SHIFTED YANGIANS AND FINITE W-ALGEBRAS

JONATHAN BRUNDAN AND ALEXANDER KLESHCHEV

Abstract. We give a presentation for the finite W-algebra associated to a nilpotent matrix inside the general linear Lie algebra over $\mathbb{C}$. In the special case that the nilpotent matrix consists of n Jordan blocks each of the same size l, the presentation is that of the *Yangian of level l* associated to the Lie algebra $\mathfrak{gl}_n$, as was first observed by Ragoucy and Sorba. In the general case, we are lead to introduce some generalizations of the Yangian which we call the *shifted Yangians*.

1. Introduction

Let $\mathfrak{g}$ be a finite dimensional reductive Lie algebra over $\mathbb{C}$ equipped with a non-degenerate invariant symmetric bilinear form $(.,.)$. Pick a nilpotent element $e \in \mathfrak{g}$, i.e. an element which acts nilpotently on every finite dimensional $\mathfrak{g}$-module. A $\mathbb{Z}$-grading $\mathfrak{g} = \bigoplus_{j\in\mathbb{Z}} \mathfrak{g}_j$ of $\mathfrak{g}$ is called a *good grading* for e if $e \in \mathfrak{g}_2$ and the linear map

$$\operatorname{ad} e : \mathfrak{g}_j \to \mathfrak{g}_{j+2}$$

is injective for $j \leq -1$, surjective for $j \geq -1$. This definition originates in [KRW] in the study of certain W-algebras defined from affine Lie algebras by quantum Hamiltonian reduction. A complete classification of all good gradings of simple Lie algebras up to conjugacy can be found in [EK].

Since $\operatorname{ad} e : \mathfrak{g}_{-1} \to \mathfrak{g}_1$ is bijective, the skew-symmetric bilinear form $\langle .,. \rangle$ on $\mathfrak{g}_{-1}$ defined by $\langle x, y\rangle := ([x,y], e)$ is non-degenerate. Pick a Lagrangian subspace $\mathfrak{l}$ of $\mathfrak{g}_{-1}$ with respect to the form $\langle .,.\rangle$ and set $\mathfrak{m} := \mathfrak{l} \oplus \bigoplus_{j \leq -2} \mathfrak{g}_j$. This is a nilpotent subalgebra of $\mathfrak{g}$, and the map $\chi : \mathfrak{m} \to \mathbb{C}, x \mapsto (x,e)$ defines a representation of $\mathfrak{m}$. Let I_χ denote the kernel of the corresponding associative algebra homomorphism $U(\mathfrak{m}) \to \mathbb{C}$, where $U(\mathfrak{m})$ denotes the universal enveloping algebra of $\mathfrak{m}$. Let

$$Q_\chi := U(\mathfrak{g}) \otimes_{U(\mathfrak{m})} \mathbb{C}_\chi \cong U(\mathfrak{g})/U(\mathfrak{g})I_\chi$$

denote the induced $\mathfrak{g}$-module, and consider the endomorphism algebra

$$W(\chi) := \operatorname{End}_{U(\mathfrak{g})}(Q_\chi)^{\mathrm{op}}.$$

Following terminology used in the mathematical physics literature, we refer to these algebras as *finite W-algebras*; see e.g. [BT]. Applying Frobenius reciprocity, it is often more convenient to view $W(\chi)$ instead as the subspace of $U(\mathfrak{g})/U(\mathfrak{g})I_\chi$ consisting of all cosets $y + U(\mathfrak{g})I_\chi$ such that $[x,y] \in U(\mathfrak{g})I_\chi$ for all $x \in \mathfrak{m}$. In this realization, the algebra structure on $W(\chi)$ is defined by the formula $(y + U(\mathfrak{g})I_\chi)(y' + U(\mathfrak{g})I_\chi) = yy' + U(\mathfrak{g})I_\chi$ for $y, y' \in U(\mathfrak{g})$ such that $[x,y], [x,y'] \in U(\mathfrak{g})I_\chi$ for all $x \in \mathfrak{m}$.

In the special case that our fixed good grading is *even*, i.e. $\mathfrak{g}_j = 0$ for all odd j, the algebras $W(\chi)$ were already well studied by the end of the 1970s by Lynch [L],

Research partially supported by the NSF (grant no. DMS-0139019).
2000 Subject Classification: 17B37.

generalizing work of Kostant [K] treating regular nilpotent elements. Of course in the even case, we have simply that $\mathfrak{m} = \bigoplus_{j \leq -2} \mathfrak{g}_j$. Letting $\mathfrak{p} := \bigoplus_{j \geq 0} \mathfrak{g}_j$, the PBW theorem implies that

$$U(\mathfrak{g}) = U(\mathfrak{p}) \oplus U(\mathfrak{g}) I_\chi.$$

The projection $\mathrm{pr}_\chi : U(\mathfrak{g}) \to U(\mathfrak{p})$ along this direct sum decomposition induces an isomorphism $U(\mathfrak{g})/U(\mathfrak{g})I_\chi \xrightarrow{\sim} U(\mathfrak{p})$ which leads to an easier definition of the algebra $W(\chi)$ in the even case as a *subalgebra* of $U(\mathfrak{p})$. To make this precise, define a twisted action of $\mathfrak{m}$ on $U(\mathfrak{p})$ by $x \cdot y := \mathrm{pr}_\chi([x, y])$ for $x \in \mathfrak{m}$ and $y \in U(\mathfrak{p})$. Then pr_χ induces an isomorphism between $W(\chi)$ and subalgebra $U(\mathfrak{p})^{\mathfrak{m}}$ of $U(\mathfrak{p})$ consisting of all twisted $\mathfrak{m}$-invariants. This is the original definition used by Kostant and Lynch.

The most important examples of good gradings arise as follows. By the Jacobson-Morozov theorem, we can embed e into an $\mathfrak{sl}_2$-triple (e, h, f), so $[e, f] = h, [h, e] = 2e$ and $[h, f] = -2f$. Then the representation theory of $\mathfrak{sl}_2$ implies that the $\mathrm{ad}\, h$-eigenspace decomposition of $\mathfrak{g}$ is a good grading for e. We refer to a good grading obtained in this way as a *Dynkin grading*. For a Dynkin grading, the module Q_χ is a *generalized Gelfand-Graev representation* in the sense of Kawanaka [Ka] and Moeglin [M]. Its endomorphism algebra $W(\chi)$ has been studied by Premet [P] as an application of results on Lie algebras in positive characteristic. Subsequently, a more direct approach has been given by Gan and Ginzburg [GG] which we follow here.

Returning to an arbitrary good grading, [EK, Lemma 1.1] shows that we can always embed the given element $e \in \mathfrak{g}_2$ into an $\mathfrak{sl}_2$-triple (e, h, f) with $h \in \mathfrak{g}_0$ and $f \in \mathfrak{g}_{-2}$. Letting $\mathfrak{c}_\mathfrak{g}(f)$ denote the centralizer of f and $\mathfrak{m}^\perp := \{y \in \mathfrak{g} \mid (x, y) = 0 \text{ for all } x \in \mathfrak{m}\}$, the following crucial formula is proved as in [GG, §2.3], using [EK, Theorem 1.4]:

$$\mathfrak{m}^\perp = [\mathfrak{m}, e] \oplus \mathfrak{c}_\mathfrak{g}(f).$$

Remarkably, given this formula, all the arguments from [GG] in the context of Dynkin gradings extend absolutely unchanged to arbitrary good gradings. Let us just state briefly here the analogue of [P, Proposition 6.3] which we regard as the fundamental structure theorem for $W(\chi)$; see also [GG, Theorem 4.1] and [L, Theorem 2.3]. Introduce the *Kazhdan filtration* $\cdots \subseteq \mathrm{F}_d U(\mathfrak{g}) \subseteq \mathrm{F}_{d+1} U(\mathfrak{g}) \subseteq \cdots$ of $U(\mathfrak{g})$ by declaring that a generator $x \in \mathfrak{g}_j$ is of degree $(j+2)$, i.e. $\mathrm{F}_d U(\mathfrak{g})$ is the span of all monomials $x_1 \cdots x_m$ for $m \geq 0$ and $x_1 \in \mathfrak{g}_{j_1}, \dots, x_m \in \mathfrak{g}_{j_m}$ with $(j_1+2) + \cdots + (j_m+2) \leq d$. Viewing $W(\chi)$ as a subspace of the quotient $U(\mathfrak{g})/U(\mathfrak{g})I_\chi$, there is an induced Kazhdan filtration on $W(\chi)$; we denote the associated graded Poisson algebra by $\mathrm{gr}\, W(\chi)$. Recall also that the *Slodowy slice* through the nilpotent orbit containing e is the affine subspace $e + \mathfrak{c}_\mathfrak{g}(f)$; see [S]. It has a natural Poisson structure which may be defined following [GG, §3.2] as the Hamiltonian reduction of the Kirillov-Kostant Poisson structure on $\mathfrak{g}$. Now the basic fact is that there is a canonical isomorphism

$$\nu : \mathrm{gr}\, W(\chi) \xrightarrow{\sim} \mathbb{C}[e + \mathfrak{c}_\mathfrak{g}(f)]$$

of Poisson algebras; see [GG, §4.4] for its precise definition. Hence, $W(\chi)$ can be viewed as a quantization of the Slodowy slice $e + \mathfrak{c}_\mathfrak{g}(f)$. Moreover, up to canonical isomorphism, the algebra $W(\chi)$ is independent of the particular choice of the Lagrangian subspace $\mathfrak{l}$ of $\mathfrak{g}_{-1}$; see [GG, §5.5].

Another fundamental result in this subject is *Skryabin's theorem* proved in [Sk] for Dynkin gradings; see also [GG, Theorem 6.1]. Again, Skryabin's proof extends unchanged to any good grading. To state the result, let $\mathcal{C}(\chi)$ be the category of all

$\mathfrak{g}$-modules on which $(x - \chi(x))$ act locally nilpotently for all $x \in \mathfrak{m}$ ("generalized Whittaker modules"). If $M \in \mathcal{C}(\chi)$ then the subspace

$$M^{\mathfrak{m}} := \{v \in M \mid (x - \chi(x))v = 0 \text{ for all } x \in \mathfrak{m}\} \cong \mathrm{Hom}_{U(\mathfrak{g})}(Q_\chi, M)$$

is a $W(\chi)$-module, hence $M \mapsto M^{\mathfrak{m}}$ is a functor F from $\mathcal{C}(\chi)$ to the category $W(\chi)$-Mod of all left $W(\chi)$-modules. Also, we have the functor $G := Q_\chi \otimes_{W(\chi)} ?$ from $W(\chi)$-Mod to $\mathcal{C}(\chi)$. Skryabin's theorem asserts that the functors F and G are quasi-inverse equivalences between $\mathcal{C}(\chi)$ and $W(\chi)$-Mod. Moreover, every $M \in \mathcal{C}(\chi)$ is an injective $\mathfrak{m}$-module and Q_χ is a free right $W(\chi)$-module. Somewhat weaker results in the even case can already be found in the work of Kostant and Lynch; see for example [L, Theorems 2.4, 4.1].

In the remainder of the article we study the algebras $W(\chi)$ in the special case that e is a nilpotent matrix of Jordan type $p_1 \leq \cdots \leq p_n$ inside the Lie algebra $\mathfrak{g} = \mathfrak{gl}_N$ over $\mathbb{C}$, taking the bilinear form $(.,.)$ to be the usual trace form. Our main result (Theorem 10.1) gives an explicit set of generators and relations for the algebra $W(\chi)$. One surprising consequence of our presentation is that in fact up to isomorphism the algebras $W(\chi)$ only depend on the conjugacy class of e, i.e. the partition $(p_1, \dots, p_n)$ of N, *not* on the particular choice of the good grading for e.

The classification of good gradings for e is described in [EK, Theorem 4.2] in terms of certain diagrams called *pyramids*. A consequence of this classification is that, in type A, it is sufficient in order to define *all* the algebras $W(\chi)$ to restrict attention just to even good gradings. More precisely, every good grading for e is *split* in the sense that it is always possible to adjust the grading to obtain a new, even good grading for e with the property that the subalgebra $\mathfrak{m}$ defined from the new grading coincides with the subalgebra $\mathfrak{m}$ defined from the original grading for some particular choice of Lagrangian subspace $\mathfrak{l}$. Hence the algebra $W(\chi)$ defined from the new grading is *equal* to the algebra $W(\chi)$ defined from the original grading using this choice of $\mathfrak{l}$. (In the language of [EK, §4], the pyramid defining such a new grading may be obtained from the pyramid defining the original grading by shifting one place to the left all the boxes whose first coordinates are of different parity to the first coordinates of the boxes on the bottom row.)

So we may assume without loss of generality that the given good grading is *even*. The even good gradings for e are classified up to conjugacy in [EK, Proposition 4.3]; see also [L, Lemma 7.2]. Again we visualize the classification in terms of some pyramids as explained in §7 below. We will explain the idea in this introduction just with one example: the diagram

$\pi =$
<table>
<tr><td></td><td></td><td>5</td><td></td></tr>
<tr><td>1</td><td>3</td><td>6</td><td></td></tr>
<tr><td>2</td><td>4</td><td>7</td><td>8</td></tr>
</table>

is a pyramid for $\mathfrak{g} = \mathfrak{gl}_8$ of height $n = 3$. Numbering the bricks $1, \dots, 8$ as indicated, the rows $1, \dots, 3$ from top to bottom and the columns $1, \dots, 4$ from left to right, we write $\mathrm{row}(i)$ and $\mathrm{col}(i)$ for the row and column numbers of the ith brick in the pyramid, respectively. Denoting the ij-matrix unit by $e_{i,j}$, the nilpotent matrix e associated to π is the matrix

$$e = e_{1,3} + e_{3,6} + e_{2,4} + e_{4,7} + e_{7,8}$$

defined by reading along the rows of the pyramid, and the even good grading associated to π is defined by declaring that $e_{i,j}$ is of degree $(\mathrm{col}(j) - \mathrm{col}(i))$; actually, according to this definition, e is of degree 1 not 2 since in the case of an even good grading we prefer from now on to divide all degrees by 2. Moreover, the Jordan type $(p_1, \dots, p_n)$ of e in this example is $(1, 3, 4)$ ("row lengths") and the parabolic subalgebra $\mathfrak{p} := \bigoplus_{j \geq 0} \mathfrak{g}_j$ is of standard Levi shape $(2, 2, 3, 1)$ ("column heights"). ¿From the pyramid π we also read off the *level* $l := p_n$ and a certain *shift matrix* $\sigma = (s_{i,j})_{1 \leq i,j \leq n}$ as explained in §7; in our example,

$$l = 4, \qquad \sigma = \left(\begin{array}{ccc} 0 & 0 & 1 \\ 2 & 0 & 1 \\ 2 & 0 & 0 \end{array} \right).$$

From now on we will denote the finite W-algebra $W(\chi)$ instead by $W(\pi)$; recall it may be defined in the even case as the subalgebra $U(\mathfrak{p})^{\mathfrak{m}}$ of all twisted $\mathfrak{m}$-invariants in $U(\mathfrak{p})$. Our main theorem (Theorem 10.1) asserts for any pyramid π that $W(\pi)$ is isomorphic to the *shifted Yangian* $Y_{n,l}(\sigma)$ *of level* l, namely, the quotient of the *shifted Yangian* $Y_n(\sigma)$ by the two-sided ideal generated by elements $\{D_1^{(r)}\}_{r > p_1}$. Here, $Y_n(\sigma)$ denotes the algebra defined by generators $\{D_i^{(r)}\}_{1 \leq i \leq n, r > 0}, \{E_i^{(r)}\}_{1 \leq i < n, r > s_{i,i+1}}$ and $\{F_i^{(r)}\}_{1 \leq i < n, r > s_{i+1,i}}$ subject to the relations (2.4)–(2.15) recorded below. In the special case that $p_1 = \cdots = p_n$, i.e. all the Jordan blocks of e are of the same size l, the pyramid π is an $n \times l$ rectangle and the shift matrix σ is the zero matrix, our presentation is a variation on Drinfeld's presentation [D] for the *Yangian* $Y_{n,l}$ *of level* l considered by Cherednik [C]. Hence in this case, $W(\chi)$ is a quotient of the Yangian Y_n associated to the Lie algebra $\mathfrak{gl}_n$, as was first noticed by Ragoucy and Sorba [RS].

The remainder of the article is organized as follows. In §2, we define the shifted Yangian $Y_n(\sigma)$ and prove a PBW theorem for it. In §3, we introduce some more elaborate parabolic presentations for $Y_n(\sigma)$ following [BK1]. These are important because they allow us in §4 to write down an explicit formula for the so-called baby comultiplications. In §5, we introduce the canonical filtration of $Y_n(\sigma)$, which eventually turns out to correspond to the Kazhdan filtration of $W(\chi)$. In §6, we prove a PBW theorem for the finitely generated quotients $Y_{n,l}(\sigma)$ of $Y_n(\sigma)$. Then we turn our attention back to the finite W-algebras $W(\chi)$, beginning in §7 by explaining the classification of even good gradings in terms of pyramids. In §8, we recall the setup of [GG] in our special case in some detail. The most important section of the article is §9, where we write down explicit formulae for elements of $U(\mathfrak{p})$ which eventually turn out to be precisely the generators $D_i^{(r)}$, $E_i^{(r)}$ and $F_i^{(r)}$ of $W(\chi)$ that we are after. The main theorem is then proved by induction in §10, the key tool for the induction step being the baby comultiplications. In §11 we discuss more general comultiplications. Finally, §12 gives a much simpler and more direct proof of the main theorem in the special case $p_1 = \cdots = p_n$, using a different description of the generators of $W(\chi)$ in this special case which is closely related to the Capelli determinant.

We finally note that the exposition in the rest of the article is pretty much self-contained, apart from appealing to the results of [BK1] and the opening lemma of [GG]. In a subsequent article [BK2], we will use the presentation for $W(\chi)$ obtained here to study its highest weight representation theory.

Acknowledgements. We are grateful to Alexander Premet for illuminating discussions with the second author which triggered this work, Pavel Etingof for a useful discussion of [RS], and Gerhard Roehrle for showing us the paper [EK]. The first author also thanks Jeremy Rickard, Joseph Chuang and the Department of Mathematics at the University of Bristol for their hospitality whilst this paper was written up.

2. The shifted Yangian

Fix $n \geq 1$ and a matrix $\sigma = (s_{i,j})_{1 \leq i,j \leq n}$ of non-negative integers ("shifts") such that

$$s_{i,j} + s_{j,k} = s_{i,k} \tag{2.1}$$

whenever $|i-j|+|j-k| = |i-k|$. Note this means that $s_{1,1} = \cdots = s_{n,n} = 0$, and the matrix σ is completely determined by the upper diagonal entries $s_{1,2}, s_{2,3}, \dots, s_{n-1,n}$ and the lower diagonal entries $s_{2,1}, s_{3,2}, \dots, s_{n,n-1}$. We also associate to σ the tuple $\delta = (d_1, \dots, d_{n-1})$ of non-negative integers ("differences") where $d_i := s_{i,i+1} + s_{i+1,i}$. Note that $d_1 + \cdots + d_{n-1} = s_{1,n} + s_{n,1}$ (the "total difference").

The *shifted Yangian* associated to the matrix σ is the algebra $Y_n(\sigma)$ over $\mathbb{C}$ defined by generators

$$\{D_i^{(r)}\}_{1 \leq i \leq n, r > 0},$$
$$\{E_i^{(r)}\}_{1 \leq i < n, r > s_{i,i+1}},$$
$$\{F_i^{(r)}\}_{1 \leq i < n, r > s_{i+1,i}}$$

subject to certain relations. In order to write down these relations, let

$$D_i(u) := \sum_{r \geq 0} D_i^{(r)} u^{-r} \in Y_n(\sigma)[[u^{-1}]] \tag{2.2}$$

where $D_i^{(0)} := 1$, and then define some new elements $\widetilde{D}_i^{(r)}$ of $Y_n(\sigma)$ from the equation

$$\widetilde{D}_i(u) = \sum_{r \geq 0} \widetilde{D}_i^{(r)} u^{-r} := -D_i(u)^{-1}. \tag{2.3}$$

For example, $\widetilde{D}_i^{(0)} = -1, \widetilde{D}_i^{(1)} = D_i^{(1)}, \widetilde{D}_i^{(2)} = D_i^{(2)} - D_i^{(1)} D_i^{(1)}, \dots$. With this notation, the relations are as follows.

$$[D_i^{(r)}, D_j^{(s)}] = 0, \tag{2.4}$$

$$[E_i^{(r)}, F_j^{(s)}] = \delta_{i,j} \sum_{t=0}^{r+s-1} \widetilde{D}_i^{(t)} D_{i+1}^{(r+s-1-t)}, \tag{2.5}$$

$$[D_i^{(r)}, E_j^{(s)}] = (\delta_{i,j} - \delta_{i,j+1}) \sum_{t=0}^{r-1} D_i^{(t)} E_j^{(r+s-1-t)}, \tag{2.6}$$

$$[D_i^{(r)}, F_j^{(s)}] = (\delta_{i,j+1} - \delta_{i,j}) \sum_{t=0}^{r-1} F_j^{(r+s-1-t)} D_i^{(t)}, \tag{2.7}$$

$$[E_i^{(r)}, E_i^{(s)}] = \sum_{t=s_{i,i+1}+1}^{s-1} E_i^{(t)} E_i^{(r+s-1-t)} - \sum_{t=s_{i,i+1}+1}^{r-1} E_i^{(t)} E_i^{(r+s-1-t)}, \tag{2.8}$$

$$[F_i^{(r)}, F_i^{(s)}] = \sum_{t=s_{i+1,i}+1}^{r-1} F_i^{(r+s-1-t)} F_i^{(t)} - \sum_{t=s_{i+1,i}+1}^{s-1} F_i^{(r+s-1-t)} F_i^{(t)}, \tag{2.9}$$

$$[E_i^{(r)}, E_{i+1}^{(s+1)}] - [E_i^{(r+1)}, E_{i+1}^{(s)}] = -E_i^{(r)} E_{i+1}^{(s)}, \tag{2.10}$$

$$[F_i^{(r+1)}, F_{i+1}^{(s)}] - [F_i^{(r)}, F_{i+1}^{(s+1)}] = -F_{i+1}^{(s)} F_i^{(r)}, \tag{2.11}$$

$$[E_i^{(r)}, E_j^{(s)}] = 0 \qquad \text{if } |i-j| > 1, \tag{2.12}$$

$$[F_i^{(r)}, F_j^{(s)}] = 0 \qquad \text{if } |i-j| > 1, \tag{2.13}$$

$$[E_i^{(r)}, [E_i^{(s)}, E_j^{(t)}]] + [E_i^{(s)}, [E_i^{(r)}, E_j^{(t)}]] = 0 \qquad \text{if } |i-j| = 1, \tag{2.14}$$

$$[F_i^{(r)}, [F_i^{(s)}, F_j^{(t)}]] + [F_i^{(s)}, [F_i^{(r)}, F_j^{(t)}]] = 0 \qquad \text{if } |i-j| = 1, \tag{2.15}$$

for all admissible r, s, t, i, j. (For an example of what we mean by "admissible" here, the relation (2.10) should be understood to hold for all $i = 1, \dots, n-2$, $r > s_{i,i+1}$ and $s > s_{i+1,i+2}$.)

If the matrix σ is the zero matrix, we denote $Y_n(\sigma)$ simply by Y_n. In this special case, the above presentation is a variation on Drinfeld's presentation [D] for the usual Yangian $Y(\mathfrak{gl}_n)$ associated to the Lie algebra $\mathfrak{gl}_n$; see [BK1, Remark 5.12]. We will prove in Corollary 2.2 below that the map sending the generators of $Y_n(\sigma)$ to the elements with the same name in Y_n is an injective algebra homomorphism. Given this fact, the algebra $Y_n(\sigma)$ is canonically a *subalgebra* of the usual Yangian Y_n. By the relations, there is an anti-automorphism $\tau : Y_n \to Y_n$ of order 2 defined by

$$\tau(D_i^{(r)}) = D_i^{(r)}, \quad \tau(E_i^{(r)}) = F_i^{(r)}, \quad \tau(F_i^{(r)}) = E_i^{(r)}. \tag{2.16}$$

This obviously interchanges the two subalgebras $Y_n(\sigma)$ and $Y_n(\sigma^t)$ of Y_n, where σ^t denotes the transpose of the matrix σ. Hence τ also induces an anti-isomorphism $\tau : Y_n(\sigma) \to Y_n(\sigma^t)$.

Suppose next that $\dot\sigma = (\dot s_{i,j})_{1 \le i,j \le n}$ is another shift matrix satisfying (2.1), such that $\dot s_{i,i+1} + \dot s_{i+1,i} = s_{i,i+1} + s_{i+1,i}$ for all $i = 1, \dots, n-1$, i.e. the differences associated to $\dot\sigma$ are the same as for σ. A check of relations shows that there is a unique algebra isomorphism $\iota : Y_n(\sigma) \to Y_n(\dot\sigma)$ defined by

$$\iota(D_i^{(r)}) = \dot D_i^{(r)}, \quad \iota(E_i^{(r)}) = \dot E_i^{(r - s_{i,i+1} + \dot s_{i,i+1})}, \quad \iota(F_i^{(r)}) = \dot F_i^{(r - s_{i+1,i} + \dot s_{i+1,i})}. \tag{2.17}$$

(Here and later on we denote the generators $D_i^{(r)}, E_i^{(r)}$ and $F_i^{(r)}$ of $Y_n(\dot\sigma)$ instead by $\dot D_i^{(r)}, \dot E_i^{(r)}$ and $\dot F_i^{(r)}$ to avoid potential confusion.) Hence the isomorphism type of the algebra $Y_n(\sigma)$ only depends on the differences δ, rather than on the shift matrix σ itself. However, the embedding of $Y_n(\sigma)$ into Y_n from in the previous paragraph clearly does depend on the particular choice of σ.

Let us now prove as promised that the canonical map $Y_n(\sigma) \to Y_n$ is injective. Introduce the *loop filtration* $\mathrm{L}_0 Y_n(\sigma) \subseteq \mathrm{L}_1 Y_n(\sigma) \subseteq \cdots$ by declaring that the generators $D_i^{(r)}, E_i^{(r)}$ and $F_i^{(r)}$ of $Y_n(\sigma)$ are of degree $(r-1)$, i.e. $\mathrm{L}_d Y_n(\sigma)$ is the span of all monomials in the generators of total degree $\leq d$. For $1 \leq i < j \leq n$ and $r > s_{i,j}$, define elements $E_{i,j}^{(r)} \in Y_n(\sigma)$ recursively by

$$E_{i,i+1}^{(r)} := E_i^{(r)}, \qquad E_{i,j}^{(r)} := [E_{i,j-1}^{(r-s_{j-1,j})}, E_{j-1}^{(s_{j-1,j}+1)}]. \tag{2.18}$$

Similarly, for $1 \leq i < j \leq n$ and $r > s_{j,i}$, define elements $F_{i,j}^{(r)} \in Y_n(\sigma)$ by

$$F_{i,i+1}^{(r)} := F_i^{(r)}, \qquad F_{i,j}^{(r)} := [F_{j-1}^{(s_{j,j-1}+1)}, F_{i,j-1}^{(r-s_{j,j-1})}]. \tag{2.19}$$

It is easy to see that $E_{i,j}^{(r)}$ and $F_{i,j}^{(r)}$ belong to $\mathrm{L}_{r-1}Y_n(\sigma)$. For $1 \leq i,j \leq n$ and $r \geq s_{i,j}$, define

$$e_{i,j;r} := \begin{cases} \mathrm{gr}_r^{\mathrm{L}}\, D_i^{(r+1)} & \text{if } i = j, \\ \mathrm{gr}_r^{\mathrm{L}}\, E_{i,j}^{(r+1)} & \text{if } i < j, \\ \mathrm{gr}_r^{\mathrm{L}}\, F_{j,i}^{(r+1)} & \text{if } i > j, \end{cases} \tag{2.20}$$

all elements of the associated graded algebra $\mathrm{gr}^{\mathrm{L}}\, Y_n(\sigma)$. Let $\mathfrak{gl}_n[t]$ denote the Lie algebra $\mathfrak{gl}_n \otimes \mathbb{C}[t]$ on basis $\{e_{i,j}t^r\}_{1 \leq i,j \leq n, r \geq 0}$, viewed as a graded Lie algebra so that $e_{i,j}t^r$ is of degree r. In view of the assumption (2.1), the vectors $\{e_{i,j}t^r\}_{1 \leq i,j \leq n, r \geq s_{i,j}}$ span a subalgebra of $\mathfrak{gl}_n[t]$ which we denote by $\mathfrak{gl}_n[t](\sigma)$ (the "shifted loop algebra"). The grading on $\mathfrak{gl}_n[t](\sigma)$ induces a grading on the universal enveloping algebra $U(\mathfrak{gl}_n[t](\sigma))$.

Theorem 2.1. *There is an isomorphism $\pi : U(\mathfrak{gl}_n[t](\sigma)) \to \mathrm{gr}^{\mathrm{L}}\, Y_n(\sigma)$ of graded algebras such that $e_{i,j}t^r \mapsto e_{i,j;r}$ for each $1 \leq i,j \leq n$ and $r \geq s_{i,j}$.*

Proof. Using the relations like in the proof of [BK1, Lemma 5.8], one shows for all $1 \leq h,i,j,k \leq n$, $r \geq s_{i,j}$ and $s \geq s_{h,k}$ that

$$[e_{i,j;r}, e_{h,k;s}] = e_{i,k;r+s}\delta_{h,j} - \delta_{i,k}e_{h,j;r+s}, \tag{2.21}$$

equality in $\mathrm{gr}^{\mathrm{L}}\, Y_n(\sigma)$. Hence there is a well-defined surjection $\pi : U(\mathfrak{gl}_n[t](\sigma)) \twoheadrightarrow \mathrm{gr}^{\mathrm{L}}\, Y_n(\sigma)$ mapping $e_{i,j}t^r \in \mathfrak{gl}_n[t](\sigma)$ to $e_{i,j;r} \in \mathrm{gr}^{\mathrm{L}}\, Y_n(\sigma)$.

In the special case $\sigma = 0$, the PBW theorem for the usual Yangian Y_n implies that the ordered monomials in the elements $\{e_{i,j;r}\}_{1 \leq i,j \leq n, r \geq 0}$ are linearly independent in $\mathrm{gr}^{\mathrm{L}}\, Y_n$; see the proof of [BK1, Lemma 5.10]. Hence π is an isomorphism in this case.

In general, the canonical map $Y_n(\sigma) \to Y_n$ is a homomorphism of filtered algebras, so induces a map $\mathrm{gr}^{\mathrm{L}}\, Y_n(\sigma) \to \mathrm{gr}^{\mathrm{L}}\, Y_n$ which sends $e_{i,j;r} \in \mathrm{gr}^{\mathrm{L}}\, Y_n(\sigma)$ to $e_{i,j;r} \in \mathrm{gr}^{\mathrm{L}}\, Y_n$ (despite the fact that it does *not* in general send $E_{i,j}^{(r+1)}, F_{i,j}^{(r+1)} \in Y_n(\sigma)$ to $E_{i,j}^{(r+1)}, F_{i,j}^{(r+1)} \in Y_n$ if $j - i > 1$). So the previous paragraph implies that the ordered monomials in the elements $\{e_{i,j;r}\}_{1 \leq i,j \leq n, r \geq s_{i,j}}$ are linearly independent in $\mathrm{gr}^{\mathrm{L}}\, Y_n(\sigma)$ too. Hence π is an isomorphism in general. □

Corollary 2.2. *The canonical map $Y_n(\sigma) \to Y_n$ is injective.*

Proof. We saw in the proof of Theorem 2.1 that the canonical map $Y_n(\sigma) \to Y_n$ is filtered and the associated graded map $\mathrm{gr}^{\mathrm{L}}\, Y_n(\sigma) \to \mathrm{gr}^{\mathrm{L}}\, Y_n$ is injective. □

The presentation of $Y_n(\sigma)$ is adapted to the natural *triangular decomposition* of this algebra. Let $Y_{(1^n)}$ denote the subalgebra of $Y_n(\sigma)$ generated by the $D_i^{(r)}$'s, let $Y_{(1^n)}^+(\sigma)$ denote the subalgebra of $Y_n(\sigma)$ generated by the $E_i^{(r)}$'s and let $Y_{(1^n)}^-(\sigma)$ denote the subalgebra generated by the $F_i^{(r)}$'s, for all admissible i, r.

