


Differentiating and Integrating ZX Diagrams

Quanlong Wang ✉
Cambridge Quantum, UK

Richie Yeung ✉ 
Cambridge Quantum, UK

Abstract

ZX-calculus has proved to be a useful tool for quantum technology with a wide range of successful applications. Most of these applications are of an algebraic nature. However, other tasks that involve differentiation and integration remain unreachable with current ZX techniques. Here we elevate ZX to an analytical perspective by realising differentiation and integration entirely within the framework of ZX-calculus. We explicitly illustrate the new analytic framework of ZX-calculus by applying it in context of quantum machine learning.

2012 ACM Subject Classification Computer systems organization → Quantum computing

Keywords and phrases ZX Calculus, Diagrammatic Differentiation, Quantum Machine Learning

Acknowledgements We would like to thank Bob Coecke and Konstantinos Meichanetzidis for useful discussions and valuable comments. We would also like to thank Emmanuel Jeandel and Simon Perdrix for the fruitful discussions after the first version of this paper.

1 Introduction

ZX-calculus is a powerful graphical rewrite system proposed by Coecke and Duncan [9] for linear maps, particularly for quantum circuits. A node with n edges in a ZX diagram, like in tensor network notation, represent an order n tensor. Moreover, it is possible to directly evaluate the tensor by performing local rewrites. Using these local rewrites, ZX-calculus has been successfully applied to circuit compilation [3, 4, 16, 31], measurement-based quantum computing [23, 17], fusion-based quantum computing [6], quantum error correction [5], quantum natural language processing [12, 26, 22], and quantum foundations [1, 11, 10, 18]. ZX-calculus can even be used as a concrete realisation of quantum theory [13]. These applications of ZX-calculus are algebraic in nature, and take advantage of *rewriting as a form of computation*: in fact ZX-calculus is a sound, universal [8] and complete [20] proof system that serves as an alternative to traditional linear algebra. However, without the analytical tools of differentiation and integration, ZX-calculus fell short of tackling variational problems such as quantum machine learning or realising a comprehensive version of quantum mechanics including quantum dynamics.

There have been previous attempts at providing semantics for differentiating and integrating ZX diagrams [38, 32, 39]. However, by explicitly using Hilbert space operations, such as addition, these attempts fall outside the realm of vanilla ZX-calculus as there are no techniques on further manipulating sums of diagrams.

In this paper we give for the first time rules for differentiating arbitrary ZX diagrams and integrating a wide class of ZX diagrams (including quantum circuits), thus paving the way for an analytical version of ZX-calculus. As an example, we apply these new techniques to analyse the barren plateau phenomenon from quantum machine learning.

$$\frac{\partial}{\partial \theta} \left[\text{Diagram with 3 green circles } \theta \text{ and yellow diamonds} \right] = \text{Diagram with green square } i, \text{ red circle } \pi, \text{ and green circles } \theta$$

Summary of results

1. Differentiation of arbitrary (algebraic) ZX diagrams, with a unified diagrammatic chain and product rule. (Theorem 14 and Theorem 15)
2. Definite integration of circuit-like ZX diagrams, with up to 3 occurrences of a parameter. (Proposition 19, Theorem 21, and Theorem 23)
3. Diagrammatic formula for the expectation and variance of a quantum circuit's gradient $\frac{\partial \langle H \rangle}{\partial \theta_i}$. (Lemma 25 and Theorem 28)
4. Step-by-step calculation for a simple quantum circuit. (Example 30)

From a general ZX-calculus perspective, this is the first paper to combine sums of ZX diagrams into a single ZX diagram in a methodical way. In particular, we highlight the importance of the W spider in ZX-calculus, which corresponds to the derivation structure of the product rule. These results required the combined power of the Z, X and W spiders, all of 3 of which can be naturally represented within algebraic ZX-calculus.

Note: For presentation purposes, the proofs of some theorems and lemmas are moved to the appendix.

2 Algebraic ZX-calculus

The generators of the original ZX-calculus [9] are chosen with the aim to conveniently represent quantum computational models using complementary observables. On the other hand, the ZW-calculus [19] is designed based on the GHZ and W states, two maximally entangled quantum states [14]. It is known that the Z and W spiders from ZW-calculus act as the multiplication and addition monoid respectively, making it possible to perform arithmetic [15, 19].

$$\begin{array}{c} \boxed{a} \quad \boxed{b} \\ \diagdown \quad \diagup \\ \text{Green Circle} \end{array} = \boxed{a \cdot b} \qquad \begin{array}{c} \boxed{a} \quad \boxed{b} \\ \diagdown \quad \diagup \\ \text{Black Triangle} \end{array} = \boxed{a + b}$$

We will see that the W state is crucial for dealing with sums of diagrams, and it is in fact closely related to the product rule used in differentiation. Conveniently, algebraic ZX-calculus [33] compactly decomposes the W spider and other gadgets such as the logical AND gate [28], into Z spiders, X spiders and triangle gates, thus giving us the benefits of ZX and ZW calculus within a single unified framework.

$$\begin{array}{c} \text{Black Triangle} \\ \diagdown \quad \diagup \end{array} = \begin{array}{c} \text{Red Circle} \\ \diagdown \quad \diagup \\ \text{Yellow Triangle} \\ \diagdown \quad \diagup \\ \text{Green Circle} \end{array} \qquad \boxed{\triangle} = \begin{array}{c} \text{Yellow Triangle} \quad \text{Yellow Triangle} \\ \diagdown \quad \diagup \\ \text{Green Circle labeled } -1 \end{array}$$

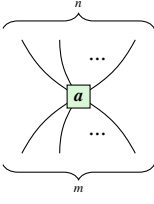







The yellow triangle of algebraic ZX-calculus is powerful as it sends the computational basis to a non-orthogonal basis, which makes diagrammatic representation and calculation of other logical gates much simpler. Conversely, the representation of the yellow triangle using other graphical calculi is

unnatural and cumbersome. Intuitively, this is because the triangle gate is a low-level primitive in comparison to the Z, X, W and H spiders [2]. Using this algebraic extension of ZX, we are able to show that it is a universal and complete language for not just complex numbers, but also commutative rings and semirings [35].

In this section, we give an introduction to the algebraic ZX-calculus, including its generators and rewriting rules. In this paper ZX diagrams are either read from left to right or top to bottom.

2.1 Generators

The diagrams in algebraic ZX-calculus are defined by freely combining the following generating objects:

$R_{Z,a}^{(n,m)} : n \rightarrow m$		$\mathbb{I} : 1 \rightarrow 1$	
$H : 1 \rightarrow 1$		$\sigma : 2 \rightarrow 2$	
$C_a : 0 \rightarrow 2$		$C_u : 2 \rightarrow 0$	
$T : 1 \rightarrow 1$		$T^{-1} : 1 \rightarrow 1$	

■ **Table 1** Generators of algebraic ZX-calculus, where $m, n \in \mathbb{N}$, $a \in \mathbb{C}$.

2.2 Additional notation

For simplicity, we introduce additional notation based on the given generators:

1. The green spider from the original ZX-calculus can be defined using the green box spider in algebraic ZX-calculus.

$$\begin{array}{ccc} \begin{array}{c} \dots \\ \diagup \quad \diagdown \\ \text{green circle} \\ \diagdown \quad \diagup \\ \dots \end{array} & := & \begin{array}{c} \dots \\ \diagup \quad \diagdown \\ \text{green box } e^{ia} \\ \diagdown \quad \diagup \\ \dots \end{array} \end{array} \quad \begin{array}{ccc} \begin{array}{c} \dots \\ \diagup \quad \diagdown \\ \text{white circle} \\ \diagdown \quad \diagup \\ \dots \end{array} & := & \begin{array}{c} \dots \\ \diagup \quad \diagdown \\ \text{green box } 1 \\ \diagdown \quad \diagup \\ \dots \end{array} \end{array}$$

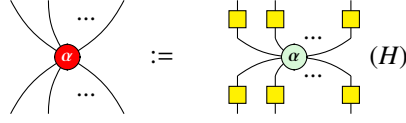
2. The whitespace around a diagram can be interpreted as an explicit horizontal composition with the empty diagram.

$$\begin{array}{c} \cdot \cdot \cdot \cdot \\ \cdot \cdot \cdot \cdot \\ \cdot \cdot \cdot \cdot \\ \cdot \cdot \cdot \cdot \end{array} :=$$

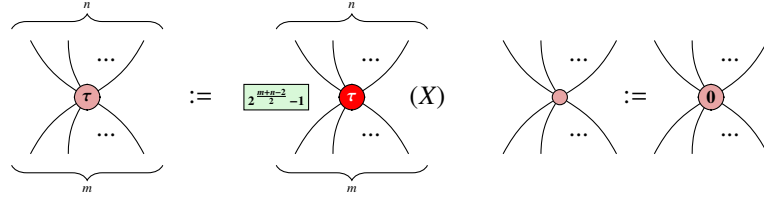
3. The transposes of the triangle and the inverse triangle can be drawn as an inverted triangle.

$$\begin{array}{c} | \\ \text{yellow triangle up} \end{array} := \begin{array}{c} \text{yellow triangle down} \\ \curvearrowright \end{array} \quad \begin{array}{c} | \\ \text{yellow triangle down}^{-1} \end{array} := \begin{array}{c} \text{yellow triangle up}^{-1} \\ \curvearrowleft \end{array}$$

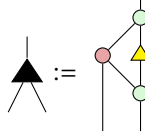
4. The red spider from the original ZX-calculus can be defined by performing Hadamard conjugation on each leg of the green spider.



5. The pink spider is the algebraic equivalent of the red spider. It is only defined for $\tau \in \{0, \pi\}$, and is rescaled to have integer components in its matrix representation.



6. The W spider from ZW calculus can be expressed as follows.



2.3 Interpretation

Although the generators in ZX-calculus are formal mathematical objects in their own right, in the context of this paper we interpret the generators as linear maps, so each ZX diagram is equivalent to a vector or matrix.

$$\begin{aligned}
 & \text{Diagram with } n \text{ legs and } m \text{ legs, labeled } a: \\
 & \quad = |0\rangle^{\otimes m} \langle 0|^{\otimes n} + a |1\rangle^{\otimes m} \langle 1|^{\otimes n}, \\
 & \text{Diagram with } n \text{ legs and } m \text{ legs, labeled } k\pi: \\
 & \quad = \sum_{\substack{0 \leq i_1, \dots, i_m, j_1, \dots, j_n \leq 1 \\ i_1 + \dots + i_m + k \equiv j_1 + \dots + j_n \pmod{2}}} |i_1, \dots, i_m\rangle \langle j_1, \dots, j_n|, k \in \{0, 1\},
 \end{aligned}$$

$$\begin{aligned}
 \text{Yellow square} &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad \text{Yellow triangle} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad \text{Yellow triangle}^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}, \\
 \text{Red circle } \alpha &= e^{i\frac{\alpha}{2}} \begin{pmatrix} \cos \frac{\alpha}{2} & -i \sin \frac{\alpha}{2} \\ -i \sin \frac{\alpha}{2} & \cos \frac{\alpha}{2} \end{pmatrix},
 \end{aligned}$$

$$\begin{aligned}
 \text{Red circle } \pi &= \text{Pink circle } \pi = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \text{Pink circle } \pi \text{ with green square } a = \begin{pmatrix} a-1 & \\ & \end{pmatrix} = a, \quad \text{Pink circle } \pi = 0, \quad \text{Pink circle } \pi = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},
 \end{aligned}$$

