POLITECNICO DI TORINO
Repository ISTITUZIONALE

On the number of solutions to a random instance of the permuted kernel problem

Original
On the number of solutions to a random instance of the permuted kernel problem / Sanna, C.. - In: JOURNAL OF COMPLEXITY. - ISSN 0885-064X. - STAMPA. - 86:(2025), pp. 1-12. [10.1016/j.jco.2024.101898]

Availability:
This version is available at: 11583/2992744 since: 2024-09-24T19:40:34Z

Publisher:
Elsevier

Published
DOI:10.1016/j.jco.2024.101898

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright
Elsevier preprint/submitted version

Preprint (submitted version) of an article published in JOURNAL OF COMPLEXITY © 2025, http://doi.org/10.1016/j.jco.2024.101898

(Article begins on next page)

25 November 2026

On the number of solutions to a random instance of the permuted kernel problem*

Carlo Sanna

Department of Mathematical Sciences, Politecnico di Torino
Corso Duca degli Abruzzi 24, 10129 Torino, Italy

carlo.sanna@polito.it

arXiv:2406.00453v1 [math.CO] 1 Jun 2024

Abstract

The *Permuted Kernel Problem* (PKP) is a problem in linear algebra that was first introduced by Shamir in 1989. Roughly speaking, given an $\ell \times m$ matrix $\boldsymbol{A}$ and an $m \times 1$ vector $\boldsymbol{b}$ over a finite field of q elements $\mathbb{F}_q$, the PKP asks to find an $m \times m$ permutation matrix $\boldsymbol{\pi}$ such that $\boldsymbol{\pi b}$ belongs to the kernel of $\boldsymbol{A}$. In recent years, several post-quantum digital signature schemes whose security can be provably reduced to the hardness of solving random instances of the PKP have been proposed. In this regard, it is important to know the expected number of solutions to a random instance of the PKP in terms of the parameters q, ℓ, m. Previous works have heuristically estimated the expected number of solutions to be $m!/q^\ell$.

We provide, and rigorously prove, exact formulas for the expected number of solutions to a random instance of the PKP and the related *Inhomogeneous Permuted Kernel Problem* (IPKP), considering two natural ways of generating random instances.

Keywords: cryptography; digital signatures; NP-complete problem; permutations; permuted kernel problem; post-quantum cryptography.

MSC2020: 05A05, 05A16, 15A99, 11T71, 68Q25.

1 Introduction

The *Permuted Kernel Problem* (PKP) is a problem in linear algebra that was first introduced by Shamir in 1989 [23]. Roughly speaking, given an $\ell \times m$ matrix $\boldsymbol{A}$ and an $m \times 1$ vector $\boldsymbol{b}$ over a finite field of q elements $\mathbb{F}_q$, the PKP asks to find an $m \times m$ permutation matrix $\boldsymbol{\pi}$ such that $\boldsymbol{\pi b}$ belongs to the kernel of $\boldsymbol{A}$, that is, $\boldsymbol{A\pi b} = \boldsymbol{0}$. Using as security assumption the computational hardness of solving random instances $(\boldsymbol{A}, \boldsymbol{b})$ of the PKP, Shamir devised an identification scheme that has a very efficient implementation on low-cost smart cards.

In recent years, for several reasons, the PKP has become a very attractive problem to build post-quantum cryptographic schemes. First, the PKP is based on simple objects and operations, which can be implemented easily and efficiently. Second, the hardness of the PKP, and of some natural variants of the PKP, has been intensively studied [1, 10, 12, 13, 16, 17, 21, 22], and despite many efforts the best known algorithms have exponential complexities. Third, the PKP is known to be NP-complete in the strong sense [23], and quantum computers are expected to have a limited advantage in solving NP-complete problems (essentially, no more than the advantage of Grover's search) [2].

*C. Sanna is a member of GNSAGA of INdAM and of CrypTO, the group of Cryptography and Number Theory of the Politecnico di Torino. This work was partially supported by project SERICS (PE00000014) under the MUR National Recovery and Resilience Plan funded by the European Union – NextGenerationEU.

Consequently, several post-quantum digital signature schemes whose security can be provably reduced to the hardness of solving random instances of the PKP have been proposed, namely: PKP-DSS (2019) [5], SUSHSYFISH (2020) [4], Bidoux and Gaborit's (2023) [6], and PERK (2024) [3]. In particular, PERK has been submitted to the NIST additional call for the post-quantum cryptography standardization process [15].

In all these schemes, the public key is a random instance of the PKP (or a variant thereof), the secret key is a solution to such an instance, and the signing and verification algorithms constitute a non-interactive zero-knowledge proof of knowledge of the solution. Furthermore, the parameters q, ℓ, m are selected in order to ensure that the known algorithms to solve the PKP require on average 2^λ or more operations, where λ is the desired security parameter.

In addition, for mainly two reasons, it is desiderable to choose q, ℓ, m so that the random instance of the PKP is likely to have exactly one solution (or at least a number of solutions that is bounded by a known constant). First, because it is natural to have exactly one secret key corresponding to the public key. Second, because the complexity of an algorithm searching a solution to the PKP is expected to be (approximately) inversely proportional to the number of solutions.

Shamir [23] stated that if $q^\ell \approx m!$ then a random instance of the PKP is likely to have a unique solution. This claim (or natural generalizations of it) is repeated in most of the subsequent works [7, 12, 13, 14, 16, 17, 19]. The reasoning behind it is the following. Assuming that $\ell \leq m$ and that $\boldsymbol{A}$ has rank equal to ℓ, the probability that a random $m \times 1$ vector $\boldsymbol{c}$ over $\mathbb{F}_q$, which is taken with uniform distribution and independently from $\boldsymbol{A}$, belongs to the kernel of $\boldsymbol{A}$ is equal to $q^{-\ell}$. Hence, assuming that, for a uniformly distributed random $m \times m$ permutation matrix $\boldsymbol{\pi}$, the vector $\boldsymbol{\pi b}$ behaves like the random vector $\boldsymbol{c}$, by the linearity of the expectation we get that the expected number of solutions to the PKP instance $(\boldsymbol{A}, \boldsymbol{b})$ is equal to $m!/q^\ell$.

Note that the previous reasoning is only a heuristic and not a rigorous mathematical proof. The main issue is that, unlike $\boldsymbol{c}$ and $\boldsymbol{A}$, the random variables $\boldsymbol{\pi b}$ and $\boldsymbol{A}$ are not independent. In fact, the probability of the event $\boldsymbol{A\pi b} = \boldsymbol{0}$ depends on how the random instance $(\boldsymbol{A}, \boldsymbol{b})$ is generated.

Moreover, for real-world choices of the parameters q, ℓ, m, the heuristic formula $m!/q^\ell$ cannot be empirically tested. In fact, since the parameters are chosen to make finding the solutions difficult, one cannot efficiently count the solutions to test the formula empirically. (Note that, since PKP is NP-complete, its search version reduces to its decision version, which in turn reduces to its counting version. Therefore, there should be no shortcuts to count the solutions much more efficiently than by actually finding them.)

In light of the previous considerations, and due to the importance of building post-quantum cryptography on solid mathematical foundations, the purpose of this paper is to provide, and rigorously prove, exact formulas for the expected number of solution to random instances of the PKP. More precisely, we consider both the PKP and the *Inhomogeneous Permuted Kernel Problem* (IPKP), and for each of these problems we study two natural ways of generating random instances.

The structure of the paper is as follows. In Section 2, we state the main notation. In Section 3, we provide the definition of the IPKP and we introduce the algorithms to generate random instances of the IPKP and the PKP. In Section 4, we state the main results of the paper, that is, the exact formulas for the expected number of solutions to the random instances of the IPKP and the PKP generated by the algorithms. Moreover, we compare these exact formulas with the heuristic formula. In Section 5, we give the proofs of the main results. Finally, in Section 6, we state some concluding remarks and some possible questions for future research.

