Imperial College London

Compact Matrix Quantum Group Equivariant Neural Networks

Quantum Groups Seminar, University of Lille

Edward Pearce-Crump

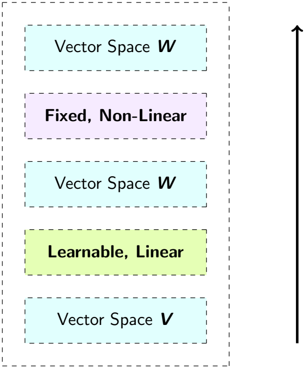
May 20, 2024

Outline

1. Neural Networks, Symmetries and Groups

Neural Networks

Neural networks consist of a **composition** of **layers**, where each layer has the following form:



The learnable, linear layer function is often given in the form of a **parameterised weight matrix** so that learning can take place.

Machine Learning, Symmetries and Groups

In machine learning, we would like to **develop principled approaches** for constructing neural networks.

One important approach is

- to identify symmetries that exist in data (e.g. permutation symmetry in a set of objects),
- view the symmetries formally as groups (e.g symmetric group), and then
- create neural network architectures that take advantage of the group symmetries in the data when performing learning.

In some cases, this has been done by using

- **layer spaces** that are tensor power representations of a subgroup $G(n)$ of $GL(n)$, that is, a group homomorphism

$$\rho_k : G(n) \rightarrow GL((\mathbb{R}^n)^{\otimes k}) \tag{1}$$

given by

$$\rho_k(g)(v_1 \otimes \cdots \otimes v_k) := gv_1 \otimes \cdots \otimes gv_k \tag{2}$$

for all $g \in G(n)$ and for all vectors $v_i \in \mathbb{R}^n$.

and

- **layer functions** that are equivariant to $G(n)$:

Definition

If $\rho_k : G(n) \to GL((\mathbb{R}^n)^{\otimes k})$ and $\rho_l : G(n) \to GL((\mathbb{R}^n)^{\otimes l})$ are two representations of $G(n)$, then $\phi : (\mathbb{R}^n)^{\otimes k} \to (\mathbb{R}^n)^{\otimes l}$ is said to be **equivariant** to $G(n)$ if

$$\phi(\rho_k(g)[v]) = \rho_l(g)[\phi(v)] \tag{3}$$

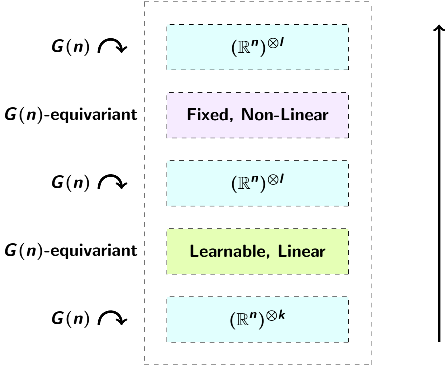
for all $g \in G(n)$ and $v \in (\mathbb{R}^n)^{\otimes k}$.

$$\begin{array}{ccc} (\mathbb{R}^n)^{\otimes k} & \xrightarrow{\rho_k(g)} & (\mathbb{R}^n)^{\otimes k} \\ \phi \downarrow & & \downarrow \phi \\ (\mathbb{R}^n)^{\otimes l} & \xrightarrow{\rho_l(g)} & (\mathbb{R}^n)^{\otimes l} \end{array} \tag{4}$$

to give ...

Group Equivariant Neural Networks

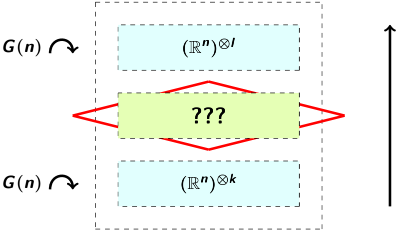
The layers have the following form:



In particular, if $G(n)$ is compact, then we call this network a **compact matrix group equivariant neural network**.

Weight Matrix Classfications

Much work has been done to **characterise** all of the possible equivariant, learnable, linear layers that appear in a group equivariant neural network for different groups $G(n)$.



- S_n: Maron et al. (2018), Ravanbakhsh (2020), Pearce-Crump (2022)
- $A_n, O(n), Sp(n), SO(n)$: Finzi (2021), Pearce-Crump (2023)
- and so on!

There is a but ...

As quantum group people, we know that:

There exist symmetries that cannot be understood formally as groups.

2. Research Question

Research Question

Can we construct neural networks that take advantage of **quantum symmetries** in data?

3. Compact Matrix Quantum Groups

The Passage from Groups to Quantum Groups

Recommended: An excellent motivation piece written by Weber (2020) titled Quantum Symmetry.

Let $G(n)$ be a compact matrix group. Consider the C*-algebra $C(G(n))$ of continuous functions on $G(n)$, which is commutative. Firstly, we can define functions $u_{i,j} : G(n) \to \mathbb{C}$ for $1 \leq i, j \leq n$ such that $u_{i,j}(g) = g_{i,j}$. Note that the $u_{i,j}$ generate $C(G(n))$ and the matrices $u := (u_{i,j})$ and $u^\top$ are invertible.

