



Bounds for spectral projectors on generic tori

Pierre Germain¹ · Simon L. Rydin Myerson²

Received: 6 July 2022 / Revised: 2 December 2022 / Accepted: 14 December 2022
© The Author(s) 2022

Abstract

We investigate norms of spectral projectors on thin spherical shells for the Laplacian on generic tori, including generic rectangular tori. We state a conjecture and partially prove it, improving on previous results concerning arbitrary tori.

Mathematics Subject Classification 42A45 · 11D75 · 11L07 · 42B15 · 11H06 · 11P21

Contents

1	Introduction	1
1.1	Boundedness of spectral projectors on Riemannian manifolds	1
1.2	Spectral projectors on tori	1
1.2.1	Different kinds of tori	1
1.2.2	The conjecture	1
1.2.3	Known results if $p = \infty$	1
1.2.4	Known results if $p < \infty$	1
1.3	A new result through harmonic analysis	1
1.4	An elementary approach for $p = \infty$ and δ very small	1
2	Notations	1
3	Bounds on Weyl sums	1
3.1	Bound for small time	1
3.2	The one-dimensional case	1
3.3	The case of generic rectangular tori	1
3.4	The case of generic tori	1
4	Proof of Theorem 1.2	1

✉ Simon L. Rydin Myerson
Simon.Rydin-Myerson@warwick.ac.uk
https://maths.fan

Pierre Germain
pgermain@cims.nyu.edu

¹ Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, NY 10012-1185, USA

² Mathematics Institute, University of Warwick, Zeeman Building, Coventry CV4 7AL, United Kingdom

- 5 Some linear algebra
- 5.1 Singular values and largest subdeterminants
- 5.2 Describing some convex bodies
- 5.3 Extending matrices with prescribed largest subdeterminants
- 6 Proof of Theorem 1.3
- 6.1 Integrating over λ and β
- 6.2 Counting matrices with prescribed subdeterminants
- 6.3 The maximization procedure
- 6.4 Borel-Cantelli and the end of the argument
- Appendix A. Borel-Cantelli lemma
- References

1 Introduction

1.1 Boundedness of spectral projectors on Riemannian manifolds

Given a Riemannian manifold M with Laplace-Beltrami operator Δ , and given some $\lambda \geq 1$ and $0 < \delta < 1$, let

$$P_{\lambda,\delta} = P_{\lambda,\delta}^\chi = \chi \left(\frac{\sqrt{-\Delta} - \lambda}{\delta} \right), \tag{1.1}$$

where χ is a cutoff function taking values in $[0, 1]$ supported in $[-1, 1]$, and equal to 1 on $[-\frac{1}{2}, \frac{1}{2}]$. This definition is understood through the functional calculus for the Laplace-Beltrami operator, which is a self-adjoint operator on (complete) Riemannian manifolds.

A general question is to estimate

$$\|P_{\lambda,\delta}^\chi\|_{L^2 \rightarrow L^p}, \quad \text{where } p \in [2, \infty],$$

the exact choice of χ being immaterial¹.

The answer to this question is known in the case of the Euclidean space: define

$$p_{ST} = \frac{2(d+1)}{d-1}, \quad \sigma(p) = d - 1 - \frac{2d}{p}.$$

Then by Stein-Tomas [20, 21] we have

$$\|P_{\lambda,\delta}\|_{L^2 \rightarrow L^p} \sim \begin{cases} \lambda^{\sigma(p)/2} \delta^{1/2} & \text{if } p \geq p_{ST}, \\ \lambda^{\frac{d-1}{2}(\frac{1}{2} - \frac{1}{p})} \delta^{\frac{(d+1)}{2}(\frac{1}{2} - \frac{1}{p})} & \text{if } 2 \leq p \leq p_{ST}, \end{cases} \tag{1.2}$$

¹ If χ and ψ are two cutoff functions, then $\|P_{\lambda,\delta}^\chi\|_{L^2 \rightarrow L^p} \lesssim \|P_{\lambda,\delta}^\psi\|_{L^2 \rightarrow L^p} + \|P_{\lambda-\delta,\delta}^\psi\|_{L^2 \rightarrow L^p} + \|P_{\lambda+\delta,\delta}^\psi\|_{L^2 \rightarrow L^p}$.

where we write $A \sim B$ if the two quantities A and B are such that $\frac{1}{C}A \leq B \leq CA$, for a constant C which depends only on the d . The answer is again known in the case of compact Riemannian manifolds when $\delta = 1$ (Sogge [19], Theorem 5.1.1), for which

$$\|P_{\lambda,1}\|_{L^2 \rightarrow L^p} \sim \begin{cases} \lambda^{\sigma(p)/2} & \text{if } p \geq p_{ST}, \\ \lambda^{\frac{d-1}{2}(\frac{1}{2}-\frac{1}{p})} & \text{if } 2 \leq p \leq p_{ST}. \end{cases} \tag{1.3}$$

1.2 Spectral projectors on tori

1.2.1 Different kinds of tori

From now on, we focus on the case of tori given by the quotient $\mathbb{R}^d / (\mathbb{Z}e_1 + \dots + \mathbb{Z}e_d)$, where e_1, \dots, e_d is a basis of \mathbb{R}^d , with the standard metric. This is equivalent to considering the operators

$$P_{\lambda,\delta} = \chi \left(\frac{\sqrt{-\mathcal{Q}(\nabla)} - \lambda}{\delta} \right) \quad \text{on } \mathbb{T}^d = \mathbb{R}^d / \mathbb{Z}^d,$$

where ∇ is the standard gradient operator, and \mathcal{Q} is a quadratic form on \mathbb{R}^d , with coefficients β_{ij} :

$$\mathcal{Q}(x) = \sum_{i,j=1}^d \beta_{ij} x^i x^j \quad \implies \quad \mathcal{Q}(\nabla) = \sum_{i,j=1}^d \beta_{ij} \partial_i \partial_j.$$

Here (β_{ij}) is a symmetric positive definite real matrix. Dispensing with factors of 2π , which can be absorbed in \mathcal{Q} , the associated Fourier multiplier has the symbol

$$\chi \left(\frac{\sqrt{\mathcal{Q}(k)} - \lambda}{\delta} \right).$$

Standard and rectangular tori correspond to the following particular cases.

- The *standard torus* corresponds to (e_i) being orthonormal, or $\beta_{ij} = \delta_{ij}$.
- A *rectangular torus* corresponds to (e_i) being orthogonal, or equivalently to a diagonal quadratic form $\beta_{ij} = \beta_i \delta_{ij}$.

We will be concerned in this article with generic tori, which for our purposes are defined as follows.

Definition 1.1

- Consider the rectangular tori with $\beta_i \in [1, 2]$ for each i ; we say a property is true for *generic rectangular tori* if it is true on a set of $(\beta_i)_{1 \leq i \leq d}$ with full Lebesgue measure in $[1, 2]^d$.
- Consider the tori with $\beta_{ij} = \delta_{ij} + h_{ij}$ for each $1 \leq i, j \leq d$ and some $h_{ij} = h_{ji} \in [-\frac{1}{10d^2}, \frac{1}{10d^2}]$; we say a property is true for *generic tori* if it is true for a set of $(h_{ij})_{1 \leq i \leq j \leq d}$ with full Lebesgue measure in $[-\frac{1}{10d^2}, \frac{1}{10d^2}]^{d(d+1)/2}$.

1.2.2 The conjecture

It was conjectured in [10] that, for an arbitrary torus,

$$\|P_{\lambda,\delta}\|_{L^2 \rightarrow L^p} \lesssim 1 + (\lambda\delta)^{\frac{(d-1)}{2}(\frac{1}{2}-\frac{1}{p})} + \lambda^{\frac{d-1}{2}-\frac{d}{p}}\delta^{1/2} \quad \text{if } \delta > \lambda^{-1}; \quad (1.4)$$

here and below we denote $A \lesssim B$ if the quantities A and B are such that $A \leq CB$ for a constant C , where C may depend on the dimension d . This paper also contains new results towards this conjecture, as well as a survey of known results

In the present paper, we turn our attention towards generic tori, for which the typical spacing between eigenvalues of $\sqrt{-\Delta}$ is λ^{1-d} . Indeed, if k ranges in $[-R, R]^d$, then $\sqrt{Q(k)}$ takes $(2R)^d$ values in $[-CR, CR]$; if the β_{ij} are chosen generically we expect these to distribute approximately uniformly. This naturally leads to replacing the above conjecture by the following: for generic tori,

$$\|P_{\lambda,\delta}\|_{L^2 \rightarrow L^p} \lesssim_{\beta,\epsilon} 1 + (\lambda\delta)^{\frac{(d-1)}{2}(\frac{1}{2}-\frac{1}{p})} + \lambda^{\frac{d-1}{2}-\frac{d}{p}}\delta^{1/2} \quad \text{if } \delta > \lambda^{1-d+\epsilon} \quad (1.5)$$

(here the notation $A \lesssim_{\alpha} B$ means that the constant C in the relation $A \leq CB$ may depend on the parameter α).

1.2.3 Known results if $p = \infty$

For $p = \infty$, the problem of bounding $\|P_{\lambda,\delta}\|_{L^1 \rightarrow L^\infty}$ is closely related to counting lattice points in ellipsoids, a classical question in number theory. Namely, choosing $\chi = \mathbf{1}_{[-1,1]}$,

$$\|P_{\lambda,\delta}\|_{L^1 \rightarrow L^\infty} = N(\lambda + \delta) - N(\lambda - \delta),$$

where $N(\lambda)$ is the counting function associated to the quadratic form Q , defined as the number of lattice points $n \in \mathbb{Z}^d$ such that $Q(n) < \lambda^2$.

To leading order, $N(\lambda)$ equals $\text{Vol}(E)\lambda^d$, where $\text{Vol}(E)$ is the ellipsoid $\{Q(x) < 1\}$; the error term is denoted $P(\lambda)$:

$$N(\lambda) = \text{Vol}(B_1)\lambda^d + P(\lambda).$$

For the state of the art regarding $P(\lambda)$ for any fixed Q we refer the reader to the comments after (1.3) in [10], and to the work of Bourgain-Watt [7] giving an improved bound for the standard two-dimensional torus. For generic quadratic forms, there are a number of additional results.

- For generic diagonal forms, Jarník [13] showed that $P(\lambda) = O(\lambda^{d/2})$ if $d \geq 4$; a weaker, but more general, result is due to Schmidt [17].
- Landau [16] showed that $P(\lambda) = \Omega(\lambda^{\frac{d-1}{2}})$ for generic forms.
- It has been shown that the average size of the error, say $[\mathbb{E}|P(\lambda)|^2]^{1/2}$ is $O(\lambda^{\frac{d-1}{2}})$, for different types of averaging: over translations of the integer lattice [15], over shears [14], and over the coefficients (β_i) of a diagonal form [12].

- When $d = 2$, Trevisan [22] has investigated in more detail the distribution of the normalised error $P(\lambda)\lambda^{-1/2}$ when \mathcal{Q} is chosen at random and λ is large.
- The quantity $P(\lambda + \delta) - P(\lambda - \delta)$ has also received attention. In particular it has average size $O(\sqrt{\delta\lambda^{d-1}})$ when averaged over translations of the integer lattice [9], provided that $\delta \leq \lambda^{-\frac{d-1}{d+1}-\epsilon}$.

These results lead to the conjecture that the correct bound for the error term $P(\lambda)$ for generic \mathcal{Q} , or for generic diagonal \mathcal{Q} , would be $O(\lambda^{\frac{d-1}{2}})$. Meanwhile for $P(\lambda + \delta) - P(\lambda - \delta)$ the corresponding conjecture would be $O(\sqrt{\delta\lambda^{d-1}})$, at least for $\delta > \lambda^{1-d+\epsilon}$.

