

ADDITION AND DIFFERENTIATION OF ZX-DIAGRAMS

EMMANUEL JEANDEL ^{a,b,c,d}, SIMON PERDRIX ^{a,b,c,d}, AND MARGARITA VESHCHEREROVA ^e

^a LORIA, Campus scientifique, BP 239, 54506, Vandoeuvre-lès-Nancy Cedex, France
e-mail address: emmanuel.jeandel@loria.fr, simon.perdrix@loria.fr

^b CNRS, Délégation Centre-Est, 17 rue Notre-Dame des Pauvres, BP 10075, 54519, Vandoeuvre-lès-Nancy Cedex, France

^c Université de Lorraine, 34 cours Léopold, 54052, Nancy Cedex, France

^d Inria MOCQUA, 615 Rue du Jardin-Botanique, 54600 Villers-lès-Nancy, France

^e Terra Quantum AG, St. Gallerstrasse 16A, 9400 Rorschach, Switzerland
e-mail address: mv@terraquantum.swiss

ABSTRACT. The ZX-calculus is a powerful framework for reasoning in quantum computing. It provides in particular a compact representation of matrices of interest. A peculiar property of the ZX-calculus is the absence of a formal sum allowing the linear combinations of arbitrary ZX-diagrams. The universality of the formalism guarantees however that for any two ZX-diagrams, the sum of their interpretations can be represented by a ZX-diagram. We introduce a general, inductive definition of the addition of ZX-diagrams, relying on the construction of controlled diagrams. Based on this addition technique, we provide an inductive differentiation of ZX-diagrams.

Indeed, given a ZX-diagram with variables in the description of its angles, one can differentiate the diagram according to one of these variables. Differentiation is ubiquitous in quantum mechanics and quantum computing (e.g. for solving optimization problems). Technically, the differentiation of ZX-diagrams is strongly related to summation as witnessed by the product rules.

We also introduce an alternative, non-inductive, differentiation technique rather based on the isolation of the variables. Finally, we apply our results to deduce a diagram for an Ising Hamiltonian.

1. INTRODUCTION

ZX-calculus, originally introduced in [CD11], is a *graphical language* that allows reasoning about quantum computing. In this language complex computations on qubits are represented with *diagrams* made out of *elementary generators*¹. Each diagram corresponds to a linear transformation between Hilbert spaces of qubit states. A compact set of *rewrite rules* allows to transform diagrams into equivalent ones. The notable advantage of ZX-calculus compared

Key words and phrases: ZX-calculus, Addition of ZX-diagrams, Differentiation of ZX-diagrams, Diagrammatic differentiation.

The work is an extended version of [JPV22].

¹We use the TikZit software <https://tikzit.github.io> to draw ZX-diagrams

to other representations (e.g. linear maps, circuits, and tensor networks) is that in this language the computations may be done *entirely graphically*. A general introduction to the language alongside the overview of the main applications is available in [vdW20].

Due to its flexibility, ZX-calculus is widely used to address different problems of quantum computing. For instance, ZX-calculus allowed the derivation of important results in the field of *measurement-based quantum computing (MBQC)* [DP10, KvdW19]. It was also successfully applied to circuit optimization [CDD⁺20, dBBW20, KvdW20, DKPvdW20], design and verification of error-correction color codes [GD18], and to the analysis and compilation of *surface codes* [Hor11, dBH20, HEMN20].

However, the applications of the ZX-calculus to the rapidly growing field of variational algorithms [CAB⁺21] such as *Quantum Approximate Optimization Algorithm (QAOA)* [FGG14], *Variational Quantum Eigensolver (VQE)* [PMS⁺14] and variational quantum machine learning are so far limited. Nevertheless, as variational algorithms do not require error correction, the incoming emergence of *NISQ* devices makes them an object of particular attention [Pre18]. We believe that the reason why they are still unexplored with the means of ZX-calculus is the absence of a convenient way to **differentiate parametrized diagrams**. Indeed, the basic building blocks of variational algorithms are parametrized circuits, and the search for optimal parameter values is a crucial part of these algorithms. In theory, finding optimal parameters may be NP-hard [BK21]. In practice, the search is usually done by classical numerical optimization methods and most of them use derivatives [GS17].

The main difficulty for differentiation of ZX-diagrams comes from the *product rule*: $\partial fg = \partial f g + f \partial g$ that involves adding two terms. This rule is crucial to evaluate the derivative of the sequential composition and the tensor product. As ZX-diagrams are defined inductively with these two compositions, the derivative of a complex diagram could be computed out of the derivatives of its parts if we were allowed to use the product rule, and therefore sum diagrams.

The works [ZG21, TYd21] use explicit sums of diagrams (usually referred to as *bags of diagrams*) to represent the derivative of diagrams with multiple occurrences of the parameter. The major disadvantage of this approach is that there are no rules to manipulate sums of ZX-diagrams. Therefore, we can't fully exploit the power of graphical computation (while it's still possible to use rewrite rules on summands [SH22]). In this work, we suggest an approach where *the derivative of a parametrized ZX-diagram is another ZX-diagram*. Hence we avoid the extension of the signature with formal sums. In order to tackle sums that appear in the product rule, we introduce an original technique to perform the addition of diagrams entirely in the ZX-calculus. For this purpose, we use special diagrams called controlled states [JPV19]. We suggest an inductive way to represent every ZX-diagram by such a state. As we know how to sum controlled states [JPV19] the addition for arbitrary diagrams follows. An inductive definition of the derivative is obtained by an explicit diagrammatic representation of the product rules.

In an attempt to give a ready-to-use toolbox for differentiation, we provide an easy and convenient way to compute the derivative for the family of linear diagrams $ZX(\beta)$ [JPV19], that is diagrams where angles might depend linearly (with integer coefficients) on one parameter β . Most of the circuits that are used for variational algorithms belong to $ZX(\beta)$ and we believe that our formulas will make their analysis much simpler.

A definition for a derivative similar to our formulas was obtained in the independent work [WY22]. This work uses *W-spiders* [Had15] to handle product rules. In contrast to our result, the diagrammatic differentiation presented in [WY22] maps to a ZX-diagram a diagram

from another language called *algebraic ZX-calculus* [Wan20]. The algebraic ZX-calculus is convenient to represent arbitrary complex numbers, therefore their differentiation procedure can handle more general families of diagrams than $ZX(\beta)$. This advantage comes with the cost of abandoning the legacy of the vanilla ZX-calculus which is by far the most popular graphical calculus for quantum computing.



In the end, we show how our result together with the Stone's theorem [Sto32] allows finding a ZX-diagram representing an *Ising Hamiltonian*. In variational algorithms Ising Hamiltonians are typically used in the definition of the loss function for parameter optimization. In practical applications Hamiltonians are usually given as a sum of local terms $H = \sum_{i=1}^k H_i$ where each H_i acts only on a limited number of qubits. Small individual terms H_i can be relatively easily represented as diagrams. For instance, for the Hamiltonian $H = \sum_{u,v} Z_u Z_v$ diagrams for individual terms are trivial: $H_{u,v} = \begin{array}{c} \text{---} \pi \text{---} \\ \text{---} \pi \text{---} \end{array}$. However, in the traditional framework there is no graphical procedure that constructs the overall sum H out of such elementary blocks.

The ability to perform addition directly leads to a procedure for differentiation. On the other direction, with Stone's theorem some specific kinds of sums might be straightforwardly expressed as diagrams if we can compute derivatives. We demonstrate how to combine our formula for the derivatives and the theorem to find a diagrammatic representation for example Hamiltonian.

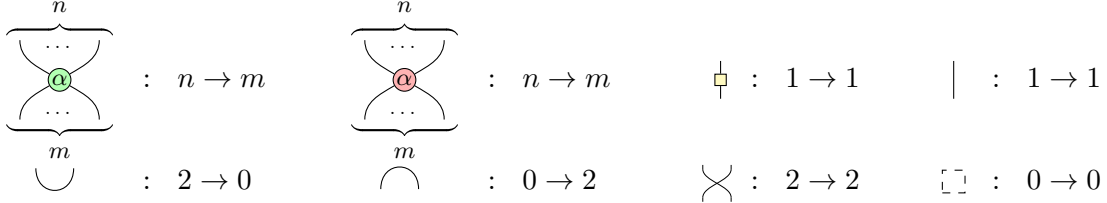
Structure of the paper. In Section 2, we give a brief introduction to the ZX-calculus followed by some useful lemmas. In Section 3, we provide a general overview of addition and differentiation in quantum computing. In particular, we show how addition may be represented in circuit notation. We continue by giving an overview of the potential benefits coming from using ZX-diagrams to reason about derivatives. This section, alongside additional examples and explanations, constitutes the major extension of the current work with respect to the previous version [JPV22].

In the next Section 4 we introduce an inductive procedure for the transformation of a diagram into a controlled form. We show how this procedure leads to an algorithm for the addition of ZX-diagrams. In Section 5, we provide two approaches for the differentiation, including an inductive definition (directly inspired by the product rules) and compact formulas for derivatives of linear diagrams $ZX(\beta)$ where β is a vector of parameters. In Section 6, we show how to apply our formulas to obtain a diagram for an Ising Hamiltonian.

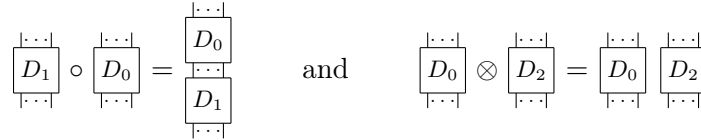
2. ZX-CALCULUS

2.1. Syntax and Semantics. The ZX-diagrams are generated by green spiders , red spiders  and Hadamard \dagger , where both kinds of spiders have an arbitrary number of inputs/outputs and are decorated with *angles*. ZX-diagrams are also made of wires: the identity $|$, the swap \times and also the possibility to bend wires with a cup \cup and a cap \cap . A wire can connect two nodes or, alternatively, have free ends that point either up or down. Wires with free ends pointing towards the top (towards the bottom) of the diagram are called *inputs* (*outputs*). We denote by $D : n \rightarrow m$ a diagram with n inputs and m outputs. Finally, the empty diagram is denoted with $[\]$.

Definition 2.1. ZX-diagrams are inductively defined as follows: for $n, m \in \mathbb{N}$ and $\alpha \in \mathbb{R}/2\pi\mathbb{Z}$,



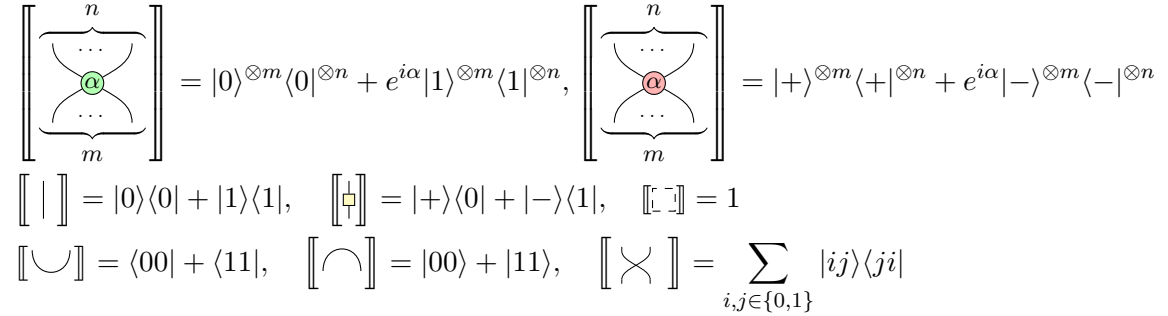
are ZX-diagrams, and for any ZX-diagrams $D_0 : a \rightarrow b$, $D_1 : b \rightarrow c$, and $D_2 : c \rightarrow d$, $D_1 \circ D_0 : a \rightarrow c$ and $D_0 \otimes D_2 : a + c \rightarrow b + d$ are ZX-diagrams. Pictorially:



As in most works on ZX-calculus, an empty spider denotes the spider with $\alpha = 0$. A diagram with no input/output is called a *scalar*. To compactly write scalar factors, we introduce syntactic sugar $[-]^{\otimes n}$. For any scalar $d : 0 \rightarrow 0$ the notation $d^{\otimes n}$ corresponds to $\underbrace{d \otimes \dots \otimes d}_n$.

Semantically, ZX-diagrams are standardly interpreted as linear maps, and thus they can be used to represent quantum evolutions.

Definition 2.2. For any ZX-diagram $D : n \rightarrow m$, let $\llbracket D \rrbracket \in \mathcal{M}_{2^m, 2^n}(\mathbb{C})$ be inductively defined as: $\llbracket D_1 \circ D_0 \rrbracket = \llbracket D_1 \rrbracket \circ \llbracket D_0 \rrbracket$, $\llbracket D_0 \otimes D_2 \rrbracket = \llbracket D_0 \rrbracket \otimes \llbracket D_2 \rrbracket$, and



where *bra-ket* notations are used: $|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $|+\rangle = \frac{|0\rangle+|1\rangle}{\sqrt{2}}$, $|-\rangle = \frac{|0\rangle-|1\rangle}{\sqrt{2}}$, $|xy\rangle = |x\rangle \otimes |y\rangle$ and $\langle x| = |x\rangle^\dagger$ is the adjoint (complex conjugate) of $|x\rangle$.

Example 2.3. Complex numbers can be represented with diagrams:

$$\llbracket \bigcirc \rrbracket = 2, \quad \llbracket \textcircled{\alpha} \rrbracket = 1 + e^{i\alpha}, \quad \llbracket \begin{array}{c} \textcircled{\pi} \\ \textcircled{\alpha} \end{array} \rrbracket = \sqrt{2}, \quad \llbracket \begin{array}{c} \textcircled{\pi} \\ \textcircled{\alpha} \end{array} \rrbracket = \sqrt{2}e^{i\alpha}, \quad \llbracket \begin{array}{c} \textcircled{\pi} \\ \textcircled{\alpha} \end{array} \rrbracket = \frac{1}{\sqrt{2}} \quad (2.1)$$

Sometimes it is meaningful to consider diagrams with angles from a restricted sub-group \mathcal{G} of $\mathbb{R}/2\pi\mathbb{Z}$. Such restrictions lead to *fragments* of the language, denoted ZX $_{\mathcal{G}}$ -calculus [JPV19]. The standard interpretation associates to each ZX $_{\mathcal{G}}$ -diagram $D : n \rightarrow m$ a matrix

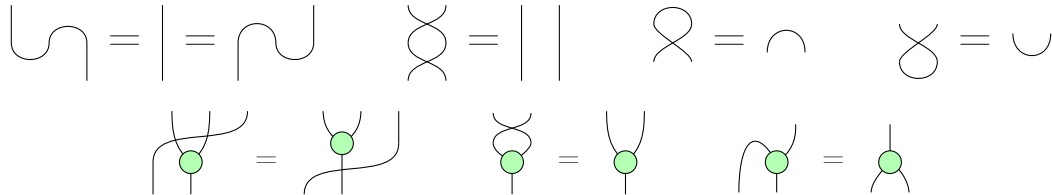
$\llbracket D \rrbracket \in \mathcal{M}_{2^m, 2^n}(\mathcal{R}_{\mathcal{G}})$ with elements in the ring $\mathcal{R}_{\mathcal{G}} = \mathbb{Z} \left[\frac{1}{\sqrt{2}}, e^{i\mathcal{G}} \right]$ - the smallest ring that contains \mathbb{Z} , $\frac{1}{\sqrt{2}}$ and $\{e^{ia} | a \in \mathcal{G}\}$ [JPV19].

In particular, the $\frac{\pi}{2}$ - (resp. π -) fragment², also called Clifford (resp. real Clifford) fragment, enjoys nice properties [Bac14, DP14] but is not universal for quantum computing, even approximately. Furthermore, any quantum computation that can be expressed in this fragment can be efficiently simulated on a classical computer. As soon as the group contains the angle $\frac{\pi}{4}$, the corresponding fragment is approximately universal for quantum computing: any $2^n \times 2^n$ unitary transformation can be approximated by a ZX-diagram from this fragment with arbitrary precision. In particular the $\frac{\pi}{4}$ -fragment, also called ‘Clifford+T’ fragment has been extensively studied [JPV18a, JPVW17, NW18]. Other finitely generated fragments have been considered in [JPV19].

Notice that for any sub-group \mathcal{G} of $\mathbb{R}/2\pi\mathbb{Z}$ that contains $\frac{\pi}{4}$, $ZX_{\mathcal{G}}$ -diagrams are *universal* [JPV19] in the sense that for any matrix $M \in \mathcal{M}_{2^m, 2^n}(\mathcal{R}_{\mathcal{G}})$ there exists a $ZX_{\mathcal{G}}$ -diagram $D : n \rightarrow m$ such that $\llbracket D \rrbracket = M$.

2.2. The calculus. Two ZX-diagrams may have the same interpretation, as a consequence the language is equipped with a set of rewrite rules (Fig. 1) that allows to transform diagrams.

In addition, ZX-diagrams can be deformed at will: all wires may be bent in any manner that keeps intact the order of inputs and outputs. It is also allowed to arbitrarily change the order of wires for green and red spiders and the Hadamard. Corresponding transformation rules are aggregated under the paradigm *Only topology matters*:



As a direct consequence of the axiom (S1), these rules also hold for non-empty spiders. They also imply that whenever a wire has no free ends, we can draw it in any way, f.e. as a horizontal line.