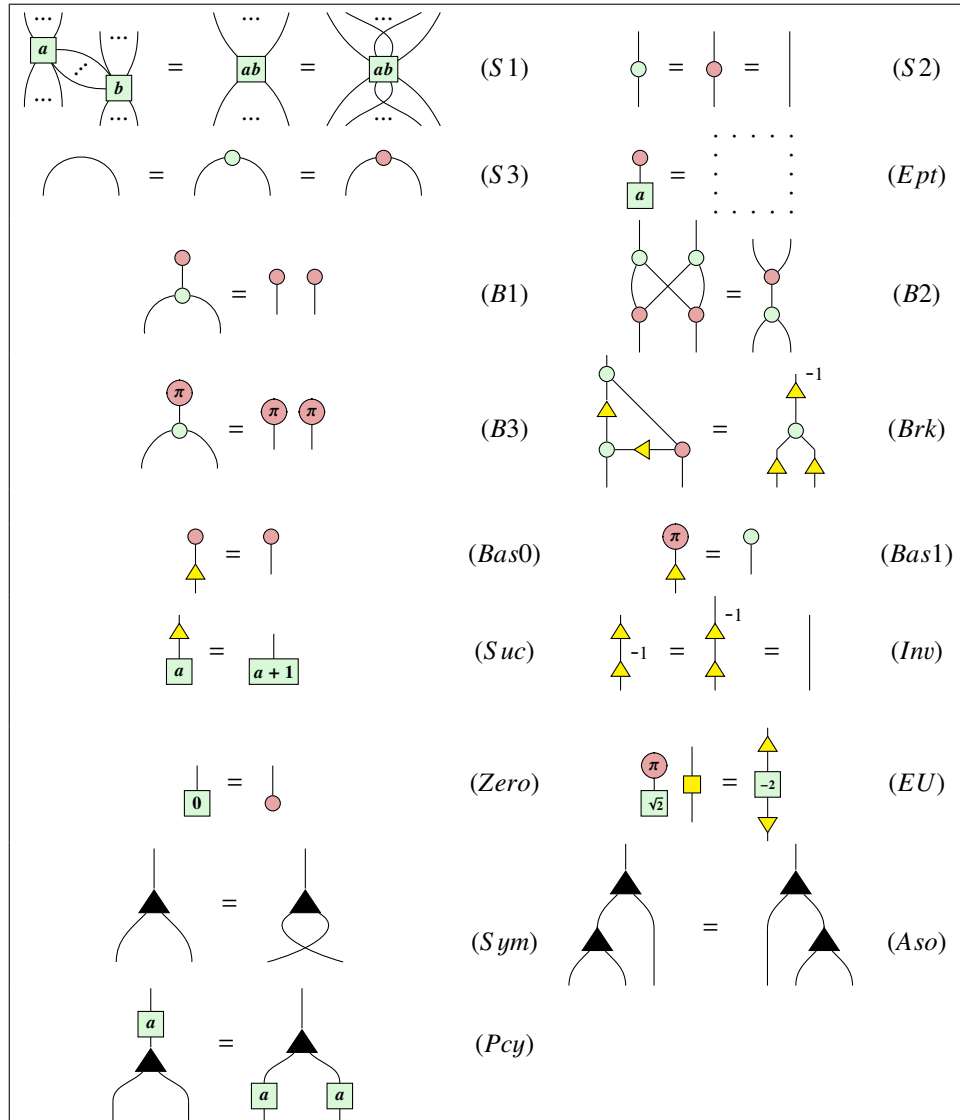
$$\begin{array}{c} \diagup \diagdown \\ \diagdown \diagup \end{array} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \frown = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad \smile = \begin{pmatrix} 1 & 0 & 0 & 1 \end{pmatrix}, \quad \begin{array}{c} \cdots \\ \cdots \\ \cdots \\ \cdots \end{array} = 1,$$

where

$$a \in \mathbb{C}, \quad \alpha \in \mathbb{R}, \quad |0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \text{red dot}, \quad \langle 0| = \begin{pmatrix} 1 & 0 \end{pmatrix} = \text{red dot}, \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \text{red dot with } \pi, \quad \langle 1| = \begin{pmatrix} 0 & 1 \end{pmatrix} = \text{red dot with } \pi.$$

2.4 Rules

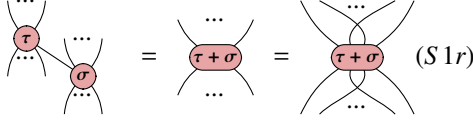
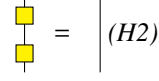
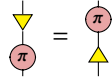

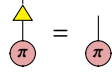
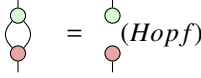
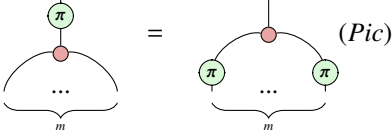
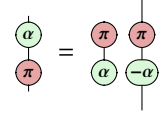
Now we give the rewriting rules of algebraic ZX-calculus.



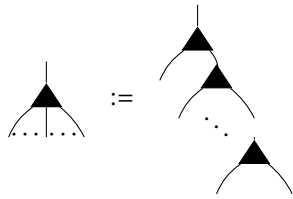
Where $a, b \in \mathbb{C}$. The vertically flipped versions of the rules are assumed to hold as well.

2.5 Useful lemmas

The following lemmas which will be used in the sequel can be derived from the rules.

<p>► Lemma 1. [34] For $\tau, \sigma \in \{0, \pi\}$, pink spiders fuse.</p> 	<p>► Lemma 2. [34] Hadamard is involutive.</p> 
<p>► Lemma 3. [34] Pink π transposes the triangle.</p> 	<p>► Lemma 4. [34] Green π inverts the triangle.</p> 
<p>► Lemma 5. [34] (triangle)^T stabilises $1\rangle$.</p> 	<p>► Lemma 6. [34] Hopf rule.</p> 
<p>► Lemma 7. [34] π copy rule. For $m \geq 0$:</p> 	<p>► Lemma 8. [34] π commutation rule.</p> 

► **Remark 9.** Due to the associative rule (Aso), we can define the W spider



and give its interpretation as follows [36]:

$$\underbrace{\text{W spider}}_m = \underbrace{|0 \cdots 0\rangle}_m \langle 0| + \sum_{k=1}^m \underbrace{|0 \cdots 0 1 0 \cdots 0\rangle}_{k-1} \langle 1|.$$

As a consequence, we have

$$\begin{aligned}
 \text{Diagram 1} &= \text{Diagram 2} \\
 \text{Diagram 3} &= \text{Diagram 4} + \text{Diagram 5} + \dots + \text{Diagram 6}
 \end{aligned} \tag{1}$$

For $n = 2$, the state $|01\rangle + |10\rangle$ can be represented as the quantum state corresponding to the Pauli X gate according to the map-state duality:

► **Lemma 10.**

$$\text{Diagram 7} = \text{Diagram 8}$$

3 Differentiating ZX diagrams

In this section, we show how to differentiate any algebraic ZX diagram within algebraic ZX-calculus, and how to represent the derivative of original ZX diagrams [9] and quantum circuits in algebraic ZX as special case. We start by differentiating the simplest parameterised generator in original ZX-calculus: one-legged green spider.

► **Lemma 11.** Suppose $f(\theta)$ is a differentiable real function of θ . Then

$$\frac{\partial}{\partial \theta} \left[\text{Diagram 9} \right] = \text{Diagram 10}$$

Note: For presentation purposes, the proofs of some theorems and lemmas are moved to the appendix.

Using derivative of the one-legged spider, we can differentiate any ZX diagram with only one occurrence of the parameter being differentiated against. Here is an example.

$$\begin{aligned}
 \text{Diagram 11} &= \text{Diagram 12} \xrightarrow{\partial} \text{Diagram 13} \stackrel{B3}{=} \text{Diagram 14} \\
 |00\rangle\langle 0| + e^{i2\theta} |11\rangle\langle 1| &\xrightarrow{\partial} 2i * e^{i2\theta} |11\rangle\langle 1|
 \end{aligned}$$

When there are multiple occurrences of the same parameter, the derivative can be expressed as a sum of ZX diagrams using the product rule. For example, the density matrix of $R_z(\theta)$ can be differentiated as follows.

$$\text{Diagram 15} \xrightarrow{\partial} \text{Diagram 16} + \text{Diagram 17}$$

Since there are no rules on how to further manipulate sums of ZX diagrams, we cannot do much with derivatives in this form. Moreover, the addition operation is discouraged in ZX-calculus as it does not have a meaningful temporal interpretation, whilst the composition of two diagrams $f \circ g$ can be interpreted as two processes happening one after the other, and the tensor of two diagram $f \otimes g$ can be interpreted as two processes happening concurrently.

For these reasons, we should try to express the derivative as a single diagram. By observing that the product rule leaves the unparameterised parts of the diagram untouched and can be “factored out”, we only need to resynthesise the derivative of the parameterised part.

After this factorisation, the diagrammatic terms in the sum (top of the diagram) can be further manipulated until we can eliminate the sum using a simple rule such as $|0\rangle + |1\rangle = \sqrt{2}|+\rangle$. (See appendix for a demonstration of this technique)

Therefore

This equation, first derived by Zhao et al. [39], is essentially the parameter shift rule by Schuld et al. [30] expressed as a single ZX diagram. Note that our goal in this paper is not to develop new parameter shift rules, as that would require us to express the differentiated circuit as a linear combination of density matrices.

The key result of the paper allows us to express the derivative of an arbitrary ZX diagram in terms of a single diagram. It is based on the observation that the product rule and the unnormalised $|W_n\rangle$ state resemble each other: in the product rule, each term has one differentiated function, and in the W state each term has one bit set to 1 in the basis state.

$$\partial(fgh) = (\partial f)gh + f(\partial g)h + fg(\partial h)$$

$$|W_3\rangle = |100\rangle + |010\rangle + |001\rangle$$

We will show in Theorem 15 that the product rule can indeed be represented using a W state supplemented with some local change of bases. The following lemma demonstrates that the difference between of a “differentiated spider” and an “undifferentiated spider” can be expressed as a change of basis from the computational basis $|0\rangle, |1\rangle$.

► **Lemma 12.** *For any complex number a , we have*

To differentiate algebraic ZX diagrams, we first differentiate its parameterised generator, the one-legged green box:

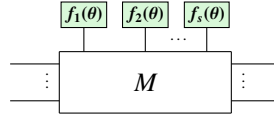
► **Lemma 13.** Suppose $f(t) = k(t) + ig(t)$, where $k(t)$ and $g(t)$ are differentiable real functions, t belongs to some interval T , $f(t) \neq 0$. Then

$$\frac{\partial}{\partial t} \left[\begin{array}{c} \boxed{f(t)} \\ | \\ \pi \end{array} \right] = \begin{array}{c} \boxed{\frac{f'(t)}{f(t)}} \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \boxed{f(t)} \end{array}$$

If $f(t_0) = 0$ for some $t_0 \in T$, then

$$\frac{\partial}{\partial t} \left[\begin{array}{c} \boxed{f(t)} \\ | \\ \pi \end{array} \right]_{t_0} = \begin{array}{c} \pi \\ | \\ \boxed{f'(t_0)} \end{array}$$

All parameterised differentiable algebraic ZX diagrams can be rewritten into the following form, where M is an unparameterised ZX diagram with respect to θ , and $\{f_i(\theta)\}_i$ are differentiable real functions of θ . Parameterised green spiders can be written as a green box with an exponentiated phase, and parameterised red spiders can be converted to parameterised green spiders via Hadamard conjugation. We emphasise that M can contain other parameterised spiders, just not with respect to θ .



► **Theorem 14.** Assume $f_j(t) = k_j(t) + ig_j(t)$, where $k_j(t)$ and $g_j(t)$ are differentiable real functions, t belongs to some interval T . Then for $f_j(t) \neq 0, \forall j \in \{1, \dots, s\}$:

$$\frac{\partial}{\partial t} \left[\begin{array}{c} \boxed{f_1(t)} \quad \boxed{f_2(t)} \quad \boxed{f_s(t)} \\ | \quad | \quad | \\ \vdots \quad \vdots \quad \vdots \\ \boxed{M} \\ | \quad | \quad | \\ \vdots \quad \vdots \quad \vdots \end{array} \right] = \begin{array}{c} \pi \\ \blacktriangle \\ \begin{array}{c} \boxed{\frac{f'_1(t)}{f_1(t)}} \quad \boxed{\frac{f'_2(t)}{f_2(t)}} \quad \boxed{\frac{f'_s(t)}{f_s(t)}} \\ \downarrow \quad \downarrow \quad \downarrow \\ \boxed{f_1(t)} \quad \boxed{f_2(t)} \quad \boxed{f_s(t)} \end{array} \\ | \quad | \quad | \\ \vdots \quad \vdots \quad \vdots \\ \boxed{M} \\ | \quad | \quad | \\ \vdots \quad \vdots \quad \vdots \end{array}$$

If $f_j(t_0) \neq 0$ for $1 \leq j \leq k$ and $f_j(t_0) = 0$ for $k+1 \leq j \leq s$ (the order of the functions can be changed without loss of generality), then

$$\frac{\partial}{\partial t} \left[\begin{array}{c} \boxed{f_1(t)} \quad \boxed{f_2(t)} \quad \boxed{f_s(t)} \\ | \quad | \quad | \\ \vdots \quad \vdots \quad \vdots \\ \boxed{M} \\ | \quad | \quad | \\ \vdots \quad \vdots \quad \vdots \end{array} \right]_{t_0} = \begin{array}{c} \pi \\ \blacktriangle \\ \begin{array}{c} \boxed{\frac{f'_1(t_0)}{f_1(t_0)}} \quad \boxed{\frac{f'_k(t_0)}{f_k(t_0)}} \quad \boxed{f'_{k+1}(t_0)} \quad \boxed{f'_s(t_0)} \\ \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \\ \boxed{f_1(t_0)} \quad \boxed{f_k(t_0)} \quad \vdots \quad \vdots \end{array} \\ | \quad | \quad | \quad | \\ \vdots \quad \vdots \quad \vdots \quad \vdots \\ \boxed{M} \\ | \quad | \quad | \quad | \\ \vdots \quad \vdots \quad \vdots \quad \vdots \end{array}$$

