

# A SHARP DIAMETER BOUND FOR UNIPOTENT GROUPS OF CLASSICAL TYPE OVER $\mathbb{Z}/p\mathbb{Z}$

JORDAN S. ELLENBERG AND JULIANNA TYMOCZKO

ABSTRACT. The unipotent subgroup of a finite group of Lie type over a prime field  $\mathbb{F}_p$  comes equipped with a natural set of generators; the properties of the Cayley graph associated to this set of generators have been much studied. In the present paper, we show that the diameter of this Cayley graph is bounded above and below by constant multiples of  $np + n^2 \log p$ , where  $n$  is the rank of the associated Lie group. This generalizes the result of [E], which treated the case of  $SL_n(\mathbb{F}_p)$ . (Keywords: diameter, Cayley graph, finite groups of Lie type. AMS classification: 20G40, 05C25)

## 1. INTRODUCTION

Given a group  $G$  endowed with a set of generators  $\Sigma = \{u_1, \dots, u_k\}$ , the *Cayley graph* of  $G$  in  $\Sigma$  is the directed graph whose vertices are elements of  $\Sigma$ , and in which  $x$  and  $y$  are joined if  $y = u_i x$  for some  $i$ . It is a long-standing problem to investigate geometric properties of the Cayley graph; especially interesting is the behavior of the random walk obtained by setting off along a randomly chosen edge at each step. In case  $G$  is a finite group, a natural invariant of the Cayley graph of  $G$  in  $\Sigma$  is its *diameter* – that is, the minimum positive integer  $N$  such that every element of  $G$  can be written as a word of length at most  $N$  in  $\Sigma$ . If  $g$  is an element of  $G$ , we define the *length* of  $g$  in  $\Sigma$  (or just the length of  $g$  if  $\Sigma$  is understood) to be the length of the shortest word in  $\Sigma$  which evaluates to  $g$ .

Diameters of Cayley graphs have been much studied: for a general discussion, see [BHKLS] or [B, §3.8]. An excellent reference for the theory of random walks on finite groups in general is Saloff-Coste's survey paper [SC]. The relationship between the diameter of  $G$  in  $\Sigma$  and the convergence of the random walk along  $\Sigma$  is discussed, for example, in [DSC1, Corollary 1]; for random walks on unipotent groups (in particular, groups of upper-triangular matrices) see [CP], [ADS], and [St]. The problem of controlling the diameters in natural families of groups is quite hard. For a general permutation group with a general set of generators, the problem of computing the diameter is NP-hard [EG]. Bounding the diameter of  $SL_2(\mathbb{F}_p)$  in standard generators, as  $p$  grows, is a well-known problem whose solution (at present) requires the theory of automorphic forms (but see [He] for a new approach via additive number theory.)

In this paper, we give upper and lower bounds for the diameter of unipotent subgroups of classical groups over prime fields; such finite groups admit a natural choice of  $\Sigma$ , which we describe in section 1.1. In case the classical group is  $SL_n$ , we recover the previously unpublished result of the first author [E] cited in [ADS], [DSC2], and [SC]. The main theorem of our paper is the following:

**Theorem 1.1.** *There exists absolute constants  $c$  and  $C$  such that, if  $p$  is an odd prime, and  $G$  is a Chevalley group of classical type of rank  $n$  over  $\mathbb{Z}/p\mathbb{Z}$ , the diameter of  $U$  in  $\Sigma$  is at most  $C(np + n^2 \log p)$  and at least  $c(np + n^2 \log p)$ .*

In the first section of this paper, we outline some of the basic facts we will need about the combinatorics of commutator relations in unipotent groups of classical type; in the second and third sections, respectively, we prove the upper and lower bounds on the diameter of  $U$  in  $\Sigma$ . In an appendix we provide more detailed descriptions of the Chevalley groups as matrix groups, decomposing along the “rows” of the matrix groups.

To begin, we summarize the algebra needed for this task. Subsection 1.1 gives a brief introduction to Chevalley groups and Subsection 1.2 details a decomposition of  $G$  upon which we rely heavily. Subsection 1.3 uses both to give generators for the subgroups we will study.

**Acknowledgments.** The authors are grateful to Persi Diaconis, under whose supervision the first author completed a special case of the present work; and to Igor Pak, for many helpful comments and suggestions. The first author is partially supported by NSF Grant DMS-0401616. The second author is partially supported by an NSF Postdoctoral Fellowship and received support from the Clay Mathematics Institute during the writing of this paper.

**1.1. Properties of Chevalley groups of classical type.** Let  $K$  be a field of characteristic other than 2, and let  $G$  be a Chevalley group of classical type over the field  $K$ : that is,  $G$  is either a special linear group, a special orthogonal group, or a symplectic group over  $K$ . Because  $G$  is a Chevalley group, its multiplicative structure is described combinatorially by a set of roots and the relations between these roots. In this section we review the combinatorics associated to  $G$  and some basic results we will need for our arguments. The Appendix provides explicit matrix representatives for all these constructions.

Let  $B$  be a Borel subgroup of  $G$ , and let  $U$  be the unipotent subgroup of  $B$ ; so  $B$  factors as a product  $TU$  where  $T$ , the maximal torus contained in  $B$ , is isomorphic to  $(K^*)^n$  for an integer  $n$  called the rank of  $G$ . For more details about the structure of Chevalley groups, see [Hu1, Chapters 8, 28, and 35].

In this paper, we will give upper and lower bounds for the diameter of  $U$  in a natural set of generators described below, in the case where  $K$  is  $\mathbb{Z}/p\mathbb{Z}$  for some odd prime  $p$ .

We begin by reviewing several properties of  $U$ . First,  $U$  is generated by a collection of one-parameter subgroups  $U_\alpha = \{\varepsilon_\alpha(c) : c \in K\}$  called *root subgroups*; here  $\alpha$  ranges over the roots of  $G$ , and the group homomorphism  $\varepsilon_\alpha : K \rightarrow G$  is obtained by exponentiating the eigenspace of the Lie algebra of  $G$  corresponding to  $\alpha$ . The set of  $\alpha$  such that  $\varepsilon_\alpha(K) \subset U$  is called the set of positive roots, and is denoted  $\Phi^+$ .

The positive roots inherit an operation from the character group which we denote addition. (Warning: the set of positive roots is not closed under this operation!) A minimal generating set of  $\Phi^+$  will have cardinality  $n$ ; denote such a generating set by  $\Delta = \{\alpha_i : 1 \leq i \leq n\}$ . These  $\alpha_i$  are called the simple roots. For example, in type  $A_n$  (i.e. when  $G = SL_{n+1}$ ) the subgroup  $U$  is the group of upper-triangular matrices with 1 along the diagonal while the simple roots correspond to the root subgroups  $U_{\alpha_i} = \{\text{Id} + cE_{i,i+1} : c \in K\}$ . More examples of explicit matrix representatives are

in the Appendix. We use the conventions of [Hu2] to index the simple roots; Figure 1 describes all the positive roots of  $G$  with respect to this choice of indexing.

The set  $\Sigma \subset G$  of elements of the form  $\varepsilon_{\alpha_i}(\pm 1)$  is a generating set for  $G$  of cardinality  $2n$ . The first part of this paper will be devoted to proving the following proposition:

**Proposition 1.** *There exists an absolute constant  $C$  such that, if  $p$  is an odd prime, and  $G$  is a Chevalley group of classical type over  $Z/pZ$ , the diameter of  $U$  in  $\Sigma$  is at most  $C(np + n^2 \log p)$ .*

*Remark 1.2.* The constant  $C$  is easy to bound explicitly, if desired.

*Remark 1.3.* The difficult part of the argument is controlling the behavior of the diameter of  $G$  as  $n$  grows; this is why we do not consider diameters of exceptional groups of Lie type, in which only the growth of  $p$  is at issue.

The key idea driving our argument is that the non-commutativity of  $U$  can be described in completely combinatorial terms by means of the Chevalley commutator relation, which is the following:

$$(1) \quad [\varepsilon_{\alpha}(s), \varepsilon_{\beta}(t)] = \prod_{\substack{i, j > 0 \\ i\alpha + j\beta \in \Phi^+}} \varepsilon_{i\alpha + j\beta}(c_{ij}s^i t^j)$$

(see [S, page 207]). Brackets denote the group commutator  $[x, y] = xyx^{-1}y^{-1}$  in  $U$ . The integers  $c_{ij}$  are determined by the roots  $\alpha$  and  $\beta$  as well as the ordering of the product; we list many of them in the Appendix. The following proposition is the only fact we will use about the  $c_{ij}$ .

**Proposition 2.** *If  $\text{char } K \neq 2$  then  $c_{ij} \neq 0$  if and only if  $i\alpha + j\beta \in \Phi^+$ .*

The additive relations between roots also play a key role in our computations. Figure 1 enumerates the positive roots of classical Chevalley groups, for  $i$  between 1 and  $n = rk(G)$ .

