Bounding the Nullities of Random Block Hankel Matrices: An Alternative Approach

Wayne Eberly *
Department of Computer Science
University of Calgary
Calgary, Alberta, Canada
eberly@cpsc.ucalgary.ca

Bradford Hovinen
Department of Mathematics
University of Toronto
Toronto, Ontario, Canada
hovinen@math.utoronto.edu

Department of Computer Science
University of Calgary
Technical Report 2005–779-10
May, 2005

Abstract

Bounds are developed for the probability that various randomly generated block Hankel matrices are rank-deficient. These bounds are potentially of use to analyze the efficiency and reliability of various randomized block Wiedemann and block Lanczos algorithms, that are either currently under development or now in use, when these are applied to solve systems of linear equations and sample from the null space of matrices over small finite fields.

The bounds that are presented here resemble ones that have previously been obtained using other arguments or that could likely be obtained by straightforward extensions of arguments that have recently been presented. The method used to obtain these bounds in this report is rather different and may be of some interest in its own right: It relies only on estimates of the number of irreducible polynomials of a given degree over a finite field and on elementary linear algebra.

1 Introduction

Krylov-based methods have recently been used (both alone and in combination with elimination-based techniques) to solve systems of linear equations whose coefficient matrices are sparse or structured matrices over finite fields and to sample from the null spaces of such matrices. An algorithm of Wiedemann [15] and various adaptations of a numerical method of Lanczos [9] have been used with considerable success.

Randomized versions of these algorithms perform computations over the Krylov space generated by the input matrix (or a conditioned matrix obtained from it) and a randomly chosen vector. More recent

[*] Research was supported in part by Natural Sciences and Engineering Research Council of Canada research grant OGP0089756.

"block algorithms" use a block consisting of a small set of independently and randomly selected vectors instead of a single vector, and work over the "block Krylov space" generated by the above-mentioned matrix and this set of vectors. Coppersmith [2, 3] has proposed block algorithms based on each of the Lanczos and Wiedemann algorithms whose scalar versions had previously been proposed, while Montgomery [12] has proposed a simpler heuristic that is based on a block Lanczos approach.

While all of these block algorithms and heuristics work well in practice, none had been fully analyzed at the time they were proposed. Kaltofen [8] has subsequently provided an analysis of the efficiency and reliability of a block Wiedemann algorithm (based on Coppersmith's) for computations over large fields. Villard [13, 14] has continued this work and provided an analysis for computations over small fields as well. Brent, Gau and Lauder [1] have subsequently obtained exact formulas for some of the values that Villard had estimated and have therefore provided improvements for this analysis.

Dumas, Gautier, Giesbrecht, Giorgi, Hovinen, Kaltofen, Saunders, Turner and Villard [4] describe a software library, LinBox, that includes block algorithms incorporating improvements suggested by these analyses. Additional information about this software and the most recent version of the library are available at the LinBox web site, http://www.linalg.org.

The analysis of these block algorithms requires the development of bounds on the probability that various randomly generated block Hankel matrices are rank-deficient. Kaltofen and Villard contributed such bounds for the large and small field cases, respectively, when the order of these matrices was slightly larger than the maximum rank possible for them, and used these bounds to complete analyses of the block Wiedemann algorithms they studied.

Something more is required if one is to analyze a block Lanczos algorithm: One must develop bounds on the probability that certain submatrices of these matrices are rank-deficient as well. This also seems to be necessary if one is to modify a block Wiedemann algorithm in order to incorporate an "early termination" mechanism of the type proposed by Lobo [11] and recently analyzed, for scalar computations over small fields, by Eberly [5].

Suitable bounds have recently been developed by Hovinen [6, 7], who has also contributed a biconditional block Lanczos algorithm and its analysis. Various ideas from the theory of commutative algebra were used here in order to adapt and apply the results that had initially been developed to study block Wiedemann and scalar Lanczos algorithms and that are summarized above.

In this report, we provide a different method to derive probability bounds of the type discussed above: Suitable bounds can be obtained using well known estimates of the number of irreducible polynomials of a given degree with coefficients in a finite field (that can be found, for example, in the text of Lidl and Neiderreiter [10]) and elementary linear algebra.

Since block Widemann and block Lanczos algorithms are still under development, it not clear precisely which of these matrices will be of interest. Rather than link this work to the analysis of any particular algorithm, bounds concerning various block Hankel matrices (and the method used to obtain them) are simply presented here in hope that they may be of some general use.

The bounds given here are somewhat more general than the ones recently published by Hovinen. However, it is quite likely that Hovinen's techniques can also be used to obtain them. Thus, this is quite probably not evidence that the technique presented here is more general than Hovinen's.

The bulk of this report includes the development and presentation of the bounds that are mentioned above. A few suggestions for future work can be found at the end.

2 Bounding Nullity: First Case

2.1 Definition and Strategy

Suppose, henceforth, that $\mathsf{F} = \mathsf{F}_q$ is a finite field with q elements. Let N be a positive integer and let $A \in \mathsf{F}^{N \times N}$ be a matrix with rank r.

Definition 2.1. Suppose that m and i are positive integers and that $v_1, v_2, \ldots, v_m \in \mathsf{F}^{N \times 1}$. Let

$$M_{A,m,i,v_1,v_2,\ldots,v_m} = \begin{bmatrix} v_1 \; Av_1 \; \cdots \; A^{\delta_1 - 1}v_1 & \cdots & v_m \; Av_m \; \cdots \; A^{\delta_m - 1}v_m \end{bmatrix} \in \mathsf{F}^{N \times i}, \tag{2.1}$$

where

$$\delta_j = \begin{cases} \lceil i/m \rceil & \text{if } j \cdot \lceil i/m \rceil + (m-j) \cdot \lfloor i/m \rfloor \le i, \\ \lfloor i/m \rfloor & \text{if } j \cdot \lceil i/m \rceil + (m-j) \cdot \lfloor i/m \rfloor > i, \end{cases}$$

for $1 \le j \le m$. Note that $\delta_j \in \{\lfloor i/m \rfloor, \lceil i/m \rceil\}$, that

$$\delta_1 \ge \delta_2 \ge \delta_3 \ge \cdots \ge \delta_m,$$

and that

$$\delta_1 + \delta_2 + \cdots + \delta_m = i.$$

Definition 2.2. Suppose that m and i are positive integers. Let $D_{A,m,i}$ be the number of choices of vectors $v_1, v_2, \ldots, v_m \in \mathsf{F}^{N \times 1}$ and scalars

$$d_{1,0}, d_{1,1}, \ldots, d_{1,\delta_1 - 1}, \ldots, d_{m,0}, d_{m,1}, \ldots, d_{m,\delta_m - 1} \in \mathsf{F},$$

for $\delta_1, \delta_2, \ldots, \delta_m$ as given in Definition 2.1, above, such that

$$M_{A,m,i,v_1,v_2,\ldots,v_m} \begin{bmatrix} d_{1,0} \\ d_{1,1} \\ \vdots \\ d_{m,\delta_m - 1} \end{bmatrix} = 0. \tag{2.2}$$

We will count this value in two ways, and compare the resulting expressions, in order to bound the probability that the matrix $M_{A,m,i,v_1,v_2,\ldots,v_m}$ has low rank when the vectors $v_1, v_2, \ldots, v_m$ are chosen uniformly and independently from $\mathsf{F}^{N \times 1}$.

In particular, bounds will be obtained for the following quantities.

Definition 2.3. Suppose that m and i are positive integers and that j is a nonnegative integer. Let $\rho_{A,m,i}(j)$ be the probability that the matrix $M_{A,m,i,v_1,v_2,\ldots,v_m}$ is rank deficient by exactly j, that is,

$$\rho_{A,m,i}(j) = \mathsf{Prob}\left[\operatorname{rank}\left(M_{A,m,i,v_1,v_2,\ldots,v_m}\right) = i - j\right] \quad \text{if } i \le N \tag{2.3}$$

and

$$\rho_{A,m,i}(j) = \mathsf{Prob}\left[\operatorname{rank}\left(M_{A,m,i,v_1,v_2,\ldots,v_m}\right) = N - j\right] \quad \text{if } i > N, \tag{2.4}$$

and let $\sigma_{A,m,i}(j)$ be the probability that this matrix is rank deficient by at least j, that is,

$$\sigma_{A,m,i}(j) = \mathsf{Prob}\left[\operatorname{rank}\left(M_{A,m,i,v_1,v_2,\ldots,v_m}\right) \le i - j\right] \quad \text{if } i \le N, \tag{2.5}$$

and

$$\sigma_{A,m,i}(j) = \mathsf{Prob}\left[\operatorname{rank}\left(M_{A,m,i,v_1,v_2,\ldots,v_m}\right) \le N - j\right] \quad \text{if } i > N, \tag{2.6}$$

when the vectors $v_1, v_2, \ldots, v_m$ are chosen uniformly and independently from $\mathsf{F}^{N \times 1}$.

2.2 Counting These Values One Way

The following lemmas clarify the relationship between the above quantities when $i \leq N$ and when $i > N$.

Lemma 2.4. *Let $C \in \mathsf{F}^{k \times \ell}$, where $\mathsf{F} = \mathsf{F}_q$ is a finite field with q elements, and where k and ℓ are positive integers such that $k \geq \ell$.*

If C has rank $\ell - j$, for an integer $j \geq 0$, then there are exactly q^j vectors $x \in \mathsf{F}^{\ell \times 1}$ such that $Cx = 0$.

Proof. Since C has rank $\ell - j$, there exists a set of $\ell - j$ columns of C that are linearly independent. Permuting the columns of C (and the entries of a vector x) as needed, we may assume without loss of generality that the leftmost $\ell - j$ columns of C are linearly independent and that the remaining j columns of C are linear combinations of the leftmost ones. That is,

$$C = \begin{bmatrix} C_1 & C_2 \end{bmatrix} \tag{2.7}$$

where $C_1 \in \mathsf{F}^{k \times (\ell - j)}$ is a matrix with full rank $\ell - j$, and $C_2 \in \mathsf{F}^{k \times j}$ is a matrix whose columns are linear combinations of the columns of C_1.

Note that, since C_1 has full rank, $C_1 z = 0$ if and only if $z = 0$ for any vector $z \in \mathsf{F}^{(\ell - j) \times 1}$.

Since the columns of C_2 are all linear combinations of the columns of C_1, there exists a matrix $Z \in \mathsf{F}^{(\ell - j) \times j}$ such that

$$C_2 = C_1 Z. \tag{2.8}$$

Notice that any vector $z \in \mathsf{F}^{\ell \times 1}$ can be written as

$$z = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \tag{2.9}$$

where $z_1 \in \mathsf{F}^{(\ell - j) \times 1}$ and $z_2 \in \mathsf{F}^{j \times 1}$.

Suppose, now that $z \in \mathsf{F}^{\ell \times 1}$. Then

$$\begin{aligned} Cz = 0 &\iff C_1 z_1 + C_2 z_2 = 0 && \text{(by (2.7) and (2.9))} \\ &\iff C_1(z_1 + Z z_2) = 0 && \text{(by (2.8))} \\ &\iff z_1 + Z z_2 = 0 && \text{(since } C_1 \text{ has full rank)} \\ &\iff z_1 = -Z z_2. \end{aligned}$$

Consequently, for any vector $z_2 \in \mathsf{F}^{j \times 1}$, there is exactly one choice of a vector $z_1 \in \mathsf{F}^{(\ell - j) \times 1}$ such that

$$Cz = 0 \qquad \text{if } z = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}.$$

Since there are exactly $|\mathsf{F}|^j = q^j$ choices of the vector z_2 the desired result now follows. □

Lemma 2.5. *Let $C \in \mathsf{F}^{k \times \ell}$ where $\mathsf{F} = \mathsf{F}_q$ is a finite field with q elements, and where k and ℓ are positive integers such that $k \leq \ell$.*

If C has rank $k - j$, for an integer $j \geq 0$, then there are exactly $q^{\ell - k + j}$ vectors $x \in \mathsf{F}^{\ell \times 1}$ such that $Cx = 0$.

Proof. This can be established by a modification of the proof of Lemma 2.4: Notice that after a permutation of columns we may write

$$C = \begin{bmatrix} C_1 & C_2 \end{bmatrix}$$

where $C_1 \in \mathsf{F}^{k \times (k-j)}$ is a matrix with full rank $k - j$ and $C_2 \in \mathsf{F}^{N \times (\ell - k + j)}$ is a matrix whose columns are linear combinations of the columns of C_1. We may conclude once again that $C_1 z = 0$ if and only if $z = 0$ for each vector $z \in \mathsf{F}^{(k-j)\times 1}$, notice that

$$C_2 = C_1 Z$$

for some matrix $Z \in \mathsf{F}^{(k-j)\times(\ell-k+j)}$, write an arbitrary vector $z \in \mathsf{F}^{\ell \times 1}$ as

$$z = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \qquad \text{for } z_1 \in \mathsf{F}^{(k-j)\times 1} \text{ and } z_2 \in \mathsf{F}^{(\ell-k+j)\times 1},$$

and then argue as above that $Cz = 0$ if and only if $z_1 = -Zz_2$. Once again, the claim then follows by a consideration of the number of choices of z_2. □

Suppose that $i \geq m$ (noting that the matrix $M_{A,m,i,v_1,v_2,\dots,v_m}$ is chosen uniformly from $\mathsf{F}^{N \times i}$ if the vectors $v_1, v_2, \dots, v_m$ are chosen uniformly and independently from $\mathsf{F}^{N\times 1}$, otherwise). Since there are q^{Nm} choices of the vectors $v_1, v_2, \dots, v_m \in \mathsf{F}^{m \times 1}$, there are also q^{Nm} ways to choose the matrix $M_{A,m,i,v_1,v_2,\dots,v_m}$. It follows that there are $q^{Nm}\rho_{A,m,i}(j)$ choices of the matrix $M_{A,m,i,v_1,v_2,\dots,v_m}$ with rank $i - j$ for any integer j such that $0 \leq j \leq i$ if $i \leq N$, and that there are $q^{Nm}\rho_{A,m,i}(j)$ choices of the matrix $M_{A,m,i,v_1,v_2,\dots v_m}$ with rank $N - j$ for any integer j such that $0 \leq j \leq N$ if $i > N$.

It follows, by Lemmas 2.4 and 2.5 that

$$D_{A,m,i} = \sum_{j=0}^{i} q^{Nm+j} \rho_{A,m,i}(j) \quad \text{if } i \leq N \tag{2.10}$$

and that

$$D_{A,m,i} = \sum_{j=0}^{N} q^{N(m-1)+i+j} \rho_{A,m,i}(j) \quad \text{if } i > N. \tag{2.11}$$

Since $\rho_{A,m,i}(j) = \sigma_{A,m,i}(j) - \sigma_{A,m,i}(j+1)$ if $j < \min(i, N)$, $\rho_{A,m,i}(j) = \sigma_{A,m,i}(j)$ if $j = \min(i, N)$, and since $\sigma_{A,m,i}(0) = 1$, the above equation can be used to establish that

$$D_{A,m,i} = q^{Nm} + (q-1) \sum_{j=1}^{i} q^{Nm+j-1} \sigma_{A,m,i}(j) \quad \text{if } i \leq N, \tag{2.12}$$

and that

$$D_{A,m,i} = q^{N(m-1)+i} + (q-1) \sum_{j=1}^{N} q^{N(m-1)+i+j-1} \sigma_{A,m,i}(j) \quad \text{if } i > N. \tag{2.13}$$

2.3 Counting These Values Another Way

A second enumeration of $D_{A,m,i}$ will also be useful. It will be helpful to consider a sequence of cases related to the structure of the matrix A.

2.3.1 First Case

Suppose first that A is similar to a companion matrix and, furthermore, that its characteristic polynomial is a power φ^n of a monic irreducible polynomial φ. In other words, suppose that

$$A = X^{-1}ZX \tag{2.14a}$$

for a nonsingular matrix $X \in \mathsf{F}^{N\times N}$, and for a matrix

$$Z = \begin{bmatrix} 0 & & & & -\alpha_0 \\ 1 & 0 & & & -\alpha_1 \\ & 1 & & & -\alpha_2 \\ & & \ddots & & \vdots \\ & & & 1 & -\alpha_{N-1} \end{bmatrix} \in \mathsf{F}^{N\times N} \tag{2.14b}$$

with $(k, \ell)^{\text{th}}$ entry 1 if $k = \ell + 1$ and with $(k, \ell)^{\text{th}}$ entry 0 otherwise when $1 \leq k \leq N$ and $1 \leq \ell \leq N - 1$, and with $(k, N)^{\text{th}}$ entry $-\alpha_{k-1}$, for $1 \leq k \leq N$, where $\varphi \in \mathsf{F}[x]$ is an irreducible polynomial with degree d, n is a positive integer such that

$$\varphi^n = x^N + \alpha_{N-1}x^{N-1} + \cdots + \alpha_1 x + \alpha_0 \in \mathsf{F}[x], \tag{2.14c}$$

and where

$$\deg(\varphi^n) = N = dn. \tag{2.14d}$$

Lemma 2.6. *If A is as given in Equations (2.14a) – (2.14d) then there exists a vector $\zeta \in \mathsf{F}^{N\times 1}$ such that*

$$\mathsf{F}^{N\times 1} = \{\, f(A)\zeta \mid f \in \mathsf{F}[x] \text{ and } \deg(f) < N \,\}$$

and such that

$$\varphi^n(A)\zeta = 0.$$

Furthermore, there is exactly one polynomial $f \in \mathsf{F}[x]$ with degree less than $N = dn$ such that

$$y = f(A)\zeta$$

for any given vector $y \in \mathsf{F}^{N\times 1}$.

Proof. Consider the k^{th} elementary vector $e_k \in \mathsf{F}^{N\times 1}$, for $1 \leq k \leq N$, whose ℓ^{th} entry is 1 if $k = \ell$ and 0 otherwise, for $1 \leq \ell \leq N$. Since

$$A = X^{-1}ZX,$$

for a nonsingular matrix $X \in \mathsf{F}^{N\times N}$, it suffices to set

$$\zeta = X^{-1}e_1 \tag{2.15}$$

in order to satisfy the conditions given in the lemma: It is easily checked, using the above relationships and Equation (2.14b), that

$$\begin{aligned} A^{i-1}\zeta &= (X^{-1}ZX)^{i-1}X^{-1}e_1 \\ &= X^{-1}Z^{i-1}XX^{-1}e_1 \\ &= X^{-1}Z^{i-1}e_i \end{aligned}$$

$$= X^{-1}e_i$$

for $1 \le i \le N$.

Since the matrix X^{-1} is nonsingular, the vectors

$$\zeta = X^{-1}e_1,\ A\zeta = X^{-1}e_2,\ \ldots,\ A^{N-1}\zeta = X^{-1}e_N$$

form a basis for $\mathsf{F}^{N\times 1}$, so that

$$\mathsf{F}^{N\times 1} = \{\, f(A)\zeta \mid f \in \mathsf{F}[x] \text{ and } \deg(f) < N \,\}$$

as claimed.

It is also clear from the definition of ζ at line (2.15), and from Equations (2.14b) and (2.14c), that

$$\varphi^n(A)\zeta = X^{-1}\varphi^n(Z)e_1 = 0.$$

Finally, if $y \in \mathsf{F}^{N\times 1}$ then it is possible to write y as

$$y = X^{-1}\begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_{N-1} \end{bmatrix} \tag{2.16}$$

for values $\alpha_0, \alpha_1, \ldots, \alpha_{N-1} \in \mathsf{F}$. In this case,

$$y = f(A)\zeta \tag{2.17}$$

for the polynomial

$$f = \alpha_0 + \alpha_1 x + \cdots + \alpha_{N-1}x^{N-1}. \tag{2.18}$$

It is also clear that, since X^{-1} is nonsingular, there is only once choice of values $\alpha_0, \alpha_1, \ldots, \alpha_{N-1} \in \mathsf{F}$ such that Equation (2.16) is satisfied, so that there is only one polynomial $f \in \mathsf{F}[x]$ with degree less than $N = dn$ such that Equations (2.17) and (2.18) are satisfied, as well. □

Corollary 2.7. *If A and ζ are as described in Lemma 2.6 and $f \in \mathsf{F}[x]$, then*

$$f(A)\zeta = 0 \qquad \textit{if and only if} \qquad f \equiv 0 \bmod \varphi^n.$$

Proof. If $f \in \mathsf{F}[x]$ then $f = f_L + \varphi^n f_H$ for polynomials $f_L, f_H \in \mathsf{F}[x]$ such that f_L has degree less than $N = dn$. It follows by the results of the lemma that

$$f(A)\zeta = f_L(A)\zeta + \varphi^n(A)f_H(A)\zeta = f_L(A)\zeta + f_H(A)(\varphi^n(A)\zeta) = f_L(A)\zeta,$$

and, furthermore, that $f_L(A)\zeta = 0$ if and only if $f_L = 0$. □

Once again, let $M_{A,m,i,v_1,v_2,\ldots,v_m}$ and let $\delta_1, \delta_2, \ldots, \delta_m$ be as given in Definition 2.1.

Lemma 2.8. *If A is as given in Equations (2.14a) – (2.14d) and m and i are positive integers such that $m \le i$, then $D_{A,m,i}$ is equal to the number of choices of polynomials*

$$f_1, f_2, \ldots, f_m \in \mathsf{F}[x]$$

where f_j has degree less than $N = dn$ for $1 \le j \le m$, and of polynomials

$$g_1, g_2, \ldots, g_m \in \mathsf{F}[x]$$

where g_j has degree less than δ_j for $1 \le j \le m$, such that

$$f_1 g_1 + f_2 g_2 + \cdots + f_m g_m \equiv 0 \bmod \varphi^n.$$

Proof. Recall that, by Definition 2.2, $D_{A,m,i}$ is the number of choices of vectors $v_1, v_2, \dots, v_m \in \mathsf{F}^{N\times 1}$ and scalars

$$d_{1,0}, d_{1,1}, \dots, d_{1,\delta_1-1}, \dots, d_{m,0}, d_{m,1}, \dots, d_{m,\delta_m-1}$$

such that Equation (2.2) is satisfied.

It follows by Lemma 2.6 that the number of choices of vectors $v_1, v_2, \dots, v_m \in \mathsf{F}^{N\times 1}$ is the same as the number of choices of polynomials $f_1, f_2, \dots, f_m \in \mathsf{F}[x]$ that each has degree less than $N = dn$. Indeed, there is exactly one such polynomial $f_j \in \mathsf{F}[x]$ with degree less than N such that

$$v_j = f_j(A)\zeta \tag{2.19}$$

for $1 \le j \le m$ and for the vector ζ described in the lemma.

There is also a straightforward one-to-one correspondence between the sequences of scalars

$$d_{1,0}, d_{1,1}, \dots, d_{1,\delta_1-1}, \dots, d_{m,0}, d_{m,1}, \dots, d_{m,\delta_m-1}$$

and sequences of polynomials $g_1, g_2, \dots, g_m \in \mathsf{F}[x]$ such that the degree of g_j is less than δ_j for $1 \le j \le m$: It suffices to set

$$g_j = d_{j,0} + d_{j,1}x + \dots, d_{j,\delta_j-1}x^{\delta_j-1} \in \mathsf{F}[x] \tag{2.20}$$

for $1 \le j \le m$ in order to obtain this correspondence.

Consider the definition of $M_{A,m,i,v_2,v_2,\dots,v_m}$ (Definition 2.1 on page 3). If the above vectors, scalars, and polynomials are related as shown in Equations (2.19) and (2.20), then it follows by the above definition that

$$M_{A,m,i,v_1,v_2,\dots,v_m} \begin{bmatrix} d_{1,0} \\ d_{1,1} \\ \vdots \\ d_{m,\delta_m-1} \end{bmatrix} = (f_1g_1 + f_2g_2 + \dots, f_mg_m)(A)\zeta.$$

It follows by Corollary 2.7 that

$$M_{A,m,i,v_1,v_2,\dots,v_m} \begin{bmatrix} d_{1,0} \\ d_{1,1} \\ \vdots \\ d_{m,\delta_m-1} \end{bmatrix} = 0 \qquad \Longleftrightarrow \qquad f_1g_1 + f_2g_2 + \dots + f_mg_m \equiv 0 \bmod \varphi^n.$$

The claim now follows by the correspondences described at the beginning of this argument. □

We are now ready to count $D_{A,m,i}$ in another way.

Lemma 2.9. *Consider a sequence of polynomials*

$$g_1, g_2, \dots, g_m \in \mathsf{F}[x]$$

such that the degree of g_j is less than δ_j for $1 \le j \le m$.

(a) *If at least one of the polynomials $g_1, g_2, \dots, g_m$ is not divisible by φ then there are exactly $q^{N(m-1)}$ choices of polynomials*

$$f_1, f_2, \dots, f_m \in \mathsf{F}[x],$$

each with degree less than $N = dn$, such that

$$f_1g_1 + f_2g_2 + \dots + f_mg_m \equiv 0 \bmod \varphi^n.$$

(b) *Let h be a positive integer such that $dh < \delta_1 = \lceil i/m \rceil$. If the polynomials $g_1, g_2, \ldots, g_m$ are all divisible by φ^h, but at least one of these polynomials is not divisible by φ^{h+1}, then there are exactly $q^{N(m-1)+dh}$ choices of polynomials*

$$f_1, f_2, \ldots, f_m \in \mathsf{F}[x],$$

each with degree less than $N = dn$, such that

$$f_1 g_1 + f_2 g_2 + \cdots + f_m g_m \equiv 0 \bmod \varphi^n.$$

(c) *Finally, if the polynomials $g_1, g_2, \ldots, g_m$ are all divisible by φ^h, where $dh \geq \delta_1 = \lceil i/m \rceil$, then there are exactly q^{Nm} choices of polynomials*

$$f_1, f_2, \ldots, f_m \in \mathsf{F}[x],$$

each with degree less than $N = dn$, such that

$$f_1 g_1 + f_2 g_2 + \cdots + f_m g_m \equiv 0 \bmod \varphi^n.$$

Proof. Once again, let $g_1, g_2, \ldots, g_m$ be polynomials in $\mathsf{F}[x]$ such that the degree of g_j is less than δ_j for $1 \leq j \leq m$, where $\mathsf{F} = \mathsf{F}_q$ is the finite field with q elements

(a) Suppose that at least one of $g_1, g_2, \ldots, g_m$ is not divisible by φ. In particular, suppose that g_ℓ is not divisible by φ where $1 \leq \ell \leq m$.

In this case, since φ is irreducible, the greatest common divisor of g_ℓ and φ^n is 1, so there exist polynomials $s, t \in \mathsf{F}[x]$ such that

$$s g_\ell + t \varphi^n = 1.$$

Clearly, s is also relatively prime to φ^n if the above equation is satisfied. Consequently, if $f_1, f_2, \ldots f_m$ are polynomials in $\mathsf{F}[x]$ with degrees less than $N = dn$, then

$$\begin{aligned} & f_1 g_1 + f_2 g_2 + \cdots + f_m g_m \equiv 0 \bmod \varphi^n \\ & \quad \Longleftrightarrow \ f_\ell g_\ell \equiv -(f_1 g_1 + f_2 g_2 + \cdots + f_{\ell-1} g_{\ell-1} + f_{\ell+1} g_{\ell+1} + \cdots + f_m g_m) \bmod \varphi^n \\ & \quad \Longleftrightarrow \ f_\ell \equiv -s(f_1 g_1 + f_2 g_2 + \cdots + f_{\ell-1} g_{\ell-1} + f_{\ell+1} g_{\ell+1} + \cdots + f_m g_m) \bmod \varphi^n, \end{aligned}$$

since $s g_\ell \equiv 1 \bmod \varphi^n$.

It follows that there is exactly one choice of a polynomial $f_\ell \in \mathsf{F}[x]$ with degree less than $N = dn$, for any choice of the polynomials $f_1, f_2, \ldots, f_{\ell-1}, f_{\ell+1}, \ldots, f_m \in \mathsf{F}[x]$ with degree less than N, such that the desired equation is satisfied. Since there are exactly $q^{N(m-1)}$ choices of the polynomials $f_1, f_2, \ldots, f_{\ell-1}, f_{\ell+1}, \ldots, f_m$, this establishes the claim.

(b) Suppose next that h is a positive integer such that $dh < \delta_1 = \lceil i/m \rceil$, that each of the polynomials $g_1, g_2, \ldots, g_m$ are divisible by φ^h, and that at least one of them is not divisible by φ^{h+1}. Suppose, in particular, that g_ℓ is not divisible by φ^{h+1} for an integer ℓ such that $1 \leq \ell \leq m$.

In this case we may write g_j as $g_j = \varphi^h \widehat{g}_j$, for $1 \leq j \leq m$, and note that $\widehat{g}_\ell$ and φ^n are relatively prime. Thus the greatest common divisor of g_ℓ and φ^n is φ^h and there exist polynomials s and t such that

$$s g_\ell + t \varphi^n = \varphi^h.$$

Note that, in this case,

$$s\widehat{g}_\ell + t\varphi^{n-h} = 1,$$

so (once again) s and φ^n are relatively prime.

It follows that if $f_1, f_2, \dots, f_m$ are polynomials in $\mathsf{F}[x]$ with degrees less than $N = dn$, then

$$\begin{aligned}
&f_1g_1 + f_2g_2 + \cdots + f_mg_m \equiv 0 \bmod \varphi^n \\
&\quad\iff f_\ell g_\ell \equiv -(f_1g_1 + f_2g_2 + \cdots + f_{\ell-1}g_{\ell-1} + f_{\ell+1}g_{\ell+1} + \cdots + f_mg_m) \bmod \varphi^n \\
&\quad\iff f_\ell \varphi^h \equiv -s(f_1g_1 + f_2g_2 + \cdots + f_{\ell-1}g_{\ell-1} + f_{\ell+1}g_{\ell+1} + \cdots + f_mg_m) \quad \bmod \varphi^n \\
&\quad\iff f_\ell \varphi^h \equiv -s(f_1\widehat{g}_1 + f_2\widehat{g}_2 + \ldots + f_{\ell-1}\widehat{g}_{\ell-1} + f_{\ell+1}\widehat{g}_{\ell+1} + \cdots + f_m\widehat{g}_m)\varphi^h \bmod \varphi^n \\
&\quad\iff f_\ell \equiv -s(f_1\widehat{g}_1 + f_2\widehat{g}_2 + \cdots + f_{\ell-1}\widehat{g}_{\ell-1} + f_{\ell+1}\widehat{g}_{\ell+1} + f_m\widehat{g}_m) \bmod \varphi^{n-h}.
\end{aligned}$$

The latter condition on f_ℓ is satisfied whenever

$$f_\ell = f_H\varphi^{n-h} + f_L,$$

where $f_L \in \mathsf{F}[x]$ is a polynomial with degree less than $\deg(\varphi^{n-h}) = d(n-h)$ such that

$$f_L \equiv -s(f_1\widehat{g}_1 + f_2\widehat{g}_2 + \cdots + f_{\ell-1}\widehat{g}_{\ell-1} + f_{\ell+1}\widehat{g}_{\ell+1} + \cdots + f_m\widehat{g}_m) \bmod \varphi^{n-h},$$

and where f_H is an arbitrarily chosen polynomial in $\mathsf{F}[x]$ with degree less than dh.

Since there is exactly one choice of f_L possible, for any choice of $f_1, f_2, \dots, f_{\ell-1}, f_{\ell+1}, \dots, f_m$ and f_H, and since there are exactly $q^{N(m-1)+dh}$ choices of $f_1, f_2, \dots, f_{\ell-1}, f_{\ell+1}, \dots, f_m$ and f_H, this establishes part (b).

(c) Finally it should be noted that if $g_1, g_2, \dots, g_m \in \mathsf{F}[x]$ are polynomials such that the degree of g_j is less than $\delta_j \leq \lceil i/m \rceil$ for $1 \leq j \leq m$, and each of these polynomials is divisible by φ^h, where $dh \geq \lceil i/m \rceil$, then

$$g_1 = g_2 = \cdots = g_m = 0.$$

In this case,

$$f_1g_1 + f_2g_2 + \cdots + f_mg_m \equiv 0 \bmod \varphi^n$$

for every choice of polynomials $f_1, f_2, \dots, f_m \in \mathsf{F}[x]$ that each has degree less than $N = dn$, and there are exactly q^{Nm} choices of these polynomials. □

Lemma 2.10. *If A is as given in Equations (2.14a) – (2.14d), and m and i are positive integers such that $i \geq m$, then*

$$D_{A,m,i} < q^{Nm} + q^{N(m-1)+i} \sum_{h \geq 0} q^{-(m-1)dh}.$$

Proof. Suppose first that h is a nonnegative integer such that $dh \leq \delta_m = \lfloor i/m \rfloor$. Then, since

$$\delta_1 + \delta_2 + \cdots + \delta_m = i,$$

there are exactly q^{i-mdh} sequences of polynomials

$$g_1, g_2, \dots, g_m \in \mathsf{F}[x]$$

such that the degree of g_j is less than δ_j for $1 \leq j \leq m$ and such that $g_1, g_2, \dots, g_m$ are all divisible by φ^h. On the other hand, if $dh > \delta_m$ then the degree of φ^h is at least $\delta_1 = \lceil i/m \rceil$, and there is exactly one sequence of polynomials $g_1, g_2, \dots, g_m \in \mathsf{F}[x]$ satisfying the given degree and divisibility constraints, namely, the sequence

$$g_1 = g_2 = \cdots = g_m = 0.$$

It now follows by the definition of $D_{A,m,i}$ (given on page 3) and Lemma 2.9 that, if A is as given in Equations (2.14a) and (2.14b), then

$$\begin{aligned} D_{A,m,i} &\leq \sum_{h=0}^{\lfloor (\lfloor i/m \rfloor /d) \rfloor} \left(q^{i-mdh} \cdot q^{N(m-1)+dh} \right) + q^{Nm} \\ &< q^{Nm} + \sum_{h \geq 0} q^{N(m-1)+i-(m-1)dh} \\ &= q^{Nm} + q^{N(m-1)+i} \sum_{h \geq 0} q^{-(m-1)dh}, \end{aligned}$$

as claimed. □

Using the closed form for a geometric series, we may now conclude that

$$D_{A,m,i} < q^{Nm} + \frac{q^{N(m-1)+i}}{1 - q^{-(m-1)d}} = q^{Nm} + \frac{q^{Nm}}{q^{N-i} - q^{N-i-(m-1)d}} \tag{2.21}$$

when A is as given in Equations (2.14a) – (2.14d).

2.3.2 Second Case

Suppose next that A is similar to a block diagonal matrix where each block is a companion matrix of the type considered in the first case:

$$A = X^{-1} \Delta X \tag{2.22a}$$

for a nonsingular matrix $X \in \mathsf{F}^{N \times N}$, and for a block diagonal matrix

$$\Delta = \begin{bmatrix} Z_1 & & & 0 \\ & Z_2 & & \\ & & \ddots & \\ 0 & & & Z_\ell \end{bmatrix} \in \mathsf{F}^{N \times N}, \tag{2.22b}$$

where each block Z_j is the companion matrix of a positive power of φ. That is, there exist positive integers $n_1, n_2, \dots, n_\ell$ such that

$$Z_j = \begin{bmatrix} 0 & & & & -\alpha_{j,0} \\ 1 & 0 & & & -\alpha_{j,1} \\ & 1 & & & -\alpha_{j,2} \\ & & \ddots & & \vdots \\ & & & 1 & -\alpha_{j,dn_j-1} \end{bmatrix} \in \mathsf{F}^{dn_j \times dn_j} \tag{2.22c}$$

is the companion matrix of the polynomial

$$\varphi^{n_j} = x^{dn_j} + \alpha_{j,dn_j-1} x^{dn_j-1} + \cdots + \alpha_{j,1} x + \alpha_{j,0} \in \mathsf{F}[x], \tag{2.22d}$$

for $1 \leq j \leq \ell$, and such that

$$n_1 \geq n_2 \geq \cdots \geq n_\ell \geq 1 \qquad \text{and} \qquad d(n_1 + n_2 + \cdots + n_\ell) = N. \tag{2.22e}$$

In this case, A has minimal polynomial φ^{n_1} and characteristic polynomial $\prod_{j=1}^{\ell} \varphi^{n_j} = \varphi^{N/d}$.

In this section we will consider the case that $\ell \leq m - 1$.

The next result generalizes Lemma 2.6 for this case.

Lemma 2.11. *If A is as given in Equations (2.22a) – (2.22e), then there exist vectors $\zeta_1, \zeta_2, \dots, \zeta_\ell \in \mathsf{F}^{N\times 1}$ such that*

$$\mathsf{F}^{N\times 1} = V_1 \oplus V_2 \oplus \cdots \oplus V_\ell,$$

where

$$V_j = \{\, f(A)\zeta_j \mid f \in \mathsf{F}[x] \text{ and } \deg(f) < dn_j \,\},$$

and such that

$$\varphi^{n_j}(A)\zeta_j = 0$$

for $1 \leq j \leq \ell$.

Furthermore, there is exactly one sequence of polynomials $f_1, f_2, \dots, f_\ell$ such that $f_j \in \mathsf{F}[x]$ has degree less than dn_j for $1 \leq j \leq \ell$ and such that

$$y = f_1(A)\zeta_1 + f_2(A)\zeta_2 + \cdots + f_\ell(A)\zeta_\ell$$

for any given vector $y \in \mathsf{F}^{N\times 1}$.

Proof. As in the proof of Lemma 2.6, it is sufficient to identify the values $\zeta_1, \zeta_2, \dots, \zeta_\ell$ that are mentioned in the statement of the lemma and to check that the given properties are all satisfied.

If A is as described in the statement of the lemma then it suffices to set

$$\zeta_j = X^{-1} e_{d(n_1+n_2+\cdots+n_{j-1})+1}$$

for $1 \leq j \leq \ell$ (so that $\zeta_1 = X^{-1}e_1$). It can then be shown that

$$A^{k-1}\zeta_j = X^{-1} e_{d(n_1+n_2+\cdots+n_{j-1})+k}$$

whenever $1 \leq j \leq \ell$ and k is a positive integer such that $1 \leq k \leq dn_j$. Consequently the vector space V_j described in the statement of the lemma has a basis

$$X^{-1}e_{d(n_1+n_2+\cdots+n_{j-1})+1}, X^{-1}e_{d(n_1+n_2+\cdots+n_{j-1})+2}, \dots, X^{-1}e_{d(n_1+n_2+\cdots+n_{j-1}+n_j)},$$

and, since the matrix X^{-1} is nonsingular, it follows immediately that

$$\mathsf{F}^{N\times 1} = V_1 \oplus V_2 \oplus \cdots \oplus V_\ell$$

as claimed.

It is also clear from the definition of ζ_j (and Equations (2.22a) – (2.22c)) that

$$\varphi^{n_j}(A)\zeta_j = X^{-1}\varphi^{n_j}(\Delta)e_{d(n_1+n_2+\cdots+n_{j-1})+1} = 0$$

for $1 \leq j \leq \ell$.

Consider a vector $y \in \mathsf{F}^{N \times 1}$. Since the matrix X^{-1} is nonsingular, it is possible to write y as

$$y = X^{-1} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_{N-1} \end{bmatrix} \tag{2.23}$$

for values $\alpha_0, \alpha_1, \dots, \alpha_{N-1} \in \mathsf{F}$. In this case,

$$y = f_1(A)\zeta_1 + f_2(A)\zeta_2 + \cdots + f_\ell(A)\zeta_\ell \tag{2.24}$$

when

$$f_j = \alpha_{d(n_1+n_2+\cdots+n_{j-1})} + \alpha_{d(n_1+n_2+\cdots+n_{j-1})+1}x + \cdots + \alpha_{d(n_1+n_2+\cdots+n_{j-1}+n_j)-1}x^{dn_j-1} \in \mathsf{F}[x] \tag{2.25}$$

for $1 \le j \le \ell$.

Since X^{-1} is nonsingular, it is clear that there is only one choice of values $\alpha_0, \alpha_1, \dots, \alpha_{N-1} \in \mathsf{F}$ such that Equation (2.23) is satisfied, so that there is only one sequence of polynomials $f_1, f_2, \dots, f_\ell \in \mathsf{F}[x]$ such that the degree of f_j is less than dn_j for $1 \le j \le \ell$ and such that Equations (2.24) and (2.25) are satisfied, as well. □

The argument used to establish Corollary 2.7 from Lemma 2.6 can be used to establish the following result from Lemma 2.11, as well.

Corollary 2.12. *If A and $\zeta_1, \zeta_2, \dots, \zeta_\ell$ are as described in Lemma 2.11, and $f_1, f_2, \dots, f_\ell \in \mathsf{F}[x]$, then*

$$f_1(A)\zeta_1 + f_2(A)\zeta_2 + \cdots + f_\ell(A)\zeta_\ell = 0$$

if and only if f_j is divisible by φ^{n_j} in $\mathsf{F}[x]$ for $1 \le j \le \ell$.

Lemma 2.13. *If A is as given in Equations (2.22a) – (2.22e) and m and i are positive integers such that $i \ge m$, then $D_{A,m,i}$ is equal to the number of choices of polynomials*

$$f_{j,k} \in \mathsf{F}[x] \qquad \text{for } 1 \le j \le \ell \text{ and } 1 \le k \le m$$

where $\deg(f_{j,k}) < dn_j$ for all j and k as above, and of polynomials

$$g_1, g_2, \dots, g_m \in \mathsf{F}[x]$$

where the degree of g_k is less than δ_k for $1 \le k \le m$, such that

$$f_{j,1}g_1 + f_{j,2}g_2 + \cdots + f_{j,m}g_m \equiv 0 \bmod \varphi^{n_j}$$

for every integer j such that $1 \le j \le \ell$.

Proof. Recall that, by Definition 2.2, $D_{A,m,i}$ is the number of choices of vectors $v_1, v_2, \dots, v_m \in \mathsf{F}^{N \times 1}$ and scalars

$$d_{1,0}, d_{1,1}, \dots, d_{1,\delta_1-1}, \dots, d_{m,0}, d_{m,1}, \dots, d_{m,\delta_m-1}$$

such that Equation (2.2) is satisfied.

It follows by Lemma 2.11 that the number of choices of vectors $v_1, v_2, \dots, v_m \in \mathsf{F}^{N\times 1}$ is the same as the number of choices of polynomials $f_{j,k} \in \mathsf{F}[x]$ such that $f_{j,k}$ has degree less than dn_j, for $1 \le j \le \ell$ and $1 \le k \le m$. Indeed, there is exactly one such choice of polynomials $f_{1,k}, f_{2,k}, \dots, f_{\ell,k}$ such that

$$v_k = f_{1,k}(A)\zeta_1 + f_{2,k}(A)\zeta_2 + \cdots + f_{\ell,k}(A)\zeta_\ell \tag{2.26}$$

for $1 \le k \le m$, and for the vectors $\zeta_1, \zeta_2, \dots, \zeta_\ell$ described in the lemma.

As noted in the proof of Lemma 2.8, there is also a straightforward one-to-one correspondence between the sequences of scalars

$$d_{1,0}, d_{1,1}, \dots, d_{1,\delta_1-1}, \dots, d_{m,0}, d_{m,1}, \dots, d_{m,\delta_m-1}$$

and sequences of polynomials $g_1, g_2, \dots, g_m \in \mathsf{F}[x]$ such that the degree of g_k is less than δ_k for $1 \le k \le m$: It is sufficient to define $g_1, g_2, \dots, g_m$ as shown in Equation (2.20) on page 8 to achieve this.

Once again, consider the definition of $M_{A,m,i,v_1,v_2,\dots,v_m}$ (Definition 2.1 on page 3). If the above vectors, scalars, and polynomials are as related in Equations (2.26) and (2.20), then it follows by the above definition that

$$M_{A,m,i,v_1,v_2,\dots,v_m} \begin{bmatrix} d_{1,0} \\ d_{1,1} \\ \vdots \\ d_{m,\delta_m-1} \end{bmatrix} = \sum_{j=1}^{\ell} (f_{j,1}g_1 + f_{j,2}g_2 + \cdots + f_{j,m}g_m)(A)\zeta_j.$$

It follows by Corollary 2.12 that

$$M_{A,m,i,v_1,v_2,\dots,v_m} \begin{bmatrix} d_{1,0} \\ d_{1,1} \\ \vdots \\ d_{m,\delta_m-1} \end{bmatrix} = 0 \qquad \Longleftrightarrow \qquad f_{j,1}g_1 + f_{j,2}g_2 + \cdots + f_{j,m}g_m \equiv 0 \bmod \varphi^{n_j} \quad \text{for } 1 \le j \le \ell.$$

The claim now follows by the definition of $D_{A,m,i}$. □

Lemma 2.14. *Suppose, once again, that $\ell \le m-1$.*

Consider a sequence of polynomials

$$g_1, g_2, \dots, g_m \in \mathsf{F}[x]$$

such that the degree of g_k less than δ_k for $1 \le k \le m$.

(a) *If at least one of the polynomials $g_1, g_2, \dots, g_m$ is not divisible by φ then there are exactly $q^{N(m-1)}$ choices of polynomials*

$$f_{j,k} \in \mathsf{F}[x] \qquad \text{for } 1 \le j \le \ell \text{ and } 1 \le k \le m$$

such that $\deg(f_{j,k}) < dn_j$ for all j and k as above, and such that

$$f_{j,1}g_1 + f_{j,2}g_2 + \cdots + f_{j,m}g_m \equiv 0 \bmod \varphi^{n_j}$$

for $1 \le j \le \ell$.

(b) *Let h be a positive integer such that $dh < \delta_1 = \lceil i/m \rceil$. If the polynomials $g_1, g_2, \ldots, g_m$ are all divisible by φ^h, but at least one of these polynomials is not divisible by φ^{h+1}, then there are at most $q^{N(m-1)+dh\ell}$ choices of polynomials*

$$f_{j,k} \in \mathsf{F}[x] \qquad \textit{for } 1 \le j \le \ell \textit{ and } 1 \le k \le m$$

such that $\deg(f_{j,k}) < dn_j$ *for all j and k as above, and such that*

$$f_{j,1}g_1 + f_{j,2}g_2 + \cdots + f_{j,m}g_m \equiv 0 \bmod \varphi^{n_j}$$

for $1 \le j \le \ell$.

(c) *Finally, if the polynomials $g_1, g_2, \ldots, g_m$ are all divisible by φ^h, where $dh \ge \delta_1 = \lceil i/m \rceil$, then there are exactly q^{Nm} choices of polynomials*

$$f_{j,k} \in \mathsf{F}[x] \qquad \textit{for } 1 \le j \le \ell \textit{ and } 1 \le k \le m$$

such that $\deg(f_{j,k}) < dn_j$ *for all j and k as above, and such that*

$$f_{j,1}g_1 + f_{j,2}g_2 + \cdots + f_{j,m}g_m \equiv 0 \bmod \varphi^{n_j}$$

for $1 \le j \le \ell$.

Proof. Suppose $g_1, g_2, \ldots, g_m$ are polynomials in $\mathsf{F}[x]$ that the degree of g_k is less than δ_k for $1 \le k \le m$.

(a) Suppose first that at least one of $g_1, g_2, \ldots, g_m$ is not divisible by φ. In particular, suppose that g_a is not divisible by φ for some integer a such that $1 \le a \le m$. In this case the greatest common divisor of g_a and φ^{n_1} is 1, since φ is irreducible, and there exist polynomials $s, t \in \mathsf{F}[x]$ such that

$$sg_a + t\varphi^{n_1} = 1.$$

It is clear that s and φ^{n_1} are relatively prime as well. Consequently, if $f_{j,k}$ are polynomials in $\mathsf{F}[x]$, for $1 \le j \le \ell$ and $1 \le k \le m$, then, since $n_1 \ge n_j$ for $1 \le j \le \ell$, the argument given in the proof of Lemma 2.9(a) can be applied once again to establish that

$$\begin{aligned} & f_{j,1}g_1 + f_{j,2}g_2 + \cdots + f_{j,m}g_m \equiv 0 \bmod \varphi^{n_j} \qquad \text{for } 1 \le j \le \ell \\ & \quad \iff f_{j,a} \equiv -s(f_{j,1}g_1 + f_{j,2}g_2 + \cdots + f_{j,a-1}g_{a-1} \\ & \qquad\qquad + f_{j,a+1}g_{a+1} + \cdots + f_{j,m}g_m) \bmod \varphi^{n_j} \qquad \text{for } 1 \le j \le \ell. \end{aligned}$$

It follows that there is exactly one choice of the polynomials

$$f_{1,a}, f_{2,a}, \ldots, f_{\ell,a}$$

for any choice of the set of polynomials $f_{j,k}$ for $1 \le j \le \ell$, $1 \le k \le m$ and $k \ne a$, such that the desired equations are satisfied. Since

$$d(n_1 + n_2 + \cdots + n_\ell) = N,$$

there are exactly $q^{N(m-1)}$ choices of the polynomials $f_{j,k}$, for $1 \le j \le \ell$ and $1 \le k \le m$ such that $k \ne a$, establishing part (a) of the claim.

(b) Suppose next that h is a positive integer such that $dh < \delta_1 = \lceil i/m \rceil$, that each of the polynomials $g_1, g_2, \dots, g_m$ are divisible by φ^h, and that at least one of them is not divisible by φ^{h+1}. Once again, suppose that g_a is not divisible by φ^{h+1} for an integer a such that $1 \le a \le m$.

Either $h < n_1$ or $h \ge n_1$; these cases will be considered separately, below.

Suppose first that $h < n_1$, so that $\gcd(\varphi^{n_1}, g_a) = \varphi^h$. Following an argument similar to the one used to prove Lemma 2.9(b), we may continue by writing g_j as $\varphi^h \widehat{g}_j$ for $1 \le j \le m$ and noting that $\widehat{g}_a$ and φ^{n_1} are relatively prime. Consequently, there exist polynomials s and t such that

$$sg_a + t\varphi^{n_1} = \varphi^h \qquad \text{and} \qquad s\widehat{g}_a + t\varphi^{n_1-h} = 1.$$

It is clear from the latter equation that s and φ^{n_1} are relatively prime. It follows, once again, since $n_1 \ge n_j$ for $1 \le j \le \ell$, that

$$\begin{aligned} &f_{j,1}g_1 + f_{j,2}g_2 + \cdots + f_{j,m}g_m \equiv 0 \bmod \varphi^{n_j} \qquad \text{for } 1 \le j \le \ell \\ &\quad \iff f_{j,a}\varphi^h \equiv -s(f_{j,1}g_1 + f_{j,2}g_2 + \cdots + f_{j,a-1}g_{a-1} \\ &\qquad\qquad + f_{j,a+1}g_{a+1} + \cdots + f_{j,m}g_m) \bmod \varphi^{n_j} \qquad \text{for } 1 \le j \le \ell. \end{aligned}$$

Now, if $h \le n_j$, then the proof proceeds as before: We continue by observing that

$$\begin{aligned} &f_{j,a}\varphi^h \equiv -s(f_{j,1}g_1 + f_{j,2}g_2 + \cdots + f_{j,a-1}g_{a-1} + f_{j,a+1}g_{a+1} + \cdots + f_{j,m}g_m) \bmod \varphi^{n_j} \\ &\quad \iff f_{j,a} \equiv -s(f_{j,1}\widehat{g}_1 + f_{j,2}\widehat{g}_2 + \cdots + f_{j,a-1}\widehat{g}_{a-1} + f_{j,a+1}\widehat{g}_{a+1} + \cdots + f_{j,m}\widehat{g}_m) \bmod \varphi^{n_j-h}, \end{aligned}$$

and that the latter equation is satisfied if and only if

$$f_{j,a} = f_{j,H}\varphi^{n_j-h} + f_{j,L}$$

where $f_{j,L}$ is a polynomial with degree less than $d(n_j - h)$ such that

$$f_{j,L} \equiv -s(f_{j,1}\widehat{g}_1 + f_{j,2}\widehat{g}_2 + \cdots + f_{j,a-1}\widehat{g}_{a-1} + f_{j,a+1}\widehat{g}_{a+1} \cdots + f_{j,m}\widehat{g}_m) \bmod \varphi^{n_j-h},$$

and where $f_{j,H}$ is an arbitrarily chosen polynomial in $\mathsf{F}[x]$ with degree less than dh. Consequently, if $1 \le j \le \ell$ and $h \le n_j$ then there are exactly $q^{dn_j(m-1)+dh}$ choices of the polynomials

$$f_{j,1}, f_{j,2}, \dots, f_{j,m} \in \mathsf{F}[x]$$

such that each of these polynomials has degree less than dn_j and such that

$$f_{j,1}g_1 + f_{j,2}g_2 + \cdots + f_{j,m}g_m \equiv 0 \bmod \varphi^{n_j}.$$

On the other hand, if $h > n_j$ then, since $g_1, g_2, \dots, g_m$ are divisible by φ^{n_j}, the equation

$$f_{j,1}g_1 + f_{j,2}g_2 + \cdots + f_{j,m}g_m \equiv 0 \bmod \varphi^{n_j}$$

is satisfied for every choice of polynomials

$$f_{j,1}, f_{j,2}, \dots, f_{j,m} \in \mathsf{F}[x]$$

such that each of the above polynomials has degree less than dn_j. Consequently there are exactly q^{dmn_j} ways to choose the above polynomials in this case. Since $n_j < h$, $dmn_j < dn_j(m-1) + dh$, so that there are at most (indeed, strictly fewer than) $q^{dn_j(m-1)+dh}$ choices of the polynomials

$$f_{j,1}, f_{j,2}, \dots, f_{j,m} \in \mathsf{F}[x]$$

such that each of these polynomials has degree less than dn_j and such that

$$f_{j,1}g_1 + f_{j,2}g_2 + \cdots + f_{j,m}g_m \equiv 0 \bmod \varphi^{n_j}$$

in this case as well.

It follows that the number of choices of polynomials $f_{j,k}$ such that $1 \le j \le \ell$, $1 \le k \le m$, and the given conditions are satisfied, is less than or equal to

$$\prod_{j=1}^{\ell} q^{dn_j(m-1)+dh} = q^{\left(\sum_{j=1}^{\ell} dn_j\right)(m-1)+dh\ell} = q^{N(m-1)+dh\ell},$$

when $h < n_1$.

It remains to consider the case that $h \ge n_1$. Since g_k is divisible by φ^h for $1 \le k \le m$, and since $n_1 \ge n_j$ for $1 \le j \le \ell$, it follows that

$$g_1 \equiv g_2 \equiv \cdots \equiv g_m \equiv 0 \bmod \varphi^{n_j}$$

for every integer j such that $1 \le j \le \ell$. Consequently

$$f_{j,1}g_1 + f_{j,2}g_2 + \cdots + f_{j,m}g_m \equiv 0 \bmod \varphi^{n_j}$$

for $1 \le j \le \ell$, for every choice of polynomials $f_{j,k}$ such that $\deg(f_{j,k}) < dn_j$, for $1 \le j \le \ell$ and $1 \le k \le m$. Since $d(n_1 + n_2 + \cdots + n_\ell) = N$, there are exactly q^{Nm} such choices of these polynomials.

Note, in this case, that

$$\begin{aligned}
Nm &= N(m-1) + N \\
&= N(m-1) + d(n_1 + n_2 + \cdots + n_\ell) && \text{(since } d(n_1 + n_2 + \cdots + n_\ell) = N\text{)} \\
&\le N(m-1) + dn_1\ell && \text{(since } n_j \le n_1 \text{ for } 1 \le j \le \ell\text{)} \\
&\le N(m-1) + dh\ell && \text{(since } n_1 \le h\text{).}
\end{aligned}$$

Consequently $q^{Nm} \le q^{N(m-1)+dh\ell}$ in this case, and there are at most $q^{N(m-1)+dh\ell}$ choice of polynomials $f_{j,k}$ satisfying the desired conditions (for $1 \le j \le \ell$ and $1 \le k \le m$) in this case, as well.

(c) The proof of part (c) of this claim is essentially the same as the proof of Lemma 2.9(c): One notes that if the given conditions are satisfied then

$$g_1 = g_2 = \cdots = g_m = 0$$

so that (in the present case) all choices of the polynomials $f_{j,k}$ such that $1 \le j \le \ell$, $1 \le k \le m$, and the degree of $f_{j,k}$ is less than dn_j for all j and k, satisfy the given conditions. Once again, there are exactly q^{Nm} ways to choose these polynomials.

□

The next result can be established from the above one in much the same way that Lemma 2.10 was established from Lemma 2.9.

Lemma 2.15. *If A is as given in Equations (2.22a) – (2.22e) and m and i are positive integers such that $i \geq m > \ell$, then*

$$D_{A,m,i} < q^{Nm} + q^{N(m-1)+i} \sum_{h \geq 0} q^{-(m-\ell)dh}.$$

Proof. Arguing as in the proof of Lemma 2.10, but using the results of Lemma 2.14 instead of those of Lemma 2.9, we find that if A, m, and i satisfy the conditions given in the statement of the lemma then

$$\begin{aligned} D_{A,m,i} &\leq \sum_{h=0}^{\lfloor (\lfloor i/m \rfloor /d) \rfloor} \left(q^{i-mdh} \cdot q^{N(m-1)+dh\ell} \right) + q^{Nm} \\ &\leq q^{Nm} + \sum_{h \geq 0} q^{N(m-1)+i-(m-\ell)dh} \\ &= q^{Nm} + q^{N(m-1)+i} \sum_{h \geq 0} q^{-(m-\ell)dh}. \end{aligned}$$ □

Using the closed form for a geometric series, we may now conclude that

$$D_{A,m,i} < q^{Nm} + \frac{q^{N(m-1)+i}}{1 - q^{-(m-\ell)d}} = q^{Nm} + \frac{q^{Nm}}{q^{N-i} - q^{N-i-(m-\ell)d}} \tag{2.27}$$

when A is as given in Equations (2.22a) – (2.22e).

2.3.3 Third Case

We next eliminate the assumption that the characteristic polynomial of A is a power of an irreducible polynomial in $\mathsf{F}[x]$. However, the assumption that the number of nontrivial invariant factors is small will be retained.

Suppose now that

$$A = X^{-1} \Lambda X \tag{2.28a}$$

for a nonsingular matrix $X \in \mathsf{F}^{N \times N}$, and for a block diagonal matrix

$$\Lambda = \begin{bmatrix} \Delta_0 & & & \\ & \Delta_1 & & \\ & & \ddots & \\ & & & \Delta_H \end{bmatrix} \in \mathsf{F}^{N \times N}. \tag{2.28b}$$

In this case, Δ_h is a matrix whose characteristic polynomial is a power of an irreducible polynomial φ_h with degree $d_h \geq 1$ in $\mathsf{F}[x]$, for $0 \leq h \leq H$, so that

$$\varphi_0, \varphi_1, \varphi_2, \ldots, \varphi_H$$

are distinct monic irreducible polynomials in $\mathsf{F}[x]$.

It will be useful, in the sequel, to consider the case that the given irreducible polynomial is x separately from other cases. To facilitate this we will suppose that $\varphi_0 = x$ (so that $d_0 = 1$), that $H \geq 0$, and that

$$\varphi_1, \varphi_2, \ldots, \varphi_H$$

are the (remaining) monic irreducible divisors of the characteristic polynomial of A.

To continue, we will suppose that Δ_h is a block diagonal matrix

$$\Delta_h = \begin{bmatrix} Z_{h,1} & & & \\ & Z_{h,2} & & \\ & & \ddots & \\ & & & Z_{h,\ell_h} \end{bmatrix} \in \mathsf{F}^{N_h \times N_h} \tag{2.28c}$$

for an integer $\ell_h \geq 0$ for $0 \leq h \leq H$, and that

$$Z_{h,j} = \begin{bmatrix} 0 & & & & -\alpha_{h,j,0} \\ 1 & 0 & & & -\alpha_{h,j,1} \\ & 1 & & & -\alpha_{h,j,2} \\ & & \ddots & & \vdots \\ & & & 1 & -\alpha_{h,j,d_h n_{h,j}-1} \end{bmatrix} \in \mathsf{F}^{d_h n_{h,j} \times d_h n_{h,j}} \tag{2.28d}$$

is the companion matrix of the polynomial

$$\varphi_h^{n_{h,j}} = x^{d_h n_{h,j}} + \alpha_{h,j,d_h n_{h,j}-1} x^{d_h n_{h,j}-1} + \cdots + \alpha_{h,j,1} x + \alpha_{h,j,0} \tag{2.28e}$$

for $0 \leq h \leq H$ and $1 \leq j \leq \ell_h$, and such that

$$\ell_0 \geq 0 \qquad \text{and} \qquad \ell_h \geq 1 \qquad \text{for } 1 \leq h \leq H, \tag{2.28f}$$

$$n_{h,1} \geq n_{h,2} \geq \cdots \geq n_{h,\ell_h} \geq 1 \qquad \text{for } 0 \leq h \leq H, \tag{2.28g}$$

$$d_h(n_{h,1} + n_{h,2} + \cdots + n_{h,\ell_h}) = N_h \qquad \text{for } 0 \leq h \leq H, \tag{2.28h}$$

and

$$N_0 + N_1 + N_2 + \cdots + N_H = N. \tag{2.28i}$$

If parameters are as defined above then A is nonsingular if and only if $\ell_0 = 0$, while A is nilpotent if and only if $H = 0$.

Finally, let

$$\ell = \max(\ell_0, \ell_1, \ell_2, \ldots, \ell_h). \tag{2.28j}$$

Since we are still interested in the case that the number of invariant factors is small, we will consider the case that $\ell \leq m - 1$.

The next result generalizes Lemma 2.11.

Lemma 2.16. *If A is as given in Equations (2.28a) – (2.28j), then there exists vectors $\zeta_{h,j} \in \mathsf{F}^{N \times 1}$, for $0 \leq h \leq H$ and $1 \leq j \leq \ell_h$, such that*

$$\mathsf{F}^{N \times 1} = \bigoplus_{\substack{0 \leq h \leq H \\ 1 \leq j \leq \ell_h}} V_{h,j}$$

where

$$V_{h,j} = \{\, f(A)\zeta_{h,j} \mid f \in \mathsf{F}[x] \text{ and } \deg(f) < d_h n_{h,j} \,\},$$

and such that

$$\varphi_h^{n_{h,j}}(A)\zeta_{h,j} = 0$$

for $0 \leq h \leq H$ *and* $1 \leq j \leq \ell_h$.

Furthermore, there is exactly one sequence of polynomials

$$f_{0,1}, f_{0,2}, \ldots, f_{0,\ell_0}, \ldots, f_{H,1}, f_{H,2}, \ldots, f_{H,\ell_H} \in \mathsf{F}[x]$$

such that $f_{h,j}$ *has degree less than* $d_h n_{h,j}$*, for* $0 \leq h \leq H$ *and* $1 \leq j \leq \ell_h$*, and such that*

$$y = \sum_{\substack{0 \leq h \leq H \\ 1 \leq j \leq \ell_h}} f_{h,j}(A)\zeta_{h,j}$$

for any given vector $y \in \mathsf{F}^{N \times 1}$.

Sketch of Proof. As in the proof of Lemma 2.11, it is sufficient to identify the values of the vectors $\zeta_{h,j}$ that are mentioned in the statement of the lemma and to check that the given properties are all satisfied.

In this case, it is sufficient to set

$$\zeta_{h,j} = X^{-1} e_{N_0 + N_1 + \cdots + N_{h-1} + d_h(n_{h,1} + n_{h,2} + \cdots + n_{h,j-1}) + 1}$$

for $0 \leq h \leq H$ and $1 \leq j \leq \ell_h$. The proof can then be completed by making a sequence of observations resembling those that were given in the proof of the above-mentioned lemma. □

Corollary 2.17. *If* A *and the vectors* $\zeta_{h,j}$ *are as described in Lemma 2.16, and* $f_{h,j} \in \mathsf{F}[x]$ *for* $0 \leq h \leq H$ *and* $1 \leq j \leq \ell_h$*, then*

$$\sum_{\substack{0 \leq h \leq H \\ 1 \leq j \leq \ell_h}} f_{h,j}(A)\zeta_{h,j} = 0$$

if and only if $f_{h,j}$ *is divisible by* $\varphi_h^{n_{h,j}}$ *in* $\mathsf{F}[x]$ *for all* h *and* j *such that* $0 \leq h \leq H$ *and* $1 \leq j \leq \ell_h$.

The next result generalizes Lemma 2.13 and can be proved by a straightforward generalization of this lemma's proof.

Lemma 2.18. *If* A *is as given in Equations* (2.28a) – (2.28j)*, and* m *and* i *are positive integers such that* $i \geq m$*, then* $D_{A,m,i}$ *is equal to the number of choices of polynomials*

$$f_{h,j,k} \in \mathsf{F}[x] \qquad \textit{for } 0 \leq h \leq H,\ 1 \leq j \leq \ell_h, \textit{ and } 1 \leq k \leq m$$

where $\deg(f_{h,j,k}) < d_h n_{h,j}$ *for all* h*,* j*, and* k *as above, and of polynomials*

$$g_1, g_2, \ldots, g_m \in \mathsf{F}[x]$$

where the degree of g_k *is less than* δ_k *for* $1 \leq k \leq m$*, such that*

$$f_{h,j,1}g_1 + f_{h,j,2}g_2 + \cdots + f_{h,j,m}g_m \equiv 0 \bmod \varphi_h^{n_{h,j}}$$

for all integers h *and* j *such that* $0 \leq h \leq H$ *and* $1 \leq j \leq \ell_h$.

The next lemma generalizes Lemma 2.14.

Lemma 2.19. *Let ψ_A be the minimal polynomial of A, so that*

$$\psi_A = \prod_{h=0}^{\ell} \varphi_h^{n_{h,1}}.$$

Suppose, once again, that $\ell \leq m - 1$.

Now consider a sequence of polynomials

$$g_1, g_2, \ldots, g_m \in \mathsf{F}[x]$$

such that the degree of g_k is less than δ_k for $1 \leq k \leq m$.

(a) *Suppose that*

$$\gcd(g_1, g_2, \ldots, g_m, \psi_A) = 1.$$

Then there are exactly $q^{N(m-1)}$ choices of polynomials

$$f_{h,j,k} \in \mathsf{F}[x] \qquad \textit{for } 0 \leq h \leq H,\ 1 \leq j \leq \ell_h, \textit{ and } 1 \leq k \leq m$$

such that $\deg(f_{h,j,k}) < d_h n_{h,j}$ for all h, j, and k and such that

$$f_{h,j,1}g_1 + h_{h,j,2}g_2 + \cdots + f_{h,j,m}g_m \equiv 0 \bmod \varphi_h^{n_{h,j}}$$

for all integers h and j such that $0 \leq h \leq H$ and $1 \leq j \leq \ell_h$.

(b) *Suppose that*

$$\gcd(g_1, g_2, \ldots, g_m, \psi_A) = \xi$$

for a polynomial $\xi \in \mathsf{F}[x]$ with degree δ, where $1 \leq \delta \leq \delta_1 = \lceil i/m \rceil$. Then there are at most $q^{N(m-1)+\delta\ell}$ choices of polynomials

$$f_{h,j,k} \in \mathsf{F}[x] \qquad \textit{for } 0 \leq h \leq H,\ 1 \leq j \leq \ell_h, \textit{ and } 1 \leq k \leq m$$

such that $\deg(f_{h,j,k}) < d_h n_{h,j}$ for all h, j, and k and such that

$$f_{h,j,1}g_1 + f_{h,j,2}g_2 + \cdots + f_{h,j,m}g_m \equiv 0 \bmod \varphi_h^{n_{h,j}}$$

for all integers h and j such that $0 \leq h \leq H$ and $1 \leq j \leq \ell_h$.

(c) *Finally, suppose that neither of the above cases applies. Then*

$$g_1 = g_2 = \cdots = g_m = 0$$

and there are exactly q^{Nm} choices of polynomials

$$f_{h,j,k} \in \mathsf{F}[x] \qquad \textit{for } 0 \leq h \leq H,\ 1 \leq j \leq \ell_h, \textit{ and } 1 \leq k \leq m$$

such that $\deg(f_{h,j,k}) < d_h n_{h,j}$ for all h, j, and k and such that

$$f_{h,j,1}g_1 + f_{h,j,2}g_2 + \cdots + f_{h,j,m}g_m \equiv 0 \bmod \varphi_h^{n_{h,j}}$$

for all integers h and j such that $0 \leq h \leq H$ and $1 \leq j \leq \ell_h$.

Proof. Let ψ_A, ℓ, and $g_0, g_1, \ldots, g_m$ be as in the statement of the lemma. We will bound the number of choices of the polynomials $f_{h,j,k}$ such that

$$f_{h,j,1}g_1 + f_{h,j,2}g_2 + \cdots + f_{h,j,m}g_m \equiv 0 \bmod \varphi_h^{n_{h,j}} \tag{2.29}$$

for $0 \le h \le H$ and $1 \le j \le \ell_h$ by an extension of the argument that was used to prove Lemma 2.14.

(a) Suppose first that

$$\gcd(g_1, g_2, \ldots, g_m, \psi_A) = 1. \tag{2.30}$$

Fix an integer h such that $0 \le h \le H$, and consider the possible choices for polynomials $f_{h,j,k}$ for $1 \le j \le \ell_h$ and $1 \le k \le m$ such that Equation (2.29) is satisfied for all j. Since there are no polynomials to be selected at all if $h = \ell_h = 0$, we may assume that either $h \ge 1$ or $\ell_0 \ge 1$. In this case, since Equation (2.30) is satisfied and φ_h is irreducible and divides ψ_A, φ_h and g_a must be relatively prime for some integer a such that $1 \le a \le m$.

By the argument used in the proof of part (a) of Lemma 2.14, there is exactly one choice of the polynomials

$$f_{h,1,a}, f_{h,2,a}, \ldots, f_{h,\ell_h,a}$$

for any choice of polynomials $f_{h,j,k}$ for $1 \le j \le \ell_h$ and $1 \le k \le m$ where $k \ne a$, such that Equation (2.29) is satisfied for all j. It follows that there are exactly $q^{N_h(m-1)}$ choices of the polynomials $f_{h,j,k}$, for $1 \le j \le \ell_h$ and $1 \le k \le m$, such that the above-mentioned equations are satisfied.

Consequently, the number of choices of polynomials $f_{h,j,k}$ for $0 \le h \le H$, $1 \le j \le \ell_h$, and $1 \le k \le m$, such that Equation (2.29) is satisfied for all h and j is

$$\prod_{h=0}^{H} q^{N_h(m-1)} = q^{\sum_{h=0}^{H} N_h(m-1)} = q^{N(m-1)},$$

as required to prove part (a).

(b) Suppose next that

$$\gcd(g_1, g_2, \ldots, g_m, \psi_A) = \xi \tag{2.31}$$

for some polynomial $\xi \in \mathsf{F}[x]$ with degree δ, where $1 \le \delta \le \delta_1 = \lceil i/m \rceil$. Since $\psi_A = \prod_{h=0}^{H} \varphi_h^{n_{h,1}}$, and $\varphi_0, \varphi_1, \ldots, \varphi_H$ are pairwise relatively prime irreducible polynomials with degrees $d_0, d_1, \ldots, d_H$ respectively, it follows that

$$\xi = \prod_{h=0}^{H} \psi_h^{m_h}$$

for some sequence of nonnegative integers $m_0, m_1, m_2, \ldots, m_H$ such that

$$0 \le m_h \le n_{h,1} \qquad \text{for } 0 \le h \le H$$

and such that

$$m_0 d_0 + m_1 d_1 + \cdots + m_H d_H = \delta.$$

Fix h such that $0 \le h \le H$. By the argument used to prove part (b) of Lemma 2.14, there are at most $q^{N_h(m-1)+m_h d_h \ell}$ choices of polynomials $f_{h,j,k}$ for $1 \le j \le \ell_h$ and $1 \le k \le m$ such that Equation (2.29) is satisfied for all j.

Consequently, the number of choices of $f_{h,j,k}$ for $0 \leq h \leq H$, $1 \leq j \leq \ell_h$ and $1 \leq k \leq m$ such that Equation (2.29) is satisfied for all h and j is at most

$$\prod_{h=0}^{H} q^{N_h(m-1)+m_h d_h \ell} = q^{\sum_{h=0}^{H}(N_h(m-1)+m_h d_h \ell)} = q^{N(m-1)+\delta\ell},$$

as required to prove part (b).

(c) Finally, suppose that neither of the cases discussed in parts (a) and (b) applies. Since the polynomials $g_1, g_2, \ldots, g_m$ each have degree less than $\delta_1 = \lceil i/m \rceil$, it follows that

$$g_1 = g_2 = \cdots = g_m = 0.$$

The result claimed in part (c) follows because there are exactly q^{Nm} choices of polynomials $f_{h,j,k}$ for $0 \leq h \leq H$, $1 \leq j \leq \ell_h$ and $1 \leq k \leq m$ that satisfy the given degree constraints, and because Equation (2.29) is satisfied for every choice of these polynomials.

□

The next result generalizes Lemma 2.15 and the analysis that follows it.

Lemma 2.20. *If A is as given in Equations (2.28a)–(2.28j), and m and i are positive integers such that $i \geq m > \ell$, then*

$$D_{A,m,i} \leq \begin{cases} q^{Nm} + 6q^{N(m-1)+i}\log_q N & \text{if } m = \ell+1, \\ q^{Nm} + 4q^{N(m-1)+i} & \text{if } m = \ell+2, \\ q^{Nm} + q^{N(m-1)+i}\left(1 + 2q^{\ell-m+1}\right) & \text{if } m \geq \ell+3. \end{cases}$$

Proof. This can be established using Lemma 2.19, just as Lemma 2.15 was established using Lemma 2.14. Let $\xi \in \mathsf{F}[x]$ be a factor of the minimal polynomial ψ_A of A with degree δ.

If $\delta < \delta_1 = \lceil i/m \rceil$ then there are at most $q^{i-m\delta}$ polynomials $g_1, g_2, \ldots, g_m \in \mathsf{F}[x]$ such that the degree of g_k is less than δ_k for $1 \leq k \leq m$ and such that

$$\gcd(g_1, g_2, \ldots, g_m, \psi_A) = \xi.$$

It follows by parts (a) and (b) of Lemma 2.19 that, for each of these choices of $g_1, g_2, \ldots, g_m$, there at most $q^{N(m-1)+\delta\ell}$ choices of the polynomials $f_{h,j,k}$ (for $0 \leq h \leq H$, $1 \leq j \leq \ell_h$, and $1 \leq k \leq m$) such that

$$f_{h,j,1}g_1 + f_{h,j,2}g_2 + \cdots + f_{h,j,m}g_m \equiv 0 \bmod \varphi_h^{n_{h,j}} \tag{2.32}$$

for all h and j such that $0 \leq h \leq H$ and $1 \leq j \leq \ell_h$.

On the other hand, if $\delta \geq \delta_1$ and $g_1, g_2, \ldots, g_m$ are polynomials such that the degree of g_k is less than $\delta_k \leq \delta_1$ for $1 \leq k \leq m$ and such that

$$\gcd(g_1, g_2, \ldots, g_m, \psi_A) = \xi$$

then $\xi = \psi_A$,

$$g_1 = g_2 \cdots = g_m = 0,$$

and all possible choices of $f_{h,j,k}$ satisfy Equation (2.32), giving an additional q^{Nm} choices of these polynomials such that the above equations are satisfied.

It follows by the above that

$$D_{A,m,i} \leq q^{Nm} + q^{N(m-1)+i} \sum_{\substack{\xi \in \mathsf{F}[x] \\ \xi \text{ divides } \psi_A \\ \deg(\xi) < \lceil i/m \rceil}} q^{(\ell - m)\deg(\xi)}. \tag{2.33}$$

Consider any polynomial ξ with degree less than $\lceil i/m \rceil$ that divides the minimal polynomial ψ_A of A. It follows by a consideration of the factorization of ψ_A that

$$\xi = \prod_{h=0}^{H} \varphi_h^{m_h},$$

for nonnegative integers $m_0, m_1, \ldots, m_h$ such that $0 \leq m_h \leq n_{h,1}$ for $0 \leq h \leq H$ and such that

$$d_0 m_0 + d_1 m_1 + \cdots + d_H m_H = \deg(\xi). \tag{2.34}$$

Combining inequality (2.33) and equation (2.34), we have that

$$\begin{aligned} D_{A,m,i} &\leq q^{Nm} + q^{N(m-1)+i} \sum_{m_0, m_1, \ldots, m_H \geq 0} q^{-(m-\ell)(d_0 m_0 + d_1 m_1 + \ldots d_H m_H)} \\ &= q^{Nm} + q^{N(m-1)+i} \prod_{h=0}^{H} \sum_{j \geq 0} \left(q^{-(m-\ell)d_h j} \right) \\ &= q^{Nm} + q^{N(m-1)+i} \prod_{h=0}^{H} \left(1 - q^{-(m-\ell)d_h} \right)^{-1}. \end{aligned}$$

It follows that an upper bound for

$$\prod_{h=0}^{H} \left(1 - q^{-(m-\ell)d_h} \right)^{-1}$$

can be used to produce an upper bound for $D_{A,m,i}$. Upper bounds that are sufficient to establish the claim are developed in the rest of this argument.

Suppose first that $m = \ell + 1$. In this case, since $\varphi_0, \varphi_1, \ldots, \varphi_H$ are distinct irreducible polynomials in $\mathsf{F}[x]$ that each divide ψ_A, and ψ_A has degree at most N, one can apply Proposition 3 of Wiedemann [15] to conclude that

$$\prod_{h=0}^{H} \left(1 - q^{-(m-\ell)d_h} \right) = \prod_{h=0}^{H} \left(1 - q^{-d_h} \right) \geq \frac{1}{6 \log_q N}.$$

Consequently

$$\prod_{h=0}^{H} \left(1 - q^{-(m-\ell)d_h} \right)^{-1} \leq 6 \log_q N,$$

as required to establish the claimed upper bound on $D_{A,m,i}$ in the case that $m = \ell + 1$.

Suppose next that $m = \ell + c$ for some integer $c \geq 2$. In this case, one can use the fact that there are at most q^j/j monic irreducible polynomials with degree j in $\mathsf{F}[x]$, for any positive integer j, to establish that

$$\begin{aligned}
\prod_{h=0}^{H} \left(1 - q^{-(m-\ell)d_h}\right) &\geq 1 - \sum_{h=0}^{H} q^{-(m-\ell)d_h} \\
&= 1 - \sum_{h=0}^{H} q^{-cd_h} \\
&\geq 1 - \sum_{j \geq 1} \frac{q^j}{j} \cdot q^{-cj} \\
&= 1 - \sum_{j \geq 1} \frac{q^{-(c-1)j}}{j} \\
&= 1 - q^{-(c-1)} - \sum_{j \geq 2} \frac{q^{-(c-1)j}}{j} \\
&\geq 1 - q^{-(c-1)} - \sum_{j \geq 2} \frac{q^{-(c-1)j}}{2} \\
&= 1 - q^{-(c-1)} - \frac{q^{-2(c-1)}}{2} \sum_{j \geq 0} q^{-(c-1)j} \\
&\geq 1 - q^{-(c-1)} - \frac{q^{-2(c-1)}}{2} \sum_{j \geq 0} 2^{-j} && \text{(since } q \geq 2 \text{ and } c \geq 2\text{)} \\
&= 1 - q^{-(c-1)} - q^{-2(c-1)}.
\end{aligned}$$

Consequently

$$\prod_{h=0}^{H} \left(1 - q^{-(m-\ell)d_h}\right)^{-1} \leq \left(1 - q^{-(c-1)} - q^{-2(c-1)}\right)^{-1}.$$

If $c = 2$ then

$$\begin{aligned}
1 - q^{-(c-1)} - q^{-2(c-1)} &= 1 - q^{-1} - q^{-2} \\
&\geq 1 - 2^{-1} - 2^{-2} && \text{(since } q \geq 2\text{)} \\
&= \tfrac{1}{4},
\end{aligned}$$

so that

$$\left(1 - q^{-(c-1)} - q^{-2(c-1)}\right)^{-1} \leq 4,$$

as required to establish the claimed upper bound in the case that $m = \ell + 2$.

Finally, suppose that $m = \ell + c$ for an integer $c \geq 3$. In this case it should be noted (using the inequalities that have already been established) that

$$\left(1 + 2q^{\ell-m+1}\right) \prod_{h=0}^{H} \left(1 - q^{(m-\ell)d_h}\right)$$

$$
\begin{aligned}
&\geq \left(1 + 2q^{-(c-1)}\right)\left(1 - q^{-(c-1)} - q^{-2(c-1)}\right) \\
&= 1 + q^{-3(c-1)}\left(q^{2(c-1)} - 3q^{c-1} - 2\right) \\
&\geq 1,
\end{aligned}
$$

since $q^{c-1} \geq 2^2 = 4$ when $c \geq 3$ and since $z^2 - 3z - 2 \geq 0$ for every real number z such that $z \geq 4$. Consequently

$$\prod_{h=0}^{H} \left(1 - q^{-(m-\ell)d_h}\right)^{-1} \leq 1 + 2q^{\ell-m+1}$$

if $m = \ell + c$ for $c \geq 3$, as required to establish the upper bound claimed for this case. □

Recall that $\sigma_{A,m,i}(j)$ is the probability that the matrix $M_{A,m,i,v_1,v_2,\dots,v_m}$ is rank deficient by at least j, assuming that the vectors $v_1, v_2, \dots, v_m$ are chosen uniformly and independently from $\mathsf{F}^{N\times 1}$.

If $i \leq N$ then it follows by Equation (2.12) (on page 5) that

$$q^{Nm} + (q-1)q^{Nm+j-1}\sigma_{A,m,i}(j) \leq D_{A,m,i}.$$

This can be used, along with the previous lemma, to establish the following.

Corollary 2.21. *Let A and ℓ be as described above. Suppose that $m \leq i \leq N$ and that j is a positive integer such that $1 \leq j \leq i$.*

$$\sigma_{A,m,i}(j) \leq \begin{cases} \frac{6\log_q N}{(q-1)\,q^{N-i+j-1}} & \text{if } m = \ell+1, \\ \frac{4}{(q-1)\,q^{N-i+j-1}} & \text{if } m = \ell+2, \\ \frac{1+2q^{\ell-m+1}}{(q-1)\,q^{N-i+j-1}} & \text{if } m \geq \ell+3. \end{cases}$$

If $i > N$ then it follows by Equation (2.13) that

$$q^{N(m-1)+i} + (q-1)q^{N(m-1)+i+j-1}\sigma_{A,m,i}(j) \leq D_{A,m,i}.$$

This can be used to establish the following as well.

Corollary 2.22. *Let A and ℓ be as described above. Suppose that $i \geq m$, $i > N$, and that j is a positive integer such that $1 \leq j \leq N$.*

$$\sigma_{A,m,i}(j) \leq \begin{cases} \frac{1}{(q-1)\,q^{i+j-N-1}} + \frac{6\log_q N}{(q-1)\,q^{j-1}} & \text{if } m = \ell+1, \\ \frac{1}{(q-1)\,q^{i+j-N-1}} + \frac{3}{(q-1)\,q^{j-1}} & \text{if } m = \ell+2, \\ \frac{1}{(q-1)\,q^{i+j-N-1}} + \frac{2q^{\ell-m+1}}{(q-1)\,q^{j-1}} & \text{if } m \geq \ell+3. \end{cases}$$

3 Two Modifications

The following modifications of the preceding analysis will be useful for the analysis of block Lanczos and block Wiedemann algorithms.

3.1 Inclusion of an Additional Factor

Once again, let us consider the matrix $M_{A,m,i,v_1,v_2,\dots,v_m}$ introduced in Definition 2.1 on page 3. It may be useful to compare the following definitions to Definitions 2.2 and 2.3.

Definition 3.1. Suppose that m and i are positive integers. Let $\widehat{D}_{A,m,i}$ be the number of choices of vectors $v_1, v_2, \dots, v_m \in \mathsf{F}^{N\times 1}$ and scalars

$$d_{1,0}, d_{1,1}, \dots, d_{1,\delta_1 - 1}, \dots, d_{m,0}, d_{m,1}, \dots, d_{m,\delta_m - 1}$$

for $\delta_1, \delta_2, \dots, \delta_m$ as given in Definition 2.1, such that

$$A \cdot M_{A,m,i,v_1,v_2,\dots,v_m} \begin{bmatrix} d_{1,0} \\ d_{1,1} \\ \vdots \\ d_{m,\delta_m - 1} \end{bmatrix} = 0. \tag{3.1}$$

Once again, let r be the rank of A; clearly, the product of A and $M_{A,m,i,v_1,v_2,\dots,v_m}$ has rank at most r. This is reflected in the definition that is given next.

Definition 3.2. Suppose that m and i are positive integers and that j is a nonnegative integer. Let $\widehat{\rho}_{A,m,i}(j)$ be the probability that the matrix $A \cdot M_{A,m,i,v_1,v_2,\dots,v_m}$ is rank deficient by j, that is,

$$\widehat{\rho}_{A,m,i}(j) = \mathsf{Prob}\left[\operatorname{rank}\left(A \cdot M_{A,m,i,v_1,v_2,\dots,v_m}\right) = i - j\,\right] \quad \text{if } i \le r, \tag{3.2}$$

and

$$\widehat{\rho}_{A,m,i}(j) = \mathsf{Prob}\left[\operatorname{rank}\left(A \cdot M_{A,m,i,v_1,v_2,\dots,v_m}\right) = r - j\,\right] \quad \text{if } i > r, \tag{3.3}$$

and let $\widehat{\sigma}_{A,m,i}(j)$ be the probability that this matrix is rank deficient by at least j, that is,

$$\widehat{\sigma}_{A,m,i}(j) = \mathsf{Prob}\left[\operatorname{rank}\left(A \cdot M_{A,m,i,v_1,v_2,\dots,v_m}\right) \le i - j\,\right] \quad \text{if } i \le j, \tag{3.4}$$

and

$$\widehat{\sigma}_{A,m,k}(j) = \mathsf{Prob}\left[\operatorname{rank}\left(A \cdot M_{A,m,i,v_1,v_2,\dots,v_m}\right) \le r - j\,\right] \quad \text{if } i > r, \tag{3.5}$$

when the vectors $v_1, v_2, \dots, v_m$ are chosen uniformly and independently from $\mathsf{F}^{N\times 1}$.

Lemmas 2.4 and 2.5 can be applied to establish the following equations, which should be compared to Equations (2.10) – (2.13): Lemma 2.4 is used to establish Equation (3.7) when $r < i \le N$, and Lemma 2.5 is used to establish this when $i > N$.

$$\widehat{D}_{A,m,i} = \sum_{j=0}^{i} q^{Nm+j}\, \widehat{\rho}_{A,m,i}(j) \quad \text{if } i \le r, \tag{3.6}$$

$$\widehat{D}_{A,m,i} = \sum_{j=0}^{r} q^{Nm-r+i+j}\, \widehat{\rho}_{A,m,i}(j) \quad \text{if } i > r, \tag{3.7}$$

$$\widehat{D}_{A,m,i} = q^{Nm} + (q-1) \sum_{j=1}^{i} q^{Nm+j-1}\, \widehat{\sigma}_{A,m,i}(j) \quad \text{if } i \le r, \tag{3.8}$$

and

$$\widehat{D}_{A,m,i} = q^{Nm-r+i} + (q-1)\sum_{j=1}^{r} q^{Nm-r+i+j-1}\,\widehat{\sigma}_{A,m,i}(j) \quad \text{if } i > r. \tag{3.9}$$

Bounds on $\widehat{\rho}_{A,m,i}(j)$ and $\widehat{\sigma}_{A,m,i}(j)$ resembling those established for $\rho_{A,m,i}(j)$ and $\sigma_{A,m,i}(j)$ in Lemma 2.20 and Corollary 2.21 follow by a modification of the analysis in Section 2.3. The required changes are summarized below.

3.1.1 Modification of the First Case

Suppose first that A is similar to a companion matrix and, furthermore, that its characteristic polynomial is a power of a monic irreducible polynomial φ. In particular, suppose that A is as described at the beginning of Section 2.3.1 and that A, Z, φ, d, and n are as in Equations (2.14a) – (2.14d).

Either $\varphi = x$ or $\varphi \neq x$. It will be useful to consider these cases separately.

Case: $\boldsymbol{\varphi = x}$. In this Lemma 2.6 and Corollary 2.7 are applicable, much as before. However, since $D_{A,m,i}$ and $\widehat{D}_{A,m,i}$ have different definitions, a result that is similar, but not identical, to Lemma 2.8 can be obtained using the argument to establish that lemma: If A is as described here, then $\widehat{D}_{A,m,i}$ is equal to the number of choices of polynomials

$$f_1, f_2, \dots, f_m \in \mathsf{F}[x]$$

where the degree of f_j is less than $N = n$ for $1 \leq j \leq m$, and of polynomials

$$g_1, g_2, \dots, g_m \in \mathsf{F}[x]$$

where the degree of g_j is less than δ_j for $1 \leq j \leq m$, such that

$$f_1 g_1 + f_2 g_2 + \cdots + f_m g_m \equiv 0 \bmod x^{n-1}. \tag{3.10}$$

One can proceed, as in the development of Lemma 2.9, by the counting the number of choices of polynomials

$$f_1, f_2, \dots, f_m$$

satisfying Equation (3.10) for any given choice of $g_1, g_2, \dots, g_m$. Slightly different numbers of choices (than in the above lemma) are obtained:

- If at least one of the polynomials $g_1, g_2, \dots, g_m$ is not divisible by x then there are exactly $q^{N(m-1)+1}$ choices of polynomials
$$f_1, f_2, \dots, f_m \in \mathsf{F}[x],$$
each with degree less than N, such that
$$f_1 g_1 + f_2 g_2 + \cdots + f_m g_m \equiv 0 \bmod x^{n-1}.$$
- Let h be a positive integer such that $h < \delta_1 = \lceil i/m \rceil$. If the polynomials $g_1, g_2, \dots, g_m$ are all divisible by x^h, but at least one of these polynomials is not divisible by x^{h+1}, then there are exactly $q^{N(m-1)+h+1}$ choices of polynomials $f_1, f_2, \dots, f_m \in \mathsf{F}[x]$, each with degree less than N, such that the above equation is satisfied.

- Finally, if the polynomials $g_1, g_2, \dots, g_m$ are all divisible by x^h, where $h \geq \delta_1 = \lceil i/m \rceil$, then there exactly q^{Nm} choices of polynomials $f_1, f_2, \dots, f_m \in \mathsf{F}[x]$, each with degree less than N, such that the above equation is satisfied (because $g_1 = g_2 = \cdots = g_m = 0$ in this case).

The argument used to establish Lemma 2.10 can now be used to show that

$$\widehat{D}_{A,m,i} < q^{Nm} + q^{N(m-1)+i+1} \sum_{h \geq 0} q^{-(m-1)h},$$

so that

$$\widehat{D}_{A,m,i} < q^{Nm} + \frac{q^{Nm}}{q^{N-i-1} - q^{N-i-m}}.$$

Case: $\boldsymbol{\varphi \neq x}$. In this case the matrix A is nonsingular. One can see by Definition 2.1 that

$$A \cdot M_{A,m,i,v_1,v_2,\dots,v_m} = M_{A,m,i,Av_1,Av_2,\dots,Av_m}$$

for any choice of vectors $v_1, v_2, \dots, v_m$. Since $Av_1, Av_2, \dots, Av_m$ are uniformly and independently selected from $\mathsf{F}^{N \times 1}$ if $v_1, v_2, \dots, v_m$ are, it can be argued in this case that

$$\widehat{D}_{A,m,i} = D_{A,m,i}.$$

Consequently it follows by Lemma 2.10 that

$$\widehat{D}_{A,m,i} < q^{Nm} + q^{N(m-1)+i} \sum_{h \geq 0} q^{-(m-1)dh} = q^{Nm} + \frac{q^{Nm}}{q^{N-i} - q^{N-i-(m-1)d}}$$

in this case.

3.1.2 Modification of the Second Case

Suppose next that A is as described at the beginning of Section 2.3.2, so that Equations (2.22a) – (2.22e) are satisfied. Once again, the minimal polynomial of A is a power of some irreducible polynomial $\varphi \in \mathsf{F}[x]$, and it is useful to consider the cases $\varphi = x$ and $\varphi \neq x$ separately.

Case: $\boldsymbol{\varphi = x}$. In this case, material in Section 2.3.2 is modified in much the same way as material in Section 2.3.1 was modified, above. Lemma 2.11 and Corollary 2.12 are unchanged, and are used to establish a result that replaces Lemma 2.13: If A is as described here, then $\widehat{D}_{A,m,i}$ is equal to the number of choices of polynomials

$$f_{j,k} \in \mathsf{F}[x] \qquad \text{for } 1 \leq j \leq \ell \text{ and } 1 \leq k \leq m$$

where $\deg(f_{j,k}) < n_j$ for all j and k as above, and of polynomials

$$g_1, g_2, \dots, g_m \in \mathsf{F}[x]$$

where the degree of g_k is less than δ_k for $1 \leq k \leq m$, such that

$$f_{j,1}g_1 + f_{j,2}g_2 + \cdots + f_{j,m}g_m \equiv 0 \bmod x^{n_j - 1}$$

for every integer j such that $1 \leq j \leq \ell$. Continuing the analysis as in Section 2.3.2, one eventually confirms that

$$\widehat{D}_{A,m,i} < q^{Nm} + q^{N(m-1)+i+\ell} \sum_{h \geq 0} q^{-(m-\ell)h} = q^{Nm} + \frac{q^{Nm}}{q^{N-i-\ell} - q^{N-i-m}}.$$

Case: $\varphi \neq x$. Once again, it can be shown in this case that

$$\widehat{D}_{A,m,i} = D_{A,m,i}$$

so that it follows by the material in Section 2.3.2 that

$$\widehat{D}_{A,m,i} < q^{Nm} + q^{N(m-1)+i} \sum_{h \geq 0} q^{-(m-\ell)dh} = q^{Nm} + \frac{q^{Nm}}{q^{N-i} - q^{N-i-(m-\ell)d}}.$$

3.1.3 Modification of the Third Case

Suppose next that Equations (2.28a) – (2.28j) are satisfied.

Special Case: No Nontrivial Invariant Factors To begin, suppose that A has no nontrivial invariant factors that all — so each invariant factor of A (in $\mathsf{F}[z]$) is either divisible by z^2 or not divisible by z. In other words, let us suppose that A does not have any nilpotent blocks with size 1.

Lemma 2.16 and Corollary 2.17 hold, as before. These can be used to establish the following result, which replaces Lemma 2.18 in this analysis: $\widehat{D}_{A,m,i}$ is equal to the number of choices of polynomials

$$f_{h,j,k} \in \mathsf{F}[x] \qquad \text{for } 0 \leq h \leq H,\ 1 \leq j \leq \ell_h, \text{ and } 1 \leq k \leq m$$

where $\deg(f_{h,j,k}) < d_h n_{h,j}$ for all h, j and k as above, and of polynomials

$$g_1, g_2, \dots, g_m \in \mathsf{F}[x]$$

where the degree of g_k is less than δ_k for $1 \leq k \leq m$, such that

$$f_{0,j,1}g_1 + f_{0,j,2}g_2 + \dots, f_{0,j,m}g_m \equiv 0 \bmod x^{n_{0,j}-1} \qquad \text{for } 1 \leq j \leq \ell_0, \tag{3.11}$$

and such that

$$f_{h,j,1}g_1 + f_{h,j,2}g_2 + \dots + f_{h,j,m}g_m \equiv 0 \bmod \varphi_h^{n_{h,j}} \qquad \text{for } 1 \leq h \leq H \text{ and } 1 \leq j \leq \ell_h. \tag{3.12}$$

This can be used to establish a result that resembles parts (a) and (b) of Lemma 2.19: If ψ_A is the minimal polynomial of A,

$$g_1, g_2, \dots, g_m \in \mathsf{F}[x]$$

is a sequence of polynomials such that the degree of g_k is less than δ_k for $1 \leq k \leq m$, and

$$\gcd(g_1, g_2, \dots, g_m, \psi_A) = \xi$$

for a polynomial with degree δ, where $0 \leq \delta < \delta_1 = \lceil i/m \rceil$, then there are at most $q^{N(m-1)+\delta\ell+\ell_0}$ choices of polynomials $f_{h,j,k} \in \mathsf{F}[x]$ for $0 \leq h \leq H$, $1 \leq j \leq \ell_h$, and $1 \leq k \leq m$, such that $\deg(f_{h,j,k}) < d_h n_{h,j}$ for all h, j, and k, and such that the conditions shown at lines (3.11) and (3.12) are satisfied. On the other hand, if the degree of

$$\gcd(g_1, g_2, \dots, g_m, \psi_A)$$

is greater than or equal to δ_1 then it must be the case (as before) that

$$g_1 = g_2 = \dots = g_m = 0,$$

so that there are exactly q^{Nm} choices of the polynomials $f_{h,j,k}$ such that the given conditions are all satisfied.

Note that if A has rank r then

$$r = N - \ell_0$$

in this case. A modification of the analysis used to prove Lemma 2.20 establishes that

$$\widehat{D}_{A,m,i} \leq q^{Nm} + q^{Nm-r+i} \prod_{h=0}^{H} \left(1 - q^{-(m-\ell)d_h}\right)^{-1},$$

so that (by a continuation of this analysis)

$$\widehat{D}_{A,m,i} \leq \begin{cases} q^{Nm} + 6q^{Nm-r+i} \log_q N & \text{if } m = \ell + 1, \\ q^{Nm} + 4q^{Nm-r+i} & \text{if } m = \ell + 2, \\ q^{Nm} + q^{Nm-r+i}(1 + 2q^{\ell-m+1}) & \text{if } m \geq \ell + 3. \end{cases}$$

This can be used along with Equation (3.4) to establish that if $i \leq r$ then

$$\widehat{\sigma}_{A,m,i}(j) \leq \begin{cases} \frac{6 \log_q N}{(q-1)\, q^{r-i+j-1}} & \text{if } m = \ell + 1, \\ \frac{4}{(q-1)\, q^{r-i+j-1}} & \text{if } m = \ell + 2, \\ \frac{1+2q^{\ell-m+1}}{(q-1)\, q^{r-i+j-1}} & \text{if } m \geq \ell + 3. \end{cases}$$

This can also be used along with Equation (3.5) to establish that if $i > r$ then

$$\widehat{\sigma}_{A,m,i}(j) \leq \begin{cases} \frac{1}{(q-1)\, q^{i+j-r-1}} + \frac{6 \log_q N}{(q-1)\, q^{j-1}} & \text{if } m = \ell + 1, \\ \frac{1}{(q-1)\, q^{i+j-r-1}} + \frac{3}{(q-1)\, q^{j-1}} & \text{if } m = \ell + 2, \\ \frac{1}{(q-1)\, q^{i+j-r-1}} + \frac{2q^{\ell-m+1}}{(q-1)\, q^{j-1}} & \text{if } m \geq \ell + 3. \end{cases}$$

General Case: In general a matrix $A \in \mathsf{F}^{N \times N}$ is similar to a block diagonal matrix, so that

$$A = X^{-1} \begin{bmatrix} \widehat{A} & 0 \\ 0 & 0 \end{bmatrix} X$$

for a nonsingular matrix $X \in \mathsf{F}^{N \times N}$, where $\widehat{A} \in \mathsf{F}^{\widehat{N} \times \widehat{N}}$ is a matrix with no nontrivial invariant factors (as considered above) and for an integer $\widehat{N}$ such that $0 \leq \widehat{N} \leq N$. The matrices $\widehat{A}$ and A clearly have the same rank in this case.

Notice that if $v \in \mathsf{F}^{N \times 1}$ then

$$v = X^{-1} \begin{bmatrix} \widehat{v} \\ \widetilde{v} \end{bmatrix}$$

for vectors $\widehat{v} \in \mathsf{F}^{\widehat{N} \times 1}$ and $\widetilde{v} \in \mathsf{F}^{(N-\widehat{N}) \times 1}$. Furthermore, if v is selected uniformly from $\mathsf{F}^{N \times 1}$ then the corresponding vectors $\widehat{v}$ and $\widetilde{v}$ are selected uniformly and independently from $\mathsf{F}^{\widehat{N} \times 1}$ and from $\mathsf{F}^{(N-\widehat{N}) \times 1}$ respectively. It is easily checked that if j is a positive integer and v and $\widehat{v}$ are as above then

$$A^j v = X^{-1} \begin{bmatrix} \widehat{A}^j \widehat{v} \\ 0 \end{bmatrix}$$

as well. Consequently, if $v_1, v_2, \ldots, v_m$ are uniformly and independently selected from $\mathsf{F}^{N\times 1}$ then

$$A \cdot M_{A,m,i,v_1,v_2,\ldots,v_n} = X^{-1} \begin{bmatrix} \widehat{A} M_{\widehat{A},m,i,\widehat{v}_1,\widehat{v}_2,\ldots,\widehat{v}_n} \\ 0 \end{bmatrix},$$

where $\widehat{v}_1, \widehat{v}_2, \ldots, \widehat{v}_m$ are chosen uniformly and independently from $\mathsf{F}^{\widehat{N}\times 1}$.

Since A and $\widehat{A}$ have the same number of *nontrivial* invariant factors, the next results follow from the inequalities that have been established for the case that A has no nontrivial invariant factors at all.

Lemma 3.3. *Let $A \in \mathsf{F}^{N\times N}$ be a matrix with rank r and with ℓ nontrivial invariant factors. Let m, i and j be positive integers such that $i \geq m > \ell$. If $i \leq r$ and $1 \leq j \leq i$ then*

$$\widehat{\sigma}_{A,m,i}(j) \leq \begin{cases} \frac{6\log_q N}{(q-1)\,q^{r-i+j-1}} & \text{if } m = \ell+1, \\ \frac{4}{(q-1)\,q^{r-i+j-1}} & \text{if } m = \ell+2, \\ \frac{1+2q^{\ell-m+1}}{(q-1)\,q^{r-i+j-1}} & \text{if } m \geq \ell+3. \end{cases}$$

and if $i > r$ and $1 \leq j \leq r$ then

$$\widehat{\sigma}_{A,m,i}(j) \leq \begin{cases} \frac{1}{(q-1)\,q^{i+j-r-1}} + \frac{6\log_q N}{(q-1)\,q^{j-1}} & \text{if } m = \ell+1, \\ \frac{1}{(q-1)\,q^{i+j-r-1}} + \frac{3}{(q-1)\,q^{j-1}} & \text{if } m = \ell+2, \\ \frac{1}{(q-1)\,q^{i+j-r-1}} + \frac{2q^{\ell-m+1}}{(q-1)\,q^{j-1}} & \text{if } m \geq \ell+3. \end{cases}$$

3.2 Avoidance of a Subspace

It will also be useful to consider the probability that the vector

$$M_{A,m,i,v_1,v_2,\ldots,v_m} \begin{bmatrix} d_{1,0} \\ d_{1,1} \\ \vdots \\ d_{m,\delta_m-1} \end{bmatrix}$$

belongs to a given subspace of $\mathsf{F}^{N\times 1}$. With that in mind we will generalize Definition 2.2 as follows.

Definition 3.4. Suppose that m and i are positive integers and that $x \in \mathsf{F}^{i\times 1}$. Let $\overline{D}_{A,m,i}(x)$ be the number of choices of vectors $v_1, v_2, \ldots, v_m \in \mathsf{F}^{N\times 1}$ and scalars

$$d_{1,0}, d_{1,1}, \ldots, d_{1,\delta_1-1}, \ldots, d_{m,0}, d_{m,1}, \ldots, d_{m,\delta_m-1}$$

such that

$$M_{A,m,i,v_1,v_2,\ldots,v_m} \begin{bmatrix} d_{1,0} \\ d_{1,1} \\ \vdots \\ d_{m,\delta_m-1} \end{bmatrix} = x.$$

Lemma 3.5. *Let m, i, and x be as above.*

If $x = 0$ then

$$\overline{D}_{A,m,i}(x) = D_{A,m,i}.$$

Otherwise

$$\overline{D}_{A,m,i}(x) \leq D_{A,m,i} - q^{Nm}.$$

Proof. If $x = 0$ then the stated equality follows by a comparison of Definitions 2.2 and 3.4, which are clearly equivalent in this case.

Suppose instead that $x \neq 0$. We may assume that A is as given in Equations (2.28a)–(2.28j), although no relationship between the values of the parameters ℓ and m should be assumed. The analysis in Section 2.3.3 can now be modified to show that $\overline{D}_{A,m,i}(x)$ is the number of choices of polynomials

$$f_{h,j,k} \in \mathsf{F}[x] \qquad \text{for } 0 \leq h \leq H,\ 1 \leq j \leq \ell_h, \text{ and } 1 \leq k \leq m$$

where $\deg(f_{h,j,k}) < d_h n_{h,j}$ for all h, j, and k as above, and of polynomials

$$g_1, g_2, \dots, g_m \in \mathsf{F}[x]$$

where the degree of g_k is less than δ_k for $1 \leq k \leq m$, such that

$$f_{h,j,1}g_1 + f_{h,j,2}g_2 + \cdots + f_{h,j,m}g_m \equiv \lambda_{h,j} \bmod \varphi_h^{n_{h,j}}$$

for all integers h and j such that $0 \leq h \leq H$ and $1 \leq j \leq \ell_h$, and where $\lambda_{h,j}$ are polynomials in $\mathsf{F}[x]$ such that $\deg(\lambda_{h,j}) < d_h n_{h,j}$, for $0 \leq h \leq H$ and $1 \leq j \leq \ell_h$, that depend on the vector x — in particular, these are polynomials whose coefficients are chosen as entries of the vector $X^{-1}x$ if Equations (2.28a)–(2.28j) are satisfied. If $x \neq 0$ then the vector $X^{-1}x$ is also nonzero, so that at least one of the polynomials $\lambda_{h,j}$ is nonzero as well.

Now consider integers h and j such that $0 \leq h \leq H$ and $1 \leq j \leq \ell_h$. If at least one of the polynomials $g_1, g_2, \dots, g_m$ is nonzero then the number of choices of polynomials $f_{h,j,k}$ (for $1 \leq k \leq m$) such that

$$f_{h,j,1}g_1 + f_{h,j,2}g_2 + \dots, +f_{h,j,m}g_m \equiv \lambda_{h,j} \bmod \varphi_h^{n_{h,j}} \tag{3.13}$$

is less than or equal to the number of choices of polynomials $f_{h,j,k}$ (for $1 \leq k \leq m$) such that

$$f_{h,j,1}g_1 + f_{h,j,2}g_2 + \dots, +f_{h,j,m}g_m \equiv 0 \bmod \varphi_h^{n_{h,j}}. \tag{3.14}$$

Indeed, either $\lambda_{h,j}$ is divisible by

$$\gcd(g_1, g_2, \dots, g_m, \varphi_h^{n_{h,j}}),$$

and there is the same number of choices of these polynomials in each case, or $\lambda_{h,j}$ is not divisible by the above greatest common divisor, and there are no choices of polynomials satisfying the condition at line (3.13) at all.

On the other hand, if $g_1 = g_2 = \cdots = g_m = 0$ then there are q^{Nm} choices of the set of polynomials $f_{h,j,k}$ for $0 \leq h \leq H$, $1 \leq j \leq \ell_h$, and $1 \leq k \leq m$ such that $\deg(f_{h,j,k}) < d_h n_{h,j}$ and the condition at line (3.14) is satisfied for all h and j, because all choices of these polynomials cause the condition to be satisfied, but there are no choices of these polynomials at all that satisfy the condition at line (3.13) to be satisfied, because $\lambda_{h,j} \neq 0$ for at least one choice of h and j. It follows that

$$\overline{D}_{A,m,i}(x) \leq D_{A,m,i} - q^{Nm}$$

if $x \neq 0$, as claimed. □

The above result will be used to analyze the probability that the following matrices have low rank.

Definition 3.6. Suppose that m, i, and d are positive integers, and that

$$v_1, v_2, \dots, v_m \in \mathsf{F}^{N \times 1} \qquad \text{and} \qquad x_1, x_2, \dots, x_d \in \mathsf{F}^{N \times 1}.$$

Let

$$\widetilde{M}_{A,m,i,v_1,v_2,\dots,v_m}(x_1, x_2, \dots, x_m) = \begin{bmatrix} M_{A,m,i,v_1,v_2,\dots,v_m} & X \end{bmatrix} \in \mathsf{F}^{N \times (i+d)},$$

where $M_{A,m,i,v_1,v_2\dots v_m} \in \mathsf{F}^{N \times i}$ is as in Definition 2.1 and where

$$X = \begin{bmatrix} x_1 & x_2 & \dots & x_d \end{bmatrix} \in \mathsf{F}^{N \times d}.$$

Definition 3.7. Suppose once again that m, i, and d are positive integers, and that

$$x_1, x_2, \dots, x_d \in \mathsf{F}^{N \times 1}.$$

Let $\widetilde{D}_{A,m,i}(x_1, x_2, \dots, x_d)$ be the number of choices vectors $v_1, v_2, \dots, v_m \in \mathsf{F}^{N \times 1}$ and scalars

$$d_{1,0}, d_{1,1}, \dots, d_{1,\delta_1 - 1}, \dots, d_{m,0}, d_{m,1}, \dots, d_{m,\delta_m - 1}, \dots, e_1, e_2, \dots, e_d \in \mathsf{F}$$

such that

$$\widetilde{M}_{A,m,i,v_1,v_2,\dots,v_m}(x_1, x_2, \dots, x_d) \begin{bmatrix} d_{1,0} \\ d_{1,1} \\ \vdots \\ d_{m,\delta_m - 1} \\ e_1 \\ e_2 \\ \vdots \\ e_d \end{bmatrix} = 0.$$

One can see by a comparison of Definitions 3.4 and 3.7 that

$$\widetilde{D}_{A,m,i}(x_1, x_2, \dots, x_d) = \sum_{e_1, e_2, \dots, e_d \in \mathsf{F}} \overline{D}_{A,m,i}(-e_1 x_1 - e_2 x_2 - \dots - e_d x_d).$$

Consequently the next result follows by Lemma 3.5 and the fact that a subspace of $\mathsf{F}^{N \times 1}$ with dimension d includes $q^d - 1$ nonzero vectors along the zero vector.

Lemma 3.8. *Suppose that the vectors $x_1, x_2, \dots, x_d \in \mathsf{F}^{N \times 1}$ are linearly independent. Then*

$$\widetilde{D}_{A,m,i}(x_1, x_2, \dots, x_d) \leq q^d D_{A,m,i} - \left(q^d - 1 \right) q^{Nm}.$$

The next result now follows by an application of Lemma 2.20.

Lemma 3.9. *If A is as given in Equations (2.28a)–(2.28j), m, i and d are positive integers such that $i \geq m > \ell$, and $\{x_1, x_2, \dots, x_d\}$ is a set of linearly independent vectors in $\mathsf{F}^{N \times 1}$, then*

$$\widetilde{D}_{A,m,i}(x_1, x_2, \dots, x_d) \leq \begin{cases} q^{Nm} + 6q^{N(m-1)+i+d} \log_q N & \text{if } m = \ell + 1, \\ q^{Nm} + 4q^{N(m-1)+i+d} & \text{if } m = \ell + 2, \\ q^{Nm} + q^{N(m-1)+i+d}(1 + 2q^{\ell - m + 1}) & \text{if } m \geq \ell + 3. \end{cases}$$

Once again, it will be useful to generalize Definition 2.3.

Definition 3.10. Suppose again that m, i, and d are positive integers, that j is a nonnegative integer, and that

$$x_1, x_2, \ldots, x_d \in \mathsf{F}^{N \times 1}.$$

Let $\widetilde{\rho}_{A,m,i}(x_1, x_2, \ldots, x_d; j)$ be the probability that the matrix $\widetilde{M}_{A,m,i,v_1,v_2,\ldots,v_m}(x_1, x_2, \ldots, x_d)$ is rank deficient by j, that is,

$$\widetilde{\rho}_{A,m,i}(x_1, x_2, \ldots, x_d; j) = \mathsf{Prob}\left[\operatorname{rank}\left(\widetilde{M}_{A,m,i,v_1,v_2,\ldots,v_m}(x_1, x_2, \ldots, x_d)\right) = i + d - j\right] \quad \text{if } i + d \leq N \tag{3.15}$$

and

$$\widetilde{\rho}_{A,m,i}(x_1, x_2, \ldots, x_d; j) = \mathsf{Prob}\left[\operatorname{rank}\left(\widetilde{M}_{A,i,k,v_1,v_2,\ldots,v_m}(x_1, x_2, \ldots, x_d)\right) = N - j\right] \quad \text{if } i + d > N, \tag{3.16}$$

and let $\widetilde{\sigma}_{A,m,i}(x_1, x_2, \ldots, x_d; j)$ be the probability that the matrix is rank deficient by at least j, that is,

$$\widetilde{\sigma}_{A,m,i}(x_1, x_2, \ldots, x_d; j) = \mathsf{Prob}\left[\operatorname{rank}\left(\widetilde{M}_{A,m,i,v_1,v_2,\ldots,v_m}(x_1, x_2, \ldots, x_d)\right) \leq i + d - j\right] \quad \text{if } i + d \leq N \tag{3.17}$$

and

$$\widetilde{\sigma}_{A,m,i}(x_1, x_2, \ldots, x_d; j) = \mathsf{Prob}\left[\operatorname{rank}\left(\widetilde{M}_{A,m,i,v_1,v_2,\ldots,v_m}(x_1, x_2, \ldots, x_d)\right) \leq N - j\right] \quad \text{if } i + d > N, \tag{3.18}$$

when the vectors $v_1, v_2, \ldots, v_m$ are chosen uniformly and independently from $\mathsf{F}^{N \times 1}$.

Lemmas 2.4 and 2.5 can be applied, once again, to relate the values that have now been defined: It follows by these lemmas that

$$\widetilde{D}_{A,m,i}(x_1, x_2, \ldots, x_d) = \sum_{j=0}^{i+d} q^{Nm+j}\, \widetilde{\rho}_{A,m,i}(x_1, x_2, \ldots, x_d; j) \quad \text{if } i + d \leq N \tag{3.19}$$

and

$$\widetilde{D}_{A,m,i}(x_1, x_2, \ldots, x_d) = \sum_{j=0}^{N} q^{N(m-1)+i+j+d}\, \widetilde{\rho}_{A,m,i}(x_1, x_2, \ldots, x_d; j) \quad \text{if } i + d > N. \tag{3.20}$$

Since $\widetilde{\rho}_{A,m,i}(x_1, x_2, \ldots, x_d; j) = \widetilde{\sigma}_{A,m,i}(x_1, x_2, \ldots, x_d; j) - \widetilde{\sigma}_{A,m,i}(x_1, x_2, \ldots, x_d; j+1)$ if $j < \min(i+d, N)$, $\widetilde{\rho}_{A,m,i}(x_1, x_2, \ldots, x_d; j) = \widetilde{\sigma}_{A,m,i}(x_1, x_2, \ldots, x_d; j)$ if $j = \min(i+d, N)$, and $\widetilde{\sigma}_{A,m,i}(x_1, x_2, \ldots, x_d; 0) = 1$, this implies that

$$\widetilde{D}_{A,m,i}(x_1, x_2, \ldots, x_d) = q^{Nm} + (q-1) \sum_{j=1}^{i+d} q^{Nm+j-1}\, \widetilde{\sigma}_{A,m,i}(x_1, x_2, \ldots, x_d; j) \quad \text{if } i + d \leq N, \tag{3.21}$$

and

$$\widetilde{D}_{A,m,i}(x_1, x_2, \ldots, x_d) = q^{N(m-1)+i+d} + \sum_{j=1}^{N} q^{N(m-1)+i+j+d-1}\, \widetilde{\sigma}_{A,m,i}(x_1, x_2, \ldots, x_d; j) \quad \text{if } i + d > N. \tag{3.22}$$

Lemma 3.9 can now be applied to obtain the following.

Lemma 3.11. *Suppose that A is as given in Equations (2.28a)–(2.28j), m, i and d are positive integers such that $i \geq m > \ell$, and that $\{x_1, x_2, \dots, x_d\}$ is a set of linearly independent vectors in $\mathsf{F}^{N\times 1}$. If $i + d \leq N$ and $1 \leq j \leq i + d$ then*

$$\widetilde{\sigma}_{A,m,i}(x_1, x_2, \dots, x_d; j) \leq \begin{cases} \frac{6 \log_q N}{(q-1)\, q^{N-(i+d)+j-1}} & \text{if } m = \ell + 1, \\ \frac{4}{(q-1)\, q^{N-(i+d)+j-1}} & \text{if } m = \ell + 2, \\ \frac{1+2q^{\ell-m+1}}{(q-1)\, q^{N-(i+d)+j-1}} & \text{if } m \geq \ell + 3. \end{cases}$$

and if $i + d > N$ and $1 \leq j \leq N$ then

$$\widetilde{\sigma}_{A,m,i}(x_1, x_2, \dots, x_d; j) \leq \begin{cases} \frac{1}{(q-1)q^{i+j+d-N-1}} + \frac{6 \log_q N}{(q-1)\, q^{j-1}} & \text{if } m = \ell + 1, \\ \frac{1}{(q-1)q^{i+j+d-N-1}} + \frac{3}{(q-1)\, q^{j-1}} & \text{if } m = \ell + 2, \\ \frac{1}{(q-1)\, q^{i+j+d-N-1}} + \frac{2q^{\ell-m+1}}{(q-1)\, q^{j-1}} & \text{if } m \geq \ell + 3. \end{cases}$$

4 Application: The Minors of a Block Hankel Matrix

4.1 A Block Hankel Matrix

Suppose once again that $A \in \mathsf{F}^{N\times N}$ is a matrix with rank r.

Definition 4.1. Suppose that m_L, m_R, i, and j are positive integers, and that

$$u_1, u_2, \dots, u_{m_L}, v_1, v_2, \dots, v_{m_R} \in \mathsf{F}^{N\times 1}.$$

Let

$$H_{A,m_L,m_R,i,j}(u_1, u_2, \dots, u_{m_L}, v_1, v_2, \dots, v_{m_R}) = M^t_{A^t,m_L,i,u_1,u_2,\dots,u_{m_L}} \cdot A \cdot M_{A,m,j,v_1,v_2,\dots,v_{m_R}} \in \mathsf{F}^{i\times j},$$

where the matrices $M_{A^t,m_L,i,u_1,u_2,\dots,u_{m_L}} \in \mathsf{F}^{N\times i}$ and $M_{A,m_R,j,v_1,v_2,\dots,v_{m_R}} \in \mathsf{F}^{N\times j}$ are as given in Definition 2.1 on page 3.

We will use the results of Sections 2 and 3 to bound the probability that this matrix is rank deficient when the vectors

$$u_1, u_2, \dots, u_{m_L}, v_1, v_2, \dots, v_{m_R}$$

are chosen uniformly and independently from $\mathsf{F}^{N\times 1}$ and when A has at most $\min(m_L, m_R) - 1$ nontrivial invariant factors.

4.2 Simplification

We may assume that A has no nontrivial invariant factors — that is, each invariant factor (in $\mathsf{F}[z]$) is either divisible by z^2 or not divisible by z. In other words, we may assume that each nilpotent block of (a rational Jordan form for) A has order at least two.

To see that this is the case, we may apply the argument used in Section 3.1.3: Note that, in general,

$$A = X^{-1} \begin{bmatrix} \widehat{A} & 0 \\ 0 & 0 \end{bmatrix} X \tag{4.1}$$

where X is a nonsingular matrix in $\mathsf{F}^{N\times N}$ and $\widehat{A} \in \mathsf{F}^{\widehat{N}\times\widehat{N}}$, for an integer $\widehat{N}$ such that $0 \le \widehat{N} \le N$, and where $\widehat{A}$ has no nontrivial invariant factors: In this case, (the rational Jordan form for) A has exactly $N - \widehat{N}$ nilpotent blocks with size one.

Suppose again that $u_1, u_2, \dots, u_{m_L}, v_1, v_2, \dots, v_{m_R} \in \mathsf{F}^{N\times 1}$. Since the matrices X and X^t are nonsingular, there exist vectors $\widehat{u}_1, \widehat{u}_2, \dots, \widehat{u}_{m_L} \in \mathsf{F}^{\widehat{N}\times 1}$ and $\widetilde{u}_1, \widetilde{u}_2, \dots, \widetilde{u}_{m_L} \in \mathsf{F}^{(N-\widehat{N})\times 1}$ such that

$$u_i = X^t \begin{bmatrix} \widehat{u}_i \\ \widetilde{u}_i \end{bmatrix} \qquad \text{for } 1 \le i \le m_L \tag{4.2}$$

and there exist vectors $\widehat{v}_1, \widehat{v}_2, \dots, \widehat{v}_{m_R} \in \mathsf{F}^{\widehat{N}\times 1}$ and there exist vectors $\widetilde{v}_1, \widetilde{v}_2, \dots, \widetilde{v}_{m_R} \in \mathsf{F}^{(N-\widehat{N})\times 1}$ such that

$$v_j = X^{-1} \begin{bmatrix} \widehat{v}_j \\ \widetilde{v}_j \end{bmatrix} \qquad \text{for } 1 \le j \le m_R. \tag{4.3}$$

It follows by Equations (4.1), (4.2), and (4.3) that if k is a positive integer then

$$u_i^t A^k v_j = \begin{bmatrix} \widehat{u}_i \\ \widetilde{u}_i \end{bmatrix}^t \cdot X \cdot X^{-1} \cdot \begin{bmatrix} \widehat{A}^k & 0 \\ 0 & 0 \end{bmatrix} \cdot X \cdot X^{-1} \cdot \begin{bmatrix} \widehat{v}_j \\ \widetilde{v}_j \end{bmatrix} = \widehat{u}_i^t \widehat{A}^k \widehat{v}_j.$$

Consequently

$$\begin{aligned} H_{A,m_L,m_R,i,j}(u_1, u_2, \dots, u_m, v_1, v_2, \dots, v_m) &= M^t_{A^t,m_L,i,u_1,u_2,\dots,u_{m_L}} \cdot A \cdot M_{A,m_R,j,v_1,v_2,\dots,v_{m_R}} \\ &= M^t_{\widehat{A}^t,m_L,i,\widehat{u}_1,\widehat{u}_2,\dots,\widehat{u}_{m_L}} \cdot \widehat{A} \cdot M_{\widehat{A},m_R,j,\widehat{v}_1,\widehat{v}_2,\dots,\widehat{v}_{m_R}} \\ &= H_{\widehat{A},m_L,m_R,i,j}(\widehat{u}_1, \widehat{u}_2, \dots, \widehat{u}_{m_L}, \widehat{v}_1, \widehat{v}_2, \dots, \widehat{v}_{m_R}). \end{aligned}$$

The claim that we may assume A has no nontrivial factors now follows by the choice of $\widehat{A}$ and the observation that (since X and X^t are nonsingular), if the vectors $u_1, u_2, \dots, u_{m_L}, v_1, v_2, \dots, v_{m_R}$ are chosen uniformly and independently from $\mathsf{F}^{N\times 1}$ then the corresponding vectors $\widehat{u}_1, \widehat{u}_2, \dots, \widehat{u}_{m_L}, \widehat{v}_1, \widehat{v}_2, \dots, \widehat{v}_{m_R}$ are chosen uniformly and independently from $\mathsf{F}^{\widehat{N}\times 1}$.

4.3 Bounding the Rank: A Useful Lemma

The rank of $H_{A,m_L,m_R,i,j}(u_1, u_2, \dots, u_{m_L}, v_1, v_2, \dots, v_{m_R})$ is the same as the rank of its transpose,

$$H_{A,m_L,m_R,i,j}(u_1, u_2, \dots, u_{m_L}, v_1, v_2, \dots, v_{m_R})^t = M^t_{A,m_R,j,v_1,v_2,\dots,v_{m_R}} \cdot A^t \cdot M_{A^t,m_L,i,u_1,v_2,\dots,v_{m_L}}.$$

It will be helpful to consider the latter matrix.

Suppose that the matrix $(A \cdot M_{A,m_R,j,v_1,v_2,\dots,v_{m_R}})^t = M^t_{A,m_R,j,v_1,v_2,\dots,v_{m_R}} \cdot A^t \in \mathsf{F}^{N\times j}$ has rank t. Then $t \le \min(r, j) \le \min(N, j)$ since the row space of this matrix is a subset of the row space of A^t. The set of vectors $w \in \mathsf{F}^{N\times 1}$ such that

$$(A \cdot M_{A,m_R,j,v_1,v_2,\dots,v_{m_R}})^t w = 0$$

is a subspace W of $\mathsf{F}^{N\times 1}$ with dimension $N - t$; let

$$w_1, w_2, \dots, w_{N-t} \in \mathsf{F}^{N\times 1}$$

be a basis for this subspace.

Notice that the vectors $w_1, w_2, \ldots, w_{N-t}$ depend on $v_1, v_2, \ldots, v_{m_R}$ but that they do depend in any way on the vectors $u_1, u_2, \ldots, u_{m_L}$. With that noted, let us consider the matrix

$$\widetilde{M}_{A^t,m_L,i,u_1,u_2,\ldots,u_{m_L}}(w_1, w_2, \ldots, w_{N-t}) \in \mathsf{F}^{N\times(N+i-t)}$$

as given in Definition 3.6 on page 34.

Let

$$S = \{x \in \mathsf{F}^{i\times 1} \mid H_{A,m_L,m_R,i,j}(u_1, u_2, \ldots, u_{m_L}, v_1, v_2, \ldots, v_{m_R})^t \cdot x = 0\} \subseteq \mathsf{F}^{i\times 1}$$

and let

$$T = \{y \in \mathsf{F}^{(N+i-t)\times 1} \mid \widetilde{M}_{A^t,m_L,i,u_1,u_2,\ldots,u_{m_L}}(w_1, w_2, \ldots, w_{N-t}) \cdot y = 0\} \subseteq \mathsf{F}^{(N+i-t)\times 1}.$$

Lemma 4.2. $|S| = |T|$.

Proof. It is sufficient to exhibit a bijection $\phi : S \to T$.

To do so, let $x \in S$. Then $x \in \mathsf{F}^{i\times 1}$ and $H_{A,m_L,m_R,i,j}(u_1, u_2, \ldots, u_{m_L}, v_1, v_2, \ldots, v_{m_R})^t \cdot x = 0$. Since

$$H_{A,m_L,m_R,i,j}(u_1, u_2, \ldots, u_{m_L}, v_1, v_2, \ldots, v_{m_R})^t = M^t_{A,m_R,j,v_1,v_2,\ldots,v_{m_R}} \cdot A^t \cdot M_{A^t,m_L,i,u_1,u_2,\ldots,u_{m_L}},$$

it follows that if $w = M_{A^t,m_L,i,u_1,u_2,\ldots,u_{m_L}} \cdot x$ then

$$\begin{aligned}(A \cdot M_{A,m_R,j,v_1,v_2,\ldots,v_{m_R}})^t w &= M^t_{A,m_R,j,v_1,v_2,\ldots,v_{m_R}} \cdot A^t \cdot w \\ &= H_{A,m_L,m_R,i,j}(u_1, u_2, \ldots, u_{m_L}, v_1, v_2, \ldots, v_{m_R})^t \cdot x \\ &= 0.\end{aligned}$$

Thus $w \in W$. Since $w_1, w_2, \ldots, w_{N-t}$ is a basis for W there exists a unique sequence of elements $c_1, c_2, \ldots, c_{N-t}$ of F such that

$$w = c_1w_1 + c_2w_2 + \cdots + c_{N-t}w_{N-t}.$$

Let us define

$$\phi(x) = \begin{bmatrix} x \\ -c_1 \\ -c_2 \\ \vdots \\ -c_{N-t} \end{bmatrix} \in \mathsf{F}^{(N+i-t)\times 1}. \tag{4.4}$$

This is well defined since the elements $c_1, c_2, \ldots, c_{N-t}$ are uniquely determined from x.

Notice that

$$\begin{aligned}&\widetilde{M}_{A^t,m_L,i,u_1,u_2,\ldots,u_{m_L}}(w_1, w_2, \ldots, w_{N-t}) \cdot \phi(x) \\ &\quad = M_{A^t,m_L,i,u_1,u_2,\ldots,u_{m_L}} \cdot x - (c_1w_1 + c_2x_2 + \cdots + c_{N-t}w_{N-t}) \\ &\qquad\qquad \text{(by the definitions of } \widetilde{M}_{A^t,m_L,i,u_1,u_2,\ldots,u_{m_L}}(w_1, w_2, \ldots, w_{N-t}) \text{ and } \phi(x)\text{)} \\ &\quad = w - w = 0,\end{aligned}$$

so that $\phi(x) \in T$ for all $x \in S$. It is also clear from Equation (4.4) that that ϕ is an injective map. All that remains is to show that it is surjective as well.

Let $y \in T$. Then we may write

$$y = \begin{bmatrix} \widehat{y} \\ \widetilde{y} \end{bmatrix}$$

for vectors $\widehat{y} \in \mathsf{F}^{i \times 1}$ and $\widetilde{y} \in \mathsf{F}^{N-t \times 1}$.

Let

$$M_w = \begin{bmatrix} w_1 & w_2 & \dots & w_{N-t} \end{bmatrix} \in \mathsf{F}^{N \times (N-t)}.$$

Since $y \in T$,

$$\widetilde{M}_{A^t, m_L, i, u_1, u_2, \dots, u_{m_L}}(w_1, w_2, \dots, w_{N-t}) \cdot y = M_{A^t, m_L, i, u_1, u_2, \dots, u_{m_L}} \cdot \widehat{y} + M_w \widetilde{y} = 0,$$

so that

$$M_{A^t, m_L, i, u_1, u_2, \dots, u_{m_L}} \cdot \widehat{y} = -M_w \widetilde{y},$$

and

$$\begin{aligned} &H_{A, m_L, m_R, i, j}(u_1, u_2, \dots, u_{m_L}, v_1, v_2, \dots, v_{m_R})^t \cdot \widehat{y} \\ &\quad = (A \cdot M_{A, m_R, j, v_1, v_2, \dots, v_{m_R}})^t \cdot M_{A^t, m_L, i, u_1, u_2, \dots, u_{m_L}} \cdot \widehat{y} \\ &\quad = -(A \cdot M_{A, m_R, j, v_1, v_2, \dots, v_{m_R}})^t M_w \widetilde{y} \\ &\quad 0 \cdot \widetilde{y} = 0, \end{aligned}$$

since the columns of M_w all belong to the subspace W, so that $(A \cdot M_{A, m_R, j, v_1, v_2, \dots, v_{m_R}})^t \cdot M_w = 0$. It follows that $\widehat{y} \in S$, so that $\phi(\widehat{y}) \in T$. Now notice that

$$y = \begin{bmatrix} \widehat{y} \\ \widetilde{y} \end{bmatrix} \in T \qquad \text{and} \qquad \varphi(\widehat{y}) = \begin{bmatrix} \widehat{y} \\ \overline{y} \end{bmatrix} \in T$$

as well, for vectors $\widetilde{y}, \overline{y} \in \mathsf{F}^{(N-t) \times 1}$. The matrix

$$y - \varphi(\widehat{y}) = \begin{bmatrix} 0 \\ \widetilde{y} - \overline{y} \end{bmatrix} \tag{4.5}$$

belongs to the subspace T as well. Consequently

$$\widetilde{M}_{A^t, m_L, i, u_1, u_2, \dots, u_{m_L}}(w_1, w_2, \dots, w_{N-t}) \cdot (y - \varphi(\widehat{y}) = 0. \tag{4.6}$$

It follows by Equations (4.5) and (4.6) that

$$M_w \cdot (\widetilde{y} - \overline{y}) = 0$$

as well. Since the matrix M_w has full rank (its columns form a basis for the subspace W),

$$\widetilde{y} - \overline{y} = 0,$$

so that $y = \varphi(\widehat{y})$. Since y was an arbitrarily chosen element of T it follows that the map φ is subjective, as claimed. □

This result can be used along with Lemmas 2.4 and 2.5, to establish the following.

Corollary 4.3. *Let m_R be a positive integer, let $v_1, v_2, \ldots, v_{m_R} \in \mathsf{F}^{N\times 1}$, and let $w_1, w_2, \ldots, w_{N-t}$ be a basis for the set of vectors*

$$W = \{w \in \mathsf{F}^{N\times 1} \mid (A \cdot M_{A,m_R,j,v_1,v_2,\ldots,v_{m_R}})^t \cdot w = 0\} \subseteq \mathsf{F}^{N\times 1},$$

so that $t \le \min(r, j)$. Let m_L be a positive integer and let $u_1, u_2, \ldots, u_{m_L} \in \mathsf{F}^{N\times 1}$. Finally, let i be a positive integer, so that either $i \le t \le j$, $t < i \le j$, or $t \le j < i$.

(a) If $i \le j$ and s is an integer such that $0 \le s \le i$ then the matrix

$$H_{A,m_L,m_R,i,j}(u_1, u_2, \ldots, u_{m_L}, v_1, v_2, \ldots, v_{m_R}) \in \mathsf{F}^{i\times j}$$

has rank $i - s$ if and only if the matrix

$$\widetilde{M}_{A^t,m_L,i,u_1,u_2,\ldots,u_{m_L}}(w_1, w_2, \ldots, w_{N-t}) \in \mathsf{F}^{N\times(N+i-t)}$$

has rank $N + i - t - s$.

(b) If $j < i$ and s is an integer such that $0 \le s \le j$, then the matrix

$$H_{A,m_L,m_R,i,j}(u_1, u_2, \ldots, u_{m_L}, v_1, v_2, \ldots, v_{m_R}) \in \mathsf{F}^{i\times j}$$

has rank $j - s$ if and only if the matrix

$$\widetilde{M}_{A^t,m_L,i,u_1,u_2,\ldots,u_{m_L}}(w_1, w_2, \ldots, w_{N-t}) \in \mathsf{F}^{N\times(N+i-t)}$$

has rank $N + j - t - s$. It must therefore be the case that $s \ge j - t$.

5 Formulas

Once again, consider the matrix $H_{A,m_L,m_R,i,j}(u_1, u_2, \ldots, u_{m_L}, v_1, v_2, \ldots, v_{m_R}) \in \mathsf{F}^{i\times j}$.

Definition 5.1. Let h, i, and j be a positive integers and let $\tau_{A,m_L,m_R,i,j}(h)$ be the probability that the block Hankel matrix $H_{A,m_L,m_R,i,j}(u_1, u_2, \ldots, u_{m_L}, v_1, v_2, \ldots, v_{m_R})$ is rank-deficient by at least h, that is,

$$\begin{aligned}
&\tau_{A,m_L,m_R,i,j}(h) \\
&\quad = \begin{cases} \mathsf{Prob}\left[\operatorname{rank}(H_{A,m_L,m_R,i,j}(u_1, u_2, \ldots, u_{m_L}, v_1, v_2, \ldots, v_{m_R})) \le i - h\right] & \text{if } i \le j \text{ and } i \le r, \\ \mathsf{Prob}\left[\operatorname{rank}(H_{A,m_L,m_R,i,j}(u_1, u_2, \ldots, u_{m_L}, v_1, v_2, \ldots, v_{m_R})) \le j - h\right] & \text{if } i > j \text{ and } j \le r, \\ \mathsf{Prob}\left[\operatorname{rank}(H_{A,m_L,m_R,i,j}(u_1, u_2, \ldots, u_{m_L}, v_1, v_2, \ldots, v_{m_R})) \le r - h\right] & \text{if } i, j > r, \end{cases}
\end{aligned}$$

when the vectors $u_1, u_2, \ldots, u_{m_L}, v_1, v_2, \ldots, v_{m_R}$ are chosen uniformly and independently from $\mathsf{F}^{N\times 1}$.

Results from previous sections can now be used to bound these quantities.

Case: $\boldsymbol{i \le j \le r}$

The matrix $AM_{A,m_R,j,v_1,v_2,\dots,v_{m_R}}\mathsf{F}^{N\times j}$ is rank-deficient with probability at most $\widehat{\sigma}_{A,m_R,j}(1)$. Suppose, instead, that this matrix has full rank j.

In this case the subspace

$$W = \{w \in \mathsf{F}^{N\times 1} \mid (AM_{A,m_R,j,v_1,v_2,\dots,v_{m_R}})^t \cdot w = 0\}$$

has dimension $N - j$ and a basis $w_1, w_2, \dots, w_{N-j} \in \mathsf{F}^{N\times 1}$. In other words, $t = j$ for the value t considered in Corollary 4.3, implying the following.

Lemma 5.2. *If $i \le j < r$, and h is a positive integer, then*

$$\tau_{A,m_L,m_R,i,j}(h) \le \widehat{\sigma}_{A,m_R,j}(1) + \widetilde{\sigma}_{A^t,m_L,i}(w_1, w_2, \dots, w_{N-j}; h),$$

where $w_1, w_2, \dots, w_{N-j}$ is a sequence of linearly independent vectors in $\mathsf{F}^{N\times 1}$.

Since the matrices A and A^t have the same (nontrivial) invariant factors, Lemmas 3.3 and 3.11 can now be used to establish the following, assuming that A has ℓ nontrivial invariant factors, and $\ell < \min(m_L, m_R)$.

Corollary 5.3. *If $i \le j < r$, $\min(m_L, m_R) < \ell$ and h is a positive integer, then*

$$\tau_{A,m_L,m_R,i,j}(h) \le \begin{cases} \frac{6\log_q N}{(q-1)\,q^{r-j}} + \frac{6\log_q N}{(q-1)\,q^{j-i+h-1}} & \text{if } \min(m_L, m_R) = \ell + 1, \\ \frac{4}{(q-1)\,q^{r-j}} + \frac{4}{(q-1)\,q^{j-i+h-1}} & \text{if } \min(m_L, m_R) = \ell + 2, \\ \frac{1+2q^{\ell-m_R+1}}{(q-1)\,q^{r-j}} + \frac{1+2q^{\ell-m_L+1}}{(q-1)\,q^{j-i+h-1}} & \text{if } \min(m_L, m_R) \ge \ell + 3. \end{cases}$$

Sharper bounds can be obtained when $m_L \ne m_R$ by applying Lemmas 3.3 and 3.11 to consider additional cases.

Case: $\boldsymbol{i \le r < j}$

The argument used in the previous case can also be applied here: The matrix $AM_{A,m_R,j,v_1,v_2,\dots,v_{m_R}}$ is rank-deficient (that is, has rank less than r) with probability at most $\widehat{\sigma}_{A,m_R,j}(1)$. If this is not the case, then the above-mentioned subspace W has dimension $N - r$ and a basis $w_1, w_2, \dots, w_{N-r} \in \mathsf{F}^{N\times 1}$.

In other words, $t = r$ for the value t considered in Corollary 4.3 and it follows that the matrix $H_{A,m_L,m_R,i,j}(u_1, u_2, \dots, u_{m_L}, v_1, v_2, \dots, v_{m_R}) \in \mathsf{F}^{i\times 1}j$ has rank $i - h$ if and only if the matrix

$$\widetilde{M}_{A^t,m_L,i,u_1,u_2,\dots,u_{m_L}}(w_1, w_2, \dots, w_{N-r}) \in \mathsf{F}^{N\times(N+i-r)}$$

has rank $N + i - r - h$. This implies the following.

Lemma 5.4. *If $i \le r < j$ and h is a positive integer then*

$$\tau_{A,m_L,m_R,i,j}(h) \le \widehat{\sigma}_{A,m_R,j}(1) + \widetilde{\sigma}_{A^t,m_L,i}(w_1, w_2, \dots, w_{N-r}; h),$$

where $w_1, w_2, \dots, w_{N-r}$ is a sequence of linearly independent vectors in $\mathsf{F}^{N\times 1}$.

Lemmas 3.3 and 3.11 can be applied to obtain useful bounds for the case that m_R is significantly greater than the number of nontrivial invariant factors of A.

Corollary 5.5. *If $i \le r < j$, $m_R \ge \ell + 3$, $m_L \ge \ell + 1$, and h is a positive integer, then*

$$\tau_{A,m_L,m_R,i,j}(h) \le \begin{cases} \frac{1}{(q-1)q^{j-r}} + \frac{2q^{\ell-m_R+1}}{q-1} + \frac{6\log_q N}{(q-1)q^{r-i+h-1}} & \text{if } m_L = \ell+1, \\ \frac{1}{(q-1)q^{j-r}} + \frac{2q^{\ell-m_R+1}}{q-1} + \frac{4}{(q-1)q^{r-i+h-1}} & \text{if } m_L = \ell+2, \\ \frac{1}{(q-1)q^{j-r}} + \frac{2q^{\ell-m_R+1}}{q-1} + \frac{1+2q^{\ell-m_L+1}}{(q-1)q^{r-i+h-1}} & \text{if } m_L \ge \ell+3. \end{cases}$$

Unfortunately this does not provide useful bounds if $m_R = \ell + 1$ or if $m_R = \ell + 2$.

In order to deal with these cases, let c be a positive integer whose value will be specified later, and notice that the matrix $AM_{A,m_R,j,v_1,v_2,\dots,v_{m_R}} \in \mathsf{F}^{N\times j}$ has rank less than $r - c$ with probability at most $\widehat{\sigma}_{A,m_R,j}(c+1)$. Lemma 3.3 can be used to show that

$$\widehat{\sigma}_{A,m_R,j}(c+1) \le \frac{1}{(q-1)q^{j-r}(\log_q N)^2} + \frac{6}{(q-1)\log_q N}$$

if $m_R = \ell + 1$ and $c \ge 2\log_q \log_q N$, and that

$$\widehat{\sigma}_{A,m_R,j}(c+1) \le \frac{1}{(q-1)q^{j-r}\log_q N} + \frac{3}{(q-1)\log_q N}$$

if $m_R = \ell + 2$ and $c \ge \log_q \log_q N$.

Suppose instead that the rank of $AM_{A,m_R,j,v_1,v_2,\dots,v_{m_R}}$ is at least $r - c$, so that the subspace W of vectors annihilated by the transpose of this matrix has dimension at most $N - r + c$ and a basis $w_1, w_2, \dots, w_{N-r+\widehat{c}}$ for some integer $\widehat{c} \le c$.

Suppose now that $c \le r - i$ so that $\widehat{c} \le r - i$ as well. It follows by Corollary 4.3 (with $t = r - \widehat{c}$, so that $t - i \ge (r - i) - c \ge 0$) that the matrix $H_{A,m_L,m_R,i,j}(u_1, u_2, \dots, u_{m_L}, v_1, v_2, \dots, v_{m_R})$ has rank $i - h$, for a positive integer h, if and only if the matrix $\widetilde{M}_{A^t,m_L,i,u_1,u_2,\dots,u_{m_L}}(w_1, w_2, \dots, w_{N-r+\widehat{c}})$ has rank $N - r + i + \widehat{c} - h$. The latter event occurs with probability

$$\widetilde{\sigma}_{A^t,m_L,i}(w_1, w_2, \dots, w_{N-r+\widehat{c}}; h) \le \begin{cases} \frac{6\log_q N}{(q-1)q^{r-i-\widehat{c}+h-1}} \le \frac{6\log_q N}{(q-1)q^{r-i-c+h-1}} & \text{if } m_L = \ell+1, \\ \frac{4}{(q-1)q^{r-i-\widehat{c}+h-1}} \le \frac{4}{(q-1)q^{r-i-c+h-1}} & \text{if } m_L = \ell+2, \\ \frac{1+2q^{\ell-m_L+1}}{(q-1)q^{r-i-\widehat{c}+h-1}} \le \frac{1+2q^{\ell-m_L+1}}{(q-1)q^{r-i-c+h-1}} & \text{if } m_L \ge \ell+3. \end{cases}$$

These provide useful bounds if $i < r - 3\lceil \log_q \log_q N \rceil$ and $m_R = \ell + 1$, or if $i < r - 2\lceil \log_q \log_q N \rceil$ and $m_R = \ell + 2$:

Corollary 5.6. *If $m_R = \ell + 1$, $m_L \ge \ell + 1$, $i < r - 3\lceil \log_q \log_q N \rceil$, and h is a positive integer such that $1 \le h \le i$, then*

$$\tau_{A,m_L,m_R,i,j}(h) \le \begin{cases} \frac{1}{(q-1)q^{j-r}(\log_q N)^2} + \frac{6}{(q-1)\log_q N} + \frac{6}{(q-1)q^{r-i-3\lceil \log_q \log_q N\rceil+h-1}} & \text{if } m_L = \ell+1, \\ \frac{1}{(q-1)q^{j-r}(\log_q N)^2} + \frac{6}{(q-1)\log_q N} + \frac{4}{(q-1)q^{r-i-2\lceil \log_q \log_q N\rceil+h-1}} & \text{if } m_L = \ell+2 \\ \frac{1}{(q-1)q^{j-r}(\log_q N)^2} + \frac{6}{(q-1)\log_q N} + \frac{1+2q^{\ell-m_L+1}}{(q-1)q^{r-i-2\lceil \log_q \log_q N\rceil+h-1}} & \text{if } m_L \ge \ell+3. \end{cases}$$

If $m_R = \ell + 2$, $m_L \ge \ell + 1$, $i < r - 2\lceil \log_q \log_q N \rceil$, and h is a positive integer such that $1 \le h \le i$, then

$$\tau_{A,m_L,m_R,i,j}(h) \le \begin{cases} \frac{1}{q^{j-r}\log_q N} + \frac{3}{(q-1)\log_q N} + \frac{6}{(q-1)q^{r-i-2\lceil \log_q \log_q N\rceil+h-1}} & \text{if } m_L = \ell+1, \\ \frac{1}{q^{j-r}\log_q N} + \frac{3}{(q-1)\log_q N} + \frac{4}{(q-1)q^{r-i-\lceil \log_q \log_q N\rceil+h-1}} & \text{if } m_L = \ell+2, \\ \frac{1}{q^{j-r}\log_q N} + \frac{3}{(q-1)\log_q N} + \frac{1+2q^{\ell-m_L+1}}{(q-1)q^{r-i-\lceil \log_q \log_q N\rceil+h-1}} & \text{if } m_L \ge \ell+3. \end{cases}$$

Note that if $\widehat{\imath} \leq i$ then the rank of the matrix $H_{A,m_L,m_R,\widehat{\imath},j}(u_1, u_2, \dots, u_{m_L}, v_1, v_2, \dots, v_{m_R})$ is less than the rank of $H_{A,m_L,m_R,i,j}(u_1, u_2, \dots, u_{m_L}, v_1, v_2, \dots, v_{m_R})$. This implies the following.

Lemma 5.7. *If $m_R = \ell + 1$, $m_L \geq \ell + 1$, $r - 3\lceil \log_q \log_q N \rceil \leq i \leq r$, and h is a positive integer such that $3\lceil \log_q \log_q N \rceil < h \leq i$, then*

$$\tau_{A,m_L,m_R,i,j}(h) \leq \tau_{A,m_L,m_R,i-3\lceil \log_q \log_q N \rceil,j}(h - 3\lceil \log_q \log_q N \rceil).$$

Similarly, if $m_R = \ell + 2$, $m_L \geq \ell + 1$, $r - 2\lceil \log_q \log_q N \rceil \leq i \leq r$, and h is a positive integer such that $2\lceil \log_q \log_q N \rceil < h \leq i$, then

$$\tau_{A,m_L,m_R,i,j}(h) \leq \tau_{A,m_L,m_R,i-2\lceil \log_q \log_q N \rceil,j}(h - 2\lceil \log_q \log_q N \rceil).$$

This can be used along with Corollary 5.6 to bounds on $\tau_{A,m_L,m_R,i,j}(h)$ for the case that $r - i$ is small and positive, $j \geq r$, and for sufficiently large h.

Case: $r < i \leq j$

The argument used at the beginning of the previous cases can be used once again: The matrix $AM_{A,m_R,j,v_1,v_2,\dots,v_{m_R}}$ is rank-deficient (that is, has rank less than r) with probability at most $\widehat{\sigma}_{A,m_R,j}(1)$. If this is not the case then the subspace W has dimension $N - r$ and a basis $w_1, w_2, \dots, w_{N-r} \in \mathsf{F}^{N \times 1}$. Once again $t = r$, for the value t considered in Corollary 4.3.

Let h be an integer such that $1 \leq h \leq r$; then the matrix $H_{A,m_L,m_R,i,j}(u_1, u_2, \dots, u_{m_L}, v_1, v_2, \dots, v_{m_R})$ has rank $r - h$ if and only if its rank is $i - s$, where $s = i - r + h$. It follows by the corollary that this is the case if and only if the matrix $\widetilde{M}_{A^t,m_L,i,u_1,u_2,\dots,u_{m_L}}(w_1, w_2, \dots, w_{N-r})$ has rank $N - (s - i + t) = N - r + i - s = N - h$.

Lemma 5.8. *If $r < i \leq j$ and h is a positive integer such that $1 \leq h \leq r$ then*

$$\tau_{A,m_L,m_R,i,j}(h) \leq \widehat{\sigma}_{A,m_R,j}(1) + \widetilde{\sigma}_{A^t,m_L,i}(w_1, w_2, \dots, w_{N-r}; h)$$

where $w_1, w_2, \dots, w_{N-r}$ is a sequence of linearly independent vectors in $\mathsf{F}^{N \times 1}$.

Lemmas 3.3 and 3.11 can be applied once again to produce useful bounds for the case that $m_R \geq \ell + 3$.

Corollary 5.9. *If $r < i \leq j$, $m_R \geq \ell + 3$, $m_L \geq \ell + 1$, and h is a positive integer, then*

$$\tau_{A,m_L,m_R,i,j}(h) \leq \begin{cases} \frac{1}{(q-1)q^{j-r}} + \frac{2q^{\ell - m_R + 1}}{q-1} + \frac{1}{(q-1)q^{i-r+h-1}} + \frac{6 \log_q N}{(q-1)q^{h-1}} & \text{if } m_L = \ell + 1, \\ \frac{1}{(q-1)q^{j-r}} + \frac{2q^{\ell - m_R + 1}}{q-1} + \frac{1}{(q-1)q^{i-r+h-1}} + \frac{3}{(q-1)q^{h-1}} & \text{if } m_L = \ell + 2, \\ \frac{1}{(q-1)q^{j-r}} + \frac{2q^{\ell - m_R + 1}}{q-1} + \frac{1}{(q-1)q^{i-r+h-1}} + \frac{2q^{\ell - m_L + 1}}{(q-1)q^{h-1}} & \text{if } m_L \geq \ell + 3. \end{cases}$$

If $m_R = \ell + 1$ or $m_R = \ell + 2$ then useful bounds can be obtained by choosing a suitable constant c, as before, and considering the case that the matrix $AM_{A,m_R,j,v_1,v_2,\dots,v_{m_R}}$ is rank-deficient by more than c; this occurs with probability $\widehat{\sigma}_{A,m_R,j}(c + 1)$.

Otherwise the rank of the above matrix is $r - \widehat{c}$ for some integer $\widehat{c} \leq c$, and the subspace W has dimension $N - r + \widehat{c}$ and a basis $w_1, w_2, \dots, w_{N-r+\widehat{c}}$. Now $t = r - \widehat{c}$, for the value t mentioned in Corollary 4.3.

Let h be an integer such that $1 \leq h \leq r$; then the matrix $H_{A,m_L,m_R,i,j}(u_1, u_2, \dots, u_{m_L}, v_1, v_2, \dots, v_{m_R})$ has rank $r - h$ if and only if its rank is $i - s$ for $s = i - r + h$. It follows by the corollary that this is the case if and only if the matrix $\widetilde{M}_{A^t,m_L,i,u_1,u_2,\dots,u_{m_L}}(w_1, w_2, \dots, w_{N-r+\widehat{c}})$ has rank $N - (s - i + t) = N - (h - \widehat{c})$. This can be used to establish the following.

Corollary 5.10. *If $m_R = \ell + 1$, $m_L \geq \ell + 1$, $r < i \leq j$, and h is a positive integer such that $3\lceil \log_q \log_q N \rceil < h \leq N$, then*

$$\tau_{A,m_L,m_R,i,j}(h) \leq \begin{cases} \frac{1}{(q-1)q^{j-r}(\log_q N)^2} + \frac{6}{(q-1)\log_q N} + \frac{1}{(q-1)q^{i-r+h-1}} + \frac{6}{(q-1)q^{h-3\lceil \log_q \log_q N \rceil}} & \text{if } m_L = \ell + 1, \\ \frac{1}{(q-1)q^{j-r}(\log_q N)^2} + \frac{6}{(q-1)\log_q N} + \frac{1}{(q-1)q^{i-r+h-1}} + \frac{3}{(q-1)q^{h-2\lceil \log_q \log_q N \rceil}} & \text{if } m_L = \ell + 2, \\ \frac{1}{(q-1)q^{j-r}(\log_q N)^2} + \frac{6}{(q-1)\log_q N} + \frac{1}{(q-1)q^{i-r+h-1}} + \frac{2q^{\ell - m_L + 1}}{(q-1)q^{h-2\lceil \log_q \log_q N \rceil}} & \text{if } m_L \geq \ell + 3. \end{cases}$$

If $m_R = \ell + 2$, $m_L \geq \ell + 1$, $r < i \leq j$, and h is a positive integer such that $2\lceil \log_q \log_q N \rceil < h \leq N$, then

$$\tau_{A,m_L,m_R,i,j}(h) \leq \begin{cases} \frac{1}{(q-1)q^{j-r}\log_q N} + \frac{3}{(q-1)\log_q N} + \frac{1}{(q-1)q^{i-r+h-1}} + \frac{6}{(q-1)q^{h-2\lceil \log_q \log_q N \rceil}} & \text{if } m_L = \ell + 1, \\ \frac{1}{(q-1)q^{j-r}\log_q N} + \frac{3}{(q-1)\log_q N} + \frac{1}{(q-1)q^{i-r+h-1}} + \frac{3}{(q-1)q^{h-\lceil \log_q \log_q N \rceil}} & \text{if } m_L = \ell + 2, \\ \frac{1}{(q-1)q^{j-r}\log_q N} + \frac{3}{(q-1)\log_q N} + \frac{1}{(q-1)q^{i-r+h-1}} + \frac{2q^{\ell - m_L + 1}}{(q-1)q^{h-\lceil \log_q \log_q N \rceil}} & \text{if } m_L \geq \ell + 3. \end{cases}$$

Case: $i > j$

Notice that

$$H_{A,m_L,m_R,i,j}(u_1, u_2, \ldots, u_{m_L}, v_1, v_2, \ldots, v_{m_R})^t = H_{A^t,m_R,m_L,j,i}(v_1, v_2, \ldots, v_{m_R}, u_1, u_2, \ldots, u_{m_L}).$$

Clearly, the vectors $v_1, v_2, \ldots, v_{m_R}, u_1, u_2, \ldots, u_{m_L}$ are chosen uniformly and independently from $\mathsf{F}^{N \times 1}$ if and only if the vectors $u_1, u_2, \ldots, u_{m_L}, v_1, v_2, \ldots, v_{m_R}$ are. This implies the following.

Lemma 5.11. *If m_L, m_R, i and j, and h are positive integers then*

$$\tau_{A,m_L,m_R,i,j}(h) = \tau_{A^t,m_R,m_L,j,i}(h).$$

Since A and A^t have the same number of nontrivial invariant factors, bounds on $\tau_{A,m_L,m_R,i,j}(h)$ for the case $i > j$ can be obtained by exchanging i and j, while simultaneously exchanging m_L and m_R, and choosing whichever of the bounds from the previous sections that is applicable.

6 Future Work

Much of the analysis of block Wiedemann and block Lanczos algorithms depends on the assumption that the blocking factor ($min(m_L, m_R)$ for m_L and m_R as used above) is greater than the number of nontrivial invariant factors of coefficient matrix A (the value given here as ℓ). On the other hand, it seems necessary to assume that the number of nontrivial nilpotent blocks (the value ℓ_0 in this report) exceeds the blocking factor as well when describing instances causing block algorithms to fail. Thus there is further work to be done to explain the behaviour of block algorithms when the blocking factor falls between the number of nontrivial nilpotent blocks and the number of nontrivial invariant factors.

All of the block algorithms that have been successfully analyzed are "biconditional:" Blocks of vectors

$$u_1, u_2, \ldots, u_{m_L}$$

and

$$v_1, v_2, \ldots, v_{m_R}$$

are chosen uniformly and independently, and the independence of these blocks is required in order to complete an analysis. On the other hand, the block Lanczos algorithms of Coppersmith [2] and Montgomery [12] are developed to be used with symmetric matrices and use the same blocks on each side: $m_L = m_R$ and $u_i = v_i$ for $1 \leq i \leq m_L$. It is clear that symmetrization of the input is unreliable in the small field case, since the matrices $A^t A$ and A can have significantly different ranks and null spaces. However, it would be interesting to understand the behaviour of algorithms that resemble Coppersmith's and Montgomery's more closely when the input matrix is symmetric.

References

[1] R. P. Brent, S. Gau, and A. G. B. Lauder. Random Krylov spaces over finite fields. *SIAM Journal on Discrete Mathematics*, 16:276–287, 2003.

[2] D. Coppersmith. Solving linear equations over GF(2): Block Lanczos algorithm. *Linear Algebra and Its Applications*, 192:33–60, 1993.

[3] D. Coppersmith. Solving homogeneous linear equations over GF(2) via block Wiedemann algorithm. *Mathematics of Computation*, 62:333–350, 1994.

[4] J.-G. Dumas, T. Gautier, M. Giesbrecht, P. Giorgi, B. Hovinen, E. Kaltofen, B. D. Saunders, W. J. Turner, and G. Villard. Linbox: A generic library for exact linear algebra. In *Proceedings, First International Congress of Mathematical Software*, pages 40–50, Beijing, China, 2002.

[5] W. Eberly. Early termination over small fields. In *Proceedings, 2003 International Symposium on Symbolic and Algebraic Computation*, pages 80–87, Philadelphia, Pennsylvania, 2003. ACM Press.

[6] B. Hovinen. Block Lanczos-style algorithms over small finite fields. Master's thesis, University of Waterloo, Waterloo, Ontario, 2004.

[7] B. Hovinen and W. Eberly. A reliable block Lanczos algorithm over small finite fields. In *Proceedings, 2005 International Symposium on Symbolic and Algebraic Computation (ISSAC '05)*, Beijing, 2005. ACM Press.

[8] E. Kaltofen. Analysis of Coppersmith's block Wiedemann algorithm for the parallel solution of sparse linear systems. *Mathematics of Computation*, 64:777–806, 1995.

[9] C. Lanczos. An iteration method for the solution of the eigenvalue problem of linear differential and integral operators. *Journal of Research. National Bureau of Standards*, 45:255–282, 1950.

[10] R. Lidl and H. Niederreiter. *Finite Fields*, volume 20 of *Encyclopedia of Mathematics and Its Applications*. Cambridge University Press, second edition, 1996.

[11] A. Lobo. *Matrix-Free Linear System Solving and Applications to Symbolic Computation.* PhD thesis, Rensselaer Polytechnic Institute, Troy, New York, 1995.

[12] P. Montgomery. A block Lanczos algorithm for finding dependencies over GF(2). In *Advances in Cryptology — EUROCRYPT '95*, volume 921 of *Lecture Notes in Computer Science*, pages 106–120. Springer, 1995.

[13] G. Villard. Further analysis of Coppersmith's block Wiedemann algorithm for the solution of sparse linear systems (extended abstract). In *Proceedings, 1997 International Symposium on Symbolic and Algebraic Computation (ISSAC '97)*, pages 32–39, Maui, Hawaii, 1997. ACM Press.

[14] G. Villard. A study of Coppersmith's block Wiedemann algorithm using matrix polynomials. Technical Report 975, Institut d'Informatique et Mathématiques Appliquées de Grenoble (IMAG), April 1997.

[15] D. H. Wiedemann. Solving sparse linear equations over finite fields. *IEEE Transactions on Information Theory*, 32:54–62, 1986.