THE p-CENTRE OF YANGIANS AND SHIFTED YANGIANS

JONATHAN BRUNDAN AND LEWIS TOPLEY

ABSTRACT. We study the Yangian Y_n associated to the general linear Lie algebra $\mathfrak{gl}_n$ over a field of positive characteristic, as well as its shifted analog $Y_n(\sigma)$. Our main result gives a description of the centre of $Y_n(\sigma)$: it is a polynomial algebra generated by its *Harish-Chandra centre* (which lifts the centre in characteristic zero) together with a large *p-centre*. Moreover, $Y_n(\sigma)$ is free as a module over its center. In future work, it will be seen that every reduced enveloping algebra $U_\chi(\mathfrak{gl}_n)$ is Morita equivalent to a quotient of an appropriate choice of shifted Yangian, and so our results will have applications in classical representation theory.

1. INTRODUCTION

The Yangian Y_n associated to the Lie algebra $\mathfrak{gl}_n$ over the complex numbers was introduced by the St. Petersburg school around 1980, and is one of the fundamental examples of quantum groups which can be defined using the *RTT formalism* of Faddeev, Reshetikhin and Takhtadzhyan [FRT]. In [D], Drinfeld introduced another presentation, allowing him to extend the notion of Yangian to all semisimple Lie algebras. In this paper, we initiate a study of the Yangian of $\mathfrak{gl}_n$, and its shifted analog in the sense of [BK2], over an arbitrary field $\Bbbk$ of positive characteristic $p > 0$.

Let us explain our motivation for doing this. In characteristic zero, the results of [BK2] show that the shifted Yangians have some truncations which are isomorphic to the *finite W-algebras* associated to nilpotent orbits in $\mathfrak{gl}_n$, as defined for example by Premet [P2]. Finite W-algebras had appeared earlier, in work of Kostant and subsequent work by mathematical physicists, but Premet's motivation came from the representation theory of Lie algebras in positive characteristic: building on his work on the Kac-Weisfeiler conjecture [P1], he also discovered some remarkable finite-dimensional *reduced finite W-algebras*, which are Morita equivalent to reduced enveloping algebras of modular reductive Lie algebras.

Recently, the second author jointly with Goodwin [GT] has introduced *modular finite W-algebras*, which are the precise analog of finite W-algebras in characteristic p. For sufficiently large p, the same algebras appeared already in Premet's foundational work [P3], where they were constructed by reducing the complex finite W-algebras modulo p. Modular finite W-algebras have a large *p-centre*, and Premet's reduced finite W-algebras may be recovered from modular finite W-algebras by specializing at a p-central character. As in characteristic zero, for the Lie algebra $\mathfrak{gl}_n$, the modular finite W-algebra is a truncation of the *modular shifted Yangian* introduced here. The reduced finite W-algebras for $\mathfrak{gl}_n$ can then be obtained from truncated modular shifted Yangians by specializing at some p-central character. For this to be useful in practice, one needs precise information about the structure of the centre of the shifted Yangian in positive characteristic. This is the main purpose of the present article.

2010 *Mathematics Subject Classification*: 17B37.

Research of J. B. supported in part by NSF grant DMS-1700905. Research of L. T. supported by European commision FP7 grant 600376 as well as EPSRC grant EP/N034449/1.

In the remainder of the introduction, we focus on the special case of the Yangian Y_n, rather than its shifted analog $Y_n(\sigma)$, and give a quick outline of our main results in this case. We define the Yangian Y_n over any field $\Bbbk$ to be the associative $\Bbbk$-algebra with generators $\big\{T_{i,j}^{(r)} \,\big|\, 1 \leqslant i,j \leqslant n, r > 0\big\}$ subject to the relations

$$\Big[T_{i,j}^{(r+1)}, T_{k,l}^{(s)}\Big] - \Big[T_{i,j}^{(r)}, T_{k,l}^{(s+1)}\Big] = T_{k,j}^{(r)} T_{i,l}^{(s)} - T_{k,j}^{(s)} T_{i,l}^{(r)} \tag{1.1}$$

for all $1 \leqslant i,j,k,l \leqslant n$ and $r, s \geqslant 0$, adopting the convention that $T_{i,j}^{(0)} := \delta_{i,j}$. When $\Bbbk$ is of characteristic zero, this is the usual RTT presentation as found e.g. in [MNO], and in this case it is well known that Y_n is a filtered deformation of the universal enveloping algebra $U(\mathfrak{gl}_n[t])$ of the polynomial current algebra $\mathfrak{gl}_n[t] := \mathfrak{gl}_n \otimes \Bbbk[t]$. In fact, the same assertion is true when $\Bbbk$ is of positive characteristic, which the reader may regard as the first evidence that this is a reasonable object to consider in the modular setting.

In characteristic zero, the centre $Z(Y_n)$ is generated freely by elements $\big\{C^{(r)} \,\big|\, r > 0\big\}$; see [MNO, Theorem 2.13]. Moreover, $C^{(r+1)}$ is a lift of the obvious central element $z_r := (e_{1,1} + \cdots + e_{n,n}) \otimes t^r$ in the centre $Z(\mathfrak{gl}_n[t])$ of the associated graded algebra $U(\mathfrak{gl}_n[t])$. The usual way to define these elements formally is to work with generating functions in $Y_n[[u^{-1}]]$ following [MNO], setting

$$T_{i,j}(u) := \sum_{r \geqslant 0} T_{i,j}^{(r)} u^{-r} \in Y_n[[u^{-1}]]. \tag{1.2}$$

Then the central elements $C^{(r)}$ are the coefficients of the *quantum determinant*

$$C(u) = \sum_{r \geqslant 0} C^{(r)} u^{-r} := \sum_{g \in \mathfrak{S}_n} \operatorname{sgn}(g) T_{g(1),1}(u) T_{g(2),2}(u-1) \cdots T_{g(n),n}(u-n+1). \tag{1.3}$$

These definitions makes sense when $\operatorname{char} \Bbbk > 0$ too, yielding elements $\big\{C^{(r)} \,\big|\, r > 0\big\}$ which freely generate a subalgebra $Z_{\mathrm{HC}}(Y_n)$ of $Z(Y_n)$, which we call the *Harish-Chandra centre*. This corresponds at the level of the associated graded algebra to the subalgebra of $Z(\mathfrak{gl}_n[t])$ generated by $\{z_r \mid r \geqslant 0\}$.

In positive characteristic, the current algebra $\mathfrak{gl}_n[t]$ admits a natural structure of *restricted Lie algebra*, with p-map $x \mapsto x^{[p]}$ which is defined on a basis of $\mathfrak{gl}_n[t]$ by the rule $(e_{i,j}t^r)^{[p]} := \delta_{i,j} e_{i,j} t^{rp}$. Consequently, $U(\mathfrak{gl}_n[t])$ also has a large *p-centre* generated by the elements

$$\big\{(e_{i,j}t^r)^p - \delta_{i,j} e_{i,j} t^{rp} \,\big|\, 1 \leqslant i,j \leqslant n, r \geqslant 0\big\}, \tag{1.4}$$

and it is natural to look for lifts of these elements in $Z(Y_n)$. We will show that such lifts are provided by the coefficients $\big\{S_{i,j}^{(pr)} \,\big|\, 1 \leqslant i,j \leqslant n, r > 0\big\}$ of the power series

$$S_{i,j}(u) = \sum_{r \geqslant 0} S_{i,j}^{(r)} u^{-r} := T_{i,j}(u) T_{i,j}(u-1) \cdots T_{i,j}(u-p+1). \tag{1.5}$$

These coefficients freely generate another subalgebra $Z_p(Y_n)$ of $Z(Y_n)$ which we call the *p-centre* of Y_n. Together, $Z_{\mathrm{HC}}(Y_n)$ and $Z_p(Y_n)$ generate the full centre $Z(Y_n)$. In fact, we will show that $Z(Y_n)$ is a free polynomial algebra generated by

$$\big\{C^{(r)} \,\big|\, r > 0\big\} \cup \big\{S_{i,j}^{(pr)} \,\big|\, 1 \leqslant i,j \leqslant n \text{ with } (i,j) \neq (1,1), r > 0\big\}. \tag{1.6}$$

Moreover, $Z_{\mathrm{HC}}(Y_n) \cap Z_p(Y_n)$ is freely generated by elements $\big\{BC^{(pr)} \,\big|\, r > 0\big\}$ defined from

$$BC(u) = \sum_{r \geqslant 0} BC^{(r)} u^{-r} := C(u) C(u-1) \cdots C(u-p+1). \tag{1.7}$$

Finally, Y_n is free (of infinite rank) as a module over its centre.

To prove these statements, we found that it was easier to work initially with the Drinfeld presentation rather than the RTT presentation of Y_n, exploiting a different choice of lifts of the elements (1.4) adapted to the Drinfeld generators. There is also a similar family of lifts which plays the analogous role for the shifted Yangians $Y_n(\sigma)$. We refer the reader to Theorem 5.11 for the precise statement of our main result in the general setup.

The remainder of the article is organized as follows. In section 2, we make some auxiliary calculations with the sorts of power series that will be used to define the central elements of $Y_n(\sigma)$ later on. In particular, the results in this section are needed to compute images of these central elements in the associated graded algebra. Then section 3 investigates the centre of this associated graded algebra. The main result here, Theorem 3.4, provides the key upper bound needed in our computation of the centre of $Y_n(\sigma)$ later on. We introduce the modular Yangians and shifted Yangians in section 4, extending some of the fundamental results such as Drinfeld-type presentation from characteristic zero to characteristic p. Section 5 contains our main results describing the centre of $Y_n(\sigma)$ in positive characteristic. Finally, in section 6, we define a modular version of the Yangian of $\mathfrak{sl}_n$, which is a certain subalgebra SY_n of Y_n. We give a presentation for SY_n valid in any characteristic; see Theorem 6.3. As we explain in the subsequent remark, this presentation is equivalent to the usual Drinfeld presentation for the Yangian of $\mathfrak{sl}_n$ taken over the field $\Bbbk$ whenever $\operatorname{char}\Bbbk \neq 2$; Drinfeld's presentation does not even make sense in characteristic 2 since it involves some halves. We also show that $Y_n \cong SY_n \otimes Z_{\mathrm{HC}}(Y_n)$ providing $p \nmid n$, and use this to deduce the results about $Z(Y_n)$ in terms of the RTT generators as formulated in this introduction.

General conventions. Let $\mathbb{N}$ denote the set of natural numbers $\{0, 1, 2, \dots\}$. Always, $\Bbbk$ will be a ground field of characteristic $p > 0$ and $\otimes$ denotes $\otimes_{\Bbbk}$. Let $\mathfrak{gl}_n := \mathfrak{gl}_n(\Bbbk)$ for some fixed $n \geqslant 1$. We denote its matrix units by $\{e_{i,j} \mid i, j = 1, \dots, n\}$ as usual.

By a *filtration* on a (unital) $\Bbbk$-algebra A, we always mean an ascending chain of subspaces

$$\mathrm{F}_0 A \subseteq \mathrm{F}_1 A \subseteq \mathrm{F}_2 A \subseteq \cdots$$

with $A = \bigcup_{r \geqslant 0} \mathrm{F}_r A$ such that $1 \in \mathrm{F}_0 A$ and $(\mathrm{F}_r A)(\mathrm{F}_s A) \subseteq \mathrm{F}_{r+s} A$ for all $r, s \in \mathbb{N}$. Setting $\mathrm{F}_{-1} A := \{0\}$ by convention, the *associated graded algebra* is the (unital) positively graded $\Bbbk$-algebra $\operatorname{gr} A = \bigoplus_{r \in \mathbb{N}} (\operatorname{gr} A)_r$ with $(\operatorname{gr} A)_r := \mathrm{F}_r A / \mathrm{F}_{r-1} A$. Let $\operatorname{gr}_r : \mathrm{F}_r A \to (\operatorname{gr} A)_r$ be the canonical quotient map. Given a subalgebra B of A, we always give B the induced filtration defined from $F_r B := B \cap F_r A$. Then the associated graded algebra $\operatorname{gr} B$ is naturally identified with a graded subalgebra of $\operatorname{gr} A$.

Given an n-tuple $\lambda = (\lambda_1, \dots, \lambda_n) \in \mathbb{N}^n$, let $\ell(\lambda) := \big|\{i = 1, \dots, n \mid \lambda_i \neq 0\}\big|$ and $|\lambda| := \sum_{i=1}^n \lambda_i$. Then set

$$\Lambda(n, r) := \big\{\lambda \in \mathbb{N}^n \;\big|\; |\lambda| = r\big\}, \qquad \Lambda^+(n, r) := \big\{\lambda \in \Lambda(n, r) \;\big|\; \lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n\big\}.$$

Elements of $\Lambda(n, r)$ and $\Lambda^+(n, r)$ are n-part *compositions* and *partitions* of r, respectively. The symmetric group $\mathfrak{S}_n$ acts on the left on $\Lambda(n, r)$ by place permutation, and $\Lambda^+(n, r)$ gives a set of orbit representatives. Also let $\mathfrak{S}_\lambda \leqslant \mathfrak{S}_n$ denote the stabilizer of $\lambda \in \Lambda^+(n, r)$, which is a parabolic subgroup of $\mathfrak{S}_n$. Let $\subseteq$ be the partial order on $\mathbb{N}^n$ defined from $\lambda \subseteq \mu$ if $\lambda_i \leqslant \mu_i$ for each $i = 1, \dots, n$.

We will use the following elementary facts several times.

Lemma 1.1. *For $\lambda \in \Lambda^+(p, r)$, $|\mathfrak{S}_p / \mathfrak{S}_\lambda|$ is non-zero in the field $\Bbbk$ if and only if $p | r$ and $\lambda_1 = \cdots = \lambda_p = r/p$, in which case it equals* 1.

Proof. Obvious. □

Lemma 1.2. *In any associative $\Bbbk$-algebra A, we have that $(\operatorname{ad} x)^p = \operatorname{ad}(x^p)$ for $x \in A$.*

Proof. We have that $\operatorname{ad} x = \lambda_x - \rho_x$, where $\lambda_x, \rho_x : A \to A$ denote the commuting operations of left and right multiplication by x, respectively. Hence $(\operatorname{ad} x)^p = (\lambda_x - \rho_x)^p = (\lambda_x)^p - (\rho_x)^p = \lambda_{x^p} - \rho_{x^p} = \operatorname{ad}(x^p)$ since we are in characteristic p. □

2. Power series

We often need to manipulate power series in $A[[u^{-1}]]$, where A is some $\Bbbk$-algebra (typically, some Yangian). When doing so, it is sometimes convenient to work in the slightly larger ring $A[[u^{-1}]][u]$ of formal Laurent series in u^{-1} with coefficients in A. In this section, we make some general observations about some particular power series which will be needed in several key places later on.

2.1. **Series of type I.** In this subsection, we assume that A_{I} is a filtered $\Bbbk$-algebra and that we are given elements $\{X^{(r)} \in \mathrm{F}_{r-1}A_{\mathrm{I}} \mid r \geqslant 0\}$ such that $\big[\operatorname{gr}_r X^{(r+1)}, \operatorname{gr}_s X^{(s+1)}\big] = 0$ for all $r, s \geqslant 0$. Since $\mathrm{F}_{-1}A_{\mathrm{I}} = \{0\}$ by our general conventions, we necessarily have that $X^{(0)} = 0$. We refer to the following power series as *a series of type* I

$$X_{\mathrm{I}}(u) = \sum_{r \geqslant 0} X_{\mathrm{I}}^{(r)} u^{-r} := \left(\sum_{r \geqslant 0} X^{(r)} u^{-r} \right)^p \in A_{\mathrm{I}}[[u^{-1}]]. \tag{2.1}$$

Lemma 2.1. *We have that $X_{\mathrm{I}}^{(r)} = 0$ for $r < p$ and $X_{\mathrm{I}}^{(p)} = \big(X^{(1)}\big)^p \in \mathrm{F}_0 A_{\mathrm{I}}$. If $r > p$ and $p \mid r$ then $X_{\mathrm{I}}^{(r)} \in \mathrm{F}_{r-p}A_{\mathrm{I}}$ and $X_{\mathrm{I}}^{(r)} \equiv (X^{(r/p)})^p \pmod{\mathrm{F}_{r-p-1}A_{\mathrm{I}}}$. Finally, if $r > p$ and $p \nmid r$ then $X_{\mathrm{I}}^{(r)} \in \mathrm{F}_{r-p-1}A_{\mathrm{I}}$.*

Proof. We obviously have that $X_{\mathrm{I}}^{(r)} = \sum_{\lambda \in \Lambda(p,r)} X^{(\lambda)}$ where $X^{(\lambda)} := X^{(\lambda_1)} X^{(\lambda_2)} \cdots X^{(\lambda_p)}$. Since $X^{(0)} = 0$, the result follows at once when $r < p$. Now suppose that $r \geqslant p$. Then we get that $X_{\mathrm{I}}^{(r)} \in \mathrm{F}_{r-p}A_{\mathrm{I}}$. The commuting assumption on the elements $\operatorname{gr}_{r-1} X^{(r)}$ gives that

$$X^{(g \cdot \lambda)} \equiv X^{(\lambda)} \pmod{\mathrm{F}_{r-p-1}A_{\mathrm{I}}}$$

for all $\lambda \in \Lambda(p,r)$ and $g \in \mathfrak{S}_p$. As $\Lambda^+(p,r)$ is a set of orbit representatives, we deduce that

$$X_{\mathrm{I}}^{(r)} = \sum_{\lambda \in \Lambda^+(p,r)} \sum_{g \mathfrak{S}_\lambda \in \mathfrak{S}_p / \mathfrak{S}_\lambda} X^{(g \cdot \lambda)} \equiv \sum_{\lambda \in \Lambda^+(p,r)} |\mathfrak{S}_p / \mathfrak{S}_\lambda| X^{(\lambda)} \pmod{\mathrm{F}_{r-p-1}A_{\mathrm{I}}}.$$

Now apply Lemma 1.1 to complete the proof. □

Remark 2.2. When we meet the series of type I later on, the elements $X^{(r)}$ will satisfy the additional relations

$$\big[X^{(r)}, X^{(s)}\big] = \sum_{t=r}^{s-1} X^{(t)} X^{(r+s-1-t)} \tag{2.2}$$

for all $1 \leqslant r < s$. Using this assumption, we will see indirectly that the elements $X_{\mathrm{I}}^{(r)}$ can be expressed as polynomials in the commuting elements $\big\{(X^{(s)})^p \,\big|\, 0 < s \leqslant \lfloor r/p \rfloor\big\}$; see Remark 5.6 below. For example, when $p = 2$, we have from the original definition that $X_{\mathrm{I}}^{(3)} = X^{(1)}X^{(2)} + X^{(2)}X^{(1)}$, which using (2.2) may be simplified to obtain $X_{\mathrm{I}}^{(3)} = (X^{(1)})^2$.

Here are some other examples when $p = 2$ under this assumption:

$$\begin{aligned} X_{\mathrm{I}}^{(0)} &= X_{\mathrm{I}}^{(1)} = 0, \\ X_{\mathrm{I}}^{(2)} = X_{\mathrm{I}}^{(3)} &= X_{\mathrm{I}}^{(5)} = (X^{(1)})^2, \\ X_{\mathrm{I}}^{(4)} &= (X^{(2)})^2 + (X^{(1)})^2, \\ X_{\mathrm{I}}^{(6)} &= (X^{(3)})^2 + (X^{(2)})^2 + (X^{(1)})^2, \\ X_{\mathrm{I}}^{(7)} &= (X^{(3)})^2 + (X^{(1)})^2, \\ X_{\mathrm{I}}^{(8)} &= (X^{(4)})^2 + (X^{(2)})^2 + (X^{(1)})^2. \end{aligned}$$

When $p > 2$ one can show that $X_{\mathrm{I}}^{(p+1)} = 0$ assuming (2.2).

2.2. **Series of type II.** Fix $n \geqslant 1$ and a sequence $(d_r)_{r \geqslant 0}$ of natural numbers with $d_0 = 0$. For $\mu \in \mathbb{N}^n$, define

$$d_\mu := \sum_{i=1}^{n} d_{\mu_i}. \tag{2.3}$$

We say that $r > 1$ is *optimal* if $d_\mu < d_r$ for all $\mu \in \mathbb{N}^n$ such that either $|\mu| < r$, or $|\mu| = r$ and $\ell(\mu) > 1$. The sole observation we need about this notion is the following:

Lemma 2.3. *Take $m \geqslant 1$ and suppose that the sequence $(d_r)_{r \geqslant 0}$ is defined from $d_r := 0$ for $r < m$ and $d_r := m\lfloor \frac{r}{m} \rfloor - m$ for $r \geqslant m$. Then mr is optimal for every $r > 1$.*

Proof. Take $r > 1$. We must show that $d_\mu < mr - m$ for $\mu \in \mathbb{N}^n$ such that either $|\mu| < mr$, or $|\mu| = mr$ and $\ell(\mu) > 1$. Take any such composition μ. Write each μ_i as $ma_i + b_i$ for $a_i \geqslant 0$ and $0 \leqslant b_i < m$. If $a_i = 0$ for all i, then $d_\mu = 0 < mr - m$. If there is only one i such that $a_i > 0$ then $d_\mu = ma_i - m$ for this i. This is less than $mr - m$ if $|\mu| < mr$ since $ma_i \leqslant |\mu|$. It is also less than $mr - m$ if $|\mu| = mr$ since in that case $ma_i < |\mu|$ as μ has at least one other non-zero part besides $ma_i + b_i$. Finally, if $a_i > 0$ for at least two different i, then

$$d_\mu = \sum_{i \text{ with } a_i > 0} (ma_i - m) < \left(\sum_{i \text{ with } a_i > 0} ma_i \right) - m \leqslant mr - m,$$

and we are done. □

For the remainder of the subsection, we assume in addition that A_{II} is some filtered $\Bbbk$-algebra, and that we are given commuting elements $\left\{ X_i^{(r)} \in \mathrm{F}_{d_r} A_{\mathrm{II}} \,\middle|\, 1 \leqslant i \leqslant n, r > 0 \right\}$. Setting $X_i^{(0)} := 1$ for all i, which belongs to $\mathrm{F}_{d_0} A_{\mathrm{II}}$ as $d_0 = 0$, consider the *power series of type* II

$$X_{\mathrm{II}}(u) = \sum_{r \geqslant 0} X_{\mathrm{II}}^{(r)} u^{-r} := \prod_{i=1}^{n} \left(\sum_{s \geqslant 0} X_i^{(s)} (u - i + 1)^{-s} \right) \in A_{\mathrm{II}}[[u^{-1}]]. \tag{2.4}$$

Lemma 2.4. *We have that*

$$X_{\mathrm{II}}^{(r)} = \sum_{\lambda \in \Lambda(n,r)} \sum_{\mu \subseteq \lambda} \left(\prod_{i=1}^{n} \binom{\lambda_i - 1}{\lambda_i - \mu_i} (i-1)^{\lambda_i - \mu_i} \right) X^{(\mu)}$$

where $X^{(\mu)} := X_1^{(\mu_1)} \cdots X_n^{(\mu_n)}$.