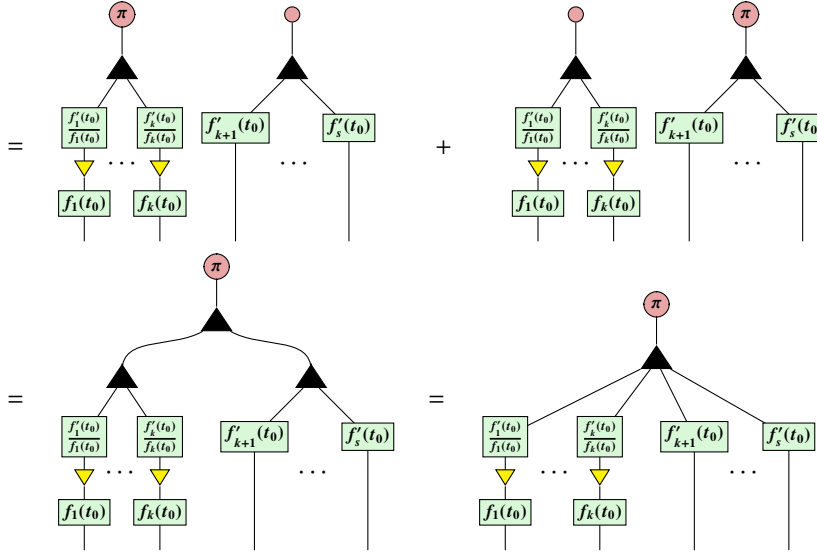
Proof. The formal semantics of diagrammatic differentiation is given in [32]. By linearity, differentiating the overall diagram amounts to differentiating the parameterised part of the diagram:

$$\begin{aligned}
 & \frac{\partial}{\partial t} \left[\begin{array}{c} \boxed{f_1(t)} \quad \boxed{f_2(t)} \quad \boxed{f_s(t)} \\ | \quad | \quad | \\ \dots \end{array} \right] \\
 &= \begin{array}{c} \boxed{\frac{f'_1(t)}{f_1(t)}} \quad \pi \quad \boxed{f_1(t)} \quad \boxed{f_2(t)} \quad \boxed{f_s(t)} \\ | \quad | \quad | \quad | \quad | \\ \pi \quad \dots \end{array} + \begin{array}{c} \boxed{\frac{f'_2(t)}{f_2(t)}} \quad \pi \quad \boxed{f_1(t)} \quad \boxed{f_2(t)} \quad \boxed{f_s(t)} \\ | \quad | \quad | \quad | \quad | \\ \pi \quad \dots \end{array} + \dots + \begin{array}{c} \boxed{\frac{f'_s(t)}{f_s(t)}} \quad \pi \quad \boxed{f_1(t)} \quad \boxed{f_2(t)} \quad \boxed{f_s(t)} \\ | \quad | \quad | \quad | \quad | \\ \pi \quad \dots \end{array} \\
 &= \begin{array}{c} \boxed{\frac{f'_1(t)}{f_1(t)}} \quad \pi \quad \boxed{f_1(t)} \quad \boxed{f_2(t)} \quad \boxed{f_s(t)} \\ | \quad | \quad | \quad | \quad | \\ \pi \quad \dots \end{array} + \begin{array}{c} \boxed{\frac{f'_2(t)}{f_2(t)}} \quad \pi \quad \boxed{f_1(t)} \quad \boxed{f_2(t)} \quad \boxed{f_s(t)} \\ | \quad | \quad | \quad | \quad | \\ \pi \quad \dots \end{array} + \dots + \begin{array}{c} \boxed{\frac{f'_s(t)}{f_s(t)}} \quad \pi \quad \boxed{f_1(t)} \quad \boxed{f_2(t)} \quad \boxed{f_s(t)} \\ | \quad | \quad | \quad | \quad | \\ \pi \quad \dots \end{array} \\
 &= \begin{array}{c} \pi \quad \pi \quad \pi \\ | \quad | \quad | \\ \boxed{\frac{f'_1(t)}{f_1(t)}} \quad \boxed{\frac{f'_2(t)}{f_2(t)}} \quad \boxed{\frac{f'_s(t)}{f_s(t)}} \\ | \quad | \quad | \\ \boxed{f_1(t)} \quad \boxed{f_2(t)} \quad \boxed{f_s(t)} \end{array} + \dots + \begin{array}{c} \pi \quad \pi \quad \pi \\ | \quad | \quad | \\ \boxed{\frac{f'_1(t)}{f_1(t)}} \quad \boxed{\frac{f'_2(t)}{f_2(t)}} \quad \boxed{\frac{f'_s(t)}{f_s(t)}} \\ | \quad | \quad | \\ \boxed{f_1(t)} \quad \boxed{f_2(t)} \quad \boxed{f_s(t)} \end{array} \\
 &= \begin{array}{c} \pi \\ | \\ \boxed{\frac{f'_1(t)}{f_1(t)}} \quad \boxed{\frac{f'_2(t)}{f_2(t)}} \quad \boxed{\frac{f'_s(t)}{f_s(t)}} \\ | \quad | \quad | \\ \boxed{f_1(t)} \quad \boxed{f_2(t)} \quad \boxed{f_s(t)} \end{array}
 \end{aligned}$$

The first step follows from Lemma 13 and the product rule, and the second step performs further factorisation of the diagram. The third step follows from Lemma 12. The final step uses the property of W spider as given in (1).

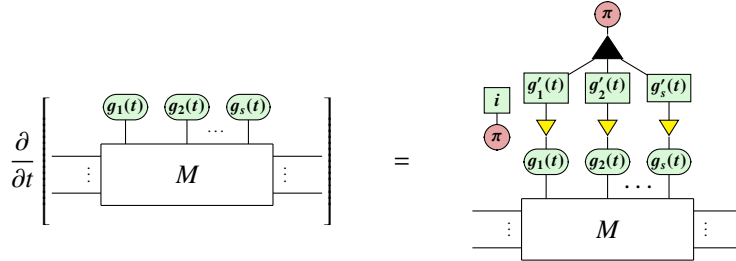
If $f_j(t_0) \neq 0$ for $1 \leq j \leq k$ and $f_j(t_0) = 0$ for $k+1 \leq j \leq s$ (the order of the functions can be changed without loss of generality), then the diagram can be differentiated in two parts using the product rule:

$$\begin{aligned}
 & \frac{\partial}{\partial t} \left[\begin{array}{c} \boxed{f_1(t)} \quad \boxed{f_2(t)} \quad \boxed{f_s(t)} \\ | \quad | \quad | \\ \dots \end{array} \right]_{t_0} \\
 &= \frac{\partial}{\partial t} \left[\begin{array}{c} \boxed{f_1(t)} \quad \boxed{f_k(t)} \\ | \quad | \\ \dots \end{array} \right]_{t_0} \left[\begin{array}{c} \boxed{f_{k+1}(t)} \quad \boxed{f_s(t)} \\ | \quad | \\ \dots \end{array} \right]_{t_0} + \left[\begin{array}{c} \boxed{f_1(t)} \quad \boxed{f_k(t)} \\ | \quad | \\ \dots \end{array} \right]_{t_0} \frac{\partial}{\partial t} \left[\begin{array}{c} \boxed{f_{k+1}(t)} \quad \boxed{f_s(t)} \\ | \quad | \\ \dots \end{array} \right]_{t_0} \\
 &= \frac{\partial}{\partial t} \left[\begin{array}{c} \boxed{f_1(t)} \quad \boxed{f_k(t)} \\ | \quad | \\ \dots \end{array} \right]_{t_0} \begin{array}{c} \pi \quad \pi \quad \pi \\ | \quad | \quad | \\ \dots \end{array} \\
 &+ \begin{array}{c} \boxed{f_1(t_0)} \quad \boxed{f_k(t_0)} \\ | \quad | \\ \dots \end{array} \left[\begin{array}{c} \pi \quad \pi \quad \pi \\ | \quad | \quad | \\ \boxed{f'_{k+1}(t_0)} \quad \boxed{f'_{k+2}(t_0)} \quad \boxed{f'_s(t_0)} \\ | \quad | \quad | \\ \dots \end{array} + \dots + \begin{array}{c} \pi \quad \pi \quad \pi \\ | \quad | \quad | \\ \boxed{f'_{k+1}(t_0)} \quad \boxed{f'_{k+2}(t_0)} \quad \boxed{f'_s(t_0)} \\ | \quad | \quad | \\ \dots \end{array} \right]
 \end{aligned}$$

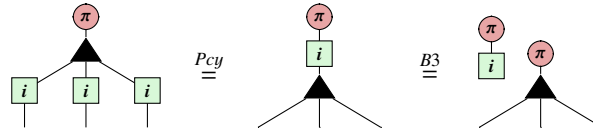


This theorem unifies the linearity, chain and product rules of differential calculus into a single diagram, without a blowup in diagram size: the number of nodes added to the diagram after differentiation is linearly proportional to the number of parameter occurrences. This makes the result practically useful for both calculations by hand and computer simulation.

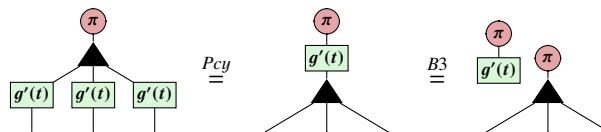
► **Theorem 15.** *The derivative of a differentiable ZX diagram can be expressed as a single ZX diagram:*



Proof. This is a special case of Theorem 14, where $f_j(t) = e^{ig_j(t)}$ and $\frac{f'_j(t)}{f_j(t)} = ig'_j(t)$. The “ i ” is common across all functions and can be factored out through the W spider.



This proof can be generalised to explicitly show the chain rule: when differentiating $f(g(t))$ with respect to t , $\frac{(f \circ g)'(t)}{(f \circ g)(t)} = g'(t) \frac{f'(g(t))}{f(g(t))}$ and so the $g'(t)$ can be dragged out using the fusion rule.



Since we can now differentiate arbitrary ZX diagrams, we can consider differentiating quantum circuits as a special case of theorem 15. Executing a circuit on a quantum computer estimates the a Hamiltonian $\langle H \rangle = \langle 0 | U^\dagger(\theta) H U(\theta) | 0 \rangle$ with an equal number of occurrences of θ and $-\theta$.

► **Corollary 16.** *The derivative of a parameterised quantum circuit can be expressed as a single ZX diagram:*

$$\frac{\partial}{\partial \theta} \left[\begin{array}{c} \theta \\ \vdots \\ \theta \end{array} \right] M \left[\begin{array}{c} -\theta \\ \vdots \\ -\theta \end{array} \right] = \begin{array}{c} i \\ \pi \end{array} \quad \text{ZX diagram with } M, \theta, -\theta, \pi$$

Proof. Note that $(-\theta)' = -1$ and $\begin{array}{c} i \\ -1 \end{array} = \begin{array}{c} i \\ \pi \end{array}$, it follows directly from Theorem 15.

► **Corollary 17.** *As a special case of 16, we obtain the parameter-shift rule from [30].*

$$\frac{\partial}{\partial \theta} \left[\theta \right] M \left[-\theta \right] \stackrel{16}{=} \begin{array}{c} i \\ \pi \end{array} \quad \text{ZX diagram with } M, \theta, -\theta, \pi$$

$$\stackrel{32}{=} \begin{array}{c} \pi \end{array} \quad \text{ZX diagram with } M, \theta + \frac{\pi}{2}, -\theta - \frac{\pi}{2}, \pi$$

► **Remark 18.** This result has been given as a theorem in [39], here we directly get it as a consequence of Corollary 16 which follows from Theorem 15.

From Corollary 16, we thus have a simple diagrammatic expression for the derivative of any parameterised quantum circuit. In general, it is not easy to obtain the derivative of a parametrised matrix in a single term bra-ket expression (with no sums, e.g. $\langle \phi | ABC | \psi \rangle$), thus showing the power of ZX-calculus and 2-dimensional diagrammatic reasoning. Similar to how the decomposition of the Pauli X gate in Corollary 17 gives us a 2 term parameter shift rule, we believe careful decomposition of the W state would yield new, more efficient gradient recipes for quantum machine learning.