Moreover, by considering the composition $\circ$ in $G(n)$, we have a map

$$\Delta : C(G(n)) \to C(G(n) \times G(n)) \tag{5}$$

$$\Delta(f)(g, h) := f(g \circ h) \tag{6}$$

We also have that $C(G(n) \times G(n)) \cong C(G(n)) \otimes C(G(n))$, which motivates the following definition ...

Compact Matrix Quantum Groups

Definition

Let A be a C^*-algebra, and let $u_{ij} \in A$, for all $i, j \in [n]$, for some $n \in \mathbb{N}$. Let u be the $n \times n$ matrix whose (i, j)-entry is u_{ij}, that is, $u \in M_n(A)$. The pair (A, u) is said to be a **compact matrix quantum group** if

1. the elements u_{ij} generate A,
2. u and $u^\top = (u_{ji})$ are invertible matrices, and
3. the comultiplication map $\Delta : A \to A \otimes A$ defined by

$$\Delta(u_{ij}) := \left(\sum_k u_{ik} \otimes u_{kj} \right) \tag{7}$$

is a $*$-homomorphism.

Convention dictates that we denote A by $C(G)$, and so we often refer to G, or sometimes the pair (G, u), as the compact matrix quantum group.

Fundamental Theorem of Compact Matrix Quantum Groups

Theorem

Let (A, u) be a compact matrix quantum group for $n \in \mathbb{N}$.

Then A is commutative if and only if $A \cong C(G(n))$ for some compact matrix subgroup $G(n)$ of $GL(n)$.

Hence if A is noncommutative, we obtain true quantum groups!

Just as groups have representations, so do quantum groups.

Definition

Let (G, u) be a compact matrix quantum group. Then an ***n*-dimensional representation** of G is an element $v \in M_n(C(G))$ such that

$$\Delta(v_{ij}) = \left(\sum_k v_{ik} \otimes v_{kj} \right) \tag{8}$$

Note that:

The matrix u given in the definition of a compact matrix quantum group (G, u) is an n-dimensional representation called the **fundamental representation** of G.

We can also define tensor product and complex conjugation, as follows.

Definition

Let $v \in M_n(C(G))$ and $w \in M_m(C(G))$.

- The **tensor product** $v \otimes w \in M_{nm}(C(G))$ is simply the Kronecker product of matrices, and
- The **complex conjugate** is $\bar{v} = (v_{ij}^*) \in M_n(C(G))$.

Similar to group representations, if v, w are representations of a compact matrix quantum group (G, u), then it can be shown that both the tensor product and complex conjugate are also representations of G.

We can also define the analogous concept of equivariance for compact matrix quantum groups.

Definition

Let $v \in M_n(C(G))$ and $w \in M_m(C(G))$ be representations of a compact matrix quantum group G.

A map $\phi : \mathbb{C}^n \to \mathbb{C}^m$ is said to be G-**equivariant** (also known as an **intertwiner**) if $\phi v = w\phi$.

We denote the set of all such *linear* maps by $\mathrm{Hom}_G(v, w)$, and it can be shown to be a vector space.

4. Woronowicz–Tannaka–Krein Duality

Woronowicz–Tannaka–Krein Duality

Intuition

The Woronowicz formulation of Tannaka–Krein duality is an important result in the theory of compact matrix quantum groups.

Informally, the duality says that we can construct a compact matrix quantum group just by knowing its fundamental representation category.

Two-Coloured Words

Consider the **two-coloured set** $\{\circ, \bullet\}$ consisting of a white point and a black point.

- For any non-negative integer k, we can construct a **word** w of length k as a string of k colours from $\{\circ, \bullet\}$.
- If $k = 0$, we define $\varnothing$ to be the empty word.

1. $w_1 := \circ \bullet \bullet \circ \bullet \circ$ is a word of length 6
2. $w_2 := \bullet \circ \bullet \bullet \circ$ is a word of length 5.
3. $w_3 := \varnothing$ is the word of length 0.

Two-Coloured Words

- The **concatenation** of two words, w_1, w_2, is the concatenation of their two strings and is written as $w_1 \cdot w_2$.

If $w_1 := \circ \bullet \bullet \circ \bullet \circ$ and $w_2 := \bullet \circ \bullet \bullet \circ$, then $w_1 \cdot w_2 = \circ \bullet \bullet \circ \bullet \circ \bullet \circ \bullet \bullet \circ$

- There is a **homomorphism on words**, $w \mapsto \bar{w}$, that is first defined on the individual colours by $\bar{\circ} := \bullet, \bar{\bullet} := \circ$, and is then applied element wise to the word w.

If $w_1 := \circ \bullet \bullet \circ \bullet \circ$, then $\bar{w}_1 := \bullet \circ \circ \bullet \circ \bullet$

- If w is a word, then its **involution** w^* is the word read backwards together with its colours inverted.

If $w_1 := \circ \bullet \bullet \circ \bullet \circ$, then $w_1^* := \bullet \circ \bullet \circ \circ \bullet$

Tensor Products of the Fundamental Representation

We can use words of colours to create tensor products of the fundamental representation of a compact matrix quantum group as follows.