1.2.4 Known results if $p < \infty$

After the pioneering work of Zygmund [23], Bourgain [1] asked for L^p bounds for eigenfunctions of the Laplacian on the (standard) torus. He conjectured that, if φ is an eigenfunction of the Laplacian with eigenvalue λ , then

$$\|\varphi\|_{L^p} \lesssim \lambda^{\frac{d}{2}-1-\frac{d}{p}} \|\varphi\|_{L^2} \quad \text{if } p \geq \frac{2d}{d-2}, \tag{1.6}$$

which is equivalent to the case $\delta = \lambda^{-1}$ of (1.4) for the standard torus. Progress on this conjecture appeared in a series of articles [2–4] culminating in the proof of the ℓ^2 -decoupling conjecture by Bourgain and Demeter [5], which implies (1.6) if $p \geq \frac{2(d-1)}{d-3}$ and $d \geq 4$.

Bounds for spectral projectors are essentially equivalent to bounds for the resolvent $(-\Delta + z)^{-1}$. This was the point of view adopted in Shen [18], Bourgain-Shao-Sogge-Yau [6], and Hickman [11]. Here the goal is to prove a sharp bound when $p^* = \frac{2d}{d-2}$ and δ is sufficiently large.

Finally, the authors of the present paper were able to prove the conjecture (1.4) when δ is sufficiently large by combining ℓ^2 -decoupling with a geometry of numbers argument [10].

To the best of our knowledge, all works concerned with $p < \infty$ address either the case of standard tori, or the general case of arbitrary tori; the generic case does not seem to have been considered specifically. This will be a focus of the present paper.

1.3 A new result through harmonic analysis

The conjecture (1.5) was proved in [10] for arbitrary tori and δ not too small, and for generic tori we can improve the range for δ as follows.

Theorem 1.2 *For generic rectangular tori and for generic tori (in the sense of Definition 1.1), the conjecture (1.5) is verified if $p > p_{ST}$ and, for some $\epsilon > 0$,*

$$\text{either } \delta > \max \left(\lambda^{-1}, \lambda^{\frac{p}{p-2} \left[1 - \frac{d}{2} + \frac{1}{p} \left(\frac{d^2-d-2}{d-1} \right) \right] + \epsilon} \right) \quad \text{or } \lambda^{1 - \frac{d}{2} + \frac{1}{p} \frac{d(d-3)}{d-1} + \epsilon} < \delta < \lambda^{-1}.$$

Namely, under these conditions, there holds for almost all choices of $(\beta_i)_{1 \leq i \leq d}$ (for generic rectangular tori) or $(\beta_{ij})_{1 \leq i, j \leq d}$ (for generic tori)

$$\|P_{\lambda, \delta}\|_{L^2 \rightarrow L^p} \lesssim_{\beta, \epsilon} \lambda^{\frac{d-1}{2} - \frac{d}{p}} \delta^{1/2}.$$

In the particularly well-behaved case when $p = \infty$ and we consider generic diagonal tori, the theorem matches the classical result of Jarník [13] mentioned in the first bullet point in Sect. 1.2.3, which even promotes the upper bound in the theorem to an asymptotic in that case.

The proof of this theorem will be given in Sect. 4. The idea of this proof is to first express the spectral projector through the Schrödinger group. First note that the operator $\chi\left(\frac{\sqrt{-\mathcal{Q}(\nabla)} - \lambda}{\delta}\right)$ can also be written $\chi\left(\frac{-\mathcal{Q}(\nabla) - \lambda^2}{\lambda\delta}\right)$ by adapting the compactly supported function χ . This in turn can be expressed as

$$P_{\lambda, \delta} = \lambda\delta \int_{\mathbb{R}} \widehat{\chi}(\lambda\delta t) e^{-2\pi i \lambda^2 t} e^{-2\pi i t \mathcal{Q}(\nabla)} dt,$$

and then split off into two pieces, corresponding to $|t| \lesssim \lambda^{-1}$ and $|t| \gtrsim \lambda^{-1}$ respectively:

$$\begin{aligned} P_{\lambda, \delta} &= \int \lambda\delta \widehat{\chi}(\lambda t) \widehat{\chi}(\lambda\delta t) e^{-2\pi i \lambda^2 t} e^{-2\pi i t \mathcal{Q}(\nabla)} dt \\ &\quad + \int \lambda\delta [1 - \widehat{\chi}(\lambda t)] \widehat{\chi}(\lambda\delta t) e^{-2\pi i \lambda^2 t} e^{-2\pi i t \mathcal{Q}(\nabla)} dt \\ &= P_{\lambda, \delta}^{\text{small}} + P_{\lambda, \delta}^{\text{large}}. \end{aligned}$$

It is easy to see that the operator $P_{\lambda, \delta}^{\text{small}}$ can be written in the form $\delta P_{\lambda, 1}$ (after adjusting the cutoff function); in other words, this corresponds to the case $\lambda = 1$, to which the universal bounds of Sogge apply.

Turning to the term $P_{\lambda, \delta}^{\text{large}}$, its operator norm will be obtained through interpolation between

- A bound $L^{p'_{ST}} \rightarrow L^{p_{ST}}$, Theorem 4.1 below. As noted in [10], is a direct consequence of ℓ^2 decoupling (and valid for any torus).
- A bound $L^1 \rightarrow L^\infty$, for which genericity will be used. Namely, we will prove in Sect. 3 that, generically in (β_{ij}) ,

$$\int_{1/N}^T \|\chi(N^{-2}\Delta) e^{it\Delta}\|_{L^1 \rightarrow L^\infty} dt \lesssim_{\beta, \epsilon} TN^{\frac{d}{2} + \epsilon}. \tag{1.7}$$

One could think of (1.7) as square-root cancellation in $L^1 L^\infty$ since $\|\chi(N^{-2}\Delta)\|_{L^1 \rightarrow L^\infty} \sim N^d$. One could also see this as a minor-arc type bound in the spirit of the circle method; indeed in the case $p = \infty$ the proof in effect reduces to an application of the Davenport-Heilbronn circle method.

1.4 An elementary approach for $p = \infty$ and δ very small

When δ is small enough a more elementary counting argument can be used.

Our main result there is Theorem 6.1 below. We first state three $L^1 \rightarrow L^\infty$ bounds proved at the start of Sect. 6 since they are particularly simple; we will then mention consequences for $L^1 \rightarrow L^p$ bounds.

Theorem 1.3 *For generic tori, and also for generic rectangular tori, the following holds. If $\delta < \lambda^{1-2d-\epsilon}$, then*

$$\|P_{\lambda,\delta}\|_{L^1 \rightarrow L^\infty} \lesssim_{\beta,\epsilon} 1. \tag{1.8}$$

If instead $a \in \mathbb{Z}$ with $d \leq a \leq 2d$ and $\lambda^{-a} \leq \delta \leq \lambda^{1-a}$, then

$$\|P_{\lambda,\delta}\|_{L^1 \rightarrow L^\infty} \lesssim_{\beta,\epsilon} \delta^{1-\frac{a+1}{d+a+1}} \lambda^{d-1+\frac{a+1}{d+a+1}+\epsilon}. \tag{1.9}$$

Finally if $a \in \mathbb{Z}$ with $\frac{d-1}{2} \leq a < d$ and $\lambda^{-a} \leq \delta \leq \lambda^{1-a}$, we have

$$\|P_{\lambda,\delta}\|_{L^1 \rightarrow L^\infty} \lesssim_{\beta,\epsilon} \delta^{1-\frac{1}{d+1-a}} \lambda^{d-1+\frac{1}{d+1-a}+\epsilon}. \tag{1.10}$$

We remark that for $\delta \leq \lambda^{-1}$, interpolation with the $L^{p_{ST}'} \rightarrow L^{p_{ST}}$ bound from [10] gives

$$\|P_{\lambda,\delta}\|_{L^2 \rightarrow L^p} \lesssim_{\beta,\epsilon} \lambda^\epsilon (\|P_{\lambda,\delta}\|_{L^1 \rightarrow L^\infty})^{\frac{1}{2}-\frac{p_{ST}}{2p}}.$$

This would always fall short of the conjecture (1.5) for $p < \infty$, even with an optimal $L^1 \rightarrow L_\infty$ bound. We highlight a few features of these bounds.

- In the setting of Theorem 1.3, conjecture (1.5) would give $\|P_{\lambda,\delta}\|_{L^1 \rightarrow L^\infty} \lesssim_{\beta,\epsilon} 1 + \delta \lambda^{d-1+\epsilon}$.
- Although (1.9) and (1.10) do not recover (1.5), they do improve on the best known bounds for $N(\lambda - \delta) - N(\lambda + \delta)$ coming from the results listed in Sect. 1.2.3.
- Both (1.10) and (1.9) are special cases of the stronger estimate (6.1) below, while (1.8) has a short self-contained proof.
- We restrict to $a > \frac{d}{2} - 1$ in (1.10) and to $a \leq 2d$ in (1.10) solely because the remaining range is already covered by Theorem 1.2 or by (1.8), see also (6.2) below.
- When $a = d$ the bound (1.10) would be trivial, and hence for $a \geq d$ the bound (1.9) takes over.

In the proof of Theorem 6.1 we will use the Borel-Cantelli lemma to reduce to estimates for moments of $\|P_{\lambda,\delta}\|$, where the moments are taken over λ and β . A short computation reduces this to the following problem.

Problem 1.4 Estimate (from above, or asymptotically) the number of matrices of the form

$$P = \begin{pmatrix} m_{11}^2 & \cdots & m_{1b}^2 \\ \vdots & \ddots & \vdots \\ m_{d1}^2 & \cdots & m_{db}^2 \\ \lambda^2 & \cdots & \lambda^2 \end{pmatrix},$$

where the m_{ij} are integers, all entries in each row lie in a specified dyadic range, and also for each k the maximal $k \times k$ subdeterminant of P lies in a specified dyadic range.

In Sect. 6.2 we give an upper bound in this counting problem using what is in effect linear algebra, relying on the rather technical Lemma 5.7 below. We are then left with a maximum over all possible choices of the various dyadic ranges, and estimating this maximum will be the most challenging part of the proof.

The bound from Lemma 5.7 could be improved. See Remark 6.2 for one path. Another route concerns the case when all β_{ij} are generic, that is the case of generic tori as opposed to generic rectangular tori. Then one could expand P ; in place of the squares $m_{1i}^2, \dots, m_{di}^2$ the i th column would contain all degree 2 monomials in m_{1i}, \dots, m_{di} . This should allow a smaller bound.

2 Notations

We adopt the following normalizations for the Fourier series on \mathbb{T}^d and Fourier transform on \mathbb{R}^d , respectively:

$$\begin{aligned} f(x) &= \sum_{k \in \mathbb{Z}^d} \widehat{f}_k e^{2\pi i k \cdot x}, & \widehat{f}_k &= \int_{\mathbb{T}^d} f(x) e^{-2\pi i k \cdot x} dx, \\ \widehat{f}(\xi) &= \int_{\mathbb{R}^d} f(x) e^{-2\pi i x \cdot \xi} dx, & f(x) &= \int_{\mathbb{R}^d} \widehat{f}(\xi) e^{2\pi i x \cdot \xi} dx. \end{aligned}$$

With this normalization,

$$\widehat{f g} = \widehat{f} * \widehat{g} \quad \text{and} \quad \widehat{f * g} = \widehat{f} \widehat{g},$$

and the Parseval and Plancherel theorems, respectively, are given by

$$\|f\|_{L^2(\mathbb{T}^d)} = \|\widehat{f}\|_{\ell^2(\mathbb{Z}^d)} \quad \text{and} \quad \|f\|_{L^2(\mathbb{R}^d)} = \|\widehat{f}\|_{L^2(\mathbb{R}^d)}.$$

The operator $m(\sqrt{-Q(\nabla)})$ can be expressed as a Fourier multiplier

$$m(\sqrt{-Q(\nabla)})f = \sum_k m(\sqrt{Q(k)}) \widehat{f}_k e^{2\pi i k \cdot x},$$

or through a convolution kernel

$$m(\sqrt{-Q(\nabla)})f(x) = \int K(x - y)f(y) dy, \quad \text{with } K(z) = \sum_k m(\sqrt{Q(k)})e^{2\pi i k \cdot z}.$$

In Sects. 5 and 6 we will often join together several matrices A_1, \dots, A_n with the same number of rows to make a larger matrix, for which we use the notation $(A_1 \mid \dots \mid A_n)$. We view column vectors as matrices with one column, so that $(A \mid \bar{v})$ is the matrix A with the vector \bar{v} added as an extra column on the right.