Proof. By the binomial expansion,

$$(u-i+1)^{-s} = u^{-s}(1+(1-i)u^{-1})^{-s} = \sum_{t\geqslant 0}\binom{-s}{t}(1-i)^t u^{-s-t}$$
$$= \sum_{t\geqslant 0}\binom{s+t-1}{t}(i-1)^t u^{-s-t} = \sum_{t\geqslant s}\binom{t-1}{t-s}(i-1)^{t-s}u^{-t}.$$

It follows that

$$\sum_{s\geqslant 0} X_i^{(s)}(u-i+1)^{-s} = \sum_{s\geqslant 0}\sum_{t\geqslant s}\binom{t-1}{t-s}(i-1)^{t-s}X_i^{(s)}u^{-t}$$
$$= \sum_{t\geqslant 0}\sum_{s=0}^{t}\binom{t-1}{t-s}(i-1)^{t-s}X_i^{(s)}u^{-t}.$$

Now take the product of this over $i = 1, \dots, n$ then expand the n parentheses, using the parts of λ and μ to index the t- and s-summations coming from each bracket, to deduce

$$X_{\mathrm{II}}(u) = \sum_{\lambda\in\mathbb{N}^n}\sum_{\mu\subseteq\lambda}\left(\prod_{i=1}^{n}\binom{\lambda_i-1}{\lambda_i-\mu_i}(i-1)^{\lambda_i-\mu_i}\right)X^{(\mu)}u^{-|\lambda|}.$$

It just remains to take the u^{-r}-coefficient. □

Lemma 2.5. *We always have that* $X_{\mathrm{II}}^{(1)} = X_1^{(1)} + \cdots + X_n^{(1)} \in \mathrm{F}_{d_1}A_{\mathrm{II}}$. *Moreover, if* $r > 1$ *is optimal then* $X_{\mathrm{II}}^{(r)} \in \mathrm{F}_{d_r}A_{\mathrm{II}}$ *and* $X_{\mathrm{II}}^{(r)} \equiv X_1^{(r)} + \cdots + X_n^{(r)} \pmod{\mathrm{F}_{d_r-1}A_{\mathrm{II}}}$.

Proof. The formula for $X_{\mathrm{II}}^{(1)}$ follows immediately from Lemma 2.4. Now assume that $r > 1$ is optimal. We have that $X^{(\mu)} \in \mathrm{F}_{d_\mu}A_{\mathrm{II}}$. So by the definition of optimal and Lemma 2.4, we get that $X_{\mathrm{II}}^{(r)} \in \mathrm{F}_{d_r}A_{\mathrm{II}}$. Moreover, the only terms in the summation from that lemma which are not contained in $\mathrm{F}_{d_r-1}A_{\mathrm{II}}$ arise when $\mu = \lambda$ and $\ell(\lambda) = 1$. □

2.3. **Series of type III.** The third type of power series we consider is defined similarly to type II, however the resulting coefficients are slightly more complicated to describe due to additional $\mathfrak{S}_p$-symmetry; this symmetry arises because the series of type III can be obtained from those of type II by setting $n = p$ and identifying $X_i^{(r)}$ with its $\mathfrak{S}_p$ -conjugates. Suppose that A_{III} is a filtered $\Bbbk$-algebra containing commuting elements $\{X^{(r)} \in \mathrm{F}_{r-1}A_{\mathrm{III}} \,|\, r > 0\}$. Setting $X^{(0)} := 1 \in \mathrm{F}_0 A_{\mathrm{III}}$, we define *power series of type* III

$$X_{\mathrm{III}}(u) = \sum_{r\geqslant 0} X_{\mathrm{III}}^{(r)}u^{-r} := \prod_{i=1}^{p}\left(\sum_{r\geqslant 0}X^{(r)}(u-i+1)^{-r}\right). \tag{2.5}$$

Like in previous subsections, our goal is to obtain information about the top degree component of the elements $X_{\mathrm{III}}^{(r)}$. Henceforth, $X^{(\lambda)}$ will denote $X^{(\lambda_1)}\cdots X^{(\lambda_p)}$ for $\lambda \in \mathbb{N}^p$; this is different to the usage in the previous subsection. Notice that $X^{(\lambda)} \in \mathrm{F}_{|\lambda|-\ell(\lambda)}A_{\mathrm{III}}$.

Lemma 2.6. *We have that*

$$X_{\mathrm{III}}^{(r)} = \sum_{s=0}^{r}\sum_{\mu\in\Lambda^+(p,s)}\sum_{\nu\in\Lambda(p,r-s)}\sum_{g\mathfrak{S}_\mu\in\mathfrak{S}_p/\mathfrak{S}_\mu}\left(\prod_{i=1}^{p}\binom{\mu_{g^{-1}(i)}+\nu_i-1}{\nu_i}(i-1)^{\nu_i}\right)X^{(\mu)}.$$

Proof. Let A_{II} be the tensor product over $\Bbbk$ of p copies of A_{III}, and define

$$X_i^{(r)} := 1^{\otimes(i-1)} \otimes X^{(r)} \otimes 1^{\otimes(p-i)} \in A_{\mathrm{II}}.$$

Applying Lemma 2.4 with $n = p$, we deduce that the power series

$$X_{\mathrm{II}}(u) = \sum_{r \geqslant 0} X_{\mathrm{II}}^{(r)} u^{-r} := \prod_{i=1}^{p} \left(\sum_{r \geqslant 0} X_i^{(r)} (u - i + 1)^{-r} \right) \in A_{\mathrm{II}}[[u^{-1}]]$$

satisfies

$$X_{\mathrm{II}}^{(r)} = \sum_{\lambda \in \Lambda(p,r)} \sum_{\mu \subseteq \lambda} \left(\prod_{i=1}^{p} \binom{\lambda_i - 1}{\lambda_i - \mu_i} (i-1)^{\lambda_i - \mu_i} \right) X^{(\mu_1)} \otimes X^{(\mu_2)} \otimes \cdots \otimes X^{(\mu_p)}.$$

Then we apply the linear map $A_{\mathrm{II}} \to A_{\mathrm{III}}$ defined by multiplying out the tensors. This maps $X_{\mathrm{II}}^{(r)} \mapsto X_{\mathrm{III}}^{(r)}$, so we deduce that

$$\begin{aligned} X_{\mathrm{III}}^{(r)} &= \sum_{\lambda \in \Lambda(p,r)} \sum_{\mu \subseteq \lambda} \left(\prod_{i=1}^{p} \binom{\lambda_i - 1}{\lambda_i - \mu_i} (i-1)^{\lambda_i - \mu_i} \right) X^{(\mu)} \\ &= \sum_{s=0}^{r} \sum_{\mu \in \Lambda(p,s)} \sum_{\nu \in \Lambda(p,r-s)} \left(\prod_{i=1}^{p} \binom{\mu_i + \nu_i - 1}{\nu_i} (i-1)^{\nu_i} \right) X^{(\mu)}, \end{aligned}$$

where to get the second equation we switched the summations then replaced λ by $\mu + \nu$. Since $X^{(\mu)} = X^{(\mu')}$ if μ and μ' are in the same $\mathfrak{S}_p$-orbit, we can simplify this further to get the final formula. □

We need to study the expression from Lemma 2.6 further. Let $\Bbbk[x_1, \dots, x_p]^{\mathfrak{S}_p}$ be the algebra of symmetric polynomials over $\Bbbk$. It is well known that this is freely generated by $\varepsilon_1, \dots, \varepsilon_p$, the *elementary symmetric polynomials* defined from

$$\varepsilon_r = \varepsilon_r(x_1, \dots, x_p) := \sum_{1 \leqslant i_1 < \cdots < i_r \leqslant p} x_{i_1} \cdots x_{i_r}.$$

We also have the *power sums*

$$\pi_r = \pi_r(x_1, \dots, x_p) := x_1^r + \cdots + x_p^r.$$

These do not generate $\Bbbk[x_1, \dots, x_p]^{\mathfrak{S}_p}$ since we are in positive characteristic, but nevertheless every homogeneous symmetric polynomial of degree $< p$ can be written as a polynomial in the power sums $\pi_1, \dots, \pi_{p-1}$. This follows by a simple inductive argument from Newton's formula which holds over the integers and hence over $\Bbbk$:

$$k\varepsilon_k = \sum_{i=1}^{k} (-1)^{i-1} \pi_i \varepsilon_{k-i}. \tag{2.6}$$

Lemma 2.7. *If $f(x_1, \dots, x_p) \in \Bbbk[x_1, \dots, x_p]^{\mathfrak{S}_p}$ is homogeneous of degree $0 < l < p - 1$ then $f(0, 1, \dots, p-1) = 0$.*

Proof. There is nothing to do if $p = 2$, so assume also that $p > 2$. By the remarks preceeding the lemma, it suffices to prove the claim that $\pi_l(0, 1, \dots, p-1) = 0$ for $l = 1, 2, \dots, p-2$. To see this, we appeal to the following famous identity, valid over $\mathbb{Z}$, which is due to Pascal:

$$\sum_{m=0}^{l} \binom{l+1}{m} (1^m + 2^m + \cdots + k^m) = (k+1)^{l+1} - 1$$

for all $k, l \in \mathbb{N}$. Pascal's identity can be easily proven by induction on k: the case $k = 1$ is just $2^{l+1} = \sum_{m=0}^{l+1} \binom{l+1}{m}$, and the inductive step follows quickly from the identity

$(k+1)^{l+1} = \sum_{m=0}^{l+1} \binom{l+1}{m} k^m$, which is a special case of the binomial theorem. Substituting $k = p-1$ and working over $\Bbbk$ we get that

$$\sum_{m=1}^{l} \binom{l+1}{m} \pi_m(0,1,\dots,p-1) = \sum_{m=1}^{l} \binom{l+1}{m} (1^m + 2^m + \cdots + (p-1)^m) = 0.$$

The claim follows easily from this by induction on $l = 1, \dots, p-2$; one needs to note that $\binom{l+1}{m}$ is non-zero in $\Bbbk$ for $1 \leqslant l \leqslant p-2$ and $m = 1, ..., l$. □

Now take $r \geqslant 0$ and $\mu \in \Lambda^+(p,s)$ for some $0 \leqslant s \leqslant r$, and define

$$\gamma_\mu^{(r)}(x_1,\dots,x_p) := \sum_{\nu \in \Lambda(p, r-|\mu|)} \sum_{g\mathfrak{S}_\mu \in \mathfrak{S}_p/\mathfrak{S}_\mu} \left(\prod_{i=1}^{p} \binom{\mu_{g^{-1}(i)} + \nu_i - 1}{\nu_i} x_i^{\nu_i} \right). \tag{2.7}$$

This is relevant because by Lemma 2.6 we have that

$$X_{\mathrm{III}}^{(r)} = \sum_{s=0}^{r} \sum_{\mu \in \Lambda^+(p,s)} \gamma_\mu^{(r)}(0,1,2,\dots,p-1) X^{(\mu)}. \tag{2.8}$$

In fact, $\gamma_\mu^{(r)}$ belongs to $\Bbbk[x_1,\dots,x_p]^{\mathfrak{S}_p}$:

Lemma 2.8. *$\gamma_\mu^{(r)}(x_1,\dots,x_p)$ is a homogeneous symmetric polynomial of degree $r - |\mu|$.*

Proof. The claim about degree is clear. For $h \in \mathfrak{S}_p$, we have that

$$\begin{aligned} h \cdot \gamma_\mu^{(r)}(x_1,\dots,x_p) &= \sum_{\nu \in \Lambda(p, r-|\mu|)} \sum_{g\mathfrak{S}_\mu \in \mathfrak{S}_p/\mathfrak{S}_\mu} \left(\prod_{i=1}^{p} \binom{\mu_{g^{-1}(i)} + \nu_i - 1}{\nu_i} x_{h(i)}^{\nu_i} \right) \\ &= \sum_{\nu \in \Lambda(p, r-|\mu|)} \sum_{g\mathfrak{S}_\mu \in \mathfrak{S}_p/\mathfrak{S}_\mu} \left(\prod_{i=1}^{p} \binom{\mu_{(hg)^{-1}(i)} + \nu_i - 1}{\nu_i} x_i^{\nu_i} \right) \end{aligned}$$

which equals $\gamma_\mu^{(r)}(x_1,\dots,x_p)$ because h permutes the coset space $\mathfrak{S}_p/\mathfrak{S}_\mu$. Hence, it is a symmetric polynomial. □

Now we can obtain our main result about the top degree component of $X_{\mathrm{III}}^{(r)}$.

Lemma 2.9. *We have $X_{\mathrm{III}}^{(0)} = 1$, $X_{\mathrm{III}}^{(r)} = 0$ for $r = 1, \dots, p-1$, and*

$$X_{\mathrm{III}}^{(p)} = (X^{(1)})^p - X^{(1)} \in \mathrm{F}_0 A_{\mathrm{III}}.$$

If $r > p$ and $p \mid r$ then $X_{\mathrm{III}}^{(r)} \in \mathrm{F}_{r-p} A_{\mathrm{III}}$ and

$$X_{\mathrm{III}}^{(r)} \equiv (X^{(r/p)})^p - X^{(r-p+1)} \pmod{\mathrm{F}_{r-p-1} A_{\mathrm{III}}}.$$

Finally, if $r > p$ and $p \nmid r$ then $X_{\mathrm{III}}^{(r)} \in \mathrm{F}_{r-p-1} A_{\mathrm{III}}$.

Proof. Putting Lemmas 2.7–2.8 together shows that

$$\gamma_\mu^{(r)}(0,1,\dots,p-1) = 0 \qquad \text{whenever} \qquad 0 < r - |\mu| < p-1.$$

Also from (2.7), it is clear that $\gamma_\mu^{(r)} = \delta_{r,0}$ in case $|\mu| = 0$. The case $|\mu| = r$ is also easy to understand: we then have simply that $\gamma_\mu^{(r)} = |\mathfrak{S}_p/\mathfrak{S}_\mu|$, which is 1 if all parts of μ are

equal and 0 otherwise, thanks to Lemma 1.1. So we can deduce from (2.8) that

$$X^{(r)}_{\mathrm{III}} = \begin{cases} 1 & \text{if } r = 0, \\ (X^{(r/p)})^p + \displaystyle\sum_{s=1}^{r-p+1} \sum_{\mu \in \Lambda^+(p,s)} \gamma^{(r)}_\mu(0,1,\dots,p-1) X^{(\mu)} & \text{if } p \mid r > 0, \\ \displaystyle\sum_{s=1}^{r-p+1} \sum_{\mu \in \Lambda^+(p,s)} \gamma^{(r)}_\mu(0,1,\dots,p-1) X^{(\mu)} & \text{if } p \nmid r > 0. \end{cases} \tag{2.9}$$

The lemma follows immediately from this in case $r < p$. Now assume that $r \geqslant p$. For $1 \leqslant s \leqslant r-p+1$ and $\mu \in \Lambda^+(p,s)$, we have that $X^{(\mu)} \in \mathrm{F}_{|\mu|-\ell(\mu)} A_{\mathrm{III}} \subseteq \mathrm{F}_{r-p} A_{\mathrm{III}}$, showing that $X^{(r)}_{\mathrm{III}} \in \mathrm{F}_{r-p} A_{\mathrm{III}}$. Moreover, $X^{(\mu)} \in \mathrm{F}_{r-p-1} A_{\mathrm{III}}$ unless $\mu = (r-p+1, 0, \dots, 0)$. To complete the proof, we show for this μ that

$$\gamma^{(r)}_\mu(0,1,\dots,p-1) = -\binom{r-1}{p-1},$$

which is 0 if $p \nmid r$ and -1 if $p \mid r$.

So $\mu = (r-p+1, 0, \dots, 0)$. A set of representatives for $\mathfrak{S}_p/\mathfrak{S}_\mu$ is given by the p distinct powers of the p-cycle $(1\ 2\ \cdots\ p)$. The product of binomial coefficients in the definition of $\gamma^{(r)}_\mu(0,1,\dots,p-1)$ is non-zero only when ν has just one non-zero part, which is necessarily equal to $p-1$. Moreover, if this non-zero part is the jth part of ν, we must have that $\mu_{g^{-1}(j)} = r-p+1$ too, i.e. there is just one choice of g that gives a non-zero contribution. We deduce that

$$\gamma^{(r)}_\mu(0,1,\dots,p-1) = \sum_{j=1}^{p} \binom{r-p+1+p-1-1}{p-1} (j-1)^{p-1} \equiv -\binom{r-1}{p-1} \pmod{p}$$

as claimed. □

Remark 2.10. We will show later on that the elements $X^{(r)}_{\mathrm{III}}$ for $p \nmid r$ can be expressed as polynomials in $\{X^{(ps)}_{\mathrm{III}} \mid 0 < s \leqslant \lfloor r/p \rfloor\}$; see Remark 5.9. It seems to be hard to give a direct proof of this. Here are some explicit examples which we computed by hand using (2.9). When $p = 2$:

$$\begin{aligned} X^{(1)}_{\mathrm{III}} &= 0, \\ X^{(2)}_{\mathrm{III}} = X^{(3)}_{\mathrm{III}} = X^{(5)}_{\mathrm{III}} &= (X^{(1)})^2 + X^{(1)}, \\ X^{(4)}_{\mathrm{III}} &= (X^{(2)})^2 + X^{(3)} + X^{(1)}X^{(2)} + X^{(2)} + X^{(2)}_{\mathrm{III}}, \\ X^{(6)}_{\mathrm{III}} &= (X^{(3)})^2 + X^{(5)} + X^{(1)}X^{(3)} + X^{(1)}X^{(4)} + X^{(2)}X^{(3)} + X^{(3)} + X^{(4)}_{\mathrm{III}}. \end{aligned}$$

When $p = 3$:

$$\begin{aligned} X^{(1)}_{\mathrm{III}} = X^{(2)}_{\mathrm{III}} = X^{(4)}_{\mathrm{III}} &= 0, \\ X^{(3)}_{\mathrm{III}} = X^{(5)}_{\mathrm{III}} &= (X^{(1)})^3 - X^{(1)}, \\ X^{(6)}_{\mathrm{III}} &= (X^{(2)})^3 - X^{(4)} + X^{(1)}X^{(3)} - (X^{(1)})^2 X^{(2)} + X^{(2)} - (X^{(2)})^2. \end{aligned}$$

Also, for all $p > 2$, we have that $X^{(p+1)}_{\mathrm{III}} = 0$.

2.4. **Series of type IV.** The fourth type is defined in almost the same way as type III. So again we assume that A_{IV} is a filtered $\Bbbk$-algebra containing commuting elements $\{X^{(r)} \in \mathrm{F}_{r-1}A_{\mathrm{IV}} \mid r > 0\}$. However now we set $X^{(0)} := 0$, before defining *power series of type* IV

$$X_{\mathrm{IV}}(u) = \sum_{r \geqslant 0} X_{\mathrm{IV}}^{(r)} u^{-r}$$

by the same formula (2.5) which we used to define $X_{\mathrm{III}}(u)$.

The elements $X_{\mathrm{IV}}^{(r)}$ are given by the same formula that was derived in Lemma 2.6 for the elements $X_{\mathrm{III}}^{(r)}$. However now for $\mu \in \Lambda(p,s)$ the monomial $X^{(\mu)} = X^{(\mu_1)} \cdots X^{(\mu_p)}$ is zero unless $\ell(\mu) = p$. The following is an immediate consequence of Lemma 2.9 using this observation.

Lemma 2.11. *We have that $X_{\mathrm{IV}}^{(r)} = 0$ for $r < p$ and $X_{\mathrm{IV}}^{(p)} = (X^{(1)})^p \in \mathrm{F}_0 A_{\mathrm{IV}}$. If $r > p$ and $p \mid r$ then $X_{\mathrm{IV}}^{(r)} \in \mathrm{F}_{r-p}A_{\mathrm{IV}}$ and $X_{\mathrm{IV}}^{(r)} \equiv (X^{(r/p)})^p \pmod{\mathrm{F}_{r-p-1}A_{\mathrm{IV}}}$. Finally, if $r > p$ and $p \nmid r$ then $X_{\mathrm{IV}}^{(r)} \in \mathrm{F}_{r-p-1}A_{\mathrm{IV}}$.*

Remark 2.12. Like in Remark 2.10, we will see later that $X_{\mathrm{IV}}^{(r)}$ for $p \nmid r$ can be expressed as a polynomial in $\{X_{\mathrm{IV}}^{(ps)} \mid 0 < s \leqslant \lfloor r/p \rfloor\}$; see Remark 6.10.