Definition

Let (G, u) be a compact matrix quantum group, where $u \in M_n(C(G))$. Let $u^\circ := u$ and $u^\bullet := \bar{u}$. Then for any word w formed from the two-coloured set $\{\circ, \bullet\}$, we define $u^{\otimes w}$ to be the corresponding tensor product of representations u° and $u^\bullet$.

If w is the word $\circ \bullet \bullet \circ$, then

$$u^{\otimes w} = u \otimes \bar{u} \otimes \bar{u} \otimes u \tag{9}$$

Space of Intertwiners for Tensor Products of the Fundamental Representation

If w_k and w_l are words of lengths k, l respectively, then we define $\text{FundRep}_G(w_k, w_l)$ to be $\text{Hom}_G(u^{\otimes w_k}, u^{\otimes w_l})$, that is, the vector space

$$\{\phi : (\mathbb{C}^n)^{\otimes k} \to (\mathbb{C}^n)^{\otimes l} \mid \phi u^{\otimes w_k} = u^{\otimes w_l} \phi\} \tag{10}$$

of linear G-equivariant maps.

Two-Coloured Representation Categories

Definition

A **two-coloured representation category** $\mathcal{C}$ is a collection of subspaces $\mathcal{C}(w_k, w_l)$, where w_k, w_l are two words constructed from $\{\circ, \bullet\}$ of lengths k, l respectively, that are subspaces of the set of all linear maps $(\mathbb{C}^n)^{\otimes k} \to (\mathbb{C}^n)^{\otimes l}$ and which satisfy the following axioms.

1. If $\phi_1 \in \mathcal{C}(w_k, w_l)$, $\phi_2 \in \mathcal{C}(w'_k, w'_l)$, then $\phi_1 \otimes \phi_2 \in \mathcal{C}(w_k \cdot w'_k, w_l \cdot w'_l)$.
2. If $\phi_1 \in \mathcal{C}(w_k, w_l)$, $\phi_2 \in \mathcal{C}(w_l, w_m)$, then $\phi_2 \circ \phi_1 \in \mathcal{C}(w_k, w_m)$.
3. If $\phi \in \mathcal{C}(w_k, w_l)$, then $\phi^* \in \mathcal{C}(w_l, w_k)$.
4. For every word w (having some length k), we have $1_n^{\otimes k} \in \mathcal{C}(w, w)$, and
5. The colours $\circ$ and $\bullet$ are dual to each other.

A **one-coloured representation category** is a two-coloured representation category where $u^\circ = u^\bullet$.

From this, we obtain two important results describing the relationship between two-coloured representation categories and compact matrix quantum groups.

Theorem

If (G, u) *is a compact matrix quantum group, then* FundRep_G *is a two-coloured representation category.*

and

Theorem (Woronowicz–Tannaka–Krein Duality)

Let $\mathcal{C}$ *be a two-coloured representation category. Then there exists a unique compact matrix quantum group* (G, u) *such that* $\text{FundRep}_G = \mathcal{C}$*.*

5. Two-Coloured Set Partitions

Idea

If we have a way of creating two-coloured representation categories, then by Woronowicz–Tannaka–Krein duality we can create many compact matrix quantum groups.

Two-coloured set partition categories form a great source of two-coloured representation categories.

Set Partitions

For non-negative integers l and k, consider the set $[l+k] := \{1, \dots, l+k\}$.

We can create a **set partition** of $[l+k]$ by partitioning it into a number of subsets. We call the subsets of a set partition **blocks**.

Let Π_{l+k} be the set of all set partitions of $[l+k]$.

We see that

$$\{1, 6 \mid 2, 3 \mid 4, 8 \mid 5, 9 \mid 7, 10 \mid 11\} \tag{11}$$

is a set partition of $[5+6]$ having six blocks.

One-Coloured Partition Diagrams

Then, for each set partition π in Π_{l+k}, we can associate to it a diagram, called a **(one-coloured)** (k, l)**–partition diagram**, consisting of two rows of vertices and edges between vertices such that there are

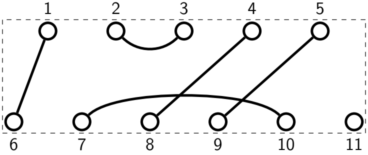
- l vertices on the top row, labelled left to right by $1, \ldots, l$
- k vertices on the bottom row, labelled left to right by $l+1, \ldots, l+k$, and
- the edges between the vertices correspond to the connected components of π.

Technically, the diagram represents the equivalence class of all diagrams with connected components equal to the blocks of π.

For the set partition

$$\pi = \{1, 6 \mid 2, 3 \mid 4, 8 \mid 5, 9 \mid 7, 10 \mid 11\} \tag{12}$$

setting $l = 5$ and $k = 6$, we have that a $(6, 5)$–partition diagram for π is



Two-Coloured Partition Diagrams

A **(two-coloured)** (w_k, w_l)**–partition diagram** d_π is a (k, l)–partition diagram together with two-coloured words w_k, w_l of lengths k, l respectively such that

- the vertices on the top row have the same colours as the word w_l, from left to right, and
- the vertices on the bottom row have the same colours as the word w_k, from left to right.