Theorem 2.3. (i) *The monomials in the elements* $\{D_i^{(r)}\}_{1\leq i\leq n, r>0}$ *taken in some fixed order form a basis for* $Y_{(1^n)}$.

(ii) *The monomials in the elements* $\{E_{i,j}^{(r)}\}_{1\leq i<j\leq n, r>s_{i,j}}$ *taken in some fixed order form a basis for* $Y_{(1^n)}^+(\sigma)$.

(iii) *The monomials in the elements* $\{F_{i,j}^{(r)}\}_{1\leq i<j\leq n, r>s_{j,i}}$ *taken in some fixed order form a basis for* $Y_{(1^n)}^-(\sigma)$.

(iv) *The monomials in the union of the elements listed in (i)–(iii) taken in some fixed order form a basis for* $Y_n(\sigma)$.

Proof. Part (iv) follows from Theorem 2.1 and the PBW theorem for $U(\mathfrak{gl}_n[t](\sigma))$. The other parts are proved similarly, going back to (2.21). □

Corollary 2.4. *The natural multiplication map* $Y_{(1^n)}^-(\sigma)\otimes Y_{(1^n)}\otimes Y_{(1^n)}^+(\sigma)\to Y_n(\sigma)$ *is a vector space isomorphism.*

Remark 2.5. Let us describe the center $Z(Y_n(\sigma))$ of $Y_n(\sigma)$. Recalling the notation (2.2), let

$$C_n(u)=\sum_{r\geq 0} C_n^{(r)}u^{-r} := D_1(u)D_2(u-1)\cdots D_n(u-n+1)\in Y_n(\sigma)[[u^{-1}]]. \tag{2.22}$$

Then, the elements $C_n^{(1)}, C_n^{(2)},\dots$ are algebraically independent and generate $Z(Y_n(\sigma))$. Indeed, exploiting the embedding $Y_n(\sigma)\hookrightarrow Y_n$, it is known by [BK1, Theorem 7.2] that the elements $C_n^{(1)}, C_n^{(2)},\dots$ are algebraically independent and generate $Z(Y_n)$, so they certainly belong to $Z(Y_n(\sigma))$. The fact that $Z(Y_n(\sigma))$ is no larger than $Z(Y_n)$ may be proved by passing to the associated graded algebra $\operatorname{gr}^{\mathrm{L}} Y_n(\sigma)$ and using a variation on the trick from the proof of [MNO, Theorem 2.13]. We will outline a different argument in Remark 11.11 below.

3. Parabolic presentations

We are going to need various more elaborate presentations of the shifted Yangian which are analogues of the parabolic presentations of [BK1]. In order to explain the relationship between all these presentations, we begin with an elementary remark about Gauss factorizations.

Let $T=(T_{i,j})_{1\leq i,j\leq n}$ be an $n\times n$ matrix with entries in some ring such that the submatrices $(T_{i,j})_{1\leq i,j\leq m}$ are invertible for all $m=1,\dots,n$. Fix also a tuple $\nu=(\nu_1,\dots,\nu_m)$ of positive integers summing to n, which we think of as the *shape* of the standard Levi subgroup $\mathfrak{gl}_\nu := \mathfrak{gl}_{\nu_1}\oplus\cdots\oplus\mathfrak{gl}_{\nu_m}$ of $\mathfrak{gl}_n$. Working with $m\times m$ block matrices so that the ab-block is of size $\nu_a\times\nu_b$, the matrix T possesses a unique Gauss factorization $T={}^\nu F\,{}^\nu D\,{}^\nu E$ where ${}^\nu D$ is a block diagonal matrix, ${}^\nu E$ is a block upper unitriangular matrix, and ${}^\nu F$ is a block lower unitriangular matrix:

$${}^\nu D=\begin{pmatrix} {}^\nu D_1 & 0 & & \cdots & 0\\ 0 & {}^\nu D_2 & & & 0\\ \vdots & & \ddots & & \vdots\\ 0 & & & {}^\nu D_{m-1} & 0\\ 0 & \cdots & & 0 & {}^\nu D_m\end{pmatrix},$$

$$
{}^{\nu}E = \begin{pmatrix} I_{\nu_1} & {}^{\nu}E_1 & * & \cdots & * \\ 0 & I_{\nu_2} & {}^{\nu}E_2 & & \vdots \\ \vdots & & I_{\nu_3} & \ddots & * \\ 0 & & & \ddots & {}^{\nu}E_{m-1} \\ 0 & \cdots & & 0 & I_{\nu_m} \end{pmatrix}, \quad {}^{\nu}F = \begin{pmatrix} I_{\nu_1} & 0 & & \cdots & 0 \\ {}^{\nu}F_1 & I_{\nu_2} & & & 0 \\ * & {}^{\nu}F_2 & I_{\nu_3} & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ * & \cdots & * & {}^{\nu}F_{m-1} & I_{\nu_m} \end{pmatrix}.
$$

The diagonal blocks of ${}^{\nu}D$ define matrices ${}^{\nu}D_1, \dots, {}^{\nu}D_m$, the upper diagonal blocks of ${}^{\nu}E$ define matrices ${}^{\nu}E_1, \dots, {}^{\nu}E_{m-1}$, and the lower diagonal blocks of ${}^{\nu}F$ define matrices ${}^{\nu}F_1, \dots, {}^{\nu}F_{m-1}$. So ${}^{\nu}D_a$ is a $\nu_a \times \nu_a$ matrix, ${}^{\nu}E_a$ is a $\nu_a \times \nu_{a+1}$ matrix, and ${}^{\nu}F_a$ is a $\nu_{a+1} \times \nu_a$ matrix. Now consider what happens when we split a block into two: suppose that $\nu_b = \alpha + \beta$ for some $1 \le b \le m$ and $\alpha, \beta \ge 1$, and let $\mu = (\nu_1, \dots, \nu_{b-1}, \alpha, \beta, \nu_{b+1}, \dots, \nu_m)$. The following lemma shows how to compute the matrices ${}^{\mu}D_1, \dots, {}^{\mu}D_{m+1}, {}^{\mu}E_1, \dots, {}^{\mu}E_m$ and ${}^{\mu}F_1, \dots, {}^{\mu}F_m$ just from knowledge of the matrices ${}^{\nu}D_1, \dots, {}^{\nu}D_m, {}^{\nu}E_1, \dots, {}^{\nu}E_{m-1}$ and ${}^{\nu}F_1, \dots, {}^{\nu}F_{m-1}$.

Lemma 3.1. *In the above notation, define an $\alpha \times \alpha$ matrix A, an $\alpha \times \beta$ matrix B, a $\beta \times \alpha$ matrix C and a $\beta \times \beta$ matrix D from the equation*

$$
{}^{\nu}D_b = \begin{pmatrix} I_\alpha & 0 \\ C & I_\beta \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix} \begin{pmatrix} I_\alpha & B \\ 0 & I_\beta \end{pmatrix}.
$$

Then,

(i) ${}^{\mu}D_a = {}^{\nu}D_a$ *for* $a < b$, ${}^{\mu}D_b = A$, ${}^{\mu}D_{b+1} = D$, *and* ${}^{\mu}D_c = {}^{\nu}D_{c-1}$ *for* $c > b+1$;

(ii) ${}^{\mu}E_a = {}^{\nu}E_a$ *for* $a < b-1$, ${}^{\mu}E_{b-1}$ *is the submatrix consisting of the first* α *columns of* ${}^{\nu}E_{b-1}$, ${}^{\mu}E_b = B$, ${}^{\mu}E_{b+1}$ *is the submatrix consisting of the last* β *rows of* ${}^{\nu}E_b$, *and* ${}^{\mu}E_c = {}^{\nu}E_{c-1}$ *for* $c > b+1$;

(iii) ${}^{\mu}F_a = {}^{\nu}F_a$ *for* $a < b-1$, ${}^{\mu}F_{b-1}$ *is the submatrix consisting of the first* α *rows of* ${}^{\nu}F_{b-1}$, ${}^{\mu}F_b = C$, ${}^{\mu}F_{b+1}$ *is the submatrix consisting of the last* β *columns of* ${}^{\nu}F_b$, *and* ${}^{\mu}F_c = {}^{\nu}F_{c-1}$ *for* $c > b+1$;

Proof. Multiply matrices. ∎

Now let us briefly recall the parabolic presentations for the Yangian Y_n from [BK1]. As in [MNO], the Yangian can be defined in terms of the RTT presentation as the algebra over $\mathbb{C}$ defined by generators $\{T_{i,j}^{(r)}\}_{1 \le i,j \le n, r > 0}$ subject just to the relations

$$
[T_{i,j}^{(r)}, T_{h,k}^{(s)}] = \sum_{t=0}^{\min(r,s)-1} \left(T_{i,k}^{(r+s-1-t)} T_{h,j}^{(t)} - T_{i,k}^{(t)} T_{h,j}^{(r+s-1-t)} \right) \tag{3.1}
$$

for every $1 \le h, i, j, k \le n$ and $r, s > 0$, where $T_{i,j}^{(0)} := \delta_{i,j}$. Let $T_{i,j}(u) := \sum_{r \ge 0} T_{i,j}^{(r)} u^{-r}$ and let $T(u)$ denote the $n \times n$ matrix $(T_{i,j}(u))_{1 \le i,j \le n}$. Given a shape $\nu = (\nu_1, \dots, \nu_m)$, consider the Gauss factorization $T(u) = F(u)D(u)E(u)$ where $D(u)$ is a block diagonal matrix, $E(u)$ is a block upper unitriangular matrix and $F(u)$ is a block lower unitriangular matrices, all block matrices being of shape ν like before. We will from now omit any extra superscript ν since it should be clear from the context which shape ν we have in mind. The diagonal blocks of $D(u)$ define matrices $D_1(u), \dots, D_m(u)$, the upper diagonal blocks of $E(u)$ define matrices $E_1(u), \dots, E_{m-1}(u)$, and the lower diagonal blocks of $F(u)$ define matrices $F_1(u), \dots, F_{m-1}(u)$. Thus $D_a(u) = (D_{a;i,j}(u))_{1 \le i,j \le \nu_a}$

is a $\nu_a \times \nu_a$ matrix, $E_a(u) = (E_{a;i,j}(u))_{1\leq i\leq\nu_a, 1\leq j\leq\nu_{a+1}}$ is a $\nu_a \times \nu_{a+1}$ matrix, and $F_a(u) = (F_{a;i,j}(u))_{1\leq i\leq\nu_{a+1}, 1\leq j\leq\nu_a}$ is a $\nu_{a+1} \times \nu_a$ matrix. Write

$$D_{a;i,j}(u) = \sum_{r\geq 0} D_{a;i,j}^{(r)} u^{-r}, \quad E_{a;i,j}(u) = \sum_{r>0} E_{a;i,j}^{(r)} u^{-r}, \quad F_{a;i,j}(u) = \sum_{r>0} F_{a;i,j}^{(r)} u^{-r},$$

thus defining elements $D_{a;i,j}^{(r)}, E_{a;i,j}^{(r)}$ and $F_{a;i,j}^{(r)}$ of Y_n, all dependent of course on the fixed choice of ν. Now [BK1, Theorem A] shows that the elements

$$\{D_{a;i,j}^{(r)}\}_{1\leq a\leq m, 1\leq i,j\leq\nu_a, r>0},$$
$$\{E_{a;i,j}^{(r)}\}_{1\leq a<m, 1\leq i\leq\nu_a, 1\leq j\leq\nu_{a+1}, r>0},$$
$$\{F_{a;i,j}^{(r)}\}_{1\leq a<m, 1\leq i\leq\nu_{a+1}, 1\leq j\leq\nu_a, r>0}$$

generate Y_n subject only to the relations (3.3)–(3.14) below (taking the shift matrix there to be the zero matrix). For example, the presentation for Y_n from §2 is the special case $\nu = (1^n)$, in which case we denote $D_{i;1,1}^{(r)}, E_{i;1,1}^{(r)}$ and $F_{i;1,1}^{(r)}$ simply by $D_i^{(r)}, E_i^{(r)}$ and $F_i^{(r)}$ respectively, while the RTT presentation from (3.1) is the special case $\nu = (n)$, in which case $D_{1;i,j}^{(r)} = T_{i,j}^{(r)}$.

We are going to adapt these parabolic presentations to the shifted Yangian $Y_n(\sigma)$. Return to the setup of §2, assuming that $\sigma = (s_{i,j})_{1\leq i,j\leq n}$ is a fixed shift matrix with associated differences $\delta = (d_1, \dots, d_{n-1})$. Suppose in addition that the shape $\nu = (\nu_1, \dots, \nu_m)$ is *admissible* for σ, meaning that $d_i = 0$ for all $\nu_1 + \cdots + \nu_{a-1} < i < \nu_1 + \cdots + \nu_a$ and $a = 1, \dots, m$. We will adopt the shorthand

$$s_{a,b}(\nu) := s_{\nu_1+\cdots+\nu_a, \nu_1+\cdots+\nu_b}. \tag{3.2}$$

The shifts $(s_{i,j})_{1\leq i,j\leq n}$ can be recovered from the "relative" shifts $(s_{a,b}(\nu))_{1\leq a,b\leq m}$ given the admissible shape ν.

Define a new algebra ${}^\nu Y_n(\sigma)$ over $\mathbb{C}$, which will shortly be identified with $Y_n(\sigma)$ from §2, by generators

$$\{D_{a;i,j}^{(r)}\}_{1\leq a\leq m, 1\leq i,j\leq\nu_a, r>0},$$
$$\{E_{a;i,j}^{(r)}\}_{1\leq a<m, 1\leq i\leq\nu_a, 1\leq j\leq\nu_{a+1}, r>s_{a,a+1}(\nu)},$$
$$\{F_{a;i,j}^{(r)}\}_{1\leq a<m, 1\leq i\leq\nu_{a+1}, 1\leq j\leq\nu_a, r>s_{a+1,a}(\nu)}$$

subject to certain relations. To write down these relations, we must introduce some further notation. Let $D_{a;i,j}^{(0)} := \delta_{i,j}$, $D_{a;i,j}(u) := \sum_{r\geq 0} D_{a;i,j}^{(r)} u^{-r}$ and introduce the matrix $D_a(u) := (D_{a;i,j}(u))_{1\leq i,j\leq\nu_a}$. Let $\widetilde{D}_a(u) = (\widetilde{D}_{a;i,j}(u))_{1\leq i,j\leq\nu_a}$ denote the matrix $-D_a(u)^{-1}$, and write $\widetilde{D}_{a;i,j}(u) = \sum_{r\geq 0} \widetilde{D}_{a;i,j}^{(r)} u^{-r}$, thus defining elements $\widetilde{D}_{a;i,j}^{(r)}$ of ${}^\nu Y_n(\sigma)$ for each $a = 1, \dots, m$, $1 \leq i, j \leq \nu_a$ and $r \geq 0$. In particular, $\widetilde{D}_{a;i,j}^{(0)} = -\delta_{i,j}$ and

$\widetilde{D}_{a;i,j}^{(1)} = D_{a;i,j}^{(1)}$. Now the relations are:

$$[D_{a;i,j}^{(r)}, D_{b;h,k}^{(s)}] = \delta_{a,b} \sum_{t=0}^{\min(r,s)-1} \left(D_{a;i,k}^{(r+s-1-t)} D_{a;h,j}^{(t)} - D_{a;i,k}^{(t)} D_{a;h,j}^{(r+s-1-t)} \right), \tag{3.3}$$

$$[E_{a;i,j}^{(r)}, F_{b;h,k}^{(s)}] = \delta_{a,b} \sum_{t=0}^{r+s-1} \widetilde{D}_{a;i,k}^{(r+s-1-t)} D_{a+1;h,j}^{(t)}, \tag{3.4}$$

$$[D_{a;i,j}^{(r)}, E_{b;h,k}^{(s)}] = \delta_{a,b} \sum_{t=0}^{r-1} \sum_{g=1}^{\nu_a} D_{a;i,g}^{(t)} E_{a;g,k}^{(r+s-1-t)} \delta_{h,j} - \delta_{a,b+1} \sum_{t=0}^{r-1} D_{b+1;i,k}^{(t)} E_{b;h,j}^{(r+s-1-t)}, \tag{3.5}$$

$$[D_{a;i,j}^{(r)}, F_{b;h,k}^{(s)}] = \delta_{a,b+1} \sum_{t=0}^{r-1} F_{b;i,k}^{(r+s-1-t)} D_{b+1;h,j}^{(t)} - \delta_{a,b} \delta_{i,k} \sum_{t=0}^{r-1} \sum_{g=1}^{\nu_a} F_{a;h,g}^{(r+s-1-t)} D_{a;g,j}^{(t)}, \tag{3.6}$$

$$[E_{a;i,j}^{(r)}, E_{a;h,k}^{(s)}] = \sum_{t=s_{a,a+1}(\nu)+1}^{s-1} E_{a;i,k}^{(t)} E_{a;h,j}^{(r+s-1-t)} - \sum_{t=s_{a,a+1}(\nu)+1}^{r-1} E_{a;i,k}^{(t)} E_{a;h,j}^{(r+s-1-t)}, \tag{3.7}$$

$$[F_{a;i,j}^{(r)}, F_{a;h,k}^{(s)}] = \sum_{t=s_{a+1,a}(\nu)+1}^{r-1} F_{a;i,k}^{(r+s-1-t)} F_{a;h,j}^{(t)} - \sum_{t=s_{a+1,a}(\nu)+1}^{s-1} F_{a;i,k}^{(r+s-1-t)} F_{a;h,j}^{(t)}, \tag{3.8}$$

$$[E_{a;i,j}^{(r)}, E_{a+1;h,k}^{(s+1)}] - [E_{a;i,j}^{(r+1)}, E_{a+1;h,k}^{(s)}] = - \sum_{g=1}^{\nu_{a+1}} E_{a;i,g}^{(r)} E_{a+1;g,k}^{(s)} \delta_{h,j}, \tag{3.9}$$

$$[F_{a;i,j}^{(r+1)}, F_{a+1;h,k}^{(s)}] - [F_{a;i,j}^{(r)}, F_{a+1;h,k}^{(s+1)}] = -\delta_{i,k} \sum_{g=1}^{\nu_{a+1}} F_{a+1;h,g}^{(s)} F_{a;g,j}^{(r)}, \tag{3.10}$$

$$[E_{a;i,j}^{(r)}, E_{b;h,k}^{(s)}] = 0 \qquad \text{if } b > a+1 \text{ or if } b = a+1 \text{ and } h \neq j, \tag{3.11}$$

$$[F_{a;i,j}^{(r)}, F_{b;h,k}^{(s)}] = 0 \qquad \text{if } b > a+1 \text{ or if } b = a+1 \text{ and } i \neq k, \tag{3.12}$$

$$[E_{a;i,j}^{(r)}, [E_{a;h,k}^{(s)}, E_{b;f,g}^{(t)}]] + [E_{a;i,j}^{(s)}, [E_{a;h,k}^{(r)}, E_{b;f,g}^{(t)}]] = 0 \qquad \text{if } |a-b| = 1, \tag{3.13}$$

$$[F_{a;i,j}^{(r)}, [F_{a;h,k}^{(s)}, F_{b;f,g}^{(t)}]] + [F_{a;i,j}^{(s)}, [F_{a;h,k}^{(r)}, F_{b;f,g}^{(t)}]] = 0 \qquad \text{if } |a-b| = 1, \tag{3.14}$$

for all admissible $a, b, f, g, h, i, j, k, r, s, t$.

Observe right away that there is a canonical homomorphism ${}^{\nu}Y_n(\sigma) \to Y_n$ mapping the generators $D_{a;i,j}^{(r)}, E_{a;i,j}^{(r)}$ and $F_{a;i,j}^{(r)}$ of ${}^{\nu}Y_n(\sigma)$ to the elements of Y_n with the same names (the ones we defined above in terms of Gauss factorizations). We are going to prove that this canonical homomorphism is injective and that its image is independent of the particular choice of the admissible shape ν. In particular this will identify ${}^{\nu}Y_n(\sigma)$ with the algebra $Y_n(\sigma)$ from §2, since that is the special case $\nu = (1^n)$ of the present definition.

The proof that the canonical map ${}^{\nu}Y_n(\sigma) \to Y_n$ is injective is an extension of the proof given in §2. For $1 \leq a < b \leq m$, $1 \leq i \leq \nu_a, 1 \leq j \leq \nu_b$ and $r > s_{a,b}(\nu)$, we define elements $E_{a,b;i,j}^{(r)}$ inductively by

$$E_{a,a+1;i,j}^{(r)} := E_{a;i,j}^{(r)}, \qquad E_{a,b;i,j}^{(r)} := [E_{a,b-1;i,k}^{(r-s_{b-1,b}(\nu))}, E_{b-1;k,j}^{(s_{b-1,b}(\nu)+1)}] \tag{3.15}$$

where $1 \le k \le \nu_{b-1}$. By the relations this definition is independent of the choice of k; see for instance [BK1, (6.9)] for a similar argument. Similarly for $1 \le a < b \le m$, $1 \le i \le \nu_b$, $1 \le j \le \nu_a$ and $r > s_{b,a}(\nu)$, we define elements $F_{a,b;i,j}^{(r)}$ by

$$F_{a,a+1;i,j}^{(r)} := F_{a;i,j}^{(r)}, \qquad F_{a,b;i,j}^{(r)} := [F_{b-1;i,k}^{(s_{b,b-1}(\nu)+1)}, F_{a,b-1;k,j}^{(r-s_{b,b-1}(\nu))}] \tag{3.16}$$

where $1 \le k \le \nu_{b-1}$. Also let Y_ν denote the subalgebra of ${}^\nu Y_n(\sigma)$ generated by the $D_{a;i,j}^{(r)}$'s, let $Y_\nu^+(\sigma)$ denote the subalgebra generated by the $E_{a;i,j}^{(r)}$'s and let $Y_\nu^-(\sigma)$ denote the subalgebra generated by the elements $F_{a;i,j}^{(r)}$'s, for all admissible a, i, j, r. The following theorem generalizes Theorem 2.3.

Theorem 3.2. (i) *The monomials in the elements* $\{D_{a;i,j}^{(r)}\}_{a=1,\dots,m,1\le i,j\le\nu_a,r>0}$ *taken in some fixed order form a basis for* Y_ν.

(ii) *The monomials in the elements* $\{E_{a,b;i,j}^{(r)}\}_{1\le a<b\le m,1\le i\le\nu_a,1\le j\le\nu_b,r>s_{a,b}(\nu)}$ *taken in some fixed order form a basis for* $Y_\nu^+(\sigma)$.

(iii) *The monomials in the elements* $\{F_{a,b;i,j}^{(r)}\}_{1\le a<b\le m,1\le i\le\nu_b,1\le j\le\nu_a,r>s_{b,a}(\nu)}$ *taken in some fixed order form a basis for* $Y_\nu^-(\sigma)$.

(iv) *The monomials in the union of the elements listed in (i)–(iii) taken in some fixed order form a basis for* ${}^\nu Y_n(\sigma)$.

Proof. Introduce the loop filtration $\mathrm{L}_0{}^\nu Y_n(\sigma) \subseteq \mathrm{L}_1{}^\nu Y_n(\sigma) \subseteq \cdots$ of ${}^\nu Y_n(\sigma)$ by declaring that the generators $D_{a;i,j}^{(r)}, E_{a;i,j}^{(r)}$ and $F_{a;i,j}^{(r)}$ are all of degree $(r-1)$. Define elements $\{e_{i,j;r}\}_{1\le i,j\le n,r\ge s_{i,j}}$ of the associated graded algebra $\mathrm{gr}^{\mathrm{L}}\,{}^\nu Y_n(\sigma)$ from the equations

$$\mathrm{gr}_r^{\mathrm{L}}\, D_{a;i,j}^{(r+1)} = e_{\nu_1+\cdots+\nu_{a-1}+i,\nu_1+\cdots+\nu_{a-1}+j;r}, \tag{3.17}$$

$$\mathrm{gr}_r^{\mathrm{L}}\, E_{a,b;i,j}^{(r+1)} = e_{\nu_1+\cdots+\nu_{a-1}+i,\nu_1+\cdots+\nu_{b-1}+j;r}, \tag{3.18}$$

$$\mathrm{gr}_r^{\mathrm{L}}\, F_{a,b;i,j}^{(r+1)} = e_{\nu_1+\cdots+\nu_{b-1}+i,\nu_1+\cdots+\nu_{a-1}+j;r}. \tag{3.19}$$

Following the proof of [BK1, Lemma 6.6], one checks that these elements satisfy the relations (2.21). Hence, there is a well-defined surjective homomorphism $\pi : U(\mathfrak{gl}_n[t](\sigma)) \twoheadrightarrow \mathrm{gr}^{\mathrm{L}}\,{}^\nu Y_n(\sigma)$ mapping $e_{i,j}t^r \in U(\mathfrak{gl}_n[t](\sigma))$ to $e_{i,j;r} \in \mathrm{gr}^{\mathrm{L}}\,{}^\nu Y_n(\sigma)$.

In the special case $\sigma = 0$, when we already know that ${}^\nu Y_n(\sigma)$ is the usual Yangian Y_n, one checks using Lemma 3.1 and induction on the length m of the shape ν that the element $e_{i,j;r}$ defined here is equal to $\mathrm{gr}_r^{\mathrm{L}}\, T_{i,j}^{(r+1)}$. In particular $e_{i,j;r}$ coincides with the element of $\mathrm{gr}_r^{\mathrm{L}}\, Y_n$ defined by (2.20). Hence like in the proof of Theorem 2.1, the PBW theorem for the usual Yangian implies that the ordered monomials in the elements $\{e_{i,j;r}\}_{1\le i,j\le n,r\ge 0}$ are linearly independent in $\mathrm{gr}^{\mathrm{L}}\, Y_n$, and π is an isomorphism.

In general, the canonical map ${}^\nu Y_n(\sigma) \to Y_n$ is a homomorphism of filtered algebras, so induces a map $\mathrm{gr}^{\mathrm{L}}\,{}^\nu Y_n(\sigma) \to \mathrm{gr}^{\mathrm{L}}\, Y_n$ which maps $e_{i,j;r} \in \mathrm{gr}^{\mathrm{L}}\,{}^\nu Y_n(\sigma)$ to $e_{i,j;r} \in \mathrm{gr}^{\mathrm{L}}\, Y_n$. So the previous paragraph implies that the ordered monomials in the elements $\{e_{i,j;r}\}_{1\le i,j\le n,r\ge s_{i,j}}$ are linearly independent in $\mathrm{gr}^{\mathrm{L}}\,{}^\nu Y_n(\sigma)$ too. Hence π is an isomorphism in general.

The theorem now follows like Theorem 2.3. □

The following two corollaries generalize Corollaries 2.2 and 2.4.

Corollary 3.3. *The canonical map* ${}^\nu Y_n(\sigma) \to Y_n$ *is injective.*

Corollary 3.4. *Multiplication* $Y_\nu^-(\sigma) \times Y_\nu \times Y_\nu^+(\sigma) \to {}^\nu Y_n(\sigma)$ *is a vector space isomorphism.*

So now for each admissible shape ν, we have defined a subalgebra ${}^\nu Y_n(\sigma)$ of Y_n. It remains to see that these subalgebras coincide for different ν. Suppose $\nu_b = \alpha + \beta$ for some $1 \leq b \leq m$ and $\alpha, \beta \geq 1$, and let $\mu = (\nu_1, \dots, \nu_{b-1}, \alpha, \beta, \nu_{b+1}, \dots, \nu_m)$. Then it suffices to show that ${}^\nu Y_n(\sigma) = {}^\mu Y_n(\sigma)$ as subalgebras of Y_n. Using Lemma 3.1, one checks that ${}^\mu Y_n(\sigma) \subseteq {}^\nu Y_n(\sigma)$. Now the equality ${}^\mu Y_n(\sigma) = {}^\nu Y_n(\sigma)$ follows easily because we have already seen in the proof of Theorem 3.2 that their associated graded algebras are equal in $\mathrm{gr}^{\mathrm{L}} Y_n$. We have now proved that the relations (3.3)–(3.14) give presentations for the shifted Yangian $Y_n(\sigma) = {}^\nu Y_n(\sigma)$ for each admissible shape ν.

Remark 3.5. As a first application of these parabolic presentations, one can introduce analogues of parabolic subalgebras of $Y_n(\sigma)$: for an admissible shape ν, define $Y_\nu^\sharp(\sigma) := Y_\nu Y_\nu^+(\sigma)$ and $Y_\nu^\flat(\sigma) := Y_\nu^-(\sigma) Y_\nu$. By the relations, these are indeed subalgebras of $Y_n(\sigma)$. Moreover, there are obvious surjective homomorphisms $Y_\nu^\sharp(\sigma) \twoheadrightarrow Y_\nu$ and $Y_\nu^\flat(\sigma) \twoheadrightarrow Y_\nu$ with kernels generated by all $E_{a;i,j}^{(r)}$ and all $F_{a;i,j}^{(r)}$ respectively.

Remark 3.6. In this remark, we describe the maps τ and ι from (2.16)–(2.17) in terms of the parabolic generators. The anti-isomorphism τ satisfies

$$\tau(D_{a;i,j}^{(r)}) = D_{a;j,i}^{(r)}, \quad \tau(E_{a;i,j}^{(r)}) = F_{a;j,i}^{(r)}, \quad \tau(F_{a;i,j}^{(r)}) = E_{a;j,i}^{(r)}; \tag{3.20}$$

cf. [BK1, (6.6)–(6.8)]. Also, in the notation of (2.17) and working with a fixed admissible shape ν (for both σ and $\dot\sigma$), the isomorphism ι satisfies

$$\iota(D_{a;i,j}^{(r)}) = \dot D_{a;i,j}^{(r)}, \qquad \iota(E_{a;i,j}^{(r)}) = \dot E_{a;i,j}^{(r - s_{a,a+1}(\nu) + \dot s_{a,a+1}(\nu))},$$
$$\iota(F_{a;i,j}^{(r)}) = \dot F_{a;i,j}^{(r - s_{a+1,a}(\nu) + \dot s_{a+1,a}(\nu))}. \tag{3.21}$$

To see this, note by the relations (3.3)–(3.14) that these formulae certainly give a well-defined isomorphism $\iota : Y_n(\sigma) \to Y_n(\dot\sigma)$. Now use Lemma 3.1 to check inductively that this is the same map as in (2.17).

Remark 3.7. Finally we wish to write down a formula for the central elements $C_n(u)$ from (2.22) in terms of the parabolic generators. For this, we need to define the determinant of an $n \times n$ matrix $A = (a_{i,j})$ with entries in a non-commutative ring. There are at least two sensible ways to do this, namely,

$$\operatorname{rdet} A = \sum_{w \in S_n} \operatorname{sgn}(\pi) a_{1,w1} \cdots a_{n,wn}, \tag{3.22}$$

$$\operatorname{cdet} A = \sum_{w \in S_n} \operatorname{sgn}(\pi) a_{w1,1} \cdots a_{wn,n}, \tag{3.23}$$

according to whether one keeps monomials in "row order" or in "column order". In the case of the Yangian Y_n itself, it is well known (see e.g. [BK1, Theorem 8.6]) that

$C_n(u)$ can be expressed in terms of the $T_{i,j}^{(r)}$'s as the *quantum determinant*

$$C_n(u) = \operatorname{rdet}\begin{pmatrix} T_{1,1}(u-n+1) & T_{1,2}(u-n+1) & \cdots & T_{1,n}(u-n+1) \\ \vdots & \vdots & \ddots & \vdots \\ T_{n-1,1}(u-1) & T_{n-1,2}(u-1) & \cdots & T_{n-1,n}(u-1) \\ T_{n,1}(u) & T_{n,2}(u) & \cdots & T_{n,n}(u) \end{pmatrix} \tag{3.24}$$

$$= \operatorname{cdet}\begin{pmatrix} T_{1,1}(u) & T_{1,2}(u-1) & \cdots & T_{1,n}(u-n+1) \\ \vdots & \vdots & \ddots & \vdots \\ T_{n-1,1}(u) & T_{n-1,2}(u-1) & \cdots & T_{n-1,n}(u-n+1) \\ T_{n,1}(u) & T_{n,2}(u-1) & \cdots & T_{n,n}(u-n+1) \end{pmatrix}. \tag{3.25}$$

The following formula expressing $C_n(u)$ in terms of parabolic generators of $Y_n(\sigma)$ for an arbitrary admissible shape $\nu = (\nu_1, \dots, \nu_m)$ is an immediate consequence:

$$C_n(u) = C_{1;\nu_1}(u)C_{2;\nu_2}(u-\nu_1)\cdots C_{m,\nu_m}(u-\nu_1-\cdots-\nu_{m-1}) \tag{3.26}$$

where

$$C_{a;\nu_a}(u) = \operatorname{rdet}\begin{pmatrix} D_{a;1,1}(u-\nu_a+1) & \cdots & D_{a;1,\nu_a}(u-\nu_a+1) \\ \vdots & \ddots & \vdots \\ D_{a;\nu_a,1}(u) & \cdots & D_{a;\nu_a,\nu_a}(u) \end{pmatrix} \tag{3.27}$$

$$= \operatorname{cdet}\begin{pmatrix} D_{a;1,1}(u) & \cdots & D_{a;1,\nu_a}(u-\nu_a+1) \\ \vdots & \ddots & \vdots \\ D_{a;\nu_a,1}(u) & \cdots & D_{a;\nu_a,\nu_a}(u-\nu_a+1) \end{pmatrix} \tag{3.28}$$

for each $a = 1, \dots, m$.