We denote $ZX \vdash D_1 = D_2$ if D_1 may be transformed to D_2 by local application of rewriting rules.

The ZX-calculus is sound, i.e. the rules preserve the semantics: if $ZX \vdash D_1 = D_2$ then $\llbracket D_1 \rrbracket = \llbracket D_2 \rrbracket$. The converse property is called completeness. The set of rules (Fig. 1) was proven complete for the $\frac{\pi}{4}$ -fragment [JPV19], and a single extra-rule makes the language complete for arbitrary diagrams [JPV18b]. Notice that alternative sets of rules have been shown to be complete for general ZX-diagrams [HNW18, Vil19]. We choose to consider the rules of Fig. 1 as they have been used to study diagrams with parameters in [JPV18b].

²I.e. the fragment of diagrams which angles are in the group generated by $\frac{\pi}{2}$ (resp. π)

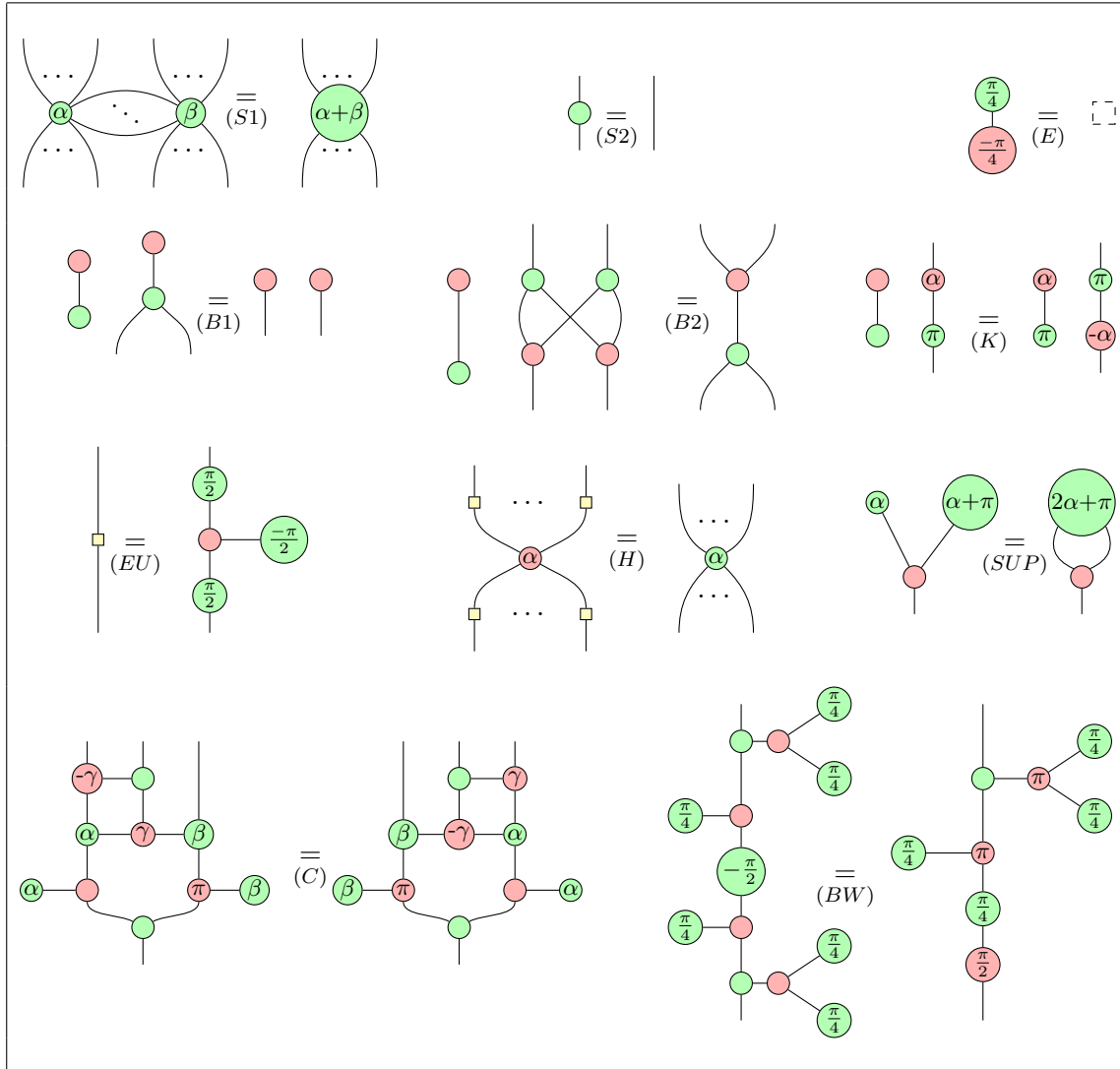


FIGURE 1. Axioms for ZX as presented in [JPV19]. All rules stay true flipped upside down and with inverted colors. Families of equations are given using ‘dots’: ... means any number of wires, ‘·’ means at least one wire.

2.3. **Useful lemmas.** Theorems and demonstrations in this work extensively use triangle:

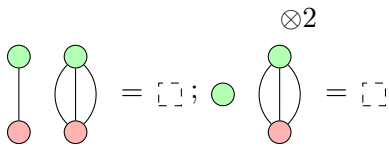
\blacktriangle - a syntactic sugar introduced in [JPV18a]. It corresponds to a non-unitary transformation:

$\left[\left[\blacktriangle \right] \right] = |0\rangle\langle 0| + |0\rangle\langle 1| + |1\rangle\langle 1|$. The triangle may be written in terms of red and green spiders as:

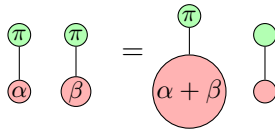
(2.2)

We also use multiple lemmas introduced in the work [JPV19]:

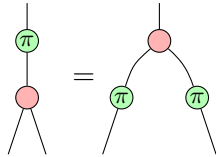
Lemma 2.4.



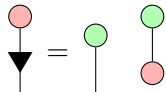
Lemma 2.5.



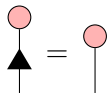
Lemma 2.6.



Lemma 2.7.



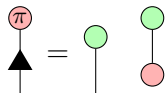
Lemma 2.8.



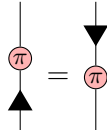
Lemma 2.9.



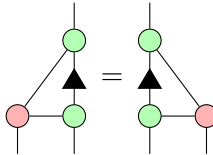
Lemma 2.10.



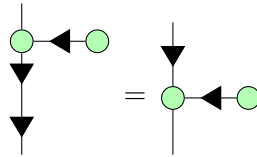
Lemma 2.11.



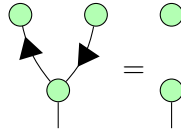
Lemma 2.12.



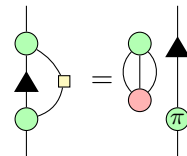
Lemma 2.13.



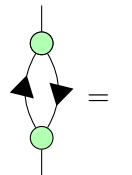
Lemma 2.14.



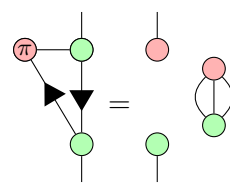
Lemma 2.15.



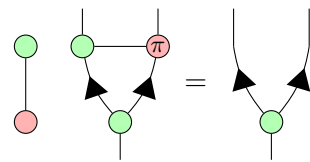
Lemma 2.16.



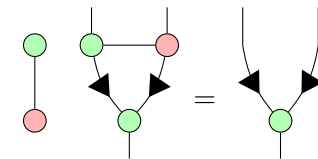
Lemma 2.17.



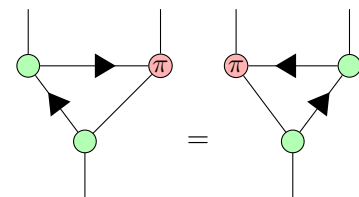
Lemma 2.18.



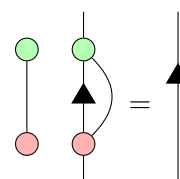
and



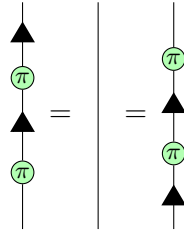
Lemma 2.19.



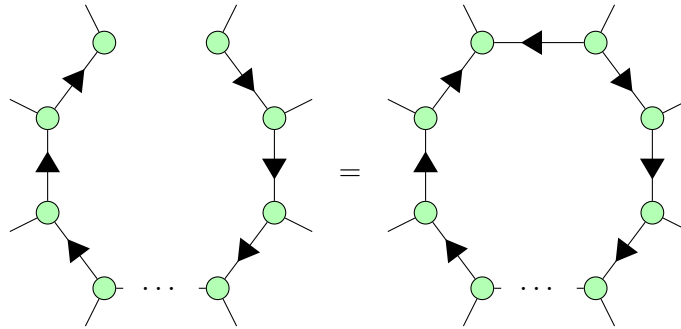
Lemma 2.20.



Lemma 2.21.

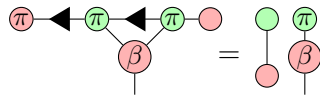


Lemma 2.22.

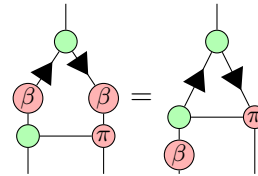


We also introduce some new lemmas that will be useful in the following sections. The proofs serve as an illustration of the graphical computation process.

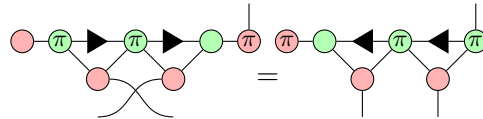
Lemma 2.23.



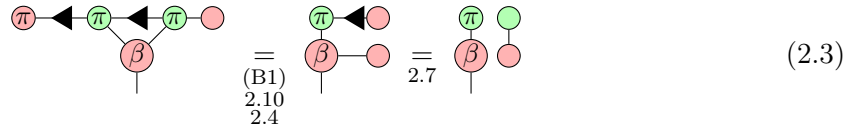
Lemma 2.24.



Lemma 2.25.

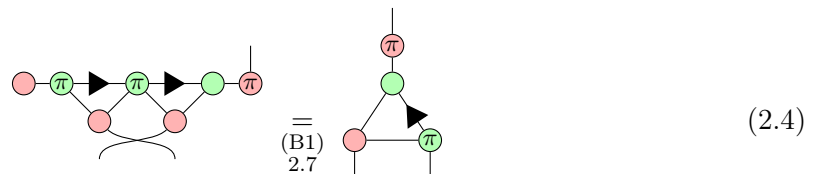


Proof (lemma 2.23).



□

Proof (lemma 2.25). The left hand side is:



For the right hand side we get:

(2.5)

□

Proof (lemma 2.24). To prove the lemma we adopt a slightly different technique than a usual sequence of rewrites. The technique consists in verification that the diagrams from the left hand side and the right hand side lead to the same result when applied on a set of *basis states* $\{|0\rangle, |1\rangle\}$ corresponding to diagrams and . As was pointed out in [CK17] for two matrices $M_1, M_2 : \mathbb{C}^{2^n} \rightarrow \mathbb{C}^{2^n}$ and a basis $\{|b_1\rangle, \dots, |b_{2^n}\rangle\}$ in \mathbb{C}^{2^n} the equality on basis input is equivalent to the equality of matrices, i.e.

$$\forall i \in [1, \dots, 2^n] : M_1|b_i\rangle = M_2|b_i\rangle \iff M_1 \equiv M_2 \tag{2.6}$$

Therefore, by demonstrating that = and = we actually

establish that $\left[\begin{array}{c} \boxed{lhs} \\ \dots \end{array} \right] = \left[\begin{array}{c} \boxed{rhs} \\ \dots \end{array} \right]$.

(2.7)

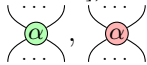
(2.8)

The lemma condition follows from the completeness of ZX-calculus. □

2.4. Linear diagrams. In ZX-diagrams the angles inside spiders are real numbers or elements of some group \mathcal{G} . However, in many equations such as (S1), (K), (H), and (SUP) we used *symbols* α and β instead of numbers. The symbolic notation was extremely handy to define *families of equalities*, i.e. equalities that hold for any real values of angles. By explicitly keeping symbols inside spiders we can draw *parameterized diagrams*. A *parameterized diagram* with n different symbols can be interpreted as a function that associates to each element of the space \mathbb{R}^n a well-defined ZX-diagram.

More precisely, we say that a diagram D is parametrized by β_1, \dots, β_k if its angles are some functions on β_1, \dots, β_k . We denote such a diagram by $D(\beta_1, \dots, \beta_k)$. It is possible to *evaluate* a parametrized diagram $D(\beta)$, $\beta \in \mathbb{R}^k$ in a point $\tilde{\beta} \in \mathbb{R}^k$ by replacing every occurrence of β_i with the respective value $\tilde{\beta}_i$. The result of the evaluation is a diagram $D(\tilde{\beta})$ from $ZX_{\mathbb{R}}$.

Among parameterized diagrams, we distinguish a family of *linear diagrams* denoted $ZX(\beta)$:

Definition 2.26 (Linear diagrams [JPV18b]). A ZX-diagram is linear in β_1, \dots, β_k with constants in $L \subset \mathbb{R}$ if it is generated by , combined by tensor product and composition such that α is of the form $\sum_i n_i \beta_i + c$ with $n_i \in \mathbb{Z}$ and $c \in L$.

In other words, in a parameterized linear diagram $D = D(\beta_1, \dots, \beta_k)$ each angle α of a (red/green) spider is of the form $\sum_i n_i \beta_i + c$ for some integer n_i and $c \in L$.

It was shown in [JPV18b] that for $L = \{\frac{n\pi}{4}\}_{n \in \mathbb{Z}}$ the Clifford+T axiomatization (Fig. 1) is complete for linear diagrams.

The family of linear diagrams may appear restricted compared to ZX-diagrams that allow angles from a more general class of functions. It is, however, sufficient for applications in variational quantum algorithms as they use circuits where parameters appear in a linear fashion [Pre18]. More importantly, for this family we are able to demonstrate simple formulas for the derivative. We believe that such formulas are not obtainable even for a slightly more general fragment $ZX_{\mathcal{A}_n}$ where angles belong to the group of affine functions $\mathcal{A}_n = \{(\beta) \rightarrow c^T \beta + c_0 | c \in \mathbb{R}^n, c_0 \in \mathbb{R}\}$ (i.e. where coefficients are *real numbers*). Intuitively, the difficulty comes from the absence of a simple representation of a general matrix over real numbers in terms of spiders. This restriction is removed in the algebraic ZX-calculus [Wan20] at the cost of an extended set of generators.

3. RELATED WORK

In this section we provide a broader context for addition and differentiation in quantum computing. In particular, we discuss how the addition of two unitary matrices may appear in quantum applications and how it can be expressed in the circuit notation [CW12]. We also review other attempts to manipulate sums of ZX-diagrams.

The differentiation in quantum computing appears in the context of variational algorithms. We review how derivatives of ZX-diagrams may be helpful for the analysis of these algorithms, in particular in the detection of *barren plateaus* and in the design of *parameter shift rules*. These ideas point out possible directions for future work.

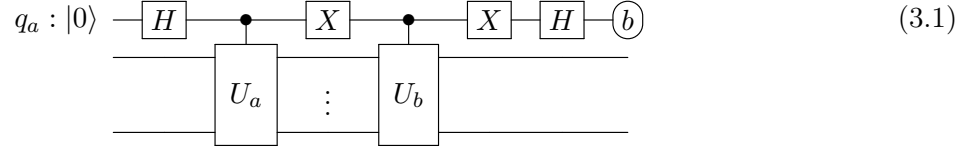
3.1. Addition in quantum computing.

Addition with circuits. In pure qubit quantum mechanics computations are performed by *unitary transformations*. Unitary transformations form a group with respect to multiplication, but the sum of two unitary transformations U_a and U_b is not unitary.

However, in some context it is meaningful to consider transformations of the form $\frac{1}{c}(U_a + U_b)$. For instance, the work [CW12] suggests to approximate the evolution of a quantum state under some Hamiltonian by a linear combination of unitary operators. It was

shown that such an approach typically outperforms the traditional technique for simulation based on *Trotterization*.

The protocol for the addition of the unitaries U_a and U_b in [CW12] is:



where q_a is an ancilla qubit initialized in a zero-state, and we have access to the controlled-version of the unitary gates U_a and U_b .

The protocol (3.1) can be easily verified. Starting in the state $|0\rangle|\psi\rangle$ we obtain:

$$|0\rangle|\psi\rangle \rightarrow \frac{|0\rangle + |1\rangle}{\sqrt{2}}|\psi\rangle \rightarrow \frac{1}{\sqrt{2}}(|0\rangle \otimes U_b|\psi\rangle + |1\rangle \otimes U_a|\psi\rangle) \quad (3.2)$$

$$\rightarrow \frac{1}{2}(|0\rangle \otimes (U_b + U_a)|\psi\rangle + |1\rangle \otimes (U_b - U_a)|\psi\rangle) \quad (3.3)$$

If the ancilla qubit is measured in the $|0\rangle$ state, the input state $|\psi\rangle$ was successfully transformed to $\frac{1}{c}(U_a + U_b)|\psi\rangle$ (c is the normalization constant).