Also in Sects. 5 and 6 we use the following notation relating to subdeterminants. If $k \leq \min(p, q)$ and M is a matrix in $\mathbb{R}^{p \times q}$, we will denote $D_k(M)$ the maximum absolute value of a $k \times k$ subdeterminant of M :

$$D_k(M) = \max_{\substack{\mathcal{I} \subset \{1, \dots, p\} \\ \mathcal{J} \subset \{1, \dots, q\} \\ \#\mathcal{I} = \#\mathcal{J} = k}} |\det(M_{ij})_{i \in \mathcal{I}, j \in \mathcal{J}}|.$$

We further define $D_0(M) = 1$ for ease of notation, and we let $D_k^{(\ell)}(M)$ denote the maximal subdeterminant, when the matrix M is restricted to its first ℓ columns:

$$D_k^{(\ell)}(M) = \max_{\substack{\mathcal{I} \subset \{1, \dots, p\} \\ \mathcal{J} \subset \{1, \dots, \ell\} \\ \#\mathcal{I} = \#\mathcal{J} = k}} |\det(M_{ij})_{i \in \mathcal{I}, j \in \mathcal{J}}|.$$

Given two quantities A and B , we denote $A \lesssim B$ if there exists a constant C such that $A \leq CB$, and $A \sim B$ if $A \lesssim B$ and $B \lesssim A$. If the implicit constant C is allowed to depend on a, b, c , the notation becomes $A \lesssim_{a,b,c} B$. In the following, it will often be the case that the implicit constant will depend on β , and on an arbitrarily small power of λ , for instance $A \lesssim_{\beta, \epsilon} \lambda^\epsilon B$. When this is clear from the context, we simply write $A \lesssim \lambda^\epsilon B$. Implicit constants in this notation may always depend on the dimension d of the torus that is the object of our study.

Finally, the Lebesgue measure of a set E is denoted $\text{mes}E$.

3 Bounds on Weyl sums

Consider the smoothly truncated Weyl sum (or regularized fundamental solution for the anisotropic Schrödinger equation)

$$K_N(t, x) = \sum_{k \in \mathbb{Z}^d} e^{-2\pi i(x \cdot k + tQ(k))} \phi\left(\frac{k_1}{N}\right) \dots \phi\left(\frac{k_d}{N}\right),$$

where ϕ is a smooth cutoff function supported on $[-1, 1]$, equal to 1 on $[-\frac{1}{2}, \frac{1}{2}]$.

In dimension one, it becomes

$$K_N^{(1)}(t, y) = \sum_{k \in \mathbb{Z}} e^{2\pi i(yk+k^2t)} \phi\left(\frac{k}{N}\right).$$

3.1 Bound for small time

For small t , the following bound holds on any torus.

Lemma 3.1 *For any non-degenerate quadratic form Q , if $|t| \lesssim \frac{1}{N}$,*

$$|K_N(t, x)| \lesssim \left(\frac{N}{tN + \frac{1}{N}}\right)^{d/2} Z\left(\frac{|x|}{tN + \frac{1}{N}}\right),$$

where Z decays super-polynomially.

Proof This is immediate on applying Poisson summation followed by stationary phase. □

3.2 The one-dimensional case

As in, for example, equation (8) of Bourgain-Demeter [3], we have:

Lemma 3.2 (Weyl bound) *If $a \in \mathbb{Z} \setminus \{0\}$, $q \in \{1, \dots, N\}$, $(a, q) = 1$ and $\left|t - \frac{a}{q}\right| \leq \frac{1}{qN}$:*

$$\forall y \in \mathbb{R}, \quad \left|K_N^{(1)}(t, y)\right| \lesssim_\epsilon \frac{N^{1+\epsilon}}{\sqrt{q}(1 + N|t - \frac{a}{q}|^{1/2})} \lesssim \frac{N^{1+\epsilon}}{\sqrt{q}}. \tag{3.1}$$

We now define a decomposition into major and minor arcs: for Q a power of 2, c_0 a constant which will be chosen small enough, and $Q \leq c_0N$,

$$\Lambda_Q(t) = \sum_{\substack{q \in \mathbb{N} \\ \frac{1}{2}Q \leq q < Q}} \sum_{\substack{a \in \mathbb{Z}^* \\ (a, q) = 1}} \mathbf{1}_{[-1, 1]} \left(NQ \left(t - \frac{a}{q} \right) \right).$$

In the definition of Λ_Q , the integer a is not allowed to be zero; it turns out to be convenient to single out the case $a = 0$, by letting

$$\Lambda_0(t) = \mathbf{1}_{[-\frac{1}{N}, \frac{1}{N}]}.$$

Observe that functions of the type $\mathbf{1}_{[-1, 1]} \left(NQ \left(t - \frac{a}{q} \right) \right)$, with $\frac{1}{2}Q_1 \leq q_1 \leq Q_1 \leq c_0N$ and $\frac{1}{2}Q_2 \leq q_2 \leq Q_2 \leq c_0N$ are disjoint if $(a_1, q_1) = (a_2, q_2) = 1$ and $\frac{a_1}{q_1} \neq \frac{a_2}{q_2}$.

Indeed, if $Q_1 \geq Q_2$,

$$\left| \frac{a_1}{q_1} - \frac{a_2}{q_2} \right| \geq \frac{1}{q_1 q_2} \geq \frac{1}{Q_1 Q_2} \geq \frac{1}{c_0 N Q_2} \geq \frac{2}{N Q_2} \geq \frac{1}{N Q_1} + \frac{1}{N Q_2}.$$

Therefore, $\Lambda_0 + \sum_{\substack{Q \in 2^{\mathbb{N}} \\ Q < c_0 N}} \Lambda_Q$ is the characteristic function of a set: the major arcs.

The minor arcs will be the complement, with characteristic function ρ . This gives the decomposition, for any $t \in \mathbb{R}$,

$$1 = \Lambda_0(t) + \sum_{\substack{Q \in 2^{\mathbb{N}} \\ Q < c_0 N}} \Lambda_Q(t) + \rho(t). \tag{3.2}$$

On the support of each of the summands above, the following bounds are available:

- On the support of Λ_0 , there holds $|t| \leq \frac{1}{N}$, and we resort to the short time bound.
- If t belongs to the support of Λ_Q , there exists $q \in [\frac{Q}{2}, Q]$ such that $|t - \frac{a}{q}| < \frac{1}{qN}$ and Weyl's bound gives $|K_N^{(1)}(t, y)| \lesssim_{\epsilon} \frac{N^{1+\epsilon}}{\sqrt{Q}}$.
- On the support of ρ , by Dirichlet's approximation theorem, there exists $a \in \mathbb{Z}$ and $q \in \{1, \dots, N\}$ relatively prime such that $|t - \frac{a}{q}| < \frac{1}{qN}$. If $q \sim N$, Weyl's bound gives $|K_N^{(1)}(t, y)| \lesssim_{\epsilon} N^{1/2+\epsilon}$. If $q \ll N$, then, since t does not belong to the support of $\sum_{Q \leq c_0 N} \Lambda_Q$, there holds $|t - \frac{a}{q}| \sim \frac{1}{qN}$, which implies again, by Weyl's bound, $|K_N^{(1)}(t, y)| \lesssim_{\epsilon} N^{1/2+\epsilon}$.

3.3 The case of generic rectangular tori

In this subsection, we assume that the tori are rectangular, or, equivalently, that the quadratic form Q is diagonal. First of all, we learn from the bounds on $K_N^{(1)}$ that, on the support of $\Lambda_{Q_1}(\beta_1 t) \dots \Lambda_{Q_k}(\beta_k t) \rho(\beta_{k+1} t) \dots \rho(\beta_d t)$, with all $Q_i \in 2^{\mathbb{N}}$ and $Q_i \leq c_0 N$,

$$|K_N(t, x)| = |K_N^{(1)}(\beta_1 t, x) \dots K_N^{(1)}(\beta_d t, x)| \lesssim_{\epsilon} \frac{N^{\frac{k+d}{2} + \epsilon}}{\sqrt{Q_1 \dots Q_k}}. \tag{3.3}$$

Lemma 3.3 *Let $\kappa, \epsilon > 0$; then for generic $\beta_1, \dots, \beta_d \in [1, 2]^d$ there holds, for $1 \leq k \leq d$, $0 \leq T \leq N^{\kappa}$, and $\epsilon > 0$, and for all Q_i, N equal to powers of 2, with $Q_i \leq N$,*

$$\int_0^T \Lambda_{Q_1}(\beta_1 t) \Lambda_{Q_2}(\beta_2 t) \dots \Lambda_{Q_k}(\beta_k t) dt \lesssim_{\epsilon, \kappa, \beta_1, \dots, \beta_d} N^{\epsilon} \frac{Q_1 \dots Q_k}{N^{\kappa}} T.$$

Proof Without loss of generality, we can choose $\beta_1 = 1$. Indeed, if $\beta_1, \dots, \beta_d, \frac{T}{\gamma}$ is changed to $\gamma\beta_1, \dots, \gamma\beta_d, T$, with $\gamma > 0$, the integral in the statement of the lemma changes by the factor γ . We claim that it suffices to prove that

$$\int_1^2 \dots \int_1^2 \int_0^T \Lambda_{Q_1}(t)\Lambda_{Q_2}(\beta_2 t) \dots \Lambda_{Q_k}(\beta_k t) dt d\beta_2 \dots d\beta_k \lesssim \mathbf{1}_{t \geq \frac{1}{4N}} \frac{Q_1 \dots Q_k}{N^k} T.$$

Indeed the case $t < \frac{1}{4N}$ of the lemma is then immediate, and the remaining case $\frac{1}{4N} \leq t \leq N^k$ follows by the Borel-Cantelli lemma as explained in Appendix A. By definition of Λ_Q ,

$$\int_1^2 \Lambda_Q(\beta t) d\beta = \frac{1}{t} \int_t^{2t} \Lambda_Q(y) dy = \frac{1}{t} \sum_{\substack{(a,q)=1 \\ \frac{1}{2}Q < q < Q}} \int_t^{2t} \mathbf{1}_{[-1,1]} \left(NQ \left(y - \frac{a}{q} \right) \right) dy.$$

To estimate this integral, we observe first that, if $[t, 2t] \cap [\frac{a}{q} - \frac{1}{NQ}, \frac{a}{q} + \frac{1}{NQ}] \neq \emptyset$, then

$$2t \geq \frac{a}{q} - \frac{1}{NQ} \geq \frac{1}{Q} \left[1 - \frac{1}{N} \right],$$

so that, in particular $t \geq \frac{1}{4Q}$. Similarly, one can show that the number of a 's such that $[t, 2t] \cap [\frac{a}{q} - \frac{1}{NQ}, \frac{a}{q} + \frac{1}{NQ}] \neq \emptyset$ is $\lesssim Qt$. Since furthermore the number of q 's in $[\frac{1}{2}Q, Q]$ is $\leq Q$, and since finally the integral of $\mathbf{1}_{[-1,1]} \left(NQ \left(y - \frac{a}{q} \right) \right)$ is $\lesssim \frac{1}{NQ}$, we obtain the estimate

$$\int_1^2 \Lambda_Q(\beta t) d\beta \lesssim \frac{1}{t} \cdot Qt \cdot Q \cdot \frac{1}{NQ} \mathbf{1}_{t \geq \frac{1}{4Q}} \lesssim \frac{Q}{N} \mathbf{1}_{t \geq \frac{1}{4Q}}.$$

Then, by Fubini's theorem

$$\begin{aligned} & \int_1^2 \dots \int_1^2 \int_0^T \Lambda_{Q_1}(t)\Lambda_{Q_2}(\beta_2 t) \dots \Lambda_{Q_k}(\beta_k t) dt d\beta_2 \dots d\beta_k \\ &= \int_0^T \Lambda_{Q_1}(t) \left[\int_1^2 \Lambda_{Q_2}(\beta_2 t) d\beta_2 \right] \dots \left[\int_1^2 \Lambda_{Q_k}(\beta_k t) d\beta_k \right] dt \\ &\lesssim \int_0^T \Lambda_{Q_1}(t) \frac{Q_2 \dots Q_k}{N^{k-1}} dt \lesssim \mathbf{1}_{t \geq \frac{1}{4Q_1}} \frac{Q_1 \dots Q_k}{N^k} T. \end{aligned}$$

□

A consequence of this lemma is an $L_t^1 L_x^\infty$ estimate on K_N .