4. Baby comultiplications

Now let us explain the real reason why the parabolic presentations of $Y_n(\sigma)$ from §3 are so important. Fix a shift matrix $\sigma = (s_{i,j})_{1\le i,j\le n}$ with associated differences $\delta = (d_1, \dots, d_{n-1})$. Recall a shape $\nu = (\nu_1, \dots, \nu_m)$ is admissible if $d_i = 0$ for all $\nu_1 + \cdots + \nu_{a-1} < i < \nu_1 + \cdots + \nu_a$ and $a = 1, \dots, m$. Throughout the section, ν will denote the *minimal admissible shape* for σ, that is, the admissible shape of smallest possible length m. For example if $\delta = (5, 5, 0, 1, 0, 0, 0, 4)$ then $\nu = (1, 1, 2, 4, 1)$. We will always work in terms of the parabolic presentation defined relative to this fixed shape ν. For example in the case of the Yangian Y_n itself, this is the RTT presentation from (3.1).

Consider first the special case that $d_1 = \cdots = d_{n-1} = 0$, i.e. the case when $Y_n(\sigma)$ is the usual Yangian Y_n. It is well known that Y_n is a Hopf algebra, with comultiplication $\Delta : Y_n \to Y_n \otimes Y_n$ which may be defined in terms of the RTT presentation by the equation

$$\Delta(T_{i,j}^{(r)}) = \sum_{s=0}^{r}\sum_{k=1}^{n} T_{i,k}^{(s)} \otimes T_{k,j}^{(r-s)}. \tag{4.1}$$

There is also the *evaluation homomorphism* $\kappa_1 : Y_n \to U(\mathfrak{gl}_n)$ defined by

$$\kappa_1(T_{i,j}^{(r)}) = \begin{cases} e_{i,j} & \text{if } r = 1, \\ 0 & \text{if } r > 1. \end{cases} \tag{4.2}$$

Let $\Delta_{\rm R} := (\mathrm{id} \otimes \kappa_1) \circ \Delta$ and $\Delta_{\rm L} := (\kappa_1 \otimes \mathrm{id}) \circ \Delta$, thus defining algebra homomorphisms

$$\Delta_{\rm R} : Y_n \to Y_n \otimes U(\mathfrak{gl}_n), \qquad T_{i,j}^{(r)} \mapsto T_{i,j}^{(r)} \otimes 1 + \sum_{k=1}^{n} T_{i,k}^{(r-1)} \otimes e_{k,j}, \tag{4.3}$$

$$\Delta_{\rm L} : Y_n \to U(\mathfrak{gl}_n) \otimes Y_n, \qquad T_{i,j}^{(r)} \mapsto 1 \otimes T_{i,j}^{(r)} + \sum_{k=1}^{n} e_{i,k} \otimes T_{k,j}^{(r-1)}. \tag{4.4}$$

The following theorem defines analogous "baby comultiplications" for the shifted Yangians in general.

Theorem 4.1. *If $d_1 = \cdots = d_{n-1} = 0$ then let $t = n$; otherwise, let t be the smallest positive integer such that $d_{n-t} \neq 0$. For $1 \leq i, j \leq t$, set $\tilde{e}_{i,j} := e_{i,j} + \delta_{i,j}(n-t) \in U(\mathfrak{gl}_t)$.*

(i) *If either $t = n$ or $s_{n-t,n-t+1} \neq 0$, define $\dot{\sigma} = (\dot{s}_{i,j})_{1 \leq i,j \leq n}$ from*

$$\dot{s}_{i,j} = \begin{cases} s_{i,j} - 1 & \text{if } i \leq n-t < j, \\ s_{i,j} & \text{otherwise.} \end{cases} \tag{4.5}$$

Then, there is a unique algebra homomorphism $\Delta_{\rm R} : Y_n(\sigma) \to Y_n(\dot{\sigma}) \otimes U(\mathfrak{gl}_t)$ such that

$$D_{a;i,j}^{(r)} \mapsto \dot{D}_{a;i,j}^{(r)} \otimes 1 + \delta_{a,m} \sum_{k=1}^{t} \dot{D}_{a;i,k}^{(r-1)} \otimes \tilde{e}_{k,j},$$

$$E_{a;i,j}^{(r)} \mapsto \dot{E}_{a;i,j}^{(r)} \otimes 1 + \delta_{a,m-1} \sum_{k=1}^{t} \dot{E}_{a;i,k}^{(r-1)} \otimes \tilde{e}_{k,j},$$

$$F_{a;i,j}^{(r)} \mapsto \dot{F}_{a;i,j}^{(r)} \otimes 1,$$

for all admissible a, i, j, r.

(ii) *If either $t = n$ or $s_{n-t+1,n-t} \neq 0$, define $\dot{\sigma} = (\dot{s}_{i,j})_{1 \leq i,j \leq n}$ from*

$$\dot{s}_{i,j} = \begin{cases} s_{i,j} - 1 & \text{if } j \leq n-t < i, \\ s_{i,j} & \text{otherwise.} \end{cases} \tag{4.6}$$

Then, there is a unique algebra homomorphism $\Delta_{\rm L} : Y_n(\sigma) \to U(\mathfrak{gl}_t) \otimes Y_n(\dot{\sigma})$ such that

$$D_{a;i,j}^{(r)} \mapsto 1 \otimes \dot{D}_{a;i,j}^{(r)} + \delta_{a,m} \sum_{k=1}^{t} \tilde{e}_{i,k} \otimes \dot{D}_{a;k,j}^{(r-1)},$$

$$E_{a;i,j}^{(r)} \mapsto 1 \otimes \dot{E}_{a;i,j}^{(r)},$$

$$F_{a;i,j}^{(r)} \mapsto 1 \otimes \dot{F}_{a;i,j}^{(r)} + \delta_{a,m-1} \sum_{k=1}^{t} \tilde{e}_{i,k} \otimes \dot{F}_{a;k,j}^{(r-1)},$$

for all admissible a, i, j, r.

(Recall we are working in terms of the parabolic presentation of shape ν, where $\nu = (\nu_1, \dots, \nu_m)$ is the minimal admissible shape for σ; this is also an admissible shape for $\dot{\sigma}$.)

Proof. Check the relations (3.3)–(3.14) (to check the relations (3.13) and (3.14) one needs to use (3.7), (3.8) and (3.9), (3.10) several times). ∎

The next lemma explains how to compute the baby comultiplications on the higher root elements $E_{a,b;i,j}^{(r)}$ and $F_{a,b;i,j}^{(r)}$, still working in terms of the minimal admissible shape ν for σ.

Lemma 4.2. (i) *Under the hypotheses of Theorem 4.1(i), we have for* $1 \leq a < b-1 < m$ *and all admissible* i, j, r *that*

$$\Delta_{\mathrm{R}}(E_{a,b;i,j}^{(r)}) = \begin{cases} \dot{E}_{a,b;i,j}^{(r)} \otimes 1 & \text{if } b < m, \\ [\dot{E}_{a,m-1;i,h}^{(r-s_{m-1,m}(\nu))}, \dot{E}_{m-1;h,j}^{(s_{m-1,m}(\nu)+1)}] \otimes 1 + \sum_{k=1}^{t} \dot{E}_{a,m;i,k}^{(r-1)} \otimes \tilde{e}_{k,j} & \text{if } b = m, \end{cases}$$

$$\Delta_{\mathrm{R}}(F_{a,b;i,j}^{(r)}) = \dot{F}_{a,b;i,j}^{(r)} \otimes 1,$$

for any $1 \leq h \leq \nu_{m-1}$.

(ii) *Under the hypotheses of Theorem 4.1(ii), we have for* $1 \leq a < b-1 < m$ *and all admissible* i, j, r *that*

$$\Delta_{\mathrm{L}}(E_{a,b;i,j}^{(r)}) = 1 \otimes \dot{E}_{a,b;i,j}^{(r)},$$

$$\Delta_{\mathrm{L}}(F_{a,b;i,j}^{(r)}) = \begin{cases} 1 \otimes \dot{F}_{a,b;i,j}^{(r)} & \text{if } b < m, \\ 1 \otimes [\dot{F}_{m-1;i,h}^{(s_{m,m-1}(\nu)+1)}, \dot{F}_{a,m-1;h,j}^{(r-s_{m,m-1}(\nu))}] + \sum_{k=1}^{t} \tilde{e}_{i,k} \otimes \dot{F}_{a,m;k,j}^{(r-1)} & \text{if } b = m, \end{cases}$$

for any $1 \leq h \leq \nu_{m-1}$.

Proof. Let us just explain how to compute $\Delta_{\mathrm{R}}(E_{a,m;i,j}^{(r)})$ for $1 \leq a < m-1$, since all the other cases are similar. By definition, $E_{a,m;i,j}^{(r)} = [E_{a,m-1;i,h}^{(r-s_{m-1,m}(\nu))}, E_{m-1;h,j}^{(s_{m-1,m}(\nu)+1)}]$ for any $1 \leq h \leq \nu_{m-1}$. Clearly $\Delta_{\mathrm{R}}(E_{a,m-1;i,h}^{(r-s_{m-1,m}(\nu))}) = \dot{E}_{a,m-1;i,h}^{(r-s_{m-1,m}(\nu))} \otimes 1$. Hence,

$$\begin{aligned} \Delta_{\mathrm{R}}(E_{a,m;i,j}^{(r)}) &= \left[\dot{E}_{a,m-1;i,h}^{(r-s_{m-1,m}(\nu))} \otimes 1, \dot{E}_{m-1;h,j}^{(s_{m-1,m}(\nu)+1)} \otimes 1 + \sum_{k=1}^{t} \dot{E}_{m-1;h,k}^{(s_{m-1,m}(\nu))} \otimes \tilde{e}_{k,j}\right] \\ &= \left[\dot{E}_{a,m-1;i,h}^{(r-s_{m-1,m}(\nu))}, \dot{E}_{m-1;h,j}^{(s_{m-1,m}(\nu)+1)}\right] \otimes 1 + \sum_{k=1}^{t} \dot{E}_{a,m;i,k}^{(r-1)} \otimes \tilde{e}_{k,j}, \end{aligned}$$

as claimed. □

Remark 4.3. In the notation of Theorem 4.1, the maps Δ_{R} and Δ_{L} are injective. One can see this as follows. Let $\varepsilon : U(\mathfrak{gl}_t) \to \mathbb{C}$ be the homomorphism with $\varepsilon(\tilde{e}_{i,j}) = 0$ for $1 \leq i, j \leq t$. By definition, $Y_n(\sigma)$ and $Y_n(\dot{\sigma})$ are subalgebras of the Yangian with $Y_n(\sigma) \subseteq Y_n(\dot{\sigma})$. Now the point is that the compositions $(\mathrm{id}\,\bar{\otimes}\varepsilon) \circ \Delta_{\mathrm{R}}$ and $(\varepsilon\bar{\otimes}\,\mathrm{id}) \circ \Delta_{\mathrm{L}}$ coincide (when defined) with the natural embedding $Y_n(\sigma) \hookrightarrow Y_n(\dot{\sigma})$.

5. The canonical filtration

In this section, we introduce another important filtration of the shifted Yangian. It is easiest to start with the Yangian Y_n itself defined by the RTT presentation as in §3. The *canonical filtration* $\mathrm{F}_0 Y_n \subseteq \mathrm{F}_1 Y_n \subseteq \cdots$ of Y_n is defined by declaring that the generators $T_{i,j}^{(r)}$ are all of degree r, i.e. $\mathrm{F}_d Y_n$ is the span of all monomials in these generators of total degree $\leq d$. It is obvious from the relations (3.1) that the associated graded algebra $\mathrm{gr}\, Y_n$ is commutative.

Suppose instead that we are given a shape $\nu = (\nu_1, \dots, \nu_m)$. Using the explicit definitions from [BK1, (6.1)–(6.4)], one checks that the parabolic generators $D_{a;i,j}^{(r)}, E_{a,b;i,j}^{(r)}$

and $F_{a,b;i,j}^{(r)}$ of Y_n are linear combinations of monomials in the $T_{i,j}^{(s)}$ of total degree r and conversely, if we define $D_{a;i,j}^{(r)}, E_{a,b;i,j}^{(r)}$ and $F_{a,b;i,j}^{(r)}$ all to be of degree r, each $T_{i,j}^{(s)}$ is a linear combination of monomials in these elements of total degree s. Hence $\mathrm{F}_d Y_n$ can also be described as the span of all monomials in the elements $D_{a;i,j}^{(r)}, E_{a,b;i,j}^{(r)}$ and $F_{a,b;i,j}^{(r)}$ of total degree $\leq d$. For $1 \leq a, b \leq m$, $1 \leq i \leq \nu_a, 1 \leq j \leq \nu_b$ and $r > 0$, define

$$e_{a,b;i,j}^{(r)} := \begin{cases} \mathrm{gr}_r\, D_{a;i,j}^{(r)} & \text{if } a = b, \\ \mathrm{gr}_r\, E_{a,b;i,j}^{(r)} & \text{if } a < b, \\ \mathrm{gr}_r\, F_{b,a;i,j}^{(r)} & \text{if } a > b, \end{cases} \tag{5.1}$$

all elements of $\mathrm{gr}_r\, Y_n$. Of course, this notation depends implicitly on the fixed shape ν. The commutativity of $\mathrm{gr}\, Y_n$ together with Theorem 3.2(iv) imply:

Theorem 5.1. *For any shape $\nu = (\nu_1, \dots, \nu_m)$, $\mathrm{gr}\, Y_n$ is the free commutative algebra on generators $\{e_{a,b;i,j}^{(r)}\}_{1 \leq a,b \leq m, 1 \leq i \leq \nu_a, 1 \leq j \leq \nu_b, r > 0}$.*

Now consider the shifted Yangian. So let $\sigma = (s_{i,j})_{1 \leq i,j \leq n}$ be a shift matrix as in §2. Viewing $Y_n(\sigma)$ as a subalgebra of Y_n, introduce the *canonical filtration* $\mathrm{F}_0 Y_n(\sigma) \subseteq \mathrm{F}_1 Y_n(\sigma) \subseteq \cdots$ of $Y_n(\sigma)$ by defining $\mathrm{F}_d Y_n(\sigma) := Y_n(\sigma) \cap \mathrm{F}_d Y_n$. Thus, the inclusion of $Y_n(\sigma)$ into Y_n is filtered and the induced map $\mathrm{gr}\, Y_n(\sigma) \to \mathrm{gr}\, Y_n$ is injective. Hence, $\mathrm{gr}\, Y_n(\sigma)$ is identified with a subalgebra of the commutative algebra $\mathrm{gr}\, Y_n$. The following theorem describes this subalgebra explicitly.

Theorem 5.2. *For an admissible shape $\nu = (\nu_1, \dots, \nu_m)$, $\mathrm{gr}\, Y_n(\sigma)$ is the subalgebra of $\mathrm{gr}\, Y_n$ generated by the elements $\{e_{a,b;i,j}^{(r)}\}_{1 \leq a,b \leq m, 1 \leq i \leq \nu_a, 1 \leq j \leq \nu_b, r > s_{a,b}(\nu)}$.*

Proof. Note in view of the relations (3.9)–(3.10) that the element $e_{a,b;i,j}^{(r)}$ of $\mathrm{gr}\, Y_n(\sigma)$ is identified with the element of the same name in $\mathrm{gr}\, Y_n$. Given this the theorem follows directly from Theorems 5.1 and 3.2(iv). □

Remark 5.3. Theorem 5.2 means that, given an admissible shape ν, we can define the canonical filtration on $Y_n(\sigma)$ intrinsically simply by declaring that the elements $D_{a;i,j}^{(r)}, E_{a,b;i,j}^{(r)}$ and $F_{a,b;i,j}^{(r)}$ of $Y_n(\sigma)$ are all of degree r, i.e. $\mathrm{F}_d Y_n(\sigma)$ is the span of all monomials in these elements of total degree $\leq d$. In particular, this definition is independent of the particular choice of admissible shape ν.

Remark 5.4. The comultiplication $\Delta : Y_n \to Y_n \otimes Y_n$ is a filtered map with respect to the canonical filtration, as follows immediately from (4.1). Similarly, the baby comultiplications $\Delta_{\mathrm{R}} : Y_n(\sigma) \to Y_n(\dot{\sigma}) \otimes U(\mathfrak{gl}_t)$ and $\Delta_{\mathrm{L}} : Y_n(\sigma) \to U(\mathfrak{gl}_t) \otimes Y_n(\dot{\sigma})$ from Theorem 4.1 are filtered maps whenever they are defined, providing we extend the canonical filtration of $Y_n(\dot{\sigma})$ to $Y_n(\dot{\sigma}) \otimes U(\mathfrak{gl}_t)$ and $U(\mathfrak{gl}_t) \otimes Y_n(\dot{\sigma})$ by declaring that the matrix units $e_{i,j} \in \mathfrak{gl}_t$ are of degree 1. This follows from Theorem 4.1 and Lemma 4.2. The argument explained in Remark 4.3 shows moreover that the associated graded maps $\mathrm{gr}\, \Delta_{\mathrm{R}}$ and $\mathrm{gr}\, \Delta_{\mathrm{L}}$ are injective.

6. Truncation

Continue with a fixed shift matrix $\sigma = (s_{i,j})_{1 \leq i,j \leq n}$ with associated differences $\delta = (d_1, \dots, d_{n-1})$. Fix also an integer $l \geq d_1 + \cdots + d_{n-1} = s_{1,n} + s_{n,1}$ (the "level"), and for $i = 1, \dots, n$ let

$$p_i := l - d_i - d_{i+1} - \cdots - d_{n-1}, \tag{6.1}$$

thus defining a tuple $(p_1, \dots, p_n)$ of integers with $0 \leq p_1 \leq \cdots \leq p_n = l$. If $\nu = (\nu_1, \dots, \nu_m)$ is an admissible shape, we will also use the shorthand

$$p_a(\nu) := p_{\nu_1 + \cdots + \nu_a} \tag{6.2}$$

for each $a = 1, \dots, m$; cf. (3.2). (A better way to keep track of all this data will be explained in the next section.)

The *shifted Yangian of level* l, denoted $Y_{n,l}(\sigma)$, is defined to be the quotient of $Y_n(\sigma)$ by the two-sided ideal generated by the elements $\{D_1^{(r)}\}_{r>p_1}$. Alternatively, in terms of the parabolic presentation relative to an admissible shape $\nu = (\nu_1, \dots, \nu_m)$, $Y_{n,l}(\sigma)$ is the quotient of $Y_n(\sigma)$ by the two-sided ideal generated by the elements $\{D_{1;i,j}^{(r)}\}_{1 \leq i,j \leq \nu_1, r > p_1}$. The equivalence of all these definitions is easy to see since $D_{1;1,1}^{(r)} = D_1^{(r)}$ and all other $D_{1;i,j}^{(r)}$ for $r > p_1$ and $1 \leq i, j \leq \nu_1$ obviously lie in the two-sided ideal generated by $\{D_{1;1,1}^{(r)}\}_{r>p_1}$ in view of the relation (3.3). For example, in the case that σ is the zero matrix, when we denote $Y_{n,l}(\sigma)$ simply by $Y_{n,l}$, this algebra is the quotient of Y_n by the two-sided ideal generated by $\{T_{i,j}^{(r)} \mid 1 \leq i, j \leq n, r > l\}$, which is precisely the definition of the *Yangian of level* l from [C]. By convention, we also write $Y_{0,0} = Y_{0,0}(\sigma)$ for the trivial algebra $\mathbb{C}$.

In the hope that it is clear from context whether we are talking about $Y_n(\sigma)$ or $Y_{n,l}(\sigma)$, we will abuse notation and use the same symbols $D_{a;i,j}^{(r)}$, $\widetilde{D}_{a;i,j}^{(r)}$, $E_{a,b;i,j}^{(r)}$ and $F_{a,b;i,j}^{(r)}$ both for the generators of $Y_n(\sigma)$ and for their canonical images in the quotient $Y_{n,l}(\sigma)$. Similarly we use the same notation $C_n^{(r)}$ for the images in $Y_{n,l}(\sigma)$ of the central elements of $Y_n(\sigma)$ from (2.22); clearly these are also central in $Y_{n,l}(\sigma)$. The anti-isomorphism τ from (2.16) factors through the quotients to induce an anti-isomorphism

$$\tau : Y_{n,l}(\sigma) \to Y_{n,l}(\sigma^t). \tag{6.3}$$

Similarly, given another shift matrix $\dot{\sigma}$ with the same differences as σ, the isomorphism ι from (2.17) induces an isomorphism

$$\iota : Y_{n,l}(\sigma) \to Y_{n,l}(\dot{\sigma}). \tag{6.4}$$

So the isomorphism type of the algebra $Y_{n,l}(\sigma)$ only depends on the tuple $(p_1, \dots, p_n)$.

We will exploit the canonical filtration $\mathrm{F}_0 Y_{n,l}(\sigma) \subseteq \mathrm{F}_1 Y_{n,l}(\sigma) \subseteq \cdots$ of $Y_{n,l}(\sigma)$ induced by the quotient map $Y_n(\sigma) \twoheadrightarrow Y_{n,l}(\sigma)$ and the canonical filtration of $Y_n(\sigma)$ from §5. Recalling Remark 5.3, this may be defined directly given an admissible shape $\nu = (\nu_1, \dots, \nu_m)$ by declaring all the elements $D_{a;i,j}^{(r)}, E_{a,b;i,j}^{(r)}$ and $F_{a,b;i,j}^{(r)}$ of $Y_{n,l}(\sigma)$ to be of degree r, and then $\mathrm{F}_d Y_{n,l}(\sigma)$ is the span of all monomials in these elements of total degree $\leq d$. For $1 \leq a, b \leq m, 1 \leq i \leq \nu_a, 1 \leq j \leq \nu_b$ and $r > s_{a,b}(\nu)$, define elements $e_{a,b;i,j}^{(r)}$ of the associated graded algebra $\operatorname{gr} Y_{n,l}(\sigma)$ by the formula (5.1). Since $\operatorname{gr} Y_{n,l}(\sigma)$ is a quotient of $\operatorname{gr} Y_n(\sigma)$, Theorem 5.2 implies that it is commutative and is generated by all $\{e_{a,b;i,j}^{(r)}\}_{1 \leq a,b \leq m, 1 \leq i \leq \nu_a, 1 \leq j \leq \nu_b, r > s_{a,b}(\nu)}$.

Lemma 6.1. *For any admissible shape $\nu = (\nu_1, \dots, \nu_m)$, $\operatorname{gr} Y_{n,l}(\sigma)$ is generated just by the elements $\{e_{a,b;i,j}^{(r)}\}_{1 \leq a,b \leq m, 1 \leq i \leq \nu_a, 1 \leq j \leq \nu_b, s_{a,b}(\nu) < r \leq s_{a,b}(\nu) + p_{\min(a,b)}(\nu)}$.*

Proof. For $1 \leq c \leq m$, let Ω_c denote the set

$$\{D_{a;i,j}^{(r)}\}_{1\leq a\leq c,1\leq i,j\leq \nu_a,0<r\leq p_a(\nu)} \cup \{E_{a,b;i,j}^{(r)}\}_{1\leq a<b\leq c,1\leq i\leq \nu_a,1\leq j\leq \nu_b,s_{a,b}(\nu)<r\leq s_{a,b}(\nu)+p_a(\nu)} \cup \{F_{a,b;i,j}^{(r)}\}_{1\leq a<b\leq c,1\leq j\leq \nu_a,1\leq i\leq \nu_b,s_{b,a}(\nu)<r\leq s_{b,a}(\nu)+p_a(\nu)},$$

and let $\widehat{\Omega}_c$ denote

$$\{D_{a;i,j}^{(r)}\}_{1\leq a\leq c,1\leq i,j\leq \nu_a,r>0} \cup \{E_{a,b;i,j}^{(r)}\}_{1\leq a<b\leq c,1\leq i\leq \nu_a,1\leq j\leq \nu_b,r>s_{a,b}(\nu)} \cup \{F_{a,b;i,j}^{(r)}\}_{1\leq a<b\leq c,1\leq j\leq \nu_a,1\leq i\leq \nu_b,r>s_{b,a}(\nu)}.$$

To prove the lemma, we show by induction on $c = 1, \dots, m$ that every element of $\widehat{\Omega}_c$ of degree r can be expressed as a linear combination of monomials in the elements of Ω_c of total degree r. The case $c = 1$ is clear, since $D_{1;i,j}^{(r)} = 0$ for $r > p_1(\nu)$ by definition. Now take $c > 1$ and assume the result has been proved for all smaller c.

For $1 \leq a < c-1$, we have by definition that $E_{a,c;i,j}^{(r)} = [E_{a,c-1;i,k}^{(r-s_{c-1,c}(\nu))}, E_{c-1;k,j}^{(s_{c-1,c}(\nu)+1)}]$ for some $1 \leq k \leq \nu_{c-1}$. By induction, $E_{a,c-1;i,k}^{(r-s_{c-1,c}(\nu))}$ is a linear combination of monomials in the elements of Ω_{c-1} of total degree $(r - s_{c-1,c}(\nu))$. By the relations, the commutator of any such monomial with $E_{c-1;k,j}^{(s_{c-1,c}(\nu)+1)}$ is a linear combination of monomials in the elements of Ω_c of total degree r. Hence, $E_{a,c;i,j}^{(r)}$ is a linear combination of monomials in the elements of Ω_c of total degree r. Similarly, so is $F_{a,c;i,j}^{(r)}$.

Next, we have by the relations that

$$E_{c-1;i,j}^{(r)} = [D_{c-1;i,i}^{(r-s_{c-1,c}(\nu))}, E_{c-1;i,j}^{(s_{c-1,c}(\nu)+1)}] - \sum_{t=1}^{r-s_{c-1,c}(\nu)-1} \sum_{h=1}^{\nu_{c-1}} D_{c-1;i,h}^{(t)} E_{c-1;h,j}^{(r-t)}. \tag{6.5}$$

By induction, $D_{c-1;i,i}^{(r-s_{c-1,c}(\nu))}$ is a linear combination of monomials in the elements of Ω_{c-1} of total degree $(r - s_{c-1,c}(\nu))$. Hence by the relations, the first term on the right hand side of (6.5) is a linear combination of monomials in the elements of $\Omega_{c-1} \cup \{E_{a,c;i,j}^{(r)}\}_{1\leq a<c,1\leq i\leq \nu_a,1\leq j\leq \nu_c,s_{a,c}(\nu)<r\leq s_{a,c}(\nu)+p_a(\nu)}$ of total degree r. Now using (6.5) and induction on r, one deduces that each $E_{c-1;i,j}^{(r)}$ is a linear combination of monomials in the elements of Ω_c of total degree r. A similar argument using the relation

$$F_{c-1;i,j}^{(r)} = [F_{c-1;i,j}^{(s_{c,c-1}(\nu)+1)}, D_{c-1;j,j}^{(r-s_{c,c-1}(\nu))}] - \sum_{t=1}^{r-s_{c,c-1}(\nu)-1} \sum_{h=1}^{\nu_{c-1}} F_{c-1;i,h}^{(r-t)} D_{c-1;h,j}^{(t)} \tag{6.6}$$

shows that each $F_{c-1;i,j}^{(r)}$ is a linear combination of monomials in the elements of Ω_c of total degree r, too.

Finally we must consider $D_{c;i,j}^{(r)}$. Recall that $p_c(\nu) = p_{c-1}(\nu) + s_{c-1,c}(\nu) + s_{c,c-1}(\nu)$. By the relations, we have for $1 \leq k \leq \nu_{c-1}$ that

$$D_{c;i,j}^{(r)} = \sum_{t=0}^{r-1} \widetilde{D}_{c-1;k,k}^{(r-t)} D_{c;i,j}^{(t)} - [E_{c-1;k,j}^{(r-s_{c,c-1}(\nu))}, F_{c-1;i,k}^{(s_{c,c-1}(\nu)+1)}]. \tag{6.7}$$

By the previous paragraph, $E_{c-1;k,j}^{(r-s_{c,c-1}(\nu))}$ is a linear combination of monomials in the elements from $\Omega_{c-1} \cup \{E_{a,c;i,j}^{(r)}\}_{1\leq a<c,1\leq i\leq \nu_a,1\leq j\leq \nu_c,s_{a,c}(\nu)<r\leq s_{a,c}(\nu)+p_a(\nu)}$ of total degree

$(r - s_{c,c-1}(\nu))$. Hence by the relations, the second term on the right hand side of (6.7) is a linear combination of monomials in the elements of Ω_c of total degree r. Now using (6.7) and induction on r one deduces that each $D_{c;i,j}^{(r)}$ is a linear combination of monomials in the elements of Ω_c of total degree r, to complete the induction step. □

Assume additionally for the next paragraph that the level l is > 0. If $d_1 = \cdots = d_{n-1} = 0$, we let $t := n$; otherwise let t be the smallest positive integer such that $d_{n-t} \neq 0$. If either $t = n$ or $s_{n-t,n-t+1} \neq 0$, it is easy to check that the baby comultiplication Δ_{R} from Theorem 4.1(i) factors through the quotients to induce a map

$$\Delta_{\mathrm{R}} : Y_{n,l}(\sigma) \to Y_{n,l-1}(\dot{\sigma}) \otimes U(\mathfrak{gl}_t), \tag{6.8}$$

where $\dot{\sigma}$ is as in (4.5). Similarly, if either $t = n$ or $s_{n-t+1,n-t} \neq 0$, the baby comultiplication Δ_{L} from Theorem 4.1(ii) induces a map

$$\Delta_{\mathrm{L}} : Y_{n,l}(\sigma) \to U(\mathfrak{gl}_t) \otimes Y_{n,l-1}(\dot{\sigma}), \tag{6.9}$$

where $\dot{\sigma}$ is as in (4.6). Recalling Remark 5.4, these maps Δ_{R} and Δ_{L} are filtered, so induce homomorphisms

$$\operatorname{gr} \Delta_{\mathrm{R}} : \operatorname{gr} Y_{n,l}(\sigma) \to \operatorname{gr}(Y_{n,l-1}(\dot{\sigma}) \otimes U(\mathfrak{gl}_t)), \tag{6.10}$$

$$\operatorname{gr} \Delta_{\mathrm{L}} : \operatorname{gr} Y_{n,l}(\sigma) \to \operatorname{gr}(U(\mathfrak{gl}_t) \otimes Y_{n,l-1}(\dot{\sigma})) \tag{6.11}$$

of graded algebras.

