Canonicalization of Bell inequalities: an overview

Denis Rosset
Email: denis.rosset@unige.ch

Jean-Daniel Bancal
Email: jdbancal.physics@gmail.com

December 22, 2013

Abstract

We give an overview of our canonicalization scheme for Bell inequalities, with sufficient details to define uniquely the canonical representative of any linear inequality with rational coefficients using finite inputs and outputs. Explicit constructions and implementation details are missing.

Canonicalization

Linear Bell inequalities are defined on joint probability distributions $P(ab...|xy...)$ using coefficients $I(ab...|xy...) \in \mathbb{R}$ and a bound $u \in \mathbb{R}$

$$\sum_{ab...xy...} I(ab...|xy...)P(ab...|xy...) \leqslant u. \tag{1}$$

In the canonicalization procedure, three redundancies in such inequalities have to be removed:

- the presence of irrelevant terms in an inequality,
- freedom in the definition of the bound, by applying a shift or multiplication with a factor,
- the transformation of the inequality under relabellings.

The procedure has to constructed such that performing the removal of one of the redundancies does not affect the removal of another redundancy: for example, transforming the inequality using relabellings should not create new irrelevant terms. Thus, the subspace corresponding to irrelevant terms has to be symmetric under relabellings. We give below an overview of the canonicalization procedure, with the minimal amount of detail to specify it in an unique way, while explicit constructions are relagated to the [TODO Appendix].

The three redundancies above can already be described at the single party level. Of course, at the single party level, the only facet of the local polytope is the *positivity*:

$$P(a|x) \geqslant 0, \tag{2}$$

for fixed a and x. We start below by describing the single party case before generalizing it to the multipartite case.

Single party correlations

Let us focus first on single party correlations described by the conditional probability distribution $P(a|x)$, where $x = 1, 2, ..., X$ is the input or measurement setting, and $a = 1, 2, ..., A_x$ represents the output or measurement outcome. The number of outputs for the x-th input is written A_x, and we impose each input to have at least two outcomes: $A_x \geqslant 2$. The tuple $(A_1, ..., A_X)$ is called the *input structure* of the party. When $A_1 = A_2 = ... = A_X$, we say that the input structure is *homogenous*.

Correlations can be transformed by changing the label of inputs or outputs. Such transformations are called *relabellings* and form a finite group H. Relabellings can be composed and form a finite group H, composed of

- relabellings of outputs, corresponding to the subgroup $\mathcal{A} \subseteq H$,
- relabellings of inputs, corresponding to the subgroup $\mathcal{X} \subset H$.

These groups are formally defined [IN APPENDIX]. For our purposes here, we only need to warn the reader that elements of $\mathcal{X}$ are restricted to those that do not change the *input structure* A_x, thus the shape of $P(a|x)$. We also note that the relabelling of outputs can be done differently for every single input.

Correlation vectors and correlation spaces

We define the *correlation vector* $\vec{P} \in \mathcal{P} = \mathbb{R}^d$ corresponding to the probability distribution $P(a|x)$ by enumerating the indices (a, x). As every coefficient $P(a|x)$ is enumerated $d = \sum_x A_x$. Correlations are not fully dimensional in $\mathcal{P}$; indeed, distributions $P(a|x)$ are normalized:

$$\forall x, \qquad \Sigma(x) \equiv \sum_{a=1}^{A_x} P(a|x) = 1. \tag{3}$$

Let $\mathcal{P}_{\text{proper}}$ be the smallest subspace of $\mathcal{P}$ containing all correlation vectors corresponding to probability distributions. The subspace $\mathcal{P}_{\text{proper}}$ is be characterized by $X - 1$ independent equations:

$$\Sigma(x) - \Sigma(x+1) = 0 \qquad \text{for } x = 1, 2, ..., X-1, \tag{4}$$

and has dimension $d_{\text{proper}} = d - X + 1$.