3. The shifted current algebra

The shifted Yangian is a filtered deformation of the universal enveloping algebra of a Lie algebra we call the *shifted current algebra*. In this section, we discuss this Lie algebra, describing the centre of its enveloping algebra. Our notation follows [BK3, ch. 2].

3.1. **The shift matrix.** A *shift matrix* is an $n \times n$ array $\sigma = (s_{i,j})_{1 \leqslant i,j \leqslant n}$ of non-negative integers satisfying

$$s_{i,j} + s_{j,k} = s_{i,k} \tag{3.1}$$

whenever $|i-j| + |j-k| = |i-k|$. It follows from the definition that $s_{i,i} = 0$ for $1 \leqslant i \leqslant n$, and so σ is entirely determined by the super-diagonal entries $s_{1,2}, s_{2,3}, ..., s_{n-1,n}$ and the sub-diagonal entries $s_{n,n-1}, s_{n-1,n-2}, ..., s_{2,1}$. We keep a choice of shift matrix fixed for the remainder of the section.

3.2. **The shifted current algebra.** The *current algebra* is the Lie algebra $\mathfrak{gl}_n[t] := \mathfrak{gl}_n \otimes \Bbbk[t]$. We will always denote this Lie algebra by $\mathfrak{g}$ and write $U(\mathfrak{g})$ for its enveloping algebra and $S(\mathfrak{g})$ for the symmetric algebra. When $x \in \mathfrak{gl}_n$ and $f \in \Bbbk[t]$ we usually abbreviate $x \otimes f = xf \in \mathfrak{g}$. As a vector space, $\mathfrak{g}$ is spanned by elements $\{e_{i,j}t^r \mid 1 \leqslant i,j \leqslant n, r \geqslant 0\}$, and the Lie bracket is given by

$$[e_{i,j}t^r, e_{k,l}t^s] = \delta_{k,j}e_{i,l}t^{r+s} - \delta_{l,i}e_{k,j}t^{r+s} \tag{3.2}$$

where $1 \leqslant i,j,k,l \leqslant n$ and $r,s \geqslant 0$.

For our fixed shift matrix σ, the *shifted current algebra* is $\mathfrak{g}_\sigma \subseteq \mathfrak{g}$ spanned by

$$\left\{e_{i,j}t^r \,\middle|\, 1 \leqslant i,j \leqslant n, r \geqslant s_{i,j}\right\}. \tag{3.3}$$

Lemma 3.1. *The shifted current algebra $\mathfrak{g}_\sigma$ is a Lie subalgebra of $\mathfrak{g}$, and it is generated as a Lie algebra by*

$$\left\{e_{i,i}t^r \,\middle|\, 1 \leqslant i \leqslant n, r \geqslant 0\right\} \cup \left\{e_{i,i+1}t^r, e_{i+1,i}t^s \,\middle|\, 1 \leqslant i < n, r \geqslant s_{i,i+1}, s \geqslant s_{i+1,i}\right\}. \tag{3.4}$$

Proof. First we show that the span of (3.3) is closed under the bracket. Let $1 \leqslant i, j, k, l \leqslant n$, $r \geqslant s_{i,j}$ and $s \geqslant s_{k,l}$. By (3.2) we only need to check that $j = k$ implies $r + s \geqslant s_{i,l}$. When $i < j < l$ or $l < j < i$ this follows from (3.1) so it remains to check the case $l < i < j$. Now we have $r + s \geqslant s \geqslant s_{j,l} = s_{j,i} + s_{i,l} \geqslant s_{i,l}$.

Now take $i < j$ and $r \geqslant s_{i,j}$. The fact that $e_{i,j}t^r$ lies in the algebra generated by (3.4) can be proven by induction on $j - i$. A similar argument treats the case that $i > j$, completing the proof. □

The adjoint action of $\mathfrak{g}_\sigma$ on itself extends uniquely to actions of $\mathfrak{g}_\sigma$ on $U(\mathfrak{g}_\sigma)$ and $S(\mathfrak{g}_\sigma)$ by derivations. The invariant subalgebras are denoted $U(\mathfrak{g}_\sigma)^{\mathfrak{g}_\sigma}$ and $S(\mathfrak{g}_\sigma)^{\mathfrak{g}_\sigma}$, and the equality $Z(\mathfrak{g}_\sigma) = U(\mathfrak{g}_\sigma)^{\mathfrak{g}_\sigma}$ follows from general principles; e.g. see [Dix, 2.4.9(i)].

There is one obvious family of central elements in $U(\mathfrak{g}_\sigma)$. For any $r \in \mathbb{N}$, we set

$$z_r := e_{1,1}t^r + \cdots + e_{n,n}t^r \in \mathfrak{g}_\sigma. \tag{3.5}$$

Then the set $\{z_r \mid r \geqslant 0\}$ forms a basis for the centre $\mathfrak{z}(\mathfrak{g}_\sigma)$ of $\mathfrak{g}_\sigma$, and $\Bbbk[z_r \mid r \geqslant 0]$ is a subalgebra of $Z(\mathfrak{g}_\sigma)$. The elements z_r also define symmetric invariants in $S(\mathfrak{g}_\sigma)^{\mathfrak{g}_\sigma}$.

3.3. **Symmetric invariants.** The current Lie algebra $\mathfrak{g}$ has an obvious grading with $e_{i,j}t^r$ in degree r, and $\mathfrak{g}_\sigma$ is a graded subalgebra. There is also a filtration

$$U(\mathfrak{g}_\sigma) = \bigcup_{r \geqslant 0} \mathrm{F}_r U(\mathfrak{g}_\sigma) \tag{3.6}$$

of the universal enveloping algebra $U(\mathfrak{g}_\sigma)$, which is defined by placing $e_{i,j}t^r$ in degree $r+1$, i.e. $\mathrm{F}_r U(\mathfrak{g}_\sigma)$ is the span of all monomials of the form $e_{i_1,j_1}t^{r_1} \cdots e_{i_m,j_m}t^{r_m}$ with total degree $(r_1 + 1) + \cdots + (r_m + 1) \leqslant r$. The associated graded algebra $\operatorname{gr} U(\mathfrak{g}_\sigma)$ is isomorphic (both as a graded algebra and as a graded $\mathfrak{g}_\sigma$-module) to $S(\mathfrak{g}_\sigma)$ graded so that $e_{i,j}t^r$ is in degree $r + 1$. We get induced an inclusion

$$\operatorname{gr} Z(\mathfrak{g}_\sigma) \subseteq S(\mathfrak{g}_\sigma)^{\mathfrak{g}_\sigma}. \tag{3.7}$$

In the remainder of the section, we are going to use this to compute $Z(\mathfrak{g}_\sigma)$, revealing in particular that equality holds in (3.7). First, we must describe $S(\mathfrak{g}_\sigma)^{\mathfrak{g}_\sigma}$ explicitly.

Lemma 3.2. *The invariant algebra $S(\mathfrak{g}_\sigma)^{\mathfrak{g}_\sigma}$ is generated by $\{z_r \mid r \geqslant 0\}$ together with $(\mathfrak{g}_\sigma)^p := \{x^p \mid x \in \mathfrak{g}_\sigma\} \subset S(\mathfrak{g}_\sigma)$. In fact, $S(\mathfrak{g}_\sigma)^{\mathfrak{g}_\sigma}$ is freely generated by*

$$\{z_r \mid r \geqslant 0\} \cup \big\{(e_{i,j}t^r)^p \,\big|\, 1 \leqslant i, j \leqslant n \text{ with } (i,j) \neq (1,1), r \geqslant s_{i,j}\big\}. \tag{3.8}$$

Proof. Since we are in characteristic $p > 0$, we have that $(\mathfrak{g}_\sigma)^p \subset S(\mathfrak{g}_\sigma)^{\mathfrak{g}_\sigma}$. Let $I(\mathfrak{g}_\sigma)$ be the subalgebra of $S(\mathfrak{g}_\sigma)^{\mathfrak{g}_\sigma}$ generated by $(\mathfrak{g}_\sigma)^p$ and $\{z_r \mid r \geqslant 0\}$. Let

$$B := \{(i, j, r) \mid 1 \leqslant i, j \leqslant n \text{ with } (i,j) \neq (1,1), r \geqslant s_{i,j}\}$$

for short. Since the elements $\{z_r \mid r \geqslant 0\} \cup \big\{e_{i,j}t^r \,\big|\, r \geqslant 0, (i,j,r) \in B\big\}$ give a basis for $\mathfrak{g}_\sigma$, it follows that

$$S(\mathfrak{g}_\sigma) = \Bbbk[z_r \mid r \geqslant 0]\big[e_{i,j}t^r \,\big|\, (i,j,r) \in B\big],$$
$$I(\mathfrak{g}_\sigma) = \Bbbk[z_r \mid r \geqslant 0]\big[(e_{i,j}t^r)^p \,\big|\, (i,j,r) \in B\big],$$

with both being free polynomial algebras. Hence, $S(\mathfrak{g}_\sigma)$ is free as an $I(\mathfrak{g}_\sigma)$-module with basis $\Big\{\prod_{(i,j,r)\in B}(e_{i,j}t^r)^{\omega(i,j,r)} \,\Big|\, \omega \in \Omega\Big\}$, where

$$\Omega := \left\{ \omega : B \to \mathbb{N} \;\middle|\; \begin{array}{l} 0 \leqslant \omega(i,j,r) < p \text{ for all } (i,j,r) \in B, \\ \omega(i,j,r) = 0 \text{ for all but finitely many } (i,j,r) \in B \end{array} \right\}.$$

To complete the proof of the lemma, we must show that $S(\mathfrak{g}_\sigma)^{\mathfrak{g}_\sigma} \subseteq I(\mathfrak{g}_\sigma)$. To do this, take $f \in S(\mathfrak{g}_\sigma)^{\mathfrak{g}_\sigma}$ and write it as

$$f = \sum_{\omega \in \Omega} c_\omega \prod_{(i,j,r) \in B} (e_{i,j} t^r)^{\omega(i,j,r)}$$

for $c_\omega \in I(\mathfrak{g}_\sigma)$, all but finitely many of which are zero. Also fix a non-zero function $\omega \in \Omega$. We must show that $c_\omega = 0$.

Suppose first that $\omega(i,j,r) > 0$ for some $(i,j,r) \in B$ with $i \neq j$. Choose $s \in \mathbb{N}$ that it is bigger than all r' such that $\omega(i',j',r') > 0$ for $(i',j',r') \in B$. Using the Leibniz rule we see that

$$\begin{aligned}\mathrm{ad}(e_{i,i}t^s)(f) = \sum_{\omega' \in \Omega} c_{\omega'} \sum_{\substack{(i',j',r') \in B \\ \omega(i',j',r') > 0}} \omega(i',j',r') (e_{i',j'} t^{r'})^{\omega(i',j',r')-1} \left[e_{i,i} t^s, e_{i',j'} t^{r'} \right] \\ \times \prod_{\substack{(i'',j'',r'') \in B \\ (i'',j'',r'') \neq (i',j',r')}} (e_{i'',j''} t^{r''})^{\omega(i'',j'',r'')}.\end{aligned}$$

It is crucial for this that the coefficients $c_{\omega'}$ belong to $I(\mathfrak{g}_\sigma) \subseteq S(\mathfrak{g}_\sigma)^{\mathfrak{g}_\sigma}$. Thanks to the choice of s, the coefficient of

$$(e_{i,j} t^r)^{\omega(i,j,r)-1} e_{i,j} t^{r+s} \prod_{\substack{(i'',j'',r'') \in B \\ (i'',j'',r'') \neq (i,j,r)}} (e_{i'',j''} t^{r''})^{\omega(i'',j'',k'')}$$

in this expression is $c_\omega \omega(i,j,r)$. But also it must be zero since $f \in S(\mathfrak{g}_\sigma)^{\mathfrak{g}_\sigma}$. Since $\omega(i,j,r)$ is non-zero in $\Bbbk$, we deduce that $c_\omega = 0$ as required.

A very similar line of reasoning treats the case that $\omega(j,j,r) > 0$ for some $(j,j,r) \in B$. This time, one picks $i \neq j$ (possible as $n > 1$) and considers the coefficient of

$$(e_{j,j} t^r)^{\omega(j,j,r)-1} e_{i,j} t^{r+s} \prod_{\substack{(i'',j'',r'') \in B \\ (i'',j'',r'') \neq (j,j,r)}} (e_{i'',j''} t^{r''})^{\omega(i'',j'',r'')}$$

in $\mathrm{ad}(e_{i,j}t^s)(f)$ for s as before. □

3.4. **The restricted structure.** If $\mathfrak{r}$ is any Lie algebra over $\Bbbk$ then a *restricted structure on* $\mathfrak{r}$ is given by defining a *p-map* $x \mapsto x^{[p]}$ on $\mathfrak{r}$ such that the map

$$\xi : \mathfrak{r} \to U(\mathfrak{r}), \qquad x \mapsto x^p - x^{[p]}$$

satisfies two properties: (i) $\xi(\mathfrak{r})$ lies in the centre of $U(\mathfrak{r})$; (ii) ξ is p-semilinear. We remind the reader that p-semilinearity means that $\xi(\lambda x) = \lambda^p \xi(x)$ and $\xi(x+y) = \xi(x) + \xi(y)$, for all $x, y \in \mathfrak{r}$, $\lambda \in \Bbbk$. It is readily seen that a p-semilinear map $\mathfrak{g} \to \mathfrak{g}$ is determined by its effect on a basis, and so the same is true of a p-map.

Lemma 3.3. *The current algebra $\mathfrak{g}$ is a restricted Lie algebra with p-map defined on the basis by the rule $(xt^r)^{[p]} := x^{[p]} t^{rp}$ where $x^{[p]}$ denotes the pth matrix power of $x \in \mathfrak{gl}_n$. Moreover, $\mathfrak{g}_\sigma$ is a restricted Lie subalgebra.*

Proof. Using Lemma 1.2, it is easy to see that when $\mathfrak{r}$ is a restricted Lie algebra with p-map $x \mapsto x^{[p]}$ and C is any commutative $\Bbbk$-algebra, the tensor product $\mathfrak{r} \otimes C$ acquires the structure of a restricted Lie algebra over $\Bbbk$ with multiplication and restricted structure given by $[x \otimes a, y \otimes b] := [x,y] \otimes ab$ and $(x \otimes a)^{[p]} := x^{[p]} \otimes a^p$ for $x, y \in \mathfrak{r}$ and $a, b \in C$. Since $\mathfrak{gl}_n$ is restricted with p-map given by the pth power of matrices the first claim of the lemma follows. Clearly $x \mapsto x^p - x^{[p]}$ sends $\mathfrak{g}_\sigma$ to $U(\mathfrak{g}_\sigma)$, and so $\mathfrak{g}_\sigma$ is a restricted subalgebra. □

3.5. **The centre of $U(\mathfrak{g}_\sigma)$.** Using the restricted structure, we can define the *p-centre* $Z_p(\mathfrak{g}_\sigma)$ of $U(\mathfrak{g}_\sigma)$ to be the subalgebra of $Z(\mathfrak{g}_\sigma)$ generated by $x^p - x^{[p]}$ for all $x \in \mathfrak{g}_\sigma$. Since the map $\xi : x \mapsto x^p - x^{[p]}$ is p-semilinear, we have that

$$Z_p(\mathfrak{g}_\sigma) = \Bbbk\left[\left(e_{i,j}t^r\right)^p - \delta_{i,j}e_{i,j}t^{rp} \;\middle|\; 1 \leqslant i,j \leqslant n, r \geqslant s_{i,j} \right] \tag{3.9}$$

as a free polynomial algebra.

Theorem 3.4. *The centre $Z(\mathfrak{g}_\sigma)$ of $U(\mathfrak{g}_\sigma)$ is generated by $\{z_r \mid r \geqslant 0\}$ and $Z_p(\mathfrak{g}_\sigma)$. In fact, $Z(\mathfrak{g}_\sigma)$ is freely generated by*

$$\{z_r \mid r \geqslant 0\} \cup \left\{ (e_{i,j}t^r)^p - \delta_{i,j}e_{i,j}t^{rp} \;\middle|\; 1 \leqslant i,j \leqslant n \text{ with } (i,j) \neq (1,1), r \geqslant s_{i,j} \right\}. \tag{3.10}$$

Proof. It suffices to prove the second statement. Let Z be the subalgebra of $Z(\mathfrak{g}_\sigma)$ generated by (3.10). For $r \geqslant 0$ we have that $z_r \in \mathrm{F}_{r+1}U(\mathfrak{g}_\sigma)$ and

$$\mathrm{gr}_{r+1}\, z_r = z_r \in S(\mathfrak{g}_\sigma).$$

For $1 \leqslant i,j \leqslant n$ and $r \geqslant s_{i,j}$ we have that $(e_{i,j}t^r)^p - \delta_{i,j}e_{i,j}t^{rp} \in \mathrm{F}_{rp+p}U(\mathfrak{g}_\sigma)$ and

$$\mathrm{gr}_{rp+p}\left[(e_{i,j}t^r)^p - \delta_{i,j}e_{i,j}t^{rp}\right] = (e_{i,j}t^r)^p \in S(\mathfrak{g}_\sigma).$$

So the elements (3.10) are lifts of the algebraically independent generators of $S(\mathfrak{g}_\sigma)^{\mathfrak{g}_\sigma}$ from (3.8). It follows that the elements (3.10) are themselves algebraically independent, and moreover $S(\mathfrak{g}_\sigma)^{\mathfrak{g}_\sigma} \subseteq \mathrm{gr}\, Z$. Thanks to (3.7), we also have $\mathrm{gr}\, Z \subseteq \mathrm{gr}\, Z(\mathfrak{g}_\sigma) \subseteq S(\mathfrak{g}_\sigma)^{\mathfrak{g}_\sigma}$, so equality must hold throughout: $\mathrm{gr}\, Z = \mathrm{gr}\, Z(\mathfrak{g}_\sigma) = S(\mathfrak{g}_\sigma)^{\mathfrak{g}_\sigma}$. This implies $Z = Z(\mathfrak{g}_\sigma)$. □

Remark 3.5. If $\Bbbk$ is algebraically closed of characteristic $p > 0$ and $\mathfrak{g} = \mathrm{Lie}(G)$ for a reductive algebraic $\Bbbk$-group G satisfying standard hypotheses, then it is well known that the centre of $U(\mathfrak{g})$ is generated by the p-centre and $U(\mathfrak{g})^G$. Theorem 3.4 can be seen as an analogue of this classical result in the context of the current algebra $\mathfrak{gl}_n[t]$.

4. Modular Yangians and shifted Yangians

In this section, we define and study the Yangian Y_n and its shifted subalgebra $Y_n(\sigma)$ in positive characteristic. In particular, we prove that the RTT presentation for Y_n from the introduction is equivalent to a slightly modified version of the Drinfeld presentation.

4.1. **The RTT generators.** We define the Yangian Y_n over $\Bbbk$ as in the introduction. So it has generators $\left\{T_{i,j}^{(r)} \;\middle|\; 1 \leqslant i,j \leqslant n, r > 0\right\}$ subject just to the relations (1.1). Recall also that $T_{i,j}^{(0)} := \delta_{i,j}$. By the same proof as [MNO, Proposition 1.2], the following give an equivalent set of defining relations:

$$\left[T_{i,j}^{(r)}, T_{k,l}^{(s)}\right] = \sum_{t=0}^{\min(r,s)-1} \left(T_{k,j}^{(t)} T_{i,l}^{(r+s-1-t)} - T_{k,j}^{(r+s-1-t)} T_{i,l}^{(t)} \right) \tag{4.1}$$

for $1 \leqslant i,j,k,l \leqslant n$ and $r,s > 0$. Using this and induction on $r+s$, it is easy to see in particular that

$$T_{i,j}^{(r)} T_{i,j}^{(s)} = T_{i,j}^{(s)} T_{i,j}^{(r)} \tag{4.2}$$

for all $r,s \geqslant 0$.

We often put the generators $T_{i,j}^{(r)}$ for all $r \geqslant 0$ together to form the power series $T_{i,j}(u)$ as in (1.2). Then these power series for all $1 \leqslant i,j \leqslant n$ can be collected together into a single matrix $T(u) := (T_{i,j}(u))_{1 \leqslant i,j \leqslant n} \in M_n(Y_n[[u^{-1}]])$. Using these matrices, the relations can be expressed in an extremely compact form, known as the RTT presentation. Since this presentation will play no role in our work, we do not bother going through the details,

rather we refer the reader to [MNO, Proposition 1.8] where the calculations apply equally well over $\Bbbk$ as they do over $\mathbb{C}$.

It will not play a significant role in this paper, but we should mention that Y_n is a Hopf algebra. Its comultiplication Δ and antipode S are given by

$$\Delta(T_{i,j}(u)) = \sum_{k=1}^{n} T_{i,k}(u) \otimes T_{k,j}(u), \qquad S(T_{i,j}(u)) = \widetilde{T}_{i,j}(u), \tag{4.3}$$

where $\widetilde{T}_{i,j}(u)$ is the (i,j)-entry of the matrix $T(u)^{-1}$. The counit sends $T_{i,j}(u) \mapsto \delta_{i,j}$.