2 Notation

Hereafter, let q be a prime power, let $\mathbb{F}_q$ be a finite field of q elements, and let $\mathbb{F}_q^*$ be the multiplicative group of $\mathbb{F}_q$. For all integers $m, n, r \geq 0$, let $\mathbb{F}_q^{m\times n}$ be the vector space of $m\times n$ matrices over $\mathbb{F}_q$, let $\mathbb{F}_q^{m\times n,r}$ be the subset of $\mathbb{F}_q^{m\times n}$ containing the matrices of rank equal to r, and let $\mathbb{F}_q^{\star m\times n,r}$ be the subset of $\mathbb{F}_q^{m\times n,r}$ containing the matrices having m pairwise distinct nonzero rows. Let $\mathbf{0}$ and $\boldsymbol{I}$ be the zero matrix and the identity matrix, respectively, where the sizes will be always clear from the context. For every matrix $\boldsymbol{A} \in \mathbb{F}_q^{m\times n}$, let $\mathrm{rk}(\boldsymbol{A})$ be the rank of $\boldsymbol{A}$, let $\boldsymbol{A}^\intercal$ be the transpose of $\boldsymbol{A}$, let $\text{left-ker}(\boldsymbol{A}) := \{\boldsymbol{x} \in \mathbb{F}_q^{1\times m} : \boldsymbol{x}\boldsymbol{A} = \mathbf{0}\}$ be the left kernel of $\boldsymbol{A}$, and for every positive integer k let $\big(\text{left-ker}(\boldsymbol{A})\big)^k$ the set of $k \times m$ matrices with each row in $\text{left-ker}(\boldsymbol{A})$.

Let $\mathbb{S}_n$ be the symmetric group of $\{1, \ldots, n\}$. For each permutation $\pi \in \mathbb{S}_n$, let $\boldsymbol{\pi}$ be the corresponding $n \times n$ permutation matrix whose entry of the ith row and jth column is equal to the Kronecker symbol $\delta_{\pi(i),j}$. For each positive integer $k \leq n$, let $\mathbb{S}_{n,k}$ be the subset of permutations in $\mathbb{S}_n$ that can be written as the product of exactly k disjoint cycles.

Let $s \leftarrow \mathcal{S}$ denote that s is taken at random with uniform distribution from the finite set $\mathcal{S}$. Let $y \leftarrow \mathsf{A}(x)$ denote running the (possibly probabilistic) algorithm A on input x and assigning the output to y.

Let $X \overset{\mathsf{d}}{=} Y$ mean that the random variables X and Y have the same distribution, let $\mathbb{P}[E]$ be the probability that the event E happens, let $\mathbb{E}[X]$ be the expected value of the random variable X, and let $\mathbb{1}[S]$ be equal to 1, respectively 0, if the statement S is true, respectively false.

Let $\varphi(\cdot)$ be the Euler function, and let $\lfloor\cdot\rfloor$ be the floor function.

3 Inhomogeneous permuted kernel problem

The *Inhomogeneous Permuted Kernel Problem* (IPKP) is defined as follows (cf. [22]).

Problem 3.1 (Inhomogeneous Permuted Kernel Problem)**.**

- Parameters: a prime power q and positive integers ℓ, m, n.
- Instance: a triple of matrices $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}) \in \mathbb{F}_q^{\ell\times m} \times \mathbb{F}_q^{m\times n} \times \mathbb{F}_q^{\ell\times n}$.
- Task: finding a permutation $\pi \in \mathbb{S}_m$ such that $\boldsymbol{A}\boldsymbol{\pi}\boldsymbol{B} = \boldsymbol{C}$.

If $\boldsymbol{C} = \mathbf{0}$ and $n = 1$, then the IPKP corresponds to the version of the PKP formulated in Section 1. In fact, the PKP is sometimes called *homogeneous* PKP. Moreover, for $n = 1$ the problem is called *monodimensional*, while for $n > 1$ it is called *multidimensional*. The heuristic formula for the expected number of solutions naturally generalizes to the multidimensional case, becoming $m!/q^{\ell n}$.

We let $N_{\mathsf{sol}}(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ and $N_{\mathsf{sol}}(\boldsymbol{A}, \boldsymbol{B})$ denote the number of solutions to the IPKP instance $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ and to the PKP instance $(\boldsymbol{A}, \boldsymbol{B})$, respectively.

In order to generate hard instances of the IPKP or the PKP, it is recommended to take $\mathrm{rk}(\boldsymbol{A}) = \ell$ and $\mathrm{rk}(\boldsymbol{B}) = n$ (see, e.g., [22, Section V.A]). Note that these conditions and the existence of a solution imply that $\max(\ell, n) \leq m$ for the IPKP, and $\ell + n \leq m$ for the PKP. Hence, a simple algorithm to generate a hard random instance of the IPKP, respectively the PKP, together with a solution of such an instance is provided by $\mathsf{GenIPKP}$ (Figure 1), respectively GenPKP (Figure 2). Indeed, the algorithm used in the implementation of PERK [18] is equivalent to $\mathsf{GenIPKP}$, with the only difference that $\boldsymbol{A}$ is sampled from $\mathbb{F}_q^{\ell\times m}$ instead of $\mathbb{F}_q^{\ell\times m,\ell}$.

Furthermore, other recommended conditions for the hardness of the instance are that $\boldsymbol{A}$ has pairwise distinct columns and that $\boldsymbol{B}$ has pairwise distinct rows [22, Section V.A]. We consider only the second condition. We let $\mathsf{GenIPKP}^\star$ and $\mathsf{GenPKP}^\star$ be defined as $\mathsf{GenIPKP}$ and GenPKP but with $\boldsymbol{B}$ taken from $\mathbb{F}_q^{\star m\times n,n}$. Note that $\mathbb{F}_q^{\star m\times n,n}$ is nonempty if and only if $n \leq m < q^n$.

$\mathsf{GenIPKP}(q, \ell, m, n)$

1 : $(\boldsymbol{A}, \pi, \boldsymbol{B}) \leftarrow \mathbb{F}_q^{\ell \times m, \ell} \times \mathbb{S}_m \times \mathbb{F}_q^{m \times n, n}$
2 : $\boldsymbol{C} \leftarrow \boldsymbol{A}\pi\boldsymbol{B}$
3 : **return** $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \pi)$

Figure 1: A generator of random instances of the IPKP, assuming that $\max(\ell, n) \leq m$.

$\mathsf{GenPKP}(q, \ell, m, n)$

1 : $(\pi, \boldsymbol{B}) \leftarrow \mathbb{S}_m \times \mathbb{F}_q^{m \times n, n}$
2 : $\boldsymbol{A} \leftarrow \big(\text{left-ker}(\boldsymbol{\pi B})\big)^{\ell} \cap \mathbb{F}_q^{\ell \times m, \ell}$
3 : **return** $(\boldsymbol{A}, \boldsymbol{B}, \pi)$

Figure 2: A generator of random instances of the PKP, assuming that $\ell + n \leq m$.

4 Expected number of solutions

Our first result is a formula for the expected number of solutions to a random instance of the IPKP generated by $\mathsf{GenIPKP}$.

Theorem 4.1. *Let ℓ, m, n be positive integers with $\max(\ell, n) \leq m$, and let $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \pi) \leftarrow \mathsf{GenIPKP}(q, \ell, m, n)$. Then*

$$\mathbb{E}\big[N_{\mathsf{sol}}(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})\big] = \sum_{k=1}^{m} |\mathbb{S}_{m,k}| \sum_{r=0}^{\min(\ell, m-k)} \frac{|\mathbb{F}_q^{\ell \times (m-k), r}|\,|\mathbb{F}_q^{(\ell-r)\times k, \ell-r}|\,q^{kr}}{|\mathbb{F}_q^{\ell \times m, \ell}|} \prod_{i=0}^{n-1} \frac{q^{m-r} - q^i}{q^m - q^i} \tag{1}$$

Regarding (1), note that $|\mathbb{S}_{m,k}|$ is a *Stirling number of the first kind*, which can be computed recursively (see, e.g., [11, Eq. 6.8]), while the cardinality of the set of matrices of a prescribed rank is given by an explicit formula (see Lemma 5.1).

