FINITE W-SUPERALGEBRAS VIA SUPER YANGIANS

YUNG-NING PENG

Abstract. Let e be an arbitrary even nilpotent element in the general linear Lie superalgebra $\mathfrak{gl}_{M|N}$ and let $\mathcal{W}_e$ be the associated finite W-superalgebra. Let $Y_{m|n}$ be the super Yangian associated to the Lie superalgebra $\mathfrak{gl}_{m|n}$. A subalgebra of $Y_{m|n}$, called the shifted super Yangian and denoted by $Y_{m|n}(\sigma)$, is defined and studied. Moreover, an explicit isomorphism between $\mathcal{W}_e$ and a quotient of $Y_{m|n}(\sigma)$ is established.

Contents

1. Introduction

A finite W-algebra is an associative algebra determined by a pair $(\mathfrak{g}, e)$, where $\mathfrak{g}$ is a finite dimensional semisimple or reductive Lie algebra and e is a nilpotent element in $\mathfrak{g}$. In the extreme case when $e = 0$, the corresponding finite W-algebra is the universal enveloping algebra $U(\mathfrak{g})$. In the other extreme case when e is the *principal* (also called *regular*) nilpotent element, Kostant [25] proved that the associated finite W-algebra is isomorphic to the center of the universal enveloping algebra.

The study of finite W-algebra for a general e was firstly developed systematically by Premet [36], in which the modern terminologies were given and a proof of the long-standing

Key words and phrases. finite W-superalgebra, shifted super Yangian.

Kac-Weisfeiler conjecture [44] was established. Moreover, finite W-algebras can be understood as quantizations of Slodowy slices [20, 37]. Since then, finite W-algebras have appeared in many branches of mathematics so that their behavior and properties can be explained from different viewpoints. In recent years, the finite W-algebras have been intensively studied by various approaches; see the survey articles [1, 26, 41] for details.

On the other hand, Yangians are certain non-commutative Hopf algebras that are important examples of quantum groups. They first appeared in physics in the work of Faddeev and his school around 80's concerning the quantum inverse scattering method. The term Yangian was given by Drinfeld [14] in honor of C.N. Yang and had been commonly used since then. They were used to provide rational solutions of the Yang-Baxter equation; see the book [27] for related topics and further applications of Yangians.

The connection between Yangians and finite W-algebras was firstly noticed by Ragoucy and Sorba [38] for type A Lie algebras. Suppose that the nilpotent element e is *rectangular*, which means that all the Jordan blocks of e are of the same size, say ℓ. They showed that the associated finite W-algebra is isomorphic to the *Yangian of level* ℓ, which is a certain quotient of the Yangian, considered by Cherednik [12, 13].

This observation is further generalized by Brundan and Kleshchev [8] to an arbitrary nilpotent $e \in \mathfrak{gl}_N$. Their main result [8, Theorem 10.1] can be shortly described as follows: the finite W-algebra associated to a nilpotent $e \in \mathfrak{gl}_N$ is isomorphic to a quotient of some subalgebra of the Yangian (called the *shifted Yangian*) associated to $\mathfrak{gl}_n$, where n is the number of Jordan blocks of e. Moreover, an explicit realization, by generators and relations, of type A finite W-algebra is obtained. This provides a powerful tool for the study of finite W-algebras, including their representations and further applications [6, 9, 10]. It is also observed recently that the shifted Yangian can also be defined by different approaches together with new generalizations and applications; see [2, 17, 18, 24].

The finite W-superalgebras are defined in a very similar way as the Lie algebra case except that the nilpotent element $e \in \mathfrak{g}$ is assumed to be *even* (with respect to the $\mathbb{Z}_2$-grading of the Lie superalgebra) with other modifications. In recent years, finite W-superalgebras and their representations have been extensively studied [4, 6, 42, 43, 45, 46, 47] with different emphases.

The super Yangian associated to $\mathfrak{gl}_{m|n}$, denoted by $Y_{m|n}$, was defined by Nazarov [28] in terms of the *RTT presentation*. It is natural to seek for connections between finite W-superalgebras and super Yangians. The very first result is obtained by Briot and Ragoucy [3], saying that if the nilpotent element $e \in \mathfrak{gl}_{M|N}$ is rectangular, then the associated finite W-superalgebra is isomorphic to a certain quotient of $Y_{m|n}$ called the *truncated super Yangian*, where m and n are the numbers of Jordan blocks of e restricted to the even and odd spaces, respectively. In recent years, there have been some results [4, 31, 32] generalizing the above observation when the nilpotent element e satisfies some assumptions, but for a general e the problem remains to be open.

The goal of this article is to give a solution to this open problem, generally establishing the connection between the finite W-superalgebras and super Yangians for type A. That is, we explicitly give a superalgebra isomorphism between the finite W-superalgebra associated to an *arbitrary* even nilpotent element $e \in \mathfrak{gl}_{M|N}$ and a quotient of a certain subalgebra of $Y_{m|n}$, obtaining a super analogue of the main result of [8] for type A Lie superalgebras in full generality.

We shortly explain our approach, which is basically generalizing the arguments in [8] to the general linear Lie superalgebras with suitable modifications and trying to overcome all of the difficulties along the way. Although there are similarities between $\mathfrak{gl}_N$ and $\mathfrak{gl}_{M|N}$ and similarities between the associated (super) Yangians, some of the earlier approaches are no longer available in the case of Lie superalgebras. Moreover, there are other technical or conceptual obstacles that did not appear in the Lie algebra case, and the messy parities make the computations rather complicated. Therefore, although the methods in [8] are well-established, such a generalization is still by no means trivial.

Our first step is to define a subalgebra of $Y_{m|n}$ which we call the *shifted super Yangian* and denote by $Y_{m|n}(\sigma)$. To obtain this subalgebra, we need to use certain presentations of $Y_{m|n}$ called the *parabolic presentations*. Similar to the Lie algebra case [7, 15], the RTT presentation and the Drinfeld's presentation can be treated as special cases of the parabolic presentations. There have been some results [22, 30] giving suitable presentations of $Y_{m|n}$, where the results [4, 32] are in fact heavily based on them. However, as noticed in [4, 32], they are no longer suitable presentations for the general case. What we need is a further generalized parabolic presentation which works for *any* 01-sequence [11, 19], which is a parametrizing set controlling the parities of elements in $Y_{m|n}$. Such a presentation was not available until the recent paper [33]. As a consequence of [33], the shifted super Yangian $Y_{m|n}(\sigma)$ can be defined as a subalgebra of $Y_{m|n}$ generated by a certain subset of the generating set for the whole $Y_{m|n}$.

However, to establish the desired connection, we need not only the subalgebra but also its *presentation*. By suitably modifying the defining relations for $Y_{m|n}$ found in [33], we obtain a set of defining relations and hence a presentation of the shifted super Yangian $Y_{m|n}(\sigma)$. It should be emphasized that there are two extra series of defining relations, (5.17) and (5.18), for $Y_{m|n}(\sigma)$ that did not appear in [4, 8, 32]. Although we are able to guess the suitable modifications, it is highly non-trivial to check that our proposed relations actually hold in $Y_{m|n}$.

Inspired from an induction argument in [9], together with a recent observation [40, Remark 2.13] which fulfills the initial step, one can eventually overcome this difficulty and a presentation of $Y_{m|n}(\sigma)$ is obtained. This further allows one to define some homomorphisms called *baby comultiplications* that will play important roles when establishing the desired connection.

We further define the *shifted super Yangian of level* ℓ, denoted by $Y^{\ell}_{m|n}(\sigma)$, as a quotient of $Y_{m|n}(\sigma)$ over some 2-sided ideal. Let's explain the meaning of the parameters in our notation. Roughly speaking, σ is a matrix recording the generating set for $Y_{m|n}(\sigma)$, while ℓ is a positive integer recording the size of the ideal we quotient out. It turns out that the data σ and ℓ can be recorded by a diagram called *pyramid* [16, 23], which we denote by π, and it makes sense to set the notation $Y_\pi := Y^{\ell}_{m|n}(\sigma)$. On the other hand, the diagram π also determines a finite W-superalgebra which we denote by $\mathcal{W}_\pi$.

In §9, we introduce the notion of *super column height* so that one may explicitly write down some distinguished elements in $\mathcal{W}_\pi$ according to the diagram π by modifying the description in [8, §9]. Our main result, Theorem 10.1, shows that the map sending the generators of Y_π to some of these distinguished elements in $\mathcal{W}_\pi$ is an isomorphism of (filtered) superalgebras, obtaining a presentation of the finite W-superalgebra $\mathcal{W}_\pi$.

It is an interesting question to generalize the results in this article to other types of Lie superalgebras. In particular, there have been some results in the case of *queer Lie superalgebras* and their associated Yangians [29] when the even nilpotent element is regular [34] or rectangular [35], but it is still open in general. We expect that the approaches in this article can be suitably modified to deal with the queer Lie superalgebra case for a general nilpotent element.

This article is organized as follows. In §2, we set up our notations and recall some necessary background knowledge about finite W-superalgebras. In particular, the notion of pyramid with respect to a 01-sequence is recalled. In §3, we recall some well-known facts about $Y_{m|n}$.

The shifted super Yangian $Y_{m|n}(\sigma)$ is defined in §4 by generators and relations, with the use of Drinfeld's presentation for $Y_{m|n}$, where some computations are relatively easier in this setting. Then we show that $Y_{m|n}(\sigma)$ can be identified as a subalgebra of $Y_{m|n}$. Some basic properties of $Y_{m|n}(\sigma)$ are also derived.

In §5 we provide a more general approach, using the parabolic presentations for $Y_{m|n}$, to define $Y_{m|n}(\sigma)$ and establish the corresponding properties obtained in §4 to parabolic case. In particular, the results in §4 serve as initial steps of some induction arguments in the parabolic case.

§6 is devoted to define the baby comultiplications that will help us establish the main result later. We explicitly write down their formulas and show that they are injective whenever they are defined.

In §7, we introduce the canonical filtration of $Y_{m|n}(\sigma)$, which eventually corresponds to the Kazhdan filtration of finite W-superalgebras. The shifted super Yangian of level ℓ is defined in §8 as a quotient of $Y_{m|n}(\sigma)$.

In §9, we explicitly define some distinguished elements in the universal enveloping algebra $U(\mathfrak{gl}_{M|N})$ that will eventually be identified as generators of our finite W-superalgebra. Our main result is stated and proved in §10.

In this article, our field is the field of complex numbers $\mathbb{C}$, which can be replaced by any algebraically closed field of characteristic zero. The terms *subalgebra* and *subspace* always mean a *sub-superalgebra* and a *sub-superspace*, respectively. For homogeneous elements x and y in an associative superalgebra A, the *supercommutator* of x and y is defined by

$$[x, y] = xy - (-1)^{|x||y|} yx,$$

where $|x|$ is the $\mathbb{Z}_2$-grading of x in A, called the *parity* of x. By convention, a homogeneous element x is called *even* (resp. *odd*) if $|x| = \overline{0}$ (resp. $\overline{1}$). We let $A_{\overline{0}}$ and $A_{\overline{1}}$ denote the set of even and odd elements in A, respectively. For associative superalgebras A and B, their tensor product $A \otimes B$ is again considered as a superalgebra by the product

$$(x \otimes y)(a \otimes b) := (-1)^{|y||a|} xa \otimes yb$$

for homogeneous $x, a \in A$ and $y, b \in B$.

Acknowledgments. The author is grateful to Shun-Jen Cheng and Weiqiang Wang for countless discussions and encouragement. A part of this article was finished during the author's visit to the RIMS (Kyoto, Japan) in 2016. The author would like to thank the RIMS for providing an excellent working environment, and also thank Naoki Genra, Ryosuke Kodera and Hiraku Nakajima for stimulating discussions during the visit. The visit is supported by the NCTS (Taipei, Taiwan), which is greatly acknowledged. The author would also like to thank Lucy Gow, Alexander Molev for communication, and particularly thank Alexander Tsymbaliuk for reading an earlier draft of the paper and providing helpful suggestions. The author is partially supported by MOST grant 105-2628-M-008-004-MY4.

2. Finite W-superalgebras and pyramids

In this section, we recall the definition of a finite W-superalgebra, which is determined by an even nilpotent element e and a semisimple element h of $\mathfrak{gl}_{M|N}$. Also, a combinatorial object called *pyramid* is recalled so that we may simultaneously encode e and h simply by a diagram π.

Throughout this section, $\mathfrak{g} = \mathfrak{gl}_{M|N}$ is identified with the set of $(M+N) \times (M+N)$ matrices with the standard $\mathbb{Z}_2$-grading $\mathfrak{g} = \mathfrak{g}_{\overline{0}} \oplus \mathfrak{g}_{\overline{1}}$ and $(\cdot\,,\cdot)$ means the non-degenerate even supersymmetric $\mathfrak{g}$-invariant bilinear form on $\mathfrak{g}$ defined by

$$(x, y) := \mathrm{str}(xy)$$

for all $x, y \in \mathfrak{g}$, where xy stands for the usual matrix product and str means the supertrace. All elements of $\mathfrak{g}$ appearing in any equations are considered homogeneous with respect to the $\mathbb{Z}_2$-grading unless specifically mentioned.

2.1. **Finite W-superalgebras of $\mathfrak{gl}_{M|N}$.** Let e be an even nilpotent element in $\mathfrak{g}$. It is well-known [23, 41] that there exists (not uniquely in general) a semisimple element $h \in \mathfrak{g}$ such that $\operatorname{ad} h : \mathfrak{g} \to \mathfrak{g}$ gives a *good $\mathbb{Z}$-grading of $\mathfrak{g}$ for e*, which means the following conditions are satisfied:

(1) $\operatorname{ad} h(e) = 2e$,
(2) $\mathfrak{g} = \bigoplus_{j\in\mathbb{Z}} \mathfrak{g}(j)$, where $\mathfrak{g}(j) := \{x \in \mathfrak{g} | \operatorname{ad} h(x) = jx\}$,
(3) the center of $\mathfrak{g}$ is contained in $\mathfrak{g}(0)$,
(4) $\operatorname{ad} e : \mathfrak{g}(j) \to \mathfrak{g}(j+2)$ is injective for all $j \leq -1$,
(5) $\operatorname{ad} e : \mathfrak{g}(j) \to \mathfrak{g}(j+2)$ is surjective for all $j \geq -1$.

In order to simplify the definition of finite W-superalgebras, throughout this article, we assume in addition that the $\mathbb{Z}$-grading is *even*; that is, $\mathfrak{g}(j) = 0$ for all $j \notin 2\mathbb{Z}$. We say $\langle e, h\rangle$ is a *good pair* if $\operatorname{ad} h$ gives an even good $\mathbb{Z}$-grading of $\mathfrak{g}$ for e.

Remark 2.1. *In general, a good pair may fail to exist in other types of classical Lie superalgebras* [23]. *But for any even nilpotent $e \in \mathfrak{gl}_{M|N}$ we can always find some h such that $\langle e, h\rangle$ is a good pair; see Theorem* 2.4.

Fix a good pair $\langle e, h\rangle$ in $\mathfrak{g}$. Define the following subalgebras of $\mathfrak{g}$ by

$$\mathfrak{p} := \bigoplus_{j\geq 0} \mathfrak{g}(j), \quad \mathfrak{m} := \bigoplus_{j<0} \mathfrak{g}(j). \tag{2.1}$$

Define $\chi \in \mathfrak{g}^*$ by

$$\chi(y) := (y, e) \qquad \forall y \in \mathfrak{g}.$$

The restriction of χ on $\mathfrak{m}$ extends to a one dimensional $U(\mathfrak{m})$-module. Let I_χ be the left ideal of $U(\mathfrak{g})$ generated by

$$\{a - \chi(a) \,|\, a \in \mathfrak{m}\}.$$

As a consequence of the PBW theorem for $U(\mathfrak{g})$, we have $U(\mathfrak{g}) = I_\chi \oplus U(\mathfrak{p})$ together with the following identification

$$U(\mathfrak{g})/I_\chi \cong U(\mathfrak{p}),$$

which is given by the natural projection $\operatorname{pr}_\chi : U(\mathfrak{g}) \to U(\mathfrak{p})$. One defines the following χ-twisted action of $\mathfrak{m}$ on $U(\mathfrak{p})$ by

$$a \cdot y := \operatorname{pr}_\chi([a, y]),$$

for all $a \in \mathfrak{m}, y \in U(\mathfrak{p})$.

The associated *finite W-superalgebra*, which we will usually omit the prefix "finite" from now on, is defined to be the space of $\mathfrak{m}$-invariants in $U(\mathfrak{p})$ under the χ-twisted action; to be explicit,

$$\begin{aligned}\mathcal{W}_{e,h} := U(\mathfrak{p})^{\mathfrak{m}} =& \{y \in U(\mathfrak{p}) \mid \operatorname{pr}_\chi([a, y]) = 0, \forall a \in \mathfrak{m}\}\\ =& \{y \in U(\mathfrak{p}) \mid \big(a - \chi(a)\big)y \in I_\chi, \forall a \in \mathfrak{m}\}.\end{aligned}$$

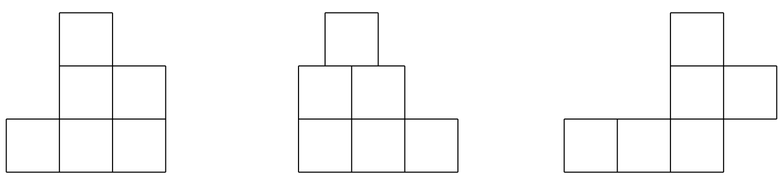
For example, if $e = 0$, then $\chi = 0$, $\mathfrak{g} = \mathfrak{g}(0) = \mathfrak{p}$ and $\mathfrak{m} = 0$, so the associated W-superalgebra is exactly $U(\mathfrak{g})$.

It seems that $\mathcal{W}_{e,h}$ depends on both of e and h from the definition. In fact, the definition is independent of the choices of h up to isomorphisms; see Remark 10.12.

2.2. **Pyramids and W-superalgebras.** We recall the notion of *pyramid* [16, 23] as a convenient tool to present a good pair $\langle e, h\rangle$. We will identify a partition $\lambda = (\lambda_1, \lambda_2, \ldots)$ with its corresponding Young diagram in French style, which means that the diagrams are left-justified and the longest row is located in the bottom.

Definition 2.2. *Let λ be a Young diagram. A pyramid is a diagram obtained by horizontally shifting the rows of λ such that each column of the shifted diagram is a connected vertical strip which starts from the bottom row.*

For example, only the left-most diagram is considered as a pyramid in this article, obtained from shifting the Young diagram of $\lambda = (3, 2, 1)$:

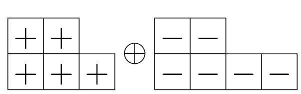


Let $V = V_{\bar{0}} \oplus V_{\bar{1}}$ be a superspace with $\dim V_{\bar{0}} = M$ and $\dim V_{\bar{1}} = N$. We identify $\mathfrak{g} = \mathfrak{gl}_{M|N}$ with $\operatorname{End} V$ and one has the following identification for $\mathfrak{g}_{\bar{0}}$

$$\mathfrak{g}_{\bar{0}} \cong \operatorname{End} V_{\bar{0}} \oplus \operatorname{End} V_{\bar{1}}.$$

As a result, an even nilpotent element $e \in \mathfrak{gl}_{M|N}$ can be thought as a sum of two nilpotent element $e = e_{\bar{0}} + e_{\bar{1}}$, where $e_i \in \operatorname{End} V_i$ for $i \in \{\bar{0}, \bar{1}\}$. Thus we may describe e by two Young diagrams μ and ν corresponding to the Jordan types of $e_{\bar{0}}$ and $e_{\bar{1}}$, respectively.

For example, the diagram



represents an even nilpotent element in $\mathfrak{gl}_{5|6}$, which is a sum of a nilpotent element in $\operatorname{End} \mathbb{C}^5$ with Jordan type $\mu = (3, 2)$ and a nilpotent element in $\operatorname{End} \mathbb{C}^6$ with Jordan type $\nu = (4, 2)$.

We put $+$ and $-$ in the boxes because we now stack the two diagrams together to obtain a new Young diagram, and we need to track from which diagram the boxes originally are.

For example, there are two possibilities if we stack the above two Young diagrams together to obtain one Young diagram:

$$\begin{array}{cccc} + & + & & \\ - & - & & \\ + & + & + & \\ - & - & - & - \end{array} \qquad\qquad \begin{array}{cccc} - & - & & \\ + & + & & \\ + & + & + & \\ - & - & - & - \end{array} \tag{2.2}$$

Remark 2.3. *The pyramids in this article correspond to certain even nilpotent elements in* $\mathfrak{gl}_{M|N}$*, hence the following condition always holds:*

all boxes in a row have the same $+$ or $-$ labeling.

As one may expect, we shift the rows of the stacked Young diagram to obtain a pyramid. For example, we take the right diagram in (2.2) and list all possibilities below:

$$\begin{array}{cccc} & - & - & \\ & + & + & \\ & + & + & + \\ - & - & - & - \end{array} \qquad \begin{array}{cccc} & & - & - \\ & & + & + \\ & + & + & + \\ - & - & - & - \end{array} \qquad \begin{array}{cccc} - & - & & \\ + & + & & \\ + & + & + & \\ - & - & - & - \end{array} \qquad \begin{array}{cccc} & - & - & \\ & + & + & \\ + & + & + & \\ - & - & - & - \end{array}$$

Soon we will see (Theorem 2.4) that each of these pyramids represents a good pair $\langle e, h\rangle$ in $\mathfrak{gl}_{5|6}$. Moreover, these are in fact *all* good pairs we could have for that given $e \in \mathfrak{gl}_{5|6}$.

Now we do the other way around: obtaining a good pair $\langle e, h\rangle$ from a given pyramid π satisfying the condition described in Remark 2.3. Assume that we have M (resp. N) boxes labeled with $+$ (resp. $-$) in π, where they came from the Young diagram of $e_{\bar{0}} \in \mathfrak{gl}_{M|0}$ (resp. $e_{\bar{1}} \in \mathfrak{gl}_{0|N}$). We enumerate those "$+$" boxes by $1, 2, \ldots, M$ down columns from left to right, and enumerate those "$-$" boxes by $\bar{1}, \bar{2}, \ldots, \overline{N}$ by the same rule. In addition, we imagine that each box of π is of size 2×2 and our pyramid is built on the x-axis, where the center of π is exactly located above the origin. For instance:

$$\pi = \begin{array}{cccc} & \bar{2} & \bar{4} & \\ & 1 & 3 & \\ & 2 & 4 & 5 \\ \bar{1} & \bar{3} & \bar{5} & \bar{6} \end{array} \tag{2.3}$$

$$x\text{-coordinates:} \quad \begin{array}{cccc} -3 & -1 & 1 & 3 \end{array}$$

Let $I = \{1 < \ldots < M < \overline{1} < \ldots < \overline{N}\}$ be an ordered index set and let $\{v_i | i \in I\}$ be the standard basis of $\mathbb{C}^{M|N}$ with respect to the following order

$$v_i < v_j \text{ if } i < j \text{ in } I.$$

Let $\{e_{i,j} \,|\, i, j \in I\}$ denote the elementary matrices in $\mathfrak{g} = \mathfrak{gl}_{M|N}$. Set $\mathrm{pa}\,(i) = 0$ if $1 \le i \le M$ and $\mathrm{pa}\,(i) = 1$ if $\overline{1} \le i \le \overline{N}$ for $i \in I$ with respect to the order given above. Then we have $e_{i,j} \in \mathfrak{g}_{\overline{0}}$ (resp. $\mathfrak{g}_{\overline{1}}$) if and only if $\mathrm{pa}\,(i) + \mathrm{pa}\,(j)$ is even (resp. odd). Moreover, their supercommutator can be explicitly given by

$$[e_{i,j}, e_{h,k}] = \delta_{j,h} e_{i,k} - (-1)^{(\mathrm{pa}\,(i)+\mathrm{pa}\,(j))(\mathrm{pa}\,(h)+\mathrm{pa}\,(k))} \delta_{k,i} e_{h,j}$$

One should note that the parity notation $\mathrm{pa}\,(i)$ used here (also used in §9) is for $\mathfrak{gl}_{M|N}$, while another widely used parity notation $|i|$ will be used later for super Yangian.

Define the element

$$e_\pi := \sum_{\boxed{i}\boxed{j} \in \pi} e_{i,j} \in \mathfrak{g}_{\overline{0}}, \tag{2.4}$$

where the sum is taken over all adjacent pairs $\boxed{i}\boxed{j}$ appeared in π.

Let $\mathrm{col}_x(i)$ denote the x-coordinate of the center of the box numbered with $i \in I$, which must be an integer by our construction. Define the following diagonal matrix

$$h_\pi := -\mathrm{diag}\big(\mathrm{col}_x(1), \ldots, \mathrm{col}_x(M), \mathrm{col}_x(\overline{1}), \ldots, \mathrm{col}_x(\overline{N})\big) \tag{2.5}$$

For example, the elements e_π and h_π associated to the pyramid π in (2.3) are

$$\begin{aligned} e_\pi &= e_{13} + e_{24} + e_{45} + e_{\overline{2}\,\overline{4}} + e_{\overline{1}\,\overline{3}} + e_{\overline{3}\,\overline{5}} + e_{\overline{5}\,\overline{6}}, \\ h_\pi &= \mathrm{diag}(1, 1, -1, -1, -3, 3, 1, 1, -1, -1, -3). \end{aligned}$$

It is easy to check that $\langle e_\pi, h_\pi \rangle$ forms a good pair.

Note that if we horizontally shift the rows of π to obtain another pyramid $\vec{\pi}$, then $e_\pi = e_{\vec{\pi}}$ but $h_\pi \neq h_{\vec{\pi}}$. The following theorem implies that all even good $\mathbb{Z}$-gradings for e_π can be obtained by shifting the rows of π.

Theorem 2.4. [23, Theorem 7.2] *Let π be a pyramid. Let $e = e_\pi$ and $h = h_\pi$ be the elements in $\mathfrak{gl}_{M|N}$ defined by (2.4) and (2.5), respectively. Then $\langle e, h \rangle$ forms a good pair for e. Moreover, any good pair for e is of the form $\langle e, h_{\vec{\pi}} \rangle$ where $\vec{\pi}$ is some pyramid obtained by shifting rows of π horizontally.*

In other words, Theorem 2.4 *classifies* all of the even good $\mathbb{Z}$-gradings of $\mathfrak{gl}_{M|N}$ for *any* even nilpotent e. (In fact, [23, Theorem 7.2] classifies *all* good $\mathbb{Z}$-gradings, not just those even good $\mathbb{Z}$-gradings considered in this article.) As a consequence, for a given pyramid π, it makes sense to denote the W-superalgebra associated to the good pair $\langle e_\pi,\, h_\pi \rangle$ simply by $\mathcal{W}_\pi := \mathcal{W}_{e_\pi, h_\pi}$.