Theorem 6.2. *For any admissible shape $\nu = (\nu_1, \dots, \nu_m)$, $\operatorname{gr} Y_{n,l}(\sigma)$ is the free commutative algebra on generators $\{e_{a,b;i,j}^{(r)}\}_{1 \le a,b \le m, 1 \le i \le \nu_a, 1 \le j \le \nu_b, s_{a,b}(\nu) < r \le s_{a,b}(\nu) + p_{\min(a,b)}(\nu)}$. Moreover, the maps $\operatorname{gr} \Delta_{\mathrm{R}}$ and $\operatorname{gr} \Delta_{\mathrm{L}}$ from (6.10)–(6.11) are injective whenever they are defined, hence so are the maps (6.8)–(6.9)*

Proof. We proceed by induction on the level l, the case $l = 0$ being trivial. Now suppose that $l > 0$ and that the first statement of the theorem has been proved for all smaller l. In view of Lemma 6.1, it suffices to check the induction step in the special case that $\nu = (\nu_1, \dots, \nu_m)$ is the minimal admissible shape for σ. At least one of the maps Δ_{R} or Δ_{L} is always defined. We will assume without loss of generality that Δ_{R} is defined; the theorem in the other case can be deduced from this one using the anti-isomorphism τ. Introduce the following elements of $\operatorname{gr}(Y_{n,l-1}(\dot{\sigma}) \otimes U(\mathfrak{gl}_t))$:

$$\dot{e}_{a,b;i,j}^{(r)} := \begin{cases} \operatorname{gr}_r \dot{D}_{a;i,j}^{(r)} \otimes 1 & \text{if } a = b, \\ \operatorname{gr}_r \dot{E}_{a,b;i,j}^{(r)} \otimes 1 & \text{if } a < b, \\ \operatorname{gr}_r \dot{F}_{a,b;i,j}^{(r)} \otimes 1 & \text{if } a > b. \end{cases}$$

Also let $x_{i,j} := \operatorname{gr}_1 1 \otimes e_{i,j}$ for $1 \le i,j \le t$. By Theorem 4.1(i), Lemma 4.2(i) and Lemma 6.1, there exist polynomials $f_{a;i,j}^{(r)}$ in all the variables $\dot{e}_{a,b;i,j}^{(r)}$ such that $\operatorname{gr} \Delta_{\mathrm{R}}$ maps

$$e_{a,b;i,j}^{(r)} \mapsto \dot{e}_{a,b;i,j}^{(r)} \tag{6.12}$$

for $1 \leq a \leq m, 1 \leq b < m, 1 \leq i \leq \nu_a, 1 \leq j \leq \nu_b$ and $s_{a,b}(\nu) < r \leq s_{a,b}(\nu)+p_{\min(a,b)}(\nu)$, and

$$e^{(r)}_{a,m;i,j} \mapsto \sum_{k=1}^{t} \dot{e}^{(r-1)}_{a,m;i,k} x_{k,j} + f^{(r)}_{a;i,j} \tag{6.13}$$

for $1 \leq a \leq m, 1 \leq i \leq \nu_a, 1 \leq j \leq \nu_m$ and $s_{a,m}(\nu) < r \leq s_{a,m}(\nu) + p_a(\nu)$, where $\dot{e}^{(0)}_{m,m;i,j} := \delta_{i,j}$. By the induction hypothesis and the PBW theorem for $U(\mathfrak{gl}_t)$, the elements

$$\{x_{i,j}\}_{1\leq i,j\leq t},$$
$$\{\dot{e}^{(r)}_{a,b;i,j}\}_{1\leq a\leq m, 1\leq b<m, 1\leq i\leq \nu_a, 1\leq j\leq \nu_b, s_{a,b}(\nu)<r\leq s_{a,b}(\nu)+p_{\min(a,b)}(\nu)},$$
$$\{\dot{e}^{(r)}_{a,m;i,j}\}_{1\leq a\leq m, 1\leq i\leq \nu_a, 1\leq j\leq \nu_m, s_{a,m}(\nu)+\delta_{a,m}-1<r\leq s_{a,m}(\nu)+p_a(\nu)-1}$$

and are algebraically independent in $\operatorname{gr}(Y_{n,l-1}(\dot{\sigma}) \otimes U(\mathfrak{gl}_t))$. Using this and (6.12)–(6.13), one verifies explicitly that the images of the generators

$$\{e^{(r)}_{a,b;i,j}\}_{1\leq a,b\leq m, 1\leq i\leq \nu_a, 1\leq j\leq \nu_b, s_{a,b}(\nu)<r\leq s_{a,b}(\nu)+p_{\min(a,b)}(\nu)}$$

of $\operatorname{gr} Y_{n,l}(\sigma)$ from Lemma 6.1 under the map $\operatorname{gr}\Delta_{\mathrm{R}}$ are algebraically independent in $\operatorname{gr} Y_{n,l-1}(\dot{\sigma}) \otimes U(\mathfrak{gl}_t)$. Hence $\operatorname{gr}\Delta_{\mathrm{R}}$ is injective and these generators must already be algebraically independent in $\operatorname{gr} Y_{n,l}(\sigma)$. This completes the proof of the induction step. □

Corollary 6.3. *For any admissible shape $\nu = (\nu_1, \dots, \nu_m)$, the monomials in the elements*

$$\{D^{(r)}_{a;i,j}\}_{1\leq a\leq m, 1\leq i,j\leq \nu_a, 0<r\leq p_a(\nu)},$$
$$\{E^{(r)}_{a,b;i,j}\}_{1\leq a<b\leq m, 1\leq i\leq \nu_a, 1\leq j\leq \nu_b, s_{a,b}(\nu)<r\leq s_{a,b}(\nu)+p_a(\nu)},$$
$$\{F^{(r)}_{a,b;i,j}\}_{1\leq a<b\leq m, 1\leq j\leq \nu_a, 1\leq i\leq \nu_b, s_{b,a}(\nu)<r\leq s_{b,a}(\nu)+p_a(\nu)}$$

taken in any fixed order form a basis for $Y_{n,l}(\sigma)$.

Remark 6.4. Obviously by their definition there is an inverse system

$$Y_{n,l}(\sigma) \twoheadleftarrow Y_{n,l+1}(\sigma) \twoheadleftarrow Y_{n,l+2}(\sigma) \twoheadleftarrow \cdots$$

Moreover, the maps are homomorphisms of filtered algebras with respect to the canonical filtrations. Comparing the basis theorems proved in Corollary 6.3 and Theorem 3.2(iv), it follows that $Y_n(\sigma) = \varprojlim Y_{n,l}(\sigma)$ where the inverse limit is taken in the category of filtered algebras. Hence we can view the shifted Yangian $Y_n(\sigma)$ as the limiting case $l \to \infty$ of the shifted Yangians $Y_{n,l}(\sigma)$ of level l.

Remark 6.5. Corollary 6.3 implies in particular that the Yangian $Y_{n,1}$ of level 1 may be identified with the universal enveloping algebra $U(\mathfrak{gl}_n)$ so that, for $1 \leq i, j \leq n$, $T^{(1)}_{i,j} \in Y_{n,1}$ is identified with the matrix unit $e_{i,j} \in \mathfrak{gl}_n$ and $T^{(r)}_{i,j} = 0$ for $r > 1$. Using (3.24)–(3.25), it is easy to describe the power series $C_n(u) = \sum_{r\geq 0} C^{(r)}_n u^{-r} \in Y_{n,1}[[u^{-1}]]$ explicitly under this identification; see also [MNO, Remark 2.11]: we have that

$$u(u-1)\cdots(u-n+1)C_n(u) = Z_n(u) \tag{6.14}$$

where

$$Z_n(u) = \mathrm{rdet}\begin{pmatrix} e_{1,1}+u-n+1 & \cdots & e_{1,n-1} & e_{1,n} \\ \vdots & \ddots & \vdots & \vdots \\ e_{n-1,1} & \cdots & e_{n-1,n-1}+u-1 & e_{n-1,n} \\ e_{n,1} & \cdots & e_{n,n-1} & e_{n,n}+u \end{pmatrix} \tag{6.15}$$

$$= \mathrm{cdet}\begin{pmatrix} e_{1,1}+u & e_{1,2} & \cdots & e_{1,n} \\ e_{2,1} & e_{2,2}+u-1 & \cdots & e_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ e_{n,1} & e_{n,2} & \cdots & e_{n,n}+u-n+1 \end{pmatrix}. \tag{6.16}$$

This proves the well known fact that the coefficients $Z_n^{(1)}, \dots, Z_n^{(n)}$ of the series $Z_n(u) = \sum_{r=0}^n Z_n^{(r)} u^{n-r} \in U(\mathfrak{gl}_n)[u]$ belong to $Z(U(\mathfrak{gl}_n))$; see also [CL, §2.2] where this is deduced from the classical Capelli identity or [HU, Appendix A.1] for a self-contained proof. By considering the images of $Z_n^{(1)}, \dots, Z_n^{(n)}$ under the Harish-Chandra homomorphism, one shows moreover that $Z_n^{(1)}, \dots, Z_n^{(n)}$ are algebraically independent and generate all of $Z(U(\mathfrak{gl}_n))$.

Remark 6.6. If $p_1 = 0$, i.e. the level l is *equal to* the total difference $d_1 + \cdots + d_{n-1}$, then we can make n smaller as follows. Let $\nu = (\nu_1, \dots, \nu_m)$ be an admissible shape for σ. Define $\dot\sigma$ to be the $(n-\nu_1)\times(n-\nu_1)$ shift matrix $(s_{i,j})_{\nu_1 < i,j \leq n}$. Note $\dot\nu := (\nu_2, \dots, \nu_m)$ is an admissible shape for $\dot\sigma$. We claim that there is an isomorphism

$$Y_{n,l}(\sigma) \xrightarrow{\sim} Y_{n-\nu_1,l}(\dot\sigma) \tag{6.17}$$

mapping $D_{1;i,j}^{(r)}, E_{1,b;i,j}^{(r)}$ and $F_{1,b;i,j}^{(r)}$ to zero for $1 < b \leq m$, $D_{a;i,j}^{(r)}$ to $\dot D_{a-1;i,j}^{(r)}$ for $1 < a \leq m$, and $E_{a,b;i,j}^{(r)}, F_{a,b;i,j}^{(r)}$ to $\dot E_{a-1,b-1;i,j}^{(r)}, \dot F_{a-1,b-1;i,j}^{(r)}$ respectively for $1 < a < b \leq m$. Here, we are working in terms of parabolic generators with respect to the shape ν in $Y_{n,l}(\sigma)$ and the shape $\dot\nu$ in $Y_{n-\nu_1,l}(\dot\sigma)$. In particular, taking ν to be the minimal admissible shape, this isomorphism allows us always to reduce to the situation that *either* $n = 0$ *or* $p_1 > 0$. Note also using (3.26) that the map (6.17) sends the generating function $C_n(u) \in Y_{n,l}(\sigma)[[u^{-1}]]$ to $\dot C_{n-\nu_1}(u - \nu_1) \in Y_{n-\nu_1,l}(\dot\sigma)[[u^{-1}]]$.

To prove the claim, it is easier to construct the inverse map. Note first using the relations (6.5)–(6.6) that $D_{1;i,j}^{(r)} = E_{1,b;i,j}^{(r)} = F_{1,b;i,j}^{(r)} = 0$ in $Y_{n,l}(\sigma)$ for $1 < b \leq m$ and all admissible i, j, r. Hence by the relation (6.7) we have that $D_{2;i,j}^{(r)} = 0$ for $r > p_2(\nu)$ and all i, j. Now there is obviously a homomorphism $Y_{n-\nu_1}(\dot\sigma) \to Y_n(\sigma)$ such that $\dot D_{a;i,j}^{(r)} \mapsto D_{a+1;i,j}^{(r)}, \dot E_{a,b;i,j}^{(r)} \mapsto E_{a+1,b+1;i,j}^{(r)}$ and $\dot F_{a,b;i,j}^{(r)} \mapsto F_{a+1,b+1;i,j}^{(r)}$. This factors through the quotients to induce the desired map $Y_{n-\nu_1,l}(\dot\sigma) \to Y_{n,l}(\sigma)$. Finally this map is an isomorphism by Corollary 6.3.

7. Pyramids

In this section we introduce the combinatorics of pyramids. (In the more general language of [EK, §4] the things we call pyramids here should be called even pyramids.) Suppose to start with that we are given a tuple $(q_1, q_2, \dots, q_l)$ of positive integers for some $l \geq 0$. We associate a diagram π consisting of q_1 bricks stacked in the first (leftmost) column, q_2 bricks stacked in the second column, ..., q_l bricks stacked in the lth (rightmost) column. For instance if $l = 4$ and $(q_1, q_2, q_3, q_4) = (4, 2, 3, 3)$, then

$$\pi = \begin{array}{cccc} 1 & & & \\ 2 & & 7 & 10 \\ 3 & 5 & 8 & 11 \\ 4 & 6 & 9 & 12 \end{array}$$

Call π a *pyramid* of *level* l and *height* $\max(q_1, \dots, q_l)$ if each row of the diagram consists of a single connected horizontal strip. Equivalently, π is a pyramid if the sequence $(q_1, q_2, \dots, q_l)$ of *column heights* is a unimodal sequence, i.e.

$$0 < q_1 \leq \cdots \leq q_k, \quad q_{k+1} \geq \cdots \geq q_l > 0 \tag{7.1}$$

for some $0 \leq k \leq l$. Of course, the above diagram is *not* a pyramid, since there is a gap between the entries 2 and 7.

Given a diagram π (not necessarily a pyramid), we pick an integer $n \geq \max(q_1, \dots, q_l)$ and number the rows of the diagram $1, 2, \dots, n$ from top to bottom. Let p_i denote the number of bricks in the ith row, thus defining the tuple $(p_1, \dots, p_n)$ of *row lengths* with

$$0 \leq p_1 \leq \cdots \leq p_n = l. \tag{7.2}$$

Usually we have in mind that n should exactly equal $\max(q_1, \dots, q_l)$, so that *either* $n = 0$ *or* $p_1 > 0$, but it is sometimes useful to allow n to be larger (cf. Remark 6.6). Fix also some numbering of the bricks of the diagram by $1, 2, \dots, N$. Usually we have in mind the numbering down columns from left to right as in the above example but any bijective numbering will do. For $i = 1, \dots, N$, let $\mathrm{row}(i)$ and $\mathrm{col}(i)$ denote the *row* and *column* numbers of the brick in which i appears, respectively.

Now let $\mathfrak{g}$ denote the Lie algebra $\mathfrak{gl}_N$ over $\mathbb{C}$ and introduce a $\mathbb{Z}$-grading $\mathfrak{g} = \bigoplus_{j \in \mathbb{Z}} \mathfrak{g}_j$ of $\mathfrak{g}$ by declaring that the ij-matrix unit $e_{i,j}$ is of degree $(\mathrm{col}(j) - \mathrm{col}(i))$. Let

$$\mathfrak{p} := \bigoplus_{j \geq 0} \mathfrak{g}_j, \qquad \mathfrak{h} := \mathfrak{g}_0, \qquad \mathfrak{m} := \bigoplus_{j < 0} \mathfrak{g}_j. \tag{7.3}$$

Thus $\mathfrak{p}$ is a parabolic subalgebra of $\mathfrak{g}$ with Levi factor $\mathfrak{h} \cong \mathfrak{gl}_{q_1} \oplus \cdots \oplus \mathfrak{gl}_{q_l}$, and $\mathfrak{m}$ is the nilradical of the opposite parabolic to $\mathfrak{p}$ with respect to $\mathfrak{h}$. Let P, H and M be the corresponding closed subgroups of the algebraic group $G := GL_N$ over $\mathbb{C}$. Also let

$$e := \sum_{\substack{1 \leq i,j \leq N \\ \mathrm{row}(i) = \mathrm{row}(j) \\ \mathrm{col}(i) = \mathrm{col}(j) - 1}} e_{i,j}. \tag{7.4}$$

By a dimension calculation, one checks that e belongs to the unique dense orbit of H on $\mathfrak{g}_1$, regardless of whether the diagram π is a pyramid or not. Moreover, the Jordan block sizes of the matrix e are precisely the lengths of the maximal connected horizontal strips in the diagram π, hence e is of Jordan type $(p_1, \dots, p_n)$ if and only if the diagram π is a pyramid.

Now recall from [Ca, §5.2] that an element e of the nilradical of $\mathfrak{p}$ is called a *Richardson element for* $\mathfrak{p}$ if its orbit under the adjoint action of P is dense in the nilradical of $\mathfrak{p}$; equivalently,

$$\dim \mathfrak{c}_{\mathfrak{g}}(e) = \dim \mathfrak{h}. \tag{7.5}$$

By another dimension calculation (as observed originally by Kraft [Kr]) the Jordan type of a Richardson element for $\mathfrak{p}$ is given by the row lengths $(p_1, \dots, p_n)$ of the diagram π, again regardless of whether π is a pyramid or not. We say that $\mathfrak{p}$ is a *good parabolic* if $\mathfrak{g}_1$ contains a Richardson element for $\mathfrak{p}$. Such an element e then clearly belongs both to the dense orbit of P on the nilradical of $\mathfrak{p}$ and to the dense orbit of H on $\mathfrak{g}_1$. Hence up to conjugacy, we may assume that e given by (7.4). Since its Jordan type must also be $(p_1, \dots, p_n)$, we obtain the following special case of [L, Lemma 7.2]:

Theorem 7.1. *$\mathfrak{p}$ is a good parabolic if and only if the diagram π is a pyramid.*

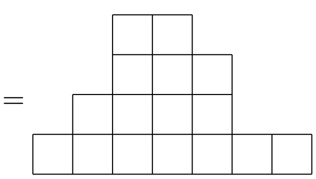
Remark 7.2. In view of [EK, Theorem 2.1], this theorem implies that there is a bijective map from pyramids to conjugacy classes of even good gradings of $\mathfrak{g} = \mathfrak{gl}_N$, as defined in the introduction. The map sends a pyramid π to twice the grading defined here, which is an even good grading for the nilpotent matrix e defined by (7.4).

To make the connection with the earlier sections, we point out that pyramids provide an extremely convenient way to visualize the data needed to define the shifted Yangian $Y_{n,l}(\sigma)$ of level l. Indeed, given a shift matrix $\sigma = (s_{i,j})_{1 \le i,j \le n}$ and a level $l \ge s_{1,n} + s_{n,1}$, we let $\delta = (d_1, \dots, d_{n-1})$ be the associated differences and define $(p_1, \dots, p_n)$ from $p_i := l - d_i - \dots - d_{n-1}$ like in (6.1). Now draw a pyramid π with p_i bricks on the ith row indented $s_{n,i}$ columns from the left hand edge for each $i = 1, \dots, n$. Conversely, given a pyramid π of height $\le n$, define a shift matrix $\sigma = (s_{i,j})_{1 \le i,j \le n}$ from the equation

$$s_{i,j} = \begin{cases} \#\{c = 1, \dots, k \mid i > n - q_c \ge j\} & \text{if } i \ge j, \\ \#\{c = k+1, \dots, l \mid i \le n - q_c < j\} & \text{if } i \le j, \end{cases} \tag{7.6}$$

where $(q_1, \dots, q_l)$ are the column heights and k is chosen as in (7.1). This definition is independent of the choice of k *only if* the pyramid π is of height *exactly* n, in which case $s_{i,j}$ is simply the number of bricks the ith row is indented from the jth row at the left edge of the diagram if $i \ge j$, at the right edge of the diagram if $i \le j$. For example,

$$l = 7, \sigma = \begin{pmatrix} 0 & 1 & 1 & 3 \\ 0 & 0 & 0 & 2 \\ 1 & 1 & 0 & 2 \\ 2 & 2 & 1 & 0 \end{pmatrix} \qquad \longleftrightarrow \qquad n = 4, \pi =$$



Now we have a convenient notation to record the explicit description of the centralizer $\mathfrak{c}_{\mathfrak{g}}(e)$ of the nilpotent element e associated to a pyramid π; see [SS, IV.1.6].

Lemma 7.3. *Let π be a pyramid of height $\le n$ with row lengths $(p_1, \dots, p_n)$ and associated shift matrix $\sigma = (s_{i,j})_{1 \le i,j \le n}$, and let e be the nilpotent matrix defined by (7.4). For $1 \le i, j \le n$ and $r \ge 0$, let*

$$c_{i,j}^{(r)} := \sum_{\substack{1 \le h,k \le N \\ \mathrm{row}(h)=i, \mathrm{row}(k)=j \\ \mathrm{col}(k)-\mathrm{col}(h)+1=r}} e_{h,k}.$$

Then, the vectors $\{c_{i,j}^{(r)}\}_{1 \le i,j \le n, s_{i,j} < r \le s_{i,j} + p_{\min(i,j)}}$ give a basis for $\mathfrak{c}_{\mathfrak{g}}(e)$.

8. Finite W-algebras

Continue with $\mathfrak{g} = \mathfrak{gl}_N$ and $G = GL_N$ acting on $\mathfrak{g}$ by the adjoint action Ad. Fixing a pyramid π with bricks numbered $1, \dots, N$, we have the associated $\mathbb{Z}$-grading on $\mathfrak{g}$ defined as in §7 so that $e_{i,j}$ is of degree $(\mathrm{col}(j) - \mathrm{col}(i))$. Also define $\mathfrak{p}, \mathfrak{h}$, $\mathfrak{m}$ and the nilpotent matrix $e \in \mathfrak{g}_1$ according to (7.3)–(7.4). Define $h \in \mathfrak{g}_0$ and $f \in \mathfrak{g}_{-1}$ to be the unique matrices so that (e, h, f) is an $\mathfrak{sl}_2$-triple in $\mathfrak{g}$. Let $\mathfrak{m}^\perp$ and $\mathfrak{p}^\perp$ denote the orthogonal complements of $\mathfrak{m}$ and $\mathfrak{p}$ with respect to the trace form $(.,.)$ on $\mathfrak{g}$, respectively, i.e. $\mathfrak{m}^\perp = \bigoplus_{j \leq 0} \mathfrak{g}_j$ and $\mathfrak{p}^\perp = \bigoplus_{j > 0} \mathfrak{g}_j$.

Lemma 8.1. *In the above notation,*

(i) $\mathfrak{m}^\perp = [\mathfrak{m}, e] \oplus \mathfrak{c}_\mathfrak{g}(f)$;
(ii) $\mathfrak{p} = [\mathfrak{p}^\perp, f] \oplus \mathfrak{c}_\mathfrak{g}(e)$;
(iii) $\mathfrak{c}_\mathfrak{g}(f)^\perp = \mathfrak{m} \oplus [\mathfrak{p}^\perp, f]$.

Proof. Since e is a Richardson element for $\mathfrak{p}$, we know that $\dim \mathfrak{c}_\mathfrak{g}(e) = \dim \mathfrak{h}$ and $\mathfrak{c}_\mathfrak{g}(e) \subseteq \mathfrak{p}$. Hence the map $\mathfrak{m} \to [\mathfrak{m}, e], x \mapsto [x, e]$ is a bijection. Similarly, $\dim \mathfrak{c}_\mathfrak{g}(f) = \dim \mathfrak{h}$ and $\mathfrak{c}_\mathfrak{g}(f) \subseteq \mathfrak{m}^\perp$. Hence the map $\mathfrak{p}^\perp \to [\mathfrak{p}^\perp, f], y \mapsto [y, f]$ is a bijection. So

$$\dim[\mathfrak{m}, e] + \dim \mathfrak{c}_\mathfrak{g}(f) = \dim \mathfrak{m} + \dim \mathfrak{h} = \dim \mathfrak{m}^\perp,$$
$$\dim[\mathfrak{p}^\perp, f] + \dim \mathfrak{c}_\mathfrak{g}(e) = \dim \mathfrak{p}^\perp + \dim \mathfrak{h} = \dim \mathfrak{p}.$$

Also by $\mathfrak{sl}_2$ theory, $[\mathfrak{m}, e] \cap \mathfrak{c}_\mathfrak{g}(f) = [\mathfrak{p}^\perp, f] \cap \mathfrak{c}_\mathfrak{g}(e) = 0$. Parts (i) and (ii) follow. For (iii), we have that

$$\dim \mathfrak{m} + \dim[\mathfrak{p}^\perp, f] = \dim \mathfrak{g} - \dim \mathfrak{h} = \dim \mathfrak{c}_\mathfrak{g}(f)^\perp,$$

and clearly $\mathfrak{m} \cap [\mathfrak{p}^\perp, f] = 0$, so we just have to check that $\mathfrak{m} \subseteq \mathfrak{c}_\mathfrak{g}(f)^\perp$ and that $[\mathfrak{p}^\perp, f] \subseteq \mathfrak{c}_\mathfrak{g}(f)^\perp$. The former statement is true since $\mathfrak{c}_\mathfrak{g}(f) \subseteq \mathfrak{m}^\perp$. For the latter, note for any $x \in \mathfrak{p}^\perp$ and $y \in \mathfrak{c}_\mathfrak{g}(f)$ that $([x, f], y) = (x, [f, y]) = 0$ by the invariance of the trace form. □

To recall the definition of the algebra $W(\pi)$ from the introduction, let $\chi : \mathfrak{m} \to \mathbb{C}$ denote the representation mapping $x \mapsto (x, e)$ for each $x \in \mathfrak{m}$. Let I_χ denote the kernel of the corresponding algebra homomorphism $U(\mathfrak{m}) \to \mathbb{C}$ and $\mathrm{pr}_\chi : U(\mathfrak{g}) \to U(\mathfrak{p})$ be the projection along the direct sum decomposition

$$U(\mathfrak{g}) = U(\mathfrak{p}) \oplus U(\mathfrak{g}) I_\chi. \tag{8.1}$$

For $x \in \mathfrak{m}$ and $y \in U(\mathfrak{p})$, we set $x \cdot y := \mathrm{pr}_\chi([x, y]) = \mathrm{pr}_\chi(xy) - \chi(x)y$, to define the *twisted action* of $\mathfrak{m}$ on $U(\mathfrak{p})$. Then, the *finite W-algebra* $W(\pi)$ associated to the pyramid π denotes the space $U(\mathfrak{p})^\mathfrak{m}$ of twisted $\mathfrak{m}$-invariants in $U(\mathfrak{p})$. Equivalently,

$$W(\pi) = \{y \in U(\mathfrak{p}) \mid (x - \chi(x))y \in U(\mathfrak{g}) I_\chi \text{ for all } x \in \mathfrak{m}\}, \tag{8.2}$$

from which it follows easily that $W(\pi)$ is actually a subalgebra of $U(\mathfrak{p})$. For example, in the special case that π consists of a single column, we obviously have that $\mathfrak{p} = \mathfrak{g}$, $\mathfrak{m} = 0$ and $e = 0$, hence $W(\pi) = U(\mathfrak{gl}_N)$.

We need the *Kazhdan filtration* $\cdots \subseteq \mathrm{F}_d U(\mathfrak{g}) \subseteq \mathrm{F}_{d+1} U(\mathfrak{g}) \subseteq \cdots$ of $U(\mathfrak{g})$ defined by declaring that the generator $e_{i,j}$ is of degree

$$\deg(e_{i,j}) := \mathrm{col}(j) - \mathrm{col}(i) + 1 \tag{8.3}$$

for each $1 \leq i, j \leq N$, i.e. $\mathrm{F}_d U(\mathfrak{g})$ is spanned by all monomials $e_{i_1,j_1} \cdots e_{i_m,j_m}$ for $m \geq 0$ and $\deg(e_{i_1,j_1}) + \cdots + \deg(e_{i_m,j_m}) \leq d$. The adjoint action of $\mathfrak{g}$ on $U(\mathfrak{g})$ is filtered in the sense that $[\mathfrak{g}_j, \mathrm{F}_d U(\mathfrak{g})] \subseteq \mathrm{F}_{d+j} U(\mathfrak{g})$ for each $j, d \in \mathbb{Z}$. Hence the associated graded algebra $\operatorname{gr} U(\mathfrak{g})$ is naturally a graded $\mathfrak{g}$-module. Of course by the PBW theorem we can identify $\operatorname{gr} U(\mathfrak{g})$ with the symmetric algebra $S(\mathfrak{g})$ with the usual $\mathfrak{g}$-action, but viewed as a graded algebra via the *Kazhdan grading* also defined by (8.3).

There are induced Kazhdan filtrations of the subalgebras $U(\mathfrak{p})$ and $W(\pi)$ defined by setting $\mathrm{F}_d U(\mathfrak{p}) := \mathrm{pr}_\chi(\mathrm{F}_d U(\mathfrak{g}))$ and $\mathrm{F}_d W(\pi) := W(\pi) \cap \mathrm{F}_d U(\mathfrak{p})$. This time we have that $\mathrm{F}_d U(\mathfrak{p}) = \mathrm{F}_d W(\pi) = 0$ for all $d < 0$. The projection $\mathrm{pr}_\chi : U(\mathfrak{g}) \to U(\mathfrak{p})$ is filtered, hence so is the twisted action of $\mathfrak{m}$ on $U(\mathfrak{p})$, and $\operatorname{gr} \mathrm{pr}_\chi : \operatorname{gr} U(\mathfrak{g}) \to \operatorname{gr} U(\mathfrak{p})$ is a graded $\mathfrak{m}$-module homomorphism. If we identify $\operatorname{gr} U(\mathfrak{p})$ with $S(\mathfrak{p})$ graded via (8.3), then we can describe the map $\operatorname{gr} \mathrm{pr}_\chi$ simply as the algebra homomorphism $p : S(\mathfrak{g}) \to S(\mathfrak{p})$ that acts as the identity on elements of $\mathfrak{p}$ and sends $x \in \mathfrak{m}$ to $\chi(x)$. This explains the top left hand quadrant of the following commutative diagram:

$$\begin{array}{ccccccc}
\operatorname{gr} U(\mathfrak{g}) & = & S(\mathfrak{g}) & \xrightarrow[\alpha]{\sim} & \mathbb{C}[\mathfrak{g}] & & \\
{\scriptstyle \operatorname{gr}\mathrm{pr}_\chi}\downarrow & & {\scriptstyle p}\downarrow & & {\scriptstyle r}\downarrow & & \\
\operatorname{gr} U(\mathfrak{p}) & = & S(\mathfrak{p}) & \xrightarrow[\beta]{\sim} & \mathbb{C}[e+\mathfrak{m}^\perp] & \hookleftarrow & \mathbb{C}[e+\mathfrak{m}^\perp]^M \\
{\scriptstyle \operatorname{gr} i}\uparrow & & {\scriptstyle q}\downarrow & & {\scriptstyle s}\downarrow & \swarrow{\scriptstyle \sim} & \\
\operatorname{gr} W(\pi) & \xrightarrow[\theta]{} & S(\mathfrak{c}_\mathfrak{g}(e)) & \xrightarrow[\gamma]{\sim} & \mathbb{C}[e+\mathfrak{c}_\mathfrak{g}(f)] & &
\end{array} \tag{8.4}$$

To make sense of the bottom left hand quadrant, let $i : W(\pi) \to U(\mathfrak{p})$ be the inclusion, and let $q : S(\mathfrak{p}) \to S(\mathfrak{c}_\mathfrak{g}(e))$ be the algebra homomorphism induced by the projection $\mathfrak{p} \to \mathfrak{c}_\mathfrak{g}(e)$ along the direct sum decomposition from Lemma 8.1(ii). Note both $[\mathfrak{p}^\perp, f]$ and $\mathfrak{c}_\mathfrak{g}(e)$ are graded subspaces of $\mathfrak{g}$; in the latter case one way to see this is by Lemma 7.3 since the basis element $c_{i,j}^{(r)}$ there is clearly of degree r with respect to the Kazhdan grading. Hence the map q is a graded map. Finally let $\theta : \operatorname{gr} W(\pi) \to S(\mathfrak{c}_\mathfrak{g}(e))$ be the composite $q \circ \operatorname{gr} i$.