Theorem 4.1 shows that, when random instances of the IPKP are generated by $\mathsf{GenIPKP}$, the heuristic formula for the expected number of solutions can be very far off from the true value. For example, let $q = 1021$, $\ell = 35$, $m = 79$, and $n = 3$, which correspond to the first parameter set of PERK [3, Table 2]. On the one hand, the heuristic formula predicts that the expected number of solutions (not counting π, which is a solution by construction) is equal to $m!/q^{\ell n} \approx 10^{-199}$. On the other hand, Theorem 4.1 yields that the true value is about $2.89 \cdot 10^{-6}$ (again, not counting π). Thus the heuristic formula underestimates by a factor of 10^{-193} the number of solutions that are not equal to π. (However, note that $2.89 \cdot 10^{-6}$ is still irrelevant and has no impact on the parameter choice of PERK.)

Our second result is a formula for the expected value of the number of solutions to a random instance of the monodimensional IPKP generated by $\mathsf{GenIPKP}^{\star}$.

Theorem 4.2. *Let ℓ and m be positive integers with $\ell \leq m < q$, and let $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \pi) \leftarrow \mathsf{GenIPKP}^{\star}(q, \ell, m, 1)$. Then*

$$\mathbb{E}\big[N_{\mathsf{sol}}(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})\big] = 1 + \frac{(m! - 1)(q^{m-\ell} - 1)}{q^m - 1}. \tag{2}$$

In the right-hand side of (2), the additive term 1 corresponds to the solution π, while the second additive term corresponds to additional solutions. Furthermore, as $m - \ell \to +\infty$, the second term is asymptotic to $m!/q^{\ell}$, in agreement with the heuristic formula.

Our third result is a formula for the expected number of solutions to a random instance of the monodimensional PKP generated by GenPKP.

Theorem 4.3. *Let ℓ, m be integers with $0 < \ell < m$. If $(\boldsymbol{A}, \boldsymbol{B}, \pi) \leftarrow \mathsf{GenPKP}(q, \ell, m, 1)$ then*

$$\mathbb{E}[N_{\mathsf{sol}}(\boldsymbol{A}, \boldsymbol{B})] = \frac{m!(q^{m-\ell} - q)}{q^m - q} + \frac{m!(q^m - q^{m-\ell})}{(q^m - 1)(q^m - q)} \left(\sum_{d \,|\, q-1} \varphi(d) \binom{\lfloor (q+m-1)/d \rfloor}{\lfloor m/d \rfloor} - q + 1 \right). \tag{3}$$

Although the first additive term on the right-hand side of (3) is asymptotic to $m!/q^\ell$ as $m - \ell \to +\infty$, the second additive term is not negligible. Hence, the expected number of solutions can be significantly larger than that predicted by the heuristic formula. For instance, let $q = 251$, $\ell = 41$, and $m = 69$, which correspond to the parameter set for the security level $\lambda = 128$ of PKP-DSS [5, Table 1]. Then the heuristic formula predicts about $m!/q^\ell \approx 0.7$ solutions, while Theorem 4.3 yields that the true value is about 5412.

Our forth and final result is a formula for the expected number of solutions to a random instance of the monodimensional PKP generated by GenPKP^*.

Theorem 4.4. *Let ℓ and m be positive integers with $\ell < m < q$. If $(\boldsymbol{A}, \boldsymbol{B}, \pi) \leftarrow \mathsf{GenPKP}^*(q, \ell, m, 1)$ then*

$$\mathbb{E}[N_{\mathsf{sol}}(\boldsymbol{A}, \boldsymbol{B})] = \frac{m!(q^{m-\ell} - q)}{q^m - q} + \frac{q^m - q^{m-\ell}}{(q^m - q)\binom{q-1}{m}} \sum_{d \,|\, \gcd(q-1, m)} \varphi(d) \binom{(q-1)/d}{m/d}. \tag{4}$$

As $m - \ell \to +\infty$ (and consequently $q \to +\infty$), the first additive in the right-hand side of (4) is asymptotic to $m!/q^\ell$, while the second additive term is not exceeding 3, unless we have that $m = q - 1$ (see Lemma 5.6). Therefore, we can say that Theorem 4.4 is in agreement with the heuristic formula.

5 Proofs

5.1 Preliminaries

We collect in this section some preliminary lemmas needed later.

Lemma 5.1. *Let $m, n, r \geq 0$ be integers. Then we have that*

$$|\mathbb{F}_q^{m \times n, r}| = \prod_{i=0}^{r-1} \frac{(q^m - q^i)(q^n - q^i)}{q^r - q^i},$$

with the usual convention that the empty product is equal to 1.

Proof. See, e.g., [8]. □

Lemma 5.2. *Let $\ell, m, m_1, m_2 \geq 0$ be integers with $\ell \leq m = m_1 + m_2$. Let $\boldsymbol{A} \leftarrow \mathbb{F}_q^{\ell \times m, \ell}$ and write $\boldsymbol{A} = (\boldsymbol{A}_1 \mid \boldsymbol{A}_2)$ where $\boldsymbol{A}_1 \in \mathbb{F}_q^{\ell \times m_1}$ and $\boldsymbol{A}_2 \in \mathbb{F}_q^{\ell \times m_2}$. Then*

$$\mathbb{P}\big[\operatorname{rk}(\boldsymbol{A}_1) = r\big] = \frac{|\mathbb{F}_q^{\ell \times m_1, r}| \, |\mathbb{F}_q^{(\ell - r) \times m_2, \ell - r}| \, q^{m_2 r}}{|\mathbb{F}_q^{\ell \times m, \ell}|},$$

for every integer r with $0 \leq r \leq \min(\ell, m_1)$.

Proof. Let us count the number of $\boldsymbol{A} = (\boldsymbol{A}_1 \mid \boldsymbol{A}_2) \in \mathbb{F}_q^{\ell \times m, \ell}$ such that $\operatorname{rk}(\boldsymbol{A}_1) = r$. Clearly, there are $|\mathbb{F}_q^{\ell \times m_1, r}|$ choices for $\boldsymbol{A}_1$. We have to determine the number of possible choices for $\boldsymbol{A}_2$. Let C and C_1 be the columnspaces of $\boldsymbol{A}$ and $\boldsymbol{A}_1$, respectively. Hence, we have that $\dim(C) = \ell$, $\dim(C_1) = r$, and $C_1 \subseteq C$. Consequently, the quotient space C/C_1 can be identified with $\mathbb{F}_q^{(\ell - r) \times 1}$ via the choice of a basis. Let $\rho \colon C \to \mathbb{F}_q^{(\ell - r) \times 1}$ be the natural projection onto the quotient space, and let $\rho(\boldsymbol{A}_2)$ be the $(\ell - r) \times m_2$ matrix obtained by

applying ρ to each column of $\boldsymbol{A}_2$. Since the columns of $\boldsymbol{A}$ generates C, it follows easily that the columns of $\rho(\boldsymbol{A}_2)$ generates $\mathbb{F}_q^{(\ell-r)\times 1}$. Hence, we have that $\mathrm{rk}(\rho(\boldsymbol{A}_2)) = \ell - r$. Moreover, since the kernel of ρ has dimension r, for each choice of $\rho(\boldsymbol{A}_2) \in \mathbb{F}_q^{(\ell-r)\times m_2,\ell-r}$ there correspond $q^{m_2 r}$ choices of $\boldsymbol{A}_2 \in \mathbb{F}_q^{\ell\times m_2}$. Thus we get that there are $|\mathbb{F}_q^{(\ell-r)\times m_2,\ell-r}|\, q^{m_2 r}$ choices for $\boldsymbol{A}_2$. The main claim follows. □