Remark 2.5. *If we permute the rows with the same length of π to obtain a new pyramid π', then we have $e_\pi = e_{\pi'}$, $h_\pi = h_{\pi'}$ and hence $\mathcal{W}_\pi = \mathcal{W}_{\pi'}$. For example, the two Young diagrams in* (2.2) *give us exactly the same list of good pairs by shifting their rows.*

We label the columns of π from left to right by $1, \ldots, \ell$. For any $i \in I$, let $\mathrm{col}(i)$ denote the column in which i appear. The *Kazhdan filtration* of $U(\mathfrak{g})$

$$\cdots \subseteq F_d U(\mathfrak{g}) \subseteq F_{d+1} U(\mathfrak{g}) \subseteq \cdots$$

is defined by setting

$$\deg(e_{i,j}) := \mathrm{col}(j) - \mathrm{col}(i) + 1 \tag{2.6}$$

for each $i, j \in I$, where $F_d U(\mathfrak{g})$ denotes the span of all supermonomials $e_{i_1,j_1} \cdots e_{i_s,j_s}$ for $s \geq 0$ with $\sum_{k=1}^{s} \deg\,(e_{i_k,j_k}) \leq d$. Let $\mathrm{gr}\, U(\mathfrak{g})$ denote the graded superalgebra associated to the Kazhdan filtration. A natural grading on $\mathcal{W}_\pi$ is induced from the projection $\mathfrak{g} \twoheadrightarrow \mathfrak{p}$ and we denote by $\mathrm{gr}\, \mathcal{W}_\pi$ the associated graded superalgebra.

Let $\mathfrak{g}^e$ denote the centralizer of e in $\mathfrak{g}$ and let $S(\mathfrak{g}^e)$ denote the associated supersymmetric superalgebra. The same setting (2.6) defines the Kazhdan filtration on $S(\mathfrak{g}^e)$. The following result still holds in our case since our pyramid π satisfies the condition in Remark 2.3.

Proposition 2.6. [47, Remark 3.11] *$S(\mathfrak{g}^e)$ and $\mathrm{gr}\, \mathcal{W}_\pi$ are isomorphic as graded superalgebras.*

2.3. **Shift matrix.** We give an alternative way to describe a pyramid. An $(m+n) \times (m+n)$ matrix $\sigma = (s_{i,j})_{1 \leq i,j \leq m+n}$ is called a *shift matrix* if its entries are non-negative integers satisfying the following condition

$$s_{i,j} + s_{j,k} = s_{i,k}, \tag{2.7}$$

whenever $|i-j| + |j-k| = |i-k|$. For example, the following matrix is a shift matrix:

$$\sigma = \begin{pmatrix} 0 & 1 & 2 & 2 & 3 & 3 \\ 0 & 0 & 1 & 1 & 2 & 2 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ 3 & 3 & 2 & 2 & 0 & 0 \\ 4 & 4 & 3 & 3 & 1 & 0 \end{pmatrix} \tag{2.8}$$

Lemma 2.7. *The following facts hold for a shift matrix $\sigma = (s_{i,j})_{1 \leq i,j \leq m+n}$.*

(1) *If the entries in the last column $\{s_{i,m+n} \,|\, 1 \leq i \leq m+n\}$ are known, then the whole upper-triangular part of σ is determined.*
(2) *If the entries in the upper-diagonal $\{s_{i,i+1} \,|\, 1 \leq i < m+n\}$ are known, then the whole upper-triangular part of σ is determined.*
(3) *If the entries in the last row $\{s_{m+n,i} \,|\, 1 \leq i \leq m+n\}$ are known, then the whole lower-triangular part of σ is determined.*

(4) *If the entries in the lower-diagonal* $\{s_{i+1,i} \,|\, 1 \le i < m+n\}$ *are known, then the whole lower-triangular part of* σ *is determined.*

Proof. By (2.7). □

In our superalgebra setting, we need to record the $\pm$-labeling of each row in our pyramid, so we introduce the following terminology. Let $m, n \in \mathbb{Z}_{\geq 0}$. A $0^m 1^n$-*sequence*, or 01-*sequence* for short, is an ordered sequence Υ consisting of m 0's and n 1's. For $1 \le i \le m+n$, the i-th digit of Υ is denoted by $|i|$.

Suppose that $\sigma \in M_{m+n}(\mathbb{Z}_{\geq 0})$ is a shift matrix. Let ℓ be an integer such that $\ell > s_{1,m+n} + s_{m+n,1}$ and let Υ be a fixed $0^m 1^n$-sequence. Then one can obtain a pyramid π, with m (resp. n) rows labeled by "+" (resp. "$-$") and the bottom row consisting of ℓ boxes, from the triple (σ, ℓ, Υ) by the following fashion.

Start with a rectangular Young diagram consisting of $m+n$ rows and ℓ columns, which we denote by Ξ. We number the rows of Ξ from top to bottom by $1, 2, \ldots, m+n$. For each $1 \le i \le m+n$, we label all boxes in the i-th row of Ξ by "$+$" if $|i| = 0$, and by "$-$" if $|i| = 1$.

Next we obtain our pyramid from this rectangle. Consider the entries in the last row and the last column of σ: $\{s_{m+n,i} \,|\, 1 \le i \le m+n\}$ and $\{s_{i,m+n} \,|\, 1 \le i \le m+n\}$. For each $1 \le j \le m+n$, we erase the leftmost $s_{m+n,j}$ boxes and the rightmost $s_{j,m+n}$ boxes in the j-th row of Ξ. By (2.7), the resulted diagram is a pyramid which has ℓ boxes in the bottom row and $\ell - s_{m+n,1} - s_{1,m+n}$ boxes in the top row. For example, take $\ell = 8$ and let σ be the one given in (2.8) with $\Upsilon = 101010$, the resulted pyramid π is

$$\begin{array}{cccccccc}
 & & & & - & & & \\
 & & & & + & + & & \\
 & & & - & - & - & - & \\
 & & & + & + & + & + & \\
 & - & - & - & - & - & - & \\
+ & + & + & + & + & + & + & +
\end{array}$$

Conversely, given a pyramid π which represents a good pair. Let ℓ be the number of boxes in the bottom of π and let m and n be the numbers of rows of π labeled by $+$ and $-$, respectively. We number the rows of π from top to bottom by $1, 2, \ldots, m+n$ as before. Since π satisfies the condition in Remark 2.3, we may obtain a $0^m 1^n$-sequence Υ by assigning the i-th digit of Υ to be 0 (resp. 1) if the boxes in the i-th row are labeled by "$+$" (resp. "$-$").

For each $1 \le i \le m+n$, define the number $s_{m+n,i}$ (resp. $s_{i,m+n}$) to be the number of missing boxes on the left-hand side (resp. right-hand side) of the i-th row of π in a rectangular diagram Ξ of size $(m+n)\times\ell$. This gives us the entries of the last row and the last column of σ and hence we are able to recover the whole σ by Lemma 2.7. The discussion above is summarized in the following proposition.

Proposition 2.8. *Let S be the set of triples (σ,ℓ,Υ) where σ is a shift matrix of size $m+n$, $\ell > s_{m+n,1}+s_{1,m+n}$ is an integer and Υ is a 0^m1^n-sequence. Let P be the set of all pyramids π such that π has m (resp. n) rows labeled by $+$ (resp. $-$) and ℓ columns. Then there exists a bijection between S and P.*

Roughly speaking, σ determines the *shape and height*, ℓ determines the *width* and Υ determines the $\pm$*-labeling* of π and vice versa.

The following proposition is a super analogue of a well-known result about $\mathfrak{g}^e$. Since our pyramid π satisfies the condition described in Remark 2.3, its proof is similar to the Lie algebra case as remarked in [4].

Proposition 2.9. [4, Lemma 4.2] *Let π be a pyramid with row lengths $\{p_i\,|\,1\leq i\leq m+n\}$, where the rows are labeled from top to bottom. Let $\sigma=(s_{i,j})_{1\leq i,j\leq m+n}$ be the associated shift matrix of π in the triple (σ,ℓ,Υ). Let $e=e_\pi$ be the nilpotent element defined by (2.4). Let M (resp. N) be the number of boxes of π labeled in $+$ (resp. $-$). For all $1\leq i,j\leq m+n$ and $r>0$, define*

$$c_{i,j}^{(r)} := \sum_{\substack{h,k\in I\\ row(h)=i,\ row(k)=j\\ col(k)-col(h)=r-1}} e_{h,k}\in\mathfrak{g}=\mathfrak{gl}_{M|N}.$$

Then $\{c_{i,j}^{(r)}\,|\,1\leq i,j\leq m+n,\ s_{i,j}<r\leq s_{i,j}+p_{min(i,j)}\}$ forms a linear basis for $\mathfrak{g}^e$.

3. The super Yangian $Y_{m|n}$

In this section, we recall some well-known facts about the super Yangian associated to the general linear Lie superalgebra.

3.1. RTT presentations of $Y_{m|n}$.

Definition 3.1. [28] *For a given 01-sequence Υ, the Yangian associated to the general linear Lie superalgebra $\mathfrak{gl}_{m|n}$, denoted by $Y_{m|n}$, is the associative $\mathbb{Z}_2$-graded algebra with unity generated over $\mathbb{C}$ by the RTT generators*

$$\left\{t_{i,j}^{(r)}\,|\,1\leq i,j\leq m+n; r\geq 1\right\}, \tag{3.1}$$

subject to following RTT relations:

$$\left[t_{i,j}^{(r)},t_{h,k}^{(s)}\right]=(-1)^{|i|\,|j|+|i|\,|h|+|j|\,|h|}\sum_{g=0}^{\min(r,s)-1}\left(t_{h,j}^{(g)}\,t_{i,k}^{(r+s-1-g)}-t_{h,j}^{(r+s-1-g)}\,t_{i,k}^{(g)}\right), \tag{3.2}$$

where the parity of $t_{i,j}^{(r)}$ is defined by $|i|+|j|\pmod 2$. By convention, we set $t_{i,j}^{(0)}:=\delta_{ij}$.

The original definition in [28] corresponds to the case when Υ is the *standard* 01-sequence, which is defined as

$$\Upsilon^{st} := \overbrace{0\ldots0}^{m}\overbrace{1\ldots1}^{n}.$$

As observed in [31, 40], up to isomorphism, the definition of $Y_{m|n}$ is independent of the choices of Υ so we often omit it in our notation when appropriate.

For each $1 \le i, j \le m+n$, define the formal series

$$t_{i,j}(u) := \sum_{r\ge 0} t_{i,j}^{(r)} u^{-r} \in Y_{m|n}[[u^{-1}]].$$

It is well-known [28] that $Y_{m|n}$ is a Hopf-superalgebra. In particular, the comultiplication $\Delta : Y_{m|n} \to Y_{m|n} \otimes Y_{m|n}$ can be nicely described as

$$\Delta(t_{i,j}^{(r)}) = \sum_{s=0}^{r} \sum_{k=1}^{m+n} t_{i,k}^{(r-s)} \otimes t_{k,j}^{(s)}. \tag{3.3}$$

Moreover, there exists a surjective homomorphism

$$\mathrm{ev} : Y_{m|n} \to U(\mathfrak{gl}_{m|n})$$

called the *evaluation homomorphism*, defined by

$$\mathrm{ev}\left(t_{i,j}(u)\right) := \delta_{ij} + (-1)^{|i|} e_{ij} u^{-1}, \tag{3.4}$$

where $e_{ij} \in \mathfrak{gl}_{m|n}$ means the elementary matrix.

The following proposition gives a PBW basis for $Y_{m|n}$ in terms of the RTT generators.

Proposition 3.2. [22, Theorem 1] *The set of supermonomials in the following elements*

$$\left\{ t_{i,j}^{(r)} \mid 1 \le i, j \le m+n, r \ge 1 \right\}$$

taken in some fixed order forms a linear basis for $Y_{m|n}$.

Define the *loop filtration* on $Y_{m|n}$

$$L_0 Y_{m|n} \subseteq L_1 Y_{m|n} \subseteq L_2 Y_{m|n} \subseteq \cdots$$

by setting $\deg t_{ij}^{(r)} = r-1$ for each $r \ge 1$ and letting $L_k Y_{m|n}$ be the span of all supermonomials of the form

$$t_{i_1 j_1}^{(r_1)} t_{i_2 j_2}^{(r_2)} \cdots t_{i_s j_s}^{(r_s)}$$

with total degree not greater than k. We denote by $\mathrm{gr}^L Y_{m|n}$ the associated graded superalgebra.

Let $\mathfrak{gl}_{m|n}[x]$ denote the *loop superalgebra* $\mathfrak{gl}_{m|n} \otimes \mathbb{C}[x]$, where a basis is given by

$$\{ e_{ij} x^r \mid 1 \le i, j \le m+n, r \ge 0 \}.$$

Let $U(\mathfrak{gl}_{m|n}[x])$ denote its universal enveloping algebra with the natural filtration and grading given by

$$\deg e_{ij}x^r := r.$$

The following corollary is a consequence of Proposition 3.2.

Corollary 3.3. [22, Corollary 1] *The assignment*

$$t_{ij}^{(r)} \mapsto (-1)^{|i|} e_{ij} x^{r-1}$$

gives rise to an isomorphism $\operatorname{gr}^L Y_{m|n} \cong U(\mathfrak{gl}_{m|n}[x])$ *of graded superalgebras.*

3.2. **Parabolic generators of $Y_{m|n}$.** In this subsection, we give another generating set for $Y_{m|n}$. Eventually it will allow us to define a certain subalgebra of $Y_{m|n}$ which can not be observed by the earlier RTT-presentation.

Firstly we introduce a convenient shorthand notation which will be frequently used in this article. Let $\mu = (\mu_1, \ldots, \mu_z)$ be a given composition of $m+n$ with length z and let Υ be a fixed $0^m 1^n$-sequence. We break Υ into z subsequences according to μ; that is,

$$\Upsilon = \Upsilon_1 \Upsilon_2 \ldots \Upsilon_z,$$

where Υ_1 is the subsequence consisting of the first μ_1 digits of Υ, Υ_2 is the subsequence consisting of the next μ_2 digits of Υ, and so on. For example, if we have $\Upsilon = 011100011$ and $\mu = (2, 4, 3)$, then

$$\Upsilon = \overbrace{01}^{\Upsilon_1}\overbrace{1100}^{\Upsilon_2}\overbrace{011}^{\Upsilon_3}.$$

For each $1 \leq a \leq z$, let p_a and q_a denote the number of 0's and 1's in Υ_a, respectively. For a fixed $1 \leq a \leq z$ and each value of $i = 1, 2, \ldots, \mu_a$, we define the *restricted parity* $|i|_a$ by

$$|i|_a := \text{the } i\text{-th digits of } \Upsilon_a,$$

or equivalently

$$|i|_a = |\sum_{j=1}^{a-1} \mu_j + i|. \tag{3.5}$$

Define the $(m+n) \times (m+n)$ matrix with entries in $Y_{m|n}[[u^{-1}]]$ by

$$T(u) := \Big(t_{i,j}(u)\Big)_{1 \leq i,j \leq m+n}$$

Remark 3.4. *Following* [22], *the matrix $T(u)$ is identified with the following operator*

$$\sum_{i,j=1}^{m+n} t_{i,j}(u) \otimes (-1)^{|j|(|i|+1)} E_{i,j} \in Y_{m|n}[[u^{-1}]] \otimes \operatorname{End} \mathbb{C}^{m|n}$$

where $E_{i,j}$ denotes the elementary matrix in $\operatorname{End} \mathbb{C}^{m|n}$. *The term $(-1)^{|j|(|i|+1)}$ ensures that the matrix product can be calculated in the usual way.*

Note that the leading minors of the matrix $T(u)$ are always invertible and hence the matrix $T(u)$ possesses a Gauss decomposition [21] with respect to μ. To be explicit, we have

$$T(u) = F(u)D(u)E(u) \tag{3.6}$$

for unique *block matrices* $D(u)$, $E(u)$ and $F(u)$ of the form

$$D(u) = \begin{pmatrix} D_1(u) & 0 & \cdots & 0 \\ 0 & D_2(u) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & D_z(u) \end{pmatrix},$$

$$E(u) = \begin{pmatrix} I_{\mu_1} & E_{1,2}(u) & \cdots & E_{1,z}(u) \\ 0 & I_{\mu_2} & \cdots & E_{2,z}(u) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & I_{\mu_z} \end{pmatrix},$$

$$F(u) = \begin{pmatrix} I_{\mu_1} & 0 & \cdots & 0 \\ F_{2,1}(u) & I_{\mu_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ F_{z,1}(u) & F_{z,2}(u) & \cdots & I_{\mu_z} \end{pmatrix},$$

where

$$D_a(u) = \big(D_{a;i,j}(u)\big)_{1\le i,j\le \mu_a}, \tag{3.7}$$
$$E_{a,b}(u) = \big(E_{a,b;h,k}(u)\big)_{1\le h\le \mu_a, 1\le k\le \mu_b}, \tag{3.8}$$
$$F_{b,a}(u) = \big(F_{b,a;k,h}(u)\big)_{1\le k\le \mu_b, 1\le h\le \mu_a}, \tag{3.9}$$

are $\mu_a \times \mu_a$, $\mu_a \times \mu_b$ and $\mu_b \times \mu_a$ matrices, respectively, for all $1 \le a \le z$ in (3.7) and all $1 \le a < b \le z$ in (3.8) and (3.9). In fact, these matrices can be explicitly obtained by *quasideterminants* (cf. [21], [33, Proposition 3.1]).

Since all of the submatrices $D_a(u)$'s are invertible, it allows one to define the $\mu_a \times \mu_a$ matrix $D'_a(u) = \big(D'_{a;i,j}(u)\big)_{1\le i,j\le \mu_a}$ by

$$D'_a(u) := \big(D_a(u)\big)^{-1}.$$

The entries of these matrices give us some formal series with coefficients in $Y_{m|n}$:

$$D_{a;i,j}(u) = \textstyle\sum_{r\ge 0} D^{(r)}_{a;i,j} u^{-r}, \qquad D'_{a;i,j}(u) = \sum_{r\ge 0} D'^{(r)}_{a;i,j} u^{-r}, \tag{3.10}$$

$$E_{a,b;h,k}(u) = \textstyle\sum_{r\ge 1} E^{(r)}_{a,b;h,k} u^{-r}, \qquad F_{b,a;k,h}(u) = \sum_{r\ge 1} F^{(r)}_{b,a;k,h} u^{-r}. \tag{3.11}$$

Actually we only need the diagonal, upper-diagonal and lower-diagonal blocks. Hence we set

$$E_{b;h,k}(u) := E_{b,b+1;h,k}(u) = \sum_{r\geq 1} E_{b;h,k}^{(r)} u^{-r}, \qquad F_{b;k,h}(u) := F_{b+1,b;k,h}(u) = \sum_{r\geq 1} F_{b;k,h}^{(r)} u^{-r}, \tag{3.12}$$

for $1 \leq b \leq z-1$. As proved in [33], these coefficients

$$\{D_{a;i,j}^{(r)}, D_{a;i,j}^{\prime(r)} \,|\, 1 \leq a \leq z, 1 \leq i,j \leq \mu_a, r \geq 0\}$$
$$\{E_{b;h,k}^{(r)} \,|\, 1 \leq b < z, 1 \leq h \leq \mu_b, 1 \leq k \leq \mu_{b+1}, r \geq 1\}$$
$$\{F_{b;k,h}^{(r)} \,|\, 1 \leq b < z, 1 \leq k \leq \mu_{b+1}, 1 \leq h \leq \mu_b, r \geq 1\}$$

form a generating set for $Y_{m|n}$, called the *parabolic generators of* $Y_{m|n}$, which will be denoted by $\mathcal{P}_\mu$. By [33, Lemma 4.2], their parities can be explicitly determined by the following rule:

$$\text{parity of } D_{a;i,j}^{(r)} = |i|_a + |j|_a \pmod 2, \tag{3.13}$$
$$\text{parity of } E_{b;h,k}^{(r)} = |h|_b + |k|_{b+1} \pmod 2, \tag{3.14}$$
$$\text{parity of } F_{b;k,h}^{(r)} = |k|_{b+1} + |h|_b \pmod 2. \tag{3.15}$$

In the special case when $\mu = (1^{m+n}) := (\overbrace{1,\ldots,1}^{m+n})$, the generating set, which will be denoted by $\mathcal{P}_D$, appeared in an analogue of the Drinfeld presentation for $Y_{m|n}$ [7, 15, 22, 33, 39, 40]. We list $\mathcal{P}_D$ explicitly here since it will be used right away:

$$\{D_a^{(r)}, D_a^{\prime(r)} \,|\, 1 \leq a \leq m+n, r \geq 0\}, \tag{3.16}$$
$$\{E_b^{(r)} \,|\, 1 \leq b < m+n, r \geq 1\}, \tag{3.17}$$
$$\{F_b^{(r)} \,|\, 1 \leq b < m+n, r \geq 1\}, \tag{3.18}$$

and their parities are given by

$$|D_a^{(r)}| = |D_a^{\prime(r)}| = 0, \qquad |E_b^{(r)}| = |F_b^{(r)}| = |b| + |b+1| \pmod 2. \tag{3.19}$$

4. Shifted super Yangian: Drinfeld's presentation

Recall from §2 that a pyramid π can be uniquely recorded by a triple (σ, ℓ, Υ) where σ is a shift matrix of size $m+n$, ℓ is a positive integer and Υ is a 01-sequence. Following [8, §2], we use σ and Υ to define the following structure, which is one of the main objects studied in this article.

Definition 4.1. *Let $m,n \in \mathbb{Z}_{\geq 0}$, $\sigma = (s_{i,j})$ be a shift matrix of size $m+n$ with a fixed 0^m1^n-sequence Υ. The shifted super Yangian of $\mathfrak{gl}_{m|n}$ associated to σ, denoted by $Y_{m|n}(\sigma)$, is* *the superalgebra over $\mathbb{C}$ generated by following symbols*

$$\big\{D_a^{(r)}, D_a'^{(r)} \,|\, 1\leq a\leq m+n,\ r\geq 0\big\},$$
$$\big\{E_b^{(r)} \,|\, 1\leq b< m+n,\ r> s_{b,b+1}\big\},$$
$$\big\{F_b^{(r)} \,|\, 1\leq b< m+n,\ r> s_{b+1,b}\big\},$$

where their parities are defined by (3.19)*, subject to the following relations:*

$$D_a^{(0)}=D_a'^{(0)} \;=\; 1\,, \tag{4.1}$$
$$\sum_{t=0}^{r} D_a^{(t)}D_a'^{(r-t)} \;=\; \delta_{r0}, \tag{4.2}$$
$$\big[D_a^{(r)}, D_b^{(s)}\big] \;=\; 0, \tag{4.3}$$

$$[D_a^{(r)}, E_b^{(s)}] \;=\; (-1)^{|a|}\big(\delta_{a,b}-\delta_{a,b+1}\big)\sum_{t=0}^{r-1} D_a^{(t)}E_b^{(r+s-1-t)}, \tag{4.4}$$
$$[D_a^{(r)}, F_b^{(s)}] \;=\; (-1)^{|a|}\big(\delta_{a,b+1}-\delta_{a,b}\big)\sum_{t=0}^{r-1} F_b^{(r+s-1-t)}D_a^{(t)}, \tag{4.5}$$
$$[E_a^{(r)}, F_b^{(s)}] \;=\; \delta_{a,b}(-1)^{|a+1|+1}\sum_{t=0}^{r+s-1} D_a'^{(r+s-1-t)}D_{a+1}^{(t)}, \tag{4.6}$$

$$[E_a^{(r)}, E_a^{(s)}] \;=\; (-1)^{|a+1|}\Big(\sum_{t=s_{a,a+1}+1}^{s-1} E_a^{(r+s-1-t)}E_a^{(t)} \;-\; \sum_{t=s_{a,a+1}+1}^{r-1} E_a^{(r+s-1-t)}E_a^{(t)}\Big), \tag{4.7}$$

$$[F_a^{(r)}, F_a^{(s)}] \;=\; (-1)^{|a|}\Big(\sum_{t=s_{a+1,a}+1}^{r-1} F_a^{(r+s-1-t)}F_a^{(t)} \;-\; \sum_{t=s_{a+1,a}+1}^{s-1} F_a^{(r+s-1-t)}F_a^{(t)}\Big), \tag{4.8}$$

$$[E_a^{(r+1)}, E_{a+1}^{(s)}]-[E_a^{(r)}, E_{a+1}^{(s+1)}] \;=\; (-1)^{|a+1|}E_a^{(r)}E_{a+1}^{(s)}\,, \tag{4.9}$$
$$[F_a^{(r+1)}, F_{a+1}^{(s)}]-[F_a^{(r)}, F_{a+1}^{(s+1)}] \;=\; (-1)^{1+|a||a+1|+|a+1||a+2|+|a||a+2|}F_{a+1}^{(s)}F_a^{(r)}\,, \tag{4.10}$$

$$[E_a^{(r)}, E_b^{(s)}]=0 \qquad \text{if } |b-a|>1, \tag{4.11}$$

$$[F_a^{(r)}, F_b^{(s)}]=0 \qquad \text{if } |b-a|>1, \tag{4.12}$$

$$\big[E_a^{(r)}, [E_a^{(s)}, E_b^{(t)}]\big]+\big[E_a^{(s)}, [E_a^{(r)}, E_b^{(t)}]\big]=0 \quad \text{if } |a-b|=1, \tag{4.13}$$

$$\big[F_a^{(r)}, [F_a^{(s)}, F_b^{(t)}]\big]+\big[F_a^{(s)}, [F_a^{(r)}, F_b^{(t)}]\big]=0 \quad \text{if } |a-b|=1, \tag{4.14}$$

$$\big[\,[E_{a-1}^{(r)}, E_a^{(t)}]\,,\,[E_a^{(t)}, E_{a+1}^{(s)}]\,\big] = 0 \;\; when \;\; m+n \geq 4 \; and \;\; |a|+|a+1| = 1, \tag{4.15}$$

$$\big[\,[F_{a-1}^{(r)}, F_a^{(t)}]\,,\,[F_a^{(t)}, F_{a+1}^{(s)}]\,\big] = 0 \;\; when \;\; m+n \geq 4 \; and \;\; |a|+|a+1| = 1, \tag{4.16}$$

for all admissible indices a, b, r, s, t. For example, (4.4) is meant to hold for all $r \geq 0$, $s > s_{b,b+1}$, $1 \leq a \leq m+n$ and $1 \leq b < m+n$.

Note that when σ is the zero matrix, the presentation above coincides[1] with the presentation of $Y_{m|n}$ given in [33] by taking $\mu = (1^{m+n})$ therein (in this special case, such a presentation for $Y_{m|n}$ is also obtained in [40]). As a result, we may identify $Y_{m|n}(0) = Y_{m|n}$.

