}$$

$$I_1(0) + I_1(1) = \mu. \tag{120}$$

We first consider the case when $I_0 = I_0(0) = 1$. In this case,

$$0 = I_0(1) = \int \mathrm{d}\theta\ \mathcal{L}(\theta;1)p(\theta). \tag{121}$$

Since $\mathcal{L}(\theta;1)$ and $p(\theta)$ are both continuous and nonnegative, it follows that

$$\mathcal{L}(\theta;1)p(\theta) = 0. \tag{122}$$

This implies that

$$\begin{aligned} I_1(0) &= \int \mathrm{d}\theta\ \theta\mathcal{L}(\theta;0)p(\theta) \\ &= \int \mathrm{d}\theta\ \theta\,[1-\mathcal{L}(\theta;1)]\,p(\theta) \\ &= \mu - \int \mathrm{d}\theta\ \theta\ \underbrace{\mathcal{L}(\theta;1)p(\theta)}_{=0,\text{ by Eq. (122)}} \\ &= \mu. \end{aligned} \tag{123}$$

Since $\Omega' = \{0\}$, it follows that

$$\begin{aligned} \mathbb{E}_{\mathrm{d}} \operatorname{Var}_{\theta|\mathrm{d}}(\theta|d) &= \sigma^2 + \mu^2 - \frac{I_1(0)^2}{I_0(0)} \\ &= \sigma^2 + \mu^2 - \mu^2/1 \\ &= \sigma^2. \end{aligned} \tag{124}$$

By symmetry, the case when $I_0 = 0$ is similar, and also gives $\mathbb{E}_{\mathrm{d}} \operatorname{Var}_{\theta|\mathrm{d}}(\theta|d) = \sigma^2$.

Finally, we consider the case when $I_0 \notin \{0,1\}$. In this case,

$$\mathbb{E}_{\mathrm{d}} \operatorname{Var}_{\theta|\mathrm{d}}(\theta|d) = \sigma^2 + \mu^2 - \sum_{d\in\{0,1\}} \frac{I_1(d)^2}{I_0(d)}. \tag{125}$$

The last term of the above expression can be simplified as

$$\begin{aligned}
\sum_{d\in\{0,1\}} \frac{I_1(d)^2}{I_0(d)} &= \frac{I_1(0)^2}{I_0(0)} + \frac{I_1(1)^2}{I_0(1)} \\
&= \frac{I_1^2}{I_0} + \frac{(\mu - I_1)^2}{1 - I_0} \\
&= \frac{I_1^2(1-I_0) + I_0\left(\mu^2 - 2\mu I_1 + I_1^2\right)}{I_0(1-I_0)} \\
&= \frac{I_1^2 - I_0 I_1^2 + \mu^2 I_0 - 2\mu I_0 I_1 + I_0 I_1^2}{I_0(1-I_0)} \\
&= \frac{\mu^2 - 2I_1\mu + I_1^2/I_0}{1-I_0}. \qquad (126)
\end{aligned}$$

Hence,

$$\begin{aligned}
\mathbb{E}_d \operatorname{Var}_{\theta|d}(\theta|d) &= \sigma^2 + \mu^2 - \frac{\mu^2 - 2I_1\mu + I_1^2/I_0}{1-I_0} \\
&= \sigma^2 - \frac{-\mu^2 + \mu^2 I_0 + \mu^2 - 2I_1\mu + I_1^2/I_0}{1-I_0} \\
&= \sigma^2 - \frac{\mu^2 I_0^2 - 2I_1 I_0\mu + I_1^2}{I_0(1-I_0)} \\
&= \sigma^2 - \frac{(I_1 - \mu I_0)^2}{I_0(1-I_0)}. \qquad (127)
\end{aligned}$$

□

It turns out that Eq. (115) can be simplified if we expressed the expected posterior variance in terms of the bias, which we now define. For a two-outcome likelihood function $\mathcal{L}(\theta; d)$, where $d \in \Omega = \{0, 1\}$, we define the *bias* to be

$$\Lambda(\theta) = 2\mathcal{L}(\theta; 0) - 1. \qquad (128)$$

This gives the following expression for the likelihood function in terms of the bias

$$\mathcal{L}(\theta; d) = \frac{1}{2}\left[1 + (-1)^d \Lambda(\theta)\right]. \qquad (129)$$

Let us now define

$$b = \int \mathrm{d}\theta\; p(\theta)\Lambda(\theta), \qquad (130)$$

$$\chi = \frac{1}{\sigma^2}\int \mathrm{d}\theta\; (\theta - \mu)p(\theta)\Lambda(\theta). \qquad (131)$$

We shall call b the *expected bias* and χ the *chi function*. It is straightforward to check that b and χ can be expressed in terms of I_0 and I_1, which were defined in Eq. (116), as follows:

$$b = 2I_0 - 1, \qquad (132)$$

$$\chi = \frac{2}{\sigma^2}(I_1 - \mu I_0). \qquad (133)$$

Substituting these expressions into Eq. (115) gives

$$\mathbb{E}_{\mathrm{d}} \operatorname{Var}_{\theta|\mathrm{d}}(\theta|d) = \sigma^2\left(1 - \sigma^2\mathcal{V}\right), \qquad (134)$$

with the variance reduction factor (113) given by

$$\mathcal{V} = \begin{cases} \frac{\chi^2}{1-b^2}, & |b| < 1, \\ 0, & |b| = 1. \end{cases} \qquad (135)$$

3.3 Two-outcome likelihood functions with a Gaussian prior

3.3.1 Variance reduction factor with a Gaussian prior

In this section, we fix the prior distribution to be the Gaussian distribution

$$p(\theta) = p(\theta;\mu,\sigma) := \frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(\theta-\mu)^2}{2\sigma^2}} \tag{136}$$

with prior mean $\mu \in \mathbb{R}$ and prior variance $\sigma^2 \in \mathbb{R}^+$. As before, we denote the bias by $\Lambda(\theta)$. The expected bias and the chi function are given by

$$b = b(\mu,\sigma) := \int_{-\infty}^{\infty} \mathrm{d}\theta\ p(\theta;\mu,\sigma)\Lambda(\theta), \tag{137}$$

$$\chi = \chi(\mu,\sigma) := \frac{1}{\sigma^2}\int_{-\infty}^{\infty} \mathrm{d}\theta\ (\theta-\mu)p(\theta;\mu,\sigma)\Lambda(\theta). \tag{138}$$

The Gaussian prior has the following nice property[4]: differentiating the expected bias with respect to the prior mean gives the chi function, i.e.

$$\chi(\mu,\sigma) = \frac{\partial}{\partial\mu}b(\mu,\sigma). \tag{139}$$

Next, note that $b(\mu,\sigma) < 1$ if and only if the bias $\Lambda(\theta)$ is neither the constant 1 function nor the constant -1 function (write this as $\Lambda \notin \{-1, 1\}$). This, together with Eq. (139), gives the following expression for the variance reduction factor defined in Eq. (135):

$$\mathcal{V} = \mathcal{V}(\mu,\sigma) := \frac{\partial_\mu b(\mu,\sigma)^2}{1-b(\mu,\sigma)^2}\mathbb{1}_{\Lambda\notin\{\pm 1\}} \tag{140}$$

where $\mathbb{1}_{\Lambda\notin\{\pm 1\}}$ denotes the indicator function which is equal to 1 when $\Lambda \notin \{\pm 1\}$ and 0 otherwise.

So far, we have assumed that the bias $\Lambda(\theta)$ is arbitrary. In the rest of this section, we will derive series expansions for $b(\mu,\sigma)$ and $\chi(\mu,\sigma)$ when $\Lambda(\theta)$ is expanded as a (Taylor or Fourier) series.

3.3.2 Bias as a Taylor series

We will now derive series expansions for $b(\mu,\sigma)$ and $\chi(\mu,\sigma)$ when $\Lambda(\theta)$ is written as a Taylor series at μ, i.e.

$$\Lambda(\theta) = \sum_{k=0}^{\infty}\Lambda_k(\theta-\mu)^k, \quad \text{where } \Lambda_k = \frac{1}{k!}\Lambda^{(k)}(\mu). \tag{141}$$

Substituting this into Eq. (137) gives

$$\begin{aligned} b\,(\mu,\sigma) &= \int_{-\infty}^{\infty} \mathrm{d}\theta\ p(\theta;\mu,\sigma)\Lambda(\theta) \\ &= \int_{-\infty}^{\infty} \mathrm{d}\theta\ \frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(\theta-\mu)^2}{2\sigma^2}}\sum_{k=0}^{\infty}\Lambda_k(\theta-\mu)^k \\ &= \sum_{k=0}^{\infty}\Lambda_k\int_{-\infty}^{\infty} \mathrm{d}\theta\ \frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(\theta-\mu)^2}{2\sigma^2}}(\theta-\mu)^k \\ &= \sum_{k=0}^{\infty}\Lambda_k(k-1)!!\sigma^k\mathbb{1}_{k\in 2\mathbb{Z}} \end{aligned} \tag{142}$$

[4] In fact, the Gaussian prior is the only (continuous) prior with this property. To see this, note that equating $\chi(\mu,\sigma) = \partial_\mu b(\mu,\sigma)$ gives the differential equation $\frac{1}{\sigma^2}(\theta-\mu)p(\theta;\mu,\sigma) = \partial_\mu p(\theta;\mu,\sigma)$. The unique solution to this equation with the normalization boundary condition $\int_{-\infty}^{\infty}\mathrm{d}\theta\ p(\theta;\mu,\sigma) = 1$ is the Gaussian distribution $p(\theta;\mu,\sigma)$ as given by Eq. (136).

where in the last line we used the following identity: for $\sigma > 0$ and $n \in \mathbb{N}$,

$$\int_{-\infty}^{\infty} \mathrm{d}\theta \ \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{(\theta-\mu)^2}{2\sigma^2}} (\theta - \mu)^n = (n-1)!!\sigma^n \mathbb{1}_{n\in 2\mathbb{Z}}, \tag{143}$$

where

$$n!! = \prod_{k=0}^{\lceil n/2 \rceil - 1} (n-2k) = n(n-2)(n-4)\dots \tag{144}$$

(with the empty product equal to 1) denotes the double factorial of n.

Substituting $\Lambda_k = \frac{1}{k!}\Lambda^{(k)}(\mu)$ into Eq. (142) gives

$$\begin{aligned} b(\mu,\sigma) &= \sum_{k=0}^{\infty} \frac{1}{k!}\Lambda^{(k)}(\mu)(k-1)!!\sigma^k \mathbb{1}_{k\in 2\mathbb{Z}} \\ &= \sum_{k=0}^{\infty} \frac{1}{k!!}\Lambda^{(k)}(\mu)\sigma^k \mathbb{1}_{k\in 2\mathbb{Z}}, \qquad \text{where we used } \frac{(k-1)!!}{k!} = \frac{1}{k!!} \\ &= \sum_{j=0}^{\infty} \Lambda^{(2j)}(\mu)\frac{\sigma^{2j}}{(2j)!!}. \end{aligned} \tag{145}$$

Differentiating this expression with respect to μ gives

$$\chi(\mu,\sigma) = \sum_{j=0}^{\infty} \Lambda^{(2j+1)}(\mu)\frac{\sigma^{2j}}{(2j)!!}. \tag{146}$$

Equivalently, the derivatives of the expected bias and the chi function are given as follows:

$$\left.\frac{\partial^k}{\partial\sigma^k} b(\mu,\sigma)\right|_{\sigma=0} = \Lambda^{(k)}(\mu)(k-1)!!\mathbb{1}_{k\in 2\mathbb{Z}}, \tag{147}$$

$$\left.\frac{\partial^k}{\partial\sigma^k} \chi(\mu,\sigma)\right|_{\sigma=0} = \Lambda^{(k+1)}(\mu)(k-1)!!\mathbb{1}_{k\in 2\mathbb{Z}}. \tag{148}$$

3.3.3 Bias as a trigonometric Fourier series

Next, we will derive series expansions for $b(\mu,\sigma)$ and $\chi(\mu,\sigma)$ when $\Lambda(\theta)$ is written as a trigonometric Fourier series with period 2π, i.e.

$$\Lambda(\theta) = \sum_{l=0}^{\infty} \left[c_l \cos(l\theta) + d_l \sin(l\theta)\right], \tag{149}$$

where

$$c_l = \frac{1}{(1+\delta_l)\,\pi} \int_{-\pi}^{\pi} \mathrm{d}\theta \ \Lambda(\theta)\cos(l\theta), \tag{150}$$

$$d_l = \frac{1}{\pi} \int_{-\pi}^{\pi} \mathrm{d}\theta \ \Lambda(\theta)\sin(l\theta), \tag{151}$$

where δ_l is the Kronecker delta which is equal to 1 if $l = 0$ and 0 otherwise.

Substituting Eq. (149) into Eq. (137) gives

$$\begin{aligned}
b(\mu,\sigma) &= \int_{-\infty}^{\infty} \mathrm{d}\theta\ p(\theta;\mu,\sigma)\Lambda(\theta) \\
&= \int_{-\infty}^{\infty} \mathrm{d}\theta\ \frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(\theta-\mu)^2}{2\sigma^2}} \sum_{l=0}^{-\infty} \left[c_l \cos(l\theta) + d_l \sin(l\theta)\right] \\
&= \sum_{l=0}^{\infty} \left[c_l \int_{\infty}^{\infty} \mathrm{d}\theta\ \frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(\theta-\mu)^2}{2\sigma^2}} \cos(l\theta) + d_l \int_{-\infty}^{\infty} \mathrm{d}\theta\ \frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(\theta-\mu)^2}{2\sigma^2}} \sin(l\theta)\right] \\
&= \sum_{l=0}^{\infty} \mathrm{e}^{-\frac{1}{2}l^2\sigma^2} \left[c_l \cos(l\mu) + d_l \sin(l\mu)\right],
\end{aligned} \tag{152}$$

where in the last line we used the following identities: for $\sigma > 0$ and $\mu, l \in \mathbb{R}$,

$$\int_{-\infty}^{\infty} \mathrm{d}\theta\ \frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(\theta-\mu)^2}{2\sigma^2}} \cos(l\theta) = \mathrm{e}^{-\frac{1}{2}l^2\sigma^2} \cos(l\mu), \tag{153}$$

$$\int_{-\infty}^{\infty} \mathrm{d}\theta\ \frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(\theta-\mu)^2}{2\sigma^2}} \sin(l\theta) = \mathrm{e}^{-\frac{1}{2}l^2\sigma^2} \sin(l\mu). \tag{154}$$

Differentiating Eq. (152) with respect to μ gives

$$\chi(\mu,\sigma) = \sum_{l=1}^{\infty} l\mathrm{e}^{-\frac{1}{2}l^2\sigma^2} \left[d_l \cos(l\mu) - c_l \sin(l\mu)\right]. \tag{155}$$

3.4 Limiting behavior of the expected posterior variance for small prior variance

3.4.1 Connection to Fisher information

Throughout this section, we will assume that the prior distribution $p(\theta;\mu,\sigma)$ is Gaussian and given by Eq. (136), and that the bias $\Lambda(\theta)$ can be expressed as the Taylor series given by Eq. (141). We will consider the limiting behavior of the expected posterior variance as σ vanishes and show the relationship between this quantity and Fisher information.