Lemma 5.3. *Let $\ell, m, m', n, s \geq 0$ be integers such that $s \leq \min(m, m')$, $\ell \leq m$, and $n \leq m'$, and let $\boldsymbol{M} \in \mathbb{F}_q^{m\times m',s}$. If $(\boldsymbol{A}, \boldsymbol{B}) \leftarrow \mathbb{F}_q^{\ell\times m,\ell} \times \mathbb{F}_q^{m'\times n,n}$ then*

(i) $\mathbb{P}[\boldsymbol{AM} = \boldsymbol{0}] = \prod_{i=0}^{\ell-1} \frac{q^{m-s}-q^i}{q^m-q^i}$;

(ii) $\mathbb{P}[\boldsymbol{MB} = \boldsymbol{0}] = \prod_{i=0}^{n-1} \frac{q^{m'-s}-q^i}{q^{m'}-q^i}$;

(iii) $\mathbb{P}[\boldsymbol{AMB} = \boldsymbol{0}] = \sum_{r=0}^{\min(\ell,s)} \frac{|\mathbb{F}_q^{\ell\times s,r}|\,|\mathbb{F}_q^{(\ell-r)\times(m-s),\ell-r}|\,q^{(m-s)r}}{|\mathbb{F}_q^{\ell\times m,\ell}|} \prod_{i=0}^{n-1} \frac{q^{m'-r}-q^i}{q^{m'}-q^i}$.

Proof. First, note that (i) follows from (ii) since $\boldsymbol{AM} = \boldsymbol{0}$ is equivalent to $\boldsymbol{M}^\intercal \boldsymbol{A}^\intercal = \boldsymbol{0}$ and transposition does not change the rank.

Let us prove (ii). We have that $\boldsymbol{MB} = \boldsymbol{0}$ is equivalent to each column of $\boldsymbol{B}$ belonging to the kernel of $\boldsymbol{M}$, which has dimension $m' - s$. Thus there are $|\mathbb{F}_q^{(m'-s)\times n,n}|$ possible choices for $\boldsymbol{B}$. Hence, using Lemma 5.1, we get that

$$\mathbb{P}[\boldsymbol{MB} = \boldsymbol{0}] = \frac{|\mathbb{F}_q^{(m'-s)\times n,n}|}{|\mathbb{F}_q^{m'\times n,n}|} = \prod_{i=0}^{n-1} \frac{q^{m'-s}-q^i}{q^{m'}-q^i}$$

which proves (ii).

It remains to prove (iii). Since $\mathrm{rk}(\boldsymbol{M}) = s$, there exist $\boldsymbol{Q} \in \mathbb{F}_q^{m\times m,m}$ and $\boldsymbol{R} \in \mathbb{F}_q^{s\times m',s}$ such that

$$\boldsymbol{M} = \boldsymbol{Q}\begin{pmatrix}\boldsymbol{R}\\ \boldsymbol{0}\end{pmatrix}.$$

Let $\widetilde{\boldsymbol{A}} \leftarrow \mathbb{F}_q^{\ell\times m,\ell}$ and write $\widetilde{\boldsymbol{A}} = (\widetilde{\boldsymbol{A}}_1 \mid \widetilde{\boldsymbol{A}}_2)$ where $\widetilde{\boldsymbol{A}}_1 \in \mathbb{F}_q^{\ell\times s}$ and $\widetilde{\boldsymbol{A}}_2 \in \mathbb{F}_q^{\ell\times(m-s)}$. Since $\boldsymbol{AQ} \overset{\mathsf{d}}{=} \widetilde{\boldsymbol{A}}$ and $\mathrm{rk}(\widetilde{\boldsymbol{A}}_1\boldsymbol{R}) = \mathrm{rk}(\widetilde{\boldsymbol{A}}_1) \leq \min(\ell, s)$, we have that

$$\begin{aligned}\mathbb{P}[\boldsymbol{AMB} = \boldsymbol{0}] &= \mathbb{P}\Big[\boldsymbol{AQ}\begin{pmatrix}\boldsymbol{R}\\ \boldsymbol{0}\end{pmatrix}\boldsymbol{B} = \boldsymbol{0}\Big] = \mathbb{P}\Big[\widetilde{\boldsymbol{A}}\begin{pmatrix}\boldsymbol{R}\\ \boldsymbol{0}\end{pmatrix}\boldsymbol{B} = \boldsymbol{0}\Big] = \mathbb{P}[\widetilde{\boldsymbol{A}}_1\boldsymbol{RB} = \boldsymbol{0}]\\ &= \sum_{r=0}^{\min(\ell,s)} \sum_{\boldsymbol{N}\in\mathbb{F}_q^{\ell\times m',r}} \mathbb{P}[\widetilde{\boldsymbol{A}}_1\boldsymbol{R} = \boldsymbol{N}]\,\mathbb{P}[\boldsymbol{NB} = \boldsymbol{0}]. \qquad (5)\end{aligned}$$

By (ii), we have that

$$\mathbb{P}[\boldsymbol{NB} = \boldsymbol{0}] = \prod_{i=0}^{n-1} \frac{q^{m'-r}-q^i}{q^{m'}-q^i}, \qquad (6)$$

while, using again the fact that $\mathrm{rk}(\widetilde{\boldsymbol{A}}_1\boldsymbol{R}) = \mathrm{rk}(\widetilde{\boldsymbol{A}}_1)$ and employing Lemma 5.2, we get that

$$\begin{aligned}\sum_{\boldsymbol{N}\in\mathbb{F}_q^{\ell\times m',r}} \mathbb{P}[\widetilde{\boldsymbol{A}}_1\boldsymbol{R} = \boldsymbol{N}] &= \mathbb{P}[\mathrm{rk}(\widetilde{\boldsymbol{A}}_1\boldsymbol{R}) = r]\\ &= \mathbb{P}[\mathrm{rk}(\widetilde{\boldsymbol{A}}_1) = r] = \frac{|\mathbb{F}_q^{\ell\times s,r}|\,|\mathbb{F}_q^{(\ell-r)\times(m-s),\ell-r}|\,q^{(m-s)r}}{|\mathbb{F}_q^{\ell\times m,\ell}|}. \qquad (7)\end{aligned}$$

Thus, putting together (5), (6), and (7), we obtain (iii). □

For each permutation $\sigma \in \mathbb{S}_m$ and for each $\lambda \in \mathbb{F}_q^*$, let $E_{\sigma,\lambda}$ be the set of $\boldsymbol{x} \in \mathbb{F}_q^{m\times 1,1}$ such that $\boldsymbol{\sigma x} = \lambda \boldsymbol{x}$. Moreover, let $E_\sigma := \bigcup_{\lambda\in\mathbb{F}_q^*} E_{\sigma,\lambda}$. In other words, we have that $E_{\sigma,\lambda}$ is the set of eigenvectors of $\boldsymbol{\sigma}$ with eigenvalue λ, if λ is an eigenvalue of $\boldsymbol{\sigma}$, otherwise $E_{\sigma,\lambda} = \varnothing$. (Note that $\lambda = 0$ is not an eigenvalue of $\boldsymbol{\sigma}$, since $\boldsymbol{\sigma}$ is invertible.) Furthermore, let $E^\star_{\sigma,\lambda} := E_{\sigma,\lambda} \cap \mathbb{F}_q^{\star m\times 1,1}$ and $E^\star_\sigma := \bigcup_{\lambda\in\mathbb{F}_q^*} E^\star_{\sigma,\lambda}$.