In the remaining part of this section, we will show that $Y_{m|n}(\sigma)$ can be identified as a subalgebra of $Y_{m|n}$ in general (Corollary 4.5). Let $\mathcal{P}_{D,\sigma}$ be the generating set of $Y_{m|n}(\sigma)$ in Definition 4.1. Let $\Gamma : \mathcal{P}_{D,\sigma} \to \mathcal{P}_D$ be the map sending elements in $\mathcal{P}_{D,\sigma}$ to the elements with the same name (3.16)–(3.18) in $\mathcal{P}_D$ obtained by Gauss decomposition.

Proposition 4.2. *The map Γ induces a canonical homomorphism $\Gamma : Y_{m|n}(\sigma) \to Y_{m|n}$.*

Proof. By setting $\mu = (1^{m+n})$ in [33, Proposition 7.1], or simply by [40, (2.2)–(2.10)], we see that the relations (4.1)–(4.14) are preserved by Γ. Setting $k = l$ in the *generalized quartic Serre relations* in [40, (2.14), (2.15)], we see that (4.15) and (4.16) are preserved by Γ as well. $\square$

It remains to show that Γ is injective. We introduce the *loop filtration* on $Y_{m|n}(\sigma)$

$$L_0 Y_{m|n}(\sigma) \subseteq L_1 Y_{m|n}(\sigma) \subseteq L_2 Y_{m|n}(\sigma) \subseteq \cdots$$

by setting the degrees of the generators $D_a^{(r)}$, $E_b^{(r)}$, and $F_b^{(r)}$ to be $(r-1)$ and setting $L_k Y_{m|n}(\sigma)$ to be the span of all supermonomials in the generators of total degree not greater than k. Let $\operatorname{gr}^L Y_{m|n}(\sigma)$ denote the associated graded superalgebra.

For $1 \leq a < b \leq m+n$, $r > s_{a,b}$ and $t > s_{b,a}$, define the following higher root elements $E_{a,b}^{(r)}, F_{b,a}^{(t)} \in Y_{m|n}(\sigma)$ recursively by

$$E_{a,a+1}^{(r)} := E_a^{(r)}, \qquad E_{a,b}^{(r)} := (-1)^{|b-1|}[E_{a,b-1}^{(r-s_{b-1,b})}, E_{b-1}^{(s_{b-1,b}+1)}], \tag{4.17}$$

$$F_{a+1,a}^{(t)} := F_a^{(t)}, \qquad F_{b,a}^{(t)} := (-1)^{|b-1|}[F_{b-1}^{(s_{b,b-1}+1)}, F_{b-1,a}^{(t-s_{b,b-1})}]. \tag{4.18}$$

By definition, we have $E_{a,b}^{(r)} \in L_{r-1} Y_{m|n}(\sigma)$ and $F_{b,a}^{(t)} \in L_{t-1} Y_{m|n}(\sigma)$.

[1] Note that some relations given in Definition 4.1 are redundant, which is fine for our purpose. In [33], which studied the case $\sigma = 0$ under our current setting, the relations (4.15)–(4.16) were given only for $t = 1$. As noticed in [40, Remark 2.13], by relations (4.1)–(4.14), the case $t = 1$ implies that they actually hold for all $t \geq 1$. In addition, (4.15)–(4.16) hold for $|a| + |a+1| = 0$ as well but they can also be deduced from (4.1)–(4.14).

Define the elements $\{e_{a,b}^{(r)} \,|\, 1 \le a, b \le m+n, r \ge s_{a,b}\} \subseteq \mathrm{gr}^L Y_{m|n}(\sigma)$ by

$$e_{a,b}^{(r)} := \begin{cases} \mathrm{gr}_r^L D_a^{(r+1)} & \text{if } a = b, \\ \mathrm{gr}_r^L E_{a,b}^{(r+1)} & \text{if } a < b, \\ \mathrm{gr}_r^L F_{a,b}^{(r+1)} & \text{if } a > b. \end{cases} \tag{4.19}$$

Using the same argument as in [33, Lemma 7.5], except that one uses the defining relations of $Y_{m|n}(\sigma)$ listed in Definition 4.1, we deduce the following result.

Proposition 4.3. [8, (2.21)][22, (51)] *For all $1 \le a, b, c, d \le m+n$, $r \ge s_{a,b}$, $t \ge s_{c,d}$, the following identity holds in $\mathrm{gr}^L Y_{m|n}(\sigma)$:*

$$[e_{a,b}^{(r)}, e_{c,d}^{(t)}] = (-1)^{|b|}\delta_{b,c} e_{a,d}^{(r+t)} - (-1)^{|a||b|+|a||c|+|b||c|}\delta_{a,d} e_{c,b}^{(r+t)} \tag{4.20}$$

Let $\mathfrak{gl}_{m|n}[x](\sigma)$ be the subalgebra of the loop superalgebra $\mathfrak{gl}_{m|n}[x]$ generated by the following elements

$$\{e_{ij}x^r \,|\, 1 \le i, j \le m+n, r \ge s_{i,j}\}.$$

By (2.7), $\mathfrak{gl}_{m|n}[x](\sigma)$ is indeed a subalgebra of $\mathfrak{gl}_{m|n}[x]$. Let the universal enveloping algebra $U\big(\mathfrak{gl}_{m|n}[x](\sigma)\big)$ be equipped with the natural grading induced by the grading on $\mathfrak{gl}_{m|n}[x]$.

Theorem 4.4. [8, Theorem 2.1] *The map*

$$\gamma : U\big(\mathfrak{gl}_{m|n}[x](\sigma)\big) \longrightarrow \mathrm{gr}^L Y_{m|n}(\sigma)$$

defined by

$$\gamma(e_{a,b}x^r) = (-1)^{|a|} e_{a,b}^{(r)},$$

for all $1 \le a, b \le m+n$, $r \ge s_{a,b}$, is an isomorphism of graded superalgebras.

Proof. γ is a homomorphism by (4.20). Since the image of γ contains the image of $\mathcal{P}_{D,\sigma}$ in $\mathrm{gr}^L Y_{m|n}(\sigma)$, γ is surjective.

It remains to show the injectivity. Consider firstly the special case when $\sigma = 0$, where we can identify $Y_{m|n}(0) = Y_{m|n}$. By [33, Proposition 7.9], the ordered supermonomials in the elements $\{e_{a,b}^{(r)} \,|\, 1 \le a, b \le m+n,\ r \ge 0\}$ are linearly independent in $\mathrm{gr}^L Y_{m|n}$. It follows that γ is injective.

For the general case, observe that the canonical map $\Gamma : Y_{m|n}(\sigma) \to Y_{m|n}$ is a homomorphism of filtered superalgebras. It induces a map $\mathrm{gr}^L Y_{m|n}(\sigma) \to \mathrm{gr}^L Y_{m|n}$, sending $e_{a,b}^{(r)} \in \mathrm{gr}^L Y_{m|n}(\sigma)$ to $e_{a,b}^{(r)} \in \mathrm{gr}^L Y_{m|n}$. By the previous paragraph, the ordered supermonomials in the elements $\{e_{a,b}^{(r)} \,|\, 1 \le a, b \le m+n,\ r \ge s_{a,b}\}$ are linearly independent in $\mathrm{gr}^L Y_{m|n}(\sigma)$ as well, which implies that γ is injective by the PBW theorem for $U\big(\mathfrak{gl}_{m|n}[x](\sigma)\big)$. $\square$

Corollary 4.5. [8, Corollary 2.2] *The canonical homomorphism $\Gamma : Y_{m|n}(\sigma) \to Y_{m|n}$ is injective. As a consequence, the structure $Y_{m|n}(\sigma)$ defined in Definition 4.1 can be identified as a subalgebra of $Y_{m|n}$.*

5. Shifted super Yangian: Parabolic presentations

In this section, we provide a more sophisticated definition for $Y_{m|n}(\sigma)$ together with corresponding results obtained in § 4. For the sake of the purpose, we first introduce some terminologies and notations.

Let $\sigma = (s_{i,j})$ be a shift matrix of size $m+n$. We say a composition $\mu = (\mu_1, \dots, \mu_z)$ of $m+n$ of length z is *admissible to* σ if

$$s_{\mu_1+\mu_2+\cdots+\mu_{a-1}+i,\mu_1+\mu_2+\cdots+\mu_{a-1}+j} = 0$$

for all $1 \leq a \leq z$, $1 \leq i, j \leq \mu_a$. In addition, μ is called *minimal admissible* if it is admissible to σ and its length is minimal among all compositions admissible to σ. Clearly, for a shift matrix σ, its minimal admissible shape uniquely exists. Moreover, (1^{m+n}) is admissible for any σ of size $m+n$.

Remark 5.1. *The notion of admissibility can be intuitively explained in terms of pyramid. Note that one can decompose a pyramid horizontally into a number of rectangles. An admissible shape μ records the heights of these rectangles from top to bottom, while the minimal admissible shape records such a decomposition with the least number of rectangles.*

When $\mu = (\mu_1, \mu_2, \dots, \mu_z)$ is admissible to σ, we will use a shorthand notation

$$s^{\mu}_{a,b} := s_{\mu_1+\ldots+\mu_a,\mu_1+\ldots+\mu_b}, \quad \forall\ 1 \leq a, b \leq z. \tag{5.1}$$

Note that one can recover the original matrix σ if an admissible shape μ and the numbers $\{s^{\mu}_{a,b} | 1 \leq a, b \leq z\}$ are known. Moreover, under the assumption (2.7), the admissible condition implies that for any $1 \leq a, b \leq z$, we have

$$s_{\mu_1+\cdots+\mu_{a-1}+i,\mu_1+\cdots+\mu_{b-1}+j} = s^{\mu}_{a,b}, \qquad \forall 1 \leq i \leq \mu_a, 1 \leq j \leq \mu_b. \tag{5.2}$$

Let Υ be a fixed 0^m1^n-sequence. We decompose Υ into z subsequences according to μ

$$\Upsilon = \Upsilon_1\Upsilon_2\cdots\Upsilon_z,$$

and define the restricted parity $|i|_a$ as in (3.5). Now we give the following presentation for $Y_{m|n}(\sigma)$, which is a super analogue of the shifted Yangian given in [8, §3].

Definition 5.2. *Let $\sigma = (s_{i,j})$ be a shift matrix of size $m+n$ with a fixed 0^m1^n-sequence Υ. Let $\mu = (\mu_1, \dots, \mu_z)$ be an admissible shape to σ. The shifted super Yangian of $\mathfrak{gl}_{m|n}$ associated to σ and μ, denoted by $Y_\mu(\sigma)$, is the superalgebra over $\mathbb{C}$ generated by the following symbols*

$$\big\{D^{(r)}_{a;i,j}, D'^{(r)}_{a;i,j} \,|\, 1 \leq a \leq z,\ 1 \leq i,j \leq \mu_a,\ r \geq 0\big\},$$

$$\big\{E^{(r)}_{b;h,k} \,|\, 1 \leq b < z,\ 1 \leq h \leq \mu_b, 1 \leq k \leq \mu_{b+1},\ r > s^{\mu}_{b,b+1}\big\},$$

$$\big\{F^{(r)}_{b;k,h} \,|\, 1 \leq b < z,\ 1 \leq h \leq \mu_b, 1 \leq k \leq \mu_{b+1},\ r > s^{\mu}_{b+1,b}\big\},$$

where their parities are defined by (3.13)–(3.15), subject to the following relations:

$$D_{a;i,j}^{(0)} = D_{a;i,j}'^{(0)} = \delta_{ij}\,, \tag{5.3}$$

$$\sum_{p=1}^{\mu_a}\sum_{t=0}^{r} D_{a;i,p}^{(t)} D_{a;p,j}'^{(r-t)} = \delta_{r0}\delta_{ij}\,, \tag{5.4}$$

$$\begin{aligned}\left[D_{a;i,j}^{(r)}, D_{b;h,k}^{(s)}\right] &= \delta_{ab}(-1)^{|i|_a|j|_a+|i|_a|h|_a+|j|_a|h|_a} \times \\ &\sum_{t=0}^{min(r,s)-1} \left(D_{a;h,j}^{(t)} D_{a;i,k}^{(r+s-1-t)} - D_{a;h,j}^{(r+s-1-t)} D_{a;i,k}^{(t)}\right),\end{aligned} \tag{5.5}$$

$$\begin{aligned}[D_{a;i,j}^{(r)}, E_{b;h,k}^{(s)}] = \delta_{a,b}\delta_{hj}(-1)^{|h|_a|j|_a} \sum_{p=1}^{\mu_a}\sum_{t=0}^{r-1} D_{a;i,p}^{(t)} E_{b;p,k}^{(r+s-1-t)} \\ - \delta_{a,b+1}(-1)^{|h|_b|k|_a+|h|_b|j|_a+|j|_a|k|_a} \sum_{t=0}^{r-1} D_{a;i,k}^{(t)} E_{b;h,j}^{(r+s-1-t)},\end{aligned} \tag{5.6}$$

$$\begin{aligned}[D_{a;i,j}^{(r)}, F_{b;h,k}^{(s)}] = -\delta_{a,b}\delta_{ik}(-1)^{|i|_a|j|_a+|h|_{a+1}|i|_a+|h|_{a+1}|j|_a} \sum_{p=1}^{\mu_a}\sum_{t=0}^{r-1} F_{b;h,p}^{(r+s-1-t)} D_{a;p,j}^{(t)} \\ + \delta_{a,b+1}(-1)^{|h|_a|k|_b+|h|_a|j|_a+|j|_a|k|_b} \sum_{t=0}^{r-1} F_{b;i,k}^{(r+s-1-t)} D_{a;h,j}^{(t)},\end{aligned} \tag{5.7}$$

$$[E_{a;i,j}^{(r)}, F_{b;h,k}^{(s)}] = \delta_{a,b}(-1)^{|h|_{a+1}|k|_a+|j|_{a+1}|k|_a+|h|_{a+1}|j|_{a+1}+1} \sum_{t=0}^{r+s-1} D_{a;i,k}'^{(r+s-1-t)} D_{a+1;h,j}^{(t)}, \tag{5.8}$$

$$\begin{aligned}[E_{a;i,j}^{(r)}, E_{a;h,k}^{(s)}] = (-1)^{|h|_a|j|_{a+1}+|j|_{a+1}|k|_{a+1}+|h|_a|k|_{a+1}} \times \\ \Big(\sum_{t=s_{a,a+1}^{\mu}+1}^{s-1} E_{a;i,k}^{(r+s-1-t)} E_{a;h,j}^{(t)} - \sum_{t=s_{a,a+1}^{\mu}+1}^{r-1} E_{a;i,k}^{(r+s-1-t)} E_{a;h,j}^{(t)}\Big),\end{aligned} \tag{5.9}$$

$$\begin{aligned}[F_{a;i,j}^{(r)}, F_{a;h,k}^{(s)}] = (-1)^{|h|_{a+1}|j|_a+|j|_a|k|_a+|h|_{a+1}|k|_a} \times \\ \Big(\sum_{t=s_{a+1,a}^{\mu}+1}^{r-1} F_{a;i,k}^{(r+s-1-t)} F_{a;h,j}^{(t)} - \sum_{t=s_{a+1,a}^{\mu}+1}^{s-1} F_{a;i,k}^{(r+s-1-t)} F_{a;h,j}^{(t)}\Big),\end{aligned} \tag{5.10}$$

$$[E_{a;i,j}^{(r+1)}, E_{a+1;h,k}^{(s)}] - [E_{a;i,j}^{(r)}, E_{a+1;h,k}^{(s+1)}] = (-1)^{|j|_{a+1}|h|_{a+1}}\delta_{h,j} \sum_{q=1}^{\mu_{a+1}} E_{a;i,q}^{(r)} E_{a+1;q,k}^{(s)}\,, \tag{5.11}$$

$$[F_{a;i,j}^{(r+1)}, F_{a+1;h,k}^{(s)}] - [F_{a;i,j}^{(r)}, F_{a+1;h,k}^{(s+1)}] =$$

$$(-1)^{|i|_{a+1}(|j|_a+|h|_{a+2})+|j|_a|h|_{a+2}+1}\delta_{i,k}\sum_{q=1}^{\mu_{a+1}} F^{(s)}_{a+1;h,q}F^{(r)}_{a;q,j}\,, \tag{5.12}$$

$$[E^{(r)}_{a;i,j},E^{(s)}_{b;h,k}]=0 \qquad \text{if} \quad |b-a|>1 \quad \text{or} \quad \text{if } b=a+1 \text{ and } h\neq j, \tag{5.13}$$

$$[F^{(r)}_{a;i,j},F^{(s)}_{b;h,k}]=0 \qquad \text{if} \quad |b-a|>1 \quad \text{or} \quad \text{if } b=a+1 \text{ and } i\neq k, \tag{5.14}$$

$$\big[E^{(r)}_{a;i,j},[E^{(s)}_{a;h,k},E^{(t)}_{b;f,g}]\big]+\big[E^{(s)}_{a;i,j},[E^{(r)}_{a;h,k},E^{(t)}_{b;f,g}]\big]=0 \quad \text{if } |a-b|\geq 1, \tag{5.15}$$

$$\big[F^{(r)}_{a;i,j},[F^{(s)}_{a;h,k},F^{(t)}_{b;f,g}]\big]+\big[F^{(s)}_{a;i,j},[F^{(r)}_{a;h,k},F^{(t)}_{b;f,g}]\big]=0 \quad \text{if } |a-b|\geq 1, \tag{5.16}$$

$$\big[\,[E^{(r)}_{a-1;i,f_1},E^{(t)}_{a;f_2,j}]\,,\,[E^{(t)}_{a;h,g_1},E^{(s)}_{a+1;g_2,k}]\,\big]=0 \;\; \text{if } z\geq 4 \text{ and } |h|_a+|j|_{a+1}=1, \tag{5.17}$$

$$\big[\,[F^{(r)}_{a-1;i,f_1},F^{(t)}_{a;f_2,j}]\,,\,[F^{(t)}_{a;h,g_1},F^{(s)}_{a+1;g_2,k}]\,\big]=0 \;\; \text{if } z\geq 4 \text{ and } |j|_a+|h|_{a+1}=1, \tag{5.18}$$

for all indices $a,b,f,f_1,f_2,g,g_1,g_2,h,i,j,k,r,s,t$ *that make sense. For example,* (5.11) *is supposed to hold for all* $1\leq a\leq z-2$, $1\leq i\leq \mu_a$, $1\leq h,j\leq \mu_{a+1}$, $1\leq k\leq \mu_{a+2}$, $r\geq s^{\mu}_{a,a+1}+1$, $s\geq s^{\mu}_{a+1,a+2}+1$.

In the special case where σ is the zero matrix, the above relations are precisely[2] the defining relations of $Y_{m|n}$ with respect to the parabolic generators $\mathcal{P}_\mu$ introduced in §3. As a result, we may simply write $Y_\mu=Y_\mu(0)$ instead of $Y_{m|n}$ to emphasize that we are using the parabolic presentation to define $Y_{m|n}$. The generators of $Y_\mu(\sigma)$, denoted by $\mathcal{P}_{\mu,\sigma}$, will be called the *parabolic generators* of $Y_\mu(\sigma)$. Later we will identify $\mathcal{P}_{\mu,\sigma}$ as subset of $\mathcal{P}_\mu$.

Remark 5.3. *As noticed in* [31, 40], *up to isomorphism, the definition of* Y_μ *is independent of the choice of the 01-sequence* Υ *since the RTT presentation of* $Y_{m|n}$ *is. For* $Y_\mu(\sigma)$, *we have a similar but slightly weaker phenomenon. Write* $Y_\mu(\sigma,\Upsilon)$ *for the shifted super Yangian to emphasize the choice of* Υ. *Let* S_{m+n} *be the symmetric group on* $m+n$ *objects, which acts on* Υ *by permutation, and let* S_μ *denote its Young subgroup associated to* μ. *We have*

$$Y_\mu(\sigma,\Upsilon)\cong Y_\mu(\sigma,\rho\cdot\Upsilon) \qquad \forall \rho\in S_\mu,$$

which is an immediate consequence of Remark 8.5 and Theorem 10.1.

Fix an admissible shape μ. Similar to §3, we will show that $Y_\mu(\sigma)$ can be identified as a subalgebra of Y_μ. Let $\Gamma:\mathcal{P}_{\mu,\sigma}\to\mathcal{P}_\mu$ be the map sending elements in $\mathcal{P}_{\mu,\sigma}$ to the elements (3.10) and (3.12) with the same name in $\mathcal{P}_\mu$ that are obtained by Gauss decomposition with respect to μ.

[2] Similar to Definition 4.1, when $\sigma=0$ the relations (5.17)–(5.18) are assumed to hold only for $t=1$ in [33], which suffices to imply that they in fact hold for all $t\geq 1$; see Proposition 5.15. In addition, (5.17)–(5.18) also hold when $|j|_a+|h|_{a+1}=1$.

Proposition 5.4. *The map* Γ *induces a canonical homomorphism* $\Gamma : Y_\mu(\sigma) \to Y_\mu$.

Proof. By [33], the relations (5.3)–(5.16) hold in Y_μ whenever the indices make sense. It remains to show that (5.17) and (5.18) also hold in Y_μ. These relations are crucial differences from the Lie algebra case in [8] and earlier partial results in [4, 32]. Checking these relations turns out to be very technical and involved. As a result, we postpone the proof to the end of this section; see Proposition 5.15. $\square$

For $1 \leq a < b \leq z$, $1 \leq i \leq \mu_a$, $1 \leq j \leq \mu_b$, $r > s^\mu_{a,b}$ and a fixed $1 \leq k \leq \mu_{b-1}$, we define the higher root elements $E^{(r)}_{a,b;i,j} \in Y_\mu(\sigma)$ recursively by

$$E^{(r)}_{a,a+1;i,j} := E^{(r)}_{a;i,j}, \qquad E^{(r)}_{a,b;i,j} := (-1)^{|k|_{b-1}}[E^{(r-s^\mu_{b-1,b})}_{a,b-1;i,k}, E^{(s^\mu_{b-1,b}+1)}_{b-1;k,j}]. \tag{5.19}$$

Similarly, using the same indices except for $r > s^\mu_{b,a}$, we define $F^{(r)}_{b,a;j,i} \in Y_\mu(\sigma)$ by

$$F^{(r)}_{a+1,a;j,i} := F^{(r)}_{a;j,i}, \qquad F^{(r)}_{b,a;j,i} := (-1)^{|k|_{b-1}}[F^{(s^\mu_{b,b-1}+1)}_{b-1;j,k}, F^{(r-s^\mu_{b,b-1})}_{b-1,a;k,i}]. \tag{5.20}$$

It turns out that the above definitions are independent of the choice of k; see Remark 5.8.

We introduce the *loop filtration* on $Y_\mu(\sigma)$

$$L_0Y_\mu(\sigma) \subseteq L_1Y_\mu(\sigma) \subseteq L_2Y_\mu(\sigma) \subseteq \cdots$$

by setting the degrees of the generators $D^{(r)}_{a;i,j}$, $E^{(r)}_{a;i,j}$, and $F^{(r)}_{a;i,j}$ to be $r-1$ and setting $L_kY_\mu(\sigma)$ to be the span of all supermonomials in the generators of total degree not greater than k. We let $\mathrm{gr}^L Y_\mu(\sigma)$ denote the associated graded superalgebra and define the elements $\{e^{(r)}_{a,b;i,j} \mid 1 \leq a, b \leq z, 1 \leq i \leq \mu_a, 1 \leq j \leq \mu_b, r \geq s^\mu_{a,b}\} \subseteq \mathrm{gr}^L Y_\mu(\sigma)$ by

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

The following is a parabolic version of Proposition 4.3, which can be proved by a similar argument.

Proposition 5.5. [7, Lemma 6.7][33, Lemma 7.5] *For all* $1 \leq a, b, c, d \leq z$, $1 \leq i \leq \mu_a$, $1 \leq j \leq \mu_b$, $r \geq s^\mu_{a,b}$, $t \geq s^\mu_{c,d}$, *the following identity holds in* $\mathrm{gr}^L Y_\mu(\sigma)$*:*

$$[e^{(r)}_{a,b;i,j}, e^{(t)}_{c,d;h,k}] = (-1)^{|j|_b}\delta_{b,c}\delta_{h,j}e^{(r+t)}_{a,d;i,k} - (-1)^{|i|_a|j|_b+|i|_a|h|_c+|j|_b|h|_c}\delta_{a,d}\delta_{i,k}e^{(r+t)}_{c,b;h,j}. \tag{5.21}$$

Theorem 5.6. *The map*

$$\gamma : U\big(\mathfrak{gl}_{m|n}[x](\sigma)\big) \longrightarrow \mathrm{gr}^L Y_\mu(\sigma)$$

defined by

$$\gamma(e_{\mu_1+\cdots+\mu_{a-1}+i,\mu_1+\cdots+\mu_{b-1}+j}x^r) = (-1)^{|i|_a}e^{(r)}_{a,b;i,j},$$

for all $1 \leq a, b \leq z$, $1 \leq i \leq \mu_a$, $1 \leq j \leq \mu_b$, $r \geq s^{\mu}_{a,b}$, is an isomorphism of graded superalgebras.

Proof. γ is a surjective homomorphism by (5.21). For injectivity, we start with the case $\sigma = 0$, where we already know that $Y_\mu(0) = Y_\mu$, and the statement follows from Corollary 3.3. For the general case, observe that the canonical map $\Gamma : Y_\mu(\sigma) \to Y_\mu$ is a homomorphism of filtered superalgebras (under loop filtration), and its induced map $\mathrm{gr}^L Y_\mu(\sigma) \to \mathrm{gr}^L Y_\mu$ sends $e^{(r)}_{a,b;i,j} \in \mathrm{gr}^L Y_\mu(\sigma)$ to $e^{(r)}_{a,b;i,j} \in \mathrm{gr}^L Y_\mu$. By the previous paragraph, the ordered supermonomials in the elements $\{e^{(r)}_{a,b;i,j} \,|\, 1 \leq a, b \leq m+n,\ r \geq s^{\mu}_{a,b}\}$ are linearly independent in $\mathrm{gr}^L Y_\mu(\sigma)$, hence γ is injective by the PBW theorem for $U\big(\mathfrak{gl}_{m|n}[x](\sigma)\big)$. □

Theorem 5.7. *Let $Y_\mu(\sigma)$ be the subalgebra of Y_μ generated by the union of the following subsets of $\mathcal{P}_\mu$:*

$$\big\{D^{(r)}_{a;i,j}, D'^{(r)}_{a;i,j} \,|\, 1 \leq a \leq z,\ 1 \leq i, j \leq \mu_a,\ r \geq 0\big\},$$
$$\big\{E^{(r)}_{b;h,k} \,|\, 1 \leq b < z,\ 1 \leq h \leq \mu_b, 1 \leq k \leq \mu_{b+1},\ r > s^{\mu}_{b,b+1}\big\},$$
$$\big\{F^{(r)}_{b;k,h} \,|\, 1 \leq b < z,\ 1 \leq k \leq \mu_{b+1}, 1 \leq h \leq \mu_b,\ r > s^{\mu}_{b+1,b}\big\}.$$

Then the relations (5.3)–(5.18) form a set of defining relations for $Y_\mu(\sigma)$. In other words, $Y_\mu(\sigma)$ defined in Definition 5.2 can be realized as a subalgebra of the super Yangian Y_μ.