The Fisher information is commonly used to capture the power of a likelihood function in the estimation process [13]. The Fisher information is defined as

$$\mathcal{I}(\theta) = \mathbb{E}_d \left(\frac{\partial}{\partial\theta} \log \mathcal{L}(\theta;d)\right)^2. \tag{156}$$

Larger Fisher information indicates that the data is expected to be more informative of the value of the unknown parameter θ. In practice, however, we cannot engineer a likelihood function to maximize the value of the Fisher information because we do not know the actual value of θ. Nevertheless, we can show that the Fisher information is closely related to the variance reduction factor $\mathcal{V}$.

In the case of the two-outcome likelihood function, the Fisher information evaluates to

$$\begin{aligned}
\mathcal{I}(\theta) &= \mathcal{L}(\theta;0)\left(\frac{\mathcal{L}'(\theta;0)}{\mathcal{L}(\theta;0)}\right)^2 + \mathcal{L}(\theta;1)\left(\frac{\mathcal{L}'(\theta;1)}{\mathcal{L}(\theta;1)}\right)^2 \\
&= \frac{\left(\mathcal{L}'(\theta;0)\right)^2}{\mathcal{L}(\theta;0)(1-\mathcal{L}(\theta;0))}.
\end{aligned} \tag{157}$$

In terms of the bias of the likelihood function (128), we can express the Fisher information as

$$\mathcal{I}(\theta) = \frac{\Lambda'(\theta)^2}{1-\Lambda(\theta)^2}. \tag{158}$$

We now show that $\mathcal{V}$ can be approximated by the Fisher information when σ is small. Expanding the expression of the expected bias given by Eq. (145) gives

$$b(\mu,\sigma) = \Lambda(\mu) + \frac{1}{2}\Lambda^{(2)}(\mu)\sigma^2 + \frac{1}{8}\Lambda^{(4)}(\mu)\sigma^4 + O(\sigma^6) \tag{159}$$
$$\to \Lambda(\mu) \quad \text{as } \sigma \to 0. \tag{160}$$

Expanding the expression of the chi function given by Eq. (145) gives

$$\chi(\mu,\sigma) = \Lambda'(\mu) + \frac{1}{2}\Lambda^{(3)}(\mu)\sigma^2 + \frac{1}{8}\Lambda^{(5)}(\mu)\sigma^4 + O(\sigma^6) \tag{161}$$
$$\to \Lambda'(\mu) \quad \text{as } \sigma \to 0. \tag{162}$$

Hence, as long as[5] $|\Lambda(\mu)| \neq 1$ (which implies that $\Lambda \notin \{\pm 1\}$), the variance reduction factor (140) as σ goes to zero is equal to the Fisher information at $\theta = \mu$,

$$\lim_{\sigma\to 0} \mathcal{V}(\mu,\sigma) = \frac{\Lambda'(\mu)^2}{1-\Lambda(\mu)^2} = \mathcal{I}(\mu). \tag{163}$$

Therefore, using Eqs. (134), the expected posterior variance when σ is small is approximately linear in the Fisher information:

$$\mathbb{E}_{\mathrm{d}} \operatorname{Var}_{\theta|\mathrm{d}}(\theta|d) \approx \sigma^2 \left(1 - \frac{\sigma^2 \Lambda'(\mu)^2}{1-\Lambda(\mu)^2}\right)$$
$$= \sigma^2 \left(1 - \sigma^2 \mathcal{I}(\mu)\right), \quad \text{for } \sigma \approx 0. \tag{164}$$

3.4.2 Applying L'Hôpital's rule

Take $\Omega = \{0,1\}$ and the prior distribution to be the Gaussian distribution $p(\theta;\mu,\sigma)$ given by Eq. (136). In Section 3.4.1, we showed that if $|\Lambda(\mu)| \neq 1$, the variance reduction factor in the limit when $\sigma \to 0$ (by Eq. (140))

$$\lim_{\sigma\to 0} \mathcal{V}(\mu,\sigma) = \lim_{\sigma\to 0} \frac{\chi(\mu,\sigma)^2}{1-b(\mu,\sigma)^2} \mathbb{1}_{\Lambda\notin\{\pm 1\}} \tag{165}$$

is equal to the Fisher information at the prior mean μ (see Eq. (163)). In this section, we shall explore the case when $|\Lambda(\mu)| = 1$. In this case, L'Hôpital's rule may be used to evaluate the limit $\lim_{\sigma\to 0} \mathcal{V}(\mu,\sigma)$ as follows:

$$\lim_{\sigma\to 0} \mathcal{V}(\mu,\sigma) = \frac{\frac{\partial^w}{\partial\sigma^w}\chi(\mu,\sigma)^2\big|_{\sigma=0}}{\frac{\partial^w}{\partial\sigma^w}\left[1-b(\mu,\sigma)^2\right]\big|_{\sigma=0}} \mathbb{1}_{\Lambda\notin\{\pm 1\}} \tag{166}$$

where w is the minimum $w' \in \mathbb{N}$ such that

$$\left.\frac{\partial^{w'}}{\partial\sigma^{w'}}\chi(\mu,\sigma)^2\right|_{\sigma=0} \neq 0 \quad \text{or} \quad \left.\frac{\partial^{w'}}{\partial\sigma^{w'}}\left[1-b(\mu,\sigma)^2\right]\right|_{\sigma=0} \neq 0. \tag{167}$$

Note that w could be calculated using Eq. (168) and (169) from the following proposition.

Proposition 14.

$$\left.\frac{\partial^n}{\partial\sigma^n}\chi(\mu,\sigma)^2\right|_{\sigma=0} = \mathbb{1}_{n\in 2\mathbb{Z}} \sum_{k=0}^{n} \frac{n!}{k!!(n-k)!!} \Lambda^{(n-k+1)}(\mu)\Lambda^{(k+1)}(\mu)\mathbb{1}_{k\in 2\mathbb{Z}} \tag{168}$$

$$\left.\frac{\partial^n}{\partial\sigma^n}\left[1-b(\mu,\sigma)^2\right]\right|_{\sigma=0} = \delta_n - \mathbb{1}_{n\in 2\mathbb{Z}} \sum_{k=0}^{n} \frac{n!}{k!!(n-k)!!} \Lambda^{(n-k)}(\mu)\Lambda^{(k)}(\mu)\mathbb{1}_{k\in 2\mathbb{Z}}. \tag{169}$$

[5] When $|\Lambda(\mu)| \neq 1$ does not hold, Eq. (163) may not either. In this case, L'Hôpital's rule may be used to evaluate the limit $\lim_{\sigma\to 0} \mathcal{V}(\mu,\sigma)$. We study this case in detail in Section 3.4.2.

n	$\left.\frac{\partial^n}{\partial\sigma^n}\chi(\mu,\sigma)^2\right\|_{\sigma=0}$	$\left.\frac{\partial^n}{\partial\sigma^n}\left[1-b(\mu,\sigma)^2\right]\right\|_{\sigma=0}$
odd	0	0
0	Λ'^2	$1-\Lambda^2$
2	$2\Lambda'\Lambda^{(3)}$	$-2\Lambda\Lambda^{(2)}$
4	$6\Lambda^{(3)2}+6\Lambda'\Lambda^{(5)}$	$-6\Lambda^{(2)2}-6\Lambda\Lambda^{(4)}$
6	$90\Lambda^{(3)}\Lambda^{(5)}+30\Lambda'\Lambda^{(7)}$	$-90\Lambda^{(2)}\Lambda^{(4)}-30\Lambda\Lambda^{(6)}$

Table 1: Table of $\left.\frac{\partial^n}{\partial\sigma^n}\chi(\mu,\sigma)^2\right|_{\sigma=0}$ and $\left.\frac{\partial^n}{\partial\sigma^n}\left[1-b(\mu,\sigma)^2\right]\right|_{\sigma=0}$ values for odd n and $n=0,2,4,6$ calculated using Eqs. (168) and (169). These expressions can be used to determine the minimum w' for which Eq. (166) holds, which in turn can be used to calculate Eq. (167).

Proof. We first state a useful identity:

$$\frac{\partial^n}{\partial x^n}f(x)^2=\sum_{k=0}^{n}\binom{n}{k}f^{(n-k)}(x)f^{(k)}(x) \tag{170}$$

where $f^{(l)}(x)=\frac{\partial^l}{\partial x^l}f(x)$ denotes the lth derivative with respect to x.

To prove Eq. (168), we use Eqs. (148) and (170) to obtain

$$\begin{aligned}\left.\frac{\partial^n}{\partial\sigma^n}\chi(\mu,\sigma)^2\right|_{\sigma=0}&=\sum_{k=0}^{n}\binom{n}{k}\left(\left.\frac{\partial^{n-k}}{\partial\sigma^{n-k}}\chi(\mu,\sigma)\right|_{\sigma=0}\right)\left(\left.\frac{\partial^{k}}{\partial\sigma^{k}}\chi(\mu,\sigma)\right|_{\sigma=0}\right)\\&=\sum_{k=0}^{n}\binom{n}{k}\Lambda^{(n-k+1)}(\mu)(n-k-1)!!\mathbb{1}_{n-k\in2\mathbb{Z}}\Lambda^{(k+1)}(\mu)(k-1)!!\mathbb{1}_{k\in2\mathbb{Z}}\\&=\mathbb{1}_{n\in2\mathbb{Z}}\sum_{k=0}^{n}\frac{n!}{k!!(n-k)!!}\Lambda^{(n-k+1)}(\mu)\Lambda^{(k+1)}(\mu)\mathbb{1}_{k\in2\mathbb{Z}}.\end{aligned} \tag{171}$$

To prove Eq. (169), we use Eqs. (147) and (170) to obtain

$$\begin{aligned}\left.\frac{\partial^n}{\partial\sigma^n}\left[1-b(\mu,\sigma)^2\right]\right|_{\sigma=0}&=\delta_n-\sum_{k=0}^{n}\binom{n}{k}\left(\left.\frac{\partial^{n-k}}{\partial\sigma^{n-k}}b(\mu,\sigma)\right|_{\sigma=0}\right)\left(\left.\frac{\partial^{k}}{\partial\sigma^{k}}b(\mu,\sigma)\right|_{\sigma=0}\right)\\&=\delta_n-\sum_{k=0}^{n}\binom{n}{k}\Lambda^{(n-k)}(\mu)(n-k-1)!!\mathbb{1}_{n-k\in2\mathbb{Z}}\Lambda^{(k)}(\mu)(k-1)!!\mathbb{1}_{k\in2\mathbb{Z}}\\&=\delta_n-\mathbb{1}_{n\in2\mathbb{Z}}\sum_{k=0}^{n}\frac{n!}{k!!(n-k)!!}\Lambda^{(n-k)}(\mu)\Lambda^{(k)}(\mu)\mathbb{1}_{k\in2\mathbb{Z}}.\end{aligned} \tag{172}$$

□

Note that Proposition 14 implies that both $\left.\frac{\partial^n}{\partial\sigma^n}\chi(\mu,\sigma)^2\right|_{\sigma=0}$ and $\left.\frac{\partial^n}{\partial\sigma^n}\left[1-b(\mu,\sigma)^2\right]\right|_{\sigma=0}$ vanish when n is odd, which implies that the statement (167) is necessarily false for these values of n. Hence, the integer w in Eq. (166) must be even. A table of $\left.\frac{\partial^n}{\partial\sigma^n}\chi(\mu,\sigma)^2\right|_{\sigma=0}$ and $\left.\frac{\partial^n}{\partial\sigma^n}\left[1-b(\mu,\sigma)^2\right]\right|_{\sigma=0}$ values for some small values of n (namely, $n=0,2,4,6$) is given in Table 1.

As a consequence of Table 1 and Eqs. (166) and (167), we obtain, for example, the following limiting behaviors of $\mathcal{V}(\mu,\sigma)$ as $\sigma\to0$:

- If $|\Lambda(\mu)|=1$, $\Lambda'(\mu)=0$, and $\Lambda''(\mu)\neq0$, then

$$\lim_{\sigma\to0}\mathcal{V}(\mu,\sigma)=0. \tag{173}$$

- If $|\Lambda(\mu)| = 1$, $\Lambda'(\mu) = \Lambda''(\mu) = 0$, and ($\Lambda'''(\mu) \neq 0$ or $-\Lambda''^2(\mu) \mp \Lambda''''(\mu) \neq 0$), then

$$\lim_{\sigma \to 0} \mathcal{V}(\mu, \sigma) = \frac{\Lambda'''(\mu)^2}{-\Lambda''^2(\mu) \mp \Lambda''''(\mu)}. \tag{174}$$

- If $|\Lambda(\mu)| = 1$, $\Lambda'(\mu) = \Lambda''(\mu) = \Lambda'''(\mu) = 0$, $-\Lambda''^2(\mu) \mp \Lambda''''(\mu) = 0$, and ($3\Lambda'''(\mu)\Lambda''''(\mu) \neq 0$ or $-3\Lambda''(\mu)\Lambda''''(\mu) \mp 30\Lambda^{(6)}(\mu) \neq 0$), then

$$\lim_{\sigma \to 0} \mathcal{V}(\mu, \sigma) = \frac{3\Lambda'''(\mu)\Lambda'''''(\mu)}{-3\Lambda''(\mu)\Lambda''''(\mu) \mp 30\Lambda^{(6)}(\mu)}. \tag{175}$$

4 Engineered likelihood functions

In this section, we apply the tools that we developed in Section 3 to the quantum-generated likelihood functions from Section 2. The problem that we wish to solve may be phrased as an optimization problem, which we will state in Section 4.1 (see Eq. (185)). In Section 4.3, we numerically solve this optimization problem to compare the performance of various engineered likelihood functions with each other and with the fixed-angle Chebyshev likelihood functions.

4.1 Minimizing the expected posterior variance

Our goal is to *engineer* quantum likelihood functions by choosing appropriate tunable parameters $\vec{x} \in \mathbb{R}^{2L}$ in the parametrized quantum-generated likelihood functions given by Eq. (10). Specifically, we will choose these tunable parameters $\vec{x}$ to minimize the expected posterior variance (105) of the unknown parameter $\theta = \arccos(\langle\bar{0}| P |\bar{0}\rangle)$ given by Eq. (2). The likelihood functions that arise from such a minimization are called *engineered likelihood functions*.

We will take the prior distribution to be the Gaussian distribution with probability density function given by $p(\theta; \mu, \sigma)$ (see Eq. (136)) and the likelihood function to be the quantum-generated likelihood function $\mathcal{L}^{\mathcal{A}}(\theta; d, \vec{x}) = \frac{1}{2}\left[1 + (-1)^d \Lambda^{\mathcal{A}}(\theta; \vec{x})\right]$ given by Eq. (10). We will consider both the ancilla-free scheme ($\mathcal{A} =$ AF) and the ancilla-based scheme ($\mathcal{A} =$ AB).