Lemma 5.4. *For each positive integer m, we have that*

$$\sum_{\sigma\in\mathbb{S}_m} |E_\sigma| = m!\left(\sum_{d\,|\,q-1} \varphi(d)\binom{\lfloor (q+m-1)/d\rfloor}{\lfloor m/d\rfloor} - q + 1\right). \tag{8}$$

Proof. Let $\sigma \in \mathbb{S}_m$, $\lambda \in \mathbb{F}_q^*$, and $\boldsymbol{x} \in \mathbb{F}_q^{m\times 1,1}$. Write $\sigma = \gamma_1\cdots\gamma_t$ and $\boldsymbol{x} = (x_1,\dots,x_m)^\intercal$, where $\gamma_1,\dots,\gamma_t$ are disjoint cycles and $x_1,\dots,x_m \in \mathbb{F}_q$. By "$\boldsymbol{x}$ over γ_k" we mean the sequence (x_j) where j runs over the orbit of the elements moved by γ_k. Also, we let $|\gamma_k|$ denote the length of the cycle γ_k. It follows easily that $\boldsymbol{x} \in E_{\sigma,\lambda}$ if and only if, for each $k \in \{1,\dots,t\}$, we have that either $\lambda^{|\gamma_k|} = 1$ and $\boldsymbol{x}$ over γ_k is a geometric progression of ratio λ, or $\lambda^{|\gamma_k|} \neq 1$ and $\boldsymbol{x}$ over γ_k is a zero sequence. Hence, we have that

$$|E_{\sigma,\lambda}| = q^{|\{k\,:\,\lambda^{|\gamma_k|}=1\}|} - 1.$$

Therefore, since the $E_\sigma = \bigcup_{\lambda\in\mathbb{F}_q^*} E_{\sigma,\lambda}$ is a disjoint union, we get that

$$|E_\sigma| = \sum_{\lambda\,\in\,\mathbb{F}_q^*}\left(q^{|\{k\,:\,\lambda^{|\gamma_k|}=1\}|} - 1\right) = \sum_{d\,|\,q-1} \varphi(d)\, q^{|\{k\,:\,d \text{ divides } |\gamma_k|\}|} - q + 1, \tag{9}$$

where we used the facts that: $\lambda^{|\gamma_k|} = 1$ if and only if the multiplicative order of λ divides $|\gamma_k|$; and there are $\varphi(d)$ elements of order d in the multiplicative (cyclic) group $\mathbb{F}_q^*$.

At this point, we employ the theory of combinatorial classes and their generating functions [9, Part A, Chapters I and II]. Letting

$$F_d(z) := \log\left(\frac{1}{1-z}\right) - \frac{1}{d}\log\left(\frac{1}{1-z^d}\right) \quad \text{and} \quad G_d(z) := \frac{1}{d}\log\left(\frac{1}{1-z^d}\right),$$

we have that $F_d(z)$, respectively $G_d(z)$, is the exponential generating function of the combinatorial class of cycles with length not divisible, respectively divisible, by d. Hence, since each permutation can be uniquely written as a product of disjoint cycles, we have that

$$\begin{aligned}
\sum_{m=0}^{\infty}\left(\sum_{\sigma=\gamma_1\cdots\gamma_k\,\in\,\mathbb{S}_m} q^{|\{k\,:\,d \text{ divides } |\gamma_k|\}|}\right)\frac{z^m}{m!} &= \exp\big(F_d(z) + G_d(z)q\big) \\
&= \frac{1}{1-z}\cdot\frac{1}{(1-z^d)^{(q-1)/d}} = \sum_{i=0}^{\infty} z^i \sum_{j=0}^{\infty}\binom{(q-1)/d+j-1}{j} z^{dj} \\
&= \sum_{m=0}^{\infty}\sum_{j=0}^{\lfloor m/d\rfloor}\binom{(q-1)/d+j-1}{j} z^m = \sum_{m=0}^{\infty}\binom{\lfloor (q+m-1)/d\rfloor}{\lfloor m/d\rfloor} z^m,
\end{aligned}$$

where we used the identity $\sum_{j=0}^{r}\binom{s+j-1}{j} = \binom{s+r}{r}$, which holds for all integers $r, s \geq 0$. Therefore, we get that

$$\sum_{\sigma=\gamma_1\cdots\gamma_k\,\in\,\mathbb{S}_m} q^{|\{k\,:\,d \text{ divides } |\gamma_k|\}|} = m!\binom{\lfloor (q+m-1)/d\rfloor}{\lfloor m/d\rfloor}. \tag{10}$$

Finally, summing (9) over $\sigma \in \mathbb{S}_m$, and employing (10), we obtain (8), as desired. □

Lemma 5.5. *For each positive integer m, we have that*

$$\sum_{\sigma \in \mathbb{S}_m} |E^\star_\sigma| = m! \sum_{d \,|\, \gcd(q-1,m)} \varphi(d) \binom{(q-1)/d}{m/d}. \tag{11}$$

Proof. The proof is similar to that of Lemma 5.4. Let $\sigma \in \mathbb{S}_m$, $\lambda \in \mathbb{F}_q^*$, and $\boldsymbol{x} \in \mathbb{F}_q^{\star m \times 1,1}$. Write $\sigma = \gamma_1 \cdots \gamma_t$, where $\gamma_1, \ldots, \gamma_t$ are disjoint cycles, and let d be the multiplicative order of λ. It follows easily that $\boldsymbol{x} \in E^\star_{\sigma,\lambda}$ if and only if: for each $k \in \{1, \ldots, t\}$ we have that $|\gamma_k| = d$ and $\boldsymbol{x}$ over γ_k is a nonzero geometric progression of ratio λ; and the aforementioned geometric progressions are disjoint. In other words, we have that $\boldsymbol{x} \in E^\star_{\sigma,\lambda}$ if and only if $\boldsymbol{x}$ over γ_k, for $k \in \{1, \ldots, t\}$, are disjoint cosets of the quotient group $\mathbb{F}_q^*/\langle\lambda\rangle$. Consequently, we get that $|E^\star_{\sigma,\lambda}| = 0$ if $|\gamma_k| \neq d$ for some k, while

$$|E^\star_{\sigma,\lambda}| = \binom{(q-1)/d}{m/d} (m/d)! \cdot d^{m/d} \tag{12}$$

if $|\gamma_k| = d$ for each $k \in \{1, \ldots, t\}$ (and so $t = m/d$). More precisely, the first factor in (12) is equal to the number of choices of one of the $(q-1)/d$ classes of $\mathbb{F}_q^*/\langle\lambda\rangle$ for each of the m/d cycles of σ, while the second factor is equal to the number of choices of one of the d elements of each class for each of the m/d cycles. Therefore, since $E^\star_\sigma = \bigcup_{\lambda \in \mathbb{F}_q^*} E^\star_{\sigma,\lambda}$ is a disjoint union, we get that

$$|E^\star_\sigma| = \sum_{d \,|\, \gcd(q-1,m)} \varphi(d) \binom{(q-1)/d}{m/d} (m/d)!\, d^{m/d}\, \mathbb{1}\big[|\gamma_k| = d \text{ for } k = 1, \ldots, t\big], \tag{13}$$

where we used the fact that there are $\varphi(d)$ elements of order d in the multiplicative (cyclic) group $\mathbb{F}_q^*$. If d divides m, then there are $m!/\big(d^{m/d}(m/d)!\big)$ permutations whose cycles have all length equal to d [20, Eq. (1.2)]. Therefore, summing (13) over $\sigma \in \mathbb{S}_m$ we get (11), as desired. □

Lemma 5.6. *For each positive integer $m \leq q-2$, we have that*

$$\sum_{d \,|\, \gcd(q-1,m)} \varphi(d) \binom{(q-1)/d}{m/d} < 3 \binom{q-1}{m}.$$