4.2. **Filtrations and the PBW theorem.** There are two distinguished filtrations on Y_n. The first one, called the *canonical filtration* in [BK1, §2], places $T_{i,j}^{(r)}$ in degree r. The associated graded algebra is commutative. Since this filtration will not play a significant role in this article, we will not reserve any special notation for it. In fact, will never mention it again after the proof of the following fundamental *PBW theorem*.

Theorem 4.1. *Ordered monomials in the generators* $\big\{T_{i,j}^{(r)} \bigm| 1 \leqslant i,j \leqslant n, r > 0\big\}$ *taken in any fixed order give a basis for* Y_n.

Proof. It is easy to see that these monomials span, e.g. one can argue by induction on degree in the canonical filtration. To show that they are linearly independent, we may assume that $\Bbbk$ is algebraically closed. Then the proof given in characteristic zero in [BK1, Theorem 3.1] works just as well in positive characteristic. □

The second filtration on Y_n, called *the loop filtration* in [BK1, §2], will be ubiquitous. We denote it by

$$Y_n = \bigcup_{r \geqslant 0} \mathrm{F}_r Y_n. \tag{4.4}$$

It is defined by placing $T_{i,j}^{(r)}$ in degree $r-1$, i.e. $\mathrm{F}_r Y_n$ is the span of all monomials of the form $T_{i_1,j_1}^{(r_1)} \cdots T_{i_m,j_m}^{(r_m)}$ with $(r_1-1)+\cdots+(r_m-1) \leqslant r$. We warn the reader that the notation $\mathrm{F}_r Y_n$ is often used elsewhere in the literature for the canonical filtration. Also (4.4) should not be confused with the filtration on $U(\mathfrak{g})$ from (3.6); the latter will never be used again.

To describe the associated graded algebra $\operatorname{gr} Y_n$, recall that $\mathfrak{g} = \mathfrak{gl}_n[t]$ is generated by $\{e_{i,j}t^r \mid 1 \leqslant i,j \leqslant n, r \geqslant 0\}$ subject to relations (3.2), and notice by the definition that $\operatorname{gr} Y_n$ is generated by elements $\big\{\operatorname{gr}_{r-1} T_{i,j}^{(r)} \bigm| 1 \leqslant i,j \leqslant n, r > 0\big\}$.

Lemma 4.2. *There is an isomorphism* $\chi : U(\mathfrak{g}) \xrightarrow{\sim} \operatorname{gr} Y_n$ *sending* $e_{i,j}t^r \mapsto \operatorname{gr}_r T_{i,j}^{(r+1)}$ *for each* $1 \leqslant i,j \leqslant n$ *and* $r \geqslant 0$.

Proof. Relation (4.1) implies that

$$\begin{aligned}
[\operatorname{gr}_r T_{i,j}^{(r+1)}, \operatorname{gr}_s T_{k,l}^{(s+1)}] &= [T_{i,j}^{(r+1)}, T_{k,l}^{(s+1)}] + \mathrm{F}_{r+s-1} Y_n \\
&= \delta_{k,j} T_{i,l}^{(r+s+1)} - \delta_{i,l} T_{k,j}^{(r+s+1)} + \mathrm{F}_{r+s-1} Y_n \\
&= \delta_{k,j} \operatorname{gr}_{r+s} T_{i,l}^{(r+s+1)} - \delta_{i,l} \operatorname{gr}_{r+s} T_{k,j}^{(r+s+1)}.
\end{aligned}$$

Comparing with (3.2), we deduce that the map in the statement of the lemma is well defined. To see that it is an isomorphism, one uses the PBW basis from Theorem 4.1 to see that a basis for $U(\mathfrak{g})$ is sent to a basis for $\operatorname{gr} Y_n$. □

4.3. **Drinfeld-type presentation.** Since the leading minors of the $n \times n$ matrix $T(u)$ are invertible, it possesses a Gauss factorization

$$T(u) = F(u)D(u)E(u) \tag{4.5}$$

for unique matrices

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

$$E(u) = \begin{pmatrix} 1 & E_{1,2}(u) & \cdots & E_{1,n}(u) \\ 0 & 1 & \cdots & E_{2,n}(u) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}, \; F(u) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ F_{1,2}(u) & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ F_{1,n}(u) & F_{2,n}(u) & \cdots & 1 \end{pmatrix}.$$

This defines power series

$$D_i(u) = \sum_{r \geqslant 0} D_i^{(r)} u^{-r}, \qquad E_{i,j}(u) = \sum_{r \geqslant 0} E_{i,j}^{(r)} u^{-r}, \qquad F_{i,j}(u) = \sum_{r \geqslant 0} F_{i,j}^{(r)} u^{-r}$$

in $Y_n[[u^{-1}]]$ with $D_i^{(0)} = 1$ and $E_{i,j}^{(0)} = F_{i,j}^{(0)} = 0$. Let

$$E_i(u) = \sum_{r \geqslant 0} E_i^{(r)} u^{-r} := E_{i,i+1}(u), \qquad F_i(u) = \sum_{r \geqslant 0} F_i^{(r)} u^{-r} := F_{i,i+1}(u)$$

for short. Also let

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

We warn the reader that our $\widetilde{D}_i(u)$ differs by a sign from one used in [BK1, BK2, BK3]. This accounts for several other sign differences in the exposition below compared to *loc. cit.*, e.g. we have that $\widetilde{D}_0^{(0)} = 1$.

In terms of quasi-determinants of [GR], we have the following more explicit descriptions, as noted already in [BK1, §5]:

$$D_i(u) = \begin{vmatrix} T_{1,1}(u) & \cdots & T_{1,i-1}(u) & T_{1,i}(u) \\ \vdots & \ddots & \vdots & \vdots \\ T_{i-1,1}(u) & \cdots & T_{i-1,i-1}(u) & T_{i-1,i}(u) \\ T_{i,1}(u) & \cdots & T_{i,i-1}(u) & \boxed{T_{i,i}(u)} \end{vmatrix}, \tag{4.6}$$

$$E_{i,j}(u) = \widetilde{D}_i(u) \begin{vmatrix} T_{1,1}(u) & \cdots & T_{1,i-1}(u) & T_{1,j}(u) \\ \vdots & \ddots & \vdots & \vdots \\ T_{i-1,1}(u) & \cdots & T_{i-1,i-1}(u) & T_{i-1,j}(u) \\ T_{i,1}(u) & \cdots & T_{i,i-1}(u) & \boxed{T_{i,j}(u)} \end{vmatrix}, \tag{4.7}$$

$$F_{i,j}(u) = \begin{vmatrix} T_{1,1}(u) & \cdots & T_{1,i-1}(u) & T_{1,i}(u) \\ \vdots & \ddots & \vdots & \vdots \\ T_{i-1,1}(u) & \cdots & T_{i-1,i-1}(u) & T_{i-1,i}(u) \\ T_{j,1}(u) & \cdots & T_{j,i-1}(u) & \boxed{T_{j,i}(u)} \end{vmatrix} \widetilde{D}_i(u). \tag{4.8}$$

Since $E_{j-1}^{(1)} = T_{j-1,j}^{(1)}$ and $F_{j-1}^{(1)} = T_{j,j-1}^{(1)}$, it follows easily that

$$E_{i,j}^{(r)} = \big[E_{i,j-1}^{(r)}, E_{j-1}^{(1)}\big], \qquad F_{i,j}^{(r)} = \big[F_{j-1}^{(1)}, F_{i,j-1}^{(r)}\big]. \tag{4.9}$$

for $i + 1 < j \leqslant n$.

Now we state the main theorem of the section, which is the modular analogue of [BK1, Theorem 5.2]. Although not written explicitly there in this form, this presentation must surely have been known to Drinfeld when writing [D].

Theorem 4.3. *The algebra Y_n is generated by the elements $\{D_i^{(r)} \mid 1 \leqslant i \leqslant n, r > 0\}$ and $\{E_i^{(r)}, F_i^{(r)} \mid 1 \leqslant i < n, r > 0\}$ subject only to the following relations:*

$$\big[D_i^{(r)}, D_j^{(s)}\big] = 0, \tag{4.10}$$

$$\big[E_i^{(r)}, F_j^{(s)}\big] = -\delta_{i,j} \sum_{t=0}^{r+s-1} D_{i+1}^{(r+s-1-t)} \widetilde{D}_i^{(t)}, \tag{4.11}$$

$$\big[D_i^{(r)}, E_j^{(s)}\big] = (\delta_{i,j} - \delta_{i,j+1}) \sum_{t=0}^{r-1} D_i^{(t)} E_j^{(r+s-1-t)}, \tag{4.12}$$

$$\big[D_i^{(r)}, F_j^{(s)}\big] = (\delta_{i,j+1} - \delta_{i,j}) \sum_{t=0}^{r-1} F_j^{(r+s-1-t)} D_i^{(t)}, \tag{4.13}$$

$$\big[E_i^{(r)}, E_i^{(s)}\big] = \sum_{t=r}^{s-1} E_i^{(t)} E_i^{(r+s-1-t)} \qquad \text{if } r < s, \tag{4.14}$$

$$\big[F_i^{(r)}, F_i^{(s)}\big] = \sum_{t=s}^{r-1} F_i^{(r+s-1-t)} F_i^{(t)} \qquad \text{if } r > s, \tag{4.15}$$

$$\big[E_i^{(r+1)}, E_{i+1}^{(s)}\big] - \big[E_i^{(r)}, E_{i+1}^{(s+1)}\big] = E_i^{(r)} E_{i+1}^{(s)}, \tag{4.16}$$

$$\big[F_i^{(r)}, F_{i+1}^{(s+1)}\big] - \big[F_i^{(r+1)}, F_{i+1}^{(s)}\big] = F_{i+1}^{(s)} F_i^{(r)}, \tag{4.17}$$

$$\big[E_i^{(r)}, E_j^{(s)}\big] = 0 \qquad \text{if } |i-j| > 1, \tag{4.18}$$

$$\big[F_i^{(r)}, F_j^{(s)}\big] = 0 \qquad \text{if } |i-j| > 1, \tag{4.19}$$

$$\Big[E_i^{(r)}, \big[E_i^{(s)}, E_j^{(t)}\big]\Big] + \Big[E_i^{(s)}, \big[E_i^{(r)}, E_j^{(t)}\big]\Big] = 0 \qquad \text{if } |i-j| = 1, r \neq s, \tag{4.20}$$

$$\Big[F_i^{(r)}, \big[F_i^{(s)}, F_j^{(t)}\big]\Big] + \Big[F_i^{(s)}, \big[F_i^{(r)}, F_j^{(t)}\big]\Big] = 0 \qquad \text{if } |i-j| = 1, r \neq s \tag{4.21}$$

$$\Big[E_i^{(r)}, \big[E_i^{(r)}, E_j^{(t)}\big]\Big] = 0 \qquad \text{if } |i-j| = 1, \tag{4.22}$$

$$\Big[F_i^{(r)}, \big[F_i^{(r)}, F_j^{(t)}\big]\Big] = 0 \qquad \text{if } |i-j| = 1, \tag{4.23}$$

for all admissible i, j, r, s, t. In these relations, we use the shorthands $D_i^{(0)} = \widetilde{D}_i^{(0)} := 1$, and the elements $\widetilde{D}_i^{(r)}$ for $r > 0$ are defined recursively by $\widetilde{D}_i^{(r)} := -\sum_{t=1}^{r} D_i^{(t)} \widetilde{D}_i^{(r-t)}$.

Remark 4.4. The relations (4.10)–(4.19) are the same as relations (5.9)–(5.18) of [BK1], however relations (5.19)–(5.20) of *loc. cit.* are expressed here as the four relations (4.20)–(4.23). This is essential in charactersitic 2.

Proof. This is very similar to the proof in the characteristic zero explained in [BK1]. We just give a brief account in order to highlight the minor differences.

To start with, we need to check that the relations (4.10)–(4.23) do indeed hold in Y_n. For all but the last two relations, this is carried out already in [BK1] over the ground field $\mathbb{C}$. The arguments there start from the power series form of the defining relation (1.1), namely, that

$$(u-v)[T_{i,j}(u), T_{k,l}(v)] = T_{k,j}(u) T_{i,l}(v) - T_{k,j}(v) T_{i,l}(u) \tag{4.24}$$

in $Y_n[[u^{-1}, v^{-1}]]$, then extract from this various relations satisfied by the power series $D_i(u), E_i(u)$ and $F_i(u)$; some of these are recorded in Lemma 4.6 below. Then the desired relations follow by computing various coefficients in these power series relations. These calculations can be performed without any difference in positive characteristic, yielding all of our relations except for (4.22)–(4.23). To establish (4.22), we use the power series relation

$$[E_i(u), [E_i(u), E_j(v)]] = 0 \tag{4.25}$$

for $|i-j| = 1$ which is proved by reducing first to the case $\{i, j\} = \{1, 2\}$ then arguing as in [BK1, Lemma 5.6]. Taking the $u^{-2r}v^{-t}$-coefficient in this identity and using also (4.20) gives (4.22). The proof of (4.23) is similar.

Now let $\widehat{Y}_n$ be the algebra with generators $\big\{D_i^{(r)}, \widetilde{D}_i^{(r)} \;\big|\; 1 \leqslant i \leqslant n, r \geqslant 0\big\}$ and $\big\{E_i^{(r)}, F_i^{(r)} \;\big|\; 1 \leqslant i < n, r > 0\big\}$ subject to all of the relations recorded in the statement of the theorem. The previous paragraph implies that there is a well-defined homomorphism $\theta : \widehat{Y}_n \to Y_n$ taking the generators of $\widehat{Y}_n$ to the elements of Y_n with the same names. Define higher root elements $E_{i,j}^{(r)}, F_{i,j}^{(r)} \in \widehat{Y}_n$ for all $1 \leqslant i < j \leqslant n$ and $r > 0$ by setting $E_{i,i+1}^{(r)} := E_i^{(r)}$ and $F_{i,i+1} := F_i^{(r)}$, then using the formula (4.9) inductively when $|j-i| > 1$. This definition ensures that θ sends $E_{i,j}^{(r)}, F_{i,j}^{(r)} \in \widehat{Y}_n$ to the elements of Y_n with the same names.

To complete the proof, we define a filtration $\widehat{Y}_n = \bigcup_{r \geqslant 0} \mathrm{F}_r \widehat{Y}_n$ by declaring that the elements $D_i^{(r)}, E_{i,j}^{(r)}$ and $F_{i,j}^{(r)}$ defined in the previous paragraph are of filtered degree $r-1$. We claim that there is a surjective graded algebra homomorphism

$$\psi : U(\mathfrak{g}) \twoheadrightarrow \operatorname{gr} \widehat{Y}_n, \quad e_{i,j}t^r \mapsto e_{i,j;r} := \begin{cases} \operatorname{gr}_r D_i^{(r+1)} & \text{if } i = j, \\ \operatorname{gr}_r E_{i,j}^{(r+1)} & \text{if } i < j, \\ \operatorname{gr}_r F_{j,i}^{(r+1)} & \text{if } i > j, \end{cases} \tag{4.26}$$

for $1 \leqslant i, j \leqslant n$ and $r \geqslant 0$. To see this, surjectivity is immediate from the way the filtration is defined, so it suffices to show that the image of the defining relation (3.2) for U holds in $\operatorname{gr} \widehat{Y}_n$, i.e. $[e_{i,j;r}, e_{k,l;s}] = \delta_{k,l} e_{i,l;r+s} - \delta_{i,l} e_{k,j;r+s}$ for $1 \leqslant i, j, k, l \leqslant n$ and $r, s \geqslant 0$. There are six cases: (a) $i = j, k = l$; (b) $i = j, k < l$; (c) $i = j, k > l$; (d) $i < j, k < l$; (e) $i > j, k > l$; (f) $i < j, k > l$. For these, (a) is immediate from (4.11), (c) is a similar argument to (b), and (e) is a similar argument to (d). So we just prove (b), (d) and (f). The proof of (d) is the same argument as in the proof of [BK1, Lemma 5.8] except when $r = s$ and $|i-j| = 1$, when one must replace the relation $[e_{i,i+1;r}, [e_{i,i+1;s}, e_{j,j+1;t}]] = -[e_{i,i+1;s}, [e_{i,i+1;r}, e_{j,j+1;t}]]$ with $[e_{i,i+1;r}, [e_{i,i+1;r}, e_{j,j+1;s}]] = 0$, which follows straight from (4.22). Also (b) follows when $|k-l| = 1$ using (4.12); then it may be deduced in general using (d) and induction on $k-l$. The argument for (f) is similar: it follows when $|i-j| = |k-l| = 1$ using (4.11); then it follows in general using (d), (e) and induction.

Using (4.6)–(4.8), one sees that $D_i^{(r)}, E_{i,j}^{(r)}, F_{i,j}^{(r)} \in \mathrm{F}_{r-1} Y_n$, so that θ is a filtered homomorphism. Moreover, the following diagram commutes:

$$\begin{array}{ccc} & U(\mathfrak{g}) & \\ {\scriptstyle\psi}\swarrow & & \searrow{\scriptstyle\chi} \\ \operatorname{gr} \widehat{Y}_n & \xrightarrow{\operatorname{gr}\theta} & \operatorname{gr} Y_n. \end{array}$$

We already showed that χ is an isomorphism in Lemma 4.2. It follows that $\operatorname{gr}\theta$ is one too. Hence, θ is an isomorphism as required. $\square$

Henceforth, we will *identify* $U(\mathfrak{g})$ and $\operatorname{gr} Y_n$ via the isomorphism χ from Lemma 4.2, i.e. we identify $e_{i,j}t^r$ with $\operatorname{gr}_r T_{i,j}^{(r+1)}$. The proof of Theorem 4.3 shows moreover that $D_i^{(r+1)}, E_{i,j}^{(r+1)}$ and $F_{i,j}^{(r+1)}$ all belong to $\mathrm{F}_r Y_n$, and under our identification we have that

$$e_{i,j}t^r = \begin{cases} \operatorname{gr}_r D_i^{(r+1)} & \text{if } i = j, \\ \operatorname{gr}_r E_{i,j}^{(r+1)} & \text{if } i < j, \\ \operatorname{gr}_r F_{j,i}^{(r+1)} & \text{if } i > j. \end{cases} \tag{4.27}$$

Using this and the PBW theorem for $U(\mathfrak{g})$, we obtain the following giving a PBW basis for Y_n in terms of the Drinfeld generators.

Theorem 4.5. *Ordered monomials in the elements*

$$\big\{D_i^{(r)} \,\big|\, 1 \leqslant i \leqslant n, r > 0\big\} \cup \big\{E_{i,j}^{(r)}, F_{i,j}^{(r)} \,\big|\, 1 \leqslant i < j \leqslant n, r > 0\big\} \tag{4.28}$$

taken in any fixed ordering form a basis for Y_n.

We will need the power series forms of some of the relations from Theorem 4.3. We record these and some consequences next.

Lemma 4.6. *The following equalities hold in $Y_n[[u^{-1}, v^{-1}]]$:*

$$(u-v)[E_i(u), F_i(v)] = D_{i+1}(u)\widetilde{D}_i(u) - D_{i+1}(v)\widetilde{D}_i(v), \tag{4.29}$$

$$(u-v)[E_i(u), E_i(v)] = (E_i(v) - E_i(u))^2, \tag{4.30}$$

$$(u-v)[E_i(u), D_i(v)] = D_i(v)(E_i(u) - E_i(v)), \tag{4.31}$$

$$(u-v)[E_i(u), \widetilde{D}_{i+1}(v)] = (E_i(u) - E_i(v))\widetilde{D}_{i+1}(v), \tag{4.32}$$

$$(u-v)[E_i(u), \widetilde{D}_i(v)] = (E_i(v) - E_i(u))\widetilde{D}_i(v), \tag{4.33}$$

$$(u-v)[E_i(u), D_{i+1}(v)] = D_{i+1}(v)(E_i(v) - E_i(u)). \tag{4.34}$$

Proof. Equations (4.29)–(4.32) were proven over $\mathbb{C}$ in [BK1, Lemma 5.4]; the same proof works here. Then (4.33)–(4.34) follow from (4.31)–(4.32) using $D_j(v)\widetilde{D}_j(v) = 1$. □

Corollary 4.7. *The following hold in $Y_n[[u^{-1}]]$:*

$$E_i(u-1)D_i(u) = D_i(u)E_i(u), \qquad \widetilde{D}_i(u)E_i(u-1) = E_i(u)\widetilde{D}_i(u), \tag{4.35}$$

$$D_{i+1}(u)E_i(u) = E_i(i+1)D_{i+1}(u), \qquad E_i(u)\widetilde{D}_{i+1}(u) = \widetilde{D}_{i+1}(u)E_i(i+1). \tag{4.36}$$

Proof. These follow from the identities in the previous lemma by specializing v. For example, to get the first relation in (4.35), set $v := u+1$ in (4.31), simplify, then replace u by $u-1$. □

Lemma 4.8. *The following relations hold in $Y_n[[u^{-1}, v^{-1}]]$ for all $m \geqslant 0$:*

$$(u-v)[E_i(u), (E_i(v) - E_i(u))^m] = m(E_i(v) - E_i(u))^{m+1}, \tag{4.37}$$

$$(u-v)[E_i(u), D_i(v)(E_i(v) - E_i(u))^m] = (m-1)D_i(v)(E_i(v) - E_i(u))^{m+1}, \tag{4.38}$$

$$(u-v)[E_i(u), D_{i+1}(v)(E_i(v) - E_i(u))^m] = (m+1)D_{i+1}(v)(E_i(v) - E_i(u))^{m+1}, \tag{4.39}$$

$$(u-v)[E_i(u), D_{i+1}(v)(E_i(v) - E_i(u))^m \widetilde{D}_i(v)] = (m+2)D_{i+1}(v)(E_i(v) - E_i(u))^{m+1}\widetilde{D}_i(v). \tag{4.40}$$

Proof. The relation (4.37) follows from (4.30) and the Leibniz rule. Then (4.38)–(4.40) follow from (4.37), (4.31), (4.33) and (4.34) using Leibniz again. □

The following relations are closely related to the ones in Lemma 4.8, but it is easier to prove them from scratch.