Root	Parameters	Type
$\sum_{j=i}^k \alpha_j$	$1 \leq i \leq k \leq n$	$A_n, B_n, C_n, D_n$
except $\alpha_{n-1} + \alpha_n \notin \Phi$ in type $D_n$		
$\sum_{j=i}^n \alpha_j + \sum_{j=k}^n \alpha_j$	$1 \leq i < k \leq n$	$B_n$
$\sum_{j=i}^n \alpha_j + \sum_{j=k}^{n-1} \alpha_j$	$1 \leq i \leq k < n$	$C_n$
$\sum_{j=i}^{n-2} \alpha_j + \alpha_n$	$1 \leq i \leq n-2$	$D_n$
$\sum_{j=i}^n \alpha_j + \sum_{j=k}^{n-2} \alpha_j$	$1 \leq i < k \leq n-2$	$D_n$

FIGURE 1. Positive roots of groups of classical type

As we can see in Figure 1, each positive root can be written uniquely as a linear combination of simple roots with nonnegative integer coefficients. The sum of these coefficients is called the height of the root and is denoted  $\text{ht}(\sum c_i \alpha_i) = \sum c_i$ . We write  $\alpha > \beta$  if  $\alpha - \beta$  is a sum of positive roots.

For a thorough introduction to algebraic groups and root systems, the reader is referred to [S] or [Hul].

**1.2. Decomposing  $U$  into  $U_i$ .** To describe the diameter of  $U$  we decompose  $U$  as a product of subgroups  $U_i$ , by analogy with the decomposition of the group of upper triangular matrices as a product of the subgroups consisting of matrices with zero entries away from the diagonal and row  $i$ . In this section we define the  $U_i$  and prove some of their properties. In later sections we will bound the diameter of each  $U_i$  using the Chevalley commutator relation.

We begin by defining a subset  $(\Phi^i)'$  of  $\Phi^+$ :

$$(\Phi^i)' = \{\alpha_i\} \cup \{\alpha \in \Phi^+ \mid \alpha > \alpha_i, \nexists j < i \text{ such that } \alpha > \alpha_j\}.$$

For instance, in type  $A_3$  these sets are  $(\Phi^1)' = \{\alpha_1, \alpha_1 + \alpha_2, \alpha_1 + \alpha_2 + \alpha_3\}$ ,  $(\Phi^2)' = \{\alpha_2, \alpha_2 + \alpha_3\}$ , and  $(\Phi^3)' = \{\alpha_3\}$ . By contrast, in type  $D_5$  the set  $(\Phi^2)'$  is  $(\Phi^2)' = \{\alpha_2, \alpha_2 + \alpha_3, \alpha_2 + \alpha_3 + \alpha_4, \alpha_2 + \alpha_3 + \alpha_5, \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5, \alpha_2 + 2\alpha_3 + \alpha_4 + \alpha_5\}$ . Note that, for fixed  $i$ , the roots given in the first column of Figure 1 are precisely those contained in  $(\Phi^i)'$ .

Let  $U_i'$  be the subgroup of  $U$  generated by  $\{U_\alpha : \alpha \in (\Phi^i)'\}$ . The Chevalley relations and Figure 1 show that  $U_i'$  is abelian in types  $A_n, B_n$ , and  $D_n$ , and Heisenberg in type  $C_n$ . In other words, since the set  $\{\alpha + \beta : \alpha \in (\Phi^i)', \beta \in (\Phi^i)', \alpha + \beta \in \Phi^+\}$  is empty in types  $A_n, B_n$ , and  $D_n$ , the elements of  $U_i'$  commute. However, in type  $C_n$  there is a unique root  $\gamma_i = 2 \sum_{j=i}^{n-1} \alpha_j + \alpha_n$  that has the property

$$\{\gamma_i\} = \{\alpha + \beta : \alpha \in (\Phi^i)', \beta \in (\Phi^i)', \alpha + \beta \in \Phi^+\}.$$

Moreover, for each  $\alpha \neq \gamma_i$  in  $(\Phi^i)'$  the root  $\gamma_i - \alpha$  is also in  $(\Phi^i)'$ . The root  $\gamma_i$  is called the long root in  $(\Phi^i)'$ .

The roots in  $(\Phi^i)'$  in type  $A_n$  have the form  $\sum_{j=i}^k \alpha_j$  for  $k = i, i+1, \dots, n$  and so are totally ordered by height. In fact, if the heights of  $\alpha$  and  $\beta$  in  $(\Phi^i)'$  differ by one then  $\alpha - \beta = \pm \alpha_j$  for some simple root  $\alpha_j$ . Moreover, if  $\alpha$  is in  $(\Phi^i)'$  and  $\alpha + \alpha_j$  is a root then  $\alpha + k\alpha_j$  is not a root for any  $k > 1$ .

For a general Chevalley group  $G$ , we will define  $\Phi^i$  and  $U_i$  so as to preserve these properties to the greatest extent possible. Our definitions are given in Figure 2. For instance, in type  $A_n$ , the group  $U_i$  consists of matrices  $\{(\text{Id} + \sum_{j=i+1}^n a_j E_{ij}) : a_j \in K\}$ . The Appendix gives detailed descriptions of the  $U_i'$  in the other types.

Note that in types  $B_n, C_n$ , and  $D_n$ , the  $U_i$  are abelian quotient groups of  $U_i'$ . The images of the root subgroups  $U_\alpha$  for  $\{\alpha \in \Phi^i\}$  in the quotient  $U_i' \rightarrow U_i$  generate  $U_i$ . By an abuse of notation, we also use  $U_\alpha$  and  $\varepsilon_\alpha$  to denote the image of  $U_\alpha$  and  $\varepsilon_\alpha$  under the homomorphism  $U_i' \rightarrow U_i$ .

Type	$U_i$	$\Phi^i$
$A_n$	$U_i'$	$(\Phi^i)'$
$B_n$	$U_i' / U_{\sum_{j=i}^n \alpha_j}$	$(\Phi^i)' - \{\sum_{j=i}^n \alpha_j\}$
$C_n$	$U_i' / U_{\gamma_i}$	$(\Phi^i)' - \{\gamma_i\}$
$D_n$	$U_i' / U_{\sum_{j=i}^{n-2} \alpha_j + \alpha_n}$	$(\Phi^i)' - \{\sum_{j=i}^{n-2} \alpha_j + \alpha_n\}$

FIGURE 2. Definition of  $U_i$  and  $\Phi^i$

For all Chevalley groups of classical type, the group  $U$  is the product

$$U = \prod_{i=1}^n U_i'$$

[S, 11.3.3 Exercise 2a]. If  $g = \prod_{\alpha \in \Phi^i} \varepsilon_\alpha(s_\alpha)$  and  $g' = \prod_{\alpha \in \Phi^i} \varepsilon_\alpha(s'_\alpha)$  are in  $U_i$  then by the commutativity of  $U_i$  their product is

$$(2) \quad gg' = \prod_{\alpha \in \Phi^i} \varepsilon_\alpha(s_\alpha + s'_\alpha).$$

The next fact motivates the choice we have made of  $\Phi^i$ . It follows by inspection of Figure 1. For each type, define  $r_i$  to be the cardinality of  $\Phi^i$ .

**Proposition 3.**  $\Phi^i$  is totally ordered by height. If  $\alpha$  is in  $\Phi^i$  and  $\beta$  is a simple root then  $j\alpha + l\beta$  is in  $\Phi^i$  for exactly one positive pair  $(j, l)$ , with  $j = 1$  and  $l \in \{1, 2\}$ .

In types  $A_n, C_n$ , and  $D_n$  the set  $\Phi^i$  contains exactly one root of each height from 1 to  $r_i$ . In type  $B_n$  the set  $\Phi^i$  contains exactly one root of each height from 1 to  $n - i$  and one root of each height from  $n - i + 2$  to  $r_i + 1$ .

**1.3. Generators for  $U_i$ .** In the rest of this section, we assume that  $i$  has been fixed, that  $\Phi^i$  has been ordered by height from smallest to largest, and that  $\beta_j$  denotes the  $j^{\text{th}}$  element of  $\Phi^i$  in this order.

For instance, in type  $A_n$  the matrices  $\varepsilon_{\beta_j}(t_j)$  have  $t_j$  in the  $(i, j + i)$  entry, 1's along the diagonal, and zeroes elsewhere.

The following proposition uses the Chevalley commutator relation to generate elements of large-height root subgroups by means of reasonably short words.

**Proposition 4.** Define functions

$$\begin{aligned} f : \{1, 2, \dots, r_i\} &\longrightarrow \{i, i + 1, \dots, \text{rk}(G)\} \text{ and} \\ m : \{1, 2, \dots, r_i\} &\longrightarrow \{1, 2\} \end{aligned}$$

by the equations  $\beta_j - \beta_{j-1} = m(j)\alpha_{f(j)}$  (when  $j \geq 2$ ), as well as  $m(1) = 1$  and  $f(1) = i$ .

If  $\beta_{j-1} \in \Phi^i$  and  $l \in \{1, 2, \dots, n\}$  then

$$[\varepsilon_{\beta_{j-1}}(s), \varepsilon_{\alpha_l}(t)] = \begin{cases} \varepsilon_{\beta_j}(c_{1,m(j)}st^{m(j)}) & \text{if } f(j) = l \\ \varepsilon_{\beta_j}(0) & \text{if } f(j) \neq l. \end{cases}$$

Here  $c_{1,m(j)}$  is the structure constant arising in Equation (1). Note that the equality in Proposition 4 holds in  $U_i$ ; the corresponding equality in  $U'_i$  does not hold in general. However, the equality *does* hold in  $U'_i$  when  $j \leq n - i$ .