Proof. If $m = q-2$ then the claim is obvious since $\gcd(q-1, m) = 1$. Hence, assume that $m \leq q-3$. Note that $\binom{a/d}{b/d}^d \leq \binom{a}{b}$ for all positive integers a, b, d such that $d \mid \gcd(a, b)$. Indeed, this claim follows by considering that we can pick b elements from $\{1, \ldots, a\}$ by picking b/d elements from $\{a(k-1)/d+1, \ldots, ak/d\}$ for each $k \in \{1, \ldots, d\}$. Consequently, we have that

$$\begin{aligned}
\sum_{\substack{d \,|\, \gcd(q-1,m) \\ d > 1}} \varphi(d) \binom{(q-1)/d}{m/d} &\leq \sum_{\substack{d \,|\, \gcd(q-1,m) \\ d > 1}} \varphi(d) \binom{q-1}{m}^{1/d} \\
&\leq \sum_{\substack{d \,|\, \gcd(q-1,m) \\ d > 1}} \varphi(d) \binom{q-1}{m}^{1/2} = \gcd(q-1, m) \binom{q-1}{m}^{1/2} \leq m \binom{q-1}{m}^{1/2},
\end{aligned}$$

where we employed the well-known identity $\sum_{d \mid n} \varphi(d) = n$, which holds for every positive integer n. Therefore, we get that

$$\binom{q-1}{m}^{-1} \sum_{d \,|\, \gcd(q-1,m)} \varphi(d) \binom{(q-1)/d}{m/d} \leq 1 + m \binom{q-1}{m}^{-1/2} \leq 1 + m \binom{m+2}{m}^{-1/2} < 3,$$

as desired. □

5.2 Proof of Theorem 4.1

Let ℓ, m, n be positive integers with $\max(\ell, n) \leq m$, and let

$$(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \pi) \leftarrow \mathsf{GenIPKP}(q, \ell, m, n).$$

We have that

$$\begin{aligned} N_{\mathsf{sol}}(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}) &= \sum_{\rho \in \mathbb{S}_m} \mathbb{1}[\boldsymbol{A}\boldsymbol{\rho}\boldsymbol{B} = \boldsymbol{C}] = \sum_{\rho \in \mathbb{S}_m} \mathbb{1}[\boldsymbol{A}\boldsymbol{\rho}\boldsymbol{B} = \boldsymbol{A}\boldsymbol{\pi}\boldsymbol{B}] = \sum_{\rho \in \mathbb{S}_m} \mathbb{1}[\boldsymbol{A}(\boldsymbol{\rho} - \boldsymbol{\pi})\boldsymbol{B} = \boldsymbol{0}] \\ &= \sum_{\sigma \in \mathbb{S}_m} \mathbb{1}[\boldsymbol{A}(\boldsymbol{\sigma} - \boldsymbol{I})\boldsymbol{\pi}\boldsymbol{B} = \boldsymbol{0}] \stackrel{\mathsf{d}}{=} \sum_{\sigma \in \mathbb{S}_m} \mathbb{1}[\boldsymbol{A}(\boldsymbol{\sigma} - \boldsymbol{I})\widetilde{\boldsymbol{B}} = \boldsymbol{0}], \end{aligned} \tag{14}$$

where $\widetilde{\boldsymbol{B}} \leftarrow \mathbb{F}_q^{m \times n, n}$ and we employed the substitution $\rho = \sigma\pi$ and the fact that $\widetilde{\boldsymbol{B}} \stackrel{\mathsf{d}}{=} \boldsymbol{\pi}\boldsymbol{B}$. Consequently, we get that

$$\mathbb{E}\big[N_{\mathsf{sol}}(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})\big] = \sum_{\sigma \in \mathbb{S}_m} \mathbb{P}[\boldsymbol{A}(\boldsymbol{\sigma} - \boldsymbol{I})\widetilde{\boldsymbol{B}} = \boldsymbol{0}]. \tag{15}$$

Note that $\boldsymbol{A}$ and $\widetilde{\boldsymbol{B}}$ are independent random variables. If $k \leq m$ is a positive integer and $\sigma \in \mathbb{S}_{m,k}$, then it is easy to prove that $\mathrm{rk}(\boldsymbol{\sigma} - \boldsymbol{I}) = m - k$. Therefore, from (15) and Lemma 5.3(iii), we have that

$$\begin{aligned} \mathbb{E}\big[N_{\mathsf{sol}}(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})\big] &= \sum_{k=1}^{m} \sum_{\sigma \in \mathbb{S}_{m,k}} \mathbb{P}[\boldsymbol{A}(\boldsymbol{\sigma} - \boldsymbol{I})\widetilde{\boldsymbol{B}} = \boldsymbol{0}] \\ &= \sum_{k=1}^{m} |\mathbb{S}_{m,k}| \sum_{r=0}^{\min(\ell, m-k)} \frac{|\mathbb{F}_q^{\ell \times (m-k), r}|\, |\mathbb{F}_q^{(\ell-r)\times k, \ell - r}|\, q^{kr}}{|\mathbb{F}_q^{\ell \times m, \ell}|} \prod_{i=0}^{n-1} \frac{q^{m-r} - q^i}{q^m - q^i}, \end{aligned}$$

as desired. The proof is complete.

5.3 Proof of Theorem 4.2

Let ℓ and m be integers with $0 \leq \ell \leq m < q$, and let $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \pi) \leftarrow \mathsf{GenIPKP}^{\star}(q, \ell, m, 1)$. By reasoning as in the beginning of the proof of Theorem 4.1, we obtain that

$$\mathbb{E}\big[N_{\mathsf{sol}}(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})\big] = \sum_{\sigma \in \mathbb{S}_m} \mathbb{P}[\boldsymbol{A}(\boldsymbol{\sigma} - \boldsymbol{I})\widetilde{\boldsymbol{B}} = \boldsymbol{0}], \tag{16}$$

where $\widetilde{\boldsymbol{B}} \leftarrow \mathbb{F}_q^{\star m \times 1, 1}$. It follows easily that, for each $\sigma \in \mathbb{S}_m$ and $\boldsymbol{x} \in \mathbb{F}_q^{\star m \times 1, 1}$, the rank of $(\boldsymbol{\sigma} - \boldsymbol{I})\boldsymbol{x}$ is equal to 0 if σ is the identity, and is equal to 1 otherwise. Hence, since $\boldsymbol{A}$ and $\widetilde{\boldsymbol{B}}$ are independent random variables, from Lemma 5.3(i) we get that

$$\mathbb{P}[\boldsymbol{A}(\boldsymbol{\sigma} - \boldsymbol{I})\widetilde{\boldsymbol{B}} = \boldsymbol{0}] = \begin{cases} 1 & \text{if } \boldsymbol{\sigma} = \boldsymbol{I}; \\ (q^{m-\ell} - 1)/(q^m - 1) & \text{if } \boldsymbol{\sigma} \neq \boldsymbol{I}. \end{cases} \tag{17}$$

Therefore, from (16) and (17) we obtain (2), as desired. The proof is complete.

5.4 Proof of Theorem 4.3

Let ℓ and m be positive integers with $\ell < m$, and let

$$(\boldsymbol{A}, \boldsymbol{B}, \pi) \leftarrow \mathsf{GenPKP}(q, \ell, m, 1) \quad \text{and} \quad (\widetilde{\boldsymbol{A}}, \widetilde{\boldsymbol{B}}, \widetilde{\boldsymbol{C}}, \widetilde{\pi}) \leftarrow \mathsf{GenIPKP}(q, \ell, m, 1).$$

Since line 2 of GenPKP is equivalent to a loop that keeps picking $\boldsymbol{A} \leftarrow \mathbb{F}_q^{\ell \times m, \ell}$ until $\boldsymbol{A}\boldsymbol{\pi}\boldsymbol{B} = \boldsymbol{0}$ (essentially, a rejection sampling), it follows that the distribution of $(\boldsymbol{A}, \boldsymbol{B})$ is equal to the conditional distribution of $(\widetilde{\boldsymbol{A}}, \widetilde{\boldsymbol{B}})$ with respect to the event $\widetilde{\boldsymbol{C}} = \boldsymbol{0}$.