To explicitly indicate dependence on the prior mean μ, the prior variance σ, the tunable parameters $\vec{x} \in \mathbb{R}^{2L}$ and the scheme $\mathcal{A}$, we shall denote the expected bias (137), the chi function (138) and the variance reduction factor (140) by

$$b^{\mathcal{A}}(\mu, \sigma; \vec{x}) = \int_{-\infty}^{\infty} \mathrm{d}\theta\ p(\theta; \mu, \sigma)\Lambda^{\mathcal{A}}(\theta; \vec{x}) \tag{176}$$

$$\chi^{\mathcal{A}}(\mu, \sigma; \vec{x}) = \frac{1}{\sigma^2}\int_{-\infty}^{\infty} \mathrm{d}\theta\ (\theta - \mu)p(\theta; \mu, \sigma)\Lambda^{\mathcal{A}}(\theta; \vec{x}) = \partial_\mu b^{\mathcal{A}}(\mu, \sigma; \vec{x}) \tag{177}$$

$$\mathcal{V}^{\mathcal{A}}(\mu, \sigma; \vec{x}) = \frac{\chi^{\mathcal{A}}(\mu, \sigma; \vec{x})^2}{1 - b^{\mathcal{A}}(\mu, \sigma; \vec{x})^2}\mathbb{1}_{\Lambda^{\mathcal{A}}(\theta; \vec{x}) \notin \{\pm 1\}} \tag{178}$$

respectively. By Eq. (134), the expected posterior variance may be expressed in terms of the variance reduction factor as

$$\mathbb{E}_{\mathrm{d}}\,\mathrm{Var}_{\theta|\mathrm{d}}(\theta|d; \mu, \sigma, \vec{x}, \mathcal{A}) = \sigma^2\left[1 - \sigma^2\mathcal{V}^{\mathcal{A}}(\mu, \sigma; \vec{x})\right]. \tag{179}$$

Next, we find series expansions for the functions (176)–(178). By Theorems 10 and 11, the biases $\Lambda^{\mathcal{A}}(\theta; \vec{x})$, for $\mathcal{A} \in \{\mathrm{AF}, \mathrm{AB}\}$, can be written as the cosine polynomials

$$\Lambda^{\mathcal{A}}(\theta; \vec{x}) = \sum_{l=0}^{q(\mathcal{A})} \mu_l^{\mathcal{A}}(\vec{x})\cos(l\theta) \tag{180}$$

where

$$q(\mathcal{A}) = \begin{cases} 2L+1, & \mathcal{A} = \mathrm{AF} \\ L, & \mathcal{A} = \mathrm{AB}, \end{cases} \tag{181}$$

and $\mu_l^{\mathcal{A}}(\vec{x})$'s are given by Eq. (65) and (90).

The series (180) allows us to use the machinery introduced in Section 3.3.3: by Eqs. (152) and (155), the bias and the chi function can be written as the sums

$$b^{\mathcal{A}}(\mu, \sigma; \vec{x}) = \sum_{l=0}^{q(\mathcal{A})} \mu_l^{\mathcal{A}}(\vec{x}) \mathrm{e}^{-l^2\sigma^2/2} \cos(l\mu), \tag{182}$$

$$\chi^{\mathcal{A}}(\mu, \sigma; \vec{x}) = -\sum_{l=1}^{q(\mathcal{A})} \mu_l^{\mathcal{A}}(\vec{x}) \mathrm{e}^{-l^2\sigma^2/2} l \sin(l\mu). \tag{183}$$

and the variance reduction factor may be written as

$$\mathcal{V}^{\mathcal{A}}(\mu, \sigma; \vec{x}) = \frac{\chi^{\mathcal{A}}(\mu, \sigma; \vec{x})^2}{1 - b^{\mathcal{A}}(\mu, \sigma; \vec{x})^2} = \frac{\left[\sum_{l=0}^{q(\mathcal{A})} \mu_l^{\mathcal{A}}(\vec{x}) \mathrm{e}^{-l^2\sigma^2/2} \cos(l\mu)\right]^2}{1 - \left[\sum_{l=1}^{q(\mathcal{A})} \mu_l^{\mathcal{A}}(\vec{x}) \mathrm{e}^{-l^2\sigma^2/2} l \sin(l\mu)\right]^2}. \tag{184}$$

Our goal is to find tunable parameters $\vec{x} = (x_1, \ldots, x_{2L}) \in \mathbb{R}^{2L}$ that minimize the expected posterior variance (179). Due to the inverse relationship between the expected posterior variance and the variance reduction factor (see Eq. (179)), minimizing the former is equivalent to maximizing the latter. Since $\mathcal{V}(\mu, \sigma; \vec{x})$ is 2π-periodic in each coordinate x_i, it suffices to restrict the search space of each x_i to $(-\pi, \pi]$. In other words, the optimization problem we wish to solve may be stated as:

$$\begin{aligned} &\texttt{Input:} \quad (\mu, \sigma, \mathcal{A}), \text{ where } \mu \in \mathbb{R}, \sigma > 0, \mathcal{A} \in \{\mathrm{AF}, \mathrm{AB}\} \\ &\texttt{Output:} \quad \underset{\vec{x} \in (-\pi, \pi]^{2L}}{\arg\max}\ \mathcal{V}^{\mathcal{A}}(\mu, \sigma; \vec{x}). \end{aligned} \tag{185}$$

Note that the input prior variance σ in the optimization problem (185) is required to be positive to guarantee that Eq. (184) is well-defined. For the case when $\sigma \to 0$, the results of Section 3.4 may be used. In particular, Eqs. (163) and (173) imply the following.

- If $|\Lambda^{\mathcal{A}}(\mu; \vec{x})| \neq 1$, then

$$\lim_{\sigma \to 0} \mathcal{V}^{\mathcal{A}}(\mu, \sigma; \vec{x}) = \frac{(\partial_\mu \Lambda^{\mathcal{A}}(\mu; \vec{x}))^2}{1 - \Lambda^{\mathcal{A}}(\mu; \vec{x})}. \tag{186}$$

- If $|\Lambda^{\mathcal{A}}(\mu; \vec{x})| = 1$, $\partial_\mu \Lambda^{\mathcal{A}}(\mu; \vec{x}) = 0$, and $\partial_\mu^2 \Lambda^{\mathcal{A}}(\mu; \vec{x}) \neq 0$, then

$$\lim_{\sigma \to 0} \mathcal{V}^{\mathcal{A}}(\mu, \sigma; \vec{x}) = 0. \tag{187}$$

We will present some numerical results obtained by solving the optimization problem (185) in Section 4.3. Before we do that, we study the behavior and properties of the variance reduction factor (184) in the special case where the angles $\vec{x}$ are chosen to give rise of Chebyshev likelihood functions.

4.2 Chebyshev variance reduction factor

The Fourier coefficients (102) allow us to compute the variance reduction factor (184) in the case when $\vec{x} = (\pi/2)^{2L}$. By substituting Eq. (102) into Eqs. (182) and (183), we find that

$$b^{\mathcal{A}}\left(\mu, \sigma; \left(\tfrac{\pi}{2}\right)^{2L}\right) = (-1)^r \mathrm{e}^{-q^2\sigma^2/2}\cos(q\mu), \tag{188}$$

$$\chi^{\mathcal{A}}\left(\mu, \sigma; \left(\tfrac{\pi}{2}\right)^{2L}\right) = -(-1)^r \mathrm{e}^{-q^2\sigma^2/2} q\sin(q\mu). \tag{189}$$

where

$$r = \begin{cases} 0, & \mathcal{A} = \mathrm{AF} \\ L, & \mathcal{A} = \mathrm{AB} \end{cases} \tag{190}$$

and $q = q(\mathcal{A})$ is given by Eq. (181).

Hence, substituting Eq. (188) and Eq. (189) into Eq. (184) gives the following expression for the variance reduction factor:

$$\mathcal{V}^{\mathcal{A}}\left(\mu, \sigma; \left(\tfrac{\pi}{2}\right)^{2L}\right) = \frac{q^2\sin^2(q\mu)}{\mathrm{e}^{q^2\sigma^2} - \cos^2(q\mu)}. \tag{191}$$

The following proposition lists a few useful properties of the variance reduction factor (191).

Proposition 15. *Let $L \in \mathbb{Z}^+$ and $\sigma > 0$. The variance reduction factor $\mathcal{V}^{\mathcal{A}}(\mu, \sigma; (\frac{\pi}{2})^{2L})$ given by Eq.* (191) *satisfies the following properties.*

1. *Periodicity (in μ) with period $\frac{\pi}{q}$: For all $\mu \in \mathbb{R}$,*

$$\mathcal{V}^{\mathcal{A}}\left(\mu + \tfrac{\pi}{q}, \sigma; (\tfrac{\pi}{2})^{2L}\right) = \mathcal{V}^{\mathcal{A}}\left(\mu, \sigma; (\tfrac{\pi}{2})^{2L}\right). \tag{192}$$

2. *Minimum: For all $\mu \in \mathbb{R}$,*

$$\mathcal{V}^{\mathcal{A}}\left(\mu, \sigma; (\tfrac{\pi}{2})^{2L}\right) \geq 0 \tag{193}$$

   with equality if and only if $\mu \in \frac{\pi}{q}\mathbb{Z}$.

3. *Maximum: For all $\mu \in \mathbb{R}$,*

$$\mathcal{V}^{\mathcal{A}}\left(\mu, \sigma; (\tfrac{\pi}{2})^{2L}\right) \leq q^2\mathrm{e}^{-q^2\sigma^2} \tag{194}$$

   with equality if and only if $\mu \in \frac{\pi}{2q}\mathbb{Z}_{\mathrm{odd}}$.

In the above, $q = q(\mathcal{A})$ is given by Eq. (181).

Proof.

1. This follows directly from Eq. (191) and the fact that $\sin^2(q\mu)$ and $\cos^2(q\mu)$ are both periodic and have period $\frac{\pi}{q}$.

2. The numerator $q^2\sin^2(q\mu)$ of Eq. (191) is clearly positive. The denominator $\mathrm{e}^{q^2\sigma^2} - \cos^2(q\mu)$ of Eq. (191) is positive since $\sigma > 0$. Equality holds if and only if the numerator $q^2\sin^2(q\mu) = 0$, which holds if and only if $\mu \in \frac{\pi}{q}\mathbb{Z}$.

3. By dividing both the numerator and denominator of Eq. (191) by $\sin^2(q\mu)$, the variance reduction factor (191) can be written as

$$\mathcal{V}^{\mathcal{A}}(\mu, \sigma; (\tfrac{\pi}{2})^{2L}) = \frac{q^2}{1 + \left(\mathrm{e}^{q^2\sigma^2} - 1\right)\csc^2(q\mu)} \tag{195}$$

whenever $\sin(q\mu) \neq 0$. Since $\csc^2(q\mu) \geq 1$,

$$\mathcal{V}^{\mathcal{A}}(\mu, \sigma; (\tfrac{\pi}{2})^{2L}) \leq \frac{q^2}{1 + e^{q^2\sigma^2} - 1} = q^2 e^{-q^2\sigma^2}. \tag{196}$$

Equality holds if and only if $\csc^2(q\mu) = 1$, which holds if and only if $\mu \in \frac{\pi}{2q}\mathbb{Z}_{\text{odd}}$.

□

Note that Eq. (194) can be used to give an upper bound for the variance reduction factor that is independent of L: since the function $q \mapsto q^2 e^{-q^2\sigma^2}$ achieves a maximum of $(e\sigma^2)^{-1}$ at $q = \frac{1}{\sigma}$, it follows that

$$\mathcal{V}^{\mathcal{A}}\left(\mu, \sigma; (\tfrac{\pi}{2})^{2L}\right) \leq \frac{1}{e\sigma^2}, \tag{197}$$

with equality if and only if $q = \frac{1}{\sigma}$ and $\mu \in \frac{\pi}{2q}\mathbb{Z}_{\text{odd}}$.

Finally, we conclude with a proposition that characterizes the limiting behavior of the Chebyshev variance reduction factor as $\sigma \to 0$:

Proposition 16. *Let $L \in \mathbb{Z}^+$ and $\mu \in \mathbb{R}$. Then,*

$$\lim_{\sigma \to 0} \mathcal{V}^{\mathcal{A}}(\mu, \sigma; (\tfrac{\pi}{2})^{2L}) = q^2 \mathbb{1}_{\mu \notin \frac{\pi}{q}\mathbb{Z}} \tag{198}$$

where $q = q(\mathcal{A})$ is given by Eq. (181).

Proof. We first note that the bias (98) and its first two derivatives are given by

$$\begin{aligned} \Lambda^{\mathcal{A}}\left(\theta; (\tfrac{\pi}{2})^{2L}\right) &= (-1)^r \cos(q\theta) \\ (\Lambda^{\mathcal{A}})'\left(\theta; (\tfrac{\pi}{2})^{2L}\right) &= -(-1)^r q \sin(q\theta) \\ (\Lambda^{\mathcal{A}})''\left(\theta; (\tfrac{\pi}{2})^{2L}\right) &= -(-1)^r q^2 \cos(q\theta) \end{aligned} \tag{199}$$

where q and r are given by Eq. (181) and (190) respectively.

If $\mu \notin \frac{\pi}{q}\mathbb{Z}$, then $\Lambda^{\mathcal{A}}(\mu; \vec{x}) \neq 1$. By using Eq. (186), we obtain

$$\lim_{\sigma \to 0} \mathcal{V}^{\mathcal{A}}(\mu, \sigma; (\tfrac{\pi}{2})^{2L}) = \frac{q^2 \sin^2(q\mu)}{1 - \cos^2(q\mu)} = q^2. \tag{200}$$

If $\mu \in \frac{\pi}{q}\mathbb{Z}$, it follows from the expressions in Eq. (199) that

$$\begin{aligned} \Lambda^{\mathcal{A}}\left(\theta; (\tfrac{\pi}{2})^{2L}\right) &\in \{1, -1\} \\ (\Lambda^{\mathcal{A}})'\left(\theta; (\tfrac{\pi}{2})^{2L}\right) &= 0 \\ (\Lambda^{\mathcal{A}})''\left(\theta; (\tfrac{\pi}{2})^{2L}\right) &\in \{q^2, -q^2\}. \end{aligned} \tag{201}$$

Therefore, in this case, we have

$$|\Lambda^{\mathcal{A}}\left(\theta; (\tfrac{\pi}{2})^{2L}\right)| = 1, \quad (\Lambda^{\mathcal{A}})'\left(\theta; (\tfrac{\pi}{2})^{2L}\right) = 0, \quad (\Lambda^{\mathcal{A}})''\left(\theta; (\tfrac{\pi}{2})^{2L}\right) \neq 0. \tag{202}$$