4 Integrating ZX diagrams

In this section we show how to integrate a wide class of ZX diagrams, including parameterised quantum circuits, using algebraic ZX-calculus. This allows us to evaluate the expectation and variance of a quantum circuit's derivative over the uniform distribution, as demonstrated in section 5.

► **Proposition 19.** *Let k be an arbitrary non-zero integer, ${}^{m+1}\left\{\begin{array}{c} \vdots \\ K \\ \vdots \end{array}\right\}^{n+1}$ be any ZX diagram with $m, n \geq 0$ which has no occurrence of α . Then*

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} {}^m\left\{\begin{array}{c} \vdots \\ K \\ \vdots \end{array}\right\}^n d\alpha = {}^m\left\{\begin{array}{c} \vdots \\ K \\ \vdots \end{array}\right\}^n$$

► **Remark 20.** This proposition has been proved in [39] for $k = 1$, here we give a new proof with our sum techniques.

► **Theorem 21.** *Let k be an arbitrary non-zero integer, ${}^{m+2}\left\{\begin{array}{c} \vdots \\ A \\ \vdots \end{array}\right\}^{n+2}$ be any ZX diagram with $m, n \geq 0$ which has no occurrence of α . Then*

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} {}^m\left\{\begin{array}{c} \vdots \\ A \\ \vdots \end{array}\right\}^n d\alpha =$$

► **Remark 22.** The corresponding result of this theorem is shown as Lemma 2 in [39] where there are three diagrammatic sum terms after integration, which results in their computation of variance of gradients becoming very complicated. Here we only obtain a single diagram after integration.

► **Theorem 23.** *Let k be an arbitrary non-zero integer, ${}^{m+3}\left\{\begin{array}{c} \vdots \\ A \\ \vdots \end{array}\right\}^{n+3}$ be any ZX diagram with $m, n \geq 0$ which has no occurrence of α . Then*

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} {}^m\left\{\begin{array}{c} \vdots \\ A \\ \vdots \end{array}\right\}^n d\alpha =$$

The proofs for Theorem 21 and Theorem 23 are similar to Proposition 19, and are given in the appendix.

5 Example Application: Quantum Machine Learning

In the NISQ era of quantum computing [29], many applications require the optimisation of parameterised quantum circuits: in quantum chemistry, variational quantum eigensolvers [25] are optimised to find the ground state of a Hamiltonian; in quantum machine learning, a circuit ansatz is optimised against a cost function [24], much alike how neural networks are optimised in classical machine learning.

However, while the approach of using gradient-based methods to optimise deep neural networks has been consistently effective [7], gradient-based optimisation of parameterised quantum circuits often suffer from barren plateaus: the training landscape of many circuit ansätze have been shown to be exponentially flat with respect to circuit size, making gradient descent impossible [27]. Therefore, it is crucial to develop techniques to detect and avoid barren plateaus.

So far, there has not been a fully diagrammatic analysis of barren plateaus using ZX-calculus. We believe the main obstacle to the analysis is the lack of techniques for manipulating sums of diagrams: an expectation of a Hamiltonian contains at least two occurrences of each circuit parameter, so the derivative of the expectation with respect to that parameter requires the product rule. Similarly, the rule for integrating diagrams in [39] introduces three terms, after n integrals there would be 3^n terms. Using these rules, the analysis of barren plateaus become exponentially costly, as the number of diagrams to be evaluated is exponential to the number of parameters in the circuit.

In this section, as a demonstration of the new differentiation and integration techniques of this paper, we show that the analysis of barren plateaus by Zhao et al. [39] can be done entirely within the framework of ZX without introducing sums of diagrams. Using the same setup, we consider an n -qubit parameterised quantum circuit $U(\theta)$ and a Hamiltonian H , then the expectation of H can be written as $\langle H \rangle = \langle 0|U^\dagger(\theta)HU(\theta)|0\rangle$, where the variables of the parameterised quantum circuit appear exactly once in $U(\theta)$. The parameters of the quantum circuit are assumed to be independently and uniformly distributed on the interval $[\pi, -\pi]$, as this is a common way to initialise parameters. If $\text{Var}\left(\frac{\partial \langle H \rangle}{\partial \theta_j}\right) \approx 0$, then the quantum circuit is likely to start in a barren plateau where the circuit gradient $\frac{\partial \langle H \rangle}{\partial \theta_i}$ is close to 0.

To avoid scalars for the diagrammatic representation of a parameterised quantum circuit, we will use the two kinds of X spiders (red and pink respectively) as defined in Section 2 which are equivalent up to some scalar.

Now we can represent the gate set without scalars.

$$R_Z(\theta) = \text{---} \textcircled{\theta} \text{---}, \quad R_X(\theta) = \text{---} \textcircled{\theta} \text{---}, \quad CNOT = \begin{array}{c} \text{---} \textcircled{\text{green}} \text{---} \\ | \\ \text{---} \textcircled{\text{red}} \text{---} \end{array}, \quad H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \text{---} \textcircled{\text{yellow}} \text{---}$$

Below we give the general form of $\langle H \rangle$.

$$\langle H \rangle = \begin{array}{ccc} \theta_1 & \text{---} & \text{---} & -\theta_1 \\ \vdots & & H & \vdots \\ \theta_m & \text{---} & \text{---} & -\theta_m \end{array} \quad (2)$$

► **Example 24.** Let $U(\theta) = \begin{array}{cc} \text{---} \theta_1 \text{---} & \text{---} \theta_3 \text{---} \\ \text{---} \theta_2 \text{---} & \text{---} \theta_4 \text{---} \end{array}$, the Hamiltonian $H = X \otimes X$. Then

$$\langle H \rangle = \begin{array}{ccccccc} \theta_1 & & \theta_3 & & \pi & & -\theta_3 & & -\theta_1 \\ \theta_2 & & \theta_4 & & \pi & & -\theta_4 & & -\theta_2 \end{array}$$

$$\frac{H}{X} = \begin{array}{ccccccc} \theta_1 & & \theta_3 & & \pi & & -\theta_3 & & -\theta_1 \\ \theta_2 & & \theta_4 & & \pi & & -\theta_4 & & -\theta_2 \\ \theta_3 & & \theta_4 & & \pi & & -\theta_3 & & -\theta_4 \\ \theta_4 & & \theta_4 & & \pi & & -\theta_4 & & -\theta_4 \end{array}$$

► **Lemma 25.** Given $\langle H \rangle$ in the form of (2), we have $\mathbf{E} \left(\frac{\partial \langle H \rangle}{\partial \theta_j} \right) = 0$, for $j = 1, \dots, m$.

As a consequence, we have

$$\text{Var} \left(\frac{\partial \langle H \rangle}{\partial \theta_j} \right) = \mathbf{E} \left(\left(\frac{\partial \langle H \rangle}{\partial \theta_j} \right)^2 \right) = \frac{1}{(2\pi)^m} \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} \left(\frac{\partial \langle H \rangle}{\partial \theta_j} \right)^2 d\theta_1 \cdots d\theta_m, j = 1, \dots, m.$$

► **Lemma 26.**

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\frac{\partial \langle H \rangle}{\partial \theta_j} \right)^2 d\theta_j = \begin{array}{c} \begin{array}{c} \pi \\ \vdots \\ \theta_m \\ \vdots \\ \theta_1 \end{array} \begin{array}{c} \vdots \\ H \\ \vdots \\ H \\ \vdots \end{array} \begin{array}{c} -\theta_1 \\ \vdots \\ -\theta_m \\ \vdots \\ -\theta_1 \end{array} \\ \pi \end{array}$$

► **Remark 27.** The corresponding result given in [39] on page 12 is different from ours as shown in Lemma 26.

► **Theorem 28.** Given $\langle H \rangle$ in the form of (2), we have

$$\text{Var} \left(\frac{\partial \langle H \rangle}{\partial \theta_j} \right) = \begin{array}{c} \begin{array}{c} \vdots \\ \pi \\ \vdots \\ \theta_m \\ \vdots \\ \theta_1 \end{array} \begin{array}{c} \vdots \\ H \\ \vdots \\ H \\ \vdots \end{array} \begin{array}{c} -\theta_1 \\ \vdots \\ -\theta_m \\ \vdots \\ -\theta_1 \end{array} \\ \pi \end{array}$$

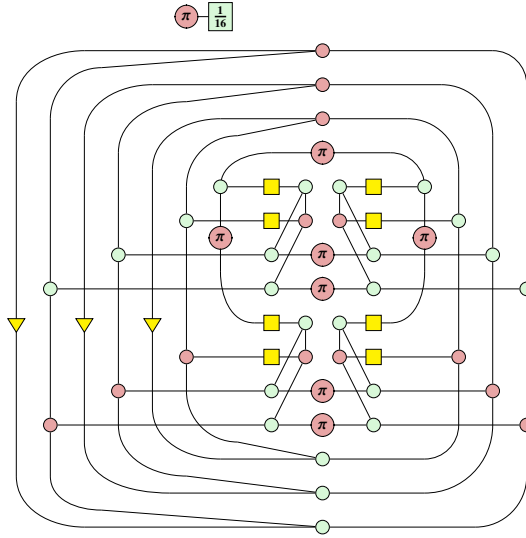
Proof. We start from the diagram as shown in Lemma 26, then drag out variables iteratively according to Theorem 21, and the result follows. ◀

► **Remark 29.** The variance computed in [39] is based on a sum over 3^{m-1} terms (m is the number of parameters in the considered circuit), so is infeasible by their method using ZX-calculus when m is large, thus they have to resort to tensor networks which goes beyond the ZX method. In contrast, we avoid this exponential explosion by integrating without sums using algebraic ZX-calculus.

► **Example 30.** Using the same $\langle H \rangle$ as Example 24, start from Lemma 26 and Theorem 28, we have

$$\text{Var}\left(\frac{\partial \langle H \rangle}{\partial \theta_1}\right) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\frac{\partial \langle H \rangle}{\partial \theta_1}\right)^2 d\theta_1 d\theta_2 d\theta_3 d\theta_4 = \frac{3}{16}$$

which requires evaluating the following diagram (full calculation in appendix C.1).



6 Conclusion and further work

We have elevated ZX-calculus from a graphical language for algebraic calculations to a new graphical tool for analytical reasoning. For example, it can now be used for tackling quantum optimisation problems and reasoning about quantum mechanics. We believe these techniques will extend the applicability of ZX-calculus to more problems related to quantum computing. With more work, ZX-calculus can become a general tool for graphical differential calculus.

There are many directions for future work:

1. **Generalisation of the results to qudit and qufinite cases [37]:** The ideas in this paper can be extended beyond qubits, giving us more diagrammatic analytical tools.

After uploading a preprint of this paper on arxiv, we were informed by Emmanuel Jeandel and Simon Perdrix that they were also working on a paper on the topic of differentiation in ZX-calculus with similar results, now available here [21]. A comparison of our results can be found in Appendix D.

2. **Integrating arbitrary ZX diagrams:** This paper only gives the definite integral for circuit-like ZX diagrams. Indefinite integration of arbitrary ZX diagrams would in a sense complete the analytical ZX-calculus.
3. **New gradient recipes for quantum machine learning:** Rewriting Lemma 16 into a linear combination of circuits would immediately give new recipes for evaluating a quantum circuit's gradient.
4. **New algorithm for detecting barren plateaus in QML ansatz:** An efficient method of the diagram in Theorem 28, which is mostly Clifford, would yield an algorithm to detect barren plateaus in an ansatz without using a quantum computer.