The finite group H has a permutation representation on distributions $P(a|x)$ and thus on correlation vectors $\vec{P}$ (a construction is given in the [TODO Appendix]). In the subspace $\mathcal{P}_{\text{proper}}$, we single out the[1] subspace $\mathcal{P}_{\mathbb{1}}$ invariant under relabellings, corresponding to the span of the uniformly random probability distribution $\vec{P}_{\mathbb{1}}$:

$$P_{\mathbb{1}}(a|x) = \frac{1}{A_x}. \tag{5}$$

Using the subspace structure $\mathcal{P}_{\mathbb{1}} \subset \mathcal{P}_{\text{proper}} \subseteq \mathcal{P}$, we want to identify spaces $\mathcal{P}_{\text{correlations}}$ and $\mathcal{P}_{\text{improper}}$ so that the following decomposition:

$$\mathcal{P} = \mathcal{P}_{\text{proper}} \oplus \mathcal{P}_{\text{improper}} = (\mathcal{P}_{\mathbb{1}} \oplus \mathcal{P}_{\text{correlations}}) \oplus \mathcal{P}_{\text{improper}}, \tag{6}$$

can be used to write any correlation vector $\vec{P}$ in orthogonal components:

$$\vec{P} = \omega_{\mathbb{1}} \vec{P}_{\mathbb{1}} \oplus \vec{P}_{\text{correlations}} \oplus \vec{P}_{\text{improper}}, \qquad \vec{P}_{\text{correlations}} \in \mathcal{P}_{\text{correlations}}, \quad \vec{P}_{\text{improper}} \in \mathcal{P}_{\text{improper}}. \tag{7}$$

Orthogonality requires a well-defined inner product between correlation vectors $\left(\vec{P}_1,\ \vec{P}_2\right)$; but such a definition does not make physical sense. Instead, we have to work with dual space $\mathcal{I} \equiv \mathcal{P}^*$ composed of linear functionals $I : \mathcal{P} \longrightarrow \mathbb{R}$.

1. The maximal subspace invariant under H has dimension 1 only when the input structure is homogenous. For the non-homogenous case, we define a unidimensional $\mathcal{P}_{\mathbb{1}}$ by analogy.

Linear functionals and Bell inequalities

Linear functionals on $\mathcal{P}$, associated with a bound u, define linear inequalities on correlations:

$$I(P) \leqslant u, \tag{8}$$

which are readily identifed with linear Bell inequalities:

$$\sum_{x=1}^{X} \sum_{a=1}^{A_x} I(a|x) P(a|x) \leqslant u. \tag{9}$$

Now, $\mathcal{P}^*$ is isomorphic to $\mathbb{R}^d$, and thus I can be written as a vector $\vec{I} \in \mathbb{R}^d$. The value $I(P)$ is then obtained by the canonical inner product in $\mathbb{R}^d$:

$$I(P) = \left(\vec{I}\,, \vec{P}\right). \tag{10}$$

We warn the reader not to be misled by the fact that $\mathcal{P} = \mathbb{R}^d$ and $\mathcal{I} = \mathbb{R}^d$; as said earlier, the products between two correlation vectors $\left(\vec{P}_1, \vec{P}_2\right)$ or between two linear functionals $\left(\vec{I}_1, \vec{I}_2\right)$ do not make sense.

We single out two subspaces of $\mathcal{I}$:

- the space $\mathcal{I}_{\text{improper}}$ which is the span of the functionals (4) taking value 0 on $\mathcal{P}_{\text{proper}}$ (e.g. the annihilator of $\mathcal{P}_{\text{proper}}$),

- the unidimensional space $\mathcal{I}_{\mathbb{1}}$ symmetric under H, whose basis consists of $\left\{\vec{I}_{\mathbb{1}}\right\}$, the *average normalization* functional:

$$I_{\mathbb{1}}(P) = \frac{1}{X} \sum_{x=1}^{X} \sum_{a=1}^{A_x} P(a|x). \tag{11}$$

