Combinatorics, Probability and Computing (2025), **34**, pp. 131–150
doi:10.1017/S0963548324000312

ARTICLE

Cokernel statistics for walk matrices of directed and weighted random graphs

Alexander Van Werde

Eindhoven University of Technology, Department of Mathematics and Computer Science, Eindhoven, Netherlands
Email: a.van.werde@tue.nl

(Received 30 January 2024; revised 23 August 2024; accepted 28 August 2024; first published online 18 October 2024)

Abstract

The *walk matrix* associated to an $n \times n$ integer matrix $\mathbf{X}$ and an integer vector b is defined by $\mathbf{W} := (b, \mathbf{X}b, \ldots, \mathbf{X}^{n-1}b)$. We study limiting laws for the cokernel of $\mathbf{W}$ in the scenario where $\mathbf{X}$ is a random matrix with independent entries and b is deterministic. Our first main result provides a formula for the distribution of the p^m-torsion part of the cokernel, as a group, when $\mathbf{X}$ has independent entries from a specific distribution. The second main result relaxes the distributional assumption and concerns the $\mathbb{Z}[x]$-module structure.

The motivation for this work arises from an open problem in spectral graph theory, which asks to show that random graphs are often determined up to isomorphism by their (generalised) spectrum. Sufficient conditions for generalised spectral determinacy can, namely, be stated in terms of the cokernel of a walk matrix. Extensions of our results could potentially be used to determine how often those conditions are satisfied. Some remaining challenges for such extensions are outlined in the paper.

Keywords: Determined by spectrum; walk matrix; cokernel; random graph

2020 MSC Codes: Primary: 05C50; Secondary: 15B52, 60B20

1. Introduction

What information about a graph is encoded in the spectrum of its adjacency matrix? A well-known conjecture by van Dam and Haemers [24] suggests that the answer to this question is *all information in the typical case* in the sense that almost all graphs are determined up to isomorphism by their spectrum. Unfortunately, progress towards that conjecture has been fairly limited due to the fact that there are essentially no known general-purpose methods to prove that a graph is determined by its spectrum. There are many more proof techniques available to prove that a graph's spectrum does *not* determine it than to show that it does [1, 2, 8, 10, 21].

In view of this, it is intriguing that a sufficient condition for *generalised* spectral determinacy was discovered in 2006 by Wang and Xu [29, 30]. The *generalised spectrum* of a graph G here refers to the ordered pair $(\mathrm{spec}(\mathbf{A}), \mathrm{spec}(\mathbf{A}^c))$ consisting of the spectra of the adjacency matrix $\mathbf{A}$ of G and of the adjacency matrix $\mathbf{A}^c$ of its complement graph. Refinements of the results of [29, 30] have given rise to an active area of research in recent years; see, for example, [12, 17, 18, 26–28]. For definiteness, let us state a refinement that is particularly insightful and motivates the results of the current paper.

Define a matrix with integer entries by $\mathbf{W} := (\mathbf{A}^{j-1}e)_{j=1}^n$ where $e = (1, \ldots, 1)^{\mathrm{T}}$ is the all-ones vector and n is the number of vertices of G. Let $\mathrm{coker}(\mathbf{W}) := \mathbb{Z}^n/\mathbf{W}(\mathbb{Z}^n)$ denote the *cokernel* of this matrix. Then, the following is an equivalent rephrasing of [27, Theorem 1.1]; see [18, p. 2].

© The Author(s), 2024. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

Theorem 1.1 (Wang, [27]). *Let G be a simple graph. Assume that there exists an odd and square-free integer m such that*

$$\operatorname{coker}(\mathbf{W}) \cong (\mathbb{Z}/2\mathbb{Z})^{\lfloor n/2 \rfloor} \oplus (\mathbb{Z}/m\mathbb{Z})$$

as an Abelian group. Then, G is determined by its generalised spectrum up to isomorphism.

It is believed that the conditions of Theorem 1.1 are satisfied for a nonvanishing fraction of all simple graphs [27, Conjecture 2]. For comparison, the best known bound on the non-generalised problem is due to Koval and Kwan [11], who recently established that there are at least e^{cn} graphs on n vertices that are determined by their spectrum. The result from [11] represents a significant improvement relative to the previous long-standing barrier of $e^{c\sqrt{n}}$ but still only yields a quickly vanishing fraction of all $(1-o(1))2^{n(n-1)/2}/n!$ simple graphs. So, a nonvanishing fraction being determined by (generalised) spectrum would signify a remarkable development.

However, even though criteria for generalised spectral determinacy have been known for almost 20 years now, it remains an open problem to prove that the criteria are indeed frequently satisfied. The current state of knowledge on the frequency of satisfaction is mostly limited to numerical studies [9, 24, 27, 28]. The lack of theoretical work is surprising given that criteria for generalised spectral determinacy are an active area of research. A possible explanation is that it is not clear what proof techniques could be used. The current paper develops a novel line of attack by making a connection to proof techniques [3, 14, 20, 33], which were historically developed in the context of Cohen–Lenstra heuristics for the class groups of number fields [5] and in the study of sandpile groups of random graphs [4].

The problem is too challenging to solve in a single step, so we direct our efforts to a variant which is more convenient for technical reasons. Instead of simple graphs, we study random directed graphs with random edge weights. Concretely, given a $\mathbb{Z}^{n\times n}$-valued random matrix $\mathbf{X}$ with independent entries and a vector $b \in \mathbb{Z}^n$, we study the cokernel of the associated *walk matrix* $\mathbf{W} := (\mathbf{X}^{j-1}b)_{j=1}^n$. To explain the terminology, note that if $\mathbf{X}$ is $\{0,1\}^{n\times n}$-valued and interpreted as the directed adjacency matrix of a directed graph and $b = \mathbb{1}_S$ is the indicator vector of a subset $S \subseteq \{1,\dots,n\}$, then $\mathbf{W}_{i,j}$ counts the number of walks of length $j-1$ starting from i with an endpoint in S. Walk matrices of this kind are also of independent interest due to applications to control theory [7, 16, 22] and graph isomorphism problems [7, 13, 25].

In future work, it would be interesting to pursue extensions of our results to the setting of simple graphs where Theorem 1.1 is applicable.[1] The adjacency matrix of an undirected random graph has to be symmetric, so its entries cannot be independent, which makes the problem more difficult. We expect that our proof techniques will still give insights provided that appropriate modifications are made, but these modifications pose a nontrivial challenge; see Section 4.

1.1 Results

Given an Abelian group H and a prime power p^m, we denote $H_{p^m} := H/p^mH$. Then, the condition of Theorem 1.1 is satisfied if and only if $\operatorname{coker}(\mathbf{W})_{p^2} \in \{0, \mathbb{Z}/p\mathbb{Z}\}$ for every odd prime p and $\operatorname{coker}(\mathbf{W})_{2^2} \cong (\mathbb{Z}/2\mathbb{Z})^{\lfloor n/2 \rfloor}$. To determine how frequently Theorem 1.1 is applicable, it hence suffices to study the joint distribution of $\operatorname{coker}(\mathbf{W})_{p^2}$ over all primes p when $\mathbf{X}$ is the adjacency matrix of an undirected Erdős–Rényi random graph and $b = e$.

[1] There is no known analogue of Theorem 1.1 for directed graphs. (See however [17].) Indeed, directed graphs are typically *not* determined by their (generalised) spectrum because the transpose of a directed adjacency matrix has the same spectrum but typically corresponds to a different directed graph.

https://doi.org/10.1017/S0963548324000312 Published online by Cambridge University Press

In a directed and weighted setting, the following result gives a remarkably simple formula for the limiting marginal distribution for a single prime p:

Theorem 1.2. *Fix a prime p and an integer $m \geq 1$. For every $n \geq 1$, let $\mathbf{X}$ be a $\mathbb{Z}^{n\times n}$-valued random matrix with independent* $\mathrm{Unif}\{0, 1, \ldots, p^m - 1\}$*-distributed entries, and let $b \in \mathbb{Z}^n$ be a deterministic vector with $b \not\equiv 0 \bmod p$.*

Fix some $\ell \geq 1$ and $0 = \lambda_0 \leq \lambda_1 \leq \cdots \leq \lambda_\ell \leq m$. Let $i_0 := \#\{i \leq \ell : \lambda_i = m\}$ and denote $\delta_j := \lambda_{\ell-j+1} - \lambda_{\ell-j}$. Then, as Abelian groups,

$$\lim_{n\to\infty} \mathbb{P}\left(\mathrm{coker}(\mathbf{W})_{p^m} \cong \bigoplus_{i=1}^{\ell} \frac{\mathbb{Z}}{p^{\lambda_i}\mathbb{Z}}\right) = \prod_{i=i_0}^{\infty}\left(1 - p^{-(i+1)}\right)\prod_{j=1}^{\ell} p^{-j\delta_j}.$$

Note in particular that Theorem 1.2 predicts that $\mathrm{coker}(\mathbf{W})_{p^2} \in \{0, \mathbb{Z}/p\mathbb{Z}\}$ with nonvanishing probability, at least for random directed and weighted graphs. For example, the limiting probabilities associated with the primes $p = 3$, 5, and 7 are approximately 0.75, 0.91, and 0.96, respectively. We here emphasise odd primes because, while Theorem 1.2 also applies when $p = 2$, it is known that the distribution of $\mathrm{coker}(\mathbf{W})_{2^m}$ is very different for simple and non-simple graphs; see Section 4. It would hence be ill-advised to use directed graphs as a model for the condition $\mathrm{coker}(\mathbf{W})_{2^2} \cong (\mathbb{Z}/2\mathbb{Z})^{\lfloor n/2\rfloor}$.

The proof of Theorem 1.2 relies on interpretable combinatorial arguments. The downside is that the distributional assumption on the entries of $\mathbf{X}$ plays a crucial role, which makes the approach unsuitable for the study of unweighted graphs. Fortunately, a different proof approach allows us to study $\mathrm{coker}(\mathbf{W})$ in a general setting, which also covers the case where $\mathbf{X}$ has $\{0, 1\}$-valued entries. Additionally, this allows us to gain insight on the joint law across different primes, and the result even applies to sparse graphs.

For this more general setting, it turns out to be essential to interact with all canonical structure on $\mathrm{coker}(\mathbf{W})$. Equip $\mathbb{Z}^n$ with the $\mathbb{Z}[x]$-module structure defined by $xv := \mathbf{X}v$, and note that $\mathbf{W}(\mathbb{Z}^n)$ is precisely the $\mathbb{Z}[x]$-submodule of $\mathbb{Z}^n$ generated by b. Hence, the quotient $\mathrm{coker}(\mathbf{W})$ is canonically equipped with the structure of a $\mathbb{Z}[x]$-module. We require some additional notation. Let $Q(x) \in \mathbb{Z}[x]$ be a monic polynomial and consider a prime power p^m. Then, given a $\mathbb{Z}[x]$-module N, we define a quotient module by $N_{p^m,Q} := N/(p^m N + Q(x)N)$. We further abbreviate $R_{p^m,Q} := \mathbb{Z}[x]/(p^m\mathbb{Z}[x] + Q(x)\mathbb{Z}[x])$.

Fix a finite collection of prime numbers $\mathscr{P}$ and consider a scalar $\alpha > 0$. Then, a $\mathbb{Z}$-valued random variable Y is said to be *α-balanced mod $\mathscr{P}$* if for all $p \in \mathscr{P}$ and $y \in \mathbb{Z}/p\mathbb{Z}$,

$$\mathbb{P}(Y \equiv y \bmod p) \leq 1 - \alpha.$$

Given a ring R, recall that $\mathrm{Ext}^1_R(N, M)$ denotes the set of extensions of an R-modules N by an R-module M [31, p. 77], and denote $\mathrm{Aut}_R(N)$ and $\mathrm{Hom}_R(N, M)$ for the sets of R-module automorphisms and homomorphisms, respectively.