Hence, by Eq. (187), we obtain

$$\lim_{\sigma \to 0} \mathcal{V}^{\mathcal{A}}(\mu, \sigma; (\tfrac{\pi}{2})^{2L}) = 0, \tag{203}$$

which completes the proof of the proposition. □

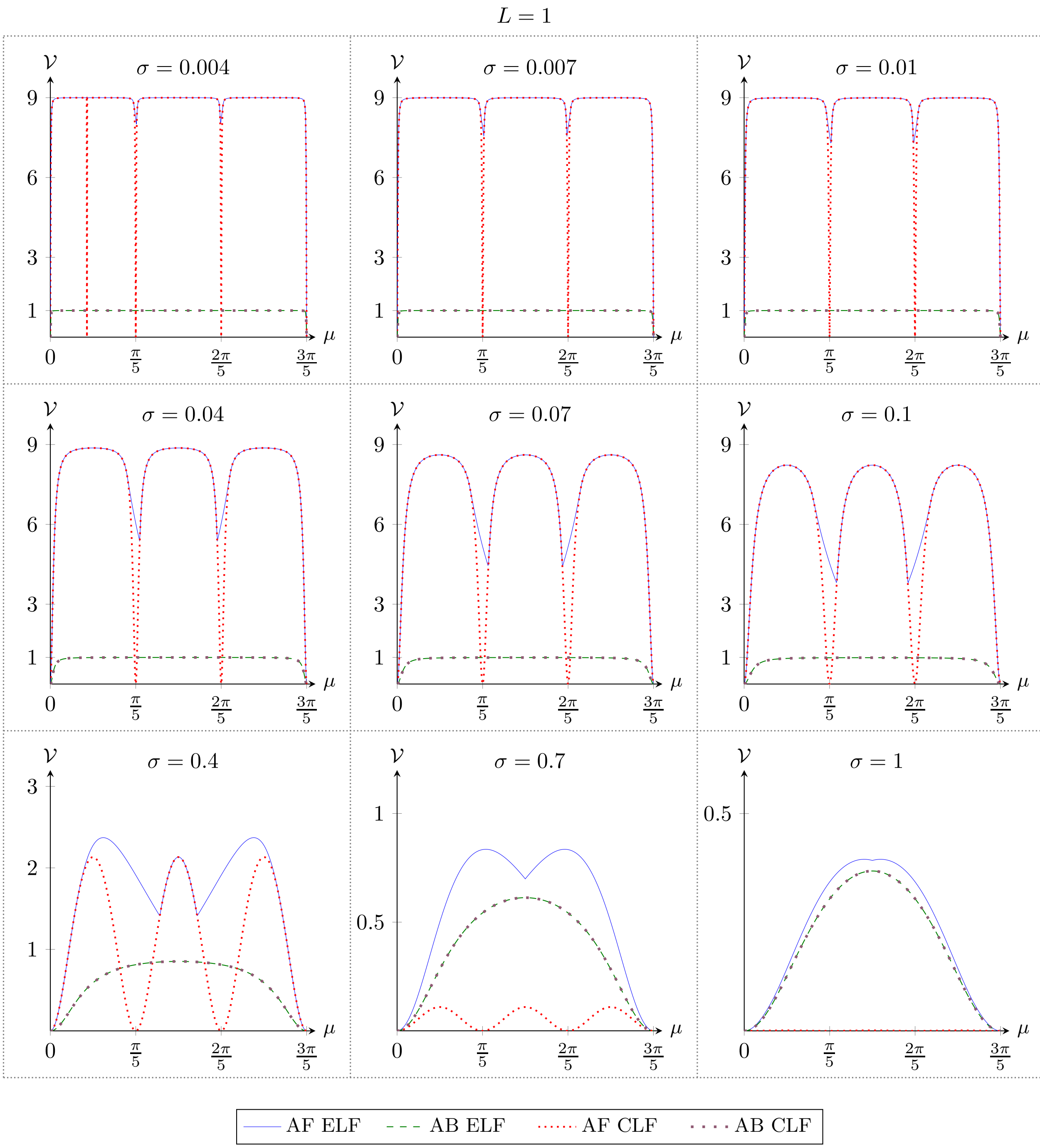


Figure 4.1: Plots of the variance reduction factor versus the prior mean μ for $L = 1$ for various prior variances σ. In each plot, the engineered likelihood function (ELF) is compared with the Chebyshev likelihood function (CLF) for both the ancilla-free (AF) and ancilla-based (AB) schemes.

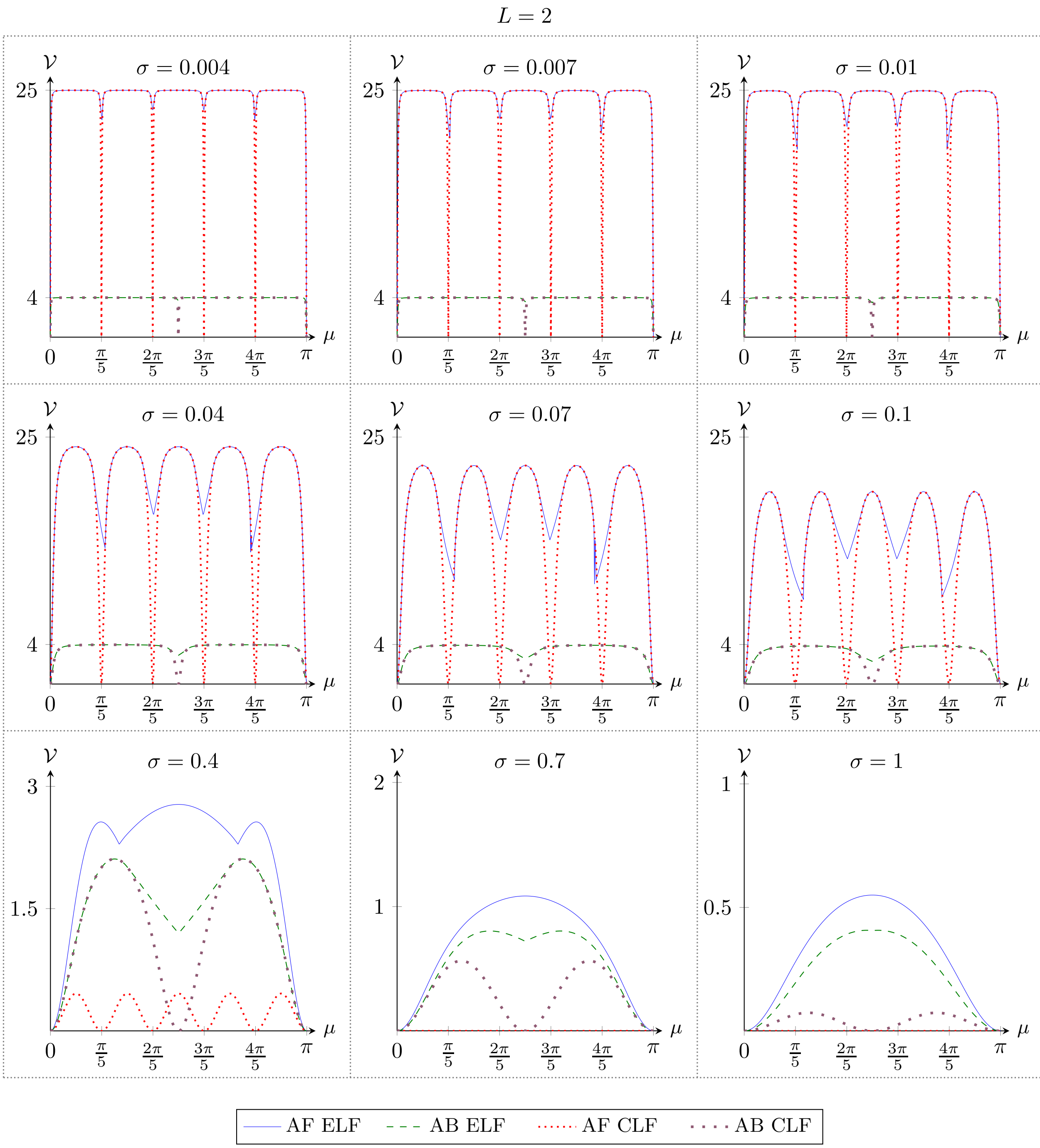


Figure 4.2: Plots of the variance reduction factor versus the prior mean μ for $L = 2$ for various prior variances σ. In each plot, the engineered likelihood function (ELF) is compared with the Chebyshev likelihood function (CLF) for both the ancilla-free (AF) and ancilla-based (AB) schemes.

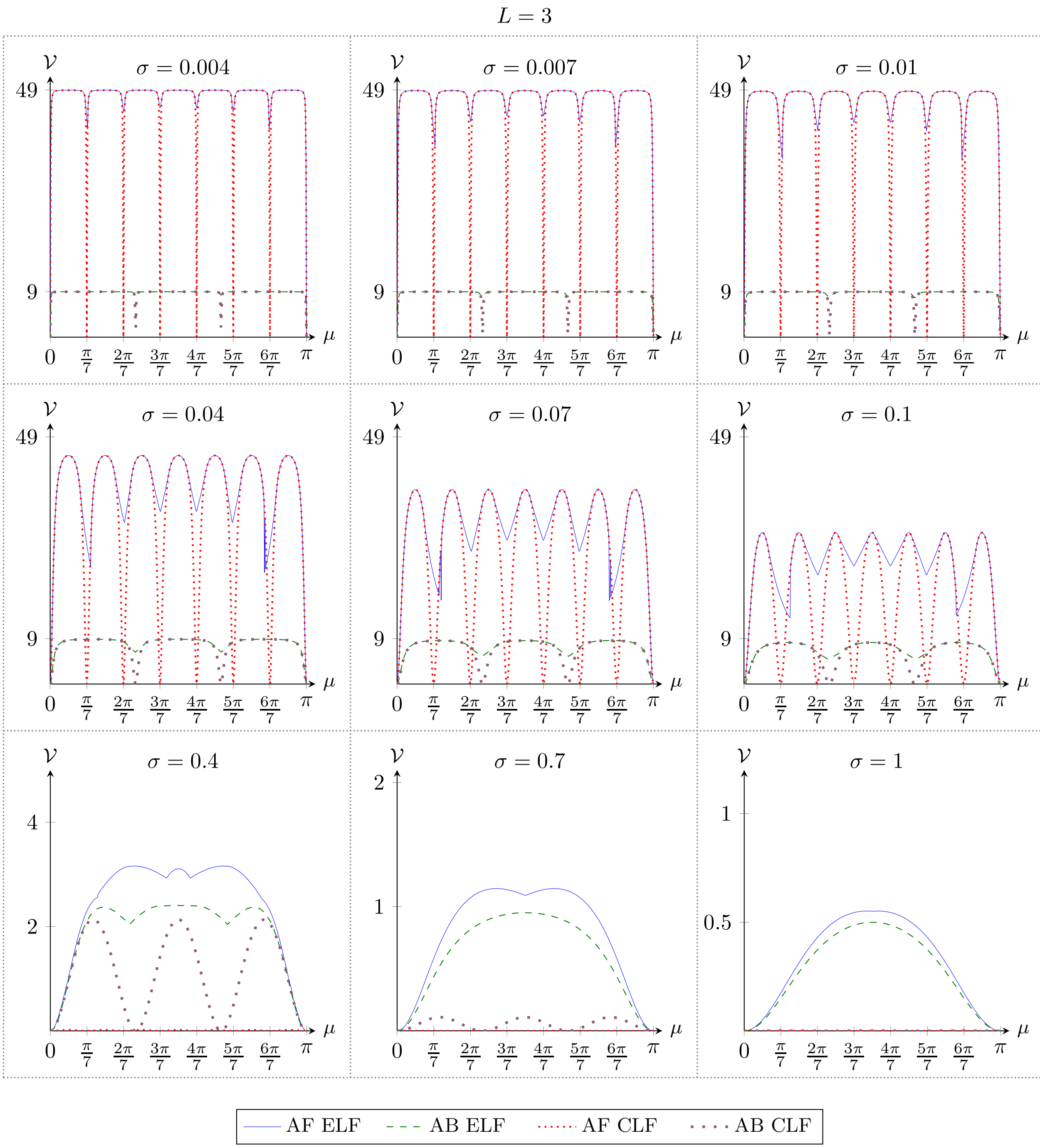


Figure 4.3: Plots of the variance reduction factor versus the prior mean μ for $L = 3$ for various prior variances σ. In each plot, the engineered likelihood function (ELF) is compared with the Chebyshev likelihood function (CLF) for both the ancilla-free (AF) and ancilla-based (AB) schemes.

4.3 Numerical simulations

The goal here is to solve the optimization problem (185). The objective function, which we seek to maximize, is the variance reduction factor, whose analytical expression is given by Eq. (184). To solve this optimization problem, we performed the following steps using Wolfram Mathematica [14]:

(i) compute the Fourier coefficients $\mu_l^{\mathcal{A}}(\vec{x})$'s using the expansion formulas given in Eq. (65) and (90).

(ii) compute the bias (182) and the chi function (183) and plug the resulting expressions into Eq. (184) for the variance reduction factor.

(iii) Feed the expression for the variance reduction factor into Mathematica's built-in optimization function `NMaximize` to find the parameters $(x_1, \ldots, x_{2L})$ that maximize $\mathcal{V}^{\mathcal{A}}(\mu, \sigma; \vec{x} = x_1, \ldots, x_{2L})$.

We present the results of this optimization in Figures 4.1–4.3, where we plot graphs of the variance reduction factor $\mathcal{V}^{\mathcal{A}}(\mu, \sigma; \vec{x})$ versus the prior mean μ for different values of σ for $L = 1, 2, 3$. In each graph, we compare the engineered likelihood function (ELF) with the Chebyshev likelihood function (CLF) for both the ancilla-free (AF) and ancilla-based (AB) schemes. As seen from the plots, in both the AF and AB schemes, there exist values of μ for which the ELFs outperform the CLFs (i.e. the ELF variance reduction factor is larger than the CLF variance reduction factor). This gap in performance decreases as σ goes to zero, suggesting that using ELFs would be most beneficial when the prior variance σ is large.

Another property of the plots is that the ancilla-free schemes yield much larger values for the expected posterior variance than for the ancilla-based case. This can be explained by the fact that as cosine polynomials, the ancilla-free bias has degree $2L + 1$ (see Eq. (64)), while the ancilla-based bias has only degree L (See Eq. (89)).

5 Concluding remarks

In this paper, we developed tools for characterizing and analyzing ELFs, focusing on the ancilla-based and ancilla-free schemes. Both these schemes involve alternate applications of generalized reflection operators, which may be visualized as rotations in a two-dimensional subspace. This visualization can be used to show that each of these schemes produces likelihood functions that can be written as cosine polynomials whose degree scales with the number of alternations. We showed that these polynomials can be used to derive analytical expressions for the expected posterior variance describing the parameter of interest. Finally, we presented simulation results to compare the performance of various ELFs with each other and to CLFs.

The results in this paper may be extended in a number of ways. Firstly, while we have limited the scope of this paper to only noiseless ELFs, the results here can be generalized to the case where the ELFs are noisy. This case—which arises when the states, transformations and measurements in circuits in Figure 2.1 are replaced by imperfect noisy versions of themselves, and which is arguably more relevant in this noisy intermediate-scale quantum [15] era, where near-term quantum devices are subject to high levels of noise—is treated in detail in a separate paper [12], which builds on the foundations laid here (see Appendix G for a brief discussion of incorporating noise into the ELF framework).

Secondly, while we have modeled the prior distribution by a Gaussian random variable (136), it might be appropriate in certain cases to model it using other distributions. We leave the sensitivity analysis on the prior distribution chosen to future work. Thirdly, while we have focused our attention on two specific schemes, namely the ancilla-based scheme and the ancilla-free scheme, we note that this framework can be extended to variants or extensions of these schemes. We leave this consideration for future work.

Acknowledgments

We thank Peter J. Love for helpful discussions.

A List of mathematical symbols

We list in this appendix some of the mathematical symbols that appear in this paper.

B A product-of-sums expansion

B.1 String reductions

Let $s, t \in \{p,q\}^*$ be strings. We say that s 1-*reduces* to t if there exist $t_1, t_2 \in \{p,q\}^*$ such that

1. $t = t_1 t_2$,
2. $s = t_1 pp t_2$ or $s = t_1 qq t_2$.

where juxtaposition of strings in the above notation indicates string concatenation.

For $k \geq 2$, we say that s k-*reduces* to t if there exist $u_1, u_2, \ldots, u_{k-1} \in \{p,q\}^*$ such that

1. s 1-reduces to u_1,
2. u_i 1-reduces to u_{i+1} for all $i = 1, \ldots, k-2$,
3. u_{k-1} 1-reduces to t.

Also, we say (trivially) that s 0-*reduces* to t if $s = t$.

Say that s *reduces* to t if there exists $k \in \mathbb{N}$ such that s k-reduces to t. In other words, s reduces to t if t can be obtained from s by repeatedly deleting substrings pp and qq. We write $s \to t$ if s reduces to t. For example, $pppqqq \to pq$.

We say that s is *irreducible* if there does not exist $r \in \{p,q\}^*$ such that $s \to r$ and $|r| < |s|$. We will make use of the following notation: for $u \in \mathbb{F}_2$, let

$$s^u = \begin{cases} \varepsilon & u = 0 \\ s & u = 1. \end{cases} \tag{204}$$

where ε is the empty string.