*Proof.* The functions  $f$  and  $m$  are well defined by inspection of the tables in Figures 1 and 2. In fact, note that  $m(j) = 2$  only when  $j = n - i + 1$  in type  $B_n$ .

By Proposition 3, there is at most one root in  $\Phi^i$  that is a linear combination of  $\beta_{j-1}$  and  $\alpha_l$  with positive integer coefficients. If it exists, this root is  $\beta_j$ .

The claim is now an application of the Chevalley commutator relation.  $\square$

This proposition permits us to define a map from  $K^r$  to  $U_i$  that generates elements of  $U_i$  via successive conjugation by an element of a simple root subgroup. It is this map that allows us quickly to generate elements of  $U$  with large matrix entries.

**Definition 1.** Define the map  $\theta_{k,r} : K^r \longrightarrow U_i$  inductively by

$$\begin{aligned} \theta_{k,1}(s) &= \varepsilon_{\beta_k}(s) \\ \theta_{k,r}(s_1, \dots, s_r) &= (\varepsilon_{\alpha_{f(k+r-1)}}(-s_r)) (\theta_{k,r-1}(s_1, \dots, s_{r-1})) (\varepsilon_{\alpha_{f(k+r-1)}}(s_r)) \end{aligned}$$

$\theta_{k,r}$  is well-defined only if  $k+r-1 \leq r_i$ . For example, if  $n \geq 3$  in any classical type then the map  $\theta_{1,3}$  is

$$\theta_{1,3}(s_1, s_2, s_3) = \varepsilon_{\alpha_3}(-s_3) (\varepsilon_{\alpha_2}(-s_2) (\varepsilon_{\alpha_1}(s_1)) \varepsilon_{\alpha_2}(s_2)) \varepsilon_{\alpha_3}(s_3).$$

The matrix for this product when  $i = 1$  in type  $A_3$  (i.e. when  $G = SL_4$ ) is

$$\begin{pmatrix} 1 & s_1 & s_1 s_2 & s_1 s_2 s_3 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Note that  $\theta_{k,r}(s_1, \dots, s_r) \theta_{k,r}(-s_1, s_2, s_3, \dots, s_r) = 1$ .

Since the image of  $\theta_{k,r}$  is in  $U_i$  it factors uniquely as a product of elements of root subgroups, described next.

**Proposition 5.** Write  $\theta_{k,r}(s_1, \dots, s_r) = \prod_{l=1}^{r_i} \varepsilon_{\beta_l}(t_l)$ .

- (1) If  $l < k$  or  $k+r-1 < l$  then  $t_l = 0$ .
- (2) Fix  $k$  and  $r$  so that either  $n-i+1 \leq k$  or  $k+r-1 \leq n-i+1$ . If  $k \leq l \leq k+r-1$  then  $t_l = s_k \prod_{j=k+1}^l c_{1,m(j)} s_j^{m(j)}$ .

*Proof.* The proof inducts on  $r$ . The claim holds by definition if  $r = 1$ . Note

$$\theta_{k,r}(s_1, \dots, s_r) = (\varepsilon_{\alpha_{k+r-1}}(-s_r)) (\theta_{k,r-1}(s_1, \dots, s_{r-1})) (\varepsilon_{\alpha_{k+r-1}}(s_r))$$

by construction. Use induction to write  $\theta_{k,r-1}(s_1, \dots, s_{r-1}) = \prod_{l=k}^{k+r-2} \varepsilon_{\beta_l}(t'_l)$ . Proposition 4 shows that

$$\begin{aligned} & (\varepsilon_{\alpha_{k+r-1}}(-s_r)) \left( \prod_{l=k}^{k+r-2} \varepsilon_{\beta_l}(t'_l) \right) (\varepsilon_{\alpha_{k+r-1}}(s_r)) \\ &= \left( \prod_{l=k}^{k+r-2} \varepsilon_{\beta_l}(t'_l) \right) \prod_{\substack{l \text{ s.t. } k \leq l-1 \leq k+r-2 \\ \text{and } f(l) = k+r-1}} \varepsilon_{\beta_l}(c_{1,m(l)} t'_{l-1} s_r^{m(l)}). \end{aligned}$$

This can be written as  $\prod_{l=k}^{k+r-1} \varepsilon_{\beta_l}(t_l)$  for some  $t_l$ . In particular, Part 1 holds.

By inspection of Figure 1, there are at most two  $\beta_l$  in  $\Phi^i$  with  $\beta_l - \beta_{l-1} = m(l)\alpha_{k+r-1}$ , of which one has  $l-1 < n-i+1$  while the other has  $l-1 \geq n-i+1$ . By hypothesis in Part 2, at most one of these roots is in  $\{\beta_k, \dots, \beta_{k+r-1}\}$ . The root  $\beta_{k+r-1}$  is such a root by definition of  $\theta_{k,r}$ . So  $t_l = t'_l$  for  $l = k, \dots, k+r-2$  and  $t_{k+r-1} = c_{1,m(k+r-1)} t_{k+r-2} s_r^{m(k+r-1)}$ . Induction completes the proof.  $\square$

## 2. PROOF OF PROPOSITION 1

In this section, we will prove Proposition 1. The main idea is to show explicitly how to generate elements of  $U_i$  with relatively short words in  $\Sigma$ . We accomplish this by further decomposing each  $U_i$  into three smaller subgroups, each generated by root subgroups whose heights lie in a specified range.

**2.1. The contribution of the simple root subgroups.** Each  $U_i$  contains the simple root subgroup  $U_{\alpha_i}$ . We can generate an element  $\varepsilon_{\alpha_i}(k)$  of  $U_{\alpha_i}$  by means of at most  $p/2$  copies of  $\varepsilon_{\alpha_i}(\pm 1)$ .

**2.2. The contribution of medium-height root subgroups.** Fix  $k \leq r_i$ . Let  $U_i^k$  be the subgroup  $\prod_{j=2}^k U_{\beta_j}$  of  $U_i$  and let  $U_i^{>k}$  be the subgroup  $U_i^{>k} = \prod_{j=k+1}^{r_i} U_{\beta_j}$ . We study  $U_i^k$  in this section and  $U_i^{>k}$  in the next.

The following lemma expresses each generator of the root subgroup  $U_{\beta_l}$  in terms of  $\Sigma$ . Recall that  $c_{1,m(j)}$  is the structure constant defined by  $[\varepsilon_{\beta_{j-1}}(1), \varepsilon_{\alpha_{f(j)}}(1)]$  in  $U_i$ , as in Proposition 4. Denote the product of the first  $l$  structure constants by

$$d_l = \prod_{j=2}^l c_{1,m(j)}.$$

(Note that  $|d_l|$  is a power of 2.)

**Lemma 1.** *For each  $\beta_l$  in  $\Phi^i$ , the generator  $\varepsilon_{\beta_l}(d_l)$  of  $U_{\beta_l}$  in  $U_i$  can be written as a word in  $\varepsilon_{\alpha_j}(\pm 1)$  with at most  $8l$  letters.*

*Proof.* We write  $\mathbf{1}^l$  for the vector in  $K^l$  each of whose entries is 1.

First, assume  $l \leq n - i + 1$  or, equivalently,  $\beta_l \leq \sum_{j=i}^{n-1} \alpha_j + 2\alpha_n$ . (The sum  $\sum_{j=i}^{n-1} \alpha_j + 2\alpha_n$  is a root only in type  $B_n$ , where it equals  $\beta_{n-i+1}$ . In the other types, the condition  $\beta_l \leq \sum_{j=i}^{n-1} \alpha_j + 2\alpha_n$  is equivalent to  $\beta_l \leq \sum_{j=i}^n \alpha_j$ .) From Proposition 5 we have that

$$\begin{aligned} \theta_{1,l}(\mathbf{1}^l) \theta_{1,l-1}(-1, \mathbf{1}^{l-2}) &= \theta_{1,l-1}(\mathbf{1}^{l-1}) \theta_{1,l-1}(-1, \mathbf{1}^{l-2}) \varepsilon_{\beta_l}(d_l) \\ &= \varepsilon_{\beta_l}(d_l). \end{aligned}$$

By construction, the left hand side is a word of length at most  $4l$  in the letters  $\varepsilon_{\alpha_j}(\pm 1)$ . The element  $\varepsilon_{\beta_l}(d_l)$  gives a generator for  $U_{\beta_l}$  in  $U_i$  since  $p$  is odd. Note that  $\varepsilon_{\beta_{n-i+1}}(-d_{n-i+1}) = (\varepsilon_{\beta_{n-i+1}}(d_{n-i+1}))^{-1}$  also has length at most  $4(n - i + 1)$ .