Proof. We slightly change the notation in this proof to avoid possible confusion. Let $\tilde{Y}_\mu(\sigma)$ denote the abstract superalgebra generated by elements in $\mathcal{P}_{\mu,\sigma}$ with defining relations given in Definition 5.2 and let $Y_\mu(\sigma)$ denote the concrete subalgebra of Y_μ as stated in the theorem.

Let $\Gamma : \tilde{Y}_\mu(\sigma) \to Y_\mu(\sigma)$ be the canonical homomorphism in Proposition 5.4. Γ is clearly surjective, and it is injective as well by Theorem 5.6. □

Remark 5.8. *By Theorem 5.7, $E^{(r)}_{b;h,k}$ and $F^{(r)}_{b;k,h}$ are now concrete elements in Y_μ. Using the same argument as in [7, (6.9)] together with the admissible condition (5.2), one can show that the higher root elements defined recursively by (5.19) and (5.20) are independent of the choices of k.*

Let $Y^0_\mu(\sigma)$ denote the subalgebra of $Y_\mu(\sigma)$ generated by all of the $D^{(r)}_{a;i,j}$'s , $Y^+_\mu(\sigma)$ denote the subalgebra generated by all of the $E^{(r)}_{b;h,k}$'s and $Y^-_\mu(\sigma)$ denote the subalgebra generated by all of the $F^{(r)}_{b;k,h}$'s. The following corollary give PBW bases for these subalgebras.

Corollary 5.9. [8, Theorem 3.2]

(1) *The set of supermonomials in the elements*

$$\{D^{(r)}_{a;i,j} \,|\, 1 \leq a \leq z, 1 \leq i, j \leq \mu_a, r > 0\}$$

taken in some fixed order forms a basis for $Y^0_\mu(\sigma)$.

(2) *The set of supermonomials in the elements*

$$\{E^{(r)}_{a,b;h,k}\,|\,1\le a<b\le z, 1\le h\le \mu_a, 1\le k\le \mu_b, r>s^{\mu}_{a,b}\}$$

taken in some fixed order forms a basis for $Y^+_\mu(\sigma)$.

(3) *The set of supermonomials in the elements*

$$\{F^{(r)}_{b,a;k,h}\,|\,1\le a<b\le z, 1\le k\le \mu_b, 1\le h\le \mu_a, r>s^{\mu}_{b,a}\}$$

taken in some fixed order forms a basis for $Y^-_\mu(\sigma)$.

(4) *The set of supermonomials in the union of the elements listed in* (1)–(3) *taken in some fixed order forms a basis for* $Y_\mu(\sigma)$.

Proof. (4) follows from Theorem 5.6 and the PBW theorem for $U\big(\mathfrak{gl}_{m|n}[x](\sigma)\big)$. The others can be deduced by the same argument together with (5.21). □

Corollary 5.10. [8, Corollary 3.4] *The multiplicative map* $Y^-_\mu(\sigma)\otimes Y^0_\mu(\sigma)\otimes Y^+_\mu(\sigma)\longrightarrow Y_\mu(\sigma)$ *is an isomorphism of superspaces.*

Now we show that the definition of $Y_\mu(\sigma)$ is independent of the choice of the admissible shape μ. It suffices to show that $Y_\mu(\sigma)=Y_{(1^{m+n})}(\sigma)$. Assume that $\mu=(\mu_1,\dots,\mu_z)$ is admissible to σ. If $\mu_j=1$ for all j, then we have done. Otherwise, suppose that $\mu_p>1$ for some $1\le p\le z$ and we decompose $\mu_p=x+y$ for some positive integers x,y.

Define a finer composition ν of length $z+1$ by setting $\nu_i=\mu_i$ for all $1\le i\le p-1$, $\nu_p=x$, $\nu_{p+1}=y$, $\nu_{j+1}=\mu_j$ for all $p+1\le j\le z$; that is,

$$\nu=(\mu_1,\dots,\mu_{p-1},x,y,\mu_{p+1},\dots,\mu_z),$$

which is also admissible to σ by definition. We claim that

$$Y_\mu(\sigma)=Y_\nu(\sigma).$$

Now we prove our claim. Consider the Gauss decomposition of the matrix $T(u)$ with respect to μ and ν, respectively:

$$T(u)={}^{\mu}E(u){}^{\mu}D(u){}^{\mu}F(u)={}^{\nu}E(u){}^{\nu}D(u){}^{\nu}F(u),$$

where the matrices are block matrices as described in §3.

Denote by ${}^{\mu}D_a$ and ${}^{\nu}D_a$ the a-th diagonal matrices in ${}^{\mu}D(u)$ and ${}^{\nu}D(u)$ with respect to the compositions μ and ν, respectively. Similarly, let ${}^{\mu}E_a$ and ${}^{\mu}F_a$ denote the matrices in the a-th upper and the a-th lower diagonal of ${}^{\mu}E(u)$ and ${}^{\mu}F(u)$, respectively; ${}^{\nu}E_a$ and ${}^{\nu}F_a$ are defined to be the matrices in the a-th upper and the a-th lower diagonal of ${}^{\nu}E(u)$ and ${}^{\nu}F(u)$, respectively.

Lemma 5.11. [8, Lemma 3.1] *Using the above notation, define an $(x \times x)$-matrix A, an $(x \times y)$-matrix B, a $(y \times x)$-matrix C and a $(y \times y)$-matrix D from the equation*

$$ {}^{\mu}D_p = \begin{pmatrix} I_x & 0 \\ C & I_y \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix} \begin{pmatrix} I_x & B \\ 0 & I_y \end{pmatrix}. $$

Then

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

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

(iii) ${}^{\nu}F_a = {}^{\mu}F_a$ *for* $a < p-1$, ${}^{\nu}F_{p-1}$ *is the submatrix consisting of the first x rows of* ${}^{\mu}F_{p-1}$, ${}^{\nu}F_p = C$, ${}^{\mu}F_{p+1}$ *is the submatrix consisting of the last y columns of* ${}^{\mu}F_p$, *and* ${}^{\nu}F_c = {}^{\mu}F_{c-1}$ *for* $c > p+1$.

Proof. Matrix multiplication. □

As a consequence of Lemma 5.11, we see that $Y_\nu(\sigma) \subseteq Y_\mu(\sigma)$. Now the equality follows from the fact that the isomorphism $U\big(\mathfrak{gl}_{m|n}[x](\sigma)\big) \cong \operatorname{gr}^L Y_\mu(\sigma)$ is independent of the choice of μ. Applying induction on the length of the admissible shape μ, we have deduced the desired result.

Corollary 5.12. *$Y_\mu(\sigma)$ is independent of the choice of the admissible shape μ.*

Let σ be a shift matrix with an admissible shape μ. Note that the transpose matrix σ^t is again a shift matrix while μ is still admissible for σ^t. On the other hand, suppose that $\vec{\sigma} = (\vec{s}_{i,j})_{1\le i,j\le m+n}$ is another shift matrix satisfying (2.7) and the condition

$$ \vec{s}_{i,i+1} + \vec{s}_{i+1,i} = s_{i,i+1} + s_{i+1,i} $$

holds for all $1 \le i \le m+n-1$. As a result, if μ is an admissible shape for σ then it is also admissible for $\vec{\sigma}$. Denote by $\vec{D}^{(r)}_{a;i,j}$, $\vec{E}^{(r)}_{b;h,k}$ and $\vec{F}^{(r)}_{b;k,h}$ the parabolic generators of $Y_\mu(\vec{\sigma})$ to avoid confusion. The following results can be easily deduced from the presentation of $Y_\mu(\sigma)$.

Proposition 5.13. *The map $\tau : Y_\mu(\sigma) \to Y_\mu(\sigma^t)$ defined by*

$$ \tau(D^{(r)}_{a;i,j}) = D^{(r)}_{a;j,i},\ \ \tau(E^{(r)}_{b;h,k}) = F^{(r)}_{b;k,h},\ \ \tau(F^{(r)}_{b;k,h}) = E^{(r)}_{b;h,k} \tag{5.22} $$

is a superalgebra anti-isomorphism of order 2.

Proposition 5.14. *The map $\iota : Y_\mu(\sigma) \to Y_\mu(\vec{\sigma})$ defined by*

$$ \iota(D^{(r)}_{a;i,j}) = \vec{D}^{(r)}_{a;i,j},\quad \iota(E^{(r)}_{b;h,k}) = \vec{E}^{(r - s^{\mu}_{b,b+1} + \vec{s}^{\mu}_{b,b+1})}_{b;h,k},\quad \iota(F^{(r)}_{b;k,h}) = \vec{F}^{(r - s^{\mu}_{b+1,b} + \vec{s}^{\mu}_{b+1,b})}_{b;k,h}, \tag{5.23} $$

is a superalgebra isomorphism.

Now we prove the missing piece in the proof of Proposition 5.4.

Proposition 5.15. *The relations (5.17) and (5.18) hold in Y_μ, where $E^{(r)}_{b;h,k}$ and $F^{(r)}_{b;k,h}$ are the elements in Y_μ defined by (3.12).*

Proof. We only provide the detail for proving (5.17), while (5.18) is similar. Inspired by [9, §2.4], the proof is given by downward induction on the length of the admissible shape μ. Our initial step is the case $\mu = (1^{m+n})$, where (5.17) reduces to (4.15), which holds due to [40, (2.14)].

Assume now $\mu = (\mu_1, \dots, \mu_z)$ with $z < m+n$. Following the same notations given in the proof of Corollary 5.12, we may choose some $1 \le p \le z$ and decompose $\mu_p = x + y$ to obtain a new composition $\nu = (\mu_1, \dots, \mu_{p-1}, x, y, \mu_{p+1}, \dots, \mu_z)$. Now (5.17) and (5.18) hold in Y_ν by induction, which further implies that Theorem 5.7 holds for ν and hence we may identify $Y_\nu(\sigma)$ as a subalgebra of Y_ν. The key idea is to describe the relations between the elements ${}^\mu E^{(r)}_{b;h,k}$ and ${}^\nu E^{(r)}_{b;h,k}$.

Recall the set $\mathcal{P}_{\mu,\sigma}$ consisting of the following elements in Y_μ

$$\begin{aligned}
&\big\{{}^\mu D^{(r)}_{a;i,j} \,|\, 1 \le a \le z,\ 1 \le i,j \le \mu_a,\ r \ge 0\big\}\\
&\big\{{}^\mu E^{(r)}_{b;h,k} \,|\, 1 \le b < z,\ 1 \le h \le \mu_b, 1 \le k \le \mu_{b+1},\ r > s^\mu_{b,b+1}\big\}\\
&\big\{{}^\mu F^{(r)}_{b;k,h} \,|\, 1 \le b < z,\ 1 \le h \le \mu_b, 1 \le k \le \mu_{b+1},\ r > s^\mu_{b+1,b}\big\}
\end{aligned}$$

obtained by applying the Gauss decomposition to $T(u)$ with respect to μ. Similarly, replacing μ by ν, we have the following elements in Y_ν as well

$$\begin{aligned}
&\big\{{}^\nu D^{(r)}_{a;i,j} \,|\, 1 \le a \le z+1,\ 1 \le i,j \le \nu_a,\ r \ge 0\big\}\\
&\big\{{}^\nu E^{(r)}_{b;h,k} \,|\, 1 \le b \le z,\ 1 \le h \le \nu_b, 1 \le k \le \nu_{b+1},\ r > s^\nu_{b,b+1}\big\}\\
&\big\{{}^\nu F^{(r)}_{b;k,h} \,|\, 1 \le b \le z,\ 1 \le h \le \nu_b, 1 \le k \le \nu_{b+1},\ r > s^\nu_{b+1,b}\big\}
\end{aligned}$$

For every $1 \le a < b \le z+1$, $1 \le i \le \nu_a$, $1 \le j \le \nu_b$, we inductively define higher root elements ${}^\nu E^{(r)}_{a,b;i,j}$ for $r > s^\nu_{a,b}$ by equation (5.19) and similarly define ${}^\nu F^{(r)}_{b,a;j,i}$ for $r > s^\nu_{b,a}$ by equation (5.20). We further define the following formal series in $Y_\nu(\sigma)[[u^{-1}]]$:

$$ {}^\nu E_{a,b;i,j}(u) := \sum_{r > s^\nu_{a,b}} {}^\nu E^{(r)}_{a,b;i,j} u^{-r}, \qquad {}^\nu F_{b,a;j,i}(u) := \sum_{r > s^\nu_{b,a}} {}^\nu F^{(r)}_{b,a;j,i} u^{-r}, \tag{5.24}$$

and let ${}^\nu D_{a;i,j}(u)$ be given as in (3.7) with respect to ν. Note that the value of k in (5.19) and (5.20) can be arbitrarily chosen between 1 and ν_{b-1} due to Remark 5.8. Moreover, one should be careful that the series (5.24) are in general different from those series in $Y_\nu[[u^{-1}]]$ given by (3.11) so that we have to slightly modify the argument in the proof of Corollary 5.12.

Using these series, one defines the following matrices

$$
\begin{aligned}
{}^{\nu}D_a(u) &= \left({}^{\nu}D_{a;i,j}(u)\right)_{1\le i,j\le \nu_a}\\
{}^{\nu}E_{a,b}(u) &= \left({}^{\nu}E_{a,b;h,k}(u)\right)_{1\le h\le \nu_a, 1\le k\le \nu_b}\\
{}^{\nu}F_{b,a}(u) &= \left({}^{\nu}F_{b,a;k,h}(u)\right)_{1\le k\le \nu_b, 1\le h\le \nu_a}
\end{aligned}
$$

One further defines the block matrices ${}^{\nu}D(u)$, ${}^{\nu}E(u)$ and ${}^{\nu}F(u)$ exactly the same way as (3.6)–(3.9), except that we use their product to define the following matrix

$$
{}^{\nu}G(u) := {}^{\nu}F(u){}^{\nu}D(u){}^{\nu}E(u)
$$

By exactly the same way, one defines the higher root elements ${}^{\mu}E^{(r)}_{a,b;i,j}$, ${}^{\mu}F^{(r)}_{b,a;j,i}$, formal series ${}^{\mu}E_{a,b;i,j}(u)$, ${}^{\mu}F_{b,a;j,i}(u)$, ${}^{\mu}D_{a;i,j}(u)$, block matrices ${}^{\mu}D(u)$, ${}^{\mu}E(u)$ and ${}^{\mu}F(u)$ and hence their product ${}^{\mu}G(u) := {}^{\mu}F(u){}^{\mu}D(u){}^{\mu}E(u)$. A key observation from [9, §2.4] is that these two matrices are in fact the same and hence we have

$$
{}^{\nu}F(u){}^{\nu}D(u){}^{\nu}E(u) = {}^{\nu}G(u) = {}^{\mu}G(u) = {}^{\mu}F(u){}^{\mu}D(u){}^{\mu}E(u)
$$

As a consequence of Lemma 5.11, for each $1 \le a < b \le z$, $1 \le i \le \mu_a$ and $1 \le j \le \mu_b$, we have the following relation

$$
{}^{\mu}E_{a,b;i,j}(u) = \begin{cases}
{}^{\nu}E_{a,b;i,j}(u) & \text{if } b < p;\\
{}^{\nu}E_{a,b;i,j}(u) & \text{if } b = p, j \le x;\\
{}^{\nu}E_{a,b+1;i,j-x}(u) & \text{if } b = p, j > x;\\
{}^{\nu}E_{a,b+1;i,j}(u) & \text{if } a < p < b;\\
{}^{\nu}E_{a,b+1;i,j}(u) & \\
\qquad - \sum_{q=1}^{y} {}^{\nu}E_{a,a+1;i,q}(u){}^{\nu}E_{a+1,b+1;q,j}(u) & \text{if } a = p, i \le x;\\
{}^{\nu}E_{a+1,b+1;i-x,j}(u) & \text{if } a = p, i > x;\\
{}^{\nu}E_{a+1,b+1;i,j}(u) & \text{if } a > p.
\end{cases} \tag{5.25}
$$

Now let us back to the proof of (5.17). We may assume that $f_1 = f_2 = f$ and $g_1 = g_2 = g$ by (5.11). Moreover, by (5.25), ${}^{\mu}E^{(r)}_{a;i,j} = {}^{\nu}E^{(r)}_{a;i,j}$ except for $a \in \{p-1, p, p+1\}$ so the general case is further reduced to the special case $\mu = (\mu_1, \mu_2, \mu_3, \mu_4)$ since (5.17) holds for ν by induction. Therefore, it suffices to check the following relation holds in Y_μ for any $t > s^{\mu}_{2,3}$:

$$
\left[\,[{}^{\mu}E^{(r)}_{1;i,f}, {}^{\mu}E^{(t)}_{2;f,j}]\,,\,[{}^{\mu}E^{(t)}_{2;h,g}, {}^{\mu}E^{(s)}_{3;g,k}]\,\right] = 0 \tag{5.26}
$$

This can be checked by a case-by-case discussion. We list all possibilities below:

$$
p = 1, \qquad 1 \le i \le x \tag{5.27}
$$

$$
p = 1, \qquad 1 \le i - x \le y \tag{5.28}
$$

$$
p = 2, \qquad 1 \le f \le x, \qquad 1 \le h \le x \tag{5.29}
$$

$$
p = 2, \qquad 1 \le f - x \le y, \qquad 1 \le h \le x \tag{5.30}
$$

$$p = 2, \qquad 1 \le f \le x, \qquad 1 \le h - x \le y \tag{5.31}$$

$$p = 2, \qquad 1 \le f - x \le y, \qquad 1 \le h - x \le y \tag{5.32}$$

$$p = 3, \qquad 1 \le g \le x, \qquad 1 \le j \le x \tag{5.33}$$

$$p = 3, \qquad 1 \le g - x \le y, \qquad 1 \le j \le x \tag{5.34}$$

$$p = 3, \qquad 1 \le g \le x, \qquad 1 \le j - x \le y \tag{5.35}$$

$$p = 3, \qquad 1 \le g - x \le y, \qquad 1 \le j - x \le y \tag{5.36}$$

$$p = 4, \qquad 1 \le k \le x \tag{5.37}$$

$$p = 4, \qquad 1 \le k - x \le y \tag{5.38}$$

We will check some of them in detail here and the remaining ones can be deduced similarly.

Suppose that (5.27) holds. By (5.25), we have

$$ {}^{\mu}E^{(r)}_{1;i,f} = {}^{\nu}E^{(r)}_{1,3;i,f} - \sum_{s^{\nu}_{2,3}<q<r} \sum_{\ell=1}^{y} {}^{\nu}E^{(r-q)}_{1;i,\ell}\, {}^{\nu}E^{(q)}_{2;\ell,f}$$

Note that the admissible condition implies $s^{\mu}_{1,2} = s^{\nu}_{2,3}$ so the indices q and $r-q$ make sense. Then relation (5.26) becomes

$$\Big[\,[{}^{\nu}E^{(r)}_{1,3;i,f} - \sum_{s^{\nu}_{2,3}<q<r} \sum_{\ell=1}^{y} {}^{\nu}E^{(r-q)}_{1;i,\ell}\, {}^{\nu}E^{(q)}_{2;\ell,f},\ {}^{\nu}E^{(t)}_{3;f,j}]\,,\,[{}^{\nu}E^{(t)}_{3;h,g},\ {}^{\nu}E^{(s)}_{4;g,k}]\,\Big] =$$

$$\Big[\,[{}^{\nu}E^{(r)}_{1,3;i,f},\ {}^{\nu}E^{(t)}_{3;f,j}]\,,\,[{}^{\nu}E^{(t)}_{3;h,g},\ {}^{\nu}E^{(s)}_{4;g,k}]\,\Big] - \Big[\,[\sum_{s^{\nu}_{2,3}<q<r} \sum_{\ell=1}^{y} {}^{\nu}E^{(r-q)}_{1;i,\ell}\, {}^{\nu}E^{(q)}_{2;\ell,f},\ {}^{\nu}E^{(t)}_{3;f,j}]\,,\,[{}^{\nu}E^{(t)}_{3;h,g},\ {}^{\nu}E^{(s)}_{4;g,k}]\,\Big]$$

We first use the relation (5.19) to rewrite ${}^{\nu}E^{(r)}_{1,3;i,f} = (-1)^{|\ell|}[{}^{\nu}E^{(r-s^{\nu}_{2,3})}_{1;i,\ell},\ {}^{\nu}E^{(s^{\nu}_{2,3}+1)}_{2;\ell,f}]$. Then we use super Jacobi identity twice together with the fact that ${}^{\nu}E^{(r-s^{\nu}_{2,3})}_{1;i,\ell}$ and ${}^{\nu}E^{(t)}_{3;f,j}$ supercommute to rewrite the first term into

$$(-1)^{|\ell|}\Big[\,{}^{\nu}E^{(r-s^{\nu}_{2,3})}_{1;i,\ell},\,\big[\,[{}^{\nu}E^{(s^{\nu}_{2,3}+1)}_{2;\ell,f},\ {}^{\nu}E^{(t)}_{3;f,j}]\,,\,[{}^{\nu}E^{(t)}_{3;h,g},\ {}^{\nu}E^{(s)}_{4;g,k}]\,\big]\,\Big]$$

Similarly, up to an irrelevant sign factor, one can rewrite the second term as

$$\sum_{s^{\nu}_{2,3}<q<r} \sum_{\ell=1}^{y} {}^{\nu}E^{(r-q)}_{1;i,\ell}\,\big[\,[\,{}^{\nu}E^{(q)}_{2;\ell,f},\ {}^{\nu}E^{(t)}_{3;f,j}]\,,\,[{}^{\nu}E^{(t)}_{3;h,g},\ {}^{\nu}E^{(s)}_{4;g,k}]\,\big]$$

Now both of them are zero since (5.17) holds for ν by induction and the case (5.27) is proved.

Suppose that (5.34) holds. Using (5.25), we rewrite (5.26) into

$$\Big[\,[{}^{\nu}E^{(r)}_{1;i,f},{}^{\nu}E^{(t)}_{2;f,j}],[{}^{\nu}E^{(t)}_{2,4;h,g},{}^{\nu}E^{(s)}_{4;g,k}]\,\Big]$$

By relation (5.19), we have

$$ {}^{\nu}E^{(t)}_{2,4;h,g} = (-1)^{|\ell|}[{}^{\nu}E^{(t)}_{2;h,\ell},\ {}^{\nu}E^{(1)}_{3;\ell,g}], \tag{5.39}$$

where it is crucial to use the fact that $s^\nu_{3,4}=0$ due to the admissible condition. Following the same argument given in the case (5.27), one easily deduces that (5.26) is indeed zero in the case (5.34).

Now we prove the case (5.35). By (5.25), equation (5.26) becomes

$$\Big[\big[{}^\nu E^{(r)}_{1;i,f}, {}^\nu E^{(t)}_{2,4;f,j}\big]\,,\,\big[{}^\nu E^{(t)}_{2;h,g}, {}^\nu E^{(s)}_{3,5;g,k}-\sum_{s^\nu_{4,5}<q<s}\sum_{\ell=1}^{\nu_4}{}^\nu E^{(s-q)}_{3;g,\ell}\,{}^\nu E^{(q)}_{4;\ell,k}\big]\Big] \tag{5.40}$$

For convenience, write

$$B={}^\nu E^{(s)}_{3,5;g,k}-\sum_{s^\nu_{4,5}<q<s}\sum_{\ell=1}^{\nu_4}{}^\nu E^{(s-q)}_{3;g,\ell}\,{}^\nu E^{(q)}_{4;\ell,k}.$$

We need an extra relation before moving on. Applying the shift map ψ_{ν_1} in [33, Lemma 4.2] to the equation [33, (6.31)], one deduces the following relation in $Y_\nu[[u^{-1},v^{-1}]]$

$$\big[E_{2,4;f,j}(u),\,E_{3,5;g,k}(v)-\sum_{\ell=1}^{\nu_4}E_{3;g,\ell}(v)E_{4;\ell,k}(v)\big]=0 \tag{5.41}$$

We emphasize again that the series $E_{2,4;f,j}(u)$ and $E_{3,5;g,k}(v)$ in (5.41) are given by (3.8) and they are in general different from ${}^\nu E_{2,4;f,j}(u)$ and ${}^\nu E_{3,5;g,k}(v)$ defined by (5.24). Fortunately, $s^\nu_{3,4}=0$ due to the admissible condition so that we do have ${}^\nu E_{2,4;f,j}(u)=E_{2,4;f,j}(u)$. By using (5.11) in the case $\sigma=0$ multiple times, one deduces that

$$E^{(s)}_{3,5;g,k}={}^\nu E^{(s)}_{3,5;g,k}+\sum_{j=1}^{s^\nu_{4,5}}\sum_{\ell=1}^{\nu_4}E^{(s+j-1)}_{3;g,\ell}E^{(j)}_{4;\ell,k}.$$

As a result, we may rewrite (5.41) into the following identity in $Y_\nu(\sigma)$

$$\big[{}^\nu E^{(t)}_{2,4;f,j},\,{}^\nu E^{(s)}_{3,5;g,k}-\sum_{s^\nu_{4,5}<q<s}\sum_{\ell=1}^{\nu_4}{}^\nu E^{(s-q)}_{3;g,\ell}\,{}^\nu E^{(q)}_{4;\ell,k}\big]=\big[{}^\nu E^{(t)}_{2,4;f,j},\,B\big]=0 \tag{5.42}$$

By super Jacobi identity and (5.39), we rewrite (5.40) into

$$\begin{aligned}&\Big[\big[{}^\nu E^{(r)}_{1;i,f}, {}^\nu E^{(t)}_{2,4;f,j}\big]\,,\,\big[{}^\nu E^{(t)}_{2;h,g},\,B\big]\Big]\\&=\Big[\big[\big[{}^\nu E^{(r)}_{1;i,f}, {}^\nu E^{(t)}_{2,4;f,j}\big]\,,\,{}^\nu E^{(t)}_{2;h,g}\big]\,,\,B\Big]\pm\Big[{}^\nu E^{(t)}_{2;h,g}\,,\,\big[\big[{}^\nu E^{(r)}_{1;i,f}, {}^\nu E^{(t)}_{2,4;f,j}\big]\,,\,B\big]\Big]\end{aligned}$$

The second term is zero due to (5.42) and the fact that ${}^\nu E^{(r)}_{1;i,f}$ supercommute with B, which is a consequence of equation (5.13). Using (5.39) and super Jacobi identity, we rewrite the term inside the bracket of the first term as follows

$$\begin{aligned}\Big[\big[{}^\nu E^{(r)}_{1;i,f}, {}^\nu E^{(t)}_{2,4;f,j}\big]\,,\,{}^\nu E^{(t)}_{2;h,g}\Big]&=\Big[\big[{}^\nu E^{(r)}_{1;i,f}, (-1)^{|\ell|}\big[{}^\nu E^{(t)}_{2;f,\ell}, {}^\nu E^{(1)}_{3;\ell,j}\big]\big]\,,\,{}^\nu E^{(t)}_{2;h,g}\Big]\\&=\pm\Big[{}^\nu E^{(r)}_{1;i,f}\,,\,\big[{}^\nu E^{(t)}_{2;h,g},\big[{}^\nu E^{(t)}_{2;f,\ell}, {}^\nu E^{(1)}_{3;\ell,j}\big]\big]\Big]\pm\Big[\big[{}^\nu E^{(r)}_{1;i,f}\,,\,{}^\nu E^{(t)}_{2;h,g}\big]\,,\,\big[{}^\nu E^{(t)}_{2;f,\ell}, {}^\nu E^{(1)}_{3;\ell,j}\big]\Big]\end{aligned}$$

The first term is zero due to equation (5.15) while the second term is zero since (5.17) holds for ν by induction. This completes the proof of (5.26) in the case (5.35).