Now we turn our attention to the right hand half of the diagram. Imitating [GG, §2.1], let $\gamma : \mathbb{C}^\times \to G$ be the group homomorphism defined by letting

$$\gamma(t) := \operatorname{diag}(t^{-\operatorname{col}(1)}, t^{-\operatorname{col}(2)}, \cdots, t^{-\operatorname{col}(N)}) \in GL_N \tag{8.5}$$

for each $t \in \mathbb{C}^\times$. So $\operatorname{Ad}\gamma(t)$ acts on $\mathfrak{g}_j$ by the scalar t^j for each $t \in \mathbb{C}^\times$ and $j \in \mathbb{Z}$. Introduce a linear action ρ of $\mathbb{C}^\times$ on the variety $\mathfrak{g}$ by letting

$$\rho(t)(x) := t^{-1} \operatorname{Ad}\gamma(t)(x) \tag{8.6}$$

for each $t \in \mathbb{C}^\times, x \in \mathfrak{g}$. Thus, $\rho(t)(e_{i,j}) = t^{-\deg(e_{j,i})} e_{i,j}$ for each $1 \leq i, j \leq N$. In particular, $\rho(t)(e) = e$ for every $t \in \mathbb{C}^\times$ and

$$\lim_{t \to \infty} \rho(t)(x) = 0 \tag{8.7}$$

for all $x \in \mathfrak{m}^\perp$. We get an induced action $\bar{\rho}$ of $\mathbb{C}^\times$ on the coordinate algebra $\mathbb{C}[\mathfrak{g}]$ with

$$(\bar{\rho}(t)(f))(x) = f(\rho(t^{-1})(x)) \tag{8.8}$$

for each $f \in \mathbb{C}[\mathfrak{g}], x \in \mathfrak{g}$. Using this, we define a grading on $\mathbb{C}[\mathfrak{g}]$ by declaring that $f \in \mathbb{C}[\mathfrak{g}]$ is of degree j if $\bar{\rho}(t)(f) = t^j f$ for each $t \in \mathbb{C}^\times$. Since e is a ρ-fixed point, it is easy to see that each $\rho(t)$ leaves both $e+\mathfrak{m}^\perp$ and the Slodowy slice $e+\mathfrak{c}_\mathfrak{g}(f)$ invariant. So just like for $\mathbb{C}[\mathfrak{g}]$, we get induced actions $\bar\rho$ of $\mathbb{C}^\times$ on the coordinate algebras $\mathbb{C}[e+\mathfrak{m}^\perp]$ and $\mathbb{C}[e+\mathfrak{c}_\mathfrak{g}(f)]$ which we use to introduce a grading on these algebras. The natural restriction maps $r:\mathbb{C}[\mathfrak{g}]\twoheadrightarrow\mathbb{C}[e+\mathfrak{m}^\perp]$ and $s:\mathbb{C}[e+\mathfrak{m}^\perp]\twoheadrightarrow\mathbb{C}[e+\mathfrak{c}_\mathfrak{g}(f)]$ are $\bar\rho$-equivariant, hence are graded.

The trace form defines an isomorphism $\alpha:S(\mathfrak{g})\to\mathbb{C}[\mathfrak{g}]$ of graded algebras. It maps the kernel of the homomorphism $p:S(\mathfrak{g})\to S(\mathfrak{p})$ isomorphically onto the annihilator in $\mathbb{C}[\mathfrak{g}]$ of the closed subvariety $e+\mathfrak{m}^\perp$. Hence α induces a graded algebra isomorphism $\beta:S(\mathfrak{p})\to\mathbb{C}[e+\mathfrak{m}^\perp]$. Similarly by Lemma 8.1(ii),(iii), β maps the kernel of the homomorphism $q:S(\mathfrak{p})\to S(\mathfrak{c}_\mathfrak{g}(e))$ isomorphically onto the annihilator in $\mathbb{C}[e+\mathfrak{m}^\perp]$ of $e+\mathfrak{c}_\mathfrak{g}(f)$, hence induces a graded algebra isomorphism $\gamma:S(\mathfrak{c}_\mathfrak{g}(e))\to\mathbb{C}[e+\mathfrak{c}_\mathfrak{g}(f)]$.

To complete the picture, we need to introduce an action of the subgroup M of G corresponding to the subalgebra $\mathfrak{m}$ of $\mathfrak{g}$. The adjoint action of G on $\mathfrak{g}$ induces an action of G on $\mathbb{C}[\mathfrak{g}]$ by algebra automorphisms, such that the derived action of $\mathfrak{g}$ corresponds under the isomorphism α to the usual $\mathfrak{g}$-action on $S(\mathfrak{g})$. The subgroup M of G leaves $e+\mathfrak{m}^\perp$ invariant, hence we get induced actions of M and $\mathfrak{m}$ on $\mathbb{C}[e+\mathfrak{m}^\perp]$, such that the action of $\mathfrak{m}$ agrees under the isomorphism β with the twisted action of $\mathfrak{m}$ on $S(\mathfrak{p})$. In particular since $S(\mathfrak{p})$ is a graded $\mathfrak{m}$-module, it follows that the space $\mathbb{C}[e+\mathfrak{m}^\perp]^\mathfrak{m}=\mathbb{C}[e+\mathfrak{m}^\perp]^M$ of $\mathfrak{m}$-invariants/M-fixed points is a *graded* subalgebra of $\mathbb{C}[e+\mathfrak{m}^\perp]$. One can see this directly by introducing an action ρ of $\mathbb{C}^\times$ on $M\times(e+\mathfrak{m}^\perp)$ defined by

$$\rho(t)(m,x):=(\gamma(t)m\gamma(t)^{-1},\rho(t)(x)) \tag{8.9}$$

for all $m\in M, x\in e+\mathfrak{m}^\perp$ and $t\in\mathbb{C}^\times$. The following calculation checks that the adjoint action $\varphi:M\times(e+\mathfrak{m}^\perp)\to e+\mathfrak{m}^\perp$ is ρ-equivariant:

$$\begin{aligned}\operatorname{Ad}(\gamma(t)m\gamma(t)^{-1})\rho(t)(x)&=t^{-1}\operatorname{Ad}\gamma(t)\operatorname{Ad}m\operatorname{Ad}\gamma(t)^{-1}\operatorname{Ad}\gamma(t)(x)\\&=t^{-1}\operatorname{Ad}\gamma(t)\operatorname{Ad}m(x)=\rho(t)(\operatorname{Ad}m(x)).\end{aligned}$$

This now implies that the space of M-fixed points in $\mathbb{C}[e+\mathfrak{m}^\perp]$ is invariant under each $\rho(t)$, hence is graded. Let us also note that

$$\lim_{t\to\infty}(\gamma(t)m\gamma(t)^{-1})=1 \tag{8.10}$$

for all $m\in M$.

It just remains to prove that the composite of the inclusion $\mathbb{C}[e+\mathfrak{m}^\perp]^M\hookrightarrow\mathbb{C}[e+\mathfrak{m}^\perp]$ and the projection $s:\mathbb{C}[e+\mathfrak{m}^\perp]\twoheadrightarrow\mathbb{C}[e+\mathfrak{c}_\mathfrak{g}(f)]$ is an isomorphism. This follows from the following key result, which is due originally to Kostant [K, Theorem 1.2], and is proved in this generality in [L, Theorem 1.2] using Zariski's Main Theorem. The alternative argument sketched here is due to Gan and Ginzburg; see [GG, Lemma 2.1].

Theorem 8.2. *The adjoint action $\varphi:M\times(e+\mathfrak{c}_\mathfrak{g}(f))\to e+\mathfrak{m}^\perp$ is an isomorphism of affine varieties.*

Proof. We just verify the hypothesis needed to apply the general result from the proof of [GG, Lemma 2.1]: *An equivariant morphism $\varphi:X_1\to X_2$ of smooth affine $\mathbb{C}^\times$-varieties with contracting $\mathbb{C}^\times$-actions which induces an isomorphism between the tangent spaces at the $\mathbb{C}^\times$-fixed points must be an isomorphism.* We have already defined actions ρ of $\mathbb{C}^\times$ on $e+\mathfrak{m}^\perp$ (8.6) and on $M\times(e+\mathfrak{c}_\mathfrak{g}(f))$ (8.9) and checked that φ is ρ-equivariant. By (8.7) and (8.10), we have that $\lim_{t\to\infty}\rho(t)(m,x)=(1,e)$ for each

$(m,x) \in M \times (e + \mathfrak{c}_{\mathfrak{g}}(f))$ and that $\lim_{t\to\infty} \rho(t)(x) = e$ for each $x \in e + \mathfrak{m}^\perp$. Hence the $\mathbb{C}^\times$-actions are both contracting. So finally we need to check that the differential $d\varphi_{(1,e)}$ is an isomorphism between the tangent spaces $T_{(1,e)}(M \times (e + \mathfrak{c}_{\mathfrak{g}}(f)))$ and $T_e(e+\mathfrak{m}^\perp)$. But if we identify the tangent spaces with $\mathfrak{m} \oplus \mathfrak{c}_{\mathfrak{g}}(f)$ and $\mathfrak{m}^\perp$ respectively, then the differential is the map $(x,y) \mapsto [x,e] + y$. Hence it is an isomorphism by Lemma 8.1(i). □

The crucial thing that we can now read off from the diagram (8.4) is the following:

Corollary 8.3. *The map $\theta : \operatorname{gr} W(\pi) \to S(\mathfrak{c}_{\mathfrak{g}}(e))$ is an injective graded algebra homomorphism.*

Proof. Clearly $\operatorname{gr} i$ maps $\operatorname{gr} W(\pi)$ injectively into the space of $\mathfrak{m}$-invariants in $\operatorname{gr} U(\mathfrak{p})$. Hence, $\beta \circ \operatorname{gr} i$ maps $\operatorname{gr} W(\pi)$ into $\mathbb{C}[e+\mathfrak{m}^\perp]^M$. Hence $s \circ \beta \circ \operatorname{gr} i$ is injective. But this is $\gamma \circ \theta$ by the commmutativity of the diagram, hence θ is injective too. □

Remark 8.4. It is known by [L, Theorem 2.3] that the map θ is actually an *isomorphism*, hence $\operatorname{gr} W(\pi)$ is isomorphic to the coordinate algebra $\mathbb{C}[e + \mathfrak{c}_{\mathfrak{g}}(f)]$ of the Slodowy slice as stated in the introduction. A quicker proof can also be given by following the arguments of [GG, §5]. However we do not need to use this fact yet, and we will be able deduce it later on as a consequence of the main result of the article; see Corollary 10.2.

Remark 8.5. The restriction of the map $\operatorname{pr}_\chi : U(\mathfrak{g}) \to U(\mathfrak{p})$ to $Z(U(\mathfrak{g}))$ defines an injective algebra homomorphism $\psi : Z(U(\mathfrak{g})) \hookrightarrow W(\pi)$ whose image is contained in the center $Z(W(\pi))$. The fact that ψ is an algebra homomorphism with image contained in $Z(W(\pi))$ is easiest to see using the definition of $W(\pi)$ as the endomorphism algebra $\operatorname{End}_{U(\mathfrak{g})}(Q_\chi)$ given in the introduction, since in those terms ψ is just the representation $Z(U(\mathfrak{g})) \to \operatorname{End}_{\mathbb{C}}(Q_\chi)$ of $Z(U(\mathfrak{g}))$ on the module Q_χ. The fact that ψ is injective is proved in [P, 6.2] or [L, Proposition 2.6] by observing that the (injective) Harish-Chandra homomorphism factors through the map ψ. In [BK2, §6] we show moreover using some basic facts about the representation theory of $W(\pi)$ (some of the proofs of which depend on knowing the main result of the article below) that the image of ψ is actually *equal to* $Z(W(\pi))$. Hence, $\psi : Z(U(\mathfrak{g})) \to Z(W(\pi))$ is actually an *isomorphism*.

9. Invariants

In this section we define some remarkable elements of $U(\mathfrak{p})$, many of which will eventually turn out to be $\mathfrak{m}$-invariant, i.e. to belong to the subalgebra $W(\pi)$. Letting $(q_1, \dots, q_l)$ denote the column heights of our fixed pyramid π, pick an integer $n \geq \max(q_1, \dots, q_l)$. Define $\rho = (\rho_1, \dots, \rho_N)$ by setting

$$\rho_i := n - q_{\operatorname{col}(i)} - q_{\operatorname{col}(i)+1} - \cdots - q_l. \tag{9.1}$$

For $1 \leq i, j \leq N$, define

$$\tilde{e}_{i,j} := (-1)^{\operatorname{col}(j)-\operatorname{col}(i)}(e_{i,j} + \delta_{i,j}\rho_i), \tag{9.2}$$

so

$$[\tilde{e}_{i,j}, \tilde{e}_{h,k}] = (\tilde{e}_{i,k} - \delta_{i,k}\rho_i)\delta_{h,j} - \delta_{i,k}(\tilde{e}_{h,j} - \delta_{h,j}\rho_j). \tag{9.3}$$

Let us also spell out the effect of the homorphism $U(\mathfrak{m}) \to \mathbb{C}$ induced by the character χ: we have that

$$\tilde{e}_{i,j} \mapsto \begin{cases} -1 & \text{if } \mathrm{row}(i) = \mathrm{row}(j) \text{ and } \mathrm{col}(i) = \mathrm{col}(j) + 1; \\ 0 & \text{otherwise.} \end{cases} \tag{9.4}$$

For $1 \leq i, j \leq n$ and signs $\sigma_1, \dots, \sigma_n \in \{\pm\}$, we let $T^{(0)}_{i,j;\sigma_1,\dots,\sigma_n} := \delta_{i,j}\sigma_i$ and for $r \geq 1$ define

$$T^{(r)}_{i,j;\sigma_1,\dots,\sigma_n} := \sum_{s=1}^{r} \sum_{\substack{i_1,\dots,i_s \\ j_1,\dots,j_s}} \sigma_{\mathrm{row}(i_2)} \cdots \sigma_{\mathrm{row}(i_s)} \tilde{e}_{i_1,j_1} \cdots \tilde{e}_{i_s,j_s} \tag{9.5}$$

where the second sum is over all $1 \leq i_1, \dots, i_s, j_1, \dots, j_s \leq N$ such that

(1) $\deg(e_{i_1,j_1}) + \cdots + \deg(e_{i_s,j_s}) = r$ (recall (8.3));
(2) $\mathrm{col}(i_t) \leq \mathrm{col}(j_t)$ for each $t = 1, \dots, s$;
(3) if $\sigma_{\mathrm{row}(j_t)} = +$ then $\mathrm{col}(j_t) < \mathrm{col}(i_{t+1})$ for each $t = 1, \dots, s-1$;
(4) if $\sigma_{\mathrm{row}(j_t)} = -$ then $\mathrm{col}(j_t) \geq \mathrm{col}(i_{t+1})$ for each $t = 1, \dots, s-1$;
(5) $\mathrm{row}(i_1) = i$, $\mathrm{row}(j_s) = j$;
(6) $\mathrm{row}(j_t) = \mathrm{row}(i_{t+1})$ for each $t = 1, \dots, s-1$.

Note the assumptions (1) and (2) imply that $T^{(r)}_{i,j;\sigma_1,\dots,\sigma_n}$ belongs to $\mathrm{F}_r U(\mathfrak{p})$. For $x = 0, 1, \dots, n$, let $T^{(r)}_{i,j;x}$ denote $T^{(r)}_{i,j;\sigma_1,\dots,\sigma_n}$ in the special case that $\sigma_1 = \cdots = \sigma_x = -$, $\sigma_{x+1} = \cdots = \sigma_n = +$. Define

$$T_{i,j;x}(u) := \sum_{r \geq 0} T^{(r)}_{i,j;x} u^{-r} \in U(\mathfrak{p})[[u^{-1}]]. \tag{9.6}$$

Since this is the most critical definition in the entire paper, let us give some simple examples.

Example 9.1. For any $1 \leq i, j \leq n$ and $x = 0, 1, \dots, n$,

$$T^{(1)}_{i,j;x} = \sum_{\substack{1 \leq h,k \leq N \\ \mathrm{row}(h)=i, \mathrm{row}(k)=j \\ \mathrm{col}(h)=\mathrm{col}(k)}} \tilde{e}_{h,k},$$

$$\begin{aligned} T^{(2)}_{i,j;x} = \sum_{\substack{1 \leq h,k \leq N \\ \mathrm{row}(h)=i, \mathrm{row}(k)=j \\ \mathrm{col}(h)=\mathrm{col}(k)-1}} \tilde{e}_{h,k} &- \sum_{\substack{1 \leq h_1,h_2,k_1,k_2 \leq N \\ \mathrm{row}(h_1)=i, \mathrm{row}(k_1)=\mathrm{row}(h_2) \leq x, \mathrm{row}(k_2)=j \\ \mathrm{col}(h_1)=\mathrm{col}(k_1) \geq \mathrm{col}(h_2)=\mathrm{col}(k_2),}} \tilde{e}_{h_1,k_1}\tilde{e}_{h_2,k_2} \\ &+ \sum_{\substack{1 \leq h_1,h_2,k_1,k_2 \leq N \\ \mathrm{row}(h_1)=i, \mathrm{row}(k_1)=\mathrm{row}(h_2) > x, \mathrm{row}(k_2)=j \\ \mathrm{col}(h_1)=\mathrm{col}(k_1) < \mathrm{col}(h_2)=\mathrm{col}(k_2),}} \tilde{e}_{h_1,k_1}\tilde{e}_{h_2,k_2}. \end{aligned}$$

Lemma 9.2. *Suppose $0 \leq x < y \leq n$.*

(i) *If $x < i \leq y$ and $y < j \leq n$ then*

$$T_{i,j;x}(u) = \sum_{k=x+1}^{y} T_{i,k;x}(u) T_{k,j;y}(u).$$

(ii) *If $y < i \leq n$ and $x < j \leq y$ then*

$$T_{i,j;x}(u) = \sum_{k=x+1}^{y} T_{i,k;y}(u)T_{k,j;x}(u).$$

(iii) *If $y < i, j \leq n$ then*

$$T_{i,j;x}(u) = T_{i,j;y}(u) + \sum_{k,l=x+1}^{y} T_{i,k;y}(u)T_{k,l;x}(u)T_{l,j;y}(u).$$

(iv) *If $x < i, j \leq y$ then*

$$\sum_{k=x+1}^{y} T_{i,k;x}(u)T_{k,j;y}(u) = -\delta_{i,j}.$$

Proof. Let $\xi_{i,j} := \tilde{e}_{i,j}u^{-\deg(e_{i,j})}$ for short.

(i) The right hand side of the formula in (i) is a sum of monomials of the form

$$\pm(\xi_{i_1,j_1} \cdots \xi_{i_r,j_r})(\xi_{k_1,l_1} \cdots \xi_{k_s,l_s}), \tag{9.7}$$

for various $r \geq 0, s \geq 1$, where $\pm\xi_{i_1,j_1} \cdots \xi_{i_r,j_r}$ appears in $T_{i,k;x}(u)$ and $\pm\xi_{k_1,l_1} \cdots \xi_{k_s,l_s}$ appears in $T_{k,j;y}(u)$ for some $x < k \leq y$. Let X be the sum of all such monomials for which $\mathrm{col}(j_r) < \mathrm{col}(k_1)$ if $r > 0$ and for which $\mathrm{row}(l_t) \notin \{x+1, \dots, y\}$ for all $1 \leq t < s$. Let Y be the sum of all remaining monomials. Thus, the right hand side of the formula in (i) is equal to $X + Y$. Now we proceed to show that $X = T_{i,j;x}(u)$ and that $Y = 0$.

First consider X. Take a monomial of the form (9.7) appearing in X, so $\mathrm{col}(j_r) < \mathrm{col}(k_1)$ if $r > 0$ and $\mathrm{row}(l_t) \notin \{x+1, \dots, y\}$ for all $1 \leq t < s$. It is easy to see that this monomial also appears in the expansion of $T_{i,j;x}(u)$, with the same sign. Moreover, the monomial (9.7) appears in X exactly once: otherwise we would be able to obtain this monomial in the expansion of X in another way by splitting it *either* as

$$\pm(\xi_{i_1,j_1} \cdots \xi_{i_r,j_r}\xi_{k_1,l_1} \cdots \xi_{k_t,l_t})(\xi_{k_{t+1},l_{t+1}} \cdots \xi_{k_s,l_s})$$

for some $1 \leq t < s$ so that, writing $h = \mathrm{row}(l_t)$, $\pm\xi_{i_1,j_1} \cdots \xi_{i_r,j_r}\xi_{k_1,l_1} \cdots \xi_{k_t,l_t}$ appears in $T_{i,h;x}(u)$ and $\pm\xi_{k_{t+1},l_{t+1}} \cdots \xi_{k_s,l_s}$ appears in $T_{h,j;y}(u)$, *or* as

$$\pm(\xi_{i_1,j_1} \cdots \xi_{i_{u-1},j_{u-1}})(\xi_{i_u,j_u} \cdots \xi_{i_r,j_r}\xi_{k_1,l_1} \cdots \xi_{k_s,l_s})$$

for some $1 \leq u \leq r$ so that, writing $h = \mathrm{row}(i_u)$, $\xi_{i_1,j_1} \cdots \xi_{i_{u-1},j_{u-1}}$ appears in $T_{i,h;x}(u)$ and $\pm\xi_{i_u,j_u} \cdots \xi_{i_r,j_r}\xi_{k_1,l_1} \cdots \xi_{k_s,l_s}$ appears in $T_{h,j;y}(u)$. The former case does not happen since we have that $h \notin \{x+1, \dots, y\}$ by assumption. The latter case does not happen since $\mathrm{col}(j_r) < \mathrm{col}(k_1)$ contrary to the definition of the monomials arising in $T_{h,j;y}(u)$. To complete the proof that $X = T_{i,j;x}(u)$, we need to show that every monomial $\pm\xi_{p_1,q_1} \cdots \xi_{p_u,q_u}$ appearing in $T_{i,j;x}(u)$ also appears in X. Take $r \geq 0$ to be the maximal index such that $\pm\xi_{p_1,q_1} \cdots \xi_{p_r,q_r}$ appears in $T_{i,k;x}(u)$ for some $x < k \leq y$; such an r exists as for $r = 0$ the monomial 1 appears in $T_{i,i;x}(u)$. It remains to observe that $\mathrm{col}(q_r) < \mathrm{col}(p_{r+1})$ and that $\mathrm{row}(q_t) \notin \{x+1, \dots, y\}$ for all $r < t < u$, for otherwise r could be made bigger. In particular this means that $\pm\xi_{p_{r+1},q_{r+1}} \cdots \xi_{p_u,q_u}$ appears in $T_{k,j;y}(u)$. Hence if we split our monomial as $\pm(\xi_{p_1,q_1} \cdots \xi_{p_r,q_r})(\xi_{p_{r+1},q_{r+1}} \cdots \xi_{p_u,q_u})$, we have something of the form (9.7) that appears in X.

Now consider Y. Take a monomial of the form (9.7) appearing in Y, i.e. either $r > 0$ and $\mathrm{col}(j_r) \geq \mathrm{col}(k_1)$ or there is some $1 \leq t < s$ such that $x < \mathrm{row}(l_t) \leq y$. We show that this monomial appears exactly twice in the expansion of Y, with opposite signs. There are two cases.

Suppose first that $r > 0$ and that $\mathrm{col}(j_r) \geq \mathrm{col}(k_1)$. Let $t \leq r$ be the maximal index such that $x < \mathrm{row}(i_t) \leq y$; such a t exists since $\mathrm{row}(i_1) = i$ and $x < i \leq y$. Let $h := \mathrm{row}(i_t)$. Then the monomial $\pm\xi_{i_1,j_1} \cdots \xi_{i_{t-1},j_{t-1}}$ appears in $T_{i,h;x}(u)$ and the monomial $\pm\xi_{i_t,j_t} \cdots \xi_{i_r,j_r}\xi_{k_1,l_1} \cdots \xi_{k_s,l_s}$ appears in $T_{h,j;y}(u)$. Moreover, using the facts that $\mathrm{col}(j_{t-1}) < \mathrm{col}(i_t)$ if $t > 1$, $\mathrm{col}(j_r) \geq \mathrm{col}(k_1)$, and the maximality of the choice of t, we see that

$$\pm(\xi_{i_1,j_1} \cdots \xi_{i_r,j_r})(\xi_{k_1,l_1} \cdots \xi_{k_s,l_s}), \;\; \pm(\xi_{i_1,j_1} \cdots \xi_{i_{t-1},j_{t-1}})(\xi_{i_t,j_t}, \cdots , \xi_{i_r,j_r}\xi_{k_1,l_1} \cdots \xi_{k_s,l_s})$$

are the only ways to split the monomial (9.7) so the left hand term occurs in $T_{i,g;x}(u)$ and the right hand term appears in $T_{g,j;y}(u)$ for some $x < g \leq y$. It just remains to check that the two have opposite signs.

Suppose instead that $\mathrm{col}(j_r) < \mathrm{col}(k_1)$ if $r > 0$ and that $x < \mathrm{row}(l_t) \leq y$ for some $1 \leq t < s$. Choose the minimal such t and let $h := \mathrm{row}(l_t)$. Then the monomial $\pm\xi_{i_1,j_1} \cdots \xi_{i_r,j_r}\xi_{k_1,l_1} \cdots \xi_{k_t,l_t}$ appears in $T_{i,h;x}(u)$ and the monomial $\pm\xi_{k_{t+1},l_{t+1}} \cdots \xi_{k_s,l_s}$ appears in $T_{h,j;y}(u)$. Moreover using the facts that $\mathrm{col}(j_r) < \mathrm{col}(k_1)$ if $r > 0$, $\mathrm{col}(l_t) \geq \mathrm{col}(k_{t+1})$, and the minimality of the choice of t, we have that

$$\pm(\xi_{i_1,j_1} \cdots \xi_{i_r,j_r})(\xi_{k_1,l_1} \cdots \xi_{k_s,l_s}), \;\; \pm(\xi_{i_1,j_1} \cdots \xi_{i_r,j_r}\xi_{k_1,l_1} \cdots \xi_{k_t,l_t})(\xi_{k_{t+1},l_{t+1}} \cdots \xi_{k_s,l_s})$$

are the only ways to split the monomial (9.7) so the first multiple occurs in $T_{i,g;x}(u)$ and the second multiple appears in $T_{g,j;y}(u)$ for some $x < g \leq y$. The two have opposite signs.

(ii) Similar.

(iii) Using (ii), we can rewrite (iii) as

$$T_{i,j;x}(u) - T_{i,j;y}(u) = \sum_{k=x+1}^{y} T_{i,k;x}(u)T_{k,j;y}(u). \tag{9.8}$$

The terms on the right hand side of (9.8) look like

$$\pm(\xi_{i_1,j_1} \cdots \xi_{i_r,j_r})(\xi_{k_1,l_1} \cdots \xi_{k_s,l_s}), \tag{9.9}$$

for various $r, s \geq 1$, where $\pm\xi_{i_1,j_1} \ldots \xi_{i_r,j_r}$ appears in $T_{i,k;x}(u)$ and $\pm\xi_{k_1,l_1} \ldots \xi_{k_s,l_s}$ appears in $T_{k,j;y}(u)$ for some $x < k \leq y$. Let X be the sum of all such monomials for which $\mathrm{col}(j_r) < \mathrm{col}(k_1)$ and $\mathrm{row}(l_t) \notin \{x+1, \ldots, y\}$ for all $1 \leq t < s$. Let Y be the sum of all such monomials for which $\mathrm{col}(j_r) \geq \mathrm{col}(k_1)$ and $\mathrm{row}(j_t) \notin \{x+1, \ldots, y\}$ for all $1 \leq t < r$. Let Z be the sum of all the remaining monomials, so the right hand side of (9.8) is equal to $X + Y + Z$. We will show that $X + Y = T_{i,j;x}(u) - T_{i,j;y}(u)$ and that $Z = 0$.

So first consider $X + Y$. By definition, $T_{i,j;x}(u)$ is a sum of monomials of the form $\pm\xi_{p_1,q_1} \cdots \xi_{p_u,q_u}$. Let A denote the sum of all these monomials with the property that $x < \mathrm{row}(q_t) \leq y$ for some $1 \leq t < u$. Let B be the sum of all the remaining monomials, so $T_{i,j;x}(u) = A + B$. Similarly, $T_{i,j;y}(u)$ is a sum of monomials $\pm\xi_{p_1,q_1} \cdots \xi_{p_u,q_u}$. Let C denote the sum of all the ones with the property that $x < \mathrm{row}(q_t) \leq y$ for some $1 \leq t < u$. Let D denote the sum of all the rest, so $T_{i,j;y}(u) = C + D$. Now one checks that $X = A, Y = -C$ and $B = D$. Hence $X + Y = (A + B) - (C + D) = T_{i,j;x}(u) - T_{i,j;y}(u)$ as claimed.

It remains to show that $Z = 0$. Recall that Z is the sum of all monomials of the form (9.9) such that *either* $\mathrm{col}(j_r) < \mathrm{col}(k_1)$ and $x < \mathrm{row}(l_t) \leq y$ for some $1 \leq t < s$, *or* $\mathrm{col}(j_r) \geq \mathrm{col}(k_1)$ and $x < \mathrm{row}(j_t) \leq y$ for some $1 \leq t < r$. In the former case, choose the minimal index $t \geq 1$ for which $x < \mathrm{row}(l_t) \leq y$. Then

$$\pm(\xi_{i_1,j_1} \cdots \xi_{i_r,j_r})(\xi_{k_1,l_1} \cdots \xi_{k_s,l_s}),\ \ \pm(\xi_{i_1,j_1} \cdots \xi_{i_r,j_r}\xi_{k_1,l_1} \cdots \xi_{k_t,l_t})(\xi_{k_{t+1},l_{t+1}} \cdots \xi_{k_s,l_s})$$

are the only two ways to split the monomial (9.9) so that the left hand term appears in $T_{i,g;x}(u)$ and the right hand term appears in $T_{g,j;y}(u)$ for some $x < g \leq y$. Moreover, they appear in the expansion of Z with opposite signs. The latter case is similar, but one uses the maximal index $t < r$ for which $x < \mathrm{row}(j_t) \leq y$.

(iv) The left hand side of the formula in (iv) is a sum of monomials of the form

$$\pm(\xi_{i_1,j_1} \cdots \xi_{i_r,j_r})(\xi_{k_1,l_1} \cdots \xi_{k_s,l_s}), \tag{9.10}$$

for various $r, s \geq 0$, where $\pm\xi_{i_1,j_1} \cdots \xi_{i_r,j_r}$ appears in $T_{i,k;x}(u)$ and $\pm\xi_{k_1,l_1} \cdots \xi_{k_s,l_s}$ appears in $T_{k,j;y}(u)$ for some $x < k \leq y$. Note that $r = 0$ is allowed only if $i = k$, and $s = 0$ is allowed only if $k = j$. So if $i = j$ we get a contribution -1 from the $r = s = 0$ term, while if $i \neq j$ then there is no $r = s = 0$ term. Now the formula in (iv) will follow if we can show that whenever $(r, s) \neq (0, 0)$, the monomial (9.10) appears exactly twice on the left hand side with opposite signs. We consider two cases.

First, suppose that $s > 0$ and that $\mathrm{col}(j_r) < \mathrm{col}(k_1)$ if $r > 0$. Let $t \geq 1$ be the minimal index such that $x < \mathrm{row}(l_t) \leq y$; such a t exists as $\mathrm{row}(l_s) = j$ and $x < j \leq y$. Let $h := \mathrm{row}(l_t)$. Then $\pm\xi_{i_1,j_1} \ldots \xi_{i_r,j_r}\xi_{k_1,l_1} \ldots \xi_{k_t,l_t}$ appears in $T_{i,h;x}(u)$, and $\pm\xi_{k_{t+1},l_{t+1}} \ldots \xi_{k_s,l_s}$ appears in $T_{h,j;y}(u)$. Moreover, using the facts that $\mathrm{col}(l_t) \geq \mathrm{col}(k_{t+1})$ if $t < s$, $\mathrm{col}(j_r) < \mathrm{col}(k_1)$ if $r > 0$, and the minimality of the choice of t, we see that

$$\pm(\xi_{i_1,j_1} \ldots \xi_{i_r,j_r})(\xi_{k_1,l_1} \ldots \xi_{k_s,l_s}),\ \ \pm(\xi_{i_1,j_1} \ldots \xi_{i_r,j_r}\xi_{k_1,l_1} \ldots \xi_{k_t,l_t})(\xi_{k_{t+1},l_{t+1}} \ldots \xi_{k_s,l_s})$$

are the only ways to split the monomial (9.10) so that the left term occurs in $T_{i,g;x}(u)$ and the right term occurs in $T_{g,j;y}(u)$ for some $x < g \leq y$. The two have opposite signs.