Lemma 3.4 (Square root cancellation in $L^1_t L^\infty_x$) *Let $\kappa, \epsilon > 0$; then for generic $\beta_1, \dots, \beta_d \in [1, 2]^d$ there holds, for N a power of 2 and $\frac{1}{N} < T < N^\kappa$,*

$$\frac{1}{T} \int_{1/N}^T \|K_N(t, \cdot)\|_{L^\infty} dt \lesssim_{\epsilon, \kappa, \beta_1 \dots \beta_d} N^{\frac{d}{2} + \epsilon}.$$

Proof The first step is to use the decomposition (3.2) in each of the variables $\beta_1 t, \dots, \beta_d t$. Note that, since $t > \frac{1}{N}$ and $\beta_i \geq 1$, the term $\Lambda_0(\beta_i t)$ is always zero. By the inequality (3.3), for almost any choice of β_1, \dots, β_d , there holds

$$\begin{aligned} & \frac{1}{T} \int_{1/N}^T \|K_N(t, \cdot)\|_{L^\infty} dt \\ & \lesssim_\epsilon \sum_{k=0}^d \sum_{\substack{Q_1, \dots, Q_k \leq c_0 N \\ Q_i \in 2^{\mathbb{N}}}} \frac{1}{T} \int_{1/N}^T \Lambda_{Q_1}(\beta_1 t) \dots \Lambda_{Q_k}(\beta_k t) \rho(\beta_{k+1} t) \dots \rho(\beta_d t) dt \frac{N^{\frac{k+d}{2} + \epsilon}}{\sqrt{Q_1 \dots Q_k}} \\ & \lesssim_{\epsilon, \kappa, \beta_1 \dots \beta_d} N^\epsilon \sum_{k=0}^d \sum_{Q_1, \dots, Q_k \leq c_0 N} \frac{Q_1 \dots Q_k}{N^k} \frac{N^{\frac{k+d}{2}}}{\sqrt{Q_1 \dots Q_k}} \\ & \lesssim N^{\frac{d}{2} + \epsilon}. \end{aligned}$$

□

3.4 The case of generic tori

We start with an averaging lemma.

Lemma 3.5 *For $A \in \mathbb{R}$, $\lambda > 0$, and $N \geq 10$,*

$$\int_{-1}^1 \min \left(\frac{1}{\|\lambda h + A\|}, N \right) dh \lesssim \begin{cases} \log N & \text{if } \lambda \geq 1 \\ \frac{\log N}{\lambda} & \text{if } \lambda \leq 1, \end{cases}$$

where, for a real number x , the notation $\|x\|$ stands for $\text{dist}(x, \mathbb{Z}) = \min_{k \in \mathbb{Z}} |k - x|$.

Proof If $\lambda \geq 1$, the left-hand side can be bounded by the average of the function $\min \left(\frac{1}{\|h\|}, N \right)$, which equals $\log N$. If $\lambda \leq 1$, the left-hand side is bounded by $\int_{-1}^1 \min \left(\frac{1}{\|\lambda h\|}, N \right) dh \lesssim \frac{\log N}{\lambda}$. □

Armed with this lemma, we can now prove the desired square root cancellation result - its proof already appeared in [8], but we include an equivalent version here for the reader’s convenience. Recall that the measure on nonsingular symmetric matrices we consider is given by $B = \text{Id} + h_{ij}$, where h is a symmetric matrix, all of whose coefficients are independent (besides the symmetry assumption) and uniformly distributed in $\left[\frac{1}{10d^2}, \frac{1}{10d^2} \right]$.

Lemma 3.6 (Square root cancellation in $L^1_x L^\infty_x$) *Let $\kappa, \epsilon > 0$; then for generic $(\beta_{i,j})$, there holds, for N a power of 2 and $\frac{1}{N} < T < N^\kappa$,*

$$\frac{1}{T} \int_{1/N}^T \|K_N(t, \cdot)\|_{L^\infty} dt \lesssim_{\epsilon, \kappa, \beta_{ij}} N^{\frac{d}{2} + \epsilon}.$$

Proof By the Borel-Cantelli argument in Appendix A, the result would follow from the bound

$$\int \frac{1}{T} \int_{1/N}^T \|K_N(t, \cdot)\|_{L^\infty} dt dB \lesssim_\epsilon N^{\frac{d}{2} + \epsilon},$$

which will now be proved. For $x \in \mathbb{T}^d$ and $t \in \mathbb{R}$, applying Weyl differencing gives

$$|K_N(t, x)|^2 = \sum_{m,n} \phi\left(\frac{m_1 + n_1}{2N}\right) \dots \phi\left(\frac{m_d + n_d}{2N}\right) \left(\frac{m_1 - n_1}{2N}\right) \dots \phi\left(\frac{m_d - n_d}{2N}\right) e^{2itQ(m,n)},$$

(where the sum is implicitly restricted to m_i, n_i having the same parity). By Abel summation, this implies that

$$\|K_N(t, \cdot)\|_{L^\infty} \lesssim \sum_{|n| \lesssim N} \prod_{i=1}^d \min\left(\frac{1}{\|tQ_i(n)\|}, N\right),$$

where $Q_i(n) = \sum_j \beta_{ij} n_j$. Combining the Cauchy-Schwarz inequality with the above yields

$$\begin{aligned} \int_{1/N}^T \|K_N(t, \cdot)\|_{L^\infty} dt dB &\lesssim \sqrt{T} \left[\int \int_{1/N}^T \|K_N(t, \cdot)\|_{L^\infty}^2 dt dB \right]^{1/2} \\ &\lesssim \sqrt{T} \left[\int \int_{1/N}^T \sum_{|n| \lesssim N} \prod_{i=1}^d \min\left(\frac{1}{\|tQ_i(n)\|}, N\right) dt dB \right]^{1/2}, \end{aligned}$$

where $dB = \prod_{i \leq j} d\beta_{i,j}$. We now exchange the order of summation and integration, performing first the integration over B . Without loss of generality, assume that $|n_1| \sim |n|$. Note that $tQ_{1,1}(n) = t \sum_j \beta_{1,j} n_j$; therefore, by Lemma 3.4, integrating first $\int \min\left(\frac{1}{\|tQ_1(n)\|}, N\right) d\beta_{1,1}$ gives $\log N \langle \frac{1}{t|n|} \rangle$. We integrate next $\int \min\left(\frac{1}{\|tQ_2(n)\|}, N\right) d\beta_{1,2}$, giving the same result, and similarly $\int \min\left(\frac{1}{\|tQ_k(n)\|}, N\right) d\beta_{1,k}$, for $3 \leq k \leq d$. Finally, $\int \prod_{2 \leq i \leq j} d\beta_{i,j}$ gives $O(1)$.

Coming back to the sequence of inequalities above,

$$\int_{1/N}^T \|K_N(t, \cdot)\|_{L^\infty} dt dB \lesssim \sqrt{T} \left[\sum_{|n| \lesssim N} (\log N)^d \int_{1/N}^T \langle \frac{1}{t|n|} \rangle^d dt \right]^{1/2} \lesssim TN^{d/2}(\log N)^{d/2}.$$

□

4 Proof of Theorem 1.2

An important element of the proof is the optimal $L^2 \rightarrow L^{pST}$ bound for spectral projectors. As observed in the previous article by the authors [10], it is a consequence of the ℓ^2 decoupling bound of Bourgain-Demeter. The statement is as follows.

Theorem 4.1 *If $\lambda > 1$ and $\delta < 1$, for any positive definite quadratic form Q ,*

$$\|P_{\lambda,\delta}\|_{L^2 \rightarrow L^{pST}} \lesssim (1 + \lambda\delta)^{\frac{1}{pST}}.$$

We now turn to the proof of Theorem 1.2.

Proof Step 1: Allowing more general cutoff functions. Define the spectral projector

$$P'_{\lambda,\delta} = \zeta \left(\frac{Q(\nabla) + \lambda^2}{\lambda\delta} \right),$$

where ζ is a Schwartz function such that $\zeta(0) > 0$ and $\widehat{\zeta}$ is compactly supported.

We claim that it suffices to prove Theorem 1.2 for the spectral projector $P'_{\lambda,\delta}$ instead of $P_{\lambda,\delta}$. Indeed, assume that $P'_{\lambda,\delta}$ enjoys the bound in this theorem. Since there exists $c > 0$ such that $\zeta(x) \geq c\mathbf{1}_{[-c,c]}(x)$, the desired bound follows² for the operator $\mathbf{1}_{[-c,c]} \left(\frac{Q(\nabla) + \lambda^2}{\lambda\delta} \right)$. This implies in turn this bound for the operator $\mathbf{1}_{[-a,a]} \left(\frac{\sqrt{-Q(\nabla)} - \lambda}{\delta} \right)$, for a constant $a > 0$. Finally, this implies the desired bound for $P_{\lambda,\delta}$ since $|\chi(x)|$ can be bounded by a finite sum of translates of $\mathbf{1}_{[-a,a]}$.

We now claim that $P'_{\lambda,1}$ enjoys the Sogge bounds (1.3), just like $P_{\lambda,1}$. This follows from writing

$$P'_{\lambda,1} = \sum_{n > -\frac{\delta}{\lambda}} P'_{\lambda,1} \mathbf{1}_{[\lambda+(n-1)\delta, \lambda+n\delta]} \left(\sqrt{-Q(\nabla)} \right)$$

² Here, we are using that, if $m_1(\Delta)$ and $m_2(\Delta)$ are two Fourier multipliers such that $|m_1| \geq |m_2|$, then $\|m_1(\Delta)\|_{L^2 \rightarrow L^p} \geq \|m_2(\Delta)\|_{L^2 \rightarrow L^p}$. This follows from Parseval's theorem.

and bounding

$$\begin{aligned} \|P'_{\lambda,1}\|_{L^2 \rightarrow L^p} &\lesssim c_n \sum_{n > -\frac{\delta}{\lambda}} \left\| \mathbf{1}_{[\lambda+(n-1)\delta, \lambda+n\delta]} \left(\sqrt{-Q(\nabla)} \right) \right\|_{L^2 \rightarrow L^p} \quad \text{with } c_n \\ &= \sup_{(n-1)\delta < x < n\delta} \left| \zeta \left(\frac{-x^2 + \lambda^2}{\lambda\delta} \right) \right|. \end{aligned}$$

The rapid decay of ζ implies that $|c_n| \lesssim n^{-N}$ for any N , while $\mathbf{1}_{[\lambda+(n-1)\delta, \lambda+n\delta]}$ enjoys the Sogge bounds. Thus, it is not hard to sum the above series, and deduce that $P'_{\lambda,1}$ also enjoys the Sogge bounds.

By a similar argument, it can be shown that Theorem 4.1 applies to $P'_{\lambda,1}$.

Step 2: splitting the spectral projector. Writing the function $x \mapsto \zeta \left(\frac{Q(x)+\lambda^2}{\lambda\delta} \right)$ as a Fourier transform, the operator $P'_{\lambda,\delta}$ becomes

$$P'_{\lambda,\delta} = \int_{\mathbb{R}} \lambda\delta \widehat{\zeta}(\lambda\delta t) e^{-2\pi i\lambda^2 t} e^{-2\pi i t Q(\nabla)} dt, \tag{4.1}$$

with the kernel

$$P'_{\lambda,\delta}(x) = \int_{\mathbb{R}} \lambda\delta \widehat{\zeta}(\lambda\delta t) e^{-2\pi i\lambda^2 t} K_N(t, x) dt;$$

here, we choose N to be a power of 2 in the range $[2\lambda, 4\lambda]$.

The basic idea is to split the integral giving $P'_{\lambda,\delta}$ into two pieces, $|t| < \lambda^{-1}$ and $|t| > \lambda^{-1}$. The former corresponds to an operator of the type $\delta P'_{\lambda,1}$, for which bounds are well-known: this corresponds to the classical Sogge theorem. The latter can be thought of as an error term, it will be bounded by interpolation between $p = p_{ST}$ and $p = \infty$, and it is for this term that genericity is used.