We remark that protocol (3.1) uses an ancilla qubit and post-selection on the measurement output and, therefore, exceeds the framework of pure qubit computations. We also emphasize that the circuit involves *controlled versions* of unitary transformations. In order to be executed on a quantum computer, the controlled unitaries have to be compiled into a sequence of elementary gates. An efficient compilation is non-trivial even for a simple controlled-controlled-not gate usually called *Toffoli gate*. Interestingly, the use of graphical representations may enhance the compilation of controlled gates [KvdWK19].

Addition with the ZX-Calculus. In the context of the ZX-calculus, there is no natural way to perform a linear combination of arbitrary ZX-diagrams. Fundamentally, this deficiency is due to the absence of a proper physical interpretation for a sum of diagrams. Indeed, ZX-calculus is a language for a process theory representation of quantum computing - i.e. each diagram is a *process* [CK17]. The composition of processes can be interpreted as their sequential application, the tensor product corresponds to the parallel application, but the addition doesn't have a meaningful physical interpretation.

Several attempts were made to address this issue. The first option selected in [TYd21, SH22] consists in an extension of the formalism from diagrams to *bags of diagrams*, i.e. formal sums of diagrams. The use of formal sums means computations and diagram simplifications cannot be done purely in the context of diagrams, but also require basic calculus.

Independently from our result, [YWS] suggested an alternative protocol for the addition of two ZX-diagrams. This article uses the algebraic ZXW-calculus instead of vanilla ZX-calculus.

There are some similarities with our work, in the sense that their result can also be seen as based on a translation to a controlled form. However, their definition of a controlled form is not equivalent to ours. In addition, instead of inductive procedure, they rely on a decomposition of a diagram on elementary matrices that have an exponential complexity in the worst-case. A more detailed comparison of the two approaches is presented in Section 4.3.

3.2. Differentiation in variational algorithms. In quantum computing a *variational algorithm (VQA)* is defined by a *parameterized quantum circuit (PQC)* (also called *ansatz*) and a *loss function* that helps to discriminate between the different values of the parameters. Usually, the *loss function* is given as an expectation of an energy operator in the state prepared by PQC. Such expectation can be represented as a scalar ZX-diagram with some angles depending on the parameters.

Many state-of-the-art methods for the optimization of the loss function such as *Quantum Analytic Descent* [KB22] and *meta-learning optimizers* [WSW⁺21] use derivatives [GS17]. For efficient training it is important to detect deficiencies in the parameter landscape. In particular, gradient-based methods fail to converge to an optimal solution in the presence of so-called *barren plateaus* - zones with exponentially vanishing gradient values. It was shown that some specific structure of the *cost function* [CSV⁺21] or, alternatively, the structure of the *ansatz itself* [MBS⁺18] may lead to a landscape containing such zones.

As was demonstrated in [ZG21], the diagrammatic representation is beneficial in the analysis of the barren plateau phenomena. Yet the work [ZG21] pointed out a crucial obstacle limiting the application of ZX-calculus to variational algorithms - notably the absence of convenient tools for the *addition* and *differentiation* of parametrized diagrams. Due to this obstacle, in [ZG21] the analysis was limited to diagrams with only two occurrences of the parameter. In [TYd21] an arbitrary number of parameter occurrences is tackled with explicitly represented product rules leading to *bags of diagrams*.

Other than in the analysis of the barren plateau, diagrammatic representation can be used to design *parameter shift rules*.

Parameter shift rules. *Parameter shift rules* appeared as a solution to the problem of gradient evaluation. Indeed, the gradient of the loss function $\frac{\partial \langle \psi_0 | U^\dagger(\boldsymbol{\alpha}) O U(\boldsymbol{\alpha}) | \psi_0 \rangle}{\partial \alpha_i}$ (where O is the energy operator) in general can't be directly evaluated on quantum hardware. However, in some cases the gradient can be expressed as a linear combination of the loss function values at different points:

$$\frac{\partial L}{\partial \alpha_{i^*}}(\hat{\boldsymbol{\alpha}}) = \sum_{k=1}^m \epsilon_k L(\boldsymbol{\alpha}^k) \quad (3.4)$$

where $\alpha_i^k = \begin{cases} \hat{\alpha}_i, & \forall i \neq i^* \\ \hat{\alpha}_i + \phi_i, & i = i^* \end{cases}$. The values $L(\boldsymbol{\alpha}^k)$ can be computed with stochastic approximation.

Decompositions of the form (3.4) are called *parameter shift rules*. We remark that the parameter shift rules lead to an unbiased estimator for the gradient compared to the *finite difference approximation* $\frac{\partial L}{\partial \alpha}(\alpha_0) = \frac{L(\alpha_0 + \delta) - L(\alpha_0)}{\delta}$ which, in turn, enhances the parameter optimization process [MBK21].

The first shift rule was derived for the case $U(\alpha) = W e^{i\alpha H} V$ where W and V don't depend on α , while H (called *generator*) has two eigenvalues $+1$ and -1 [MNKF18]. It was further extended to generators with arbitrary symmetric eigenvalues $+r$ and $-r$ [SBG⁺19]:

$$\frac{\partial L}{\partial \alpha}(\alpha) = \frac{r}{4} \left(L\left(\alpha + \frac{\pi}{4r}\right) - L\left(\alpha - \frac{\pi}{4r}\right) \right) \quad (3.5)$$

For more complex $U(\alpha)$ we can decompose $U(\alpha)$ on a product of elementary matrices $U(\alpha) = E_1(\alpha) \times \cdots \times E_M(\alpha)$ and use product rules. Alternatively, more sophisticated techniques [WIWL22, KE21, ILY21] involve a *spectral decomposition* of the generator H . These advanced rules imply a smaller number of elements in the linear combination (3.4) compared to the decomposition, leading to an economy in the number of calls made to the quantum computer.

We remark that simply having a ZX-diagram for the derivative of the loss function is not enough to design new shift rules. We need to further extend the toolbox of ZX-calculus, for instance, we may need a procedure that compiles a diagram into a linear combination of other diagrams corresponding to valid quantum expectations. This extension constitutes one of the possible directions for future work.

4. ADDITION OF ZX-DIAGRAMS

The addition is a natural operation on matrices from a Hilbert space. However, it doesn't correspond to a physical process, hence it is not reflected in the standard ZX-calculus [CK17].

It is not hard to show that it is not possible directly, from two diagrams D_1 and D_2 corresponding to two matrices M_1 and M_2 , to obtain a diagram corresponding to the sum $M_1 + M_2$ by just plugging D_1 and D_2 into a bigger diagram. Indeed, any diagram that contains a subdiagram representing the 0 matrix also represents a 0 matrix. So if D_1 represents a 0 matrix and D_2 represents a nonzero matrix, then plugging D_1 and D_2 into a bigger diagram will give a zero matrix, and therefore will not represent the sum of the two matrices.

On the other hand, for any two diagrams $D_1, D_2 : n \rightarrow m$, the universality of the ZX-calculus guarantees that there exists a diagram $D : n \rightarrow m$ such that $\llbracket D \rrbracket = \llbracket D_1 \rrbracket + \llbracket D_2 \rrbracket$.

We provide in this section a general construction for such a diagram.

The key ingredient is to use controlled versions of the original diagrams, rather than the original diagrams. As pointed out in [JPV19], one can inductively define the addition on 'controlled' versions of the diagrams. A controlled version of a diagram D_0 is roughly speaking a diagram with an extra input such that when this extra input is set to $|1\rangle$ the diagram behaves as D_0 and when it is set to $|0\rangle$ the diagram behaves as a neutral diagram. Roughly speaking, the addition procedure puts controlled versions of two diagrams aside and plugs a diagram corresponding to $\text{const}(|10\rangle + |01\rangle)$ to the inputs.

To build a controlled version of a diagram our construction uses *controlled states* originally introduced in [JPV19].

4.1. Controlled states.

Definition 4.1 (Controlled state [JPV19]). A ZX-diagram $C : 1 \rightarrow n$ is a controlled state if

$$\llbracket C \rrbracket |0\rangle = \sum_{x \in \{0,1\}^n} |x\rangle = \left\llbracket \underbrace{\begin{array}{c} \textcircled{0} \quad \textcircled{0} \\ \vdots \\ \textcircled{0} \end{array}}_n \right\rrbracket.$$

Intuitively, a controlled state is a way to encode the state $\llbracket C \rrbracket |1\rangle$. A controlled state with no outputs is called *controlled scalar*. For instance, the scalar 0 could be encoded by the controlled scalar $C_0 = \begin{array}{c} \textcircled{0} \\ \textcircled{0} \end{array} \downarrow$. Indeed, we can explicitly check that $\llbracket C_0 \rrbracket |0\rangle = 1$ and

$\llbracket C_0 \rrbracket |1\rangle = 0$:

$$\left[\begin{array}{c} \otimes 2 \\ \text{green} \text{ red} \\ \text{red} \end{array} \right] \stackrel{2.4}{=} \llbracket \cdot \rrbracket = 1, \quad \left[\begin{array}{c} \otimes 2 \\ \text{green} \text{ red} \\ \text{red} \end{array} \right] \times \llbracket \pi \rrbracket \stackrel{2.1}{=} 0 \quad (4.1)$$

It will be more useful in what follows to perform *graphical proofs* instead of matrix computation to manipulate controlled states. With these rules, we verify the condition

$\llbracket C \rrbracket |0\rangle = \sum_{x \in \{0,1\}^n} |x\rangle$ by checking if $ZX \vdash \begin{array}{c} \text{green} \text{ red} \\ \text{red} \\ \boxed{C} \\ \vdots \end{array} = \underbrace{\text{green} \dots \text{green}}_n$. The diagram corresponding to

the encoded state is simply $\begin{array}{c} \text{green} \text{ red} \\ \text{red} \\ \boxed{D} \\ \vdots \end{array} = \begin{array}{c} \text{green} \text{ red} \\ \text{red} \\ \boxed{C} \\ \vdots \end{array}$.

Example 4.2 (Proven in [JPV19]). Diagrams $C_{1/\sqrt{2}} = \begin{array}{c} \text{green} \text{ red} \\ \text{red} \\ \pi/4 \end{array}$ and $C_2 = \begin{array}{c} \blacktriangle \\ \text{green} \end{array}$ are controlled scalars:

$$\begin{array}{c} \text{green} \text{ red} \\ \text{red} \\ \boxed{C_{1/\sqrt{2}}} \end{array} = \begin{array}{c} \text{green} \text{ red} \\ \text{red} \\ \pi/4 \end{array} = \begin{array}{c} \text{green} \text{ red} \\ \text{red} \\ \pi/4 \end{array} \quad \text{and} \quad \begin{array}{c} \text{green} \text{ red} \\ \text{red} \\ \boxed{C_2} \end{array} = \begin{array}{c} \text{green} \text{ red} \\ \text{red} \\ \blacktriangle \end{array} = \begin{array}{c} \text{green} \end{array} \quad (4.2)$$

representing respectively $\frac{1}{\sqrt{2}}$ and 2.

Controlled states have nice properties that allows to perform element-wise addition and tensor product of corresponding vectors:

Lemma 4.3 (Sum and tensor product [JPV19]). *For any controlled states $C_1, C_2 : 1 \rightarrow n$ and $C_3 : 1 \rightarrow m$ the diagrams:*

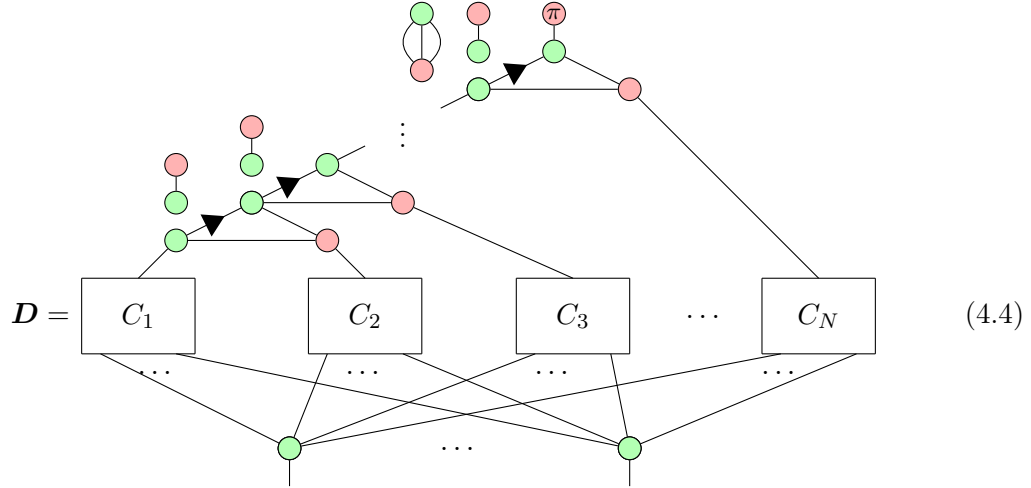
$$C_+ = \begin{array}{c} \text{green} \text{ red} \\ \text{green} \text{ red} \\ \text{red} \\ \boxed{C_1} \quad \boxed{C_2} \\ \vdots \quad \vdots \\ \text{green} \end{array} \quad \text{and} \quad C_\otimes = \begin{array}{c} \text{green} \\ \boxed{C_1} \quad \boxed{C_3} \\ \vdots \quad \vdots \end{array} \quad (4.3)$$

are controlled states such that $\llbracket C_+ \rrbracket |1\rangle = \llbracket C_1 \rrbracket |1\rangle + \llbracket C_2 \rrbracket |1\rangle$ and $\llbracket C_\otimes \rrbracket |1\rangle = \llbracket C_1 \rrbracket |1\rangle \otimes \llbracket C_3 \rrbracket |1\rangle$.

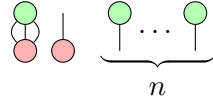
The previous lemma can be used to add two controlled states. More precisely, if one of the controlled states encodes a state diagram $D_1 : 0 \rightarrow n$ and the second controlled state

encodes $D_2 : 0 \rightarrow n$ the diagram $\begin{array}{c} \text{green} \text{ red} \\ \text{red} \\ \boxed{D_+} \\ \vdots \end{array} = \begin{array}{c} \text{green} \text{ red} \\ \text{red} \\ \boxed{C_+} \\ \vdots \end{array}$ correspond to the vector $\llbracket D_1 \rrbracket + \llbracket D_2 \rrbracket$. Moreover, as the construction returns a controlled state we can directly proceed to the addition of new

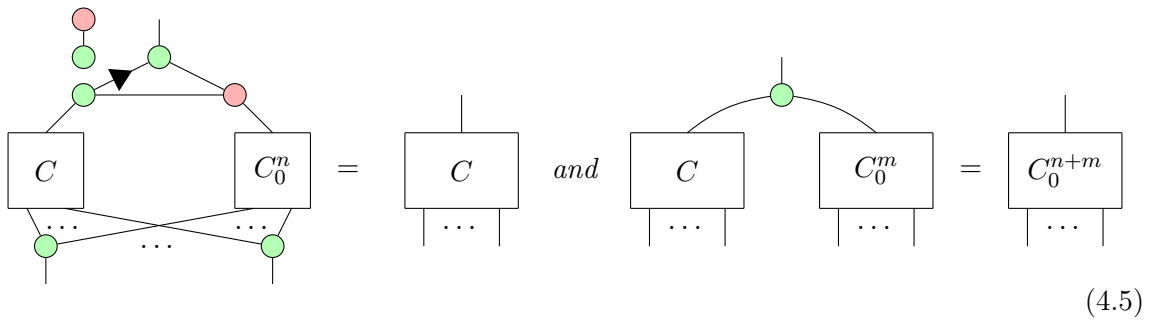
terms: for N controlled versions C_1, \dots, C_N of D_1, \dots, D_N the diagram:



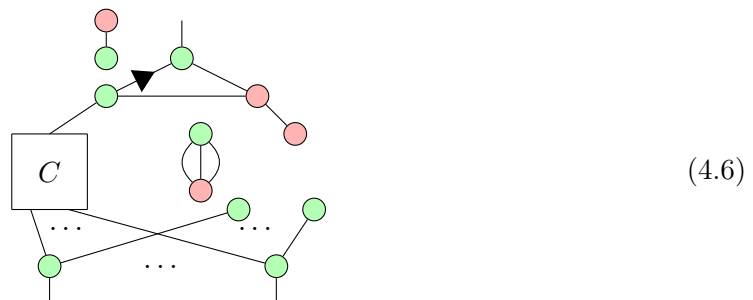
correspond to the state $[[D]] = [[D_1]] + \dots + [[D_N]]$.

We remark that for each number of output wires n the zero-element $C_0^n : 1 \rightarrow n$ is a controlled state . Indeed, we can prove the following lemma:

Lemma 4.4. For any controlled state $C : 1 \rightarrow n$



Proof. We start by proving the first equality. The left-hand side is:



The top part of the diagram can be transformed as follows:

$$(4.7)$$

The statement follows from (S2) and (S1).

For the second equality we have:

$$(4.8)$$

The lemma follows from the definition of the controlled state. \square

We follow in this work the definition of a controlled state in [JPV19]. There exist other definitions, see e.g. [YWS], which essentially differ in the output of the zero-case.

4.2. Controlizer. The lemma 4.3 provides a way to obtain the sum of two diagrams in a controlled state form.

The definition and the lemma technically allow only to solve the problem for states rather than arbitrary matrices, but notice that going from a matrix $n \rightarrow m$ to a state $0 \rightarrow n + m$ is just a matter of bending wires. We remark that bending wires alone can't lead to a procedure for addition, as there is no obvious way to directly sum state diagrams without controlled versions.