Define

$$t^{(ukv)} := p^u (qp)^k q. \tag{205}$$

It is straightforward to check that $t^{(ukv)}$ satisfies the following properties:

1. For all $u, v \in \mathbb{F}_2$ and for all $k \in \mathbb{N}$, the string $t^{(ukv)}$ is irreducible.
2. For all $s \in \{p,q\}^*$, there exist unique $u, v \in \mathbb{F}_2$ and $k \in \mathbb{N}$ such that $s \to t^{(ukv)}$.

Note that $|t^{(ukv)}| = u + v + 2k$. Write

$$s \sim t, \text{ if there exists } u \in \{p,q\}^* \text{ such that } s \to u \text{ and } t \to u. \tag{206}$$

It is easy to see that for each $n \in \mathbb{N}$, the relation $\sim$ is an equivalence relation on the set of strings $\{p,q\}^n$. For example, $pqqpq \sim ppqpp$ since they both reduce to q. Note that $\sim$ is preserved by string reversal, i.e.

$$s \sim t \iff s^R \sim t^R. \tag{207}$$

Let

$$\mathcal{M}_{ukv} = \{s \in \{p,q\}^* : s \sim t^{(ukv)}\} \tag{208}$$

denote the set of strings s that reduce to $t^{(ukv)}$. Note that $\mathcal{M}_{ukv}$ forms a partition of $\{p,q\}^*$.

Let

$$r^{(pq)} : \mathbb{F}_2^n \to \{p,q\}^*$$
$$(x_1,\dots,x_n) \mapsto r_n r_{n-1}\dots r_2 r_1 \tag{209}$$

where

$$r_i = \begin{cases} \varepsilon & x_i = 0 \\ p & x_i = 1, i \text{ even} \\ q & x_i = 1, i \text{ odd} \end{cases} \tag{210}$$
$$= \begin{cases} p^{x_i} & i \text{ even}, \\ q^{x_i} & i \text{ odd}. \end{cases} \tag{211}$$

In other words,

$$r^{(pq)}(x_1,\dots,x_n) = r^{x_n}\dots p^{x_2}q^{x_1}, \tag{212}$$

where $r = p$ if k is even, and $r = q$ otherwise. Note that $|r^{(pq)}(x)| = \mathrm{wt}(x)$.

B.2 The set Θ^n_{ukv}

We are now ready to define the set Θ^n_{ukv}. Let $n \in \mathbb{Z}^+$, $k \in \mathbb{N}$ and $u,v \in \mathbb{F}_2$. Define

$$\Theta^n_{ukv} = \{\vec{x} \in \mathbb{F}_2^n : r^{(pq)}(x) \in \mathcal{M}_{ukv}\} \tag{213}$$
$$= \{\vec{x} \in \mathbb{F}_2^n : r^{(pq)}(x) \to t^{(ukv)}\}. \tag{214}$$

In other words, Θ^n_{ukv} is the set of strings $\vec{x} = x_1x_2\dots x_n \in \mathbb{F}_2^n$ for which the string $r^{x_n}\dots p^{x_4}q^{x_3}p^{x_2}q^{x_1}$ reduces to $p^u(qp)^kq^v$, where $r = p$ if n is even and $r = q$ if n is odd. By convention, for $k \notin \mathbb{N}$, we take $\Theta^n_{ukv} = \emptyset$.

The following theorem characterizes the set of k values for which Θ^n_{ukv} is nonempty.

Theorem 17. *Let $n \in \mathbb{Z}^+$, $u,v \in \mathbb{F}_2$ and $k \in \mathbb{N}$. Then,*

$$\Theta^n_{ukv} \neq \emptyset \iff k \leq \left\lfloor \frac{n-1}{2} \right\rfloor - u\mathbb{1}_{n\in 2\mathbb{Z}+1}$$
$$= \begin{cases} \frac{n}{2} - 1 & n \text{ even}, \\ \frac{n-1}{2} - u & n \text{ odd}. \end{cases} \tag{215}$$

Proof. We will consider the even and odd cases of n separately.

Case 1: $n = 2m$ is even, where $m \in \mathbb{Z}^+$

($\Rightarrow$) Assume that $\Theta^n_{ukv} \neq \emptyset$. Then let $\vec{x} \in \Theta^{2m}_{ukv}$, i.e.

$$p^{x_{2m}}q^{x_{2m-1}}\dots p^{x_2}q^{x_1} \to p^u(qp)^kq^v, \tag{216}$$

which implies that

$$p^{x_{2m}+u}q^{x_{2m-1}}\dots p^{x_2}q^{x_1+v} \to (qp)^k. \tag{217}$$

- First, we show that $k \leq m$. By Eq. (217),

$$|(qp)^k| \leq |p^{x_{2m}+u}q^{x_{2,-1}}\dots p^{x_2}q^{x_1+v}| \tag{218}$$
$$\implies 2k \leq 2m \tag{219}$$
$$\implies k \leq m. \tag{220}$$

- Second, we show that $k \neq m$. Suppose, for the sake of contradiction, that $k = m$. By Eq. (217),

$$p^{x_2 m+u} q^{x_{2m-1}} \dots p^{x_2} q^{x_1+v} \to (qp)^m \tag{221}$$

which implies that

$$|p^{x_{2m}+u} q^{x_{2m-1}} \dots p^{x_2} q^{x_1+v}| \geq 2m. \tag{222}$$

But

$$\begin{aligned} &|p^{x_{2m}+u} q^{x_{2m-1}} \dots p^{x_2} q^{x_1+v}| \\ &= |\{i \in [2m] : x_{2m} + u = x_1 + v = 1, x_{2m=1} = \dots = x_2 = 1\}| \\ &\leq 2m \end{aligned} \tag{223}$$

By Eqs. (222) and (223), we get

$$|p^{x_{2m}+u} q^{x_{2m-1}} \dots p^{x_2} q^{x_1+v}| = 2m \tag{224}$$

which implies that

$$x_{2m} \neq u, x_1 \neq v, x_{2m-1} = \dots = x_2 = 1 \tag{225}$$

Substituting this into Eq. (217) gives

$$(pq)^m \to (qp)^m, \tag{226}$$

which is a contradiction. Hence, $k \neq m$.

The above two bullet points imply that

$$k \leq m - 1 = \frac{n}{2} - 1. \tag{227}$$

($\Leftarrow$) Let $k \leq \frac{n}{2} - 1 = m - 1$, and consider $\vec{x} = v1^{2k}u0^{2(m-k-1)} \in \{0,1\}^{2m}$. Note that $\vec{x}$ is a well-defined string since $m - k - 1 \geq 0$. Also, note that

$$r^{(pq)}(\vec{x}) = p^u (qp)^k q^v = t^{(ukv)} \tag{228}$$

is irreducible. Hence, $\vec{x} \in \Theta^{2m}_{ukv} = \Theta^n_{ukv}$, which implies that

$$\Theta^n_{ukv} \neq \emptyset. \tag{229}$$

<u>Case 2: $n = 2m + 1$ is odd, where $m \in \mathbb{N}$</u>

($\Rightarrow$) Assume that $\Theta^n_{ukv} \neq \emptyset$. Then there exists some $x \in \Theta^{2m+1}_{ukv}$, i.e.

$$q^{x_{2m+1}} p^{x_{2m}} \dots p^{x_2} q^{x_1} \to p^u (qp)^k q^v \tag{230}$$

- <u>Case: $u = 0$</u>
  By Eq. (230),

$$\begin{aligned} &q^{x_{2m+1}} p^{x_{2m}} \dots p^{x_2} q^{x_1} \to (qp)^k q^v \\ \implies &q^{x_{2m+1}} p^{x_{2m}} \dots p^{x_2} q^{x_1+v} \to (qp)^k \\ \implies &|q^{x_{2m+1}} p^{x_{2m}} \dots p^{x_2} q^{x_1+v}| \to |(qp)^k| \\ \implies &2m + 1 \geq 2k \\ \implies &k \leq m + \frac{1}{2}. \end{aligned} \tag{231}$$

  Since $m \in \mathbb{Z}^+$, $k \leq m = \frac{n-1}{2}$.

- Case: $u = 1$
  - First, we show that $k \leq m$. By Eq. (230),

$$\begin{aligned} & q^{x_{2m+1}} p^{x_{2m}} \dots p^{x_2} q^{x_1} \to p(qp)^k q^v \\ \implies & q^{x_{2m+1}} p^{x_{2m}} \dots p^{x_2} q^{x_1+v} \to p(qp)^k \\ \implies & |q^{x_{2m+1}} p^{x_{2m}} \dots p^{x_2} q^{x_1+v}| \geq |p(qp)^k| \\ \implies & 2m+1 \geq 2k+1 \\ \implies & k \leq m. \end{aligned} \tag{232}$$

  - Second, we show that $k \neq m$. Suppose, for the sake of contradiction, $k = m$. Then, By Eq. (230),

$$q^{x_{2m+1}} p^{x_{2m}} \dots p^{x_2} q^{x_1+v} \to p(qp)^m \tag{233}$$

$$\implies |q^{x_{2m+1}} p^{x_{2m}} \dots p^{x_2} q^{x_1+v}| = 2m+1 \tag{234}$$

    which implies that

$$x_{2m+1} = x_{2m} = \dots = x_2 = x_1 + v = 1 \tag{235}$$

    Substituting this into Eq. (233), we obtain

$$q(pq)^m \to p(qp)^m, \tag{236}$$

    which is a contradiction. Hence, $k \leq m - 1 = \frac{n-1}{2} - 1$.

($\Leftarrow$) Let $k \leq \frac{n-1}{2} - u = m - u$. Define the string

$$\vec{x} = \begin{cases} v1^{2k}0^{2(m-k)} & u = 0 \\ v1^{2k}u0^{2(m-k)-1} & u = 1. \end{cases} \tag{237}$$

Note that $\vec{x}$ is a well-defined string in $\{0,1\}^{2m+1}$, i.e. each of the substrings in the definition (237) have non-negative length. To see this, note that when $u = 0$, $m - k \geq m - (m - u) = 0$; and when $u = 1$, $2(m-k) - 1 \geq 2(m - (m-u)) - 1 = 2u - 1 = 1$.

Therefore,

$$r^{(pq)}(\vec{x}) = \begin{cases} (qp)^k q^v & u = 0 \\ p(qp)^k q^v & u = 1 \end{cases} \tag{238}$$

$$= p^u (qp)^k q^v \tag{239}$$

$$\to p^u (qp)^k q^v. \tag{240}$$

Hence,

$$\vec{x} \in \Theta^{2m+1}_{ukv} = \Theta^n_{ukv}. \tag{241}$$

This implies that $\Theta^n_{ukv} \neq \emptyset$.

□

B.3 Product-of-sums expansion formula

We start by proving the following lemma.

Lemma 18. *Let $n \in \mathbb{Z}^+$ and $P^2 = Q^2 = I$. Let*

$$R_y = \begin{cases} P & y \text{ even} \\ Q & y \text{ odd} \end{cases} \tag{242}$$

Then for all $\vec{x} \in \Theta^n_{ukv}$,

$$R_n^{x_n} R_{n-1}^{x_{n-1}} \dots R_1^{x_1} = P^u (QP)^k Q^v. \tag{243}$$

Proof. Consider

$$R_n^{x_n} R_{n-1}^{x_{n-1}} \dots R_1^{x_1} = R_n^{x_n} \dots P^{x_4} Q^{x_3} P^{x_2} Q^{x_1} = r^{(PQ)}(\vec{x}). \tag{244}$$

□

Lemma 19. *Let $n \in \mathbb{Z}^+$ and $P^2 = Q^2 = I$. Let $\{a^y_x : x \in \mathbb{F}_2,\ y \in [n]\} \subset \mathbb{C}$. Then,*

$$\begin{aligned} &(a^n_0 + a^n_1 R) \dots \left(a^4_0 + a^4_1 P\right)\left(a^3_0 + a^3_1 Q\right)\left(a^2_0 + a^2_1 P\right)\left(a^1_0 + a^1_1 Q\right) \\ &= \sum_{k=0}^{\infty} \sum_{u,v \in \mathbb{F}_2} \left(\sum_{\vec{x} \in \Theta^n_{ukv}} a^1_{x_1} a^2_{x_2} \dots a^n_{x_n} \right) P^u (QP)^k Q^v \end{aligned} \tag{245}$$

where $R = P$ if n is even and $R = Q$ if n is odd.

Proof. Consider

$$\begin{aligned} &(a^n_0 + a^n_1 R) \dots \left(a^4_0 + a^4_1 P\right)\left(a^3_0 + a^3_1 Q\right)\left(a^2_0 + a^2_1 P\right)\left(a^1_0 + a^1_1 Q\right) \\ &= \prod_{y=n}^{1} (a^y_0 + a^y_1 R_y), \quad R_y = \begin{cases} P & y \text{ even} \\ Q & y \text{ odd} \end{cases} \\ &= \prod_{y=n}^{1} \sum_{x \in \mathbb{F}_2} a^y_x R^x_y \\ &= \left(\sum_{x_n \in \mathbb{F}_2} a^n_{x_n} R_n^{x_n} \right) \dots \left(\sum_{x_1 \in \mathbb{F}_2} a^1_{x_1} R_1^{x_1} \right) \\ &= \sum_{\vec{x} \in \mathbb{F}_2^n} a^1_{x_1} \dots a^n_{x_n} R_n^{x_n} \dots R_x^{x_1} \\ &= \sum_{k=0}^{\infty} \sum_{u,v \in \mathbb{F}_2} \sum_{\vec{x} \in \Theta^n_{ukv}} a^1_{x_1} a^2_{x_2} \dots a^n_{x_n} R_n^{x_n} \dots R_1^{x_1} \\ &= \sum_{k=0}^{\infty} \sum_{u,v \in \mathbb{F}_2} \left(\sum_{\vec{x} \in \Theta^n_{ukv}} a^1_{x_1} a^2_{x_2} \dots a^n_{x_n} \right) P^u (QP)^k Q^v. \end{aligned} \tag{246}$$

□

Note that by Theorem 17, Eq. (245) can also be written as

$$\begin{aligned} &(a^n_0 + a^n_1 R) \dots \left(a^4_0 + a^4_1 P\right)\left(a^3_0 + a^3_1 Q\right)\left(a^2_0 + a^2_1 P\right)\left(a^1_0 + a^1_1 Q\right) \\ &= \sum_{u,v \in \mathbb{F}_2} \sum_{k=0}^{\lfloor (n-1)/2 \rfloor - u \mathbb{1}_{n \in 2\mathbb{Z}+1}} \left(\sum_{\vec{x} \in \Theta^n_{ukv}} a^1_{x_1} a^2_{x_2} \dots a^n_{x_n} \right) P^u (QP)^k Q^v. \end{aligned} \tag{247}$$

C Cardinalities of Θ^n_{ukv} and Ξ^α_l

What is the space complexity of storing each of the Fourier coefficients, of say, Eq. (51) and (65)? To address this question, it suffices to find the cardinality of the sets Ξ^α_l.

To begin, we prove the following theorem, which gives the cardinality of the set Θ^n_{ukv}.