Now suppose  $l > n - i + 1$ . The generator  $\varepsilon_{\beta_l}(d_l)$  can be written as

$$\theta_{n-i+1, l-n+i}(d_{n-i+1}, \mathbf{1}^{1-n+i-1}) \theta_{n-i+1, l-n+i-1}(-d_{n-i+1}, \mathbf{1}^{1-n+i-2}).$$

The length of this word is at most  $8(n - i + 1) + 4(l - n + i - 1) \leq 8l$ .  $\square$

We now bound the length of  $\varepsilon_{\beta_l}(sd_l)$  for arbitrary  $s$  in  $K$ .

**Lemma 2.** *The length of  $\varepsilon_{\beta_l}(sd_l)$  in  $U_i$  is at most  $48l\sqrt{s}$ .*

*Proof.* By abuse of notation we also use  $s$  to denote the integer in the real interval  $[0, p - 1]$  which reduces to  $s \bmod p$ . Let  $t$  be the largest integer less than or equal to  $\sqrt{s}$ :

$$t = \lfloor \sqrt{s} \rfloor.$$

Since the  $\varepsilon_{\beta_l}$  are additive homomorphisms, we know that

$$\varepsilon_{\beta_l}(sd_l) = \varepsilon_{\beta_l}(t^2 d_l) \varepsilon_{\beta_l}((s - t^2)d_l).$$

By the commutator relationship in  $U_i$ ,

$$\begin{aligned} \varepsilon_{\beta_l}(t^2 d_l) &= [\varepsilon_{\beta_{l-1}}(td_{l-1}), \varepsilon_{\alpha_{f(l)}}(t)] \text{ and} \\ \varepsilon_{\beta_l}((s - t^2)d_l) &= [\varepsilon_{\beta_{l-1}}(d_{l-1}), \varepsilon_{\alpha_{f(l)}}(s - t^2)]. \end{aligned}$$

When  $l = n - i + 1$  in type  $B_n$ , we use the relations

$$\begin{aligned} \varepsilon_{\beta_l}(t^2 d_l) &= [\varepsilon_{\beta_{l-1}}(d_{l-1}), \varepsilon_{\alpha_{f(l)}}(t)] \text{ and} \\ \varepsilon_{\beta_l}((s - t^2)d_l) &= [\varepsilon_{\beta_{l-1}}((s - t^2)d_l), \varepsilon_{\alpha_{f(l)}}(1)]. \end{aligned}$$

Using the inequalities

$$s - t^2 \leq ((t + 1)^2 - 1) - t^2 \leq 2t \leq 2\sqrt{s},$$

and the previous lemma we conclude that the elements  $\varepsilon_{\beta_{l-1}}(td_{l-1})$  and  $\varepsilon_{\beta_{l-1}}(d_{l-1})$  have length at most  $(8l - 8)\sqrt{s}$  and  $8l - 8$ . In the type  $B_n$  calculation, note that the element  $\varepsilon_{\beta_{l-1}}((s - t^2)d_{l-1})$  has length at most  $(8l - 8)(2\sqrt{s})$ . In all types, the elements  $\varepsilon_{\alpha_{f(l)}}(t)$  and  $\varepsilon_{\alpha_{f(l)}}(s - t^2)$  have length at most  $\sqrt{s}$  and  $2\sqrt{s}$ , respectively.

Consequently, the total length of  $\varepsilon_{\beta_l}(sd_l)$  is at most  $48l\sqrt{s}$ .  $\square$

**Proposition 6.** *Let  $k = \lfloor \frac{\log p}{\log 2} \rfloor$ . Every element of  $U_i^k$  can be expressed by a word in  $\Sigma$  of length at most*

$$\frac{48}{\log^2 2} \sqrt{p} \log^2 p.$$

*Proof.* Since  $s$  is at most  $p$ , the diameter of  $U_i^k$  is at most

$$\sum_{l=2}^k 48l\sqrt{p} < 48k^2\sqrt{p} \leq \frac{48}{\log^2 2} \sqrt{p} \log^2 p.$$

$\square$

Note that the constant here is not meant to be optimal; in particular, one can do better by specifying the type of the Chevalley group.

**2.3. The contribution of large-height root subgroups.** The next step is to bound the diameter of  $U_i^{>k}$ , which we do using a kind of binary expansion as described below.

Fix an element  $g$  in  $U_i^{>k}$  and write  $g = \prod_{l=k+1}^{r_i} \varepsilon_{\beta_l}(s_l)$ . Define the vector

$$\mathbf{u} = (s_{k+1}, s_{k+2}, \dots, s_{r_i}) \in K^{r_i - k}.$$

Recall that  $d_l = \prod_{j=2}^l c_{1,m(j)}$  is the product of structure constants. Each entry  $\frac{s_l}{2d_l}$  is in  $K$  and so (considered as an integer between 0 and  $p - 1$ ) has a binary decomposition. Define a function  $b(l, j)$  by

$$b(l, j) = \begin{cases} 0 & \text{if } 2^j \text{ is not in the binary decomposition of } \frac{s_l}{2d_l} \\ 1 & \text{if } 2^j \text{ is in the binary decomposition of } \frac{s_l}{2d_l}. \end{cases}$$

Define

$$\mathbf{u}_j = 2^{j+1} (b(k+1, j)d_{k+1}, \dots, b(r_i, j)d_{r_i}).$$

so that  $\sum_j \mathbf{u}_j = \mathbf{u}$ . We may think of the  $\mathbf{u}_j$  as making up a “binary decomposition” of  $\mathbf{u}$ . If  $s_{j,l}$  denotes the  $l^{\text{th}}$  entry of  $\mathbf{u}_j$ , then

$$\prod_{j=0}^{\lfloor \frac{\log p}{\log 2} \rfloor} \prod_{l=k+1}^{r_i} \varepsilon_{\beta_l}(s_{j,l}) = \prod_{l=k+1}^{r_i} \varepsilon_{\beta_l}(s_l) = g.$$

Thus the length of  $g$  is bounded by the combined lengths of the words  $\prod_{l=k+1}^{r_i} \varepsilon_{\beta_l}(s_{j,l})$ .

We first study words whose vectors  $\mathbf{u}_j$  are zero in the last  $r_i - n + i - 1$  entries.

**Lemma 3.** *Let  $\mathbf{u} \in K^{n-i+1-j}$  be a vector with entries  $s_l \in \{0, d_{l+j}2^{j+1}\}$ . The word  $\prod_{l=j+1}^{n-i+1} \varepsilon_{\beta_l}(s_{l-j})$  has length at most  $8(n - i + 1)$ .*



*Proof.* Define vectors  $\mathbf{a}_1$  and  $\mathbf{a}_2$  by

$$a_{1,l} = \begin{cases} 2 & 1 \leq l \leq j \\ 1 & j < l \leq n - i + 1 \end{cases}$$

and

$$a_{2,l} = \begin{cases} -2 & l = 1 \\ 2 & 1 < l \leq j \\ (-1)^{\frac{s_{l-j}}{d_l 2^{j+1}} + \frac{s_{l-j-1}}{d_{l-1} 2^{j+1}}} & j < l \leq n - i + 1 \end{cases}$$

More concretely,  $a_{2,l}$  is  $+1$  whenever  $s_{l-j} = s_{l-j-1}$ , and is  $-1$  otherwise.

By construction,  $\theta_{1,n-i+1}(\mathbf{a}_1)$  and  $\theta_{1,n-i+1}(\mathbf{a}_2)$  each have length no greater than  $4(n - i + 1)$ . The vectors were also constructed so that

$$\theta_{1,l}(a_{1,1}, \dots, a_{1,l})\theta_{1,l}(a_{2,1}, \dots, a_{2,l}) = 1$$

for each  $l \leq j$ . When  $l > j$ , we use Proposition 5 to compute

$$\begin{aligned} & \theta_{1,l}(a_{1,1}, \dots, a_{1,l})\theta_{1,l}(a_{2,1}, \dots, a_{2,l}) \\ &= \theta_{1,l-1}(a_{1,1}, \dots, a_{1,l-1})\theta_{1,l-1}(a_{2,1}, \dots, a_{2,l-1})\varepsilon_{\beta_l} \left( d_l \left( 2^j - 2^j (-1)^{\frac{s_{l-j}}{d_l 2^{j+1}}} \right) \right) \\ &= \theta_{1,l-1}(a_{1,1}, \dots, a_{1,l-1})\theta_{1,l-1}(a_{2,1}, \dots, a_{2,l-1})\varepsilon_{\beta_l}(s_{l-j}). \end{aligned}$$

By induction  $\theta_{1,n-i+1}(\mathbf{a}_1)\theta_{1,n-i+1}(\mathbf{a}_2) = \prod_{l=j+1}^{n-i+1} \varepsilon_{\beta_l}(s_{l-j})$ , which proves the claim.  $\square$

We now consider the case where  $\mathbf{u}_j$  is zero in all but the last  $r_i - n + i - 1$  coordinates. This can only happen in types  $B_n$ ,  $C_n$ , and  $D_n$ .

**Lemma 4.** Fix  $j \leq r_i - 1$  and denote  $\max\{j, n - i + 1\}$  by  $q$ . Let  $\mathbf{u} \in K^{r_i - q}$  be a vector whose entries have each  $s_l \in \{0, d_{l+q} 2^{j+1}\}$ . The word  $\prod_{l=q+1}^{r_i} \varepsilon_{\beta_l}(s_{l-q})$  has length at most  $24r_i$ .