Lemma 4.9. *For any $i = 1, \dots, n-1$, $m \geqslant 0$ and $r, s > 0$, we have that*

$$\left[E_i^{(r)}, \sum_{\substack{s_1,\dots,s_m \geqslant r \\ s_1+\dots+s_m=(m-1)(r-1)+s}} E_i^{(s_1)} \cdots E_i^{(s_m)}\right] = m \sum_{\substack{s_1,\dots,s_{m+1} \geqslant r \\ s_1+\dots+s_{m+1}=m(r-1)+s}} E_i^{(s_1)} \cdots E_i^{(s_{m+1})}, \tag{4.41}$$

$$\left[E_i^{(r)}, \sum_{\substack{s_1,\dots,s_m \leqslant r-1 \\ s_1+\dots+s_m=(m-1)(r-1)+s}} E_i^{(s_1)} \cdots E_i^{(s_m)}\right] = -m \sum_{\substack{s_1,\dots,s_{m+1} \leqslant r-1 \\ s_1+\dots+s_{m+1}=m(r-1)+s}} E_i^{(s_1)} \cdots E_i^{(s_{m+1})}, \tag{4.42}$$

$$\left[E_i^{(r)}, \sum_{\substack{s_1,\dots,s_m \geqslant r, t \geqslant 0 \\ s_1+\dots+s_m+t=m(r-1)+s}} D_i^{(t)} E_i^{(s_1)} \cdots E_i^{(s_m)}\right] = (m-1) \sum_{\substack{s_1,\dots,s_{m+1} \geqslant r, t \geqslant 0 \\ s_1+\dots+s_{m+1}+t=(m+1)(r-1)+s}} D_i^{(t)} E_i^{(s_1)} \cdots E_i^{(s_{m+1})}, \tag{4.43}$$

$$\left[E_i^{(r)}, \sum_{\substack{s_1,\dots,s_m \geqslant r, t \geqslant 0 \\ s_1+\dots+s_m+t=m(r-1)+s}} D_{i+1}^{(t)} E_i^{(s_1)} \cdots E_i^{(s_m)}\right] = (m+1) \sum_{\substack{s_1,\dots,s_{m+1} \geqslant r, t \geqslant 0 \\ s_1+\dots+s_{m+1}+t=(m+1)(r-1)+s}} D_{i+1}^{(t)} E_i^{(s_1)} \cdots E_i^{(s_{m+1})}, \tag{4.44}$$

$$\left[E_i^{(r)}, \sum_{\substack{s_1,\dots,s_m \geqslant r, t \geqslant 0, u \geqslant 0 \\ s_1+\dots+s_m+t+u=m(r-1)+s}} D_{i+1}^{(t)} E_i^{(s_1)} \cdots E_i^{(s_m)} \tilde{D}_i^{(u)}\right] = (m+2) \sum_{\substack{s_1,\dots,s_{m+1} \geqslant r, t \geqslant 0, u \geqslant 0 \\ s_1+\dots+s_{m+1}+t+u=(m+1)(r-1)+s}} D_{i+1}^{(t)} E_i^{(s_1)} \cdots E_i^{(s_{m+1})} \tilde{D}_i^{(u)}. \tag{4.45}$$

Proof. For (4.42), we note in the summation on the left that $s_1, \dots, s_m \leqslant r-1$ and $s_1 + \dots + s_m = (m-1)(r-1)+s$ implies that $r-1 \geqslant s_1, \dots, s_m \geqslant s > 0$, so the expression makes sense. By (4.14), we have that

$$[E_i^{(r)}, E_i^{(s_k)}] = - \sum_{\substack{s_k', s_k'' \leqslant r-1 \\ s_k'+s_k''=s_k+r-1}} E_i^{(s_k')} E_i^{(s_k'')}$$

for $0 < s_k \leqslant r-1$. Using this and the Leibniz rule, we deduce that the left hand side of (4.42) equals

$$-\sum_{k=1}^{m} \sum_{\substack{s_1,\dots,s_m \leqslant r-1 \\ s_1+\dots+s_m=(m-1)(r-1)+s}} \sum_{\substack{s_k', s_k'' \leqslant r-1 \\ s_k'+s_k''=s_k+r-1}} E_i^{(s_1)} \cdots E_i^{(s_{k-1})} E_i^{(s_k')} E_i^{(s_k'')} E_i^{(s_{k+1})} \cdots E_i^{(s_m)}$$

$$= -\sum_{k=1}^{m} \sum_{\substack{s_1,\dots,s_{m+1} \leqslant r-1 \\ s_1+\dots+s_{m+1}=m(r-1)+s}} E_i^{(s_1)} \cdots E_i^{(s_{m+1})},$$

which gives the right hand side of (4.42). The proof of (4.41) is similar. Then (4.43)–(4.44) follow from (4.12) and (4.41) with one more application of Leibniz. Finally (4.45) follows similarly from (4.44) together with

$$\left[E_i^{(r)}, \tilde{D}_i^{(s)}\right] = \sum_{t=0}^{s-1} E_i^{(r+s-1-t)} \tilde{D}_i^{(t)}. \tag{4.46}$$

This may be deduced from (4.33) by dividing by $(u-v)$ then equating coefficients in exactly the manner explained in the paragraph following [BK1, (5.23)]. □

The relations in the next lemma involve the expressions

$$D_{i\downarrow m}(u) := D_i(u)D_i(u-1)\cdots D_i(u-m+1), \tag{4.47}$$
$$D_{i\uparrow m}(u) := D_i(u)D_i(u+1)\cdots D_i(u+m-1). \tag{4.48}$$

In view of (4.10), the order of the products on the right hand sides here is irrelevant.

Lemma 4.10. *The following relations hold for all $m \geqslant 1$:*

$$(u-v)[D_{i\downarrow m}(u), E_i(v)] = mD_{i\downarrow m}(u)(E_i(v) - E_i(u)), \tag{4.49}$$
$$(u-v)[D_{i\uparrow m}(u), E_{i-1}(v)] = mD_{i\uparrow m}(u)(E_{i-1}(u) - E_{i-1}(v)). \tag{4.50}$$

Proof. For (4.49), we actually prove it in the following equivalent form:

$$(u-v-m)D_{i\downarrow m}(u)E_i(v) = (u-v)E_i(v)D_{i\downarrow m}(u) - mD_{i\downarrow m}(u)E_i(u). \tag{4.51}$$

This follows when $m = 1$ from (4.31). To prove (4.51) in general, proceed by induction: given (4.51) for some $m \geqslant 1$, multiply both sides on the left by $(u-v-m-1)D_i(u-m)$ then simplify using the $m = 1$ case already proved plus (4.35) to obtain the analogous formula with m replaced by $m+1$.

The proof of (4.50) is similar. One actually shows equivalently that

$$(u-v+m)D_{i\uparrow m}(u)E_{i-1}(v) = (u-v)E_{i-1}(v)D_{i\uparrow m}(u) + mD_{i\uparrow m}(u)E_{i-1}(u). \tag{4.52}$$

This follows when $m = 1$ from (4.34), then the general case follows by a similar induction to the previous paragraph. □

For our final relations, we let

$$H_i(u) = \sum_{r\geqslant 0} H_i^{(r)}u^{-r} := -D_{i+1}(u)\widetilde{D}_i(u) \tag{4.53}$$

assuming $1 \leqslant i < n$. In particular, $H_i^{(0)} = -1$. It is straightforward to see from (4.27) that $H_i^{(r+1)} \in \mathrm{F}_r Y_n$ and

$$\mathrm{gr}_r\, H_i^{(r+1)} = e_{i,i}t^r - e_{i+1,i+1}t^r. \tag{4.54}$$

Note also by Corollary 4.7 that

$$H_i(u)E_i(u-1) = E_i(u+1)H_i(u). \tag{4.55}$$

Lemma 4.11. *The following relations hold in $Y_n[[u^{-1}, v^{-1}]]$:*

$$(u-v-1)[H_i(u), E_i(v)] = 2H_i(u)(E_i(u-1) - E_i(v)), \tag{4.56}$$
$$(u-v+1)[H_i(u), E_i(v)] = 2(E_i(u+1) - E_i(v))H_i(u), \tag{4.57}$$
$$(u-v)[H_{i-1}(u), E_i(v)] = -H_{i-1}(u)(E_i(u) - E_i(v)), \tag{4.58}$$
$$(u-v-1)[H_{i-1}(u), E_i(v)] = -(E_i(u-1) - E_i(v))H_{i-1}(u), \tag{4.59}$$
$$(u-v)[H_{i+1}(u), E_i(v)] = -(E_i(u) - E_i(v))H_{i+1}(u), \tag{4.60}$$
$$(u-v+1)[H_{i+1}(u), E_i(v)] = -H_{i+1}(u)(E_i(u+1) - E_i(v)). \tag{4.61}$$

Proof. We just go through (4.58), (4.61) and (4.56), since the other three are similar. The identity (4.58) follows easily from (4.31) and Leibniz, using that $\widetilde{D}_{i-1}(v)$ commutes with $E_i(u)$ by (4.10). For (4.61), we have by (4.32) that

$$(u-v-1)E_i(u)\widetilde{D}_{i+1}(v) = (u-v)\widetilde{D}_{i+1}(v)E_i(u) - E_i(v)\widetilde{D}_{i+1}(v).$$

Using also (4.36) gives that

$$(u-v-1)[E_i(u), \widetilde{D}_{i+1}(v)] = \widetilde{D}_{i+1}(v)E_i(u) - \widetilde{D}_{i+1}(v)E_i(v+1).$$

Now multiply by $D_{i+2}(v)$ (which commutes with $E_i(u)$) then interchange u and v to get the conclusion. Finally, for (4.56), we have by (4.40) that

$$(u-v)[E_i(u), D_{i+1}(v)\widetilde{D}_i(v)] = 2D_{i+1}(v)(E_i(v) - E_i(u))\widetilde{D}_i(v). \tag{4.62}$$

From (4.33), we have that

$$(u-v+1)E_i(u)\widetilde{D}_i(v) = (u-v)\widetilde{D}_i(v)E_i(u) + \widetilde{D}_i(v)E_i(v-1).$$

Then we multiply (4.62) by $(u-v+1)$ and use these identities plus (4.35) to obtain

$$(u-v)(u-v+1)[E_i(u), H_i(v)] = 2(u-v)H_i(v)(E_i(v-1) - E_i(u)).$$

Dividing through by $(u-v)$ and interchanging u and v gives the result. □

Corollary 4.12. *The following hold in* $Y_n[[u^{-1}, v^{-1}]]$*:*

$$(u-v)[H_i(u), E_i(v)] = -H_i(u)E_i(v) - E_i(v)H_i(u) + 2H_i(u)E_i(u-1), \tag{4.63}$$

$$\begin{aligned}(u-v)\Big[H_{i-1}\big(u+\tfrac{1}{2}\big), E_i(v)\Big] &= \tfrac{1}{2}\Big(H_{i-1}(u+\tfrac{1}{2})E_i(v) + E_i(v)H_{i-1}\big(u+\tfrac{1}{2}\big)\Big)\\ &\quad - \tfrac{1}{2}\Big(H_{i-1}\big(u+\tfrac{1}{2}\big)E_i\big(u+\tfrac{1}{2}\big) + E_i\big(u-\tfrac{1}{2}\big)H_{i-1}\big(u+\tfrac{1}{2}\big)\Big),\end{aligned} \tag{4.64}$$

$$\begin{aligned}(u-v)\Big[H_{i+1}\big(u-\tfrac{1}{2}\big), E_i(v)\Big] &= \tfrac{1}{2}\Big(H_{i+1}(u-\tfrac{1}{2})E_i(v) + E_i(v)H_{i+1}\big(u-\tfrac{1}{2}\big)\Big)\\ &\quad - \tfrac{1}{2}\Big(H_{i+1}\big(u-\tfrac{1}{2}\big)E_i\big(u+\tfrac{1}{2}\big) + E_i\big(u-\tfrac{1}{2}\big)H_{i+1}\big(u-\tfrac{1}{2}\big)\Big),\end{aligned} \tag{4.65}$$

assuming that $\operatorname{char}\Bbbk \neq 2$ *for the last two (so that* $\frac{1}{2}$ *makes sense).*

Proof. When $\operatorname{char}\Bbbk \neq 2$, these follow by averaging the corresponding pairs identities from Lemma 4.11, e.g. (4.64) is ((4.58)+(4.59))/2 with u replaced by $u+\frac{1}{2}$. For (4.63), one also needs to use (4.55). To establish (4.63) when $\operatorname{char}\Bbbk = 2$, we observe by (4.56) that $\big[H_i^{(r)}, E_i^{(s)}\big] = 0$ for all $r, s > 0$, which easily implies the desired identity. □

4.4. **The modular shifted Yangian.** Pick a shift matrix σ as in §3.1. The *shifted Yangian* is the subalgebra $Y_n(\sigma) \subseteq Y_n$ generated by the following elements:

$$\big\{D_i^{(r)} \;\big|\; 1 \leqslant i \leqslant n, r > 0\big\} \cup \big\{E_i^{(r)}, F_i^{(s)} \;\big|\; 1 \leqslant i < n, r > s_{i,i+1}, s > s_{i+1,i}\big\} \tag{4.66}$$

Notice that when σ is the zero matrix we have $Y_n(\sigma) = Y_n$. Neither the PBW basis nor the centre of $Y_n(\sigma)$ can be described without introducing higher root elements, however the elements $E_{i,j}^{(r)} \in Y_n$ do not lie in the subalgebra $Y_n(\sigma)$ for a general shift matrix. Instead, following [BK2, (2.18)–(2.19)], we recursively define

$${}^{\sigma}E_{i,j}^{(r)} := \Big[{}^{\sigma}E_{i,j-1}^{(r-s_{j-1,j})}, E_{j-1}^{(s_{j-1,j}+1)}\Big], \qquad {}^{\sigma}F_{i,j}^{(s)} := \Big[F_{j-1}^{(s_{j,j-1}+1)}, {}^{\sigma}F_{i,j-1}^{(s-s_{j,j-1})}\Big] \tag{4.67}$$

for $1 \leqslant i < j \leqslant n$ and $r > s_{i,j}, s > s_{j,i}$. As always, the filtration on Y_n induces a filtration on its subalgebra $Y_n(\sigma)$ so that $\operatorname{gr} Y_n(\sigma) \hookrightarrow \operatorname{gr} Y_n$.

Lemma 4.13. *For any shift matrix* σ*,* $1 \leqslant i < j \leqslant n$*, and* $r \geqslant s_{i,j}, s \geqslant s_{j,i}$*, we have that* ${}^{\sigma}E_{i,j}^{(r+1)} \in \mathrm{F}_r Y_n$ *and* ${}^{\sigma}F_{i,j}^{(s+1)} \in \mathrm{F}_s Y_n$*. Moreover, recalling that* $\operatorname{gr} Y_n$ *is identified with* $U(\mathfrak{g})$*, we have that*

$$\operatorname{gr}_r {}^{\sigma}E_{i,j}^{(r+1)} = e_{i,j}t^r, \qquad \operatorname{gr}_s {}^{\sigma}F_{i,j}^{(s+1)} = e_{j,i}t^s. \tag{4.68}$$

Hence, $\operatorname{gr} Y_n(\sigma)$ *is identified with the subalgebra* $U(\mathfrak{g}_\sigma)$ *of* $U(\mathfrak{g})$*.*

Proof. We just prove the statements about ${}^\sigma E_{i,j}^{(r+1)}$. When $j = i+1$ the result is immediate from (4.27). To deduce it in general, we proceed by induction on $j-i$. For the induction step, ${}^\sigma E_{i,j}^{(r+1)} = \left[{}^\sigma E_{i,j-1}^{(r+1-s_{j-1,j})}, E_{j-1}^{(s_{j-1,j}+1)}\right]$ so by induction it lies in $\mathrm{F}_{r-s_{j-1,j}+s_{j-1,j}} Y_n$ as required. Moreover, by induction again, its image in the associated graded algebra is

$$[e_{i,j-1} t^{r-s_{j-1,j}}, e_{j-1,j} t^{s_{j-1,j}}] = e_{i,j} t^r$$

using (3.2). □

From this and the PBW theorem for $U(\mathfrak{g}_\sigma)$, we also get the following PBW theorem for $Y_n(\sigma)$:

Theorem 4.14. *Ordered monomials in the elements*

$$\left\{D_i^{(r)} \;\middle|\; 1 \leqslant i \leqslant n, r > 0\right\} \cup \left\{{}^\sigma E_{i,j}^{(r)}, {}^\sigma F_{i,j}^{(s)} \;\middle|\; 1 \leqslant i < j \leqslant n, r > s_{i,j}, s > s_{j,i}\right\} \tag{4.69}$$

taken in any fixed ordering form a basis for $Y_n(\sigma)$.

We have defined $Y_n(\sigma)$ as a subalgebra of Y_n. It can also be defined by generators and relations: the following theorem shows that it has its own Drinfeld presentation.

Theorem 4.15. *The shifted Yangian $Y_n(\sigma)$ is generated by the elements* (4.66) *subject to the relations (4.10)–(4.23), interpreting "admissible i, j, r, s, t" so that the left hand sides of these relations only involve generators of $Y_n(\sigma)$.*

Proof. Let $\widehat{Y}_n(\sigma)$ be the algebra defined by these generators and relations. Since all of the relations hold in $Y_n(\sigma)$, there is a homomorphism $\theta : \widehat{Y}_n(\sigma) \to Y_n(\sigma)$ taking the generators to the elements of $Y_n(\sigma)$ with the same name. Introduce higher root elements ${}^\sigma E_{i,j}^{(r)}$ and ${}^\sigma F_{i,j}^{(r)}$ in $\widehat{Y}_n(\sigma)$ by repeating (4.67), so that θ takes these to the elements of $Y_n(\sigma)$ the same name too. Then define a filtration on $\widehat{Y}_n(\sigma)$ by placing $D_i^{(r)}, {}^\sigma E_{i,j}^{(r)}$ and ${}^\sigma F_{i,j}^{(r)}$ into filtered degree $r-1$. Arguing in the same way as the paragraph following (4.26), we see that there is a surjective graded algebra homomorphism

$$\psi : U(\mathfrak{g}_\sigma) \twoheadrightarrow \operatorname{gr} \widehat{Y}_n(\sigma), \quad e_{i,j} t^r \mapsto \begin{cases} \operatorname{gr}_r D_i^{(r+1)} & \text{if } i = j, \\ \operatorname{gr}_r E_{i,j}^{(r+1)} & \text{if } i < j, \\ \operatorname{gr}_r F_{j,i}^{(r+1)} & \text{if } i > j, \end{cases}$$

for $1 \leqslant i, j \leqslant n$ and $r \geqslant s_{i,j}$. Finally, we observe that θ is filtered, and

$$(\operatorname{gr} \theta) \circ \psi : U(\mathfrak{g}_\sigma) \to \operatorname{gr} Y_n(\sigma)$$

is the identity by Lemma 4.13. This implies that $\operatorname{gr} \theta$ is an isomorphism, hence, so is θ. □

4.5. **Automorphisms.** In the next section, we will also need to exploit various automorphisms/isomorphisms of Yangians and shifted Yangians. We briefly list the ones that we need below. In all cases, existence follows easily from the defining relations; see also [BK1, §2] and [BK2, (2.16)–(2.17)].

(1) ("Translation") For $c \in \Bbbk$, there is an automorphism $\eta_c : Y_n \to Y_n$ defined from $\eta_c(T_{i,j}(u)) = T_{i,j}(u-c)$, i.e. $\eta_c(T_{i,j}^{(r)}) = \sum_{s=1}^r \binom{r-1}{r-s} c^{r-s} T_{i,j}^{(s)}$. In terms of Drinfeld generators, η_c sends $D_i(u) \mapsto D_i(u-c), E_{i,j}(u) \mapsto E_{i,j}(u-c)$ and $F_{i,j}(u) \mapsto F_{i,j}(u-c)$, from which one sees that η_c does not leave $Y_n(\sigma)$ invariant in general.

(2) ("Multiplication by a power series") For any power series $f(u) \in 1 + u^{-1}\Bbbk[[u^{-1}]]$, there is an automorphism $\mu_f : Y_n \to Y_n$ defined from $\mu_f(T_{i,j}(u)) = f(u) T_{i,j}(u)$, i.e. $\mu_f(T_{i,j}^{(r)}) = \sum_{s=0}^r a_s T_{i,j}^{(r-s)}$ if $f(u) = \sum_{s \geqslant 0} a_s u^{-s}$. On Drinfeld generators, we have that $\mu_f(D_i(u)) = f(u) D_i(u)$, $\mu_f(E_i(u)) = E_i(u)$ and $\mu_f(F_i(u)) = F_i(u)$. So

this time μ_f restricts to an automorphism of each shifted Yangian $Y_n(\sigma)$ fixing the off-diagonal generators (4.67).