Theorem 1.3. *Fix a finite set of primes $\mathscr{P}$. For every $n \geq 1$, let $\mathbf{X}$ be a $\mathbb{Z}^{n\times n}$-valued random matrix with independent entries, not necessarily identically distributed, such that each entry $\mathbf{X}_{i,j}$ is α_n-balanced mod $\mathscr{P}$. Further, let $b \in \mathbb{Z}^n$ be deterministic with $b \not\equiv 0 \bmod p$ for every $p \in \mathscr{P}$.*

Assume that $\lim_{n\to\infty} n\alpha_n/\ln(n) = \infty$. Then, for every integer $m \geq 1$, monic polynomial $Q \in \mathbb{Z}[x]$ of degree ≥ 1, and collection of finite $R_{p^m,Q}$-modules $N_{p^m,Q}$:

(1) The quotients $\mathrm{coker}(\mathbf{W})_{p^m,Q}$ associated with different primes $p \in \mathscr{P}$ are asymptotically independent. More precisely, as $\mathbb{Z}[x]$-modules,

$$\lim_{n\to\infty} \mathbb{P}\big(\forall p \in \mathscr{P} : \mathrm{coker}(\mathbf{W})_{p^m,Q} \cong N_{p^m,Q}\big) = \prod_{p\in\mathscr{P}} \mu_{p^m,Q}(N_{p^m,Q})$$

for certain probability measures $\mu_{p^m,Q}$ supported on finite $R_{p^m,Q}$-modules.

https://doi.org/10.1017/S0963548324000312 Published online by Cambridge University Press

(2) *Suppose that* $Q(x) \equiv \prod_{i=1}^{r_p} Q_{i,p}(x)^{q_{i,p}} \bmod p$ *is the unique factorisation of* $Q \bmod p$ *into powers of distinct monic irreducible polynomials* $Q_{i,p} \in \mathbb{F}_p[x]$. *Let* $d_{i,p} := \deg Q_{i,p}$ *denote the degree of* $Q_{i,p}$. *Then, the measure* $\mu_{p^m,Q}$ *is given by the following identity:*

$$\mu_{p^m,Q}(N_{p^m,Q}) = \frac{1}{\#N_{p^m,Q}\#\mathrm{Aut}_{R_{p^m,Q}}(N_{p^m,Q})} \prod_{i=1}^{r_p}\prod_{j=1}^{\infty} \times \left(1 - \frac{\#\,\mathrm{Ext}^1_{R_{p^m,Q}}\big(N_{p^m,Q}, \mathbb{F}_p[x]/(Q_{i,p}(x)\mathbb{F}_p[x])\big)}{\#\mathrm{Hom}_{R_{p^m,Q}}(N_{p^m,Q}, \mathbb{F}_p[x]/(Q_{i,p}(x)\mathbb{F}_p[x])} p^{-(1+j)d_{i,p}}\right).$$

So far as it pertains to the group structure, the conclusion of Theorem 1.3 with $\mathscr{P} = \{p\}$ is weaker than Theorem 1.2, but only slightly. To study $\mathrm{coker}(\mathbf{W})_{p^m}$ itself, without the additional quotient by $Q(x)$, it would, namely, be sufficient to have a tightness condition stating that $\lim_{C\to\infty} \liminf_{n\to\infty} \mathbb{P}(\#\mathrm{coker}(\mathbf{W})_{p^m} \le C) = 1$. Such a condition would yield the limiting law as a $\mathbb{Z}[x]$-module,[2] and the limiting law as a group then follows by summing over all $\mathbb{Z}[x]$-module structures on $\oplus_{i=1}^{\ell} \mathbb{Z}/p^{\lambda_i}\mathbb{Z}$.

We are not aware of a direct proof that the aforementioned sum recovers the formula in Theorem 1.2, but this would follow indirectly since Theorem 1.3 also applies to a matrix with uniform entries as in Theorem 1.2. So, if the tightness condition holds, then the distribution of $\mathrm{coker}(\mathbf{W})_{p^m}$ converges to the same limit as in Theorem 1.2, and one has asymptotic independence for any fixed finite set of primes. Hence, additionally assuming that the restriction to finite sets of primes can also be removed, we are led to the following:

Conjecture 1.4. *For every* $n \ge 1$, *let* $\mathbf{X}$ *be a* $\{0,1\}^{n\times n}$*-valued random matrix with independent entries, and let* $b = e$ *be the all-ones vector. Assume that there exists a sequence* α_n *satisfying* $\lim_{n\to\infty} n\alpha_n/\ln(n) = \infty$ *such that* $\mathbb{P}(\mathbf{X}_{i,j} = 0), \mathbb{P}(\mathbf{X}_{i,j} = 1) \le 1 - \alpha_n$ *for each entry. Then,*

$$\lim_{n\to\infty} \mathbb{P}\big(\mathrm{coker}(\mathbf{W})_{p^2} \in \{0, \mathbb{Z}/p\mathbb{Z}\} \text{ for every odd prime p}\big) = \prod_{\text{odd primes } p} \big(1 + p^{-1}\big) \prod_{i=0}^{\infty} \big(1 - p^{-(i+1)}\big).$$

Remark 1.5. The restriction that $\alpha_n \gg \ln(n)/n$ in Theorem 1.3, which limits how sparse the matrices are allowed to be, is close to optimal. The conclusion of Theorem 1.3, namely, has to fail when $\mathbf{X}$ has independent entries satisfying $\mathbb{P}(\mathbf{X}_{i,j} = 0) \ge 1 - (1-\varepsilon)\ln(n)/n$ for $\varepsilon > 0$.

Indeed, the coupon collector theorem then implies that $\mathbf{X}$ has many rows equal to zero so that $\mathrm{rank}(\mathbf{X}) \le n - 2$ with high probability. The latter implies that $\mathbf{X}(\mathbb{F}_p^n) + \mathbb{F}_p b \ne \mathbb{F}_p^n$. Hence, since $\mathrm{coker}(\mathbf{W})_{p,x} \cong \mathbb{Z}^n/(\mathbf{W}(\mathbb{Z}^n) + p\mathbb{Z}^n + x\mathbb{Z}^n)$ is isomorphic to $\mathbb{F}_p^n/(\mathbf{X}(\mathbb{F}_p^n) + \mathbb{F}_p b)$, it would follow that $\lim_{n\to\infty} \mathbb{P}(\mathrm{coker}(\mathbf{W})_{p,x} = \{0\}) = 0$. This is incompatible with the conclusion of Theorem 1.3 since $\mu_{p,x}(\{0\}) = \prod_{j=1}^{\infty}(1 - p^{-1-j}) \ne 0$.

Remark 1.6. The limiting distribution of the $\mathbb{Z}[x]$-module $\mathrm{coker}(Q(\mathbf{X}))$ with $Q(x)$ a fixed polynomial was recently studied by Cheong and Yu [3]. Interestingly, the formulas found in Theorem 1.3 and [3, Theorem 1.3] bear a close resemblance, although they are not identical. This resemblance is not entirely surprising since the studied objects can be related. Indeed, note that $\mathrm{coker}(Q(\mathbf{X}))_{p^m} \cong \mathbb{Z}^n/(Q(x)\mathbb{Z}^n + p^m\mathbb{Z}^n)$ whereas $\mathrm{coker}(\mathbf{W})_{p^m,Q} \cong \mathbb{Z}^n/(Q(x)\mathbb{Z}^n + p^m\mathbb{Z}^n + \mathbb{Z}[x]b)$.

One can however not directly recover Theorem 1.3 from [3]. For one thing, [3] only considers the case $\mathscr{P} = \{p\}$ and does not allow sparse settings where $\alpha_n \to 0$. Further, $\mathrm{coker}(\mathbf{W})_{p^m,Q}$ not

[2] Indeed, note that the tightness condition would imply that $\mathrm{coker}(\mathbf{W})_{p^m,Q} \cong \mathrm{coker}(\mathbf{W})_{p^m}$ with high probability for $Q(x) := \prod_{i\in\{1,\dots,r\}} \prod_{j\in\{1,\dots,r\}\setminus\{i\}} (t^i - t^j)$ with r sufficiently large because a finite group can only admit finitely many distinct endomorphisms.

https://doi.org/10.1017/S0963548324000312 Published online by Cambridge University Press

only depends on the isomorphism class of $\operatorname{coker}(Q(\mathbf{X}))_{p^m}$ but also on how the reduction of b lies in $\operatorname{coker}(Q(\mathbf{X}))_{p^m}$, which is information we do not have access to.

1.2 Proof techniques

The proof of Theorem 1.2 relies on an analysis of the sequence of random vectors $(\mathbf{X}^{t-1}b)_{t=1}^n$, viewed as a stochastic process. More precisely, we show in Corollary 2.2 that the group structure of $\operatorname{coker}(\mathbf{W})_{p^m}$ can be computed in terms of the sequence of random variables $0 = U_1 \leq U_2 \leq \ldots \leq U_n \leq \infty$ defined by

$$U_t := \sup\left\{j \geq 0 : \mathbf{X}^{t-1}b \in \operatorname{span}_{\mathbb{Z}}\left(\mathbf{X}^{i-1}b : 1 \leq i \leq t-1\right) + p^j\mathbb{Z}^n\right\}. \tag{1.1}$$

Here, given a ring R and $v_1, \ldots, v_t \in R^n$, we write $\operatorname{span}_R(v_1, \ldots, v_t) := \left\{\sum_{i=1}^t c_i v_i : c_i \in R\right\}$. The remaining difficulty is then to study the joint law of the U_j. The distributional assumption in Theorem 1.2 plays an important role for the latter task: the independence and equidistribution of the entries imply that $\mathbf{X}$ induces a uniform random endomorphism of $(\mathbb{Z}/p^m\mathbb{Z})^n$, which can be used to establish Markovian dynamics for $\min\{U_t, m\}$; see Lemma 2.4.

This proof yields a direct combinatorial interpretation for the formula in Theorem 1.2. Given U_t, a counting argument implies that the probability that $U_{t+1} \geq U_t + \delta$ is $p^{-\delta(n-t)}$ for any δ satisfying $U_t + \delta \leq m$. Taking $j = n - t$ then explains the factors of the form $p^{-j\delta_j}$ in Theorem 1.2. Factors of the form $1 - p^{-j}$ arise when we additionally have to enforce that $U_{t+1} \leq U_t + \delta$. A further benefit of the approach is that it can also be used to establish the law of $\operatorname{coker}(\mathbf{W})_{p^m}$ for finite n; see Proposition 2.5.

The proof of Theorem 1.3 relies on more sophisticated techniques. In particular, we employ the *category-theoretic moment method*. In classical probability theory, the moment method allows one to establish convergence in distribution of a sequence of $\mathbb{R}$-valued random variables $(Y_i)_{i=1}^\infty$ by showing that the moments $\mathbb{E}[Y_i^n]$ converge to those of the desired limiting law provided that some mild conditions are satisfied. Results of Sawin and Wood [20] similarly allow one to establish limiting laws for random algebraic objects such as groups and modules by showing that category-theoretic moments converge. Here, given a ring R and a deterministic R-module N, the *N-moment*[3] of a random R-module Y is given by $\mathbb{E}[\#\operatorname{Sur}_R(Y, N)]$ where $\operatorname{Sur}_R(Y, N)$ denotes the set of surjective R-module morphisms from Y to N.

The main challenge is hence to estimate the N-moments of the random $\mathbb{Z}[x]$-module $\operatorname{coker}(\mathbf{W})$. To this end, we employ a strategy developed by Wood [32, 33] and Nguyen and Wood [15] for the estimation of moments of random algebraic objects associated with random matrices with α-balanced entries. A key difference between our setting and the one in [15, 32, 33] is that we have little control over the joint law of the entries $\mathbf{W}_{i,j}$ of our matrix of interest since these are nontrivial algebraic combinations of the entries of $\mathbf{X}$. This is why it is essential to view $\operatorname{coker}(\mathbf{W})$ as a $\mathbb{Z}[x]$-module, not only as a group. That is, the $\mathbb{Z}[x]$-module-theoretic viewpoint allows us to untangle the algebraic dependencies and hence estimate the category-theoretic moments; see (3.3) and the subsequent remarks.

The advantage of the category-theoretic approach is that it is robust, as is demonstrated by the general distributional assumptions in Theorem 1.3. For comparison, the proof approach for Theorem 1.2 is highly non-robust. Indeed, as mentioned above, that proof uses that $\mathbf{X}$ induces a uniform random endomorphism of $(\mathbb{Z}/p^m\mathbb{Z})^n$: a property which is only satisfied when $\mathbf{X} \bmod p^m$ has independent and uniformly distributed entries. The robustness of the category-theoretic approach makes us hopeful that we will be able to generalise it in future work, although this remains a nontrivial task; see Section 4.

[3] An explanation for the terminology *moment* may be found in [4, Section 3.3].

https://doi.org/10.1017/S0963548324000312 Published online by Cambridge University Press

1.3 Structure of this paper

The proof of Theorem 1.2 is given in Section 2. We there also give a non-asymptotic variant of Theorem 1.2. The proof of Theorem 1.3 is given in Section 3. Directions for future work are outlined in Section 4.

2. Proof of Theorem 1.2

Throughout this section, we fix a prime p and a vector $b \in \mathbb{Z}^n$ with $b \not\equiv 0 \bmod p$. Recall from Section 1.2 that the proof has two main ingredients. The first ingredient is Corollary 2.2, which shows that the group structure of $\operatorname{coker}(\mathbf{W})_{p^m}$ can be computed in terms of $U_1, \ldots, U_n$. The second ingredient is Lemma 2.4, which concerns the joint law of the U_t when $\mathbf{X}$ is random as in Theorem 1.2. We combine these results to establish Theorem 1.2 in Section 2.3.