With these tools, we want the following decomposition:

$$\mathcal{P} = \mathcal{P}_{\mathbb{1}} \oplus \mathcal{P}_{\text{correlations}} \oplus \mathcal{P}_{\text{improper}}, \qquad \mathcal{I} = \mathcal{I}_{\mathbb{1}} \oplus \mathcal{I}_{\text{correlations}} \oplus \mathcal{I}_{\text{improper}}, \tag{12}$$

with dimension $d_{\mathbb{1}} = 1$, $d_{\text{correlations}} = d - X$, $d_{\text{improper}} = X - 1$, such that elements of different subspaces $s \neq s' \in \{\mathbb{1}, \text{correlations}, \text{improper}\}$ are biorthogonal:

$$\forall \vec{I}_s \in \mathcal{I}_s, \vec{P}_{s'} \in \mathcal{P}_{s'}, \qquad \left(\vec{P}_s, \vec{I}_{s'}\right) = 0. \tag{13}$$

For this decomposition to be unique, we impose the restrictions:

— the subspaces $\mathcal{P}_{\text{improper}}$ and $\mathcal{I}_{\text{correlations}}$ are invariant under H,

— each element $\vec{P}_{\text{improper}} \in \mathcal{P}_{\text{improper}}$ is invariant under relabelling of outputs.

Given this decomposition, any correlation vector can written as $\vec{P} = \omega_{\mathbb{1}}\, \vec{P}_{\mathbb{1}} + \vec{P}_{\text{correlations}} + \vec{P}_{\text{improper}}$, with $\omega_{\mathbb{1}} \in \mathbb{R}$, $\vec{P}_{\text{correlations}} \in \mathcal{P}_{\text{correlations}}$ and $\vec{P}_{\text{improper}} \in \mathcal{P}_{\text{improper}}$. For vectors $\vec{P}$ corresponding to probability distributions $P(a|x)$, $\omega_{\mathbb{1}} = \left(I_{\mathbb{1}}, \vec{P}\right) = 1$ and $\vec{P}_{\text{improper}} = 0$. Thus, probability distributions are fully specified by $\vec{P}_{\text{correlations}}$.

Any linear Bell inequality can be decomposed as $\vec{I} = \nu_1 \vec{I}_{\mathbb{1}} + \vec{I}_{\text{correlations}} + \vec{I}_{\text{improper}}$. The value of $I(P)$ on probability distributions P is:

$$\begin{aligned} I(P) &= \left(\nu_1 \vec{I}_{\mathbb{1}}, \omega_{\mathbb{1}} \vec{P}_{\mathbb{1}}\right) + \left(\vec{I}_{\text{correlations}}, \vec{P}_{\text{correlations}}\right) + \left(\vec{I}_{\text{improper}}, 0\right) \\ &= \nu_1 + \left(\vec{I}_{\text{correlations}}, \vec{P}_{\text{correlations}}\right). \end{aligned}$$

Redundancies in an inequality $I(P) \leqslant u$ can then be removed by transforming:

$$\left(\nu_1 \vec{I}_{\mathbb{1}} + \vec{I}_{\text{correlations}} + \vec{I}_{\text{improper}}, \vec{P}\right) \leqslant u, \tag{14}$$

into

$$\left(\vec{I}_{\text{correlations}}, \vec{P}\right) \leqslant u - \nu_1. \tag{15}$$

Thus, any linear Bell inequality can be described by a vector $\vec{I}_{\text{correlations}} \in \mathcal{I}_{\text{correlations}}$ in a space of dimension $d_{\text{correlations}} = d - X$ and a bound $u' = u - \nu_1 \in \mathbb{R}$.

Bases and notations

Both probability distributions and linear Bell inequalities can be described by the coefficients of their vectors $\vec{P}$ and $\vec{I}$; as they can be restricted to the subspaces $\mathcal{P}_{\text{correlations}}$, $\mathcal{I}_{\text{correlations}}$, two notations have been used in the litterature:

— a notation using *correlators* for binary outputs [TODO: cite Pitowsky, Sliwa, and before?],

— a notation which omits the coefficients for the last output $a = A_x$, as these coefficients can always be recovered using the normalization [TODO: cite Collins-Gisin].