We define the **two-coloured partition vector space** $P_{w_k}^{w_l}(n)$ to be the $\mathbb{C}$-linear span of the set of all (w_k, w_l)–partition diagrams.

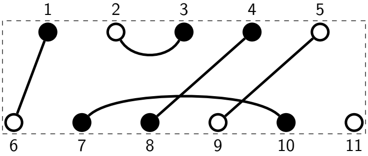
Again, for the set partition

$$\pi = \{1, 6 \mid 2, 3 \mid 4, 8 \mid 5, 9 \mid 7, 10 \mid 11\} \tag{13}$$

setting $l = 5$ and $k = 6$, and choosing

- w_5 to be the word $\bullet \circ \bullet \bullet \circ$, and
- w_6 to be the word $\circ \bullet \bullet \circ \bullet \circ$,

we have that a two-coloured (w_6, w_5)–partition diagram for π is



Operations on Two-Coloured Set Partition Diagrams

We can define three $\mathbb{C}$-(bi)linear operations on two-coloured set partition diagrams:

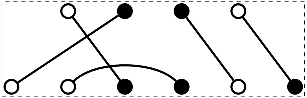
$$\textbf{composition:} \bullet : P_{w_l}^{w_m}(n) \times P_{w_k}^{w_l}(n) \rightarrow P_{w_k}^{w_m}(n) \tag{14}$$

$$\textbf{tensor product:} \otimes : P_{w_k}^{w_l}(n) \times P_{w_q}^{w_m}(n) \rightarrow P_{w_k \cdot w_q}^{w_l \cdot w_m}(n) \tag{15}$$

$$\textbf{involution:} * : P_{w_k}^{w_l}(n) \rightarrow P_{w_l}^{w_k}(n) \tag{16}$$

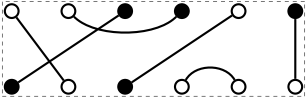
Composition

If d_{π_2} is the $(\circ\circ\bullet\bullet\circ\bullet, \circ\bullet\bullet\circ)$–partition diagram



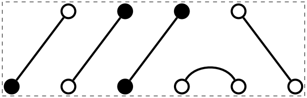
(17)

and d_{π_1} is the two-coloured $(\bullet\circ\bullet\circ\circ\circ, \circ\circ\bullet\bullet\circ\bullet)$–partition diagram



(18)

then $d_{\pi_2} \circ d_{\pi_1}$ is the $(\bullet\circ\bullet\circ\circ\circ, \circ\bullet\bullet\circ)$–partition diagram

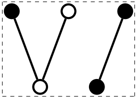


(19)

Then $d_{\pi_2} \bullet d_{\pi_1}$ is the diagram (19) multiplied by n, since one connected component was removed from the middle row of $d_{\pi_2} \circ d_{\pi_1}$.

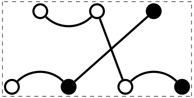
Tensor Product

If d_{π_1} is the two-coloured $(\circ \bullet, \bullet \circ \bullet)$–partition diagram



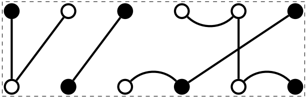
(20)

and d_{π_2} is the two-coloured $(\circ \bullet \circ \bullet, \circ \circ \bullet)$–partition diagram



(21)

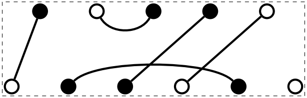
then $d_{\pi_1} \otimes d_{\pi_2}$ is the $(\circ \bullet \circ \bullet \circ \bullet, \bullet \circ \bullet \circ \circ \bullet)$–partition diagram



(22)

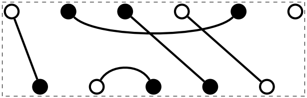
Involution

If d_π is the two-coloured $(\circ \bullet \bullet \circ \bullet \circ, \bullet \circ \bullet \bullet \circ)$–partition diagram



(23)

then d_π^* is the $(\bullet \circ \bullet \bullet \circ, \circ \bullet \bullet \circ \bullet \circ)$–partition diagram



(24)

Two-Coloured Partition Categories

Definition

Using these operations, we form the **two-coloured partition category** $\mathcal{P}(n)$, whose

- set of objects is the set of two-coloured words, and
- whose set of morphisms between words w_k and w_l, $\mathcal{P}(n)(w_k, w_l)$, is the set of all (w_k, w_l)–partition diagrams.

This category can be shown to be a strict monoidal involutive category.

We also have the following important definition.

Definition

A **two-coloured category of partitions** $\mathcal{K}(n)$ is any subcategory of $\mathcal{P}(n)$ such that

1. the collection of sets $\mathcal{K}(n)(w_k, w_l)$, for any words w_k, w_l, are a subset of $\mathcal{P}(n)(w_k, w_l)$,
2. the collection of sets $\mathcal{K}(n)(w_k, w_l)$ is closed under the composition, tensor product and involution operations,
3. the identity partition diagram is an element of $\mathcal{K}(n)(\circ, \circ)$,
4. the (top-row) pair partition diagram corresponding to the set partition $\{1, 2\}$ of $\{1, 2\}$ superimposed with the word $w_2 := \circ\bullet$ is in $\mathcal{K}(n)(\varnothing, w_2)$.