2.1 Computing $\operatorname{coker}(\mathbf{W})_{p^m}$ in terms of $U_1, \ldots, U_n$

Recall the definition of U_t from (1.1). Fix an integer $m \geq 1$ and abbreviate $R := \mathbb{Z}/p^m\mathbb{Z}$. The following lemma then produces an R-module basis for R^n which is well-adapted to the computation of $\operatorname{coker}(\mathbf{W})_{p^m}$.

Lemma 2.1. *Write $\Lambda_t := \min\{U_t, m\}$. Then, there exist $v_1, \ldots, v_n \in R^n$ such that for every $t \leq n$ the following properties are satisfied:*

(i) The reduction of the matrix $(v_1, \ldots, v_t)$ modulo p has rank t over $\mathbb{F}_p$.

(ii) It holds that $\operatorname{span}_R\left(\mathbf{X}^{i-1}b \bmod p^m : 1 \leq i \leq t\right) = \operatorname{span}_R\left(p^{\Lambda_i} v_i : 1 \leq i \leq t\right)$.

Proof. We proceed by induction on t. If $t = 1$, then $U_1 = 0$, so $v_1 \equiv b \bmod p^m$ satisfies both properties. Now suppose that $t > 1$ and assume that there exist $v_1, \ldots, v_{t-1}$ such that both properties are satisfied. We prove the existence of some v_t.

First, suppose that $U_t \geq m$. Then, the definition (1.1) yields $\operatorname{span}_R\left(\mathbf{X}^{i-1}b \bmod p^m : i \leq t\right) = \operatorname{span}_R\left(\mathbf{X}^{i-1}b \bmod p^m : i \leq t-1\right)$ and $p^{\Lambda_t} v = p^m v = 0$ for every $v \in R^n$. Consequently, due to the induction hypothesis, both properties are satisfied if we let $v_t \in R^n$ be an arbitrary vector, which is not in $\operatorname{span}_R(v_1, \ldots, v_{t-1}) + pR^n$. Such a vector exists because $t - 1 < n$.

Now suppose that $U_t < m$. By definition of U_t, there then exist $w \in p^{U_t}R^n$ and $r \in \operatorname{span}_R\left(\mathbf{X}^{i-1}b \bmod p^m : i \leq t-1\right)$ such that $\mathbf{X}^{t-1}b \equiv w + r \bmod p^m$. Pick some $v_t \in R^n$ with $p^{U_t} v_t = w$. Then, due to the induction hypothesis, item (ii) is satisfied, and item (i) is equivalent to the statement that $v_t \notin \operatorname{span}_R(v_1, \ldots, v_{t-1}) + pR^n$. The latter statement is true. Indeed, if not, then $w \in \operatorname{span}_R\left(p^{U_t}v_1, \ldots, p^{U_t}v_{t-1}\right) + p^{U_t+1}R^n$. Then, considering that $\operatorname{span}_R(p^{U_t}v_i : i \leq t-1) \subseteq \operatorname{span}_R(\mathbf{X}^{i-1}b \bmod p^m : i \leq t-1)$ by item (ii) of the induction hypothesis and the fact that the U_i are nondecreasing, it follows from $\mathbf{X}^{t-1}b \equiv w + r \bmod p^m$ that $\mathbf{X}^{t-1}b \in \operatorname{span}_{\mathbb{Z}}(\mathbf{X}^{i-1}b : i \leq t-1) + p^{U_t+1}\mathbb{Z}^n$ contradicting the maximality of U_t in (1.1). This shows that both properties are satisfied. □

Corollary 2.2. *Adopt the notation of Lemma 2.1. Then,* $\operatorname{coker}(\mathbf{W})_{p^m} \cong \oplus_{t=1}^n \mathbb{Z}/p^{\Lambda_t}\mathbb{Z}$.

Proof. The case with $t = n$ in item (i) of Lemma 2.1 implies that the vectors $v_1, \ldots, v_n \in R^n$ determine an R-module basis for R^n. Further, item (ii) yields that $\mathbf{W}(R^n) = \operatorname{span}_R\left(p^{\Lambda_1}v_1, \ldots, p^{\Lambda_n}v_n\right)$. The claim is hence immediate since $\operatorname{coker}(\mathbf{W})_{p^m} \cong R^n/\mathbf{W}(R^n)$. □

2.2 Markovian dynamics

Given a matrix $\mathbf{M}$, abbreviate $\operatorname{rank}_p(\mathbf{M})$ for the rank of $\mathbf{M}$ mod p over $\mathbb{F}_p$. Recall that $R = \mathbb{Z}/p^m\mathbb{Z}$. The following lemma provides a partial converse for Lemma 2.1 in the case $U_t < m$:

https://doi.org/10.1017/S0963548324000312 Published online by Cambridge University Press

Lemma 2.3. *Consider some $t \le n$ and integers $0 = u_1 \le u_2 \le \ldots \le u_t < m$. Then, it holds that $U_i = u_i$ for every $i \le t$ if and only if there exist $v_1, \ldots, v_t \in R^n$ with $\mathrm{rank}_p(v_1, \ldots, v_t) = t$ such that for every $i \le t$ one has $\mathrm{span}_R\,(\mathbf{X}^{j-1} b \bmod p^m : j \le i) = \mathrm{span}_R\,(p^{u_j} v_j : j \le i)$.*

Proof. If $U_i = u_i$ for every $i \le t$, then the existence of $v_1, \ldots, v_t$ with the claimed properties follows from Lemma 2.1. Conversely, assume that such $v_1, \ldots, v_t$ exist. Then, for every $i \le t$,

$$\begin{aligned}\mathrm{span}_R\big(\mathbf{X}^{j-1} b \bmod p^m : j \le i\big) &= \mathrm{span}_R\big(p^{u_j} v_j : j < i\big) + \mathrm{span}_R(p^{u_i} v_i) \\ &\subseteq \mathrm{span}_R\big(\mathbf{X}^{j-1} b \bmod p^m : j < i\big) + p^{u_i} R^n.\end{aligned} \tag{2.1}$$

Considering that $R = \mathbb{Z}/p^m\mathbb{Z}$ with $m \ge u_i$, it follows that $\mathbf{X}^{i-1} b \in \mathrm{span}_{\mathbb{Z}}(\mathbf{X}^{j-1} b : j < i) + p^{u_i}\mathbb{Z}^n$. The definition (1.1) hence yields $U_i \ge u_i$. On the other hand, since $\mathrm{rank}_p(v_1, \ldots, v_t) = t$ and $u_i < m$, we have $p^{u_i} v_i \notin \mathrm{span}_R(p^{u_j} v_j : j < i) + p^{u_i+1} R^n$. Hence,

$$\begin{aligned}\mathrm{span}_R\big(p^{u_j} v_j : j < i\big) + \mathrm{span}_R(p^{u_i} v_i) &\not\subseteq \mathrm{span}_R\big(p^{u_j} v_j : j < i\big) + p^{u_i+1} R^n \\ &= \mathrm{span}_R\big(\mathbf{X}^{j-1} b \bmod p^m : j < i\big) + p^{u_i+1} R^n.\end{aligned} \tag{2.2}$$

Given the equality in (2.1) and the fact that $u_i + 1 \le m$, this implies that $\mathbf{X}^{i-1} b \notin \mathrm{span}_{\mathbb{Z}}(\mathbf{X}^{j-1} b : j < i) + p^{u_i+1}\mathbb{Z}^n$. This means that $U_i \le u_i$. Combine the inequalities $U_i \ge u_i$ and $U_i \le u_i$ to conclude the proof. □

Lemma 2.4. *Assume that* $\mathbf{X}$ *has independent and* $\mathrm{Unif}\{0, 1, \ldots, p^m - 1\}$*-distributed entries. Consider some $t \le n - 1$. Then, for every $0 = u_1 \le \ldots \le u_t < m$ and $u_{t+1} \le m$,*

$$\mathbb{P}\big(U_{t+1} \ge u_{t+1} \mid U_i = u_i,\ \forall i \in \{1, \ldots, t\}\big) = p^{-(n-t)(u_{t+1} - u_t)}. \tag{2.3}$$

In particular, if additionally $u_{t+1} < m$,

$$\mathbb{P}\big(U_{t+1} = u_{t+1} \mid U_i = u_i,\ \forall i \in \{1, \ldots, t\}\big) = p^{-(n-t)(u_{t+1} - u_t)}\big(1 - p^{-(n-t)}\big). \tag{2.4}$$

Proof. Two ordered sets of vectors $v_1, \ldots, v_t \in R^n$ and $w_1, \ldots, w_t \in R^n$ are said to be *equivalent* if $\mathrm{span}_R(p^{u_j} v_j : j \le i) = \mathrm{span}_R(p^{u_j} w_j : j \le i)$ for every $i \le t$.

Lemma 2.3 implies that the event $\{U_i = u_i : \forall i \le t\}$ can be written as a union of events of the form $\{\mathrm{span}_R(\mathbf{X}^{j-1} b \bmod p^m : j \le i) = \mathrm{span}_R(p^{u_j} v_j : j \le i),\ \forall i \le t\}$ indexed by vectors $v_1, \ldots, v_t \in R^n$ with $\mathrm{rank}_p(v_1, \ldots, v_t) = t$ and $\mathrm{span}_R(v_1) = \mathrm{span}_R(b \bmod p^m)$. Two such events are equal if the corresponding sets of vectors are equivalent and mutually exclusive otherwise. Hence, by conditioning on the equivalence class, (2.3) follows if we show that for every $v_1, \ldots, v_t \in R^n$ with $\mathrm{rank}_p(v_1, \ldots, v_t) = t$ and $\mathrm{span}_R(v_1) = \mathrm{span}_R(b \bmod p^m)$,

$$\begin{aligned}\mathbb{P}\big(U_{t+1} \ge u_{t+1} \mid \mathrm{span}_R\big(\mathbf{X}^{j-1} b \bmod p^m : j \le i\big) &= \mathrm{span}_R(p^{u_j} v_j : j \le i),\ \forall i \le t\big) \\ &= p^{-(n-t)(u_{t+1} - u_t)}.\end{aligned} \tag{2.5}$$

Fix $v_1, \ldots, v_t$ and let $E := \{\mathrm{span}_R\big(\mathbf{X}^{j-1} b \bmod p^m : j \le i\big) = \mathrm{span}_R(p^{u_j} v_j : j \le i),\ \forall i \le t\}$ denote the event in the condition of (2.5).

It follows from the definition that $U_{t+1} \ge u_{t+1}$ if and only if $\mathbf{X}(\mathrm{span}_{\mathbb{Z}}(\mathbf{X}^{i-1} b : i \le t)) \subseteq \mathrm{span}_{\mathbb{Z}}(\mathbf{X}^{i-1} b : i \le t) + p^{u_{t+1}}\mathbb{Z}^n$. Hence, conditional on E, one has $U_{t+1} \ge u_{t+1}$ if and only if $\mathbf{X}(p^{u_t} v_t) \in \mathrm{span}_R(p^{u_j} v_j : j \le t) + p^{u_{t+1}} R^n$. Considering that $\mathrm{rank}_p(v_1, \ldots, v_t) = t$ and that the u_i are nondecreasing, the latter occurs if and only if $\mathbf{X}(v_t) \in \mathrm{span}_R(v_1, \ldots, v_t) + p^{u_{t+1} - u_t} R^n$. Hence,

$$\mathbb{P}\big(U_{t+1} \ge u_{t+1} \mid E\big) = \mathbb{P}\big(\mathbf{X}(v_t) \in \mathrm{span}_R(v_1, \ldots, v_t) + p^{u_{t+1} - u_t} R^n \mid E\big). \tag{2.6}$$

The assumption $\mathrm{span}_R(v_1) = \mathrm{span}_R(b \bmod p^m)$ implies that the event $\mathrm{span}_R(\mathbf{X}^{j-1} b \bmod p^m : j \le 2) = \mathrm{span}_R(p^{u_j} v_1 : j \le 2)$ only depends on $\mathbf{X}(v_1)$. Similarly, continuing in an inductive fashion, the

https://doi.org/10.1017/S0963548324000312 Published online by Cambridge University Press

event E only depends on $\mathbf{X}(v_1), \ldots, \mathbf{X}(v_{t-1})$. Hence, by the law of total probability,

$$\begin{aligned}&\mathbb{P}\big(\mathbf{X}(v_t) \in \operatorname{span}_R(v_1, \ldots, v_t) + p^{u_{t+1}-u_t} R^n \mid E\big) \\ &\quad = \mathbb{E}\Big[\mathbb{P}\big(\mathbf{X}(v_t) \in \operatorname{span}_R(v_1, \ldots, v_t) + p^{u_{t+1}-u_t} R^n \mid \mathbf{X}(v_1), \ldots, \mathbf{X}(v_{t-1})\big) \mid E\Big].\end{aligned} \tag{2.7}$$