What remains to be done is to explain how to compute algorithmically, from a state (diagram) D , a controlled version of the same state (diagram).

We introduce for this the notion of a *controlizer* - a map that associates diagrams with the corresponding controlled states.

Notation 4.5. $ZX[n, m]$ denotes the family of all ZX-diagrams with n inputs and m outputs.

The formal definition of the controlizer is:

Definition 4.6 (Controlizer). We say that a map $C : ZX[n, m] \rightarrow ZX[1, n + m]$ that associates to every diagram $D : n \rightarrow m$ a diagram $C(D) : 1 \rightarrow n + m$ is a controlizer if the following conditions hold for any ZX-diagram D :

- $C(D)$ is a controlled state
- The state represents D up to some scalars in the sense that:

$$(4.9)$$

Several maps satisfy the definition of the controlizer. For instance, the map presented in [JPV19] that associates to each diagram its *normal form* is in fact a controlizer. In [YWS] an alternative controlled state form is constructed from the 2^n coefficients of the algebraic normal form.

In our approach the controlizer is defined by induction over the composition and the tensor product and doesn't rely on normal forms. It means that the shape of the controlizer will follow closely the shape of the original diagram. In terms of complexity, this also implies that the size of the controlizer output is polynomial in the size of the diagram.

Definition 4.7 (Inductive controlizer). We define the map $C : ZX[n, m] \rightarrow ZX[1, n + m]$ that associates to each diagram $D : n \rightarrow m$ a diagram $C(D) : 1 \rightarrow n + m$ as follows:

- For the generators β , α , \square , \cap , \cup , \times :

$$C(\beta) = \begin{array}{c} \bullet \\ \pi \\ \beta \\ \bullet \end{array}, \quad C(\alpha) = \begin{array}{c} \bullet \\ \pi \\ \bullet \end{array}, \quad C(\square) = \begin{array}{c} \bullet \\ \pi \\ \square \\ \bullet \end{array}$$

$$C(\cup) = C(\cap) = \begin{array}{c} \bullet \\ \pi \\ \bullet \end{array}, \quad C(\times) = \begin{array}{c} \bullet \\ \pi \\ \bullet \end{array} \quad (4.10)$$

- Generators α and α can be decomposed as follows using the above generators:

$$C\left(\begin{array}{c} n \\ \dots \\ \alpha \\ \dots \\ m \end{array}\right) = C\left(\underbrace{\alpha \dots \alpha}_{n+m}\right), \quad C\left(\begin{array}{c} n \\ \dots \\ \alpha \\ \dots \\ m \end{array}\right) = C\left(\underbrace{\alpha \dots \alpha}_{n+m}\right)$$

- For tensor product $D_{\otimes} = D_2 \otimes D_1$ and composition $D_{\circ} = D_3 \circ D_1$ where $D_1 : n \rightarrow m$, $D_2 : k \rightarrow l$ and $D_3 : m \rightarrow k$:

$$C(D_{\otimes}) = \begin{array}{c} \bullet \\ \begin{array}{cc} C(D_2) & C(D_1) \\ \dots & \dots \end{array} \\ \begin{array}{cccc} \dots & \dots & \dots & \dots \\ k & n & l & m \end{array} \end{array}, \quad C(D_{\circ}) = \begin{array}{c} \bullet \\ \begin{array}{cc} C(D_1) & C(D_3) \\ \dots & \dots \end{array} \\ \begin{array}{cc} \dots & \dots \\ n & k \end{array} \end{array} \quad (4.11)$$

We remark that the output of the inductive controlizer is not unique as it depends on the decomposition order. For instance, two diagrams that are equivalent up to deformation (the *topology matters* rule) will produce different outputs, that are not equivalent up to deformation. However, the semantics is preserved: both representations will of course represent the same matrix.

We now verify that our inductive controlizer 4.7 satisfies the definition 4.6:

Proof. We prove the claim for the generators β , \circlearrowleft , \square , \cap , \cup , \times and both compositions. The case of α and α' is a direct consequence of the axiom (S1).

First, we inductively prove that for all diagrams D the diagram $C(D)$ is controlled state.

• **Generators:**

$$C(\beta) \circ \text{diagram} = \text{diagram} \stackrel{(B1)}{=} \text{diagram} \stackrel{4.2, 2.10}{=} \text{diagram} \stackrel{2.4}{=} \text{diagram} \quad (4.12)$$

$$C(\circlearrowleft) \circ \text{diagram} = \text{diagram} \stackrel{(B1)}{=} \text{diagram} \stackrel{4.2, 2.7}{=} \text{diagram} \stackrel{(B1), 2.4}{=} \text{diagram} \quad (4.13)$$

$$C(\square) \circ \text{diagram} = \text{diagram} \stackrel{(B1), 2.4}{=} \text{diagram} \stackrel{4.2, 2.17}{=} \text{diagram} \stackrel{(H), 2.4}{=} \text{diagram} \quad (4.14)$$

$$C(\cap) \circ \text{diagram} = \text{diagram} \stackrel{(B1)}{=} \text{diagram} \stackrel{4.2, 2.7}{=} \text{diagram} \stackrel{(B1), 2.4}{=} \text{diagram} \quad (4.15)$$

$$C(\times) \circ \text{diagram} = \text{diagram} \stackrel{(B1)}{=} \text{diagram} \stackrel{4.2, 2.6, (B1)}{=} \text{diagram} \stackrel{2.10}{=} \text{diagram} \stackrel{(B1), 2.4}{=} \text{diagram} \quad (4.16)$$

• **Tensor product:** it directly follows from lemma (4.3) that if $C(D_1)$ and $C(D_2)$ are controlled states, then the diagram $C(D_1 \otimes D_2)$ from (4.11) is controlled state.

- **Composition:** let's assume that $C(D_1)$ and $C(D_2)$ are controlled states. We show that $C(D_1 \circ D_2)$ is also a controlled state:

$$\begin{aligned}
 C(D_1 \circ D_2) \circ \begin{array}{c} \text{red} \\ \text{green} \end{array} &= \begin{array}{c} \text{green} \text{ red} \\ \text{red} \text{ green} \end{array} \otimes 2m \\
 &= \begin{array}{c} \text{green} \text{ red} \\ \text{red} \text{ green} \end{array} \otimes 2m \\
 &\stackrel{(B1)}{=} \begin{array}{c} \text{green} \text{ red} \\ \text{red} \text{ green} \end{array} \otimes 2m \\
 &\stackrel{2.4}{=} \begin{array}{c} \text{green} \text{ red} \\ \text{red} \text{ green} \end{array} \otimes 2m \\
 &\stackrel{2.3}{=} \begin{array}{c} \text{green} \text{ red} \\ \text{red} \text{ green} \end{array} \otimes 2m
 \end{aligned} \tag{4.17}$$

Next we show that the map C satisfy the property (4.9), i.e. it correctly encodes the input diagram.

- **Generators:**

$$C(\beta) \circ \begin{array}{c} \text{red} \\ \text{green} \end{array} \circ \pi = \begin{array}{c} \text{red} \pi \\ \text{green} \pi \\ \beta \end{array} \stackrel{(B1)}{=} \begin{array}{c} \text{red} \pi \\ \text{green} \pi \\ \beta \end{array} \stackrel{4.2}{=} \begin{array}{c} \text{red} \\ \text{green} \end{array} \circ \beta \stackrel{2.4}{=} \begin{array}{c} \text{red} \\ \text{green} \end{array} \circ \beta \tag{4.18}$$

$$C(\pi) \circ \begin{array}{c} \text{red} \\ \text{green} \end{array} \circ \pi = \begin{array}{c} \text{red} \pi \\ \text{green} \pi \\ \pi \end{array} \stackrel{(B1)}{=} \begin{array}{c} \text{red} \pi \\ \text{green} \pi \\ \pi \end{array} \stackrel{4.2}{=} \begin{array}{c} \text{red} \\ \text{green} \end{array} \circ \pi \stackrel{2.9}{=} \begin{array}{c} \text{red} \\ \text{green} \end{array} \circ \pi \tag{4.19}$$

$$C(\frac{\pi}{4}) \circ \begin{array}{c} \text{red} \\ \text{green} \end{array} \circ \pi = \begin{array}{c} \text{red} \pi \\ \text{green} \pi \\ \pi \\ \frac{\pi}{4} \end{array} \stackrel{(B1)}{=} \begin{array}{c} \text{red} \pi \\ \text{green} \pi \\ \pi \\ \frac{\pi}{4} \end{array} \stackrel{2.4}{=} \begin{array}{c} \text{red} \pi \\ \text{green} \pi \\ \pi \\ \frac{\pi}{4} \end{array} \stackrel{4.2}{=} \begin{array}{c} \text{red} \\ \text{green} \end{array} \circ \frac{\pi}{4} \stackrel{2.16}{=} \begin{array}{c} \text{red} \\ \text{green} \end{array} \circ \frac{\pi}{4} \stackrel{(S2)}{=} \begin{array}{c} \text{red} \\ \text{green} \end{array} \circ \frac{\pi}{4} \tag{4.20}$$

$$C(\text{cap}) \circ \text{proj} = \text{diagram} \stackrel{(B1)}{=} \text{diagram} \stackrel{4.2}{=} \text{diagram} \stackrel{2.9}{=} \text{diagram} \stackrel{(S2)}{=} \text{cap} \quad (4.21)$$

$$C(\text{crossing}) \circ \text{proj} = \text{diagram} \stackrel{(B1)}{=} \text{diagram} \stackrel{4.2}{=} \text{diagram} \stackrel{2.8}{=} \text{diagram} \stackrel{(S2)}{=} \text{crossing} \quad (4.22)$$

- **Composition:** Let's assume that the property (4.9) holds for diagrams $D_1 : n \rightarrow m$ and $D_2 : m \rightarrow k$. The controller for their composition $D_1 \circ D_2$ is given by:

$$C(D_1 \circ D_2) = \text{diagram} = \text{diagram} = \text{diagram} = D_2 \circ D_1 \quad (4.23)$$

- **Tensor product:** Let's assume that the property (4.9) holds for diagrams $D_1 : n \rightarrow m$ and $D_2 : k \rightarrow l$. The controller for their tensor $D_1 \otimes D_2$ is given by:

(4.24)

The diagram (4.24) is equal to $\begin{matrix} \overbrace{\dots}^k & \overbrace{\dots}^n \\ \boxed{D_2} & \boxed{D_1} \\ \underbrace{\dots}_l & \underbrace{\dots}_m \end{matrix}$ by the induction assumption.

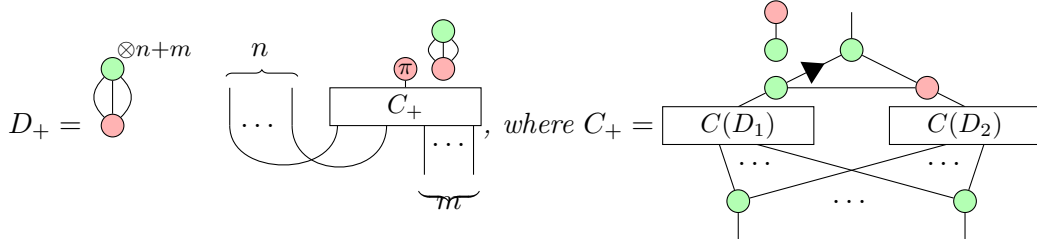
□

To give a flavor of our computation process we show how to obtain a controlled version of π :

(4.25)

Addition of arbitrary diagrams. The controlizer allows to map every two ZX-diagrams $D_1 : n \rightarrow m$ and $D_2 : n \rightarrow m$ to controlled states $C_1 : 1 \rightarrow n + m$ and $C_2 : 1 \rightarrow n + m$. The lemma 4.3 provides a way to obtain a sum of resulting controlled states. By combining this two results we can recover the diagram for addition of D_1 and D_2 . The addition protocol is resumed in the following theorem:

Theorem 4.8. For diagrams $D_1 : n \rightarrow m$ and $D_2 : n \rightarrow m$ the diagram

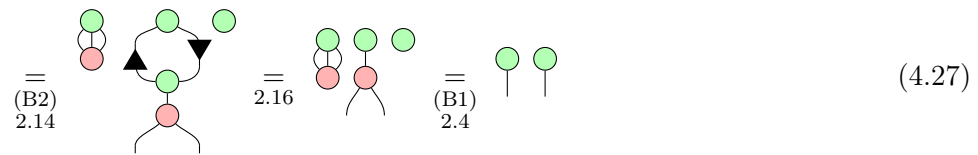
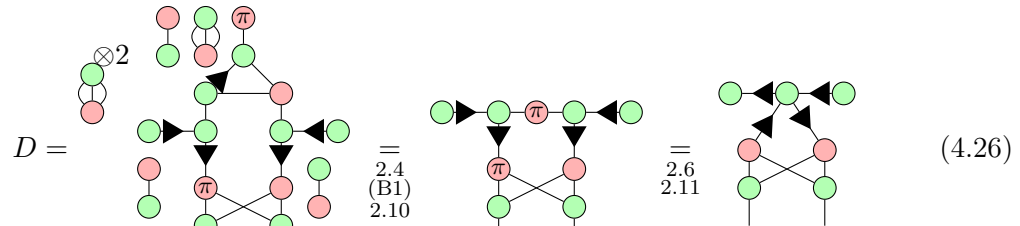


is such that $\llbracket D_+ \rrbracket = \llbracket D_1 \rrbracket + \llbracket D_2 \rrbracket$.

Proof (theorem 4.8). The theorem follows from the definition of the controlizer and the lemma 4.3. □

We illustrate the diagrammatic addition with a simple example.

Example 4.9. Using 4.8, we construct a diagram D corresponding to the addition of \frown and π .



In this example we simplified the diagram with graphical rewriting to obtain a known result:

$$\llbracket \frown \rrbracket + \llbracket \pi \rrbracket = (|00\rangle + |11\rangle) + (|01\rangle + |10\rangle) = 2|++\rangle = \llbracket \begin{array}{c} \bullet \\ | \\ \bullet \end{array} \begin{array}{c} \bullet \\ | \\ \bullet \end{array} \rrbracket \quad (4.28)$$

4.3. Discussion. The previous example shows that our inductive procedure leads to large diagrams, even when the initial diagrams are fairly simple. The resulting diagrams roughly follow the shape of the initial diagrams, but are still quite different. This defect is however consistent with intuition. Indeed, contrary to sequential and parallel compositions, the addition is not a physical operation, hence it is not surprising that can't be easily incorporated into the framework of ZX-diagrams. On the positive side, we can rely on the powerful equation theory of the ZX-calculus to simplify, when it is possible, the diagrams representing the sum of diagrams.

An alternative approach for the addition in the algebraic ZXW-calculus was presented in [YWS]. It also relies on the controlled forms of the diagram, but their controlled forms are different from ours. Notably, contrary to our approach that accepts diagrams with an arbitrary number of input and output wires, the controlled forms in [YWS] are defined only for two families of diagrams: state diagrams $ZX[0, n]$ and diagrams $ZX[m, m]$ that correspond

to *square matrices*. Controlled matrices [YWS] are built from the decomposition of the initial diagram D on a product of diagrams for so-called *elementary matrices* [WY21]. In the general case, the decomposition on elementary matrices seems to require an explicit computation of the matrix corresponding to the diagram. Therefore, although for given controlled forms the addition from [YWS] is done diagrammatically with the help of *W-spider*, the computation of the controlled forms uses the semantics. In contrast, in our inductive procedure the entire process is maintained in the graphical framework.

However, in some special cases of practical importance the approach suggested in [YWS] is beneficial. For instance, as was shown in [YWS] that for local Hamiltonians made out of Pauli tensors the decomposition on elementary matrices is straightforward. As a consequence, the controlled versions are easy to compute and the overall procedure leads to better-looking diagrams.

5. DIFFERENTIATION OF ZX-DIAGRAMS

A basic definition of the *derivative* is the derivative for a smooth function $f : \mathbb{R} \rightarrow \mathbb{R}$:

$$\frac{\partial f}{\partial x}(x_0) = \lim_{\delta \rightarrow 0} \frac{f(x_0 + \delta) - f(x_0)}{\delta} \quad (5.1)$$

In traditional calculus the derivative usually is computed from the decomposition of the function $f : \mathbb{R} \rightarrow \mathbb{R}$ on a sum, product, and composition of elementary functions. Derivatives for compositions of function functions are computed using linearity and the product rules: $\partial(f \circ g) = f \circ \partial g + \partial f \circ g$ and similarly for the tensor product.

A parameterized ZX-diagram $D(\beta)$ can be understood as a function $f : \mathbb{R} \rightarrow \text{ZX}$ that associates to each real value $\beta \in \mathbb{R}$ a ZX-diagram. In the context of variational algorithms, we consider the derivatives over the *parameters*.

In order to keep the notations simple in what follows we restrict our attention to the set of linear diagrams over one variable, denoted by $\text{ZX}(\beta)$. We denote the corresponding matrices by $\mathcal{M}(\beta)$. The elements of matrices in $\mathcal{M}(\beta)$ belong to the ring generated by complex numbers and the smooth functions $e^{i\mathcal{A}}$ where $\mathcal{A} = \{k\beta \mid k \in \mathbb{Z}\}$ is the group of affine functions with integer coefficients. We remark that even if we restrict the consideration to diagrams over one variable, all results are easily extended to the case of partial derivatives ∂_{β_i} for linear diagrams with an arbitrary number of variables *mutatis mutandis*.