Hence, we have that

$$\begin{aligned}\mathbb{E}[N_{\mathsf{sol}}(\boldsymbol{A},\boldsymbol{B})] &= \sum_{\rho\in\mathbb{S}_m} \mathbb{P}[\widetilde{\boldsymbol{A}}\boldsymbol{\rho}\widetilde{\boldsymbol{B}} = \mathbf{0} \mid \widetilde{\boldsymbol{C}} = \mathbf{0}] \\ &= \frac{1}{\mathbb{P}[\widetilde{\boldsymbol{C}} = \mathbf{0}]} \sum_{\rho\in\mathbb{S}_m} \mathbb{P}[\widetilde{\boldsymbol{A}}\boldsymbol{\rho}\widetilde{\boldsymbol{B}} = \mathbf{0} \text{ and } \widetilde{\boldsymbol{C}} = \mathbf{0}]. \end{aligned} \tag{18}$$

Let $\widehat{\boldsymbol{B}} \leftarrow \mathbb{F}_q^{m\times 1,1}$, so that $\widehat{\boldsymbol{B}} \overset{\mathsf{d}}{=} \widetilde{\pi}\widetilde{\boldsymbol{B}}$. Since $\widetilde{\boldsymbol{A}}$ and $\widehat{\boldsymbol{B}}$ are independent random variables, by Lemma 5.3(i) we have that

$$\mathbb{P}[\widetilde{\boldsymbol{C}} = \mathbf{0}] = \mathbb{P}[\widetilde{\boldsymbol{A}}\widetilde{\pi}\widetilde{\boldsymbol{B}} = \mathbf{0}] = \mathbb{P}[\widetilde{\boldsymbol{A}}\widehat{\boldsymbol{B}} = \mathbf{0}] = \frac{q^{m-\ell}-1}{q^m-1}. \tag{19}$$

Moreover, we have that

$$\begin{aligned}\sum_{\rho\in\mathbb{S}_m} \mathbb{P}[\widetilde{\boldsymbol{A}}\boldsymbol{\rho}\widetilde{\boldsymbol{B}} = \mathbf{0} \text{ and } \widetilde{\boldsymbol{C}} = \mathbf{0}] &= \sum_{\rho\in\mathbb{S}_m} \mathbb{P}[\widetilde{\boldsymbol{A}}\boldsymbol{\rho}\widetilde{\boldsymbol{B}} = \mathbf{0} \text{ and } \widetilde{\boldsymbol{A}}\widetilde{\pi}\widetilde{\boldsymbol{B}} = \mathbf{0}] \\ &= \sum_{\sigma\in\mathbb{S}_m} \mathbb{P}[\widetilde{\boldsymbol{A}}\boldsymbol{\sigma}\widetilde{\pi}\widetilde{\boldsymbol{B}} = \mathbf{0} \text{ and } \widetilde{\boldsymbol{A}}\widetilde{\pi}\widetilde{\boldsymbol{B}} = \mathbf{0}] = \sum_{\sigma\in\mathbb{S}_m} \mathbb{P}[\widetilde{\boldsymbol{A}}\boldsymbol{\sigma}\widehat{\boldsymbol{B}} = \mathbf{0} \text{ and } \widetilde{\boldsymbol{A}}\widehat{\boldsymbol{B}} = \mathbf{0}] \\ &= \sum_{\sigma\in\mathbb{S}_m} \mathbb{P}[\widetilde{\boldsymbol{A}}(\boldsymbol{\sigma}\widehat{\boldsymbol{B}} \mid \widehat{\boldsymbol{B}}) = \mathbf{0}], \end{aligned} \tag{20}$$

where we employed the substitution $\rho = \sigma\widetilde{\pi}$.

For each $\sigma \in \mathbb{S}_m$ and $\boldsymbol{x} \in \mathbb{F}_q^{m\times 1,1}$, we have that $\mathrm{rk}(\boldsymbol{\sigma}\boldsymbol{x} \mid \boldsymbol{x})$ is equal to 1 if $\boldsymbol{x} \in E_\sigma$, and is equal to 2 otherwise. Hence, by Lemma 5.3(i), we get that

$$\mathbb{P}[\widetilde{\boldsymbol{A}}(\boldsymbol{\sigma}\boldsymbol{x} \mid \boldsymbol{x}) = \mathbf{0}] = \frac{q^{m-\ell}-1}{q^m-1} \begin{cases} 1 & \text{if } \boldsymbol{x} \in E_\sigma; \\ (q^{m-\ell}-q)/(q^m-q) & \text{if } \boldsymbol{x} \notin E_\sigma. \end{cases}$$

Consequently, we obtain that

$$\begin{aligned}\mathbb{P}[\widetilde{\boldsymbol{A}}(\boldsymbol{\sigma}\widehat{\boldsymbol{B}} \mid \widehat{\boldsymbol{B}}) = \mathbf{0}] &= \mathbb{P}[\widehat{\boldsymbol{B}} \in E_\sigma]\,\mathbb{P}\big[\widetilde{\boldsymbol{A}}(\boldsymbol{\sigma}\widehat{\boldsymbol{B}} \mid \widehat{\boldsymbol{B}}) = \mathbf{0} \mid \widehat{\boldsymbol{B}} \in E_\sigma\big] \\ &\quad + \mathbb{P}[\widehat{\boldsymbol{B}} \notin E_\sigma]\,\mathbb{P}\big[\widetilde{\boldsymbol{A}}(\boldsymbol{\sigma}\widehat{\boldsymbol{B}} \mid \widehat{\boldsymbol{B}}) = \mathbf{0} \mid \widehat{\boldsymbol{B}} \notin E_\sigma\big] \\ &= \frac{q^{m-\ell}-1}{q^m-1}\left(\frac{|E_\sigma|}{q^m-1} + \left(1-\frac{|E_\sigma|}{q^m-1}\right)\frac{q^{m-\ell}-q}{q^m-q}\right) \\ &= \frac{q^{m-\ell}-1}{q^m-1}\left(\frac{q^{m-\ell}-q}{q^m-q} + \frac{q^m-q^{m-\ell}}{(q^m-1)(q^m-q)}\,|E_\sigma|\right) \end{aligned} \tag{21}$$

Therefore, from (18), (19), (20), and (21), we have that

$$\mathbb{E}[N_{\mathsf{sol}}(\boldsymbol{A},\boldsymbol{B})] = \frac{m!(q^{m-\ell}-q)}{q^m-q} + \frac{q^m-q^{m-\ell}}{(q^m-1)(q^m-q)} \sum_{\sigma\in\mathbb{S}_m} |E_\sigma|.$$

At this point, the claim follows from Lemma 5.4. The proof is complete.

5.5 Proof of Theorem 4.4

Let ℓ and m be positive integers with $\ell < m < q$, and let

$$(\boldsymbol{A},\boldsymbol{B},\pi) \leftarrow \mathsf{GenPKP}^*(q,\ell,m,1) \quad \text{and} \quad (\widetilde{\boldsymbol{A}},\widetilde{\boldsymbol{B}},\widetilde{\boldsymbol{C}},\widetilde{\pi}) \leftarrow \mathsf{GenIPKP}^*(q,\ell,m,1).$$

The proof proceeds *mutatis mutandis* as the proof of Theorem 4.3. The only differences are that: we pick $\widehat{\boldsymbol{B}} \leftarrow \mathbb{F}_q^{\star m\times 1,1}$, we have $E_\sigma^\star$ instead of E_σ (so that, in particular, it holds $\mathbb{P}[\widehat{\boldsymbol{B}} \in E_\sigma^\star] = |E_\sigma^\star|/(m!\binom{q-1}{m})$), and we apply Lemma 5.5 instead of Lemma 5.4.