*Proof.* Define vectors  $\mathbf{a}_1$  and  $\mathbf{a}_2$  by

$$a_{1,l} = \begin{cases} 2 & 1 \leq l \leq j - n + i - 1 \\ 1 & j - n + i - 1 < l \leq r_i - n + i - 1 \end{cases}$$

(in particular, if  $j \leq n - i + 1$  then  $a_{1,l} = 1$  for all  $l$ ) and

$$a_{2,l} = \begin{cases} 2 & 1 \leq l \leq j - n + i - 1 \\ (-1)^{\frac{s_{l-q}}{d_l 2^{j+1}} + \frac{s_{l-q-1}}{d_{l-1} 2^{j+1}}} & j - n + i - 1 < l \leq r_i - n + i - 1. \end{cases}$$

Write  $t = d_{n-i+1} 2^j$  if  $j < n - i + 1$  and  $t = d_{n-i+1} 2^{n-i+1}$  otherwise. The previous lemma showed that  $\varepsilon_{\beta_{n-i+1}}(\pm t)$  is a word of length at most  $8(n - i + 1)$ .

By construction  $\theta_{n-i+1, r_i - n + i}(t, \mathbf{a}_1)$  and  $\theta_{n-i+1, r_i - n + i}(-t, \mathbf{a}_2)$  each have length at most  $2(8(n - i + 1) + 4(r_i - n + i - 1))$ . Since  $r_i \geq 2(n - i)$  in types  $B_n$ ,  $C_n$ , and  $D_n$ , this length is at most  $24r_i$ . As in the previous proof,

$$\theta_{n-i+1, r_i - n + i}(t, \mathbf{a}_1)\theta_{n-i+1, r_i - n + i}(-t, \mathbf{a}_2) = \prod_{l=q+1}^{r_i} \varepsilon_{\beta_l}(s_{l-q}).$$

$\square$

The words  $\prod_{l=k+1}^{r_i} \varepsilon_{\beta_l}(s_{j,l})$  can be constructed by Lemmas 3 and 4 as long as  $k \geq j$ . Since  $j$  ranges only up to  $\lfloor \frac{\log p}{\log 2} \rfloor$ , we have shown that if we take  $k = \lfloor \frac{\log p}{\log 2} \rfloor$ , each word  $\prod_{l=k+1}^{r_i} \varepsilon_{\beta_l}(s_{j,l})$  can be expressed by a word of length at most  $32r_i$ . We thus obtain:

**Proposition 7.** *Every element of  $U_i^{>k}$  can be expressed by a word in  $\Sigma$  of length at most  $32r_i \left( \frac{\log p}{\log 2} + 1 \right)$ .*

**2.4. Generating the kernel of the projection to  $U_i$ .** In type  $A_n$  the groups  $U_i$  and  $U'_i$  are identical. In the other types, the kernel  $V_i$  of the projection from  $U'_i$  to  $U_i$  is a single root subgroup  $\varepsilon_\beta(\mathbb{Z}/p\mathbb{Z})$ . In order to describe completely the generation of  $U'_i$  by words in  $\Sigma$ , it remains only to treat this subgroup.

We now prove that this root subgroup can be generated using the method of Section 2.2 if  $n - i + 1 \leq \lfloor \frac{\log p}{\log 2} \rfloor$  and that of Section 2.3 if not.

Let  $\beta$  be the root whose root subgroup is  $V_i$ . In type  $B_n$  the root  $\beta$  is  $\sum_{j=i}^n \alpha_j$ , while in type  $C_n$  the root  $\beta$  is  $\sum_{j=i}^{n-1} 2\alpha_j + \alpha_n$ , and in type  $D_n$  the root  $\beta$  is  $\sum_{j=i}^{n-2} \alpha_j + \alpha_n$ .

We will show below that the subgroup  $V_i$  can be generated by explicitly given commutators. Define an integer  $d$  by

$$d = \begin{cases} 2d_{n-i}c_{11} & \text{in type } B_n, \\ 2d_{n-i}^2c_{21} & \text{in type } C_n, \text{ and} \\ d_{n-i-1}c_{11} & \text{in type } D_n \end{cases}$$

where the structure constants  $c_{ij}$  are computed with respect to the commutator  $[\varepsilon_{\beta_{n-i}}(1), \varepsilon_{\alpha_n}(1)]$  in type  $B_n$  and  $C_n$ , and with respect to  $[\varepsilon_{\beta_{n-i-1}}(1), \varepsilon_{\alpha_n}(1)]$  in type  $D_n$ , and  $d_l$  is the product of the first  $l$  structure constants as defined above. Note that  $d$  is a power of 2 and in particular is a unit in  $\mathbb{Z}/p\mathbb{Z}$ .

**Lemma 5.** *Every element of  $V_i$  can be written as a word in  $\Sigma$  of length at most  $32(n - i + 1)(\sqrt{p} + 1)$*

*Proof.* As in the proof of Lemma 2, we generate an arbitrary element  $\varepsilon_\beta(sd)$  by means of commutators. As there, set  $t = \lfloor \sqrt{s} \rfloor$ . By contrast to Lemma 2, our commutator identities here lie in  $U'_i$ , not the quotient  $U_i$ . We will use the fact that the root subgroups attached to  $\beta$  and  $\beta_{n-i+1}$  commute with each other, as can be seen by another application of the Chevalley commutator relation.

In type  $B_n$  we have that

$$[\varepsilon_{\beta_{n-i}}(ud_{n-i}), \varepsilon_{\alpha_n}(v)] = \varepsilon_\beta \left( uv \frac{d}{2} \right) \varepsilon_{\beta_{n-i+1}}(uv^2d_{n-i+1})$$

and

$$[\varepsilon_{\beta_{n-i}}(-ud_{n-i}), \varepsilon_{\alpha_n}(-v)] = \varepsilon_\beta \left( uv \frac{d}{2} \right) \varepsilon_{\beta_{n-i+1}}(-uv^2d_{n-i+1}).$$

It follows that

$$\begin{aligned} \varepsilon_\beta(t^2d) &= [\varepsilon_{\beta_{n-i}}(td_{n-i}), \varepsilon_{\alpha_n}(t)][\varepsilon_{\beta_{n-i}}(-td_{n-i}), \varepsilon_{\alpha_n}(-t)] \text{ and} \\ \varepsilon_\beta((s-t^2)d) &= [\varepsilon_{\beta_{n-i}}(d_{n-i}), \varepsilon_{\alpha_n}(s-t^2)][\varepsilon_{\beta_{n-i}}(-d_{n-i}), \varepsilon_{\alpha_n}(t^2-s)]. \end{aligned}$$

Similarly, in type  $C_n$  we have that

$$[\varepsilon_{\beta_{n-i}}(ud_{n-i}), \varepsilon_{\alpha_n}(v)] = \varepsilon_\beta \left( u^2v \frac{d}{2} \right) \varepsilon_{\beta_{n-i+1}}(uvd_{n-i+1})$$

and so the following relations hold:

$$\begin{aligned} \varepsilon_\beta(t^2d) &= [\varepsilon_{\beta_{n-i}}(td_{n-i}), \varepsilon_{\alpha_n}(1)][\varepsilon_{\beta_{n-i}}(-td_{n-i}), \varepsilon_{\alpha_n}(1)] \text{ and} \\ \varepsilon_\beta((s-t^2)d) &= [\varepsilon_{\beta_{n-i}}(d_{n-i}), \varepsilon_{\alpha_n}(s-t^2)][\varepsilon_{\beta_{n-i}}(-d_{n-i}), \varepsilon_{\alpha_n}(s-t^2)]. \end{aligned}$$

In type  $D_n$  we use the relations:

$$\begin{aligned}\varepsilon_\beta(t^2d) &= [\varepsilon_{\beta_{n-i-1}}(td_{n-i-1}), \varepsilon_{\alpha_n}(t)] \text{ and} \\ \varepsilon_\beta((s-t^2)d) &= [\varepsilon_{\beta_{n-i-1}}(d_{n-i-1}), \varepsilon_{\alpha_n}(s-t^2)].\end{aligned}$$

By the same analysis as in Lemma 2, the length of  $\varepsilon_{\beta_{n-i}}(td_{n-i})$  is at most  $8(n-i)\sqrt{s}$  and the length of  $\varepsilon_{\beta_{n-i}}(d_{n-i})$  is at most  $8(n-i)$ . (Note that we are using here the fact that the identity in Proposition 4 is correct in  $U'_i$ , not merely  $U_i$ , whenever  $j \leq n-i$ .) By definition, the length of  $\varepsilon_{\alpha_n}(s-t^2)$  is at most  $2\sqrt{s}$  and the length of  $\varepsilon_{\alpha_n}(t)$  is at most  $\sqrt{s}$ . The total length of  $\varepsilon_\beta(sd)$  is at most  $32(n-i+1)(\sqrt{s}+1)$ , and since  $s$  is no greater than  $p$ , the lemma is proven.  $\square$