(3) ("Transposition") Let $\tau : Y_n \to Y_n$ be the anti-automorphism defined by $\tau(T_{i,j}^{(r)}) := T_{j,i}^{(r)}$ or, on Drinfeld generators, $\tau(D_i^{(r)}) = D_i^{(r)}, \tau(E_{i,j}^{(r)}) = F_{i,j}^{(r)}, \tau(F_{i,j}^{(r)}) = E_{i,j}^{(r)}$. This restricts to define an anti-isomorphism $\tau : Y_n(\sigma) \to Y_n(\sigma^T)$ between shifted Yangians, where σ^T is the transposed shift matrix.

(4) ("Change of shift matrix") Suppose that σ is a shift matrix as usual and $\dot{\sigma} = (\dot{s}_{i,j})_{1\leqslant i,j\leqslant n}$ is another shift matrix satisfying $s_{i,i+1} + s_{i+1,i} = \dot{s}_{i,i+1} + \dot{s}_{i+1,i}$ for all $1 \leqslant i < n$. Then, as a consequence of Theorem 4.15, there is an isomorphism $\iota : Y_n(\sigma) \xrightarrow{\sim} Y_n(\dot{\sigma})$, $D_i^{(r)} \mapsto D_i^{(r)}$, ${}^{\sigma}E_{i,j}^{(r)} \mapsto {}^{\dot{\sigma}}E_{i,j}^{(r-s_{i,j}+\dot{s}_{i,j})}$, ${}^{\sigma}F_{i,j}^{(r)} \mapsto {}^{\dot{\sigma}}F_{i,j}^{(r-s_{j,i}+\dot{s}_{j,i})}$.

(5) ("Permutation") For each $w \in \mathfrak{S}_n$, there is an automorphism $w : Y_n \to Y_n$ sending $T_{i,j}^{(r)} \mapsto T_{w(i),w(j)}^{(r)}$. This is clear from the RTT presentation. It does not leave the subalgebra $Y_n(\sigma)$ invariant in general.

Lemma 4.16. *For any $1 \leqslant i < j \leqslant n$, the permutation automorphism of Y_n defined by the transposition $(i+1\ j)$ maps $E_i(u) \mapsto E_{i,j}(u)$ and $F_i(u) \mapsto F_{i,j}(u)$.*

Proof. Since $E_i(u) = E_{i,i+1}(u)$ and $F_i(u) = F_{i,i+1}(u)$, this follows from (4.6)–(4.8). □

5. Centres of Y_n and $Y_n(\sigma)$

In this section, we describe the centre of the modular shifted Yangian, giving precise formulas for the generators. Unlike the previous section, most of the results presented here are not analogues of statements regarding the shifted Yangian defined over the complex numbers.

5.1. Harish–Chandra centre.

We define a power series by the rule

$$C(u) = \sum_{r\geqslant 0} C^{(r)} u^{-r} := D_1(u)D_2(u-1)\cdots D_n(u-n+1) \in Y_n[[u^{-1}]]. \tag{5.1}$$

The algebra generated by the coefficients $\left\{C^{(r)} \;\middle|\; r > 0\right\}$ will be denoted $Z_{\mathrm{HC}}(Y_n)$. We call it the *Harish-Chandra centre* of Y_n. The following theorem shows that the associated graded algebra $\operatorname{gr} Z_{\mathrm{HC}}(Y_n)$ is identified with $\Bbbk[z_r \mid r \geqslant 0] \subseteq Z(\mathfrak{g})$.

Theorem 5.1. *The elements $C^{(r)}$ lie in the centre of Y_n. Furthermore, we have that $C^{(r+1)} \in \mathrm{F}_r Y_n$ and*

$$\operatorname{gr}_r C^{(r+1)} = z_r \in U(\mathfrak{g}). \tag{5.2}$$

Hence, $C^{(1)}, C^{(2)}, C^{(3)}, \ldots$ are algebraically independent.

Proof. To prove that $C^{(r)}$ is central, using (4.10) and the anti-automorphism τ from §4.5, we are reduced to checking that $[E_i^{(r)}, C^{(s)}] = 0$ for all i, r, s. This can be proven in the same manner as [BK1, Theorem 7.2] using the power series relations (4.31) and (4.34). The second claim is noted in the proof of [BK1, Theorem 7.2]. It can also be deduced using Lemma 2.5, taking $A_{\mathrm{II}} := Y_n$ and $X_i^{(r)} := D_i^{(r)}$ so that $d_r = r - 1$ for $r > 0$, and noting that every $r > 1$ is optimal in this situation by Lemma 2.3 with $m = 1$. The final assertion follows because $z_0, z_1, \ldots$ are algebraically independent in $\operatorname{gr} Y_n$. □

Notice also that $Z_{\mathrm{HC}}(Y_n) \subseteq Z(Y_n(\sigma))$ for any choice of shift matrix and so we may also denote it $Z_{\mathrm{HC}}(Y_n(\sigma))$ and call it the *Harish-Chandra centre* of $Y_n(\sigma)$.

5.2. **Off-diagonal p-central elements.** We are ready to exhibit our first p-central elements. In this subsection, we investigate the ones that lie in the "root subalgebras" $Y^{\pm}_{i,j} \subset Y_n$ for $1 \leqslant i < j \leqslant n$, that is, the subalgebras generated by $\{E^{(r)}_{i,j} \mid r > 0\}$ and $\{F^{(r)}_{i,j} \mid r > 0\}$, respectively. In fact, we will give two different expressions for central elements in $Y^{\pm}_{i,j}$ in the first two lemmas. The first of these involving power series produces more complicated central elements, but this form often seems to be useful in practice.

Lemma 5.2. *For $1 \leqslant i < j \leqslant n$, all coefficients in the power series $(E_{i,j}(u))^p$ and $(F_{i,j}(u))^p$ belong to $Z(Y_n)$.*

Proof. Using Lemma 4.16 plus the anti-automorphism τ, the proof of this reduces to checking that the coefficients of $(E_i(u))^p$ are central in Y_n for each $i = 1, \dots, n-1$. To prove this, using Lemma 1.2, it suffices to establish the following identities in $Y_n[[u^{-1}, v^{-1}]]$ for all admissible j:

$$(\operatorname{ad} E_i(u))^p(E_j(v)) = 0, \tag{5.3}$$

$$(\operatorname{ad} E_i(u))^p(D_j(v)) = 0, \tag{5.4}$$

$$(\operatorname{ad} E_i(u))^p(F_j(v)) = 0. \tag{5.5}$$

In this paragraph we check (5.3). To show that $[E_i(u)^p, E_i(v)] = 0$, we use (4.30) and (4.37) repeatedly:

$$\begin{aligned}(u-v)^p(\operatorname{ad} E_i(u))^{p-1}([E_i(u), E_i(v)]) &= (u-v)^{p-1}(\operatorname{ad} E_i(u))^{p-1}((E_i(v) - E_i(u))^2)\\ &= (u-v)^{p-2}(\operatorname{ad} E_i(u))^{p-2}(2(E_i(v) - E_i(u))^3)\\ &= \cdots = p!(E_i(v) - E_i(u))^{p+1} = 0.\end{aligned}$$

Dividing by $(u-v)^p$ (which we may do since $Y_n[[u^{-1}, v^{-1}]][u, v]$ has no zero divisors) gives the desired identity. To see that $(\operatorname{ad} E_i(u))^p(E_j(v)) = 0$ when $|i-j| = 1$, we actually already have that $(\operatorname{ad} E_i(u))^2(E_j(v)) = 0$ by (4.25). When $|i-j| > 1$, the identity is clear because $[E_i(u), E_j(v)] = 0$ by (4.18).

For (5.4), it is immediate from (4.12) if $j < i$ or $j > i+1$. For the case $j = i$, we have by (4.31) and (4.38) with $k = 1$ that $(u-v)^2(\operatorname{ad} E_i(u))^2(D_i(v)) = 0$. Hence, on dividing by $(u-v)^2$, we get that $(\operatorname{ad} E_i(u))^p(D_i(v)) = 0$. Finally, when $j = i+1$, we use (4.34) and (4.39) repeatedly:

$$\begin{aligned}(u-v)^p(\operatorname{ad} E_i(u))^p(D_{i+1}(v)) &= (u-v)^{p-1}(\operatorname{ad} E_i(u))^{p-1}(D_{i+1}(v)(E_i(v) - E_i(u)))\\ &= (u-v)^{p-2}(\operatorname{ad} E_i(u))^{p-2}(2(E_i(v) - E_i(u))^2)\\ &= \cdots = p!(E_i(v) - E_i(u))^p = 0.\end{aligned}$$

Dividing by $(u-v)^p$ completes the proof of (5.4).

Finally, for (5.5), it follows when $i \neq j$ immediately from (4.11). When $i = j$, we observe using (4.40) repeatedly that

$$\begin{aligned}&(u-v)^{p-1}(\operatorname{ad} E_i(u))^{p-1}(D_{i+1}(v)\widetilde{D}_i(v))\\ &\qquad= (u-v)^{p-2}(\operatorname{ad} E_i(u))^{p-2}(2D_{i+1}(v)(E_i(v) - E_i(u))\widetilde{D}_i(v))\\ &\qquad= \cdots = p!D_{i+1}(v)(E_i(v) - E_i(u))^{p-1}\widetilde{D}_i(v) = 0.\end{aligned}$$

Hence, $(\operatorname{ad} E_i(u))^{p-1}(D_{i+1}(v)\widetilde{D}_i(v)) = 0$. We can also set $v = u$ in this identity to see that $(\operatorname{ad} E_i(u))^{p-1}(D_{i+1}(u)\widetilde{D}_i(u)) = 0$. Then using (4.29) we get the conclusion:

$$(u-v)(\operatorname{ad} E_i(u))^p(F_i(v)) = (\operatorname{ad} E_i(u))^{p-1}(D_{i+1}(u)\widetilde{D}_i(u) - D_{i+1}(v)\widetilde{D}_i(v)) = 0.$$

□

Lemma 5.3. *For $1 \leqslant i < j \leqslant n$ and $r > 0$, we have that $\big(E_{i,j}^{(r)}\big)^p, \big(F_{i,j}^{(r)}\big)^p \in Z(Y_n)$.*

Proof. Using Lemmas 4.16 and 1.2, this reduces to checking

$$\big(\operatorname{ad} E_i^{(r)}\big)^p \left(E_j^{(s)}\right) = 0, \tag{5.6}$$

$$\big(\operatorname{ad} E_i^{(r)}\big)^p \left(D_j^{(s)}\right) = 0, \tag{5.7}$$

$$\big(\operatorname{ad} E_i^{(r)}\big)^p \left(F_j^{(s)}\right) = 0. \tag{5.8}$$

These may all be proved in a very similar way to (5.3)–(5.5), using the identities from Lemma 4.9 in place of the ones from Lemma 4.8 used before, and also (4.11) and (4.22). For example, to prove (5.6) when $i = j$, we use (4.41) if $s \geqslant r$ or (4.42) if $s \leqslant r-1$, taking $m = 1, 2, \dots, p$ in turn. Here is the calculation in the latter case:

$$\begin{aligned}
\big(\operatorname{ad} E_i^{(r)}\big)^p \left(E_i^{(s)}\right) &= -\big(\operatorname{ad} E_i^{(r)}\big)^{p-1} \left(\sum_{\substack{s_1, s_2 \leqslant r-1 \\ s_1+s_2 = r-1+s}} E_i^{(s_1)} E_i^{(s_2)} \right) \\
&= 2\big(\operatorname{ad} E_i^{(r)}\big)^{p-2} \left(\sum_{\substack{s_1, s_2, s_3 \leqslant r-1 \\ s_1+s_2+s_3 = 2(r-1)+s}} E_i^{(s_1)} E_i^{(s_2)} E_i^{(s_3)} \right) \\
&= \cdots = (-1)^p p! \sum_{\substack{s_1, \dots, s_{p+1} \leqslant r-1 \\ s_1+\cdots+s_{p+1} = p(r-1)+s}} E_i^{(s_1)} \cdots E_i^{(s_{p+1})} = 0.
\end{aligned}$$

□

For use in the next theorem, we let

$$P_{i,j}(u) = \sum_{r \geqslant p} P_{i,j}^{(r)} u^{-r} := E_{i,j}(u)^p, \qquad Q_{i,j}(u) = \sum_{r \geqslant p} Q_{i,j}^{(r)} u^{-r} := F_{i,j}(u)^p. \tag{5.9}$$

Theorem 5.4. *For $1 \leqslant i < j \leqslant n$, the algebras $Z(Y_n) \cap Y_{i,j}^+$ and $Z(Y_n) \cap Y_{i,j}^-$ are infinite rank polynomial algebras freely generated by the central elements $\left\{ \big(E_{i,j}^{(r)}\big)^p \;\middle|\; r > 0 \right\}$ and $\left\{ \big(F_{i,j}^{(r)}\big)^p \;\middle|\; r > 0 \right\}$, respectively. We have that $\big(E_{i,j}^{(r)}\big)^p, \big(F_{i,j}^{(r)}\big)^p \in \mathrm{F}_{rp-p} Y_n$ and*

$$\operatorname{gr}_{rp-p} \big(E_{i,j}^{(r)}\big)^p = \big(e_{i,j} t^{r-1}\big)^p, \qquad \operatorname{gr}_{rp-p} \big(F_{i,j}^{(r)}\big)^p = \big(e_{j,i} t^{r-1}\big)^p. \tag{5.10}$$

For $r \geqslant p$ we have that

$$P_{i,j}^{(r)} = \begin{cases} \left(E_{i,j}^{(r/p)}\right)^p + (*) & \text{if } p \mid r, \\ (*) & \text{if } p \nmid r, \end{cases} \tag{5.11}$$

where $(*) \in \mathrm{F}_{r-p-1} Y_n$ is a polynomial in the elements $\left(E_{i,j}^{(s)}\right)^p$ for $1 \leqslant s < \lfloor r/p \rfloor$. Hence, the central elements $\left\{ P_{i,j}^{(rp)} \,\middle|\, r > 0 \right\}$ give another algebraically independent set of generators for $Z(Y_n) \cap Y_{i,j}^+$ lifting the central elements $\left\{ \big(e_{i,j} t^{r-1}\big)^p \;\middle|\; r > 0 \right\}$ of $\operatorname{gr} Y_n$. Analogous statements with $Y_{i,j}^+, E, P$ and $e_{i,j} t^{r-1}$ replaced by $Y_{i,j}^-, F, Q$ and $e_{j,i} t^{r-1}$ also hold.

Proof. It suffices to prove all of the statements for $Y_{i,j}^+$; then they follow for $Y_{i,j}^-$ using the anti-automorphism τ from §4.5. Let $\mathfrak{g}_{i,j}^+$ be the commutative subalgebra of $\mathfrak{g}$ spanned by

$\{e_{i,j}t^r \mid r \geqslant 0\}$. By Theorem 3.4, it is easy to see that

$$Z(\mathfrak{g}) \cap U(\mathfrak{g}^+_{i,j}) = \Bbbk\Big[\big(e_{i,j}t^r\big)^p \;\Big|\; r \geqslant 0\Big]. \tag{5.12}$$

Note that $\mathrm{gr}(Z(Y_n) \cap Y^+_{i,j}) \subseteq Z(\mathfrak{g}) \cap U(\mathfrak{g}^+_{i,j})$. By Lemma 5.3, we know that $\big(E^{(r)}_{i,j}\big)^p$ belongs to $Z(Y_n) \cap Y^+_{i,j}$. Moreover, by (4.27), it is clear that $\big(E^{(r)}_{i,j}\big)^p$ is of filtered degree $rp-p$ with $\mathrm{gr}_{rp-p}\big(E^{(r)}_{i,j}\big)^p = \big(e_{i,j}t^{r-1}\big)^p$. It follows that $\mathrm{gr}(Z(Y_n) \cap Y^+_{i,j}) = Z(\mathfrak{g}) \cap U(\mathfrak{g}^+_{i,j})$ and the elements $\Big\{\big(E^{(r)}_{i,j}\big)^p \;\Big|\; r > 0\Big\}$ are algebraically independent generators for $Z(Y_n) \cap Y^+_{i,j}$.

Now consider (5.11). By Lemma 5.2, each $P^{(r)}_{i,j}$ belongs to $Z(Y_n) \cap Y^+_{i,j}$, hence, it is a polynomial in the elements $\Big\{\big(E^{(s)}_{i,j}\big)^p \;\Big|\; s > 0\Big\}$. It remains to show that $P^{(r)}_{i,j} \in \mathrm{F}_{r-p-1}Y_n$ if $p \nmid r$, or that $P^{(r)}_{i,j} \equiv \Big(E^{(r/p)}_{i,j}\Big)^p \pmod{\mathrm{F}_{r-p}Y_n}$ if $r \mid p$. This follows by an application of Lemma 2.1, taking $A_{\mathrm{I}} := Y^+_{i,j}$ and $X^{(r)} := E^{(r)}_{i,j}$ so that $X_{\mathrm{I}}(u) = E_{i,j}(u)$. □

Corollary 5.5. *For any σ, the elements $\Big\{\big({}^\sigma E^{(r)}_{i,j}\big)^p \;\Big|\; r > s_{i,j}\Big\}$ and $\Big\{\big({}^\sigma F^{(r)}_{i,j}\big)^p \;\Big|\; r > s_{j,i}\Big\}$ are central in $Y_n(\sigma)$. Moreover, they belong to $\mathrm{F}_{rp-p}Y_n(\sigma)$ and*

$$\mathrm{gr}_{rp-p}\big({}^\sigma E^{(r)}_{i,j}\big)^p = \big(e_{i,j}t^{r-1}\big)^p, \qquad \mathrm{gr}_{rp-p}\big({}^\sigma F^{(r)}_{i,j}\big)^p = \big(e_{j,i}t^{r-1}\big)^p. \tag{5.13}$$

Proof. The assertion (5.13) is immediate from (4.68), so we just need to establish the centrality. In case σ is upper triangular, we have that ${}^\sigma E^{(r)}_{i,j} = E^{(r)}_{i,j}$, which is central in Y_n by Lemma 5.3 so certainly central in the subalgebra $Y_n(\sigma)$. The centrality of ${}^\sigma E^{(r)}_{i,j}$ in general then follows using the change of shift matrix isomorphism from §4.5. The centrality of ${}^\sigma F^{(r)}_{i,j}$ is proved similarly. □

Remark 5.6. The algebra $Y^+_{i,i+1}$ is generated by $\big\{E^{(r)}_i \mid r > 0\big\}$ subject just to the relations (4.14); these give enough relations because they suffice to establish that the ordered monomials in the generators span $Y^+_{i,i+1}$. Thus, we are in the situation of Remark 2.2 with $X^{(r)} := E^{(r)}_i \in \mathrm{F}_{r-1}Y^+_{i,j}$. Theorem 5.4 shows that $P^{(r)}_{i,i+1}$ is a polynomial in $\Big\{\big(E^{(s)}_i\big)^p \mid 0 < s \leqslant \lfloor r/p \rfloor\Big\}$. This establishes the claim made in Remark 2.2.

5.3. **Diagonal p-central elements.** Next we introduce the p-central elements that belong to the diagonal subalgebras

$$Y^0_i := \Bbbk\big[D^{(r)}_i \mid r > 0\big] \tag{5.14}$$

of Y_n. Note Y^0_i is also a subalgebra of $Y_n(\sigma)$ for any shift matrix σ. For $i = 1, \dots, n$, we define

$$B_i(u) = \sum_{r \geqslant 0} B^{(r)}_i u^{-r} := D_i(u)D_i(u-1)\cdots D_i(u-p+1) \in Y^0_i[[u^{-1}]]. \tag{5.15}$$

Lemma 5.7. *For all $i = 1, \dots, n$ and $r > 0$, the element $B^{(r)}_i$ belongs to $Z(Y_n)$.*

Proof. In view of (4.10), it suffices to check for $1 \leqslant j < n$ that

$$[B_i(u), E_j(v)] = 0 = [B_i(u), F_j(v)]$$

in $Y_n[[u^{-1}, v^{-1}]]$. By applying the anti-automorphism τ from §4.5, it actually suffices to check just the first equality. This is clear when $j \notin \{i-1, i\}$ by (4.12). When $j = i-1$ or $j = i$, it follows from the identities (4.49)–(4.50) taking $m := p$, noting that $B_i(u) = D_{i\downarrow p}(u) = D_{i\uparrow p}(u)$ by (4.10). □

Theorem 5.8. *Assume that $n \geqslant 2$. For $1 \leqslant i \leqslant n$, the algebra $Z(Y_n) \cap Y_i^0$ is an infinite rank polynomial algebra freely generated by the central elements $\big\{B_i^{(rp)} \,\big|\, r > 0\big\}$. This statement also describes the algebras $Z(Y_n(\sigma)) \cap Y_i^0$ for any shift matrix σ. We have that $B_i^{(rp)} \in \mathrm{F}_{rp-p} Y_n$ and*

$$\mathrm{gr}_{rp-p}\, B_i^{(rp)} = (e_{i,i}t^{r-1})^p - e_{i,i}t^{rp-p}. \tag{5.16}$$

For $0 < r < p$, we have that $B_i^{(r)} = 0$. For $r \geqslant p$ with $p \nmid r$, we have that $B_i^{(r)} \in \mathrm{F}_{r-p-1} Y_n$ and it is a polynomial in the elements $\big\{B_i^{(sp)} \,\big|\, 0 < s \leqslant \lfloor r/p \rfloor\big\}$.