References

- 1 Miriam Backens & Ali Nabi Duman (2015): *A Complete Graphical Calculus for Spekkens' Toy Bit Theory*. *Foundations of Physics* 46(1), p. 70–103, doi:10.1007/s10701-015-9957-7. Available at <http://dx.doi.org/10.1007/s10701-015-9957-7>.
- 2 Miriam Backens & Aleks Kissinger (2018): *ZH: A Complete Graphical Calculus for Quantum Computations Involving Classical Non-linearity*. In: *Proceedings of the 15th International Workshop on Quantum Physics and Logic, QPL 2018, Halifax, Canada, 3-7th June 2018.*, pp. 23–42. doi:10.4204/EPTCS.287.2.
- 3 Niel de Beaudrap, Xiaoning Bian & Quanlong Wang (2020): *Fast and Effective Techniques for T-Count Reduction via Spider Nest Identities*. In Steven T. Flammia, editor: *15th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2020), Leibniz International Proceedings in Informatics (LIPIcs)* 158, Schloss Dagstuhl–Leibniz-Zentrum für Informatik, Dagstuhl, Germany, pp. 11:1–11:23, doi:10.4230/LIPIcs.TQC.2020.11.
- 4 Niel de Beaudrap, Xiaoning Bian & Quanlong Wang (2020): *Techniques to reduce $\pi/4$ -parity-phase circuits, motivated by the ZX calculus*. *Proceedings of the 16th International Conference on Quantum Physics and Logic 2019, EPTCS* 318, pp. 131–149. ArXiv:1911.09039.
- 5 Niel de Beaudrap & Dominic Horsman (2020): *The ZX calculus is a language for surface code lattice surgery*. *Quantum* 4, doi:10.22331/q-2020-01-09-218.
- 6 Hector Bombin, Chris Dawson, Ryan V Mishmash, Naomi Nickerson, Fernando Pastawski & Sam Roberts (2021): *Logical blocks for fault-tolerant topological quantum computation*. *arXiv preprint arXiv:2112.12160*.
- 7 Anna Choromanska, Mikael Henaff, Michael Mathieu, Gérard Ben Arous & Yann LeCun (2015): *The loss surfaces of multilayer networks*. In: *Artificial intelligence and statistics*, PMLR, pp. 192–204.
- 8 Bob Coecke & Ross Duncan (2008): *Interacting Quantum Observables*. In: *Automata, Languages and Programming*, 5126, Springer Berlin Heidelberg, pp. 298–310. https://doi.org/10.1007/978-3-540-70583-3_25.
- 9 Bob Coecke & Ross Duncan (2011): *Interacting quantum observables: categorical algebra and diagrammatics*. *New Journal of Physics* 13(4), p. 043016. Available at <http://stacks.iop.org/1367-2630/13/i=4/a=043016>. doi:10.1088/1367-2630/13/4/043016.
- 10 Bob Coecke, Ross Duncan, Aleks Kissinger & Quanlong Wang (2012): *Strong Complementarity and Non-locality in Categorical Quantum Mechanics*. In: *Proceedings of the 2012 27th Annual IEEE/ACM Symposium on Logic in Computer Science, LICS '12*, IEEE Computer Society, pp. 245–254. doi:10.1109/LICS.2012.35.
- 11 Bob Coecke & Bill Edwards (2012): *Spekkens's toy theory as a category of processes*. *Proceedings of Symposia in Applied Mathematics* 71, pp. 61–88.

- 12 Bob Coecke, Giovanni de Felice, Konstantinos Meichanetzidis & Alexis Toumi (2020): *Foundations for Near-Term Quantum Natural Language Processing*. Available at <https://arxiv.org/abs/2012.03755>.
- 13 Bob Coecke, Dominic Horsman, Aleks Kissinger & Quanlong Wang (2022): *Kindergarden quantum mechanics graduates ...or how I learned to stop gluing LEGO together and love the ZX-calculus*. *Theoretical Computer Science* 897, pp. 1–22, doi:<https://doi.org/10.1016/j.tcs.2021.07.024>. Available at <https://www.sciencedirect.com/science/article/pii/S0304397521004308>.
- 14 Bob Coecke & Aleks Kissinger (2010): *The Compositional Structure of Multipartite Quantum Entanglement*. In Samson Abramsky, Cyril Gavioille, Claude Kirchner, Friedhelm Meyer auf der Heide & Paul G. Spirakis, editors: *Automata, Languages and Programming*, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 297–308. doi:10.1007/978-3-642-14162-1_25.
- 15 Bob Coecke, Aleks Kissinger, Alex Merry & Shibdas Roy (2011): *The GHZ/W-calculus contains rational arithmetic*. *Electronic Proceedings in Theoretical Computer Science* 52, p. 34–48, doi:10.4204/eptcs.52.4. Available at <http://dx.doi.org/10.4204/EPTCS.52.4>.
- 16 Ross Duncan, Aleks Kissinger, Simon Perdrix & John Van De Wetering (2020): *Graph-theoretic Simplification of Quantum Circuits with the ZX-calculus*. *Quantum* 4, p. 279.
- 17 Ross Duncan & Simon Perdrix (2010): *Rewriting measurement-based quantum computations with generalised flow*. In: *International Colloquium on Automata, Languages, and Programming*, Springer, pp. 285–296, doi:10.1007/978-3-642-14162-1_24.
- 18 Stefano Gogioso & William Zeng (2019): *Generalised Mermin-type non-locality arguments*. *Logical Methods in Computer Science* Volume 15, Issue 2, doi:10.23638/LMCS-15(2:3)2019. Available at <https://lmcs.episciences.org/5402>.
- 19 Amar Hadzihasanovic (2015): *A Diagrammatic Axiomatisation for Qubit Entanglement*. In: *2015 30th Annual ACM/IEEE Symposium on Logic in Computer Science*, pp. 573–584. doi:10.1109/LICS.2015.59.
- 20 Amar Hadzihasanovic, Kang Feng Ng & Quanlong Wang (2018): *Two Complete Axiomatisations of Pure-state Qubit Quantum Computing*. In: *Proceedings of the 33rd Annual ACM/IEEE Symposium on Logic in Computer Science, LICS '18*, ACM, pp. 502–511. doi:10.1145/3209108.3209128.
- 21 Emmanuel Jeandel, Simon Perdrix & Margarita Veshchezerova (2022): *Addition and Differentiation of ZX-diagrams*. *arXiv preprint arXiv:2202.11386*.
- 22 Dimitri Kartsaklis, Ian Fan, Richie Yeung, Anna Pearson, Robin Lorenz, Alexis Toumi, Giovanni de Felice, Konstantinos Meichanetzidis, Stephen Clark & Bob Coecke (2021): *lambeq: An Efficient High-Level Python Library for Quantum NLP*. *arXiv preprint arXiv:2110.04236*.
- 23 Aleks Kissinger & John van de Wetering (2019): *Universal MBQC with generalised parity-phase interactions and Pauli measurements*. *Quantum* 3, doi:10.22331/q-2019-04-26-134.
- 24 Maciej Lewenstein (1994): *Quantum perceptrons*. *Journal of Modern Optics* 41(12), pp. 2491–2501.
- 25 Jin-Guo Liu, Yi-Hong Zhang, Yuan Wan & Lei Wang (2019): *Variational quantum eigensolver with fewer qubits*. *Phys. Rev. Research* 1, p. 023025, doi:10.1103/PhysRevResearch.1.023025. Available at <https://link.aps.org/doi/10.1103/PhysRevResearch.1.023025>.
- 26 Robin Lorenz, Anna Pearson, Konstantinos Meichanetzidis, Dimitri Kartsaklis & Bob Coecke (2021): *Qnlp in practice: Running compositional models of meaning on a quantum computer*. *arXiv preprint arXiv:2102.12846*.
- 27 Jarrod R McClean, Sergio Boixo, Vadim N Smelyanskiy, Ryan Babbush & Hartmut Neven (2018): *Barren plateaus in quantum neural network training landscapes*. *Nature communications* 9(1), pp. 1–6.

- 28 Anthony Munson, Bob Coecke & Quanlong Wang (2020): *AND-gates in ZX-calculus: spider nest identities and QBC-completeness*. *Proceedings of the 17th International Conference on Quantum Physics and Logic (QPL) 2020*. arXiv:1910.06818.
- 29 John Preskill (2018): *Quantum Computing in the NISQ era and beyond*. *Quantum* 2, p. 79, doi:10.22331/q-2018-08-06-79. Available at <https://doi.org/10.22331/q-2018-08-06-79>.
- 30 Maria Schuld, Ville Bergholm, Christian Gogolin, Josh Izaac & Nathan Killoran (2019): *Evaluating analytic gradients on quantum hardware*. *Physical Review A* 99(3), p. 032331.
- 31 Seyon Sivarajah, Silas Dilkes, Alexander Cowtan, Will Simmons, Alec Edgington & Ross Duncan (2020): *t|ket>: a retargetable compiler for NISQ devices*. *Quantum Science and Technology* 6(1), p. 014003.
- 32 Alexis Toumi, Richie Yeung & Giovanni de Felice (2021): *Diagrammatic Differentiation for Quantum Machine Learning*. In Chris Heunen & Miriam Backens, editors: *Proceedings 18th International Conference on Quantum Physics and Logic, QPL 2021, Gdansk, Poland, and online, 7-11 June 2021, EPTCS 343*, pp. 132–144, doi:10.4204/EPTCS.343.7. Available at <https://doi.org/10.4204/EPTCS.343.7>.
- 33 Quanlong Wang (2020): *An algebraic axiomatisation of ZX-calculus*. *Proceedings of the 17th International Conference on Quantum Physics and Logic (QPL) 2020*. arXiv:1911.06752.
- 34 Quanlong Wang (2020): *Algebraic complete axiomatisation of ZX-calculus with a normal form via elementary matrix operations*. arXiv:2007.13739v3.
- 35 Quanlong Wang (2020): *Completeness of algebraic ZX-calculus over arbitrary commutative rings and semirings*. arXiv:1912.01003v3.
- 36 Quanlong Wang (2021): *A non-anyonic qudit ZW-calculus*. arXiv:2109.11285.
- 37 Quanlong Wang (2021): *Qufinite ZX-calculus: a unified framework of qudit ZX-calculi*. arXiv:2104.06429.
- 38 Richie Yeung (2020): *Diagrammatic Design and Study of Ansatzes for Quantum Machine Learning*. arXiv preprint arXiv:2011.11073.
- 39 Chen Zhao & Xiao-Shan Gao (2021): *Analyzing the barren plateau phenomenon in training quantum neural networks with the ZX-calculus*. *Quantum* 5, p. 466, doi:10.22331/q-2021-06-04-466. Available at <https://doi.org/10.22331/q-2021-06-04-466>.

A Proofs and Lemmas

In this appendix, we include all the lemmas with their proofs which have been essentially existed (up to scalars) in previous papers. The lemmas are given in the order which they appear in this paper.

► Lemma 10.

Proof.

◀

► **Lemma 11.** Suppose $f(\theta)$ is a differentiable real function of θ . Then

$$\frac{\partial}{\partial \theta} \left[\begin{array}{c} \text{f}(\theta) \\ | \end{array} \right] = \begin{array}{c} \boxed{if'(\theta)} \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \text{f}(\theta) \end{array}$$

Proof.

$$\begin{aligned} \frac{\partial}{\partial \theta} \left[\begin{array}{c} \text{f}(\theta) \\ | \end{array} \right] &= \frac{\partial}{\partial \theta} \left[|0\rangle + e^{if(\theta)} |1\rangle \right] = if'(\theta) e^{if(\theta)} |1\rangle \\ &= \begin{array}{c} \boxed{if'(\theta)} \\ | \\ \pi \end{array} \begin{array}{c} \text{f}(\theta) \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \pi \end{array} \stackrel{B3}{=} \begin{array}{c} \boxed{if'(\theta)} \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \text{f}(\theta) \end{array} \end{aligned}$$

► **Lemma 31.** $\begin{array}{c} \boxed{i} \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \pi \end{array} + \begin{array}{c} \boxed{-i} \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \pi \end{array} \stackrel{31}{=} \begin{array}{c} \frac{\pi}{2} \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \pi \end{array}$

Proof. We use the triangle to do the change of basis from $|+\rangle$ and $|1\rangle$ to $|0\rangle$ and $|1\rangle$.

$$\begin{aligned} &\begin{array}{c} \boxed{i} \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \pi \end{array} + \begin{array}{c} \boxed{-i} \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \pi \end{array} \\ &\stackrel{5, Suc}{\stackrel{Zero}{=}} \begin{array}{c} \boxed{i} \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \pi \end{array} + \begin{array}{c} \boxed{-i} \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \pi \end{array} = \begin{array}{c} \boxed{i} \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \pi \end{array} + \begin{array}{c} \boxed{i} \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \pi \end{array} \\ &\stackrel{B1}{\stackrel{B3}{=}} \begin{array}{c} \boxed{i} \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \pi \end{array} + \begin{array}{c} \boxed{i} \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \pi \end{array} = \begin{array}{c} \boxed{i} \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \pi \end{array} \end{aligned}$$

The rest of the proof is identical to the proof of Lemma 32.