Recall that the entries of $\mathbf{X}$ are independent and $\operatorname{Unif}\{0, 1, \ldots, p^m - 1\}$-distributed. This implies that $\mathbf{X}$ induces a uniform random endomorphism of R^n. Hence, since it was assumed that $\operatorname{rank}_p(v_1, \ldots, v_t) = t$, it holds that $\mathbf{X}(v_t)$ has a uniform distribution on R^n and is independent of $\mathbf{X}(v_1), \ldots, \mathbf{X}(v_{t-1})$. Consequently, a counting argument yields that

$$\mathbb{P}\big(\mathbf{X}(v_t) \in \operatorname{span}_R(v_1, \ldots, v_t) + p^{u_{t+1}-u_t} R^n \mid \mathbf{X}(v_1), \ldots, \mathbf{X}(v_{t-1})\big) = p^{-(n-t)(u_{t+1}-u_t)}. \tag{2.8}$$

Combine (2.6)–(2.8) to establish (2.5). This proves (2.3). Further, (2.4) is an immediate consequence of (2.3) since $U_{t+1} = u_{t+1}$ if and only if $U_{t+1} \geq u_{t+1}$ and $U_{t+1} < u_{t+1} + 1$. □

2.3 *The law of* $\operatorname{coker}(\mathbf{W})_{p^m}$

It now only remains to combine Corollary 2.2 and Lemma 2.4. This allows us to also determine the law of $\operatorname{coker}(\mathbf{W})_{p^m}$ when n is finite:

Proposition 2.5. *Fix some $n \geq 1$, let $\mathbf{X}$ be a $\mathbb{Z}^{n\times n}$-valued random matrix with independent $\operatorname{Unif}\{0, 1, \ldots, p^m - 1\}$-distributed entries, and let $b \in \mathbb{Z}^n$ be a deterministic vector with $b \not\equiv 0 \bmod p$. Fix an integer $0 \leq i_0 \leq n - 1$.*

Pick integers $0 = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n \leq m$ and denote $\delta_i = \lambda_{n-i+1} - \lambda_{n-i}$. Assume that $\lambda_i < m$ if and only if $i \leq n - i_0$. Then, as Abelian groups,

$$\mathbb{P}\left(\operatorname{coker}(\mathbf{W})_{p^m} \cong \bigoplus_{i=1}^{n} \frac{\mathbb{Z}}{p^{\lambda_i}\mathbb{Z}}\right) = \prod_{i=i_0}^{n-2} \left(1 - p^{-(i+1)}\right) \prod_{j=1}^{n-1} p^{-j\delta_j}.$$

Proof. By Corollary 2.2 and the assumption that $\lambda_i < m$ for every $i \leq n - i_0$ and $\lambda_i = m$ for every $i > n - i_0$,

$$\begin{aligned}\mathbb{P}\big(\operatorname{coker}(\mathbf{W})_{p^m} \cong \oplus_{i=1}^{n} \mathbb{Z}/p^{\lambda_i}\mathbb{Z}\big) &= \mathbb{P}\big(U_i = \lambda_i,\ \forall i \in \{1, \ldots, n - i_0\}\big) \\ &\quad \times \mathbb{P}\big(U_i \geq m,\ \forall i \in \{n - i_0 + 1, \ldots, n\} \mid U_i = \lambda_i,\ \forall i \in \{1, \ldots, n - i_0\}\big).\end{aligned} \tag{2.9}$$

Recall that $\delta_i = \lambda_{n-i+1} - \lambda_{n-i}$. Hence, using that $U_1 = 0$ together with (2.4) from Lemma 2.4, which is applicable due to the assumption that $\lambda_i < m$ for every $i \leq n - i_0$,

$$\begin{aligned}\mathbb{P}\big(U_i = \lambda_i,\ \forall i \in \{1, \ldots, n - i_0\}\big) &= \prod_{i=i_0}^{n-2} \mathbb{P}(U_{n-i} = U_{n-i-1} + \delta_{i+1} \mid U_j = \lambda_j,\ \forall j < n - i) \\ &= \prod_{i=i_0}^{n-2} \left(1 - p^{-(i+1)}\right) p^{-(i+1)\delta_{i+1}}.\end{aligned} \tag{2.10}$$

If $i_0 = 0$, then the second probability in (2.9) is equal to one since there is no i satisfying $n + 1 \leq i \leq n$. In this case, the combination of (2.9) and (2.10) concludes the proof.

Now suppose that $i_0 > 0$. Then, it holds that $U_i \geq m$ for all $i > n - i_0$ if and only if $U_{n-i_0+1} \geq m$. Hence, using (2.3) from Lemma 2.4 and recalling that $m = \lambda_{n-i_0} + \delta_{i_0}$,

$$\begin{aligned}&\mathbb{P}\big(U_i \geq m,\ \forall i \in \{n - i_0 + 1, \ldots, n\} \mid U_i = \lambda_i,\ \forall i \in \{1, \ldots, n - i_0\}\big) \\ &\qquad = \mathbb{P}(U_{n-i_0+1} \geq m \mid U_i = \lambda_i,\ \forall i \in \{1, \ldots, n - i_0\}\big) \\ &\qquad = p^{-i_0\delta_{i_0}}.\end{aligned} \tag{2.11}$$

https://doi.org/10.1017/S0963548324000312 Published online by Cambridge University Press

Remark $p^{-i_0\delta_{i_0}} = \prod_{i=1}^{i_0} p^{-i\delta_i}$ since the assumption that $\lambda_i = m = \lambda_{i+1}$ for all $i > n - i_0$ ensures that $\delta_i = 0$ for every $i < i_0$. The combination of (2.9)–(2.11) hence concludes the proof. □

Proof of Theorem 1.2. Let $\tilde{\lambda}_i := 0$ for $i \in \{1, \dots, n-\ell\}$ and let $\tilde{\lambda}_i := \lambda_{i-(n-\ell)}$ for $i \geq n - \ell + 1$. The result then follows by considering the probability that $\mathrm{coker}(\mathbf{W})_{p^m} \cong \oplus_{i=1}^n \mathbb{Z}/p^{\tilde{\lambda}_i}\mathbb{Z}$ in Proposition 2.5 and taking the limit $n \to \infty$. □

3. Proof of Theorem 1.3

We establish a more general result than Theorem 1.3 and study the limiting law of the $\mathbb{Z}[x]$-module $\mathrm{coker}(\tilde{\mathbf{W}})$ where $\tilde{\mathbf{W}} := (\mathbf{X}^{j-1}\mathbf{B})_{j=1}^n$ is the $n \times nk$ matrix associated to a deterministic $n \times k$ matrix $\mathbf{B}$ for some fixed k. For future reference, let us here state all relevant assumptions:

(A1) For every $n \geq 1$, let $\mathbf{X}$ be a $\mathbb{Z}^{n\times n}$-valued random matrix with independent entries such that each entry is α_n-balanced mod $\mathscr{P}$.

(A2) Fix some $k \geq 1$. For every $n \geq k$, let $\mathbf{B} \in \mathbb{Z}^{n\times k}$ be a deterministic matrix such that $\mathbf{B}$ mod p has rank k over $\mathbb{F}_p$ for every $p \in \mathscr{P}$. We denote $\tilde{\mathbf{W}} := (\mathbf{X}^{j-1}\mathbf{B})_{j=1}^n$ and write $\mathrm{coker}(\tilde{\mathbf{W}}) := \mathbb{Z}^n/\tilde{\mathbf{W}}(\mathbb{Z}^{nk})$.

(A3) Assume that $\lim_{n\to\infty} n\alpha_n/\ln(n) = \infty$.

The desired result, describing the limiting distribution of $\mathrm{coker}(\tilde{\mathbf{W}})_{p^m,Q}$ under these assumptions, is given in Proposition 3.13.

As was outlined in Section 1.2, we rely on the category-theoretic moment method. The main ingredient required for the proof is correspondingly an estimate on the moments of $\mathrm{coker}(\tilde{\mathbf{W}})$:

Proposition 3.1. *Adopt assumptions* (A1) *to* (A3). *Then, for every finite $\mathbb{Z}[x]$-module N such that all prime divisors of $\#N$ are in $\mathscr{P}$,*

$$\lim_{n\to\infty} \mathbb{E}[\#\,\mathrm{Sur}_{\mathbb{Z}[x]}(\mathrm{coker}(\tilde{\mathbf{W}}), N)] = (\#N)^{-k}. \tag{3.1}$$

We prove Proposition 3.1 in Section 3.1 and then use a general-purpose result of Sawin and Wood [20, Lemma 6.3] to solve the associated moment problem in Section 3.2.

Remark 3.2. The assumption in Proposition 3.1 that all prime divisors of $\#N$ are in $\mathscr{P}$ cannot be removed. Indeed, recall that (A1) and (A2) only make assumptions regarding $\mathrm{rank}_p(\mathbf{B})$ and the balanced nature of the entries of $\mathbf{X}$ at primes $p \in \mathscr{P}$.

So, for instance, at $p \notin \mathscr{P}$ it could occur that $\mathbb{P}(\mathbf{X} \equiv 0 \bmod p) = 1$ and $\mathbf{B} = 0 \bmod p$ in which case $\mathbb{E}[\#\mathrm{Sur}_{\mathbb{Z}[x]}(\mathrm{coker}(\tilde{\mathbf{W}}), N)] = p^n - 1$ with $N = \mathbb{F}_p[x]/x\mathbb{F}_p[x]$. In particular, it then holds that $\lim_{n\to\infty} \mathbb{E}[\#\mathrm{Sur}_{\mathbb{Z}[x]}(\mathrm{coker}(\tilde{\mathbf{W}}), N)] = \infty$, which is incompatible with the conclusion of Proposition 3.1.

Remark 3.3. There is a sense in which $\mathrm{coker}(\tilde{\mathbf{W}})$ is a fairly natural random algebraic object to study. Note that $\mathbf{W}(\mathbb{Z}^{nk})$ is exactly the $\mathbb{Z}[x]$-submodule of $\mathbb{Z}^n$ generated by the columns of $\mathbf{B}$. Hence, introducing formal symbols $e_1, \dots, e_n$, we have

$$\mathrm{coker}(\tilde{\mathbf{W}}) \cong \left(e_1, \dots, e_n : xe_j = \sum_{i=1}^n \mathbf{X}_{i,j} e_i,\ \sum_{i=1}^n \mathbf{B}_{i,r} e_i = 0,\ \forall i \leq n,\ \forall r \leq k \right) \tag{3.2}$$

as $\mathbb{Z}[x]$-modules. So, $\mathrm{coker}(\tilde{\mathbf{W}})$ corresponds to the finitely presented $\mathbb{Z}[x]$-module, which is found when one considers n generators, imposes a random action for x specified by $\mathbf{X}$, and imposes $k \geq 1$ additional deterministic constraints specified by $\mathbf{B}$.

https://doi.org/10.1017/S0963548324000312 Published online by Cambridge University Press

3.1 Computing the limiting N-moments

Let N be a deterministic $\mathbb{Z}[x]$-module such that all prime divisors of $\#N$ are in $\mathscr{P}$. We may consider $\mathbf{X}$ as a random element of $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}^n, \mathbb{Z}^n)$, consider $\mathbf{B}$ as a deterministic element of $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}^k, \mathbb{Z}^n)$, and consider x as inducing an element of $\mathrm{Hom}_{\mathbb{Z}}(N, N)$. Here, note that $\mathrm{Hom}_{\mathbb{Z}}(\cdot, \cdot)$ simply returns the set of group morphisms since a $\mathbb{Z}$-module and an Abelian group are the same thing.

Now observe that a morphism of Abelian groups $F : \mathbb{Z}^n \to N$ descends to a morphism of $\mathbb{Z}[x]$-modules $\overline{F} : \mathrm{coker}(\tilde{\mathbf{W}}) \to N$ if and only if the compositions of F with $\mathbf{B}$, $\mathbf{X}$, and x satisfy $F\mathbf{B} = 0$ and $F\mathbf{X} = xF$. Moreover, every $\mathbb{Z}[x]$-module morphism $\overline{F} : \mathrm{coker}(\tilde{\mathbf{W}}) \to N$ arises from some unique $F : \mathbb{Z}^n \to N$ in this fashion. Consequently, since surjectivity is conserved,

$$\mathbb{E}[\#\mathrm{Sur}_{\mathbb{Z}[x]}(\mathrm{coker}(\tilde{\mathbf{W}}), N)] = \sum_{F \in \mathrm{Sur}_{\mathbb{Z}}(\mathbb{Z}^n, N) \,:\, F\mathbf{B} = 0} \mathbb{P}(F\mathbf{X} = xF). \tag{3.3}$$

Let us here emphasise that, while (3.3) was relatively easy to prove, the simplification which this step offers is significant. Indeed, observe that the joint law of the entries of $\tilde{\mathbf{W}}$ is not easy to understand since these entries are nontrivial algebraic combinations of the entries of $\mathbf{X}$ and $\mathbf{B}$. On the other hand, $F\mathbf{X} = xF$ is a linear equation in terms of $\mathbf{X}$ and hence fairly explicit.