We now create a map that takes

- two-coloured (w_k, w_l)–partition diagrams d_π living in a two-coloured category of partitions to

to

- linear maps $\phi_\pi : (\mathbb{C}^n)^{\otimes k} \to (\mathbb{C}^n)^{\otimes l}$ living in a two-coloured representation category.

By choosing the standard basis for $\mathbb{C}^n$, we obtain $n^l \times n^k$ matrices instead, which are denoted by E_π in what follows.

Suppose that d_π is a (w_k, w_l)–partition diagram. We define E_π as follows.

Associate the indices $i_1, i_2, \dots, i_l$ with the vertices in the top row of d_π and $j_1, j_2, \dots, j_k$ with the vertices in the bottom row of d_π. Then, if $S_\pi((I,J))$ is defined to be the set

$$\left\{(I,J) \in [n]^{l+k} \;\middle|\; \text{if } x, y \text{ are in the same block of } \pi, \text{then } i_x = i_y \right\} \quad (25)$$

(where we have momentarily replaced the elements of J by $i_{l+m} := j_m$ for all $m \in [k]$), we have that

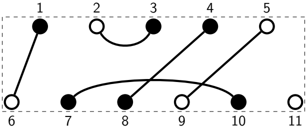
$$E_\pi := \sum_{I \in [n]^l, J \in [n]^k} \delta_{\pi,(I,J)} E_{I,J} \quad (26)$$

where

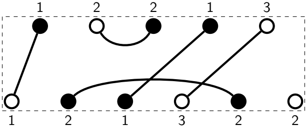
$$\delta_{\pi,(I,J)} := \begin{cases} 1 & \text{if } (I,J) \in S_\pi((I,J)) \\ 0 & \text{otherwise} \end{cases} \quad (27)$$

and $E_{I,J}$ is the $n^l \times n^k$ matrix having a 1 in the (I,J) position and is 0 elsewhere. We extend the definition of the map $d_\pi \mapsto E_\pi$ linearly to $P_{w_k}^{w_l}(n)$.

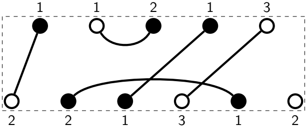
If d_π is the two-coloured $(\circ \bullet \bullet \circ \bullet \circ, \bullet \circ \bullet \bullet \circ)$–partition diagram



then, if $n = 3$, for example, we see from



that the $(1, 2, 2, 1, 3 \mid 1, 2, 1, 3, 2, 2)$-entry of E_π is 1, whereas, from



we see that the $(1, 1, 2, 1, 3 \mid 2, 2, 1, 3, 1, 2)$-entry of E_π is 0.

Monoidal Functor

The map $d_\pi \mapsto E_\pi$ defines a monoidal functor between categories:

Theorem

The map $d_\pi \mapsto E_\pi$ defines a strict $\mathbb{C}$–linear monoidal functor

$$\Theta : \mathcal{P}(n) \rightarrow \mathsf{Mat} \tag{28}$$

In fact, if $\mathcal{K}(n)$ is any two-coloured subcategory of partitions of $\mathcal{P}(n)$, then, letting $\mathcal{C}(n)$ be the image of $\mathcal{K}(n)$ under $d_\pi \mapsto E_\pi$, that is

$$\mathcal{C}(n)(w_k, w_l) = \{E_\pi \mid d_\pi \in \mathcal{K}(n)(w_k, w_l)\} \tag{29}$$

we get that $\mathcal{C}(n)$ is a two-coloured representation category.

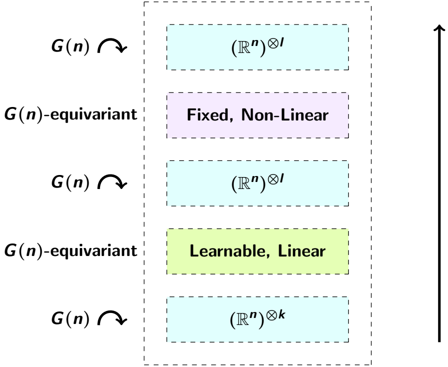
Hence we can use Woronowicz-Tannaka-Krein duality together with the previous result to obtain what we set out to show, namely that two-coloured categories of partitions are a great source of new compact matrix quantum groups. Said precisely, we have

A two-coloured category of partitions $\mathcal{K}(n)$ determines a unique compact matrix quantum group (G, u), where u is the n-dimensional fundamental representation of G.

We call these compact matrix quantum groups **easy**.