For the second case, suppose either that $s = 0$, or that $r, s > 0$ and $\mathrm{col}(j_r) \geq \mathrm{col}(k_1)$. Let $t \leq r$ be the maximal index such that $x < \mathrm{row}(i_t) \leq y$ and argue in a similar fashion. □

To explain the significance of this lemma, let $T(u) := (T_{i,j;0}(u))_{1 \leq i,j \leq n}$, an $n \times n$ matrix with entries in $U(\mathfrak{p})[[u^{-1}]]$. Also let $\nu = (\nu_1, \ldots, \nu_m)$ be a fixed shape. Consider the Gauss factorization $T(u) = F(u)D(u)E(u)$ where $D(u)$ is a block diagonal matrix, $E(u)$ is a block upper unitriangular matrix, and $F(u)$ is a block lower unitriangular matrix, all block matrices being of shape ν. The diagonal blocks of $D(u)$ define matrices $D_1(u), \ldots, D_m(u)$, the upper diagonal blocks of $E(u)$ define matrices $E_1(u), \ldots, E_{m-1}(u)$, and the lower diagonal matrices of $F(u)$ define matrices $F_1(u), \ldots, F_{m-1}(u)$. Also let $\widetilde{D}_a(u) := -D_a(u)^{-1}$. Thus $D_a(u) = (D_{a;i,j}(u))_{1 \leq i,j \leq \nu_a}$ and $\widetilde{D}_a(u) = (\widetilde{D}_{a;i,j}(u))_{1 \leq i,j \leq \nu_a}$ are $\nu_a \times \nu_a$-matrices, $E_a(u) = (E_{a;i,j}(u))_{1 \leq i \leq \nu_a, 1 \leq j \leq \nu_{a+1}}$

is a $\nu_a \times \nu_{a+1}$-matrix, and $F_a(u) = (F_{a;i,j}(u))_{1 \leq i \leq \nu_{a+1}, 1 \leq j \leq \nu_a}$ is a $\nu_{a+1} \times \nu_a$-matrix. Write

$$D_{a;i,j}(u) = \sum_{r \geq 0} D_{a;i,j}^{(r)} u^{-r}, \qquad \widetilde{D}_{a;i,j}(u) = \sum_{r \geq 0} \widetilde{D}_{a;i,j}^{(r)} u^{-r},$$
$$E_{a;i,j}(u) = \sum_{r > 0} E_{a;i,j}^{(r)} u^{-r}, \qquad F_{a;i,j}(u) = \sum_{r > 0} F_{a;i,j}^{(r)} u^{-r},$$

thus defining elements $D_{a;i,j}^{(r)}, E_{a;i,j}^{(r)}$ and $F_{a;i,j}^{(r)}$ of $U(\mathfrak{p})$, all dependent of course on the fixed choice of ν. All this parallels the definition of the elements of Y_n with the same names in the paragraph following Lemma 3.1.

Theorem 9.3. *With $\nu = (\nu_1, \dots, \nu_m)$ fixed as above and all admissible a, i, j, we have that*

$$\begin{aligned}
D_{a;i,j}(u) &= T_{\nu_1+\cdots+\nu_{a-1}+i, \nu_1+\cdots+\nu_{a-1}+j; \nu_1+\cdots+\nu_{a-1}}(u),\\
\widetilde{D}_{a;i,j}(u) &= T_{\nu_1+\cdots+\nu_{a-1}+i, \nu_1+\cdots+\nu_{a-1}+j; \nu_1+\cdots+\nu_a}(u),\\
E_{a;i,j}(u) &= T_{\nu_1+\cdots+\nu_{a-1}+i, \nu_1+\cdots+\nu_a+j; \nu_1+\cdots+\nu_a}(u),\\
F_{a;i,j}(u) &= T_{\nu_1+\cdots+\nu_a+i, \nu_1+\cdots+\nu_{a-1}+j; \nu_1+\cdots+\nu_a}(u).
\end{aligned}$$

Proof. Note it suffices to prove the formulae for D, E and F, since the one for $\widetilde{D}$ follows from the one for D by Lemma 9.2(iv) taking $x = \nu_1 + \cdots + \nu_{a-1}$ and $y = \nu_1 + \cdots + \nu_a$. Now we proceed by induction on m, the base case $m = 1$ being trivial. For the induction step, suppose the theorem has been proved for the shape $\nu = (\nu_1, \dots, \nu_m)$. So, in terms of matrices, we have that

$$\begin{aligned}
{}^{\nu}D_a(u) &= \left(T_{\nu_1+\cdots+\nu_{a-1}+i, \nu_1+\cdots+\nu_{a-1}+j; \nu_1+\cdots+\nu_{a-1}}(u)\right)_{1 \leq i,j \leq \nu_a},\\
{}^{\nu}E_a(u) &= \left(T_{\nu_1+\cdots+\nu_{a-1}+i, \nu_1+\cdots+\nu_a+j; \nu_1+\cdots+\nu_a}(u)\right)_{1 \leq i \leq \nu_a, 1 \leq j \leq \nu_{a+1}},\\
{}^{\nu}F_a(u) &= \left(T_{\nu_1+\cdots+\nu_a+i, \nu_1+\cdots+\nu_{a-1}+j; \nu_1+\cdots+\nu_a}(u)\right)_{1 \leq i \leq \nu_{a+1}, 1 \leq j \leq \nu_a},
\end{aligned}$$

where we have added a superscript ν for clarity. Write $\nu_b = \alpha + \beta$ for some $1 \leq b \leq m$ and $\alpha, \beta \geq 1$, and let $\mu = (\nu_1, \dots, \nu_{b-1}, \alpha, \beta, \nu_{b+1}, \dots, \nu_m)$. Define matrices A, B, C and D by

$$\begin{aligned}
A &= \left(T_{\nu_1+\cdots+\nu_{b-1}+i, \nu_1+\cdots+\nu_{b-1}+j; \nu_1+\cdots+\nu_{b-1}}(u)\right)_{1 \leq i,j \leq \alpha},\\
B &= \left(T_{\nu_1+\cdots+\nu_{b-1}+i, \nu_1+\cdots+\nu_{b-1}+\alpha+j; \nu_1+\cdots+\nu_{b-1}+\alpha}(u)\right)_{1 \leq i \leq \alpha, 1 \leq j \leq \beta},\\
C &= \left(T_{\nu_1+\cdots+\nu_{b-1}+\alpha+i, \nu_1+\cdots+\nu_{b-1}+j; \nu_1+\cdots+\nu_{b-1}+\alpha}(u)\right)_{1 \leq i \leq \beta, 1 \leq j \leq \alpha},\\
D &= \left(T_{\nu_1+\cdots+\nu_{b-1}+\alpha+i, \nu_1+\cdots+\nu_{b-1}+\alpha+j; \nu_1+\cdots+\nu_{b-1}+\alpha}(u)\right)_{1 \leq i,j \leq \beta}.
\end{aligned}$$

Then Lemma 9.2(i)–(iii) with $x = \nu_1 + \cdots + \nu_{b-1}$ and $y = \nu_1 + \cdots + \nu_{b-1} + \alpha$ tell us that

$${}^{\nu}D_b(u) = \begin{pmatrix} A & AB \\ CA & D + CAB \end{pmatrix} = \begin{pmatrix} I_\alpha & 0 \\ C & I_\beta \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix} \begin{pmatrix} I_\alpha & B \\ 0 & I_\beta \end{pmatrix}.$$

Lemma 3.1 explains how to read off the matrices ${}^\mu D_a(u), {}^\mu E_a(u)$ and ${}^\mu F_a(u)$ from this factorization to get that

$$\begin{aligned}
{}^\mu D_a(u) &= \left(T_{\mu_1+\cdots+\mu_{a-1}+i,\mu_1+\cdots+\mu_{a-1}+j,\mu_1+\cdots+\mu_{a-1}}(u)\right)_{1\le i,j\le \mu_a},\\
{}^\mu E_a(u) &= \left(T_{\mu_1+\cdots+\mu_{a-1}+i,\mu_1+\cdots+\mu_a+j,\mu_1+\cdots+\mu_a}(u)\right)_{1\le i\le \mu_a, 1\le j\le \mu_{a+1}},\\
{}^\mu F_a(u) &= \left(T_{\mu_1+\cdots+\mu_a+i,\mu_1+\cdots+\mu_{a-1}+j,\mu_1+\cdots+\mu_a}(u)\right)_{1\le i\le \mu_{a+1}, 1\le j\le \mu_a}.
\end{aligned}$$

This completes the induction step. □

In the extreme case that $\nu = (1^n)$, we write simply $D_i^{(r)}, \widetilde{D}_i^{(r)}, E_i^{(r)}$ and $F_i^{(r)}$ for the elements $D_{i;1,1}^{(r)}$, $\widetilde{D}_{i;1,1}^{(r)}$, $E_{i;1,1}^{(r)}$ and $F_{i;1,1}^{(r)}$ of $U(\mathfrak{p})$, respectively.

Corollary 9.4. $D_i^{(r)} = T_{i,i;i-1}^{(r)}$, $\widetilde{D}_i^{(r)} = T_{i,i;i}^{(r)}$, $E_i^{(r)} = T_{i,i+1;i}^{(r)}$ *and* $F_i^{(r)} = T_{i+1,i;i}^{(r)}$.

10. Main theorem

Let π be a pyramid of level l with column heights $(q_1,\dots,q_l)$. Pick an integer $n \ge \max(q_1,\dots,q_l)$ and read off from the pyramid π a shift matrix $\sigma = (s_{i,j})_{1\le i,j\le n}$ according to (7.6). Define the finite W-algebra $W(\pi) = U(\mathfrak{p})^{\mathfrak{m}}$ viewed as a filtered algebra via the Kazhdan filtration as in §8. Define the shifted Yangian $Y_{n,l}(\sigma)$ of level l viewed as a filtered algebra via the canonical filtration as in §§2, 5 and 6. Suppose also that $\nu = (\nu_1,\dots,\nu_m)$ is an admissible shape for σ, and recall the notation $s_{a,b}(\nu)$ and $p_a(\nu)$ from (3.2) and (6.2). We have the elements $D_{a;i,j}^{(r)}$, $\widetilde{D}_{a;i,j}^{(r)}$, $E_{a;i,j}^{(r)}$ and $F_{a;i,j}^{(r)}$ of $U(\mathfrak{p})$ defined as in §9 relative to this fixed shape; see Theorem 9.3 for the explicit formulae. We also have the parabolic generators $D_{a;i,j}^{(r)}$, $\widetilde{D}_{a;i,j}^{(r)}$, $E_{a;i,j}^{(r)}$ and $F_{a;i,j}^{(r)}$ of $Y_{n,l}(\sigma)$ as in §3. The main result of the article is as follows.

Theorem 10.1. *There is a unique isomorphism $Y_{n,l}(\sigma) \xrightarrow{\sim} W(\pi)$ of filtered algebras such that for any admissible shape $\nu = (\nu_1,\dots,\nu_m)$ the generators*

$$\begin{aligned}
&\{D_{a;i,j}^{(r)}\}_{1\le a\le m, 1\le i,j\le \nu_a, r>0},\\
&\{E_{a;i,j}^{(r)}\}_{1\le a< m, 1\le i\le \nu_a, 1\le j\le \nu_{a+1}, r>s_{a,b}(\nu)},\\
&\{F_{a;i,j}^{(r)}\}_{1\le a< m, 1\le i\le \nu_{a+1}, 1\le j\le \nu_a, r>s_{b,a}(\nu)}
\end{aligned}$$

of $Y_{n,l}(\sigma)$ map to the elements of $U(\mathfrak{p})$ with the same names. In particular, these elements of $U(\mathfrak{p})$ are $\mathfrak{m}$-invariants and they generate $W(\pi)$.

The proof will take up most of the rest of the section. First however let us record a couple of corollaries of the theorem.

Corollary 10.2. *The map $\theta : \operatorname{gr} W(\pi) \to S(\mathfrak{c}_{\mathfrak{g}}(e))$ from (8.4) is an isomorphism.*

Proof. Since we already know by Corollary 8.3 that the map θ is an injective, graded homomorphism, it suffices given Theorem 10.1 to show that $\operatorname{gr} Y_{n,l}(\sigma)$ and $S(\mathfrak{c}_{\mathfrak{g}}(e))$ have the same dimension in each degree. This follows from Theorem 6.2 and Lemma 7.3 (the element $c_{i,j}^{(r)}$ there is of degree r with respect to the Kazhdan grading). □

For the next corollary, let $\dot\pi$ be another pyramid with the same row lengths as π, i.e. the nilpotent matrix $\dot e$ defined from $\dot\pi$ is conjugate to the nilpotent matrix e defined from π. Let $\dot\sigma = (\dot s_{i,j})_{1\le i,j\le n}$ be a shift matrix corresponding to the pyramid $\dot\pi$. By

Theorem 10.1, the formulae in §9 define generators $\dot{D}_{a;i,j}^{(r)}$, $\dot{E}_{a;i,j}^{(r)}$ and $\dot{F}_{a;i,j}^{(r)}$ of $W(\dot{\pi})$ for each admissible shape ν.

Corollary 10.3. *There is an algebra isomorphism $\iota : W(\pi) \to W(\dot{\pi})$ defined on parabolic generators with respect to an admissible shape ν by the formulae (3.21).*

Proof. This follows from Theorem 10.1 and (6.4). □

So now we must prove Theorem 10.1. The first reduction is to observe that it suffices to prove it in the special case that ν is the minimal admissible shape for σ. It then follows for all other admissible shapes by Lemma 3.1 and induction on the length of the shape. So we assume from now on that $\nu = (\nu_1, \dots, \nu_m)$ is the minimal admissible shape for σ, and let $t := \nu_m = \min(q_1, q_l)$. We proceed to prove Theorem 10.1 by induction on the level l.

Consider first the base case $l = 1$, i.e. π consists of a single column of height t. Since the nilpotent element e from (7.4) is zero in this case, we have by definition that $W(\pi) = U(\mathfrak{gl}_t)$. Moreover, according to Theorem 9.3, we have that $D_{m;i,j}^{(1)} = \tilde{e}_{i,j} = e_{i,j} + \delta_{i,j}(n-t)$ for $1 \leq i,j \leq t$, and all other elements $D_{a;i,j}^{(r)}, E_{a;i,j}^{(r)}$ and $F_{a;i,j}^{(r)}$ of $W(\pi)$ are equal to zero. On the other hand, Remarks 6.6 and 6.5 imply that there is an isomorphism $Y_{n,1}(\sigma) \xrightarrow{\sim} Y_{t,1} = U(\mathfrak{gl}_t)$ such that $D_{m,i,j}^{(1)} \mapsto e_{i,j}$, with all the other parabolic $D_{a;i,j}^{(r)}, E_{a;i,j}^{(r)}$ and $F_{a;i,j}^{(r)}$ of $Y_{n,1}(\sigma)$ mapping to zero. Composing with the automorphism $e_{i,j} \mapsto \tilde{e}_{i,j}$ of $U(\mathfrak{gl}_t)$ gives the required isomorphism $Y_{n,1}(\sigma) \xrightarrow{\sim} W(\pi)$.

Assume from now on that $l > 1$ and that the theorem has been proved for all smaller levels. The verification of the induction step splits into two cases corresponding to the two possible baby comultiplications Δ_{R} and Δ_{L} that were defined in (6.8)–(6.9):
Case Δ_{R}*:* either $t = n$ or $s_{n-t,n-t+1} \neq 0$;
Case Δ_{L}*:* either $t = n$ or $s_{n-t+1,n-t} \neq 0$.
The argument in the two cases is quite similar. We will explain it in detail in case Δ_{R}, then briefly indicate the changes in case Δ_{L}.

So assume that either $t = n$ or $s_{n-t,n-t+1} = s_{m-1,m}(\nu) \neq 0$; in particular, $q_1 \geq q_l$. For notational convenience, assume that the numbering of the bricks of the pyramid π is the standard numbering down columns from left to right. Let $\dot{\pi}$ be the pyramid obtained from π by removing the rightmost column, i.e. the bricks numbered $(N - t + 1), (N - t + 2), \dots, N$, from the pyramid π. Let $\dot{\sigma} = (\dot{s}_{i,j})_{1 \leq i,j \leq n}$ be the shift matrix corresponding to the pyramid $\dot{\pi}$ defined by (4.5). Define $\dot{\mathfrak{p}}, \dot{\mathfrak{m}}$ and $\dot{e}$ in $\dot{\mathfrak{g}} = \mathfrak{gl}_{N-t}$ according to (7.3)–(7.4) and let $\dot{\chi} : \dot{\mathfrak{m}} \to \mathbb{C}$ be the character $x \mapsto (x, \dot{e})$. Let $\dot{D}_{a;i,j}^{(r)}$, $\dot{\tilde{D}}_{a;i,j}^{(r)}$, $\dot{E}_{a;i,j}^{(r)}$ and $\dot{F}_{a;i,j}^{(r)}$ denote the elements of $U(\dot{\mathfrak{p}})$ as defined in §9 relative to the shape ν. By the induction hypothesis, Theorem 10.1 holds for $\dot{\pi}$, so we know already that the following elements of $U(\dot{\mathfrak{p}})$ are invariant under the twisted action of $\dot{\mathfrak{m}}$, i.e. they belong to finite W-algebra $W(\dot{\pi}) = U(\dot{\mathfrak{p}})^{\dot{\mathfrak{m}}}$:

(i) $\dot{D}_{a;i,j}^{(r)}$ and $\dot{\tilde{D}}_{a;i,j}^{(r)}$ for $1 \leq a \leq m$, $1 \leq i,j \leq \nu_a$ and $r > 0$;
(ii) $\dot{E}_{a;i,j}^{(r)}$ for $1 \leq a < m$, $1 \leq i \leq \nu_a$, $1 \leq j \leq \nu_{a+1}$ and $r > s_{a,a+1}(\nu) - \delta_{a,m-1}$;
(iii) $\dot{F}_{a;i,j}^{(r)}$ for $1 \leq a < m$, $1 \leq i \leq \nu_{a+1}$, $1 \leq j \leq \nu_a$ and $r > s_{a+1,a}(\nu)$.

Recalling (9.2), we must work now with the non-standard embedding of $U(\dot{\mathfrak{g}})$ into $U(\mathfrak{g})$ under which the generators $\tilde{e}_{i,j}$ of $U(\dot{\mathfrak{g}})$ defined from the pyramid $\dot{\pi}$ map to the generators $\tilde{e}_{i,j}$ of $U(\mathfrak{g})$ defined from the pyramid π, for $1 \le i,j \le N-t$. This also embeds $U(\dot{\mathfrak{p}})$ into $U(\mathfrak{p})$ and $\dot{\mathfrak{m}}$ into $\mathfrak{m}$. Moreover, recalling (9.4), the character $\dot{\chi}$ of $\dot{\mathfrak{m}}$ is the restriction of the character χ of $\mathfrak{m}$, hence the twisted action of $\dot{\mathfrak{m}}$ on $U(\dot{\mathfrak{p}})$ is the restriction of the twisted action of $\mathfrak{m}$ on $U(\mathfrak{p})$. The following crucial lemma gives us an inductive description of the elements $D^{(r)}_{a;i,j}$, $E^{(r)}_{a;i,j}$ and $F^{(r)}_{a;i,j}$ of $U(\mathfrak{p})$ in terms of the elements (i)–(iii) of $U(\dot{\mathfrak{p}})$; this should be compared with Theorem 4.1(i).

Lemma 10.4. *The following equations hold for $r > 0$, all admissible a, i, j and any fixed $1 \le h \le t$:*

$$D^{(r)}_{a;i,j} = \dot{D}^{(r)}_{a;i,j} + \delta_{a,m}\left(\sum_{k=1}^{t} \dot{D}^{(r-1)}_{a;i,k}\tilde{e}_{N-t+k,N-t+j} + [\dot{D}^{(r-1)}_{a;i,h}, \tilde{e}_{N-2t+h,N-t+j}]\right), \tag{10.1}$$

$$E^{(r)}_{a;i,j} = \dot{E}^{(r)}_{a;i,j} + \delta_{a,m-1}\left(\sum_{k=1}^{t} \dot{E}^{(r-1)}_{a;i,k}\tilde{e}_{N-t+k,N-t+j} + [\dot{E}^{(r-1)}_{a;i,h}, \tilde{e}_{N-2t+h,N-t+j}]\right), \tag{10.2}$$

$$F^{(r)}_{a;i,j} = \dot{F}^{(r)}_{a;i,j}, \tag{10.3}$$

where for (10.2) we are assuming that $r > 1$ if $a = m-1$.

Proof. This follows using Theorem 9.3 and the explicit form of the elements $T^{(r)}_{i,j;x}$ from (9.5). □

In the next few lemmas we will use these inductive descriptions to show that the elements $D^{(r)}_{a;i,j}, E^{(r)}_{a;i,j}$ and $F^{(r)}_{a;i,j}$ of $U(\mathfrak{p})$ are $\mathfrak{m}$-invariants for the appropriate r.

Lemma 10.5. *The following elements of $U(\mathfrak{p})$ are invariant under the twisted action of $\mathfrak{m}$:*

- (i) $D^{(r)}_{a;i,j}$ *and* $\widetilde{D}^{(r)}_{a;i,j}$ *for* $1 \le a \le m-1$, $1 \le i,j \le \nu_a$ *and* $r > 0$;
- (ii) $E^{(r)}_{a;i,j}$ *for* $1 \le a < m-1$, $1 \le i \le \nu_a$, $1 \le j \le \nu_{a+1}$ *and* $r > s_{a,a+1}(\nu)$;
- (iii) $F^{(r)}_{a;i,j}$ *for* $a = 1 \le a < m$, $1 \le i \le \nu_{a+1}$, $1 \le j \le \nu_a$ *and* $r > s_{a+1,a}(\nu)$.

Proof. By Lemma 10.4 and the definitions of $\widetilde{D}^{(r)}_{a;i,j}$ and $\dot{\widetilde{D}}^{(r)}_{a;i,j}$, all these elements of $U(\mathfrak{p})$ coincide with the corresponding elements of $U(\dot{\mathfrak{p}})$. Hence by the induction hypothesis we already know they are invariant under the twisted action of $\dot{\mathfrak{m}}$. It remains to show that the elements are invariant under the twisted action of all $\tilde{e}_{f,g}$ with $1 \le g \le N-t < f \le N$. By Theorem 9.3 and the explicit form of (9.5), all the elements we are considering are linear combinations of monomials of the form $\tilde{e}_{i_1,j_1}\cdots\tilde{e}_{i_r,j_r} \in U(\dot{\mathfrak{p}})$ with $1 \le i_s \le N-t$ and $1 \le j_s \le N-2t$ for all $s = 1,\dots,r$. Using just the fact that $\chi(e_{f,g}) = 0$ for all $1 \le g \le N-2t$ and $N-t < f \le N$, it is easy to see that all such monomials are invariant under the twisted action of all $\tilde{e}_{f,g}$ with $1 \le g \le N-t < f \le N$. □

Lemma 10.6. *The following elements of $U(\mathfrak{p})$ are invariant under the twisted action of $\dot{\mathfrak{m}}$:*

- (i) $D^{(r)}_{m;i,j}$ *for* $1 \le i,j \le \nu_m$ *and* $r > 0$;
- (ii) *(assuming $t < n$)* $E^{(r)}_{m-1;i,j}$ *for* $1 \le i \le \nu_{m-1}$, $1 \le j \le \nu_m$ *and* $r > s_{m-1,m}(\nu)$.

Proof. (i) Take $x \in \dot{\mathfrak{m}}$. According to Lemma 10.4, we have that

$$D_{m;i,j}^{(r)} = \dot{D}_{m;i,j}^{(r)} + \sum_{k=1}^{t} \dot{D}_{m;i,k}^{(r-1)} \tilde{e}_{N-t+k,N-t+j} + [\dot{D}_{m;i,h}^{(r-1)}, \tilde{e}_{N-2t+h,N-t+j}].$$

Noting that $[x, \tilde{e}_{N-t+k,N-t+j}] = [x, \tilde{e}_{N-2t+h,N-t+j}] = 0$ and using the induction hypothesis, one deduces easily from this equation that $\mathrm{pr}_\chi([x, D_{m;i,j}^{(r)}]) = 0$ as required.

(ii) Similar. □

Lemma 10.7. *The following elements of $U(\mathfrak{p})$ are invariant under the twisted action of $\tilde{e}_{f,g}$ for all $1 \le g \le N-t < f \le N$:*

(i) $D_{m;i,j}^{(1)}$ *for $1 \le i,j \le \nu_m$;*

(ii) *(assuming $t < n$ and $s_{m-1,m}(\nu) = 1$) $D_{m;i,j}^{(2)}$ for $1 \le i,j \le \nu_m$;*

(iii) *(assuming $t < n$ and $s_{m-1,m}(\nu) = 1$) $E_{m-1;i,j}^{(2)}$ for $1 \le i \le \nu_{m-1}$, $1 \le j \le \nu_m$.*

Proof. Part (i) is easily checked directly using (9.3), (9.4) and the explicit formula for $D_{m;i,j}^{(1)}$ given by Theorem 9.3 and Example 9.1. The proofs of (ii) and (iii) are similar, though the calculations are not so easy. □

Lemma 10.8. *Suppose that $t < n$ and $s_{m-1,m}(\nu) = 1$. Then, the following equations hold in $U(\mathfrak{p})$ for $r > 1$, all admissible i,j and any fixed $1 \le g \le \nu_{m-1}$:*

$$D_{m;i,j}^{(r+1)} = [F_{m-1;i,g}^{(2)}, E_{m-1;g,j}^{(r)}] + \sum_{s=0}^{r} \widetilde{D}_{m-1;g,g}^{(r+1-s)} D_{m;i,j}^{(s)}, \tag{10.4}$$

$$E_{m-1;i,j}^{(r+1)} = [D_{m-1;i,g}^{(2)}, E_{m-1;g,j}^{(r)}] - \sum_{f=1}^{\nu_{m-1}} D_{m-1;i,f}^{(1)} E_{m-1;f,j}^{(r)}. \tag{10.5}$$

Proof. We prove (10.5), the first equation being a similar trick. By the induction hypothesis and the relation (3.5), we know for all $r > 0$ that

$$[\dot{D}_{m-1;i,g}^{(2)}, \dot{E}_{m-1;g,j}^{(r)}] = \dot{E}_{m-1;i,j}^{(r+1)} + \sum_{f=1}^{\nu_{m-1}} \dot{D}_{m-1;i,f}^{(1)} \dot{E}_{m-1;f,j}^{(r)}.$$

By Lemma 10.4, we have for $r > 1$ that

$$E_{m-1;g,j}^{(r)} = \dot{E}_{m-1;g,j}^{(r)} + \sum_{k=1}^{t} \dot{E}_{m-1;g,k}^{(r-1)} \tilde{e}_{N-t+k,N-t+j} + [\dot{E}_{m-1;g,h}^{(r-1)}, \tilde{e}_{N-2t+h,N-t+j}].$$

Obviously, $[\dot{D}_{m-1;i,g}^{(2)}, \tilde{e}_{N-t+k,N-t+j}] = 0$. Moreover, by Theorem 9.3 and the form of (9.5), no monomial in the expansion of $\dot{D}_{m-1;i,g}^{(2)}$ involves any matrix unit of the form $\tilde{e}_{?,N-2t+h}$, hence $[\dot{D}_{m-1;i,g}^{(2)}, \tilde{e}_{N-2t+h,N-t+j}] = 0$ too. Now we can commute with

$\dot{D}^{(2)}_{m-1;i,g} = D^{(2)}_{m-1;i,g}$ to deduce that

$$[D^{(2)}_{m-1;i,g}, E^{(r)}_{m-1;g,j}] = \dot{E}^{(r+1)}_{m-1;i,j} + \sum_{f=1}^{\nu_{m-1}} \dot{D}^{(1)}_{m-1;i,f}\dot{E}^{(r)}_{m-1;f,j}$$
$$+ \sum_{k=1}^{t}\left\{\dot{E}^{(r)}_{m-1;i,k} + \sum_{f=1}^{\nu_{m-1}} \dot{D}^{(1)}_{m-1;i,f}\dot{E}^{(r-1)}_{m-1;f,k}\right\}\tilde{e}_{N-t+k,N-t+j}$$
$$+ \left[\dot{E}^{(r)}_{m-1;i,h} + \sum_{f=1}^{\nu_{m-1}} \dot{D}^{(1)}_{m-1;i,f}\dot{E}^{(r-1)}_{m-1;f,h}, \tilde{e}_{N-2t+h,N-t+j}\right].$$

Finally rewrite the right hand side using Lemma 10.4 again to see that it equals $E^{(r+1)}_{m-1;i,j} + \sum_{f=1}^{\nu_{m-1}} D^{(1)}_{m-1;i,f}E^{(r)}_{m-1;f,j}$. □

Lemma 10.9. *Suppose that $t = n$ and $l > 1$ or that $t < n$ and $s_{m-1,m}(\nu) > 1$. Then the following elements of $U(\mathfrak{p})$ are invariant under the twisted action of $\tilde{e}_{N-t+f,N-2t+g}$ for all $1 \le f, g \le t$.*

(i) *$D^{(r)}_{m;i,j}$ for $1 \le i, j \le \nu_m$ and $r > 1$;*

(ii) *(assuming $t < n$) $E^{(r)}_{m-1;i,j}$ for $1 \le i \le \nu_{m-1}$, $1 \le j \le \nu_m$ and $r > s_{m-1,m}(\nu)$.*

Proof. (i) Let $\ddot{\pi}$ denote the pyramid obtained by removing the rightmost column from the pyramid $\dot{\pi}$, i.e. the bricks numbered $(N-2t+1), (N-2t+2), \dots, (N-t)$. Define $\ddot{\mathfrak{p}}, \ddot{\mathfrak{m}}$ and $\ddot{e}$ in $\ddot{\mathfrak{g}} = \mathfrak{gl}_{N-2t}$ according to (7.3)–(7.4), and embed $U(\ddot{\mathfrak{g}})$ into $U(\dot{\mathfrak{g}})$ in exactly the same (non-standard) way as we embedded $U(\dot{\mathfrak{g}})$ into $U(\mathfrak{g})$. We will make use of the elements $\ddot{D}^{(r)}_{a;i,j}$ of $W(\ddot{\pi})$. By Lemma 10.4 applied to $\dot{\pi}$, the following equation holds for $r > 0$, $1 \le i, j \le \nu_m$ and any fixed $1 \le h \le t$:

$$\dot{D}^{(r)}_{m;i,j} = \ddot{D}^{(r)}_{m;i,j} + \sum_{c=1}^{t}\ddot{D}^{(r-1)}_{m;i,c}\tilde{e}_{N-2t+c,N-2t+j} + [\ddot{D}^{(r-1)}_{m;i,h}, \tilde{e}_{N-3t+h,N-2t+j}].$$

Substituting this into (10.1) and simplifying a little using (9.3) we deduce for $r > 1$ that $D^{(r)}_{m;i,j} = A + B + C + D + E + F + G + H$ where

$$A = \ddot{D}^{(r)}_{m;i,j}, \qquad E = \sum_{k,c=1}^{t}\ddot{D}^{(r-2)}_{m;i,c}\tilde{e}_{N-2t+c,N-2t+k}\tilde{e}_{N-t+k,N-t+j},$$

$$B = \sum_{c=1}^{t}\ddot{D}^{(r-1)}_{m;i,c}\tilde{e}_{N-2t+c,N-2t+j}, \qquad F = \sum_{k=1}^{t}[\ddot{D}^{(r-2)}_{m;i,h}, \tilde{e}_{N-3t+h,N-2t+k}]\tilde{e}_{N-t+k,N-t+j},$$

$$C = [\ddot{D}^{(r-1)}_{m;i,h}, \tilde{e}_{N-3t+h,N-2t+j}], \qquad G = \sum_{c=1}^{t}\ddot{D}^{(r-2)}_{m;i,c}\tilde{e}_{N-2t+c,N-t+j},$$

$$D = \sum_{k=1}^{t}\ddot{D}^{(r-1)}_{m;i,k}\tilde{e}_{N-t+k,N-t+j}, \qquad H = [\ddot{D}^{(r-2)}_{m;i,h}, \tilde{e}_{N-3t+h,N-t+j}].$$

Now commute each of these elements in turn with $x := \tilde{e}_{N-t+f,N-2t+g}$ then apply pr_χ, using (9.1), (9.3), (9.4) and the observation that x commutes with all elements of $U(\ddot{\mathfrak{p}})$, to deduce that

$$\begin{aligned}
&\mathrm{pr}_\chi([x,A]) = 0, && \mathrm{pr}_\chi([x,C]) = 0,\\
&\mathrm{pr}_\chi([x,B]) = -\delta_{f,j}\ddot{D}^{(r-1)}_{m;i,g} && \mathrm{pr}_\chi([x,D]) = \delta_{f,j}\ddot{D}^{(r-1)}_{m;i,g},
\end{aligned}$$

$$\mathrm{pr}_\chi([x,E]) = t\delta_{f,j}\ddot{D}^{(r-2)}_{m;i,g} - \ddot{D}^{(r-2)}_{m;i,g}\tilde{e}_{N-t+f,N-t+j} + \delta_{f,j}\sum_{c=1}^{t}\ddot{D}^{(r-2)}_{m;i,c}\tilde{e}_{N-2t+c,N-2t+g},$$

$$\mathrm{pr}_\chi([x,F]) = \delta_{f,j}[\ddot{D}^{(r-2)}_{m;i,h}, \tilde{e}_{N-3t+h,N-2t+g}],$$

$$\mathrm{pr}_\chi([x,G]) = \ddot{D}^{(r-2)}_{m;i,g}\tilde{e}_{N-t+f,N-t+j} - t\delta_{f,j}\ddot{D}^{(r-2)}_{m;i,g} - \delta_{f,j}\sum_{c=1}^{t}\ddot{D}^{(r-2)}_{m;i,c}\tilde{e}_{N-2t+c,N-2t+g},$$

$$\mathrm{pr}_\chi([x,H]) = -\delta_{f,j}[\ddot{D}^{(r-2)}_{m;i,h}, \tilde{e}_{N-3t+h,N-2t+g}].$$

These sum to zero, hence $\mathrm{pr}_\chi([x, D^{(r)}_{m;i,j}]) = 0$.