► **Lemma 12.** For any complex number a , we have

$$\begin{array}{c} \pi \\ | \\ \boxed{a} \\ | \\ \triangledown \end{array} = \begin{array}{c} \pi \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \boxed{a} \end{array} \begin{array}{c} \pi \\ | \\ \pi \end{array}$$

Proof.

$$\begin{aligned} &\begin{array}{c} \pi \\ | \\ \boxed{a} \\ | \\ \triangledown \end{array} \stackrel{S1}{=} \begin{array}{c} \pi \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \boxed{a} \end{array} \begin{array}{c} \pi \\ | \\ \pi \end{array} \stackrel{B1}{=} \begin{array}{c} \pi \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \boxed{a} \end{array} \begin{array}{c} \pi \\ | \\ \pi \end{array} \stackrel{Zero}{\stackrel{Suc, Ept}{=}} \begin{array}{c} \pi \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \pi \end{array} \\ &\begin{array}{c} \pi \\ | \\ \boxed{a} \\ | \\ \triangledown \end{array} \stackrel{S1}{=} \begin{array}{c} \pi \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \boxed{a} \end{array} \begin{array}{c} \pi \\ | \\ \pi \end{array} \stackrel{B3}{=} \begin{array}{c} \pi \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \pi \end{array} \stackrel{5}{=} \begin{array}{c} \pi \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \pi \end{array} \begin{array}{c} \pi \\ | \\ \pi \end{array} \end{aligned}$$

► **Lemma 13.** Suppose $f(t) = k(t) + ig(t)$, where $k(t)$ and $g(t)$ are differentiable real functions, t belongs to some interval T , $f(t) \neq 0$. Then

$$\frac{\partial}{\partial t} \left[\begin{array}{c} \boxed{f(t)} \\ | \end{array} \right] = \begin{array}{c} \boxed{\frac{f'(t)}{f(t)}} \quad \pi \\ \pi \quad \boxed{f(t)} \end{array}$$

If $f(t_0) = 0$ for some $t_0 \in T$, then

$$\frac{\partial}{\partial t} \left[\begin{array}{c} \boxed{f(t)} \\ | \end{array} \right]_{t_0} = \begin{array}{c} \pi \\ \boxed{f'(t_0)} \end{array}$$

Proof.

$$\begin{aligned} \frac{\partial}{\partial t} \left[\begin{array}{c} \boxed{f(t)} \\ | \end{array} \right] &= \frac{\partial}{\partial t} [|0\rangle + f(t)|1\rangle] = f'(t)|1\rangle \\ &= \begin{array}{c} \boxed{f'(t)} \quad \pi \\ \pi \end{array} \end{aligned}$$

If $f(t) \neq 0$ for all $t \in T$, then

$$\begin{array}{c} \boxed{f'(t)} \quad \pi \\ \pi \end{array} = \begin{array}{c} \boxed{\frac{f'(t)}{f(t)}} \quad \boxed{f(t)} \quad \pi \\ \pi \quad \pi \end{array} = \begin{array}{c} \boxed{\frac{f'(t)}{f(t)}} \quad \pi \\ \pi \quad \boxed{f(t)} \end{array}$$

If $f(t_0) = 0$ for some $t_0 \in T$, then

$$\begin{array}{c} \boxed{f'(t_0)} \quad \pi \\ \pi \end{array} = \begin{array}{c} \pi \\ \boxed{f'(t_0)} \end{array}$$

► **Lemma 32.**

$$\begin{array}{c} \pi \\ \swarrow \quad \searrow \\ \text{triangle} \quad \pi \\ \swarrow \quad \searrow \\ \text{yellow triangle} \quad \text{yellow triangle} \end{array} = \begin{array}{c} -\frac{\pi}{2} \quad \pi \\ \swarrow \quad \searrow \\ \pi \quad -\frac{\pi}{2} \end{array}$$

Proof.

$$\begin{aligned} &\begin{array}{c} \pi \\ \swarrow \quad \searrow \\ \text{triangle} \quad \pi \\ \swarrow \quad \searrow \\ \text{yellow triangle} \quad \text{yellow triangle} \end{array} \stackrel{10}{=} \begin{array}{c} \pi \\ \swarrow \quad \searrow \\ \text{yellow triangle} \quad \pi \\ \swarrow \quad \searrow \\ \text{yellow triangle} \quad \text{yellow triangle} \end{array} \stackrel{3}{=} \begin{array}{c} \pi \quad \pi \\ \swarrow \quad \searrow \\ \text{yellow triangle} \quad \text{yellow triangle} \end{array} \stackrel{4}{=} \begin{array}{c} \pi \quad \pi \\ \swarrow \quad \searrow \\ \text{yellow triangle} \quad \text{yellow triangle} \end{array} \stackrel{Inv}{=} \begin{array}{c} \pi \quad \pi \\ \swarrow \quad \searrow \\ \text{yellow triangle} \quad \text{yellow triangle} \end{array} \\ &\stackrel{8}{=} \begin{array}{c} \pi \quad -\pi \\ \swarrow \quad \searrow \\ \text{yellow triangle} \quad \text{yellow triangle} \end{array} = \begin{array}{c} -\frac{\pi}{2} \quad -\frac{\pi}{2} \\ \swarrow \quad \searrow \\ \pi \quad \pi \end{array} \stackrel{8}{=} \begin{array}{c} \pi \quad -\frac{\pi}{2} \\ \swarrow \quad \searrow \\ \pi \quad \pi \end{array} \end{aligned}$$

► **Lemma 33.**

$$\text{Diagram with two green vertices connected by two arcs and a red circle labeled } \pi \text{ in the middle} = \text{Diagram with a red circle labeled } \pi \text{ on a horizontal line}$$

Proof.

► **Lemma 34.**

The diagrammatic equation (3.10) shows two diagrams separated by an equals sign. The left diagram features a central red circle labeled π connected to two green circles. Each green circle is part of a loop structure with a vertical line extending upwards. The right diagram shows a similar structure, but the top vertical line is now connected to a red circle labeled π , which is in turn connected to the central red circle labeled π .

Proof.

The figure shows a sequence of six diagrams representing the reduction of a Feynman diagram. The diagrams are connected by equivalence symbols: $S1$, Pic , $S1$, and $Hopf$. The diagrams consist of vertices (green and red) and lines (curved and straight) with labels like π and ellipsis.

► **Lemma 35.**

The diagram shows an equality between two graph structures. On the left, two green nodes are connected by a horizontal line. Above the left green node is a red node labeled π , and above the right green node is another red node labeled π . Curved lines extend downwards from each green node, representing connections to other parts of the graph. On the right, the two red π nodes have been contracted into a single red node labeled π positioned in the middle of the horizontal line connecting the two green nodes. The curved lines extending from the green nodes remain the same.

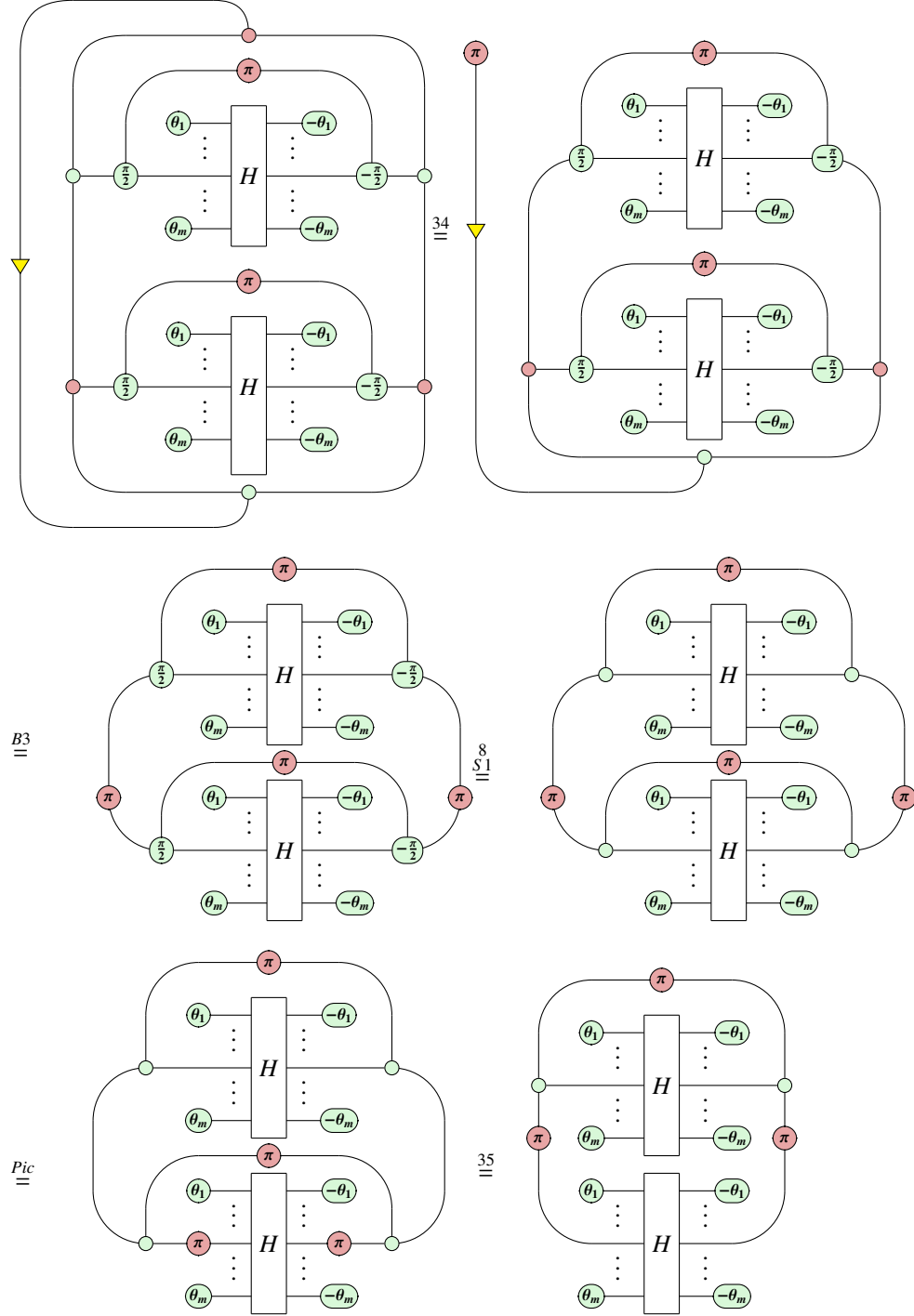
Proof.

The figure shows a sequence of five graphs connected by reduction rules. The first graph has two green nodes at the bottom connected by a horizontal line, with two red nodes labeled π above them. The second graph, reached via rule $S1$, has a green node on the left, a red node π above it, and a green node on the right with a red node π above it. The third graph, reached via rule Pic , has a green node on the left, a red node π above it, and a green node on the right with a red node π below it. The fourth graph, reached via rule $S1$, has a green node on the left, a red node π above it, and a green node on the right with a red node π below it. The fifth graph, reached via rule $S1$, has a green node on the left, a red node π above it, and a green node on the right with a red node π above it. Ellipses (...) are used to indicate additional nodes and edges in the first and fifth graphs.

► **Lemma 26.**

Proof. By theorem 21,

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\frac{\partial \langle H \rangle}{\partial \theta_j} \right)^2 d\theta_j =$$



B

 Integration proofs

► **Proposition 19.** Let k be an arbitrary non-zero integer, ${}^{m+1}\left\{ \begin{array}{c} \vdots \\ K \\ \vdots \end{array} \right\}^{n+1}$ be any ZX diagram with $m, n \geq 0$ which has no occurrence of α . Then

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} {}^m \left\{ \begin{array}{c} \overset{k\alpha}{\vdots} \\ K \\ \underset{-k\alpha}{\vdots} \end{array} \right\}^n d\alpha = {}^m \left\{ \begin{array}{c} \vdots \\ K \\ \vdots \end{array} \right\}^n$$

Proof.

$$\begin{aligned} {}^m \left\{ \begin{array}{c} \overset{k\alpha}{\vdots} \\ K \\ \underset{-k\alpha}{\vdots} \end{array} \right\}^n &= \left((\langle 0| + e^{-ik\alpha} \langle 1|) \otimes I^{\otimes n} \right) K \left((|0\rangle + e^{ik\alpha} |1\rangle) \otimes I^{\otimes m} \right) \\ &= (\langle 0| \otimes I^{\otimes n}) K (|0\rangle \otimes I^{\otimes m}) + e^{ik\alpha} (\langle 0| \otimes I^{\otimes n}) K (|1\rangle \otimes I^{\otimes m}) \\ &\quad + e^{-ik\alpha} (\langle 1| \otimes I^{\otimes n}) K (|0\rangle \otimes I^{\otimes m}) + (\langle 1| \otimes I^{\otimes n}) K (|1\rangle \otimes I^{\otimes m}) \end{aligned}$$