The strategy that we use to estimate the N-moments from here on is as follows. We show that there are approximately $(\#N)^{n-k}$ surjections $F : \mathbb{Z}^n \to N$ with $F\mathbf{B} = 0$ in Section 3.1.1. Subsequently, we show that $\mathbb{P}(F\mathbf{X} = xF) \approx (\#N)^{-n}$ for most terms in (3.3) in Section 3.1.2, and we show that the remaining terms give a negligible contribution in Section 3.1.3. Finally, we combine these ingredients to conclude the proof of Proposition 3.1 in Section 3.1.4.

3.1.1. Estimate on the number of surjections satisfying $F\mathbf{B} = 0$

The *exponent* of a finite Abelian group G is the smallest positive integer $\exp(G) \geq 1$ such that $\exp(G)G = 0$. Note that $p \mid \exp(G)$ if and only if $p \mid \#G$.

Lemma 3.4. *Let G be a finite Abelian group and let $\mathbf{B} \in \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}^k, \mathbb{Z}^n)$ be such that $\mathrm{rank}_p(\mathbf{B}) = k$ for every prime divisor p of $\exp(G)$. Then, there exists a constant $C > 0$ depending only on G such that for all $n \geq k$*

$$|\#\{F \in \mathrm{Sur}_{\mathbb{Z}}(\mathbb{Z}^n, G) : F\mathbf{B} = 0\} - (\#G)^{n-k}| \leq C\Big(\frac{\#G}{2}\Big)^n.$$

Proof. We first argue that we can replace $\mathrm{Sur}_{\mathbb{Z}}(\mathbb{Z}^n, G)$ by $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}^n, G)$ up to a negligible error. If a morphism $F : \mathbb{Z}^n \to G$ is not surjective then there exists some proper subgroup $H \subsetneq G$ such that $F(\mathbb{Z}^n) = H$. Hence, since $\#\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}^n, H) = (\#H)^n$,

$$\#(\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}^n, G) \setminus \mathrm{Sur}_{\mathbb{Z}}(\mathbb{Z}^n, G)) \leq \sum_{H \subsetneq G} \#\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}^n, H) = \sum_{H \subsetneq G} (\#H)^n. \tag{3.4}$$

Denote S_G for the number of proper subgroups of G. Then, since $\#H \leq \#G/2$ by H being a proper subgroup,

$$\begin{aligned} &|\#\{F \in \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}^n, G) : F\mathbf{B} = 0\} - \#\{F \in \mathrm{Sur}_{\mathbb{Z}}(\mathbb{Z}^n, G) : F\mathbf{B} = 0\}| \\ &\qquad \leq \#(\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}^n, G) \setminus \mathrm{Sur}_{\mathbb{Z}}(\mathbb{Z}^n, G)) \leq S_G\Big(\frac{\#G}{2}\Big)^n. \end{aligned} \tag{3.5}$$

We next argue that $\#\{F \in \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}^n, G) : F\mathbf{B} = 0\} = (\#G)^{n-k}$. (This is not immediate because the columns $b_1, \ldots, b_k$ of $\mathbf{B}$ are not necessarily part of a $\mathbb{Z}$-module basis for $\mathbb{Z}^n$.)

Let $\mathbf{B} = \mathbf{UDV}$ be the Smith normal form of $\mathbf{B}$. This means that $\mathbf{U} \in \mathbb{Z}^{n \times n}$ and $\mathbf{V} \in \mathbb{Z}^{k \times k}$ are matrices with $\det(\mathbf{U}), \det(\mathbf{V}) \in \{-1, 1\}$ and $\mathbf{D}$ is an $n \times k$ diagonal matrix with integer diagonal

https://doi.org/10.1017/S0963548324000312 Published online by Cambridge University Press

entries satisfying $d_1 \mid d_2 \mid \ldots \mid d_k$. For brevity, denote $a := \exp(G)$. For every $p \mid a$, the assumption that $\mathrm{rank}_p(\mathbf{B}) = k$ implies that $d_i \not\equiv 0 \bmod p$. It follows that the d_i are multiplicative units for $\mathbb{Z}/a\mathbb{Z}$. Hence, the matrix $\mathbf{D}' := \mathrm{diag}(d_1, \ldots, d_k)$ is invertible in $(\mathbb{Z}/a\mathbb{Z})^{k\times k}$. Further, note that $\mathbf{U}$ and $\mathbf{V}$ are invertible over $\mathbb{Z}$. Hence, if $u_1, \ldots, u_n$ are the columns of $\mathbf{U}$, then the reductions to $(\mathbb{Z}/a\mathbb{Z})^n$ determine a $(\mathbb{Z}/a\mathbb{Z})$-module basis, and the reduction of $\mathbf{D}'\mathbf{V}$ to $(\mathbb{Z}/a\mathbb{Z})^{k\times k}$ is invertible. Consequently, since $\mathbf{B} = (u_1, \ldots, u_k)\mathbf{D}'\mathbf{V}$, the reductions of $b_1, \ldots, b_k$ together with the reductions of the u_i with $i \geq k+1$ determine a $(\mathbb{Z}/a\mathbb{Z})$-module basis for $(\mathbb{Z}/a\mathbb{Z})^n$.

Denote $\pi : \mathbb{Z}^n \to (\mathbb{Z}/a\mathbb{Z})^n$ for the reduction map. Then, since $a = \exp(G)$, it holds for every $F \in \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}^n, G)$ that there is some unique $\overline{F} \in \mathrm{Hom}_{\mathbb{Z}/a\mathbb{Z}}((\mathbb{Z}/a\mathbb{Z})^n, G)$ with $F = \overline{F} \circ \pi$. Recall that for any ring R an R-module morphism from a free R-module to an arbitrary R-module may be specified uniquely by arbitrarily specifying the images of the basis elements. Consequently, since the $\pi(b_i)$ are part of a $(\mathbb{Z}/a\mathbb{Z})$-module basis,

$$\begin{aligned} \#\left\{F \in \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}^n, G) : F\mathbf{B} = 0\right\} &= \#\left\{\overline{F} \in \mathrm{Hom}_{\mathbb{Z}/a\mathbb{Z}}((\mathbb{Z}/a\mathbb{Z})^n, G) : \overline{F} \circ \pi \circ \mathbf{B} = 0\right\} \\ &= (\#G)^{n-k}. \end{aligned} \tag{3.6}$$

Combine (3.5) with (3.6) and set $C := S_G$ to conclude the proof. □

We next estimate $\mathbb{P}(F\mathbf{X} = xF)$. The quality of the estimates will be better when F is 'very surjective'. To make this precise, we rely on a notion of *codes* which is due to Wood [32] and a notion of *robust morphisms* which is due to Nguyen and Wood [15].

3.1.2. Estimate for codes

Let $e_1, \ldots, e_n \in \mathbb{Z}^n$ be the standard basis vectors. For any $\sigma \subseteq \{1, \ldots, n\}$, write $V_\sigma := \mathrm{span}_{\mathbb{Z}}(e_i : i \in \sigma)$ for the $\mathbb{Z}$-submodule of $\mathbb{Z}^n$ consisting of vectors whose nonzero coordinates are in σ. We abbreviate $V_{\setminus\sigma} := V_{\{1,\ldots,n\}\setminus\sigma}$.

Definition 3.5. *Let G be an Abelian group. Then, $F \in \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}^n, G)$ is called a code of distance $w \geq 1$ if for every $\sigma \subseteq \{1, \ldots, n\}$ with $\#\sigma < w$ one has $F(V_{\setminus\sigma}) = G$.*

The foregoing definition may also be applied to $\mathbb{Z}[x]$-modules since these can be viewed as Abelian groups through the $\mathbb{Z}$-module structure.

Lemma 3.6. *Adopt assumptions (A1) and (A3), and fix a $\mathbb{Z}[x]$-module N such that all prime divisors of $\#N$ are in $\mathscr{P}$. Then, for every $\delta > 0$, there exist constants $C, c > 0$ such that for all $n \geq 1$*

$$\sum_{\substack{F \in \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}^n, N) \\ F \text{ a code of distance } \delta n}} \left|\mathbb{P}(F\mathbf{X} = xF) - (\#N)^{-n}\right| \leq Cn^{-c}. \tag{3.7}$$

Proof. For any code F and any $A \in \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}^n, N)$, it is shown in [15, Lemma 4.7] that $|\mathbb{P}(F\mathbf{X} = A) - (\#N)^{-n}| \leq C(\#N)^{-n}n^{-c}$. Actually, strictly speaking, [15, Lemma 4.7] is stated for matrices that are α_n-balanced at all primes, but only balancedness at primes dividing $\#N$ is necessary for its proof. (Indeed, [15, Lemma 4.7] follows from [15, Lemma 4.5] whose proof may be found in [33, Lemma 2.1] and only requires the weaker condition; see [33, Definition 1].) The result now follows since there are at most $\#\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}^n, N) = (\#N)^n$ summands in (3.7). □

3.1.3. Estimate for non-codes

The contribution of terms in (3.3) corresponding to non-codes turns out to be negligible. It is however delicate to make this rigorous. The estimate that can be achieved on $\mathbb{P}(F\mathbf{X} = xF)$ for a generic non-code F is, namely, insufficient to beat the combinatorial factor corresponding to the

https://doi.org/10.1017/S0963548324000312 Published online by Cambridge University Press

number of non-codes. Hence, a subdivision of the non-codes is required to balance the quality of the estimates against the combinatorial costs.

For an integer d with prime factorisation $d = \prod_i p_i^{e_i}$, denote $\ell(d) := \sum_i e_i$. Given a subgroup $H \subseteq G$, let $[G:H] := \#G/\#H$ denote the index of H in G.

Definition 3.7. *Let G be a finite Abelian group and let $\delta > 0$ be a scalar. Then, $F \in \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}^n, G)$ is called δ-robust for a subgroup $H \subseteq G$ if H is minimal with the property that*

$$\#\big\{i \in \{1,\dots,n\} : F(e_i) \notin H\big\} \le \ell([G:H])\delta n \tag{3.8}$$

That is, H satisfies (3.8), and no strict subgroup $H' \subsetneq H$ satisfies (3.8).

The main motivation for Definition 3.7 is the following property: if F is δ-robust for H, then the restriction of F to V_σ with $\sigma := \{i : F(e_i) \in H\}$ is a code of distance δn when H is viewed as the codomain of this restriction. Indeed, suppose this were not the case. Then, there exists $\mu \subseteq \sigma$ with $\#\mu < \delta n$ such that $H' := F(V_{\sigma\setminus\mu})$ is a strict subgroup of H. So, since $[G:H'] \ge [G:H] + 1$,

$$\begin{aligned}\#\{i \in \{1,\dots,n\} : F(e_i) \notin H'\} &\le \#\{i \in \{1,\dots,n\} : F(e_i) \notin H\} + \#\mu \\ &\le \ell([G:H'])\delta n\end{aligned} \tag{3.9}$$

contradicting the minimality of H.

In particular, any $F \in \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}^n, G)$ that is not a code of distance δn is not δ-robust for G. However, (3.8) is always satisfied when $G = H$. This implies that any non-code has to be δ-robust for some, not necessarily unique, proper subgroup of G. Hence,

$$\begin{aligned}&\{F \in \mathrm{Sur}_{\mathbb{Z}}(\mathbb{Z}^n, G) : F \text{ not a code of distance } \delta n\} \\ &\qquad \subseteq \bigcup_{H \subsetneq G} \{F \in \mathrm{Sur}_{\mathbb{Z}}(\mathbb{Z}^n, G) : F \text{ is } \delta\text{-robust for } H\}.\end{aligned} \tag{3.10}$$

We next establish an estimate on $\mathbb{P}(F\mathbf{X} = xF)$ when F is δ-robust for some H. The following lemma generalises [15, Lemma 4.11], which concerns a similar bound for $\mathbb{P}(F(Y) = 0)$.