These notations correspond to different bases of $\mathcal{P}_{\text{correlations}}$ and $\mathcal{I}_{\text{correlations}}$, constructed explicitly in the [TODO: Appendix].

Multipartite correlations

To develop the multipartite canonicalization scheme, we consider the bipartite joint probability distribution $P(ab|xy)$. The input structure is described by $A_x \geqslant 2$ and $B_y \geqslant 2$ for $x = 1, ..., X$ and $y = 1, ..., Y$.

Considering the single party marginals $P^{\text{A}}(a|x)$ and $P^{\text{B}}(b|y)$ and the associated vector spaces $\mathcal{P}^{\text{A}}$ and $\mathcal{P}^{\text{B}}$, the correlation vector $\vec{P}$ associated to $P(ab|xy)$ is an element of the tensor space $\mathcal{P} = \mathcal{P}^{\text{A}} \otimes \mathcal{P}^{\text{B}}$. Linear inequalities are associated with elements of the dual space $\mathcal{I} = \mathcal{P}^* = \mathcal{I}^{\text{A}} \otimes \mathcal{I}^{\text{B}}$.

We write below a decomposition of $\mathcal{P}$ and $\mathcal{I}$:

$$\mathcal{P} = \mathcal{P}_{\mathbb{1}} \oplus \mathcal{P}_{\text{correlations}} \oplus \mathcal{P}_{\text{improper}}, \qquad \mathcal{I} = \mathcal{I}_{\mathbb{1}} \oplus \mathcal{I}_{\text{correlations}} \oplus \mathcal{I}_{\text{improper}}, \tag{16}$$

such that $P_{\mathbb{1}}(ab|xy) = \frac{1}{A_x B_y}$, $I_{\mathbb{1}}(ab|xy) = \frac{1}{XYZ}$, and $\mathcal{P}_{\mathbb{1}} \oplus \mathcal{P}_{\text{correlations}}$ is the minimal subspace of $\mathcal{P}$ containing all correlation vectors corresponding to probability distributions $P(ab|xy)$.

Because $\mathcal{P} = \mathcal{P}^{\text{A}} \otimes \mathcal{P}^{\text{B}}$, we have:

$$\mathcal{P} = (\mathcal{P}^{\text{A}}_{\mathbb{1}} \oplus \mathcal{P}^{\text{A}}_{\text{correlations}} \oplus \mathcal{P}^{\text{A}}_{\text{improper}}) \otimes (\mathcal{P}^{\text{B}}_{\mathbb{1}} \oplus \mathcal{P}^{\text{B}}_{\text{correlations}} \oplus \mathcal{P}^{\text{B}}_{\text{improper}}). \tag{17}$$

Note: the reader should think that bases can be constructed for all single party subspaces [ref TODO: Appendix]. A basis for the subspace $\mathcal{P}^A_s \otimes \mathcal{P}^B_{s'}$ is given by taking tensor products of all pairs of basis elements for $\mathcal{P}^A_s$ and $\mathcal{P}^B_{s'}$.

The terms in the expanded tensor product are interpreted as follows:

Combined	Subspace	Notation	Interpretation
$\mathcal{P}_{\mathbb{1}}$	$\mathcal{P}^{\mathrm{A}}_{\mathbb{1}} \otimes \mathcal{P}^{\mathrm{B}}_{\mathbb{1}}$	$\mathcal{P}_{\mathbb{1}}$	distribution with uniformly random outputs
$\mathcal{P}_{\text{correlations}}$	$\mathcal{P}^{\mathrm{A}}_{\text{correlations}} \otimes \mathcal{P}^{\mathrm{B}}_{\mathbb{1}}$	$\mathcal{P}_{\text{correlations:A}}$	correlations of Alice marginal distribution
	$\mathcal{P}^{\mathrm{A}}_{\mathbb{1}} \otimes \mathcal{P}^{\mathrm{B}}_{\text{correlations}}$	$\mathcal{P}_{\text{correlations:B}}$	correlations of Bob marginal distribution
	$\mathcal{P}^{\mathrm{A}}_{\text{correlations}} \otimes \mathcal{P}^{\mathrm{B}}_{\text{correlations}}$	$\mathcal{P}_{\text{correlations:AB}}$	fully bipartite correlations
$\mathcal{P}_{\text{signaling}}$	$\mathcal{P}^{\mathrm{A}}_{\text{improper}} \otimes \mathcal{P}^{\mathrm{B}}_{\text{correlations}}$	$\mathcal{P}_{\mathrm{A}\to\mathrm{B}}$	Alice→Bob signaling correlations
	$\mathcal{P}^{\mathrm{A}}_{\text{correlations}} \otimes \mathcal{P}^{\mathrm{B}}_{\text{improper}}$	$\mathcal{P}_{\mathrm{B}\to\mathrm{A}}$	Bob→Alice signaling correlations
$\mathcal{P}_{\text{norm}}$	$\mathcal{P}^{\mathrm{A}}_{\text{improper}} \otimes \mathcal{P}^{\mathrm{B}}_{\mathbb{1}}$	$\mathcal{P}_{\text{norm:A}}$	improperly normalized marginals for Alice
	$\mathcal{P}^{\mathrm{A}}_{\mathbb{1}} \otimes \mathcal{P}^{\mathrm{B}}_{\text{improper}}$	$\mathcal{P}_{\text{norm:B}}$	improperly normalized marginals for Bob
	$\mathcal{P}^{\mathrm{A}}_{\text{improper}} \otimes \mathcal{P}^{\mathrm{B}}_{\text{improper}}$	$\mathcal{P}_{\text{norm:AB}}$	improperly normalized distributions with properly normalized marginals

Table 1. Decomposition of the bipartite space $\mathcal{P} = \mathcal{P}^{\mathrm{A}} \otimes \mathcal{P}^{\mathrm{B}}$ into $9 = 1 + 3 + 2 + 3$ subspaces. See [TODO Appendix] for the motivation of the interpretation.

We identify $\mathcal{P}_{\text{improper}} = \mathcal{P}_{\text{signaling}} \oplus \mathcal{P}_{\text{norm}}$, as correlations obtained by performing space-like separated measurements are non-signaling. The same decomposition is done on $\mathcal{I}$, and thus the procedure of Eqs. (14) and (15) can be applied to any bipartite linear Bell inequality.

For general multipartite scenarios, a similar decomposition is performed, with the following interpretation:

— the tensor product $\mathcal{P}^{\mathrm{A}}_{\mathbb{1}} \otimes \mathcal{P}^{\mathrm{B}}_{\mathbb{1}} \otimes \mathcal{P}^{\mathrm{C}}_{\mathbb{1}} \otimes \ldots$ is identified with $\mathcal{P}_{\mathbb{1}}$,

— products containing at least one $\mathcal{P}^{\mathrm{A,B,C,\ldots}}_{\text{correlations}}$ and possibly some $\mathcal{P}^{\mathrm{A,B,C,\ldots}}_{\mathbb{1}}$ are collected in $\mathcal{P}_{\text{correlations}}$,

— products containing at least one $\mathcal{P}^{\mathrm{A,B,C,\ldots}}_{\text{improper}}$ and possibly some $\mathcal{P}^{\mathrm{A,B,C,\ldots}}_{\mathbb{1}}$ are collected in $\mathcal{P}_{\text{norm}}$,

— products containing at least one $\mathcal{P}^{\mathrm{A,B,C,\ldots}}_{\text{correlations}}$ and at least one $\mathcal{P}^{\mathrm{A,B,C,\ldots}}_{\text{improper}}$ are collected in $\mathcal{P}_{\text{signaling}}$ (the indices of the correlations and improper terms will indicate the direction of signaling).