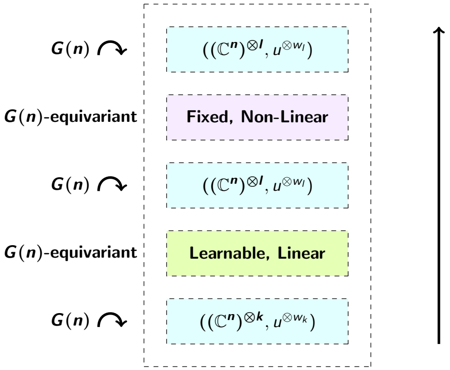
6. Compact Matrix Quantum Group Equivariant Neural Networks

Recall from before a Group Equivariant Neural Network for matrix groups $G(n)$:



Compact Matrix Quantum Group Equivariant Neural Networks

A Compact Matrix Quantum Group Equivariant Neural Network for a CMQG $(G(n), u)$ consists of layers of the form:



Note that the learnable, linear, $G(n)$-equivariant function is in $\text{FundRep}_G(w_k, w_l)$.

The formulation of a compact matrix quantum group equivariant neural network is well-defined since the fundamental representation category of a compact matrix quantum group is a two-coloured representation category.

Moreover, we get:

Theorem

Let f_{NN} be a compact matrix quantum group equivariant neural network for a compact matrix **group** $(G(n), u)$*. If all of the words used in the network only consist of the white point* $\circ$*, then* f_{NN} *is, in fact, a compact matrix group equivariant neural network for* $G(n)$*.*

Hence

Every compact matrix group equivariant neural network is a compact matrix quantum group equivariant neural network.

7. Weight Matrix Classification

We can go one step further with the following intermediate theorem, whose proof is immediate.

Theorem

Suppose that $(G(n), u)$ *is a compact matrix quantum group that has been obtained under Woronowicz–Tannaka–Krein duality from a two-coloured representation category (which is* $\text{FundRep}_{G(n)}$ *by the statement of the duality). Let* W *be a weight matrix of size* $n^l \times n^k$ *appearing in a compact matrix quantum group equivariant neural network for* $G(n)$*, coming from words* w_k, w_l *of lengths* k *and* l *respectively. If* $\{M_1, \dots, M_p\}$ *is a spanning set for* $\text{FundRep}_{G(n)}(w_{l-1}, w_l)$*, then*

$$W = \sum_{i=1}^{p} w_i M_i \tag{30}$$

But if the two-coloured representation category FundRep_G is the image under $d_\pi \mapsto E_\pi$ of a two-coloured category of partitions $\mathcal{K}(n)$, we know what the spanning set is!

Consequently, we can characterise the weight matrices that appear in any easy compact matrix quantum group equivariant neural network, as follows.

Theorem

With the same setup as before, if $(G(n), u)$ is an easy compact matrix quantum group coming from a two-coloured category of partitions $\mathcal{K}(n)$, then

$$W = \sum_{\pi \mid d_\pi \in \mathcal{K}(n)(w_k, w_l)} w_\pi E_\pi \tag{31}$$

Furthermore, we benefit from the fact that characterisations of two-coloured categories of partitions and their associated easy compact matrix quantum groups have been studied extensively in the literature.

We give some important examples in the following slides.

One-Coloured Partition Categories

One-coloured partition categories were introduced initially by Banica and Speicher (2009), studied further in Banica et al. (2010), Weber (2013), Raum and Weber (2014, 2015), and characterised fully by Raum and Weber (2016).

They can be divided into four cases: group, non-crossing, half-liberated, and the rest.

We focus on the first two cases.

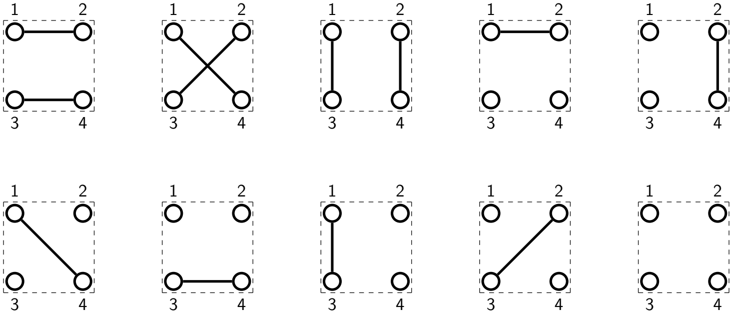
One-Coloured Partition Categories: The Group Case

Banica and Speicher (2009) provided a full characterisation. The image of all (k, l)–partition diagrams satisfying the appropriate conditions on the blocks determines the weight matrices.

Group	**Conditions on the blocks of the (k, l)–partition diagrams**
Symmetric group S_n	—
Orthogonal group $O(n)$	Blocks come in pairs.
Hyperoctahedral group H_n	Blocks are of even size.
Bistochastic group B_n	Blocks are of size one or two.
Modified symmetric group $S_n' := \mathbb{Z}_2 \times S_n$	The number of blocks of odd size is even.
Modified bistochastic group $B_n' := \mathbb{Z}_2 \times B_n$	Has an even number of blocks of size one and any number of blocks of size two.