**Lemma 6.** *If  $n-i+1 > \lfloor \frac{\log p}{\log 2} \rfloor$  then every element of  $V_i$  can be written as a word in  $\Sigma$  of length at most  $8(n-i+1) \left( \lfloor \frac{\log p}{\log 2} \rfloor + 1 \right)$ .*

*Proof.* The following formulas hold by induction on the Chevalley commutator relation: in type  $B_n$ ,

$$\begin{aligned}\theta_{1,n-i+1}(t_1, t_2, \dots, t_{n-i}, t_{n-i+1})\theta_{1,n-i+1}(-t_1, t_2, \dots, t_{n-i}, -t_{n-i+1}) \\ = \varepsilon_\beta(d \prod_{l=1}^{n-i+1} t_l),\end{aligned}$$

in type  $C_n$ ,

$$\begin{aligned}\theta_{1,n-i+1}(t_1, t_2, \dots, t_{n-i+1})\theta_{1,n-i+1}(-t_1, t_2, \dots, t_{n-i+1}) \\ = \varepsilon_\beta(dt_{n-i+1} \prod_{l=1}^{n-i} t_l^2),\end{aligned}$$

and in type  $D_n$ ,

$$\begin{aligned}\theta_{1,n-i+1}(t_1, t_2, \dots, t_{n-i-1}, 0, t_{n-i+1})\theta_{1,n-i+1}(-t_1, t_2, \dots, t_{n-i-1}, 0, 0) \\ = \varepsilon_\beta(dt_{n-i+1} \prod_{l=1}^{n-i-1} t_l).\end{aligned}$$

To generate  $\varepsilon_\beta(2^j d)$  in types  $B_n$  and  $D_n$ , let  $t_l = 2$  for  $l = 1$  through  $j$  and  $t_l = 1$  otherwise. To generate  $\varepsilon_\beta(2^{2j+1} d)$  in type  $C_n$  let  $t_l = 2$  for  $l = 1$  through  $j$ , let  $t_{n-i+1} = 2$ , and let  $t_l = 1$  otherwise. To generate  $\varepsilon_\beta(2^{2j} d)$  in type  $C_n$ , let  $t_l = 2$  for  $l = 1$  through  $j$ , and  $t_l = 1$  otherwise.

This shows that  $\varepsilon_\beta(2^j d)$  can be written as a word of length at most  $8(n-i+1)$  in types  $B_n$ ,  $C_n$ , and  $D_n$  for each  $j$  between 0 and  $\lfloor \frac{\log p}{\log 2} \rfloor$ . Thus, by means of binary expansion, we can write  $\varepsilon_\beta(sd)$  for any  $s \in \mathbb{Z}/p\mathbb{Z}$  as a word of length at most  $8(n-i+1) \left( \lfloor \frac{\log p}{\log 2} \rfloor + 1 \right)$ .  $\square$

## 2.5. The diameter of $U$ .

**Theorem 1.** *There exists a constant  $C$ , independent of  $n$  and  $p$ , such that the diameter of  $U$  in  $\Sigma$  is at most*

$$C(np + n^2 \log p).$$

*Proof.* We have already observed that  $U$  is the product of the  $U'_i$ , so it suffices to show that every element  $u$  of  $U'_i$  can be written as a short word in  $\Sigma$ .

The results of Sections 2.1, 2.2 and 2.3 show that we can find an element  $\tilde{u}$  of  $U'_i$  such that  $\tilde{u}$  and  $u$  project to the same element of  $U_i$ , and  $\tilde{u}$  is a word in  $\Sigma$  of length at most

$$\frac{p}{2} + \frac{48}{\log^2 2} \sqrt{p} \log^2 p + 32r_i \left( \frac{\log p}{\log 2} + 1 \right).$$

Since  $\tilde{u}u^{-1}$  lies in  $V_i$ , we have from Section 2.4 that  $\tilde{u}u^{-1}$  can be expressed by a word of length  $32(n-i+1)(\sqrt{p}+1)$  in general, and of length  $8(n-i+1)\left(\frac{\log p}{\log 2}+1\right)$  once  $n-i+1 > \left\lfloor \frac{\log p}{\log 2} \right\rfloor$ . In either case,  $u$  can be written as a word of length at most

$$\frac{p}{2} + \frac{48}{\log^2 2} \sqrt{p} \log^2 p + \frac{32}{\log 2} \log p (\sqrt{p}+1) + 32r_i \left( \frac{\log p}{\log 2} + 1 \right) + 8(n-i+1) \left( \frac{\log p}{\log 2} + 1 \right).$$

Now the diameter of  $U$  in  $\Sigma$  is bounded by the sum of the above expression as  $i$  ranges from 1 to  $n$ ; this sum is evidently bounded by

$$\frac{np}{2} + n \left[ \frac{48}{\log^2 2} \sqrt{p} \log^2 p + \frac{32}{\log 2} \log p (\sqrt{p}+1) \right] + 36n^2 + 4n \left( \frac{\log p}{\log 2} + 1 \right)$$

using the fact that  $\sum_{i=1}^n r_i$  is at most  $n^2$  by [Hu2, section 12.2]. This expression is evidently bounded above by a constant multiple of  $np + n^2 \log p$ .  $\square$

### 3. A LOWER BOUND FOR THE DIAMETER

In this section we prove the following lower bound on the diameter of  $G$  in  $\Sigma$ .

**Proposition 8.** *There exists an absolute constant  $c$  such that, if  $p$  is an odd prime, and  $G$  is a Chevalley group of classical type over  $Z/pZ$ , the diameter of  $U$  in  $\Sigma$  is at least  $c(np + n^2 \log p)$ .*

The main idea is to exploit the fact that each generator in  $\Sigma$  commutes with almost all the others: this means that there are many identities among the words in  $\Sigma$  of some fixed length  $\ell$ ; from an upper bound on the number of distinct words of length  $\ell$  we can derive a lower bound on the diameter of  $U$ .

Since the size of our generating set for  $U$  has cardinality  $2n$ , there are at most  $(2n)^m$  words in  $G$  of length  $m$ . Since  $\log |U|$  is on order of  $n^2 \log p$ , this implies immediately that

$$d(U, \Sigma) \geq \log_n |U| = O(n^2 (\log n)^{-1} \log p).$$

We will show by means of commutativity between the elements of  $\Sigma$  that the  $(\log n)^{-1}$  factor can be removed. Note also that the diameter of  $U$  is at least as great as the diameter of the abelianization of  $U$ , which is on order  $np$ . Together, these facts prove Proposition 8.

For any group  $G$  with a generating set  $S$ , let  $v_{G,S}(m)$  be the number of elements of  $G$  expressible as a word of length exactly  $m$  in the elements of  $S$ . We organize these values into a generating function

$$V_{G,S}(t) = \sum_{m=0}^{\infty} v_{G,S}(m) t^m.$$

We take  $\tilde{U}$  to be a group defined by generators and relations as follows: let  $\tilde{U}$  be generated by a set of elements  $\tilde{\Sigma} = \{\tilde{e}_1, \dots, \tilde{e}_n, \tilde{e}'_1, \dots, \tilde{e}'_n\}$ , subject to the relations

$$(3) \quad [\tilde{e}_i, \tilde{e}_j] = [\tilde{e}'_i, \tilde{e}'_j] = [\tilde{e}_i, \tilde{e}'_j] = [\tilde{e}'_i, \tilde{e}_j] = 1$$

for each pair  $i, j$  such that the root subgroups of  $U$  attached to  $\alpha_i$  and  $\alpha_j$  commute with each other. There is a natural surjection  $\phi: \tilde{U} \rightarrow U$  sending  $\tilde{e}_i$  to  $\varepsilon_{\alpha_i}(1)$  and  $\tilde{e}'_i$  to  $\varepsilon_{\alpha_i}(-1)$ . It follows that

$$v_{U,\Sigma}(m) \leq v_{\tilde{U},\tilde{\Sigma}}(m)$$

for all nonnegative integers  $m$ .

**Proposition 9.**

$$v_{\tilde{U}, \tilde{\Sigma}}(m) \leq 2^{n+2}(4 + 2\sqrt{3})^m.$$

*Proof.* The main tool is a result of Cartier and Foata [CF], which shows that for any group  $G$  with generating set  $S = \{s_1, \dots, s_r\}$  whose only relations are commutations between the  $s_i$ ,

$$[V_{G,S}(t)]^{-1} = \sum_T (-t)^{|T|}$$

where  $T$  ranges over all subsets of  $S$  whose elements commute pairwise. For instance, if  $G$  is the free abelian group on  $S$ , then  $T$  ranges over all subsets of  $S$ , and

$$V_{G,S}(t)^{-1} = (1-t)^r.$$

More precisely, we may define  $V_{G,S}(x_1, \dots, x_r)$  to be the generating function whose  $x_1^{m_1} \dots x_r^{m_r}$  coefficient is the number of distinct words in  $G$  composed of  $m_1$  copies of  $s_1$ ,  $m_2$  copies of  $s_2$ , and so on. Then

$$[V_{G,S}(x_1, \dots, x_r)]^{-1} = \sum_T (-1)^{|T|} \prod_{i \in T} x_i.$$

Let  $G$  be the subgroup of  $\tilde{U}$  generated by  $S = \{\tilde{e}_1, \dots, \tilde{e}_n\}$ . We begin by bounding the volume growth of  $G$  with respect to  $S$ .