Canonicalization of Bell inequalities

With this construction, we perform the canonicalization of an inequality using these steps:

– we transform any lower bound to an upper bound:

$$\left(\vec{I}\,,\vec{P}\right) \geqslant l \qquad \longrightarrow \qquad \left(-\vec{I}\,,\vec{P}\right) \leqslant -l, \tag{18}$$

– we decompose the inequality, remove the irrelevant part and shift the bound:

$$\left(\nu_1 \vec{I}_{\mathbb{1}} + \vec{I}_{\text{correlations}} + \vec{I}_{\text{improper}}, \vec{P}\right) \leqslant u \qquad \longrightarrow \qquad \left(\vec{I}_{\text{correlations}}, \vec{P}\right) \leqslant u - \nu_1, \tag{19}$$

– if $\vec{I}$ has rational coefficients, we multiply $\vec{I}$ by a factor such that the resulting vector has integer coefficients with greatest common divisor 1,

– we enumerate all representatives of $\vec{I}$ under relabellings and pick the minimal representative under lexicographic ordering.

This procedure removes the three redundancies explained above. The only arbitrary part in the definition of this procedure is the selection of the minimal lexicographic representative, which we make explicit below.

Let $f\colon k \mapsto (a, x, b, y, ...)$ be an enumeration of the indices $(a, x, b, y, ...)$, for $k = 1, ..., d$. The lexicographic ordering is such that:

$$\vec{I} <_{\text{lex}} \vec{J} \quad \Leftrightarrow \quad \exists \ell \text{ s.t. } I_{f(k)} = J_{f(k)} \text{ for } k < \ell \text{ and } I_{f(\ell)} < J_{f(\ell)}. \tag{20}$$

We now define our enumeration f. Let $f^{\mathrm{A}}\colon k^{\mathrm{A}} \mapsto (a, x)$, $f^{\mathrm{B}}\colon k^{\mathrm{B}} \mapsto (b, y)$, ... be enumerations for the single party marginals $\vec{P}^{\,\mathrm{A}} \in \mathbb{R}^{d^{\mathrm{A}}}$, $\vec{P}^{\,\mathrm{B}} \in \mathbb{R}^{d^{\mathrm{B}}}$, ... To define f, we decompose the index k as:

$$k = k^{\mathrm{A}} + d^{\mathrm{A}}(k^{\mathrm{B}} - 1) + d^{\mathrm{A}} d^{\mathrm{B}}(k^{\mathrm{C}} - 1) + ..., \tag{21}$$

and attribute $(a, x) = f^{\mathrm{A}}(k^{\mathrm{A}})$, $(b, y) = f^{\mathrm{B}}(k^{\mathrm{B}})$.

We define now the single party enumeration f^{A}, f^{B}, ... such that (here for f^{A}):

- for each input x, outputs $a = 1, 2, ..., A_x$ are contiguous and enumerated in increasing order,
- inputs are then enumerated in increasing order.

Appendix: relabellings (incomplete)

Relabellings

Two types of relabellings can be applied to single party marginals $P(a|x)$:

- relabelling of outputs h_{out}, described by index $\xi = 1, ..., X$ and a permutation π, such that:

$$a' = a^{h_{\text{out}}} = \begin{cases} a^{\pi} & \text{if} \quad x = \xi, \\ a & \text{otherwise.} \end{cases} \tag{22}$$

- relabelling of inputs h_{in}, described by a permutation σ preserving the structure of outcomes $(A_{\sigma(x)} = A_x)$, such that $x' = x^{h_{\text{in}}} = x^{\sigma}$.

The action of a relabelling h on $P(a|x)$ is such that $P' = P^h$ has coefficients $P'(a'|x') = P(a|x)$:

$$P^{h_{\text{out}}}(a|x) = \begin{cases} P\big(a^{\pi^{-1}}\big|x\big) & \text{if } x = \xi, \\ P(a|x) & \text{otherwise,} \end{cases} \qquad P^{h_{\text{in}}}(a|x) = P\big(a\big|x^{\sigma^{-1}}\big). \tag{23}$$

We write id the identity relabelling which leaves P invariant: $P^{\text{id}} = P$.