The cases (5.32) and (5.36) are similar to (5.34); the cases (5.28), (5.37) and (5.38) are immediate results of the induction hypothesis; the cases (5.31) and (5.33) are similar to (5.27); the cases (5.29) and (5.30) are similar to (5.35). □

Remark 5.16. *During the proof, one may observe that the index t in the middle two terms of (5.17) and (5.18) must be the same. In fact, for the special case $\mu = (1^{m+n})$ and $\sigma = 0$, their equivalent relations were firstly proposed in [39] with mistakes, allowing different t. However, such relations are too strong and the resulted structure collapses to trivial. The relations were noticed and corrected by Gow [22] and were generalized in [33, 40], eventually suggested the current forms of (5.17) and (5.18).*

6. Baby comultiplications

Although $Y_{m|n}$ is a Hopf-superalgebra, the shifted super Yangian $Y_\mu(\sigma)$ is not closed under the comultiplication defined by (3.3) in general; that is,

$$\Delta(Y_\mu(\sigma)) \nsubseteq Y_\mu(\sigma) \otimes Y_\mu(\sigma).$$

As compensation, we define some comultiplication-like maps on $Y_\mu(\sigma)$ as in [8, §4].

We first set up our assumptions and notations throughout this section. Let σ be a non-zero shift matrix of size $m+n$ with minimal admissible shape $\mu = (\mu_1, \ldots, \mu_z)$. Let Υ be a fixed 0^m1^n-sequence and let $Y_\mu(\sigma)$ be the shifted super Yangian defined in §5. Suppose that there are p 0's and q 1's in the very last μ_z digits of Υ; that is, Υ_z is a 0^p1^q-sequence and $\mu_z = p+q$. Since μ is minimal admissible and $\sigma \neq 0$, we have that $1 \leq \mu_z < m+n$ and either $s_{m+n-\mu_z, m+n+1-\mu_z} \neq 0$ or $s_{m+n+1-\mu_z, m+n-\mu_z} \neq 0$.

Theorem 6.1. *Let $\mu = (\mu_1, \mu_2, \ldots, \mu_z)$ be minimal admissible to σ. For $1 \leq i, j \leq \mu_z$, define*

$$\tilde{e}_{i,j} := e_{i,j} + \delta_{i,j}((m-p) - (n-q)) \in U(\mathfrak{gl}_{p|q}).$$

Here $e_{i,j}$ is the elementary matrix identified with the element in $\mathfrak{gl}_{p|q}$ and its parity is determined by the 0^p1^q-sequence Υ_z.

(1) *Suppose that $s_{m+n-\mu_z, m+n+1-\mu_z} \neq 0$. Define $\dot{\sigma} = (\dot{s}_{i,j})_{1 \leq i,j \leq m+n}$ by*

$$\dot{s}_{i,j} = \begin{cases} s_{i,j} - 1 & \text{if } i \leq m+n-\mu_z < j, \\ s_{i,j} & \text{otherwise.} \end{cases} \tag{6.1}$$

Then the map $\Delta_R : Y_\mu(\sigma) \to Y_\mu(\dot\sigma) \otimes U(\mathfrak{gl}_{p|q})$ *defined by*

$$\begin{aligned}
D^{(r)}_{a;i,j} &\mapsto \dot D^{(r)}_{a;i,j} \otimes 1 + \delta_{a,z} \sum_{f=1}^{\mu_z} (-1)^{|f|_z} \dot D^{(r-1)}_{a;i,f} \otimes \tilde e_{f,j},\\
E^{(r)}_{b;h,k} &\mapsto \dot E^{(r)}_{b;h,k} \otimes 1 + \delta_{b,z-1} \sum_{f=1}^{\mu_z} (-1)^{|f|_z} \dot E^{(r-1)}_{b;h,f} \otimes \tilde e_{f,k},\\
F^{(r)}_{b;k,h} &\mapsto \dot F^{(r)}_{b;k,h} \otimes 1,
\end{aligned}$$

is a superalgebra homomorphism.

(2) *Suppose that* $s_{m+n+1-\mu_z,m+n-\mu_z} \neq 0$. *Define* $\dot\sigma = (\dot s_{i,j})_{1\le i,j\le m+n}$ *by*

$$\dot s_{i,j} = \begin{cases} s_{i,j} - 1 & \text{if } j \le m+n-\mu_z < i,\\ s_{i,j} & \text{otherwise.}\end{cases} \tag{6.2}$$

Then the map $\Delta_L : Y_\mu(\sigma) \to U(\mathfrak{gl}_{p|q}) \otimes Y_\mu(\dot\sigma)$ *defined by*

$$\begin{aligned}
D^{(r)}_{a;i,j} &\mapsto 1 \otimes \dot D^{(r)}_{a;i,j} + \delta_{a,z}(-1)^{|i|_z} \sum_{k=1}^{\mu_z} \tilde e_{i,k} \otimes \dot D^{(r-1)}_{a;k,j},\\
E^{(r)}_{b;h,k} &\mapsto 1 \otimes \dot E^{(r)}_{b;h,k},\\
F^{(r)}_{b;k,h} &\mapsto 1 \otimes \dot F^{(r)}_{b;k,h} + \delta_{b,z-1}(-1)^{|k|_z} \sum_{f=1}^{\mu_z} \tilde e_{k,f} \otimes \dot F^{(r-1)}_{b;f,h},
\end{aligned}$$

is a superalgebra homomorphism.

To avoid possible confusion, in the above description and hereafter, the parabolic generators of $Y_\mu(\dot\sigma)$ *are denoted by* $\dot D^{(r)}_{a;i,j}$, $\dot E^{(r)}_{a;i,j}$, *and* $\dot F^{(r)}_{a;i,j}$, *where* $\dot\sigma$ *is the shift matrix defined by either* (6.1) *or* (6.2), *with respect to the same shape* μ *which is also admissible to* $\dot\sigma$.

Proof. It is straightforward to check that Δ_R and Δ_L preserve the defining relations in Definition 5.2. Note that it suffices to check the special case $z = 4$ since the non-trivial situations only happen in the very last block. Similar to [8, Theorem 4.2], to check (5.15) and (5.16), one needs to use (5.9), (5.10), (5.11) and (5.12) multiple times.

We check (5.18) here as an illustrating example since it is a super phenomenon which does not appear in [8]. Assume $z = 4$ and (6.2) holds. Applying Δ_L to the left-hand-side of (5.18), we have

$$\big[\,[1\otimes \dot F^{(r)}_{1;i,f_1}, 1\otimes \dot F^{(t)}_{2;f_2,j}]\,,\,[1\otimes \dot F^{(t)}_{2;h,g_1}, 1\otimes \dot F^{(s)}_{3;g_2,k} + (-1)^{|g_2|_4} \sum_{x=1}^{\mu_4} \tilde e_{g_2,x} \otimes \dot F^{(s-1)}_{3;x,k}]\,\big] \tag{6.3}$$

Recall that for associative superalgebras A and B, their tensor product $A\otimes B$ is naturally a superalgebra where $|a\otimes b| := |a|+|b|$ for homogeneous $a\in A$, $b\in B$. Given $x\otimes y$ and $a\otimes b$ in $A\otimes B$, their supercommutator is explicitly given by

$$\begin{aligned}[x\otimes y, a\otimes b] &= (x\otimes y)(a\otimes b) - (-1)^{(|x|+|y|)(|a|+|b|)}(a\otimes b)(x\otimes y)\\ &= (-1)^{|a||y|}(xa\otimes yb) - (-1)^{(|x|+|y|)(|a|+|b|)+|x||b|}(ax\otimes by)\end{aligned}$$

By the formula above, (6.3) equals to

$$1\ \otimes\ \big[\,[\dot F^{(r)}_{1;i,f_1}, \dot F^{(t)}_{2;f_2,j}]\,,\,[\dot F^{(t)}_{2;h,g_1}, \dot F^{(s)}_{3;g_2,k}]\,\big]\ +\ \theta\sum_{x=1}^{\mu_4}\tilde e_{g_2,x}\ \otimes\ \big[\,[\dot F^{(r)}_{1;i,f_1}, \dot F^{(t)}_{2;f_2,j}]\,,\,[\dot F^{(t)}_{2;h,g_1}, \dot F^{(s-1)}_{3;x,k}]\big],$$

where $\theta=\pm1$ is an irrelevant sign. It vanishes due to (5.18) in $Y_{m|n}(\dot\sigma)$. □

The next lemma computes the images of higher root elements $E^{(r)}_{a,b;i,j}$ and $F^{(r)}_{b,a;i,j}$ under Δ_R and Δ_L.

Lemma 6.2. *(1) Suppose the assumption of Theorem 6.1(1) holds. For all admissible indices i,j,r and $1\le a<b-1<z$, we have*

$$\begin{aligned}\Delta_R(F^{(r)}_{b,a;i,j}) &= \dot F^{(r)}_{b,a;i,j}\otimes 1,\\ \Delta_R(E^{(r)}_{a,b;i,j}) &= \dot E^{(r)}_{a,b;i,j}\otimes 1 \qquad \text{if } b<z,\end{aligned}$$

and

$$\Delta_R(E^{(r)}_{a,z;i,j}) \;=\; (-1)^{|h|_{z-1}}[\dot E^{(r-s^{\mu}_{z-1,z})}_{a,z-1;i,h}, \dot E^{(s^{\mu}_{z-1,z}+1)}_{z-1;h,j}]\ \otimes\ 1\ +\ \sum_{k=1}^{\mu_z}(-1)^{|k|_z}\dot E^{(r-1)}_{a,z;i,k}\ \otimes\ \tilde e_{k,j},$$

for any $1\le h\le \mu_{z-1}$.

(2) Suppose the assumption of Theorem 6.1(2) holds. For all admissible indices i,j,r and $1\le a<b-1<z$, we have

$$\begin{aligned}\Delta_L(E^{(r)}_{a,b;i,j}) &= 1\otimes \dot E^{(r)}_{a,b;i,j},\\ \Delta_L(F^{(r)}_{b,a;i,j}) &= 1\otimes \dot F^{(r)}_{b,a;i,j} \qquad \text{if } b<z,\end{aligned}$$

and

$$\Delta_L(F^{(r)}_{z,a;i,j}) \;=\; (-1)^{|h|_{z-1}}\big(1\ \otimes\ [\dot F^{(s^{\mu}_{z,z-1}+1)}_{z-1;i,h}, \dot F^{(r-s^{\mu}_{z,z-1})}_{z-1,a;h,j}]\big)\ +\ (-1)^{|i|_z}\sum_{k=1}^{\mu_z}\tilde e_{i,k}\ \otimes\ \dot F^{(r-1)}_{z-1,a;k,j},$$

for any $1\le h\le \mu_{z-1}$.

Proof. We compute $\Delta_R(E^{(r)}_{a,z;i,j})$ for $1\le a<z-1$ in detail here, while others are similar. By definition, for any $1\le h\le\mu_{z-1}$, we have

$$E^{(r)}_{a,z;i,j} = (-1)^{|h|_{z-1}}[E^{(r-s^{\mu}_{z-1,z})}_{a,z-1;i,h}, E^{(s^{\mu}_{z-1,z}+1)}_{z-1;h,j}].$$

Also, $\Delta_R(E^{(r-s^\mu_{z-1,z})}_{a,z-1;i,h}) = \dot{E}^{(r-s^\mu_{z-1,z})}_{a,z-1;i,h} \otimes 1$. Hence

$$\begin{aligned}\Delta_R(E^{(r)}_{a,z;i,j}) &= (-1)^{|h|_{z-1}} \left[\dot{E}^{(r-s^\mu_{z-1,z})}_{a,z-1;i,h} \otimes 1, \dot{E}^{(s^\mu_{z-1,z}+1)}_{z-1;h,j} \otimes 1\right] \\ &\quad + (-1)^{|h|_{z-1}} \left[\dot{E}^{(r-s^\mu_{z-1,z})}_{a,z-1;i,h} \otimes 1, \sum_{k=1}^{\mu_z} (-1)^{|k|_z} \dot{E}^{(s^\mu_{z-1,z})}_{z-1;h,k} \otimes \tilde{e}_{k,j}\right] \\ &= (-1)^{|h|_{z-1}} \left[\dot{E}^{(r-s^\mu_{z-1,z})}_{a,z-1;i,h}, \dot{E}^{(s^\mu_{z-1,z}+1)}_{z-1;h,j}\right] \otimes 1 + \sum_{k=1}^{\mu_z} (-1)^{|k|_z} \dot{E}^{(r-1)}_{a,z;i,k} \otimes \tilde{e}_{k,j}.\end{aligned}$$

□

Proposition 6.3. *If the assumption of Theorem 6.1(1) holds, then Δ_R is injective. Similarly, if the assumption of Theorem 6.1(2) holds, then Δ_L is injective.*

Proof. Let $\epsilon : U(\mathfrak{gl}_{p|q}) \to \mathbb{C}$ be the homomorphism such that

$$\epsilon(\tilde{e}_{i,j}) = 0$$

for $1 \le i, j \le \mu_z$. By definition, $Y_\mu(\sigma) \subseteq Y_\mu(\dot{\sigma}) \subseteq Y_\mu$ is a chain of subalgebras. Note that the compositions $m \circ (\mathrm{id} \otimes \epsilon) \circ \Delta_R$ and $m \circ (\epsilon \otimes \mathrm{id}) \circ \Delta_L$ coincide with the natural embedding $Y_\mu(\sigma) \hookrightarrow Y_\mu(\dot{\sigma})$, where $m(a \otimes b) := ab$ is the usual multiplication map. This implies that the maps Δ_R and Δ_L are injective whenever they are defined. □

7. Canonical filtration

There is another filtration on $Y_{m|n}$, called the *canonical filtration*

$$F_0 Y_{m|n} \subseteq F_1 Y_{m|n} \subset F_2 Y_{m|n} \subseteq \cdots$$

defined by $\deg t^{(r)}_{ij} := r$ where $F_d Y_{m|n}$ is defined to be the span of all supermonomials in $t^{(r)}_{ij}$ of total degree not greater than d. Let $\mathrm{gr}\, Y_{m|n}$ denote the associated superalgebra, which is supercommutative by (3.2).

Now we describe the canonical filtration using parabolic presentations. Let $\mu = (\mu_1, \ldots, \mu_z)$ be a composition of $m+n$. By [33, Proposition 3.1], the parabolic generators $D^{(r)}_{a;i,j}$ $E^{(r)}_{a,b;i,j}$ and $F^{(r)}_{b,a;i,j}$ of Y_μ are linear combinations of supermonomials in $t^{(s)}_{i,j}$ of total degree r.

On the other hand, if we set $D^{(r)}_{a;i,j}$, $E^{(r)}_{a,b;i,j}$ and $F^{(r)}_{b,a;i,j}$ all to be of degree r, by multiplying the matrix equation $T(u) = F(u)D(u)E(u)$, each $t^{(r)}_{ij}$ is a linear combination of supermonomials in the parabolic generators of total degree r as well. Thus $F_d Y_{m|n}$ can be alternatively defined as the span of all supermonomials in the parabolic generators $D^{(r)}_{a;i,j}$ $E^{(r)}_{a,b;i,j}$ and $F^{(r)}_{b,a;i,j}$ of total degree $\le d$.

For $1 \le a, b \le z$, $1 \le i \le \mu_a$, $1 \le j \le \mu_b$ and $r > 0$, define the following elements in $\operatorname{gr} Y_\mu$ by

$$e^{(r)}_{a,b;i,j} := \begin{cases} \operatorname{gr}_r D^{(r)}_{a;i,j} & \text{if } a = b, \\ \operatorname{gr}_r E^{(r)}_{a,b;i,j} & \text{if } a < b, \\ \operatorname{gr}_r F^{(r)}_{a,b;i,j} & \text{if } a > b. \end{cases} \tag{7.1}$$

Since $\operatorname{gr} Y_\mu$ is supercommutative, together with Corollary 5.9 (4), the following result can be deduced immediately.

Proposition 7.1. [8, Theorem 5.1] *For any shape $\mu = (\mu_1, \dots, \mu_z)$, $\operatorname{gr} Y_\mu$ is the free supercommutative superalgebra on generators $\{e^{(r)}_{a,b;i,j} \,|\, 1 \le a, b \le z, 1 \le i \le \mu_a, 1 \le j \le \mu_b, r > 0\}$.*

Suppose now σ is a shift matrix of size $m + n$ and $\mu = (\mu_1, \dots, \mu_z)$ is an admissible shape to σ. We induce the canonical filtration of Y_μ to the subalgebra $Y_\mu(\sigma)$ by defining

$$F_d Y_\mu(\sigma) := F_d Y_\mu \cap Y_\mu(\sigma).$$

The natural embedding $Y_\mu(\sigma) \hookrightarrow Y_\mu$ is a filtered map and the induced map $\operatorname{gr} Y_\mu(\sigma) \to \operatorname{gr} Y_\mu$ is injective as well, so that we may identify $\operatorname{gr} Y_\mu(\sigma)$ as a subalgebra of $\operatorname{gr} Y_\mu$. The next theorem gives a set of generators of $\operatorname{gr} Y_\mu(\sigma)$.

Theorem 7.2. [8, Theorem 5.2] *For an admissible shape $\mu = (\mu_1, \dots, \mu_z)$, $\operatorname{gr} Y_\mu(\sigma)$ is the subalgebra of $\operatorname{gr} Y_\mu$ generated by the elements*

$$\{e^{(r)}_{a,b;i,j} \,|\, 1 \le a, b \le z, 1 \le i \le \mu_a, 1 \le j \le \mu_b, r > s^\mu_{a,b}\}.$$

Proof. By relations (5.11) and (5.12), the elements $e^{(r)}_{a,b;i,j}$ of $\operatorname{gr} Y_\mu(\sigma)$ can be identified as the elements of the same notation in $\operatorname{gr} Y_\mu$ defined in (7.1) by the embedding $\operatorname{gr} Y_\mu(\sigma) \to \operatorname{gr} Y_\mu$. Now the statement follows from Corollary 5.9 (4) and Proposition 7.1. $\square$

Similar to [8, Remark 5.3], one consequence of Theorem 7.2 is that we may define the canonical filtration on $Y_\mu(\sigma)$ intrinsically by setting the degree of the elements $D^{(r)}_{a;i,j}$, $E^{(r)}_{a,b;i,j}$ and $F^{(r)}_{ab,a;j,i}$ in $Y_\mu(\sigma)$ to be r. By Corollary 5.12, such a definition is independent of the choice of admissible shape μ.

By definition, the comultiplication $\Delta : Y_\mu \to Y_\mu \otimes Y_\mu$ is a filtered map with respect to the canonical filtration. If we extend the canonical filtration of $Y_\mu(\dot\sigma)$ to $Y_\mu(\dot\sigma) \otimes U(\mathfrak{gl}_{p|q})$ by declaring the degree of the matrix unit $e_{ij} \in \mathfrak{gl}_{p|q}$ to be 1, then the baby comultiplications Δ_R and Δ_L defined in Theorem 6.1, as long as they are defined, are filtered maps as well. Moreover, the same argument as in the proof of Proposition 6.3 implies that the associated graded maps

$$\operatorname{gr} \Delta_L : \operatorname{gr} Y_\mu(\dot\sigma) \to \operatorname{gr}\big(Y_\mu(\dot\sigma) \otimes U(\mathfrak{gl}_{p|q})\big) \qquad \operatorname{gr} \Delta_R : \operatorname{gr} Y_\mu(\dot\sigma) \to \operatorname{gr}\big(U(\mathfrak{gl}_{p|q}) \otimes Y_\mu(\dot\sigma)\big)$$

are injective as well. We state this fact as a proposition.

Proposition 7.3. [8, Remark 5.4] *The induced maps* $\operatorname{gr}\Delta_R$ *and* $\operatorname{gr}\Delta_L$ *are injective whenever they are defined,*

8. Truncation

Let σ be a fixed shift matrix of size $m+n$. Choose an integer $\ell > s_{1,m+n}+s_{m+n,1}$, which will be called *level* later. For each $1 \le i \le m+n$, set

$$p_i := \ell - s_{i,m+n} - s_{m+n,i}. \tag{8.1}$$

This defines a tuple $(p_1, \dots, p_{m+n})$ of integers such that $0 < p_1 \le \cdots \le p_{m+n} = \ell$. Let $\mu = (\mu_1, \dots, \mu_z)$ be an admissible shape for σ. For each $1 \le a \le z$, set

$$p^\mu_a := p_{\mu_1+\ldots+\mu_a}. \tag{8.2}$$

Since μ is admissible, together with (2.7), for any $1 \le a \le z$, we have $p_i = p^\mu_a$ for any value of i such that $1 \le i - \sum_{k=1}^{a-1} \mu_k \le \mu_a$.

Following [8, §6], we define the *shifted super Yangian of level* ℓ, denoted by $Y^\ell_\mu(\sigma)$, to be the quotient of $Y_\mu(\sigma)$ by the two-side ideal of $Y_\mu(\sigma)$ generated by

$$\{D^{(r)}_{1;i,j} \,|\, 1 \le i,j \le \mu_1, r > p_1\}.$$

We claim that the definition of $Y^\ell_\mu(\sigma)$ is independent of the choice of the admissible shape μ so that we may simply write $Y^\ell_{m|n}(\sigma)$ when appropriate. Let I_μ denote the two-sided ideal associated to μ as in the definition. Since $\nu = (1^{m+n})$ is admissible for any σ, it suffices to prove that $I_\mu = I_\nu$.

By definition, we have ${}^\nu D^{(r)}_1 = t^{(r)}_{1,1}$. Assume μ is an arbitrary admissible shape. By [33, (3.10)], we have ${}^\mu D^{(r)}_{1;1,1} = t^{(r)}_{1,1} = {}^\nu D^{(r)}_1$ and hence $I_\nu \subseteq I_\mu$. On the other hand, one may deduce from (5.5) that ${}^\mu D^{(r)}_{1;i,j} \in I_\nu$ for all $1 \le i,j \le \mu_1$, $r > p_1$, and our claim follows.

When $\sigma = 0$, the two-sided ideal is generated by $\{t^{(r)}_{i,j} \,|\, 1 \le i,j \le m+n,\, r > \ell\}$. In this special case, the quotient is exactly the *truncated super Yangian* in [3, 31], which is a super analogy of *Yangian of level* ℓ due to Cherednik [12, 13]. It should be clear from the context that we are dealing with $Y_\mu(\sigma)$ or the quotient $Y^\ell_\mu(\sigma)$ and hence, by abusing notation, we will use the same symbols $D^{(r)}_{a;i,j}$, $E^{(r)}_{a,b;i,j}$ and $F^{(r)}_{b,a;i,j}$ to denote the elements in $Y_\mu(\sigma)$ and their images in the quotient $Y^\ell_\mu(\sigma)$.

It is obvious that the anti-isomorphism τ defined in (5.22) factors through the quotient and induces an anti-isomorphism

$$\tau : Y^\ell_\mu(\sigma) \to Y^\ell_\mu(\sigma^t). \tag{8.3}$$

Similarly, let $\vec{\sigma}$ be another shift matrix satisfying that $\vec{s}_{i,i+1}+\vec{s}_{i+1,i}=s_{i,i+1}+s_{i+1,i}$ for all $1\leq i\leq m+n-1$. Then the isomorphism ι defined by (5.23) also induces an isomorphism

$$\iota: Y_\mu^\ell(\sigma)\to Y_\mu^\ell(\vec{\sigma}). \tag{8.4}$$

Recall the canonical filtration defined in §7. We obtain a filtration

$$F_0Y_\mu^\ell(\sigma)\subseteq F_1Y_\mu^\ell(\sigma)\subseteq\cdots$$

induced from the quotient map $Y_\mu(\sigma)\to Y_\mu^\ell(\sigma)$, where we define the elements $D_{a;i,j}^{(r)}$, $E_{a,b;i,j}^{(r)}$ and $F_{b,a;i,j}^{(r)}$ of $Y_\mu^\ell(\sigma)$ to be of degree r and $F_dY_\mu^\ell(\sigma)$ is the span of all supermonomials in these elements of total degree $\leq d$.

For $1\leq a,b\leq z$, $1\leq i\leq\mu_a$, $1\leq j\leq\mu_b$ and $r>s_{a,b}^\mu$, define element $e_{a,b;i,j}^{(r)}$ (by abusing notation again) in the associative graded superalgebra $\operatorname{gr}Y_\mu^\ell(\sigma)$ according to exactly the same formula (7.1), except that now our D's, E's and F's here are in the quotient. By Proposition 7.1 and Theorem 7.2, $\operatorname{gr}Y_\mu^\ell(\sigma)$ is also supercommutative and is generated by the elements

$$\{e_{a,b;i,j}^{(r)}\in\operatorname{gr}Y_\mu^\ell(\sigma)\,|\,1\leq a,b\leq z,1\leq i\leq\mu_a,1\leq j\leq\mu_b,r>s_{a,b}^\mu\}$$

Following the same argument in [8, Lemma 6.1], one may deduce that $\operatorname{gr}Y_\mu^\ell(\sigma)$ is in fact *finitely generated*.

Lemma 8.1. *For any admissible shape $\mu=(\mu_1,\ldots,\mu_z)$, $\operatorname{gr}Y_\mu^\ell(\sigma)$ is generated only by the elements*

$$\{e_{a,b;i,j}^{(r)}\,|\,1\leq a,b\leq z,1\leq i\leq\mu_a,1\leq j\leq\mu_b,s_{a,b}^\mu<r\leq s_{a,b}^\mu+p_{min(a,b)}^\mu\}.$$

Let $\sigma=(s_{ij})_{1\leq i,j\leq m+n}$ be a non-zero shift matrix with minimal admissible shape $\mu=(\mu_1,\ldots,\mu_z)$ and let Υ be a 0^m1^n-sequence. Then μ_z equals to the size of the largest zero square matrix in the southeastern corner of σ. Hence we have $1\leq\mu_z<m+n$ and either $s_{m+n-\mu_z,m+n+1-\mu_z}\neq0$ or $s_{m+n+1-\mu_z,m+n-\mu_z}\neq0$. Let p and q denote the the number of 0's and 1's respectively in the last μ_z digits of the 0^m1^n-sequence Υ.