Since $k \neq 0$, we have $\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ik\alpha} d\alpha = 0$ for all k , therefore

$$\begin{aligned} \frac{1}{2\pi} \int_{-\pi}^{\pi} {}^m \left\{ \begin{array}{c} \overset{k\alpha}{\vdots} \\ K \\ \underset{-k\alpha}{\vdots} \end{array} \right\}^n d\alpha &= (\langle 0| \otimes I^{\otimes n}) K (|0\rangle \otimes I^{\otimes m}) + (\langle 1| \otimes I^{\otimes n}) K (|1\rangle \otimes I^{\otimes m}) \\ &= {}^m \left\{ \begin{array}{c} \bullet \\ \vdots \\ K \\ \vdots \\ \bullet \end{array} \right\}^n + {}^m \left\{ \begin{array}{c} \pi \\ \vdots \\ K \\ \vdots \\ \pi \end{array} \right\}^n = {}^m \left\{ \begin{array}{c} \bullet \\ \vdots \\ K \\ \vdots \\ \bullet \end{array} \right\}^n + {}^m \left\{ \begin{array}{c} \pi \\ \vdots \\ K \\ \vdots \\ \pi \end{array} \right\}^n \\ &= {}^m \left\{ \begin{array}{c} \bullet \\ \vdots \\ K \\ \vdots \\ \bullet \end{array} \right\}^n + {}^m \left\{ \begin{array}{c} \pi \\ \vdots \\ K \\ \vdots \\ \pi \end{array} \right\}^n \\ &= {}^m \left\{ \begin{array}{c} \bullet \\ \vdots \\ K \\ \vdots \\ \bullet \end{array} \right\}^n = {}^m \left\{ \begin{array}{c} \vdots \\ K \\ \vdots \end{array} \right\}^n \end{aligned}$$

► **Lemma 36.**

$$\begin{array}{c} \bullet \\ \bullet \end{array} + \begin{array}{c} \pi \\ \bullet \end{array} = \begin{array}{c} \bullet \\ \bullet \end{array}$$

Proof.

$$\text{Diagram 1} + \text{Diagram 2} \stackrel{\text{Bas0}}{\stackrel{\text{Bas1}}{=}} \text{Diagram 3} + \text{Diagram 4} = \text{Diagram 5} = \text{Diagram 6}$$

► **Theorem 21.** Let k be an arbitrary non-zero integer, ${}_{m+2} \left\{ \begin{array}{c} \vdots \\ A \\ \vdots \end{array} \right\}_{n+2}$ be any ZX diagram with $m, n \geq 0$ which has no occurrence of α . Then

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} {}_m \left\{ \begin{array}{c} k\alpha \\ k\alpha \\ \vdots \\ \vdots \end{array} \right\} A \left\{ \begin{array}{c} -k\alpha \\ -k\alpha \\ \vdots \\ \vdots \end{array} \right\}_n d\alpha = \text{Diagram with } \pi \text{ and } A$$

Proof.

$$\begin{aligned} F &= {}_m \left\{ \begin{array}{c} k\alpha \\ k\alpha \\ \vdots \\ \vdots \end{array} \right\} A \left\{ \begin{array}{c} -k\alpha \\ -k\alpha \\ \vdots \\ \vdots \end{array} \right\}_n \\ &= \left[(\langle 0| + e^{-ik\alpha} \langle 1|) \otimes (\langle 0| + e^{-ik\alpha} \langle 1|) \otimes I^{\otimes n} \right] A \left[(|0\rangle + e^{ik\alpha} |1\rangle) \otimes (|0\rangle + e^{ik\alpha} |1\rangle) \otimes I^{\otimes m} \right] \\ &= \left[(\langle 00| + e^{-ik\alpha} (\langle 01| + \langle 10|) + e^{-i2k\alpha} \langle 11|) \otimes I^{\otimes n} \right] A \left[(|00\rangle + e^{ik\alpha} (|01\rangle + |10\rangle) + e^{i2k\alpha} |11\rangle) \otimes I^{\otimes m} \right] \\ &= (\langle 00| \otimes I^{\otimes n}) A (|00\rangle \otimes I^{\otimes m}) + (\langle 11| \otimes I^{\otimes n}) A (|11\rangle \otimes I^{\otimes m}) \\ &\quad + ((\langle 01| + \langle 10|) \otimes I^{\otimes n}) A ((|01\rangle + |10\rangle) \otimes I^{\otimes m}) \\ &\quad + e^{ik\alpha} (\langle 00| \otimes I^{\otimes n}) A ((|01\rangle + |10\rangle) \otimes I^{\otimes m}) + e^{i2k\alpha} (\langle 00| \otimes I^{\otimes n}) A (|11\rangle \otimes I^{\otimes m}) \\ &\quad + e^{-ik\alpha} ((\langle 01| + \langle 10|) \otimes I^{\otimes n}) A (|00\rangle \otimes I^{\otimes m}) + e^{-ik\alpha} (\langle 11| \otimes I^{\otimes n}) A ((|01\rangle + |10\rangle) \otimes I^{\otimes m}) \\ &\quad + e^{-i2k\alpha} (\langle 11| \otimes I^{\otimes n}) A (|00\rangle \otimes I^{\otimes m}) + e^{ik\alpha} ((\langle 01| + \langle 10|) \otimes I^{\otimes n}) A (|11\rangle \otimes I^{\otimes m}) \end{aligned}$$

Since $k \neq 0$, we have $\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ik\alpha} d\alpha = 0$ for all k , therefore

$$\begin{aligned} \frac{1}{2\pi} \int_{-\pi}^{\pi} F d\alpha &= (\langle 00| \otimes I^{\otimes n}) A (|00\rangle \otimes I^{\otimes m}) + (\langle 11| \otimes I^{\otimes n}) A (|11\rangle \otimes I^{\otimes m}) \\ &\quad + ((\langle 01| + \langle 10|) \otimes I^{\otimes n}) A ((|01\rangle + |10\rangle) \otimes I^{\otimes m}) \end{aligned}$$

$$\begin{aligned}
&= \left\{ \begin{array}{c} \text{red dots} \\ \vdots \\ \text{red dots} \end{array} \right\}_m \left\{ \begin{array}{c} \vdots \\ A \\ \vdots \end{array} \right\}_n + \left\{ \begin{array}{c} \pi \\ \vdots \\ \pi \end{array} \right\}_m \left\{ \begin{array}{c} \vdots \\ A \\ \vdots \end{array} \right\}_n + \left\{ \begin{array}{c} \pi \\ \vdots \\ \pi \end{array} \right\}_m \left\{ \begin{array}{c} \vdots \\ A \\ \vdots \end{array} \right\}_n \\
&= \left\{ \begin{array}{c} \text{red dots} \\ \vdots \\ \text{red dots} \end{array} \right\}_m \left\{ \begin{array}{c} \vdots \\ A \\ \vdots \end{array} \right\}_n + \left\{ \begin{array}{c} \pi \\ \vdots \\ \pi \end{array} \right\}_m \left\{ \begin{array}{c} \vdots \\ A \\ \vdots \end{array} \right\}_n + \left\{ \begin{array}{c} \pi \\ \vdots \\ \pi \end{array} \right\}_m \left\{ \begin{array}{c} \vdots \\ A \\ \vdots \end{array} \right\}_n \\
&= \left\{ \begin{array}{c} \text{red dots} \\ \vdots \\ \text{red dots} \end{array} \right\}_m \left\{ \begin{array}{c} \vdots \\ A \\ \vdots \end{array} \right\}_n + \left\{ \begin{array}{c} \pi \\ \vdots \\ \pi \end{array} \right\}_m \left\{ \begin{array}{c} \vdots \\ A \\ \vdots \end{array} \right\}_n + \left\{ \begin{array}{c} \pi \\ \vdots \\ \pi \end{array} \right\}_m \left\{ \begin{array}{c} \vdots \\ A \\ \vdots \end{array} \right\}_n \\
&= \left\{ \begin{array}{c} \text{red dots} \\ \vdots \\ \text{red dots} \end{array} \right\}_m \left\{ \begin{array}{c} \vdots \\ A \\ \vdots \end{array} \right\}_n + \left\{ \begin{array}{c} \pi \\ \vdots \\ \pi \end{array} \right\}_m \left\{ \begin{array}{c} \vdots \\ A \\ \vdots \end{array} \right\}_n \\
&= \left\{ \begin{array}{c} \text{red dots} \\ \vdots \\ \text{red dots} \end{array} \right\}_m \left\{ \begin{array}{c} \vdots \\ A \\ \vdots \end{array} \right\}_n + \left\{ \begin{array}{c} \pi \\ \vdots \\ \pi \end{array} \right\}_m \left\{ \begin{array}{c} \vdots \\ A \\ \vdots \end{array} \right\}_n \stackrel{36}{=} \left\{ \begin{array}{c} \text{red dots} \\ \vdots \\ \text{red dots} \end{array} \right\}_m \left\{ \begin{array}{c} \vdots \\ A \\ \vdots \end{array} \right\}_n
\end{aligned}$$

► **Theorem 23.** Let k be an arbitrary non-zero integer, $\left\{ \begin{array}{c} \vdots \\ A \\ \vdots \end{array} \right\}_{m+3}$ be any ZX diagram with $m, n \geq 0$ which has no occurrence of α . Then

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \left\{ \begin{array}{c} k\alpha \\ k\alpha \\ k\alpha \\ \vdots \end{array} \right\}_m \left\{ \begin{array}{c} \vdots \\ A \\ \vdots \end{array} \right\}_n d\alpha = \left\{ \begin{array}{c} \text{red dots} \\ \vdots \\ \text{red dots} \end{array} \right\}_m \left\{ \begin{array}{c} \vdots \\ A \\ \vdots \end{array} \right\}_n$$

Proof. Similar to the proof of Theorem 21, we have

$$\begin{aligned}
 & \frac{1}{2\pi} \int_{-\pi}^{\pi} d\alpha = \left(\langle 000 | \otimes I^{\otimes n} \right) A \left(|000\rangle \otimes I^{\otimes m} \right) + \left(\langle 111 | \otimes I^{\otimes n} \right) A \left(|111\rangle \otimes I^{\otimes m} \right) \\
 & + \left((\langle 001 | + \langle 010 | + \langle 100 |) \otimes I^{\otimes n} \right) A \left((|001\rangle + |010\rangle + |100\rangle) \otimes I^{\otimes m} \right) \\
 & + \left((\langle 011 | + \langle 101 | + \langle 110 |) \otimes I^{\otimes n} \right) A \left((|011\rangle + |101\rangle + |110\rangle) \otimes I^{\otimes m} \right) \\
 & = \text{[Diagrammatic expansion showing 16 terms with various gate configurations and ancilla qubits]} \\
 & = \text{[Diagrammatic expansion showing 4 terms with more complex gate configurations]} \\
 & = \text{[Diagrammatic expansion showing 2 terms with even more complex gate configurations]} \\
 & = \text{[Final diagrammatic representation of the integral]}
 \end{aligned}$$

C Barren Plateau Analysis

► **Lemma 25.** Given $\langle H \rangle$ in the form of (2), we have $\mathbf{E} \left(\frac{\partial \langle H \rangle}{\partial \theta_j} \right) = 0$, for $j = 1, \dots, m$.