Lemma 3.8. *Fix scalars $\delta, \alpha > 0$, an integer $n \ge 1$, and a finite Abelian group G. Fix a proper subgroup $H \subsetneq G$, denote $d := [G:H]$, and consider a maximal chain of proper subgroups*

$$H = G_{\ell(d)} \subsetneq \cdots \subsetneq G_2 \subsetneq G_1 \subsetneq G_0 = G. \tag{3.11}$$

Consider a δ-robust morphism $F \in \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}^n, G)$ for H. For every $1 \le j \le \ell(d)$ denote $p_j := [G_{j-1} : G_j]$ and

$$w_j := \#\big\{i \in \{1,\dots,n\} : F(e_i) \in G_{j-1} \setminus G_j\big\}. \tag{3.12}$$

Abbreviate $a := \exp(G)$ for the exponent of G. Then, for every $g \in G$ and every $\mathbb{Z}^n$-valued random vector Y whose entries are independent and α-balanced modulo all prime divisors of a,

$$\mathbb{P}\big(F(Y) = g\big) \le \big((\#G)^{-1} d + \exp\big(-\alpha\delta n/a^2\big)\big) \prod_{j=1}^{\ell(d)} \left(p_j^{-1} + \frac{p_j - 1}{p_j} \exp\left(-\alpha w_j / p_j^2\right)\right). \tag{3.13}$$

Proof. The strategy in this proof is to reduce to the case of codes where estimates are available from [15, Lemma 4.5]. Again, similar to the remarks in the proof of Lemma 3.6, that lemma is stated for the case where Y is balanced at all primes, but the case where Y is merely balanced at prime divisors of a follows from its proof.

For every $j \in \{1, 2, \dots, \ell(d)\}$, define a set of indices by

$$\sigma_j := \big\{i \in \{1,\dots,n\} : F(e_i) \in G_{j-1} \setminus G_j\big\}. \tag{3.14}$$

Then, for every $r \le \ell(d)$, the set of indices i with $F(e_i) \notin G_r$ is given by $\Sigma_r := \cup_{j=1}^r \sigma_j$. Write $Y = (y_1, \dots, y_n)$ and observe that $\sum_{i \notin \Sigma_r} y_i F(e_i) \in G_r$ for any $r \le \ell(d)$. Consequently, since $F(Y) =$

https://doi.org/10.1017/S0963548324000312 Published online by Cambridge University Press

$\sum_{i=1}^{n} y_i F(e_i)$, it is only possible to have $F(Y) = g$ if $\sum_{i\in\Sigma_r} y_i F(e_i) - g \in G_r$ for all $r \le \ell(d)$. Hence, by definition of conditional probability,

$$\begin{aligned}\mathbb{P}\left(F(Y)=g\right) &= \mathbb{P}\left(\sum_{i\in\Sigma_1} y_i F(y_i) - g \in G_1\right)\mathbb{P}\left(F(Y)=g \,\middle|\, \sum_{i\in\Sigma_1} y_i F(y_i) - g \in G_1\right)\\ &= \prod_{j=1}^{\ell(d)} \mathbb{P}\left(\sum_{i\in\Sigma_j} y_i F(y_i) - g \in G_j \,\middle|\, \forall r<j\colon \sum_{i\in\Sigma_r} y_i F(y_i) - g \in G_r\right)\\ &\quad\times \mathbb{P}\left(F(Y)=g \,\middle|\, \forall r \le \ell(d)\colon \sum_{i\in\Sigma_r} y_i F(y_i) - g \in G_r\right).\end{aligned} \tag{3.15}$$

We next bound the probabilities occurring in (3.15). Recall that the y_i are independent. Hence, if we fix some $j \le \ell(d)$ and condition on the values achieved by the y_i with $i \in \Sigma_{j-1}$, then

$$\begin{aligned}&\mathbb{P}\left(\sum_{i\in\Sigma_j} y_i F(e_i) - g \in G_j \,\middle|\, \forall r<j\colon \sum_{i\in\Sigma_r} y_i F(y_i) - g \in G_r\right)\\ &= \mathbb{E}\left[\mathbb{P}\Big(\sum_{i\in\Sigma_j} y_i F(e_i) - g \in G_j \,\Big|\, y_i : i \in \Sigma_{j-1}\Big) \,\middle|\, \forall r<j\colon \sum_{i\in\Sigma_r} y_i F(y_i) - g \in G_r\right]\\ &\le \max_{h\in G_{j-1}} \mathbb{P}\left(\sum_{i\in\sigma_j} y_i F(e_i) - h \in G_j\right).\end{aligned} \tag{3.16}$$

Here, the final step used that $\Sigma_j \setminus \Sigma_{j-1} = \sigma_j$, and $\sum_{i\in\Sigma_{j-1}} y_i F(e_i) - g$ was identified with h. Denote $F_j : V_{\sigma_j} \to G_{j-1}/G_j$ for the map found by restricting F to V_{σ_j} and reducing modulo G_j. Recall (3.12) and note that $w_j = \#\sigma_j$. We claim that F_j is a code of distance w_j; recall Definition 3.5. Indeed, the maximality of (3.11) ensures that G_{j-1}/G_j is a cyclic group of prime order and consequently, for every $i \in \sigma_j$, $F_j(e_i)$ generates G_{j-1}/G_j since $F_j(e_i) \not\equiv 0 \bmod G_j$ by definition of σ_j; recall (3.14). Now apply [15, Lemma 4.5] to F_j and use that $p_j = \#(G_{j-1}/G_j)$ to find

$$\mathbb{P}\left(\sum_{i\in\sigma_j} y_i F(e_i) - h \in G_j\right) \le \frac{1}{p_j} + \frac{p_j - 1}{p_j}\exp\left(-\frac{\alpha w_j}{p_j^2}\right). \tag{3.17}$$

Using this bound on the product in (3.15) yields the product in (3.13). It remains to bound the remaining factor. Here, similarly to (3.16), we have

$$\mathbb{P}\left(F(Y)=g \,\middle|\, \forall r \le \ell(d)\colon \sum_{i\in\Sigma_r} y_i F(y_i) - g \in G_r\right) \le \max_{h\in G_{\ell(d)}} \mathbb{P}\left(\sum_{i\notin\Sigma_{\ell(d)}} y_i F(e_i) = h\right).$$

By the argument preceding (3.9), the restriction of F to $V_{\setminus\Sigma_{\ell(d)}}$ defines a code of distance δn. Hence, by [15, Lemma 4.5] and the fact that $\exp(H) \mid \exp(G)$,

$$\mathbb{P}\left(\sum_{i\notin\Sigma_{\ell(d)}} y_i F(e_i) = h\right) \le (\#H)^{-1} + \exp\left(-\alpha\delta n/a^2\right). \tag{3.18}$$

It was here used that $(\#H - 1)/\#H \le 1$. Use that $\#H = \#G/d$ to conclude the proof. □

https://doi.org/10.1017/S0963548324000312 Published online by Cambridge University Press

Corollary 3.9. *Adopt assumption (A1), and let N be a $\mathbb{Z}[x]$-module such that all prime divisors of $\#N$ are in $\mathscr{P}$. Then, for every subgroup $H \subseteq N$ and every $F \in \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}^n, N)$, which is δ-robust for H,*

$$\mathbb{P}(F\mathbf{X} = xF) \leq \left((\#N)^{-1}d + \exp\left(-\alpha_n \delta n/a^2\right)\right)^n \prod_{j=1}^{\ell(d)} \left(p_j^{-1} + \frac{p_j - 1}{p_j} \exp\left(-\alpha_n w_j/p_j^2\right)\right)^n$$

where d, p_j, w_j, and a are defined as in Lemma 3.8 with $G = N$.

Proof. Since the entries of $\mathbf{X}$ are assumed to be independent, one has that $\mathbb{P}(F\mathbf{X} = xF) = \prod_{i=1}^n \mathbb{P}(F(\mathbf{X}e_i) = xF(e_i))$. The result is hence immediate from Lemma 3.8 applied with $Y := \mathbf{X}e_i$ and $g := xF(e_i)$. □

Lemma 3.10. *Let N be a $\mathbb{Z}[x]$-module such that all prime divisors of $\#N$ are in $\mathscr{P}$ and adopt assumptions 1 and 3. Then, there exists $\delta_0 > 0$ such that for every $\delta < \delta_0$, there exist constants $C, c > 0$ such that for all $n \geq 1$*

$$\sum_{\substack{F \in \mathrm{Sur}_{\mathbb{Z}}(\mathbb{Z}^n, N) \\ F \text{ not a code of distance } \delta n}} \mathbb{P}(F\mathbf{X} = xF) \leq Cn^{-c}.$$

Proof. By (3.10), one may upper bound the sum as

$$\sum_{\substack{F \in \mathrm{Sur}_{\mathbb{Z}}(\mathbb{Z}^n, N) \\ F \text{ not a code of distance } \delta n}} \mathbb{P}(F\mathbf{X} = xF) \leq \sum_{H \subsetneq N} \sum_{\substack{F \in \mathrm{Sur}_{\mathbb{Z}}(\mathbb{Z}^n, N) \\ F \text{ is } \delta\text{-robust for } H}} \mathbb{P}(F\mathbf{X} = xF). \tag{3.19}$$

Fix some proper subgroup $H \subsetneq N$ and pick a maximal chain of subgroups G_j as in (3.11) with $G = N$. We denote $d := [N : H]$.

By [15, Lemma 4.10], the number of $F \in \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}^n, N)$ that are δ-robust for H and satisfy that there are exactly w_j indices $i \leq n$ with $F(e_i) \in G_{j-1} \setminus G_j$ for $1 \leq j \leq \ell(d)$ is at most $(\#H)^{n - \sum_{j=1}^{\ell(d)} w_j} \prod_{j=1}^{\ell(d)} \binom{n}{w_j} (\#G_{j-1})^{w_j}$. When F is surjective, we have $w_1 \neq 0$. Further, when F is δ-robust for H, we have $w_j \leq \ell(d)\delta n$ for all $j \leq \ell(d)$. Indeed, if this were not the case, then we would have $\#\{i \leq n : F(e_i) \notin H\} > \ell(d)\delta n$, which contradicts (3.8). Now, by the combination of Corollary 3.9 with the aforementioned count on the number of δ-robust morphisms,

$$\begin{aligned} \sum_{\substack{F \in \mathrm{Sur}_{\mathbb{Z}}(\mathbb{Z}^n, N) \\ F \text{ is } \delta\text{-robust for } H}} \mathbb{P}(F\mathbf{X} = xF) \leq & \sum_{\substack{0 \leq w_1, \ldots, w_{\ell(d)} \leq \ell(d)\delta n \\ w_1 \neq 0}} (\#H)^{n - \sum_{j=1}^{\ell(d)} w_j} \prod_{j=1}^{\ell(d)} \binom{n}{w_j} (\#G_{j-1})^{w_j} \\ & \times \left((\#N)^{-1}d + \exp\left(-\alpha_n \delta n/a^2\right)\right)^n \prod_{j=1}^{\ell(d)} \left(p_j^{-1} + \frac{p_j - 1}{p_j} \exp\left(-\alpha_n w_j/p_j^2\right)\right)^n. \end{aligned} \tag{3.20}$$

It remains a nontrivial task to compute the right-hand side of (3.20). Fortunately, a related sum was considered by Nguyen and Wood [15], and we can extract the relevant estimate from their proofs. More precisely, the sum in (3.20) is a special case of the sum which occurs in the first centred equation of the proof of [15, Theorem 4.12]: take $u = 0$ in their notation. Following the arguments word for word up to the centred inequality at the end of page 23 in [15] now yields the desired result. □

https://doi.org/10.1017/S0963548324000312 Published online by Cambridge University Press

Lemma 3.11. *Adopt assumptions (A1) and (A3), and fix a $\mathbb{Z}[x]$-module N such that all prime divisors of $\#N$ are in $\mathscr{P}$. Then, there exists $\delta_0 > 0$ such that for every $\delta < \delta_0$, there exist constants $C, c > 0$ such that for all $n \geq 1$*

$$\sum_{\substack{F \in \mathrm{Sur}_{\mathbb{Z}}(\mathbb{Z}^n, N) \\ F \text{ not a code of distance } \delta n}} |\mathbb{P}(F\mathbf{X} = xF) - (\#N)^{-n}| \leq Cn^{-c}.$$

Proof. By Definition 3.5, if F is not a code of distance δn, then we can find some $\sigma \subseteq \{1, \ldots, n\}$ and a proper subgroup $H \subsetneq N$ such that $F(V_{\setminus\sigma}) \subseteq H$ and $\#\sigma = \lfloor \delta n \rfloor$. Hence, since F is uniquely determined by the images of the basis elements of $\mathbb{Z}^n$,

$$\sum_{\substack{F \in \mathrm{Sur}_{\mathbb{Z}}(\mathbb{Z}^n, N) \\ F \text{ not a code of distance } \delta n}} (\#N)^{-n} \leq \sum_{H \subsetneq N} \binom{n}{\lfloor \delta n \rfloor} (\#H)^{n - \lfloor \delta n \rfloor} (\#N)^{-n + \lfloor \delta n \rfloor}. \tag{3.21}$$

The binary entropy bound [6, Eq.(7.14), p. 151] implies that $\binom{n}{\lfloor \delta n \rfloor} \leq 2^{nE(\lfloor \delta n \rfloor / n)}$ where $E(y) := -y \log_2(y) + (1-y)\log_2(1-y)$. Note that $\lim_{y \to 0} E(y) = 0$. We can hence find some sufficiently small δ_0 such that Lemma 3.10 is applicable and $\delta_0 + E(\lfloor \delta_0 n \rfloor / n) < 1/2$ for all $n \geq 1$. Then, since $\#H \leq \#N/2$, we have for every $\delta < \delta_0$ that

$$\sum_{\substack{F \in \mathrm{Sur}_{\mathbb{Z}}(\mathbb{Z}^n, N) \\ F \text{ not a code of distance } \delta n}} (\#N)^{-n} \leq S_N 2^{(E(\lfloor \delta n \rfloor / n) + \delta - 1)n} \leq S_N 2^{-n/2} \tag{3.22}$$

where S_N is the number of proper subgroups of N. Let $C', c' > 0$ be such that $S_N 2^{-n/2} \leq C' n^{-c'}$ for all $n \geq 1$, and use Lemma 3.10 together with the triangle inequality to conclude the proof. □

3.1.4. Combining the estimates

We finally combine all preceding estimates to complete the proof of Proposition 3.1.