Suppose that $s_{m+n-\mu_z,m+n+1-\mu_z}\neq0$. By definition, for all $1\leq i,j\leq\mu_1$, we have $\Delta_R(D_{1;i,j}^{(r)})=\dot{D}_{1;i,j}^{(r)}\otimes1$. If in addition $r>p_1$, then clearly $\dot{D}_{1;i,j}^{(r)}$ equals to zero in the quotient $Y_\mu^{\ell-1}(\dot{\sigma})$. It implies that the baby comultiplication Δ_R defined in Theorem 6.1 factors through the quotient and we obtain an induced map

$$\Delta_R: Y_\mu^\ell(\sigma)\to Y_\mu^{\ell-1}(\dot{\sigma})\otimes U(\mathfrak{gl}_{p|q}) \tag{8.5}$$

where $\dot{\sigma}$ is given by (6.1).

Similarly, if $s_{m+n+1-\mu_z,m+n-\mu_z}\neq0$, then Δ_L induces a map

$$\Delta_L: Y_\mu^\ell(\sigma)\to U(\mathfrak{gl}_{p|q})\otimes Y_\mu^{\ell-1}(\dot{\sigma}) \tag{8.6}$$

where $\dot{\sigma}$ is given by (6.2).

Recall that Δ_R and Δ_L are filtered maps with respect to the canonical filtration, so they induce the following homomorphisms of graded superalgebras

$$\operatorname{gr}\Delta_R : \operatorname{gr} Y_\mu^\ell(\sigma) \to \operatorname{gr}\big(Y_\mu^{\ell-1}(\dot\sigma) \otimes U(\mathfrak{gl}_{p|q})\big), \tag{8.7}$$

$$\operatorname{gr}\Delta_L : \operatorname{gr} Y_\mu^\ell(\sigma) \to \operatorname{gr}\big(U(\mathfrak{gl}_{p|q}) \otimes Y_\mu^{\ell-1}(\dot\sigma)\big). \tag{8.8}$$

Theorem 8.2. *For any admissible shape $\mu = (\mu_1, \ldots, \mu_z)$, $\operatorname{gr} Y_\mu^\ell(\sigma)$ is the free supercommutative superalgebra on generators*

$$\{e_{a,b;i,j}^{(r)} \,|\, 1 \leq a, b \leq z, 1 \leq i \leq \mu_a, 1 \leq j \leq \mu_b, s_{a,b}^\mu < r \leq s_{a,b}^\mu + p_{min(a,b)}^\mu\}.$$

Also, the maps $\operatorname{gr}\Delta_R$ and $\operatorname{gr}\Delta_L$ in (8.7) and (8.8) are injective whenever they are defined, and so are the maps Δ_R and Δ_L in (8.5) and (8.6).

Proof. Similar to the argument in [8, Theorem 6.2], except that our induction starts from $\ell = 1$. In that case, the assertion follows from [31, Proposition 2.3]. □

As a corollary, we obtain a PBW basis for $Y_{m|n}^\ell(\sigma)$.

Corollary 8.3. *For any admissible shape $\mu = (\mu_1, \ldots, \mu_z)$, the supermonomials in the elements*

$$\{D_{a;i,j}^{(r)} | 1 \leq a \leq z, 1 \leq i, j \leq \mu_a, 0 < r \leq p_a^\mu\},$$
$$\{E_{a,b;i,j}^{(r)} | 1 \leq a < b \leq z, 1 \leq i \leq \mu_a, 1 \leq j \leq \mu_b, s_{a,b}^\mu < r \leq s_{a,b}^\mu + p_a^\mu\},$$
$$\{F_{b,a;i,j}^{(r)} | 1 \leq a < b \leq z, 1 \leq i \leq \mu_b, 1 \leq j \leq \mu_a, s_{b,a}^\mu < r \leq s_{b,a}^\mu + p_a^\mu\},$$

taken in any fixed order forms a basis for $Y_{m|n}^\ell(\sigma)$.

Another corollary is obtained by counting.

Corollary 8.4. *Consider $Y_{m|n}^\ell(\sigma)$ together with the canonical filtration and some fixed Υ. Let $S(\mathfrak{g}^e)$ be the supersymmetric superalgebra of $\mathfrak{g}^e$ with the Kazhdan filtration, where e is the nilpotent element corresponding to the triple (σ, ℓ, Υ) as explained in §2. Denote by $F_d Y_{m|n}^\ell(\sigma)$ and $F_d S(\mathfrak{g}^e)$ the superspaces with total degree not greater than d in the associated filtered superalgebras respectively. Then for each $d \geq 0$, we have $\dim F_d Y_{m|n}^\ell(\sigma) = \dim F_d S(\mathfrak{g}^e)$.*

Proof. Take $\mu = (1^{m+n})$ in Theorem 8.2. Then the statement follows from Proposition 2.9 and induction on d. □

Remark 8.5. *Consider the following inverse system*

$$Y_{m|n}^\ell(\sigma) \twoheadleftarrow Y_{m|n}^{\ell+1}(\sigma) \twoheadleftarrow Y_{m|n}^{\ell+2}(\sigma) \twoheadleftarrow \cdots$$

where the maps are homomorphisms of filtered superalgebras with respect to the canonical filtration. As an observation from Corollary 5.9 (4) and Corollary 8.3, we have

$$Y_{m|n}(\sigma) = \lim_{\leftarrow} Y_{m|n}^\ell(\sigma)$$

where the inverse limit is taken in the category of filtered superalgebras. Thus, similar to [8, Remark 6.4], *we may view $Y_{m|n}(\sigma)$ as the inverse limit $\ell \to \infty$ of the shifted super Yangians of level ℓ.*

9. Invariants

Let π be a given pyramid of height $m+n$ associated to a 0^m1^n-sequence Υ. Let M and N be the number of boxes in π labeled by " $+$ " and " $-$ ", respectively. Let $\mathfrak{p}$ and $\mathfrak{m}$ be the subalgebras of $\mathfrak{gl}_{M|N}$ associated to the good pair (e_π, h_π). Generalizing [8, §9], we will define some distinguished elements in $U(\mathfrak{p})$. In the next section we will show that many of them are $\mathfrak{m}$-invariant (under the χ-twisted action) and hence they are elements in $\mathcal{W}_\pi$.

We number the columns of π from left to right by $1, \ldots, \ell$. Let $h = m - n$ and let $(\check{q}_1, \ldots, \check{q}_\ell)$ denote the *super column heights* of π, where each $\check{q}_i$ is defined to be the number of boxes in the i-th column of π labeled with " $+$ " subtract the number of boxes labeled with " $-$ " in the same column.

Define $\rho = (\rho_1, \ldots, \rho_\ell)$, where ρ_r is given by

$$\rho_r := h - \check{q}_r - \check{q}_{r+1} - \cdots - \check{q}_\ell \tag{9.1}$$

for each $r = 1, \ldots, \ell$.

Recall the ordered index set $I := \{1 < \ldots < M < \overline{1} < \ldots < \overline{N}\}$. For all $i, j \in I$, define

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

where $\mathrm{pa}\,(i) := 0$ if $i \in \{1, \ldots, M\}$ and $\mathrm{pa}\,(i) := 1$ otherwise, as defined in §2.

By calculation one easily shows that

$$\begin{aligned}[]
[\tilde{e}_{i,j}, \tilde{e}_{h,k}] = (\tilde{e}_{i,k} - \delta_{i,k}(-1)^{\mathrm{pa}(i)}\rho_{\mathrm{col}(i)})\delta_{h,j}& \\
- (-1)^{(\mathrm{pa}(i)+\mathrm{pa}(j))(\mathrm{pa}(h)+\mathrm{pa}(k))}\delta_{i,k}(\tilde{e}_{h,j} - \delta_{h,j}(-1)^{\mathrm{pa}(j)}\rho_{\mathrm{col}(j)}).&
\end{aligned} \tag{9.3}$$

The effect of the homomorphism $U(\mathfrak{m}) \to \mathbb{C}$ induced by the character χ can be obtained easily by definition. We explicitly give the result here since it will be frequently used later. For any $i, j \in I$, we have

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

Now we are going to define certain crucial elements in the universal enveloping algebra $U(\mathfrak{gl}_{M|N})$. For $1 \le i, j \le m+n$ and signs $\sigma_i \in \{\pm\}$, we firstly set

$$T^{(0)}_{i,j;\sigma_1,\ldots,\sigma_{m+n}} := \sigma_i \delta_{i,j}$$

and then for $r \geq 1$ we define

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

where the second sum is taken over all $i_1,\dots,i_s,j_1,\dots,j_s \in I$ such that

(1) $\deg(e_{i_1,j_1}) + \cdots + \deg(e_{i_s,j_s}) = r$;
(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 $t = 1,\dots,s-1$.

Due to conditions (1) and (2), $T^{(r)}_{i,j;\sigma_1,\dots,\sigma_{m+n}}$ belongs to $\mathrm{F}_r U(\mathfrak{p})$.

For an integer $0 \leq x \leq m+n$, we set the shorthand notation

$$T^{(r)}_{i,j;x} := T^{(r)}_{i,j;\sigma_1,\dots,\sigma_{m+n}}$$

where

$$\sigma_i = \begin{cases} - & \text{if } i \leq x, \\ + & \text{if } i > x. \end{cases}$$

We further define the following series for all $1 \leq i,j \leq m+n$:

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

The following lemma can be established by exactly the same approach as [8, Lemma 9.2]. We omit the details since the argument there is quite formal and does not depend on the underlying associative superalgebra in which the calculations are performed.

Lemma 9.1. [8, Lemma 9.2] *Let $0 \leq i,j,x,y \leq m+n$ be integers with $x < y$.*

(1) *If $x < i \leq y < j \leq m+n$ then*

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

(2) *If $x < j \leq y < i \leq m+n$ then*

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

(3) *If $x < y < i \leq m+n$ and $y < j \leq m+n$, then*

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

(4) *If $x < i \leq y \leq m+n$ and $x < j \leq y$, then*

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

Define an invertible $(m+n)\times(m+n)$ matrix with entries in $U(\mathfrak{p})[[u^{-1}]]$ by

$$T(u) := \big(T_{i,j;0}(u)\big)_{1\leq i,j\leq m+n}$$

Fix a composition $\mu = (\mu_1, \mu_2, \ldots, \mu_z)$ of $m+n$. Applying the Gauss decomposition of §3, we have

$$T(u) = F(u)D(u)E(u)$$

where $D(u)$ is a diagonal block matrix, $E(u)$ is an upper unitriangular block matrix, and $F(u)$ is a lower unitriangular block matrix, with respect to μ.

The diagonal blocks of $D(u)$ define matrices $D_1(u), \ldots, D_z(u)$, the upper diagonal blocks of $E(u)$ define matrices $E_{1,2}(u), \ldots, E_{z-1,z}(u)$, and the lower diagonal matrices of $F(u)$ define matrices $F_{2,1}(u), \ldots, F_{z,z-1}(u)$, respectively. Set $E_b(u) = E_{b,b+1}(u)$, $F_b(u) = F_{b+1,b}(u)$ for $1 \leq b \leq z-1$ and $D'_a(u) := D_a(u)^{-1}$ for all $1 \leq a \leq z$. The entries of these matrices in turn define the following series:

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

$$E_{b;h,k}(u) = \sum_{r\geq 1} E^{(r)}_{b;h,k} u^{-r}, \qquad F_{b;k,h}(u) = \sum_{r\geq 1} F^{(r)}_{b;k,h} u^{-r},$$

for all $1 \leq a \leq z$, $1 \leq b \leq z-1$, $1 \leq i,j \leq \mu_a$, $1 \leq h \leq \mu_b$, $1 \leq k \leq \mu_{b+1}$.

Nevertheless, all of these elements, depending on the fixed choice of μ, are parallel to the elements in $Y_{m|n}$ with the same notations given in §3, except that the elements defined here belong to $U(\mathfrak{p})$.

Theorem 9.2. [8, Theorem 9.3] *Let $\mu = (\mu_1, \ldots, \mu_z)$ be fixed as above. For any admissible indices a, b, i, j, h, k, we have*

$$\begin{aligned}
D_{a;i,j}(u) &= T_{\mu_1+\cdots+\mu_{a-1}+i,\mu_1+\cdots+\mu_{a-1}+j;\mu_1+\cdots+\mu_{a-1}}(u),\\
D'_{a;i,j}(u) &= -T_{\mu_1+\cdots+\mu_{a-1}+i,\mu_1+\cdots+\mu_{a-1}+j;\mu_1+\cdots+\mu_a}(u),\\
E_{b;h,k}(u) &= T_{\mu_1+\cdots+\mu_{b-1}+h,\mu_1+\cdots+\mu_b+k;\mu_1+\cdots+\mu_b}(u),\\
F_{b;k,h}(u) &= T_{\mu_1+\cdots+\mu_b+k,\mu_1+\cdots+\mu_{b-1}+h;\mu_1+\cdots+\mu_b}(u).
\end{aligned}$$

Proof. Note that it suffices to show the identities for D, E and F, since the one for D' follows from the one for D and Lemma 9.1(4). We prove our statement by induction on the length of μ. The initial case is $\mu = (m+n)$, which is trivial since $T(u) = D_1(u)$.

Now let $\mu = (\mu_1, \ldots, \mu_z)$ be a composition of length $z \geq 2$. Define a new composition $\nu = (\nu_1, \ldots, \nu_{z-1})$ of length $z-1$ by setting $\nu_i = \mu_i$ for all $1 \leq i \leq z-2$ and $\nu_{z-1} = \mu_{z-1} + \mu_z$; that is, merge the last two parts of μ. By the induction hypothesis, we have

$$\begin{aligned}
{}^{\nu}D_a(u) &= \left(T_{\nu_1+\cdots+\nu_{a-1}+i,\nu_1+\cdots+\nu_{a-1}+j;\nu_1+\cdots+\nu_{a-1}}(u)\right)_{1\leq i,j\leq \nu_a}, \ \forall\, 1 \leq a \leq z-1,\\
{}^{\nu}E_b(u) &= \left(T_{\nu_1+\cdots+\nu_{b-1}+h,\nu_1+\cdots+\nu_b+k;\nu_1+\cdots+\nu_b}(u)\right)_{1\leq h\leq \nu_b,1\leq k\leq \nu_{b+1}}, \ \forall\, 1 \leq b \leq z-2,\\
{}^{\nu}F_b(u) &= \left(T_{\nu_1+\cdots+\nu_b+k,\nu_1+\cdots+\nu_{b-1}+h;\nu_1+\cdots+\nu_b}(u)\right)_{1\leq k\leq \nu_{b+1},1\leq h\leq \nu_b}, \ \forall\, 1 \leq b \leq z-2,
\end{aligned}$$

where we add a superscript ν to emphasize that these elements are defined with respect to ν. Note that ${}^{\nu}D_a(u) = {}^{\mu}D_a(u)$ for all $1 \leq a \leq z-2$ and ${}^{\nu}E_b(u) = {}^{\mu}E_b(u)$, ${}^{\nu}F_b(u) = {}^{\mu}F_b(u)$ for all $1 \leq b \leq z-3$.

Moreover, by Lemma 5.11, ${}^{\mu}E_{z-2}(u)$ equals to the submatrix consisting of the first μ_{z-1} columns of ${}^{\nu}E_{z-2}(u)$, while ${}^{\mu}F_{z-2}(u)$ equals to the submatrix consisting of the top μ_{z-1} rows of ${}^{\nu}F_{z-2}(u)$. Both of them are of the form described in the theorem. It remains to check the identities for ${}^{\mu}D_{z-1}(u)$, ${}^{\mu}D_z(u)$, ${}^{\mu}E_{z-1}(u)$ and ${}^{\mu}F_{z-1}(u)$.

Define matrices P, Q, R and S by

$$\begin{aligned}
P &= \left(T_{\mu_1+\cdots+\mu_{z-2}+i,\mu_1+\cdots+\mu_{z-2}+j;\mu_1+\cdots+\mu_{z-2}}(u)\right)_{1\leq i,j\leq \mu_{z-1}},\\
Q &= \left(T_{\mu_1+\cdots+\mu_{z-2}+i,\mu_1+\cdots+\mu_{z-2}+\mu_{z-1}+j;\mu_1+\cdots+\mu_{z-2}+\mu_{z-1}}(u)\right)_{1\leq i\leq \mu_{z-1},1\leq j\leq \mu_z},\\
R &= \left(T_{\mu_1+\cdots+\mu_{z-2}+\mu_{z-1}+i,\mu_1+\cdots+\mu_{z-2}+j;\mu_1+\cdots+\mu_{z-2}+\mu_{z-1}}(u)\right)_{1\leq i\leq \mu_z,1\leq j\leq \mu_{z-1}},\\
S &= \left(T_{\mu_1+\cdots+\mu_{z-2}+\mu_{z-1}+i,\mu_1+\cdots+\mu_{z-2}+\mu_{z-1}+j;\mu_1+\cdots+\mu_{z-2}+\mu_{z-1}}(u)\right)_{1\leq i,j\leq \mu_z}.
\end{aligned}$$

By Lemma 9.1 with $x = \mu_1 + \ldots + \mu_{z-2}$ and $y = \mu_1 + \ldots + \mu_{z-1}$, we have

$${}^{\nu}D_{z-1}(u) = \begin{pmatrix} I_{\mu_{z-1}} & 0 \\ R & I_{\mu_z} \end{pmatrix} \begin{pmatrix} P & 0 \\ 0 & S \end{pmatrix} \begin{pmatrix} I_{\mu_{z-1}} & Q \\ 0 & I_{\mu_z} \end{pmatrix} = \begin{pmatrix} P & PQ \\ RP & S+RPQ \end{pmatrix}.$$

Now the explicit descriptions of the matrices ${}^{\mu}D_{z-1}(u)$, ${}^{\mu}D_z(u)$, ${}^{\mu}E_{z-1}(u)$ and ${}^{\mu}F_{z-1}(u)$ follows from Lemma 5.11, which completes the induction argument. □

In the extreme case that $\mu = (1^{m+n})$, we write simply $D_i^{(r)}, D_i'^{(r)}, E_j^{(r)}$ and $F_j^{(r)}$ for the elements $D_{i;1,1}^{(r)}$, $D_{i;1,1}'^{(r)}$, $E_{j;1,1}^{(r)}$ and $F_{j;1,1}^{(r)}$ of $U(\mathfrak{p})$ for all $1 \leq i \leq m+n$, $1 \leq j \leq m+n-1$, $r \geq 1$, respectively.

Corollary 9.3. [8, Corollary 9.4] *$D_i^{(r)} = T_{i,i;i-1}^{(r)}$, $E_j^{(r)} = T_{j,j+1;j}^{(r)}$, $F_j^{(r)} = T_{j+1,j;j}^{(r)}$ and $D_i'^{(r)} = -T_{i,i;i}^{(r)}$.*

10. Main theorem

Let π be a pyramid associated with a 0^m1^n-sequence Υ which corresponds to a good pair in $\mathfrak{gl}_{M|N}$ and let (σ, ℓ, Υ) be the triple associated to π given by Proposition 2.8. Let $Y_{m|n}^{\ell}(\sigma)$ denote the shifted super Yangian of level ℓ associated to π equipped with the canonical filtration and let $\mathcal{W}_\pi$ denote the finite W-superalgebra associated to π equipped with the Kazhdan filtration .

Suppose also that $\mu = (\mu_1, \ldots, \mu_z)$ is an admissible shape for σ, and recall the shorthand notations $s^\mu_{a,b}$ and p^μ_a from (5.1) and (8.2). We have the elements $D^{(r)}_{a;i,j}$, $D'^{(r)}_{a;i,j}$, $E^{(r)}_{b;h,k}$ and $F^{(r)}_{b;k,h}$ of $U(\mathfrak{p})$ defined by Theorem 9.2 according to this fixed shape μ. On the other hand, we also have the parabolic generators $D^{(r)}_{a;i,j}$, $D'^{(r)}_{a;i,j}$, $E^{(r)}_{b;h,k}$ and $F^{(r)}_{b;k,h}$ in $Y^\ell_\mu(\sigma)$ as defined in § 8. We are ready to present the main result of this article.

Theorem 10.1. *Let π be a pyramid and let (σ, ℓ, Υ) be the corresponding triple given by Proposition 2.8. For any shape $\mu = (\mu_1, \ldots, \mu_z)$ admissible to σ, there exists a unique isomorphism $Y^\ell_\mu(\sigma) \xrightarrow{\sim} \mathcal{W}_\pi$ of filtered superalgebras such that the generators*

$$\{D^{(r)}_{a;i,j} \,|\, 1 \le a \le z, 1 \le i, j \le \mu_a, r > 0\},$$
$$\{E^{(r)}_{b;h,k} \,|\, 1 \le b < z, 1 \le h \le \mu_b, 1 \le k \le \mu_{b+1}, r > s^\mu_{b,b+1}\},$$
$$\{F^{(r)}_{b;k,h} \,|\, 1 \le b < z, 1 \le k \le \mu_{b+1}, 1 \le h \le \mu_b, r > s^\mu_{b+1,b}\}$$

of $Y^\ell_\mu(\sigma)$ are mapped to corresponding elements of $U(\mathfrak{p})$ denoted by the same symbols. In particular, these elements of $U(\mathfrak{p})$ are $\mathfrak{m}$-invariants and they form a generating set for $\mathcal{W}_\pi$.

Similar to the argument in [8], the proof of Theorem 10.1 is processed by induction on the number $\ell - t$, where ℓ is the length of the bottom row and t is the length of the top row of π. Our initial case is $\ell = t$. In this case, the pyramid is of rectangular shape so the associated shift matrix is the zero matrix. Hence the shifted super Yangian is the whole $Y_{m|n}$ itself, and its quotient is exactly the truncated super Yangian $Y^\ell_{m|n}$. As mentioned in §1 , the statement of the theorem in this special case was firstly established in [3]; see also [31] for an approach similar to our setting here.

Assume that our pyramid π is not of rectangular shape so that $\ell \ge 2$ and $\ell - t > 0$. By induction on the length of the shape and Lemma 5.11, it suffices to prove the special case when μ is the minimal admissible shape for σ.

Let H denote the absolute height of the shortest column of π. Since π is a pyramid, either $H = |q_1|$ or $H = |q_\ell|$. There are two cases:

- Case R: $H = |q_\ell| \le |q_1|$.
- Case L: $H = |q_1| < |q_\ell|$.

We will explain the proof of Case R in detail and only sketch the proof of Case L, which can be obtained by a very similar argument with mild modifications.

From now on we assume that Case R holds. Recall that we numbered the boxes of π using the index set

$$I := \{1 < \cdots < M < \overline{1} < \ldots < \overline{N}\}$$

in the standard way: down columns from left to right, where i (respectively, $\bar{i}$) stands for the boxes labeled with $+$ (respectively, $-$). Suppose that there are p (respectively, q) boxes labeled with $+$ (respectively, $-$) in the right-most column of π. Since μ is minimal admissible, we have $H = p + q = \mu_z$.

Let $\dot{\pi}$ be the pyramid obtained by removing the right-most column of π. We know that the removed boxes of π are numbered with

$$M-p+1, M-p+2, \ldots, M, \overline{N-q+1}, \overline{N-q+2}, \ldots, \overline{N},$$

and their order in the right-most column is determined by Υ_z, the last H digits of the 0^m1^n-sequence Υ.

By our assumption, the bottom H rows of π form a rectangle, call it π_H. A key observation [31, Remark 3.5] is that permuting the rows of the rectangle π_H will not change the corresponding even good pair (e_π, h_π); see also Remark 2.5. Although our argument in fact works in general, for convenience, we assume that the last H digits of Υ are the standard one:

$$\Upsilon_z = \overbrace{0\cdots0}^{p}\overbrace{1\cdots1}^{q}.$$

As a result, the right-most two columns of π are of the form

$\vdots$	
$M-2p+1$	$M-p+1$
$M-2p+2$	$M-p+2$
$\vdots$	$\vdots$
$M-p$	M
$\overline{N-2q+1}$	$\overline{N-q+1}$
$\overline{N-2q+2}$	$\overline{N-q+2}$
$\vdots$	$\vdots$
$\overline{N-q}$	$\overline{N}$

Let $\dot{\sigma} = (\dot{s}_{i,j})_{1\le i,j\le m+n}$ be the shift matrix defined by (6.1) where its associated pyramid is $\dot{\pi}$. Define $\dot{\mathfrak{p}}, \dot{\mathfrak{m}}$ and $\dot{e}$ in $\dot{\mathfrak{g}} = \mathfrak{gl}_{M-p|N-q}$ according to (2.1) and (2.4) and let $\dot{\chi} : \dot{\mathfrak{m}} \to \mathbb{C}$ be the character $x \mapsto (x, \dot{e})$.

Let $\dot{D}^{(r)}_{a;i,j}, \dot{D}'^{(r)}_{a;i,j}, \dot{E}^{(r)}_{b;h,k}$ and $\dot{F}^{(r)}_{b;k,h}$ denote the elements of $U(\dot{\mathfrak{p}})$ as defined in §9 associated to the same shape μ, which is admissible for both of σ and $\dot{\sigma}$. By the induction hypothesis, Theorem 10.1 holds for $\dot{\pi}$, so the following elements of $U(\dot{\mathfrak{p}})$ are invariant under the $\dot{\chi}$-twisted action of $\dot{\mathfrak{m}}$; in other words, they belong to the finite W-superalgebra $\mathcal{W}_{\dot{\pi}}$:

$$\{\dot{D}_{a;i,j}^{(r)},\ \dot{D}_{a;i,j}^{\prime(r)}\} \text{ for } 1 \leq a \leq z,\ 1 \leq i,j \leq \mu_a \text{ and } r > 0;$$
$$\{\dot{E}_{b;h,k}^{(r)}\} \text{ for } 1 \leq b \leq z-1,\ 1 \leq h \leq \mu_a,\ 1 \leq k \leq \mu_{a+1} \text{ and } r > s_{b,b+1}^{\mu} - \delta_{b+1,z};$$
$$\{\dot{F}_{b;k,h}^{(r)}\} \text{ for } 1 \leq b \leq z-1,\ 1 \leq k \leq \mu_{a+1},\ 1 \leq h \leq \mu_a \text{ and } r > s_{b+1,b}^{\mu}.$$

We introduce the following non-standard embedding of $U(\dot{\mathfrak{g}})$ into $U(\mathfrak{g})$. For all i,j in the index set

$$\dot{I} := \{1, \ldots, M-p, \overline{1}, \ldots, \overline{N-q}\},$$

the generators $\tilde{e}_{ij}$ of $U(\dot{\mathfrak{g}})$ defined by (9.2) with respect to the pyramid $\dot{\pi}$ are identified with the elements $\tilde{e}_{ij}$ in $U(\mathfrak{g})$ defined by (9.2) with respect to π. One notes that such an identification in turns embeds $U(\dot{\mathfrak{p}})$ into $U(\mathfrak{p})$ and $\dot{\mathfrak{m}}$ into $\mathfrak{m}$, respectively. Moreover, the character $\dot{\chi}$ of $\dot{\mathfrak{m}}$ is precisely the restriction of the character χ of $\mathfrak{m}$. As a consequence, the $\dot{\chi}$-twisted action of $\dot{\mathfrak{m}}$ on $U(\dot{\mathfrak{p}})$ is precisely the restriction of the χ-twist action of $\mathfrak{m}$ on $U(\mathfrak{p})$.