The sum for the matrices in $\mathcal{M}(\beta)$ is naturally defined by entry-wise addition. Therefore, we can define the derivative in $\mathcal{M}(\beta)$ by specifying the general definition for monoidal categories with sums given in [TYd21]:

Definition 5.1 (Derivative in $\mathcal{M}(\beta)$). A derivative ∂_M in $\mathcal{M}(\beta)$ is a linear unary operator on $\mathcal{M}(\beta)$ that associates to a matrix $L \in \mathcal{M}(\beta)$, $L : n \rightarrow m$ another matrix $L' \in \mathcal{M}(\beta)$, $L' : n \rightarrow m$ of the same dimensions. It satisfies the following axioms:

◦-product rule : $\partial_M[A \circ B] = \partial_M[A] \circ B + A \circ \partial_M[B]$

⊗-product rule : $\partial_M[A \otimes B] = \partial_M[A] \otimes B + A \otimes \partial_M[B]$

Rather than defining the derivative as a limit of finite difference and deriving the product rules, the work [TYd21] suggests an axiomatic definition that proclaims *product rules* as the fundamental property of the derivative operator. We remark that the entry-wise differentiation of matrix elements satisfies this definition.

Although our theorem 4.8 provides a fully diagrammatic way for the addition of diagrams, we avoid extending the language with a formal introduction of a sum operator in order

to circumvent unnecessary complications. Therefore, the general definition from [TYd21] doesn't apply. Instead, we suggest an alternative semantics for the derivative in the ZX-calculus requiring the coherence between the derivative of a diagram and the derivative of the corresponding linear map:

Definition 5.2 (ZX-derivative). A derivative $\partial_{ZX} : ZX[n, m] \rightarrow ZX[n, m]$ is a unary operator that commutes with the standard interpretation:

$$[[\partial_{ZX} D]] = \partial_M [[D]] \tag{5.2}$$

where ∂_M is an (arbitrary) fixed derivative in $\mathcal{M}(\beta)$ satisfying the definition (5.1).

The property (5.2) is called *diagrammatic differentiation*.

5.1. Diagrammatic differentiation with controlizers. The derivative in $\mathcal{M}(\beta)$ is defined through product rules that involve sums that will be translated to the diagrammatic framework using *controlizers*.

Notation 5.3. In what follows we denote by C any map that satisfies the definition 4.6 of controlizer, for instance the one of 4.7.

Definition 5.4. Given a controlizer C , the C -derivative is the map $\Delta : ZX(\beta) \rightarrow ZX(\beta)$ that associates to a parameterized diagram $D : n \rightarrow m$ another diagram $\Delta(D) : 1 \rightarrow n + m$ defined as follows:

- **Generators:** For parametrized spiders:

$$\Delta \left(\begin{array}{c} \beta \\ \text{---} \\ \text{---} \end{array} \right) = \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \end{array}, \quad \Delta \left(\begin{array}{c} -\beta \\ \text{---} \\ \text{---} \end{array} \right) = \begin{array}{c} \pi \\ \text{---} \\ \text{---} \end{array} \tag{5.3}$$

$$\Delta \left(\begin{array}{c} n \\ \dots \\ k\beta \\ \dots \\ m \end{array} \right) = \Delta \left(\begin{array}{c} n \\ \dots \\ \beta \\ \vdots \\ \beta \\ \dots \\ m \end{array} \right), \quad \Delta \left(\begin{array}{c} n \\ \dots \\ k\beta \\ \dots \\ m \end{array} \right) = \Delta \left(\begin{array}{c} n \\ \dots \\ \square \\ \dots \\ \square \\ \dots \\ m \end{array} \right) \tag{5.4}$$

For all other generators $g : n \rightarrow m$ that do not contain an instance of the variable β , we define $\Delta(g) = \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \end{array}$.

- **Tensor product:** for $D_1 : n \rightarrow m$ and $D_2 : k \rightarrow l$ the diagram $\Delta(D_1 \otimes D_2)$ is:

$$\Delta(D_1 \otimes D_2) = \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \end{array} \tag{5.5}$$

- **Composition:** for $D_1 : n \rightarrow m$ and $D_2 : m \rightarrow k$ the diagram $\Delta(D_2 \circ D_1)$ is:

$$\Delta(D_2 \circ D_1) = \begin{array}{c} \begin{array}{c} \otimes 2m \\ \begin{array}{c} \text{green} \\ \text{red} \end{array} \end{array} \\ \begin{array}{c} \Delta(D_1) \quad C(D_2) \quad C(D_1) \quad \Delta(D_2) \\ \dots \quad \dots \quad \dots \quad \dots \\ \underbrace{\quad \quad \quad}_n \quad \quad \quad \underbrace{\quad \quad \quad}_k \end{array} \end{array} \quad (5.6)$$

As should be clear from the definition, the C -derivative of a diagram is not derivative of a diagram. Rather, it corresponds to a controlled-version of the derivative.

A simple verification shows that for any generator G the C -derivative $\Delta(G)$ is a controlled state. From lemma 4.3 it follows that for every diagram $D : n \rightarrow m$ the C -derivative $\Delta(D) : 1 \rightarrow n + m$ is a controlled state. Moreover, as was proven in lemma 4.4 the C -derivative $\begin{array}{c} \text{green} \\ \text{red} \end{array} \dots \begin{array}{c} \text{green} \\ \text{red} \end{array}$ of generators that are independent on β is precisely the zero-element C_0^{m+m} for the addition of controlled states.

The same lemma implies, in particular, that the C -derivative of any diagram $D : n \rightarrow m$ that is independent on β is $\begin{array}{c} \text{green} \\ \text{red} \end{array} \dots \begin{array}{c} \text{green} \\ \text{red} \end{array}$.

Similarly to controlizers, a step-by-step application of the map Δ may lead to different diagrams depending on the order of decomposition on tensor products and compositions. However, all possible outputs are semantically equivalent and, by the completeness of ZX-calculus, are equivalent as diagrams.

By combining C -derivatives and controlizers we can reproduce the semantics of product rules in a graphical representation:

Definition 5.5. Given a C -derivative Δ , let $\partial_C : ZX(\beta) \rightarrow ZX(\beta)$ be the unary operator such that for any diagram $D : n \rightarrow m$

$$\partial_C[D] = \begin{array}{c} \begin{array}{c} \pi \\ \text{green} \\ \text{red} \end{array} \dots \begin{array}{c} \text{green} \\ \text{red} \end{array} \otimes n+m \\ \dots \quad \underbrace{\quad \quad \quad}_n \quad \begin{array}{c} \text{green} \\ \text{red} \\ \pi \\ \Delta(D) \\ \dots \end{array} \quad \underbrace{\quad \quad \quad}_m \end{array} \quad (5.7)$$

Theorem 5.6. The operator ∂_C satisfies the definition (5.2) of diagrammatic differentiation.

Proof (theorem 5.6). We give the proof by induction over tensor product and composition.

- **Generators:** First, we show that $\partial_M \left[\left[\begin{array}{c} \beta \\ | \end{array} \right] \right] = ie^{i\beta} | - \rangle$ and $\partial_M \left[\left[\begin{array}{c} -\beta \\ | \end{array} \right] \right] = -ie^{-i\beta} | - \rangle$.

Indeed, we have:

$$\partial_M \left[\left[\begin{array}{c} \beta \\ | \end{array} \right] \right] = \partial_M (| + \rangle + e^{i\beta} | - \rangle) = ie^{i\beta} | - \rangle \quad (5.8)$$

$$\partial_M \left[\left[\begin{array}{c} -\beta \\ | \end{array} \right] \right] = \partial_M (| + \rangle + e^{-i\beta} | - \rangle) = -ie^{-i\beta} | - \rangle \quad (5.9)$$

For the diagrams $\partial_C \left(\begin{smallmatrix} \beta \\ \uparrow \end{smallmatrix} \right)$ and $\partial_C \left(\begin{smallmatrix} -\beta \\ \uparrow \end{smallmatrix} \right)$ we verify the property of diagrammatic differentiation:

$$\left[\begin{array}{c} \pi \\ \pi \text{ (green)} \\ \pi \text{ (red)} \\ \pi \text{ (red)} \\ \pi \text{ (green)} \\ \pi \text{ (green)} \\ \beta \\ \uparrow \end{array} \right] = \left[\begin{array}{c} \pi \\ \pi \text{ (green)} \\ \pi \text{ (red)} \\ \pi \text{ (red)} \\ \pi \text{ (green)} \\ \beta \\ \uparrow \end{array} \right] = ie^{i\beta} | - \rangle \quad (5.10)$$

$$\left[\begin{array}{c} \pi \\ \pi \text{ (green)} \\ \pi \text{ (red)} \\ \pi \text{ (red)} \\ \pi \text{ (green)} \\ -\beta \\ \uparrow \end{array} \right] = \left[\begin{array}{c} \pi \\ \pi \text{ (green)} \\ \pi \text{ (red)} \\ \pi \text{ (red)} \\ \pi \text{ (green)} \\ -\beta \\ \uparrow \end{array} \right] = -ie^{-i\beta} | - \rangle \quad (5.11)$$

For a generator $g : n \rightarrow m$ that doesn't depend on β :

$$\llbracket \partial_C(g) \rrbracket = \left[\begin{array}{c} \pi \\ \pi \text{ (green)} \\ \pi \text{ (red)} \\ \pi \text{ (red)} \\ \pi \text{ (green)} \\ \pi \text{ (green)} \\ \uparrow \end{array} \right] \otimes \left[\begin{array}{c} n \\ \dots \\ \dots \\ m \end{array} \right] \stackrel{(S1)}{=} \llbracket \pi \otimes \dots \rrbracket = (0)_{n \times m} \quad (5.12)$$

- **Tensor product:** Let's assume that $\llbracket \partial_C(D_1) \rrbracket = \partial_M \llbracket D_1 \rrbracket$ and $\llbracket \partial_C(D_2) \rrbracket = \partial_M \llbracket D_2 \rrbracket$ for $D_1 : n \rightarrow m$ and $D_2 : l \rightarrow k$. For $D = D_1 \otimes D_2$:

$$\llbracket \partial_C(D) \rrbracket = \left[\begin{array}{c} \pi \\ \pi \text{ (green)} \\ \pi \text{ (red)} \\ \pi \text{ (red)} \\ \pi \text{ (green)} \\ \pi \text{ (green)} \\ \otimes \\ \dots \\ \dots \\ \Delta(D) \\ \uparrow \end{array} \right] \quad (5.13)$$

By applying the definition 5.4 of the C -derivative for the tensor product we obtain the diagram:

$$\left[\begin{array}{c} \pi \\ \pi \text{ (green)} \\ \pi \text{ (red)} \\ \pi \text{ (red)} \\ \pi \text{ (green)} \\ \pi \text{ (green)} \\ \otimes \\ \dots \\ \dots \\ \Delta(D_1) \quad C(D_2) \quad C(D_1) \quad \Delta(D_2) \\ \dots \\ \dots \\ m \quad l \end{array} \right] \quad (5.14)$$

We denote by $\underbrace{\left[\begin{array}{c} T \\ \dots \\ \dots \end{array} \right]}_{\substack{n+m \quad k+l}}$ the part of the diagram surrounded by the red frame.

According to the definition, C -derivatives $\Delta(D_1)$, $\Delta(D_2)$ and controlizer outputs $C(D_1)$, $C(D_2)$ are controlled states. Therefore, from lemma 4.3 we get:

$$\left[\begin{array}{c} T \\ \dots \\ \dots \end{array} \right] = \left[\begin{array}{c} \text{Controlizer} \\ C(D_1) \quad \Delta(D_2) \\ \dots \end{array} \right] + \left[\begin{array}{c} \text{Controlizer} \\ \Delta(D_1) \quad C(D_2) \\ \dots \end{array} \right] \quad (5.15)$$

Meanwhile, by bending wires we see that the resting part of the diagram (5.14) is:

$$B = \left[\begin{array}{c} \text{Controlizer} \\ \dots \end{array} \right] \quad (5.16)$$

The sum in $\mathcal{M}(\beta)$ is distributive over the composition. Therefore,

$$\left[\begin{array}{c} \text{Controlizer} \\ \dots \end{array} \right] = \quad (5.17)$$

$$\left[\begin{array}{c} \text{Controlizer} \\ \dots \end{array} \right] \quad (5.18)$$

$$+ \left[\begin{array}{c} \text{Controlizer} \\ \dots \end{array} \right] \quad (5.19)$$

The claim follows from the induction assumption and the definition of controlizer.

- **Composition:** Following the same reasoning as for the tensor product we obtain for $D_1 : n \rightarrow m$ and $D_2 : m \rightarrow l$

$$\partial_C[D_2 \circ D_1] = \left[\begin{array}{c} \text{Controlizer} \\ \dots \end{array} \right] \quad (5.20)$$

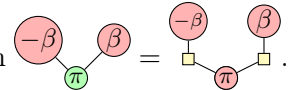
where the diagram T is exactly the same as in (5.15). On more time, by distributivity of the sum in $\mathcal{M}(\beta)$ over the composition we obtain

$$[[\partial_C[D_2 \circ D_1]]] = \text{Diagram (5.21)} \tag{5.21}$$

$$+ \text{Diagram (5.22)} \tag{5.22}$$

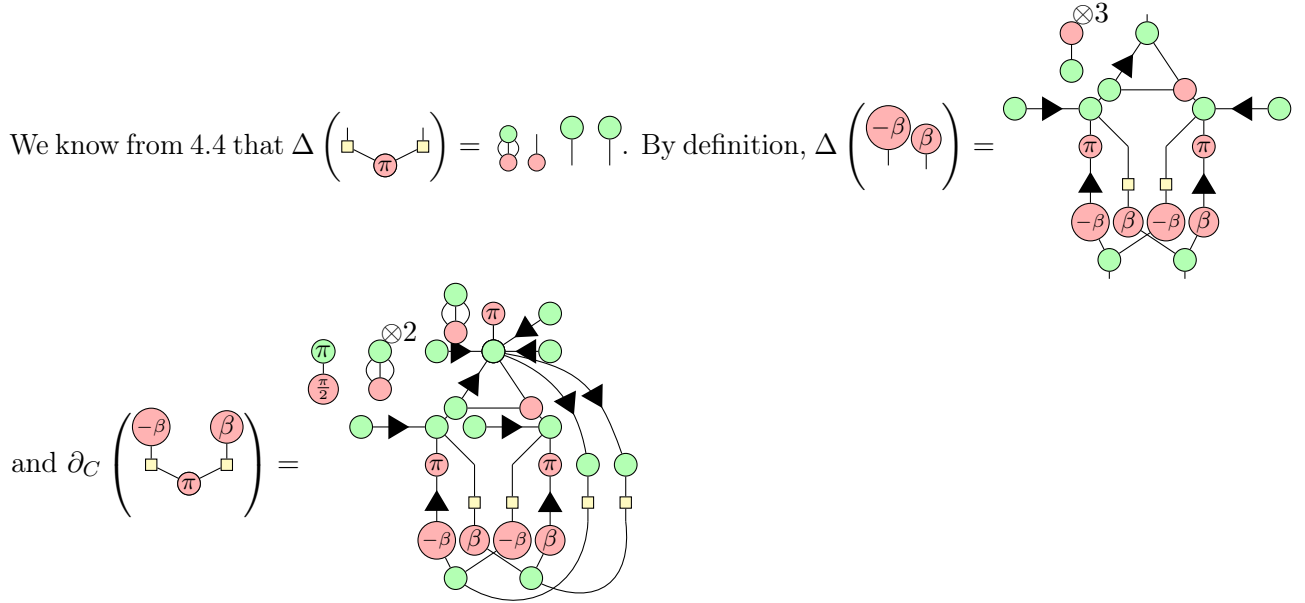
The claim follows from the definition of controlizer an the induction assumption. □

We consider a simple example to illustrate how to perform the inductive procedure.

Example 5.7. We apply the definition 5.5 to the simple scalar diagram .

Notice that $C(\text{Diagram}) = \text{Diagram} \stackrel{2.12}{=} \text{Diagram} \stackrel{(2.2)}{=} \text{Diagram} \stackrel{2.6}{=} \text{Diagram}$, moreover $C(\text{Diagram})$ was already found in the example 4.25. We obtain a diagram for $C(\text{Diagram})$:

$$\text{Diagram 1} \stackrel{2.18}{=} \text{Diagram 2} \stackrel{2.18}{=} \text{Diagram 3} \stackrel{2.16}{=} \text{Diagram 4} \tag{5.23}$$



5.2. Formula for derivatives in ZX(β). Although perfectly correct, the differentiation procedure described above leads to very puzzling output even for small diagrams (see example 5.7). In this section we provide a simpler approach to obtain the derivative of a diagram in ZX(β) written in a special form. We formalize it in definitions ∂_{ZX} and ∂_P of unary operators that satisfy the property of diagrammatic differentiation (5.2).