Proof. By integrating over the uniform distribution, we have

$$\mathbf{E}\left(\frac{\partial \langle H \rangle}{\partial \theta_j}\right) = \frac{1}{(2\pi)^m} \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} \frac{\partial \langle H \rangle}{\partial \theta_j} d\theta_1 \dots d\theta_m.$$

We prove the theorem by showing that $\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\partial \langle H \rangle}{\partial \theta_j} d\theta_j = 0$.

$$\begin{aligned} \frac{\partial \langle H \rangle}{\partial \theta_j} &\stackrel{17}{=} \text{Diagram 1} \\ \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\partial \langle H \rangle}{\partial \theta_j} d\theta_j &\stackrel{25}{=} \text{Diagram 2} \\ &\stackrel{33}{=} \text{Diagram 3} = 0 \end{aligned}$$

Since the unconnected pink π spider is equal to the zero scalar, the entire diagram evaluates to zero. \blacktriangleleft

C.1 Variance of Gradient Calculation

We start with some simple lemmas that will aid with this specific calculation:

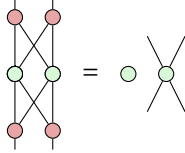
► **Lemma 37.**

$$\text{Diagram A} = \text{Diagram B}$$

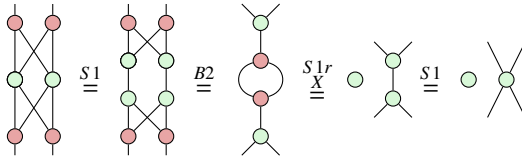
Proof.

$$\begin{aligned} &\text{Diagram 1} \stackrel{Pic}{=} \text{Diagram 2} \stackrel{Pic}{=} \text{Diagram 3} \\ &\stackrel{S1}{=} \text{Diagram 4} \stackrel{S1}{=} \text{Diagram 5} \stackrel{S1r}{=} \text{Diagram 6} \stackrel{S1.H}{=} \text{Diagram 7} \stackrel{Pic}{=} \text{Diagram 8} \end{aligned}$$

► **Lemma 38.**

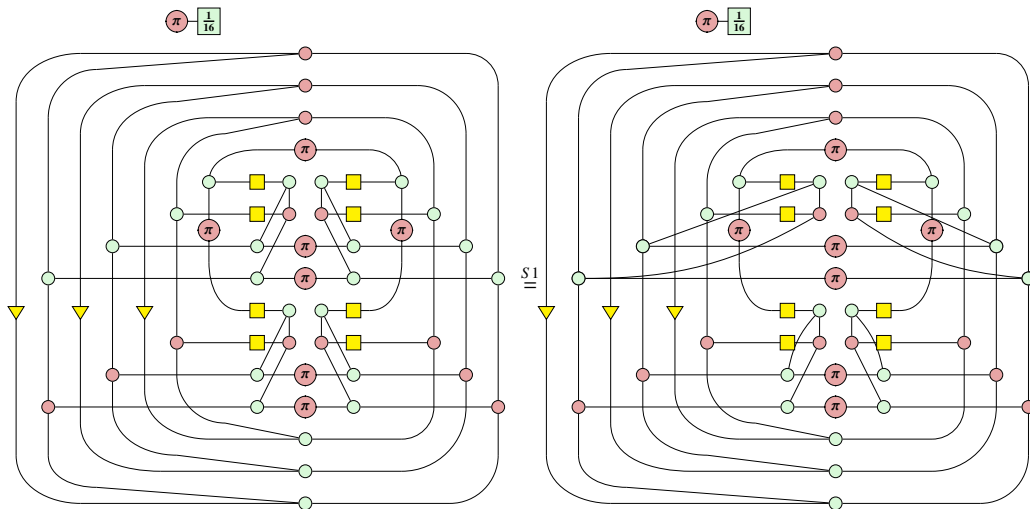


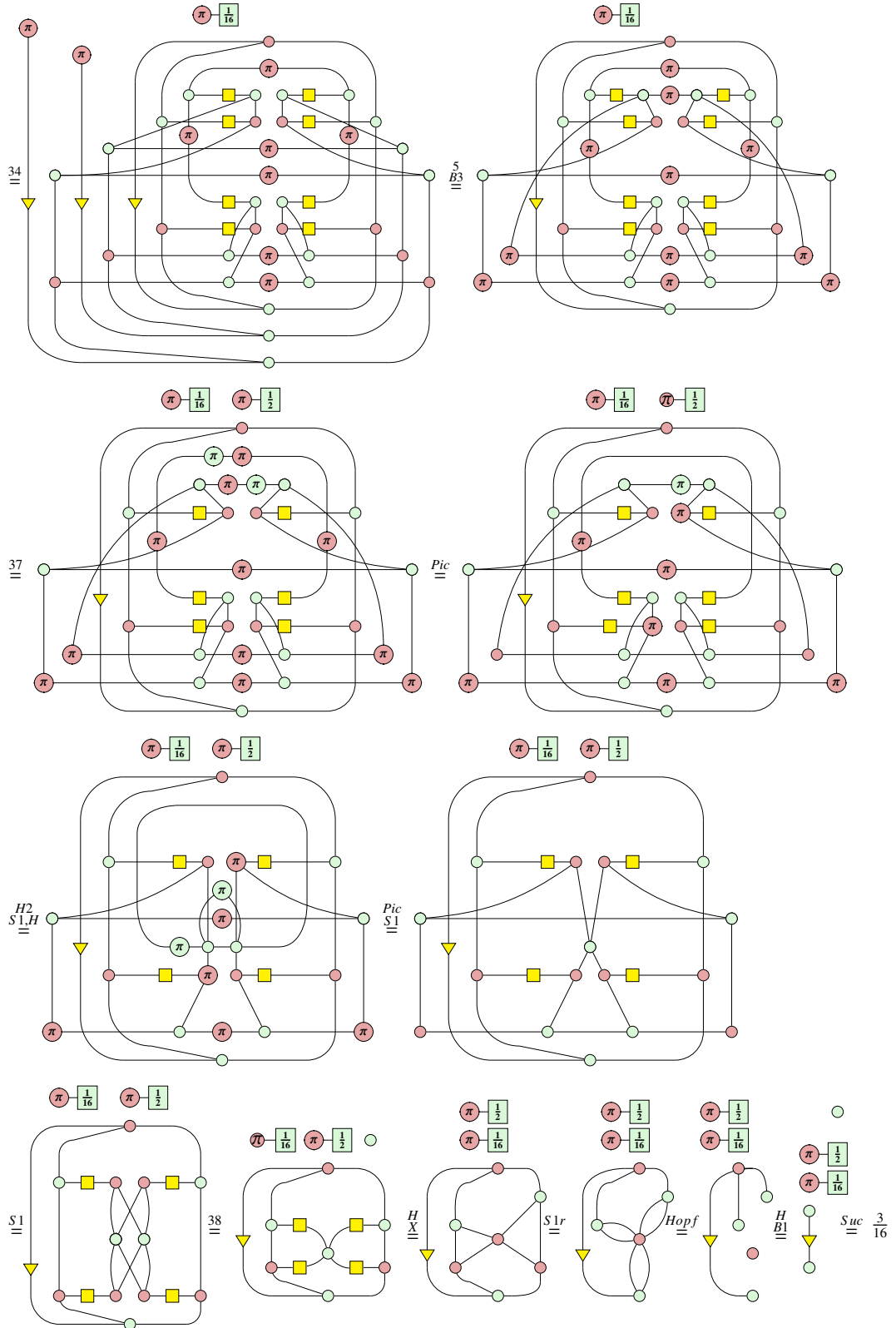
Proof.



Start from Lemma 26 and Theorem 28, we have

$$\text{Var}\left(\frac{\partial \langle H \rangle}{\partial \theta_1}\right) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\frac{\partial \langle H \rangle}{\partial \theta_1}\right)^2 d\theta_1 d\theta_2 d\theta_3 d\theta_4 =$$





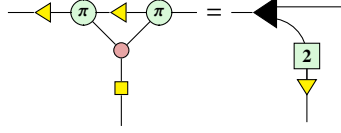
► Remark 39. The corresponding result in [39] on page 30 shows that $\mathbf{Var}\left(\frac{\partial(H)}{\partial\theta_1}\right) = \frac{3}{64}$ which is different from our result here.

D Comparison with Jeandel et al.

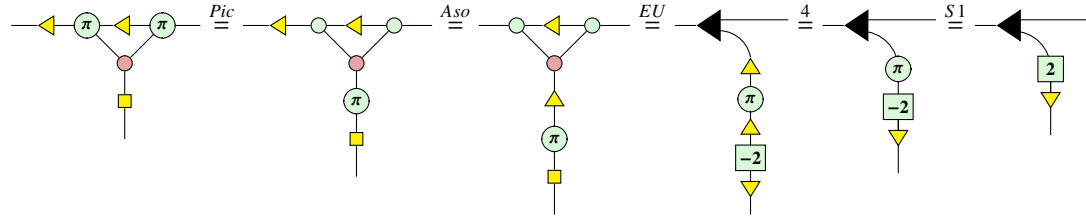
We want to show that the two papers arrive at similar results through very different techniques, and that some results in [21] can be more compactly represented by explicitly using the W spider. The nodes used in [21] are similar to ours: the circle green nodes, the Hadamard node and the red nodes are exactly the same, but we additionally have pink nodes which are different to the red nodes up to a variational scalar, and have green box nodes which have the circle green nodes as special cases while can be turned into circle green nodes with the help of the yellow triangle node [20]. Also note that their black triangle corresponds to our yellow triangle and our black triangle corresponds to the W spider.

First, we show that the triangular-shaped diagram that features in their differentiation result can be represented compactly using the W spider:

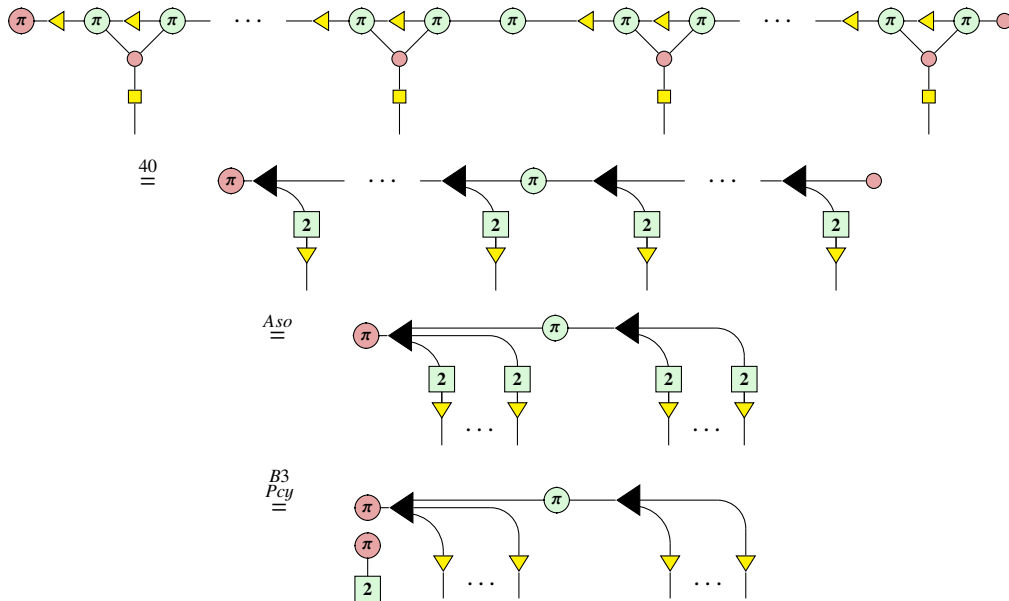
► **Lemma 40.**

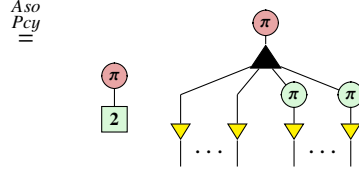


Proof.



► **Remark 41.** The differentiation of “linear” ZX diagrams result obtained by [21] is equivalent to our lemma 16.



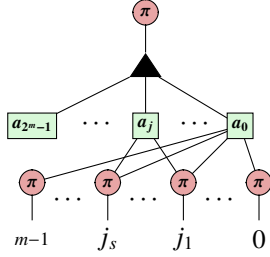


Although the end results are equivalent, we emphasise that their result is obtained through their theory of summing controlled diagrams, whilst our result is obtained through our arbitrary differentiation result.

► **Remark 42.** The general differentiation result in [21] requires an inductive conversion to controlled diagrams, and is obtained through combining diagrammatic addition and the Leibniz product rule. Because of this, the resulting diagram will not resemble the original diagram.

In comparison, our Theorem 14 does not affect the topology of the original diagram and can be calculated almost instantly.

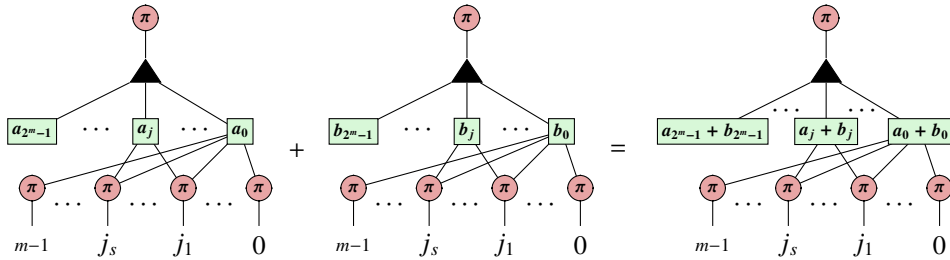
► **Remark 43.** Both the general results on addition of diagrams and on differentiation of diagrams shown in [21] are based on induction on generators. If such induction methods are allowed, then there is an alternative way to generally add two ZX diagrams or differentiate a ZX diagram: first rewrite inductively the diagrams into the compressed normal form shown in [34] as follows



which corresponding to the vector

$$\begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{2^m-2} \\ a_{2^m-1} \end{pmatrix},$$

then the sum of two diagrams can be obtained by adding up the corresponding parameters in the two diagrams:



and the differentiation can be obtained element-wisely:

