Sparse Matrix Computations over Small Fields: A Simpler Block Lanczos Algorithm and Its Analysis

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

ABSTRACT

A simplified "block Lanczos" algorithm is presented and its correctness established. While its efficiency in the general case is not proved, preconditioning used for similar algorithms is also sufficient here. Results concerning reliability and efficiency may be of more general interest because they may serve to (somewhat) better explain the performance of other block algorithms, including block Wiedemann algorithms and algorithms that use rectangular blocking.

Categories and Subject Descriptors

F.2.1 [**Analysis of Algorithms and Problem Complexity**]: Numerical Algorithms and Problems—*computations in finite fields, computations in matrices*; I.1.2 [**Symbolic and Algebraic Manipulation**]: Algorithms—*algebraic algorithms, analysis of algorithms*

General Terms

Algorithms, Performance, Reliability

Keywords

Block Lanczos algorithms, computations in finite fields, worst-case expected performance and reliability

1. INTRODUCTION

Block "Krylov-based" algorithms — including block Lanczos and block Wiedemann algorithms — have been used in sieve-based factorization algorithms and various other number-theoretic computations since Coppersmith's development of a block Lanczos algorithm [1].

Unfortunately the block Lanczos algorithms that were originally developed for these applications are provably unreliable in the worst case: They begin with a symmetrization of the input matrix that is provably correct for computations over the real numbers but that can significantly reduce the rank of the input matrix and, furthermore, fail to achieve the matrix conditions needed to ensure reliability (the beginning of the report [3] provides further details). Consequently subsequent work (including the development of all block Wiedemann algorithms, including the one proposed by Coppersmith [2]) has concerned biconditional algorithms which do not require the input matrix to be symmetric and double the number of vectors to be managed.

A biconditional Block Lanczos algorithm, whose efficiency and reliability could be proved for suitably conditioned input matrices over small finite fields, was first presented and analyzed by Bradford Hovinen in his Master's thesis [7]; a subsequent paper [8] summarizes the key results and is more readily available. A rectangular variant (which uses different block sizes on the left and right) has subsequently been proposed by the author of this report [3].

As noted above, one can prove that block Lanczos algorithms that symmetrize the input matrix are provably unreliable, in the worst case, for computations over finite fields. Nevertheless, a review of the development of these algorithms suggests improvements that might be made to the biconditional algorithms that have more recently been developed. In particular, Montgomery [10] proposed a considerable simplification of the management of vectors, in a block Lanczos algorithm, that has inspired the present work. An algorithm that simplifies the management of vectors in a biconditional block Lanczos computation in a similar way — and that, therefore, might be more easily implemented and maintained than its predecessors — is described in Section 2. Since block sizes can be selected for a variety of reasons, including efficient use of storage, it might also be of interest that all restrictions on the block sizes that can be used have been removed. The algorithm is provably correct for any block size $k \geq 2$ and (like Hovinen's) it is provably efficient as long as the block size exceeds the number of nontrivial invariant factors of the input matrix.

Now, unless these block Krylov matrices are used in significantly different (and presently unknown) ways than they currently are, some sort of conditioning of the input matrix is necessary if the computations are to be reliable in the worst case: The number of nontrivial invariant factors should be less than the block size if the algorithm is being used to determine the rank of the input matrix, in order to ensure that the Krylov space being generated is equal to the column (or row) space of the matrix. If one wishes to solve a linear system (or certify that it is inconsistent) or sample

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

uniformly from the null space, then the block size should exceed the number of invariant factors that are divisible by x^2.

On the other hand, some computations — including the solution of a linear system when the the minimal polynomial of the input matrix is not divisible by x^2, or the computation of a nonzero vector in the null space (making no guarantees about its distribution) — do not necessarily require any additional conditions on the input matrix, or the use of preconditioners, at all. Indeed, both of the above problems can be solved for arbitrary input matrices using Weidemann's scalar algorithm [12]. This suggests the question (still, to my knowledge, open) of whether conditioning of the input is required when block algorithms are used to solve these problems. Some modest progress on this is reported in Section 6. In particular, the efficiency of the algorithm is also established when it is applied to input matrices having arbitrarily many nontrivial invariant factors, provided that (for block size k) the degree of the k^{th} invariant factor is small.

This part of the work being reported is quite "curiosity based:" Algorithms that use different sizes on the left and right are available, have been more completely analyzed, and (at least, for block Wiedemann algorithms) may be more efficient because the number of applications of the input matrix can be reduced — see Kaltofen [9] and Villard [11] for details. That noted, the results of Section 6 are also applicable to block Wiedemann algorithms (rectangular or otherwise) that employ "early termination" heuristics and might therefore be of more general interest — they establish that, with high probability, breakdowns can only occur near the end of a computation. Regardless of badly "conditioning" has failed (and, no matter how pathological the input matrix might be) things will go well until the number of vectors generated, in a Krylov space being traversed, approaches the sum of the degrees of the first $k-1$ invariant factors of the input matrix.

This report omits proofs of a variety of claims. A more complete report that includes these proofs is now available [4].

2. A BLOCK LANCZOS ALGORITHM

Suppose one wishes to solve a system $Ax = b$ for a given matrix $A \in \mathsf{F}_q^{n\times n}$ and vector $b \in \mathsf{F}_q^{n\times 1}$.

2.1 Objectives

Consider a positive integer k and a sequence of vectors $\vec{v} = v_1, v_2, \ldots, v_k \in \mathsf{F}_q^{n\times 1}$; let $\mathcal{KS}_{\vec{v}}$ denote the "Krylov space" defined using these vectors with A as an operator — that is, $\mathcal{KS}_{\vec{v}}$ is the subspace of $\mathsf{F}_q^{n\times 1}$ spanned by the vectors $A^r v_s$ for $r \geq 0$ and $1 \leq s \leq k$. Similarly, for k as above and vectors $\vec{u} = u_1, u_2, \ldots, u_k \in \mathsf{F}_q^{n\times 1}$, let $\widetilde{\mathcal{KS}}_{\vec{u}}$ denote the Krylov space defined using these vectors with A^T as an operator.

A secondary — but key— objective of virtually any "block Lanczos" algorithm is to construct a basis for $\mathcal{KS}_{\vec{v}}$ for a given set of vectors $v_1, v_2, \ldots, v_k$. If the goal is to solve a system $Ax = b$ then one searches for a solution χ as a linear combination of $v_1, v_2, \ldots, v_k$. Modifications needed to certify the inconsistency of systems, sample from the null space, and solve various related problems have been described elsewhere (see, for example, Eberly [3] for details) and will not be discussed further here.

2.2 Details of the Lanczos Phase

2.2.1 Objectives and Invariants

The computation begins with a "Lanczos phase:" Vectors $u_1, u_2, \ldots, u_k, w, w_2, w_3, \ldots, w_k$ are chosen uniformly and independently from $\mathsf{F}_q^{n\times 1}$, w_1 is set to be $A \cdot w + b$, and Gram-Schmidt orthogonalization is applied to try to construct dual orthogonal bases for $\widetilde{\mathcal{KS}}_{\vec{u}}$ and $\mathcal{KS}_{\vec{v}}$, where $v_r = A \cdot w_r$ for $1 \leq r \leq k$ and $\vec{v} = v_1, v_2, \ldots, v_k$.

The Lanczos phase will proceed in a sequence of "stages" (beginning with a "stage 0"). By the end of stage i, for $i \geq 0$, vectors $u_{r,s}, v_{r,s} \in \mathsf{F}_q^{n\times 1}$ will have been constructed, for $0 \leq r \leq i$ and $1 \leq s \leq k$, such that the following property is satisfied.

- *Invariant #1* The vectors $u_{r,s}$ such that $0 \leq r \leq i$ and $1 \leq s \leq k$ span the same subspace of F_q as the vectors $(A^T)^a u_b$ such that $0 \leq a \leq i$ and $1 \leq b \leq k$. The vectors $v_{r,s}$ such that $0 \leq r \leq i$ and $1 \leq s \leq k$ span the same subspace of F_q as the vectors $A^a v_b$ such that $0 \leq a \leq i$ and $1 \leq b \leq k$, as well.

At each point of the computation one has constructed a sequence of vectors

$$\mu_1, \mu_2, \ldots, \mu_\ell \in \{u_{r,s} \mid 0 \leq r \leq i \text{ and } 1 \leq s \leq k\} \qquad (1)$$

and

$$\nu_1, \nu_2, \ldots, \nu_\ell \in \{v_{r,s} \mid 0 \leq r \leq i \text{ and } 1 \leq s \leq k\} \qquad (2)$$

such that the following properties are satisfied at the end of each stage of the Lanczos phase:

- *Invariant #2:* For $1 \leq r, s \leq \ell$, $\mu_r^T \cdot \nu_s = 1$ if $r = s$ and $\mu_r^T \cdot \nu_s = 0$ otherwise.

Henceforth we will say that a vector $u_{r,s}$ (respectively, $v_{r,s}$) is *matched* if $u_{r,s} \in \{\mu_1, \mu_2, \ldots, \mu_\ell\}$ (respectively, if $\nu_{r,s} \in \{\nu_1, \nu_2, \ldots, \nu_\ell\}$), and we will say that $u_{r,s}$ (respectively, $v_{r,s}$) is *unmatched*, otherwise. The following additional invariants should be satisfied at the end of each stage of the Lanczos phase as well.

- *Invariant #3:* If $0 \leq r \leq i$, $1 \leq s \leq k$, and $u_{r,s}$ is unmatched, then $u_{r,s}^T \cdot \nu_t = 0$ for $1 \leq t \leq \ell$, and if $0 \leq r' \leq i$, $1 \leq s' \leq k$, and $v_{r',s'}$ is unmatched, then $\mu_t^T \cdot v_{r',s'} = 0$ for $1 \leq t \leq \ell$ as well.
- *Invariant #4:* If $0 \leq r, r' \leq i$, $1 \leq s, s' \leq k$, and $u_{r,s}$ and $v_{r',s'}$ are both unmatched, then $u_{r,s}^T \cdot v_{r',s'} = 0$.

The next property concerns a positive integer $\Delta_{n,k}$ depending only on the order n of the input matrix and the block size k in use. The Lanczos phase will be terminated at the end of the first stage where the following property does *not* hold.

- *Invariant #5:* If $i \geq \Delta_{n,k}$ then the vectors $u_{r,s}$ and $v_{r,s}$ are both matched, for all r and s such that $0 \leq r \leq i - \Delta_{n,k}$ and $1 \leq s \leq k$.

A final pair of invariants are needed for a version of this algorithm that solves a linear system. The first of these concerns an additional set of vectors $w_{r,s} \in \mathsf{F}_q^{n\times 1}$, for $0 \leq r \leq i$ and $1 \leq s \leq k$, that will be maintained.

- *Invariant #6*: $v_{r,s} = A \cdot w_{r,s}$ for all integers r and s such that $0 \leq r \leq i$ and $1 \leq s \leq k$.

The second concerns four vectors $\sigma, w, \chi, \rho \in \mathsf{F}_q^{n\times 1}$:

- *Invariant #7*: $\sigma = A\cdot w + b$, χ is a linear combination of $\nu_1, \nu_2, \ldots, \nu_\ell$, $\mu_r^T \cdot A \cdot \chi = \mu_r^T \cdot \sigma$ for $1 \le r \le \ell$, and $A\cdot\chi + \rho = \sigma$.

The bulk of the processing during the Lanczos phase concerns ***matching*** and ***orthogonalization*** steps.

In a ***matching*** step, one begins with a subset $\widehat{u}_1, \widehat{u}_2, \ldots, \widehat{u}_a$ of the currently unmatched vectors $u_{r,s}$ as well as a subset $\widehat{v}_1, \widehat{v}_2, \ldots, \widehat{v}_b$ of the currently unmatched vectors $v_{r,s}$ — along with corresponding vectors $\widehat{w}_1, \widehat{w}_2, \ldots, \widehat{w}_b$ from the set of vectors $w_{r,s}$ such that $A\cdot\widehat{w}_t = \widehat{v}_t$ for $1 \le t \le b$.

One then — by some means — determines the rank r of the Hankel matrix

$$H = \left[\widehat{u}_1\widehat{u}_2\ldots\widehat{u}_a\right]^T \cdot \left[\widehat{v}_1\widehat{v}_2\ldots\widehat{v}_b\right] \in \mathsf{F}_q^{a\times b} \tag{3}$$

as well as integer indices $\sigma_1, \sigma_2, \ldots, \sigma_r$ such that $1 \le \sigma_1 < \sigma_2 < \cdots < \sigma_r \le a$, and integer indices $\tau_1, \tau_2, \ldots, \tau_r$ such that $1 \le \tau_1 < \tau_2 < \cdots < \tau_r \le b$, and where the matrix

$$H_{\vec{\sigma},\vec{\tau}} = \left[\widehat{u}_{\sigma_1}\widehat{u}_{\sigma_2}\ldots\widehat{u}_{\sigma_r}\right]^T \cdot \left[\widehat{v}_{\tau_1}\widehat{v}_{\tau_2}\ldots\widehat{v}_{\tau_r}\right] \in \mathsf{F}_q^{r\times r}$$

is a maximal nonsingular submatrix of H. To continue one somehow finds a pair of invertible matrices $X_L, X_R \in \mathsf{F}_q^{r\times r}$ such that $X_R \cdot X_L = H_{\vec{\sigma},\vec{\tau}}^{-1}$. Now one should set

$$\mathcal{M}_L := \left[\widehat{u}_{\sigma_1}\widehat{u}_{\sigma_2}\ldots\widehat{u}_{\sigma_r}\right] \cdot X_L^T \tag{4}$$

— updating the values of $\widehat{u}_{\sigma_1}, \widehat{u}_{\sigma_2}, \ldots, \widehat{u}_{\sigma_r}$ to be the corresponding columns of $\mathcal{M}_L$ in the process — set

$$\mathcal{M}_R := \left[\widehat{v}_{\tau_1}\widehat{v}_{\tau_2}\ldots\widehat{v}_{\tau_r}\right] \cdot X_R \tag{5}$$

and, finally, set

$$\mathcal{M}_R^{pre} := \left[\widehat{w}_{\tau_1}\widehat{w}_{\tau_2}\ldots\widehat{w}_{\tau_r}\right] \cdot X_R \tag{6}$$

— updating the values of $\widehat{v}_{\tau_1}, \widehat{v}_{\tau_2}, \ldots \widehat{v}_{\tau_r}$ (respectively, the values of $\widehat{w}_{\tau_1}, \widehat{w}_{\tau_2}, \ldots, \widehat{w}_{\tau_r}$) to be the corresponding columns of $\mathcal{M}_R$ (respectively, $\mathcal{M}_R^{pre}$) in the process.

Following these updates

$$\mathcal{M}_L^T \cdot \mathcal{M}_R \cdot X_L = X_L \cdot H_{\vec{\sigma},\vec{\tau}} \cdot X_R \cdot X_L = X_L$$

since $X_R \cdot X_L = H_{\vec{\sigma},\vec{\tau}}^{-1}$ — and, since X_L is a nonsingular matrix, $\mathcal{M}_L^T \cdot \mathcal{M}_R = I_r$. The vector $\widehat{u}_{\sigma_s}$ has now been "matched" with the vector $\widehat{v}_{\sigma_s}$, for $1 \le s \le r$, so that the sequences at lines (1) and (2) can be extended by setting $\mu_{\ell+s}$ (respectively, $\nu_{\ell+s}$) to be $\widehat{u}_{\sigma_s}$ (respectively, $\widehat{v}_{\tau_s}$) for $1 \le s \le r$ and adding r to the value of ℓ.

With that noted, the next set of operations are required: For $D := \mathcal{M}_L^T \cdot (A\cdot w + b)$, one should now set

$$\rho := \rho - \mathcal{M}_R \cdot D \quad \text{and} \quad \chi := \chi + \mathcal{M}_R^{pre} \cdot D$$

in order to ensure that Invariant #7 is satisfied after these modifications if it was satisfied before them.

Note. In order to reduce running time, storage space, or rely upon existence code, one might also consider versions of these updates that also modify *unmatched* vectors $\widehat{u}_s$ or $\widehat{v}_t$, for $1 \le s \le a$ or $1 \le t \le b$. If the vector spaces spanned by the vectors $\widehat{u}_s$ for $1 \le s \le a$ (respectively, $\widehat{v}_t$ for $1 \le t \le b$) are unchanged by these updates, and if Invariant #6 is preserved, then the claims in the rest of this report will still be correct and provable, in essentially the same way, even if such an alternative "matching" step is used.

One also needs a pair of ***orthogonalization*** steps. For a "left orthogonalization step," one requires an integer sequence $\gamma_1, \gamma_2, \ldots, \gamma_t \le \ell$ such that

$$1 \le \gamma_1 < \gamma_2 < \cdots < \gamma_t \le \ell$$

and one sets matrices $\mathcal{M}_L, \mathcal{M}_R, \mathcal{M}_R^{pre} \in \mathsf{F}_q^{n\times t}$ to be the matrices with columns $\mu_{\gamma_1}, \mu_{\gamma_2}, \ldots, \mu_{\gamma_t}$, $\nu_{\gamma_1}, \nu_{\gamma_2}, \ldots, \nu_{\gamma_t}$ and $\omega_{\gamma_1}, \omega_{\gamma_2}, \ldots, \omega_{\gamma_t}$ respectively, where $\omega_{\gamma_i} = w_{r,s}$ if $\nu_{\gamma_i} = v_{r,s}$, so that $A\cdot\omega_{\gamma_i} = \nu_{\gamma_i}$ for $1 \le i \le t$. Given a set $\widehat{u}_1, \widehat{u}_2, \ldots, \widehat{u}_m$ of the set of vectors $u_{r,s}$, set $\widehat{U} \in \mathsf{F}_q^{n\times m}$ to be the matrix with columns $\widehat{u}_1, \widehat{u}_2, \ldots, \widehat{u}_m$. Now the update

$$\widehat{U} := \widehat{U} - \mathcal{M}_L \cdot D \qquad \text{for } D = \mathcal{M}_R^T \cdot \widehat{U} \in \mathsf{F}_q^{t\times m} \tag{7}$$

(with corresponding updates to $\widehat{u}_1, \widehat{u}_2, \ldots, \widehat{u}_m$) is sufficient to ensure that $\widehat{u}_r^T \cdot \nu_{\gamma_s} = 0$ for $1 \le r \le m$ and $1 \le s \le t$.

For a "right orthogonalization step," let $\gamma_1, \gamma_2, \ldots, \gamma_t$ and matrices $\mathcal{M}_L, \mathcal{M}_R, \mathcal{M}_R^{pre} \in \mathsf{F}_q^{n\times t}$ be as above. Given a set $\widehat{v}_1, \widehat{v}_2, \ldots, \widehat{v}_m$ of the vectors $v_{r,s}$, and vectors $\widehat{w}_1, \widehat{w}_2, \ldots, \widehat{w}_m$ such that $A\cdot\widehat{w}_r = \widehat{v}_r$ for $1 \le r \le m$, set $\widehat{V}$ and $\widehat{W}$ to be the matrices in $\mathsf{F}_q^{n\times m}$ with columns $\widehat{v}_1, \widehat{v}_2, \ldots, \widehat{v}_m$ and $\widehat{w}_1, \widehat{w}_2, \ldots, \widehat{w}_m$ respectively. Now the updates

$$\widehat{V} := \widehat{V} - \mathcal{M}_R \cdot D \quad \text{and} \quad \widehat{W} := \widehat{W} - \mathcal{M}_R^{pre} \cdot D \qquad \text{for } D = \mathcal{M}_L^T \cdot \widehat{V} \in \mathsf{F}_q^{t\times m} \tag{8}$$

(with updates to $\widehat{v}_1, \widehat{v}_2, \ldots, \widehat{v}_m$ and $\widehat{w}_1, \widehat{w}_2, \ldots, \widehat{w}_m$) ensure that $\mu_{\gamma_r}^T \cdot \widehat{v}_s = 0$ and $A\cdot\widehat{w}_s = \widehat{v}_s$ for $1 \le r \le t$ and $1 \le s \le m$.

Stage 0: To begin one should set $\ell = 0$, $\chi = 0$, $\sigma = \rho = A\cdot w + b$, and one should initialize vectors by setting $u_{0,s}$ to be u_s, setting $w_{0,s}$ to be w_s, and setting $v_{0,s}$ to be $A\cdot w_s$ for $1 \le s \le k$. One should continue by matching $u_{0,1}, u_{0,2}, \ldots, u_{0,k}$ with $v_{0,1}, v_{0,2}, \ldots, v_{0,k}$ as described above. For this first stage, all matched vectors should be orthogonalized against all unmatched vectors. That is, one should perform left and right orthogonalizations with $t = \ell$ and $\gamma_r = r$ for $1 \le r \le t$, setting $\widehat{u}_1, \widehat{u}_2, \ldots, \widehat{u}_m$ (respectively, $\widehat{v}_1, \widehat{v}_2, \ldots, \widehat{v}_m$) to be the set of all vectors $u_{0,r}$ (respectively, $v_{0,r}$) that were not matched during the initial matching step.

Stage i, for $i \ge 1$: Each later stage should begin by applying A or A^T as an operator:

$$u_{i,s} := A^T \cdot u_{i-1,s}, \quad v_{i,s} : A\cdot v_{i-1,s} \quad \text{and} \quad w_{i,s} := v_{i-1,s} \tag{9}$$

for each integer s such that $1 \le s \le k$

Now, while it is necessary to apply orthogonalization steps similar to the ones at lines (7) and (8) in order to establish Invariant #3 once again, Invariant #5 serves to limit the number of these that are required.

LEMMA 2.1. *Suppose there are at least $i+1$ stages of the Lanczos phase of the algorithm, and Invariant #5 is satisfied at the end of each of the first i stages. Then, at the beginning of stage i (that is, the $i+1^{st}$ stage),*

$$(A^T \cdot u_{i-1,s}) \cdot \nu_t = \mu_t^T \cdot (A\cdot v_{i-1,s}) = 0$$

for $1 \le s \le k$ and for every integer t such that $1 \le t \le \ell$ and either $\mu_t = u_{g,h}$ or $\nu_t = v_{g,h}$ for integers g and h such that $0 \le g \le i - 2\cdot\Delta_{n,k} - 3$ and $1 \le h \le k$.

Consequently, in order to re-establish Invariant #3 it now suffices to employ updates as shown at lines (7) and (8), above, where $t = k$, $\widehat{u}_r = u_{i,r}$ and $\widehat{v}_r = v_{i,r}$ for $1 \le r \le k$ and where matched vectors μ_a and ν_a such that $\mu_a = u_{c,d}$

and $\nu_a = v_{c',d'}$ such that $i - 2 \cdot \Delta_{n,k} - 2 \leq c, c' \leq i$ and $1 \leq d, d' \leq k$.

Indeed, these are the only matched vectors that will be needed after this point in the computation. Consequently, if all other matched vectors are deleted at this point a circular queue (of arrays) can be used to store all of the vectors $u_{r,s}$, $v_{r,s}$ and $w_{r,s}$ that might still be needed, a circular queue (of linked lists) can be used to maintain the indices of currently unmatched vectors, and a linked list (whose length will not exceed $(2 \cdot \Delta_{n,k} + 3) \cdot k$) can be used to store the indices of pairs of matched vectors that are still required.

This stage should continue with a process of pairing (and "matching") previously unmatched vectors $u_{r,s}$ and $v_{r',s'}$ in order to extend the sequences at lines (1) and (2). It turns out that a simple "greedy" strategy will suffice (and, this the primary algorithmic contribution here): Whenever possible, "older" vectors should be matched before "newer" ones.

In particular, since Invariant #5 was satisfied at the end of stage $i-1$ the only unmatched vectors at the beginning of stage i are vectors $u_{r,s}$ and $v_{r,s}$ such that $i - \Delta_{n,k} \leq r \leq i$ and $1 \leq s \leq k$. The matching process should include a sequence of rounds $1, 2, 3, \ldots, 2 \cdot \Delta_{n,k} + 1$, as follows.

- *Round $2j+1$, for $0 \leq j \leq \Delta_{n,k} - 1$:* The unmatched vectors $u_{i-\Delta_{n,k}+j,r}$ such that $1 \leq r \leq k$ are matched with the unmatched vectors $v_{i,s}$ such that $1 \leq s \leq k$, in order to obtain additional pairs of matched vectors. A left orthogonalization step is carried out using the newly matched vectors and all (still) unmatched vectors $u_{r,s}$ such that $i - \Delta_{n,k} + j \leq r \leq i$ and $1 \leq s \leq k$, and a right orthogonalization step is carried out using the newly matched vectors and all (still) unmatched vectors $v_{i,s}$ such that $1 \leq s \leq k$.

- *Round $2j+2$, for $0 \leq j \leq \Delta_{n,k} - 1$:* The unmatched vectors $u_{i,r}$ such that $1 \leq r \leq k$ are matched with the unmatched vectors $v_{i-\Delta_{n,k}+j,s}$ such that $1 \leq s \leq k$, in order to obtain additional pairs of matched vectors. A left orthogonalization step is carried out using the newly matched vectors and all (still) unmatched vectors $u_{i,s}$ such that $1 \leq s \leq k$, and a right orthogonalization step is carried out using the newly matched vectors and all (still) unmatched vectors $v_{r,s}$ such that $i - \Delta_{n,k} + j \leq r \leq i$ and $1 \leq s \leq k$.

- *Round $2 \cdot \Delta_{n,k} + 1$:* The unmatched vectors $u_{i,r}$ such that $1 \leq r \leq k$ are matched with the unmatched vectors $v_{i,s}$ such that $1 \leq s \leq k$ in order to obtain additional pairs of matched vectors. A left orthogonalization step is carried out using the newly matched vectors and all (still) unmatched vectors $u_{i,s}$ such that $1 \leq s \leq k$, and a right orthogonalization step is carried out using the newly matched vectors and all (still) unmatched vectors $v_{i,s}$ such that $1 \leq s \leq k$.

Quite a few orthogonalization steps have been left out above. Nevertheless, the following can be proved.

LEMMA 2.2. *Consider the updates of vectors $u_{r,s}$ and $v_{r,s}$. given above in stage i of the Lanczos phase for $i \geq 1$.*

(a) *If j is an integer such that $0 \leq j \leq i-1$ then the subspace of $\mathsf{F}_q^{n \times 1}$ spanned by the set of vectors $u_{r,s}$ (respectively, $v_{r,s}$) such that $0 \leq r \leq j$ and $1 \leq s \leq k$ is unchanged by the updates included in stage i of the Lanczos phase.*

(b) *If h is an integer such that $0 \leq h \leq \Delta_{n,k} - 1$, then the only vectors whose* values *can be changed by updates, after round $2h+2$ of the Lanczos phase, are vectors $u_{r,s}$ and $v_{r,s}$ such that $i - \Delta_{n,k} + h < r \leq i$ and $1 \leq s \leq k$.*

(c) *If the steps described above and Invariants #1–#5 were satisfied before stage i of the Lanczos phase of the computation then Invariants #1–#4 are satisfied at the end of stage i as well.*

2.3 Details of the Elimination Phase

2.3.1 Objectives and Invariants

The computation will end with an "elimination phase" which will also proceed in a series of stages. The following data is accumulated:

- A sequence of vectors
$$\lambda_1, \lambda_2, \ldots, \lambda_m \in \mathsf{F}_q^{n \times 1}, \tag{10}$$
will be stored as the columns of a matrix $M_\lambda \in \mathsf{F}_q^{n \times m}$.

- Another sequence of vectors of the same length m $\kappa_1, \kappa_2, \ldots, \kappa_m \in \mathsf{F}_q^{n \times 1}$ such that $A \cdot \kappa_r = \lambda_r$ for $1 \leq r \leq m$. These will be stored as the columns of a matrix $M_\kappa \in \mathsf{F}_q^{n \times m}$ such that $A \cdot M_\kappa = M_\lambda$.

- A permutation matrix $P \in \mathsf{F}_q^{n \times n}$ will be maintained.

- Another sequence of vectors
$$\varphi_1, \varphi_2, \ldots, \varphi_g \in \mathsf{F}_q^{n \times 1}, \tag{11}$$
will be stored as the columns of a matrix $M_\varphi \in \mathsf{F}_q^{n \times g}$.

The following properties will be satisfied at the end of stage j of the elimination phase (for $j \geq 0$) if $j+1$ or more stages are included in the computation — assuming that the Lanczos phase ended with stage i.

- *Invariant #8:* The vectors $\nu_1, \nu_2, \ldots, \nu_\ell, \lambda_1, \lambda_2, \ldots, \lambda_m$ span the same subspace of $\mathsf{F}_q^{n \times 1}$ as the vectors $A^a v_b$ such that $1 \leq a \leq i + j$ and $1 \leq b \leq k$.

- *Invariant #9:* $\mu_r^T \cdot \lambda_s = 0$ for all integers r and s such that $1 \leq r \leq \ell$ and $1 \leq s \leq m$.

- *Invariant #10:*
$$P \cdot M_\lambda = \begin{bmatrix} L_\lambda \\ X_\lambda \end{bmatrix} \tag{12}$$
for a lower triangular matrix $L_\lambda \in \mathsf{F}_q^{m \times m}$ with ones on its diagonal and for a matrix $X_\lambda \in \mathsf{F}_q^{(n-m) \times m}$.

- *Invariant #11:* The vectors $\varphi_1, \varphi_2, \ldots, \varphi_g$ are linearly independent and the sequence of vectors
$$\nu_1, \nu_2, \ldots, \nu_\ell, \lambda_1, \lambda_2, \ldots, \lambda_m, \\ A \cdot \varphi_1, A \cdot \varphi_2, \ldots, A \cdot \varphi_g$$
span the same subspace of $\mathsf{F}_q^{n \times 1}$ as the vectors $A^a v_b$ such that $0 \leq a \leq i + j + 1$ and $1 \leq b \leq k$.

Another pair of invariants are needed for a version of the algorithm that solves a linear system.

- *Invariant #12*: $A \cdot M_\kappa = M_\lambda$.

- *Invariant #13:* $A \cdot \chi + \rho = A \cdot w + b$. Furthermore, χ is a linear combination of

$$\nu_1, \nu_2, \ldots, \nu_\ell, \lambda_1, \lambda_2, \ldots, \lambda_m,$$

$\mu_h^T \cdot A \cdot \chi = \mu_h^T \cdot (A \cdot w + b)$ for $1 \leq h \leq \ell$, and if

$$P \cdot \rho = \left[\rho_1 \rho_2 \ldots \rho_n\right]^T \in \mathsf{F}_q^{n \times 1}$$

for P as shown at line (12) then $\rho_t = 0$ for $1 \leq t \leq m$.

The algorithm will continue until $g = 0$ (so that the sequence of vectors shown at line (11), above is empty) at the end of a stage.

2.3.2 Details of Stages

The stages will also maintain a pair of matrices $\mathcal{M}_{new} \in \mathsf{F}_q^{n \times r}$ and $\mathcal{M}_{new}^{pre} \in \mathsf{F}_q^{n \times r}$ such that $A \cdot \mathcal{M}_{new}^{pre} = \mathcal{M}_{new}$; the following operations will be performed.

- An ***orthogonalization step***: Performing a right orthogonalization step, with the matrices $\mathcal{M}_L$, $\mathcal{M}_R$ and $\mathcal{M}_R^{pre}$ from the final stage of the Lanczos phase and using $\mathcal{M}_{new}$ and $\mathcal{M}_{new}^{pre}$ as the matrices $\widehat{V}$ and $\widehat{W}$, respectively (in updates at line (8)), will suffice to ensure that the columns of $\mathcal{M}_{new}$ are orthogonal to $\mu_1, \mu_2, \ldots, \mu_\ell$.

- An ***elimination step*** will also be needed: If

$$P \cdot \mathcal{M}_{new} = \begin{bmatrix} Y \\ Z \end{bmatrix}$$

for $Y \in \mathsf{F}_q^{m \times r}$ and $Z \in \mathsf{F}_q^{(n-m) \times r}$, then performing the updates

$$\mathcal{M}_{new} := \mathcal{M}_{new} - M_\lambda \cdot L_\lambda^{-1} \cdot Y$$

and

$$\mathcal{M}_{new}^{pre} := \mathcal{M}_{new}^{pre} - M_\kappa \cdot L_\lambda^{-1} \cdot Y$$

will suffice to ensure that

$$P \cdot \mathcal{M}_{new} = \begin{bmatrix} 0_{m \times h} \\ \widehat{\mathcal{M}}_{new} \end{bmatrix}$$

for a matrix $\widehat{\mathcal{M}}_{new} \in \mathsf{F}_q^{(n-m) \times h}$.

- In a ***compression step*** one should determine the rank s of the above matrix $\widehat{\mathcal{M}}_{new}$ as well as a matrix $X \in \mathsf{F}_q^{r \times s}$ such that $\widehat{\mathcal{M}}_{new} \cdot X \in \mathsf{F}_q^{(n-m) \times s}$ has rank s as well. Matrices should be updated by setting

$$\mathcal{M}_{new} := \mathcal{M}_{new} \cdot X \quad \text{and} \quad \mathcal{M}_{new}^{pre} := \mathcal{M}_{new}^{pre} \cdot X$$

(effectively replacing $\widehat{\mathcal{M}}_{new}$ with $\widehat{\mathcal{M}}_{new} \cdot X$ as well and setting r to be s).

- In a ***triangularization*** step one should compute a permutation matrix $\widehat{P} \in \mathsf{F}_q^{(n-m) \times (n-m)}$, a lower triangular matrix $\widehat{L} \in \mathsf{F}_q^{(n-m) \times r}$ with ones on the diagonal, and a nonsingular upper triangular matrix $\widehat{U} \in \mathsf{F}_q^{r \times r}$, such that $\widehat{\mathcal{M}}_{new} = \widehat{P} \cdot \widehat{L} \cdot \widehat{U}$. Another pair of updates

$$\mathcal{M}_{new} := \mathcal{M}_{new} \cdot \widehat{U}^{-1} \quad \text{and} \quad \widehat{\mathcal{M}}_{new} := \widehat{\mathcal{M}}_{new} \cdot \widehat{U}^{-1}$$

effectively replaces $\widehat{\mathcal{M}}_{new}$ with $\widehat{P} \cdot \widehat{L}$.

- Suppose next that $\widetilde{L} \in \mathsf{F}_q^{r \times r}$ consists of the top r rows of the above matrix $\widehat{L}$, so that $\widetilde{L}$ is a nonsingular lower triangular matrix with ones on its diagonal and recall that, by invariant #13, above,

$$P \cdot \rho = \begin{bmatrix} 0_m \\ \widehat{\rho} \end{bmatrix}$$

for a vector $\rho \in \mathsf{F}_q^{(n-m) \times 1}$. Set $\widehat{\rho}_1 \in \mathsf{F}_q^{r \times 1}$ to be the vector containing the top r entries of $\widehat{P}^T \cdot \widehat{\rho}$. Then, in a ***solution*** step, one should update ρ and χ by setting

$$\rho := \rho - \mathcal{M}_{new} \cdot \widetilde{L}^{-1} \cdot \widehat{\rho}_1 \quad \text{and} \quad \chi := \chi + \mathcal{M}_{new}^{pre} \widetilde{L}^{-1} \cdot \widehat{\rho}_1.$$

- Finally, in an ***update step***, another pair of updates should be performed:

$$P := P \cdot \begin{bmatrix} I_m & 0 \\ 0 & \widehat{P}^T \end{bmatrix} \quad \text{and} \quad M_\varphi := \mathcal{M}_{new},$$

and the columns of $\mathcal{M}_{new}$ and $\mathcal{M}_{new}^{pre}$ should be appended to the matrices M_λ and M_κ respectively.

Stage 0: Initially $m = 0$, and P is the identity matrix. Set $\mathcal{M}_{new}$ and $\mathcal{M}_{new}^{pre}$ to include as columns all vectors $v_{r,s}$ (respectively, $w_{r,s}$) such that $0 \leq r \leq i-1$, $1 \leq s \leq k$, and $v_{r,s}$ was unmatched at the end of the final stage of the Lanczos phase. The *compression*, *triangularization*, *solution* and *update* steps should be carried out — and all columns of M_φ removed.

The vectors $v_{i,s}$ (and, respectively, $w_{i,s}$) such that $1 \leq s \leq k$ and $v_{i,s}$ was unmatched at the end of stage i of the Lanczos phase should then be used as the initial columns of $\mathcal{M}_{new}$ (respectively, $\mathcal{M}_{new}^{pre}$). The *elimination*, *compression*, *triangularization*, *solution* and *update* steps should be carried out. Finally, the vectors $v_{i,s'}$ such that $1 \leq s' \leq k$ and $v_{i,s'}$ was matched at the end of stage i should be appended as columns of M_φ as well.

Stage j for $j \geq 1$: The matrices $\mathcal{M}_{new}$ and $\mathcal{M}_{new}^{pre}$ should be initialized to be $A \cdot M_\varphi$ and M_φ, respectively. The *orthogonalization*, *elimination*, *compression*, *triangularization*, *solution* and *update* steps should be carried out.

As noted above, this phase of the algorithm will end as soon as $g = 0$ at the end of a stage. Following this, one should check whether $\rho = 0$. If it is, then $\chi - w$ can be returned a solution for the system $Ax = b$. Otherwise, a solution has not been found and, indeed, no vector χ such that $A \cdot \chi = A \cdot w + b$ is contained in $\mathcal{KS}_{\vec{v}}$.

2.4 Correctness and Efficiency

THEOREM 2.3. *If the above algorithm is executed then, on termination, the vectors*

$$\nu_1, \nu_2, \ldots, \nu_\ell, \lambda_1, \lambda_2, \ldots, \lambda_m \tag{13}$$

form a basis for the Krylov space $\mathcal{KS}_{\vec{v}}$ and either a vector $x \in \mathsf{F}_q^{n \times 1}$ such that $A \cdot x = b$ has returned, or $\mathcal{KS}_{\vec{v}}$ does not include any vector χ such that $A \cdot \chi = A \cdot w + b$ and it has been reported that a solution has not been found.

THEOREM 2.4. *Suppose that $\mathcal{KS}_{\vec{v}}$ has dimension $d = \ell + m$. Then the above algorithm can be applied to produce a basis as shown at line* (13) *using the selection of $2k$ vectors uniformly and independently from $\mathsf{F}_q^{n \times 1}$, at most $d + (\Delta_{n,k} + 2)k + 1$ multiplications of A by vectors, at most $d + (\Delta_{n,k} + 1)k$*

multiplications of A^T by vectors, $O(d(\Delta_{n,k}\cdot k)^2 n + (m+\Delta_{n,k})mn)$ additional operations over F_q and space required to store $O(mn + \Delta_{n,k}\cdot kn)$ elements of F_q.

It follows that if both $\Delta_{n,k}\cdot k$ and the length of elimination phase is reasonably short (so that m is as well) then this algorithm is asymptotically efficient. We will be considering cases where $\Delta_{n,k}\cdot k, m \in O(\log n)$ — in which case it follows by the above that the algorithm requires $d + O(\log n)$ multiplications of A and A^T by vectors, $O(dn\log^2 n)$ additional operations over F_q, and space required to store $O(n\log n)$ elements of F_q.

2.5 Matrices of Interest

For a positive integer k_L, vectors $\vec{u} = u_1, u_2, \ldots, u_{k_L} \in \mathsf{F}_q^{n\times 1}$, and a positive integer r, let $\widehat{\mathcal{K}}_{A,\vec{u},r} \in \mathsf{F}_q^{n\times rk_L}$ be the matrix with columns

$$\begin{array}{c} u_1,\ A^T\cdot u_1,\ (A^T)^2\cdot u_1,\ \ldots,\ (A^T)^{r-1}\cdot u_1,\\ u_2,\ A^T\cdot u_2,\ (A^T)^2\cdot u_2,\ \ldots,\ (A^T)^{r-1}\cdot u_2,\\ \ldots\ u_{k_L},\ A^T\cdot u_{k_L},\ (A^T)^2\cdot u_{k_L},\ \ldots,\ (A^T)^{r-1}\cdot u_{k_L}\end{array}$$

— that is, the vectors $(A^T)^h u_j$ h and j such that $0 \le h \le r-1$ and $1 \le j \le k_L$. Similarly, for positive integers k_R and s, and vectors $\vec{v} = v_1, v_2, \ldots, v_{k_R} \in \mathsf{F}_q^{n\times 1}$, let $\mathcal{K}_{A,\vec{v},s} \in \mathsf{F}_q^{n\times sk_R}$ be the matrix with columns

$$\begin{array}{c} v_1,\ A\cdot v_1,\ A^2\cdot v_1,\ \ldots,\ A^{s-1}\cdot v_1,\\ v_2,\ A\cdot v_2,\ A^2\cdot v_2,\ \ldots,\ A^{s-1}\cdot v_2,\\ \ldots\ v_{k_R},\ A\cdot v_{k_R},\ A^2\cdot v_{k_R},\ \ldots,\ A^{s-1}\cdot v_{k_R}\end{array}$$

— that is, the vectors $A^h v_j$ for h and j such that $0 \le h \le s-1$ and $1 \le j \le k_R$. Finally, for positive integers k_L and k_R, vectors $u_1, u_2, \ldots, u_{k_L} \in \mathsf{F}_q^{n\times 1}$ and $v_1, v_2, \ldots, v_{k_R} \in \mathsf{F}_q^{n\times 1}$, and positive integers r and s, let

$$\mathcal{H}_{A,\vec{u},\vec{v},r,s} = \widehat{\mathcal{K}}^T_{A,\vec{u},r}\cdot\mathcal{K}_{A,\vec{v},s} \in \mathsf{F}_q^{rk_L\times sk_R}. \tag{14}$$

This is a block Hankel matrix; in particular, it has the form

$$\begin{bmatrix} H_{1,1} & H_{1,2} & \ldots & H_{1,k_R}\\ H_{2,1} & H_{2,2} & \ldots & H_{2,k_R}\\ \vdots & \vdots & \ddots & \vdots\\ H_{k_L,1} & H_{k_L,2} & \ldots & H_{k_L,k_R}\end{bmatrix}$$

where each submatrix $H_{a,b}$ is an $r\times s$ Hankel matrix: For $1 \le c \le r$ and $1 \le d \le s$, its entry in row c and column d is the value α_{c+d-2}, where $\alpha_t = u_a^T\cdot A^t\cdot v_b$. for $0 \le t \le r+s-2$.

It suffices to consider the case that $k_L = k_R = k$ in order to analyze the algorithm given here.

LEMMA 2.5. *Let k be a positive integer and let*

$$\vec{u} = u_1, u_2, \ldots, u_k, \in \mathsf{F}_q^{n\times 1} \text{ and } \vec{v} = v_1, v_2, \ldots, v_k \in \mathsf{F}_q^{n\times 1}.$$

If $i \ge \Delta_{n,k}-1$ and $\mathcal{H}_{A,\vec{u},\vec{v},a,a+\Delta_{n,k}}$ and $\mathcal{H}_{A,\vec{u},\vec{v},a+\Delta_{n,k},a}$ each have maximal rank ak for $0 \le a \le i-\Delta_{n,k}+1$ then Invariant #5 is satisfied at the end of each of the first i stages of the Lanczos phase of the computation, (so that there will be at least $i+1$ stages in the Lanczos phase).

Note: It will generally be clear that the input matrix $A \in \mathsf{F}_q^{n\times n}$ has been used to define the above matrices $\widehat{\mathcal{K}}_{A,\vec{u},r}$, $\mathcal{K}_{A,\vec{v},t}$, and $\mathcal{H}_{A,\vec{u},\vec{v},s,t}$ — so they will generally be denoted more succinctly as $\widehat{\mathcal{K}}_{\vec{u},s}$, $\mathcal{K}_{\vec{v},t}$, and $\mathcal{H}_{\vec{u},\vec{v},s,t}$ respectively.

3. A USEFUL MATRIX NORMAL FORM

Let $f = x^d + \alpha_{d-1}x^{d-1} + \alpha_{d-2}x^{d-2} + \cdots + \alpha_0$ be a monic polynomial with degree d in $\mathsf{F}_q[x]$. Then the *companion matrix* of f is the $d\times d$ matrix C_f with a one below the diagonal and zeroes elsewhere in the first $d-1$ columns and whose final column is

$$\begin{bmatrix}-\alpha_0 & -\alpha_1 & -\alpha_2 & \ldots & -\alpha_{d-1}\end{bmatrix}^T.$$

it is well-known that the minimal polynomial and characteristic polynomial of C_f are both equal to f. Every matrix $A \in \mathsf{F}_q^{n\times n}$ is similar to a unique block-diagonal matrix $C_{\vec{f}} = C_{f_1,f_2,\ldots,f_\ell} \in \mathsf{F}_q^{n\times n}$ whose diagonal blocks are companion matrices $C_{f_1}, C_{f_2}, \ldots, C_{f_\ell}$ for monic polynomials

$$f_1, f_2, \ldots, f_\ell \in \mathsf{F}_q[x]$$

such that f_i is divisible by f_{i+1} in $\mathsf{F}_q[x]$ for $1 \le i \le \ell-1$. The polynomials $f_1, f_2, \ldots, f_\ell$ are called the *invariant factors* of A (so that, in particular, f_i will be called the "i^{th} invariant factor" of A for $1 \le i \le \ell$) and these are also unique. For a proof of the uniqueness of both the Frobenius normal form and the invariant factors see, for example, Gantmacher [6] (who refers to the above matrix $C_{f_1,f_2,\ldots,f_\ell}$ as the "first natural normal form of A," instead).

We will say that "A has ℓ invariant factors" if its Frobenius normal form is as described above. We will say that "A has h *nontrivial* invariant factors" if $h \le \ell$ and $f_h \ne x = f_{h+1} = f_{h+2} = \cdots = f_\ell$.

4. EXPONENTIAL NULLITY

Consider a matrix $B \in \mathsf{F}_q^{s\times t}$. Let us define the *left exponential nullity* of B, $\mathsf{xnull}_L(B)$, to be the number of vectors $x \in \mathsf{F}_q^{s\times 1}$ such that $x^T\cdot B = 0$, and the *right exponential nullity* of B, $\mathsf{xnull}_R(B)$, to be the number of vectors $y \in \mathsf{F}_q^{t\times 1}$ such that $B\cdot y = 0$.

It is easily shown that if B has rank r then $\mathsf{xnull}_L(B) = q^{s-r}$ and $\mathsf{xnull}(B) = q^{t-r}$ — so that the left and right exponential nullifies are the same if $s = t$. We will call this common value the *exponential nullity* and denote it as $\mathsf{xnull}(B)$ in this case.

While this notation was not used, various properties of exponential nullities, including one similar to the following, were explored in the report [5].

LEMMA 4.1. *Let $A \in \mathsf{F}_q^{n\times n}$, $s,t \ge 1$, and consider the matrix $H = \mathcal{H}_{k,\vec{u},\vec{v},s,t} \in \mathsf{F}_q^{sk\times tk}$ where vectors $\vec{u} = u_1, u_2, \ldots, u_k$ and $\vec{w} = w, w_2, w_3, \ldots, w_k$ are chosen uniformly and independently from $\mathsf{F}_q^{n\times 1}$, $w_1 = A\cdot w + b$, $v_h = A\cdot w_h$ for $1 \le h \le k$, and $\vec{v} = v_1, v_2, \ldots, v_k$.*

Then, if r is an integer such that $0 \le r \le \min(s,t)$ then the probability that the rank of the above matrix is less than or equal to r is less than or equal to each of $\mathsf{E}[\mathsf{xnull}_L(H)]/q^{s-r}$ and $\mathsf{E}[\mathsf{xnull}_R(H)]/q^{t-r}$.

Furthermore, if $s \le t$ then the probability that the rank of this matrix ls less than s is at most $(\mathsf{E}[\mathsf{xnull}_L(H)]-1)/(q-1)$ and, if $t \le s$ then the probability that the rank of this matrix is less than t is at most $(\mathsf{E}[\mathsf{xnull}_R(H)]-1)/(q-1)$.

Claims about the expected performance of the algorithm described in Section 2 will be established by bounding the expected value of the left or right exponential nullity of matrices $\mathcal{H}_{k,\vec{u},\vec{v},s,t}$, where either $s = t+\Delta_{n,k}$ or $t = s+\Delta_{n,k}$, and applying the above lemma along with Lemma 2.5.

A second (new) technical lemma will also be of use in explaining the behaviour of the algorithm presented in Section 2 and, indeed, a variety of other Krylov-based algorithms that are presently under development.

LEMMA 4.2. *Let $A \in \mathsf{F}_q^{n\times n}$, $s,t \geq 1$, and consider the matrix $H = \mathcal{H}_{k,\vec{u},\vec{v},s,t} \in \mathsf{F}_q^{sk\times tk}$ where vectors $\vec{u} = u_1, u_2, \ldots, u_k$ and $\vec{w} = w, w_2, w_3, \ldots, w_k$ are chosen uniformly and independently from $\mathsf{F}_q^{n\times 1}$, $w_1 = A\cdot w$, $v_h = A\cdot w_h$ for $1 \leq h \leq k$, and $\vec{v} = v_1, v_2, \ldots, v_k$.*

If another matrix $W \in \mathsf{F}_q^{sk\times tk}$ is selected (from some subset of $\mathsf{F}_q^{sk\times tk}$) independently from the above vectors $\vec{u}$ and $\vec{w}$, then the expected value of the left exponential nullity of the matrix $H+W$ is less than or equal to that of H, and the expected value of the right exponential nullity of $H+W$ is less than or equal to that of H as well.

Question: Is there a similar provable result that relates the expected values of the (left or right) *nullities* of matrices H and $H+W$, for H and W as above? It seems likely that a result along these lines could establish the efficiency of the algorithm of Section 2 without any assumptions about the input matrix A at all. Unfortunately, the techniques used to prove Lemma 4.2 do not seem to be applicable when "nullites" are considered instead of "exponential nullities."

5. MATRICES FOR WHICH BREAKDOWN IS PROVABLY UNLIKELY

The reliability of a block Lanczos algorithm using block size k, when applied to a matrix $A \in \mathsf{F}_q^{n\times n}$ such that the number h of nontrivial invariant factors is less than k, was first established by Bradford Hovinen in his Master's Thesis [7]. One can also establish this by bounding the exponential nullities of the associated matrices considered in Lemma 2.5, above. The following is a reasonably straightforward extension of bounds presented in [5].

LEMMA 5.1. *Suppose that vectors*

$$u_1, u_2, \ldots, u_k, w, w_2, w_3, \ldots, w_k$$

are chosen uniformly and independently from $\mathsf{F}_q^{n\times 1}$, and that $w_1 = A\cdot w + b$, and that $v_a = A\cdot w_a$ for $1 \leq a \leq h$. Let $\vec{u} = u_1, u_2, \ldots, u_k$ and $\vec{v} = v_1, v_2, \ldots, v_k$.

Suppose, as well, that A has h nontrivial invariant factors, for $h \leq k-1$, and let r be the rank of A.

Then if a is an integer such that $1 \leq a \leq \lfloor r/k\rfloor - \Delta_{n,k} - 1$,

$$\mathsf{E}[\mathsf{xnull}_L(\mathcal{H}_{\vec{u},\vec{v},a,a+\Delta_{n,k}})] \leq 1 + q^{(2-\Delta_{n,k})k}\cdot(f(h,k) + f(h,k)^2\cdot q^{1-k})$$

and

$$\mathsf{E}[\mathsf{xnull}_R(\mathcal{H}_{\vec{u},\vec{v},a+\Delta_{n,k},a})] \leq 1 + q^{(2-\Delta_{n,k})k}\cdot(f(h,k) + f(h,k)^2\cdot q^{1-k})$$

as well, where

$$f(h,k) = \begin{cases} 6\cdot\log_q n & \text{if } k = h+1,\\ 4 & \text{if } k = h+2,\\ 1+2q^{h-k+1} & \text{if } k \geq h+3.\end{cases}$$

Let $c > 0$; then (assuming that $k \geq 2$), if

$$\Delta_{n,k} \geq \lceil((1+c)\log_q n + 2\log_q\log_q n + 7)/k\rceil \qquad (15)$$

then

$$q^{(\Delta_{n,k}-2)k} \geq 32n^{1+c}\log_q^2 n \geq n^{1+c}\cdot(f(h,k) + f(h,k)^2\cdot q^{1-k})$$

implying that the expected values of $\mathsf{xnull}_L(\mathcal{H}_{\vec{u},\vec{v},a,a+\Delta_{n,k}})$ and $\mathsf{xnull}_R(\mathcal{H}_{\vec{u},\vec{v},a+\Delta_{n,k}})$ are each at most $1+n^{-(c+1)}$. Now, it follows by the inequalities at the end of Lemma 4.1 that the probability that either of the matrices $\mathcal{H}_{\vec{u},\vec{v},a,a+\Delta_{n,k}}$ or $\mathcal{H}_{\vec{u},\vec{v},a+\Delta_{n,k},a}$ is rank-deficient is at most $2n^{-(1+c)}$.

Summing failure probabilities, one can establish that there are at least $\lfloor r/k\rfloor - 1$ stages of the Lanczos phase of the algorithm with probability at least $1-2n^{-c}$. In this case $\ell \geq (\lfloor r/k\rfloor - \Delta_{n,k} - 2)\cdot k \geq r - (\Delta_{n,k}+3)k$ (since Invariant #5 was satisfied at the end of stage $\lfloor n/k\rfloor - 3$), so that $m \leq r - \ell \leq (\Delta_{n,k}+3)\cdot k \in O(\log n)$. It follows by Theorem 2.4 and the remarks following it that, with high probability, the algorithm is efficient in this case.

6. BREAKDOWN IS NOT EARLY

Something positive can also be said about the performance of the algorithm when it is applied to certain matrices with k or more nontrivial invariant factors: Suppose the first k invariant factors are $f_1, f_2, \ldots, f_k$, let

$$r_1 = \begin{cases}\deg(f_1) & \text{if } x \text{ does not divide } f_1,\\ \deg(f_1)-1 & \text{otherwise,}\end{cases} \qquad (16)$$

and, for $2 \leq s \leq k$, let

$$r_s = \begin{cases} r_{s-1} + \deg(f_s) & \text{if } x \text{ does not divide } f_s,\\ r_{s-1} + \deg(f_s) - 1 & \text{otherwise.}\end{cases} \qquad (17)$$

If $\vec{v} = v_1, v_2, \ldots, v_k$ is as described in previous sections then the dimension of the Krylov space $\mathcal{KS}_{\vec{v}}$ is at most r_k.

Let h be an integer such that $1 \leq h \leq k-1$. Then it is possible express A as a sum $A = A_1 + A_2$ for $A_1, A_2 \in \mathsf{F}_q^{n\times n}$ such that $A_1\cdot A_2 = A_2\cdot A_1 = 0$ and A_1 has h nontrivial invariant factors $f_1, f_2, \ldots, f_h$.

Furthermore, it turns out that if the vectors

$$u_1, u_2, \ldots, u_k, w, w_2, w_3, \ldots, w_k$$

are chosen uniformly and independently from $\mathsf{F}_q^{n\times 1}$, $w_1 = A\cdot w + b$, $v_a = A\cdot w_a$, $y_a = A_1\cdot w_a$ $z_a = A_2\cdot w_a$ for $1 \leq a \leq k$, $\vec{y} = y_1, y_2, \ldots, y_k$ and $\vec{z} = z_1, z_2, \ldots, z_k$ then

$$\mathcal{H}_{A,\vec{u},\vec{v}s,t} = \mathcal{H}_{A_1,\vec{u},\vec{y},s,t} + \mathcal{H}_{A_2,\vec{u},\vec{z},s,t} \qquad (18)$$

and, furthermore, the matrices $\mathcal{H}_{A_1,\vec{u},\vec{y},s,t}$ and $\mathcal{H}_{A_2,\vec{u},\vec{z},s,t}$ are independently distributed. This can be used, along with Lemmas 4.2 and 5.1, to establish the following.

LEMMA 6.1. *Suppose that the vectors $u_1, u_2, \ldots, u_k$ and $w, w_2, w_3, \ldots, w_k$ are chosen uniformly and independently from $\mathsf{F}_q^{n\times 1}$, that $w_1 = A\cdot w + b$, and that $v_a = A\cdot w_a$ for $1 \leq a \leq k$. Let $\vec{u} = u_1, u_2, \ldots, u_k$ and $\vec{v} = v_1, v_2, \ldots, v_k$. Suppose that A has k or ore nontrivial invariant factors, that the first k invariant factors are $f_1, f_2, \ldots, f_k$, and that $r_1, r_2, \ldots, r_k$ are as defined at lines (16) and (17).*

Let h be an integer such that $1 \leq h \leq k-1$. Then, if a is an integer such that $1 \leq a \leq \lfloor r_h/k\rfloor - \Delta_{n,k} - 1$,

$$\mathsf{E}[\mathsf{xnull}_L(\mathcal{H}_{\vec{u},\vec{v},a,a+\Delta_{n,k}})] \leq 1 + q^{(2-\Delta_{n,k})\cdot k}\cdot(f(h,k) + f(h,k)^2\cdot q^{1-k})$$

and

$$\mathsf{E}[\mathsf{xnull}_R(\mathcal{H}_{\vec{u},\vec{v},a+\Delta_{n,k},a})] \leq 1 + q^{(2-\Delta_{n,k})\cdot k} \cdot (f(h,k) + f(h,k)^2 \cdot q^{1-k})$$

as well, where $f(h,k)$ is as given in Lemma 5.1.

Summing failure probabilities one can establish that there are at most $\lfloor r_{k-1}/k \rfloor - 1$ stages of the Lanczos phase with probability at least $1 - 2n^{-c}$, if $\Delta_{n,k}$ is as shown at line (15), above. In this case $\ell \geq r_{k-1} - (\Delta_{n,k} + 3)k$ at the end of the Lanczos phase, so that

$$m \leq r_k - r_{k-1} + (\Delta_{n,k} + 3)k \leq \deg(f_k) + (\Delta_{n,k} + 3)k.$$

It follows that, with high probability, the algorithm is also efficient if the input matrix has k or more nontrivial invariant factors but the degree of the k^{th} invariant factor is small. In particular, the asymptotic bound on performance match those given in the remark following Theorem 2.4, above.

7. FUTURE DIRECTIONS

While the design of the algorithm described in Section 2 is now complete, an implementation is not yet available.

Variants of the algorithm might merit consideration. For example, one might continue a modified Lanczos phase after Invariant #5 is violated by looking for linear dependencies in the set of vectors $u_{r,s}$ (respectively, $v_{r,s}$ and reducing the block size on the left or right accordingly. It is possible that such an algorithm would have better expected performance on arbitrary input matrices $A \in \mathsf{F}_q^{n\times n}$ than the one presented here. That noted, it seems unlikely that either the description of such an algorithm or its analysis would be "simple."

It is also possible that the rectangular block Lanczos algorithm of [3] can now be improved by removing unnecessary orthogonalizations in light of the analysis of the algorithm outlined above.

It seems likely that the results of Section 6 can be extended. In particular, it appears that the length of the elimination phase (and value of m) is also logarithmic in n, with high probability, if the the k^{th} invariant factor has high degree but has a small number of distinct irreducible factors in $\mathsf{F}_q[x]$ A proof of this is in progress.

Unfortunately the techniques used here are of little help in other cases — the bounds that one can obtain for the expected "exponential nullities" of matrices, when the k^{th} invariant factor is an arbitrary matrix in $\mathsf{F}_q[x]$, are too high to be of value. It is also possible (and provable) that matrices $C_{f_1,f_2,\ldots,f_k}$ such that $f_1 = f_2 = \cdots = f_k = f$, for an arbitrary polynomial $f \in \mathsf{F}_q[x]$ with degree at most n/k, are a "hardest case" here — that is, if the algorithm's performance on such matrices is good with high probability then it will also work well on arbitrary matrices $A \in \mathsf{F}_q^{n\times n}$. An attempt to develop a more precise statement along the above lines and prove it is also in progress.

All that said, there is also the question of whether conditioning is required, at all, when one wishes to use a block Lanczos algorithm — or block Wiedemann algorithm incorporating some form of "early termination" — to solve a system of linear equations, when the minimal polynomial of the input matrix is not divisible by x^2, or to obtain a nonzero element of the null space of a singular input matrix. I presently suspect that conditioning is *not* needed here, but (to my knowledge) the question remains open.

8. REFERENCES

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

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

[3] W. Eberly. Yet another block Lanczos algorithm: How to simplify the computation and reduce reliance on preconditioners in the small field case. In *Proceedings, 2010 International Symposium on Symbolic and Algebraic Computation*, pages 289–296, 2010.

[4] W. Eberly. Sparse matrix computations over small finite fields: A simpler block Lanczos algorithm and its analysis. Technical Report 2013–1036–03, Department of Computer Science, University of Calgary, 2013. Available online at `http://www.cpsc.ucalgary.ca/~eberly/simple_lanczos.pdf`.

[5] W. Eberly and B. Hovinen. Bounding the nullities of random block Hankel matrices: An alternative approach. Technical Report 2005–779–10, Department of Computer Science, University of Calgary, 2005. Available online at `http://www.cpsc.ucalgary.ca/~eberly/block_termination_report.pdf`.

[6] F. R. Gantmacher. *The Theory of Matrices*, volume one. Chelsea Publishing Company, second edition, 1959.

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

[8] B. Hovinen and W. Eberly. A reliable block Lanczos algorithm over small finite fields. In *Proceedings, 2005 International Symposium on Symbolic and Algebraic Computation*, pages 177–184, 2005.

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

[10] P. Montgomery. A block Lanczos algorithm for finding dependences over GF(2). In *Lecture Notes in Computer Science*, volume 921, pages 106–120. Springer-Verlag, 1995.

[11] G. Villard. A study of Ccoppersmith's block Wiedemann algorithm using matrix polynomials. Technical Report 975, LMC-IMAG, 1997.

[12] D. Wiedemann. Solving sparse linear equations over finite fields. *IEEE Transactions on Information Theory*, 33:54–62, 1986.

APPENDIX

A. DETAILS OF THE ALGORITHM

I regret that I have yet to discover and provide a succinct proof of the correctness of a "block Lanczos algorithm."

With that noted, Lemma A.4 really is "key" to this and worth consideration as time allows: It is used to establish Lemma A.5–A.7, which establish that the orthogonalizations used in the initialization of vectors $u_{i,s}$ and $v_{i,s}$ (for $i \geq 1$) are sufficient, Lemma A.12, which is useful in showing that the orthogonalizations in matching rounds of the Lanczos phase are adequate, Lemma A.24, which establishes that Invariants #8 and 11 hold after each stage of the elimination

phase, and Lemmas A.31 and A.32, which imply Lemma 2.5.

Finally, I will apologize for any typographical errors and poor writing found in these appendices. The online technical report [4] is currently just a copy of this submission, including the appendices. While I do not intend to change the first eight pages of this I hope to post one or more updates that improve the appendices following the conference submission deadline.

A.1 A More Detailed Description of the Lanczos Phase

A.1.1 Data Structures

Once again, consider an integer $i \geq 0$ such that the Lanczos phase of the computation includes at least $i+1$ stages (ending with stage i). For $0 \leq j \leq i$, let ℓ_j be the length ℓ of the sequences shown at lines (1) and (2) at the end of stage j of the Lanczos phase of the algorithm.

The algorithm can be implemented to make use of the following data structures.

- An array V_L, with indices (r, s) such that $0 \leq r \leq 2 \cdot \Delta_{n,k} + 3$ and $1 \leq s \leq k$, can be used to store the vectors $u_{a,b}$ that are currently required. In particular, at the end of stage i of the Lanczos phase, $u_{a,b}$ will be stored at location $V_L[a \bmod 2 \cdot \Delta_{n,k} + 4, b]$ for all integers a and b such that $\max(0, i-2\cdot\Delta_{n,k}-3) \leq a \leq i$ and $1 \leq b \leq k$.
- An array V_R, with indices (r, s) such that $0 \leq r \leq 2\cdot\Delta_{n,k}+3$ and $1 \leq s \leq k$, can be used to store the vectors $v_{a,b}$ that are currently required — and is organized as described for V_L above.
- An array V_P, with indices (r, s) such that $0 \leq r \leq 2\cdot\Delta_{n,k}+3$ and $1 \leq s \leq k$, can be used to store the vectors $w_{a,b}$ that are currently required — and is organized as described for V_L, above, as well.
- A linked list M will store pairs of indices of "matched" vectors. In particular, at the end of stage i, M will include an entry $((r, s), (r', s'))$ whenever $i-2\cdot\Delta-3 \leq r, r' \leq i$, $1 \leq s, s' \leq k$, $\mu_h = u_{r,s}$ and $\nu_h = v_{r',s'}$ for an integer h such that $1 \leq h \leq \ell$.
- A pair of arrays U_L and U_R with indices r such that $0 \leq r \leq \Delta_{n,k}$ will be used to store the indices of unmatched vectors. In particular, at the end of stage i the indices of all unmatched vectors $u_{s,t}$ (respectively, $v_{s,t}$) such that

  $$\max(0, i - \Delta_{n,k} + 1) \leq s \leq i$$

  will be stored on the linked list $U_L[s \bmod \Delta_{n,k} + 1]$ (respectively, $U_R[s \bmod \Delta_{n,k} + 1]$).

A.1.2 Two "Orthogonalization" Subroutines

A routine orthogL will be used perform updates as shown at line (7). This routine receives the following inputs.

- A matrix $\mathcal{M}_L \in \mathsf{F}_q^{n\times m}$ whose columns are the vectors $\mu_{\sigma_1}, \mu_{\sigma_2}, \ldots, \mu_{\sigma_m}$, for a nonnegative integer $m \leq \ell$, and for distinct integers $\sigma_1, \sigma_2, \ldots, \sigma_m$ such that $1 \leq \sigma_a \leq \ell$ for $1 \leq a \leq m$.
- A matrix $\mathcal{M}_R \in \mathsf{F}_q^{n\times m}$ whose columns are the vectors $\nu_{\sigma_1}, \mu_{\sigma_2}, \ldots, \nu_{\sigma_m}$ for the nonnegative integer m and integers $\sigma_1, \sigma_2, \ldots, \sigma_m$ as above.
- An integer h such that $0 \leq h_L \leq \Delta_{n,k} + 1$.

Pseudocode is as follows; to simplify this, vectors $u_{r,s}$ (respectively, $v_{r,s}$ and $w_{r,s}$) are named instead of their locations $V_L[r \bmod 2 \cdot \Delta_{n,k} + 4, s]$ ($V_R[r \bmod 2 \cdot \Delta_{n,k} + 4, s]$ and $V_P[r \bmod 2 \cdot \Delta_{n,k} + 4, s]$, respectively) in the data structures that have been described above.

procedure orthogL($\mathcal{M}_L$, $\mathcal{M}_R$, h)

1. Set t to be the length of the linked list $U_L[h]$ and, if $U_L[h]$ has entries

   $$(r, s_1), (r, s_2), \ldots, (r, s_t),$$

   set $\mathcal{U} \in \mathsf{F}_q^{n\times t}$ to be the matrix with columns

   $$u_{r,s_1}, u_{r,s_2}, \ldots, u_{r,s_t}.$$

2. if $(m > 0$ and $t > 0)$ then
3. $\quad D := \mathcal{M}_R^T \cdot \mathcal{U} \in \mathsf{F}_q^{m\times t}$
4. $\quad \mathcal{U} := \mathcal{U} - \mathcal{M}_L \cdot D$
5. $\quad$ for $1 \leq a \leq t$ do
6. $\quad\quad$ Set u_{r,s_a} to be column a of $\mathcal{U}$
   $\quad$ end for
   end if

end procedure

Lemma A.1, below, follows by inspection of the code and the fact that if Invariant #2 is satisfied then $\mathcal{M}_L^T \cdot \mathcal{M}_R = I_m$ so that, if D is as shown at line 3 above, then (before line 4)

$$\begin{aligned}(\mathcal{U} - \mathcal{M}_L \cdot D)^T \cdot \mathcal{M}_R &= \mathcal{U}^T \cdot \mathcal{M}_R - D^T \cdot (\mathcal{M}_L^T \cdot \mathcal{M}_R) \\ &= \mathcal{U}^T \cdot \mathcal{M}_R - D^T \\ &= 0,\end{aligned}$$

and

$$(\mathcal{M}_L \cdot D)^T \cdot \nu_\beta = D^T \cdot \mathcal{M}_L^T \cdot \nu_\beta = 0$$

if $1 \leq \beta \leq \ell$ and $\beta \notin \{\sigma_1, \sigma_2, \ldots, \sigma_m\}$.

LEMMA A.1. *Suppose that the procedure* orthogL *is executed when Invariant #2 is satisfied and that the vectors u_{r,s_a} are unmatched at the time when this procedure is called for $1 \leq a \leq t$. Then, on termination,*

$$u_{r,s_a}^T \cdot \nu_{\sigma_b} = 0$$

for $1 \leq a \leq t$ and $1 \leq b \leq m$. If β is an integer such that $\leq \beta \leq \ell$ and

$$\beta \notin \{\sigma_1, \sigma_2, \ldots, \sigma_m\},$$

then $u_{r,s_a}^T \cdot \nu_\beta = 0$ on termination if and only if $u_{r,s_a}^T \cot \nu_\beta = 0$ when the procedure was executed as well. None of the vectors $u_{a,b}$ or $v_{a,b}$ (such that $0 \leq a \leq i$ and $1 \leq b \leq k$) except for the vectors $u_{r,s_1}, u_{r,s_2}, \ldots, u_{r,s_t}$ are modified by the execution of this procedure.

The procedure can be implemented to use $O(mtn)$ operations over F_q using standard arithmetic.

A similar routine orthogR will be used to perform updates as shown at line (8) while also ensuring that Invariant A, above, remains satisfied. This routine receives the following inputs.

- A matrix $\mathcal{M}_L \in \mathsf{F}_q^{n\times m}$ whose columns are the vectors $\mu_{\sigma_1}, \mu_{\sigma_2}, \ldots, \mu_{\sigma_m}$, for a nonnegative integer $m \leq \ell$, and for distinct integers $\sigma_1, \sigma_2, \ldots, \sigma_m$ such that $1 \leq \sigma_a \leq \ell$ for $1 \leq a \leq m$.

- A matrix $\mathcal{M}_R \in \mathsf{F}_q^{n\times m}$ whose columns are the vectors $\nu_{\sigma_1}, \nu_{\sigma_2}, \ldots, \nu_{\sigma_m}$ for the nonnegative integer m and integers $\sigma_1, \sigma_2, \ldots, \sigma_m$ as above.
- A matrix $\mathcal{M}_R^{pre} \in \mathsf{F}_q^{n\times m}$ whose columns are vectors

  $$\omega_{\sigma_1}, \omega_{\sigma_2}, \ldots, \omega_{\sigma_m} \in \mathsf{F}_q^{n\times 1}$$

  such that $A \cdot \omega_{\sigma_a} = \nu_{\sigma_a}$ for $1 \leq a \leq m$ — so that $A \cdot \mathcal{M}_R^{pre} = \mathcal{M}_R$.
- An integer h such that $0 \leq h_L \leq \Delta_{n,k} + 1$.

Pseudocode for this routine is as follows.

procedure orthogR($\mathcal{M}_L$, $\mathcal{M}_R$, $\mathcal{M}_R^{pre}$, h)

1. Set t to be the length of the linked list $U_R[h]$ and, if $U_R[h]$ has entries

   $$(r, s_1), (r, s_2), \ldots, (r, s_t),$$

   set $\mathcal{U} \in \mathsf{F}_q^{n\times t}$ to be the matrix with columns

   $$v_{r,s_1}, v_{r,s_2}, \ldots, v_{r,s_t}$$

   and set $\mathcal{W} \in \mathsf{F}_q^{n\times t}$ to be the matrix with columns

   $$w_{r,s_1}, w_{r,s_2}, \ldots, w_{r,s_t}$$

   — so that $A \cdot \mathcal{W} = \mathcal{U}$.
2. if ($m > 0$ and $t > 0$) then
3. $\quad D := \mathcal{M}_L^T \cdot \mathcal{U} \in \mathsf{F}_q^{m\times t}$
4. $\quad \mathcal{U} := \mathcal{U} - \mathcal{M}_R \cdot D$
5. $\quad \mathcal{W} := \mathcal{W} - \mathcal{M}_R^{pre} \cdot D$
6. $\quad$ for $1 \leq a \leq t$ do
7. $\quad\quad$ Set u_{r,s_a} to be column a of $\mathcal{U}$
8. $\quad\quad$ Set w_{r,s_a} to be column a of $\mathcal{W}$

   $\quad$ end for

   end if

end procedure

The proof of Lemma A.2, below, is almost the same as that of Lemma A.1. It also depends on the assumption that $A \cdot \mathcal{M}_R^{pre} = \mathcal{M}_R$, so that $A \cdot \mathcal{W} = \mathcal{U}$ after the executions of steps 4 and 5 if this relationship held before this.

LEMMA A.2. *Suppose that the procedure* orthogR *is executed with Invariants #2 and #6 satisfied and with* $A \cdot \mathcal{M}_R^{pre} = \mathcal{M}_R$*, and that the vectors* v_{r,s_a} *are unmatched at the time when that this procedure is called for* $1 \leq a \leq t$*. Then, on termination,*

$$\mu_{\sigma_b}^T \cdot v_{r,s_a} = 0$$

for $1 \leq a \leq t$ *and* $1 \leq b \leq m$*, and Invariant #6 is satisfied once again. If* β *is an integer such that* $1 \leq \beta \leq \ell$ *and*

$$\beta \notin \{\sigma_1, \sigma_2, \ldots, \sigma_m\}$$

then $\mu_\beta^T \cdot v_{r,s_a} = 0$ *on termination if and only if* $\mu_\beta^T \cdot v_{r,s_a} = 0$ *when the procedure was executed as well. None of the vectors* $u_{a,b}$ *or* $v_{a,b}$ *(such that* $0 \leq a \leq i$ *and* $1 \leq b \leq k$*) except for the vectors* $v_{r,s_1}, v_{r,s_2}, \ldots, v_{r,s_t}$ *are modified by the execution of this procedure.*

The procedure can be implemented to use $O(mtn)$ *operations over* F_q *using standard arithmetic.*

A.1.3 A "Matching" Subroutine

A routine match will be used to matched a given pair of sequences of unmatched vectors, perform updates as shown at lines (4) and (5), extend the sequences at lines (1) and (2), and ensure that Invariant #7 is satisfied once again. This routine receives as inputs a pair of integers h_L and h_R such that $0 \leq h_L, h_R \leq \Delta_{n,k} + 1$ and attempts to match vectors $u_{r,s}$ such that (r, s) is an index in the linked list $U_L[h_L]$ with vectors $v_{r',s'}$ such (r', s') is an entry in the linked list $U_R[h_R]$.

The routine will return a sequence of three matrices, $\mathcal{M}_L$, $\mathcal{M}_R$, and $\mathcal{M}_R^{pre}$: The columns of $\mathcal{M}_L$ and $\mathcal{M}_R$ will be the vectors μ_h (respectively, ν_h) that have been added to the sequence at line (1) (respectively, at line (2)), while $\mathcal{M}_R^{pre}$ will be a matrix such that $A \cdot \mathcal{M}_R^{pre} = \mathcal{M}_R$.

Pseudocode for this routine is as follows.

procedure match(h_L, h_R)

1. Set t_L and t_R to be the lengths of the linked lists $U_L[h_L]$ and $U_R[h_R]$ respectively. If $U_L[h_L]$ has entries

   $$(r, s_1), (r, s_2), \ldots, (r, s_{t_L}),$$

   set $\widehat{\mathcal{K}}_L \in \mathsf{F}_q^{n\times t_L}$ to be the matrix with columns

   $$u_{r,s_1}, u_{r,s_2}, \ldots, u_{r,s_{t_L}}.$$

   Similarly, if $U_R[h_R]$ has entries

   $$(r', s_1'), (r', s_2') \ldots, (r', s_{t_R}'),$$

   set $\mathcal{K}_R \in \mathsf{F}_q^{n\times t_R}$ to be the matrix with columns

   $$v_{r',s_1'}, v_{r',s_2'}, \ldots, v_{r',s_{t_R}'}$$

   and set $\mathcal{K}_R^{pre} \in \mathsf{F}_q^{n\times t_R}$ to be the matrix with columns

   $$w_{r',s_1'}, w_{r',s_2'}, \ldots, w_{r',s_{t_R}'}$$

   — so that $A \cdot \mathcal{K}_R^{pre} = \mathcal{K}_R$.
2. if ($t_L > 0$ and $t_R > 0$) then
3. $\quad \mathcal{H} := \widehat{\mathcal{K}}_L^T \cdot \mathcal{K}_R \in \mathsf{F}_q^{t_L\times t_R}$
4. $\quad$ Compute the rank $\widehat{r}$ of $\mathcal{H}$ as well as a sequence of integer indices $\sigma_1, \sigma_2, \ldots, \sigma_{\widehat{r}}$ where

   $$1 \leq \sigma_1 < \sigma_2 < \cdots < \sigma_{\widehat{r}} \leq t_L$$

   along with a sequence of integer indices $\tau_1, \tau_2, \ldots, \tau_{\widehat{r}}$ where

   $$1 \leq \tau_1 < \tau_2 < \cdots < \tau_{\widehat{r}} \leq t_R,$$

   so that the submatrix $\widehat{\mathcal{H}} \in \mathsf{F}_q^{\widehat{r}\times\widehat{r}}$ of $\mathcal{H}$ that includes rows $\sigma_1, \sigma_2, \ldots, \sigma_{\widehat{r}}$ and columns $\tau_1, \tau_2, \ldots, \tau_{\widehat{r}}$ is a maximal nonsingular submatrix of $\mathcal{H}$.
5. $\quad$ (In any way that is convenient) compute invertible matrices $X_L, X_R \in \mathsf{F}_q^{\widehat{r}\times\widehat{r}}$ such that $X_R \cdot X_L = \widehat{\mathcal{H}}^{-1}$, for $\widehat{\mathcal{H}}$ as above.
6. $\quad$ Set $\widetilde{\mathcal{K}_L} \in \mathsf{F}_q^{n\times\widehat{r}}$ to be the matrix with columns

   $$u_{r,s_{\sigma_1}}, u_{r,s_{\sigma_2}}, \ldots, u_{r,s_{\sigma_{\widehat{r}}}},$$

   set $\widetilde{\mathcal{K}_R} \in \mathsf{F}_q^{n\times\widehat{r}}$ to be the matrix with columns

   $$v_{r',s_{\tau_1}'}, v_{r',s_{\tau_2}'}, \ldots, v_{r',s_{\tau_{\widehat{r}}}'}$$

   and set $\widetilde{\mathcal{K}}_R^{pre} \in \mathsf{F}_q^{n\times\widehat{r}}$ to be the matrix with columns

   $$w_{r',s_{\tau_1}'}, w_{r',s_{\tau_2}'}, \ldots, w_{r',s_{\tau_{\widehat{r}}}'}$$

   — so that $A \cdot \widetilde{\mathcal{K}}_R^{pre} = \widetilde{\mathcal{K}_R}$ and $\widetilde{\mathcal{K}_L}^T \cdot \widetilde{\mathcal{K}_R}$ is equal to the matrix $\widehat{\mathcal{H}}$ mentioned in the previous steps.

7. $\mathcal{M}_L := \widetilde{\mathcal{K}_L} \cdot X_L^T$;
8. $\mathcal{M}_R := \widetilde{\mathcal{K}_R} \cdot X_R$; $\quad \mathcal{M}_R^{pre} := \widetilde{\mathcal{K}}_R^{pre} \cdot X_R$
9. for $1 \le g \le \widehat{r}$ do
10. Set $u_{r,s_{\sigma_g}}$ to be column g of $\mathcal{M}_L$
11. Set $v_{r',s'_{\tau_g}}$ to be column g of $\mathcal{M}_R$
12. Set $w_{r',s'_{\tau_g}}$ to b column g of $\mathcal{M}_R^{pre}$
13. Append entry $((r, s_{\sigma_g}), (r', s'_{\tau_g}))$ onto the list M

    end for
14. Remove the entries

$$(r, s_{\sigma_1}), (r, s_{\sigma_2}), \ldots, (r, s_{\sigma_g})$$

    from the list $U_L[h_L]$
15. Remove the entries

$$(r', s'_{\tau_1}), (r', s'_{\tau_2}), \ldots, (r', s'_{\tau_g})$$

    from the list $U_R[h_R]$
16. $z := \mathcal{M}_L^T \cdot \sigma \in \mathsf{F}_q^{\widehat{r} \times 1}$
17. $\chi := \chi + \mathcal{M}_R^{pre} \cdot z$; $\quad \rho := \rho - \mathcal{M}_R \cdot z$
18. return $(\mathcal{M}_L, \mathcal{M}_R, \mathcal{M}_R^{pre})$

    end if

end procedure

As discussed in Section 2, if $\mathcal{M}_L$ and $\mathcal{M}_R$ are computed as shown at lines 3–8, above, then $\mathcal{M}_L^T \cdot \mathcal{M}_R = I_{\widehat{r}}$. Furthermore one can see by inspection of the code that the columns of $\mathcal{M}_L$ are linear combinations of

$$u_{r,s_1}, u_{r,s_2}, \ldots, u_{r,s_{t_L}}$$

and that the columns of $\mathcal{M}_R$ are linear combinations of

$$v_{r',s'_1}, v_{r',s'_2}, \ldots, v_{r',s'_{t_R}},$$

so that Invariant #2 will hold once again after the execution of this procedure. An inspection of line 16 and 17 should confirm that $\mathcal{M}_L^T \cdot (A \cdot \chi + \rho)$ is unchanged after the execution of these lines but that $\mathcal{M}_L^T \cdot A \cdot \chi = \mathcal{M}_L^T \cdot \sigma = \mathcal{M}_L^T \cdot (A \cdot w + b)$, as needed to re-establish Invariant #7.

The following claim can now be verified by inspection of the code.

LEMMA A.3. *Suppose procedure* match *is executed with Invariants #2, 6 and 7 satisfied and when* $u_{r,s_a}^T \cdot \nu_b = 0$ *for* $1 \le a \le t_L$ *and* $1 \le b \le \ell$ *and* $\mu_b^T \cdot v_{r',s'_c}$ *for* $1 \le b \le \ell$ *and* $1 \le c \le t_R$*, for the vectors*

$$u_{r,s_1}, u_{r,s_2}, \ldots, u_{r,s_{t_L}}$$

and

$$v_{r',s'_1}, v_{r',s'_2} \ldots, v_{r',s'_{t_R}}$$

as shown in step 1.

Then Invariants #2, 6 and 7 are satisfied again on termination of the procedure.

Furthermore, if $\widehat{\mathcal{K}}_L \in \mathsf{F}_q^{n \times t_L}$ *(respectively,* $\mathcal{K}_R \in \mathsf{F}_q^{n \times t_R}$*) is the matrix whose columns are the vectors* $u_{r,s}$ *(respectively,* $v_{r',s}$*) such that* $1 \le s \le k$ *and these vectors were unmatched when the procedure was executed, then on termination, the sequences of vectors at lines* (1) *and* (2) *have been extended with vectors*

$$\mu_{\ell+1}, \mu_{\ell+2}, \ldots, \mu_{\ell+m} \tag{19}$$

and

$$\nu_{\ell+1}, \nu_{\ell+2}, \ldots, \nu_{\ell+m} \tag{20}$$

respectively, where $\mu_{\ell+a}$ *(respectively,* $\nu_{\ell+a}$*) is a linear combination of the columns of* $\widehat{\mathcal{K}}_L$ *(respectively,* $\mathcal{K}_R$*) for* $1 \le a \le b$ *and where* m *is the rank of the matrix* $\widehat{\mathcal{K}}_L^T \cdot \mathcal{K}_R$*. It returns, as output, matrices* $\mathcal{M}_L, \mathcal{M}_R, \mathcal{M}_R^{pre} \in \mathsf{F}_q^{n \times m}$ *such that the columns of* $\mathcal{M}_L$ *are the vectors at line* (19)*, the columns of* $\mathcal{M}_R$ *are the vectors at line* (20)*, and* $A \cdot \mathcal{M}_R^{pre} = \mathcal{M}_R$*.*

None of the vectors $u_{a,b}$ *or* $v_{a,b}$ *have been modified by this procedure except for the* m *pairs of vectors that have been newly matched.*

The procedure can be implemented to use $O(n \cdot t_L \cdot t_R)$ *operations over* F_q *using standard arithmetic,*

As noted in Section 2 it is acceptable to use a version of this that also modifies the vectors

$$u_{r,s_1}, u_{r,s_2}, \ldots, u_{r,s_L}$$

and

$$v_{r',s'_1}, v_{r',s'_2} \ldots, v_{r',s'_L}$$

provided that the column spaces spanned by each of the above sequences is unchanged. The above lemma would need to be modified (to allow for such a change) if such a "matching" procedure was used, but it could still be used to establish the results that follow.

A.1.4 The Main Method

Pseudocode for the Lanczos phase is now as follows. As noted above, this version of the algorithm receives a matrix $A \in \mathsf{F}_q^{n \times n}$ and a vector b as input and attempts to provide a solution for the system $Ax = b$.

// Initialization

1. Select vectors

$$u_1, u_2, \ldots, u_k, w, w_2, \ldots, w_k$$

   uniformly and independently from $\mathsf{F}_q^{n \times 1}$. Set w_1 to be $A \cdot w + b$.
2. for $1 \le j \le k$ do
3. $v_j := A \cdot w_j$

   end for
4. Initialize the list M to be empty
5. for $0 \le r \le \Delta_{n,k}$ do
6. Initialize the lists $U_L[r]$ and $U_R[r]$ to be empty

   end for
7. $\chi := 0 \in \mathsf{F}_q^{n \times 1}$; $\rho := \sigma := w_1$

// Stage 0

8. for $1 \le j \le k$ do
9. $u_{0,j} := u_j$; $v_{0,j} := v_j$; $w_{0,j} := w_j$

   end for
10. Insert the entries $(0,1), (0,2), \ldots, (0,k)$ into each of the lists $U_L[0]$ and $U_R[0]$
11. $(\mathcal{M}_L, \mathcal{M}_R, \mathcal{M}_R^{pre}) :=$ match$(0,0)$
12. orthogL$(\mathcal{M}_L, \mathcal{M}_R, 0)$
13. orthogR$(\mathcal{M}_L, \mathcal{M}_R, \mathcal{M}_R^{pre}, 0)$

// Stage i for $i \ge 1$

14. $i := 1$
15. while ($i < \Delta_{n,k}$ or ($U_L[i \bmod \Delta_{n,k} + 1]$ and $U_R[i \bmod \Delta_{n,k} + 1]$ are both empty)) do

    // $i \bmod (\Delta_{n,k} + 1) = i - \Delta_{n,k} - 1 \bmod (\Delta_{n,k} + 1)$,

    // so Invariant #5 was satisfied at the end of

    // the previous round.

16. Remove all entries $((r,s),(r',s'))$ such that either $r \le i - 2\cdot\Delta_{n,k} - 3$ or $r' \le i - 2\cdot\Delta_{n,k} - 3$ from the list M
17. for $1 \le j \le k$ do
18. $u_{i,j} := A^T \cdot u_{i-1,j}$; $v_{i,j} := A \cdot v_{i-1,j}$; $w_{i,j} := v_{i-1,j}$
end for
19. Append the entries

$$(i,1),(i,2),\ldots,(i,k)$$

onto each of the lists $U_L[i \bmod \Delta_{n,k}+1]$ and $U_R[i \bmod \Delta_{n,k}+1]$
20. if (M is nonempty) then
21. If the entries of M are

$$((r_1,s_1),(r'_1,s'_1)),((r_2,s_2),(r'_2,s'_2)),\ldots,((r_h,s_h),(r'_h,s'_h))$$

then set $\mathcal{M}_L \in \mathsf{F}_q^{n\times h}$ to be the matrix with columns

$$u_{r_1,s_1}, u_{r_2,s_2}, \ldots, u_{r_h,s_h},$$

set $\mathcal{M}_R \in \mathsf{F}_q^{n\times h}$ to be the matrix with columns

$$v_{r'_1,s'_1}, v_{r'_2,s'_2}, \ldots, v_{r'_h,s'_h},$$

and set $\mathcal{M}_R^{pre} \in \mathsf{F}_q^{n\times h}$ to be the matrix with columns

$$w_{r'_1,s'_1}, w_{r'_2,s'_2}, \ldots, w_{r'_h,s'_h}$$

— so that $A\cdot\mathcal{M}_R^{pre} = \mathcal{M}_R$.
22. orthogL$(\mathcal{M}_L,\mathcal{M}_R,i \bmod \Delta_{n,k}+1)$
23. orthogR$(\mathcal{M}_L,\mathcal{M}_R,\mathcal{M}_R^{pre},i \bmod \Delta_{n,k}+1)$
end if
24. for $j = 0,1,2,\ldots,\Delta_{n,k}-1$ do
// Matching Round $\#2j+1$
25. $(\mathcal{M}_L,\mathcal{M}_R,\mathcal{M}_R^{pre}) :=$ match$(i-\Delta_{n,k}+j \bmod \Delta_{n,k}+1, i \bmod \Delta_{n,k}+1)$
26. for $j \le h \le \Delta_{n,k}$ do
27. orthogL$(\mathcal{M}_L,\mathcal{M}_R,i-\Delta_{n,k}+h \bmod \Delta_{n,k}+1)$
end for
28. orthogR$(\mathcal{M}_L,\mathcal{M}_R,\mathcal{M}_R^{pre},i \bmod \Delta_{n,k}+1)$
// Matching Round $\#2j+2$
29. $(\mathcal{M}_L,\mathcal{M}_R,\mathcal{M}_R^{pre}) :=$ match$(i \bmod \Delta_{n,k}+1, i-\Delta_{n,k}+j \bmod \Delta_{n,k}+1)$
30.. for $j \le h \le \Delta_{n,k}$ do
31. orthogR$(\mathcal{M}_L,\mathcal{M}_R,\mathcal{M}_R^{pre}, i-\Delta_{n,k}+h \bmod \Delta_{n,k}+1)$
end for
32. orthogL$(\mathcal{M}_L,\mathcal{M}_R,i \bmod \Delta_{n,k}+1)$
end for
// Matching Round $\#2\cdot\Delta_{n,k}+1$
33. $(\mathcal{M}_L,\mathcal{M}_R,\mathcal{M}_R^{pre}) :=$ match$(i \bmod \Delta_{n,k}+1, i \bmod \Delta_{n,k}+1)$
34. orthogL$(\mathcal{M}_L,\mathcal{M}_R,i \bmod \Delta_{n,k}+1)$
35. orthogR$(\mathcal{M}_L,\mathcal{M}_R,\mathcal{M}_R^{pre},i \bmod \Delta_{n,k}+1)$
end while

A.2 On the Correctness and Efficiency of the Lanczos Phase

A.2.1 A Key Lemma

The next lemma is instrumental in proving the orthogonalizations included in the Lanczos phase of the algorithm are sufficient to establish Invariants #1–5 and for the proof of Lemma 2.5.

LEMMA A.4. *Let i be an integer such that $i \ge 0$ and there are at least $i+1$ stages of the Lanczos phase. Then the following properties are satisfied at the end of stage i.*

(a) *Invariant #1 is satisfied.*

(b) *Suppose j is an integer such that $1 \le j \le i+1$ and $\mathcal{M}_{L,i,j}, \mathcal{M}_{R,i,j} \in \mathsf{F}_q^{n\times jk}$ be the matrices with columns $u_{r,s}$ and $v_{r,s}$, respectively, for $0 \le r \le j-1$ and $1 \le s \le k$.*
Then there exist nonsingular matrices $X_{i,j}, Y_{i,j} \in \mathsf{F}_q^{jk\times jk}$ such that

$$\mathcal{M}_{L,i,j} = \widehat{\mathcal{K}}_{\vec{u},j}\cdot X_{i,j} \quad \text{and} \quad \mathcal{M}_{R,i,j} = \mathcal{K}_{\vec{v},j}\cdot Y_{i,j} \tag{21}$$

where matrices $\widehat{\mathcal{K}}_{\vec{u},j}$ and $\mathcal{K}_{\vec{v},j}$ are as defined in Subsection 2.5.

(c) *If r and s are integers such that $0 \le r \le i-1$ and $1 \le s \le k$ then there exist elements $\alpha_{r,s,a,b}$ and $\beta_{r,s,a,b}$ of F_q, for $0 \le a \le r+1$ and $1 \le b \le k$, such that*

$$A^T\cdot u_{r,s} = \sum_{a=0}^{r+1}\sum_{b=1}^{k}\alpha_{r,s,a,b}\cdot u_{a,b} \tag{22}$$

and

$$A\cdot v_{r,s} = \sum_{a=0}^{r+1}\sum_{b=1}^{k}\beta_{r,s,a,b}\cdot v_{a,b}. \tag{23}$$

PROOF. The claims can be established by induction on i.

Notice first that, since $u_{0,r}$ and $v_{0,s}$ are initialized to be u_r and v_r, respectively, Invariant #1 is certainly satisfied at the *beginning* of stage #0, that is, immediately after these vectors have been defined but before any have been matched.

The equations at line (21) are also satisfied for $j=1$ (the only case to be considered here). In particular, they are satisfied when one chooses both $X_{0,1}$ and $Y_{0,1}$ to be the identity matrix in $\mathsf{F}_q^{k\times k}$.

To continue, one should notice that the updates included in matchings and orthogonalizations are all invertible linear transformations that effectively update the above matrices using updates of the form

$$\mathcal{M}_{L,0,1} := \mathcal{M}_{L,0,1}\cdot X_{update}$$

and

$$\mathcal{M}_{R,0,1} := \mathcal{M}_{R,0,1}\cdot Y_{update}$$

for invertible matrices $X_{update}, Y_{update} \in \mathsf{F}_q^{k\times k}$. Now, if one also updates the matrices $X_{1,0}$ and $X_{1,0}$ by setting

$$X_{1,0} := X_{1,0}\cdot X_{update}$$

and

$$Y_{1,0} := Y_{1,0}\cdot Y_{update},$$

then the equations at line (21) are satisfied after the update if they were satisfied before — and the matrices $X_{1,0}, Y_{1,0} \in \mathsf{F}_q^{k\times k}$ are still nonsingular as well.

It follows by a straightforward induction on the number of updates performed that equations as shown at line (21) at the end of stage #0, as desired. This implies that Invariant #1 is satisfied at the end of stage #0 as well.

Since the claim in part (c) is vacuous when $i = 0$, this establishes the basis.

For the inductive step suppose that $i \geq 0$, there are at least $i+2$ stages of the Lanczos phase, and that the above three properties hold at the end of stage i. It is necessary and sufficient to prove that they hold at the end of stage $i+1$ as well.

It follows by the inductive hypothesis that equations as shown at line (21) hold for $j = i+1$. Now, multiplying both sides of the first equation by A^T and multiplying both sides of the second by A one has that

$$A^T \cdot \mathcal{M}_{L,i,i+1} = A^T \cdot \widehat{\mathcal{K}}_{\vec{u},i+1} \cdot X_{i,i+1}.$$

Now consider the columns of the matrix $A^T \cdot \mathcal{M}_{L,i,i+1}$: These are either vectors $A^T \cdot u_{i,s}$ for $1 \leq s \leq k$ or they are vectors $A^T \cdot u_{r,s}$ for $0 \leq r \leq i-1$ and $1 \leq s \leq k$. Now, $A^T \cdot u_{i,s}$ is the initial value of $u_{i+1,s}$ for $1 \leq s \leq k$ and it follows by part (c) of the inductive hypothesis that each vector $A^T \cdot u_{r,s}$ such that $0 \leq r \leq i-1$ and $1 \leq s \leq k$ is a linear combination of the vectors $u_{a,b}$ such that $0 \leq a \leq i$ and $1 \leq b \leq k$ at the end of stage i. Consequently each of the columns of the matrix $A^T \cdot \mathcal{M}_{L,i,i+1}$ is a linear combination of the vectors $u_{a,b}$ such that $0 \leq a \leq i+1$ and $1 \leq b \leq k$ — when the values of these vectors immediately after the vectors $u_{i+1,s}$ and $v_{i+1,s}$ have been initialized are considered. Now, since the matrix $X_{i,i+1}$ is invertible it follows that the columns of $A^T \cdot \widehat{\mathcal{K}}_{\vec{u},i+1}$ are linear combinations of these vectors as well. The columns of $\widehat{\mathcal{K}}_{\vec{u},i+1}$ are certainly all linear combinations of these vectors too, since these columns are linear combinations of vectors $u_{a,b}$ such that $0 \leq a \leq i$ and $1 \leq b \leq k$. It follows that the columns of $\widehat{\mathcal{K}}_{\vec{u},i+2}$ are all linear combinations of the vectors $u_{a,b}$ such that $0 \leq a \leq i+1$ and $1 \leq b \leq k$, because each column of $\widehat{\mathcal{K}}_{\vec{u},i+2}$ is either a column of $\widehat{\mathcal{K}}_{\vec{u},i+1}$ or of $A^T \cdot \widehat{\mathcal{K}}_{\vec{u},i+1}$.

On the other hand, since

$$A^T \cdot \mathcal{M}_{L,i,i+1} = A^T \cdot \widehat{\mathcal{K}}_{\vec{u},i+1} \cdot X_{i,i+1}$$

since $A^T \cdot u_{i,s}$ is a column of $A^T \cdot \mathcal{M}_{L,i,i+1}$ and since every column of $A^T \cdot \widehat{\mathcal{K}}_{\vec{u},i+1}$ is also a column of $\widehat{\mathcal{K}}_{\vec{u},i+2}$, each vector $u_{i+1,s}$ is a linear combination of the columns of $\widehat{\mathcal{K}}_{\vec{u},i+2}$ when the vectors $u_{i+1,s}$ are first initialized. It follows by the inductive hypothesis that each vector $u_{r,s}$ such that $1 \leq r \leq i$ and $1 \leq s \leq k$ is a linear combination of the columns of $\widehat{\mathcal{K}}_{\vec{u},i+2}$ at this point as well, since the columns of $\widehat{\mathcal{K}}_{\vec{u},i+2}$ include all of the columns of $\widehat{\mathcal{K}}_{\vec{u},i+1}$.

The same consideration of vectors $v_{r,s}$ establishes the same property for these vectors, as needed to establish Invariant #1 at the beginning of stage $i+1$, that is, immediately after $u_{i+1,s}$ (respectively, $v_{i+1,s}$) have been set to be $A^T \cdot u_{i,s}$ (respectively, $A \cdot v_{i,s}$) for $1 \leq s \leq k$.

It now suffices to note that all matching and orthogonalization updates in stage $i+1$ that follow are invertible (and argue as in the basis) to conclude that Invariant #1 is satisfied at the *end* of stage $i+1$ as well.

In order to establish part (b) of the claim, let us suppose that the matrix $\widehat{\mathcal{K}}_{\vec{u},i+2}$ has rank s (so that $0 \leq s \leq (i+2) \cdot k$). Notice that, since Invariant #1 is satisfied at the end of stage $\#i+1$, $\mathcal{M}_{L,i+1,i+2}$ has rank s as well and, furthermore, each column of either of these matrices is a linear combination of the columns of the other. Consider the following additional matrices.

- There are permutation matrices
$$P_1, P_2 \in \mathsf{F}_q^{(i+2)\cdot k \times (i+2)\cdot k}$$
such that the first s columns of $\widehat{\mathcal{K}}_{\vec{u},i+2} \cdot P_1$ (respectively, of $\mathcal{M}_{L,i+1,i+2} \cdot P_2$) are linearly independent and such that the remaining $(i+2) \cdot k - s$ columns are linear combinations of the first s.
- Consequently there exist nonsingular upper triangular matrices
$$U_1 = \begin{bmatrix} I_s & X_1 \\ 0 & I_{(i+2)\cdot k} \end{bmatrix} \quad \text{and} \quad U_2 = \begin{bmatrix} I_s & X_2 \\ 0 & I_{(k+2)\cdot k} \end{bmatrix}$$
in $\mathsf{F}_q^{(i+2)\cdot k \times (i+2)\cdot k}$ such that
$$\widehat{\mathcal{K}}_{\vec{u},i+2} \cdot P_1 \cdot U_1 = \begin{bmatrix} A & 0 \end{bmatrix}$$
and
$$\mathcal{M}_{L,i+1,i+2} \cdot P_2 \cdot U_2 = \begin{bmatrix} B & 0 \end{bmatrix}$$
where $A, B \in \mathsf{F}_q^{n \times s}$ each has full rank s (and so that the last $(i+2) \cdot k - s$ columns of each of $\widehat{\mathcal{K}}_{\vec{u},i+2} \cdot P_1 \cdot U_1$ and $\mathcal{M}_{L,i_1,i_2} \cdot P_2 \cdot U_2$ are equal to zero).
- Now, the columns of A are linear combinations of the columns of B, and vice-versa. Consequently there exists a nonsingular matrix $C \in \mathsf{F}_q^{s \times s}$ such that $A \cdot C = B$. Furthermore the matrix
$$\widehat{C} = \begin{bmatrix} C & 0 \\ 0 & I_{(i+2)\cdot k - s} \end{bmatrix} \in \mathsf{F}_q^{(i+2)\cdot k \times (i+2)\cdot k}$$
is also nonsingular — and
$$\widehat{\mathcal{K}}_{\vec{u},i+2} \cdot P_1 \cdot U_1 \cdot \widehat{C} = \mathcal{M}_{L,i+1,i+2} \cdot P_2 \cdot U_2.$$
- We now have that
$$\mathcal{M}_{L,i+1,i+2} = \widehat{\mathcal{K}}_{\vec{u},i+2} \cdot X_{i+1,i+2}$$
for the invertible matrix
$$X_{i+1,i+2} = P_1 \cdot U_1 \cdot \widehat{C} \cdot U_2^{-1} \cdot P_1^{-1} \in \mathsf{F}_q^{(i+2)\cdot k \times (i+2)\cdot k}.$$

It follows by the same argument that there exists a nonsingular matrix $Y_{i+1,i+2} \in \mathsf{F}_q^{(i+2)\cdot k \times (i+2)\cdot k}$ such that

$$\mathcal{M}_{R,i+1,i+2} = \mathcal{K}_{\vec{v},i+2} \cdot Y_{i+1,i+2}$$

as well.

Now, to establish that

$$\mathcal{M}_{L,i+1,j} = \widehat{\mathcal{K}}_{\vec{u},j} \cdot X_{i+1,j} \quad \text{and} \quad \mathcal{M}_{R,i+1,j} = \mathcal{K}_{\vec{v},j} \cdot Y_{i+1,j}$$

for nonsingular matrices $X_{i+1,j}, Y_{i+1,j} \in \mathsf{F}_q^{j\cdot k \times j \cdot k}$ at the end of stage $i+1$ of the Lanczos phase as well, for $1 \leq j \leq i+1$, one should notice that it follows by the inductive hypothesis that such relationships hold at the end of stage i and the beginning of stage $i+1$. Furthermore, a close examination of the updates included in matching and orthogonalization steps during stage $i+1$ (noting, in particular, the limitation in orthogonalization steps after matchings) is sufficient to establish that every such update is invertible and modifies the value of a vector $u_{r,s}$ (respectively, $v_{r,s}$) by replacing it with a linear combination of the values of vectors $u_{r',s'}$

(respectively, $v_{r',s'}$) such that $0 \leq r' \leq r$ and $1 \leq s' \leq k$. Consequently, these updates modify matrices $\mathcal{M}_{L,i+1,j}$ and $\mathcal{M}_{R,i+1,j}$ using updates of the form

$$\mathcal{M}_{L,i+1,j} := \mathcal{M}_{L,i+1,j} \cdot X_{update}$$

and

$$\mathcal{M}_{R,i+1,j} := \mathcal{M}_{R,i+1,j} \cdot Y_{update}$$

for nonsingular matrices $X_{update}, Y_{update} \in \mathsf{F}_q^{j\cdot k \times j\cdot k}$. Updates

$$X_{i+1,j} := X_{i+1,j} \cdot X_{update}$$

and

$$Y_{i+1,j} := Y_{i+1,j} \cdot Y_{update}$$

suffice to ensure that the equations at lines (21) are satisfied after these updates if they were satisfied before them. Part (b) of the claim now follows by a straightforward induction on the number of updates made in stage $i+1$.

Finally, part (c) of the claim is a consequence of part (b): If $0 \leq r \leq i$ and $1 \leq s \leq k$ then, at the end of stage $i+1$, $A^T \cdot u_{r,s}$ is a column of $A^T \cdot \mathcal{M}_{L,i+1,r}$, so that it is a linear combination of the columns of $A^T \cdot \widehat{\mathcal{K}}_{\vec{u},r}$. However, all such columns are also columns of $\widehat{\mathcal{K}}_{\vec{u},r+1}$, so $\widehat{A} \cdot u_{r,s}$ is a linear combination of the columns of $\widehat{\mathcal{K}}_{\vec{u},r+1}$. That is, there exists a vector $\gamma \in \mathsf{F}_q^{(r+1)\cdot k \times 1}$ such that

$$A^T \cdot u_{r,s} = \widehat{\mathcal{K}}_{\vec{u},r+1} \cdot \gamma.$$

However, as noted above,

$$\mathcal{M}_{L,i+1,r+1} = \widehat{\mathcal{K}}_{\vec{u},r+1} \cdot X_{i+1,r+1}$$

for a a nonsingular matrix $X_{i+1,r+1} \in \mathsf{F}_q^{(r+1)\cdot k \times (r+1)\cdot k}$. Consequently, if we set $\widehat{\gamma}$ to be $X_{i+1,r+1}^{-1} \cdot \gamma$ then

$$A^T u_{r,s} = \mathcal{M}_{L,i+1,r+1} \cdot \widehat{\gamma}.$$

It follows by the same argument that there exists a vector $\widehat{\delta} \in \mathsf{F}_q^{(r+1)\cdot k \times 1}$ such that

$$A \cdot v_{r,s} = \mathcal{M}_{R,i+1,r+1} \cdot \widehat{\delta}$$

at the end of stage $i+1$ as well. Part (c) now follows — because the columns of the matrix $\mathcal{M}_{L,i+1,r+1}$ (respectively, $\mathcal{M}_{R,i+1,r+1}$) are the vectors $u_{a,b}$ (respectively, $v_{a,b}$) such that $0 \leq a \leq r$ and $1 \leq b \leq k$. □

A.2.2 Proofs of Lemmas 2.1 and 2.2

The next lemma is useful for the proof of Lemma 2.1.

LEMMA A.5. *Let i be an integer such that $i \geq 0$ and there are at least $i+1$ stages in the Lanczos phase of the algorithm. Suppose, as well, that Invariants #2–5 are satisfied at the beginning of the first i stages of the algorithm. Then if t is an integer such that $1 \leq t \leq \ell$ and either $\mu_t = u_{g,h}$ or $\nu_t = v_{g,h}$ such that $0 \leq g \leq i - 2\cdot\Delta_{n,k} - 3$ and $1 \leq h \leq k$ then there exist elements $\alpha_{t,a}, \beta_{t,a}$ of F_q such that $1 \leq a \leq \ell$ and such that*

$$A^T \cdot \mu_t = \sum_{a=1}^{\ell} \alpha_{t,a} \cdot \mu_a \quad \textit{and} \quad A \cdot \nu_t = \sum_{a=1}^{\ell} \beta_{t,a} \cdot \nu_a$$

at the end of stage $i-2$ of the Lanczos phase of the computation.

PROOF. Suppose, as in the statement of the lemma, that i is an integer such that $i \geq 0$, there are at least $i+1$ stages included in the Lanczos phase of the algorithm, and that Invariants #2–5 are satisfied at the end of the first i stages. Let t be an integer such that $1 \leq t \leq \ell$ and either $\mu_t = u_{g,h}$ or $\nu_t = v_{g,h}$ for integers g and h such that $0 \leq g \leq i - 2\cdot\Delta_{n,k} - 3$ and $1 \leq h \leq k$.

Then, since Invariant #5 was satisfied at the end of each of the first i stages (and, in particular, at the end of stage $g+\Delta_{n,k}$), vectors μ_t and μ_t were matched at or before the end of stage $g+\Delta_{n,k}$. Consequently $\mu_t = u_{a,b}$ and $\nu_t = v_{c,d}$ for integers a, b, c, d such that $0 \leq a, c \leq g+\Delta_{n,k} \leq i-\Delta_{n,k}-3$ and $1 \leq b, c \leq k$. It now follows by part (c) of Lemma A.4 that, at the end of stage $i-2$,

$$A^T \cdot \mu_t = \sum_{c=0}^{i-\Delta_{n,k}-2} \sum_{d=1}^{k} \alpha_{t,c,d} \cdot u_{c,d}$$

and

$$A \cdot \nu_t = \sum_{c=0}^{i-\Delta_{n,k}-2} \sum_{d=1}^{k} \beta_{t,c,d} \cdot v_{c,d}$$

where $\alpha_{t,c,d}, \beta_{t,c,d} \in \mathsf{F}_q$ for $0 \leq c \leq i - \Delta_{n,k} - 2$ and $1 \leq d \leq k$.

Now, since Invariant #5 is satisfied at the end of stage $i-2$, each value $u_{c,d}$ and $v_{c,d}$ such that $0 \leq i - \Delta_{n,k} - 2$ has been matched at this point so that

$$A^T \cdot \mu_t = \sum_{a=1}^{\ell} \alpha_{t,a} \cdot \mu_a \quad \text{and} \quad A \cdot \nu_t = \sum_{a=1}^{\ell} \beta_{t,a} \cdot \nu_a$$

where $\alpha_{t,a}, \beta_{t,a} \in \mathsf{F}_q$ for $1 \leq a \leq \ell$ as well. □

LEMMA A.6. *Let i be an integer such that $i \geq 0$ and there are at least $i+1$ stages included in the Lanczos phase of the algorithm. Suppose, as well, that Invariants #2–5 are satisfied at the end of each of the first i stages. Then, at the beginning of the stage i (that is, the $i+1^{st}$ stage),*

$$(A^T \cdot u_{i-1,s}) \cdot \nu_t = \mu_t^T \cdot (A \cdot v_{i-1,s}) = 0$$

for $1 \leq s \leq k$ and for every integer t such that $1 \leq t \leq \ell$ and either $\mu_t = u_{g,h}$ or $\nu_t = v_{g,h}$ for integers g and h such that $0 \leq g \leq i - 2\cdot\Delta_{n,k} - 3$ and $1 \leq h \leq k$.

PROOF. Suppose, as in the statement of the lemma, that i is an integer such that $i \geq 0$, there are at least $i+1$ stages included in the Lanczos phase of the algorithm, and that Invariants #2–5 are satisfied at the end of the first i stages. Let s be an integer such that $1 \leq s \leq k$ and let t be an integer such that $1 \leq t \leq \ell$ and either $\mu_t = u_{g,h}$ or $\nu_t = v_{g,h}$ for integers g and h such that $0 \leq g \leq i - 2\cdot\Delta_{n,k} - 3$ and $1 \leq h \leq k$.

Then it follows by Lemma A.5, above,

$$A^T \cdot \mu_t = \sum_{a=1}^{\ell_{i-2}} \alpha_{t,a} \cdot \mu_a \quad \text{and} \quad A \cdot \nu_t = \sum_{a=1}^{\ell_{i-2}} \beta_{t,a} \cdot \nu_a$$

where $\gamma_{t,a}, \delta_{t,a} \in \mathsf{F}_q$ for $1 \leq a \leq \ell_{i-2}$ as well, where ℓ_{i-2} is the length of the sequences of matched vectors at lines (1) and (2) at the end of stage $i-2$ of the Lanczos phase.

It follows that, at the end of stage $i-1$ (and the beginning

of stage i),

$$(A^T \cdot u_{i-1,s}) \cdot \nu_t = u_{i-1,s}^T \cdot (A \cdot \nu_t) = \sum_{a=1}^{\ell_{i-2}} \beta_{t,a} \cdot u_{i-1,s}^T \cdot \nu_a.$$

However, $u_{i-1,s}^T \cdot \nu_a = 0$ for $1 \leq a \leq \ell_{i-2}$ — for either $u_{i-1,s}$ was matched during stage $i-1$, or it was not. If it was matched then $u_{i-1,s} = \mu_b$ for an integer $b \geq \ell_{i-2}+1$, so that $b \neq a$, and it follows by Invariant #2 that $u_{i-1,s}^T \cdot \nu_a = 0$. On the other hand, if $u_{i-1,s}$ was not matched at the end of stage $i-1$ then $u_{i-1,s}^T \cdot \nu_a = 0$ by Invariant #3, instead.

It now follows that $(A^T \cdot u_{i-1,s}) \cdot \nu_t = 0$. The same argument establishes that $\mu_t^T \cdot (A \cdot v_{i-1,s}) = 0$ as well, as required. □

The next lemma follows from the previous one and an inspection of the algorithm that has been defined.

LEMMA A.7. *Let i be an integer such that $i \geq 0$ and there are at least $i+1$ stages in the Lanczos phase of the algorithm. Suppose, as well, that Invariants #2–5 are satisfied at the beginning of the first i stages of the algorithm. Then, if the vectors $u_{i,s}$ and $v_{i,s}$ are initialized and the orthogonalizations described after Lemma 2.1 are performed, then $u_{i,s}^T \cdot \nu_a = 0 = \nu_a^T \cdot v_{i,s}$ for $1 \leq a \leq \ell$ and $1 \leq s \leq k$, so that Invariants #2 and #3 are both satisfied at this point.*

PROOF. If the conditions in the lemma are satisfied then Invariant #2 is satisfied after the above-mentioned operations are performed because this invariant was satisfied at the end of stage $i-1$, and none of the vectors mentioned in this invariant have been changed by these operations.

Now it follows by Lemma A.6 that, following the initialization of $u_{i,s}$ (respectively, $v_{i,s}$) as $A^T \cdot u_{i-1,s}$ (respectively, $A \cdot v_{i-1,s}$), $\mu_t^T \cdot v_{i,s} = u_{i,s}^T \cdot \nu_t = 0$ for $1 \leq s \leq k$ and every integer t such that either $\mu_t = u_{g,h}$ or $\nu_t = v_{g,h}$ where $0 \leq g \leq i - 2 \cdot \Delta_{n,k} - 3$ and $1 \leq h \leq k$.

On the other hand, if $1 \leq t \leq \ell$ and t does not satisfy this condition then $\mu_t = u_{c,d}$ and $\nu_t = v_{c',d'}$ for integers c, c', d, d' such that $i - 2 \cdot \Delta_{n,k} - 2 \leq c, c' \leq i$ and $1 \leq d, d' \leq k$, so that an orthogonalization step including μ_t (and ν_t) has been included as part of the initialization of $u_{i,s}$ and $v_{i,s}$. It therefore follows by the correctness of the orthogonalization step (see Lemmas A.1 and A.2, above) that

$$u_{i,s}^T \cdot \nu_a = 0 = \mu_a^T \cdot v_{i,s}$$

after these operations, for every integer s such that $1 \leq s \leq k$ and for every integer a such that $1 \leq a \leq \ell$.

Note as well that if $0 \leq r \leq i-1$ and $u_{r,s}$ (respectively, $v_{r,s}$) is a vector that is unmatched at this point then

$$u_{r,s}^T \cdot \nu_a = 0$$

(respectively, $\mu_a^T \cdot v_{r,s} = 0$) for $1 \leq a \leq \ell$ as well, because Invariant #3 was satisfied at the end of stage $i-1$, and none of the vectors $u_{r,s}$ or $v_{r,s}$ such that $0 \leq r \leq i-1$ and $1 \leq s \leq k$, and none of the vectors $\mu_1, \mu_2, \ldots, \mu_\ell$ or $\nu_1, \nu_2, \ldots, \nu_\ell$, have been changed by the operations at the beginning of stage i being considered above.

Thus Invariant #3 is also satisfied, as claimed. □

Lemma A.8–A.12 help to establish that Invariants #2–5 are satisfied at the end of round i, for $i \geq 1$.

LEMMA A.8. *Let i be an integer such that $i \geq 1$ and there are at least $i+1$ stages of the Lanczos phase of the algorithm (ending with stage i). Let g be an odd integer such that $1 \leq g \leq 2 \cdot \Delta_{n,i} + 1$.*

Recall that round g of stage i begins with a matching step, during which unmatched vectors $u_{i-\Delta_{n,k}+j,s}$ are matched with unmatched vectors $v_{i,t}$ for $j = \lfloor g/2 \rfloor$ and $1 \leq s, t \leq k$.

For $0 \leq r \leq i$ and $1 \leq s \leq k$,

- *let $a_{r,s}$ be the value of $u_{r,s}$ immediately before round g of stage i,*
- *let $b_{r,s}$ be the value of $u_{r,s}$ immediately after round g of stage i,*
- *let $c_{r,s}$ be the value of $v_{r,s}$ immediately before round g of stage i, and*
- *let $d_{r,s}$ be the value of $v_{r,s}$ immediately after round g of stage i.*

Then the following relationships hold.

(a) *If $0 \leq r < i + \Delta_{n,k} + j$ then $a_{r,s} = b_{r,s}$ for $1 \leq s \leq k$.*

(b) *Suppose that $r = i + \Delta_{n,k} + j$ and that $\sigma_1, \sigma_2, \ldots, \sigma_h$ are integers such that*

$$1 \leq \sigma_1 < \sigma_2 < \cdots < \sigma_h \leq k$$

and $u_{r,\sigma_1}, u_{r,\sigma_2}, \ldots, u_{r,\sigma_h}$ are the vectors $u_{r,s}$ (for $1 \leq s \leq k$) that were already matched before round g of stage i. Suppose $\tau_1, \tau_2, \ldots, \tau_{k-h}$ are integers such that

$$1 \leq \tau_1 < \tau_2 < \cdots < \tau_{k-h}$$

and $u_{r,\tau_1}, u_{r,\tau_2}, \ldots, u_{r,\tau_{k-h}}$ are the vectors $u_{r,s}$ (for $1 \leq s \leq k$) that were not *matched, yet, at the beginning of round g of stage i.*

Then $a_{r,\sigma_w} = b_{r,\sigma_w}$ for $1 \leq w \leq h$, and

$$\begin{bmatrix} b_{r,\tau_1} & b_{r,\tau_2} & \cdots & b_{r,\tau_{k-h}} \end{bmatrix} = \begin{bmatrix} a_{r,\tau_1} & a_{r,\tau_2} & \cdots & a_{r,\tau_{k-h}} \end{bmatrix} \cdot X_L$$

for a nonsingular matrix $X_L \in \mathsf{F}_q^{(k-h)\times(k-h)}$.

(c) *If $i - \Delta_{n,k} + j < r \leq i$ then*

$$\begin{bmatrix} b_{r,1} & b_{r,2} & \cdots & b_{r,k} \end{bmatrix} = \begin{bmatrix} a_{r,1} & a_{r,2} & \cdots & a_{r,k} \end{bmatrix} + \begin{bmatrix} a_{s,\tau_1} & a_{s,\tau_2} & \cdots & a_{s,\tau_{k-h}} \end{bmatrix} \cdot Y_{s,r}$$

for $s = i - \Delta_{n,k} + j < r$, $\tau_1, \tau_2, \ldots, \tau_{k-h}$ as above, and for a matrix $Y_{s,r} \in \mathsf{F}_q^{(k-h)\times k}$.

(d) *If $1 \leq r \leq i-1$ then $c_{r,s} = d_{r,s}$ for $1 \leq s \leq k$.*

(e) *Suppose that $\theta_1, \theta_2, \ldots, \theta_{\widehat{h}}$ are integers such that*

$$1 \leq \theta_1 < \theta_2 < \cdots < \theta_{\widehat{h}} \leq k$$

and $v_{i,\theta_1}, v_{i,\theta_2}, \ldots, v_{i,\theta_{\widehat{h}}}$ are the vectors $v_{i,s}$ (for $1 \leq s \leq k$) that were already matched before round g of stage i. Suppose $\iota_1, \iota_2, \ldots, \iota_{k-\widehat{h}}$ are integers such that

$$1 \leq \iota_1 < \iota_2 < \cdots < \iota_{k-\widehat{h}}$$

and $v_{i,\iota_1}, v_{i,\iota_2}, \ldots, v_{i,\iota_{k-\widehat{h}}}$ are the vectors $v_{i,s}$ (for $1 \leq s \leq k$) that were not *matched, yet, at the beginning of round g of stage i.*

Then $c_{i,\theta_w} = d_{i,\theta_w}$ for $1 \leq w \leq \widehat{h}$ and

$$\begin{bmatrix} d_{i,\iota_1} & d_{i,\iota_2} & \cdots & d_{i,\iota_{k-\widehat{h}}} \end{bmatrix} = \begin{bmatrix} c_{i,\iota_1} & c_{i,\iota_2} & \cdots & c_{i,\iota_{k-\widehat{h}}} \end{bmatrix} \cdot X_R$$

for a nonsingular matrix $X_R \in \mathsf{F}_q^{(k-\widehat{h})\times(k-\widehat{h})}$.

PROOF. Parts (a) and (d) of the claim can be established by an inspection of the details of round g of stage i of the Lanczos phase, when g is odd: If $j = \lfloor g/2 \rfloor$, as above, then $u_{r,s}$ is not accessed during this round for $0 \le r < i - \Delta_{n,k} + j$, and $v_{r,s}$ is not accessed or modified for $0 \le r \le i-1$ and $1 \le s \le k$, either.

Part (b) can be established by noticing that no vectors that have been matched before this round are accessed or modified during it (so that $a_{r,\sigma_w} = b_{r,\sigma_w}$ for $r = i - \Delta_{n,k} + j$ and $1 \le w \le h$, as claimed) and by noticing that the previously unmatched vectors

$$u_{r,\tau_1}, u_{r,\tau_2}, \ldots, u_{r,\tau_{k-h}}$$

are involved in two operations during round g:

- a matching operation matches these with the unmatched vectors

$$v_{i,\iota_1}, v_{i,\iota_2}, \ldots, v_{i,\iota_{k-\widehat{h}}};$$

- the vectors u_{r,τ_w} that are still unmatched are then orthogonalized, to ensure that each is orthogonal to each newly matched vector v_{i,ι_y}.

Let e_{r,τ_w} be the value of the vector u_{r,τ_w} after the above matching step but before the orthogonalization, for $1 \le w \le h$. Then it follows by the correctness of the matching procedure (see Lemma A.3 for details) that

$$\begin{bmatrix} e_{r,\tau_1} & e_{r,\tau_2} & \cdots & e_{r,\tau_h} \end{bmatrix} = \begin{bmatrix} a_{r,\tau_1} & a_{r,\tau_2} & \cdots & a_{r,\tau_h} \end{bmatrix} \cdot Y_L$$

for a nonsingular matrix $Y_L \in \mathsf{F}_q^{(k-h)\times(k-h)}$. Similarly, since the orthogonalization step subtracts a linear combination of the values of newly matrices vectors u_{r,τ_w} from each of these vectors that is still unmatched,

$$\begin{bmatrix} b_{r,\tau_1} & b_{r,\tau_2} & \cdots & b_{r,\tau_h} \end{bmatrix} = \begin{bmatrix} e_{r,\tau_1} & e_{r,\tau_2} & \cdots & e_{r,\tau_h} \end{bmatrix} \cdot Z_L$$

for a nonsingular matrix $Z_L \in \mathsf{F}_q^{(k-h)\times(k-h)}$. It now suffices to set $X_L = Y_L \cdot Z_L$ (noting that this is also a nonsingular matrix in $\mathsf{F}_q^{(k-h)\times 1}$) to establish this part of the claim.

Part (e) can be established using the argument given above to establish part (b).

Finally, note that if $i - \Delta_{n,k} + j < r \le i$ then then the vectors $u_{r,a}$ are included in an orthogonalization step to ensure that they are orthogonal to newly matched vectors v_{ι_y} but otherwise unchanged. Since the above vectors v_{ι_y} were matched with (some of) the vectors

$$u_{s,\tau_1}, u_{s,\tau_2}, \ldots, u_{s,\tau_{k-h}}$$

for $s = i - \Delta_{n,k} + j$, a consideration of the details of the orthogonalization process suffices to establish that the values of $u_{r,1}, u_{r,2}, \ldots, u_{r,k}$ using an update with the form shown in part (c), as well, as required to establish the claim. □

LEMMA A.9. *Let i be an integer such that $i \ge 1$ and there are at least $i+1$ stages of the Lanczos phase of the algorithm (ending with stage i). Let g be an even integer such that $1 \le g \le 2 \cdot \Delta_{n,i}$.*

Recall that round g of stage i begins with a matching step, during which unmatched vectors $u_{i,s}$ are matched with unmatched vectors $v_{i-\Delta_{n,k}+j,t}$ for $j = (g-2)/2$ and $1 \le s, t \le k$.

For $0 \le r \le i$ and $1 \le s \le k$,

- *let $a_{r,s}$ be the value of $u_{r,s}$ immediately before round g of stage i,*
- *let $b_{r,s}$ be the value of $u_{r,s}$ immediately* after *round g of stage i,*
- *let $c_{r,s}$ be the value of $v_{r,s}$ immediately before round g of stage i, and*
- *let $d_{r,s}$ be the value of $v_{r,s}$ immediately after round g of stage i.*

Then the following relationships hold.

(a) *If $0 \le r < i + \Delta_{n,k} + j$ then $c_{r,s} = d_{r,s}$ for $1 \le s \le k$.*

(b) *Suppose that $r = i\Delta_{n,k} + j$ and that $\sigma_1, \sigma_2, \ldots, \sigma_h$ are integers such that*

$$1 \le \sigma_1 < \sigma_2 < \cdots < \sigma_h \le k$$

and $v_{r,\sigma_1}, v_{r,\sigma_2}, \ldots, v_{r,\sigma_h}$ are the vectors $v_{r,s}$ (for $1 \le s \le k$) that were already matched before round g of stage i. Suppose $\tau_1, \tau_2, \ldots, \tau_{k-h}$ are integers such that

$$1 \le \tau_1 < \tau_2 < \cdots < \tau_{k-h}$$

and $v_{r,\tau_1}, v_{r,\tau_2}, \ldots, v_{r,\tau_{k-h}}$ are the vectors $v_{r,s}$ (for $1 \le s \le k$) that were not *matched, yet, at the beginning of round g of stage i.*

Then $c_{r,\sigma_w} = d_{r,\sigma_w}$ for $1 \le w \le h$ and

$$\begin{bmatrix} d_{r,\tau_1} & d_{r,\tau_2} & \cdots & d_{r,\tau_{k-h}} \end{bmatrix} = \begin{bmatrix} c_{r,\tau_1} & c_{r,\tau_2} & \cdots & c_{r,\tau_{k-h}} \end{bmatrix} \cdot X_R$$

for a nonsingular matrix $X_R \in \mathsf{F}_q^{(k-h)\times(k-h)}$.

(c) *If $i - \Delta_{n,k} + j < r \le i$ then*

$$\begin{bmatrix} d_{r,1} & d_{r,2} & \cdots & d_{r,k} \end{bmatrix} = \begin{bmatrix} c_{r,1} & c_{r,2} & \cdots & c_{r,k} \end{bmatrix} + \begin{bmatrix} c_{s,\tau_1} & c_{s,\tau_2} & \cdots & c_{s,\tau_{k-h}} \end{bmatrix} \cdot Y_{s,r}$$

for $\tau_1, \tau_2, \ldots, \tau_{k-h}$ as above and for a matrix $Y_{s,r} \in \mathsf{F}_q^{(k-h)\times k}$.

(d) *If $1 \le r \le i-1$ then $a_{r,s} = b_{r,s}$ for $1 \le s \le k$.*

(e) *Suppose that $\theta_1, \theta_2, \ldots, \theta_{\widehat{h}}$ are integers such that*

$$1 \le \theta_1 < \theta_2 < \cdots < \theta_{\widehat{h}} \le k$$

and $u_{i,\theta_1}, u_{i,\theta_2}, \ldots, u_{i,\theta_{\widehat{h}}}$ are the vectors $u_{i,s}$ (for $1 \le s \le k$) that were already matched before round g of stage i. Suppose $\iota_1, \iota_2, \ldots, \iota_{k-\widehat{h}}$ are integers such that

$$1 \le \iota_1 < \iota_2 < \cdots < \iota_{k-\widehat{h}}$$

and $u_{i,\iota_1}, u_{i,\iota_2}, \ldots, u_{i,\iota_{k-\widehat{h}}}$ are the vectors $u_{i,s}$ (for $1 \le s \le k$) that were not *matched, yet, at the beginning of round g of stage i.*

Then $a_{i,\theta_w} = b_{i,\theta_w}$ for $1 \le w \le \widehat{h}$ and

$$\begin{bmatrix} b_{i,\iota_1} & b_{i,\iota_2} & \cdots & b_{i,\iota_{k-\widehat{h}}} \end{bmatrix} = \begin{bmatrix} a_{i,\iota_1} & a_{i,\iota_2} & \cdots & a_{i,\iota_{k-\widehat{h}}} \end{bmatrix} \cdot X_L$$

for a nonsingular matrix $X_L \in \mathsf{F}_q^{(k-\widehat{h})\times(k-\widehat{h})}$.

PROOF. This can be proved in the same way as the lemma that preceded it: Notice that, for g as described in the claim, the matchings and orthogonalizations are exactly as described for round $g-1$ (an odd integer between 0 and $2 \cdot \Delta_{n,k}$) except that the roles of $u_{r,s}$ and $v_{r,s}$ are reversed. Consequently, reversing the roles of $u_{r,s}$ and $v_{r,s}$ in the previous proof provides a proof of the current lemma as well. □

LEMMA A.10. *Let i be an integer such that $i \geq 1$ and there are at least $i+1$ stages of the Lanczos phase of the computation (ending with stage i). Let g be an integer such that $1 \leq g \leq 2 \cdot \Delta_{n,k} + 1$.*

Suppose that r and s are integers such that $1 \leq r, s \leq k$ and, during round g of stage i of the Lanczos phase, the currently unmatched vectors $u_{r,a}$ (for $1 \leq a \leq k$) are matched with the currently unmatched vectors $v_{s,b}$ (for $1 \leq b \leq k$).

If Invariants #2 and 3 are satisfied at the end of round g then $u_{r,a}^T \cdot v_{s,b} = 0$ for all integers a and b such that $1 \leq a, b \leq k$ and $u_{r,a}$ and $v_{s,b}$ are still unmatched at this point.

PROOF. Let $\sigma_1, \sigma_2, \ldots, \sigma_h$ be integers such that

$$1 \leq \sigma_1 < \sigma_2 < \cdots < \sigma_h \leq k$$

and $u_{r,\sigma_1}, u_{r,\sigma_2}, \ldots, u_{r,\sigma_h}$ are the vectors $u_{r,a}$ (for $1 \leq a \leq k$) that are unmatched at the beginning of round g. Similarly, let $\tau_1, \tau_2, \ldots, \tau_m$ be integers such that

$$1 \leq \tau_1 < \tau_2 < \cdots < \tau_m \leq k$$

and $v_{s,\tau_1}, v_{s,\tau_2}, \ldots, v_{s,\tau_m}$ are the vectors $v_{s,b}$ (for $1 \leq b \leq k$) that are unmatched at the beginning of round g.

Let $a_{r,t}$ and $b_{r,t}$ be the values of $u_{r,t}$ before and after round g, respectively, for $1 \leq t \leq k$, and let $c_{s,w}$ and $d_{s,w}$ be the values of $v_{s,w}$ before and after round g, respectively, for $1 \leq w \leq k$, as well.

Consider the matrices

$$K_L = \begin{bmatrix} a_{r,\sigma_1} & a_{r,\sigma_2} & \cdots & a_{r,\sigma_h} \end{bmatrix} \in \mathsf{F}_q^{n \times h}$$

and

$$K'_L = \begin{bmatrix} b_{r,\sigma_1} & b_{r,\sigma_2} & \cdots & b_{r,\sigma_h} \end{bmatrix} \in \mathsf{F}_q^{n \times h}$$

whose columns are the values of the vectors

$$u_{r,\sigma_1}, u_{r,\sigma_2}, \ldots, u_{r,\sigma_h}$$

before and after round g, respectively. It follows by Lemmas A.8 and A.9 that

$$K'_L = K_L \cdot X_L$$

for a nonsingular matrix $X_L \in \mathsf{F}_q^{h \times h}$. Consider also the matrices

$$K_R = \begin{bmatrix} c_{s,\tau_1} & c_{s,\tau_2} & \cdots & c_{s,\tau_m} \end{bmatrix} \in \mathsf{F}_q^{n \times m}$$

and

$$K'_R = \begin{bmatrix} d_{s,\tau_1} & d_{s,\tau_2} & \cdots & d_{s,\tau_m} \end{bmatrix} \in \mathsf{F}_q^{n \times m}$$

whose columns are the values of the vectors

$$v_{s,\tau_1}, v_{s,\tau_2}, \ldots, v_{s,\tau_m}$$

before and after round g, respectively. It also follows by Lemmas A.8 and A.9 that

$$K'_R = K_R \cdot Y_R$$

for a nonsingular matrix $Y_R \in \mathsf{F}_q^{m \times m}$.

Let t be the number of pairs of vectors that are matched during round g. It follows by the correctness of the match procedure (see, in particular, Lemma A.3) that t is also the rank of the matrix $K_L^T \cdot K_R \in \mathsf{F}_q^{h \times m}$. Now, since the above matrices $X_L \in \mathsf{F}_q^{h \times h}$ and $Y_R \in \mathsf{F}_q^{m \times m}$ are nonsingular (so that X_L^T is nonsingular as well), and

$$(K'_L)^T \cdot K'_R = X_L^T \cdot (K_L^T \cdot K_R) \cdot Y_R,$$

it follows that the matrix $(K'_L)^T \cdot K'_R$ has rank t as well.

Now there exist permutation matrices $P_L \in \mathsf{F}_q^{h \times h}$ and $P_R \in \mathsf{F}_q^{m \times m}$ such that the first t columns of $K'_L \cdot P_L$ and of $K'_R \cdot P_R$ are the values (after round g) of the t pairs of vectors matched during this round — so that (since Invariant #2 is satisfied after round g) the top left submatrix of the matrix $P_L^T \cdot ({K'_L}^T \cdot K'_R) \cdot P_R$ is the identity matrix I_t.

Furthermore, since Invariant #3 is satisfied at the end of round g, the final $h-t$ columns of $K'_L \cdot P_L$ are the unmatched vectors $u_{r,w}$ at the end of this round (for $1 \leq w \leq k$), and the final $m-t$ columns of $K'_R \cdot P_R$ are the unmatched vectors $v_{s,w}$ at the end of this round (for $1 \leq w \leq k$) as well. Consequently, since Invariant #3 is also satisfied at the end of this round,

$$P_L^T \cdot ({K'_L}^T \cdot K'_R) \cdot K_R = \begin{bmatrix} I_t & 0 \\ 0 & Z \end{bmatrix}$$

for a matrix $Z \in \mathsf{F}_q^{(k-t) \times (m-t)}$ whose entries are the inner products $u_{r,a}^T \cdot v_{s,b}$ of the vectors $u_{r,a}$ and $v_{s,b}$ that remain unmatched. Now, since ${K'_L}^T \cdot K_R$ has rank t, the matrix $P_L^T \cdot ({K'_L}^T \cdot K'_R) \cdot P_R$ has rank t as well, so that $Z = 0$ — as needed to establish the claim. □

LEMMA A.11. *Invariants #2–4, 6 and 7 are satisfied at the end of stage 0 of the Lanczos phase of the computation.*

PROOF. Note first that Invariants #2–4, 6 and 7 all hold at the beginning of stage 0 because the claims in Invariants #2–4 and 6 are vacuous, and Invariant #7 is satisfied because $\chi = 0$ and $\rho = \sigma = A \cdot w + b$.

Invariants #2 and 3 are still satisfied after the initialization of $u_{0,r}$, $w_{0,r}$ and $v_{0,r}$, for $1 \leq r \leq k$, because these claims are still vacuous. Invariant #6 still holds at this point since $v_{0,r} = A \cdot w_r = A \cdot w_{0,r}$ for $1 \leq r \leq k$. Invariant #7 still holds because none of the values mentioned in it have been changed.

It follows by the correctness of the match procedure (see Lemma A.3, above) that Invariants #2, 6 and 7 hold again after the matching step in stage 0. These invariants are also satisfied again after the orthogonalization steps that complete stage 0, because none of the values mentioned in these invariants are changed by these final operations.

Invariant #3 has been re-established as well, at the end of stage 0, because all unmatched vectors $u_{0,r}$ (respectively, $v_{0,r}$) have been orthogonalized against all matched vectors $v_{0,s}$ (respectively, $u_{0,s}$) by the end of this stage.

Finally, the proof that Invariant #4 has also been re-established is the same as the proof given for Lemma A.10, above. □

LEMMA A.12. *Let i be an integer such that $i \geq 1$ and there are at least $i+1$ stages of the Lanczos phase of the computation. Suppose, as well, that Invariants #2–#7 are satisfied at the end of stage h of the Lanczos phase for $0 \leq h \leq i-1$.*

Consider the various updates of vectors $u_{r,s}$ and $v_{r,s}$ included in stage i of the Lanczos phase of the computation, for integers r and s such that $0 \leq r \leq i$ and $1 \leq s \leq k$ that follow the initialization of $u_{i,s}$ and $v_{i,s}$.

(a) *If g is an integer such that $0 \leq g \leq 2 \cdot \Delta_{n,k} + 1$ then Invariants #2, 3, 6 and 7 are all satisfied after round g of stage i of the Lanczos phase of the computation.*

(b) *If g is an integer such that $0 \leq g \leq 2 \cdot \Delta_{n,k} + 1$ and vectors $u_{r,s}$ and $v_{r',s'}$ are unmatched after round g of stage i of the Lanczos phase, where $0 \leq r, r' \leq i-1$ and $1 \leq s, s' \leq k$, then $u_{r,s}^T \cdot v_{r',s'} = 0$ at this point in the computation.*

(c) *If g is an even integer such that $0 \leq g \leq 2 \cdot \Delta_{n,k}$ and r and s are integers such that $0 \leq r < i - \Delta_{n,k} + g/2$ and $1 \leq s \leq k$, and $u_{r,s}$ (respectively, $v_{r,s}$) is unmatched after round g, then $u_{r,s}^T \cdot v_{a,b} = 0$ (respectively, $u_{a,b}^T \cdot v_{r,s} = 0$) after round g for all integers a and b such that $0 \leq a \leq i$ and $1 \leq b \leq k$.*

PROOF. The above claims will be established by induction on g. The strong form of induction will be useful here, since part (c) includes conditions that are only satisfied when g is even.

Basis: It follows by Lemma A.7, above, that Invariants #2 and #3 hold once again after the initialization of $u_{i,s}$ and $v_{i,s}$ for $1 \leq s \leq k$.

If $0 \leq r \leq i-1$ and $1 \leq s \leq k$ then $A \cdot w_{r,s} = v_{r,s}$ after the initialization of the vectors $u_{i,s'}$ and $v_{i,s'}$ because Invariant #6 was satisfied at the end of stage $i-1$ and none of the vectors $v_{r,s}$ or $w_{r,s}$ such that $0 \leq r \leq i-1$ and $1 \leq s \leq k$ were changed by this initialization. Since $w_{i,s} = v_{i-1,s}$ and $A \cdot v_{i-s,s}$, $v_{i,s} = A \cdot w_{i,s}$ at the beginning of the initialization process, re-establishing Invariant #6 at that point. It follows by the correctness of the orthogonalization process (see Lemma A.2, above) that Invariant #6 holds at the end of the initialization of $u_{i,s}$ and $v_{i,s}$ as well.

Invariant #7 also holds at the end of the initialization of $u_{i,s}$ and $v_{i,s}$, as needed to establish part (a) of the claim, because it held at the end of stage $i-1$ and none of the values referred to in this invariant have been modified.

Now, since none of the vectors $u_{r,s}$ or $v_{r,s}$ such that $0 \leq r \leq i-1$ and $1 \leq s \leq k$ are changed by these initial operations, and since Invariants #3–5 were also satisfied at the end of stage $i-1$, part (b) of the claim follows because Invariant #4 was satisfied at the end of stage $i-1$.

Part (c) holds because Invariant #5 also held at the end of stage $i-1$: There are no vectors $u_{r,s}$ or $v_{r,s}$ such that $0 \leq r < i - \Delta_{n,k}$ that are unmatched at this point at all, so the claim is vacuous.

Inductive Step: Suppose $0 \leq g \leq 2 \cdot \Delta_{n,k}$ and that the above conditions are satisfied for every integer h such that $0 \leq h \leq g$. It is necessary and sufficient to establish that they are satisfied for $g+1$ as well.

Let us first consider part (a) of the claim. Round $g+1$ begins with a *matching* update that extends the sequence of vectors shown at line (1) and (2). It follows by the correctness of the match procedure (see, in particular, Lemma A.3) that Invariants #2, 6 and 7 are satisfied once again after this update.

Invariants #2 and 7 are satisfied at the end of round $g+1$ because none of the remaining (orthogonalizaton) operations modify any of the values mentioned in them. It follows the correctness of the orthogonalization procedure (see Lemma A.2, once again) that Invariant #6 is satisfied at the end of this round as well.

In remains to re-establish Invariant #3. The cases $g \leq 2 \cdot \Delta_{n,k} - 1$ and $g = 2 \cdot \Delta_{n,k}$ are considered separately below.

Case: $0 \leq g \leq 2 \cdot \Delta_{n,k} - 1$. Once again, round $g+1$ begins with a *matching* update.that extends the sequences of vectors shown at lines (1) and (2). In particular, either

1. unmatched vectors $u_{r,s}$ are matched with unmatched vectors $v_{i,s'}$, for some integer r such that $0 \leq r \leq i-1$, or
2. unmatched vectors $u_{i,s}$ are matched with unmatched vectors $v_{r,s'}$ for some integer r such that $0 \leq r \leq i-1$, instead.

In particular, this is the case for $r = i - \Delta_{n,k} + \lfloor g/2 \rfloor$. The round ended with a series of orthogonalization steps (described above and, as needed again, below).

Subcase 1: In this case g is odd and Lemma A.8 is applicable; it now suffices to show that if $0 \leq a \leq i$, $1 \leq b \leq k$, and $u_{a,b}$ (respectively, $v_{a,b}$) is unmatched at the end of round $g+1$ then $u_{a,b}^T \cdot \nu_c = 0$ (respectively, $\mu_c^T \cdot v_{a,b} = 0$) for $1 \leq c \leq \ell$ at the end of round $g+1$.

Now it follows by the inductive hypothesis that if μ_c and ν_c were already matched at the end of round g then $u_{a,b}^T \cdot \nu_c = 0$ (respectively, $\mu_c^T \cdot v_{a,b} = 0$) at the end of round $g+1$ — Invariant #3 was satisfied at the end of round g, the values of μ_c and ν_c have not been changed during round $g+1$, and the value of $u_{a,b}$ (respectively, $v_{a,b}$) at the end of round $g+1$ is a linear combination of the values of vectors $u_{e,f}$ (respectively, $v_{e,f}$) that were unmatched at the beginning of round $g+1$, and at the end of round g — see Lemma A.8, above.

Similarly, if $0 \leq a < r$ then $a < i - \Delta_{n,k} + g/2$ as well, and it follows by part (c) of the inductive hypothesis that, at the end of round g, $u_{a,b}^T \cdot v_{d,e} = 0$ for all integers d and e such that $0 \leq d \leq i$ and $1 \leq e \leq k$.

Once again, it follows by Lemma A.8 that the value of $u_{a,b}$ is a linear combination of the values of vectors $u_{a',b'}$ at the end of round g, for a' and b' such that $0 \leq a' \leq a$, $1 \leq b' \leq k$, and $u_{a',b'}$ was also unmatched at the end of round g. It also follows by this lemma that the value of $v_{d,e}$ is now a linear combination of the values of vectors $v_{d',e'}$ (for $0 \leq d' \leq i$ and $1 \leq e' \leq k$) as they were defined at the end of round g. This suffices to establish that $u_{a,b}^T \cdot v_{d,e} = 0$ at the end of round $g+1$ as well. In particular, it establishes that $u_{a,b}^T \cdot \nu_c = 0$ at the end of round $g+1$ if μ_c was unmatched at the end of round g but matched at the end of round $g+1$.

On the other hand, if $r \leq a \leq i$ then the matching step in round $g+1$ is followed by an orthogonalization step orthogonalizing still-unmatched vectors $u_{a,b}$ and newly matched vectors $\nu_c = v_{i,s'}$, so it follows by the correctness of the orthogonalization process that $u_{a,b} \cdot \nu_c = 0$ at the end of round $g+1$ in this case, as well.

Consider still-unmatched vectors $v_{a,b}$ and newly matched vectors μ_c. Now, if $0 \leq a \leq i-1$ and $1 \leq b \leq k$ then, since $\mu_c = u_{r,s}$ for $0 \leq r \leq i-1$, and $u_{r,s}$ and $v_{a,b}$ were each unmatched at the end of round g, it follows by part (b) of the inductive hypothesis that that $\mu_c^T \cdot v_{a,b} = u_{r,s}^T \cdot v_{a,b} = 0$ at the end of round g — and, indeed, that $u_{r',s'} \cdot v_{a',b'} = 0$ at the end of round g as well, for all integers a', b', r', s' such that $0 \leq a', r' \leq i-1$, $1 \leq b', s' \leq k$, and $u_{r',s'}$ and $v_{a',b'}$ were each unmatched at the end of round g.

Once again, it follows by Lemma A.8 that the value of $v_{a,b}$ at the end of round $g+1$ is a linear combination of the values of unmatched vectors $v_{a',b'}$, for $0 \leq a' \leq a$ and $1 \leq b' \leq k$, as these values were defined at the end of round i. The lemma implies that the value of $u_{r,s}$ at the end of round $g+1$ is a linear combination of values $u_{r',s'}$ for r' and s' such that $0 \leq r' \leq r$ and $1 \leq s' \leq k$ and $u_{r',s'}$ was unmatched at the end of round g, as these values were defined at the end of round g as well. It follows that $u_{r,s}^T \cdot v_{a,b} = \mu_c^T \cdot v_{a,b} = 0$ at

the end of round $g+1$ once again.

It remains only to consider the case that $a = i$. In this case one should note that the matching step in round $g+1$ is followed by an orthogonalization step, orthogonalizing still-unmatched vectors $v_{a,b}$ with newly matched vectors μ_c, and ensuring that $\mu_c^T \cdot v_{a,b} = 0$ as well.

It follows (finally) that Invariant #3 is satisfied at the end of round $g+1$ if this subcase is applicable.

Subcase 2: The proof for this other subcase is almost identical — the roles of vectors $u_{a,b}$ and $v_{a,b}$ must simply be interchanged in the argument, and Lemma A.9 must be applied instead of Lemma A.8.

Case: $g = 2 \cdot \Delta_{n,k}$. In this case the argument needed to re-establish Invariant #3 is similar to, but simpler, than the above.

In particular, in this case, the matching phase matches vectors $u_{i,s}$ with vectors $v_{i,s'}$.

Now, if $0 \leq r \leq i-1$, $1 \leq s \leq k$, and $u_{r,s}$ (respectively, $v_{r,s}$) was unmatched at the end of round g then an application of part (c) of the inductive hypothesis establishes that $u_{r,s}^T \cdot v_{i,t} = 0$ for every unmatched vector $v_{i,t}$ (respectively, $u_{i,t}^T \cdot v_{r,s} = 0$ for every unmatched vector $u_{i,t}$) at the end of round g. It follows by Lemma A.8 that, at the end of this round, the value of an unmatched vector $u_{r,s}$ (respectively, $v_{r,s}$) is a linear combination of the values of vectors $u_{r',w}$ (respectively, $v_{r',w}$) such that $0 \leq r' \leq r$, $1 \leq w \leq k$, and $u_{r',w}$ (respectively, $v_{r',w}$) was unmatched at the end of round g — as these values were defined after round g. Thus $u_{r,s}^T \cdot v_{i,s'} = 0$ (respectively, $u_{i,s'} \cdot v_{r,s} = 0$), so that $u_{r,s}^T \cdot \nu_c = 0$ (respectively, $\mu_c^T \cdot v_{r,s} = 0$) whenever $u_{r,s}$ (respectively, $v_{r,s}$) is still unmatched after round $g+1$ and μ_c and ν_c were newly matched in this final round.

On the other hand, if $r = i$, then the matching step is followed by orthogonalization steps that orthogonalize still-unmatched vectors $u_{r,s}$ with newly matched vectors $\nu_c = v_{i,t}$ and that orthogonalize still unmatched vectors $v_{r,s}$ with newly matched vectors $\mu_c = u_{i,t}$, as well. The correctness of the orthogonalization process suffices to complete the proof (at long last) that Invariant #3 is also established at the end of this final round, as needed to complete the proof that part (a) of the claim holds.

The proof that part (b) holds after round $g+1$, if it held after round g, is more straightforward: If r, s, r' and s' are integers such that $0 \leq r, r' \leq i-1$, $1 \leq s, s' \leq k$, and $u_{r,s}$ and $v_{r',s'}$ are also both unmatched at the end of round $g+1$ then they were unmatched at the end of round g as well. Once again, Lemma A.8 establishes that the value of $u_{r,s}$ at the end of round $g+1$ is a linear combination of the values of vectors $u_{a,b}$ such that $0 \leq a \leq r$, $1 \leq b \leq k$, and $u_{a,b}$ was unmatched at the end of round g, as these values were defined at the end of round g. Similarly, the value of $v_{r',s'}$ at the end of round $g+1$ is a linear combination of the values of vectors $v_{a',b'}$ such that $0 \leq a' \leq r'$, $1 \leq b' \leq k$, and $v_{a',b'}$ was unmatched at the end of round g, as these values were defined at the end of round g, as well. Since (by part (b) of the inductive hypothesis) $u_{a,b}^T \cdot v_{a',b'}$ at the end of round g for a, b, a' and b' as above, $u_{r,s}^T \cdot v_{r',s'} = 0$ at the end of round $g+1$ as well, as required to establish part (b).

Finally, let us consider part (c) of the claim.

If $g+1$ is an odd integer then there is nothing to be done to establish this, because (when $g+1$ is being considered) the claim is vacuous.

With that noted, suppose, instead, that $g+1$ is an even integer. Then $g \geq 1$ and $g-1$ is a nonnegative even integer — so that the inductive hypothesis can be applied to make conclusions about the state of vectors after round $g-1$.

Consider a vector $u_{r,s}$ that is unmatched after round $g+1$, where $0 \leq r < i - \Delta_{n,k} + (g+1)/2$ and $1 \leq s \leq k$. Either $0 \leq r < i - \Delta_{n,k} + (g-1)/2$ as well, or $r = i - \Delta_{n,k} + (g-1)/2$; these subcases are considered separately below.

Subcase: $0 \leq r < i - \Delta_{n,k} + (g-1)/2$. In this case, since $g-1$ is also a nonnegative integer, it follows by the inductive hypothesis that $u_{r,s}^T \cdot v_{a,b} = 0$ for every vector $v_{a,b}$ such that $0 \leq a \leq i$ and $1 \leq b \leq k$. Furthermore, $u_{r',s'}^T \cdot v_{a',b'} = 0$ for every vector $u_{r',s'}$ such that $0 \leq r' \leq r$, $1 \leq s' \leq k$, and $u_{r',s'}$ was also unmatched at the end after round $g-1$, and for every vector $v_{a',b'}$ such that $0 \leq a' \leq i$ and $1 \leq b' \leq k$, as well.

Once again, Lemmas A.8 and A.9 establish that the updates in rounds g and $g+1$ only update the value of $u_{r,s}$, for r and s as above, by replacing it with a linear combination of the values of vectors $u_{r',s'}$ such that $0 \leq r' \leq r$, $1 \leq s' \leq k$, and $u_{r',s'}$ was also unmatched at the end of round $g-1$ — as these values were defined at the end of round $g-1$. They only update the value of $v_{a,b}$ by replacing it with a linear combination of the values of vectors $v_{a',b'}$ such that $0 \leq a' \leq i$ and $1 \leq b' \leq k$, as these values were defined at the end of round $g-1$, as well. Consequently $u_{r,s}^T \cdot v_{a,b} = 0$ at the end of round $g+1$.

The argument needed to establish that $u_{a,b}^T \cdot v_{r,s}$ if $0 \leq r < i - \Delta_{n,k} + (g-1)/2$, $1 \leq s \leq k$, $v_{r,s}$ is unmatched at the end of round $g+1$, $0 \leq a \leq i$ and $1 \leq b \leq k$, is the same — the roles of vectors $u_{r,s}$ and $v_{r,s}$ need only be exchanged.

Subcase: $r = i - \Delta_{n,k} + (g-1)/2$. In this case, part (c) can be established by noticing that rounds g and $g+1$ include matching phases in which unmatched vectors $u_{r,s}$ (respectively, $v_{r,s}$) are matched with unmatched vectors $v_{i,s'}$ (respectively, $u_{i,s'}$).

Now, the unmatched vectors $u_{r,s}$ are orthogonal to all *matched* vectors $v_{r',s'}$ such that $0 \leq r' \leq i$ and $1 \leq s \leq k$, because Invariant #3 has been re-established, as noted above.

They are orthogonal to all *unmatched* vectors $v_{r',s'}$ such that $0 \leq r' \leq i-1$ and $1 \leq s' \leq k$ because property (b) holds once again (as established above), as well.

Finally, to see that they are also orthogonal to all *unmatched* vectors $v_{i,s'}$, recall that rounds g and $g+1$ including a matching step in which the vectors $u_{r,s}$, that are unmatched at the beginning of round g, are matched with the unmatched vectors $v_{i,s'}$ that are unmatched at the beginning of this round. Since Invariants #2 and #3 hold at the end of this round (as noted above), this is now a consequence of Lemma A.10, above.

The corresponding result for unmatched vectors $v_{r,s}$, that is needed to establish part (c), follows by the same argument.

The claim now follows by induction on g. □

LEMMA A.13. *Let i be an integer such that $i \geq 1$ and there are at least $i+1$ stages of the Lanczos phase of the computation (ending with stage i). Suppose, as well, that Invariants #2–7 are satisfied at the end of stage $i-1$ of the computation.*

Then Invariants #2–4, 6 and 7 are satisfied at the end of the end of stage i of the Lanczos phase as well.

PROOF. It follows by Lemma A.12 that Invariants #2, 3, 6 and 7 are satisfied at the end of stage i, so it remains only

to establish Invariant #4.

It also follows by the above lemma that, at the end of round $2 \cdot \Delta_{n,k}$ of stage i, $u_{a,b}^T \cdot v_{c,d} = 0$ for all integers a, b, c and d such that $0 \leq a \leq i-1$, $1 \leq b \leq k$, $u_{a,b}$ is unmatched after round $2 \cdot \Delta_{n,k}$, $0 \leq c \leq i$ and $1 \leq d \leq k$. It follows by this lemma, as well, that, at the end of round $2 \cdot \Delta_{n,k}$, $u_{a,b}^T \cdot v_{c,d} = 0$ for all integers a, b, c and d such that $0 \leq a \leq i$, $1 \leq b \leq k$, $0 \leq c \leq i-1$, $1 \leq d \leq k$, and $v_{c,d}$ is unmatched after this round.

Lemmas A.8 and A.9 imply that the above orthogonality conditions are still satisfied at the end of round $2 \cdot \Delta_{n,k}+1$ — that is, at the end of stage i of the Lanczos phase — as well: For, after round $2 \cdot \Delta_{n,k} + 1$, the value of each unmatched vector $u_{a,b}$ (respectively, $v_{a,b}$) for $0 \leq a \leq i-1$ and $1 \leq b \leq k$ is a linear combination of the values of the vectors $u_{a',b'}$ (respectively, $v_{a',b'}$) such that $0 \leq a' \leq a$ and $1 \leq b' \leq k$, as defined at the beginning of this round. Similarly, the value of each vector $v_{c,d}$ (respectively, $u_{c,d}$) after this round is a linear combination of the values of vectors $v_{c',d'}$ (respectively, $u_{c',d'}$) such that $0 \leq c' \leq c$ and $1 \leq d' \leq k$ as defined at the beginning of this round as well.

It remains only to establish that if $1 \leq a, b \leq k$ and $u_{i,a}$ and $v_{i,b}$ are both unmatched after round $2 \cdot \Delta_{n,k} + 1$, then $u_{i,a}^T \cdot v_{i,b} = 0$ as well. Since round $2 \cdot \Delta_{n,k} + 1$ includes a matching step in which unmatched vectors $u_{i,a}$ are matched with unmatched vectors $v_{i,b}$, and Invariants #2 and 3 are satisfied after this round, this follows by Lemma A.10. □

LEMMA A.14. *Let i be an integer such that $i \geq 0$ and there are at least $i+1$ stages of the Lanczos phase of the algorithm. Then Invariants #2–#7 are satisfied at the end of each of the first i stages and Invariants #2–#4, 6 and 7 are satisfied at the end of stage i (that is, the $i+1^{st}$ stage) as well.*

PROOF. The claim can now be established by induction on i.

It follows by Lemma A.11 that Invariants #2–4, 6 and 7 are satisfied at the end of stage 0 of the Lanczos phase, as required to establish the basis.

Suppose (for the inductive step) that $i \geq 0$, there are at least $i+2$ stages of the Lanczos phase, ending with with stage $i+1$, and that Invariants #2–4, 6 and 7 are satisfied at the end of stage i. Invariant #5 must also be satisfied at this point, because the Lanczos phase of the computation does not end with stage i. Lemma A.13 now implies that Invariants #2–4, 6 and 7 are established after stage $i+1$ as well, as needed to complete the proof. □

PROOF OF LEMMA 2.1. This is now a straightforward consequence of Lemmas A.4, A.6 and A.14, above. □

PROOF OF LEMMA 2.2. This is now a straightforward corollary of Lemmas A.4 and A.14. □

A.2.3 Bounding the Cost of the Lanczos Phase

LEMMA A.15. *Suppose $i \geq 0$ and there are at least $i+1$ stages of the Lanczos phase (ending with stage i). Then there are at most $\Delta_{n,k} \cdot k$ unmatched vectors $u_{r,s}$ such that $0 \leq r \leq i-1$ and $1 \leq s \leq k$, and at most $\Delta_{n,k} \cdot k$ unmatched vectors $v_{r,s}$ such that $0 \leq r \leq i-1$ and $1 \leq s \leq k$, at the end of stage $i-1$.*

PROOF. Since the Lanczos phase did not end at stage $i-1$, Invariant #5 was satisfied at the end of this stage — so that the vectors $u_{r,s}$ and $v_{r,s}$ such that $0 \leq r \leq i-1-\Delta_{n,k}$ and $1 \leq s \leq k$ were all unmatched. Consequently the only vectors that could be unmatched at this point are the vectors $u_{r,s}$ (respectively, $v_{r,s}$ such that $i - \Delta_{n,k} \leq r \leq i-1$ and $1 \leq s \leq k$. The claim now follows since only $\Delta_{n,k} \cdot k$ such vectors exist. □

The following lemma is now easily established using Lemmas A.1, A.2, A.3, the fact (by inspection of the code) that $O(\Delta_{n,k} \cdot k)$ matched vectors are stored at any point, and fact (established by the lemma above) that there are at most $\Delta_{n,k} \cdot k$ unmatched vectors $u_{r,s}$ (respectively, $v_{r,s}$) at the beginning of any stage of the Lanczos phase, as well.

LEMMA A.16. *The cost of the initialization step and each stage of the Lanczos phase can be bounded as follows.*

(a) *The initialization step requires $k+1$ multiplications of vectors by A, the selection of $2k$ vectors uniformly and independently from $\mathsf{F}_q^{n \times 1}$, and $O(n)$ additional operations over F_q.*

(b) *Stage 0 requires $O(n \cdot k^2)$ additional operations over F_q.*

(c) *If $i \geq 1$ and there are at least $i+1$ stages of the Lanczos phase (ending with stage i), then stage i requires k multiplications of vectors by A, k multiplications of vectors by A^T, and $O(\Delta_{n,k}^2 \cdot k^2 \cdot n)$ additional operations over F_q.*

(d) *$O(\Delta_{n,k} \cdot k)$ vectors and additional values are stored at any time, so the algorithm requires space required to store $O(\Delta_{n,k} \cdot k \cdot n)$ elements of F_q as well as $O(\Delta_{n,k} \cdot k)$ nonnegative integers with values between 0 and n.*

A.3 A More Detailed Description of the Elimination Phase

A.3.1 Initialization

The following procedure defines matrices $\mathcal{M}_L, \mathcal{M}_R, \mathcal{M}_R^{pre} \in \mathsf{F}_q^{n \times t}$ that will be used (as global data) for orthogonalization steps during the Lanczos phase.

procedure setup

1. If the entries of the list M (used during the Lanczos phase) are

$$((r_1, s_1), (r_1', s_1')), ((r_2, s_2), (r_2', s_2')), \ldots, ((r_t, s_t), (r_t', s_t'))$$

then set $\mathcal{M}_L \in \mathsf{F}_q^{n \times t}$ to be the matrix with columns

$$u_{r_1,s_1}, u_{r_2,s_2}, \ldots, u_{r_t,s_t},$$

set $\mathcal{M}_R \in \mathsf{F}_q^{n \times t}$ to be the matrix with columns

$$v_{r_1',s_1'}, v_{r_2',s_2'}, \ldots, v_{r_t',s_t'},$$

and set $\mathcal{M}_R^{pre} \in \mathsf{F}_q^{n \times t}$ to be the matrix with columns

$$w_{r_1',s_1'}, w_{r_2',s_2'}, \ldots, w_{r_t',s_t'}$$

— so that $A \cdot \mathcal{M}_R^{pre} = \mathcal{M}_R$ and $\mathcal{M}_L^T \cdot \mathcal{M}_R = I_t$.

end procedure

The following lemma is a consequence of the fact that Invariants #2 and #6 were satisfied at the end of the Lanczos phase of the computation (and inspection of the code).

LEMMA A.17. *Procedure* setup*, above, produces matrices* $\mathcal{M}_L, \mathcal{M}_R, \mathcal{M}_R^{pre} \in \mathsf{F}_q^{n\times t}$ *such that* $\mathcal{M}_L^T \cdot \mathcal{M}_R = I_t \in \mathsf{F}_q^{t\times t}$ *and such that* $A \cdot \mathcal{M}_R^{pre} = \mathcal{M}_R$*.*

A.3.2 An "Orthogonalization" Step

A routine orthogonalize will receive as inputs a pair of matrices $\mathcal{M}_{new}, \mathcal{M}_{new}^{pre} \in \mathsf{F}_q^{n\times h}$ such that $A \cdot \mathcal{M}_R^{pre} = \mathcal{M}_R$ and will ensure that the columns of $\mathcal{M}_{new}$ are orthogonal to the vectors $\mu_1, \mu_2, \ldots, \mu_\ell$ that had been produced at the end of the Lanczos phase of the algorithm.

procedure orthogonalize($\mathcal{M}_{new}, \mathcal{M}_{new}^{pre}$)
1. $D := \mathcal{M}_L^T \cdot \mathcal{M}_{new} \in \mathsf{F}_q^{t\times h}$
2. $\mathcal{M}_{new} := \mathcal{M}_{new} - \mathcal{M}_R \cdot D$
3. $\mathcal{M}_{new}^{pre} := \mathcal{M}_{new}^{pre} - \mathcal{M}_R^{pre} \cdot D$
4. return($\mathcal{M}_{new}$, $\mathcal{M}_{new}^{pre}$)

The following lemma is of use in establishing the correctness of this procedure.

LEMMA A.18. *Suppose that* $j \geq 1$*, there are at least* $j+1$ *stages of the elimination phase, and that Invariant #9 is satisfied at the beginning of stage* j*. Then, if* t *is an integer such that either* $\mu_t = u_{g,h}$ *or* $\nu_t = v_{g,h}$ *where* $0 \leq g \leq i - 2 \cdot \Delta_{n,k} - 3$ *and* $1 \leq h \leq k$*, then* $\mu_t^T \cdot (A \cdot \varphi_s) = 0$ *for every integer* s *such that* $1 \leq s \leq g$ *at the beginning of stage* j *of the elimination phase.*

PROOF. Note first that, since Invariant #5 was satisfied at the end of round $i-1$ of the Lanczos phase, it follows by Lemma A.5 that if t is as above then

$$A^T \cdot \mu_t = \sum_{t=1}^{\ell} \alpha_{a,t} \cdot \mu_a \tag{24}$$

where $\alpha_{a,t} \in \mathsf{F}_q$ for $1 \leq a \leq \ell$ and, furthermore, $\alpha_{a,t} = 0$ unless the vector μ_a was matched at or before the end of round $i-2$ of the Lanczos phase.

The cases $j = 1$ and $j \geq 2$ are considered separately below.

Case: $j = 1$. Examining the details of stage 0 of the elimination phase one should note that, at the beginning of stage 1, every vector φ_r such that $1 \leq r \leq g$ is a linear combination of

- vectors $v_{s,t}$ such that $0 \leq s \leq i$, $1 \leq t \leq k$, and the vector $v_{s,t}$ was unmatched at the end of stage i of the Lanczos phase, and
- vectors $v_{i,t}$ such that $1 \leq t \leq k$ and $v_{i,t}$ was matched at the end of stage i of the Lanczos phase.

Now, if $v_{s,t}$ was unmatched at the end of stage i of the Lanczos phase then it follows by Invariant #4 (which was satisfied at that point) that $\mu_a^T \cdot v_{s,t} = 0$ for $1 \leq a \leq \ell$. On the other hand, if $v_{i,t}$ was matched at the end of stage i and μ_a was matched at or before the end of stage $i-2$ of the Lanczos phase then $v_{i,t} = \nu_w$ for an integer $w > a$, since $v_{i,t}$ was necessarily matched during round i. Invariant #2 (which was also satisfied at the end of stage i of the Lanczos phase) implies that $\mu_a^T \cdot v_{i,t} = \mu_a^T \cdot \nu_w = 0$ as well.

It follows that $\mu_a^T \cdot \varphi_r$ for every integer a such that $\alpha_{t,a} \neq 0$, and the equation shown above at line (24) implies that

$$\mu_t^T \cdot (A \cdot \varphi_r) = (A^T \cdot \mu_t)^T \cdot \varphi_r = 0.$$

Case: $j \geq 2$. The argument for this case is similar but simpler. Notice that, after the application of the update step at the end of stage $j-1$ (and the beginning of stage j), each vector φ_r (such that $1 \leq r \leq g$) is equal to one of the vectors λ_s that was included during stage $j-1$ — so that Invariant #9 implies that $\mu_a^T \cdot \varphi_r = \mu_a^T \cdot \lambda_s = 0$ for $1 \leq a \leq \ell$.

Once again, the equation at line (24) can now be used to establish that

$$\mu_t^T \cdot (A \cdot \varphi_r) = (A^T \cdot \mu_t)^T \cdot \varphi_r = 0$$

as required. □

Since the above procedure is not called during stage 0 of the elimination phase at all, this suffices to establish (the harder part of) the following as well — note that, before step 2 of the procedure orthogonalize,

$$\mathcal{M}_L^T \cdot (\mathcal{M}_{new} - \mathcal{M}_R \cdot D) = \mathcal{M}_L^T \cdot \mathcal{M}_{new} - D = 0 \in \mathsf{F}_q^{t\times h},$$

so that $\mathcal{M}_L^T \cdot \mathcal{M}_{new} = 0$ after step 2 has been carried out.

LEMMA A.19. *Suppose that procedure* orthogonalize *is executed with a pair of matrices* $\mathcal{M}_{new}, \mathcal{M}_{new}^{pre} \in \mathsf{F}_q^{n\times h}$ *such that* $A \cdot \mathcal{M}_{new}^{pre} = \mathcal{M}_{new}$ *as its inputs. Then, on termination of this procedure, the following properties are satisfied.*

(a) *The subspace of* $\mathsf{F}_q^{n\times 1}$ *spanned by the vectors*

$$\nu_1, \nu_2, \ldots, \nu_\ell$$

and the columns of $\mathcal{M}_{new}$ *will not have been changed.*

(b) $\mu_r^T \cdot \mathcal{M}_{new} = 0$ *for* $1 \leq r \leq \ell$*.*

(c) $A \cdot \mathcal{M}_{new}^{pre} = \mathcal{M}_{new}$*.*

(d) *No matrices (or other data) except for the matrices* $\mathcal{M}_{new}$ *and* $\mathcal{M}_{new}^{pre}$ *will have been changed.*

The procedure uses $O(\Delta_{n,k} \cdot k \cdot h \cdot n)$ *operations over* F_q *using standard arithmetic.*

A.3.3 An "Elimination" Step

A routine eliminate will receive matrices $\mathcal{M}_{new}, \mathcal{M}_{new}^{pre} \in \mathsf{F}_q^{n\times h}$ and will ensure that, following the update,

$$P \cdot \mathcal{M}_{new} = \begin{bmatrix} 0 \\ \widehat{\mathcal{M}}_{new} \end{bmatrix}$$

for a matrix $\widehat{\mathcal{M}}_{new} \in \mathsf{F}_q^{(n-m)\times h}$. The algorithm makes use of the permutation P and matrices M_κ and M_λ as defined for the elimination stage (and considered in Invariants #12 and #13).

procedure eliminate($\mathcal{M}_{new}, \mathcal{M}_{new}^{pre}$)
1. Let $Y \in \mathsf{F}_q^{m\times h}$ and let L_λ be the nonsingular lower triangular matrix in $\mathsf{F}_q^{m\times m}$ such that

$$P \cdot \mathcal{M}_{new} = \begin{bmatrix} Y \\ Z \end{bmatrix} \text{ and } P \cdot M_\lambda = \begin{bmatrix} L_\lambda \\ X_\lambda \end{bmatrix}$$

for matrices $Z \in \mathsf{F}_q^{(n-m)\times h}$ and $X_\lambda \in \mathsf{F}_q^{(n-m)\times m}$.
2. $W := L_\lambda^{-1} \cdot Y \in \mathsf{F}_q^{m\times h}$
3. $\mathcal{M}_{new} := \mathcal{M}_{new} - M_\lambda \cdot W$
4. $\mathcal{M}_{new}^{pre} := \mathcal{M}_{new}^{pre} - M_\kappa \cdot W$
end procedure

Lemma A.20, below, follows by inspection of the code: Note that, after step 2,

$$P \cdot M_\lambda \cdot W = \begin{bmatrix} L_\lambda \\ X_\lambda \end{bmatrix} \cdot W = \begin{bmatrix} Y \\ X_\lambda \cdot W \end{bmatrix}$$

so it will be true that

$$P \cdot \mathcal{M}_{new} = \begin{bmatrix} 0 \\ \widehat{\mathcal{M}}_{new} \end{bmatrix}$$

after step 3, as required. Since $A \cdot M_\kappa = M_\lambda$, it will be true that $A \cdot \mathcal{M}^{pre}_{new} = \mathcal{M}_{new}$ after step 4, as well.

LEMMA A.20. *Suppose that procedure* eliminate *is executed with a pair of matrices $\mathcal{M}_{new}, \mathcal{M}^{pre}_{new} \in \mathsf{F}_q^{n \times h}$, such that $\mu_r^T \cdot \mathcal{M}_{new} = 0$ for $1 \le r \le \ell$ and $A \cdot \mathcal{M}^{pre}_{new} = \mathcal{M}_{new}$, as its inputs.*

Suppose as well, that Invariants #10–12 are satisfied when the procedure is executed.

Then, on termination, the following properties are satisfied.

(a) *The subspace of $\mathsf{F}_q^{n \times 1}$ spanned by the columns of M_λ and $\mathcal{M}_{new}$ has not been changed.*

(b) *$\mu_r^T \cdot \mathcal{M}_{new} = 0$ for $1 \le r \le \ell$.*

(c) *$P \cdot \mathcal{M}_{new} = \begin{bmatrix} 0 \\ \widehat{\mathcal{M}}_{new} \end{bmatrix}$ for a matrix $\widehat{\mathcal{M}}_{new} \in \mathsf{F}_q^{(n-m) \times h}$.*

(d) *$A \cdot \mathcal{M}^{pre}_{new} = \mathcal{M}_{new}$.*

(e) *No matrices (or other data) except for the matrices $\mathcal{M}_{new}$ and $\mathcal{M}^{pre}_{new}$ will have been changed.*

The procedure uses $O(nmh)$ operations over F_q using standard arithmetic.

A.3.4 A "Compression" Step

The next procedure receives, as input, a pair of matrices $V \in \mathsf{F}_q^{r \times s}$ and $W \in \mathsf{F}_q^{n \times s}$ (for $r \le n$) such that

$$A \cdot W = P^T \cdot \begin{bmatrix} 0 \\ V \end{bmatrix} \tag{25}$$

and, reducing r as needed, modifies V in order to ensure that the column space of the matrix has been unchanged but the columns are now linearly independent — while ensuring that an equation as shown at line (25) is satisfied, once again.

procedure compress(V, W)
1. Compute the rank t of V as well as a matrix $X \in \mathsf{F}_q^{s \times t}$ such that the matrix $V \cdot X$ has full rank t.
2. $V := V \cdot X$
3. $W := W \cdot X$
4. $s := t$
5. return(V, W)

end procedure

Lemma A.21, below, follows by inspection of the above code.

LEMMA A.21. *Suppose that procedure* compress *is executed with a pair of matrices $V \in \mathsf{F}_q^{r \times s}$ and $W \in \mathsf{F}_q^{n \times s}$ such the equation at line (25), above is satisfied.*

Then, on termination of the procedure, the following properties are satisfied.

(a) *The column space of V has not been changed (but the number s of columns in each of the matrices V and W may have been reduced).*

(b) *An equation of the form shown at line (25) is satisfied, once again.*

(c) *The columns of V are linearly independent.*

(d) *No matrices (or other data) except for V and W have been changed.*

The procedure uses $O(n \cdot s^2)$ operations over F_q using standard arithmetic.

A.3.5 A "Triangularization" Step

The next procedure receives, as inputs a pair of matrices $V \in \mathsf{F}_q^{r \times s}$ and $W \in \mathsf{F}_q^{n \times s}$ (for $r \le n$) such that the equation at line (25) is satisfied, and the columns of V are linearly independent.

It modifies V and W in such a way that the column space and dimension of V have not been changed, and the equation at line (25) is once again satisfied, but there exists a permutation matrix $\widehat{P} \in \mathsf{F}_q^{r \times r}$ such that the top r rows of $\widehat{P} \cdot V$ is a lower triangular matrix with ones on its diagonal. The permutation matrix $\widehat{P}$ is also returned as output.

procedure triangularize(V, W)
1. Compute a $\widetilde{P}$-$\widehat{L}$-$\widehat{U}$ factorization of V —that is, compute a permutation matrix $\widetilde{P} \in \mathsf{F}_q^{r \times r}$, a lower triangular matrix $\widehat{U} \in \mathsf{F}_q^{r \times s}$ with ones on its diagonal, and a nonsingular upper triangular matrix $\widehat{U} \in \mathsf{F}_q^{s \times s}$ such that $V = \widetilde{P} \cdot \widehat{L} \cdot \widehat{U}$.
2. $\widehat{U}^{inv} := \widehat{U}^{-1}$
3. $V := V \cdot \widehat{U}^{inv}$
4. $W : W \cdot \widehat{U}^{inv}$
5. $\widehat{P} := \widetilde{P}^T$
6. return $(V, W, \widehat{P})$

end procedure

Lemma A.22 follows by inspection of the above code.

LEMMA A.22. *Suppose the procedure* triangularize *is executed with a pair of matrices $V \in \mathsf{F}_q^{r \times s}$ and $W \in \mathsf{F}_q^{n \times s}$ as inputs such that the equation at line (25) is satisfied, and the columns of V are linearly independent,.*

Then, on termination of the procedure, the following properties are satisfied.

(a) *The column space of V has not been changed.*

(b) *An equation of the form shown at line (25) is satisfied, once again.*

(c) *$\widehat{P} \cdot V = \widehat{L}$, where $\widehat{P} \in \mathsf{F}_q^{r \times r}$ is a permutation matrix and $\widehat{L} \in \mathsf{F}_q^{r \times s}$ is a lower triangular matrix with ones on its diagonal.*

The procedure uses $O(n \cdot s^2)$ operations over F_q using standard arithmetic.

A.3.6 "Solution" and "Update" Steps

The final procedure completes a stage of the elimination phase. It also receives as input a permutation matrix $\widehat{P} \in \mathsf{F}_q^{(n-m) \times (n-m)}$ and a matrix $\widehat{\mathcal{M}}_{new} \in \mathsf{F}_q^{(n-m) \times h}$ such that

$$P \cdot \mathcal{M}_{new} = \begin{bmatrix} 0 \\ \widehat{\mathcal{M}}_{new} \end{bmatrix}$$

and such that

$$\widehat{P} \cdot \widehat{\mathcal{M}}_{new} = \begin{bmatrix} \widetilde{L} \\ X \end{bmatrix}$$

where $\widetilde{L} \in \mathsf{F}_q^{h \times h}$ is lower triangular with ones on its diagonal and where $X \in \mathsf{F}_q^{(n-m-h) \times h}$.

procedure solve_and_update($\widehat{P}, \mathcal{M}_{new}, \widehat{\mathcal{M}}_{new}, \mathcal{M}_{new}^{pre}$)

1. Set $\widehat{\rho} \in \mathsf{F}_q^{(n-m) \times 1}$ to be the vector such that

$$P \cdot \rho = \begin{bmatrix} 0 \\ \widehat{\rho} \end{bmatrix}$$

(see Invariant #13, above).

2. Set $\widetilde{L} \in \mathsf{F}_q^{h \times h}$ to be the lower triangular matrix with ones on its diagonal such that

$$\widehat{P} \cdot \widehat{\mathcal{M}}_{new} = \begin{bmatrix} \widetilde{L} \\ X \end{bmatrix}$$

for $X \in \mathsf{F}_q^{(n-m-h) \times h}$ and set $\widehat{\rho}_1 \in \mathsf{F}_q^{h \times 1}$ to be the vector such that

$$\widehat{P} \cdot \widehat{\rho} = \begin{bmatrix} \widehat{\rho}_1 \\ y \end{bmatrix}$$

for a vector $y \in \mathsf{F}_q^{(n-m-h) \times 1}$.

3. $\eta := \widetilde{L}^{-1} \cdot \widehat{\rho}_1 \in \mathsf{F}_q^{h \times 1}$
4. $\rho := \rho - \mathcal{M}_{new} \cdot \eta$
5. $\chi := \chi + \mathcal{M}_{new}^{pre} \cdot \eta$
6. $P := \begin{bmatrix} I_m & \\ & \widehat{P} \end{bmatrix} \cdot P \in \mathsf{F}_q^{n \times n}$
7. $M_\varphi := \mathcal{M}_{new}$ (so that g is now equal to the number of columns in $\mathcal{M}_{new}$)
8. Append the columns of the matrices $\mathcal{M}_{new}$ and $\mathcal{M}_{new}^{pre}$ onto the matrices V_λ and V_κ respectively (adding h to the value of m in the process).

end procedure

LEMMA A.23. *Suppose the procedure* solve_and_update *is executed with inputs $\widehat{P} \in \mathsf{F}_q^{m \times m}$, $\mathcal{M}_{new} \in \mathsf{F}_q^{n \times h}$, $\widehat{\mathcal{M}}_{new} \in \mathsf{F}_q^{(n-m) \times h}$, and $\mathcal{M}_{new}^{pre} \in \mathsf{F}_q^{n \times h}$, such that the following properties are satisfied.*

(a) *Invariants #8–10 and #12–13 are satisfied.*

(b) $P \cdot \mathcal{M}_{new} = \begin{bmatrix} 0 \\ \widehat{\mathcal{M}}_{new} \end{bmatrix}$.

(c) $\widehat{P} \cdot \widehat{\mathcal{M}}_{new}$ *is lower triangular with ones on its diagonal.*

(d) $\mu_a^T \cdot \mathcal{M}_{new} = 0$ *for* $1 \leq a \leq \ell$.

(e) $A \cdot \mathcal{M}_{new}^{pre} = \mathcal{M}_{new}$.

Let $\mathcal{X} \subseteq \mathsf{F}_q^{n \times 1}$ be the column space of the matrix

$$\begin{bmatrix} M_\lambda & \mathcal{M}_{new} \end{bmatrix} \in \mathsf{F}_q^{n \times (m+r)}$$

and let $\mathcal{Y}$ be the columnspace of the matrix $\mathcal{M}_{new}$ (for the matrix M_λ as defined at the beginning of the execution of the procedure). Then the following properties are satisfied on termination of the procedure.

(a) M_λ *now has column space* $\mathcal{X}$.

(b) *The columns of M_φ are linearly independent and M_φ has column space $\mathcal{Y}$.*

(c) *Invariants #9, 10, 12 and #13 are satisfied once again.*

The procedure uses $O(n \cdot h)$ operations over F_q using standard arithmetic.

PROOF. Since the columns of $\mathcal{M}_{new}$ have been appended as new columns of M_λ at step 8, above, and these matrices are otherwise unchanged, part (a) of the claim follows by inspection of the code.

Following step 7 (and the rest of the procedure) the columns of M_φ are those of the above matrix $\mathcal{M}_{new}$. Now, as noted above (and using the original value of the matrix P here),

$$\begin{bmatrix} I_m & \\ & \widehat{P} \end{bmatrix} \cdot P \cdot \mathcal{M}_{new} = \begin{bmatrix} 0 \\ \widetilde{L} \\ X \end{bmatrix} \tag{26}$$

where $\widetilde{L} \in \mathsf{F}_q^{h \times h}$ is a lower triangular matrix with ones on its diagonal — establishing the linear independence of the columns of $\mathcal{M}_{new}$ (and of M_φ).

Since $\mathcal{Y}$ is the column space of the matrix $\mathcal{M}_{new}$ it follows by inspection of the code (specifically, step 7) that M_φ now has column space $\mathcal{Y}$, as needed to establish part (b).

Since $\mu_a^T \cdot \mathcal{M}_{new} = 0$ for $1 \leq a \leq \ell$, it follows by inspection of the code (specifically, step 8) that Invariant #9 is satisfied on termination of this procedure if it was satisfied before it.

Since P is modified to have value

$$\begin{bmatrix} I_m & \\ & \widehat{P} \end{bmatrix} \cdot P$$

at step 6, and the columns of $\mathcal{M}_{new}$ are appended to M_λ at step 8, the equation at line (26), above, establishes that Invariant #10 will be satisfied on termination of this procedure if it was initially satisfied, as well.

Since $A \cdot \mathcal{M}_{new}^{pre} = \mathcal{M}_{new}$, one can see by inspection of the code (specifically, step 8) that Invariant #12 is satisfied at the end of the execution of this procedure if it was satisfied before it too.

In order to consider Invariant #13 let us denote by χ and ρ the values of these vectors before the execution of this procedure, and denote by χ' and ρ' the values of these vectors after it. Then

$$\chi' = \chi + \mathcal{M}_{new}^{pre} \cdot \eta \qquad \text{and} \qquad \rho' = \rho - \mathcal{M}_{new} \cdot \eta$$

for the vector $\eta \in \mathsf{F}_q^{h \times 1}$ defined at step 3, above. Then

$$\begin{aligned} A \cdot \chi' + \rho' &= A \cdot (\chi + \mathcal{M}_{new}^{pre} \cdot \eta) + (\rho - \mathcal{M}_{new} \cdot \eta) \\ &= (A \cdot \chi + \rho) + (\mathcal{M}_{new} \cdot \eta - \mathcal{M}_{new} \cdot \eta) \\ &\qquad\qquad \text{(since } A \cdot \mathcal{M}_{new}^{pre} = \mathcal{M}_{new}\text{)} \\ &= A \cdot \chi + \rho \\ &= A \cdot w + b, \end{aligned}$$

since Invariant #13 was initially satisfied. Furthermore, if $1 \leq a \leq \ell$ then

$$\begin{aligned} \mu_a^T \cdot \rho' &= \mu_a^T \cdot (\rho - \mathcal{M}_{new} \cdot \eta) \\ &= \mu_a^T \cdot \rho \qquad\qquad \text{(since } \mu_a^T \cdot \mathcal{M}_{new} = 0\text{)} \\ &= 0, \end{aligned}$$

once again, because Invariant #13 was initially satisfied.

Now let P and P' denote the values of the permutation matrix P before and after the execution of this procedure,

respectively. Then

$$\begin{aligned} P' \cdot \rho' &= \begin{bmatrix} I_m & \\ & \widehat{P} \end{bmatrix} \cdot P \cdot (\rho - \mathcal{M}_{new} \cdot \eta) \\ &= \begin{bmatrix} I_m & \\ & \widehat{P} \end{bmatrix} \cdot \left(\begin{bmatrix} 0_m \\ \widehat{\rho} \end{bmatrix} - \begin{bmatrix} 0_m \\ \widehat{\mathcal{M}}_{new} \end{bmatrix} \cdot \eta \right) \\ &= \begin{bmatrix} 0_m \\ \widehat{\rho}_1 \\ y \end{bmatrix} - \begin{bmatrix} 0_m \\ \widetilde{L} \\ X \end{bmatrix} \cdot \widetilde{L}^{-1} \cdot \widehat{\rho}_1 \\ &= \begin{bmatrix} 0_m \\ \widehat{\rho}_1 \\ y \end{bmatrix} - \begin{bmatrix} 0_m \\ \widehat{\rho}_1 \\ X \cdot \widetilde{L}^{-1} \cdot \widehat{\rho}_1 \end{bmatrix} \\ &= \begin{bmatrix} 0_{m+h} \\ y - X \cdot \widetilde{L}^{-1} \cdot \widehat{\rho}_1 \end{bmatrix} \end{aligned}$$

as needed to re-establish Invariant #13 and to establish part (c) of the claim.

The claimed bound on the cost of this procedure follows by inspection of the code — noting that, since $\widetilde{L} \in \mathsf{F}_q^{h \times h}$ is lower triangular with ones on its diagonal, the cost to compute the vector η at step 3 is in $O(h^2)$. □

A.3.7 The Main Method

// Initialization

1. setup

// Stage #0

2. Set h to be the number of vectors $v_{r,s}$ such that $0 \leq r \leq i-1$, $1 \leq s \leq k$, and $v_{r,s}$ was unmatched at the end of (the final) stage i of the Lanczos phase. Set $\mathcal{M}_{new}, \mathcal{M}^{pre}_{new} \in \mathsf{F}_q^{h \times h}$ to be the matrices whose columns are the above vectors $v_{r,s}$ and corresponding vectors $w_{r,s}$, so that $A \cdot \mathcal{M}^{pre}_{new} = \mathcal{M}_{new}$.
3. Set m and g to be zero (since M_λ, M_κ and M_φ have zero columns) and set P to be the identity matrix I_n.
4. if $(h > 0)$ then
5. $\quad (\mathcal{M}_{new}, \mathcal{M}^{pre}_{new}) := \mathsf{compress}(\mathcal{M}_{new}, \mathcal{M}^{pre}_{new})$
6. $\quad$ if ($\mathcal{M}_{new}$ has at least one column) then
7. $\quad\quad (\mathcal{M}_{new}, \mathcal{M}^{pre}_{new}, \widehat{P}) := \mathsf{triangularize}(\mathcal{M}_{new}, \mathcal{M}^{pre}_{new})$
8. $\quad\quad \mathsf{solve\_and\_update}(\widehat{P}, \mathcal{M}_{new}, \mathcal{M}_{new}, \mathcal{M}^{pre}_{new})$
   $\quad$ end if
   end if
9. Once again, set g to be 0 (removing all columns from M_φ). Set h to be the number of vectors $v_{i,s}$ such that $1 \leq s \leq k$ and $v_{i,s}$ was unmatched at the end of (the final) stage i of the Lanczos phase. Set $\mathcal{M}_{new}, \mathcal{M}^{pre}_{new} \in \mathsf{F}_q^{n \times h}$ to be the matrices whose columns are the above vectors $v_{i,s}$ (and corresponding vectors $w_{i,s}$), so that $A \cdot \mathcal{M}^{pre}_{new} = \mathcal{M}_{new}$, once again.
10. if $(h > 0)$ then
11. $\quad (\mathcal{M}_{new}, \mathcal{M}^{pre}_{new}) := \mathsf{eliminate}(\mathcal{M}_{new}, \mathcal{M}^{pre}_{new})$
12. $\quad$ Set $\widehat{\mathcal{M}}_{new} \in \mathsf{F}_q^{(n-m) \times h}$ to be the matrix such that

$$P \cdot \mathcal{M}_{new} = \begin{bmatrix} 0 \\ \widehat{\mathcal{M}}_{new} \end{bmatrix}.$$

13. $\quad (\widehat{\mathcal{M}}_{new}, \mathcal{M}^{pre}_{new}) := \mathsf{compress}(\widehat{\mathcal{M}}_{new}, \mathcal{M}^{pre}_{new})$
14. $\quad$ if ($\widehat{\mathcal{M}}_{new}$ has at least one column) then
15. $\quad\quad (\widehat{\mathcal{M}}_{new}, \mathcal{M}^{pre}_{new}, \widehat{P}) := \mathsf{triangularize}(\widehat{\mathcal{M}}_{new}, \mathcal{M}^{pre}_{new})$
16. $\quad\quad \mathcal{M}_{new} := P^T \cdot \begin{bmatrix} 0 \\ \widehat{\mathcal{M}}_{new} \end{bmatrix}$
17. $\quad\quad \mathsf{solve\_and\_update}(\widehat{P}, \mathcal{M}_{new}, \widehat{\mathcal{M}}_{new}, \mathcal{M}^{pre}_{new})$
    $\quad$ end if
    end if
18. Append each vector $v_{i,s}$ such that $1 \leq s \leq k$ and $v_{i,s}$ was matched at the end of (the final) stage i of the Lanczos phase as a column of M_φ (increasing g by the number of such vectors).

// Stage #j for $j \geq 1$

19. while $(g > 0)$ do
20. $\quad h := g$; $\mathcal{M}^{pre}_{new} := M_\varphi$; $\mathcal{M}_{new} := A \cdot \mathcal{M}^{pre}_{new}$
21. $\quad$ Remove all of the columns from M_φ and set g to be 0
22. $\quad (\mathcal{M}_{new}, \mathcal{M}^{pre}_{new}) := \mathsf{orthogonalize}(\mathcal{M}_{new}, \mathcal{M}^{pre}_{new})$
23. $\quad (\mathcal{M}_{new}, \mathcal{M}^{pre}_{new}) := \mathsf{eliminate}(\mathcal{M}_{new}, \mathcal{M}^{pre}_{new})$
24. $\quad$ Set $\widehat{\mathcal{M}}_{new} \in \mathsf{F}_q^{(n-m) \times h}$ to be the matrix such that

$$P \cdot \mathcal{M}_{new} = \begin{bmatrix} 0 \\ \widehat{\mathcal{M}}_{new} \end{bmatrix}.$$

25. $\quad (\widehat{\mathcal{M}}_{new}, \mathcal{M}^{pre}_{new}) := \mathsf{compress}(\widehat{\mathcal{M}}_{new}, \mathcal{M}^{pre}_{new})$
26. $\quad$ if ($\widehat{\mathcal{M}}_{new}$ has at least one column) then
27. $\quad\quad (\widehat{\mathcal{M}}_{new}, \mathcal{M}^{pre}_{new}, \widehat{P}) := \mathsf{triangularize}(\widehat{\mathcal{M}}_{new}, \mathcal{M}^{pre}_{new})$
28. $\quad\quad \mathcal{M}_{new} := P^T \cdot \begin{bmatrix} 0 \\ \widehat{\mathcal{M}}_{new} \end{bmatrix}$
29. $\quad\quad \mathsf{solve\_and\_update}(\widehat{P}, \mathcal{M}_{new}, \widehat{\mathcal{M}}_{new}, \mathcal{M}^{pre}_{new})$
    $\quad$ end if
    end while

// Recovery of Solution

30. if $(\rho == 0)$ then
31. $\quad x := \chi - w$
32. $\quad$ Return x as a vector such that $A \cdot x = b$
    else
33. $\quad$ Report that no vector x such that $A \cdot x = b$ was found.
    end if

A.4 On the Correctness and Efficiency of the Elimination Phase

A.4.1 Establishing the Invariants after Each Stage

LEMMA A.24. *Suppose that $i, j \geq 0$, there were exactly $i+1$ stages of the Lanczos phase, and that there are at least $j+1$ stages of the elimination phase. Then Invariants #8 and 11 are both satisfied at the end of stage #j of the elimination phase.*

PROOF. The claim will be established by induction on j.

Basis: It will be helpful to begin by establishing the following properties, which hold at the end of the Lanczos phase of the computation.

(a) The subspace spanned by $\nu_1, \nu_2, \ldots, \nu_\ell$, and the vectors $v_{r,s}$ such that $0 \leq r \leq i$, $1 \leq s \leq k$, and $v_{r,s}$ was unmatched at the end of (the final) stage i of the Lanczos phase, is the same as the subspace spanned by the vectors $A^a \cdot v_b$ such that $0 \leq a \leq i$ and $1 \leq b \leq k$.

(b) The subspace spanned by $\nu_1, \nu_2, \ldots, \nu_\ell$, the vectors $v_{r,s}$ such that $0 \leq r \leq i$, $1 \leq s \leq k$, and $v_{r,s}$ was unmatched

at the end of stage i of the Lanczos phase, and the vectors $A \cdot v_{i,t}$ such that $1 \leq t \leq k$, is the same as the subspace spanned by the vectors $A^a \cdot v_b$ such that $0 \leq a \leq i+1$ and $1 \leq b \leq k$.

To see that this is the case note, first, that the vectors $\nu_1, \nu_2, \ldots, \nu_\ell$ and the unmatched vectors $v_{r,s}$ such that $0 \leq r \leq i$ and $1 \leq s \leq k$ are just the vectors $v_{c,d}$ such that $0 \leq c \leq i$ and $1 \leq d \leq k$. Consequently the first claim is implied by the fact that Invariant #1 is satisfied at the end of stage i of the Lanczos phase — see Lemma A.4, above.

Now consider a vector $A^a \cdot v_b$ such that $0 \leq a \leq i+1$ and $1 \leq b \leq k$. If $a \leq i$ then it follows by the above that $A^a \cdot v_b$ is a linear combination of the vectors $v_{c,d}$ such that $1 \leq c \leq r$ and $1 \leq d \leq k$. On the other hand, if $a = i+1$ then

$$\begin{aligned} A^{i+1} \cdot v_b &= A \cdot (A^i \cdot v_b) \\ &= A \cdot \left(\sum_{c=0}^{i} \sum_{d=1}^{k} \alpha_{c,d} \cdot v_{c,d} \right) \qquad \text{(by part (a), above)} \\ &= \zeta + \sum_{d=1}^{k} \alpha_{i,d} \cdot A \cdot v_{i,d} \end{aligned}$$

where

$$\zeta = \sum_{c=0}^{i-1} \sum_{d=1}^{k} \alpha_{c,d} \cdot A \cdot v_{c,d}$$

and where $\alpha_{c,d} \in \mathsf{F}_q$ for $0 \leq c \leq i$ and $1 \leq d \leq k$.

Lemma A.4 implies that $A \cdot v_{c,d}$ is a linear combination of the vectors $v_{r,s}$ such that $0 \leq r \leq i$ and $1 \leq s \leq k$ whenever $0 \leq c \leq i-1$ and $1 \leq d \leq k$, so it follows that ζ is a linear combination of these vectors as well. $A^{i+1} \cdot v_b$ has now been expressed as a linear combination of these vectors and the vectors $A \cdot v_{i,t}$ such that $1 \leq t \leq k$ — as needed to establish that the subspace spanned by the vectors $A^a \cdot v_b$ such that $0 \leq a \leq i+1$ and $1 \leq b \leq k$ is contained in the subspace spanned by $\nu_1, \nu_2, \ldots, \nu_\ell$, the unmatched vectors $v_{r,s}$ such that $0 \leq r \leq i$ and $1 \leq b \leq k$, and the vectors $A \cdot v_{i,t}$ such that $1 \leq t \leq k$.

The opposite containment is easier to prove: It follows by part (a), above, that each vector ν_a for $1 \leq a \leq \ell$ and each unmatched vector $v_{r,s}$ for $0 \leq r \leq i$ and $1 \leq s \leq k$ is a linear combination of the vectors $A^c \cdot v_d$ such that $0 \leq c \leq i$ and $1 \leq d \leq k$. Furthermore, since $v_{i,t}$ is in the subspace spanned by the vectors $A^c \cdot v_d$ such that $0 \leq c \leq i$ and $1 \leq d \leq k$, for $1 \leq t \leq k$, $A \cdot v_{i,t}$ is in the subspace spanned by the vectors $A^c \cdot v_d$ such that $0 \leq c \leq i+1$ and $1 \leq d \leq k$. This suffices to confirm that the two subspaces mentioned in part (b) above the same, as required.

It should next be noticed that stage 0 of the elimination phase consists of a pair of rounds — including the steps at lines 2–8 and 9–18, respectively.

We first claim that, after the first of these rounds,, the sequence of vectors

$$\lambda_1, \lambda_2, \ldots \lambda_m$$

spans the same subspace of $\mathsf{F}_q^{n \times 1}$ as the set of all vectors $v_{r,s}$ such that $0 \leq r \leq i-1$, $1 \leq s \leq k$ and $v_{r,s}$ was unmatched at the end of (the final) stage i of the Lanczos phase of the computation. To see that this is the case, let us consider the following alternative claim, which is certainly satisfied after the initialization of the matrix $\mathcal{M}_{new}$ at step 2.

(c) The subspace of $\mathsf{F}_q^{n \times 1}$ spanned by the columns of $\mathcal{M}_{new}$ is equal to the subspace of $\mathsf{F}_q^{n \times 1}$ spanned by the unmatched vectors $v_{r,s}$ such that $0 \leq r \leq i-1$ and $1 \leq s \leq k$.

Let us first consider the case that there are no unmatched vectors $v_{r,s}$ such that $0 \leq r \leq i-1$ and $1 \leq s \leq k$ at all. In this case the test at line 4 fails and this first round ends — with $m = 0$ — and the above claim (concerning the vectors $\lambda_1, \lambda_2, \ldots, \lambda_m$) is trivially satisfied.

Suppose, instead, that there is at least one unmatched vector $v_{r,s}$ such that $0 \leq r \leq i-1$ and $1 \leq s \leq k$ — but that all such vectors are equal to 0. In this case the test at line 4 is passed, and procedure compress is executed at line 5. The matrix $\mathcal{M}_{new}$ is modified in such a way that its column space is unchanged but its columns are linearly independent — so that the number h of its columns is equal to 0 after line 5 (see Lemma A.21, above). Consequently the test at line 6 fails and the first round ends with $m = 0$ — as suffices to establish the above claim, concerning $\lambda_1, \lambda_2, \ldots, \lambda_m$, once again.

Finally let us consider the case that there is at least one nonzero vector $v_{r,s}$ such that $0 \leq r \leq i-1$ and $1 \leq s \leq k$ and $v_{r,s}$ is unmatched at the end of the final stage of the Lanczos phase. Note, first, that claim (c), above, is satisfied after the execution of procedure compress at line 5 because this does not change the column space of $\mathcal{M}_{new}$ (see Lemma A.21, above). Now $\mathcal{M}_{new}$ must still include at least column so that $h > 0$ and the test at line 6 is passed, and lines 8 and 9 are also executed before the completion of the first round.

Claim (c) is also satisfied after the execution of the procedure triangularize at line 7 because this procedure does not change the column space of $\mathcal{M}_{new}$ or modify other data either (see Lemma A.22, above).

Since $m = 0$ at this point it now follows by part (a) of Lemma A.23 that the subspace spanned by $\lambda_1, \lambda_2, \ldots, \lambda_m$ after the execution of solve_and_update is equal to the subspace spanned by the unmatched vectors $v_{r,s}$ such that $0 \leq r \leq i-1$ and $1 \leq s \leq k$ — for $\lambda_1, \lambda_2, \ldots, \lambda_m$ will span the same subspace as the columns of $\mathcal{M}_{new}$ after this step if m was equal to 0 before it.

We next claim that, after the second round (that is, the end of stage 0), Invariants #8 and #11 are both satisfied. It will be helpful to consider the following claims, which are satisfied after line 9, instead.

(d) The subspace of $\mathsf{F}_q^{n \times 1}$ spanned by the vectors

$$\lambda_1, \lambda_2, \ldots, \lambda_m$$

is the same as the subspace spanned by the vectors $v_{r,s}$ such that $0 \leq r \leq i-1$, $1 \leq s \leq k$, and $v_{r,s}$ was unmatched at the end of (the final) stage i of the Lanczos phase.

(e) The subspace of $\mathsf{F}_q^{n \times 1}$ spanned by the vectors

$$\lambda_1, \lambda_2, \ldots, \lambda_m$$

and the columns of $\mathcal{M}_{new}$ is the same as the subspace spanned by the vectors $v_{r,s}$ such that $0 \leq r \leq i$, $1 \leq s \leq k$, and $v_{r,s}$ was unmatched at the end of stage i of the Lanczos phase.

(f) The subspace of $\mathsf{F}_q^{n \times 1}$ spanned by the vectors

$$\nu_1, \nu_2, \ldots, \nu_\ell, \lambda_1, \lambda_2, \ldots, \lambda_m,$$

and the columns of $\mathcal{M}_{new}$ and $A \cdot \mathcal{M}_{new}$ is the same as the subspace spanned by the vectors $A^a \cdot v_b$ such that $0 \leq r \leq i$ and $1 \leq s \leq k$ and the vectors $A \cdot v_{i,t}$ such that $1 \leq t \leq k$ and $v_{i,t}$ was unmatched at the end of stage i of the Lanczos phase.

Clam (d) is satisfied because it was satisfied at the end of the first round and none of the vectors $\lambda_1, \lambda_2, \ldots, \lambda_m$ have been changed by the initialization of $\mathcal{M}_{new}$. Part (e) is a consequence of the initialization of $\mathcal{M}_{new}$ as describe at line 9.

In order to see that part (f) is also correct at this point, one should notice that it follows by Lemma A.4 and claim (e) that the subspace spanned by the vectors

$$\nu_1, \nu_2, \ldots, \nu_\ell, \lambda_1, \lambda_2, \ldots, \lambda_m$$

and the columns of $\mathcal{M}_{new}$ is the is the same as that spanned by the vectors $A^a \cdot v_b$ such that $0 \leq a \leq i$ and $1 \leq b \leq k$. The claim now follows because, after line 9, the columns of $\mathcal{M}_{new}$ are the vectors $v_{i,t}$ such that $1 \leq t \leq k$ and $v_{i,t}$ was unmatched at the end of the final stage of the Lanczos phase.

Suppose, now, that there were no vectors $v_{i,t}$ such that $1 \leq t \leq k$ and $v_{i,t}$ was unmatched at the end of stage i of the Lanczos phase at all. In this case $h = 0$ after line 9 so that the test at line 10 fails, and the only other step executed during stage 0 of the elimination phase is step 18. In this case, Invariant #8 is satisfied at the end of stage 0 because it is implied by claim (e), above, when the matrix $\mathcal{M}_{new}$ has no columns, and because the vectors $\lambda_1, \lambda_2, \ldots, \lambda_m$ are not changed by the execution of step 18.

Invariant #11 is satisfied as well because all of the vectors $v_{i,t}$ such that $1 \leq t \leq k$ are included as columns of M_φ during the execution of step 18 in this case, so that the invariant is now a consequence of Invariant #8 and property (b), above.

Suppose next that there is at least one unmatched vector $v_{i,t}$ such that $1 \leq t \leq k$. In this case, $h > 0$ when this is checked at line 10, so that steps 11 and 12 will be executed.

Consider the effects of the execution of procedure eliminate at line 11. Property (d), above, will still be satisfied because it was satisfied after step 9 and none of the vectors mentioned in it have been changed. Property (e) will also hold because the subspace spanned by $\lambda_1, \lambda_2, \ldots, \lambda_m$ and the columns of $\mathcal{M}_{new}$ has not been changed by the execution of eliminate (see Lemma A.20, above).

In order to see that property (f) is also satisfied after this step, notice that the transformation applied by the execution of procedure eliminate was a linear transformation in which the columns $\gamma_1, \gamma_2, \ldots, \gamma_h$ of $\mathcal{M}_{new}$ were replaced by linear combinations of these columns and the vectors $\lambda_1, \lambda_2, \ldots, \lambda_m$:

$$\gamma_a := \sum_{b=1}^{n} \alpha_{a,b} \cdot \gamma_b + \sum_{c=1}^{m} \beta_{a,c} \cdot \lambda_c \tag{27}$$

and, as noted above, this does not change the subspace spanned by $\gamma_1, \gamma_2, \ldots, \gamma_h, \lambda_1, \lambda_2, \ldots, \lambda_m$.

It follows that $A \cdot \gamma_a$ is being updated in a similar way, for $1 \leq a \leq h$ —

$$A \cdot \gamma_a = \sum_{b=1}^{h} \alpha_{a,b} \cdot (A \cdot \gamma_b) + \sum_{c=1} \beta_{a,c} \cdot (A \cdot \lambda_c) \tag{28}$$

and the subspace spanned by $A \cdot \gamma_1, A \cdot \gamma_2, \ldots, A \cdot \gamma_h$, and $A \cdot \lambda_1, A \cdot \lambda_2, \ldots, A \cdot \lambda_m$ has not been changed, either.

Now suppose that $A \cdot \lambda_c$ is a linear combination of the vectors $\nu_1, \nu_2, \ldots, \nu_\ell$, $\lambda_1, \lambda_2, \ldots, \lambda_m$, and columns $\gamma_1, \gamma_2, \ldots, \gamma_h$ of $\mathcal{M}_{new}$ for $1 \leq c \leq m$. Then it would follow that the subspace spanned by $\nu_1, \nu_2, \ldots, \nu_\ell$, $\lambda_1, \lambda_2, \ldots, \lambda_m$ and the columns of $\mathcal{M}_{new}$ and $A \cdot \mathcal{M}_{new}$ was also unchanged by the execution of the procedure eliminate — for, considering the updates at lines (27) and (28) in sequence, one could demonstrate that a given vector $\zeta \in \mathsf{F}_q^{n \times 1}$ was a linear combination of $\nu_1, \nu_2, \ldots, \nu_\ell, \lambda_1, \lambda_2, \ldots, \lambda_m$, and the columns of $\mathcal{M}_{new}$ and $A \cdot \mathcal{M}_{new}$ before the update if and only if it can be expressed as a linear combination of these vectors after it.

It now suffices to note that λ_c is a linear combination of the vectors $v_{r,s}$ such that $0 \leq r \leq i-1$ and $1 \leq s \leq k$, by property (e) — so that Lemma A.4 implies that $A \cdot \lambda_c$ is a linear combination of the vectors $v_{r,s}$ such that $0 \leq r \leq i$ and $1 \leq s \leq k$. It now follows by property (e) that $A \cdot \lambda_c$ is a linear combination of $\nu_1, \nu_2, \ldots, \nu_\ell$, $\lambda_1, \lambda_2, \ldots, \lambda_m$ and the columns of $\mathcal{M}_{new}$ — as required to show that property (f) is also satisfied after the execution of procedure eliminate at line 11 because it was satisfied before it.

Suppose, now, that each unmatched vector $v_{i,t}$ was a linear combination of the vectors $\lambda_1, \lambda_2, \ldots, \lambda_m$. Then each of the columns of $\mathcal{M}_{new}$ would be equal to 0 after the execution of eliminate at line 11. The matrix $\widehat{\mathcal{M}}_{new}$ defined at line 12 would thus be the zero matrix in $\mathsf{F}_q^{(n-m) \times h}$, so that it would have no columns at all as a result of the execution of procedure compress at line 13 (see Lemma A.21, above). The test at line 14 would therefore fail, and step 18 would be the only remaining step to be executed during stage 0 of the elimination phase.

In this case it follows by property (e), above, that the subspace of $\mathsf{F}_q^{n \times 1}$ spanned by the vectors $\lambda_1, \lambda_2, \ldots, \lambda_m$ is the same as the subspace spanned by the unmatched vectors $v_{r,s}$ such that $0 \leq r \leq i$ and $1 \leq s \leq k$ — and property (a), above, implies that Invariant #8 has now been satisfied.

Furthermore, if $v_{i,t}$ is an unmatched vector then $A \cdot v_{i,t}$ is a linear combination of $\nu_1, \nu_2, \ldots, \nu_\ell$ and $\lambda_1, \lambda_2, \ldots, \lambda_m$ in this case — for it follows by property (d), above, that $v_{i,t}$ is a linear combination of vectors $v_{r,s}$ such that $0 \leq r \leq i-1$ and $1 \leq s \leq k$, so that Lemma A.4 implies that $A \cdot v_{i,t}$ is a linear combination of $v_{r,s}$ such that $0 \leq r \leq i$ and $1 \leq s \leq k$ — so that this now follows by property (a), above, and Invariant #8.

Consequently Invariant #11 will be satisfied, in this case, after the execution of step #18, which sets the columns of M_φ to be the matched vectors $v_{i,t}$ such that $1 \leq t \leq k$.

Finally, suppose that there is at least one unmatched vector $v_{i,t}$ that is not a linear combination of $\lambda_1, \lambda_2, \ldots, \lambda_m$ after step 9. Once again, properties (d), (e) and (f) will be satisfied after the execution of step 11, as argued above, and it will follow by property (e) that the matrix $\mathcal{M}_{new}$ will have at least one column that is nonzero at this point. Consequently the matrix $\widehat{\mathcal{M}}_{new}$ defined at line 12 will be nonzero. The column space of this matrix is not changed by the execution of compress at line 12 (see Lemma A.21, above), so $\widehat{\mathcal{M}}_{new}$ will still have at least one column when this is checked at line 14 and steps 15–18 will be executed before the end of stage 0 of the elimination phase.

Now, the column space of $\widehat{\mathcal{M}}_{new}$ will not have been changed by the execution of procedure compress at line 13 (see, again,

Lemma A.21) or the execution of procedure triangularize at line 15 (see Lemma A.22) and, since

$$P \cdot \mathcal{M}_{new} = \begin{bmatrix} 0 \\ \widehat{\mathcal{M}}_{new} \end{bmatrix}$$

both before the execution of step 12 and after the execution of step 16, the column space of $\mathcal{M}_{new}$ will be the same, after steps 11 and 16, as well. Since none of the vectors

$$\nu_1, \nu_2, \ldots, \nu_\ell, \lambda_1, \lambda_2, \ldots, \lambda_m$$

are changed by the execution of steps 12–16 it follows that properties (d), (e) and (f) hold after the execution of step 16 because they held before the execution of step 12.

It now follows by Lemma A.23 that the execution of procedure solve_and_update at line 17 establishes Invariant #8: For the column space of the vectors $\lambda_1, \lambda_2, \ldots, \lambda_m$ after this operation will be the same as the subspace of $\mathsf{F}_q^{n\times 1}$ spanned by the vectors $\lambda_1, \lambda_2, \ldots, \lambda_m$ and the columns of $\mathcal{M}_{new}$ before it, and this follows because property (e) was satisfied after the execution of line 16. Furthermore, the column space of the matrix M_φ after step 16 will be the same as the column space of the matrix $\mathcal{M}_{new}$ before it (see, again, the above lemma), so property (f) can now be applied to establish that, after step 17, the subspace of $\mathsf{F}_q^{n\times 1}$ spanned by the vectors

$$\nu_1, \nu_2, \ldots, \nu_\ell, \lambda_1, \lambda_2, \ldots, \lambda_m$$

and the columns of $A \cdot M_\varphi$ will be the same as the subspace spanned by the vectors $v_{r,s}$ such that $0 \leq r \leq i$ and $1 \leq s \leq k$ and the vectors $A \cdot v_{i,t}$ such that $1 \leq t \leq k$ and $v_{i,t}$ was unmatched at the end of stage i of the Lanczos phase.

The addition of columns to M_φ in the execution of step 18 therefore ensures that the subspace spanned the subspace spanned by

$$\nu_1, \nu_2, \ldots, \nu_\ell, \lambda_1, \lambda_2, \ldots, \lambda_m$$

and the columns of $A \cdot M_\varphi$ is the same as the subspace spanned by $v_{r,s}$ such that $0 \leq r \leq i$ and $1 \leq s \leq k$ and $A \cdot v_{i,t}$ such that $1 \leq t \leq k$. Since Invariant #1 was satisfied at the end of the Lanczos phase and Invariant #8 is established now, another application of property (b), above, now suffices to establish Invariant #11 at the end of the stage 0 of the elimination phase as well.

Inductive Step: Suppose that $j \geq 0$, there are at least $j+2$ stages of the elimination phase, and that Invariants #8 and 11 are all satisfied at the end of stage j. It is necessary and sufficient to establish that these are satisfied at the end of stage $j+1$ as well.

To begin, it will be helpful to establish yet another property.

(g) The subspace of $\mathsf{F}_q^{n\times 1}$ spanned by the vectors

$$\begin{aligned} &\nu_1, \nu_2, \ldots, \nu_\ell, \lambda_1, \lambda_2, \ldots, \lambda_m, \\ &\qquad A \cdot \varphi_1, A \cdot \varphi_2, \ldots, A \cdot \varphi_g, \\ &\qquad \ldots A^2\varphi_1, A^2 \cdot \varphi_2, \ldots, A^2 \cdot \varphi_g \end{aligned} \quad (29)$$

is the same as the subspace spanned by the vectors $A^a \cdot v_b$ such that $0 \leq a \leq i+j+2$ and $1 \leq b \leq k$.

To see that this is the case one should note the following.

- Each vector $A^a \cdot v_b$ such that $0 \leq a \leq i+j+1$ and $1 \leq b \leq k$ is a linear combination of the vectors at line (29), above, because it follows by the inductive hypothesis that Invariant #11 is satisfied at the end of stage j of the elimination phase, so that this vector is a linear combination of

  $$\nu_1, \nu_2, \ldots, \nu_\ell, \lambda_1, \lambda_2, \ldots, \lambda_m, A \cdot \varphi_1, A \cdot \varphi_2, \ldots, A \cdot \varphi_g.$$

- It now follows that if $1 \leq b \leq k$ then

  $$\begin{aligned} A^{i+j+2} \cdot v_b &= A \cdot (A^{j+1} \cdot v_b) \\ &= A \cdot \left(\sum_{r=1}^{\ell} \alpha_r \cdot \nu_r + \sum_{s=1}^{m} \beta_s \cdot \lambda_b + \sum_{t=1}^{g} \gamma_t \cdot A \cdot \varphi_t \right) \\ &= A \cdot \zeta + \sum_{t=1}^{g} \gamma_t \cdot A^2 \cdot \varphi_t \end{aligned}$$

  where $\alpha_r \in \mathsf{F}_q$ for $1 \leq r \leq \ell$, $\beta_s \in \mathsf{F}_q$ for $1 \leq s \leq m$, $\gamma_t \in \mathsf{F}_q$ for $1 \leq t \leq g$, and where

  $$\zeta = \sum_{r=1}^{\ell} \alpha_r \cdot \nu_r + \sum_{s=1}^{m} \beta_s \cdot \lambda_s.$$

  Now it also follows by the inductive hypothesis that Invariant #8 was satisfied at the end of stage j of the elimination phase, so that ζ is a linear combination of the vectors $A^c \cdot v_d$ for $0 \leq c \leq i+j$ and $1 \leq d \leq k$. It follows that $A \cdot \zeta$ is a linear combination of the vectors $A^c \cdot v_d$ such that $0 \leq c \leq i+j+1$ and $1 \leq d \leq k$ — implying, by Invariant #11 once again, that $A \cdot \zeta$ is a linear combination of

  $$\nu_1, \nu_2, \ldots, \nu_\ell, \lambda_1, \lambda_2, \ldots, \lambda_m, A \cdot \varphi_1, A \cdot \varphi_2, \ldots, A \cdot \varphi_g.$$

  It now follows that $A^{i+j+2} \cdot v_b$ is a linear combination of the vectors shown at line (29) as well.

- It has now been established that the subspace of $\mathsf{F}_q^{n\times 1}$ spanned by the vectors $A^a \cdot v_b$ such that $0 \leq a \leq i+j+2$ and $1 \leq b \leq k$ is contained in the subspace spanned by the vectors listed at one (29).

  Note, on the other hand, that it follows by Invariant #11 that each of the vectors

  $$\nu_1, \nu_2, \ldots, \nu_\ell, \lambda_1, \lambda_2, \ldots, \lambda_m, A \cdot \varphi_1, A \cdot \varphi_2, \ldots, A \cdot \varphi_g$$

  is contained in the subspace spanned by $A^a \cdot v_b$ for $0 \leq a \leq i+j+2$ and $1 \leq b \leq k$ — because they are all contained in the subspace spanned by $A^a \cdot v_b$ for $0 \leq a \leq i+j+1$ and $1 \leq b \leq k$.

  Furthermore, since $A \cdot \varphi_c$ is in the subspace spanned by $A^a \cdot b_b$ such that $0 \leq a \leq i+j+1$ and $1 \leq b \leq k$, for $1 \leq c \leq g$, $A^2 \cdot \varphi_c = A \cdot (A \cdot \varphi_c)$ is certainly in the subspace spanned by $A^d \cdot v_e$ such that $0 \leq d \leq i+j+2$ and $1 \leq e \leq k$ — as needed to establish the opposite containment and prove that the subspaces of $\mathsf{F}_q^{n\times 1}$ being considered are the same.

Notice that the matrix $\mathcal{M}_{new}$ is initialized to have the vectors

$$A \cdot \varphi_1, A \cdot \varphi_2, \ldots, A \cdot \varphi_g$$

as its entries, at line 20. It follows by the above, that, at this point in the computation, the following properties are satisfied.

(h) The subspace of $\mathsf{F}_q^{n\times 1}$ spanned by the vectors

$$\nu_1, \nu_2, \ldots, \nu_\ell, \lambda_1, \lambda_2, \ldots, \lambda_m$$

and the columns of $\mathcal{M}_{new}$ is equal to the subspace of $\mathsf{F}_q^{n\times 1}$ spanned by the vector $A^a \cdot v_b$ such that $0 \leq a \leq i+j+1$ and $1 \leq b \leq k$.

(i) The subspace of $\mathsf{F}_q^{n\times 1}$ spanned by the vectors

$$\nu_1, \nu_2, \ldots, \nu_\ell, \lambda_1, \lambda_2, \ldots, \lambda_m,$$

the columns of $\mathcal{M}_{new}$, and the columns of $A \cdot \mathcal{M}_{new}$, is equal to the subspace of $\mathsf{F}_q^{n\times 1}$ spanned by the vectors $A^a \cdot v_b$ such that $0 \leq a \leq i+j+2$ and $1 \leq b \leq k$.

(j) Consequently, it follows by the inductive hypothesis (which included the fact that Invariant #8 was satisfied at the end of stage j) that each of the vectors

$$A \cdot \nu_1, A \cdot \nu_2, \ldots, A \cdot \nu_\ell, A \cdot \lambda_1, A \cdot \lambda_2, \ldots, A \cdot \lambda_m \quad (30)$$

is in the subspace of $\mathsf{F}_q^{n\times 1}$ spanned by the vectors

$$\nu_1, \nu_2, \ldots, \nu_\ell, \lambda_1, \lambda_2, \ldots, \lambda_m$$

and the columns of $\mathcal{M}_{new}$.

Let us consider these subspaces as they are defined after the execution of the procedure orthogonalize at line 22.

Property (h), above, is still satisfied, because the subspace of $\mathsf{F}_q^{n\times 1}$ spanned by $\nu_1, \nu_2, \ldots, \nu_\ell$ and the columns of $\mathcal{M}_{new}$ has not been modified and none of the vectors $\lambda_1, \lambda_2, \ldots, \lambda_m$ have been changed (see Lemma A.19, above).

Property (j) also continues to hold because the subspace spanned by the vectors $\nu_1, \nu_2, \ldots, \nu_\ell$ and columns of $\mathcal{M}_{new}$ has not been changed, and it is only the matrix $\mathcal{M}_{new}$ that has been modified (see Lemma A.19, once again). Consequently the subspace spanned by $\nu_1, \nu_2, \ldots, \nu_\ell, \lambda_1, \lambda_2, \ldots, \lambda_m$ and the columns of $\mathcal{M}_{new}$ has not been changed either — and none of the vectors shown at line (30) have been modified.

As noted above the subspace of $\mathsf{F}_q^{n\times 1}$ spanned by the vectors $\nu_1, \nu_2, \ldots, \nu_\ell$ and the columns of $\mathcal{M}_{new}$ has not been changed. It follows that the subspace spanned by $A \cdot \nu_1, A \cdot \nu_2, \ldots, A \cdot \nu_\ell$ and the columns of $A \cdot \mathcal{M}_{new}$ has not been changed, either. Once again, the vectors $\lambda_1, \lambda_2, \ldots, \lambda_m$ have not been modified. It follows that the subspace spanned by $\nu_1, \nu_2, \ldots, \nu_\ell$ $A \cdot \nu_1, A \cdot \nu_2, \ldots, A \cdot \nu_\ell$, $\lambda_1, \lambda_2, \ldots, \lambda_m$, and the columns of $\mathcal{M}_{new}$ and of $A \cdot \mathcal{M}_{new}$ has also not been changed — and, since $A \cdot \nu_1, A \cdot \nu_2, \ldots, A \cdot \nu_\ell$ are all in the subspace spanned by $\nu_1, \nu_2, \ldots, \nu_\ell$, $\lambda_1, \lambda_2, \ldots, \lambda_m$ and the columns of $\mathcal{M}_{new}$ (by property (j)), this implies that property (i) of the claim holds after the orthogonalize operation if it did before.

A similar analysis establishes that the above claims are all satisfied after the execution of the procedure eliminate at line 23. In this case one uses the fact that only the columns of $\mathcal{M}_{new}$ (and $A \cdot \mathcal{M}_{new}$) have been changed and the subspace spanned by $\lambda_1, \lambda_2, \ldots, \lambda_m$ and the columns of $\mathcal{M}_{new}$ has not been modified (see Lemma A.20 above), so the roles of $\nu_1, \nu_2, \ldots, \nu_\ell$ and $\lambda_1, \lambda_2, \ldots, \lambda_m$ in the argument are reversed.

Now suppose that, after the above application of eliminate, the columns of $\mathcal{M}_{new}$ are all equal to 0. In this case the application of procedure compress at line 25 removes all columns from $\widehat{\mathcal{M}}_{new}$ so that the test at line 26 fails and this stage of the elimination phase ends. However, Invariants #8 and #11 have both been established: Since the columns of $\mathcal{M}_{new}$ were all zero at this point and the matrix M_φ does not have any columns either, property (h) implies Invariant #8 and property (i) implies Invariant #11.

Suppose, on the other hand, that there is at least one nonzero column of $\mathcal{M}_{new}$ after the application of procedure eliminate at line 23. Then the matrix $\widehat{\mathcal{M}}_{new}$ will still have at least one column after the application of compress at line 25, so that the test at line 26 will be passed and lines 27–29 will be executed before the end of this stage of the elimination phase.

The column space of $\widehat{\mathcal{M}}_{new}$ has not been changed by the application of procedures compress and triangularize at lines 25 and 27 (see, again, Lemmas A.21 and A.22 above), and

$$P \cdot \mathcal{M}_{new} = \begin{bmatrix} 0 \\ \widehat{\mathcal{M}}_{new} \end{bmatrix}$$

both before the execution of step 25 and after the execution of step 28, so the column space of $\mathcal{M}_{new}$ has not been changed by the execution of steps 25–28, either. Properties (h)–(j) therefore hold after the execution of step 28 because they held before the step 25. Once again, the application of solve_and_update at line 29 will establish Invariants #8 and #11, as needed to complete to complete the proof — for it follows by Lemma A.23 that, after this operation, the subspace of $\mathsf{F}_q^{n\times 1}$ spanned by $\lambda_1, \lambda_2, \ldots, \lambda_m$ will be the same as the subspace that was spanned by $\lambda_1, \lambda_2, \ldots, \lambda_m$ and the columns of $\mathcal{M}_{new}$ (so that property (h) can be used to establish Invariant #8), and the column space of M_φ after the operation will be the same as the column space of $\mathcal{M}_{new}$ before it as well (so that property (i) can be used to establish Invariant #11). □

LEMMA A.25. *Invariant #9 is satisfied at the end of every stage of the elimination phase.*

PROOF. Let j be an integer such that $j \geq 0$ and there are at least $j+1$ stages of the elimination phase. We will show that Invariant #9 is satisfied at the end of stage j by induction on j.

Basis: Note first that Invariant #9 is trivially satisfied at the beginning of stage 0 because $m = 0$ and the claim is vacuous at that point.

Notice as well that the columns of the matrix $\mathcal{M}_{new}$ are linear combinations of unmatched vectors $v_{r,s}$ such that $0 \leq r \leq i$ and $1 \leq s \leq k$ throughout the first round (steps 2–8) of stage 0 of the elimination phase, to the fact that Invariant #3 was satisfied at the end of the Lanczos phase implies that $\nu_a^T \cdot \mathcal{M}_{new} = 0$ for $1 \leq a \leq \ell$. It also follows by Lemma A.23 that $\mu_a^T \cdot \lambda_b = 0$ for $1 \leq a \leq \ell$ and $1 \leq b \leq m$ after the procedure solve_and_update is applied at step 8.

The argument needed to establish this for the second round (steps 9–18) of stage 0 is almost the same. One should note, here, as well, that it follows by the correctness of the procedure eliminate (Lemma A.20) that $\mu_a^T \cdot \mathcal{M}_{new} = 0$ for $1 \leq a \leq \ell$ after procedure eliminate is applied at line 11. Now, either step 17 is never reached at all — in which case the vectors $\lambda_1, \lambda_2, \ldots, \lambda_m$ are not changed during the execution of the second round (establishing the result), or the column space of $\mathcal{M}_{new}$ is unchanged by steps 12–16 and $\mu_a^T \cdot \mathcal{M}_{new} = 0$ immediately before the application of procedure solve_and_update at step 17. In the latter case, the correctness of this procedure (as described in Lemma A.23) suffices to establish the desired result once again.

Inductive Step: Suppose that $j \geq 0$, there are at least $j+2$ stages of the elimination phase, and that Invariant #9 is satisfied at the end of stage j of the elimination phase. It is necessary and sufficient to show that it is satisfied at the end of stage $j+1$ as well.

To begin, one should notice that it follows by the correctness of the orthogonalize procedure (Lemma A.19) that $\mu_a^T \cdot \mathcal{M}_{new} = 0$ for $1 \leq a \leq \ell$ after this procedure is applied at step 22. The argument needed to establish the result is now the same as the argument used to establish it for the second round of stage 0. □

LEMMA A.26. *Invariants #10, 12 and 13 are satisfied at the end of every stage of the elimination phase.*

PROOF. This can be established using another straightforward proof by induction on the number of stages that have been carried out already — using the fact that Invariant #7 was satisfied at the end of the Lanczos phase and that, at the beginning of stage 0 of the elimination phase, $m = 0$ — so that Invariant #7 implies Invariant #13, and Invariants #10 and 13 are vacuous claims at this point. □

The following lemma will be of use in bounding the length of the elimination phase of the algorithm as well as the cost of each of its stages and, therefore in bounding the cost of the algorithm.

LEMMA A.27. *Suppose $i \geq 0$ and there are $i+1$ stages of the Lanczos phase of the algorithm (ending with stage i). Then $\ell \geq (i - \Delta_{n,k}) \cdot k$ and there are at most $(\Delta_{n,k}+1) \cdot k$ unmatched vectors $v_{r,s}$ such that $0 \leq r \leq i$ and $1 \leq s \leq k$ at the end of the Lanczos phase of the algorithm.*

PROOF. It follows by Lemma A.15 that there are at most $\Delta_{n,k} \cdot k$ unmatched vectors $v_{r,s}$ such that $0 \leq r \leq i-1$ and $1 \leq s \leq k$ at the end of stage $i-1$ of the Lanczos phase. Consequently $\ell \geq (i - \Delta_{n,k}) \cdot k$ at this point, because ℓ and the number of unmatched vectors $v_{r,s}$ such that $0 \leq r \leq i-1$ and $1 \leq s \leq k$ must be equal to $i \cdot k$ at this point.

Since at most k additional unmatched vectors $v_{r,s}$ such that $0 \leq r \leq i$ and $1 \leq s \leq k$ could be added during stage i (namely, the vectors $v_{i,s}$ such that $1 \leq s \leq k$), this suffices to establish the bound on the number of unmatched vectors that is claimed above, It also implies that $\ell \geq (i - \Delta_{n,k}) \cdot k$ at the end of the Lanczos phase, because ℓ could not have been reduced during the final stage. □

Suppose there are $j+1$ stages of the elimination phase (ending with stage j). For $0 \leq r \leq j$, let m_r and g_r be the values of m and g (that is, the number of columns in the matrices M_λ and M_φ), respectively, at the end of stage r. The following is easily established using Lemma A.27 and by inspection of the algorithm.

LEMMA A.28. *Suppose, as above, that there are $j+1$ stages of the elimination phase (ending with stage j).*

(a) $m_0 \leq (\Delta_{n,k}+1) \cdot k$.

(b) $k \geq g_0 \geq g_1 \geq g_2 \geq \cdots \geq g_j = 0$.

(c) *If $0 \leq r \leq j-1$ then $m_{r+1} = m_r + g_{r+1} = m_0 + \sum_{s=1}^{r+1} g_s$.*

The following is now a straightforward consequence of Lemmas A.19–A.23, Lemma A.27, and inspection of the algorithm.

LEMMA A.29. *The costs of the stages of the elimination phase are bounded as follows.*

(a) *The initialization step requires $O(\Delta_{n,k} \cdot kn)$ operations over F_q.*

(b) *Stage 0 requires $O(\Delta_{n,k}^2 \cdot k^2 n)$ operations over F_q.*

(c) *If $1 \leq r \leq j$ then stage r requires g_{r-1} multiplications of vectors by A and $O(\Delta_{n,k} \cdot k \cdot g_{r-1} \cdot n + g_{r-1} \cdot mn)$ operations over F_q.*

(d) *Storage space required is that needed to store $O(m + (\Delta_{n,k}+2) \cdot k)$ vectors in $\mathsf{F}_q^{n \times 1}$, that is to store $O(mn + (\Delta_{n,k}+2) \cdot kn)$ elements of F_q.*

The following bounds are now a straightforward consequence of Lemma A.28 — which establishes that $\sum_{r=0}^{j} g_r \leq m+k$ — and the bounds given in Lemma A.29, above.

LEMMA A.30. *The elimination phase requires at most $m+k$ multiplication of vectors by A and $O(\Delta_{n,k}^2 \cdot k^2 n + \Delta_{n,k} \cdot kmn + m^2 n)$ operations over F_q.*

A.5 On the Correctness and Efficiency of the Algorithm

PROOF OF THEOREM 2.3. Consider the sequence

$$\nu_1, \nu_2, \ldots, \nu_\ell, \lambda_1, \lambda_2, \ldots, \lambda_m$$

of vectors in $\mathsf{F}_q^{n \times 1}$ that have been generated on termination of this algorithm — recalling that, by Lemma A.24, Invariants #8 and #11 hold at the end of the final stage of the elimination phase and that the matrix M_φ has no columns at this point.

Since Invariant #8 was satisfied on termination of the algorithm, the subspace of $\mathsf{F}_q^{n \times 1}$ spanned by these vectors is a subspace of $\mathcal{KS}_{\vec{v}}$ that includes the vectors $v_1, v_2, \ldots, v_k$ — in particular, if the Lanczos phase included $i+1$ stages (ending with stage #i) and the elimination phase included $j+1$ stages (ending with stage #j) then Invariant #8 implies that this is the space spanned by the vectors $A^r \cdot v_s$ such that $0 \leq r \leq i+j$ and $1 \leq s \leq k$.

This space is also closed under multiplication by A — for if $0 \leq r \leq i+j-1$ and $1 \leq s \leq k$ then it follows by the above that $A \cdot (A^r \cdot v_s) = A^{r+1} \cdot v_s$ is also an element of this space. On the other hand, ir $r = i+j$ then it follows by Invariant #11 that $A \cdot (A^r \cdot v_s)$ is a linear combination of $\nu_1, \nu_2, \ldots, \nu_\ell$, $\lambda_1, \lambda_2, \ldots, \lambda_m$, and the columns of the matrix M_φ. However the matrix M_φ has no columns at this point — for, otherwise, the elimination phase would not have ended after stage j. Consequently, $A \cdot (A^r \cdot v_s)$ is a linear combination of

$$\nu_1, \nu_2, \ldots, \nu_\ell, \lambda_1, \lambda_2, \ldots, \lambda_m$$

as well, as required to establish the closure property that has been claimed.

It now follows that the above vectors span the Krylov space $\mathcal{KS}_{\vec{v}}$, because this is the smallest subspace of $\mathsf{F}_q^{n \times 1}$ that includes the vectors $v_1, v_2, \ldots, v_k$ and that is closed under multiplication by A.

Next consider elements α_a and β_b of F_q, for $1 \leq a \leq \ell$ and $1 \leq b \leq m$, such that

$$\sum_{a=1}^{\ell} \alpha_a \cdot \nu_a + \sum_{b=1}^{m} \beta_m \cdot \lambda_m = 0.$$

It follows by Lemma A.14 that Invariant #2 was satisfied at the end of the Lanczos phase of the computation and, since none of the vectors μ_r or ν_r are modified after this, for $1 \leq r \leq \ell$, this invariant is still satisfied on termination of the algorithm. Lemma A.25 establishes that Invariant #9 is satisfied on termination of the algorithm as well, so that if $1 \leq a \leq \ell$ then

$$0 = \mu_a^T \cdot \left(\sum_{a=1}^{\ell} \alpha_a \cdot \nu_a + \sum_{b=1}^{m} \beta_m \cdot \lambda_m \right) = \alpha_a.$$

in other words, $\alpha_1 = \alpha_2 = \cdots = \alpha_\ell = 0$, so that

$$\sum_{b=1}^{m} \beta_b \cdot \lambda_b = 0$$

as well. Now, as stated in Lemma A.26, Invariant #10 is also satisfied on termination of the algorithm, and this implies that the vectors $\lambda_1, \lambda_2, \ldots, \lambda_m$ are linearly independent — so that $\beta_1 = \beta_2 = \cdots = \beta_m = 0$ as well. It now follows that the vectors

$$\nu_1, \nu_2, \ldots, \nu_\ell, \lambda_1, \lambda_2, \ldots, \lambda_m$$

are linearly independent — and that they form a basis for the Krylov space $\mathcal{KS}_{\vec{v}}$, as claimed.

Suppose, next, that this Krylov space includes a vector ζ such that $A \cdot \zeta = A \cdot w + b$ and note that, since Invariant #13 is satisfied on termination of the algorithm (by Lemma A.26, once again), $\rho \in \mathsf{F}_q^{n \times 1}$, $\chi \in \mathcal{KS}_{\vec{v}}$, and $A \cdot \chi + \rho = A \cdot w + b$ as well.

Consequently $\rho = A \cdot \zeta - A \cdot \chi$, so that $\rho \in \mathcal{KS}_{\vec{v}}$ too — and

$$\rho = \sum_{a=1}^{\ell} \gamma_1 \cdot \nu_a + \sum_{b=1}^{m} \delta_b \cdot \lambda_b$$

where $\gamma_a, \beta_b \in \mathsf{F}_q$ for $1 \leq a \leq \ell$ and $1 \leq b \leq m$.

However, it also follows by Invariant #13 that $\mu_a^T \cdot \rho = 0$, and

$$0 = \mu_a^T \cdot \left(\sum_{a=1}^{\ell} \gamma_1 \cdot \nu_a + \sum_{b=1}^{m} \delta_b \cdot \lambda_b \right) = \gamma_a$$

— so that $\gamma_1 = \gamma_2 = \cdots = \gamma_\ell = 0$ and

$$\rho = \sum_{b=1}^{m} \delta_b \cdot \lambda_b.$$

Notice, next, that it also follows by Invariant #13 that

$$P \cdot \rho = \begin{bmatrix} 0 \\ \widehat{\rho} \end{bmatrix} \tag{31}$$

for a matrix $\widehat{\rho} \in \mathsf{F}_q^{(n-m) \times 1}$. However, Invariant #10 is also satisfied on termination of the algorithm, and this implies that $\lambda_1, \lambda_2, \ldots, \lambda_m$ are the columns of a matrix M_λ such that

$$P \cdot M_\lambda = \begin{bmatrix} L_\lambda \\ X_\lambda \end{bmatrix}$$

where $L_\lambda \in \mathsf{F}_q^{m \times m}$ is nonsingular (indeed, it is lower triangular with ones on its diagonal) and where $X_\lambda \in \mathsf{F}_q^{(n-m) \times m}$. Consequently the equation at line (31) can only be satisfied if $b_1 = b_2 = \cdots = b_m = 0$ and $\rho = 0$ as well — in which case the test at line 30 is passed and $A \cdot x = b$ for the vector $x = \chi - w$ that is returned at line 32.

On the other hand, if the test at line 30 fails — that is, $\rho \neq 0$ — then it follows by the above that the Krylov space $\mathcal{KS}_{\vec{v}}$ does not contain a vector χ such that $A \cdot \chi = A \cdot w + b$, as needed to establish the rest of the claim. □

Proof of Theorem 2.4. This is now a straightforward consequence of Lemma A.16, Lemma A.27 (which establishes that $\ell \geq i - \Delta_{n,k} \cdot k$), and Lemma A.30, above. □

Lemma A.31. *Let i and j be integers such that $0 \leq j \leq i$ and there are at least $i+1$ stages of the Lanczos phase (ending with stage i). Suppose that the matrix $\mathcal{H}_{\vec{u},\vec{v},j+1,i+1}$ has rank r. Then exactly r of the vectors $u_{a,b}$ such that $0 \leq a \leq j$ and $1 \leq b \leq k$ have been matched at the end of stage i of the Lanczos phase of the computation.*

Proof. This is a reasonably straightforward consequence of part (b) of Lemma A.4: As in the statement of this lemma, let $\mathcal{M}_{L,i,j+1} \in \mathsf{F}_q^{n \times k(j+1)}$ be the matrix with columns $u_{a,b}$ such that $0 \leq a \leq j$ and $1 \leq b \leq k$, as these are defined at the end of stage i of the Lanczos phase of the computation, and let $\mathcal{M}_{R,i,i+1} \in \mathsf{F}_q^{n \times k(i+1)}$ be the matrix with columns $v_{c,d}$ such that $0 \leq c \leq i$ and $1 \leq d \leq k$, as these are defined at the end of stage i of the Lanczos phase of the computation as well.

Now suppose that exactly s of the vectors $u_{a,b}$ such that $0 \leq a \leq j$ and $1 \leq b \leq k$ that have been matched at the end of stage i of the Lanczos phase. In particular, suppose that these are the vectors $\mu_{\sigma_1}, \mu_{\sigma_2}, \ldots, \mu_{\sigma_s}$ for integers $\sigma_1, \sigma_2, \ldots, \sigma_s$ such that $1 \leq \sigma_1 < \sigma_2 < \cdots < \sigma_s \leq \ell$.

Suppose, as well, that $s+t$ of the vectors $v_{a,b}$ such that $0 \leq a \leq i$ have been matched at the end of stage i. In particular, suppose that these are the vectors

$$\nu_{\sigma_1}, \nu_{\sigma_2}, \ldots, \nu_{\sigma_s}, \nu_{\tau_1}, \nu_{\tau_2}, \ldots, \nu_{\tau_t}$$

for integers $\tau_1, \tau_2, \ldots, \tau_t$ such that $1 \leq \tau_1 < \tau_2 < \cdots < \tau_t \leq \ell$.

Applying permutations $P_L \in \mathsf{F}_q^{k(j+1) \times k(j+1)}$ and $P_R \in \mathsf{F}_q^{k(i+1) \times k(i+1)}$ we may obtain matrices

$$\widehat{\mathcal{M}}_{L,i,j+1} = \mathcal{M}_{L,i,j+1} \cdot P_L \text{ and } \widehat{\mathcal{M}}_{R,i,i+1} = \mathcal{M}_{R,i,i+1} \cdot P_R$$

such that $\widehat{\mathcal{M}}_{L,i,j+1}$ has as its columns the above vectors

$$\mu_{\sigma_1}, \mu_{\sigma_2}, \ldots, \mu_{\sigma_s},$$

followed by the $k(j+1) - s$ vectors $u_{a,b}$ such that $0 \leq a \leq j$, $1 \leq b \leq k$, and $u_{a,b}$ is unmatched at the end of stage i of the Lanczos phase, and such that $\widehat{\mathcal{M}}_{R,i,i+1}$ has as its columns the above vectors

$$\nu_{\sigma_1}, \nu_{\sigma_2}, \ldots, \nu_{\sigma_s}, \nu_{\tau_1}, \nu_{\tau_2}, \ldots, \nu_{\tau_t},$$

followed by the $k(i+1) - s - t$ vectors $v_{c,d}$ such that $0 \leq c \leq i$, $1 \leq d \leq k$, and $v_{c,d}$ is unmatched at the end of stage i of the Lanczos phase. Since Invariants #2 and #3 are satisfied at the end of the stage i of the Lanczos phase,

$$\widehat{\mathcal{M}}_{L,i,j+1}^T \cdot \widehat{\mathcal{M}}_{R,i,i+1} = \begin{bmatrix} I_s & 0 \\ 0 & 0 \end{bmatrix} \in \mathsf{F}_q^{k(j+1) \times k(i+1)}$$

so that the rank of $\widehat{\mathcal{M}}_{L,i,j}^T \cdot \widehat{\mathcal{M}}_{R,i,i+1}$ is equal to s.

However, Lemma A.4 implies that there exist nonsingular matrices $X_{i,j+1} \in \mathsf{F}_q^{(j+1)k \times (j+1)k}$ and $Y_{i,1+1} \in \mathsf{F}_q^{(i+1)k \times (i+1)k}$ such that

$$\mathcal{M}_{L,i,j+1} = \widehat{\mathcal{K}}_{\vec{u},j+1} \cdot X_{i,j+1}$$

and

$$\mathcal{M}_{R,i,i+1} = \mathcal{K}_{\vec{v},i+1} \cdot Y_{i,i+1}.$$

Now the matrix $X^T_{i,j+1}$ is nonsingular as well and

$$\begin{aligned}
&\mathcal{H}_{\vec{u},\vec{v},j+1,i+1} \\
&\quad = \widehat{\mathcal{K}}^T_{\vec{u},j+1} \cdot \mathcal{K}_{\vec{v},i+1} \\
&\quad = (X^T_{i,j+1})^{-1} \cdot \mathcal{M}^T_{L,i,j+1} \cdot \mathcal{M}_{R,i,i+1} \cdot Y^{-1}_{i,i+1} \\
&\quad = (X^T_{i,j+1})^{-1} \cdot P_L \cdot \widehat{\mathcal{M}}^T_{L,i,j+1} \cdot \widehat{\mathcal{M}}_{R,i,i+1} \cdot P^T_R \cdot Y^{-1}_{i,i+1}
\end{aligned}$$

so that the matrices $\mathcal{H}_{\vec{u},\vec{v},j+1,i+1}$ has rank s, since $\widehat{\mathcal{M}}^T_{L,i,j+1} \cdot \widehat{\mathcal{M}}_{R,i,i+1}$ does. That is, the rank r of $\mathcal{H}_{\vec{u},\vec{v},j+1,i+1}$ is equal to the number of matched vectors s stated in the claim. □

The proof of the next lemma is almost identical to that of the above one.

LEMMA A.32. *Let i and j be integers such that $0 \leq j \leq i$ and there are at least $i+1$ stages of the Lanczos phase (ending with stage #i). Suppose that the matrix $\mathcal{H}_{\vec{u},\vec{v},i+1,j+1}$ has rank r. Then exactly r of the vectors $v_{a,b}$ such that $0 \leq a \leq j$ and $1 \leq b \leq k$ have been matched at the end of stage i of the Lanczos phase of the computation.*

PROOF OF LEMMA 2.5. This can now be established by induction on i.

Basis: If $0 \leq i \leq \Delta_{n,k} - 2$ then both Invariant #5 and the claim are trivially satisfied because both claims are vacuous. Invariant #5 is also trivially satisfied if $i = \Delta_{n,k} - 1$, and the condition in the claim is trivially satisfied as well, because the matrices $\mathcal{H}_{\vec{u},\vec{v},0,\Delta n,k}$ and $\mathcal{H}_{\vec{u},\vec{v},\Delta_{n,k},0}$ each have zero columns (and certainly do have nonnegative renk).

Inductive Step: Suppose $i \leq \Delta_{n,k} - 1$ and that the claim is satisfied for i; it is necessary and sufficient to establish it for $i+1$ as well. Suppose, therefore, that the matrices $\mathcal{H}_{\vec{u},\vec{v},a,a+\Delta_{n,k}}$ and $\mathcal{H}_{\vec{u},\vec{v},a+\Delta_{n,k},a}$ both have maximal rank ak for $0 \leq i - \Delta_{n,k} + 2$. Then, since this condition is satisfied for $0 \leq i \leq \Delta_{n,k} + 1$ it follows by the inductive hypothesis that there Invariant #5 was satisfied at the end of the first i stages, so that there will are at least $i+1$ stages of the Lanczos phase (ending with stage i).

It now follows by Lemmas A.31 and A.32, above, that Invariant #5 is satisfied at the end of stage i as well —for these imply that if both of the matrices $\mathcal{H}_{\vec{u},\vec{v},i-\Delta_{n,k}+1,i+1}$ and $\mathcal{H}_{\vec{u},\vec{v},i+1,i-\Delta_{n,k}+1}$ have full rank $(i - \Delta_{n,k} + 1)k$, then all $(i - \Delta_{n,k} + 1)k$ of the vectors $u_{r,s}$ (respectively, $v_{r,s}$) such that $0 \leq r \leq i - \Delta_{n,k}$ and $1 \leq s \leq k$ have been matched at or before the end of stage i, as required. □

B. PROOFS OF RESULTS IN SECTION 4

B.1 Proof of Lemma 4.1

While the notation used in the report [5] is different, proofs of the next elementary result and Lemma 4.1 can also be found there.

LEMMA B.1. *If a matrix $B \in \mathsf{F}_q^{s \times t}$ has rank r then*

$$\mathsf{xnull}_L(B) = q^{s-r} \text{ and } \mathsf{xnull}_R(B) = q^{t-r}.$$

PROOF. If $B \in \mathsf{F}_q^{s \times t}$ has rank r then there are permutation matrices $P_L \in \mathsf{F}_q^{s \times s}$ and $P_R \in \mathsf{F}_q^{t \times t}$ such that

$$P_L \cdot B \cdot P_R = \begin{bmatrix} B_{1,1} & B_{1,2} \\ B_{2,1} & B_{2,2} \end{bmatrix} \tag{32}$$

where $B_{1,1} \in \mathsf{F}_q^{r \times r}$, $B_{1,2} \in \mathsf{F}_q^{r \times (t-r)}$, $B_{2,1} \in \mathsf{F}_q^{(s-r) \times r}$, $B_{2,2} \in \mathsf{F}_q^{(s-r) \times (t-r)}$, and $B_{1,1}$ is nonsingular. Furthermore, since P_L and P_R are both nonsingular, the matrix shown on the right hand side above also has rank r, so that

$$\begin{bmatrix} B_{1,1} & B_{1,2} \\ B_{2,1} & B_{22} \end{bmatrix} = \begin{bmatrix} I_r & 0 \\ B_{2,1} \cdot B_{1,1}^{-1} & I_{s-r} \end{bmatrix} \cdot \begin{bmatrix} B_{1,1} & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} I_r & B_{1,1}^{-1} \cdot B_{1,2} \\ 0 & I_{t-r} \end{bmatrix} \tag{33}$$

— where the matrices shown above on the right side of the equation are in $\mathsf{F}_q^{s \times s}$, $\mathsf{F}_q^{s \times s} t$, and $\mathsf{F}_q^{t \times t}$ respectively — for an expansion of the product on the right confirms that this is equal to

$$\begin{bmatrix} B_{1,1} & B_{1,2} \\ B_{2,1} & B_{2,1} \cdot B_{1,1}^{-1} \cdot B_{1,2} \end{bmatrix}$$

and this could only be different from

$$\begin{bmatrix} B_{1,1} & B_{1,2} \\ B_{2,1} & B_{2,2} \end{bmatrix}$$

if the right $t-r$ columns of the above matrix were not linear combinations of the left r ones — in which case the rank of the above matrix, and the rank of B, would both exceed r. Note as well that the lower triangular matrix shown at the beginning of the right hand side of the equation at line (33) and the upper triangular matrix at the end of this equation are each nonsingular.

With that noted, consider a vector $x \in \mathsf{F}_q^{r \times 1}$. A consideration of the equations at lines (32) and (33) confirms that $x^T \cdot B = 0$ if and only if

$$P_L \cdot x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

for vectors $x_1 \in \mathsf{F}_q^{r \times 1}$ and $x_2 \in \mathsf{F}_q^{(s-r) \times 1}$ such that

$$\begin{bmatrix} x_1^T & x_2^T \end{bmatrix} \cdot \begin{bmatrix} I_r & 0 \\ B_{2,1} \cdot B_{1,1}^{-1} & I_{s-r} \end{bmatrix} \cdot \begin{bmatrix} B_{1,1} & 0 \\ 0 & 0 \end{bmatrix} = 0$$

(note that, since P_L is a permutation matrix, $P_L^T \cdot P_L = I_s$). Now, since $B_{1,1}$ is nonsingular, an expansion of the above equation confirms that this equation is satisfied for any vector $x_2 \in \mathsf{F}_q^{(s-r) \times 1}$ provided that

$$x_1^T = -x_2^T \cdot B_{2,1} \cdot B_{1,1}^{-1}.$$

Since there are q^{s-r} choices of x_2 and one choice of x_1 for each, for which the above equation is satisfied, it follows that the left exponential nullity of B is q^{s-r} as claimed.

Similarly, if $y \in \mathsf{F}_q^{t \times 1}$ then $B \cdot y = 0$ if and only if

$$P_R^T \cdot y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

where $y_1 \in \mathsf{F}_q^{r \times 1}$, $y_2 \in \mathsf{F}_q^{(t-r) \times 1}$, and

$$\begin{bmatrix} B_{1,1} & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} I_r & B_{1,1}^{-1} \cdot B_{1,2} \\ 0 & I_{t-r} \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = 0.$$

An expansion of the above confirms that this equation is satisfied for any vector $y_2 \in \mathsf{F}_q^{(t-r) \times 1}$ provided that

$$y_1 = -B_{1,1}^{-1} \cdot B_{1,2} \cdot y_2.$$

Since there are q^{t-r} choices of y_2 and one choice of y_1 for each, for which the above equation is satisfied, it follows that the right exponential nullity of B is q^{t-r} as well. □

PROOF OF LEMMA 4.1. Consider $H = \mathcal{H}_{k,\vec{u},\vec{v},s,t} \in \mathsf{F}_q^{sk\times tk}$ where the vectors $\vec{u} = u_1, u_2, \ldots, u_k$, $\vec{w} = w, , w_2, \ldots, w_k$ are chosen uniformly and independently from $\mathsf{F}_q^{n\times 1}$, $w_1 = A \cdot w = b$, and where $v_h = A \cdot w_h$ for $1 \leq h \leq k$, and $\vec{v} = v_1, v_2, \ldots, v_k$. For $0 \leq w \leq \min(s,t)$ set ρ_w to be the probability that H has rank w and set σ_w to be the probability that H as rank at most w. Then it follows by Lemma B.1, above, that

$$\begin{aligned}
\mathsf{E}[\mathsf{xnull}_L(H)] &= \sum_{w=0}^{\min(s,t)} \rho_w \cdot q^{s-w} \\
&= \sum_{w=0}^{r} \rho_w \cdot q^{s-w} + \sum_{w=r+1}^{\min(s,t)} \rho_w \cdot q^{s-w} \\
&\geq \sum_{w=0}^{r} \rho_w \cdot q^{s-r} + \sum_{w=r+1}^{\min(s,t)} \rho_w \cdot 0 \\
&= q^{s-r} \cdot \sum_{w=0}^{r} \rho_w \\
&= q^{s-r} \cdot \sigma_r.
\end{aligned}$$

Dividing both sides of the above inequality by q^{s-r} one can see that the probability that H has rank at most r is less than or equal to $\mathsf{E}[\mathsf{xnull}_L(H)]/q^{s-r}$, as claimed.

The proof that this probability is also less than or equal to $\mathsf{E}[\mathsf{xnull}_R(H)]/q^{t-r}$ follows by another application of the above lemma: It follows by Lemma B.1 that

$$\mathsf{xnull}_R(H) = q^{t-s} \cdot \mathsf{xnull}_L(H)$$

for every matrix H as defined above, so that

$$\mathsf{E}[\mathsf{xnull}_R(H)] = q^{t-s} \cdot \mathsf{E}[\mathsf{xnull}_L(H)]$$

as well — and this suffices to establish the second inequality in the claim.

Finally, if $s \leq t$ then

$$\begin{aligned}
\mathsf{E}[\mathsf{xnull}_L(H)] - 1 &= \sum_{w=0}^{s} \rho_w (q^{s-w} - 1) \\
&= \sum_{w=0}^{s-1} \rho_w (q^{s-w} - 1) \\
&\geq \sum_{w=0}^{s-1} \rho_w (q - 1) \\
&= (q-1) \sum_{w=0}^{s-1} \rho_w \\
&= (q-1) \cdot \mathsf{Pr}[\mathsf{rank}(H) < s],
\end{aligned}$$

so that H has rank less than s with probability at most $(\mathsf{E}[\mathsf{xnull}_L(H)]-1)/(q-1)$ as claimed. It follows by essentially the same argument that if $t \leq s$ then H has rank less than t with probability $(\mathsf{E}[\mathsf{xnull}_R(H)] - 1)/(q-1)$ as well. □

B.2 Proof of Lemma 4.2

LEMMA B.2. *Let H be as described in Lemma 4.2, and let W be a fixed matrix in $\mathsf{F}_q^{sk\times tk}$. Then $\mathsf{E}[\mathsf{xnull}_L(H+W)] \leq \mathsf{E}[\mathsf{xnull}_L(H)]$ and $\mathsf{E}[\mathsf{xnull}_R(H+W)] \leq \mathsf{E}[\mathsf{xnull}_R(H)]$.*

PROOF. Recall that $H = \widehat{K}_L^T \cdot K_R \in \mathsf{F}_q^{ks\times kt}$, where $\widehat{K}_L = \widehat{\mathcal{K}}_{k,\vec{u},s}$ and $K_R = \mathcal{K}_{k,\vec{v},t}$ where vectors $\vec{u} = u_1, u_2, \ldots, u_k$, $\vec{w} = w, w_2, \ldots, w_k$ are chosen uniformly and independently from $\mathsf{F}_q^{n\times 1}$ where $w_1 = A \cdot w$, and $v_h = A \cdot w_h$ for $1 \leq h \leq k$, and $\vec{v} = v_1, v_2, \ldots, v_k$. Since the above vectors $\vec{u}$ and $\vec{w}$ chosen uniformly and independently from $\mathsf{F}_q^{n\times 1}$ we may consider an experiment in which the vectors $\vec{w}$ (and corresponding vectors $\vec{v}$) are fixed, with the vectors $\vec{u}$ being chosen after that.

With that noted, consider the following pair of indicator random variables — which are defined after vectors $\vec{u}$ have been selected and are functions of the vectors $\vec{w}$. For a given vector $x \in \mathsf{F}_q^{s\times 1}$,

- $I_{\vec{u},x}$ depends on the vectors $\vec{w}$ and is equal to one if $x^T \cdot H = 0$, for $H = \widehat{K}_L \cdot K_R$ as above, and is equal to 0 otherwise.
- $\widehat{I}_{W,\vec{u},x}$ depends on the vectors $\vec{w}$ and is equal to one if $x^T \cdot (H + W) = 0$ for H as above, and is equal to 0 otherwise.

Now, next, that

$$\mathsf{E}[\mathsf{xnull}_L(H)] = \sum_{x \in \mathsf{F}_q^{r\times 1}} S_L(x) \tag{34}$$

where

$$S_L(x) = q^{-kn} \cdot \sum_{\vec{u}=u_1,u_2,\ldots,u_k \in \mathsf{F}_q^{n\times 1}} \mathsf{E}[I_{\vec{u},x}] \tag{35}$$

and

$$\mathsf{E}[\mathsf{xnull}_L(H+W)] = \sum_{x \in \mathsf{F}_q^{r\times 1}} \widehat{S}_L(W,x) \tag{36}$$

where

$$\widehat{S}_L(W,x) = q^{-kn} \cdot \sum_{\vec{u}=u_1,u_2,\ldots,u_k \in \mathsf{F}_q^{n\times 1}} \mathsf{E}[\widehat{I}_{W,\vec{u},x}] \tag{37}$$

and where the expectations shown on the right sides of equations (35) and (37), above, concern probability spaces in which the matrix W and vectors $\vec{u}$ and x are fixed, with vectors $\vec{w}$ to be selected uniformly and independently from $\mathsf{F}_q^{n\times 1}$. Note that if

$$\mathsf{Pr}[\widehat{I}_{W,\vec{u}x} = 1] \leq \mathsf{Pr}[I_{\vec{u},x} = 1]$$

for all vectors $\vec{u} = u_1, u_2, \ldots, w_k \in \mathsf{F}_q^{n\times 1}$ and $x \in \mathsf{F}_q^{r\times 1}$, then it would follow (by an inspection of the equations at lines (35) and (37), above) that

$$\widehat{S}_L(W,x) \leq S_L(x)$$

for all $x \in \mathsf{F}_q^{r\times 1}$ as well. This would also imply (by a consideration of the equations at lines (34) and (36)) that

$$\mathsf{E}[\mathsf{xnull}_L(H+W)] \leq \mathsf{E}[\mathsf{xnull}(H)],$$

as required to prove the first part of the lemma.

With that noted, two cases should be considered.

Case (i): There exist vectors $\widehat{w}, \widehat{w}_2, \widehat{w}_3, \widehat{w}_3, \ldots, \widehat{w}_k \in \mathsf{F}_q^{n\times 1}$ such that, if $\widehat{w}_1 = A \cdot \widehat{w}$ and $\widehat{v}_h = A \cdot \widehat{w}_h$ for $1 \leq h \leq k$, and $\vec{z} = \widehat{v}_1, \widehat{v}_2, \ldots, \widehat{v}_k$, then

$$x^T \cdot \widehat{\mathcal{K}}_{k,\vec{u},s} \cdot \mathcal{K}_{k,\vec{z},t} = x^T \cdot W$$

Case (ii) No such vectors $\widehat{w}, \widehat{w}_2, \ldots, \widehat{w}_k \in \mathsf{F}_q^{n\times 1}$ exists.

Case (i): In this case $\mathsf{Pr}[\widehat{I}_{W,\vec{u},x} = 1] = \mathsf{Pr}[I_{\vec{u},x} = 1]$: For vectors $w, w_2, w_3, \ldots, w_k \in \mathsf{F}_q^{n\times 1}$ are chosen with the same probability as $w - \widehat{w}, w_2 - \widehat{w}_2, w_3 - \widehat{w}_3, \ldots, w_k - \widehat{w}_k$ — and

if $\widehat{H} = \widehat{\mathcal{K}}_{k,\vec{u},s} \cdot \mathcal{K}_{k,\vec{y},t}$ for $\vec{y} = v_1 - \widehat{v}_1, v_2 - \widehat{v}_2, \ldots, v_k - \widehat{v}_k$, and $H = \widehat{\mathcal{K}}_{k,\vec{u},s} \cdot \mathcal{K}_{k,\vec{v},t}$ then $x^T \cdot (\widehat{H} + W) = x^T \cdot H$, so that $x^T \cdot (\widehat{H} + W) = 0$ if and only if $x^T \cdot H = 0$.

Case (ii): In this case there are no vectors

$$\vec{w} = w, w_2, w_3, \ldots, w_k \in \mathsf{F}_q^{n \times 1}$$

such that $x^T \cdot (\widehat{\mathcal{K}}_{k,\vec{u},s} \cdot \mathcal{K}_{k,\vec{v},t} + W) = 0$ at all. Consequently,

$$\mathsf{Pr}[\widehat{I}_{W,\vec{u},x} = 1] = 0,$$

and the desired inequality is trivial.

The second part of the lemma now follows by additional applications of Lemma B.1: For any matrix

$$H = \widehat{\mathcal{K}}^T_{k,\vec{u},s} \cdot \mathcal{K}_{k,\vec{v},t} \in \mathsf{F}_q^{sk \times tk},$$

it follows by the relationship between left and right exponential nullities and rank that

$$\mathsf{xnull}_R(H) = q^{(t-s)k} \cdot \mathsf{xnull}_L(H)$$

and

$$\mathsf{xnull}_R(H + W) = q^{(t-s)k} \cdot \mathsf{xnull}_L(H + W)$$

as well. Consequently if the matrix H is chosen using the distribution considered here then

$$\mathsf{E}[\mathsf{xnull}_R(H)] = q^{(t-s)k} \cdot \mathsf{E}[\mathsf{xnull}_L(H)],$$

and

$$\mathsf{E}[\mathsf{xnull}_R(H + W)] = q^{(t-s)k} \cdot \mathsf{E}[\mathsf{xnull}_L(H + W)],$$

so that fact that

$$\mathsf{E}[\mathsf{xnull}_L(H + W)] \leq \mathsf{E}[\mathsf{xnull}_L(H)]$$

(established above) implies that

$$\mathsf{E}[\mathsf{xnull}_R(H + W)] \leq \mathsf{E}[\mathsf{xnull}_R(H)]$$

as well. □

The lemma establishes a result that is, in some sense, intermediate between the above and Lemma 4.2. It will be of use in proving results from Section 6.

LEMMA B.3. *Suppose that vectors*

$$u_1, u_2, \ldots, u_k, w, w_2, w_3, \ldots, w_k$$

are chosen uniformly and independently from $\mathsf{F}_q^{n \times 1}$*, and that* $\vec{z} = z_1, z_2, \ldots, z_k$ *is a fixed sequence of vectors in* $\mathsf{F}_q^{n \times 1}$*. If* $w_1 = A \cdot w$*,* $v_a = A \cdot w_a$ *for* $1 \leq a \leq k$*, and* $\vec{y} = v_1, v_2, \ldots, v_k$*, and* s *and* t *are positive integers, then*

$$\mathsf{E}[\mathsf{xnull}_L(\mathcal{H}_{\vec{u},\vec{y},s,t} + \mathcal{H}_{\vec{u},\vec{z},s,t})] \leq \mathsf{E}[\mathsf{xnull}_L(\mathcal{H}_{\vec{u},\vec{y},s,t})]$$

and

$$\mathsf{E}[\mathsf{xnull}_R(\mathcal{H}_{\vec{u},\vec{y},s,t} + \mathcal{H}_{\vec{u},\vec{z},s,t})] \leq \mathsf{E}[\mathsf{xnull}_R(\mathcal{H}_{\vec{u},\vec{y},s,t})].$$

PROOF. As in the previous proof, one should consider indicator random variables that are defined after the vectors $\vec{u} = u_1, u_2, \ldots, u_k$ have been selected and that are functions of the vectors $\vec{w} = w, w_2, w_3, \ldots, w_k$ for a given vector $x \in \mathsf{F}_q^{tk \times 1}$: For a given vector $x \in \mathsf{F}_q^{tk \times 1}$,

- $I_{\vec{u},x}$ depends on the vectors $\vec{w}$, is equal to one if

$$\mathcal{H}_{\vec{u},\vec{y},s,t} \cdot x = 0,$$

and is equal to zero otherwise, and

- $\widehat{I}_{\vec{u},x}$ depends on the vectors $\vec{w}$ and is equal to one if

$$(\mathcal{H}_{\vec{u},\vec{y},s,t} + \mathcal{H}_{\vec{u},\vec{z},s,t}) \cdot x = 0$$

and is equal to one, otherwise.

As in the previous proof one should consider two cases.

- *Case:* There exist $\widehat{w}, \widehat{w}_2, \widehat{w}_3, \ldots, \widehat{w}_k \in \mathsf{F}_q^{n \times 1}$ such that if $\widehat{w}_1 = A \cdot \widehat{w}$, $\widehat{v}_a = A \cdot \widehat{w}_a$ for $1 \leq a \leq k$, and $\vec{y}' = \widehat{v}_1, \widehat{v}_2, \ldots, \widehat{v}_k$, then

$$\mathcal{H}_{\vec{u},\vec{y}',s,t} \cdot x = \mathcal{H}_{\vec{u},\vec{z},s,t} \cdot x$$

- No such vectors $\widehat{w}, \widehat{w}_2, \widehat{w}_3, \ldots, \widehat{w}_k$ exist.

In the first case one argues as in the first case for the previous proof that

$$\mathsf{E}[\widehat{I}_{\vec{u},x}] = \mathsf{E}[I_{\vec{u},x}],$$

which certainly implies that $\mathsf{E}[\widehat{I}_{\vec{u},x}] \leq \mathsf{E}[I_{\vec{u},x}]$. In the second case one should observe (as for the second case in the previous proof) that $\mathsf{E}[\widehat{I}_{\vec{u},x}] = 0$, so that $\mathsf{E}[\widehat{I}_{\vec{u},x}] \leq \mathsf{E}[I_{\vec{u},x}]$ in this case as well.

The rest of the proof proceeds just as before. □

PROOF OF LEMMA 4.2. Let $\mathcal{H}_{k,r,s}$ be the set of matrices $H = \widehat{\mathcal{K}}^T_{k,\vec{u},s} \cdot \mathcal{K}_{k,\vec{v},t} \in \mathsf{F}_q^{sk \times tk}$ defined from vectors $\vec{u}$ and $\vec{w}$ (where $w_1 = A \cdot w$, $v_h = A \cdot w_h$ for $1 \leq h \leq k$ and $\vec{v} = v_1, v_2, \ldots, v_k$) as described in the claim. Similarly let $\mathcal{W}_{k,r,s} \subseteq \mathsf{F}_q^{sk \times tk}$ be the set from which the matrices W described in the claim are selected. In order to establish the claim in the lemma about left exponential nullities it is necessary and sufficient to show that

$$\sum_{W \in \mathcal{W}_{k,r,s}} \sum_{H \in \mathcal{H}_{k,r,s}} \rho(H, W) \cdot \mathsf{xnull}_L(H + W) \tag{38}$$

is less than or equal to

$$\sum_{W \in \mathcal{W}_{k,r,s}} \sum_{H \in \mathcal{H}_{k,r,s}} \rho(H, W) \cdot \mathsf{xnull}_L(H) \tag{39}$$

where $\rho(H, W)$ is the probability that the matrix H is selected from $\mathcal{H}_{k,r,s}$ and W is selected from $\mathcal{W}_{k,r,s}$, for $H \in \mathcal{H}_{k,r,s}$ and $W \in \mathcal{W}_{k,r,s}$. Now, since the matrices H and W are chosen independently,

$$\rho(H, W) = \rho_1(H) \cdot \rho_2(W)$$

where $\rho_1(H)$ is the probability that H is selected from $\mathcal{H}_{k,r,s}$ and $\rho_2(W)$ is the probability that W is chosen from $\mathcal{W}_{k,r,s}$.

Consequently

$$\begin{aligned}
&\sum_{W \in \mathcal{W}_{k,r,s}} \sum_{H \in \mathcal{H}_{k,r,s}} \rho(H, W) \cdot \mathsf{xnull}_L(H + W) \\
&= \sum_{W \in \mathcal{W}_{k,r,s}} \rho_2(W) \cdot \left(\sum_{H \in \mathcal{H}_{k,r,s}} \rho_1(H) \cdot \mathsf{xnull}_L(H + W) \right) \\
&\leq \sum_{W \in \mathcal{W}_{k,r,s}} \rho_2(W) \cdot \left(\sum_{H - \mathcal{H}_{k,r,s}} \rho_1(H) \cdot \mathsf{xnull}_L(H) \right).
\end{aligned}$$

The final inequality shown here is a consequence of Lemma B.2, above, since

$$\sum_{H \in \mathcal{H}_{k,r,s}} \rho_1(H) \cdot \mathsf{xnull}_L(H + W)$$

and

$$\sum_{H \in \mathcal{H}_{k,r,s}} \rho_1(H) \cdot \mathsf{xnull}_L(H)$$

are, respectively, the expected values of the left exponential nullities of $H+W$ and of H, for a fixed matrix $W \in \mathsf{F}_q^{sk \times tk}$ when H is randomly selected as described as these claims. Now, since

$$\sum_{W \in \mathcal{W}_{k,r,s}} \rho_2(W) \cdot \left(\sum_{H \in \mathcal{H}_{k,r,s}} \rho_1(H) \cdot \mathsf{xnull}_L(H) \right) = \sum_{W \in \mathcal{W}_{k,r,s}} \sum_{H \in \mathcal{H}_{k,r,s}} \rho(H,W) \cdot \mathsf{xnull}_L(H),$$

it follows that the sum at line (38) is less than or equal to the sum at line (39), as required to complete the proof. □

C. PROOFS OF RESULTS IN SECTION 5

C.1 Summary of Prior Results

We will begin with a summary of results from the technical report [5]. Unfortunately the notation used in that report differs from what is being used here, so a translation is also being provided.

Note that for every matrix $A \in \mathsf{F}_q^{n \times n}$ there exists an integer $N \leq n$ and a matrix $\widehat{A} \in \mathsf{F}_q^{N \times N}$ such that

$$X \cdot A \cdot X^{-1} = \begin{bmatrix} \widehat{A} & 0 \\ 0 & 0 \end{bmatrix} \tag{40}$$

for a nonsingular matrix $X \in \mathsf{F}_q^{n \times n}$, $\widehat{A}$ has the same number of nontrivial invariant factors as A, and every invariant factor of $\widehat{A}$ (in $\mathsf{F}_q[x]$) is either divisible by x^2 or not divisible by x. Furthermore, for vectors $u, w \in \mathsf{F}_q^{n \times 1}$, if

$$u = X^T \cdot \begin{bmatrix} \widehat{u} \\ \widetilde{u} \end{bmatrix} \quad \text{and} \quad w = X^{-1} \cdot \begin{bmatrix} \widehat{w} \\ \widetilde{w} \end{bmatrix}$$

where $\widehat{u}, \widehat{w} \in \mathsf{F}_q^{N \times 1}$ and $\widetilde{u}, \widetilde{w} \in \mathsf{F}_q^{(n-N) \times 1}$, and $s \geq 1$, then

$$\begin{aligned} & u^T \cdot A^s \cdot w \\ &= \begin{bmatrix} \widehat{u} \\ \widetilde{u} \end{bmatrix}^T \cdot X \cdot A^s \cdot X^{-1} \cdot \begin{bmatrix} \widehat{w} \\ \widetilde{w} \end{bmatrix} \\ &= \begin{bmatrix} \widehat{u} \\ \widetilde{u} \end{bmatrix}^T \cdot (X \cdot A \cdot X^{-1})^s \cdot \begin{bmatrix} \widehat{w} \\ \widetilde{w} \end{bmatrix} \\ &= \begin{bmatrix} \widehat{u} \\ \widetilde{u} \end{bmatrix} \cdot \begin{bmatrix} \widehat{A} & 0 \\ 0 & 0 \end{bmatrix}^s \cdot \begin{bmatrix} \widehat{w} \\ \widetilde{w} \end{bmatrix} \\ &= \widehat{u}^T \cdot \widehat{A}^s \cdot \widehat{w}. \end{aligned}$$

Note as well that if $u_1, u_2, \ldots, u_k, w_1, w_2, \ldots, w_k$ are chosen uniformly and independently from $\mathsf{F}_q^{n \times 1}$ then the corresponding vectors $\widehat{u}_1, \widehat{u}_2, \ldots, \widehat{u}_k, \widehat{w}_1, \widehat{w}_2, \ldots, \widehat{w}_k$ are chosen uniformly and independently from $\mathsf{F}_q^{N \times 1}$.

The matrices A and $\widehat{A}$ have the same rank and the same number h of nontrivial invariant factors. Furthermore, since each invariant factor of $\widehat{A}$ is either divisible by x^2 or not divisible by x at all, it can be shown that

$$N - h \leq r \leq N$$

if A has h nontrivial invariant factors and rank r. Indeed, r is equal to the difference between N and the number of invariant factors of $\widehat{A}$ that are divisible by x^2.

As in Section 5, let

$$f(h,k) = \begin{cases} 6 \cdot \log_q N & \text{if } k = h+1, \\ 4 & \text{if } k = h+2, \\ 1 + 2q^{h-k+1} & \text{if } k \geq h+3. \end{cases} \tag{41}$$

LEMMA C.1. *Suppose A, $\widehat{A}$ and N are as above, $k > h$, and vectors*

$$\vec{u} = u_1, u_2, \ldots, u_k$$

(respectively, $\vec{w} = w_1, w_2, \ldots, w_k$) are chosen uniformly and independently from $\mathsf{F}_q^{n \times 1}$. If s is an integer such that $1 \leq s \leq \lfloor N/k \rfloor$, then

$$\mathsf{E}[\mathsf{xnull}_R(\widehat{\mathcal{K}}_{\vec{u},s})], \mathsf{E}[\mathsf{xnull}_R(\mathcal{K}_{\vec{w},s})] \leq 1 + f(h,k) \cdot q^{sk-N}.$$

PROOF. The bound given here for the expected right exponential nullity of $\mathcal{K}_{\vec{w},s}$ is a consequence of Lemma 2.20 of [5], which presents an upper bound for the product the expected value of $\mathsf{xnull}_R(\mathcal{K}_{\vec{w},r})$ and the number q^{kN} of choices of the vectors $w_1, w_2, \ldots, w_k$ when $n = N$. In the report, m is being used as the block size instead of k, ℓ is the number of nontrivial invariant factors instead of h, and i is the number sk of columns of the matrix being considered.

Since A^T has the same number of nontrivial invariant factors as A the claimed bound for the expected exponential nullity of $\widehat{\mathcal{K}}_{\vec{u},r}$ follows by the same argument. □

When sampling from the null space one is generally working with a matrix $\mathcal{K}_{\vec{v},r}$ for $\vec{v} = v_1, v_2, \ldots, v_k$ where $v_s = A \cdot w_s$ for $1 \leq s \leq k$ instead.

LEMMA C.2. *Suppose A, $\widehat{A}$ and N are as above, $k > h$, A has rank r, vectors $\vec{w} = w_1, w_2, \ldots, w_k$ are chosen uniformly and independently from $\mathsf{F}_q^{n \times 1}$, and $\vec{v} = v_1, v_2, \ldots, v_k$ where $v_s = A \cdot w_s$ for $1 \leq s \leq k$. If s is an integer such that $1 \leq s \leq \lfloor r/k \rfloor$ then*

$$\mathsf{E}[\mathsf{xnull}_R(\mathcal{K}_{\vec{v},s})] \leq 1 + f(h,k) \cdot q^{sk+1-r}.$$

PROOF. See Section 3.1 of [5], noting that the expected value discussed here is the ratio of the value "$\widehat{D}_{A,m,i}$" defined in this report to q^{Nm} when (once again) m is being used as the block size instead of k and i is the number sk of columns in the matrix being considered. □

A generalization of "exponential nullity" is useful in order to consider the block Hankel matrices $\mathcal{H}_{\vec{u},\vec{v},a,b}$ being considered in this report. Once again, consider an arbitrary matrix $B \in \mathsf{F}_q^{s \times t}$ for positive integers s and t, and let $\mathcal{X}$ and $\mathcal{Y}$ be subspaces of $\mathsf{F}_q^{s \times 1}$ and $\mathsf{F}_q^{t \times 1}$, respectively. Let us define the *right-$\mathcal{X}$-nullity* of B, $\mathsf{xnull}_R(\mathcal{X}, B)$, to be the number of vectors $y \in \mathsf{F}_q^{t \times 1}$ such that $B \cdot y \in \mathcal{X}$, and let us define the *left-$\mathcal{Y}$-nullity* of B, $\mathsf{xnull}_L(\mathcal{Y}, B)$, to be the number of vectors $x \in \mathsf{F}_q^{s \times 1}$ such that $x^T \cdot B \in \mathcal{Y}$.

Since A^T has the same number of nontrivial invariant factors as A, the following is easily established from the results in [5].

LEMMA C.3. *Suppose A, $\widehat{A}$ and N are as above. Let $\mathcal{X}$ be a subspace of $\mathsf{F}_q^{n \times 1}$ with dimension d, let vectors*

$$\vec{u} = u_1, u_2, \ldots, u_k$$

be chosen uniformly and independently from $\mathsf{F}_q^{n\times 1}$*, and let* s *be an integer such that* $1 \le s \le \lfloor (N-d)/k \rfloor$*. Then*

$$\mathsf{E}[\mathsf{xnull}_R(\mathcal{X}, \widehat{\mathcal{K}}_{\vec{u},s})] \le 1 + f(h,k)\cdot q^{sk+d-N}$$

PROOF. Recalling that the matrices A^T and A have the same number of nontrivial invariant factors, this follows by the results presented in Section 3.2 of [5] — note, in particular, Lemma 3.5. □

This remaining extension was not explored in the previous report — the arguments in Sections 3.1 and 3.2 were never combined. However, with the benefit of hindsight it is clear that this is not difficult, and that the following result can be established from Lemma C.1 along with an extremely minor modification of Lemma 3.5 of [5] and its proof.

LEMMA C.4. *Suppose* A*,* $\widehat{A}$ *and* N *are as above. Let* $\mathcal{X}$ *be a subspace of* $\mathsf{F}_q^{n\times 1}$ *with dimension* d*, let vectors*

$$\vec{w} = w_1, w_2, \ldots, w_k$$

be chosen uniformly and independently from $\mathsf{F}_q^{n\times 1}$ *and let* $\vec{v} = v_1, v_2, \ldots, v_k$ *where* $v_a = A\cdot w_a$ *for* $1 \le a \le k$*, and let* s *be an integer such that* $1 \le s \le \lfloor (r-d)/k \rfloor$*. Then*

$$\mathsf{E}[\mathsf{xnull}_R(\mathcal{X}, \mathcal{K}_{\vec{v},s})] \le 1 + f(h,k)\cdot q^{sk+d+1-r}.$$

C.2 Modifying the Initial Vectors

Recall that in the algorithm described in Section 2, vectors $u_1, u_2, \ldots, u_k$, w, and $w_2, w_3, \ldots, w_k$ are chosen uniformly and independently from $\mathsf{F}_q^{n\times 1}$, w_1 is set to $A\cdot w + b$, and v_a is set to be $A\cdot w_a$ for $1 \le a \le k$.

An intermediate situation will be considered here first: Suppose, as above, that $u_1, u_2, \ldots, u_k$, w and $w_2, w_3, \ldots, w_k$ are chosen uniformly and independently from $\mathsf{F}_q^{n\times 1}$, w_1 is set to be $A\cdot w$, instead, and v_a is set to be $A\cdot w_a$ for $1 \le a \le k$, once again. In this case the following is easily established.

LEMMA C.5. *Let*

$$\vec{v} = v_1, v_2, \ldots, v_k \quad \textit{and} \quad \vec{z} = w_1, v_2, v_3, \ldots, v_k$$

for the factors described above. Let t *be an integer such that* $t \le \lfloor r/k \rfloor$ *where* r *is the rank of* A*. Then*

$$\mathsf{xnull}_R(\mathcal{K}_{\vec{v},t}) \le \mathsf{xnull}_R(\mathcal{K}_{\vec{z},t+1})$$

and, if $\mathcal{X}$ *is a subspace of* $\mathsf{F}_q^{n\times 1}$ *then*

$$\mathsf{xnull}_R(\mathcal{X}, \mathcal{K}_{\vec{v},t}) \le \mathsf{xnull}_R(\mathcal{X}, \mathcal{K}_{\vec{z},t+1})$$

as well.

PROOF. Note that, since $v_1 = A\cdot w_1$, the columns of $\mathcal{K}_{\vec{v},t}$ are also columns of $\mathcal{K}_{\vec{z},t+1}$, so there exists a permutation matrix $P \in \mathsf{F}_q^{k(t+1)\times k(t+1)}$ such that

$$\mathcal{K}_{\vec{z},t+1}\cdot P = \begin{bmatrix}\mathcal{K}_{\vec{v},t} & Y\end{bmatrix}$$

for a matrix $Y \in \mathsf{F}_q^{n\times k}$. Now, for an arbitrary vector $y \in \mathsf{F}_q^{kt\times 1}$, let

$$\widehat{y} = P\cdot \begin{bmatrix} y \\ 0 \end{bmatrix} \in \mathsf{F}_q^{k(t+1)\times 1}$$

noticing that if $y_1, y_2 \in \mathsf{F}_q^{kt\times 1}$ and $y_1 \ne y_2$ then $\widehat{y}_1 \ne \widehat{y}_2$ as well. It now suffices to notice that

$$\begin{aligned}\mathcal{K}_{\vec{z},t+1}\cdot\widehat{y} &= (\mathcal{K}_{\vec{z},t+1}\cdot P)\cdot\begin{bmatrix} y \\ 0\end{bmatrix} \\ &= \begin{bmatrix}\mathcal{K}_{\vec{v},t} & Y\end{bmatrix}\cdot\begin{bmatrix} y \\ 0\end{bmatrix} \\ &= \mathcal{K}_{\vec{v},t}\cdot y.\end{aligned}$$

Consequently $\mathcal{K}_{\vec{z},t+1}\cdot\widehat{y} = 0$ if and only if $\mathcal{K}_{\vec{v},t}\cdot y = 0$ and $\mathcal{K}_{\vec{z},t+1}\cdot\widehat{y} \in \mathcal{X}$ if and only if $\mathcal{K}_{\vec{v},t}\cdot y \in \mathcal{X}$. The claim now follows by an application of the definitions of "right exponential nullity" and "right-$\mathcal{X}$-nullity." □

LEMMA C.6. *Let* $\vec{v}$ *and* t *be as in the previous lemma and let* H *be a fixed matrix in* $\mathsf{F}_q^{n\times tk}$*. Then, if* $\mathcal{X}$ *is a subspace of* $\mathsf{F}_q^{n\times 1}$*, then*

$$\mathsf{E}[\mathsf{xnull}_R(\mathcal{X}, \mathcal{K}_{\vec{v},t} + H)] \le \mathsf{E}[\mathsf{xnull}_R(\mathcal{X}, \mathcal{K}_{\vec{v},t})].$$

PROOF. The proof is similar to that of Lemma B.2 (which establishes the claim if $\mathcal{X} = \{0\}$): For a fixed vector $x \in \mathsf{F}_q^{tk\times 1}$, consider a pair of indicator random variables (that may be thought of as functions of the vectors $\vec{v}$):

$$I_{\mathcal{X},x}(\vec{v}) = \begin{cases} 1 & \text{if } \mathcal{K}_{\vec{v},t}\cdot x \in \mathcal{X}, \\ 0 & \text{otherwise,}\end{cases}$$

and

$$\widehat{I}_{\mathcal{X},x}(\vec{v}) = \begin{cases} 1 & \text{if } (\mathcal{K}_{\vec{v},t} + H)\cdot x \in \mathcal{X}, \\ 0 & \text{otherwise.}\end{cases}$$

Now, for a given vector x there are two cases to be considered.

Case (i): There exist vectors $\widehat{w}, \widehat{w_2}, \widehat{w}_3, \ldots, \widehat{w}_k \in \mathsf{F}_q^{n\times 1}$ such that if $\widehat{w}_1 = A\cdot\widehat{w}$, $\widehat{v}_h = A\cdot\widehat{w}_h$, and $\vec{z} = \widehat{v}_1, \widehat{v}_2, \ldots, \widehat{v}_k$, then $(\mathcal{K}_{\vec{z},t} + H)\cdot x \in \mathcal{X}$

In this case one can observe that, since vectors $w - \widehat{w}, w_2 - \widehat{w}_2, w_3 - \widehat{w}_3, \ldots, w_k - \widehat{w}_k$ are chosen uniformly and independently from $\mathsf{F}_q^{n\times 1}$ if $\widehat{w}, \widehat{w}_2, \widehat{w}_3, \ldots, \widehat{w}_k$ are,

$$\mathsf{E}[\widehat{I}_{\mathcal{X},x}] = \mathsf{E}[I_{\mathcal{X},x}].$$

Case: There are no such vectors $\widehat{w}, \widehat{w}_2, \widehat{w}_3, \ldots, \widehat{w}_k$. In this case $\widehat{I}_{\mathcal{X},x}(\vec{v}) = 0$ for all choices of the vectors $\widehat{v}$, so that

$$\mathsf{E}[\widehat{I}_{\mathcal{X},x}] = 0 \le \mathsf{E}[I_{\mathcal{X},x}].$$

Now it suffices to note that

$$\mathsf{E}[\mathsf{xnull}_R(\mathcal{X}, \mathcal{K}_{\vec{v},t} + H)] = \sum_{x\in\mathsf{F}_q^{tk\times 1}} \mathsf{E}[\widehat{I}_{\mathcal{X},x}]$$

and

$$\mathsf{E}[\mathsf{xnull}_R(\mathcal{X}, \mathcal{K}_{\vec{v},t})] = \sum_{x\in\mathsf{F}_q^{tk\times 1}} \mathsf{E}[I_{\mathcal{X},x}]$$

in order to use the above to establish that

$$\mathsf{E}[\mathsf{xnull}_R(\mathcal{X}, \mathcal{K}_{\vec{v},t} + H)] \le \mathsf{E}[\mathsf{xnull}_R(\mathcal{X}, \mathcal{K}_{\vec{v},t})]$$

as claimed. □

Now, to obtain bounds on the values

$$\mathsf{E}[\mathsf{xnull}_R(\mathcal{K}_{\vec{v},t})] \text{ and } \mathsf{E}[\mathsf{xnull}_R(\mathcal{X}, \mathcal{K}_{\vec{v},t})]$$

when $w, w_2, w_3, \ldots, w_k$ are chosen uniformly and independently from $\mathsf{F}_q^{n\times 1}$, $w_1 = A\cdot w + b$, $v_h = A\cdot w_h$ for $1 \le h \le k$, and $\vec{v} = v_1, v_2, \ldots, v_k$, it suffices to notice that

$$\mathcal{K}_{\vec{v},t} = \mathcal{K}_{\vec{y},t} + \mathcal{K}_{\vec{z},t}$$

where $\widehat{w}_1 = A\cdot w$, $y_1 = A\cdot \widehat{w}_1$, $y_h = A\cdot w_h$ for $2 \le h \le k$, and $\vec{y} = y_1, y_2, \ldots, y_k$ — so that Lemma C.5 can be applied to bound

$$\mathsf{E}[\mathsf{xnull}_R(\mathcal{K}_{\vec{y},t})] \text{ and } \mathsf{E}[\mathcal{X}, \mathsf{xnull}_R(\mathcal{K}_{\vec{y},t})]$$

— and where $\vec{z} = A\cdot b, 0, 0, \ldots, 0$ — so that $K_{\vec{z},t}$ is a fixed matrix in $\mathsf{F}_q^{n\times tk}$, and Lemmas B.2 and C.6 can be applied to establish that

$$\mathsf{E}[\mathsf{xnull}_R(\mathcal{K}_{\vec{v},t})] \le \mathsf{E}[\mathsf{xnull}_R(\mathcal{K}_{\vec{y},t})]$$

and

$$\mathsf{E}[\mathsf{xnull}_R(\mathcal{X}, \mathcal{K}_{\vec{v},t})] \le \mathsf{E}[\mathsf{xnull}_R(\mathcal{X}, \mathcal{K}_{\vec{y},t})]$$

as well.

Indeed, applying Lemmas C.2 and C.4 along with the above, we can now obtain the following bounds.

LEMMA C.7. *Suppose vectors $w, w_2, w_3, \ldots, w_k$ are chosen uniformly and independently from $\mathsf{F}_q^{n\times 1}$, $w_1 = A\cdot w + b$, $v_a = A\cdot w_a$ for $1 \le a \le k$, and $\vec{v} = v_1, v_2, \ldots, v_k$. It t is an integer such that $1 \le t < \lfloor r/k \rfloor$ where r is the rank of A then*

$$\mathsf{E}[\mathsf{xnull}_R(\mathcal{K}_{\vec{v},t})] \le 1 + f(h,k)\cdot q^{(t+1)k+1-r}$$

and if $\mathcal{X}$ is a subspace of $\mathsf{F}_q^{n\times 1}$ with dimension $d < r$ and $1 \le t < \lfloor (r-d)/k \rfloor$ then

$$\mathsf{E}[\mathsf{xnull}_R(\mathcal{X}, \mathcal{K}_{\vec{v},t})] \le 1 + f(h,k)\cdot q^{(t+1)k+d+1-r}.$$

C.3 Proof of Lemma 5.1

As suggested at the beginning of Appendix C, it suffices to consider the case that $n = N$ and $\widehat{A} = A$, for $\widehat{A}$ as shown at line (40) above — for if $u_1, u_2, \ldots, u_k, w, w_2, \ldots, w_k$ are chosen uniformly and independently from $\mathsf{F}_q^{n\times 1}$ then the corresponding vectors $\widehat{u}_1, \widehat{u}_2, \ldots, \widehat{u}_k, \widehat{w}, \widehat{w}_2, \ldots, \widehat{w}_k$ are chosen uniformly and independently from $\mathsf{F}_q^{N\times 1}$.

Furthermore, if one sets $\vec{y} = \widehat{u}_1, \widehat{u}_2, \ldots, \widehat{u}_k$ sets $\widehat{w}_1 = \widehat{A}\cdot \widehat{w} + \widehat{b}$ where

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

for $\widehat{b} \in \mathsf{F}_q^{N\times 1}$ and $\widetilde{b} \in \mathsf{F}_q^{(n-N)\times 1}$, sets $\widehat{v}_a = \widehat{A}\cdot \widehat{w}_a$ for $1 \le a \le k$ and, finally, sets $\vec{z} = \widehat{v}_1, \widehat{v}_2, \ldots, \widehat{v}_k$ then, since $u_a^T \cdot A^c \cdot w_b = \widehat{u}_a^T \cdot \widehat{A}^c \cdot \widehat{w}_b$ for $1 \le a, b \le k$ and $c \ge 1$,

$$\mathcal{H}_{\vec{u},\vec{v},s,t} = \mathcal{H}_{\vec{y},\vec{z},s,t}$$

for all integers $s, t \ge 1$.

We will therefore assume, for the rest of this proof, that $n = N$ and $A = \widehat{A}$.

Suppose first that $s < t < \lfloor r/k \rfloor$. Notice that if $z \in \mathsf{F}_q^{sk\times 1}$, then

$$\begin{aligned} z^T \cdot \mathcal{H}_{\vec{u},\vec{v},s,t} = 0 &\iff \mathcal{H}_{\vec{u},\vec{v},s,t}^T \cdot z = 0\\ &\iff \mathcal{K}_{\vec{v},t}^T \cdot (\widehat{\mathcal{K}}_{\vec{u},s}\cdot z) = 0\\ &\iff \widehat{\mathcal{K}}_{\vec{u},s}\cdot z \in \mathcal{X}\end{aligned}$$

where $\mathcal{X}$ is the subspace of $\mathsf{F}_q^{n\times 1}$ consisting of the vectors $y \in \mathsf{F}_q^{n\times 1}$ such that $y^T \cdot \mathcal{K}_{\vec{v},t} = 0$. Thus

$$\mathsf{xnull}_L(\mathcal{H}_{\vec{u},\vec{v},s,t}) = \mathsf{xnull}_R(\mathcal{X}, \widehat{\mathcal{K}}_{\vec{u},s}).$$

Notice next that, for $0 \le a \le tk$, if $\mathsf{xnull}_R(\mathcal{K}_{\vec{v},t}) = q^a$ then $\mathsf{xnull}_L(\mathcal{K}_{\vec{v},t}) = q^{(N-tk)+a}$ by Lemma B.1, so that $\mathcal{X}$ has dimension $N - tk + a$ and it follows by Lemma C.3 that

$$\mathsf{E}[\mathsf{xnull}_R(\mathcal{X}, \widehat{\mathcal{K}}_{\vec{u},s})] \le 1 + f(h,k)\cdot q^{(s-t)k+a}.$$

Consequently

$$\begin{aligned}&\mathsf{E}[\mathsf{xnull}_L(\mathcal{H}_{\vec{u},\vec{v},s,t})]\\ &\quad\le \sum_{a=0}^{tk} \mathsf{Pr}[\mathsf{xnull}_R(\mathcal{K}_{\vec{v},t}) = q^a]\cdot\left(1 + f(h,k)\cdot q^{(s-t)k+a}\right)\\ &\quad= \sum_{a=0}^{tk} \mathsf{Pr}[\mathsf{xnull}_R(\mathcal{K}_{\vec{v},t}) = q^a]\\ &\qquad + f(h,k)\cdot q^{(s-t)k}\cdot \sum_{a=0}^{tk} \mathsf{Pr}[\mathsf{xnull}_R(\mathcal{K}_{\vec{v},t}) = q^a]\cdot q^a\\ &\quad= 1 + f(h,k)\cdot q^{(s-t)k}\cdot \mathsf{E}[\mathsf{xnull}_R(\mathcal{K}_{\vec{v},t})]\\ &\quad\le 1 + f(h,k)\cdot q^{(s-t)k}\cdot\left(1 + f(h,k)\cdot q^{(t+1)k+1-r}\right)\\ &\qquad\qquad\text{(by Lemma C.7)}\\ &\quad= 1 + q^{(s-t)k}\left(f(h,k) + f(h,k)^2\cdot q^{(t+1)k+1-r}\right)\\ &\quad\le 1 + q^{(s-t)k}\cdot\left(f(h,k) + q\cdot f(h,k)^2\right)\\ &\qquad\qquad\text{(since } t \le \lfloor r/k\rfloor - 1).\end{aligned}$$

In particular, it follows from the above that if $1 \le s < \lfloor r/k \rfloor - \Delta_{n,k}$ then

$$\begin{aligned}&\mathsf{E}[\mathsf{xnull}_L(\mathcal{H}_{\vec{u},\vec{v}s,s+\Delta_{n,k}})]\\ &\quad\le 1 + q^{-\Delta_{n,k}\cdot k}\cdot\left(f(h,k) + q\cdot f(h,k)^2\right)\\ &\quad\le 1 + q^{(2-\Delta_{n,k})\cdot k}\cdot\left(f(h,k) + f(h,k)\cdot q^{1-k}\right)\end{aligned}$$

as claimed.

Suppose, next, that $t < s < \lfloor r/k \rfloor$ instead, and notice that, for $z \in \mathsf{F}_q^{tk\times 1}$,

$$\begin{aligned}\mathcal{H}_{\vec{u},\vec{v},r,s}\cdot z = 0 &\iff \widehat{\mathcal{K}}_{\vec{u},s}^T\cdot(\mathcal{K}_{\vec{v},t}\cdot z) = 0\\ &\iff \mathcal{K}_{\vec{v},t}\cdot z \in \mathcal{X}\end{aligned}$$

where $\mathcal{X}$ is the subspace of $\mathsf{F}_q^{n\times 1}$ consisting of vectors $y \in \mathsf{F}_q^{n\times 1}$ such that $y^T \cdot \widehat{\mathcal{K}}_{\vec{v},s} = 0$. Thus

$$\mathsf{xnull}_R(\mathcal{H}_{\vec{u},\vec{v},s,t}) = \mathsf{xnull}_R(\mathcal{X}, \mathcal{K}_{\vec{v},t}).$$

Notice next that, for $0 \le a \le sk$, if $\mathsf{xnull}_R(\widehat{\mathcal{K}}_{\vec{u},s}) = q^a$ then $\mathsf{xnull}_L(\widehat{\mathcal{K}}_{\vec{u},s}) = q^{(N-sk)+a}$ by Lemma B.1, so that $\mathcal{X}$ has dimension $N - sk + a$ and it follows by Lemma C.7 that

$$\mathsf{E}[\mathsf{xnull}_R(\mathcal{X}, \mathcal{K}_{\vec{v},t})] \le 1 + f(h,k)\cdot q^{(t-s+1)k+N-r+1+a}.$$

Consequently

$$\begin{aligned}
&\mathsf{E}[\mathsf{xnull}_R(\mathcal{H}_{\vec{u},\vec{v},s,t})] \\
&\quad\le \sum_{a=0}^{sk} \mathsf{Pr}[\mathsf{xnull}_R(\widehat{\mathcal{K}}_{\vec{u},s}) = q^a] \\
&\qquad \cdot \left(1 + f(h,k)\cdot q^{(t-s+1)k+N-r+1+a}\right) \\
&\quad= \sum_{a=0}^{sk} \mathsf{Pr}[\mathsf{xnull}_R(\widehat{\mathcal{K}}_{\vec{u},s}) = q^a] \\
&\qquad + f(h,k)\cdot q^{(t-s+1)k+N-r+1}\cdot \\
&\qquad\quad \sum_{a=0}^{sk} \mathsf{Pr}[\mathsf{xnull}_R(\widehat{\mathcal{K}}_{\vec{u},s}) = q^a]\cdot q^a \\
&\quad= 1 + f(h,k)\cdot q^{(t-s+1)k+N-r+1}\cdot \mathsf{E}[\mathsf{xnull}_R(\widehat{\mathcal{K}}_{\vec{u},s})] \\
&\quad\le 1 + f(h,k)\cdot q^{(t-s+2)k}\cdot \mathsf{E}[\mathsf{xnull}_R(\widehat{\mathcal{K}}_{\vec{u},s})] \\
&\qquad\qquad \text{(since } N - r \le h \le k-1\text{)} \\
&\quad\le 1 + f(h,k)\cdot q^{(t-s+2)k}\cdot \left(1 + f(h,k)\cdot q^{sk-r}\right) \\
&\qquad\qquad \text{(by Lemma C.1, since } r \le N\text{)} \\
&\quad\le 1 + f(h,k)\cdot q^{(t-s+2)k}\cdot \left(1 + f(h,k)\cdot q^{-k}\right) \\
&\qquad\qquad \text{(since } s \le \lfloor r/k \rfloor - 1\text{)} \\
&\quad= 1 + q^{(t-s+2)k}\cdot \left(f(h,k) + f(h,k)^2\cdot q^{-k}\right).
\end{aligned}$$

In particular, when $s = t + \Delta_{n,k} < \lfloor n/k \rfloor$,

$$\begin{aligned}
&\mathsf{E}[\mathsf{xnull}_R(\mathcal{H}_{\vec{u},\vec{v},t+\Delta_{n,k},t})] \\
&\qquad \le 1 + q^{(2-\Delta_{n,k})k}\cdot \left(f(h,k) + f(h,k)^2\cdot q^{-k}\right),
\end{aligned}$$

which suffices to establish the claim.

D. PROOFS OF RESULTS IN SECTION 6

Suppose once again that $A \in \mathsf{F}_q^{n\times n}$ has Frobenius normal form $C_{f_1,f_2,\ldots,f_\ell}$ and that at least k of the invariant factors of A are nontrivial. By the definition of "Frobenius normal form," there exists a nonsingular matrix $Y \in \mathsf{F}_q^{n\times n}$ such that

$$A = Y^{-1}\cdot \begin{bmatrix} C_{f_1} & & & 0 \\ & C_{f_2} & & \\ & & \ddots & \\ 0 & & & C_{f_\ell} \end{bmatrix} \cdot Y. \tag{42}$$

Now, as in Section 6, let h be an integer $1 \le h \le k-1$ and let

$$N = \sum_{a=1}^{h} \deg(f_a),$$

the sum of the first h invariant factors of A. Let

$$A_1 = A\cdot Y^{-1}\cdot \begin{bmatrix} I_N & 0 \\ 0 & 0 \end{bmatrix}\cdot Y \tag{43}$$

and let

$$A_2 = A\cdot Y^{-1}\cdot \begin{bmatrix} 0 & 0 \\ 0 & I_{n-N} \end{bmatrix}\cdot Y. \tag{44}$$

Then it is easily checked that, since

$$A = Y^{-1}\cdot \begin{bmatrix} C_{f_1,f_2,\ldots,f_h} & 0 \\ 0 & C_{f_{h+1},f_{h+2},\ldots,f_\ell} \end{bmatrix}\cdot Y,$$

$$A_1 = Y^{-1}\cdot \begin{bmatrix} C_{f_1,f_2,\ldots,f_h} & 0 \\ 0 & 0_{n-N} \end{bmatrix}\cdot Y,$$

and

$$A_2 = Y^{-1}\cdot \begin{bmatrix} 0_N & 0 \\ 0 & C_{f_{h+1},f_{h+2},\ldots,f_\ell} \end{bmatrix}\cdot Y,$$

so that $A = A_1 + A_2$, $A_1\cdot A_2 = A_2\cdot A_1 = 0$, and A_1 has h nontrivial invariant factors — namely, $f_1, f_2, \ldots, f_h$.

Suppose now that vectors $u_1, u_2, \ldots, u_k, w, w_2, w_3, \ldots, w_k$ are chosen uniformly and independently from $\mathsf{F}_q^{n\times 1}$. Then, for $1 \le a \le k$,

$$u_a = Y^T\cdot \begin{bmatrix} \widehat{u}_a \\ \widetilde{u}_a \end{bmatrix}$$

where $\widehat{u}_a \in \mathsf{F}_q^{N\times 1}$ and $\widetilde{u}_a \in \mathsf{F}_q^{(n-N)\times 1}$, and

$$w = Y^{-1}\cdot \begin{bmatrix} \widehat{w} \\ \widetilde{w} \end{bmatrix} \quad\text{and}\quad w_b = Y^{-1}\cdot \begin{bmatrix} \widehat{w}_b \\ \widetilde{w}_a \end{bmatrix}$$

for $2 \le b \le k$, where

$$\widehat{w}, \widehat{w}_2, \widehat{w}_3, \ldots, \widehat{w}_k \in \mathsf{F}_q^{N\times 1}$$

and

$$\widetilde{w}, \widetilde{w}_2, \widetilde{w}_3, \ldots, \widetilde{w}_k \in \mathsf{F}_q^{(n-N)\times 1}.$$

Furthermore, since Y (and Y^T) is nonsingular, it is not difficult to argue that the vectors

$$\begin{aligned}
&\widehat{u}_1, \widehat{u}_2, \ldots, \widehat{u}_k, \widetilde{u}_1, \widetilde{u}_2, \ldots, \widetilde{u}_k, \\
&\qquad\qquad \widehat{w}, \widehat{w}_2, \widehat{w}_3, \ldots, \widehat{w}_k, \widetilde{w}, \widetilde{w}_2, \widetilde{w}_3, \ldots, \widetilde{w}_k
\end{aligned}$$

are uniformly and *independently* selected from their respective vector spaces.

Suppose next that we set

$$u_{a,1} = Y^T\cdot \begin{bmatrix} \widehat{u}_a \\ 0 \end{bmatrix} \quad\text{and}\quad u_{a,2} = Y^T\cdot \begin{bmatrix} 0 \\ \widetilde{u}_a \end{bmatrix}$$

for $1 \le a \le k$, so that $u_a = u_{a,1} + u_{a,2}$, and we set

$$w_{0,1} = Y^{-1}\cdot \begin{bmatrix} \widehat{w} \\ 0 \end{bmatrix}, \qquad w_{0,2} = Y^{-1}\cdot \begin{bmatrix} 0 \\ \widetilde{w} \end{bmatrix},$$

and

$$w_{a,1} = Y^{-1}\cdot \begin{bmatrix} \widehat{w}_a \\ 0 \end{bmatrix} \quad\text{and}\quad w_{a,2} = Y^{-1}\cdot \begin{bmatrix} 0 \\ \widetilde{w}_a \end{bmatrix}$$

for $2 \le a \le k$, so that $w = w_{0,1} + w_{0,2}$ and $w_a = w_{a,1} + w_{a,2}$ for $2 \le a \le k$ as well. Note evaluation (using the above expressions) suffices to confirm that

$$u_{a,1}^T\cdot w_{0,2} = 0 \quad\text{and}\quad u_{a,1}^T\cdot w_{c,2} = 0$$

for $1 \le a \le k$ and $2 \le c \le k$, and that

$$u_{a,2}^T\cdot w_{0,1} = 0 \quad\text{and}\quad u_{a,2}^T\cdot w_{c,1} = 0$$

for $1 \le a \le k$ and $2 \le c \le k$ as well. It follows that

$$u_a^T\cdot w = u_{a,1}^T\cdot w_{0,1} + u_{a,2}^T\cdot w_{0,2}$$

and that

$$u_a^T\cdot w_c = u_{a,1}^T\cdot w_{c,1} + u_{a,2}^T\cdot w_{c,2}$$

for $1 \le a \le k$ and $2 \le c \le k$.

Next consider the vector $y_0 = A\cdot w$. As above, note that

$$y_0 = Y^{-1}\cdot \begin{bmatrix} \widehat{y}_0 \\ \widetilde{y}_0 \end{bmatrix}$$

for $\widehat{y}_0 \in \mathsf{F}_q^{N\times 1}$ and $\widetilde{y}_0 \in \mathsf{F}_q^{(n-N)\times 1}$ once again. As above, set

$$y_{0,1} = Y^{-1}\cdot\begin{bmatrix}\widehat{y}_0\\0\end{bmatrix} \quad \text{and} \quad y_{0,2} = Y^{-1}\cdot\begin{bmatrix}0\\\widetilde{y}_0\end{bmatrix};$$

then, as above,

$$u_{a,1}^T\cdot y_{0,2} = u_{a,2}^T\cdot y_{0,1} = 0$$

so that

$$u_a^T\cdot y_0 = u_{a,1}^T\cdot y_{0,1} + u_{a,2}^T\cdot y_{0,2}.$$

It is now reasonably easy to establish that

$$u_{a,1}^T\cdot A_2 = u_{a,2}^T\cdot A_1 = 0$$

for $1\le a\le k$, that

$$A_2\cdot y_{0,1} = A_1\cdot y_{0,2} = 0,$$

and that

$$A_2\cdot w_{a,1} = A_1\cdot w_{a,2} = 0$$

for $2\le a\le k$ as well. Since $A_1\cdot A_2 = A_2\cdot A_1 = 0$, it now follows that if $y_1 = A\cdot y_0$ and $v_a = A\cdot w_a$ for $2\le a\le k$, then

$$y_1 = A_1\cdot y_{0,1} + A_2\cdot y_{0,2} \quad \text{and} \quad v_a = A_1\cdot w_{a,1} + A_2\cdot w_{a,2}$$

for $2\le a\le k$. Indeed, if we set $y_{1,1} = A_1\cdot y_{0,1}$, $y_{1,2} = A_2\cdot y_{0,2}$, $v_{a,1} = A_1\cdot w_{a,1}$ and $v_{a,2} = A_2\cdot w_{a,2}$ for $2\le a\le k$, then it can be established from the above that

$$A^s\cdot y_1 = A_1^s\cdot y_{1,1} + A_2^s\cdot y_{1,2} \quad \text{and} \quad A^s\cdot v_a = A_1^s\cdot v_{a,1} + A_2^s\cdot v_{a,2}$$

for $2\le a\le k$ and for every integer $s\ge 0$. Finally,

$$u_a^T\cdot A^s\cdot y_0 = u_{a,1}\cdot A_1^s\cdot y_{0,1} + u_{a,2}\cdot A_2^s\cdot y_{0,2}$$

and

$$u_a^T\cdot A^s\cdot v_c = u_{a,1}^T\cdot A_1^s\cdot v_{c,1} + u_{a,2}^T\cdot A_2^s\cdot v_{c,2}$$

for $1\le a\le k$, $2\le c\le k$, and for every integer $s\ge 0$ as well.

With all that noted, set

$$\vec{u} = u_1, u_2, \ldots, u_k \tag{45}$$

as usual and set

$$\vec{u}_1 = u_{1,1}, u_{2,1}, \ldots, u_{k,1} \text{ and } \vec{u}_2 = u_{1,2}, u_{2,2}\ldots, u_{k,2}. \tag{46}$$

Set

$$\vec{c} = y_1, v_2, v_3, \ldots, v_k, \tag{47}$$

$$\vec{c}_1 = y_{1,1}, v_{2,1}, v_{3,1}, \ldots, v_{k,1} \tag{48}$$

and

$$\vec{c}_2 = y_{1,2}, v_{2,2}, v_{3,2}, \ldots, v_{k,2}. \tag{49}$$

It now follows from the above that

$$\mathcal{H}_{A,\vec{u},\vec{c},s,t} = \mathcal{H}_{A_1,\vec{u}_1,\vec{c}_1,s,t} + \mathcal{H}_{A_2,\vec{u}_2,\vec{c}_2,s,t} \tag{50}$$

for all positive integers s and t.

Next consider the given vector b. Set $b_1 = A_1\cdot b$ and $b_2 = A_2\cdot b$, so that $A\cdot b = b_1 + b_2$. Define another three sequences of vectors, each of length k:

$$\vec{d} = A\cdot b, 0, 0, \ldots, 0, \tag{51}$$

$$\vec{d}_1 = b_1, 0, 0, \ldots, 0 \text{ and } \vec{d}_2 = b_2, 0, 0\ldots, 0. \tag{52}$$

Note that if $w_1 = A\cdot w + b$, $v_1 = A\cdot w_1$, $v_{1,1} = A_1\cdot w_1$ and $v_{1,2} = A_2\cdot w_1$ (so that $v_1 = v_{1,1} + v_{1,2}$), and if we set

$$\vec{v} = v_1, v_2, \ldots, v_k \tag{53}$$

as usual, and

$$\vec{v}_1 = v_{1,1}, v_{2,1}, \ldots, v_{k,1} \text{ and } \vec{v}_2 = v_{1,2}, v_{2,2}, \ldots, v_{k,2}, \tag{54}$$

then

$$\mathcal{H}_{A,\vec{u},\vec{v},s,t} = \mathcal{H}_{A,\vec{u},\vec{c},s,t} + \mathcal{H}_{A,\vec{u},\vec{d},s,t} \tag{55}$$

for all positive integers s and t. Recalling the equation at line (18), and the definitions of the sequences $\vec{y}$ and $\vec{z}$ of vectors immediately preceding it, note as well that

$$\mathcal{H}_{A_1,\vec{u},\vec{y},s,t} = \mathcal{H}_{A_1,\vec{u}_1,\vec{c}_1,s,t} + \mathcal{H}_{A_1,\vec{u}_1,\vec{d}_1,s,t}$$

and

$$\mathcal{H}_{A_2,\vec{u},\vec{z},s,t} = \mathcal{H}_{A_2,\vec{u}_2,\vec{c}_2,s,t} + \mathcal{H}_{A_2,\vec{u}_2,\vec{d}_2,s,t}.$$

Proof of Lemma 6.1. One can show (with a reasonably straightforward argument) that r_h is the rank of the matrix $A_1\in\mathsf{F}_q^{n\times n}$ that has been described above, and that A_1 has h nontrivial invariant factors. Since $h<k$, and the matrix A_1 has only h nontrivial invariant factors, if follows by Lemma 5.1 (using $b=0$ in this case) that if a is an integer such that $1\le a\le \lfloor r_h/k\rfloor - \Delta_{n,k} - 1$, vectors $u_1, u_2, \ldots, u_k, w, w_2, w_3, \ldots, w_k$ are chosen uniformly and independently from $\mathsf{F}_q^{n\times 1}$, $y_{0,1} = A_1\cdot w$, $y_{1,1} = A_1\cdot y_0$, $v_{a,1} = A_1\cdot w_a$ for $2\le a\le k$, and if $\vec{u}$ and $\vec{c}_1$ are the sequences of vectors shown at lines (45) and (48), above, then

$$\begin{aligned}&\mathsf{E}[\mathsf{xnull}_L(\mathcal{H}_{A_1,\vec{u},\vec{c}_1,a,a+\Delta_{n,k}})]\\&\qquad\le 1 + q^{(2-\Delta_{n,k})k}\cdot(f(h,k) + f(h,k)^2\cdot q^{1-k})\end{aligned}$$

and that

$$\begin{aligned}&\mathsf{E}[\mathsf{xnull}_R(\mathcal{H}_{A_1,\vec{u},\vec{c}_1,a+\Delta_{n,k},a})]\\&\qquad\le 1 + q^{(2-\Delta_{n,k})k}\cdot(f(h,k) + f(h,k)^2\cdot q^{1-k})\end{aligned}$$

as well, for $f(h,k)$ as given in Lemma 5.1.

Note next, that, since $u_a = u_{a,1} + u_{a,2}$ for $1\le a\le k$, if the sequences $\vec{u}_1$ and $\vec{u}_2$ are as shown at line (46), then

$$\begin{aligned}\mathcal{H}_{A_1,\vec{u},\vec{c}_1,s,t} &= \mathcal{H}_{A_1,\vec{u}_1,\vec{c}_1,s,t} + \mathcal{H}_{A_2\vec{u}_2,\vec{c}_1,s,t}\\&= \mathcal{H}_{A_1,\vec{u},1\vec{c}_1,s,t}\end{aligned}$$

since $u_{a,2}^T\cdot A_1^s\cdot y_{1,1} = u_{a,2}^T\cdot A_1^s\cdot v_{c,1}$ for $1\le a\le k$, $2\le c\le k$, and every integer $s\ge 0$. Consequently it follows from the above that

$$\begin{aligned}&\mathsf{E}[\mathsf{xnull}_L(\mathcal{H}_{A_1,\vec{u}_1,\vec{c}_1,a,a+\Delta_{n,k}})]\\&\qquad\le 1 + q^{(2-\Delta_{n,k})k}\cdot(f(h,k) + f(h,k)^2\cdot q^{1-k})\end{aligned}$$

and that

$$\begin{aligned}&\mathsf{E}[\mathsf{xnull}_R(\mathcal{H}_{A_1,\vec{u}_1,\vec{c}_1,a+\Delta_{n,k},a})]\\&\qquad\le 1 + q^{(2-\Delta_{n,k})k}\cdot(f(h,k) + f(h,k)^2\cdot q^{1-k})\end{aligned}$$

as well.

Next recall that if $u_1, u_2, \ldots, u_k, w, w_2, w_3, \ldots, w_k$ are chosen uniformly and independently from $\mathsf{F}_q^{n\times 1}$ and the vectors in the sequences $\vec{u}_1$, $\vec{u}_2$, $\vec{c}_1$ and $\vec{c}_2$ are defined as described above, then the vectors in the latter four sequences are chosen independently as well —so that matrices $\mathcal{H}_{A_1,\vec{u}_1,\vec{c}_1,s,t}$

and $\mathcal{H}_{A_2,\vec{u}_2,\vec{c}_2,s,t}$ are selected independently too. It now follows by an application of Lemma 4.2 and the equation at line (50), above, that if a is an integer such that $1 \leq \lfloor r_h/k \rfloor - \Delta_{n,k} - 1$ then

$$\mathsf{E}[\mathsf{xnull}_L(\mathcal{H}_{A,\vec{u},\vec{c},a,a+\Delta_{n,k}})] \leq 1 + q^{(2-\Delta_{n,k})k} \cdot (f(h,k) + f(h,k)^2 \cdot q^{1-k})$$

and

$$\mathsf{E}[\mathsf{xnull}_R(\mathcal{H}_{A,\vec{u},\vec{c},a+\Delta_{n,k},a})] \leq 1 + q^{(2-\Delta_{n,k})k} \cdot (f(h,k) + f(h,k)^2 \cdot q^{1-k}).$$

Finally, since the vector $\vec{d}$ shown at line (51) is fixed (it depends only on the inputs A and b), and application of Lemma B.3 and the equation at line (55), above, are sufficient to establish that if a is an integer such that $1 \leq a \leq \lfloor r_h/k \rfloor - \Delta_{n,k} - 1$ then

$$\mathsf{E}[\mathsf{xnull}_L(\mathcal{H}_{A,\vec{u},\vec{v},a,a+\Delta_{n,k}})] \leq 1 + q^{(2-\Delta_{n,k})k} \cdot (f(h,k) + f(h,k)^2 \cdot q^{1-k})$$

and that

$$\mathsf{E}[\mathsf{xnull}_R(\mathcal{H}_{A,\vec{u},\vec{v},a+\Delta_{n,k},a})] \leq 1 + q^{(2-\Delta_{n,k})k} \cdot (f(h,k) + f(h,k)^2 \cdot q^{1-k})$$

as well, as needed to establish the claim. $\square$