Proof. Let $\mathfrak{g}_i^0$ be the abelian subalgebra of $\mathfrak{g}$ spanned by $\big\{e_{i,i}t^r \,\big|\, r \geqslant 0\big\}$. By Theorem 3.4 and the assumption $n \geqslant 2$, one sees that

$$Z(\mathfrak{g}) \cap U(\mathfrak{g}_i^0) = \Bbbk\Big[\big(e_{i,i}t^r\big)^p - e_{i,i}t^{rp} \,\Big|\, r \geqslant 0\Big]. \tag{5.17}$$

We have that $\mathrm{gr}(Z(Y_n) \cap Y_i^0) \subseteq Z(\mathfrak{g}) \cap U(\mathfrak{g}_i^0)$. By Lemma 5.7, we know that $B_i^{(rp+p)}$ belongs to $Z(Y_n) \cap Y_i^0$. Moreover, applying Lemma 2.9 with $A_{\mathrm{III}} = Y_i^0$ and $X^{(r)} = D_i^{(r)}$, we see that $B_i^{(rp+p)} \in \mathrm{F}_{rp} Y_i^0$ and $\mathrm{gr}_{rp}\, B_i^{(rp+p)} = (e_{i,i}t^r)^p - e_{i,i}t^{rp}$. We deduce that $\mathrm{gr}(Z(Y_n) \cap Y_i^0) = Z(\mathfrak{g}) \cap U(\mathfrak{g}_i^0)$ and the elements $\big\{B_i^{(rp)} \,\big|\, r > 0\big\}$ are algebraically independent generators for it. The same argument works in any $Y_n(\sigma)$.

Lemma 2.9 also implies that $B_i^{(r)} = 0$ for $0 < r < p$ and that $B_i^{(r)} \in F_{r-p-1} Y_i^0$ if $r \geqslant p$ with $p \nmid r$. In this case, since it is central by Lemma 5.7, it must be a polynomial in the elements $\big\{B_i^{(sp)} \,\big|\, 0 < s \leqslant \lfloor r/p \rfloor\big\}$. □

Remark 5.9. By Theorem 4.14, Y_i^0 is a free polynomial algebra generated by the elements $D_i^{(1)}, D_i^{(2)}, \ldots$. So Theorem 5.8 also establishes the claim made in Remark 2.10.

5.4. **Main Theorem.** Now we can state and prove our main results. Let σ be any shift matrix. We have already defined the Harish-Chandra centre $Z_{\mathrm{HC}}(Y_n(\sigma))$ at the end of §5.1. Also define the *p-centre* $Z_p(Y_n(\sigma))$ of $Y_n(\sigma)$ to be the subalgebra generated by

$$\big\{B_i^{(rp)} \,\big|\, 1 \leqslant i \leqslant n, r > 0\big\} \cup \Big\{\big({}^{\sigma}E_{i,j}^{(r)}\big)^p, \big({}^{\sigma}F_{i,j}^{(r)}\big)^p \,\Big|\, 1 \leqslant i < j \leqslant n, r > s_{i,j}, s > s_{j,i}\Big\}. \tag{5.18}$$

We have shown that both $Z_{\mathrm{HC}}(Y_n(\sigma))$ and $Z_p(Y_n(\sigma))$ are subalgebras of $Z(Y_n(\sigma))$; see Corollary 5.5 and Lemma 5.7. Note also by (4.68) and (5.16) that $\mathrm{gr}\, Z_p(Y_n(\sigma))$ may be identified with the p-centre $Z_p(\mathfrak{g}_\sigma)$ of $U(\mathfrak{g}_\sigma)$ from (3.9).

We also need one more family of elements: recalling (5.1) and (5.15), we let

$$BC(u) := \sum_{r \geqslant 0} BC^{(r)} u^{-r} := B_1(u)B_2(u-1)\cdots B_n(u-n+1) = C(u)C(u-1)\cdots C(u-p+1). \tag{5.19}$$

From this definition, it follows that each $BC^{(r)}$ can be expressed as a polynomial in the elements $\big\{B_i^{(s)} \,\big|\, 1 \leqslant i \leqslant n, s > 0\big\}$, so that it belongs to $Z_p(Y_n(\sigma))$ by Theorem 5.8. Moreover, it is also a polynomial in the elements $\big\{C^{(s)} \,\big|\, s > 0\big\}$, so that it belongs to $Z_{\mathrm{HC}}(Y_n(\sigma))$. We have just shown that $BC^{(r)} \in Z_{\mathrm{HC}}(Y_n(\sigma)) \cap Z_p(Y_n(\sigma))$.

Lemma 5.10. *For $r > 0$, we have that $BC^{(rp)} \in F_{rp-p} Y_n(\sigma)$ and*

$$\mathrm{gr}_{rp}\, BC^{(rp)} = z_{r-1}^p - z_{rp-p}. \tag{5.20}$$

Proof. Let $d_r := 0$ for $r < p$ and $d_r := p\lfloor r/p \rfloor - p$ for $r \geqslant p$. Theorem 5.8 implies that $B_i^{(r)} \in \mathrm{F}_{d_r} Y_n(\sigma)$ for every $r > 0$. Now apply Lemma 2.5 with $A_{\mathrm{II}} = Y_n(\sigma)$ and

$X_i^{(r)} = B_i^{(r)}$, noting that pr is optimal for every $r > 1$ by Lemma 2.3, to deduce that $BC^{(rp)} \in F_{rp-p}Y_n(\sigma)$ and

$$BC^{(rp)} \equiv B_1^{(rp)} + \cdots + B_n^{(rp)} \pmod{F_{rp-p-1}Y_n(\sigma)}.$$

We are now done thanks to (5.16) once again. □

Theorem 5.11. *The centre $Z(Y_n(\sigma))$ is generated by $Z_{\mathrm{HC}}(Y_n(\sigma))$ and $Z_p(Y_n(\sigma))$. Moreover:*

(1) *$Z_{\mathrm{HC}}(Y_n(\sigma))$ is the free polynomial algebra generated by $\{C^{(r)} \mid r > 0\}$;*
(2) *$Z_p(Y_n(\sigma))$ is the free polynomial algebra generated by*

$$\{B_i^{(rp)} \mid 1 \leqslant i \leqslant n, r > 0\} \cup \left\{ \left({}^{\sigma}E_{i,j}^{(r)}\right)^p, \left({}^{\sigma}F_{i,j}^{(s)}\right)^p \;\middle|\; 1 \leqslant i < j \leqslant n, r > s_{i,j}, s > s_{j,i} \right\}; \quad (5.21)$$

(3) *$Z(Y_n(\sigma))$ is the free polynomial algebra generated by*

$$\{B_i^{(rp)}, C^{(r)} \mid 2 \leqslant i \leqslant n, r > 0\} \cup \left\{ \left({}^{\sigma}E_{i,j}^{(r)}\right)^p, \left({}^{\sigma}F_{i,j}^{(s)}\right)^p \;\middle|\; 1 \leqslant i < j \leqslant n, r > s_{i,j}, s > s_{j,i} \right\}; \quad (5.22)$$

(4) *$Z_{\mathrm{HC}}(Y_n(\sigma)) \cap Z_p(Y_n(\sigma))$ is the free polynomial algebra generated by $\{BC^{(rp)} | r > 0\}$.*

Proof. (1) This is Theorem 5.1.

(2) The given elements generate $Z_p(Y_n(\sigma))$ by the definition. We just need to observe that they are algebraically independent. This follows because by (5.13) and (5.16) they are lifts of the algebraically independent generators of the p-centre of the associated graded algebra from (3.9).

(3) Let Z be the subalgebra of $Z(Y_n(\sigma))$ generated by the given elements. We have that

$$\operatorname{gr} Z \subseteq \operatorname{gr} Z(Y_n(\sigma)) \subseteq Z(\operatorname{gr} Y_n(\sigma)) = Z(\mathfrak{g}_\sigma). \quad (5.23)$$

We have seen already that generators of Z are lifts of the algebraically independent generators of $Z(\mathfrak{g}_\sigma)$ from (3.10). Hence, they are algebraically independent and equality holds everywhere in (5.23). This implies that $Z = Z(Y_n(\sigma))$.

(4) We have already observed that all $BC^{(rp)}$ belong to $Z_{\mathrm{HC}}(Y_n(\sigma)) \cap Z_p(Y_n(\sigma))$. Also they are algebraically independent as they are lifts of algebraically independent elements of $U(\mathfrak{g}_\sigma)$ by (5.20). We claim that $Z_p(Y_n(\sigma))$ is freely generated by the elements

$$\{BC^{(rp)}, B_i^{(rp)} \mid 2 \leqslant i \leqslant n, r > 0\} \cup \left\{ \left({}^{\sigma}E_{i,j}^{(r)}\right)^p, \left({}^{\sigma}F_{i,j}^{(s)}\right)^p \;\middle|\; 1 \leqslant i < j \leqslant n, r > s_{i,j}, s > s_{j,i} \right\}.$$

The result follows from the claim since we know already from (3) that all of these elements different from $BC^{(rp)}$ are algebraically independent of anything in $Z_{\mathrm{HC}}(Y_n(\sigma))$.

To prove the claim, we use (5.13) (5.16) and (5.20) to pass to the associated graded algebra, thereby reducing to showing that

$$\{z_r^p - z_{rp}, (e_{i,i}t^r)^p - e_{i,i}t^{rp} \mid 2 \leqslant i \leqslant n, r \geqslant 0\} \cup \left\{ \left(e_{i,j}t^r\right)^p \;\middle|\; 1 \leqslant i \neq j \leqslant n, r \geqslant s_{i,j} \right\}$$

freely generate $Z_p(\mathfrak{g}_\sigma)$. This is easily seen by comparing them to the algebraically independent generators from (3.9). □

Corollary 5.12. *The shifted Yangian $Y_n(\sigma)$ is free as a module over its centre, with basis given by the ordered monomials in*

$$\{D_i^{(r)} \mid 2 \leqslant i \leqslant n, r > 0\} \cup \{{}^{\sigma}E_{i,j}^{(r)}, {}^{\sigma}F_{i,j}^{(s)} \mid 1 \leqslant i < j \leqslant n, r > s_{i,j}, s > s_{j,i}\} \quad (5.24)$$

in which no exponent is p or more.

Proof. It suffices to show that the set consisting of ordered monomials in (5.22) multiplied by ordered monomials in (5.24) with all exponents $< p$ gives a basis for $Y_n(\sigma)$. To see this, we pass to the associated graded algebra using (4.27), (4.68), (5.2), (5.13) and (5.16) to reduce to showing that the monomials

$$\prod_{r\geqslant 0} z_r^{a_{1,1,r}} \prod_{\substack{2\leqslant i\leqslant n\\ r\geqslant 0}} \left((e_{i,i}t^r)^p - e_{i,i}t^{rp}\right)^{a_{i,i;r}} \prod_{\substack{1\leqslant i\neq j\leqslant n\\ r>s_{i,j}}} (e_{i,j}t^r)^{pa_{i,j,r}} \prod_{\substack{2\leqslant i\leqslant n\\ r\geqslant 0}} (e_{i,i}t^r)^{b_{i,i;r}} \prod_{\substack{1\leqslant i\neq j\leqslant n\\ r>s_{i,j}}} (e_{i,j}t^r)^{b_{i,j,r}}$$

for $a_{i,j;r} \geqslant 0$ and $0 \leqslant b_{i,j,r} < p$ form a basis for $U(\mathfrak{g}_\sigma)$. This is quite straightforward: these monomials are related to a PBW basis by a uni-triangular transition matrix. □

Corollary 5.13. *The shifted Yangian $Y_n(\sigma)$ is free as a module over $Z_p(Y_n(\sigma))$ with basis given by the ordered monomials in*

$$\left\{D_i^{(r)} \,\middle|\, 1 \leqslant i \leqslant n, r > 0\right\} \cup \left\{{}^\sigma E_{i,j}^{(r)}, {}^\sigma F_{i,j}^{(s)} \,\middle|\, 1 \leqslant i < j \leqslant n, r > s_{i,j}, s > s_{j,i}\right\}$$

in which no exponent is p or more.

Proof. Similar to the previous corollary. □

Remark 5.14. Assume in this remark that $n = 2$ and $p = 2$. Recalling (4.53), we denote $E_1^{(r)}, F_1^{(r)}$ and $H_1^{(r)}$ simply by $E^{(r)}, F^{(r)}$ and $H^{(r)}$. Since $H(u) = -B_2(u)C(u)^{-1}$, the elements $H^{(r)}$ are all central; this also follows from (4.56). Consider $\overline{Y}_2 := Y_2/I$ where I is the two-sided ideal generated by the central elements $(E^{(r)})^2$ and $(F^{(r)})^2$ for all $r > 0$. An induction exercise using (4.14)–(4.15) shows that the following relations hold in $\overline{Y}_2$: $[E^{(r)}, E^{(s)}] = [F^{(r)}, F^{(s)}] = 0$ for $r, s > 0$. Comparing with the presentation obtained in [G, Theorem 3], we see that $\overline{Y}_2$ may be identified with the Yangian of the Lie superalgebra $\mathfrak{gl}_{1|1}$ in characteristic 2.

6. Modular Yangian of $\mathfrak{sl}_n$

In this section, we define a subalgebra $SY_n < Y_n$ which we call the *special Yangian*, and show that this may be viewed as the modular version of the Yangian for the Lie algebra $\mathfrak{sl}_n$ rather than $\mathfrak{gl}_n$. Then we use this connection to establish the results about $Z(Y_n)$ formulated in terms of the RTT generators in the introduction.

6.1. **The special Yangian.** We would like to mimic Drinfeld's definition of SY_n in characteristic zero from [MNO]; see [MNO, §2.24] for its history. Unfortunately, the approach in *loc. cit.* only works verbatim when the ground field is infinite. Rather than insisting on that here, we will modify the definition slightly by incorporating base change. For any field extension $\mathbb{K} \supseteq \Bbbk$, we can define the Yangian over $\mathbb{K}$ by generators and relations in the same way as Y_n was defined over $\Bbbk$. The resulting $\mathbb{K}$-algebra may be identified with $Y_n \otimes \mathbb{K}$ in an obvious way. Then the automorphisms μ_f defined as in §4.5 can be viewed as $\mathbb{K}$-linear automorphisms of $Y_n \otimes \mathbb{K}$ for all $f(u) \in 1 + u^{-1}\mathbb{K}[[u^{-1}]]$. Define the *special Yangian* to be

$$SY_n := \left\{ x \in Y_n \,\middle|\, \begin{array}{l} \mu_f(x \otimes 1) = x \otimes 1 \text{ in } Y_n \otimes \mathbb{K} \\ \text{for all } f(u) \in 1 + u^{-1}\mathbb{K}[[u^{-1}]] \\ \text{and all field extensions } \mathbb{K} \supseteq \Bbbk \end{array} \right\}. \tag{6.1}$$

In fact, as is clear from the proof of the next theorem, it is enough just to take one infinite field $\mathbb{K}$ here.

Our first task is to identify the associated graded algebra $\operatorname{gr} SY_n \subseteq \operatorname{gr} Y_n = U(\mathfrak{g})$ with $U(\mathfrak{g}')$, where $\mathfrak{g}' := \mathfrak{sl}_n[t]$ is the current algebra associated to $\mathfrak{sl}_n$. Recall the elements $H_i^{(r)}$ from (4.53).

Theorem 6.1. *The algebra SY_n has a basis consisting of ordered monomials in*

$$\big\{H_i^{(r)} \,\big|\, 1 \leqslant i < n, r > 0\big\} \cup \big\{E_{i,j}^{(r)}, F_{i,j}^{(r)} \,\big|\, 1 \leqslant i < j \leqslant n, r > 0\big\} \tag{6.2}$$

taken in any fixed order. Hence, $\operatorname{gr} SY_n = U(\mathfrak{g}')$*, and multiplication defines a* vector space *isomorphism*

$$SY_n \otimes Y_1 \xrightarrow{\sim} Y_n \tag{6.3}$$

where Y_1 is identified with the subalgebra of Y_n generated by the elements $D_1^{(r)}$ in the obvious way. If we assume in addition that $p \nmid n$ then multiplication defines an algebra *isomorphism*

$$SY_n \otimes Z_{\mathrm{HC}}(Y_n) \xrightarrow{\sim} Y_n. \tag{6.4}$$

Proof. For any $\mathbb{K} \supseteq \Bbbk$ and $f(u) \in 1 + u^{-1}\mathbb{K}[[u^{-1}]]$, we have that $\mu_f(D_i(u) \otimes 1) = f(u)D_i(u) \otimes 1$ by definition, hence, $\mu_f(\widetilde{D}_i(u) \otimes 1) = f(u)^{-1}\widetilde{D}_i(u) \otimes 1$. It follows that $H_i^{(r)} \in SY_n$, and of course all $E_{i,j}^{(r)}$ and $F_{i,j}^{(r)}$ lie in SY_n too. Hence, ordered monomials in the elements (6.2) span a subspace $\overline{SY}_n$ of SY_n. We are shortly going to prove that $\overline{SY}_n = SY_n$, so that $\overline{SY}_n$ is a subalgebra; one could also prove this right away using Lemma 4.11 but actually we do not need it for the proof below.

The filtration on Y_n induces a vector space filtration on $\overline{SY}_n$ so that $\operatorname{gr}\overline{SY}_n$ is a graded subspace of $U(\mathfrak{g})$. It is easy to see from (4.27) and (4.54) that this subspace is the subalgebra $U(\mathfrak{g}')$. In particular, this implies that the ordered monomials that span $\overline{SY}_n$ are linearly independent too. Furthermore, multiplying them by ordered monomials in $\big\{D_1^{(r)} \,\big|\, r > 0\big\}$ gives a basis for Y_n. This shows that the linear map

$$\begin{aligned}\overline{SY}_n \otimes \Bbbk[x_1, x_2, \dots] &\to Y_n,\\ A(x_1, \dots, x_n) = \sum_i A_i \otimes a_i(x_1, \dots, x_n) &\mapsto A(D_1^{(1)}, \dots, D_1^{(n)}) = \sum_i A_i a_i\big(D_1^{(1)}, \dots D_1^{(n)}\big)\end{aligned}$$

is a vector space isomorphism.

Now we can show that $SY_n \subseteq \overline{SY}_n$. Take any $B \in SY_n$. By the previous paragraph, we can write it as $A\big(D_1^{(1)}, \dots, D_1^{(n)}\big)$ for a unique $A(x_1, \dots, x_n) \in \overline{SY}_n \otimes \Bbbk[x_1, \dots, x_n]$ and $n \geqslant 1$. Taking $f(u) := 1 + cu^{-n}$ for $c \in \mathbb{K}$ and an infinite field $\mathbb{K} \supseteq \Bbbk$, we have that

$$A\big(D_1^{(1)}, \dots, D_1^{(n)}\big) \otimes 1 = \mu_f\Big(A\big(D_1^{(1)}, \dots, D_1^{(n)}\big) \otimes 1\Big) = A\big(D_1^{(1)}, \dots, D_1^{(n-1)}, D_1^{(n)} + c\big) \otimes 1.$$

This implies that $A(x_1, \dots, x_n) = A(x_1, \dots, x_{n-1}, x_n + c)$ for infinitely many c. Hence, $A(x_1, \dots, x_n)$ is independent of x_n. Similarly, it is independent of $x_{n-1}, \dots, x_1$. This shows $B \in \overline{SY}_n$ as required.

We have now shown that the ordered monomials in the elements (6.2) give a basis for SY_n, that $\operatorname{gr} SY_n = U(\mathfrak{g}')$, and that the map (6.3) is a vector space isomorphism. Finally, we must prove (6.4) assuming $p \nmid n$. The vector space $SY_n \otimes Z_{\mathrm{HC}}(Y_n)$ has a basis given by the ordered monomials in (6.2) tensored with ordered monomials in $\big\{C^{(r)} \,\big|\, r > 0\big\}$. It suffices to show that the images of these give a basis for Y_n. By passing to the associated graded algebra as usual, this reduces to the observation that

$$\big\{e_{i,i}t^r - e_{i+1,i+1}t^r \,\big|\, i = 1, \dots, n-1, r \geqslant 0\big\} \cup \big\{e_{i,j}t^r \,\big|\, 1 \leqslant i \neq j \leqslant n, r \geqslant 0\big\} \cup \{z_r \mid r \geqslant 0\}$$

is a basis for $\mathfrak{g}'$. □

Remark 6.2. Recall that $\widetilde{T}_{i,j}(u) = S(T_{i,j}(u))$. Since $\mu_f(\widetilde{T}_{i,j}(u)) = f(u)^{-1}\widetilde{T}_{i,j}(u)$, the definition (6.1) implies that all coefficients of $T_{i,j}(u)\widetilde{T}_{k,l}(u)$ belong to SY_n. By passing to $\operatorname{gr} Y_n$ and using Theorem 6.1, one sees that these coefficients also generate SY_n. Using

this and (4.3), it follows that $\Delta(SY_n) \subseteq SY_n \otimes SY_n$, so that SY_n is a Hopf subalgebra of Y_n; cf. [MNO, Proposition 2.21].