First, we remark that even in the inductive procedure the effort can be significantly reduced if we separate the parts that depend on the parameter from the rest of the diagram that is constant on β . Indeed, for constant diagrams $A : k \rightarrow l$ and $B : m \rightarrow k$ that do not depend on β , the C -derivative is the controlled version of the zero diagrams C_0^{k+l} and C_0^{m+k} . Therefore, for a parameterized diagram $D(\beta) : n \rightarrow m$ the lemma 4.4 implies :

$$\Delta(D(\beta) \otimes A) = \begin{array}{c} \bullet \\ \Delta(D(\beta)) \quad C(A) \\ \dots \quad \dots \quad \dots \quad \dots \\ n \quad k \quad m \quad l \end{array}, \quad \Delta(D(\beta) \circ A) = \begin{array}{c} \bullet \otimes 2m \\ \Delta(D(\beta)) \quad C(B) \\ \dots \quad \dots \\ n \quad k \end{array} \tag{5.24}$$

These expressions are significantly simpler than the general case (5.4) where both parts of the compositions depend on β .

We want to further explore this property to find simplified formulas for the derivative.

Let's denote by $X_\beta(n, m)$ the family of diagrams $\underbrace{\beta \dots \beta}_n \underbrace{-\beta \dots -\beta}_m$ from ZX(β).

Claim 5.8. Using the rules of ZX-calculus, each diagram $D(\beta)$ from $ZX(\beta)$ may be transformed into the form

$$D(\beta) = \begin{array}{c} \dots \\ \boxed{D_1} \\ \dots \\ \boxed{X_\beta(n, m)} \\ \dots \\ \boxed{D_2} \\ \dots \end{array} = \begin{array}{c} \dots \\ \boxed{D_1} \\ \dots \\ \textcircled{\beta} \dots \textcircled{\beta} \textcircled{-\beta} \textcircled{-\beta} \\ \dots \\ \boxed{D_2} \\ \dots \end{array} \quad (5.25)$$

where n, m are some integer numbers and D_1, D_2 are constant with respect to β . We call diagrams in this form β -factored.

A rigorous demonstration of claim 5.8 may be found in [JPV18b]. It essentially follows from the (H) rule to transform any green spider into a red spider, the (S1) rule to separate the β part from the rest of the diagram, and the paradigm *Only topology matters* to rearrange the diagram.

The diagram $X_\beta(n, m)$ has a particularly regular form, so we expect it to have a nice-looking derivative. We would like to develop the intuition from the semantics $\partial_M \llbracket X_\beta(n, m) \rrbracket$.

5.2.1. *Intuition.* Due to the form obtained in the previous claim, it remains to find the derivative of $X_\beta(n, m)$ to obtain the desired results.

For reasons that will become more clear later, we first focus on a variant of $X_\beta(n, m)$ with green spiders instead of red spiders.

We denote by $|\beta\rangle = |0\rangle + e^{i\beta}|1\rangle$ the state $\llbracket \textcircled{\beta} \rrbracket$ and by $|-\beta\rangle = |0\rangle + e^{-i\beta}|1\rangle$ the state $\llbracket \textcircled{-\beta} \rrbracket$.

We also introduce the notation:

$$|\beta(n, m)\rangle = |\beta\rangle^{\otimes n} \otimes |-\beta\rangle^{\otimes m} = \sum_{x \in \{0,1\}^{n+m}} e^{i\beta(|x^+| - |x^-|)} |x^+\rangle \otimes |x^-\rangle \quad (5.26)$$

where $x^+ \in \{0,1\}^n$ corresponds to the first n qubits that are in the state $|\beta\rangle^{\otimes n}$ and $x^- \in \{0,1\}^m$ corresponds to the last m qubits that are in the state $|-\beta\rangle^{\otimes m}$. The notation $|x| : \{0,1\}^k \rightarrow \mathbb{Z}_+$ stands for the Hamming weight of the binary string x .

By applying entrywise differentiation to the vector $|\beta(n, m)\rangle$, we obtain the following result:

$$\partial_M (|\beta(n, m)\rangle) = \sum_{x \in \{0,1\}^{n+m}} i(|x^+| - |x^-|) e^{i\beta(|x^+| - |x^-|)} |x^+\rangle \otimes |x^-\rangle \quad (5.27)$$

We observe that $\partial_M (|\beta(n, m)\rangle) = iM_\Delta |\beta(n, m)\rangle$ where $M_\Delta : 2^{n+m} \rightarrow 2^{n+m}$ is a linear map such that $M_\Delta : |x\rangle \rightarrow (|x^+| - |x^-|)|x\rangle$.

Therefore, obtaining the derivative of $(|\beta(n, m)\rangle)$ requires the ability to produce the matrix M_Δ .

The first step is finding a diagram for the controlled-triangle $D_\lambda : 2 \rightarrow 2$. Semantically, we want the diagram D_λ to apply the triangle \blacktriangle to the second input if the first one is in the $|1\rangle$ state and do nothing otherwise. We denote the matrix corresponding to D_λ by

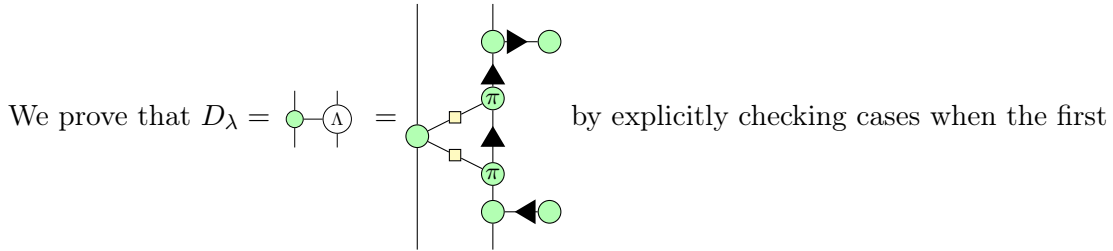
M_λ . Crucially, when the matrix M_λ is applied to a vector $|b\rangle \otimes (k|0\rangle + |1\rangle)$ for $b \in \{0, 1\}$ the coefficient k is modified depending on the value of b :

$$|b\rangle \otimes (k|0\rangle + |1\rangle) \xrightarrow{M_\lambda} |b\rangle \otimes ((k + b)|0\rangle + |1\rangle) \tag{5.28}$$

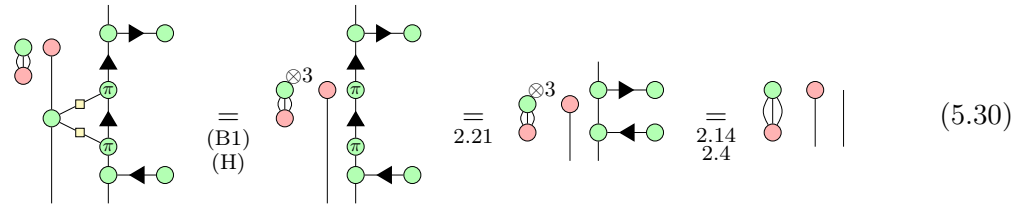
Similarly, when the input vector is $|b\rangle \otimes (k|0\rangle - |1\rangle)$ the value of b is subtracted from k :

$$|b\rangle \otimes (k|0\rangle - |1\rangle) \xrightarrow{M_\lambda} |b\rangle \otimes ((k - b)|0\rangle - |1\rangle) \tag{5.29}$$

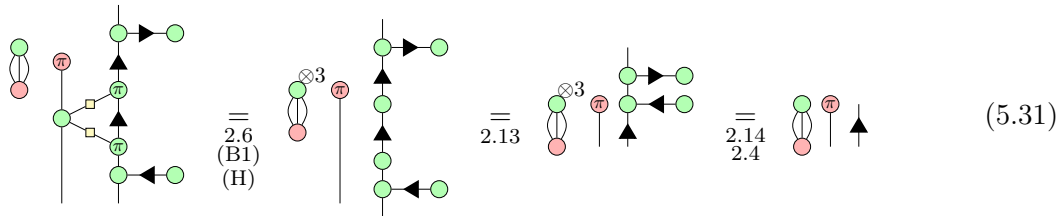
We remark the matrix M_λ is not norm-preserving, so it doesn't correspond to a valid quantum gate. However, it doesn't matter as ZX-diagrams are universal for general linear maps of dimensions $2^n \times 2^m$.



input is in $|0\rangle$ and in $|1\rangle$ states.
For the $|0\rangle$ -case we have:



When the first input is in the $|1\rangle$ state we get:



It turns out that the matrix M_Δ can be written as an evolution in $n + m + 1$ space with one ancilla input followed by a projection on $n + m$ dimensional space. The ancilla input is used to *accumulate* the coefficient α_x of the basis state $|x\rangle$. We suggest a construction in which the coefficient α_x is recovered from the projection $\langle 0|\psi_a\rangle$ of the state $|\psi_a\rangle$ of the ancilla input.

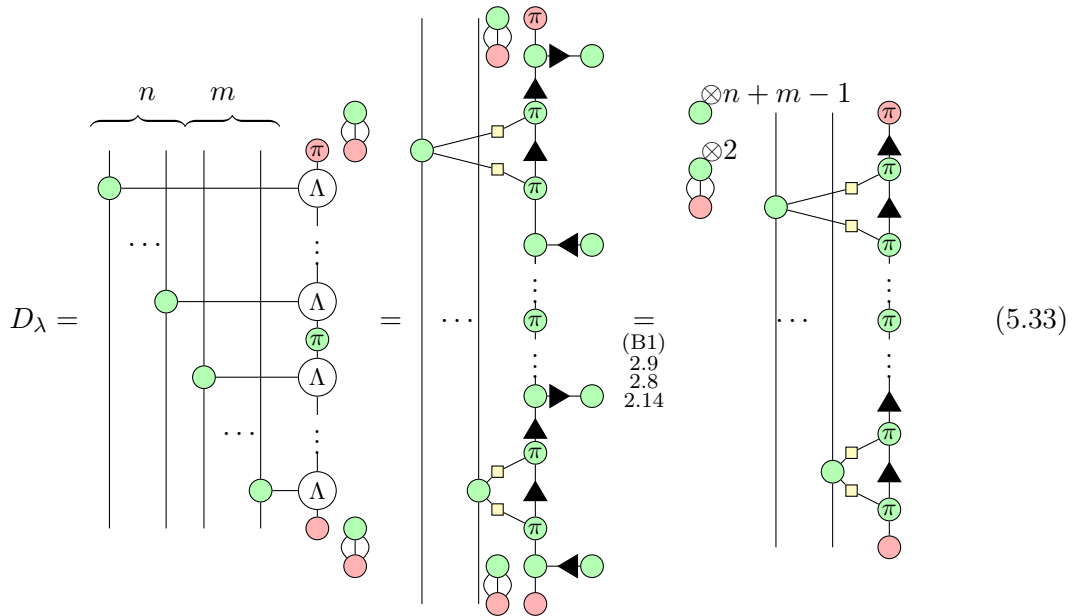
To be exact, we design an iterative procedure that multiplies the input vector $|b_0, \dots, b_{m+1}\rangle$ by the matrix M_Δ . Initially, we set $|\psi_a\rangle = |1\rangle$. Then, for $i \in [1, \dots, n]$ we apply the transformation M_λ to $|b_i\rangle|\psi_a\rangle$ controlled by the input i . After n applications the state $|\psi_a\rangle$ becomes $|\psi_a\rangle = (\sum_{i=1}^n b_i) |0\rangle + |1\rangle$. To proceed further we apply Z matrix to $|\psi_a\rangle$ and obtain $|\psi_a\rangle = (\sum_{i=1}^n b_i) |0\rangle - |1\rangle$. Then the state $|\psi_a\rangle$ is transformed to $|\psi_a\rangle = (\sum_{i=1}^n b_i - \sum_{i=n+1}^m b_i) |0\rangle - |1\rangle$ by iterative application of M_λ to $|b_i\rangle|\psi_a\rangle$. Finally, the

coefficient α_x is:

$$\alpha_x = \langle 0 | \psi_a \rangle = \left(\sum_{i=1}^n b_i - \sum_{i=n+1}^m b_i \right) \tag{5.32}$$

As the original input wires are used only in control, the output of such a procedure is $\alpha_x |b_0, \dots, b_{n+m}\rangle$.

This procedure can be directly translated in a diagram D_λ ;



The diagram for the derivative of $X_\beta(n, m) = \begin{matrix} \beta & \dots & \beta & -\beta & \dots & -\beta \end{matrix}$ can be obtained by conjugating every input with Hadamard boxes:

$$\partial_M \left[\begin{matrix} \beta & \dots & \beta & -\beta & \dots & -\beta \end{matrix} \right] = \partial_M \left[\begin{matrix} \beta & \dots & \beta & -\beta & \dots & -\beta \\ \square & \dots & \square & \square & \dots & \square \end{matrix} \right] = \left[\begin{matrix} \square & \dots & \square \end{matrix} \right] \circ \partial_M \left[\begin{matrix} \beta & \dots & \beta & -\beta & \dots & -\beta \end{matrix} \right] \tag{5.34}$$

$$= \begin{matrix} \pi \\ \pi/2 \end{matrix} \left[\begin{matrix} \square & \dots & \square \end{matrix} \right] \circ \left[\begin{matrix} \beta & \dots & \beta & -\beta & \dots & -\beta \\ \dots & & & & & \end{matrix} \right] = \begin{matrix} \pi \\ \pi/2 \end{matrix} \left[\begin{matrix} \beta & \dots & \beta & -\beta & \dots & -\beta \\ \dots & & & & & \end{matrix} \right] \tag{5.35}$$

By applying the rule (H) to the final we obtain

$$\partial_M \llbracket X_\beta(n, m) \rrbracket = \left[\left[\begin{array}{c} \text{Diagram with } n \text{ and } m \text{ segments} \end{array} \right] \right] \quad (5.36)$$

Notice that the Hadamard boxes from D_λ have disappeared in the expression of $\partial_M \llbracket X_\beta(n, m) \rrbracket$, which explains why we chose initially red spiders instead of green spiders.

5.2.2. *Diagrammatic differentiation of β -factored forms.* Motivated by the simplicity of the result (5.36), we suggest an alternative non-inductive definition for derivatives of diagrams in β -factored forms:

Definition 5.9. Given a diagram $D(\beta) = D_2 \circ (D_1 \otimes X_\beta(n, m))$ in β -factored form, let $\partial_{ZX}[D] = D_2 \circ (D_1 \otimes \partial_{ZX}[X_\beta(n, m)])$ where

$$\partial_{ZX}[X_\beta(n, m)] = \partial_{ZX} \left[\underbrace{\beta \dots \beta}_n \underbrace{-\beta \dots -\beta}_m \right] \quad (5.37)$$

$$= \left[\left[\begin{array}{c} \text{Diagram with } n \text{ and } m \text{ segments} \end{array} \right] \right] \quad (5.38)$$

Although our definition is inspired by the intuitive reflection presented above, we provide a rigorous diagrammatic demonstration of the soundness of our definition. More precisely, we prove the following theorem:

Theorem 5.10. *The operator $\partial_{ZX}[-]$ defined at (5.9) satisfies the property of diagrammatic differentiation: for any diagram $D(\beta) \in ZX(\beta)$ in β -factored form,*

$$\llbracket \partial_{ZX} D(\beta) \rrbracket = \partial_M \llbracket D(\beta) \rrbracket \quad (5.39)$$

We remark that according to the definition 5.2 the derivative $D' : n \rightarrow m$ of a diagram $D : n \rightarrow m$ that is constant on β is such that $\llbracket D' \rrbracket = \partial_M \llbracket D \rrbracket = (0)_{n \times m}$. Therefore, theorem 5.10 is a direct consequence of the following lemma:

Lemma 5.11. *For any n, m :*

$$\llbracket \partial_{ZX} X_\beta(n, m) \rrbracket = \partial_M \llbracket X_\beta(n, m) \rrbracket \quad (5.40)$$

Proof. We prove the lemma by induction. We provide the demonstration for $n = n + 1$, the proof for $m = m + 1$ is directly obtainable in the same way.

Base: We already know that $\partial_M \llbracket \beta \rrbracket = ie^{i\beta} |-\rangle$. Thus, for the case $n = 1, m = 0$ the lemma statement follows from:

$$\llbracket \partial_{ZX} X_\beta(1, 0) \rrbracket = \left[\left[\begin{array}{c} \text{Diagram with } n=1, m=0 \end{array} \right] \right] \stackrel{2.23}{=} \left[\left[\begin{array}{c} \text{Diagram with } n=1, m=0 \end{array} \right] \right] = ie^{i\beta} |-\rangle \quad (5.41)$$

Step: By induction, we assume that the equation (5.40) holds for some n and m . We show that under this assumption $[[\partial_{ZX} X_\beta(n+1, m)]] = \partial_M [[X_\beta(n+1, m)]]$.

In order to prove the induction we will proceed in three steps. Firstly we apply the product rule to the matrix $\partial_M [[X_\beta(n+1, m)]]$:

$$\begin{aligned} \partial_M [[X_\beta(n+1, m)]] &= \partial_M [[X_\beta(1, 0)]] \otimes [[X_\beta(n, m)]] + [[X_\beta(1, 0)]] \otimes \partial_M [[X_\beta(n, m)]] \\ &= [[\partial_{ZX} [X_\beta(1, 0)] \otimes X_\beta(n, m)]] + [[X_\beta(1, 0) \otimes \partial_{ZX} [X_\beta(n, m)]]] \end{aligned} \quad (5.42)$$

where the second equality follows from the inductive assumption.