Theorem 20. *Let $n \in \mathbb{Z}^+$, $u, v \in \mathbb{F}_2$, $k \in \{0, 1, 2, \ldots, \lfloor (n-1)/2 \rfloor - u\mathbb{1}_{n\in 2\mathbb{Z}+1}\}$. Then,*

$$|\Theta^n_{ukv}| = \binom{n-1}{\lfloor \frac{n-1}{2} \rfloor - u\mathbb{1}_{n\in 2\mathbb{Z}+1} - k}, \tag{248}$$

i.e. for $m \in \mathbb{Z}^+$,

$$\left|\Theta^{2m}_{ukv}\right| = \binom{2m-1}{m-1-k} = \binom{2m-1}{m+k}, \tag{249}$$

$$\left|\Theta^{2m+1}_{ukv}\right| = \binom{2m}{m-u-k} = \binom{2m}{m+u+k}. \tag{250}$$

Proof. We consider the odd and even cases separately.

- Case: $n = 2m$ even

  We first prove the following identity: for $m \in \mathbb{Z}^+$, if $P^2 = Q^2 = I$, then

$$[(I+P)(I+Q)]^m = \sum_{k=0}^{m-1} \binom{2m-1}{m+k} \left[\sum_{u,v\in\mathbb{F}_2} P^u (QP)^k Q^v \right]. \tag{251}$$

  To prove Eq. (251), we make use of mathematical induction. The base case $m = 1$ is clearly true. Assume that Eq. (251) holds for m. Then, by the induction hypothesis,

$$\begin{aligned}
[(I+P)(I+Q)]^{m+1} &= (I+P+Q+PQ) \sum_{k=0}^{m-1} \binom{2m-1}{m+k} \left[\sum_{u,v\in\mathbb{F}_2} P^u (QP)^k Q^v \right] \\
&= \sum_{k=0}^{m-1} \binom{2m-1}{m+k} (I+P+Q+PQ) \left[(QP)^k + P(QP)^k + (QP)^k Q + P(QP)^k Q\right] \\
&= \sum_{k=1}^{m-1} 2\binom{2m-1}{m+k} \left((QP)^k + P(QP)^k + (QP)^k Q + P(QP)^k Q\right) \\
&\quad + \sum_{k=1}^{m-1} \binom{2m-1}{m+k} \left((QP)^{k-1} + P(QP)^{k-1} + (QP)^{k-1} Q + P(QP)^{k-1} Q\right) \\
&\quad + \sum_{k=1}^{m-1} \binom{2m-1}{m+k} \left((QP)^{k+1} + P(QP)^{k+1} + (QP)^{k+1} Q + P(QP)^{k+1} Q\right) \\
&\quad + \binom{2m-1}{m} [3(I+P+Q+PQ) + (QP + PQP + QPQ + PQPQ)] \\
&= \sum_{k=0}^{m-2} \left[2\binom{2m-1}{m+k} + \binom{2m-1}{m+k+1} + \binom{2m-1}{m+k-1} \right] \\
&\quad \times \left((QP)^k + P(QP)^k + (QP)^k Q + P(QP)^k Q\right) \\
&\quad + (2n+3) \left((QP)^{m-1} + P(QP)^{m-1} + (QP)^{m-1} Q + P(QP)^{m-1} Q\right) \\
&\quad + (QP)^m + P(QP)^m + (QP)^m Q + P(QP)^m Q.
\end{aligned} \tag{252}$$

By Pascal's rule,

$$2\binom{2m-1}{m+k} + \binom{2m-1}{m+k+1} + \binom{2m-1}{m+k-1} = \binom{2m+1}{m+k+1}, \tag{253}$$

and hence,

$$\begin{aligned}
[(I+P)(I+Q)]^{m+1} &= \sum_{k=0}^{m} \binom{2m+1}{m+k+1} ((QP)^k + P(QP)^k + (QP)^k Q + P(QP)^k Q \\
&= \sum_{k=0}^{m} \binom{2m+1}{m+1+k} \left[\sum_{u,v\in\mathbb{F}_2} P^u (QP)^k Q^v \right],
\end{aligned} \tag{254}$$

which completes the inductive step and the proof of Eq. (251).

Next, by setting $n = 2m$ and $a^x_y = 1$ for all x, y in Eq. (247), we get

$$[(I+P)(I+Q)]^m = \sum_{k=0}^{m-1} \left[\sum_{u,v\in\mathbb{F}_2} \left|\Theta^{2m}_{ukv}\right| P^u (QP)^k Q^v \right]. \tag{255}$$

Comparing Eqs. (251) and (255) gives

$$\left|\Theta^{2m}_{ukv}\right| = \binom{2m-1}{m+k}. \tag{256}$$

- <u>Case: $n = 2m+1$ odd</u>

  We first prove the following identity: for $m \in \mathbb{N}$, if $P^2 = Q^2 = I$, then

$$(I+Q)\,[(I+P)(I+Q)]^m = \sum_{u,v\in\mathbb{F}_2} \sum_{k=0}^{m-u} \binom{2m}{m+u+k} P^u (QP)^k Q^v. \tag{257}$$

To prove Eq. (257), our starting point is Eq. (251). By multiplying $I+Q$ on the left of Eq. (251), we obtain

$$\begin{aligned}
&(I+Q)\,[(I+P)(I+Q)]^m \\
&= (I+Q) \sum_{k=0}^{m-1} \binom{2m-1}{m+k} \left[\sum_{u,v\in\mathbb{F}_2} P^u (QP)^k Q^v \right] \\
&= \sum_{k=0}^{m-1} \binom{2m-1}{m+k} \left[\sum_{u,v\in\mathbb{F}_2} P^u (QP)^k Q^v \right] + \underbrace{\sum_{k=0}^{m-1} \binom{2m-1}{m+k} \underbrace{\sum_{u,v\in\mathbb{F}_2} Q P^u (QP)^k Q^v}_{\mathscr{J}_k}}_{①}.
\end{aligned} \tag{258}$$

Now, by expanding the sum $\mathscr{J}_k$, we obtain

$$\mathscr{J}_k = \sum_{v\in\mathbb{F}_2} Q(QP)^k Q^v + (QP)^{k+1} Q^v. \tag{259}$$

When $k = 0$, Eq. (259) evaluates to

$$\mathscr{J}_0 = \sum_{v\in\mathbb{F}_2} Q^v + (QP) Q^v. \tag{260}$$

When $k \geq 1$, Eq. (259) evaluates to

$$\mathscr{I}_k = \sum_{v \in \mathbb{F}_2} P(QP)^{k-1} Q^v + (QP)^{k+1} Q^v. \tag{261}$$

Substituting these expressions for $\mathscr{I}_k$ into ① gives

$$\begin{aligned}
① &= \sum_{k=0}^{m-1} \binom{2m-1}{m+k} \mathscr{I}_k \\
&= \binom{2m-1}{m} \mathscr{I}_0 + \sum_{k=1}^{m-1} \binom{2m-1}{m+k} \mathscr{I}_k \\
&= \binom{2m-1}{m} \sum_{v \in \mathbb{F}_2} [Q^v + (QP)Q^v] + \sum_{k=1}^{m-1} \binom{2m-1}{m+k} \sum_{v \in \mathbb{F}_2} [P(QP)^{k-1} Q^v + (QP)^{k+1} Q^v] \\
&= \sum_{v \in \mathbb{F}_2} \left\{ \binom{2m-1}{m} (Q^v + QPQ^v) + \sum_{k=0}^{m-2} \binom{2m-1}{m+k+1} P(QP)^k Q^v + \sum_{k=2}^{m} \binom{2m-1}{m+k-1} (QP)^k Q^v \right\} \\
&= \sum_{v \in \mathbb{F}_2} \left[\sum_{k=0}^{m-2} \binom{2m-1}{m+k+1} P(QP)^k Q^v + \sum_{k=0}^{m} \binom{2m-1}{m+k-1} (QP)^k Q^v \right],
\end{aligned} \tag{262}$$

where the last line follows from the fact that

$$\binom{2m-1}{m} (Q^v + QPQ^v) = \binom{2m-1}{m-1} (QP)^0 Q^v + \binom{2m-1}{m} (QP) Q^v. \tag{263}$$

Substituting Eq. (262) into Eq. (258) gives

$$\begin{aligned}
&(I+Q)\left[(I+P)(I+Q)\right]^m \\
&= \sum_{v \in \mathbb{F}_2} \left\{ \sum_{k=0}^{m-1} \binom{2m-1}{m+k} \left[(QP)^k Q^v + P(QP)^k Q^v\right] \right. \\
&\quad \left. + \sum_{k=0}^{m-2} \binom{2m-1}{m+k+1} P(QP)^k Q^v + \sum_{k=0}^{m} \binom{2m-1}{m+k-1} (QP)^k Q^v \right\} \\
&= \sum_{v \in \mathbb{F}_2} \left\{ \sum_{k=0}^{m-1} \left[\binom{2m-1}{m+k} + \binom{2m-1}{m+k-1} \right] P(QP)^k Q^v \right. \\
&\quad \left. + \sum_{k=0}^{m} \left[\binom{2m-1}{m+k} + \binom{2m-1}{m+k+1} \right] P(QP)^k Q^v \right\} \\
&= \sum_{v \in \mathbb{F}_2} \left\{ \sum_{k=0}^{m-1} \binom{2m}{m+k} P(QP)^k Q^v + \sum_{k=0}^{m} \left[\binom{2m}{m+k+1} \right] P(QP)^k Q^v \right\} \\
&= \sum_{u,v \in \mathbb{F}_2} \sum_{k=0}^{m-u} \binom{2m}{m+u+k} P^u (QP)^k Q^v,
\end{aligned} \tag{264}$$

which completes the proof of Eq. (257).

Next, by setting $n = 2m+1$ and $a_y^x = 1$ for all x, y in Eq. (247), we get

$$(I+Q)\left[(I+P)(I+Q)\right]^m = \sum_{u,v \in \mathbb{F}_2} \sum_{k=0}^{m-u} \left|\Theta_{ukv}^{2m+1}\right| P^u (QP)^k Q^v. \tag{265}$$

Comparing Eqs. (257) and (265) gives

$$\left|\Theta_{ukv}^{2m+1}\right| = \binom{2m}{m+u+k}. \tag{266}$$

□

The following theorem gives the cardinality of the set Ξ_l^α.

Theorem 21. *Let $\alpha \in \mathbb{Z}^+$ and $l \in \mathbb{N}$. Then,*

$$\left|\Xi_l^{2m}\right| = \begin{cases} 2\left(\begin{array}{c}\alpha-\delta_l \\ \frac{\alpha}{2}-l-\delta_l\end{array}\right) & \alpha \text{ even}, \\ \left(\begin{array}{c}\alpha+1-\delta_l \\ \frac{\alpha+1}{2}-l-\delta_l\end{array}\right) & \alpha \text{ odd}, \end{cases} \tag{267}$$

i.e. for $m \in \mathbb{Z}^+$,

$$\left|\Xi_l^{2m}\right| = \begin{cases} 2\left(\begin{array}{c}2m-1 \\ m-1\end{array}\right) & l=0, \\ 2\left(\begin{array}{c}2m \\ m-l\end{array}\right) & l \geq 1. \end{cases} \tag{268}$$

$$\left|\Xi_l^{2m+1}\right| = \begin{cases} \left(\begin{array}{c}2m+1 \\ m\end{array}\right) & l=0, \\ 2\left(\begin{array}{c}2m+2 \\ m-l+1\end{array}\right) & l \geq 1. \end{cases} \tag{269}$$

Proof. We first consider the case when $\alpha = 2m$ is even. Since the sets in the union represented by Eq. (27) are disjoint, it follows that

$$\left|\Xi_l^{2m}\right| = \underbrace{\left|\Theta_{0l0}^{2m}\right| + \left|\Theta_{1l0}^{2m}\right|}_{①} + \underbrace{\left|\Theta_{0,l-1,1}^{2m}\right| + \left|\Theta_{1,l-1,1}^{2m}\right|}_{②}. \tag{270}$$

Here,

$$① = \binom{2m-1}{m-1-l} + \binom{2m-1}{m-1-l} = 2\binom{2m-1}{m-1-l} \tag{271}$$

and

$$② = \left[\binom{2m-1}{m-l} + \binom{2m-1}{m-l}\right] \mathbb{1}_{l\geq 1} = 2\binom{2m-1}{m-l}\mathbb{1}_{l\geq 1}. \tag{272}$$

Therefore, if $l = 0$,

$$\left|\Xi_l^{2m}\right| = 2\binom{2m-1}{m-1}. \tag{273}$$

And if $l \geq 1$,

$$\begin{aligned} \left|\Xi_l^{2m}\right| &= 2\left[\binom{2m-1}{m-1-l} + \binom{2m-1}{m-l}\right] \\ &= 2\binom{2m}{m-l}, \end{aligned} \tag{274}$$

by Pascal's rule. Next, we consider the case when $\alpha = 2m + 1$ is odd. As before, since the sets in the union represented by Eq. (27) are disjoint, it follows that

$$\left|\Xi_l^{2m+1}\right| = \underbrace{\left|\Theta_{0l0}^{2m+1}\right| + \left|\Theta_{1l0}^{2m+1}\right|}_{③} + \underbrace{\left|\Theta_{0,l-1,1}^{2m+1}\right| + \left|\Theta_{1,l-1,1}^{2m+1}\right|}_{④}. \tag{275}$$

Here,

$$③ = \binom{2m}{m-l} + \binom{2m}{m-1-l} = \binom{2m+1}{m-l} \tag{276}$$

and

$$④ = \left[\binom{2m}{m-l+1} + \binom{2m}{m-l}\right]\mathbb{1}_{l\geq 1} = \binom{2m+1}{m-l+1}\mathbb{1}_{l\geq 1}. \tag{277}$$

Therefore, if $l = 0$,

$$\left|\Xi_l^{2m}\right| = \binom{2m+1}{m}. \tag{278}$$

And if $l \geq 1$,

$$\begin{aligned}\left|\Xi_l^{2m}\right| &= \binom{2m+1}{m-l} + \binom{2m+1}{m-l+1} \\ &= \binom{2m+2}{m-l+1},\end{aligned} \tag{279}$$

which completes the proof of the theorem. □

D Trigono-multivariate polynomial functions

Let $k, d \in \mathbb{Z}^+$. A k-ary function $f : \mathbb{R}^k \to \mathbb{C}$ is a *trigono-multivariate polynomial function* of degree d if for all $\vec{y} \in \{0,1\}^{kd}$, there exists $\xi_{\vec{y}} \in \mathbb{C}$ such that for all $\vec{x} \in \mathbb{R}^k$,

$$f(\vec{x}) = \sum_{\vec{y}\in\{0,1\}^{kd}} \xi_{\vec{y}}\zeta_{\vec{y}}(\underbrace{\vec{x}, \vec{x}, \ldots, \vec{x}}_{d \text{ times}}). \tag{280}$$

Denote the set of k-ary trigono-multivariate polynomial functions of degree d by $\mathscr{T}_d^k$.

Note that the definition (280) generalizes the notions of trigono-multilinearity and trigono-multiquadraticity: from Eqs. (33) and (34), $\mathscr{T}_1^k$ and $\mathscr{T}_2^k$ are the sets of k-ary trigono-multilinear and trigono-multiquadratic functions respectively.

It is easy to check that the sets $\mathscr{T}_d^k$ satisfy the following simple closure properties.

Proposition 22. *Let $k, d, e \in \mathbb{Z}^+$. The sets $\mathscr{T}_d^k$ satisfy the following properties.*

1. *If $f, g \in \mathscr{T}_d^k$, then $f + g \in \mathscr{T}_d^k$.*
2. *If $f \in \mathscr{T}_d^k$ and $g \in \mathscr{T}_e^k$, then $fg \in \mathscr{T}_{d+e}^k$.*

In particular, the product of two trigono-multilinear functions is a trigono-multiquadratic function.