Group of output relabellings

We write $\mathcal{A}_x$ be the group of all output relabellings for the x-th input, with identity id and the composition as product. We define the action of the product $\alpha = \alpha_1 \alpha_2$ on P by the *right action*: $P^{\alpha} = (P^{\alpha_1})^{\alpha_2}$. Using the same exponential notation, we write the action of a permutation π on a as a^{π} with $a^{\pi_1 \pi_2} = (a^{\pi_1})^{\pi_2}$.

Let us consider the group of relabellings of outputs for a fixed input ξ. This group can be identified with the symmetric group acting on A_ξ elements S_{A_ξ}.

Take $\alpha_1, \alpha_2 \in \mathcal{A}_\xi$, with corresponding permutations $\pi_1, \pi_2 \in S_{A_\xi}$. Then:

$$(P^{\alpha_1})^{\alpha_2}(a|\xi) = P^{\alpha_1}\Big(a^{\pi_2^{-1}}\Big|\xi\Big) = P\Big(a^{\pi_2^{-1}\pi_1^{-1}}\Big|\xi\Big) = P\big(a^{\pi^{-1}}\big|\xi\big) = P^{\alpha}(a|x) \tag{24}$$

and α has corresponding permutation $\pi = (\pi_2^{-1}\pi_1^{-1}) = \pi_1\pi_2$.

Let us now consider the group $\mathcal{A}$ of all output relabellings, composed of products of elements of its subgroups $\{\mathcal{A}_x\}_{x=1}^X$. When $\alpha_1 \in \mathcal{A}_\xi$ and $\alpha_2 \in \mathcal{A}_{\xi'}$ for $\xi = \xi'$, they act on different parts of $P(a|x)$ and thus commute: $\alpha_1\alpha_2 = \alpha_2\alpha_1$. Then, any element of $\alpha \in \mathcal{A}$ can be decomposed as:

$$\alpha = \alpha^1\alpha^2...\alpha^X, \qquad \alpha^x \in \mathcal{A}_x, \tag{25}$$

and $\mathcal{A}$ is the direct product $\mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2 \times ... \times \mathcal{A}_X$.

Group of input relabellings

We write $\mathcal{X}$ for the group of input relabellings. We identify the elements of $\mathcal{X}$ with permutations σ acting on X elements, and identify $\mathcal{X}$ with the subgroup of S_X that leaves the number of outputs A_x invariant. Let us verify that the product of $\sigma = \sigma_1\sigma_2$ is the usual product of permutations:

$$(P^{\sigma_1})^{\sigma_2}(a|x) = P^{\sigma_1}\Big(a\Big|x^{\sigma_2^{-1}}\Big) = P\Big(a\Big|x^{\sigma_2^{-1}\sigma_1^{-1}}\Big) = P\big(a\big|x^{\sigma^{-1}}\big) = P^{\sigma}(a|x). \tag{26}$$

Group of relabellings

Let H be the group of input and outputs relabellings. We have to see if a product $h = \sigma\,\alpha = \sigma\,\alpha^1...\alpha^X$ can be reordered into $h = \beta\,\tau = \beta^1...\beta^X\tau$. We have:

$$(P^{\sigma})^{\alpha}(a|x) = P^{\sigma}\big(a^{(\alpha^x)^{-1}}\big|x\big) = P\Big(a^{\left(\alpha^{x^{\sigma^{-1}}}\right)^{-1}}\Big|x^{\sigma^{-1}}\Big), \tag{27}$$

and

$$(P^{\beta})^{\tau}(a|x) = P^{\beta}\big(a\big|x^{\tau^{-1}}\big) = P\big(a^{(\beta^x)^{-1}}\big|x^{\tau^{-1}}\big), \tag{28}$$

and thus we identify $\tau = \sigma$ and $\beta^x = \alpha^{\left(x^{\sigma^{-1}}\right)}$. Thus, any element $h \in H$ can be written as a product $h = \alpha\,\sigma$ with $\alpha \in \mathcal{A}$ and $\sigma \in \mathcal{X}$.

When the output structure is homogenous ($A_x = n$), the group H is identified with the wreath product $S_n \wr S_X$.