Secondly, we show the last sum in the equation (5.42) can be represented diagrammatically with a controlled state, i.e. the following claim holds:

Claim 5.12. We can find a controlled state $\tilde{X} : 1 \rightarrow n+1+m$ and a constant scalar $c \in ZX_{\frac{\pi}{2}}$ such that

$$\left[\left[c \begin{array}{c} \text{green circle} \text{ } \pi \\ \text{red circle} \\ \tilde{X} \\ \dots \end{array} \right] \right] = [[\partial_{ZX} [X_\beta(1, 0)] \otimes X_\beta(n, m)]] + [[X_\beta(1, 0) \otimes \partial_{ZX} [X_\beta(n, m)]]] \quad (5.43)$$

Finally, we will show that we can transform the diagram $\partial_{ZX} X_\beta(n+1, m)$ to $\left[\left[c \begin{array}{c} \text{green circle} \text{ } \pi \\ \text{red circle} \\ \tilde{X} \\ \dots \end{array} \right] \right]$ with a sequence of graphical rewrites:

Claim 5.13.

$$\partial_{ZX} X_\beta(n+1, m) = c \begin{array}{c} \text{green circle} \text{ } \pi \\ \text{red circle} \\ \tilde{X} \\ \dots \end{array} \quad (5.44)$$

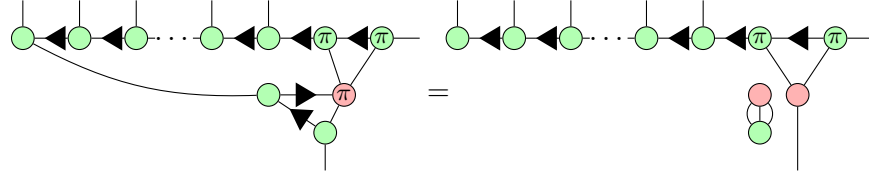
As the rewrite rules preserve the semantics, it follows from the claims above that:

$$\begin{aligned} [[\partial_{ZX} X_\beta(n+1, m)]] &\stackrel{5.13}{=} \left[\left[c \begin{array}{c} \text{green circle} \text{ } \pi \\ \text{red circle} \\ \tilde{X} \\ \dots \end{array} \right] \right] \\ &\stackrel{5.12}{=} [[\partial_{ZX} [X_\beta(1, 0)] \otimes X_\beta(n, m)]] + [[X_\beta(1, 0) \otimes \partial_{ZX} [X_\beta(n, m)]]] \\ &\stackrel{5.42}{=} \partial_M [[X_\beta(n+1, m)]] \end{aligned} \quad (5.45)$$

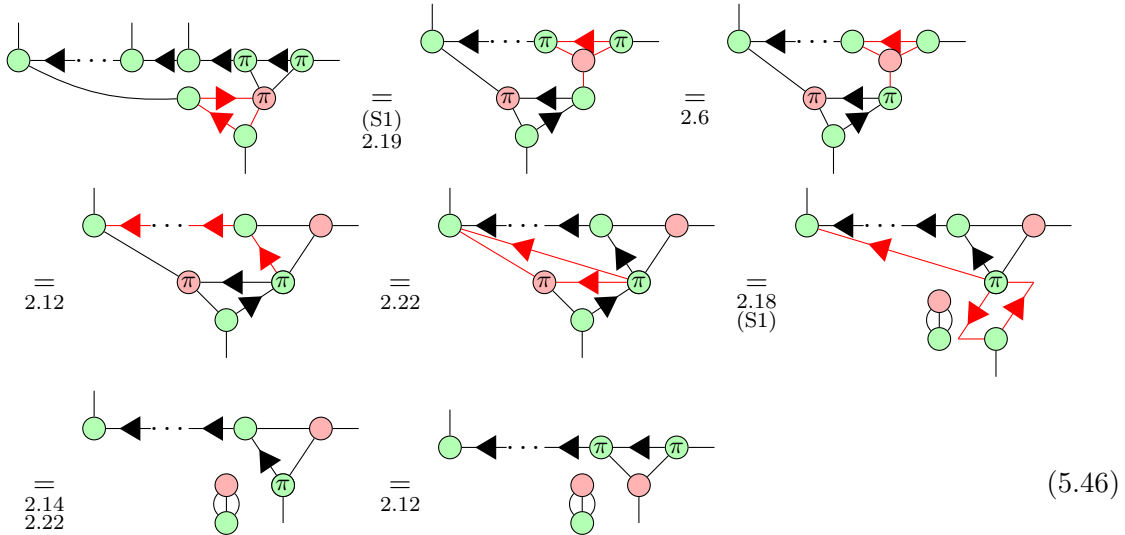
□

Proofs. While proving the claims 5.12 and 5.13 we will repeatedly use the following lemmas:

Lemma 5.14.



Proof. The right hand side transformation:



□

Lemma 5.15. *Diagrams:*

$$C_{\pm\beta} = \begin{array}{c} \pi \\ \blacktriangle \\ \pm\beta \end{array}, \quad C_{\beta}^{(n,m)} = \begin{array}{c} \overbrace{\pi \leftarrow \pi \leftarrow \dots \leftarrow \pi}^n \quad \overbrace{\pi \leftarrow \pi \leftarrow \dots \leftarrow \pi}^m \\ \begin{array}{c} \pi \\ \blacktriangle \\ \beta \end{array} \quad \begin{array}{c} \pi \\ \blacktriangle \\ \beta \end{array} \quad \dots \quad \begin{array}{c} \pi \\ \blacktriangle \\ -\beta \end{array} \quad \begin{array}{c} \pi \\ \blacktriangle \\ -\beta \end{array} \\ \otimes^{n+m} \end{array} \quad (5.47)$$

are controlled states such that

$$\llbracket C_{\pm\beta} \rrbracket |1\rangle = \left[\begin{array}{c} \pi \\ \blacktriangle \\ \pm\beta \end{array} \right], \quad \llbracket C_{\beta}^{(n,m)} \rrbracket |1\rangle = \left[\begin{array}{c} \pi \\ \blacktriangle \\ -\frac{\pi}{2} \end{array} \otimes^{(n+m-1)} \partial_{ZX}[X(n,m)] \right] \quad (5.48)$$

Proof. For the diagram $C_{\pi\beta}$ the claim follows directly from the definition. For the diagram $C_{\beta}^{(n,m)}$ we check two cases:

Therefore the claim holds for the controlled state (5.49) and the scalar $c = \begin{matrix} \otimes(n+m-2) \\ \pi \\ \beta \\ \beta \end{matrix}$ □

claim 5.13. Firstly, we simplify the expression $\tilde{X} \circ \begin{matrix} \pi \\ \beta \\ \beta \end{matrix}$:

(5.51)

We apply the lemma 5.14 to parts in the red frame and obtain:

(5.52)

The claim holds for the scalar $c = \begin{matrix} \otimes(n+m-2) \\ \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \end{matrix}$ □

5.3. Simplified formula for paired spiders. Variational quantum algorithms use gradients in the search for optimal parameter values. The objective minimized by these algorithms can be expressed as $\langle \psi(\beta) | H | \psi(\beta) \rangle$ where the diagram for $\langle \psi(\beta) | = (|\psi(\beta)\rangle)^\dagger$ is obtained out of the diagram for $|\psi(\beta)\rangle$ by flipping upside down followed by the change of signs in spiders.

Therefore, parameters in the diagram for $\langle \psi(\beta) | H | \psi(\beta) \rangle$ appear in pairs $\begin{matrix} -\beta \\ \beta \end{matrix}$.

We suggest a more compact formula for diagrams in what we call *pair-factored form*: $D_2 \circ (D_1 \otimes Y(n))$. In this expression $Y_\beta(n) = \underbrace{\begin{matrix} -\beta & \beta \\ \beta & -\beta \end{matrix} \dots \begin{matrix} -\beta & \beta \\ \beta & -\beta \end{matrix}}_n$.

Lemma 5.16. *The diagram:*

$$\partial_P(Y_\beta(n)) = \begin{matrix} \otimes 3 \\ \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \end{matrix} \begin{matrix} \otimes 2n-1 \\ \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \end{matrix} \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \dots \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \begin{matrix} -\beta \\ \beta \end{matrix} \begin{matrix} -\beta \\ \beta \end{matrix} \begin{matrix} -\beta \\ \beta \end{matrix} \begin{matrix} -\beta \\ \beta \end{matrix} \dots \begin{matrix} -\beta \\ \beta \end{matrix} \begin{matrix} -\beta \\ \beta \end{matrix} \quad (5.53)$$

satisfies $[[\partial_P(Y_\beta(n))]] = \partial_M [[Y_\beta(n)]]$.

The lemma (5.16) can be proven by applying the same approach as in the proof of the formula (5.11) for individual spiders. We leave the exact proof to the reader.

It is possible to extend 5.16 to find the derivative for $X_\beta(n, m)$ when $n \neq m$. Indeed, it is always possible to transform a diagram into an equivalent diagram that has the same

number of occurrences of β and $-\beta$, using the fact that $\begin{matrix} \pm\beta \\ \beta \end{matrix} = \begin{matrix} \square \\ \square \end{matrix}$. For instance, if $n > m$:

$$\partial_P(X_\beta(n, m)) = \sigma \circ \left[\begin{matrix} \otimes n-m \\ \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \end{matrix} \begin{matrix} \otimes 2n-1 \\ \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \end{matrix} \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \dots \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \begin{matrix} -\beta \\ \beta \end{matrix} \begin{matrix} -\beta \\ \beta \end{matrix} \dots \begin{matrix} -\beta \\ \beta \end{matrix} \begin{matrix} -\beta \\ \beta \end{matrix} \right] \quad \text{where } \sigma \text{ is some wire permutation and}$$

$$\begin{matrix} \otimes n-m \\ \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \end{matrix} \begin{matrix} \otimes 2n-1 \\ \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \end{matrix} \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \dots \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \begin{matrix} -\beta \\ \beta \end{matrix} \begin{matrix} -\beta \\ \beta \end{matrix} \dots \begin{matrix} -\beta \\ \beta \end{matrix} \begin{matrix} -\beta \\ \beta \end{matrix} \quad (5.54)$$

Example 5.17. We apply 5.16 to the same diagram as in 5.7:

$$\partial_P \left(\begin{matrix} -\beta & \beta \\ \beta & -\beta \end{matrix} \right) = \begin{matrix} \otimes 3 \\ \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \end{matrix} \begin{matrix} \otimes 2n-1 \\ \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \end{matrix} \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \begin{matrix} -\beta \\ \beta \end{matrix} \begin{matrix} -\beta \\ \beta \end{matrix} \begin{matrix} -\beta \\ \beta \end{matrix} \begin{matrix} -\beta \\ \beta \end{matrix} \dots \begin{matrix} -\beta \\ \beta \end{matrix} \begin{matrix} -\beta \\ \beta \end{matrix} = \begin{matrix} \otimes 3 \\ \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \end{matrix} \begin{matrix} \otimes 2n-1 \\ \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \end{matrix} \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \begin{matrix} -\beta \\ \beta \end{matrix} \begin{matrix} -\beta \\ \beta \end{matrix} \begin{matrix} -\beta \\ \beta \end{matrix} \begin{matrix} -\beta \\ \beta \end{matrix} \dots \begin{matrix} -\beta \\ \beta \end{matrix} \begin{matrix} -\beta \\ \beta \end{matrix} = \begin{matrix} \otimes 2 \\ \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \end{matrix} \begin{matrix} \otimes 2n-1 \\ \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \end{matrix} \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \begin{matrix} \pi & \text{green} \\ \frac{\pi}{2} & \text{red} \end{matrix} \begin{matrix} -\beta \\ \beta \end{matrix} \begin{matrix} -\beta \\ \beta \end{matrix} \begin{matrix} -\beta \\ \beta \end{matrix} \begin{matrix} -\beta \\ \beta \end{matrix} \dots \begin{matrix} -\beta \\ \beta \end{matrix} \begin{matrix} -\beta \\ \beta \end{matrix} \quad (5.55)$$

We observe that using the formula for the pair-factored form (5.53) we obtain a much more compact result than with the inductive procedure (see the example 5.7). Even if both approaches lead to diagrams with the same semantics, in practice the diagrams obtained with the formulas (5.53) and (5.38) are less verbose. For this reason they are easier to manipulate.

5.4. **Discussion.** Contrary to derivatives defined in [TYd21] and [ZG21] in our approaches a derivative of a ZX-diagram is another ZX-diagram.

In the inductive approach we proceed by integration of the product rules using *controlizers*. We also observe that as the derivative of a constant diagram is trivial, it is beneficial to "factor-out" the part that depends on the parameter. We adopted this approach in formulas for diagrams in β -factored form and pair-factored forms. To derive the formulas, we got the intuition in the desired semantics of the diagrams. Thenceforth, our formulas were rigorously proven by induction.

A result similar to our simplified formula for β -factored forms was independently derived in [WY22]. The major difference between our formula and the method shown in [WY22] is the *considered language*. Indeed, in our work we operate ZX-diagrams while Wang and Yeung consider diagrams from the more expressive *algebraic ZX-calculus*.

The difficulty to represent derivatives for non-linear diagrams follows from the fact that there is no simple way to represent real numbers in the vanilla ZX-calculus. In the algebraic ZX-calculus this restriction is removed. As a consequence, when an algebraic ZX-diagram is parameterized by an arbitrary derivable function $f(x)$, the differentiated algebraic ZX-diagram is parametrized by $f'(x)$.

6. DIAGRAMMATIC REPRESENTATION OF ISING HAMILTONIANS

The addition of diagrams naturally appears in the representation of Hamiltonians. Usually, Hamiltonians are given as a weighted sum of Pauli tensors (see [Ves22] for an introduction to Hamiltonians):

$$H = \sum_k \alpha_k \sigma_k^0 \otimes \cdots \otimes \sigma_k^n, \quad \alpha_k \in \mathbb{R}, \sigma_k^q = I \text{ or } X \text{ or } Y \text{ or } Z \quad (6.1)$$

Pauli tensors have a particularly simple decomposition on elementary matrices. Therefore, the procedure described in [YWS] is well-suited for the construction of diagrammatic representation for such Hamiltonians.

In addition, the work [YWS] also suggests a way to *exponentiate* diagrams representing Hamiltonians. The exponentiate of a Hamiltonian H is a unitary $e^{i\theta H}$. The exponentiation is done with the aid of the *Cayley-Hamilton theorem*. The Cayley-Hamilton theorem allows to decompose the exponent $e^{i\theta H}$ on a linear combination of $H^k = \underbrace{H \times \cdots \times H}_k$ for

$k \in [1, \dots, 2^n]$.

We remark that in a particular case when $H = \sum_i H_i$ is a sum of commuting terms it is usually straightforward to obtain $e^{i\theta H}$ as a product of individual $e^{i\theta H_i}$. However, if some local terms in H don't commute the product decomposition doesn't apply. For instance, this is the case for *modified mixer Hamiltonians* used in the extension of QAOA to constrained optimization problems [HWO⁺19].

In this section, we will also consider an inverse problem. Using *Stone's theorem*, we demonstrate how to obtain a diagram for H given the *derivative* of the diagram for $e^{i\theta H}$.

The Stone theorem relating the derivative of a unitary group $e^{i\gamma H}$ to its generator H was used before in [TYd21]. However, in [TYd21] the derivatives were defined as formal sums of diagrams. As a consequence, the work [TYd21] didn't suggest a single diagram representation for a Hamiltonian but rather a representation as a sum of diagrams.

We remark that for an Ising Hamiltonian H the diagram $D_U(\beta)$ of the linear map $U(\beta) = e^{i\beta H}$ is easy to find. For Hamiltonians with integer coefficients the matrix $U(\beta) = e^{i\beta H}$ belongs to $\mathcal{M}(\beta)$. It satisfies the definition of a strongly continuous one-parameter unitary group:

Definition 6.1 (Unitary group [TYd21]). A one-parameter unitary group is a unitary matrix $U : n \rightarrow n$ in $\mathcal{M}(\beta)$ with $U(0) = id_n$ and $U(\beta)U(\beta') = U(\beta + \beta')$ for all $\beta, \beta' \in \mathbb{R}$. It is strongly continuous when $\lim_{\beta \rightarrow \beta_0} U(\beta) = U(\beta_0)$ for all $\beta_0 \in \mathbb{R}$.

Theorem 6.2 (Stone ([Sto32])). *There is a one-to-one correspondence between strongly continuous one-parameter unitary groups $U : n \rightarrow n$ in $\mathcal{M}(\beta)$ and self-adjoint matrices $H : n \rightarrow n$ in \mathcal{M} . The bijection is given explicitly by $U(\beta) = e^{i\beta H}$ and $H = -i(\partial_M U)(0)$.*

We use the bijection from Stone’s theorem to find the diagram $h \in ZX_{\mathbb{R}}$ such that $\llbracket h \rrbracket = H$. Using the property $U(0) = id_n$ we obtain:

$$\begin{aligned} H &= -i[\partial_M U(\beta)](0) = -i \otimes \llbracket [\partial_{ZX} D_U](\beta) \rrbracket(0) = -i \otimes \llbracket [\partial_{ZX} D_U](0) \rrbracket \\ &= \left[\left[\begin{array}{c} \pi \\ -\frac{\pi}{2} \end{array} \right] [\partial_{ZX} D_U](0) \right] = \llbracket h \rrbracket \end{aligned} \tag{6.2}$$

where the third equality is due to the fact that the evaluation commutes with the standard interpretation.

We give an example of a diagram for an Ising Hamiltonian obtained via our approach.