Next we describe some properties of trigono-multilinear and trigono-multiquadratic functions. Recall that we have defined these functions by expressing them as the sum of exponentially many terms as in Eqs. (33) and (34), respectively. But sometimes it is more convenient to work with the following equivalent definitions of these functions.

Proposition 23. *Let $k \in \mathbb{Z}^+$. A k-ary function $f : \mathbb{R}^k \to \mathbb{C}$ is trigono-multilinear if and only if for all $j \in [k]$, there exist $(k-1)$-ary functions $C_j, S_j : \mathbb{R}^{k-1} \to \mathbb{C}$ such that*

$$f(\vec{x}) = C_j(\vec{x}_{\neg j}) \cos(x_j) + S_j(\vec{x}_{\neg j}) \sin(x_j), \tag{281}$$

where $\vec{x} = (x_1, \dots, x_k)$ and $\vec{x}_{\neg j} = (x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_k)$. We call C_j and S_j the cosine-sine-decomposition (CSD) functions of f with respect to x_j.

Proof. The necessity of the given condition is easy to prove. Suppose $f(\vec{x}) = \sum_{\vec{y} \in \{0,1\}^k} \xi_{\vec{y}} \zeta_{\vec{y}}(\vec{x})$. Then we set

$$C_j(\vec{x}_{\neg j}) = \sum_{\vec{z} \in \{0,1\}^{k-1}} \xi_{\vec{z} \circ 0} \zeta_{\vec{z}}(\vec{x}_{\neg j}), \tag{282}$$

$$S_j(\vec{x}_{\neg j}) = \sum_{\vec{z} \in \{0,1\}^{k-1}} \xi_{\vec{z} \circ 1} \zeta_{\vec{z}}(\vec{x}_{\neg j}), \tag{283}$$

where $\vec{z} \circ 0 = z_1 \dots z_{k-1} 0$ and $\vec{z} \circ 1 = z_1 \dots z_{k-1} 1$, and obtain Eq. (281).

Next, we prove the sufficiency of the given condition by induction on k. It is obvious for $k = 1$. Suppose f satisfies the given condition, i.e. $f(\vec{x}) = C_j(\vec{x}_{\neg j}) \cos(x_j) + S_j(\vec{x}_{\neg j}) \sin(x_j)$ for some C_j and S_j, for all $j \in [k]$. Let $\vec{w} = (w_1, \dots, w_{k-1}) \in \mathbb{R}^{k-1}$ be arbitrary. Let $\vec{a} = \vec{w} \circ 0 = (w_1, \dots, w_{k-1}, 0)$ and $\vec{b} = \vec{w} \circ \pi/2 = (w_1, \dots, w_{k-1}, \pi/2)$. Then we have

$$C_k(\vec{w}) = f(\vec{a}) = \hat{C}_j(\vec{w}_{\neg j}) \cos(w_j) + \hat{S}_j(\vec{w}_{\neg j}) \sin(w_j), \quad \forall j \in [k-1], \tag{284}$$

where $\hat{C}_j(\vec{w}_{\neg j}) = C_j(\vec{a}_{\neg j})$ and $\hat{S}_j(\vec{w}_{\neg j}) = S_j(\vec{a}_{\neg j})$, and

$$S_k(\vec{w}) = f(\vec{b}) = \bar{C}_j(\vec{w}_{\neg j}) \cos(w_j) + \bar{S}_j(\vec{w}_{\neg j}) \sin(w_j), \quad \forall j \in [k-1]. \tag{285}$$

where $\bar{C}_j(\vec{w}_{\neg j}) = C_j(\vec{b}_{\neg j})$ and $\bar{S}_j(\vec{w}_{\neg j}) = S_j(\vec{b}_{\neg j})$. This means that both C_k and S_k satisfy the given condition for $(k-1)$-ary functions. So by induction hypothesis, we know that both C_k and S_k are $(k-1)$-ary trigono-multilinear functions, i.e. they can be expressed as in Eq. (33). It follows that $f(\vec{x}) = C_k(\vec{x}_{\neg k}) \cos(x_k) + S_k(\vec{x}_{\neg k}) \sin(x_k)$ can be also expressed as in Eq. (33), i.e. it is a k-ary trigono-multilinear function. □

Proposition 24. *Let $k \in \mathbb{Z}^+$. A k-ary function $f : \mathbb{R}^k \to \mathbb{C}$ is trigono-multiquadratic if and only if for all $j \in [k]$, there exist $(k-1)$-ary functions $C_j, S_j, B_j : \mathbb{R}^{k-1} \to \mathbb{C}$ such that*

$$f(\vec{x}) = C_j(\vec{x}_{\neg j}) \cos(2x_j) + S_j(\vec{x}_{\neg j}) \sin(2x_j) + B_j(\vec{x}_{\neg j}), \tag{286}$$

where $\vec{x} = (x_1, \dots, x_k)$ and $\vec{x}_{\neg j} = (x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_k)$. We call C_j, S_j and B_j the cosine-sine-bias-decomposition (CSBD) functions of f with respect to x_j.

Proof. Since $\cos(2x) = \cos^2(x) - \sin^2(x)$, $\sin(2x) = 2\cos(x)\sin(x)$ and $\cos^2(x) + \sin^2(x) = 1$, it suffices to show that f is trigono-multilquadratic if and only if for all $j \in [k]$, there exist $(k-1)$-ary functions $E_j, F_j, G_j : \mathbb{R}^{k-1} \to \mathbb{C}$ such that

$$f(\vec{x}) = E_j(\vec{x}_{\neg j}) \cos^2(x_j) + F_j(\vec{x}_{\neg j}) \sin^2(x_j) + G_j(\vec{x}_{\neg j}) \cos(x_j) \sin(x_j), \tag{287}$$

where $\vec{x} = (x_1, \dots, x_k)$ and $\vec{x}_{\neg j} = (x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_k)$.

The necessity of this condition is easy to prove. Suppose $f(\vec{x}) = \sum_{\vec{y}, \vec{z} \in \{0,1\}^k} \xi_{\vec{y}\vec{z}} \zeta_{\vec{y}\vec{z}}(\vec{x}, \vec{x})$. Then we set

$$E_j(\vec{x}_{\neg j}) = \sum_{\vec{u}, \vec{v} \in \{0,1\}^{k-1}} \xi_{\vec{u} \circ 0, \vec{v} \circ 0} \zeta_{\vec{u}, \vec{v}}(\vec{x}_{\neg j}, \vec{x}_{\neg j}), \tag{288}$$

$$F_j(\vec{x}_{\neg j}) = \sum_{\vec{u}, \vec{v} \in \{0,1\}^{k-1}} \xi_{\vec{u} \circ 1, \vec{v} \circ 1} \zeta_{\vec{u}, \vec{v}}(\vec{x}_{\neg j}, \vec{x}_{\neg j}), \tag{289}$$

$$G_j(\vec{x}_{\neg j}) = \sum_{\vec{u}, \vec{v} \in \{0,1\}^{k-1}} \xi_{\vec{u} \circ 1, \vec{v} \circ 0} \zeta_{\vec{u}, \vec{v}}(\vec{x}_{\neg j}, \vec{x}_{\neg j}) + \sum_{\vec{u}, \vec{v} \in \{0,1\}^{k-1}} \xi_{\vec{u} \circ 0, \vec{v} \circ 1} \zeta_{\vec{u}, \vec{v}}(\vec{x}_{\neg j}, \vec{x}_{\neg j}), \tag{290}$$

where $\vec{u} \circ 0 = u_1 \dots u_{k-1}0$ and $\vec{u} \circ 1 = u_1 \dots u_{k-1}1$, and similarly for $\vec{v} \circ 0$ and $\vec{v} \circ 1$, and obtain Eq. (287).

Next, we prove the sufficiency of the above condition by induction on k. It is obvious for $k = 1$. Suppose f satisfies the given condition, i.e. $f(\vec{x}) = E_j(\vec{x}_{\neg j})\cos^2(x_j) + F_j(\vec{x}_{\neg j})\sin^2(x_j) + G_j(\vec{x}_{\neg j})\cos(x_j)\sin(x_j)$, for some E_j, F_j and G_j, for all $j \in [k]$. Let $\vec{w} = (w_1, \dots, w_{k-1}) \in \mathbb{R}^{k-1}$ be arbitrary. Let $\vec{a} = \vec{w} \circ 0 = (w_1, \dots, w_{k-1}, 0)$, $\vec{b} = \vec{w} \circ \pi/2 = (w_1, \dots, w_{k-1}, \pi/2)$ and $\vec{c} = \vec{w} \circ \pi/4 = (w_1, \dots, w_{k-1}, \pi/4)$. Then we have

$$E_k(\vec{w}) = f(\vec{a}) = \hat{E}_j(\vec{w}_{\neg j})\cos^2(w_j) + \hat{F}_j(\vec{w}_{\neg j})\sin^2(w_j) + \hat{G}_j(\vec{w}_{\neg j})\cos(w_j)\sin(w_j), \quad \forall j \in [k-1], \tag{291}$$

where $\hat{E}_j(\vec{w}_{\neg j}) = E_j(\vec{a}_{\neg j})$, $\hat{F}_j(\vec{w}_{\neg j}) = F_j(\vec{a}_{\neg j})$, and $\hat{G}_j(\vec{w}_{\neg j}) = G_j(\vec{a}_{\neg j})$, and

$$F_k(\vec{w}) = f(\vec{b}) = \bar{E}_j(\vec{w}_{\neg j})\cos^2(w_j) + \bar{F}_j(\vec{w}_{\neg j})\sin^2(w_j) + \bar{G}_j(\vec{w}_{\neg j})\cos(w_j)\sin(w_j), \quad \forall j \in [k-1], \tag{292}$$

where $\bar{E}_j(\vec{w}_{\neg j}) = E_j(\vec{b}_{\neg j})$, $\bar{F}_j(\vec{w}_{\neg j}) = F_j(\vec{b}_{\neg j})$, and $\bar{G}_j(\vec{w}_{\neg j}) = G_j(\vec{b}_{\neg j})$, and

$$G_k(\vec{w}) = 2f(\vec{c}) - f(\vec{a}) - f(\vec{b}) \tag{293}$$

$$= \tilde{E}_j(\vec{w}_{\neg j})\cos^2(w_j) + \tilde{F}_j(\vec{w}_{\neg j})\sin^2(w_j) + \tilde{G}_j(\vec{w}_{\neg j})\cos(w_j)\sin(w_j), \quad \forall j \in [k-1], \tag{294}$$

where $\tilde{E}_j(\vec{w}_{\neg j}) = 2E_j(\vec{c}_{\neg j}) - E_j(\vec{a}_{\neg j}) - E_j(\vec{b}_{\neg j})$, $\tilde{F}_j(\vec{w}_{\neg j}) = 2F_j(\vec{c}_{\neg j}) - F_j(\vec{a}_{\neg j}) - F_j(\vec{b}_{\neg j})$, and $\tilde{G}_j(\vec{w}_{\neg j}) = 2G_j(\vec{c}_{\neg j}) - G_j(\vec{a}_{\neg j}) - G_j(\vec{b}_{\neg j})$. This means that E_k, F_k and G_k all satisfy the above condition for $(k-1)$-ary functions. So by induction hypothesis, we know that E_k, F_k and G_k are all $(k-1)$-ary trigono-multiquadratic functions, i.e. they can be expressed as in Eq. (34). It follows that $f(\vec{x}) = E_k(\vec{x}_{\neg k})\cos^2(x_k) + F_k(\vec{x}_{\neg k})\sin^2(x_k) + G_k(\vec{x}_{\neg k})\cos(x_k)\sin(x_k)$ can be also expressed as in Eq. (34), i.e. it is a k-ary trigono-multiquadratic function. □

We say that $f \in \mathscr{T}_d^k$ is *real* if its range is contained in $\mathbb{R}$, i.e. $f(\vec{x}) \in \mathbb{R}$ for all $\vec{x} \in \mathbb{R}^k$. It turns out that for real trigono-multilinear and trigono-multiquadratic functions, if we fix the values of all variables except x_j, then we can easily determine the value of x_j that maximizes (or minimizes) the function, provided that we can efficiently evaluate the CSD or CSBD coefficient functions of the function with respect to x_j.

Specifically, suppose $f : \mathbb{R}^k \to \mathbb{C}$ satisfies the condition in Proposition 23. Then

$$\arg\max_y f(x_1, \dots, x_{j-1}, y, x_{j+1}, \dots, x_k) = \mathrm{Arg}[C_j(\vec{x}_{\neg j}) + \mathrm{i}S_j(\vec{x}_{\neg j})], \tag{295}$$

and

$$\max_y f(x_1, \dots, x_{j-1}, y, x_{j+1}, \dots, x_k) = \sqrt{C_j(\vec{x}_{\neg j})^2 + S_j(\vec{x}_{\neg j})^2}, \tag{296}$$

where $\mathrm{Arg}(x + iy) = \mathrm{atan2}(y, x)$ is the 2-argument arctangent defined by

$$\mathrm{atan2}(y, x) = \begin{cases} \arctan(y/x), & x > 0, \\ \arctan(y/x) + \pi, & x < 0, y \geq 0, \\ \arctan(y/x) - \pi, & x > 0, y < 0, \\ \pi/2, & x = 0, y > 0, \\ -\pi/2, & x = 0, y < 0, \\ \text{undefined}, & x = y = 0. \end{cases} \tag{297}$$

Note that if $C_j(\vec{x}_{\neg j}) = S_j(\vec{x}_{\neg j}) = 0$, then $f(\vec{x}) = 0$ regardless of the value of x_j.

Similarly, suppose $f : \mathbb{R}^k \to \mathbb{C}$ satisfies the condition in Proposition 24. Then

$$\arg\max_y |f(x_1, \dots, x_{j-1}, y, x_{j+1}, \dots, x_k)| = \frac{\mathrm{Arg}[\mathrm{sgn}(B_j)(C_j(\vec{x}_{\neg j}) + \mathrm{i}S_j(\vec{x}_{\neg j}))]}{2}, \tag{298}$$

and

$$\max_y |f(x_1, \dots, x_{j-1}, y, x_{j+1}, \dots, x_k)| = \sqrt{C_j(\vec{x}_{\neg j})^2 + S_j(\vec{x}_{\neg j})^2} + |B_j(\vec{x}_{\neg j})|, \tag{299}$$

where $\mathrm{sgn}(x) = 1$ if $x \geq 0$ and -1 otherwise. Note that if $C_j(\vec{x}_{\neg j}) = S_j(\vec{x}_{\neg j}) = 0$, then $f(\vec{x}) = B_j(\vec{x}_{\neg j})$ does not depend on x_j. So we can pick arbitrary $x_j \in \mathbb{R}$ to maximize $f(\vec{x})$ in this case.

The above properties of trigono-multilinear and trigono-multiquadratic functions have proved to be useful in [12] for the design of quantum algorithms for amplitude estimation based on Bayesian inference with engineered likelihood functions.

E Example: $L = 1$

In this appendix, we consider the special case when $L = 1$. In particular, we show how the results in Sections 2 and 4 specialize in this case.