(ii) Similar. □

Lemma 10.10. *The following elements of $U(\mathfrak{p})$ are invariant under the twisted action of $\mathfrak{m}$:*

(i) $D^{(r)}_{a;i,j}$ *for* $1 \le a \le m$, $1 \le i,j \le \nu_a$ *and* $r > 0$;

(ii) $E^{(r)}_{a;i,j}$ *for* $1 \le a < m$, $1 \le i \le \nu_a$, $1 \le j \le \nu_{a+1}$ *and* $r > s_{a,a+1}(\nu)$;

(iii) $F^{(r)}_{a;i,j}$ *for* $1 \le a < m$, $1 \le i \le \nu_{a+1}$, $1 \le j \le \nu_a$ *and* $r > s_{a+1,a}(\nu)$.

Proof. This is just a matter of assembling the pieces. Lemma 10.5 covers all the elements except $D^{(r)}_{m;i,j}$ and (assuming $t < n$) $E^{(r)}_{m-1;i,j}$. Since $\mathfrak{m}$ is generated by $\dot{\mathfrak{m}}$ and the elements $\tilde{e}_{f,g}$ for all $1 \le f,g \le N$ with $\mathrm{col}(f) = l, \mathrm{col}(g) = l-1$, Lemma 10.6 reduces the problem to showing that the elements $D^{(r)}_{m;i,j}$ for $r > 0$ and (assuming $t < n$) $E^{(r)}_{m-1;i,j}$ for $r > s_{m-1,m}(\nu)$ are invariant under all such $\tilde{e}_{f,g}$. Suppose first that $t = n$. Then the required invariance is checked in Lemma 10.7(i) and Lemma 10.9(i). Now assume that $t < n$. If $s_{m-1,m}(\nu) = 1$, then the invariance of $D^{(1)}_{m;i,j}, D^{(2)}_{m;i,j}$ and $E^{(2)}_{m-1;i,j}$ is checked in Lemma 10.7. The invariance of all higher $D^{(r)}_{m;i,j}$ and $E^{(r)}_{m-1;i,j}$ then follows by Lemma 10.5, Lemma 10.8 and induction on r. Finally if $s_{m-1,m}(\nu) > 1$ then the invariance of $D^{(1)}_{m;i,j}$ is checked in Lemma 10.7(i), and the remaining elements are covered by Lemma 10.9. □

Now we complete the proof of the induction step in case Δ_{R}. By the induction hypothesis, we can identify the shifted Yangian $Y_{n,l-1}(\dot{\sigma})$ with $W(\dot{\pi}) \subseteq U(\dot{\mathfrak{p}})$, so that the generators $\dot{D}^{(r)}_{a;i,j}, \dot{E}^{(r)}_{a;i,j}$ and $\dot{F}^{(r)}_{a;i,j}$ of $Y_{n,l-1}(\dot{\sigma})$ coincide with the elements of $W(\dot{\pi})$ with the same name. Theorem 6.2 then shows that there is an injective algebra homomorphism $\Delta_{\mathrm{R}} : Y_{n,l}(\sigma) \to U(\dot{\mathfrak{p}}) \otimes U(\mathfrak{gl}_t)$. Observe moreover comparing Theorem 6.2 with Lemma 7.3 (like in the proof of Corollary 10.2) that for each $d \ge 0$,

$$\dim \Delta_{\mathrm{R}}(\mathrm{F}_d Y_{n,l}(\sigma)) = \dim \mathrm{F}_d Y_{n,l}(\sigma) = \dim \mathrm{F}_d S(\mathfrak{c}_{\mathfrak{g}}(e)), \tag{10.6}$$

where $\mathrm{F}_d S(\mathfrak{c}_{\mathfrak{g}}(e))$ denotes the sum of all the graded pieces of $S(\mathfrak{c}_{\mathfrak{g}}(e))$ of degree $\le d$ in the Kazhdan grading. Define elements $E^{(r)}_{a,b;i,j}$ and $F^{(r)}_{a,b;i,j}$ of $\mathrm{F}_r U(\mathfrak{p})$ recursively by the formulae (3.15)–(3.16). Let X_d denote the subspace of $U(\mathfrak{p})$ spanned by all monomials in

$$\{D_{a;i,j}^{(r)}\}_{1\leq a\leq m, 1\leq i,j\leq \nu_a, 0<r\leq s_a(\nu)},$$
$$\{E_{a,b;i,j}^{(r)}\}_{1\leq a<b\leq m, 1\leq i\leq \nu_a, 1\leq j\leq \nu_b, s_{a,b}(\nu)<r\leq s_{a,b}(\nu)+p_a(\nu)},$$
$$\{F_{a,b;i,j}^{(r)}\}_{1\leq a<b\leq m, 1\leq i\leq \nu_b, 1\leq j\leq \nu_a, s_{b,a}(\nu)<r\leq s_{b,a}(\nu)+p_a(\nu)}$$

taken in some fixed order and of total degree $\leq d$. By Lemma 10.10, X_d is actually a subspace of $\mathrm{F}_d W(\pi)$. Define an algebra homomorphism $\varphi_{\mathrm{R}} : U(\mathfrak{p}) \to U(\dot{\mathfrak{p}}) \otimes U(\mathfrak{gl}_t)$ by

$$\varphi_{\mathrm{R}}(\tilde{e}_{i,j}) = \begin{cases} \tilde{e}_{i,j} \otimes 1 & \text{if } \mathrm{col}(i) \leq \mathrm{col}(j) \leq l-1, \\ 0 & \text{if } \mathrm{col}(i) \leq l-1, l = \mathrm{col}(j), \\ 1 \otimes \tilde{e}_{i-N+t,j-N+t} & \text{if } l = \mathrm{col}(i) = \mathrm{col}(j), \end{cases} \tag{10.7}$$

where for the rightmost tensor $\tilde{e}_{i-N+t,j-N+t} \in U(\mathfrak{gl}_t)$ denotes $e_{i-N+t,j-N+t} + \delta_{i,j}(n-t)$ like in Theorem 4.1. By Lemma 10.4, we have that

$$\varphi_{\mathrm{R}}(D_{a;i,j}^{(r)}) = \dot{D}_{a;i,j}^{(r)} \otimes 1 + \delta_{a,m} \sum_{k=1}^{t} \dot{D}_{a;i,k}^{(r-1)} \otimes \tilde{e}_{k,j}, \tag{10.8}$$

$$\varphi_{\mathrm{R}}(E_{a;i,j}^{(r)}) = \dot{E}_{a;i,j}^{(r)} \otimes 1 + \delta_{a,m-1} \sum_{k=1}^{t} \dot{E}_{a;i,k}^{(r-1)} \otimes \tilde{e}_{k,j}, \tag{10.9}$$

$$\varphi_{\mathrm{R}}(F_{a;i,j}^{(r)}) = \dot{F}_{a;i,j}^{(r)} \otimes 1. \tag{10.10}$$

Comparing this with Theorem 4.1(i) and recalling the PBW basis for $\mathrm{F}_d Y_{n,l}(\sigma)$ from Theorem 6.2 and Corollary 6.3, we see that $\varphi_{\mathrm{R}}(X_d) = \Delta_{\mathrm{R}}(\mathrm{F}_d Y_{n,l}(\sigma))$. Combining this with (10.6) and Corollary 8.3, we deduce that

$$\dim \mathrm{F}_d S(\mathfrak{c}_{\mathfrak{g}}(e)) = \dim \varphi_{\mathrm{R}}(X_d) \leq \dim X_d \leq \dim \mathrm{F}_d W(\pi) \leq \dim \mathrm{F}_d S(\mathfrak{c}_{\mathfrak{g}}(e)).$$

Hence equality holds everywhere, so we get that $X_d = \mathrm{F}_d W(\pi)$ for each $d \geq 0$, and the map $\varphi_{\mathrm{R}} : W(\pi) \to U(\dot{\mathfrak{p}}) \otimes U(\mathfrak{gl}_t)$ is injective with the same image as $\Delta_{\mathrm{R}} : Y_{n,l}(\sigma) \to U(\dot{\mathfrak{p}}) \otimes U(\mathfrak{gl}_t)$. In particular this shows that the elements listed in Theorem 10.1 generate $W(\pi)$, and the map $\varphi_{\mathrm{R}}^{-1} \circ \Delta_{\mathrm{R}} : Y_{n,l} \to W(\pi)$ is exactly the filtered algebra isomorphism described in the statement of Theorem 10.1. This completes the proof of the induction step in the case Δ_{R}.

Finally we must sketch the proof of the induction step in the case Δ_{L}. So assume that either $t = n$ or $s_{n-t+1,n-t} = s_{m,m-1}(\nu) \neq 0$; in particular, $q_1 \leq q_l$. This time it is notationally convenient to number of the bricks of the pyramid π down columns from *right to left*, not from left to right as before. For instance, the entries down the first (leftmost) column of the pyramid are $(N-t+1), (N-t+2), \dots, N$ in this new numbering. Let $\dot{\pi}$ be the pyramid obtained from π by removing this column from the pyramid π. Let $\dot{\sigma} = (\dot{s}_{i,j})_{1\leq i,j\leq n}$ be the shift matrix corresponding to the pyramid $\dot{\pi}$ defined by (4.6). Define $\dot{\mathfrak{p}}, \dot{\mathfrak{m}}$ and $\dot{e}$ in $\dot{\mathfrak{g}} = \mathfrak{gl}_{N-t} \subset \mathfrak{g}$ according to (7.3)–(7.4). This time, it happens that the natural embedding of $U(\dot{\mathfrak{g}})$ into $U(\mathfrak{g})$ induced by the embedding of $\dot{\mathfrak{g}}$ into $\mathfrak{g}$ already maps the elements $\tilde{e}_{i,j}$ of $U(\dot{\mathfrak{g}})$ to the elements $\tilde{e}_{i,j}$ of $U(\mathfrak{g})$ for $1 \leq i,j \leq N-t$. The algebra $W(\dot{\pi}) = U(\dot{\mathfrak{p}})^{\dot{\mathfrak{m}}}$ is a subalgebra of $U(\dot{\mathfrak{p}}) \subset U(\mathfrak{p})$, $\dot{\mathfrak{m}}$ is a subalgebra of $\mathfrak{m}$, and the twisted action of $\dot{\mathfrak{m}}$ on $U(\dot{\mathfrak{p}})$ is the restriction of the twisted action of $\mathfrak{m}$ on $U(\mathfrak{p})$. Let $\dot{D}_{a;i,j}^{(r)}$, $\dot{\tilde{D}}_{a;i,j}^{(r)}$, $\dot{E}_{a;i,j}^{(r)}$ and $\dot{F}_{a;i,j}^{(r)}$ denote the elements of $U(\dot{\mathfrak{p}})$ as defined in §9 relative to the shape ν. The all-important analogue of Lemma 10.4 is as follows.

Lemma 10.11. *The following equations hold for $r > 0$, all admissible a, i, j and any fixed $1 \le l \le t$:*

$$D_{a;i,j}^{(r)} = \dot{D}_{a;i,j}^{(r)} + \delta_{a,m}\left(\sum_{k=1}^{t} \tilde{e}_{N-t+i,N-t+k}\dot{D}_{a;k,j}^{(r-1)} + [\tilde{e}_{N-t+i,N-2t+h}, \dot{D}_{a;h,j}^{(r-1)}]\right), \quad (10.11)$$

$$E_{a;i,j}^{(r)} = \dot{E}_{a;i,j}^{(r)}, \quad (10.12)$$

$$F_{a;i,j}^{(r)} = \dot{F}_{a;i,j}^{(r)} + \delta_{a,m-1}\left(\sum_{k=1}^{t} \tilde{e}_{N-t+i,N-t+k}\dot{F}_{a;k,j}^{(r-1)} + [\tilde{e}_{N-t+i,N-2t+h}, \dot{F}_{a;h,j}^{(r-1)}]\right), \quad (10.13)$$

where for (10.13) we are assuming that $r > 1$ if $a = m - 1$.

Combining this with the induction hypothesis and imitating the arguments in Lemmas 10.5–10.10 one can now check:

Lemma 10.12. *The following elements of $U(\mathfrak{p})$ are invariant under the twisted action of $\mathfrak{m}$:*

(i) $D_{a;i,j}^{(r)}$ *for $1 \le a \le m$, $1 \le i, j \le \nu_a$ and $r > 0$;*

(ii) $E_{a;i,j}^{(r)}$ *for $1 \le a < m$, $1 \le i \le \nu_a$, $1 \le j \le \nu_{a+1}$ and $r > s_{a,a+1}(\nu)$;*

(iii) $F_{a;i,j}^{(r)}$ *for $1 \le a < m$, $1 \le i \le \nu_{a+1}$, $1 \le j \le \nu_a$ and $r > s_{a+1,a}(\nu)$.*

Finally, define an algebra homomorphism $\varphi_{\mathrm{L}} : U(\mathfrak{p}) \to U(\mathfrak{gl}_t) \otimes U(\dot{\mathfrak{p}})$ by

$$\varphi_{\mathrm{L}}(\tilde{e}_{i,j}) = \begin{cases} \tilde{e}_{i-N+t,j-N+t} \otimes 1 & \text{if } \mathrm{col}(i) = \mathrm{col}(j) = 1, \\ 0 & \text{if } \mathrm{col}(i) = 1, 2 \le \mathrm{col}(j), \\ 1 \otimes \tilde{e}_{i,j} & \text{if } 2 \le \mathrm{col}(i) \le \mathrm{col}(j), \end{cases} \quad (10.14)$$

where for the leftmost tensor $\tilde{e}_{i-N+t,j-N+t} \in U(\mathfrak{gl}_t)$ denotes $e_{i-N+t,j-N+t} + \delta_{i,j}(n-t)$ like in Theorem 4.1. By Lemma 10.11, we have that

$$\varphi_{\mathrm{L}}(D_{a;i,j}^{(r)}) = 1 \otimes \dot{D}_{a;i,j}^{(r)} + \delta_{a,m}\sum_{k=1}^{t} \tilde{e}_{i,k} \otimes \dot{D}_{a;k,j}^{(r-1)}, \quad (10.15)$$

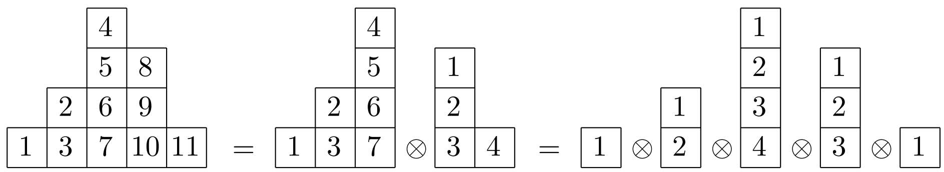
$$\varphi_{\mathrm{L}}(E_{a;i,j}^{(r)}) = 1 \otimes \dot{E}_{a;i,j}^{(r)} + \delta_{a,m-1}\sum_{k=1}^{t} \tilde{e}_{i,k} \otimes \dot{E}_{a;k,j}^{(r-1)}, \quad (10.16)$$

$$\varphi_{\mathrm{L}}(F_{a;i,j}^{(r)}) = 1 \otimes \dot{F}_{a;i,j}^{(r)}. \quad (10.17)$$

If we identify $Y_{n,l-1}(\dot{\sigma})$ with $W(\dot{\pi}) \subseteq U(\dot{\mathfrak{p}})$ using the induction hypothesis, these are the same as the images of the corresponding elements of $Y_{n,l}(\sigma)$ under the baby comultiplication Δ_{L} from Theorem 4.1(ii). So now the proof of the induction step in the case Δ_{L} can be completed like before. Theorem 10.1 is proved.

11. Grown-up comultiplication

Fix a pyramid π of height $\le n$ with column heights $(q_1, \dots, q_l)$. Throughout the section, we will work with the numbering of the bricks of π down columns from left to right. Define $\mathfrak{p}, \mathfrak{h}$, $\mathfrak{m}$ and e from (7.3), and let $W(\pi) := U(\mathfrak{p})^{\mathfrak{m}}$ be the corresponding finite W-algebra. Suppose we are given $l', l'' \geq 0$ with $l' + l'' = l$. Let π' and π'' denote the pyramids consisting just of the leftmost l' and the rightmost l'' columns of π, respectively. We write $\pi = \pi' \otimes \pi''$ whenever a pyramid is split in this way; for example,



Let $\mathfrak{p}', \mathfrak{m}', e'$ and $\mathfrak{p}'', \mathfrak{m}'', e''$ be defined from the pyramids π' and π'' inside the Lie algebras $\mathfrak{g}' = \mathfrak{gl}_{N'}$ and $\mathfrak{g}'' = \mathfrak{gl}_{N''}$, respectively. So $N = N' + N''$. Let $W(\pi') = U(\mathfrak{p}')^{\mathfrak{m}'}$ and $W(\pi'') = U(\mathfrak{p}'')^{\mathfrak{m}''}$. Recall also the elements $\tilde{e}_{i,j}$ of $U(\mathfrak{p}), U(\mathfrak{p}')$ and $U(\mathfrak{p}'')$ defined by (9.2) but working from the pyramids π, π' and π'', respectively. Define a homomorphism $\Delta_{l',l''} : U(\mathfrak{p}) \to U(\mathfrak{p}') \otimes U(\mathfrak{p}'')$ by declaring that

$$\Delta_{l',l''}(\tilde{e}_{i,j}) = \begin{cases} \tilde{e}_{i,j} \otimes 1 & \text{if } \mathrm{col}(i) \leq \mathrm{col}(j) \leq l', \\ 0 & \text{if } \mathrm{col}(i) \leq l', l'+1 \leq \mathrm{col}(j), \\ 1 \otimes \tilde{e}_{i-N',j-N'} & \text{if } l'+1 \leq \mathrm{col}(i) \leq \mathrm{col}(j), \end{cases} \tag{11.1}$$

for all $e_{i,j} \in \mathfrak{p}$. This map is obviously filtered with respect to the Kazhdan filtrations. The following lemma describes the effect of $\Delta_{l',l''}$ in terms of the elements $T_{i,j;0}^{(r)}$ of $U(\mathfrak{p})$, $U(\mathfrak{p}')$ and $U(\mathfrak{p}'')$ defined as in §9, but again working from the pyramids π, π' and π'' respectively. This should be compared with (4.1).

Lemma 11.1. *For $1 \leq i, j \leq n$ and $r > 0$,* $\Delta_{l',l''}(T_{i,j;0}^{(r)}) = \sum_{s=0}^{r} \sum_{k=1}^{n} T_{i,k;0}^{(s)} \otimes T_{k,j;0}^{(r-s)}$.

Proof. Clear from (9.5). □

One should visualize the map $\Delta_{l',l''}$ as follows. The standard embedding of $\mathfrak{g}' \oplus \mathfrak{g}''$ into $\mathfrak{g}$ also embeds $\mathfrak{p}' \oplus \mathfrak{p}''$ into $\mathfrak{p}$ and $\mathfrak{m}' \oplus \mathfrak{m}''$ into $\mathfrak{m}$. Then the map $\Delta_{l',l''}$ is just induced by the obvious projection $\mathfrak{p} \twoheadrightarrow \mathfrak{p}' \oplus \mathfrak{p}''$ followed by a constant shift. The character $\chi' \oplus \chi''$ of $\mathfrak{m}' \oplus \mathfrak{m}''$ defined by taking the trace form with $e' + e''$ is the restriction of the character χ of $\mathfrak{m}$. So the map $\Delta_{l',l''}$ sends twisted $\mathfrak{m}$-invariants in $U(\mathfrak{p})$ to twisted $(\mathfrak{m}' \oplus \mathfrak{m}'')$-invariants in $U(\mathfrak{p}') \otimes U(\mathfrak{p}'')$. This shows that the restriction of $\Delta_{l',l''}$ defines an algebra homomorphism

$$\Delta_{l',l''} : W(\pi) \to W(\pi') \otimes W(\pi''), \tag{11.2}$$

which is again a filtered map with respect to the Kazhdan filtrations. This defines a comultiplication $\Delta_{l',l''}$ between the finite W-algebras which is coassociative in the following sense:

Lemma 11.2. *If $\pi = \pi' \otimes \pi'' \otimes \pi'''$ is a pyramid, where π', π'' and π''' are of levels l', l'' and l''' respectively, then the following diagram commutes:*

$$\begin{array}{ccc} W(\pi' \otimes \pi'' \otimes \pi'') & \xrightarrow{\Delta_{l'+l'',l'''}} & W(\pi' \otimes \pi'') \otimes W(\pi''') \\ {\scriptstyle \Delta_{l',l''+l'''}} \big\downarrow & & \big\downarrow {\scriptstyle \Delta_{l',l''} \otimes 1} \\ W(\pi') \otimes W(\pi'' \otimes \pi''') & \xrightarrow{1 \otimes \Delta_{l'',l'''}} & W(\pi') \otimes W(\pi'') \otimes W(\pi''') \end{array}$$

Proof. Obvious from (11.1). □

Read off a shift matrix $\sigma = (s_{i,j})_{1 \leq i,j \leq n}$ from the pyramid π according to (7.6). ¿From now on we will identify the algebra $W(\pi)$ with the shifted Yangian $Y_{n,l}(\sigma)$ of level l using the isomorphism from Theorem 10.1. Let $t := \min(q_1, q_l)$. Suppose first that the baby comultiplication $\Delta_{\mathrm{R}} : Y_{n,l}(\sigma) \to Y_{n,l-1}(\dot{\sigma}) \otimes U(\mathfrak{gl}_t)$ from (6.8) is defined, i.e. either $t = n$ or $s_{n-t,n-t+1} \neq 0$. The pyramid corresponding to the level $(l-1)$ and the shift matrix $\dot{\sigma}$ here (which is defined by (4.5)) is simply the pyramid π' obtained from π by removing the rightmost column. So, identifying $Y_{n,l-1}(\dot{\sigma})$ with $W(\pi')$ according to Theorem 10.1 again, the map Δ_{R} is therefore identified with a map $\Delta_{\mathrm{R}} : W(\pi) \to W(\pi') \otimes U(\mathfrak{gl}_t)$, just like the map $\Delta_{l-1,1}$. Suppose instead that the baby comultiplication $\Delta_{\mathrm{L}} : Y_{n,l}(\sigma) \to U(\mathfrak{gl}_t) \otimes Y_{n,l-1}(\dot{\sigma})$ from (6.9) is defined, i.e. either $t = n$ or $s_{n-t+1,n-t} \neq 0$. This time, the pyramid corresponding to the level $(l-1)$ and the shift matrix $\dot{\sigma}$ (defined now by (4.6)) is the pyramid π'' obtained from π by removing the leftmost column. So Δ_{L} is identified with a map $W(\pi) \to U(\mathfrak{gl}_t) \otimes W(\pi'')$, just like the map $\Delta_{1,l-1}$.

Lemma 11.3. *Whenever the baby comultiplications Δ_{R} and Δ_{L} are defined, they are equal to $\Delta_{l-1,1}$ and $\Delta_{1,l-1}$, respectively.*

Proof. This is obvious in the case $l = 1$, so assume that $l > 1$. Assume the map Δ_{R} is defined. Comparing (11.1) with (10.7), it is clear that the map $\Delta_{l-1,1}$ coincides with the map φ_{R} defined in the proof of Theorem 10.1. Comparing (10.8)–(10.10) with Theorem 4.1(i), the map φ_{R} coincides with Δ_{R} under the identifications of $W(\pi)$ with $Y_{n,l}(\sigma)$ and $W(\pi')$ with $Y_{n,l-1}(\dot{\sigma})$. Hence, $\Delta_{l-1,1} = \Delta_{\mathrm{R}}$. The proof that $\Delta_{1,l-1} = \Delta_{\mathrm{L}}$ is similar, using (10.14), (10.15)–(10.17) and Theorem 4.1(ii). □

If we iterate the comultiplication (in any order by coassociativity) a total of $(l-1)$ times to split the pyramid π into its individual columns, we obtain a homomorphism

$$\mu : W(\pi) \to U(\mathfrak{gl}_{q_1}) \otimes \cdots \otimes U(\mathfrak{gl}_{q_l}). \tag{11.3}$$

Let us give a direct description of this map. The elements $\{e_{i,j}^{[r]}\}_{1 \leq r \leq l, 1 \leq i,j \leq q_r}$ defined from $e_{i,j}^{[r]} := e_{q_1+\cdots+q_{r-1}+i, q_1+\cdots+q_{r-1}+j}$ form a basis for the Levi subalgebra $\mathfrak{h}$ of $\mathfrak{p}$. Identify $U(\mathfrak{h})$ with $U(\mathfrak{gl}_{q_1}) \otimes \cdots \otimes U(\mathfrak{gl}_{q_l})$ so that $e_{i,j}^{[r]}$ is identified with $1^{\otimes(r-1)} \otimes e_{i,j} \otimes 1^{\otimes(l-r)}$. The map (11.3) is then identified with a homomorphism

$$\mu : W(\pi) \to U(\mathfrak{h}) \tag{11.4}$$

which is a filtered map with respect to the Kazhdan filtration of $W(\pi)$ and the standard filtration of $\mathrm{F}_0 U(\mathfrak{h}) \subseteq \mathrm{F}_1 U(\mathfrak{h}) \subseteq \cdots$ of $U(\mathfrak{h})$; we will write $\operatorname{gr} U(\mathfrak{h})$ for the associated graded algebra here. Let

$$\eta : U(\mathfrak{h}) \to U(\mathfrak{h}), \quad e_{i,j}^{[r]} \mapsto e_{i,j}^{[r]} + \delta_{i,j}(q_{r+1} + \cdots + q_l). \tag{11.5}$$

Let $\pi : U(\mathfrak{p}) \twoheadrightarrow U(\mathfrak{h})$ be the algebra homomorphism induced by the natural projection $\mathfrak{p} \twoheadrightarrow \mathfrak{h}$. Then, it is easy to see from (11.1) that μ is precisely the restriction of the map $\eta \circ \pi$ to $W(\pi)$. By analogy with the language used in [BT], we call μ the *Miura transform*; in [L] it is called the *generalized Harish-Chandra homomorphism*. The following result is due to Lynch [L, Corollary 2.3.2].

Theorem 11.4. *The map* $\operatorname{gr}\mu : \operatorname{gr} W(\pi) \to \operatorname{gr} U(\mathfrak{h})$ *is injective, hence so is the Miura transform* $\mu : W(\pi) \to U(\mathfrak{h})$ *itself.*

Proof. By Lemma 11.3 and the definition (11.3), we can factor μ as a composition of $(l-1)$ maps of the form Δ_{R} or Δ_{L}. Now the theorem follows from the injectivity of $\operatorname{gr}\Delta_{\mathrm{R}}$ and $\operatorname{gr}\Delta_{\mathrm{L}}$ proved in Theorem 6.2. □

Corollary 11.5. *The map* $\operatorname{gr}\Delta_{l',l''} : \operatorname{gr} W(\pi) \to \operatorname{gr} W(\pi') \otimes W(\pi'')$ *is injective for any* $l' + l'' = l$. *Hence so is the comultiplication* $\Delta_{l',l''} : W(\pi) \to W(\pi') \otimes W(\pi'')$.

Proof. We can factor μ as a composition of $\Delta_{l',l''}$ followed by $(l-2)$ more maps. □

For the next lemma, we recall from Corollary 10.3 that if π and $\dot\pi$ are two pyramids with the same row lengths, then there is a canonical isomorphism $\iota : W(\pi) \to W(\dot\pi)$. The following lemma shows that the comultiplication is compatible with these isomorphisms.

Lemma 11.6. *Suppose that* $\pi = \pi' \otimes \pi''$ *and* $\dot\pi = \dot\pi' \otimes \dot\pi''$ *are pyramids such that* $\dot\pi'$ *and* $\dot\pi''$ *have the same row lengths as* π' *and* π''*, respectively. Then the following diagram commutes:*

$$\begin{array}{ccc} W(\pi) & \xrightarrow{\Delta_{l',l''}} & W(\pi') \otimes W(\pi'') \\ {\scriptstyle\iota}\big\downarrow & & \big\downarrow{\scriptstyle\iota\otimes\iota} \\ W(\dot\pi) & \xrightarrow{\Delta_{l',l''}} & W(\dot\pi') \otimes W(\dot\pi''), \end{array}$$

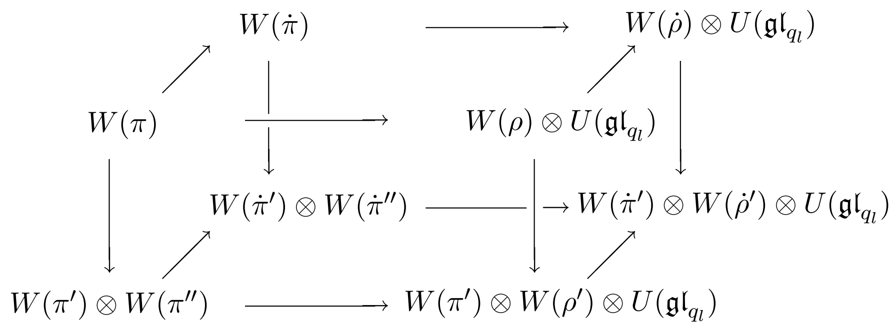
where π' *and* π'' *are of levels* l' *and* l''*, respectively.*

Proof. Proceed by induction on $l = l' + l''$. If either $l' = 0$ or $l'' = 0$, the statement of the lemma is vacuous, so the base case $l = 1$ is trivial. Now suppose that $l', l'' > 0$, so $l > 1$, and that the lemma has been proved for all smaller levels. Read off a shift matrix σ from the pyramid π and identify $W(\pi)$ with $Y_{n,l}(\sigma)$ as usual. At least one of the baby comultiplications Δ_{R} and Δ_{L} from (6.8)–(6.9) is always defined; we explain the proof of the induction step just in the case that Δ_{R} is defined, the argument being entirely similar in the other case. Also read off a shift matrix $\dot\sigma$ from the pyramid $\dot\pi$ and identify $W(\dot\pi)$ with $Y_{n,l}(\dot\sigma)$; we can ensure in doing this that the baby comultiplication Δ_{R} is defined for $Y_{n,l}(\dot\sigma)$ too.

Suppose first that $l'' = 1$. Then by Lemma 11.3, $\Delta_{l',l''}$ is equal in either case to the map Δ_{R}, and the commutativity of the diagram is easy to check explicitly on the generators of $W(\pi)$, using (3.21) and Theorem 4.1(i).