6 Concluding remarks

We proved exact formulas for the expected number of solutions to random instances of the IPKP and the PKP generated by the natural algorithms $\mathsf{GenIPKP}$, $\mathsf{GenIPKP}^\star$, GenPKP, and $\mathsf{GenPKP}^\star$. In addition, we compared these exact formulas with the heuristic formula that has been extensively used in previous works. For instances generated by $\mathsf{GenIPKP}$ and GenPKP, we showed that the heuristic formula can be far off from the exact values, while for instances generated by $\mathsf{GenIPKP}^\star$ and $\mathsf{GenPKP}^\star$ the heuristic formula provides good approximations to the exact values.

Let us briefly outline some possible directions for future research. First, it might be interesting to generalize Theorems 4.2, 4.3, and 4.4 to the multidimensional case $n > 1$. For Theorem 4.3, respectively Theorem 4.4, this seems related to the study of the rank of matrices of the form $(\boldsymbol{\sigma X} \mid \boldsymbol{X})$, where $\sigma \in \mathbb{S}_m$ and $\boldsymbol{X} \in \mathbb{F}_q^{m\times n,n}$, respectively $\boldsymbol{X} \in \mathbb{F}_q^{\star m\times n,n}$.

Second, in light of the last paragraph of Section 3, one could study the expected number of solutions to random instances of the IPKP and the PKP when $\boldsymbol{A}$ has pairwise distinct columns and $\boldsymbol{B}$ has pairwise distinct rows.

Third, one could try to compute, or just upper bound, the second moment, or more generally the higher moments, of the number of solutions N_{sol} to the random instances of the IPKP or the PKP. This could make it possible to obtain upper bounds for the tail probability $\mathbb{P}[N_{\mathsf{sol}} \geq N]$, where $N \geq 2$ is an integer, which are better than bounds based only on the expected value of N_{sol}, such as the Markov inequality.

Finally, from a more abstract perspective, one could investigate if an appropriate normalization of N_{sol} converges in law to some known random variable (for example, a standard normal random variable) when ℓ, m, n go to infinity in a way to be made precise.

References

[1] T. Baritaud, M. Campana, P. Chauvaud, and H. Gilbert, *On the security of the permuted kernel identification scheme*, Advances in Cryptology – CRYPTO '92 (Santa Barbara, CA, USA) (E. F. Brickell, ed.), Lecture Notes in Computer Science, vol. 740, Springer, Heidelberg, Germany, August 16–20, 1993, pp. 305–311.

[2] C. H. Bennett, E. Bernstein, G. Brassard, and U. Vazirani, *Strengths and weaknesses of quantum computing*, SIAM J. Comput. **26** (1997), no. 5, 1510–1523.

[3] S. Bettaieb, L. Bidoux, V. Dyseryn, A. Esser, P. Gaborit, M. Kulkarni, and M. Palumbi, *PERK: compact signature scheme based on a new variant of the permuted kernel problem*, Designs, Codes and Cryptography (2024), online first.

[4] W. Beullens, *Sigma protocols for MQ, PKP and SIS, and Fishy signature schemes*, Advances in Cryptology – EUROCRYPT 2020, Part III (Zagreb, Croatia) (A. Canteaut and Y. Ishai, eds.), Lecture Notes in Computer Science, vol. 12107, Springer, Heidelberg, Germany, May 10–14, 2020, pp. 183–211.

[5] W. Beullens, J.-C. Faugère, E. Koussa, G. Macario-Rat, J. Patarin, and L. Perret, *PKP-based signature scheme*, Progress in Cryptology - INDOCRYPT 2019: 20th International Conference in Cryptology in India (Hyderabad, India) (F. Hao, S. Ruj, and S. Sen Gupta, eds.), Lecture Notes in Computer Science, vol. 11898, Springer, Heidelberg, Germany, December 15–18, 2019, pp. 3–22.

[6] L. Bidoux and P. Gaborit, *Compact post-quantum signatures from proofs of knowledge leveraging structure for the PKP, SD and RSD problems*, Codes, Cryptology and Information Security (Cham) (S. El Hajji, S. Mesnager, and E. M. Souidi, eds.), Springer Nature Switzerland, 2023, pp. 10–42.

[7] A. Esser, J. Verbel, F. Zweydinger, and E. Bellini, *CryptographicEstimators: a software library for cryptographic hardness estimation*, Cryptology ePrint Archive, Paper 2023/589, 2023, https://eprint.iacr.org/2023/589.

[8] S. D. Fisher and M. N. Alexander, *Classroom notes: Matrices over a finite field*, Amer. Math. Monthly **73** (1966), no. 6, 639–641.

[9] P. Flajolet and R. Sedgewick, *Analytic combinatorics*, Cambridge University Press, Cambridge, 2009.

[10] J. Georgiades, *Some remarks on the security of the identification scheme based on permuted kernels*, Journal of Cryptology **5** (1992), no. 2, 133–137.

[11] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete mathematics*, second ed., Addison-Wesley Publishing Company, Reading, MA, 1994, A foundation for computer science.

[12] É. Jaulmes and A. Joux, *Cryptanalysis of PKP: A new approach*, PKC 2001: 4th International Workshop on Theory and Practice in Public Key Cryptography (Cheju Island, South Korea) (Kwangjo Kim, ed.), Lecture Notes in Computer Science, vol. 1992, Springer, Heidelberg, Germany, February 13–15, 2001, pp. 165–172.

[13] E. Koussa, G. Macario-Rat, and J. Patarin, *On the complexity of the permuted kernel problem*, Cryptology ePrint Archive, Paper 2019/412, 2019, https://eprint.iacr.org/2019/412.

[14] R. Lampe and J. Patarin, *Analysis of some natural variants of the PKP algorithm*, Cryptology ePrint Archive, Report 2011/686, 2011, https://eprint.iacr.org/2011/686.

[15] NIST, *NIST announces additional digital signature candidates for the PQC standardization process*, online, July 2023, https://csrc.nist.gov/news/2023/additional-pqc-digital-signature-candidates.

[16] T. B. Paiva and R. Terada, *Cryptanalysis of the binary permuted kernel problem*, ACNS 21: 19th International Conference on Applied Cryptography and Network Security, Part II (Kamakura, Japan) (K. Sako and N. O. Tippenhauer, eds.), Lecture Notes in Computer Science, vol. 12727, Springer, Heidelberg, Germany, June 21–24, 2021, pp. 396–423.

[17] J. Patarin and P. Chauvaud, *Improved algorithms for the permuted kernel problem*, Advances in Cryptology – CRYPTO '93 (Santa Barbara, CA, USA) (D. R. Stinson, ed.), Lecture Notes in Computer Science, vol. 773, Springer, Heidelberg, Germany, August 22–26, 1994, pp. 391–402.

[18] PERK Team, *PERK Official Website*, online, May 2023, https://pqc-perk.org.

[19] G. Poupard, *A realistic security analysis of identification schemes based on combinatorial problems*, European Transactions on Telecommunications **8** (1997), no. 5, 471–480.

[20] B. E. Sagan, *The symmetric group*, second ed., Graduate Texts in Mathematics, vol. 203, Springer-Verlag, New York, 2001, Representations, combinatorial algorithms, and symmetric functions.

[21] P. Santini, M. Baldi, and F. Chiaraluce, *A novel attack to the permuted kernel problem*, 2022 IEEE International Symposium on Information Theory (ISIT), 2022, pp. 1441–1446.

[22] P. Santini, M. Baldi, and F. Chiaraluce, *Computational hardness of the permuted kernel and subcode equivalence problems*, IEEE Transactions on Information Theory **70** (2023), no. 3, 2254–2270.

[23] A. Shamir, *An efficient identification scheme based on permuted kernels*, Advances in Cryptology – CRYPTO '89 (Santa Barbara, CA, USA) (G. Brassard, ed.), Lecture Notes in Computer Science, vol. 435, Springer, Heidelberg, Germany, August 20–24, 1990, pp. 606–609.