Turning to the implementation of this plan, we write

$$\begin{aligned} P'_{\lambda,\delta} &= \int_{\mathbb{R}} \lambda\delta \widehat{\zeta}(\lambda t) \widehat{\zeta}(\lambda\delta t) e^{-2\pi i\lambda^2 t} e^{-2\pi i t Q(\nabla)} dt \\ &\quad + \int_{\mathbb{R}} \lambda\delta [1 - \widehat{\zeta}(\lambda t)] \widehat{\zeta}(\lambda\delta t) e^{-2\pi i\lambda^2 t} e^{-2\pi i t Q(\nabla)} dt \\ &= P_{\lambda,\delta}^{\text{small}} + P_{\lambda,\delta}^{\text{large}} \end{aligned}$$

Step 3: Bounding the term corresponding to small t . Observe that $P_{\lambda,\delta}^{\text{small}}$ can be written $\delta P''_{\lambda,1}$, where $P''_{\lambda,1}$ is a variation on $P'_{\lambda,1}$; this can be compared to the definition of $P_{\lambda,\delta}^{\text{small}}$ and (4.1). We saw in Step 1 that $P'_{\lambda,1}$ enjoys the Sogge bounds, and this remains true for $P''_{\lambda,1}$.

Furthermore, by a classical TT^* argument, the operator norm of the spectral projector $L^{p'} \rightarrow L^p$ is the square of the operator norm of the spectral projector $L^2 \rightarrow L^p$ (once again, up to redefining the cutoff function χ).

Therefore, it enjoys the bound

$$\|P_{\lambda,\delta}^{\text{small}}\|_{L^{p'} \rightarrow L^p} \lesssim \delta \|P_{\lambda,1}''\|_{L^2 \rightarrow L^p}^2 \lesssim \lambda^{d-1-\frac{2d}{p}} \delta \quad \text{for } p \geq p_{ST}. \quad (4.2)$$

Step 4: Bounding the term corresponding to large t . In order to bound this term, we will interpolate between

- The case $p = p_{ST}$: in this case, we resort to Theorem 4.1. We saw in Step 1 that it applies to $P'_{\lambda,\delta}$, and, by the same argument, it applies to $P''_{\lambda,\delta}$. This gives

$$\begin{aligned} \|P_{\lambda,\delta}^{\text{large}}\|_{L^{p'_{ST}} \rightarrow L^{p_{ST}}} &\lesssim \|P'_{\lambda,\delta}\|_{L^{p'_{ST}} \rightarrow L^{p_{ST}}} + \|P_{\lambda,\delta}^{\text{small}}\|_{L^{p'_{ST}} \rightarrow L^{p_{ST}}} \\ &\lesssim \|P'_{\lambda,\delta}\|_{L^{p'_{ST}} \rightarrow L^{p_{ST}}} + \delta \|P''_{\lambda,\delta}\|_{L^{p'_{ST}} \rightarrow L^{p_{ST}}} \\ &\lesssim_{\epsilon} \lambda^{\epsilon} [1 + (\lambda\delta)^{2/p_{ST}} + \delta\lambda^{2/p_{ST}}] \lesssim_{\epsilon} \lambda^{\epsilon} [1 + (\lambda\delta)^{2/p_{ST}}]. \end{aligned}$$

- The case $p = \infty$: in this case, we resort to Lemma 3.4 (generic rectangular tori) and Lemma 3.6 (generic tori). In order for these lemmas to apply, we add a further requirement on ζ , namely that its Fourier transform be 1 in a neighbourhood of zero. Then, for almost any choice of (β_{ij}) ,

$$\|P_{\lambda,\delta}^{\text{large}}\|_{L^1 \rightarrow L^{\infty}} \lesssim \int_{\mathbb{R}} \lambda\delta |1 - \widehat{\zeta}(\lambda t)| |\widehat{\zeta}(\lambda\delta t)| \|K_N(t, \cdot)\|_{L^{\infty}} dt \lesssim_{\beta,\epsilon} \lambda^{\frac{d}{2}+\epsilon}.$$

Interpolating between these two bounds gives for almost any choice of (β_{ij}) ,

$$\|P_{\lambda,\delta}^{\text{large}}\|_{L^{p'} \rightarrow L^p} \lesssim_{\beta,\epsilon} (1 + \lambda\delta)^{2/p} \lambda^{\frac{d}{2}(1-\frac{p_{ST}}{p})+\epsilon} \quad \text{for } p \geq p_{ST}. \quad (4.3)$$

Step 5: conclusion. Finally, combining (4.2) and (4.3), and using that $\|P_{\lambda,\delta}\|_{L^2 \rightarrow L^p}^2 \sim \|P_{\lambda,\delta}^2\|_{L^{p'} \rightarrow L^p}$ (by the classical TT^* argument) gives

$$\|P_{\lambda,\delta}\|_{L^2 \rightarrow L^p} \lesssim_{\beta,\epsilon} \lambda^{\frac{d-1-d}{2}-\frac{d}{p}} \delta^{1/2} + (1 + \lambda\delta)^{\frac{1}{p}} \lambda^{\frac{d}{4}(1-\frac{p_{ST}}{p})+\epsilon},$$

from which the desired result follows. □

5 Some linear algebra

In this section we assemble technical tools to attack Problem 1.4. Recall that the goal is to count the number of matrices

$$\begin{pmatrix} m_{11}^2 & \cdots & m_{1b}^2 \\ \vdots & \ddots & \vdots \\ m_{d1}^2 & \cdots & m_{db}^2 \\ \lambda^2 & \cdots & \lambda^2 \end{pmatrix},$$

where the m_{ij} are integers in given dyadic intervals and the maximal subdeterminants of P also lie in some specified dyadic intervals. The idea is to add the columns one by one, so that we count the number of possible (m_{11}, \dots, m_{d1}) , and for each possibility we count the number of (m_{12}, \dots, m_{d2}) , and so on. The main goal of this section is Lemma 5.7, which can be understood as an estimate for the *measure* of the *real* vectors (m_{1k}, \dots, m_{dk}) which are within a distance $O(1)$ of a vector satisfying the required conditions, given the previous columns.

In this and the next section we will often use the notation $D_k(M), D_k^{(\ell)}(M)$ defined in Sect. 2.

5.1 Singular values and largest subdeterminants

We begin with a number of general statements about the size of the subdeterminants of a $p \times q$ matrix, and their relation to the *singular value decomposition*, a type of canonical form for matrices. Throughout this subsection, implicit constants in \lesssim and \sim notation may depend on p and q .

Lemma 5.1 (Singular value decomposition) *Let $M \in \mathbb{R}^{p \times q}$ and let $m = \min(p, q)$. Then there are $U \in O(p), V \in O(q)$ and (uniquely defined) singular values $\sigma_1(M) \geq \dots \geq \sigma_m(M) \geq 0$ such that*

$$M = U \begin{pmatrix} \Sigma \\ 0 \end{pmatrix} V \text{ or } U (\Sigma \ 0) V, \text{ where } \Sigma = \text{diag}(\sigma_1(M), \dots, \sigma_m(M)), \quad (5.1)$$

and where 0 is a matrix of zeroes (possibly empty).

Lemma 5.2 (Stability of σ_k and D_k under multiplication by orthogonal matrices) *If $k \leq \min(p, q)$ and M is a matrix in $\mathbb{R}^{p \times q}$,*

- (i) *If $U \in O(p)$, then $D_k(UM) \sim D_k(M)$ and $\sigma_k(UM) = \sigma_k(M)$.*
- (ii) *If $U \in O(q)$, then $D_k(MU) \sim D_k(M)$ and $\sigma_k(MU) = \sigma_k(M)$.*

Proof The statements (i) and (ii) are symmetric, so that we will only focus on (i). Let $\mathcal{I} \subset \{1, \dots, p\}, \mathcal{J} \subset \{1, \dots, q\}, \#\mathcal{I} = \#\mathcal{J} = k$. The Cauchy-Binet identity gives

$$\det (UM)_{i \in \mathcal{I}, j \in \mathcal{J}} = \det U_{i \in \mathcal{I}, \ell} \cdot M_{\ell, j \in \mathcal{J}} = \sum_{\substack{\mathcal{K} \subset \{1, \dots, p\} \\ \#\mathcal{K} = k}} \det (U_{i \ell})_{i \in \mathcal{I}, \ell \in \mathcal{K}} \det (M_{\ell j})_{\ell \in \mathcal{K}, j \in \mathcal{J}}.$$

Hence $D_k(UM) \lesssim D_k(M)$ and repeating the argument with $U^{-1}M$ in place of M shows that $D_k(M) \lesssim D_k(UM)$ as well.

Finally, it follows from the uniqueness of the (σ_i) in (5.1) that $\sigma_k(UM) = \sigma_k(M)$. □

Corollary (Relation between the D_k and σ_k) *If $k \leq \min(p, q)$, the singular values and the maximal subdeterminants are such that*

$$\sigma_k(M) \sim D_{k-1}(M)^{-1} D_k(M),$$

where we use the convention that $0^{-1}0 = 0$, or equivalently

$$D_k(M) \sim \sigma_1(M) \dots \sigma_k(M).$$

Proof By lemmas 5.1 and 5.2, it suffices to prove these formulas for a rectangular diagonal matrix; but then they are obvious. \square

Lemma 5.3 *Given a matrix $M \in \mathbb{R}^{p \times q}$, we can change the order of its columns so that for each $k \leq \ell \leq \min(p, q)$,*

$$D_k^{(\ell)}(M) \sim D_k(M). \tag{5.2}$$

Proof We claim first that it suffices to prove the result for the matrix UM , where U is orthogonal. Indeed, denoting $M^{(\ell)}$ for the restriction of M to its first ℓ columns, this implies, in combination with Lemma 5.2,

$$D_k^{(\ell)}(M) = D_k(M^{(\ell)}) \sim D_k(UM^{(\ell)}) = D_k^{(\ell)}(UM) \sim D_k(UM) \sim D_k(M),$$

which is the desired result. For the remainder of the proof, we write for simplicity $\sigma_i = \sigma_i(M)$.

We can choose U as in Lemma 5.1, in which case, assuming for instance $p \geq q$, it suffices to deal with the case

$$M = (\sigma_1 L_1, \dots, \sigma_q L_q, 0, \dots, 0)^T,$$

where $L_i \cdot L_j = \delta_{ij}$. The 0 entries are irrelevant, so we can assume that

$$M = (\sigma_1 L_1, \dots, \sigma_q L_q)^T = (M_{ij})_{1 \leq i, j \leq q}.$$

We now claim that, after permuting the columns of M , it can be ensured that, for any k , the top left square matrix of dimension $k \times k$ has nearly maximal subdeterminant:

$$\det(M_{ij})_{1 \leq i, j \leq k} \sim D_k(M). \tag{5.3}$$

The construction of the matrix permutation is iterative and proceeds as follows: expanding the determinant of M with respect to the last row, we see that

$$\sigma_1 \dots \sigma_q = \det M = \sum_{i=1}^p (-1)^{q+i} M_{q,i} \det M^{(q,i)},$$

where $M^{(q,i)}$ is the matrix obtained from M by removing the q -th row and the i -th column. Since $|M_{q,i}| \leq \sigma_q$ and $\det M^{(q,i)} \lesssim \sigma_1 \dots \sigma_{q-1}$ for all i , we can find i_0 such that $|M_{q,i_0}| \sim \sigma_q$, and $\det M^{(q,i_0)} \sim \sigma_1 \dots \sigma_{q-1}$. Exchanging the columns i_0 and q , the resulting matrix satisfies (5.3) for $k = q - 1$.

We now consider the matrix $N = (M_{ij})_{1 \leq i, j \leq q-1}$, which was denoted $M^{(q, i_0)}$ before columns were permuted. It is such that entries in the last row are $\leq \sigma_{q-1}$, and subdeterminants of size $q - 2$ are bounded by $\sigma_1 \dots \sigma_{q-2}$. Therefore, the same argument as above can be applied, and it proves (5.3) for $k = q - 2$. An obvious induction leads to the desired statement. \square

5.2 Describing some convex bodies

We can use the subdeterminants studied above to describe certain convex bodies. Our first result concerns the measure of a neighbourhood of a convex hull.