For convenience, we define the index sets

$$\mathsf{J}_1 = \{M-p+i \,|\, 1 \leq i \leq p\} \cup \{\overline{N-q+j} \,|\, 1 \leq j \leq q\},$$

$$\mathsf{J}_2 = \{M-2p+i \,|\, 1 \leq i \leq p\} \cup \{\overline{N-2q+j} \,|\, 1 \leq j \leq q\}.$$

Note that they are the numbers appearing in the right-most and the second right-most columns of the rectangle π_H, respectively.

Define the bijection $R_1 : \{1, 2, \ldots, p+q\} \to \mathsf{J}_1$ by setting $R_1(f)$ to be the number assigned to the f-th box in the right-most column of the rectangle π_H. Similarly, define the bijection $R_2 : \{1, 2, \ldots, p+q\} \to \mathsf{J}_2$ which assigns $R_2(f)$ to be the number appearing to the left of $R_1(f)$. For example, $R_1(1) = M-p+1$, $R_1(p+q) = \overline{N}$ and $R_2(p+q) = \overline{N-q}$. In particular, define

$$\eta : \mathsf{J}_1 \to \{1, 2, \ldots, p+q\} \tag{10.1}$$

to be the inverse map of R_1.

The relations between the elements $D_{a;i,j}^{(r)}$, $E_{b;h,k}^{(r)}$, $F_{b;k,h}^{(r)}$ of $U(\mathfrak{p})$ given by π and the elements $\dot{D}_{a;i,j}^{(r)}$, $\dot{E}_{b;h,k}^{(r)}$, $\dot{F}_{b;k,h}^{(r)}$ of $U(\dot{\mathfrak{p}})$ given by $\dot{\pi}$ are described in the following lemma, which is probably the most crucial step in the proof of our main theorem.

Lemma 10.2. [8, Lemma 10.4] *The following equations hold for all* $1 \leq a \leq z$, $1 \leq b \leq z-1$, $1 \leq i,j \leq \mu_a$, $1 \leq h \leq \mu_b$, $1 \leq k \leq \mu_{b+1}$, *all* $r > 0$ *that makes sense, and any fixed*

$1 \le g \le H$:

$$D^{(r)}_{a;i,j} = \dot{D}^{(r)}_{a;i,j} + \delta_{a,z}\left(\sum_{f=1}^{H}(-1)^{|f|_z}\dot{D}^{(r-1)}_{a;i,f}\tilde{e}_{R_1(f),R_1(j)} + [\dot{D}^{(r-1)}_{a;i,g}, \tilde{e}_{R_2(g),R_1(j)}]\right), \tag{10.2}$$

$$E^{(r)}_{b;h,k} = \dot{E}^{(r)}_{b;h,k} + \delta_{b+1,z}\left(\sum_{f=1}^{H}(-1)^{|f|_z}\dot{E}^{(r-1)}_{b;h,f}\tilde{e}_{R_1(f),R_1(k)} + [\dot{E}^{(r-1)}_{b;h,g}, \tilde{e}_{R_2(g),R_1(k)}]\right), \tag{10.3}$$

$$F^{(r)}_{b;k,h} = \dot{F}^{(r)}_{b;k,h}, \tag{10.4}$$

where for (10.3) we are assuming that $r > s^{\mu}_{z-1,z}$ *if* $b+1 = z$.

Proof. It can be observed from the explicit description of the elements $T^{(r)}_{i,j;x}$ in (9.5) with the help from Theorem 9.2 together with our assumption on the right-most two columns of the rectangle π_H. □

The inductive descriptions provided in Lemma 10.2, together with the induction hypothesis, allow us to deduce the following several lemmas and eventually to show that the elements $D^{(r)}_{a;i,j}$, $E^{(r)}_{b;h,k}$ and $F^{(r)}_{b;k,h}$ of $U(\mathfrak{p})$ are $\mathfrak{m}$-invariants.

Lemma 10.3. [8, Lemma 10.5] *The following elements of* $U(\mathfrak{p})$ *are* $\mathfrak{m}$*-invariant:*

(i) $D^{(r)}_{a;i,j}$ *and* $D'^{(r)}_{a;i,j}$ *for* $1 \le a \le z-1$, $1 \le i,j \le \mu_a$ *and* $r > 0$;
(ii) $E^{(r)}_{b;h,k}$ *for* $1 \le b \le z-2$, $1 \le h \le \mu_b$, $1 \le k \le \mu_{b+1}$ *and* $r > s^{\mu}_{b,b+1}$;
(iii) $F^{(r)}_{b;k,h}$ *for* $1 \le b \le z-1$, $1 \le k \le \mu_{b+1}$, $1 \le h \le \mu_b$ *and* $r > s^{\mu}_{b+1,b}$.

Proof. All of these elements in $U(\mathfrak{p})$ coincide with the elements with the same name in $U(\dot{\mathfrak{p}})$ by Lemma 10.2. Hence they are $\dot{\mathfrak{m}}$-invariant by the induction hypothesis. Define $\dot{\mathfrak{m}}^{\mathfrak{c}}$ to be the vector space complement of $\dot{\mathfrak{m}}$ in $\mathfrak{m}$. It remains to show that these elements are invariant under the χ-twisted action for all $\widetilde{e}_{f,g}$ in $\dot{\mathfrak{m}}^{\mathfrak{c}}$ only. Note that $\widetilde{e}_{f,g} \in \dot{\mathfrak{m}}^{\mathfrak{c}}$ if and only if $g \in \dot{I}$ and $f \in \mathtt{J}_1$.

By Theorem 9.2 and (9.5) again, all elements in the description of the lemma are linear combinations of supermonomials of the form $\tilde{e}_{i_1,j_1}\cdots\tilde{e}_{i_r,j_r}$ in $U(\dot{\mathfrak{p}})$ with $i_s \in \dot{I}$ and $j_s \in \dot{I}\backslash\mathtt{J}_2$ for all $1 \le s \le r$.

By (9.4), $\chi(\tilde{e}_{f,g}) = 0$ for all $g \in \dot{I}\backslash\mathtt{J}_2$ and $f \in \mathtt{J}_1$. This implies that all such supermonomials are invariant under the χ-twisted action of all $\widetilde{e}_{f,g} \in \dot{\mathfrak{m}}^{\mathfrak{c}}$ and our lemma follows. □

It remains to show that $D^{(r)}_{z;i,j}$ and $E^{(r)}_{z-1;h,k}$ are $\mathfrak{m}$-invariant. We first show that they are $\dot{\mathfrak{m}}$-invariant.

Lemma 10.4. [8, Lemma 10.6] *The following elements of* $U(\mathfrak{p})$ *are* $\dot{\mathfrak{m}}$*-invariant:*

(1) $D^{(r)}_{z;i,j}$ *for* $1 \le i,j \le \mu_z$ *and* $r > 0$.

(2) $E^{(r)}_{z-1;h,k}$ *for* $1 \leq h \leq \mu_{z-1}$, $1 \leq k \leq \mu_z$ *and* $r > s^{\mu}_{z-1,z}$.

Proof. (1) By (10.2), we obtain

$$D^{(r)}_{z;i,j} = \dot{D}^{(r)}_{z;i,j} + \sum_{f=1}^{H} (-1)^{|f|_z} \dot{D}^{(r-1)}_{z;i,f} \tilde{e}_{R_1(f),R_1(j)} + [\dot{D}^{(r-1)}_{z;i,g}, \tilde{e}_{R_2(g),R_1(j)}].$$

For any $x \in \dot{\mathfrak{m}}$, we have $[x, \widetilde{e}_{R_1(f),R_1(j)}] = 0 = [x, \widetilde{e}_{R_2(g),R_1(j)}]$. Using this result together with the induction hypothesis, one deduces that $\mathrm{pr}_\chi([x, D^{(r)}_{z;i,j}]) = 0$. The proof of (2) is similar by starting with (10.3). $\square$

Next we show that $D^{(r)}_{z;i,j}$ and $E^{(r)}_{z-1;h,k}$ are $\dot{\mathfrak{m}}^{\mathfrak{c}}$-invariant by induction on r. If $s^{\mu}_{z-1,z} > 1$, this can be deduced in a uniform way, see Lemma 10.8, so we will focus on the case for $s^{\mu}_{z-1,z} = 1$. The following lemma establishes the initial step of the induction.

Lemma 10.5. [8, Lemma 10.7]

(1) $D^{(1)}_{z;i,j}$ *is* $\dot{\mathfrak{m}}^{\mathfrak{c}}$*-invariant for all* $1 \leq i, j \leq \mu_z$.

(2) *Suppose* $s^{\mu}_{z-1,z} = 1$. *Then* $D^{(2)}_{z;i,j}$ *is* $\dot{\mathfrak{m}}^{\mathfrak{c}}$*-invariant for all* $1 \leq i, j \leq \mu_z$.

(3) *Suppose* $s^{\mu}_{z-1,z} = 1$. *Then* $E^{(2)}_{z-1;h,k}$ *is* $\dot{\mathfrak{m}}^{\mathfrak{c}}$*-invariant for all* $1 \leq h \leq \mu_{z-1}$ *and* $1 \leq k \leq \mu_z$.

Proof. We only give the detail of the proof of (1) here, where (2) and (3) can be deduced in a similar fashion.

By Theorem 9.2, (9.5) and (10.2), we have

$$D^{(1)}_{z;i,j} = \dot{D}^{(1)}_{z;i,j} + (-1)^{|i|_z} \tilde{e}_{R_1(i),R_1(j)} = \sum_{1 \leq k \leq \ell-1} \Big(\sum_{p_k,q_k} (-1)^{|i|_z} \tilde{e}_{p_k,q_k} \Big) + (-1)^{|i|_z} \tilde{e}_{R_1(i),R_1(j)},$$

where the second sum is taken over all p_k, $q_k \in \dot{I}$ satisfying the following conditions

(i) $\mathrm{col}(p_k) = \mathrm{col}(q_k) = k$,

(ii) $\mathrm{row}(p_k) = \mu_1 + \cdots + \mu_{z-1} + i$,

(iii) $\mathrm{row}(q_k) = \mu_1 + \cdots + \mu_{z-1} + j$.

Let $\tilde{e}_{f,g} \in \dot{\mathfrak{m}}^{\mathfrak{c}}$ be arbitrary given so that we have $g \in \dot{I}$ and $f \in \mathtt{J}_1$.

Suppose first that $\mathrm{row}(g) \neq \mu_1 + \ldots + \mu_{z-1} + i$. Then we have $[\tilde{e}_{f,g}, \tilde{e}_{p_k,q_k}] = 0$ for any p_k, q_k appearing in the sum. Moreover, $[\tilde{e}_{f,g}, \tilde{e}_{R_1(i),R_1(j)}] = \pm\delta_{f,R_1(j)} \tilde{e}_{R_1(i),g}$, which belongs to the kernel of χ by (9.4). It follows that $\mathrm{pr}_\chi([\tilde{e}_{f,g}, D^{(1)}_{z;i,j}]) = 0$.

Assume now that $\mathrm{row}(g) = \mu_1 + \ldots + \mu_{z-1} + i$. Then g equals exactly one p_k appearing in the sum and hence

$$\big[\tilde{e}_{f,g}, \sum_{1 \leq k \leq \ell-1} \Big(\sum_{p_k,q_k \in \dot{I}} (-1)^{|i|_z} \tilde{e}_{p_k,q_k} \Big)\big] = (-1)^{|i|_z} \tilde{e}_{f,q_k}$$

for a certain $1 \leq k \leq \ell - 1$.

Suppose in addition that $\mathrm{col}(q_k) \neq \ell - 1$. Then $\tilde{e}_{f,q_k}$ belongs to $\ker \chi$ by (9.4). Also, since $g = p_k$ and $\mathrm{col}(q_k) = \mathrm{col}(p_k) \neq \ell - 1$, the term

$$[\tilde{e}_{f,g}, \tilde{e}_{R_1(i),R_1(j)}] = \pm\delta_{f,R_1(j)}\tilde{e}_{R_1(i),g}$$

belongs to $\ker \chi$. Then we have $\mathrm{pr}_\chi[\tilde{e}_{f,g}, D^{(1)}_{z;i,j}] = 0$.

Finally, assume that $\mathrm{row}(g) = \mu_1 + \ldots + \mu_{z-1} + i$ and $\mathrm{col}(q_k) = \ell - 1$. It implies that $g = p_k = R_2(i)$. By definition, we have

$$[\tilde{e}_{f,R_2(i)}, D^{(1)}_{z;i,j}] = (-1)^{|i|_z}\tilde{e}_{f,R_2(j)} + \delta_{f,R_1(j)}(-1)^{1+|j|_z}\tilde{e}_{R_1(i),R_2(i)},$$

which belongs to the kernel of χ by (9.4). This completes the proof of (1). □

To apply induction on r, we need to find the relations between $E^{(r+1)}_{z-1;h,k}$ and $E^{(r)}_{z-1;h,k}$ and that between $D^{(r+1)}_{z;i,j}$ and $D^{(r)}_{z;i,j}$.

Lemma 10.6. [8, Lemma 10.8] *Suppose that $s^\mu_{z-1,z} = 1$. The following identities hold in $U(\mathfrak{p})$ for $r \geq 2$:*

(1)

$$E^{(r+1)}_{z-1;h,k} = (-1)^{|g|_{z-1}}[D^{(2)}_{z-1;h,g}, E^{(r)}_{z-1;g,k}] - \sum_{f=1}^{\mu_{z-1}} D^{(1)}_{z-1;h,f}E^{(r)}_{z-1;f,k}\,,$$

(2)

$$D^{(r+1)}_{z;i,j} = (-1)^{|g|_{z-1}}[F^{(2)}_{z-1;i,g}, E^{(r)}_{z-1;g,j}] - \sum_{t=1}^{r+1} D^{(r+1-t)}_{z;i,j}D'^{(r)}_{z-1;g,g}\,.$$

Proof. By the induction hypothesis and (5.6), for any $r > 0$ and any $1 \leq g \leq \mu_{z-1}$, we have

$$[\dot{D}^{(2)}_{z-1;h,g}, \dot{E}^{(r)}_{z-1;g,k}] = (-1)^{|g|_{z-1}}\dot{E}^{(r+1)}_{z-1;h,k} + (-1)^{|g|_{z-1}}\sum_{p=1}^{\mu_{z-1}}\dot{D}^{(1)}_{z-1;h,p}\dot{E}^{(r)}_{z-1;p,k}. \tag{10.5}$$

Also, (10.3) implies that for $r \geq 2$, we have

$$E^{(r)}_{z-1;g,k} = \dot{E}^{(r)}_{z-1;g,k} + \sum_{f=1}^{H}(-1)^{|f|_z}\dot{E}^{(r-1)}_{z-1;g,f}\tilde{e}_{R_1(f),R_1(k)} + [\dot{E}^{(r-1)}_{z-1;g,j}, \tilde{e}_{R_2(j),R_1(k)}] \tag{10.6}$$

It is clear that $[\dot{D}^{(2)}_{z-1;h,g}, \tilde{e}_{R_1(f),R_1(k)}] = 0$. Also, due to (9.5) and Theorem 9.2, the expansion of $\dot{D}^{(2)}_{z-1;h,g}$ into supermonomials will never involve any matrix unit of the form $\tilde{e}_{?,R_2(j)}$ and it follows that $[\dot{D}^{(2)}_{z-1;h,g}, \tilde{e}_{R_2(j),R_1(k)}] = 0$. Computing the supercommutator of (10.6) with

$D^{(2)}_{z-1;h,g} = \dot{D}^{(2)}_{z-1;h,g}$ and using (10.5), we have

$$\begin{aligned}
[D^{(2)}_{z-1;h,g}, E^{(r)}_{z-1;g,k}] &= [\dot{D}^{(2)}_{z-1;h,g}, \dot{E}^{(r)}_{z-1;g,k}] + \sum_{f=1}^{H}(-1)^{|f|_z}[\dot{D}^{(2)}_{z-1;h,g}, \dot{E}^{(r-1)}_{z-1;g,f}]\tilde{e}_{R_1(f),R_1(k)} \\
&+ \big[[D^{(2)}_{z-1;h,g}, E^{(r-1)}_{z-1;g,j}], \tilde{e}_{R_2(j),R_1(k)}\big] \\
&= (-1)^{|g|_{z-1}}\dot{E}^{(r+1)}_{z-1;h,k} + (-1)^{|g|_{z-1}}\sum_{p=1}^{\mu_{z-1}}\dot{D}^{(1)}_{z-1;h,p}\dot{E}^{(r)}_{z-1;p,k} \\
&+ \sum_{f=1}^{H}(-1)^{|f|_z}\Big((-1)^{|g|_{z-1}}\dot{E}^{(r+1)}_{z-1;h,f} + (-1)^{|g|_{z-1}}\sum_{p=1}^{\mu_{z-1}}\dot{D}^{(1)}_{z-1;h,p}\dot{E}^{(r)}_{z-1;p,f}\Big)\tilde{e}_{R_1(f),R_1(k)} \\
&+ \Big[(-1)^{|g|_{z-1}}\dot{E}^{(r+1)}_{z-1;h,j} + (-1)^{|g|_{z-1}}\sum_{p=1}^{\mu_{z-1}}\dot{D}^{(1)}_{z-1;h,p}\dot{E}^{(r)}_{z-1;p,j}, \tilde{e}_{R_2(j),R_1(k)}\Big].
\end{aligned}$$

Using (10.3) a few times, one shows that the above equals to

$$(-1)^{|g|_{z-1}}\big(E^{(r+1)}_{z-1;h,k} + \sum_{p=1}^{\mu_{z-1}} D^{(1)}_{z-1;h,p}E^{(r)}_{z-1;p,k}\big)$$

and the equality (1) is established.

Now we deal with (2). By the induction hypothesis and (5.8), we have

$$\begin{aligned}
[\dot{F}^{(2)}_{z-1;i,g}, \dot{E}^{(r)}_{z-1;g,j}] &= (-1)^{|g|_{z-1}}(\sum_{t=0}^{r+1}\dot{D}^{(r+1-t)}_{z;i,j}\dot{D}'^{(t)}_{z-1;g,g}) \\
&= (-1)^{|g|_{z-1}}\dot{D}^{(r+1)}_{z;i,j} + (-1)^{|g|_{z-1}}\sum_{t=1}^{r+1}\dot{D}^{(r+1-t)}_{z;i,j}\dot{D}'^{(t)}_{z-1;g,g}.
\end{aligned} \tag{10.7}$$

Changing the indices in equation (10.6), we have

$$E^{(r)}_{z-1;g,j} = \dot{E}^{(r)}_{z-1;g,j} + \sum_{f=1}^{H}(-1)^{|f|_z}\dot{E}^{(r-1)}_{z-1;g,f}\tilde{e}_{R_1(f),R_1(j)} + [\dot{E}^{(r-1)}_{z-1;g,h}, \tilde{e}_{R_2(h),R_1(j)}] \tag{10.8}$$

Note that the expansion of $\dot{F}^{(2)}_{z-1;i,g}$ into supermonomials will never involve any matrix unit of the forms $\tilde{e}_{?,R_1(h)}$, $\tilde{e}_{R_1(h),?}$ or $\tilde{e}_{R_2(h),?}$, and hence $[\dot{F}^{(2)}_{z-1;i,g}, \tilde{e}_{R_1(f),R_1(j)}] = [\dot{F}^{(2)}_{z-1;i,g}, \tilde{e}_{R_2(h),R_1(j)}] = 0$. As a consequence, we perform the following calculation using the fact that $F^{(2)}_{z-1;i,g} =$

$\dot{F}^{(2)}_{z-1;i,g}$ together with (10.8):

$$\begin{aligned}
[F^{(2)}_{z-1;i,g}, E^{(r)}_{z-1;g,j}] &= [\dot{F}^{(2)}_{z-1;i,g}, \dot{E}^{(r)}_{z-1;g,j}] + \sum_{f=1}^{H}(-1)^{|f|_z}[\dot{F}^{(2)}_{z-1;i,g}, \dot{E}^{(r-1)}_{z-1;g,f}]\tilde{e}_{R_1(f),R_1(j)} \\
&\quad + \big[[\dot{F}^{(2)}_{z-1;i,g}, \dot{E}^{(r-1)}_{z-1;g,h}], \tilde{e}_{R_2(h),R_1(j)}\big] \\
&= (-1)^{|g|_{z-1}}\dot{D}^{(r+1)}_{z;i,j} + (-1)^{|g|_{z-1}}\sum_{t=1}^{r+1}\dot{D}^{(r+1-t)}_{z;i,j}\dot{D}'^{(t)}_{z-1;g,g} \\
&\quad + \sum_{f=1}^{H}(-1)^{|f|_z}\Big((-1)^{|g|_{z-1}}\dot{D}^{(r)}_{z;i,f} + (-1)^{|g|_{z-1}}\sum_{t=1}^{r}\dot{D}^{(r-t)}_{z;i,f}\dot{D}'^{(t)}_{z-1;g,g}\Big)\tilde{e}_{R_1(f),R_1(j)} \\
&\quad + \Big[(-1)^{|g|_{z-1}}\dot{D}^{(r)}_{z;i,h} + (-1)^{|g|_{z-1}}\sum_{t=1}^{r}\dot{D}^{(r-t)}_{z;i,h}\dot{D}'^{(t)}_{z-1;g,g}, \tilde{e}_{R_2(h),R_1(j)}\Big]
\end{aligned}$$

Using (10.2) a few times, the above can be rewritten as

$$(-1)^{|g|_{z-1}}D^{(r+1)}_{z;i,j} + (-1)^{|g|_{z-1}}\sum_{t=1}^{r+1}D^{(r+1-t)}_{z;i,j}\dot{D}'^{(t)}_{z-1;g,g}$$

and our assertion (2) follows. □

Lemma 10.7. *Suppose $s^{\mu}_{z-1,z} = 1$. Then*

(1) $D^{(r)}_{z;i,j}$ *are $\mathfrak{m}$-invariant for all $r \geq 0$ and $1 \leq i, j \leq \mu_z$.*
(2) $E^{(r)}_{z-1;h,k}$ *are $\mathfrak{m}$-invariant for all $r > 1$ and $1 \leq h \leq \mu_{z-1}$, $1 \leq k \leq \mu_z$.*

Proof. By Lemma 10.4, these elements are $\dot{\mathfrak{m}}$-invariant. It remains to check that they are $\dot{\mathfrak{m}}^{\mathfrak{c}}$-invariant, but that follows from Lemma 10.5, Lemma 10.6 and induction on r. □

Lemma 10.8. [8, Lemma 10.9] *Suppose that $s^{\mu}_{z-1,z} > 1$. Then the following elements are invariant under the χ-twisted action of $\tilde{e}_{R_1(x),R_2(y)}$ for all $1 \leq x, y \leq H$.*

(1) $D^{(r)}_{z;i,j}$ *for all $r \geq 2$ and $1 \leq i, j \leq \mu_z$.*
(2) $E^{(r)}_{z-1;h,k}$ *for all $r > s^{\mu}_{z-1,z}$ and $1 \leq h \leq \mu_{z-1}$, $1 \leq k \leq \mu_z$.*

Proof. Let $\ddot{\pi}$ be the pyramid obtained by deleting the right-most two columns of π. Define $\ddot{\mathfrak{p}}$, $\ddot{\mathfrak{m}}$ and $\ddot{e} \in \mathfrak{gl}_{M-2p|N-2q}$ as before, and embed $U(\ddot{\mathfrak{g}})$ into $U(\dot{\mathfrak{g}})$ as how we embed $U(\dot{\mathfrak{g}})$ into $U(\mathfrak{g})$. Note that the assumption $s^{\mu}_{z-1,z} > 1$ implies that $|q_{\ell-1}| = |q_\ell| \leq |q_1|$. As a result, the induction hypothesis applies to the pyramid $\ddot{\pi}$ and hence we know that the elements $\ddot{D}^{(r)}_{z;i,j}$ in $\mathcal{W}_{\ddot{\pi}}$ are $\ddot{\mathfrak{m}}$-invariant under the $\dot{\chi}$-twisted action.

Applying Lemma 10.2 to π and $\dot{\pi}$, we have

$$D^{(r)}_{z;i,j} = \dot{D}^{(r)}_{z;i,j} + \sum_{f=1}^{H}(-1)^{|f|_z}\dot{D}^{(r-1)}_{z;i,f}\tilde{e}_{R_1(f),R_1(j)} + [\dot{D}^{(r-1)}_{z;i,g}, \tilde{e}_{R_2(g),R_1(j)}] \tag{10.9}$$

and

$$\dot{D}^{(r)}_{z;i,j} = \ddot{D}^{(r)}_{z;i,j} + \sum_{f=1}^{H} (-1)^{|f|_z} \ddot{D}^{(r-1)}_{z;i,f} \tilde{e}_{R_2(f),R_2(j)} + [\ddot{D}^{(r-1)}_{z;i,g}, \tilde{e}_{R_3(g),R_2(j)}] \tag{10.10}$$

where $R_3(g)$ is defined to be the number assigned to g-th box in the third right-most column of the rectangle π_H.