Now suppose that $l'' > 1$. Let $\rho, \dot\rho, \rho'$ and $\dot\rho'$ denote the pyramids obtained from $\pi, \dot\pi, \pi''$ and $\dot\pi''$, respectively, by removing the rightmost column (which is of height q_l

in all cases). Consider the following cube



where the maps on the front and back faces are defined from the comultiplications, and the remaining maps are isomorphisms built from ι. The front and back faces commute by Lemma 11.2. The top and bottom faces commute by the special case considered in the preceeding paragraph. The right hand face commutes by the induction hypothesis. Since the comultiplication map $W(\dot\pi') \otimes W(\dot\pi'') \to W(\dot\pi') \otimes W(\dot\rho') \otimes U(\mathfrak{gl}_{q_l})$ is injective by Corollary 11.5, it therefore follows that the left face commutes too. This completes the proof of the induction step. □

In the remainder of the section, we are going to lift the comultiplication to the shifted Yangian $Y_n(\sigma)$ itself. First, we must explain one other basic operation on pyramids, that of *column removal*. Let $1 \leq i_1 < \cdots < i_{\dot l} \leq l$ be some subset of the columns of the pyramid π, and let $\dot\pi$ be the pyramid with column heights $(q_{i_1}, \dots, q_{i_{\dot l}})$. We can read off shift matrices $\sigma = (s_{i,j})_{1 \leq i,j \leq n}$ and $\dot\sigma = (\dot s_{i,j})_{1 \leq i,j \leq n}$ from the pyramids π and π' respectively in such a way as to ensure $\dot s_{i,j} \leq s_{i,j}$ for all $1 \leq i, j \leq n$. Hence embedding $Y_n(\sigma)$ and $Y_n(\dot\sigma)$ into Y_n in the canonical way, we have that $Y_n(\sigma) \subseteq Y_n(\dot\sigma)$. This inclusion $Y_n(\sigma) \hookrightarrow Y_n(\dot\sigma)$ factors through the quotients to induce a map

$$\zeta : Y_{n,l}(\sigma) \to Y_{n,\dot l}(\dot\sigma). \tag{11.6}$$

Equivalently, identifying $Y_{n,l}(\sigma)$ with $W(\pi)$ and $Y_{n,\dot l}(\dot\sigma)$ with $W(\dot\pi)$ by Theorem 10.1, this defines a homomorphism

$$\zeta : W(\pi) \to W(\dot\pi) \tag{11.7}$$

sending the generators $D_i^{(r)}, E_i^{(r)}$ and $F_i^{(r)}$ of $W(\pi)$ for all admissible i, r to the elements $\dot D_i^{(r)}, \dot E_i^{(r)}$ and $\dot F_i^{(r)}$ of $W(\dot\pi)$, respectively.

In order to understand the relationship between this "column removal homomorphism" ζ and the comultiplication, we need another description of ζ. Let $\mu : W(\pi) \to U(\mathfrak{h})$ and $\dot\mu : W(\dot\pi) \to U(\dot{\mathfrak{h}})$ be the Miura transforms defined by (11.4). There is an obvious projection $\hat\zeta : \mathfrak{h} \twoheadrightarrow \dot{\mathfrak{h}}$ defined by

$$\hat\zeta(e_{i,j}^{[r]}) = \begin{cases} e_{i,j}^{[s]} & \text{if } r = i_s \text{ for some } s = 1, \dots, \dot l, \\ 0 & \text{otherwise.} \end{cases}$$

Lemma 11.7. *With the above notation, the following diagram commutes*

$$\begin{array}{ccc} W(\pi) & \xrightarrow{\mu} & U(\mathfrak{h}) \\ \zeta\big\downarrow & & \big\downarrow\hat\zeta \\ W(\dot\pi) & \xrightarrow{\dot\mu} & U(\dot{\mathfrak{h}}) \end{array}$$

Proof. One checks explicitly from Corollary 9.4 and the definition (9.5) that $\hat\zeta$ maps the elements $\mu(D_i^{(r)})$, $\mu(E_i^{(r)})$ and $\mu(F_i^{(r)})$ to $\dot\mu(\dot D_i^{(r)})$, $\dot\mu(\dot E_i^{(r)})$ and $\dot\mu(\dot F_i^{(r)})$. □

Corollary 11.8. *Suppose that $\pi = \pi' \otimes \pi''$ and $\dot\pi = \dot\pi' \otimes \dot\pi''$, where $\dot\pi'$ and $\dot\pi''$ are obtained by removing columns from the pyramids π' and π'', respectively. Then, the following diagram commutes*

$$\begin{array}{ccc} W(\pi) & \xrightarrow{\Delta_{l',l''}} & W(\pi') \otimes W(\pi'') \\ \zeta\big\downarrow & & \big\downarrow\zeta\otimes\zeta \\ W(\dot\pi) & \xrightarrow{\Delta_{\dot l',\dot l''}} & W(\dot\pi') \otimes W(\dot\pi'') \end{array}$$

where π', π'' are of widths l', l'' and $\dot\pi', \dot\pi''$ are of widths $\dot l', \dot l''$.

Proof. In view of Lemma 11.7 and the injectivity of the Miura transforms, this follows from the commutativity of the following diagram:

$$\begin{array}{ccc} U(\mathfrak{h}) & \xrightarrow{\sim} & U(\mathfrak{h}') \otimes U(\mathfrak{h}'') \\ \hat\zeta\big\downarrow & & \big\downarrow\hat\zeta\otimes\hat\zeta \\ U(\dot{\mathfrak{h}}) & \xrightarrow{\sim} & U(\dot{\mathfrak{h}}') \otimes U(\dot{\mathfrak{h}}'') \end{array}$$

where the horizontal maps are induced by the obvious isomorphisms $\mathfrak{h} \cong \mathfrak{h}' \oplus \mathfrak{h}''$ and $\dot{\mathfrak{h}} \cong \dot{\mathfrak{h}}' \oplus \dot{\mathfrak{h}}''$. □

Now we can prove the main theorem of the section:

Theorem 11.9. *Let σ be a shift matrix and write $\sigma = \sigma' + \sigma''$ where σ' is strictly lower triangular and σ'' is strictly upper triangular. Embedding $Y_n(\sigma), Y_n(\sigma')$ and $Y_n(\sigma'')$ into Y_n in the standard way, the restriction of the comultiplication $\Delta : Y_n \to Y_n \otimes Y_n$ gives a homomorphism*

$$\Delta : Y_n(\sigma) \to Y_n(\sigma') \otimes Y_n(\sigma'').$$

Moreover, for $l' \geq s_{n,1}, l'' \geq s_{1,n}$ and $l = l' + l''$, this map Δ factors through the quotients to define a homomorphism $Y_{n,l}(\sigma) \to Y_{n,l'}(\sigma') \otimes Y_{n,l''}(\sigma'')$ which, on identifying $Y_{n,l}(\sigma)$ with $W(\pi)$, $Y_{n,l'}(\sigma')$ with $W(\pi')$ and $Y_{n,l''}(\sigma'')$ with $W(\pi'')$ as usual, is precisely the comultiplication $\Delta_{l',l''}$ from (11.2).

Proof. Start from the map $\Delta_{l',l''} : Y_{n,l}(\sigma) \to Y_{n,l'}(\sigma') \otimes Y_{n,l''}(\sigma'')$ from (11.2). Recall from Remark 6.4 how $Y_n(\sigma), Y_n(\sigma')$ and $Y_n(\sigma'')$ are identified with the inverse limits $\varprojlim Y_{n,l}(\sigma), \varprojlim Y_{n,l'}(\sigma')$ and $\varprojlim Y_{n,l''}(\sigma'')$, respectively. Corollary 11.8 is exactly what is needed to ensure that the maps $\Delta_{l',l''}$ are stable as $l', l'' \to \infty$. Hence there is an induced homomorphism $\varprojlim \Delta_{l',l''} : Y_n(\sigma) \to Y_n(\sigma') \otimes Y_n(\sigma'')$ lifting the maps $\Delta_{l',l''}$ for all $l' \geq s_{n,1}, l'' \geq s_{1,n}$. Now we just need to show that this map $\varprojlim \Delta_{l',l''}$ agrees with the restriction of the comultiplication Δ on Y_n.

Let X be some generator $D_i^{(r)}, E_i^{(r)}$ or $F_i^{(r)}$ of $Y_n(\sigma) \subseteq Y_n$. Consider how one computes $\Delta(X)$ in practice: first one expresses X as a linear combination of monomials in the $T_{i,j}^{(r)}$ as explained in §3 (see also the explicit formulae [BK1, (5.2)–(5.4)]); then one computes $\Delta(X)$ in terms of $T_{h,k}^{(s)}$'s using the definition (4.1); finally one rewrites all $T_{h,k}^{(s)}$ appearing in the resulting expression back in terms of the generators $D_i^{(r)}, E_i^{(r)}$ and $F_i^{(r)}$. But in view of Lemma 11.1, exactly the same procedure may be used to compute the effect of $\Delta_{l',l''}$ on the image of X in $Y_{n,l}(\sigma)$ for each l. The point is that the formula expressing $T_{i,j;0}^{(r)}$ in terms of the elements $D_i^{(r)}, E_i^{(r)}$ and $F_i^{(r)}$ of $U(\mathfrak{p})$ explained in §9 is identical to the formula doing the same thing in Y_n explained in §3. It follows that $\Delta = \varprojlim \Delta_{l',l''}$. □

Remark 11.10. Using Theorem 11.9 and Lemma 11.6, one can express *all* of the comultiplications $\Delta_{l',l''} : W(\pi) \to W(\pi') \otimes W(\pi'')$ in terms of the comultiplication Δ of Y_n, as follows. Let $\dot\pi'$ denote the right-justified pyramid with the same row lengths as π', and let $\dot\pi''$ denote the left-justified pyramid with the same row lengths as π''. Let $\dot\pi := \dot\pi' \otimes \dot\pi''$. By Lemma 11.6, the comultiplication $\Delta_{l',l''} : W(\pi) \to W(\pi') \otimes W(\pi'')$ can be recovered from $\Delta_{l',l''} : W(\dot\pi) \to W(\dot\pi') \otimes W(\dot\pi'')$ using the isomorphisms ι. Finally $\Delta_{l',l''} : W(\dot\pi) \to W(\dot\pi') \otimes W(\dot\pi'')$ is one of the comultiplications described by Theorem 11.9.

Remark 11.11. We can now complete our discussion of the center $Z(W(\pi))$. Let π be a pyramid with row lengths $(p_1, \dots, p_n)$ and column heights $(q_1, \dots, q_l)$ as usual. Recall the definition of $Z_N(u) = \sum_{r=0}^N Z_N^{(r)} u^{N-r} \in U(\mathfrak{gl}_N)[u]$ from (6.15)–(6.16). ¿From Remark 8.5, there is an isomorphism $\psi : Z(U(\mathfrak{gl}_N)) \to Z(W(\pi))$ (the proof of surjectivity being deferred to [BK2]). Hence the elements $\psi(Z_N^{(1)}), \dots, \psi(Z_N^{(N)})$ are algebraically independent and generate $Z(W(\pi))$. Using the rdet formulation of the definition of $Z_N(u)$ and the description of the Miura transform μ as the composite $\eta \circ \pi$ given after (11.5), one can compute the image of $Z_N(u)$ under the map $\mu \circ \psi : Z(U(\mathfrak{gl}_N)) \hookrightarrow U(\mathfrak{gl}_{q_1}) \otimes \cdots \otimes U(\mathfrak{gl}_{q_w})$ explicitly:

$$\mu \circ \psi(Z_N(u)) = Z_{q_1}(u) \otimes Z_{q_2}(u) \otimes \cdots \otimes Z_{q_l}(u). \tag{11.8}$$

Now let σ be a shift matrix associated to π and identify $W(\pi) = Y_{n,l}(\sigma)$ according to Theorem 10.1. Consider the power series $C_n(u) \in Y_n(\sigma)[[u^{-1}]]$ from (2.22); let us also write $C_n(u)$ for the corresponding power series in the quotient $W(\pi)[[u^{-1}]]$. It is well known that the comultiplication $\Delta : Y_n \to Y_n \otimes Y_n$ maps $C_n(u)$ to $C_n(u) \otimes C_n(u)$; see e.g. [BK1, Lemma 8.1] or [NT, Proposition 1.11]. Applying Theorem 11.9 and Remark 11.10, we deduce that the comultiplication $\Delta_{l',l''} : W(\pi) \to W(\pi') \otimes W(\pi'')$ from (11.2) also has the property that

$$\Delta_{l',l''}(C_n(u)) = C_n(u) \otimes C_n(u), \tag{11.9}$$

identifying $W(\pi')$ resp. $W(\pi'')$ with $Y_{n,l'}(\sigma')$ resp. $Y_{n,l''}(\sigma'')$ for suitable shift matrices σ' resp. σ''. Iterating (11.9) a total of $(l-1)$ times and using the definition (11.3) of the Miura transform combined with (6.14) and Remarks 6.6 and 6.5, we deduce that

$$(u+n-1)^{p_1}(u+n-2)^{p_2} \cdots u^{p_n} \mu(C_n(u)) = Z_{q_1}(u) \otimes Z_{q_2}(u) \otimes \cdots \otimes Z_{q_l}(u). \tag{11.10}$$

Comparing with (11.8), we have proved the following identity written in $W(\pi)[u]$:

$$(u+n-1)^{p_1}(u+n-2)^{p_2} \cdots u^{p_n} C_n(u) = \psi(Z_N(u)). \tag{11.11}$$

Hence the elements $C_n^{(1)}, C_n^{(2)}, \dots$ of $Y_{n,l}(\sigma)$ must also generate $Z(Y_{n,l}(\sigma))$. Since this is true for all levels l, Remark 6.4 implies easily that the elements $C_n^{(1)}, C_n^{(2)}, \dots$ of $Y_n(\sigma)$ itself generate $Z(Y_n(\sigma))$, and the quotient map $Y_n(\sigma) \twoheadrightarrow Y_{n,l}(\sigma)$ maps $Z(Y_n(\sigma))$ surjectively onto $Z(Y_{n,l}(\sigma))$.

12. A special case

In this section, we give a much more direct proof of the main results of the article in the special case that the nilpotent matrix e consists of n Jordan blocks all of the same size l, i.e. the pyramid π is an $n \times l$ rectangle and $N = nl$. The main theorem in this case was first noticed by Ragoucy and Sorba [RS]; see also [BR] which is closer to the present exposition. We will identify $\mathfrak{g} = \mathfrak{gl}_N$ with the tensor product $\mathfrak{gl}_l \otimes \mathfrak{gl}_n$ so that the matrix unit $e_{l(r-1)+i,l(s-1)+j} \in \mathfrak{gl}_N$ is identified with $e_{r,s} \otimes e_{i,j} \in \mathfrak{gl}_l \otimes \mathfrak{gl}_n$ for $1 \leq r, s \leq l, 1 \leq i, j \leq n$. Numbering the bricks of the pyramid π down columns from left to right as usual, the $\mathfrak{sl}_2$-triple (e, h, f) introduced at the beginning of §8 is given explicitly in this case by

$$e = \sum_{r=1}^{l-1} e_{r,r+1} \otimes I_n, \quad h = \sum_{r=1}^{l} (l+1-2r) e_{r,r} \otimes I_n, \quad f = \sum_{r=1}^{l-1} r(l-r) e_{r+1,r} \otimes I_n.$$

Let M_n denote the algebra of $n \times n$ matrices over $\mathbb{C}$ and let $T(\mathfrak{gl}_l)$ denote the tensor algebra on the vector space $\mathfrak{gl}_l$. Define an algebra homomorphism

$$T : T(\mathfrak{gl}_l) \to M_n \otimes U(\mathfrak{gl}_N) \tag{12.1}$$

sending a generator $x \in \mathfrak{gl}_l$ to $T(x) = \sum_{i,j=1}^n e_{i,j} \otimes (x \otimes e_{i,j}) \in M_n \otimes U(\mathfrak{gl}_N)$. Also define maps

$$T_{i,j} : T(\mathfrak{gl}_l) \to U(\mathfrak{gl}_N) \tag{12.2}$$

for each $1 \leq i, j \leq n$ from the equation $T(x) = \sum_{i,j=1}^n e_{i,j} \otimes T_{i,j}(x)$, for any $x \in T(\mathfrak{gl}_l)$. Thus, thinking of $T(x)$ as an $n \times n$ matrix with entries in $U(\mathfrak{gl}_N)$, $T_{i,j}(x)$ is the ij-entry of $T(x)$. We have by definition that $T(xy) = T(x)T(y)$ for any $x, y \in T(\mathfrak{gl}_l)$, which implies that

$$T_{i,j}(xy) = \sum_{k=1}^n T_{i,k}(x) T_{k,j}(y). \tag{12.3}$$

In particular, $T_{i,j}(1) = \delta_{i,j}$ and

$$T_{i,j}(x_1 \otimes \cdots \otimes x_r) = \sum_{\substack{1 \leq i_0, i_1, \dots, i_r \leq n \\ i_0 = i, i_r = j}} (x_1 \otimes e_{i_0,i_1})(x_2 \otimes e_{i_1,i_2}) \cdots (x_r \otimes e_{i_{r-1},i_r}) \tag{12.4}$$

for arbitrary $x_1, \dots, x_r \in \mathfrak{gl}_l$. If u is an indeterminate, we will also write $T_{i,j}$ for the map obtained from $T_{i,j}$ by extending scalars from $\mathbb{C}$ to $\mathbb{C}[u]$.

Lemma 12.1. *For $1 \leq h, i, j, k \leq n$ and $x, y_1, \dots, y_r \in \mathfrak{gl}_l$,*

$$\begin{aligned}[T_{i,j}(x), T_{h,k}(y_1 \otimes \cdots \otimes y_r)] = &\sum_{s=1}^r T_{h,j}(y_1 \otimes \cdots \otimes y_{s-1}) T_{i,k}(x y_s \otimes y_{s+1} \otimes \cdots \otimes y_r) \\ &- \sum_{s=1}^r T_{h,j}(y_1 \otimes \cdots \otimes y_{s-1} \otimes y_s x) T_{i,k}(y_{s+1} \otimes \cdots \otimes y_r),\end{aligned}$$

where xy_s and $y_s x$ denote the usual products of matrices in M_l.

Proof. Use (12.3) and induction on r. □

Define $\Omega_l(u)$ to be the following $l \times l$ matrix with entries in $T(\mathfrak{gl}_l)[u]$:

$$\begin{pmatrix} e_{1,1}+u-(l-1)n & e_{1,2} & e_{1,3} & \cdots & e_{1,l} \\ 1 & e_{2,2}+u-(l-2)n & & & \vdots \\ 0 & & \ddots & & e_{l-2,l} \\ \vdots & & 1 & e_{l-1,l-1}+u-n & e_{l-1,l} \\ 0 & \cdots & 0 & 1 & e_{l,l}+u \end{pmatrix}.$$

This matrix should be compared with (6.15). Recalling the definitions (3.22) and (12.2), define

$$T_{i,j}(u) = \sum_{r=0}^{l} T_{i,j}^{(r)} u^{l-r} := T_{i,j}(\operatorname{rdet}\Omega_l(u)) \tag{12.5}$$

for each $i,j = 1,\dots,n$. It is easy to see for $r = 0,1,\dots,l$ that $T_{i,j}^{(r)}$ is precisely the element $T_{i,j;0}^{(r)}$ of $\mathrm{F}_r U(\mathfrak{p})$ defined in (9.6). We also set $T_{i,j}^{(r)} := 0$ for $r > l$. In the next lemma, we check directly that these elements belong to the algebra $W(\pi) = U(\mathfrak{p})^{\mathfrak{m}}$ introduced in §8.

Lemma 12.2. *For each $1 \le i,j \le n$ and $r > 0$, $T_{i,j}^{(r)}$ is invariant under the twisted action of $\mathfrak{m}$.*

Proof. Since $\mathfrak{m}$ is generated by elements of the form $e_{r+1,r} \otimes e_{i,j}$, it suffices to show that $\mathrm{pr}_\chi([T_{i,j}(e_{r+1,r}), T_{h,k}(\operatorname{rdet}\Omega_l(u))]) = 0$ for every $1 \le h,i,j,k \le n$ and $r = 1,\dots,l-1$. Let us write $\Omega_{[r,s]}(u)$ for the submatrix of $\Omega_l(u)$ consisting only of rows and columns numbered $r,\dots,s$. We compute using Lemma 12.1 to get that

$$\begin{aligned} &[T_{i,j}(e_{r+1,r}), T_{h,k}(\operatorname{rdet}\Omega_l(u))] = T_{h,j}(\operatorname{rdet}\Omega_{[1,r-1]}(u)) \times \\ &\qquad T_{i,k}\left(\operatorname{rdet}\begin{pmatrix} e_{r+1,r} & e_{r+1,r+1} & \cdots & e_{r+1,l} \\ 1 & e_{r+1,r+1}+u-(l-r-1)n & \cdots & e_{r+1,l} \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 1 & e_{l,l}+u \end{pmatrix}\right) \\ &\qquad - T_{h,j}\left(\operatorname{rdet}\begin{pmatrix} e_{1,1}+u-(l-1)n & \cdots & e_{1,r} & e_{1,r} \\ 1 & \ddots & & \vdots \\ \vdots & & e_{r,r}+u-(l-r)n & e_{r,r} \\ 0 & \cdots & 1 & e_{r+1,r} \end{pmatrix}\right) \\ &\qquad\qquad \times T_{i,k}(\operatorname{rdet}\Omega_{[r+2,l]}(u)). \end{aligned}$$

In order to apply pr_χ to the right hand side, we observe that for any $1 \le m \le n$,

$$\mathrm{pr}_\chi\left(T_{i,m}\left(e_{r+1,r}\left(e_{r+1,r+1}+u-(l-r-1)n\right)\right)\right) = T_{i,m}\left(e_{r+1,r+1}+u-(l-r)n\right).$$

Hence, we get that

$$\begin{aligned}\mathrm{pr}_\chi([T_{i,j}(e_{r+1,r}), T_{h,k}(\mathrm{rdet}\,\Omega_l(u))]) = T_{h,j}(\mathrm{rdet}\,\Omega_{[1,r-1]}(u))\times\\ T_{i,k}\left(\mathrm{rdet}\begin{pmatrix} 1 & e_{r+1,r+1} & \cdots & e_{r+1,l}\\ 1 & e_{r+1,r+1}+u-(l-r)n & \cdots & e_{r+1,l}\\ \vdots & & \ddots & \vdots\\ 0 & \cdots & 1 & e_{l,l}+u\end{pmatrix}\right)\\ - T_{h,j}\left(\mathrm{rdet}\begin{pmatrix} e_{1,1}+u-(l-1)n & \cdots & e_{1,r} & e_{1,r}\\ 1 & \ddots & & \vdots\\ \vdots & & e_{r,r}+u-(l-r)n & e_{r,r}\\ 0 & \cdots & 1 & 1\end{pmatrix}\right)\\ \times T_{i,k}(\mathrm{rdet}\,\Omega_{[r+2,l]}(u)).\end{aligned}$$

Making the obvious row and column operations gives that

$$\begin{aligned}\mathrm{rdet}\begin{pmatrix} 1 & e_{r+1,r+1} & \cdots & e_{r+1,l}\\ 1 & e_{r+1,r+1}+u-(l-r)n & \cdots & e_{r+1,l}\\ \vdots & & \ddots & \vdots\\ 0 & \cdots & 1 & e_{l,l}+u\end{pmatrix}\\ = (u-(l-r)n)\,\mathrm{rdet}\,\Omega_{[r+2,l]}(u),\\ \mathrm{rdet}\begin{pmatrix} e_{1,1}+u-(l-1)n & \cdots & e_{1,r} & e_{1,r}\\ 1 & \ddots & & \vdots\\ \vdots & & e_{r,r}+u-(l-r)n & e_{r,r}\\ 0 & \cdots & 1 & 1\end{pmatrix}\\ = (u-(l-r)n)\,\mathrm{rdet}\,\Omega_{[1,r-1]}(u).\end{aligned}$$

Now substituting these into the above formula for $\mathrm{pr}_\chi([T_{i,j}(e_{r+1,r}), T_{h,k}(\mathrm{rdet}\,\Omega_l(u))])$ shows that it is zero. □

Denoting the matrix unit $e_{r,r}\otimes e_{i,j}$ instead by $e_{i,j}^{[r]}$, the subalgebra $\mathfrak{h}\cong\mathfrak{gl}_n^{\oplus l}$ of $\mathfrak{g}$ has basis $\{e_{i,j}^{[r]}\}_{1\le r\le l, 1\le i,j\le n}$. Let

$$\eta: U(\mathfrak{h})\to U(\mathfrak{h}),\quad e_{i,j}^{[r]}\mapsto e_{i,j}^{[r]}+\delta_{i,j}(l-r)n. \tag{12.6}$$

Let $\pi: U(\mathfrak{p})\twoheadrightarrow U(\mathfrak{h})$ be the algebra homomorphism induced by the natural projection $\mathfrak{p}\twoheadrightarrow\mathfrak{h}$. The composite $\mu := \eta\circ\pi$ is precisely the Miura transform from (11.4).

Lemma 12.3. *For $1\le i,j\le n$ and $r>0$, the Miura transform μ maps the element $T_{i,j}^{(r)}$ of $U(\mathfrak{p})$ to the element*

$$\sum_{1\le s_1<\cdots<s_r\le l}\ \sum_{\substack{1\le i_0,\cdots,i_r\le n\\ i_0=i, i_r=j}} e_{i_0,i_1}^{[s_1]}e_{i_1,i_2}^{[s_2]}\cdots e_{i_{r-1},i_r}^{[s_r]} \tag{12.7}$$

of $U(\mathfrak{h})$.

Proof. Applying the map $e_{r,s} \mapsto \delta_{r,s} e_{r,r} + (l-r)n$ to the matrix $\Omega_l(u)$ gives a diagonal matrix with determinant $(u+e_{1,1})(u+e_{2,2})\cdots(u+e_{r,r})$. The u^{l-r}-coefficient of this is the (non-commutative) elementary symmetric function $\sum_{1\leq s_1<\cdots<s_r\leq l} e_{s_1,s_1}\cdots e_{s_r,s_r}$. Now the lemma follows from (12.4). □

Finally let $Y_{n,l}$ denote the Yangian of level l. This may be defined as the algebra on generators $\{T_{i,j}^{(r)}\}_{1\leq i,j\leq n, r>0}$ subject to the relations (3.1) and in addition $T_{i,j}^{(r)} = 0$ for all $1 \leq i,j \leq n$ and $r > l$. We will work now with the canonical filtration on $Y_{n,l}$ and the Kazhdan filtration of $W(\pi)$, as defined in sections 5 and 8 respectively. The main theorem in our special case is as follows.

Theorem 12.4. *There is an isomorphism $Y_{n,l} \xrightarrow{\sim} W(\pi)$ of filtered algebras such that the generators $\{T_{i,j}^{(r)}\}_{1\leq i,j\leq n,r>0}$ of $Y_{n,l}$ map to the elements of $W(\pi)$ with the same names. Moreover, the Miura transform $\mu : W(\pi) \to U(\mathfrak{h})$ is injective, and the map θ from the diagram (8.4) is an isomorphism.*

Proof. By the PBW theorem for $Y_{n,l}$ proved in [BK1, Theorem 3.1], the set of all monomials in the elements $\{T_{i,j}^{(r)}\}_{1\leq i,j\leq n, 1\leq r\leq l}$ taken in some fixed order and of total degree $\leq d$ form a basis for $\mathrm{F}_d Y_{n,l}$. Moreover, *loc. cit.* shows that there is an injective algebra homomorphism $\kappa_l : Y_{n,l} \hookrightarrow U(\mathfrak{h})$ mapping $T_{i,j}^{(r)} \in Y_{n,l}$ to precisely the element (12.7). We also note from Lemma 7.3 that for each $d \geq 0$,

$$\dim \kappa_l(\mathrm{F}_d Y_{n,l}) = \dim \mathrm{F}_d Y_{n,l} = \dim \mathrm{F}_d S(\mathfrak{c}_{\mathfrak{g}}(e)), \tag{12.8}$$

where $\mathrm{F}_d S(\mathfrak{c}_{\mathfrak{g}}(e))$ denotes the sum of all the graded pieces of $S(\mathfrak{c}_{\mathfrak{g}}(e))$ of degree $\leq d$ for the Kazhdan grading.

Let X_d denote the subspace of $U(\mathfrak{p})$ spanned by all monomials in the elements $\{T_{i,j}^{(r)}\}_{1\leq i,j\leq n, 0<r\leq l}$ taken in some fixed order and of total degree $\leq d$. By Lemma 12.3 and the previous paragraph, $\mu(X_d) = \kappa_l(\mathrm{F}_d Y_{n,l})$. Hence, since we know by Lemma 12.2 that $X_d \subseteq \mathrm{F}_d W(\pi)$, we get by (12.8) and Corollary 8.3 that

$$\dim \mathrm{F}_d S(\mathfrak{c}_{\mathfrak{g}}(e)) = \dim \mu(X_d) \leq \dim X_d \leq \dim \mathrm{F}_d W(\pi) \leq \dim \mathrm{F}_d S(\mathfrak{c}_{\mathfrak{g}}(e)).$$

Hence equality holds everywhere, which means that $X_d = \mathrm{F}_d W(\pi)$, the Miura transform μ is injective with the same image as κ_l, and the map θ is an isomorphism. Hence the composite $\mu^{-1} \circ \kappa_l$ gives the required isomorphism $Y_{n,l} \to W(\pi)$ of filtered algebras. □

References

[BT] J. de Boer and T. Tjin, Quantization and representation theory of finite W-algebras, *Comm. Math. Phys.* **158** (1993), 485–516.

[BR] C. Briot and E. Ragoucy, RTT presentation of finite W-algebras, *J. Phys. A* **34** (2001), 7287-7310.

[BK1] J. Brundan and A. Kleshchev, Parabolic presentations of the Yangian $Y(\mathfrak{gl}_n)$, preprint, University of Oregon, 2004.

[BK2] J. Brundan and A. Kleshchev, Representations of shifted Yangians, preprint, University of Oregon, 2004.

[Ca] R. W. Carter, *Finite groups of Lie type: conjugacy classes and complex characters*, Wiley, 1985.

[CL] R. W. Carter and G. Lusztig, On the modular representations of the general linear and symmetric groups, *Math. Z.* **136** (1974), 193–242.

[C] I. Cherednik, Quantum groups as hidden symmetries of classic representation theory, in: "Differential geometric methods in theoretical physics (Chester, 1988)", pp. 47–54, World Sci. Publishing, 1989.

[D] V. Drinfeld, A new realization of Yangians and quantized affine algebras, *Soviet Math. Dokl.* **36** (1988), 212–216.

[EK] P. Elashvili and V. Kac, Classification of good gradings of simple Lie algebras, *Amer. Math. Soc. Transl.*, to appear.

[GG] W. L. Gan and V. Ginzburg, Quantization of Slodowy slices, *Int. Math. Res. Notices* **5** (2002) 243–255.

[HU] R. Howe and T. Umeda, The Capelli identity, the double commutant theorem and multiplicity-free actions, *Math. Ann.* **290** (1991), 569–619.

[KRW] V. Kac, S. Roan and M. Wakimoto, Quantum reduction for affine superalgebras, *Comm. Math. Phys.* **241** (2003), 307–342.

[Ka] N. Kawanaka, Generalized Gelfand-Graev representations and Ennola duality, in: "Algebraic groups and related topics", *Adv. Studies in Pure Math.* **6** (1985), 175–206.

[K] B. Kostant, On Whittaker modules and representation theory, *Invent. Math.* **48** (1978), 101–184.

[Kr] H. Kraft, Parametrisierung von Konjugationklassen in $\mathfrak{sl}_n$, *Math. Ann.* **234** (1978), 209-220.

[L] T. E. Lynch, *Generalized Whittaker vectors and representation theory*, PhD thesis, M.I.T., 1979.

[M] C. Moeglin, Modèles de Whittaker et idéaux primitifs complètement premier dans les algèbres enveloppantes I, *C. R. Acad. Sci. Paris* **303** (1986), 845–848; II, *Math. Scand.* **63** (1988), 5–35.

[MNO] A. Molev, M. Nazarov and G. Olshanskii, Yangians and classical Lie algebras, *Russian Math. Surveys* **51** (1996), 205–282.

[NT] M. Nazarov and V. Tarasov, Representations of Yangians with Gelfand-Zetlin bases, *J. reine angew. Math.* **496** (1998), 181–212.

[P] A. Premet, Special transverse slices and their enveloping algebras, *Advances Math.* **170** (2002), 1–55.

[RS] E. Ragoucy and P. Sorba, Yangian realisations from finite W-algebras, *Comm. Math. Phys.* **203** (1999), 551–572.

[Sk] S. Skryabin, A category equivalence, appendix to [P].

[S] P. Slodowy, Simple singularities and simple algebraic groups, Lecture Notes in Math., vol. 815, Springer, 1980.

[SS] T. Springer and R. Steinberg, "Conjugacy classes", in: *Seminar on algebraic groups and related topics*, eds: A. Borel et al., Lecture Notes in Math. 131, Springer, 1970.

Department of Mathematics, University of Oregon, Eugene, OR 97403, USA

E-mail address: `brundan@darkwing.uoregon.edu, klesh@math.uoregon.edu`