Proof of Proposition 3.1. By (3.3) and the triangle inequality,

$$\begin{aligned} |\mathbb{E}[\#\mathrm{Sur}_{\mathbb{Z}[x]}(\mathrm{coker}(\tilde{\mathbf{W}}), N)] - (\#N)^{-k}| \leq & \sum_{F \in \mathrm{Sur}_{\mathbb{Z}}(\mathbb{Z}^n, N) : F\mathbf{B} = 0} \left| (\#N)^{-n} - \mathbb{P}(F\mathbf{X} = xF) \right| \\ & + \left| (\#N)^{-n} \#\{F \in \mathrm{Sur}_{\mathbb{Z}}(\mathbb{Z}^n, N) : F\mathbf{B} = 0\} - (\#N)^{-k} \right|. \end{aligned} \tag{3.23}$$

Pick some $\delta > 0$, which is sufficiently small to ensure that Lemma 3.11 is applicable. Then, by the triangle inequality,

$$\begin{aligned} & \sum_{F \in \mathrm{Sur}_{\mathbb{Z}}(\mathbb{Z}^n, N) : F\mathbf{B} = 0} |(\#N)^{-n} - \mathbb{P}(F\mathbf{X} = xF)| \\ & \leq \sum_{\substack{F \in \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}^n, N) \\ F \text{ a code of distance } \delta n}} |(\#N)^{-n} - \mathbb{P}(F\mathbf{X} = xF)| + \sum_{\substack{F \in \mathrm{Sur}_{\mathbb{Z}}(\mathbb{Z}^n, N) \\ F \text{ not a code of distance } \delta n}} |(\#N)^{-n} - \mathbb{P}(F\mathbf{X} = xF)|. \end{aligned} \tag{3.24}$$

Let $c, C >$ be the constants from Lemma 3.6, and let $c', C' > 0$ be the constants from Lemma 3.11. Then, the right-hand side of (3.24) is at most $Cn^{-c} + C'n^{-c'}$ and hence tends to zero as n tends to infinity. Further, by Lemma 3.4, there exists a constant $C'' > 0$ such that

$$\left| (\#N)^{-n} \#\{F \in \mathrm{Sur}_{\mathbb{Z}}(\mathbb{Z}^n, N) : F\mathbf{B} = 0\} - (\#N)^{-k} \right| \leq C'' 2^{-n}. \tag{3.25}$$

Remark that the right-hand side of (3.25) tends to zero as n tends to infinity to conclude the proof. □

https://doi.org/10.1017/S0963548324000312 Published online by Cambridge University Press

3.2 Solving the moment problem

We next apply a general result concerning measures on categories of [20, Theorem 1.6] to invert the moment problem. Using [20, Lemma 6.1] and [20, Lemma 6.3], that result may be specialised to our context – namely, to limiting measures on the category of finite modules with N-moments given by $(\#N)^{-k}$. Let us state this specialisation explicitly for the sake of definiteness:

Lemma 3.12 (Special case of [20, Theorem 1.6]). *Let R be a ring and consider a sequence of random finite R-modules X_n such that for every fixed finite R-module N*

$$\lim_{n\to\infty} \mathbb{E}[\# \operatorname{Sur}_R(X_n, N)] = (\#N)^{-k}.$$

Let S be quotient ring of R with $\#S < \infty$, and let $L_1, \dots, L_r$ be representatives of the isomorphism classes of finite simple S-modules. Further, denote q_i for the cardinality of the endomorphism field of L_i for every $i \le r$. Then, for every finite S-module N,

$$\lim_{n\to\infty} \mathbb{P}(X_n \otimes_R S \cong N) = \frac{1}{(\#N)^k \#\operatorname{Aut}_S(N)} \prod_{i=1}^{r} \prod_{j=1}^{\infty} \left(1 - \frac{q_i^{-j} \# \operatorname{Ext}^1_S(N, L_i)}{\#\operatorname{Hom}_S(N, L_i)(\#L_i)^k}\right).$$

Let us further remark that it follows from the statement of [20, Theorem 1.6] that the limiting measure in Lemma 3.12 has N-moments given by $(\#N)^{-k}$. In particular, the N-moment for $N = \{0\}$ is equal to one, which implies that the limiting measure is a probability measure.

It now remains to combine Proposition 3.1 with Lemma 3.12 and to simplify the results. This can be accomplished with a direct computation. Note that the following result implies Theorem 1.3 as the special case with $k = 1$.

Proposition 3.13. *Adopt assumptions* (A1) *to (A3). Fix a positive integer $m \ge 1$, a monic polynomial $Q \in \mathbb{Z}[x]$ of degree ≥ 1, and introduce $Q_{i,p} \in \mathbb{F}_p[x]$, $r_p \ge 1$, and $d_{i,p} = \deg Q_{i,p}$ using the factorisation of Q modulo p as in Theorem 1.3.*

For every $p \in \mathscr{P}$, denote $S_{p^m} := \mathbb{Z}[x]/(p^m\mathbb{Z}[x] + Q(x)\mathbb{Z}[x])$. Then, given a finite S_{p^m}-module $N_{p^m,Q}$ for every $p \in \mathscr{P}$, it holds that

$$\lim_{n\to\infty} \mathbb{P}(\forall p \in \mathscr{P} : \operatorname{coker}(\tilde{\mathbf{W}})_{p^m,Q} \cong N_{p^m,Q}) = \prod_{p\in\mathscr{P}} \left(\frac{1}{(\#N_{p^m,Q})^k \#\operatorname{Aut}_{S_{p^m}}(N_{p^m,Q})} \right.$$
$$\left. \times \prod_{i=1}^{r_p} \prod_{j=1}^{\infty} \left(1 - \frac{\# \operatorname{Ext}^1_{S_{p^m}}\left(N_{p^m,Q}, \mathbb{F}_p[x]/(Q_{i,p}(x)\mathbb{F}_p[x])\right)}{\#\operatorname{Hom}_{S_{p^m}}\left(N_{p^m,Q}, \mathbb{F}_p[x]/(Q_{i,p}(x)\mathbb{F}_p[x])\right)} p^{-(k+j)d_{i,p}}\right)\right).$$

Proof. Denote $R := \mathbb{Z}[x]/\left(\left(\prod_{p\in\mathscr{P}} p^m\right)\mathbb{Z}[x]\right)$ and note that for any $\mathbb{Z}[x]$-module M and any R-module N, one has a bijection between $\operatorname{Sur}_R(M \otimes_{\mathbb{Z}[x]} R, N)$ and $\operatorname{Sur}_{\mathbb{Z}[x]}(M, N)$. Consequently, by Proposition 3.1, we have that for every finite R-module N

$$\lim_{n\to\infty} \mathbb{E}[\#\operatorname{Sur}_R(\operatorname{coker}(\tilde{\mathbf{W}}) \otimes_{\mathbb{Z}[x]} R, N)] = (\#N)^{-k}. \tag{3.26}$$

Let $S := R/(Q(t)R)$ and define a finite S-module by $N := \oplus_{p\in\mathscr{P}} N_{p^m,Q}$. The Chinese remainder theorem yields $R \cong \oplus_{p\in\mathscr{P}} \mathbb{Z}[x]/(p^m\mathbb{Z}[x])$, which implies that $S = \oplus_{p\in\mathscr{P}} S_{p^m}$ and $\operatorname{coker}(\tilde{\mathbf{W}}) \otimes_{\mathbb{Z}[x]} S \cong \oplus_{p\in\mathscr{P}} \operatorname{coker}(\tilde{\mathbf{W}})_{p^m,Q}$. Consequently, by Lemma 3.12 and the fact that isomorphism occurs if and only if the corresponding summands are isomorphic,

https://doi.org/10.1017/S0963548324000312 Published online by Cambridge University Press

$$\lim_{n\to\infty} \mathbb{P}\big(\forall p \in \mathscr{P} : \operatorname{coker}(\tilde{\mathbf{W}})_{p^m,Q} \cong N_{p^m,Q}\big) = \lim_{n\to\infty} \mathbb{P}\big((\operatorname{coker}(\tilde{\mathbf{W}}) \otimes_{\mathbb{Z}[x]} R) \otimes_R S \cong N\big)$$
$$= \frac{1}{(\#N)^k \#\operatorname{Aut}_S(N)} \prod_{i=1}^{r} \prod_{j=1}^{\infty} \left(1 - \frac{\#\operatorname{Ext}^1_S(N, L_i)}{\#\operatorname{Hom}_S(N, L_i)(\#L_i)^k} q_i^{-j}\right). \tag{3.27}$$

Simple modules over S are precisely the modules of the form S/m with m a maximal ideal of S. Further, maximal ideals of $\mathbb{Z}[x]$ are of the form $m = p\mathbb{Z}[x] + f(x)\mathbb{Z}[x]$ with p a prime and $f(x) \in \mathbb{Z}[x]$ irreducible modulo p of degree ≥ 1 [19, p. 22]. Hence, since maximal ideals of S are in one-to-one correspondence with maximal ideals of $\mathbb{Z}[x]$, which contain $\prod_{p\in\mathscr{P}} p^m$ and Q, the modules L_i in (3.27) are of the form $\mathbb{F}_p[x]/(Q_{i,p}(x)\mathbb{F}_p[x])$. Consequently, by (3.27) and the fact that the endomorphism field of $\mathbb{F}_p[x]/(Q_{i,p}(x)\mathbb{F}_p[x]))$ is isomorphic to $\mathbb{F}_p[x]/(Q_{i,p}(x)\mathbb{F}_p[x])$, which is a finite field of order $p^{d_{i,p}}$,

$$\lim_{n\to\infty} \mathbb{P}\big(\forall p \in \mathscr{P} : \operatorname{coker}(\tilde{\mathbf{W}})_{p^m,Q} \cong N_{p^m,Q}\big) \tag{3.28}$$
$$= \frac{1}{(\#N)^k \#\operatorname{Aut}_S(N)} \prod_{p\in\mathscr{P}} \prod_{i=1}^{r_p} \prod_{j=1}^{\infty} \left(1 - \frac{\#\operatorname{Ext}^1_S\big(N, \mathbb{F}_p[x]/(Q_{i,p}(x)\mathbb{F}_p[x])\big)}{\#\operatorname{Hom}_S\big(N, \mathbb{F}_p[x]/(Q_{i,p}(x)\mathbb{F}_p[x])\big)} p^{-(k+j)d_{i,p}}\right).$$

Here, observe that $\#N = \prod_{p\in\mathscr{P}} \#N_{p^m,Q}$, $\#\operatorname{Aut}_S(N) = \prod_{p\in\mathscr{P}} \#\operatorname{Aut}_{S_{p^m}}(N_{p^m,Q})$, and note that $\#\operatorname{Hom}_S\big(N, \mathbb{F}_p[x]/(Q_{i,p}(x)\mathbb{F}_p[x])\big) = \#\operatorname{Hom}_{S_{p^m}}(N_{p^m,Q}, \mathbb{F}_p[x]/(Q_{i,p}(x)\mathbb{F}_p[x]))$. Further, since Ext takes direct sums in the first argument to products [31, Proposition 3.3.4],

$$\#\operatorname{Ext}^1_S\left(N, \mathbb{F}_p[x]/\left(Q_{i,p}(x)\mathbb{F}_p[x]\right)\right) = \prod_{q\in\mathscr{P}} \#\operatorname{Ext}^1_S\left(N_{q^m,Q}, \mathbb{F}_p[x]/\left(Q_{i,p}(x)\mathbb{F}_p[x]\right)\right). \tag{3.29}$$

Finally, using the definition of Ext^1 in terms of short exact sequences [31, Theorem 3.4.3] together with the fact that finite S-modules correspond to tuples of S_{p^m}-modules, one can verify that

$$\#\operatorname{Ext}^1_S(N_{q^m,Q}, \mathbb{F}_p[x]/(Q_{i,p}(x)\mathbb{F}_p[x])) = \begin{cases} \#\operatorname{Ext}^1_{S_{p^m}}(N_{p^m,Q}, \mathbb{F}_p[x]/(Q_{i,p}(x)\mathbb{F}_p[x])) & \text{if } p = q, \\ 1 & \text{else.} \end{cases} \tag{3.30}$$

Combine (3.28)–(3.30) to conclude the proof. □

4. Future work

Ultimately, we would like to show that the conditions of Theorem 1.1 are satisfied with nonvanishing probability. The current paper makes progress in this direction: we now have concrete proof techniques that can be used to study cokernels of walk matrices. There are however still a number of interesting open problems.