Theorem 6.3. *The algebra SY_n is generated by the elements*

$$\big\{H_i^{(r)}, E_i^{(r)}, F_i^{(r)} \,\big|\, 1 \leqslant i < n, r > 0\big\} \tag{6.5}$$

subject only to the relations (4.14)–(4.23) plus the following:

$$\big[H_i^{(r)}, H_j^{(s)}\big] = 0, \tag{6.6}$$

$$\big[E_i^{(r)}, F_j^{(s)}\big] = \delta_{i,j} H_i^{(r+s-1)}, \tag{6.7}$$

$$\big[H_i^{(r)}, E_j^{(s)}\big] = 0 \qquad \text{if } |i-j| > 1, \tag{6.8}$$

$$\big[H_i^{(r)}, F_j^{(s)}\big] = 0 \qquad \text{if } |i-j| > 1, \tag{6.9}$$

$$\big[H_{i-1}^{(r+1)}, E_i^{(s)}\big] - \big[H_{i-1}^{(r)}, E_i^{(s+1)}\big] = H_{i-1}^{(r)} E_i^{(s)}, \tag{6.10}$$

$$\big[H_{i-1}^{(r)}, F_i^{(s+1)}\big] - \big[H_{i-1}^{(r+1)}, F_i^{(s)}\big] = F_i^{(s)} H_{i-1}^{(r)}, \tag{6.11}$$

$$\big[H_i^{(r+1)}, E_i^{(s)}\big] - \big[H_i^{(r)}, E_i^{(s+1)}\big] = -H_i^{(r)} E_i^{(s)} - E_i^{(s)} H_i^{(r)}, \tag{6.12}$$

$$\big[H_i^{(r)}, F_i^{(s+1)}\big] - \big[H_i^{(r+1)}, F_i^{(s)}\big] = -F_i^{(s)} H_i^{(r)} - H_i^{(r)} F_i^{(s)}, \tag{6.13}$$

$$\big[H_{i+1}^{(r+1)}, E_i^{(s)}\big] - \big[H_{i+1}^{(r)}, E_i^{(s+1)}\big] = E_i^{(s)} H_{i+1}^{(r)}, \tag{6.14}$$

$$\big[H_{i+1}^{(r)}, F_i^{(s+1)}\big] - \big[H_{i+1}^{(r+1)}, F_i^{(s)}\big] = H_{i+1}^{(r)} F_i^{(s)}, \tag{6.15}$$

for all admissible i, j, r, s, t including $r = 0$ in (6.10)–(6.15); remember also $H_i^{(0)} = -1$.

Proof. In view of (4.9), Theorem 6.1 implies that SY_n is generated by the elements (6.5). Let us also show that all of the relations in the theorem are satisfied. Of course (4.14)–(4.23) hold, and the relations (6.6)–(6.7) follow from (4.10)–(4.11). For the remaining relations, (6.10), (6.12) and (6.14) follow by equating $u^{-r}v^{-s}$-coefficients in (4.58), (4.63) and (4.61), respectively, taking $r \geqslant 0$ and $s > 0$. Then (6.11), (6.13) and (6.15) follow by applying the anti-automorphism τ.

Now let $\widehat{SY}_n$ be the algebra defined by the generators and relations from the theorem. The previous paragraph shows that there is an algebra homomorphism $\widehat{SY}_n \twoheadrightarrow SY_n$ taking generators to generators. To show that it is an isomorphism, define elements $E_{i,j}^{(r)}, F_{i,j}^{(r)} \in \widehat{SY}_n$ by (4.9). Using the basis for SY_n constructed in Theorem 6.1, it suffices to show that ordered monomials in the elements (6.2) span $\widehat{SY}_n$. Moreover, we can choose the order so that the elements $F_{i,j}^{(r)}$ come first, followed by the elements $H_i^{(r}$, followed by the elements $E_{i,j}^{(r)}$. Of course, $\widehat{SY}_n$ is spanned by unordered monomials in $F_i^{(r)}, H_i^{(r)}$ and $E_i^{(r)}$. The relations allow us to inductively commute all $F_i^{(r)}$ to the beginning, all $H_i^{(r)}$ to the middle and in the chosen order due to (6.6), and all $E_i^{(r)}$ to the end. Then we get done because, in the subalgebras generated by just the $F_i^{(r)}$ or the $E_i^{(r)}$, we have available exactly the same relations as in Y_n, and there we have already established that the given ordered monomials span these subalgebras using the relations (4.14)–(4.23) and (4.9). □

Remark 6.4. Assume $\operatorname{char} \Bbbk \neq 2$. Then there is an even more efficient presentation for SY_n, namely, the usual Drinfeld presentation for the Yangian of $\mathfrak{sl}_n$ from [D] naïvely taken over the field $\Bbbk$ rather than over the complex numbers; actually, we use the "opposite"

presentation like in [BK1, Remark 5.12]. In more detail, we define new elements

$$\kappa_i(u) = \sum_{k\geqslant 0} \kappa_{i,k} u^{-k-1} := 1 + \eta_{(i-1)/2}(H_i(u)), \tag{6.16}$$

$$\xi_i^+(u) = \sum_{k\geqslant 0} \xi_{i,k}^+ u^{-k-1} := \eta_{(i-1)/2}(E_i(u)), \tag{6.17}$$

$$\xi_i^-(u) = \sum_{k\geqslant 0} \xi_{i,k}^- u^{-k-1} := \eta_{(i-1)/2}\left(F_i(u)\right), \tag{6.18}$$

where η_c is the automorphism from §4.5 (which leaves SY_n invariant). Then, the algebra SY_n is generated by $\big\{\kappa_{i,k}, \xi_{i,k}^\pm \mid 1 \leqslant i < n, k \geqslant 0\big\}$ subject only to the Drinfeld relations:

$$\big[\kappa_{i,k}, \kappa_{j,l}\big] = 0, \tag{6.19}$$

$$\big[\xi_{i,k}^+, \xi_{j,l}^-\big] = \delta_{i,j}\kappa_{i,k+l}, \tag{6.20}$$

$$\big[\kappa_{i,0}, \xi_{j,l}^\pm\big] = \pm a_{i,j}\xi_{j,l}^\pm, \tag{6.21}$$

$$\big[\kappa_{i,k}, \xi_{j,l+1}^\pm\big] - \big[\kappa_{i,k+1}, \xi_{j,l}^\pm\big] = \pm\frac{a_{i,j}}{2}(\kappa_{i,k}\xi_{j,l}^\pm + \xi_{j,l}^\pm\kappa_{i,k}), \tag{6.22}$$

$$\big[\xi_{i,k}^\pm, \xi_{j,l+1}^\pm\big] - \big[\xi_{i,k+1}^\pm, \xi_{j,l}^\pm\big] = \pm\frac{a_{i,j}}{2}(\xi_{i,k}^\pm\xi_{j,l}^\pm + \xi_{j,l}^\pm\xi_{i,k}^\pm), \tag{6.23}$$

$$\Big[\xi_{i,k_1}^\pm, \big[\xi_{i,k_2}^\pm, \xi_{j,l}^\pm\big]\Big] + \Big[\xi_{i,k_2}^\pm, \big[\xi_{i,k_1}^\pm, \xi_{j,l}^\pm\big]\Big] = 0 \text{ if } |i-j| = 1,\ k_1 \neq k_2, \tag{6.24}$$

$$\Big[\xi_{i,k}^\pm, \big[\xi_{i,k}^\pm, \xi_{j,l}^\pm\big]\Big] = 0 \text{ if } |i-j| = 1, \tag{6.25}$$

$$\big[\xi_{i,k}^\pm, \xi_{j,l}^\pm\big] = 0 \text{ if } |i-j| > 1, \tag{6.26}$$

for $a_{i,j} := 2\delta_{i,j} - \delta_{i,j+1} - \delta_{i,j-1}$ (the Cartan matrix of type A_{n-1}). This assertion is just a rephrasing of Theorem 6.3 for these modified generators. For example, the relation (6.23) is deduced in [BK1, Remark 5.12]; the relations (6.21)–(6.22) follow from (4.64)–(4.65) suitably shifted.

Remark 6.5. There is a shifted analogue $SY_n(\sigma)$ of SY_n. This may be realized as a subalgebra of $Y_n(\sigma)$ via a similar definition to (6.1). The presentation in Theorem 6.3 can also be modified to give a presentation for $SY_n(\sigma)$, in just the same way that Theorem 4.15 modifies Theorem 4.3; we leave the details of this to the reader. The Drinfeld presentation from the previous remark does not immediately make sense for $SY_n(\sigma)$, but see [WWY] for a closely related result (in characteristic zero).

6.2. **The p-centre of SY_n.** Let

$$\begin{aligned} A_i(u) = \sum_{r\geqslant 0} A_i^{(r)} u^{-r} &:= H_i(u)H_i(u-1)\cdots H_i(u-p+1) \\ &= -B_{i+1}(u)B_i(u)^{-1} \in SY_n[[u^{-1}]]. \end{aligned} \tag{6.27}$$

In view of Theorem 5.8, each $A_i^{(r)}$ belongs to $Z(SY_n)$. We define the *p-centre* of SY_n to be the subalgebra $Z_p(SY_n)$ of $Z(SY_n)$ generated by

$$\big\{A_i^{(rp)} \,\big|\, 1 \leqslant i < n, r > 0\big\} \cup \Big\{\big(E_{i,j}^{(r)}\big)^p, \big(F_{i,j}^{(r)}\big)^p \,\Big|\, 1 \leqslant i < j \leqslant n, r > 0\Big\}. \tag{6.28}$$

Also let $Z_p(\mathfrak{g}')$ be the p-centre of $U(\mathfrak{g}')$, i.e. the subalgebra of $Z(\mathfrak{g}')$ generated by $x^p - x^{[p]}$ for all $x \in \mathfrak{g}'$.

Theorem 6.6. *The generators (6.28) of $Z_p(SY_n)$ are algebraically independent, and we have that* $\operatorname{gr} Z_p(SY_n) = Z_p(\mathfrak{g}')$. *Moreover, SY_n is free as a module over $Z_p(SY_n)$ with basis*

given by the ordered monomials in

$$\big\{H_i^{(r)} \,\big|\, 1 \leqslant i < n, r > 0\big\} \cup \big\{E_{i,j}^{(r)}, F_{i,j}^{(r)} \,\big|\, 1 \leqslant i < j \leqslant n, r > 0\big\} \tag{6.29}$$

in which no exponent is p or more.

Proof. From the formula $A_i(u) = -B_{i+1}(u)B_i(u)^{-1}$, we get that $A_i^{(r)} = B_i^{(r)} - B_{i+1}^{(r)}$ plus a linear combination of monomials $B_i^{(r_1)} \cdots B_i^{(r_k)} B_{i+1}^{(s_1)} \cdots B_{i+1}^{(s_l)}$ with $r_1 + \cdots + r_k + s_1 + \cdots + s_l = r$. Combined with (5.16), it follows that $A_i^{(rp)} \in \mathrm{F}_{rp-p} Y_n$ and

$$\mathrm{gr}_{rp-p} A_i^{(rp)} = \big(e_{i,i}t^{r-1} - e_{i+1,i+1}t^{r-1}\big)^p - \big(e_{i,i}t^{rp-p} - e_{i+1,i+1}t^{rp-p}\big). \tag{6.30}$$

Then from (6.30) and (5.10), we see that the generators (6.28) of $Z_p(SY_n)$ are lifts of generators for $Z_p(\mathfrak{g}')$ coming from a basis for $\mathfrak{g}'$. This establishes the algebraic independence and that $\mathrm{gr}\, Z_p(SY_n) = Z_p(\mathfrak{g}')$. The final part of the theorem follows by similar argument to the proof of Corollary 5.13, using the PBW basis for SY_n from Theorem 6.1. □

In fact, when $p \nmid n$, the p-center of SY_n is the full center, thanks to the following theorem. This is the positive characteristic counterpart of the observation that $Z(SY_n)$ is trivial in characteristic zero from [MNO, Proposition 2.16].

Theorem 6.7. *If $p \nmid n$ then $Z_p(SY_n) = Z(SY_n)$.*

Proof. We have that $Z_p(SY_n) \subseteq Z(SY_n)$, hence, $\mathrm{gr}\, Z_p(SY_n) \subseteq \mathrm{gr}\, Z(SY_n) \subseteq Z(\mathfrak{g}')$. In the next paragraph, we show that $Z_p(\mathfrak{g}') = Z(\mathfrak{g}')$. We also know that $\mathrm{gr}\, Z_p(SY_n) = Z_p(\mathfrak{g}')$ from the previous theorem. Then we get that $\mathrm{gr}\, Z_p(SY_n) = \mathrm{gr}\, Z(SY_n)$ implying that $Z_p(SY_n) = Z(SY_n)$.

To show that $Z_p(\mathfrak{g}') = Z(\mathfrak{g}')$, the assumption $p \nmid n$ implies that $\mathfrak{g} = \mathfrak{g}' \oplus \mathfrak{z}(\mathfrak{g})$. Hence, $Z(\mathfrak{g}) \cong Z(\mathfrak{g}') \otimes \Bbbk[z_r \mid r \geqslant 0]$. It remains to observe that the elements $\{x^p - x^{[p]} \mid x \in \mathfrak{g}'\} \cup \{z_r \mid r \geqslant 0\}$ generate $Z(\mathfrak{g})$. This follows from Theorem 3.4 using the assumption $p \nmid n$. □

6.3. **Another description of the p-centre of Y_n.** Recall by the definition (5.18) and Theorems 5.4 and 5.8 that the p-centre $Z_p(Y_n)$ is the subalgebra of $Z(Y_n)$ generated by the coefficients of the power series $B_i(u)$, $P_{i,j}(u)$ and $Q_{i,j}(u)$. Let

$$S_{i,j}(u) = \sum_{r \geqslant 0} S_{i,j}^{(r)} u^{-r} := T_{i,j}(u)T_{i,j}(u-1)\cdots T_{i,j}(u-p+1) \in Y_n[[u^{-1}]]. \tag{6.31}$$

In view of (4.2), the order of the product on the right hand side here is irrelevant.

Lemma 6.8. *All of the elements $S_{i,j}^{(r)}$ belong to the p-centre $Z_p(Y_n)$.*

Proof. First we show that each $S_{i,j}^{(r)}$ belongs to $Z(Y_n)$. To see this, using the conjugation automorphism from §4.5 which sends $S_{i,j}(u)$ to $S_{w(i),w(j)}(u)$, we reduce to proving that all coefficients of $S_{1,1}(u)$ and of $S_{1,2}(u)$ are central. The latter assertions follow because

$$S_{1,1}(u) = B_1(u), \tag{6.32}$$

$$S_{1,2}(u) = B_1(u)P_{1,2}(u). \tag{6.33}$$

The first identity (6.32) here is immediate as $T_{1,1}(u) = D_1(u)$. To prove (6.33), we set $v = u - m$ in (4.51) to deduce that

$$E_i(u-m)D_i(u-m+1)\cdots D_i(u-1)D_i(u) = D_i(u-m+1)\cdots D_i(u-1)D_i(u)E_i(u)$$

for each $m = 1, \dots, p-1$. The $(1,2)$-entry of (4.5) gives that $T_{1,2}(u) = D_1(u)E_1(u)$. Hence, we get that

$$\begin{aligned} S_{1,2}(u) &= T_{1,2}(u-p+1)\cdots T_{1,2}(u-1)T_{1,2}(u) \\ &= D_1(u-p+1)E_1(u-p+1)\cdots D_1(u-1)E_1(u-1)D_1(u)E_1(u) \\ &= D_1(u-p+1)\cdots D_1(u-1)D_1(u)E_1(u)^p = B_1(u)P_{1,2}(u). \end{aligned}$$

This establishes (6.33).

The delicate point now is to show that $S_{i,j}^{(r)}$ actually lies in $Z_p(Y_n)$ not just $Z(Y_n)$. By Theorem 5.11(2), we have that $Z_p(Y_n) = Y_n \cap Z_p(Y_{n+1})$, where we are implicitly using the natural embedding $Y_n \hookrightarrow Y_{n+1}, T_{i,j}^{(r)} \mapsto T_{i,j}^{(r)}$. Hence, in order to prove that $S_{i,j}^{(r)} \in Z_p(Y_n)$, we may assume that $p \nmid n$.

So finally we assume $p \nmid n$ and show that $S_{i,j}(u) \in Z_p(Y_n)[[u^{-1}]]$. This is immediate by (6.32) in case $i = j = 1$. In general, we show equivalently that $S_{i,j}(u)S_{1,1}(u)^{-1} \in Z_p(Y_n)[[u^{-1}]]$. Using the definition (6.1), we get that $S_{i,j}(u)S_{1,1}(u)^{-1} \in SY_n[[u^{-1}]]$. Since we have shown its coefficients are central already, it therefore lies in $Z(SY_n)[[u^{-1}]]$, which by Theorem 6.7 and the definitions is $Z_p(SY_n)[[u^{-1}]] \subset Z_p(Y_n)[[u^{-1}]]$. □

Theorem 6.9. *The p-centre $Z_p(Y_n)$ is freely generated by $\{S_{i,j}^{(rp)} \mid 1 \leqslant i,j \leqslant n, r > 0\}$. We have that $S_{i,j}^{(rp)} \in \mathrm{F}_{rp-p}Y_n$ and*

$$\operatorname{gr}_{rp-p} S_{i,j}^{(rp)} = \left(e_{i,j}t^{r-1}\right)^p - \delta_{i,j}e_{i,j}t^{rp-p}. \tag{6.34}$$

For $0 < r < p$, we have that $S_{i,j}^{(r)} = 0$. For $r \geqslant p$ with $p \nmid r$, the central element $S_{i,j}^{(r)}$ belongs to $\mathrm{F}_{rp-p-1}Y_n$ and it may be expressed as a polynomial in the elements $\{S_{i,j}^{(ps)} \mid 0 < s \leqslant \lfloor r/p \rfloor\}$.

Proof. When $n = 1$, the first statement follows immediately from the definition of $Z_p(Y_n)$, remembering (6.32). The remaining statements follow too if we can prove them for larger n. So we assume from now on that $n \geqslant 2$. To prove (6.34), we apply Lemma 2.9 if $i = j$ or Lemma 2.11 if $i \neq j$, taking $X^{(r)} := T_{i,j}^{(r)}$. These lemmas also show that $S_{i,j}^{(r)} = 0$ for $0 < r < p$ and that $S_{i,j}^{(r)} \in \mathrm{F}_{rp-p-1}Y_n$ when $p \nmid r \geqslant p$.

Let $Y_{i,j}$ be the subalgebra of Y_n generated by the elements $\{T_{i,j}^{(r)} \mid r > 0\}$ and $\mathfrak{g}_{i,j}$ be the subalgebra of $\mathfrak{g}$ spanned by $\{e_{i,j}t^r \mid r \geqslant 0\}$. We have that $\operatorname{gr} Y_{i,j} = U(\mathfrak{g}_{i,j})$ and

$$Z(\mathfrak{g}) \cap U(\mathfrak{g}_{i,j}) = \Bbbk\left[\left(e_{i,j}t^r\right)^p - \delta_{i,j}e_{i,j}t^{rp} \,\middle|\, r \geqslant 0\right] \tag{6.35}$$

just like in (5.12) and (5.17). Combined with (6.34), it follows that $Z(Y_n) \cap Y_{i,j}$ is freely generated by $\{S_{i,j}^{(rp)} \mid r > 0\}$; this is exactly the same argument as used in the proofs of Theorems 5.4 and 5.8. The last assertion in the statement of the theorem follows.

Finally, we must prove the first assertion. Lemma 6.8 shows that $S_{i,j}^{(rp)}$ lies in $Z_p(Y_n)$. To show that $\{S_{i,j}^{(rp)} \mid 1 \leqslant i,j \leqslant n, r > 0\}$ are algebraically independent and generate $Z_p(Y_n)$, we pass to the associated graded algebra using (6.34), to see that they are lifts of the generators of $\operatorname{gr} Z_p(Y_n) = Z_p(\mathfrak{g})$ from (3.9). □

Remark 6.10. Similar to Remark 5.9, Theorem 6.9 justifies Remark 2.12.

When combined with Theorem 5.11 and Corollary 5.12, Theorem 6.9 finally establishes all of the statements about $Z(Y_n)$ that we formulated in the introduction. We should also note for this that the central elements $C^{(r)}$ defined by the quantum determinant in the

introduction are the same as the ones arising from (5.1). This is a non-trivial observation which is proved in characteristic zero in [BK1, Theorem 8.6]; the argument there works over $\mathbb{Z}$, hence, also in positive characteristic.

References

[BK1] J. Brundan & A. Kleshchev, Parabolic presentations of the Yangian $Y(\mathfrak{gl}_n)$, *Comm. Math. Phys.* **254** (2005), no. 1, 191–220.

[BK2] ———, Shifted Yangians and finite W-algebras, *Advances Math.* **200** (2006), 136–195.

[BK3] ———, Representations of shifted Yangians and finite W-algebras, *Mem. Amer. Math. Soc.* **196** (2008), no. 918, 107 pages.

[Dix] J. Dixmier, "Enveloping Algebras", Graduate Studies in Mathematics 11, American Mathematical Society, Providence, RI, 1966.

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

[FRT] L. Faddeev, N. Reshetikhin & L. Takhtadzhyan, Quantization of Lie groups and Lie algebras, *Leningrad Math. J.* **1** (1990), 193–225.

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

[GT] S. Goodwin & L. Topley, Modular finite W-algebras; `arXiv:1705.06223`.

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

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

[P1] A. Premet, Irreducible representations of Lie algebras of reductive groups and the Kac-Weisfeiler conjecture, *Invent. Math.* **121** (1995), 79–117.

[P2] ———Special transverse slices and their enveloping algebras, *Advances Math.* **170** (2002), 1–55.

[P3] ———, Commutative quotients of finite W-algebras, *Advances Math.* **225** (2010), 269–306.

[WWY] B. Webster, A. Weekes & O. Yacobi, A quantum Mirković-Vybornov isomorphism; `arXiv:1706.03841`.