If  $\mathcal{D}$  is a Dynkin diagram, write  $U(\mathcal{D})$  for the unipotent group attached to  $\mathcal{D}$ , write  $\Sigma(\mathcal{D})$  for its standard set of generators, and write

$$P(\mathcal{D}) = [V_{G,S}(t)]^{-1}$$

where  $G \subset \tilde{U}$  and  $S$  are described as above.

Recall that two generators  $\varepsilon_{\alpha_i}(1)$  and  $\varepsilon_{\alpha_j}(1)$  in  $\Sigma(\mathcal{D})$  commute if and only if the corresponding vertices of  $\mathcal{D}$  are not connected by an edge.

The pairwise-commuting subsets of  $S$  naturally split into those which contain the  $n^{\text{th}}$  generator and those which do not; this yields a recursion relation

$$(4) \quad P(A_n) = P(A_{n-1}) - tP(A_{n-2})$$

for  $n \geq 3$ , which implies that

$$P(A_n) = c_1 \left( \frac{1 + \sqrt{1-4t}}{2} \right)^n + c_2 \left( \frac{1 - \sqrt{1-4t}}{2} \right)^n$$

for some constants  $c_1, c_2$  depending on  $t$ . Write  $u = \sqrt{1-4t}$ .

The initial terms  $P(A_1) = 1-t = (1/4)(3+u^2)$  and  $P(A_2) = 1-2t = (1/2)(1+u^2)$  show that

$$c_1 = \frac{(1+u)^2}{u}, c_2 = \frac{(1-u)^2}{u}.$$

So

$$(5) \quad P(A_n) = \frac{1}{u} \left[ \left( \frac{1+u}{2} \right)^{n+2} - \left( \frac{1-u}{2} \right)^{n+2} \right].$$

In particular, when  $t < 1/4$ , we have  $u > 0$ ; it follows from (5) that  $P(A_n)(t) > 0$ . So all roots of  $P(A_n)$  are at least  $1/4$ . Note that if we write  $P(A_0) = 1, P(A_{-1}) = 1, P(A_{-2}) = 0$ , the recursion (4) is still satisfied.

Since  $B_n, C_n$  have the same edges as  $A_n$ , we have  $P(A_n) = P(B_n) = P(C_n)$ . The polynomial  $P(D_n)$  satisfies the recurrence (4), but has initial values

$$P(D_3) = 1 - 3t + t^2 = P(A_3), P(D_4) = 1 - 4t + 3t^2 - t^3 = P(A_4) - t^3.$$

It follows that

$$P(D_n) = P(A_n) - t^3 P(A_{n-5})$$

for all  $n \geq 3$ , since it is true for  $n = 3, 4$  and since  $P(D_n)$  satisfies (4). In terms of  $u$ , we get

$$P(D_n) = \frac{1}{8u} [f(u) + f(-u)]$$

where

$$f(u) = (1 + 7u - u^2 - u^3) \left( \frac{1+u}{2} \right)^n.$$

Now  $|1 + 7u - u^2 - u^3| \geq |1 - 7u - u^2 + u^3|$  for  $0 \leq u \leq 1$ ; it follows that  $f(u) + f(-u) > 0$  for  $0 \leq u \leq 1$ . Furthermore, if  $u > 1$ , then  $t < 0$ , and it is obvious from the recursion (4) that  $P(D_n)(t) > 0$  for all  $n$ . We conclude that  $\sqrt{1 - 4t} \leq 0$  for all  $t$  such that  $P(D_n)(t) = 0$ .

We have now shown that the roots of  $P(\mathcal{D})$  are all at least  $1/4$ .

Now define

$$Q(\mathcal{D})(t) = [V_{\bar{U}, \bar{\Sigma}}(t)]^{-1}$$

and

$$Q(\mathcal{D})(x_1, \dots, x_n, x'_1, \dots, x'_n) = [V_{\bar{U}, \bar{\Sigma}}(x_1, \dots, x_n, x'_1, \dots, x'_n)]^{-1}.$$

The relations (3) tell us that

$$Q(\mathcal{D})(x_1, \dots, x_n, x'_1, \dots, x'_n) = P(\mathcal{D})(x_1 + x'_1 - x_1 x'_1, \dots, x_n + x'_n - x_n x'_n)$$

from which we have

$$Q(\mathcal{D})(t) = P(\mathcal{D})(2t - t^2).$$

So the roots of  $Q(\mathcal{D})$  are all greater than  $1 - (\sqrt{3}/2)$ . It follows that

$$v_{\bar{U}, \bar{\Sigma}}(m) \leq C(4 + 2\sqrt{3})^m$$

for some constant  $C$ . To bound this constant roughly, write

$$\begin{aligned} C &= \max_m (4 + 2\sqrt{3})^{-m} v_{\bar{U}, \bar{\Sigma}}(m) \\ &\leq \sum_m (4 + 2\sqrt{3})^{-m} v_{\bar{U}, \bar{\Sigma}}(m) = [Q(\mathcal{D})(1 - (\sqrt{3}/2))]^{-1} = [P(\mathcal{D})(1/4)]^{-1} \end{aligned}$$

We can rewrite (5) as

$$\begin{aligned} P(A_n) &= \frac{1}{2^{n+2}} \left[ \frac{1}{u}(1+u)^{n+1} + (1+u)^{n+1} - \frac{1}{u}(1-u)^{n+1} + (1-u)^{n+1} \right] \\ &= \dots = \frac{1}{2^{n+2}} \left( \sum_{i=0}^{n+1} (1+u)^i + (1-u)^i \right). \end{aligned}$$

Setting  $u = 0$  in this expression (that is, evaluating  $P(A_n)(t)$  at  $t = 1/4$ ) we see that

$$P(A_n)(1/4) = 2^{-n-1}(n+2)$$

and

$$P(D_n)(1/4) = P(A_n)(1/4) - (1/4)^3 P(A_{n-5})(1/4) = (n+7)2^{-n-2}.$$

In either case, we have  $[P(\mathcal{D})(1/4)]^{-1} \leq 2^{n+2}$ , which yields the desired result.  $\square$

We have now shown that

$$v_{U,\Sigma}(m) \leq v_{\tilde{U},\tilde{\Sigma}}(m) \leq 2^{n+2}(4 + 2\sqrt{3})^m.$$

So

$$\sum_{m=0}^D v_{U,\Sigma}(m) \leq (1/3)2^{n+3}\sqrt{3}(4 + 2\sqrt{3})^D.$$

On the other hand, if  $D = d(U, \Sigma)$ , it must be the case that

$$\sum_{m=0}^D v_{U,\Sigma}(m) = |U| \geq p^{(1/2)n^2}.$$

We conclude that

$$d(U, \Sigma) \geq [(1/2)n^2 \log(p) - (n + 3) \log 2 + (1/2) \log 3] / \log(4 + 2\sqrt{3})$$

which is evidently bounded below by a constant multiple of  $n^2 \log p$ .

*Remark 3.1.* The discussion here applies equally well to any group endowed with a set of generators obeying the same commutation relations as  $\Sigma$ . For instance, the number of words of length  $m$  in the standard  $n - 1$  generators of the  $n$ -strand Artin braid group is of order at most  $4^m$ . This suggests that the study of random walks on the Artin braid group might resemble that of random walks on the free group on 4 letters; this theme is explored in detail in [VNB].

#### 4. APPENDIX

In this appendix, we give explicit matrix descriptions of the different Chevalley groups discussed in the paper. To begin, in type  $A_n$ , the group  $G$  is  $SL_{n+1}(K)$ ; in type  $B_n$ , the group  $G$  is  $SO_{2n+1}(K)$ ; in type  $C_n$ , the group  $G$  is  $Sp_{2n}(K)$ ; and in type  $D_n$ , the group  $G$  is  $SO_{2n}(K)$ . In all these cases, we will describe a natural embedding of  $G$  into a general linear group  $GL_N(K)$  and describe matrix representatives for generators of the  $U_\alpha$ .

The group  $G$  is the subgroup of  $GL_N(K)$  that preserves an alternating bilinear form in type  $C_n$  and a symmetric bilinear form in types  $B_n$  and  $D_n$ , so  $N$  is either  $2n$  or  $2n + 1$ . We use the bilinear form defined by  $\langle e_i, e_{N+1-i} \rangle = 1$  when  $1 \leq i \leq N/2$  and  $\pm 1$  otherwise, depending on whether the form is symmetric or alternating, to write  $G$  explicitly. We denote by  $E_{ij}$  the matrix having  $ij$ th entry 1 and all other entries 0.