Example: Bistochastic group B_n, $n = k = l = 2$

We need all $(2, 2)$–partition diagrams whose blocks are of size one or two. They are:



Now apply $d_\pi \mapsto E_\pi$ to each diagram:

$$\begin{array}{r|cc|cc|} & 1,1 & 1,2 & 2,1 & 2,2 \\ 1,1 & 1 & 0 & 0 & 1 \\ 1,2 & 0 & 0 & 0 & 0 \\ \hline 2,1 & 0 & 0 & 0 & 0 \\ 2,2 & 1 & 0 & 0 & 1 \end{array} \quad \begin{array}{r|cc|cc|} & 1,1 & 1,2 & 2,1 & 2,2 \\ 1,1 & 1 & 0 & 0 & 0 \\ 1,2 & 0 & 0 & 1 & 0 \\ \hline 2,1 & 0 & 1 & 0 & 0 \\ 2,2 & 0 & 0 & 0 & 1 \end{array} \quad \begin{array}{r|cc|cc|} & 1,1 & 1,2 & 2,1 & 2,2 \\ 1,1 & 1 & 0 & 0 & 0 \\ 1,2 & 0 & 1 & 0 & 0 \\ \hline 2,1 & 0 & 0 & 1 & 0 \\ 2,2 & 0 & 0 & 0 & 1 \end{array} \quad \begin{array}{r|cc|cc|} & 1,1 & 1,2 & 2,1 & 2,2 \\ 1,1 & 1 & 1 & 1 & 1 \\ 1,2 & 0 & 0 & 0 & 0 \\ \hline 2,1 & 0 & 0 & 0 & 0 \\ 2,2 & 1 & 1 & 1 & 1 \end{array} \quad \begin{array}{r|cc|cc|} & 1,1 & 1,2 & 2,1 & 2,2 \\ 1,1 & 1 & 0 & 1 & 0 \\ 1,2 & 0 & 1 & 0 & 1 \\ \hline 2,1 & 1 & 0 & 1 & 0 \\ 2,2 & 0 & 1 & 0 & 1 \end{array}$$

$$\begin{array}{r|cc|cc|} & 1,1 & 1,2 & 2,1 & 2,2 \\ 1,1 & 1 & 0 & 1 & 0 \\ 1,2 & 1 & 0 & 1 & 0 \\ \hline 2,1 & 0 & 1 & 0 & 1 \\ 2,2 & 0 & 1 & 0 & 1 \end{array} \quad \begin{array}{r|cc|cc|} & 1,1 & 1,2 & 2,1 & 2,2 \\ 1,1 & 1 & 0 & 0 & 1 \\ 1,2 & 1 & 0 & 0 & 1 \\ \hline 2,1 & 1 & 0 & 0 & 1 \\ 2,2 & 1 & 0 & 0 & 1 \end{array} \quad \begin{array}{r|cc|cc|} & 1,1 & 1,2 & 2,1 & 2,2 \\ 1,1 & 1 & 1 & 0 & 0 \\ 1,2 & 1 & 1 & 0 & 0 \\ \hline 2,1 & 0 & 0 & 1 & 1 \\ 2,2 & 0 & 0 & 1 & 1 \end{array} \quad \begin{array}{r|cc|cc|} & 1,1 & 1,2 & 2,1 & 2,2 \\ 1,1 & 1 & 1 & 0 & 0 \\ 1,2 & 0 & 0 & 1 & 1 \\ \hline 2,1 & 1 & 1 & 0 & 0 \\ 2,2 & 0 & 0 & 1 & 1 \end{array} \quad \begin{array}{r|cc|cc|} & 1,1 & 1,2 & 2,1 & 2,2 \\ 1,1 & 1 & 1 & 1 & 1 \\ 1,2 & 1 & 1 & 1 & 1 \\ \hline 2,1 & 1 & 1 & 1 & 1 \\ 2,2 & 1 & 1 & 1 & 1 \end{array}$$

Finally, assign a weight to each matrix, and then add them together to obtain the weight matrix:

$$\begin{array}{r|cc|cc|} & 1,1 & 1,2 & 2,1 & 2,2 \\ 1,1 & w_{1,\ldots,10} & w_{4,8,9,10} & w_{4,5,6,10} & w_{1,4,7,10} \\ 1,2 & w_{6,7,8,10} & w_{3,5,8,10} & w_{2,6,9,10} & w_{5,7,9,10} \\ \hline 2,1 & w_{5,7,9,10} & w_{2,6,9,10} & w_{3,5,8,10} & w_{6,7,8,10} \\ 2,2 & w_{1,4,7,10} & w_{4,5,6,10} & w_{4,8,9,10} & w_{1,\ldots,10} \end{array} \tag{32}$$

One-Coloured Partition Categories: The Non-Crossing Case

In order to state the classifications for the non-crossing case, we first need the following definition.

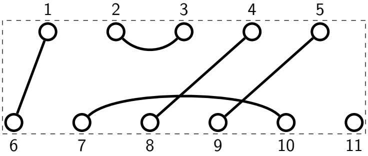
Definition

A set partition diagram d_π corresponding to a set partition π of $[l + k]$ is said to be **crossing** if there exist four integers $1 \leq x_1 < x_2 < x_3 < x_4 \leq l + k$ satisfying:

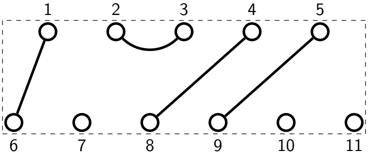
1. x_1 and x_3 are in the same block
2. x_2 and x_4 are in the same block, and
3. x_1 and x_2 are not in the same block.