Lemma 5.4 *Let X denote the $d \times d$ matrix with columns $x^{(i)}$. Then*

$$\text{mes} \left\{ \sum_{i=1}^d t_i x^{(i)} + w, |t_i| \leq 1, |w| \leq 1 \right\} \lesssim 1 + \sum_{k=1}^d D_k(X).$$

Proof Let M be the $d \times 2d$ matrix $(X|\text{Id})$, so that the set whose measure we want to estimate can be written (up to a multiplicative constant) $MB(0, 1)$. By Lemma 5.1 (singular value decomposition), we can write $M = U(\Sigma|0)V$; then

$$\begin{aligned} \text{mes}MB(0, 1) &= \text{mes}U(\Sigma|0)VB(0, 1) \\ &= \text{mes}(\Sigma|0)B(0, 1) \lesssim \text{mes}\Sigma B(0, 1) = \det \Sigma \sim D_d(M). \end{aligned}$$

There remains to evaluate $D_d(M)$; owing to the specific structure of M ,

$$D_d(M) = D_d(X|\text{Id}) \lesssim \sum_{k=1}^d D_k(X).$$

\square

We can also describe a subset of a convex hull cut out by linear inequalities, showing that it is contained in a potentially smaller convex hull.

Lemma 5.5 *Given linearly independent $v^{(1)}, \dots, v^{(d)} \in (\mathbb{R}^d)^d$, and $Y_i > 0, Z_i > 0$ there are $w^{(1)}, \dots, w^{(d)} \in (\mathbb{R}^d)^d$, with*

$$w_j^{(i)} \lesssim \min(Y_i |v^{(i)}|, Z_j), \tag{5.4}$$

such that

$$\left\{ z \in \mathbb{R}^d : z = \sum_{i=1}^d y_i v^{(i)}, |y_i| \leq Y_i, |z_i| \leq Z_i \right\} \subset \left\{ \sum_{i=1}^d t_i w^{(i)} : |t_i| \lesssim 1 \right\}. \tag{5.5}$$

Proof Let Y be the matrix with columns $Y_i v^{(i)}$, which, without loss of generality, can be assumed to have nondecreasing norms. We claim that its singular values, τ_i , are such that

$$\tau_i \lesssim Y_i |v^{(i)}|. \tag{5.6}$$

Indeed, by the Courant minimax principle, τ_k can be characterized as follows

$$\begin{aligned} \tau_k = \sigma_k(Y) &= \min_{\dim E=d+1-k} \max_{\substack{x \in E \\ |x|=1}} |Yx| \leq \max_{\substack{x_{d-k+2}=\dots=x_d=0 \\ |x|=1}} |Yx| \\ &\lesssim \max_{j=1, \dots, d-k+1} |Y_j| |v^{(j)}| = |Y_{d-k+1}| |v^{(d-k+1)}|. \end{aligned}$$

Let Z be the matrix $\text{diag}(Z_1, \dots, Z_d)$, and let $M = (Y^{-T} | Z^{-T})^T \in \mathbb{R}^{2d \times d}$. Then the set on the left-hand side of (5.5) is contained in $\{z : |Mz| \lesssim 1\}$. By Lemma 5.1, we can write $M = U(\Sigma^T | 0)^T V$, so that the set $\{z : |Mz| \lesssim 1\}$ can now be written (up to a multiplicative constant) as $WB(0, 1)$, with

$$W = V^{-1} \Sigma^{-1}.$$

We can now define the $w^{(i)}$ to be the columns of W ; in order to establish the lemma, it suffices to prove the inequality (5.4). Note first that

$$|w^{(i)}| \lesssim \sigma_i(M)^{-1}. \tag{5.7}$$

Next, denoting $U = \begin{pmatrix} U_1 & U_2 \\ U_3 & U_4 \end{pmatrix}$ where each U_i is a $d \times d$ matrix, we have $Y^{-1} = U_1 \Sigma V$. Therefore,

$$(\tau_{d+1-i})^{-1} = \sigma_i(Y^{-1}) = \sigma_i(U_1 \Sigma V) \lesssim \sigma_i(\Sigma) = \sigma_i(M). \tag{5.8}$$

Combining (5.6), (5.7) and (5.8),

$$|\vec{w}^{(i)}| \leq \sigma_i(M)^{-1} \lesssim \tau_{d+1-i} \lesssim Y_i |v^{(i)}|.$$

Finally, $W = ZU_3$, which gives $|w_j^{(i)}| \lesssim Z_j$. □

5.3 Extending matrices with prescribed largest subdeterminants

We now start to describe the columns which may be added to a given $p \times k$ matrix, with a prescribed effect on its singular values.

Lemma 5.6 *Let M be a $p \times k$ matrix, which admits a singular value decomposition as in (5.1). For some fixed $C > 0$, let*

$$\begin{aligned} S(M, R) &= \{x \in \mathbb{R}^p : D_j(M|x) \leq C D_j(M) \\ &\quad (1 \leq j \leq \min(p, k)), D_{k+1}(M|x) \leq R \ (p \geq k + 1)\} \end{aligned}$$

and set

$$\tau_i = \begin{cases} \sigma_i(M) & \text{if } i \leq k, \\ \min(\sigma_k(M), \frac{R}{D_k(M)}) & \text{if } i \geq k + 1 \text{ and } p \geq k + 1. \end{cases}$$

Then, denoting $U^{(i)}$ for the columns of the matrix U from the singular value decomposition of M ,

$$S(M, R) \subset \left\{ \sum_{i=1}^p y_i U^{(i)} : |y_i| \lesssim_{C,p,k} \tau_i \right\}.$$

Proof In the proof we allow all implicit constants in \lesssim, \sim notation to depend on C, p, k .

Step 1: $p \geq k + 1$ and $U = \text{Id}$. Then the singular value decomposition of M is $M = \begin{pmatrix} \Sigma \\ 0 \end{pmatrix} V$ and

$$(M|x) = \begin{pmatrix} \Sigma V & x' \\ 0 & x'' \end{pmatrix}, \quad x = \begin{pmatrix} x' \\ x'' \end{pmatrix}.$$

If $x \in S(M, R)$, it is immediate that $|x''| \lesssim \frac{R}{D_k(M)}$, by considering submatrices consisting of the first k rows, together with one of the last $p - k$ lines). Furthermore, by considering submatrices consisting of a $(k - 1) \times (k - 1)$ submatrix of ΣV , one of the $p - k$ last rows, and the last column, we have

$$|x''| D_{k-1}(M) \lesssim D_k(M).$$

It follows that $|x''| \lesssim \sigma_k(M)$.

We will now focus on the $k \times (k + 1)$ matrix made up of the first k rows of $(M|x)$, namely

$$\begin{pmatrix} \sigma_1(M)V^{(1)} & x_1 \\ \vdots & \vdots \\ \sigma_k(M)V^{(k)} & x_k \end{pmatrix}$$

(where $V^{(i)}$ stands for the i -th row of V). We now prove by induction on n that $|x_i| \leq \sigma_i(M)$ if $i \leq n$; this assertion for $n = k$ is the desired result. The case $n = 1$ being immediate, we can assume the assertion holds at rank n , and aim at proving it at rank $n + 1$.

The n first rows of V are orthogonal, therefore we can delete the last $n - k$ rows and some $k - n$ columns of V to get an $n \times n$ matrix with a determinant ~ 1 ; denote this matrix \tilde{V} and its rows $\tilde{V}^{(1)}, \dots, \tilde{V}^{(n)}$.

Note that the $n \times n$ matrix with rows $\sigma_1(M)\tilde{V}^{(1)}, \dots, \sigma_n(M)\tilde{V}^{(n)}$ has determinant $\sim D_n(M)$.

We now consider the submatrix of M obtained by deleting the last $n - k - 1$ rows and the same columns that were deleted from V to make \tilde{V} . That is,

$$\tilde{M} = \begin{pmatrix} \sigma_1(M)\tilde{V}^{(1)} & x^1 \\ \vdots & \vdots \\ \sigma_n(M)\tilde{V}^{(n)} & x^n \\ \sigma_{n+1}(M)\tilde{V}^{(n+1)} & x^{n+1} \end{pmatrix}.$$

We further write $\tilde{M}^{(i,n+1)}$ for the matrix \tilde{M} with i -th row and last column removed. Expanding the determinant of \tilde{M} with respect to the last column, we find that

$$\det(\tilde{M}) = \sum_{i=1}^{n+1} (-1)^{i+n+1} x^i \det \tilde{M}^{(i,n+1)} + x^{n+1} \det \tilde{M}^{(n+1,n+1)}.$$

By the induction assumption, $|x_i| \leq \sigma_i(M)$ for $1 \leq i \leq n$. Furthermore, $\det \tilde{M}^{(i,n+1)} \lesssim \frac{D_{n+1}(M)}{\sigma_i(M)}$ for $1 \leq i \leq n$, and we saw that $\det \tilde{M}^{(n+1,n+1)} \sim D_n(M)$. Finally, the definition of $S(M, R)$ requires that $\det(\tilde{M}) \lesssim D_{n+1}(M)$. Combining these observations and the above equality implies that, if $x \in S(M, R)$,

$$|x^{n+1}| D_n(M) \lesssim D_{n+1}(M), \quad \text{i.e.} \quad |x^{n+1}| \lesssim \sigma_{n+1}(M).$$

Step 2: general case $p \geq k + 1$. Then the singular value decomposition of M is $M = U \begin{pmatrix} \Sigma \\ 0 \end{pmatrix} V$. Setting $y = U^{-1}x$, we can write

$$D_j(UDV|x) = D_j(U(DV|y)) \sim D_j(DV|y), \quad \text{where} \quad D = \begin{pmatrix} \Sigma \\ 0 \end{pmatrix}.$$

Then $x \in S(M, R)$ if and only if $y \in S(DV, R) \subset \{y : |y_i| \lesssim \tau_i\}$. The desired result follows for $x = Uy$.

Step 3: the case $p \leq k$. Similarly to the case $p \geq k + 1$, one deals first with $U = \text{Id}$. Then

$$M = (\Sigma|0)V = (\Sigma V_1)$$

(where V_1 is the $p \times k$ upper submatrix of V). Then

$$(M|x) = (\Sigma V_1|x).$$

Proceeding as in Step 1, one can deduce that $|x_i| \lesssim \tau_i$ if $1 \leq i \leq p$, and the desired conclusion follows as in Step 2. □

We can apply the last lemma to Problem 1.4, with some technical complexity coming from the constant entries in the last row of the matrix P appearing there. In

the following lemma one should think of M as the first k columns of P , and x as being a column $(m_{1(k+1)}^2, \dots, m_{d(k+1)}^2, \lambda^2)^T$ to be adjoined to the matrix M . As the $m_{i(k+1)}$ range over integers of size $\sim \mu_i$, the vector $\mathcal{M}x = (m_{1(k+1)}^2 \mu_1^{-1}, \dots, m_{d(k+1)}^2 \mu_d^{-1})^T$ then takes values which are separated from each other by distances $\gtrsim 1$. In Sect. 6.2 we will use this to bound the number of integral $m_{i(k+1)}$ by the measure of a neighbourhood of the permissible real vectors $\mathcal{M}x$. It is this measure that is estimated in (5.9).