The  $U_\alpha$  may now be described in terms of these explicit matrix representatives. In type  $A_n$ , the subgroups  $U_\alpha$  correspond to the groups  $\{\text{Id} + cE_{ij} : c \in K, i < j\}$ . By contrast, in type  $C$  the  $U_\alpha$  are either of the form

$$\{\text{Id} + c(E_{ij} - E_{2n+1-j, 2n+1-i}) : i < j \leq n, c \in K\}$$

or of the form

$$\{\text{Id} + c(E_{i, 2n+1-j} + E_{j, 2n+1-i}) : i \leq j \leq n, c \in K\}.$$

In type  $B_n$  the  $U_\alpha$  are of the form

$$\{\text{Id} + c(E_{ij} - E_{2n+2-j, 2n+2-i}) : i < j < 2n + 2 - i, j \neq n + 1, i < n + 1, c \in K\}$$

or of the form

$$\{\text{Id} + c(E_{i, n+1} - E_{n+1, 2n+2-i}) - \frac{c^2}{2} E_{i, 2n+2-i} : i < n + 1, c \in K\}.$$

The  $U_\alpha$  in type  $D_n$  are always of the form

$$\{\text{Id} + c(E_{ij} - E_{2n+1-j, 2n+1-i}) : i < j < 2n+1-i, i < n, c \in K\}.$$

A different bilinear form could have been used and other generators could have been chosen. This would only change the specific constants in Figure 3. We chose as we did so that  $U$  is a subgroup of upper triangular matrices for all types.

In type  $A_n$  the group  $U'_i = \{(\text{Id} + \sum_{j=i+1}^n m_{ij} E_{ij}) : m_{ij} \in K\}$  is the group of unipotent upper-triangular matrices with nonzero off-diagonal entries only on the  $i^{\text{th}}$  row. In the other types,  $U'_i$  also has entries on the  $i^{\text{th}}$  row, but the Chevalley type of the group requires additional nonzero entries. For instance, in type  $B_n$  the group  $U'_i$  is

$$\left\{ \text{Id} + \sum_{j=i+1}^{2n+1-i} m_{ij} (E_{ij} - E_{2n+2-j, 2n+2-i}) + \left( \frac{-m_{i, n+1}^2}{2} - \sum_{j=i+1}^n m_{ij} m_{i, 2n+2-j} \right) E_{i, 2n+2-i} : m_{ij} \in K \right\}.$$

In type  $C_n$  the group  $U'_i$  is

$$\left\{ \text{Id} + \sum_{j=i+1}^n m_{ij} (E_{ij} - E_{2n+1-j, 2n+1-i}) + \sum_{j=n+1}^{2n-i} m_{ij} (E_{ij} + E_{2n+1-j, 2n+1-i}) + \left( m_{i, 2n+1-i} + \sum_{j=i+1}^n m_{ij} m_{i, 2n+1-j} \right) E_{i, 2n+1-i} : m_{ij} \in K \right\}$$

and in type  $D_n$  the group  $U'_i$  is

$$\left\{ \text{Id} + \sum_{j=i+1}^{2n-i} m_{ij} (E_{ij} - E_{2n+1-j, 2n+1-i}) : m_{ij} \in K \right\}.$$

If  $i$  is fixed then as  $j > i$  varies, the root subgroups  $U_\alpha$  given above are precisely the root subgroups for each  $\alpha$  in  $(\Phi^i)'$ . With one exception, we take  $\varepsilon_\alpha(c)$  for each  $c$  to be the matrix given in our description of  $U_\alpha$ . For instance, if  $\alpha$  is the root corresponding to a pair  $i, j \leq n$ , then in type  $A_n$  the matrix  $\varepsilon_\alpha(1) = \text{Id} + E_{ij}$  while in type  $B_n$  the matrix  $\varepsilon_\alpha(1) = \text{Id} + E_{ij} - E_{2n+2-j, 2n+2-i}$ . (We do not consider the root subgroups  $\varepsilon_\alpha(1) = \text{Id} + E_{i, 2n+1-i} + E_{i, 2n+1-i} = \text{Id} + 2E_{i, 2n+1-i}$  in type  $C_n$  to be an exception to this rule.) The one exception to this rule is for the simple root  $\alpha_n$  in type  $D_n$ . In this case, we take

$$\varepsilon_{\alpha_n}(1) = \text{Id} - E_{n, n+2} + E_{n-1, n+1},$$

and in general in type  $D_n$  the matrix  $\varepsilon_{\alpha_n}(c) = \text{Id} + c(-E_{n, n+2} + E_{n-1, n+1})$ .

Fix  $\alpha$  in  $(\Phi^i)'$  and take  $j \geq i$ . Assume that  $\alpha + \alpha_j \in \Phi^+$ . The commutator  $[\varepsilon_\alpha(s), \varepsilon_{\alpha_j}(t)]$  always lies in exactly one root subgroup in types  $A_n$  and  $D_n$ , and almost always in types  $B_n$  and  $C_n$ . Using the matrices which generate the  $U_\alpha$  we compute the structure constants  $c_{ij}$  explicitly. The following table gives the values of  $c_{ij}$  in the Chevalley commutator relation for all of the cases we will consider. The row with the ordered pair of roots  $(\alpha, \alpha_j)$  contains the nonzero coefficients in the Chevalley commutator expansion of  $[\varepsilon_\alpha(s), \varepsilon_{\alpha_j}(t)]$ .



Ordered pair	Nonzero coefficients	Conditions
$(\alpha, \alpha_j)$	$c_{11} = 1$	$\alpha \not\propto \alpha_j$ , $j \neq n$ in types $B_n, C_n, D_n$ , $j \neq n-1$ in type $D_n$
$(\alpha, \alpha_j)$	$c_{11} = -1$	$\alpha > \alpha_j$
$(\sum_{j=i}^{n-1} \alpha_j, \alpha_n)$	$c_{11} = -1, c_{12} = \frac{1}{2}$	type $B_n$
$(\sum_{j=i}^{n-1} \alpha_j, \alpha_n)$	$c_{11} = -2, c_{21} = 2$	type $C_n$
$(\sum_{j=i}^{n-2} \alpha_j, \alpha_{n-1})$	$c_{11} = 1$	type $D_n$
$(\sum_{j=i}^{n-2} \alpha_j + \alpha_n, \alpha_{n-1})$	$c_{11} = -1$	type $D_n$
$(\sum_{j=i}^{n-2} \alpha_j, \alpha_n)$	$c_{11} = 1$	type $D_n$
$(\sum_{j=i}^{n-1} \alpha_j, \alpha_n)$	$c_{11} = -1$	type $D_n$

FIGURE 3. Structure constants for  $U$ 

## REFERENCES

- [ADS] E. Arias-Castro, P. Diaconis, and R. Stanley, “A super-class walk on upper triangular matrices” *J. Algebra* 278 (2004), no. 2, 739–765.
- [B] L. Babai, “Automorphism groups, isomorphism, reconstruction.” *Handbook of combinatorics, Vol. 1*, 2 1447–1540, Elsevier, Amsterdam, 1995.
- [BHKLS] L. Babai, G. Heteyi, W. M. Kantor, A. Lubotzky, A. Seress, “On the diameter of finite groups,” *31st Annual Symposium on Foundations of Computer Science, Vol. I, II (St. Louis, MO, 1990)* 857–865, IEEE Comput. Soc. Press, Los Alamitos, CA, 1990.
- [CF] P. Cartier and D. Foata, *Problèmes combinatoires de commutation et réarrangements*, Lecture Notes in Mathematics, No.85, Springer-Verlag, New York, 1969.
- [CP] D. Coppersmith and I. Pak, “Random walk on upper triangular matrices mixes rapidly.” *Probab. Theory Related Fields* 117 (2000), no. 3, 407–417.
- [DSC1] P. Diaconis and L. Saloff-Coste, “Comparison techniques for random walk on finite groups.” *Ann. Probab.* 21 (1993), no. 4, 2131–2156.
- [DSC2] P. Diaconis and L. Saloff-Coste, “Moderate growth and random walk on finite groups.” *Geom. Funct. Anal.* 4 (1994), no. 1, 1–36.
- [E] J. Ellenberg, “A sharp diameter bound for an upper triangular matrix group,” senior honors thesis, Harvard University, 1993.
- [EG] S. Even and O. Goldreich, “Minimum-length generator sequence problem is NP-hard,” *J. Algorithms*, Vol. 2, No. 3, pp.311–313, 1981.
- [He] H. Helfgott, “Growth and generation in  $SL_2(\mathbb{Z}/p\mathbb{Z})$ ”, preprint, available as [arXiv:math.GR/0509024](https://arxiv.org/abs/math/0509024).
- [Hu1] J. Humphreys, *Linear Algebraic Groups*, Springer-Verlag, New York, 1975.
- [Hu2] J. Humphreys, *Introduction to Lie Algebras and Representation Theory*, Springer-Verlag, New York, 1972.
- [SC] L. Saloff-Coste, “Random Walks on Finite Groups”, in *Probability on Discrete Structures*, Encyclopaedia Math. Sci., 110, Springer, Berlin, , 363–346, 2004.
- [S] T. Springer, *Linear Algebraic Groups*, Birkhauser, Boston, 1981.
- [St] R. Stong, “Random walks on the groups of upper triangular matrices,” *Ann. Probab.* 23 (1995), no. 4, 1939–1949.
- [VNB] A. M. Vershik, S. Nechaev, and R. Bikbov, “Statistical properties of locally free groups with applications to braid groups and growth of random heaps,” *Commun.Mah.Phys.* 212, 469–501 (2000).