Example 6.3. Let $H : 2 \rightarrow 2$ be the Hamiltonian given by $H = Z_1 - Z_2 + Z_1 Z_2$. The diagram $D_U(\beta)$ for $U(\beta) = e^{i\beta H}$ is:

$$D_U(\beta) = \begin{array}{c} \beta \\ \pi \end{array} \begin{array}{c} -\beta \\ -\beta \end{array} \begin{array}{c} \beta \\ -\beta \end{array} \begin{array}{c} -2\beta \end{array} = \begin{array}{c} \beta \\ \pi \end{array} \begin{array}{c} \beta \\ -\beta \end{array} \begin{array}{c} -\beta \\ -\beta \end{array} \begin{array}{c} -\beta \\ -\beta \end{array} \tag{6.3}$$

Using the formula (5.38) we find the derivative of $D_U(\beta)$:

$$\partial_{ZX} D_U(\beta) = \begin{array}{c} \pi \\ \frac{\pi}{2} \end{array} \begin{array}{c} \pi \\ \pi \end{array} \begin{array}{c} \otimes 3 \\ \otimes 2 \end{array} \begin{array}{c} \pi \\ \beta \\ \pi \end{array} \begin{array}{c} \pi \\ \beta \\ \pi \end{array} \begin{array}{c} \pi \\ -\beta \\ \pi \end{array} \begin{array}{c} \pi \\ -\beta \\ \pi \end{array} \begin{array}{c} \pi \\ -\beta \\ \pi \end{array} \tag{6.4}$$

$$h = \begin{array}{c} \pi \\ -\frac{\pi}{2} \end{array} \begin{array}{c} \pi \\ \pi \end{array} \llbracket [\partial_{ZX} D_U(\beta)]_{\beta \rightarrow 0} \rrbracket = \begin{array}{c} \pi \\ \pi \end{array} \begin{array}{c} \pi \\ \pi \end{array} \begin{array}{c} \pi \\ \pi \end{array} \begin{array}{c} \pi \\ \pi \end{array} \begin{array}{c} \pi \\ \pi \end{array} \begin{array}{c} \pi \\ \pi \end{array} \begin{array}{c} \pi \\ \pi \end{array} \begin{array}{c} \pi \\ \pi \end{array} \tag{6.5}$$

It is of course possible to obtain a diagram for H using the addition procedure we obtained earlier, rather than going through computing its exponential and then differentiating. However, in the case when H is Ising, the exponential of H is typically easier to obtain than H itself.

7. DISCUSSIONS

In this work, we have introduced an inductive definition for addition of ZX-diagrams, that we have then used to introduce an inductive definition of the differentiation of ZX-diagrams. Addition and differentiation are essential tools for the development and the study of quantum algorithms, but, as a matter of fact, both of them are leading to large diagrams, even when the initial diagrams are fairly simple.

In sections 5.2 and 5.3, we have shown that instead of simplifying the resulting diagrams *a posteriori*, one can *a priori* put the initial diagrams in an appropriate form. While this approach is not inductive anymore, it seems to ease the differentiation of diagrams in practice. As an application we have shown that our result allows the construction of diagrams for Ising Hamiltonians and for derivatives of parametrized circuits. Therefore, it becomes possible to study variational algorithms entirely within the ZX-calculus. In particular, we can use rewrite rules to simplify such expressions as $\langle \psi(\hat{\beta}) | H_f | \psi(\hat{\beta}) \rangle$ and $\frac{\partial \langle \psi(\hat{\beta}) | H_f | \psi(\hat{\beta}) \rangle}{\partial \beta}$. We believe that it will lead to a better understanding of the potential of variational algorithms and of their applications to real-world problems.

ACKNOWLEDGEMENTS

This work was supported in part by the CIFRE EDF/Loria Quantum Computing for Combinatorial Optimisation, the French National Research Agency (ANR) under the research projects SoftQPro ANR-17-CE25-0009-02 and VanQuTe ANR-17-CE24-0035, the PEPR integrated project EPiQ ANR-22-PETQ-0007 part of Plan France 2030, by the STIC-AmSud project Qapla' 21-STIC-10, and by the European projects NEASQC (funded from the European Union's Horizon 2020 research and innovation programme grant agreement No 951821) and HPCQS (European High-Performance Computing Joint Undertaking under grant agreement No 101018180).

REFERENCES

- [Bac14] Miriam Backens. The ZX-calculus is complete for stabilizer quantum mechanics. *New Journal of Physics*, 16(9):093021, sep 2014. doi:10.1088/1367-2630/16/9/093021.
- [BK21] Lennart Bittel and Martin Kliesch. Training variational quantum algorithms is NP-hard. *Physical Review Letters*, 127(12), sep 2021. doi:10.1103/physrevlett.127.120502.
- [CAB⁺21] M. Cerezo, Andrew Arrasmith, Ryan Babbush, Simon C. Benjamin, Suguru Endo, Keisuke Fujii, Jarrod R. McClean, Kosuke Mitarai, Xiao Yuan, Lukasz Cincio, and Patrick J. Coles. Variational quantum algorithms. *Nature Reviews Physics*, 3(9):625–644, Aug 2021. doi:10.1038/s42254-021-00348-9.
- [CD11] Bob Coecke and Ross Duncan. Interacting quantum observables: categorical algebra and diagrammatics. *New Journal of Physics*, 13(4):043016, apr 2011. doi:10.1088/1367-2630/13/4/043016.
- [CDD⁺20] Alexander Cowtan, Silas Dilkes, Ross Duncan, Will Simmons, and Seyon Sivarajah. Phase gadget synthesis for shallow circuits. *Electronic Proceedings in Theoretical Computer Science*, 318:213–228, may 2020. doi:10.4204/eptcs.318.13.
- [CK17] Bob Coecke and Aleks Kissinger. *Picturing Phases and Complementarity*, pages 510–623. Cambridge University Press, 2017. doi:10.1017/9781316219317.010.
- [CSV⁺21] M. Cerezo, Akira Sone, Tyler Volkoff, Lukasz Cincio, and Patrick J. Coles. Cost function dependent barren plateaus in shallow parametrized quantum circuits. *Nature Communications*, 12(1), mar 2021. doi:10.1038/s41467-021-21728-w.

- [CW12] Andrew M. Childs and Nathan Wiebe. Hamiltonian Simulation Using Linear Combinations of Unitary Operations. *arXiv e-prints*, page arXiv:1202.5822, February 2012. arXiv:1202.5822.
- [dBBW20] Niel de Beaudrap, Xiaoning Bian, and Quanlong Wang. Techniques to reduce $-/4$ -parity-phase circuits, motivated by the ZX calculus. *Electronic Proceedings in Theoretical Computer Science*, 318:131–149, may 2020. doi:10.4204/eptcs.318.9.
- [dBH20] Niel de Beaudrap and Dominic Horsman. The ZX calculus is a language for surface code lattice surgery. *Quantum*, 4:218, January 2020. doi:10.22331/q-2020-01-09-218.
- [DKPvdW20] Ross Duncan, Aleks Kissinger, Simon Perdrix, and John van de Wetering. Graph-theoretic Simplification of Quantum Circuits with the ZX-calculus. *Quantum*, 4:279, jun 2020. doi:10.22331/q-2020-06-04-279.
- [DP10] Ross Duncan and Simon Perdrix. Rewriting measurement-based quantum computations with generalised flow. In Samson Abramsky, Cyril Gavoille, Claude Kirchner, Friedhelm Meyer auf der Heide, and Paul G. Spirakis, editors, *Automata, Languages and Programming*, pages 285–296, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg.
- [DP14] Ross Duncan and Simon Perdrix. Pivoting makes the ZX-calculus complete for real stabilizers. *Electronic Proceedings in Theoretical Computer Science*, 171:50–62, Dec 2014. doi:10.4204/eptcs.171.5.
- [FGG14] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. A quantum approximate optimization algorithm, 2014. doi:10.48550/ARXIV.1411.4028.
- [GD18] Liam Garvie and Ross Duncan. Verifying the smallest interesting colour code with quantomatic. *Electronic Proceedings in Theoretical Computer Science*, 266:147–163, feb 2018. doi:10.4204/eptcs.266.10.
- [GS17] Gian Giacomo Guerreschi and Mikhail Smelyanskiy. Practical optimization for hybrid quantum-classical algorithms, 2017. arXiv:1701.01450.
- [Had15] Amar Hadzihasanovic. A diagrammatic axiomatisation for qubit entanglement. In *2015 30th Annual ACM/IEEE Symposium on Logic in Computer Science*, pages 573–584, 2015. doi:10.1109/LICS.2015.59.
- [HEMN20] Michael Hanks, Marta P. Estarellas, William J. Munro, and Kae Nemoto. Effective compression of quantum braided circuits aided by ZX-calculus. *Physical Review X*, 10(4), nov 2020. doi:10.1103/physrevx.10.041030.
- [HNW18] Amar Hadzihasanovic, Kang Feng Ng, and Quanlong Wang. 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, pages 502–511, New York, NY, USA, 2018. ACM. doi:10.1145/3209108.3209128.
- [Hor11] Clare Horsman. Quantum pictorialism for topological cluster-state computing. *New Journal of Physics*, 13(9):095011, sep 2011. doi:10.1088/1367-2630/13/9/095011.
- [HWO⁺19] Stuart Hadfield, Zhihui Wang, Bryan O’Gorman, Eleanor Rieffel, Davide Venturelli, and Rupak Biswas. From the quantum approximate optimization algorithm to a quantum alternating operator ansatz. *Algorithms*, 12(2):34, feb 2019. doi:10.3390/a12020034.
- [ILY21] Artur F. Izmaylov, Robert A. Lang, and Tzu-Ching Yen. Analytic gradients in variational quantum algorithms: Algebraic extensions of the parameter-shift rule to general unitary transformations. *Physical Review A*, 104(6), dec 2021. doi:10.1103/physreva.104.062443.
- [JPV18a] Emmanuel Jeandel, Simon Perdrix, and Renaud Vilmart. A Complete Axiomatisation of the ZX-Calculus for Clifford+T Quantum Mechanics. In *The 33rd Annual {ACM/IEEE} Symposium on Logic in Computer Science, {LICS} 2018*, Proceedings of the 33rd Annual ACM/IEEE Symposium on Logic in Computer Science, pages 559–568, Oxford, United Kingdom, July 2018. doi:10.1145/3209108.3209131.
- [JPV18b] Emmanuel Jeandel, Simon Perdrix, and Renaud Vilmart. Diagrammatic Reasoning beyond Clifford+T Quantum Mechanics. In *The 33rd Annual Symposium on Logic in Computer Science*, Proceedings of the 33rd Annual ACM/IEEE Symposium on Logic in Computer Science, pages 569–578, Oxford, United Kingdom, July 2018. doi:10.1145/3209108.3209139.
- [JPV19] Emmanuel Jeandel, Simon Perdrix, and Renaud Vilmart. A generic normal form for ZX-diagrams and application to the rational angle completeness. In *Proceedings of the 34th Annual ACM/IEEE Symposium on Logic in Computer Science*, LICS '19. IEEE Press, 2019.

- [JPV22] Emmanuel Jeandel, Simon Perdrix, and Margarita Veshchezerova. Addition and Differentiation of ZX-diagrams. *arXiv e-prints*, page arXiv:2202.11386, February 2022. arXiv:2202.11386.
- [JPVW17] Emmanuel Jeandel, Simon Perdrix, Renaud Vilmart, and Quanlong Wang. ZX-Calculus: Cyclotomic Supplementarity and Incompleteness for Clifford+T Quantum Mechanics. In Kim G. Larsen, Hans L. Bodlaender, and Jean-Francois Raskin, editors, *42nd International Symposium on Mathematical Foundations of Computer Science (MFCS 2017)*, volume 83 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 11:1–11:13, Dagstuhl, Germany, 2017. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. doi:10.4230/LIPIcs.MFCS.2017.11.
- [KB22] Bálint Koczor and Simon C. Benjamin. Quantum analytic descent. *Physical Review Research*, 4(2), apr 2022. doi:10.1103/physrevresearch.4.023017.
- [KE21] Oleksandr Kyriienko and Vincent E. Elfving. Generalized quantum circuit differentiation rules. *Phys. Rev. A*, 104:052417, Nov 2021. doi:10.1103/PhysRevA.104.052417.
- [KvdW19] Aleks Kissinger and John van de Wetering. Universal MBQC with generalised parity-phase interactions and pauli measurements. *Quantum*, 3:134, apr 2019. doi:10.22331/q-2019-04-26-134.
- [KvdW20] Aleks Kissinger and John van de Wetering. Reducing the number of non-Clifford gates in quantum circuits. *Physical Review A*, 102(2), aug 2020. doi:10.1103/physreva.102.022406.
- [KvdWK19] Stach Kuijpers, John van de Wetering, and Aleks Kissinger. Graphical fourier theory and the cost of quantum addition, 2019. doi:10.48550/ARXIV.1904.07551.
- [MBK21] Andrea Mari, Thomas R. Bromley, and Nathan Killoran. Estimating the gradient and higher-order derivatives on quantum hardware. *Physical Review A*, 103(1), jan 2021. doi:10.1103/physreva.103.012405.
- [MBS⁺18] Jarrod R. McClean, Sergio Boixo, Vadim N. Smelyanskiy, Ryan Babbush, and Hartmut Neven. Barren plateaus in quantum neural network training landscapes. *Nature Communications*, 9(1), nov 2018. doi:10.1038/s41467-018-07090-4.
- [MNKF18] K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii. Quantum circuit learning. *Physical Review A*, 98(3), sep 2018. doi:10.1103/physreva.98.032309.
- [NW18] Kang Feng Ng and Quanlong Wang. Completeness of the ZX-calculus for pure qubit Clifford+T quantum mechanics, 2018. arXiv:1801.07993.
- [PMS⁺14] Alberto Peruzzo, Jarrod McClean, Peter Shadbolt, Man-Hong Yung, Xiao-Qi Zhou, Peter J. Love, Alán Aspuru-Guzik, and Jeremy L. O’Brien. A variational eigenvalue solver on a photonic quantum processor. *Nature Communications*, 5(1), jul 2014. doi:10.1038/ncomms5213.
- [Pre18] John Preskill. Quantum computing in the NISQ era and beyond. *Quantum*, 2:79, aug 2018. doi:10.22331/q-2018-08-06-79.
- [SBG⁺19] Maria Schuld, Ville Bergholm, Christian Gogolin, Josh Izaac, and Nathan Killoran. Evaluating analytic gradients on quantum hardware. *Physical Review A*, 99(3), mar 2019. doi:10.1103/physreva.99.032331.
- [SH22] Tobias Stollenwerk and Stuart Hadfield. Diagrammatic Analysis for Parameterized Quantum Circuits. *arXiv e-prints*, page arXiv:2204.01307, April 2022. arXiv:2204.01307.
- [Sto32] M. H. Stone. On one-parameter unitary groups in hilbert space. *Annals of Mathematics*, 33(3):643–648, 1932. URL: <http://www.jstor.org/stable/1968538>.
- [TYd21] Alexis Toumi, Richie Yeung, and Giovanni de Felice. Diagrammatic differentiation for quantum machine learning. *arXiv e-prints*, page arXiv:2103.07960, March 2021. arXiv:2103.07960.
- [vdW20] John van de Wetering. ZX-calculus for the working quantum computer scientist, 2020. arXiv:2012.13966.
- [Ves22] Margarita Veshchezerova. *Quantum algorithms for energy management optimization problems*. Theses, Université de Lorraine, December 2022. URL: <https://hal.univ-lorraine.fr/te1-04105922>.
- [Vil19] Renaud Vilmart. A near-optimal axiomatisation of ZX-calculus for pure qubit quantum mechanics. In *Proceedings of the 34th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*, 2019. URL: <https://arxiv.org/abs/1812.09114>, arXiv:arXiv:1812.09114.
- [Wan20] Quanlong Wang. Algebraic complete axiomatisation of ZX-calculus with a normal form via elementary matrix operations, 2020. doi:10.48550/ARXIV.2007.13739.
- [WIWL22] David Wierichs, Josh Izaac, Cody Wang, and Cedric Yen-Yu Lin. General parameter-shift rules for quantum gradients. *Quantum*, 6:677, March 2022. doi:10.22331/q-2022-03-30-677.

- [WSW⁺21] Max Wilson, Rachel Stromswold, Filip Wudarski, Stuart Hadfield, Norm Tubman, and Eleanor Rieffel. Optimizing quantum heuristics with meta-learning. *Quantum Machine Intelligence*, 3, 06 2021. doi:10.1007/s42484-020-00022-w.
- [WY21] Quanlong Wang and Richie Yeung. Representing and implementing matrices using algebraic ZX-calculus, 2021. doi:10.48550/ARXIV.2110.06898.
- [WY22] Quanlong Wang and Richie Yeung. Differentiating and integrating ZX diagrams, 2022. arXiv:2201.13250.
- [YWS] Richie Yeung, Quanlong Wang, and Razin A. Shaikh. *How to sum and exponentiate Hamiltonians in ZXW calculus*. URL: https://www.qplconference.org/proceedings2022/QPL_2022_paper_85.pdf.
- [ZG21] Chen Zhao and Xiao-Shan Gao. Analyzing the barren plateau phenomenon in training quantum neural networks with the ZX-calculus. *Quantum*, 5:466, June 2021. doi:10.22331/q-2021-06-04-466.