Substituting (10.10) into (10.9) and simplifying the result by (9.3), one deduces that for all $r \geq 2$, $D^{(r)}_{z;i,j} = A + B + C + D + E + F + G + Y$, where

$$A = \ddot{D}^{(r)}_{z;i,j}, \qquad B = \sum_{k=1}^{H} (-1)^{|k|_z} \ddot{D}^{(r-1)}_{z;i,k} \tilde{e}_{R_2(k),R_2(j)},$$

$$C = [\ddot{D}^{(r-1)}_{z;i,g}, \tilde{e}_{R_3(g),R_2(j)}], \qquad D = \sum_{k=1}^{H} (-1)^{|k|_z} \ddot{D}^{(r-1)}_{z;i,k} \tilde{e}_{R_1(k),R_1(j)}$$

$$E = \sum_{h,k=1}^{H} (-1)^{|h|_z+|k|_z} \ddot{D}^{(r-2)}_{z;i,h} \tilde{e}_{R_2(h),R_2(k)} \tilde{e}_{R_1(k),R_1(j)}, \qquad F = \sum_{k=1}^{H} (-1)^{|k|_z} \ddot{D}^{(r-2)}_{z;i,k} \tilde{e}_{R_2(k),R_1(j)},$$

$$G = \sum_{k=1}^{H} [\ddot{D}^{(r-2)}_{z;i,g}, \tilde{e}_{R_3(g),R_2(k)}] \tilde{e}_{R_1(k),R_1(j)}, \qquad Y = [\ddot{D}^{(r-2)}_{z;i,g}, \tilde{e}_{R_3(g),R_1(j)}].$$

Let $X = \tilde{e}_{R_1(x),R_2(y)}$ for some $1 \leq x, y \leq H$. Note that X supercommutes with all elements in $U(\ddot{\mathfrak{p}})$. Using (9.1), (9.3) and (9.4), we can explicitly compute their images under the composition $\mathrm{pr}_\chi \circ \mathrm{ad} X$ as follows:

$$\mathrm{pr}_\chi([X, A]) = 0, \qquad \mathrm{pr}_\chi([X, B]) = \delta_{xj} (-1)^{1+|x|_z+|y|_z} \ddot{D}^{(r-1)}_{z;i,y},$$

$$\mathrm{pr}_\chi([X, C]) = 0, \qquad \mathrm{pr}_\chi([X, D]) = \delta_{xj} (-1)^{|x|_z+|y|_z} \ddot{D}^{(r-1)}_{z;i,y},$$

$$\begin{aligned}
\mathrm{pr}_\chi([X, E]) = (-1)^{|x|_z+|y|_z} \delta_{xj} (p-q) \ddot{D}^{(r-2)}_{z;i,y} &+ (-1)^{|y|_z+1} \ddot{D}^{(r-2)}_{z;i,y} \tilde{e}_{R_1(x),R_1(j)} \\
&+ \delta_{xj} \sum_{k=1}^{H} (-1)^{(|x|_z+|y|_z)(|k|_z+|j|_z)+|k|_z} \ddot{D}^{(r-2)}_{z;i,k} \tilde{e}_{R_2(k),R_2(y)},
\end{aligned}$$

$$\begin{aligned}
\mathrm{pr}_\chi([X, F]) = -(-1)^{|x|_z+|y|_z} \delta_{xj} (p-q) \ddot{D}^{(r-2)}_{z;i,y} &+ (-1)^{|y|_z} \ddot{D}^{(r-2)}_{z;i,y} \tilde{e}_{R_1(x),R_1(j)} \\
&- \delta_{xj} \sum_{k=1}^{H} (-1)^{(|x|_z+|y|_z)(|k|_z+|j|_z)+|k|_z} \ddot{D}^{(r-2)}_{z;i,k} \tilde{e}_{R_2(k),R_2(y)},
\end{aligned}$$

$$\mathrm{pr}_\chi([X, G]) = (-1)^{|y|_z+|j|_z} \delta_{xj} [\ddot{D}^{(r-2)}_{z;i,g_1}, \tilde{e}_{R_3(g_1),R_2(f)}],$$

$$\mathrm{pr}_\chi([X, Y]) = -(-1)^{|y|_z+|j|_z} \delta_{xj} [\ddot{D}^{(r-2)}_{z;i,g_1}, \tilde{e}_{R_3(g_1),R_2(f)}].$$

As a consequence, $\mathrm{pr}_\chi([X, D^{(r)}_{z;i,j}]) = 0$. The proof of (2) is similar. $\square$

Proposition 10.9. [8, Lemma 10.10] *The following elements of* $U(\mathfrak{p})$ *are* $\mathfrak{m}$*-invariant under the* χ*-twisted action:*

$$\{D^{(r)}_{a;i,j}\}_{1\le a\le z, 1\le i,j\le \mu_a, r>0},$$
$$\{E^{(r)}_{b;h,k}\}_{1\le b<z, 1\le h\le \mu_a, 1\le k\le \mu_{a+1}, r>s^{\mu}_{a,b}},$$
$$\{F^{(r)}_{b;k,h}\}_{1\le b<z, 1\le k\le \mu_{a+1}, 1\le h\le \mu_a, r>s^{\mu}_{b,a}}.$$

Proof. It follows from the induction hypothesis and Lemma 10.3–Lemma 10.8. □

A consequence of Proposition 10.9 is that the elements in the description of Theorem 10.1 are actually elements of $\mathcal{W}_\pi$. Furthermore, by the induction hypothesis, we may identify $Y^{\ell-1}_{\mu}(\dot\sigma) = Y^{\ell-1}_{m|n}(\dot\sigma)$ with $\mathcal{W}_{\dot\pi} \subseteq U(\dot{\mathfrak{p}})$, where the generators $\dot D^{(r)}_{a:i,j}$, $\dot E^{(r)}_{b;h,k}$ and $\dot F^{(r)}_{b;k,h}$ in $Y^{\ell-1}_{\mu}(\dot\sigma)$ are identified with the elements of $\mathcal{W}_{\dot\pi}$ denoted by the same notations. Now we are going to make use of the monomorphism $\Delta_R : Y^{\ell}_{m|n}(\sigma) \to U(\dot{\mathfrak{p}}) \otimes U(\mathfrak{gl}_{p|q})$ obtained in Theorem 8.2.

By Corollary 8.4, for each $d \ge 0$, we have

$$\dim \Delta_R(F_d Y^{\ell}_{m|n}(\sigma)) = \dim F_d Y^{\ell}_{m|n}(\sigma) = \dim F_d S(\mathfrak{g}^e), \tag{10.11}$$

where $F_d S(\mathfrak{g}^e)$ is the sum of all graded elements in $S(\mathfrak{g}^e)$ of degree $\le d$ with respect to the Kazhdan grading.

Define the higher root elements $E^{(r)}_{a,b;i,j}$ and $F^{(r)}_{b,a;j,i}$ in $F_r U(\mathfrak{p})$ by equations (5.19) and (5.20) recursively, where the index k could be chosen arbitrarily there. Let X_d denote the subspace of $U(\mathfrak{p})$ spanned by all supermonomials in the elements

$$\{D^{(r)}_{a;i,j}\}_{1\le a\le z, 1\le i,j\le \mu_a, 0\le r\le s^{\mu}_{a,a}},$$
$$\{E^{(r)}_{a,b;h,k}\}_{1\le a<b\le z, 1\le h\le \mu_a, 1\le k\le \mu_b, s^{\mu}_{a,b}<r\le s^{\mu}_{a,b}+p^{\mu}_a},$$
$$\{F^{(r)}_{b,a;k,h}\}_{1\le a<b\le z, 1\le k\le \mu_b, 1\le h\le \mu_a, s^{\mu}_{b,a}<r\le s^{\mu}_{b,a}+p^{\mu}_a}.$$

taken in some fixed order with total degree $\le d$. It follows from Proposition 10.9 that X_d is a subspace of $F_d\mathcal{W}_\pi$.

Define a superalgebra homomorphism $\psi_R : U(\mathfrak{p}) \to U(\dot{\mathfrak{p}}) \otimes U(\mathfrak{gl}_{p|q})$ by

$$\psi_R(\tilde e_{i,j}) := \begin{cases} \tilde e_{i,j} \otimes 1 & \text{if } \mathrm{col}(i) \le \mathrm{col}(j) \le \ell - 1, \\ 0 & \text{if } \mathrm{col}(i) \le \ell - 1, \mathrm{col}(j) = \ell, \\ 1 \otimes \tilde e_{\eta(i),\eta(j)} & \text{if } \mathrm{col}(i) = \mathrm{col}(j) = \ell, \end{cases}$$

where the map η is defined in (10.1). By Lemma 10.2, we have

$$\psi_R(D_{a;i,j}^{(r)}) = \dot{D}_{a;i,j}^{(r)} \otimes 1 + \delta_{a,z} \sum_{f=1}^{H} (-1)^{|f|_z} \dot{D}_{a;i,f}^{(r-1)} \otimes \tilde{e}_{f,j},$$
$$\psi_R(E_{b;h,k}^{(r)}) = \dot{E}_{b;h,k}^{(r)} \otimes 1 + \delta_{b+1,z} \sum_{f=1}^{H} (-1)^{|f|_z} \dot{E}_{b;h,f}^{(r-1)} \otimes \tilde{e}_{f,k},$$
$$\psi_R(F_{b;k,h}^{(r)}) = \dot{F}_{b;k,h}^{(r)} \otimes 1.$$

Comparing this with Theorem 6.1(1) and recalling the PBW basis for $Y_{m|n}^{\ell}(\sigma)$ obtained in Corollary 8.3, we deduce that $\psi_R(X_d) = \Delta_R(F_d Y_{m|n}^{\ell}(\sigma))$. Combining this with (10.11) and Corollary 8.4, we obtain

$$\dim F_d S(\mathfrak{g}^e) = \dim \psi_R(X_d) \leq \dim X_d \leq \dim F_d \mathcal{W}_\pi \leq \dim F_d S(\mathfrak{g}^e).$$

Hence equalities hold everywhere so we have $X_d = F_d \mathcal{W}_\pi$ for each $d \geq 0$. In particular, $\psi_R : \mathcal{W}_\pi \to U(\dot{\mathfrak{p}}) \otimes \mathfrak{gl}_{p|q}$ is an injective homomorphism. Comparing ψ_R with the map Δ_R defined in Theorem 6.1(1), we see that $\psi_R(D_{a;i,j}^{(r)}) = \Delta_R(D_{a;i,j}^{(r)})$, where the elements $D_{a;i,j}^{(r)}$ on the left-hand side are the elements of $\mathcal{W}_\pi$ and the elements $D_{a;i,j}^{(r)}$ on the right-hand side are the generators of $Y_{m|n}^{\ell}(\sigma)$. Similarly, $\psi_R(E_{b;h,k}^{(r)}) = \Delta_R(E_{b;h,k}^{(r)})$ and $\psi_R(F_{b;k,h}^{(r)}) = \Delta_R(F_{b;k,h}^{(r)})$ for all admissible indices b, h, k, r.

Finally, the composition $\psi_R^{-1} \circ \Delta_R : Y_{m|n}^{\ell}(\sigma) \to \mathcal{W}_\pi$ is exactly the filtered superalgebra isomorphism described in Theorem 10.1 and the elements listed in Theorem 10.1 indeed generate $\mathcal{W}_\pi$. This completes the induction step of our main theorem under the assumption of Case R.

Next we sketch how to complete the induction step under the assumption of Case L. In this case, we enumerate the bricks of π down columns *from right to left*. Note that different ways of enumerating are just choosing different bases to describe $\mathfrak{gl}_{M|N} \cong \mathrm{End}(\mathbb{C}^{M|N})$ so we may choose the way most suitable for our purpose.

Let $\dot{\pi}$ denote the pyramid obtained from π by deleting the *left-most* column of π. Let I, $\dot{I}$, $\mathtt{J}_1$ and $\mathtt{J}_2$ be the same index sets as defined in Case R. It is clear that the deleted bricks are still numbered with elements in $\mathtt{J}_1$. Moreover, we may again assume that the left-most two columns of π are of the form

	$\vdots$
$M-p+1$	$M-2p+1$
$M-p+2$	$M-2p+2$
$\vdots$	$\vdots$
M	$M-p$
$\overline{N-q+1}$	$\overline{N-2q+1}$
$\overline{N-q+2}$	$\overline{N-2q+2}$
$\vdots$	$\vdots$
$\overline{N}$	$\overline{N-q}$

Similarly, we define the bijection $L_1 : \{1,2,\ldots,p+q\} \to \mathtt{J}_1$ by setting $L_1(f)$ to be the number assigned to the f-th box in the *left-most* column of the rectangle π_H, and define the bijection $L_2 : \{1,2,\ldots,p+q\} \to \mathtt{J}_2$ by assigning $L_2(f)$ to be the number appearing to the *right* of $L_1(f)$. In particular, denote by

$$\xi : \mathtt{J}_1 \to \{1,2,\ldots,p+q\} \tag{10.12}$$

the inverse map of L_1.

Let $\dot{\sigma}$ be the shift matrix obtained from (6.2), where the corresponding pyramid is exactly $\dot{\pi}$, and define $\dot{\mathfrak{p}}, \dot{\mathfrak{m}}, \dot{e} \in \dot{\mathfrak{g}} := \mathfrak{gl}_{M-p|N-q}$ via (2.1) and (2.4) with respect to $\dot{\pi}$. Different from Case R, note that in Case L we embed $U(\dot{\mathfrak{g}})$ into $U(\mathfrak{g})$ by the *natural embedding*, which already sends the elements $\tilde{e}_{ij}$ of $U(\dot{\mathfrak{g}})$ to the elements $\tilde{e}_{ij}$ of $U(\mathfrak{g})$ for all $i,j \in \dot{I}$.

Under the natural embedding, the superalgebra $\mathcal{W}_{\dot{\pi}} = U(\dot{\mathfrak{p}})^{\dot{\mathfrak{m}}}$ is a subalgebra of $U(\dot{\mathfrak{p}}) \subset U(\mathfrak{p})$ and the $\dot{\chi}$-twisted action of $\dot{\mathfrak{m}}$ on $U(\dot{\mathfrak{p}})$ is exactly the same with the restriction of the χ-twisted action of $\mathfrak{m}$ on $U(\mathfrak{p})$. Let $\dot{D}^{(r)}_{a;i,j}, \dot{D}'^{(r)}_{a;i,j}, \dot{E}^{(r)}_{b;h,k}$ and $\dot{F}^{(r)}_{b;k,h}$ denote the elements of $U(\dot{\mathfrak{p}})$ as defined in §9 associated to the shape μ which is the minimal admissible shape of σ and also admissible for $\dot{\sigma}$. By the induction hypothesis, all of these elements are $\dot{\mathfrak{m}}$-invariant.

From now we follow exactly the same idea in Case R to complete the proof. By the following crucial lemma, which is the analogue of Lemma 10.2, we may express the elements $D^{(r)}_{a;i,j}, D'^{(r)}_{a;i,j}, E^{(r)}_{b;h,k}$ and $F^{(r)}_{b;k,h}$ in $U(\mathfrak{p})$ in terms of $\dot{D}^{(r)}_{a;i,j}, \dot{D}'^{(r)}_{a;i,j}, \dot{E}^{(r)}_{b;h,k}$ and $\dot{F}^{(r)}_{b;k,h}$. Then by similar case-by-case discussions and computations as before, we can prove that all of the elements $D^{(r)}_{a;i,j}, D'^{(r)}_{a;i,j}, E^{(r)}_{b;h,k}$ and $F^{(r)}_{b;k,h}$ are indeed $\mathfrak{m}$-invariant under our current setting in Case L. We provide only the most crucial lemma below since its proof and other arguments are almost identical as in the earlier case.

Lemma 10.10. [8, Lemma 10.11] *The following identities hold for all admissible a, b, i, j, h, k, r and any fixed $1 \leq g \leq H$:*

$$D_{a;i,j}^{(r)} = \dot{D}_{a;i,j}^{(r)} + \delta_{a,z}(-1)^{|i|_z}\left(\sum_{f=1}^{H} \tilde{e}_{L_1(i),L_1(f)}\dot{D}_{z;f,j}^{(r-1)} + [\tilde{e}_{L_1(i),L_2(g)}, \dot{D}_{z;g,j}^{(r-1)}]\right), \tag{10.13}$$

$$E_{b;h,k}^{(r)} = \dot{E}_{b;h,k}^{(r)}, \tag{10.14}$$

$$F_{b;k,h}^{(r)} = \dot{F}_{b;k,h}^{(r)} + \delta_{b,z-1}(-1)^{|k|_z}\left(\sum_{f=1}^{H} \tilde{e}_{L_1(k),L_1(f)}\dot{F}_{z-1;f,h}^{(r-1)} + [\tilde{e}_{L_1(k),L_2(g)}, \dot{F}_{z-1;g,h}^{(r-1)}]\right), \tag{10.15}$$

where for (10.15) we are assuming that $r > s_{z,z-1}^{\mu}$ if $b = z-1$.

With the help of Lemma 10.10, one can deduce that the statement of Proposition 10.9 still holds in Case L. Finally, define a superalgebra homomorphism $\psi_L : U(\mathfrak{p}) \to U(\mathfrak{gl}_{p|q}) \otimes U(\dot{\mathfrak{p}})$ by

$$\psi_L(\tilde{e}_{i,j}) := \begin{cases} \tilde{e}_{\xi(i),\xi(j)} \otimes 1 & \text{if } \mathrm{col}(i) = \mathrm{col}(j) = 1, \\ 0 & \text{if } \mathrm{col}(i) = 1, \mathrm{col}(j) \geq 2, \\ 1 \otimes \tilde{e}_{i,j} & \text{if } 2 \leq \mathrm{col}(i) \leq \mathrm{col}(j), \end{cases}$$

where the function ξ is defined by (10.12). Using Lemma 10.10 again, we have that

$$\psi_L(D_{a;i,j}^{(r)}) = 1 \otimes \dot{D}_{a;i,j}^{(r)} + \delta_{a,z}\sum_{f=1}^{H}(-1)^{|f|_z}\tilde{e}_{i,f} \otimes \dot{D}_{a;f,j}^{(r-1)}$$

$$\psi_L(E_{b;h,k}^{(r)}) = 1 \otimes \dot{E}_{b;h,k}^{(r)},$$

$$\psi_L(F_{b,k,h}^{(r)}) = 1 \otimes \dot{F}_{b;k,h}^{(r)} + \delta_{b+1,z}\sum_{f=1}^{H}(-1)^{|f|_z}\tilde{e}_{k,f} \otimes \dot{F}_{b;f,h}^{(r-1)}.$$

Using exactly the same argument as in Case R, one shows that the map ψ_L is injective and the composition $\psi_L^{-1} \circ \Delta_L : Y_{m|n}^{\ell}(\sigma) \to \mathcal{W}_\pi$ gives the required isomorphism of filtered superalgebras. This completes the proof of Theorem 10.1.

Corollary 10.11. *Let π be a pyramid corresponding to an even good pair and $\vec{\pi}$ be a pyramid obtained by horizontally shifting rows of π. Let $\mathcal{W}_\pi$ and $\mathcal{W}_{\vec{\pi}}$ denote the associated finite W-superalgebras, respectively. Then there exists a superalgebra isomorphism $\iota : \mathcal{W}_\pi \to \mathcal{W}_{\vec{\pi}}$ defined on parabolic generators with respect to an admissible shape μ by (5.23). In other words, the definition of a finite W-superalgebra associated to an even good pair depends only on e up to isomorphism.*

Proof. This is an immediate consequence of (8.4) and the isomorphism in Theorem 10.1. □

Remark 10.12. *A more general result of Corollary 10.11 was obtained in [47] by a very different approach. It is proved that the definition of type A finite W-superalgebra is independent of the choices of the good $\mathbb{Z}$-grading (which may not be even) up to isomorphism, generalizing the results of [5, 20].*

References

[1] T. Arakawa, Introduction to W-algebras and their representation theory, *Perspectives in Lie Theory*, Springer INdAM Series **19** (2017), 179–250.

[2] A. Braverman, M. Finkelberg and H. Nakajima, Coulomb branches of $3d$ $\mathcal{N} = 4$ quiver gauge theories and slices in the affine Grassmannian (with appendices by A. Braverman, M. Finkelberg, J. Kamnitzer, R. Kodera, H. Nakajima, B. Webster, A. Weekes), *Adv. Theor. Math. Phys.* **23** (2019), no. 1, 75–166.

[3] C. Briot and E. Ragoucy, $\mathcal{W}$-superalgebras as truncations of super-Yangians, *J. Phys. A* **36** (2003), 1057-1081.

[4] J. Brown, J. Brundan and S. Goodwin, Principal W-algebras for $\mathrm{GL}(m|n)$, *Algebra and Number Theory.* **7** (2013), 1849-1882.

[5] J. Brundan and S. Goodwin, Good grading polytopes, *Proc. Lond. Math. Soc.* (3) **94**, No. 1, (2007), 155-180.

[6] J. Brundan, S. Goodwin and A. Kleshchev, Highest weight theory for finite W-algebras, *Int. Math. Res. Not. IMRN* **2008** (2008), No. 15, Art. ID rnn051.

[7] J. Brundan and A. Kleshchev, Parabolic Presentations of the Yangian $Y(\mathfrak{gl}_n)$, *Comm. Math. Phys.* **254** (2005), 191-220.

[8] J. Brundan and A. Kleshchev, Shifted Yangians and finite W-algebras, *Adv. Math.* **200** (2006), 136-195.

[9] J. Brundan and A. Kleshchev, Representations of Shifted Yangians and Finite W-algebras, *Mem. Amer. Math. Soc.* **196** (2008).

[10] J. Brundan and A. Kleshchev, Schur-Weyl duality for higher levels, *Sel. Math.* **14** (2008), 1-57.

[11] S.-J. Cheng and W. Wang, *Dualities and Representations of Lie Superalgebras*, Graduate Studies in Mathematics **144**, AMS, 2013.

[12] I. Cherednik, A new interpretation of Gelfand-Tzetlin bases, *Duke Math. J.* **54** (1987), 563-577.

[13] I. Cherednik, Quantum groups as hidden symmetries of classic representation theory, *Differential Geometric Methods in Theoretical Physics* (Chester, 1988), World Sci. Publishing, Teaneck, NJ, 1989, pp. 47-54.

[14] V. Drinfeld, Hopf algebras and the quantum Yang-Baxter equation, *Soviet Math. Dokl.* **32** (1985), 254-258.

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

[16] A. Elashvili and V. Kac, Classification of good gradings of simple Lie algebras, Lie groups and invariant theory (E. B. Vinberg ed.), *Amer. Math. Socl. Transl.* **213** (2005), 85–104.

[17] M. Finkelberg, J. Kamnitzer, K. Pham, L. Rybnikov and A. Weekes, Comultiplication for shifted Yangians and quantum open Toda lattice, *Adv. Math.* **327** (2018), 349–389.

[18] R. Frassek, V. Pestun and A. Tsymbaliuk, Lax matrices from antidominantly shifted Yangians and quantum affine algebras, preprint, arXiv:2001.04929

[19] L. Frappat, A. Sciarrino and P. Sorba, Structure of basic Lie superalgebras and of their affine extensions. *Comm. Math. Phys.* **121** (1989), 457–500.

[20] W. Gan and V. Ginzburg, Quantization of Slodowy slices, *Int. Math. Res. Not. IMRN* **5** (2002), 243–255.

[21] I. Gelfand and V. Retakh, Quasideterminants, I, *Selecta Math.* **3** (1997), 517–546.

[22] L. Gow, Gauss Decomposition of the Yangian $Y(\mathfrak{gl}_{m|n})$, *Comm. Math. Phys.* **276** (2007), 799-825.

[23] C. Hoyt, Good gradings of basic Lie superalgebras, *Israel J. Math.* **192** (2012), 251-280.

[24] J. Kamnitzer, B. Webster, A. Weekes and O. Yacobi, Yangians and quantizations of slices in the affine Grassmannian, *Algebra Number Theory* **8** (2014), no. 4, 857–893.

[25] B. Kostant, On Whittaker vectors and representation theory, *Invent. Math.* **48** (1978), 101-184.

[26] I. Losev, Finite W-algebras, *Proceedings of the International Congress of Mathematicians*, Vol. III, pp. 1281-1307, Hindustan Book Agency, New Delhi, 2010.

[27] A. Molev, *Yangians and classical Lie algebras*, Mathematical Surveys and Monographs, **143** American Mathematical Society, Providence, RI, 2007.

[28] M. Nazarov, Quantum Berezinian and the classical Capelli identity, *Lett. Math. Phys.* **21** (1991), 123-131.

[29] M. Nazarov, Yangian of the Queer Lie Superalgebra, *Comm. Math. Phys.* **208** (1999), 195-223.

[30] Y. Peng, Parabolic presentations of the super Yangian $Y(\mathfrak{gl}_{M|N})$, *Comm. Math. Phys.* **307** (2011), 229-259.

[31] Y. Peng, Finite W-superalgebras and truncated super Yangians, *Lett. Math. Phys.* **104** (2014), 89-102.

[32] Y. Peng, On shifted super Yangians and a class of finite W-superalgebras, *J. Algebra.* **422** (2015), 520-562.

[33] Y. Peng, Parabolic presentations of the super Yangian $Y(\mathfrak{gl}_{M|N})$ associated with arbitrary 01-sequences, *Comm. Math. Phys.* **346** (2016), 313-347.

[34] E. Poletaeva and V. Serganova, On Kostant's theorem for the Lie superalgebra Q(n), *Adv. Math.* **300** (2016), 320-359.

[35] E. Poletaeva and V. Serganova, On the finite W-algebra for the Lie superalgebra Q(n) in the non-regular case, *J. Math. Phys.* **58** (2017), no. 11, 111701.

[36] A. Premet, Irreducible representations of Lie algebras of reductive groups and the Kac-Weisfeiler conjecture, *Invent. Math.* **121** (1995), 79-117.

[37] A. Premet, Special transverse slices and their enveloping algebras, *Adv. Math.* **170** (2002), 1-55.

[38] E. Ragoucy and P. Sorba, Yangian realisations from finite $\mathcal{W}$-algebras, *Comm. Math. Phys.* **203** (1999), 551-572.

[39] V. Stukopin, Yangians of Lie superalgebras of type A(m, n), (in Russian) *Funktsional. Anal. i Prilozhen.* **28** (1994), no. 3, 85-88; translation in *Funct. Anal. Appl.* **28** (1994), no. 3, 217-219.

[40] A. Tsymbaliuk, Shuffle algebra realizations of type A super Yangians and quantum affine superalgebras for all Cartan data, *Lett. Math. Phys.* **110** (2020), 2083-2111.

[41] W. Wang, Nilpotent orbits and finite W-algebras, *Fields Institute Communications Series* **59** (2011), 71-105.

[42] W. Wang and L. Zhao, Representations of Lie superalgebras in prime characteristic I, *Proc. London Math. Soc.* **99** (2009), 145–167.

[43] W. Wang and L. Zhao, Representations of Lie superalgebras in prime characteristic II: the queer series, *J. Pure Appl. Algebra.* **215** (2011), 2515-2532.

[44] B. Weisfeiler and V. Kac, On irreducible representations of Lie p-algebras, *Func. Anal. Appl.* **5** (1971), 111–117.

[45] Y. Zeng and B. Shu, Finite W-superalgebras for basic Lie superalgebras, *J. Algebra.* **438** (2015), 188-234.

[46] Y. Zeng and B. Shu, On Kac-Weisfeiler modules for general and special linear Lie superalgebras, *Israel J. Math.* **214** (2016), 471-490.

[47] L. Zhao, Finite W-superalgebras for queer Lie superalgebras, *J. Pure Appl. Algebra.* **218** (2014), 1184-1194.

Department of Mathematics, National Central University, Chung-Li, Taiwan, 32054
E-mail address: ynp@math.ncu.edu.tw