Lemma 5.7 *Adopting the notation of Lemma 5.6, let $\mu_1 \geq \dots \geq \mu_{p-1} > 0$ and let \mathcal{M} be the $(p - 1) \times p$ matrix defined by*

$$\mathcal{M} = (\text{diag}(\mu_1^{-1}, \dots, \mu_{p-1}^{-1})|0).$$

As in Lemma 5.6 let M be a $p \times k$ matrix, fix $C > 0$ and put

$$S(M, R) = \{x \in \mathbb{R}^p : D_j(M|x) \leq CD_j(M) \\ (1 \leq j \leq \min(p, k)), D_{k+1}(M|x) \leq R \ (p \geq k + 1)\},$$

and let $\tau_i = \sigma_i(M)$ or $\min(\sigma_k(M), \frac{R}{D_i(M)})$ for $i \leq k$ or $i > k$ respectively. Then, for any $A > 0$, if $M_{p1} > \epsilon \sigma_1(M)$ for some $\epsilon > 0$, then

$$\text{mes} \left\{ \mathcal{M}x + w : x \in S(M, R), x_p = A, |x_i| \in \left[\frac{\mu_i^2}{C}, C\mu_i^2 \right] \ (i < p), |w| \leq 1 \right\} \\ \lesssim_{\epsilon, p, k, C} 1 + \sum_{k=1}^{p-1} D_k(\tilde{W}), \tag{5.9}$$

where \tilde{W} is a $(p - 1) \times (p - 1)$ matrix with entries such that

$$|\tilde{W}_{ij}| \lesssim_{p, k, C} \mu_j^{-1} \min(\tau_{i+1}, \mu_j^2),$$

so that $D_k(\tilde{W}) \lesssim_{p, k, C} \max_{\substack{i_1, \dots, i_k \\ j_1, \dots, j_k \\ \text{distinct}}} \prod_{\ell=1}^k \mu_{j_\ell}^{-1} \min(\tau_{i_\ell+1}, \mu_{j_\ell}^2)$.

Proof In the proof we allow all implicit constants in \lesssim, \sim notation to depend on C, p, k . Taking the difference of two vectors in the set on the right-hand side of (5.9), we see that it suffices to prove the desired statement for $A = 0$, and the condition $|x_i| \sim \mu_i^2$ replaced by $|x_i| \lesssim \mu_i^2$. In other words, it suffices to prove

$$\text{mes}\{ \mathcal{M}x + w : x \in S(M, R), x_p = 0, |x_i| \lesssim \mu_i^2 \ (i < p), |w| \leq 1 \} \lesssim_\epsilon 1 + \sum_{k=1}^{p-1} D_k(\tilde{W}).$$

Define the projector P on the first $p - 1$ coordinates of a vector of \mathbb{R}^p :

$$P((x_1, \dots, x_p)^T) = (x_1, \dots, x_{p-1}, 0)^T.$$

Let U, V be matrices as in (5.1) and let $U^{(i)}$ be the i th column of U . Since $M_{p1} > \epsilon \sigma_1(M)$ there is i_0 such that $[MV^{-1}]_{pi_0} \gtrsim \epsilon \sigma_1(M)$, that is to say $U_p^{(i_0)} \sigma_{i_0}(M) \gtrsim \epsilon \sigma_1(M)$, whence $U_p^{(i_0)} \gtrsim_{\epsilon} 1$ and $\sigma_{i_0}(M) \gtrsim_{\epsilon} \sigma_1(M)$.

By Lemma 5.6,

$$\{x \in S(M, R) : x_p = 0\} \subset \left\{ x : x = \sum_{i=1}^p y_i U^{(i)} : |y_i| \lesssim \tau_i, x_p = 0 \right\}.$$

Since the p -th coordinate of $\sum_{i=1}^p y_i U^{(i)}$ is 0, we find that

$$P \left[\sum_{i=1}^p y_i U^{(i)} \right] = \sum_{i=1}^{p-1} \tilde{y}_i \tilde{U}^{(i)},$$

where

$$\begin{aligned} \tilde{U}^{(i)} &= P \left[U^{(i)} - \frac{U_p^{(i)}}{U_p^{(i_0)}} U^{(i_0)} \right], \quad \tilde{y}_i = y_i \quad (i < i_0), \\ \tilde{U}^{(i)} &= P \left[U^{(i+1)} - \frac{U_p^{(i+1)}}{U_p^{(i_0)}} U^{(i_0)} \right], \quad \tilde{y}_i = y_{i+1} \quad (i \geq i_0), \end{aligned}$$

and our choice of i_0 above ensures that $|\tilde{U}^{(i)}| \lesssim_{\epsilon} 1$. Therefore,

$$\{Px : x \in S(M, R), x_p = 0\} \subset \left\{ x : x = \sum_{i=1}^{p-1} \tilde{y}_i \tilde{U}^{(i)}, |\tilde{y}_i| \lesssim \tau_{i+1} \right\}.$$

We now add the condition $|x_i| \lesssim \mu_i^2$ to obtain

$$\begin{aligned} &\{Px : x \in S(M, R), |x_i| \lesssim \mu_i^2 \text{ if } 1 \leq i \leq p-1, x_p = 0\} \\ &\subset \left\{ x : x = \sum_{i=1}^{p-1} \tilde{y}_i \tilde{U}^{(i)}, |\tilde{y}_i| \lesssim \tau_{i+1}, |x_i| \lesssim \mu_i^2 \right\} \\ &\subset \left\{ \sum_{t=1}^{p-1} t_i w^{(i)} : |t_i| \leq 1 \right\} \quad \text{with } |w_j^{(i)}| \lesssim \min(\tau_{i+1}, \mu_j^2), \end{aligned}$$

where the last inclusion is a consequence of Lemma 5.5. Applying the matrix \mathcal{M} , we see that

$$\{\mathcal{M}x : x \in S(M, R), |x_i| \lesssim \mu_i^2 \text{ if } 1 \leq i \leq p-1, x_p = 0\}$$

$$\subset \left\{ \sum_{t=1}^{p-1} t_i \tilde{w}^{(i)} : |t_i| \leq 1 \right\} \quad \text{with} \quad |\tilde{w}_j^{(i)}| = |[\mathcal{M}w^{(i)}]_j| \lesssim \mu_j^{-1} \min(\tau_{i+1}, \mu_j^2).$$

Finally, Lemma 5.4 gives the desired conclusion. □

6 Proof of Theorem 1.3

In this section we will prove the following result.

Theorem 6.1 *The following holds for any fixed off-diagonal coefficients $\beta_{ij} = \beta_{ji} \in [-\frac{1}{10d^2}, \frac{1}{10d^2}]$ ($i < j$) and almost all $(\beta_{11}, \dots, \beta_{dd})^T \in [1, 2]^d$. Moreover it also holds for almost all symmetric matrices with $\beta_{ji} \in [-\frac{1}{10d^2}, \frac{1}{10d^2}]$ for $i \neq j$ and $\beta_{11}, \dots, \beta_{dd} \in [1, 2]$.*

For every $b \in \mathbb{N}$, every $\delta < 1 < \lambda$ and any $\epsilon > 0$, we have

$$\|P_{\lambda, \delta}\|_{L^1 \rightarrow L^\infty} \lesssim_{\beta, \epsilon} \delta^{-1/b} \lambda^\epsilon \left[\max_{0 \leq b_2 \leq \min(b, d+1)} \delta^{b_2} \lambda^{-b_2 + (b+d)b_2 + 1 - b} \right]^{1/b}. \tag{6.1}$$

It follows in particular that, for every $a \in \mathbb{N}$ and whenever $\lambda^{-a} \leq \delta \leq \lambda^{1-a}$, we have

$$\|P_{\lambda, \delta}\|_{L^1 \rightarrow L^\infty} \lesssim_{\beta, \epsilon} \begin{cases} \delta^{1 - \frac{1}{d+1-a}} \lambda^{d-1 + \frac{1}{d+1-a} + \epsilon} & (a < d), \\ \delta^{1 - \frac{1+a}{d+1+a}} \lambda^{d-1 + \frac{1+a}{d+1+a} + \epsilon} & (a \geq d). \end{cases} \tag{6.2}$$

We emphasize that (6.2) is intended as illustrative; one could prove a stronger but less tidy result just by making a more careful choice of the parameter b at the end of the proof.

By contrast it seems more challenging to improve (6.1) using our methods. See Remark 6.2 for one idea.

Before proceeding to the proof we deduce Theorem 1.3.

Proof of Theorem 1.3 Equations (1.10) and (1.9) are cases of (6.2). For (1.8), observe that if $\|P_{\lambda, \delta}\|$ is sufficiently large in terms of the function χ , then there are two integer vectors with

$$\mathcal{Q}(x^{(1)})^{1/2}, \mathcal{Q}(x^{(2)})^{1/2} \in [\lambda - \delta, \lambda + \delta], \quad \text{and} \quad |x_i^{(1)}| \neq |x_i^{(2)}| \quad \text{for some } i.$$

Letting $y_{ij} = x_i^{(1)}x_j^{(1)} - x_i^{(2)}x_j^{(2)}$ we find that

$$y_{ij} \in \mathbb{Z}, \quad |y_{ij}| \lesssim \lambda^2, \quad y_{ii} \neq 0 \text{ for some } i, \quad \left| \sum_{i,j} \beta_{ij} y_{ij} \right| \lesssim \delta \lambda. \tag{6.3}$$

For a given matrix y_{ij} and for any off-diagonal coefficients $\beta_{ij} = \beta_{ji}$ ($i < j$) we have

$$\text{mes}\left\{(\beta_{11}, \dots, \beta_{dd})^T \in [1, 2]^d : \left| \sum_{i,j} \beta_{ij} y_{ij} \right| \leq \delta \lambda\right\} \lesssim \frac{\delta \lambda}{\max\{|y_{ij}|\}}.$$

We claim that

$$\#\{(y_{ij}) : y_{ij} = x_i^{(1)} x_j^{(1)} - x_i^{(2)} x_j^{(2)} \text{ for some } x^{(i)} \in \mathbb{Z}^d, \text{ and } 0 < \max\{|y_{ij}|\} \leq Y\} \lesssim_\epsilon Y^{d+\epsilon}.$$

For the proof, note that we cannot have every $y_{ii} = 0$ or else (y_{ij}) would vanish. We assume without loss of generality that there is $i_0 \in \{1, \dots, d\}$ such that $y_{ii} = 0$ iff $i > i_0$. There are $O(Y^{i_0})$ possible values of y_{ii} for $i \leq i_0$, and once these are chosen the identity $y_{ii} = (x_i^{(1)} + x_i^{(2)})(x_i^{(1)} - x_i^{(2)})$ determines $x_i^{(1)}, x_i^{(2)}$ for $i \leq i_0$ up to Y^ϵ possibilities. Next for $i > i_0$ we have $x_i^{(1)} = \pm x_i^{(2)}$, and so up to finitely many choices both of these are determined by the values of $y_{1i} = (\pm x_1^{(1)} \pm x_1^{(2)})x_i^{(1)}$, for which there are $O(Y^{d-i_0})$ possibilities. We conclude that there are $\lesssim_\epsilon Y^{d+\epsilon}$ choices for the $x^{(i)}$ and hence for (y_{ij}) .

We can now conclude that, for a suitably large constant C depending only on the cutoff function χ , and for any fixed values of the off-diagonal entries $|\beta_{ij}| \leq \frac{1}{10d^2}$ ($i < j$), we have

$$\begin{aligned} &\text{mes}\{(\beta_{11}, \dots, \beta_{dd})^T \in [1, 2]^d : \|P_{\lambda,\delta}\| > C \text{ for some } \lambda \sim \lambda_0, \delta < \delta_0\} \\ &\lesssim \sum_{\substack{(y_{ij}) \\ 0 < \max\{|y_{ij}|\} \lesssim \lambda_0^2}} \frac{\delta_0 \lambda_0}{\max\{|y_{ij}|\}} \\ &\lesssim_\epsilon \sum_{\substack{Y=2^k \\ k \in \mathbb{Z} \\ Y \lesssim \lambda_0^2}} \delta_0 \lambda_0 Y^{d-1+\epsilon}. \end{aligned}$$

This is $O_\epsilon(\delta_0 \lambda_0^{2d-1})$. Applying the Borel-Cantelli lemma (Lemma A.1) proves (1.8). □

We now begin the proof of Theorem 6.1. Throughout the rest of this section we write β_i for β_{ii} , put $\beta' = (\beta_1, \dots, \beta_d)^T$, and given $b, d \in \mathbb{N}$ and $M = (m_{ij})_{1 \leq i \leq d, 1 \leq j \leq b}$, we put

$$P(M) = (p_{ij}(M))_{1 \leq i \leq d+1, 1 \leq j \leq b} = \begin{pmatrix} m_{11}^2 & \dots & m_{1b}^2 \\ \vdots & \ddots & \vdots \\ m_{d1}^2 & \dots & m_{db}^2 \\ \lambda_0^2 & \dots & \lambda_0^2 \end{pmatrix}.$$