Otherwise, d_π is said to be **non-crossing**.

While



is crossing,



is non-crossing.

One-Coloured Partition Categories: The Non-Crossing Case

Banica and Speicher (2009) found six via **liberation**; Weber (2013) found the last one (not shown).

Quantum Group	**Conditions on the blocks of the (k, l)–partition diagrams**
Symmetric quantum group S_n^+	**non-crossing**
Orthogonal quantum group $O(n)^+$	Blocks come in pairs, **non-crossing**.
Hyperoctahedral quantum group H_n^+	Blocks are of even size, **non-crossing**.
Bistochastic quantum group B_n^+	Blocks are of size one or two, **non-crossing**.
Modified symmetric quantum group $S_n'^+$	The number of blocks of odd size is even, **non-crossing**.
Modified bistochastic quantum group $B_n'^+$	Has an even number of blocks of size one and any number of blocks of size two, **non-crossing**.

In fact, we obtain a stronger classification for the non-crossing case due to the following theorem by Banica and Speicher (2009).

Theorem

The spanning sets for the seven compact matrix quantum groups in the non-crossing case are in fact bases for $n \geq 4$.

Two-Coloured Partition Categories

Two-coloured partition categories are much richer than one-coloured partition categories because the generators u_{ij} of the corresponding compact matrix quantum groups are no longer self-adjoint.

A full classification is unknown, but Tarrago and Weber (2016; 2018) classified all two-coloured partitions in the group and non-crossing case.

We only discuss $U(n)$ and $U(n)^+$ here for time!

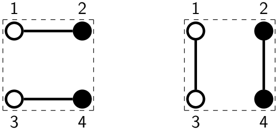
The Unitary Group $U(n)$

Compact matrix quantum group equivariant neural networks for the unitary group $U(n)$ consist of $n^l \times n^k$ weight matrices that are determined by the spanning set consisting of the image of all (w_k, w_l)–partition diagrams whose blocks come in pairs such that

- if two vertices of a block are in the same row, then they have different colours,
- if two vertices of a block are in different rows, then they have the same colours.

Note that if we consider compact matrix *group* equivariant neural networks for $U(n)$, then only $n^k \times n^k$ weight matrices (i.e $k = l$) exist, and they are determined by a spanning set that is the image of all permutations in the symmetric group S_k, expressed as (k, k)–partition diagrams. This is the classic version of Schur–Weyl duality.

Consider the weight matrix in a compact matrix quantum group equivariant neural network for $U(2)$, where the linear layer is $((\mathbb{C}^2)^{\otimes 2}, u^{\otimes w_k}) \to ((\mathbb{C}^2)^{\otimes 2}, u^{\otimes w_l})$ for $w_k = w_l = \circ\bullet$. Then, by the previous slide, the only valid (w_k, w_l)–partition diagrams are

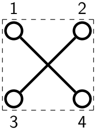


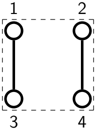
giving matrices

	1,1	1,2	2,1	2,2
1,1	1	0	0	1
1,2	0	0	0	0
2,1	0	0	0	0
2,2	1	0	0	1

	1,1	1,2	2,1	2,2
1,1	1	0	0	0
1,2	0	1	0	0
2,1	0	0	1	0
2,2	0	0	0	1

However, if we consider the weight matrix in a compact matrix **group** equivariant neural network for $U(2)$, where the linear layer is $((\mathbb{C}^2)^{\otimes 2}, u^{\otimes w_k}) \to ((\mathbb{C}^2)^{\otimes 2}, u^{\otimes w_l})$, then $w_k = w_l = \circ\circ$, and the only valid (w_k, w_l)–partition diagrams are





giving matrices

	1,1	1,2	2,1	2,2
1,1	1	0	0	0
1,2	0	0	1	0
2,1	0	1	0	0
2,2	0	0	0	1

	1,1	1,2	2,1	2,2
1,1	1	0	0	0
1,2	0	1	0	0
2,1	0	0	1	0
2,2	0	0	0	1

instead.

The Unitary Quantum Group $U(n)^+$

Unsurprisingly:

Compact matrix quantum group equivariant neural networks for the unitary quantum group $U(n)^+$ consist of $n^l \times n^k$ weight matrices that are determined by the same spanning set as for $U(n)$ but with the image of all crossing (w_k, w_l)–partition diagrams removed.

8. Summary

Conclusion

- Encoding symmetries into machine learning models is a hot topic!
- There exist symmetries that cannot be understood formally as groups.
- We have constructed new machine learning models, called compact matrix quantum group equivariant neural networks, to learn from data that has a quantum symmetry.
- We have used Woronowicz–Tannaka–Krein duality to characterise the weight matrices for a number of quantum groups.

Contact Details

Edward Pearce-Crump

Imperial College London

ep1011@ic.ac.uk

https://www.epearcecrump.co.uk