When $L = 1$, Eq. (9) becomes

$$Q(\theta; x_1, x_2) = V(x_2)U(\theta; x_1). \tag{300}$$

It is straightforward to check that the only nonempty sets $\Omega^4_{l,0}$, $\Omega^4_{l,2}$ and Γ^2_l defined by Eq. (86) are

$$\Omega^4_{0,0} = \{0110\}, \tag{301}$$
$$\Omega^4_{2,2} = \{0011\}, \tag{302}$$
$$\Gamma^2_1 = \{00, 01, 10\}, \tag{303}$$
$$\Gamma^2_3 = \{11\}. \tag{304}$$

Upon substituting these into Eq. (84), we get the following Fourier series expansion of the $L = 1$ ancilla-free bias:

$$\Lambda^{\mathrm{AF}}(\theta; x_1, x_2) = \sum_{l=0}^{3} \mu_l^{\mathrm{AF}}(x_1, x_2)\cos(l\theta), \tag{305}$$

where

$$\begin{aligned}
\mu_0^{\mathrm{AF}}(x_1, x_2) &= 2\cos(x_1)\sin(x_2)\sin(x_1)\cos(x_2), \\
\mu_1^{\mathrm{AF}}(x_1, x_2) &= \cos^2(x_1)\cos^2(x_2) + \cos^2(x_1)\sin^2(x_2) + \cos^2(x_1)\sin^2(x_2), \\
\mu_2^{\mathrm{AF}}(x_1, x_2) &= -2\cos(x_1)\cos(x_2)\sin(x_1)\sin(x_2), \\
\mu_3^{\mathrm{AF}}(x_1, x_2) &= \sin^2(x_1)\sin^2(x_2).
\end{aligned} \tag{306}$$

Therefore, using Eq. (182) and (183), the variance reduction factor (184) becomes

$$\mathcal{V}^{\mathrm{AF}}(\mu, \sigma; x_1, x_2) = \frac{\chi^{\mathrm{AF}}(\mu, \sigma; x_1, x_2)^2}{1 - b^{\mathrm{AF}}(\mu, \sigma; x_1, x_2)^2}, \tag{307}$$

where

$$\begin{aligned}
b^{\mathrm{AF}}(\mu, \sigma; x_1, x_2) = {} & 2\cos(x_1)\sin(x_2)\sin(x_1)\cos(x_2) \\
& + (\cos^2(x_1)\cos^2(x_2) + \cos^2(x_1)\sin^2(x_2) + \cos^2(x_1)\sin^2(x_2))\mathrm{e}^{-\sigma^2/2}\cos(\mu) \\
& - 2\cos(x_1)\cos(x_2)\sin(x_1)\sin(x_2)\mathrm{e}^{-2\sigma^2/2}\cos(2\mu) \\
& + \sin^2(x_1)\sin^2(x_2)\mathrm{e}^{-9\sigma^2/2}\cos(3\mu)
\end{aligned} \tag{308}$$

and

$$\begin{aligned}
\chi^{\mathrm{AF}}(\mu, \sigma; x_1, x_2) = {} & -(\cos^2(x_1)\cos^2(x_2) + \cos^2(x_1)\sin^2(x_2) + \cos^2(x_1)\sin^2(x_2))\mathrm{e}^{-\sigma^2/2}\sin(\mu) \\
& + 4\cos(x_1)\cos(x_2)\sin(x_1)\sin(x_2)\mathrm{e}^{-2\sigma^2/2}\sin(2\mu) \\
& - 3\sin^2(x_1)\sin^2(x_2)\mathrm{e}^{-9\sigma^2/2}\sin(3\mu).
\end{aligned} \tag{309}$$

In the ancilla-based case, Eq. (90) evaluates to

$$\Lambda^{\mathrm{AB}}(\theta; x_1, x_2) = \sum_{l=0}^{1} \mu_l^{\mathrm{AB}}(x_1, x_2)\cos(l\theta), \tag{310}$$

where

$$\mu_0^{\mathrm{AB}}(x_1,x_2)=\cos(x_1)\cos(x_2),$$
$$\mu_1^{\mathrm{AB}}(x_1,x_2)=-\sin(x_1)\sin(x_2). \tag{311}$$

Therefore, using Eq. (182) and (183), the variance reduction factor (184) becomes

$$\mathcal{V}^{\mathrm{AB}}(\mu,\sigma;x_1,x_2)=\frac{\chi^{\mathrm{AB}}(\mu,\sigma;x_1,x_2)^2}{1-b^{\mathrm{AB}}(\mu,\sigma;x_1,x_2)^2}, \tag{312}$$

where

$$b^{\mathrm{AB}}(\mu,\sigma;x_1,x_2)=\cos(x_1)\cos(x_2)-\sin(x_1)\sin(x_2)\mathrm{e}^{-\sigma^2/2}\cos(\mu) \tag{313}$$

and

$$\chi^{\mathrm{AB}}(\mu,\sigma;x_1,x_2)=\sin(x_1)\sin(x_2)\mathrm{e}^{-\sigma^2/2}\sin(\mu). \tag{314}$$

F Leading terms in the cosine series expansions of the biases

In this appendix, we use the expansion formulas in Section 2.3.3 to show that the leading terms of the cosine expansions (65) and (90) can be written as products of sine functions.

Proposition 25. *Let $\vec{x}\in\mathbb{R}^{2L}$. Then the leading terms of the cosine expansions* (65) *and* (90) *are given by*

$$\mu_{2L+1}^{\mathrm{AF}}(\vec{x})=\prod_{i=1}^{2L}\sin^2(x_i), \tag{315}$$

$$\mu_L^{\mathrm{AB}}(\vec{x})=(-1)^L\prod_{i=1}^{2L}\sin(x_i). \tag{316}$$

Proof. To prove Eq. (315), we first recall the expression for the leading coefficient of Eq. (65)

$$\mu_{2L+1}^{\mathrm{AF}}(\vec{x})=\left(\sum_{\substack{\vec{a}1\vec{c}^R\in\Xi_{2L+1}^{4L+1}\\ \mathrm{wt}(\vec{c})-\mathrm{wt}(\vec{a})\equiv_4 0}}-\sum_{\substack{\vec{a}1\vec{c}^R\in\Xi_{2L+1}^{4L+1}\\ \mathrm{wt}(\vec{c})-\mathrm{wt}(\vec{a})\equiv_4 2}}\right)\zeta_{\vec{a}\vec{c}}(\vec{x},\vec{x}). \tag{317}$$

By Eq. (27),

$$\Xi_{2L+1}^{4L+1}=\Theta_{0,2L+1,0}^{4L+1}\cup\Theta_{1,2L+1,0}^{4L+1}\cup\Theta_{0,2L,1}^{4L+1}\cup\Theta_{1,2L,1}^{4L+1}. \tag{318}$$

Now, setting $n=4L+1$ in Theorem 17 gives

$$k>2L-u\implies\Theta_{ukv}^{4L+1}=\emptyset. \tag{319}$$

Since $2L+1>2L-u$ for $u\in\{0,1\}$, Eq. (319) implies that

$$\Theta_{0,2L+1,0}^{4L+1}=\Theta_{1,2L+1,0}^{4L+1}=\emptyset. \tag{320}$$

Since $2L>2L-1$, Eq. (319) implies that

$$\Theta_{1,2L,1}^{4L+1}=\emptyset. \tag{321}$$

Therefore,

$$\begin{aligned}\Xi_{2L+1}^{4L+1}&=\Theta_{0,2L,1}^{4L+1}\\&=\{\vec{x}\in\{0,1\}^{4L+1}:q^{4L+1}p^{4L}\dots q^{x_3}p^{x_2}q^{x_1}\sim(qp)^{2L}q^v\}\\&=\{1^{4L+1}\}\\&=\{1^{2L}\cdot1\cdot1^{2L}\},\end{aligned} \tag{322}$$

where $\cdot$ means string concatenation.

Hence, the set of strings $\vec{a}\vec{c}$ satisfying $\vec{a}1\vec{c}^R \in \Xi_{2L+1}^{4L+1}$ and $\mathrm{wt}(\vec{c}) - \mathrm{wt}(\vec{a}) \equiv_4 0$ consists of only the element 1^{4L} and the set of strings $\vec{a}\vec{c}$ satisfying $\vec{a}1\vec{c}^R \in \Xi_{2L+1}^{4L+1}$ and $\mathrm{wt}(\vec{c}) - \mathrm{wt}(\vec{a}) \equiv_4 2$ is the empty set. Therefore,

$$\begin{aligned}\mu_{2L+1}^{\mathrm{AF}}(\vec{x}) &= \zeta_{1^{4L}}(\vec{x},\vec{x}) \\ &= \prod_{i=1}^{2L} \zeta_i(x_i)^2 \\ &= \prod_{i=1}^{2L} \sin^2(x_i).\end{aligned} \tag{323}$$

To prove Eq. (316), we recall the expression for the leading coefficient of Eq. (90)

$$\mu_L^{\mathrm{AB}}(\vec{x}) = \left(\sum_{\substack{\vec{y}\in\Xi_L^{2L} \\ \mathrm{wt}(\vec{y})\equiv_4 0}} - \sum_{\substack{\vec{y}\in\Xi_L^{2L} \\ \mathrm{wt}(\vec{y})\equiv_4 2}} \right) \zeta_{\vec{y}}(\vec{x}). \tag{324}$$

But

$$\begin{aligned}\Xi_L^{2L} &= \underbrace{\Theta_{0,L,0}^{2L}}_{=\emptyset} \cup \underbrace{\Theta_{1,L,0}^{2L}}_{=\emptyset} \cup \Theta_{0,L-1,1}^{2L} \cup \Theta_{1,L-1,1}^{2L} \\ &= \{\vec{x}\in\{0,1\}^{2L} : p^{x_{2L}} q^{x_{2L-1}} \dots p^{x_2} q^{x_1} \sim p^u (qp)^{L-1} q, u\in\{0,1\}\} \\ &= \{u1^{2L-1} : u \in \{0,1\}\} \\ &= \{01^{2L-1}, 1^{2L}\}.\end{aligned} \tag{325}$$

Now,

$$\begin{aligned}\mathrm{wt}(01^{2L-1} &= 2L-1 \neq 0 \text{ or } 2 \pmod 4 \\ \mathrm{wt}(1^{2L}) = 2L &= \begin{cases} 0 \pmod 4 & L \text{ even}, \\ 2 \pmod 4 & L \text{ odd}. \end{cases}\end{aligned} \tag{326}$$

Hence, the values of the sets $\{\vec{y}\in\Xi_L^{2L} : \mathrm{wt}(\vec{y}) \equiv_4 k\}$, for $k=0,2$, are described by the following table:

	$\{\vec{y}\in\Xi_L^{2L} : \mathrm{wt}(\vec{y}) \equiv_4 0$	$\{\vec{y}\in\Xi_L^{2L} : \mathrm{wt}(\vec{y}) \equiv_4 2$
L even	$\{1^{2L}\}$	$\emptyset$
L odd	$\emptyset$	$\{1^{2L}\}$

Consequently,

$$\begin{aligned}\mu_L^{\mathrm{AB}}(\vec{x}) &= \begin{cases} \zeta_{1^{2L}}(\vec{x}) & L \text{ even} \\ -\zeta_{1^{2L}}(\vec{x}) & L \text{ odd} \end{cases} \\ &= (-1)^L \zeta_{1^{2L}}(\vec{x}) \\ &= (-1)^L \prod_{i=1}^{2L} \sin(x_i).\end{aligned} \tag{327}$$

□

G On noisy likelihood functions

As we show in [12], depolarizing noise that occurs after each rotation operator $V(\cdot)$ in the circuits in Figure 2.1 and/or depolarizing noise during measurement lead to likelihood functions that are of the form[6]

$$\mathcal{L}_{\mathrm{noisy}}^{\mathcal{A}}(\theta; d, \vec{x}) = \frac{1}{2}\left[1 + (-1)^d f \Lambda^{\mathcal{A}}(\theta;\vec{x})\right] \tag{328}$$

[6] As described in [12], in the context of randomized benchmarking, such noise could arise from state preparation and measurement (SPAM) errors [16, 17].

for some *fidelity* parameter $f \in [0, 1)$ [12]. In other words, the effect of noise transforms the bias as

$$\Lambda^{\mathcal{A}} \to f\Lambda^{\mathcal{A}}. \tag{329}$$

Since the bias (176) and the chi function (177) are linear in $\Lambda^{\mathcal{A}}$, they transform as

$$b^{\mathcal{A}} \to fb^{\mathcal{A}}, \tag{330}$$

$$\chi^{\mathcal{A}} \to f\chi^{\mathcal{A}}. \tag{331}$$

Consequently, the variance reduction factor (178) takes the form

$$\mathcal{V}^{\mathcal{A}}_{\text{noisy}}(\mu, \sigma; \vec{x}) = \frac{f^2\chi^{\mathcal{A}}(\mu, \sigma; \vec{x})^2}{1 - f^2 b^{\mathcal{A}}(\mu, \sigma; \vec{x})^2}. \tag{332}$$

References

[1] Zhengfeng Ji, Guoming Wang, Runyao Duan, Yuan Feng, and Mingsheng Ying. Parameter estimation of quantum channels. *IEEE Transactions on Information Theory*, 54(11):5172–5185, 2008.

[2] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone. Quantum metrology. *Phys. Rev. Lett.*, 96:010401, 2006.

[3] Krysta M Svore, Matthew B Hastings, and Michael Freedman. Faster phase estimation. *Quantum Information & Computation*, 14(3&4):306–328, 2013.

[4] Nathan Wiebe and Chris Granade. Efficient bayesian phase estimation. *Phys. Rev. Lett.*, 117:010503, 2016.

[5] Ilia Zintchenko and Nathan Wiebe. Randomized gap and amplitude estimation. *Phys. Rev. A*, 93:062306, 2016.

[6] Yohichi Suzuki, Shumpei Uno, Rudy Raymond, Tomoki Tanaka, Tamiya Onodera, and Naoki Yamamoto. Amplitude estimation without phase estimation. *Quantum Information Processing*, 19(2), 2020.

[7] Maria Schuld and Nathan Killoran. Quantum machine learning in feature Hilbert spaces. *Phys. Rev. Lett.*, 122:040504, 2019.

[8] Emanuel Knill, Gerardo Ortiz, and Rolando D. Somma. Optimal quantum measurements of expectation values of observables. *Phys. Rev. A*, 75:012328, 2007.

[9] Daochen Wang, Oscar Higgott, and Stephen Brierley. Accelerated variational quantum eigensolver. *Phys. Rev. Lett.*, 122:140504, 2019.

[10] A Yu Kitaev. Quantum measurements and the abelian stabilizer problem. *arXiv preprint quant-ph/9511026*, 1995.

[11] Thomas E O'Brien, Brian Tarasinski, and Barbara M Terhal. Quantum phase estimation of multiple eigenvalues for small-scale (noisy) experiments. *New Journal of Physics*, 21(2):023022, 2019.

[12] Guoming Wang, Dax Enshan Koh, Peter D. Johnson, and Yudong Cao. Bayesian inference with engineered likelihood functions for robust amplitude estimation. *In preparation*, 2020.

[13] Ronald Aylmer Fisher. Theory of statistical estimation. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 22, pages 700–725. Cambridge University Press, 1925.

[14] Wolfram Research, Inc. Mathematica, Version 12.0. Champaign, IL, 2019.

[15] John Preskill. Quantum Computing in the NISQ era and beyond. *Quantum*, 2:79, 2018.

[16] Jay M. Gambetta, A. D. Córcoles, S. T. Merkel, B. R. Johnson, John A. Smolin, Jerry M. Chow, Colm A. Ryan, Chad Rigetti, S. Poletto, Thomas A. Ohki, Mark B. Ketchen, and M. Steffen. Characterization of addressability by simultaneous randomized benchmarking. *Phys. Rev. Lett.*, 109:240504, 2012.

[17] Mingyu Sun and Michael R Geller. Efficient characterization of correlated spam errors. *arXiv preprint arXiv:1810.10523*, 2018.