For instance, it remains entirely open to understand whether the condition $\operatorname{coker}(\mathbf{W})_{2^2} \cong (\mathbb{Z}/2\mathbb{Z})^{\lfloor n/2 \rfloor}$ is often satisfied, even heuristically. This is because the distribution of $\operatorname{coker}(\mathbf{W})_{2^m}$ is highly sensitive to the graph being simple, which makes approximation by results for directed graphs inadequate. Indeed, when $\mathbf{X}$ is the adjacency matrix of a simple graph and $b = e$ is the all-ones vector, then [26, Lemma 14] implies that $\operatorname{rank}_2 \mathbf{W} \leq \lceil n/2 \rceil$. Equivalently, it holds that $\operatorname{coker}(\mathbf{W})_2 \cong (\mathbb{Z}/2\mathbb{Z})^\ell$ for some $\ell \geq \lfloor n/2 \rfloor$. This is very different behaviour from the distribution for directed graphs in Theorem 1.2 with $p = 2$ since the latter is concentrated on small groups.

For odd primes, numerical evidence suggest that the difference is not as severe. Table 1, namely, suggest that the distribution of $\operatorname{coker}(\mathbf{W})_{p^m}$ has qualitatively similar behaviour for simple and directed graphs. Quantitatively, however, a close inspection shows that there is a small but nonzero difference that does not seem to disappear when n grows large, suggesting that the limiting distribution for simple graphs differs from the one for random directed and weighted graphs.

https://doi.org/10.1017/S0963548324000312 Published online by Cambridge University Press

Table 1. Probability that $\mathrm{coker}(\mathrm{W})_{p^2} \in \{0, \mathbb{Z}/p\mathbb{Z}\}$ when X is the adjacency matrix of an undirected Erdős–Rényi random graph on n nodes and $b = e$, estimated based on 10^5 independent samples. The estimated values have an uncertainty of ± 0.002. Also displayed is the limiting probability in the case of directed and weighted random graphs, which follows from Theorem 1.2 with $m = 2$. Computation of the group structure of $\mathrm{coker}(\mathrm{W})_{p^2}$ was done using the algorithm `smith_form` in `SageMath` [23]

p	$n = 10$	$n = 12$	$n = 15$	$n = 20$	$n = 30$	$n = 40$	Theorem 1.2
3	0.495	0.625	0.726	0.757	0.756	0.758	0.746834 . . .
5	0.549	0.725	0.869	0.913	0.914	0.915	0.912399 . . .
7	0.563	0.750	0.906	0.953	0.956	0.957	0.956337 . . .
11	0.571	0.765	0.930	0.981	0.983	0.983	0.982726 . . .

Table 2. Probability that $\mathrm{coker}(\mathrm{W})_{p^2} \in \{0, \mathbb{Z}/p\mathbb{Z}\}$ when $\mathrm{X} \sim \mathrm{Unif}\{0,1\}^{n \times n}$ is the adjacency matrix of an unweighted directed random graph and $b = e$. The same comments as in the caption of Table 1 apply: the estimation used 10^5 independent samples, there is an uncertainty of ± 0.002, and `SageMath` [23] was used

p	$n = 10$	$n = 12$	$n = 15$	$n = 20$	$n = 30$	$n = 40$	Theorem 1.2
3	0.650	0.707	0.737	0.746	0.749	0.747	0.746834 . . .
5	0.759	0.844	0.898	0.911	0.911	0.912	0.912399 . . .
7	0.786	0.881	0.940	0.956	0.957	0.956	0.956337 . . .
11	0.802	0.901	0.965	0.982	0.983	0.983	0.982726 . . .

For instance, the estimated probabilities that $\mathrm{coker}(\mathbf{W})_{p^2} \in \{0, \mathbb{Z}/p\mathbb{Z}\}$ for $p = 3$ is 0.758 ± 0.002, whereas Theorem 1.2 predicts 0.747.

So, the extension of our results to the setting of simple graphs poses an interesting problem, both for odd and even primes. We intend to pursue this in future work. Let us finally recall Conjecture 1.4, and note that a proof of this conjecture would also be valuable since the underlying challenges are also likely to show up in the study of simple graphs. In support of this conjecture, we present numerical data in Table 2, which suggests that the conclusion of Theorem 1.2 remains valid for unweighted directed graphs.

Supplementary material

The supplementary material for this article can be found at https://doi.org/10.1017/S0963548324000312.

Acknowledgements

I would like to thank Nils van de Berg, Jaron Sanders, and Haodong Zhu for providing helpful feedback on a draft of this manuscript. I further thank Aida Abiad for an inspiring talk, which motivated my interest in the spectral determinacy of graphs.

This work is part of the project Clustering and Spectral Concentration in Markov Chains with project number OCENW.KLEIN.324 of the research programme Open Competition Domain Science – M, which is partly financed by the Dutch Research Council (NWO).

Competing interests

The author declares none.

https://doi.org/10.1017/S0963548324000312 Published online by Cambridge University Press

Data availability statement

The code used to produce Tables 2 and 1 may be found in the supplementary materials.

References

[1] Abiad, A. and Haemers, W. H. (2012) Cospectral graphs and regular orthogonal matrices of level 2. *Electron. J. Combin.* **19**(3) P13. DOI: 10.37236/2383.

[2] Butler, S. (2010) A note about cospectral graphs for the adjacency and normalized Laplacian matrices. *Linear Multilin. Algeb.* **58**(3) 387–390. DOI: 10.1080/03081080902722741.

[3] Cheong, G. and Yu, M. (2023) The distribution of the cokernel of a polynomial evaluated at a random integral matrix, 10.48550/arXiv.2303.09125 arXiv preprint arXiv: 2303.09125v3.

[4] Clancy, J., Kaplan, N., Leake, T., Payne, S. and Wood, M. M. (2015) On a Cohen–Lenstra heuristic for Jacobians of random graphs. *J. Algebr. Comb.* **42**(3) 701–723. DOI: 10.1007/s10801-015-0598-x.

[5] Cohen, H. and Lenstra, H. W. (1984) Heuristics on class groups of number fields. In: *Number Theory Noordwijkerhout 1983*, Springer.

[6] Cover, T. M. and Thomas, J. A. (1999) *Elements of Information Theory*. John Wiley & Sons.

[7] Godsil, C. D. (2012) Controllable subsets in graphs. *Ann. Comb.* **16**(4) 733–744. DOI: 10.1007/s00026-012-0156-3.

[8] Godsil, C. D. and McKay, B. D. (1982) Constructing cospectral graphs. *Aequationes Math.* **25**(1) 257–268. DOI: 10.1007/BF02189621.

[9] Haemers, W. H. and Spence, E. (2004) Enumeration of cospectral graphs. *Eur. J. Combin.* **25**(2) 199–211. DOI: 10.1016/S0195-6698(03)00100-8.

[10] Halbeisen, L. and Hungerbühler, N. (1999) Generation of isospectral graphs. *J. Graph Theor.* **31**(3) 255–265.

[11] Koval, I. and Kwan, M. (2023) Exponentially many graphs are determined by their spectrum. *Q. J. Math.* **75**(3) 869–899. DOI: 10.1093/qmath/haae030.

[12] Li, S. and Sun, W. (2021) An arithmetic criterion for graphs being determined by their generalized A_α-spectra. *Discrete Math* **344**(8) 112469. DOI: 10.1016/j.disc.2021.112469.

[13] Liu, F. and Siemons, J. (2022) Unlocking the walk matrix of a graph. *J. Algebr. Comb.* **55**(3) 663–690. DOI: 10.1007/s10801-021-01065-3.

[14] Nguyen, H. H. and Paquette, E. (2020) Surjectivity of near-square random matrices. *Comb. Prob. Comp.* **29**(2) 267–292. DOI: 10.1017/S0963548319000348.

[15] Nguyen, H. H. and Wood, M. M. (2022) Random integral matrices: universality of surjectivity and the cokernel. *Invent. Math.* **228**(1) 1–76. DOI: 10.1007/s00222-021-01082-w.

[16] O'Rourke, S. and Touri, B. (2016) On a conjecture of Godsil concerning controllable random graphs. *SIAM J. Control Optim.* **54**(6) 3347–3378. DOI: 10.1137/15M1049622.

[17] Qiu, L., Wang, W. and Wang, W. (2021) Oriented graphs determined by their generalized skew spectrum. *Linear Algebra Appl.* **622** 316–332. DOI: 10.1016/j.laa.2021.03.033.

[18] Qiu, L., Wang, W. and Zhang, H. (2023) Smith normal form and the generalized spectral characterization of graphs. *Discrete Math.* **346** 113177. DOI: 10.1016/j.disc.2022.

[19] Reid, M. (1995) *Undergraduate Commutative Algebra*. Cambridge University Press.

[20] Sawin, W. and Wood, M. M. (2022) The moment problem for random objects in a category, 10.48550/arXiv.2210.06279 arXiv preprint arXiv: 2210.06279v2.

[21] Schwenk, A. J. (1973) Almost all trees are cospectral. In: *New Directions in the Theory of Graphs*, Academic Press.

[22] Sundaram, S. and Hadjicostis, C. N. (2012) Structural controllability and observability of linear systems over finite fields with applications to multi-agent systems. *IEEE Trans. Automat. Contr.* **58** 60–73. DOI: 10.1109/TAC.2012.

[23] The Sage Developers (2021) SageMath, the Sage Mathematics Software System (Version 9.3). Available at: URL: https://www.sagemath.org.

[24] van Dam, E. R. and Haemers, W. H. (2003) Which graphs are determined by their spectrum? *Linear Algebra Appl.* **373** 241–272. DOI: 10.1016/S0024-3795(03)00483-X.

[25] Verbitsky, O. and Zhukovskii, M. (2024) Canonization of a Random Circulant Graph by Counting Walks. Springer, pp. 319–334. International Conference and Workshops on Algorithms and Computation. DOI: 10.1007/978-981-97-0566-5_23.

[26] Wang, W. (2013) Generalized spectral characterization of graphs revisited. *Electron. J. Combin.* **20**(4) P4. DOI: 10.37236/3748.

[27] Wang, W. (2017) A simple arithmetic criterion for graphs being determined by their generalized spectra. *J. Combin. Theory Ser. B* **122** 438–451. DOI: 10.1016/j.jctb.2016.07.004.

[28] Wang, W. and Wang, W. (2024) Haemers' conjecture: an algorithmic perspective. *Exp. Math.*1–28. DOI: 10.1080/10586458.2024.2337229.

[29] Wang, W. and Xu, C.-X. (2006) An excluding algorithm for testing whether a family of graphs are determined by their generalized spectra. *Linear Algebra Appl.* **418**(1) 62–74. DOI: 10.1016/j.laa.2006.01.016.

https://doi.org/10.1017/S0963548324000312 Published online by Cambridge University Press

[30] Wang, W. and Xu, C.-X. (2006) A sufficient condition for a family of graphs being determined by their generalized spectra. *Eur. J. Combin.* **27**(6) 826–840. DOI: 10.1016/j.ejc.2005.05.004.
[31] Weibel, C. A. (1994) *An Introduction to Homological Algebra*. Cambridge University Press.
[32] Wood, M. M. (2017) The distribution of sandpile groups of random graphs. *J. Am. Math. Soc.* **30**(4) 915–958. DOI: 10.1090/jams/866.
[33] Wood, M. M. (2019) Random integral matrices and the Cohen–Lenstra heuristics. *Am. J. Math.* **141**(2) 383–398. DOI: 10.1353/ajm.2019.0008.

Cite this article: Van Werde A (2025). Cokernel statistics for walk matrices of directed and weighted random graphs. *Combinatorics, Probability and Computing* **34**, 131–150. https://doi.org/10.1017/S0963548324000312

https://doi.org/10.1017/S0963548324000312 Published online by Cambridge University Press