6.1 Integrating over λ and β

Our key observation is as follows. Since, for $m \in \mathbb{Z}$, we have $1 \leq \sum_{\mu \in 2^{\mathbb{N}} \cup \{0\}} \mathbf{1}_{\mu \leq 2m \leq 2\mu}$, and since χ takes non-negative values, we have for any $\lambda, \delta > 0$ that

$$\begin{aligned} & \int_{[1,2]^d} \int_{\lambda_0/2}^{\lambda_0} \|P_{\lambda,\delta}\|_{L^1 \rightarrow L^\infty}^b d\lambda d\beta' \\ &= \int_{[1,2]^d} \int_{\lambda_0/2}^{\lambda_0} \left(\sum_{\substack{\lambda_0 \geq \mu_1, \dots, \mu_d \geq 0 \\ \mu_i \in 2^{\mathbb{N}} \cup \{0\}}} \sum_{\substack{m \in \mathbb{Z}^d \\ |m_i| \in [\frac{\mu_i}{2}, \mu_i]}} \chi \left(\frac{Q(m) - \lambda^2}{\delta \lambda} \right) \right)^b d\lambda d\beta' \\ &\lesssim \int_{[1,2]^d} \int_{\lambda_0/2}^{\lambda_0} \log^{bd}(\lambda_0) \max_{\substack{\lambda_0 \geq \mu_1 \geq \dots \geq \mu_d \geq 0 \\ \mu_i \in 2^{\mathbb{N}} \cup \{0\}}} \left(\sum_{\substack{m \in \mathbb{Z}^d \\ |m_i| \in [\frac{\mu_i}{2}, \mu_i]}} \chi \left(\frac{Q(m) - \lambda^2}{\delta \lambda} \right) \right)^b d\lambda d\beta' \\ &\lesssim \int_{[1,2]^d} \int_{\lambda_0/2}^{\lambda_0} \log^{bd}(\lambda_0) \max_{\substack{\lambda_0 \geq \mu_1 \geq \dots \geq \mu_d \geq 0 \\ \mu_i \in 2^{\mathbb{N}} \cup \{0\}}} \prod_{j=1}^b \sum_{\substack{m_{1j}, \dots, m_{dj} \in \mathbb{Z} \\ |m_{ij}| \in [\frac{\mu_i}{2}, \mu_i]}} \mathbf{1}_{|Q(m_{1j}, \dots, m_{dj}) - \lambda^2| \leq \delta \lambda} d\lambda d\beta', \end{aligned}$$

and if we temporarily write the off-diagonal parts of $Q(m_{1j}, \dots, m_{dj})$ using the row vector

$$q = \left(\sum_{\substack{1 \leq i, j \leq d \\ i \neq j}} \beta_{ij} m_{i1} m_{j1}, \dots, \sum_{\substack{1 \leq i, j \leq d \\ i \neq j}} \beta_{ij} m_{ib} m_{jb} \right).$$

then this becomes

$$\begin{aligned} & \int_{[1,2]^d} \int_{\lambda_0/2}^{\lambda_0} \|P_{\lambda,\delta}\|_{L^1 \rightarrow L^\infty}^b d\lambda d\beta' \\ &\lesssim \int_{[1,2]^d} \int_{\lambda_0/2}^{\lambda_0} \log^{bd}(\lambda_0) \max_{\substack{\lambda_0 \geq \mu_1 \geq \dots \geq \mu_d \geq 0 \\ \mu_i \in 2^{\mathbb{N}} \cup \{0\}}} \prod_{j=1}^b \sum_{\substack{m_{1j}, \dots, m_{dj} \in \mathbb{Z} \\ |m_{ij}| \in [\frac{\mu_i}{2}, \mu_i]}} \mathbf{1}_{|(\beta_1, \dots, \beta_d, -\lambda^2/\lambda_0^2)P(M) - q| \leq \delta \lambda_0} \\ & d\lambda d\beta' = \log^{bd}(\lambda_0) \max_{\substack{\lambda_0 \geq \mu_1 \geq \dots \geq \mu_d \geq 0 \\ \mu_i \in 2^{\mathbb{N}} \cup \{0\}}} \sum_{\substack{M \in \mathbb{Z}^{d \times b} \\ |m_{ij}| \in [\frac{\mu_i}{2}, \mu_i]}} \text{mes} \left\{ (\beta', \lambda) \in [1, 2]^d \times \left[\frac{\lambda_0}{2}, \lambda_0 \right] : |(\beta_1, \dots, \beta_d, -\lambda^2/\lambda_0^2)P(M) - q| \leq \delta \lambda_0 \right\}, \end{aligned} \tag{6.4}$$

We can estimate the measure inside the last sum in (6.4) as follows. Notice first that

$$D_1(P(M)) = \max_{\substack{i=1, \dots, d+1 \\ j=1, \dots, b}} |p_{ij}(M)| \sim \lambda_0^2. \tag{6.5}$$

By Lemma 5.1, we have

$$\begin{aligned} & \text{mes} \left\{ \gamma \in \mathbb{R}^{d+1} : |\gamma| \leq 1, \left| \sum_i \gamma_i p_{ik}(M) - \sum_{i \neq j} \beta_{ij} m_{ik} m_{jk} \right| \leq \delta \lambda_0 \right\} \\ & \lesssim \prod_{k=1}^{\min(b,d+1)} \min \left(1, \frac{\delta \lambda_0}{\sigma_k(P(M))} \right) = \prod_{\substack{1 \leq i \leq \min(b,d+1) \\ \sigma_i(P(M)) > \delta \lambda_0}} \frac{\delta \lambda_0}{\sigma_i(P(M))}. \end{aligned}$$

Together with (6.4) and the fact that $P(M)$ does not depend on the signs of the m_{ij} , we find

$$\begin{aligned} & \int_{[1,2]^d} \int_{\lambda_0/2}^{\lambda_0} \|P_{\lambda,\delta}\|_{L^1 \rightarrow L^\infty}^b d\lambda d\beta' \\ & \lesssim (\log \lambda_0)^{bd+d} \max_{\substack{\lambda_0 \geq \mu_1 \geq \dots \geq \mu_d \geq 0 \\ \mu_i \in 2^{\mathbb{N}} \cup \{0\}}} \sum_{\substack{M \in \mathbb{Z}^{d \times b} \\ m_{ij} \in [\mu_i/2, \mu_i]}} \lambda_0 \prod_{\substack{1 \leq i \leq \min(b,d+1) \\ \sigma_i(P(M)) > \delta \lambda_0}} \frac{\delta \lambda_0}{\sigma_i(P(M))}, \end{aligned}$$

where the m_{ij} are now non-negative since $m_{ij} \in [\mu_i/2, \mu_i]$. Combining this with (6.5) and the Corollary to Lemma 5.2 yields

$$\begin{aligned} & \int_{[1,2]^d} \int_{\lambda_0/2}^{\lambda_0} \|P_{\lambda,\delta}\|_{L^1 \rightarrow L^\infty}^b d\lambda d\beta' \\ & \lesssim (\log \lambda_0)^{bd+d} \max_{\substack{\lambda_0^2 \sim L_1 \geq \dots \geq L_{\min(b,d+1)} \geq 0 \\ \lambda_0 \geq \mu_1 \geq \dots \geq \mu_d \geq 0 \\ \mu_i \in 2^{\mathbb{N}} \cup \{0\}}} \lambda_0 \prod_{\substack{1 \leq i \leq \min(b,d+1) \\ L_i > \delta \lambda_0}} \frac{\delta \lambda_0}{L_i} Z_{d,b}(\vec{\mu}, \vec{L}), \quad (6.6) \end{aligned}$$

where

$$\begin{aligned} Z_{d,b}(\vec{\mu}, \vec{L}) &= \#\left\{ M \in \mathbb{Z}^{d \times b} : \frac{m_{ij}}{\mu_i} \in \left[\frac{1}{2}, 1\right], \frac{D_k(P(M))}{L_1 \cdots L_k} \in \left[\frac{1}{2}, 1\right] \right. \\ & \quad \left. \text{for all } 1 \leq i \leq d, 1 \leq j \leq b, 1 \leq k \leq \min(b, d+1) \right\}. \quad (6.7) \end{aligned}$$

In (6.6) we may assume that $(\mu_i = 0 \implies L_{i+1} = 0)$, since otherwise $Z_{d,b}(\vec{\mu}, \vec{L})$ would be zero (there are no such M). In particular allowing μ_i to be zero is the same as allowing the dimension d to drop, in the sense that

$$\max_{\substack{\lambda_0^2 \sim L_1 \geq \dots \geq L_{\min(b,d+1)} \geq 0 \\ \lambda_0 \geq \mu_1 \geq \dots \geq \mu_d \geq 0 \\ \mu_i \in 2^{\mathbb{N}} \cup \{0\}}} \lambda_0 \left[\prod_{\substack{1 \leq i \leq \min(b,d+1) \\ L_i > \delta \lambda_0}} \frac{\delta \lambda_0}{L_i} \right] Z_{d,b}(\vec{\mu}, \vec{L})$$

$$= \max_{\substack{0 \leq d' \leq d \\ \lambda_0^2 \sim L_1 \geq \dots \geq L_{\min(b, d'+1)} \geq 0 \\ \lambda_0 \geq \mu_1 \geq \dots \geq \mu_{d'} \geq 1, \mu_i \in 2^{\mathbb{N}}}} \lambda_0 \left[\prod_{\substack{1 \leq i \leq \min(b, d'+1) \\ L_i > \delta \lambda_0}} \frac{\delta \lambda_0}{L_i} \right] Z_{d', b}(\vec{\mu}, \vec{L}).$$

6.2 Counting matrices with prescribed subdeterminants

We want to estimate the right-hand side of (6.6), under the assumption that $\mu_i \neq 0$ for any i , or in other words $\mu_i \in 2^{\mathbb{N}}$.

Our first object is to estimate from above the number of matrices M counted by the function $Z_{d, b}(\vec{\mu}, \vec{L})$ from (6.7). By Lemma 5.3, it suffices to count those M satisfying an additional condition

$$D_k(P(M)) \sim D_k^{(k)}(P(M)) \text{ for all } k \leq \min(b, d + 1),$$

since permuting the columns of these recovers all the matrices in $Z_{d, b}(\vec{\mu}, \vec{L})$

For $j = 1, \dots, b$ define the vectors $m^{(j)}$ and $n^{(j)} \in \mathbb{R}^d$ by

$$m^{(j)} = (m_{1j}, \dots, m_{dj})^T, \quad n^{(j)} = (m_{1j}^2/\mu_1, \dots, m_{dj}^2/\mu_d)^T.$$

That is the vectors $m^{(j)}$ are the columns of M . Meanwhile the vectors $n^{(j)}$ are the columns of $P(M)$ with the last element dropped and the others rescaled so that $n^{(j)}$ belongs to the set S defined by

$$S = \left\{ (u_1^2/\mu_1, \dots, u_d^2/\mu_d)^T : u_i \in \mathbb{Z}, u_i \in \left[\frac{\mu_i}{2}, \mu_i \right] \right\},$$

whose elements are separated by gaps of size ~ 1 . If M is counted in the right-hand side of (6.7), then the vector $n^{(1)}$ can be chosen arbitrarily from S ; there are $\lesssim \prod_{i=1}^d \mu_i$ choices.

Suppose now that the first k columns of $P(M)$ are given, and they satisfy

$$D_\ell(P(m^{(1)} | \dots | m^{(k)})) \sim L_1 \dots L_\ell \quad (1 \leq \ell \leq \min(k, d + 1)).$$

We want to select $m^{(k+1)}$, or equivalently $n^{(k+1)}$. We can first use that S is 1-separated to replace our counting problem by a volume estimate: letting

$$\mathcal{N}^{k+1} = \{n^{(k+1)} \in S : m_i^{(k)} \in [\mu_i/2, \mu_i], D_\ell(P(m^{(1)}, \dots, m^{(k+1)})) \sim L_1 \dots L_\ell \\ (1 \leq i \leq d, 1 \leq \ell \leq \min(k + 1, d + 1))\},$$

then

$$\#\mathcal{N}^{k+1} \lesssim \text{mes} \left[\mathcal{N}^{k+1} + B(0, 1) \right].$$

Remark 6.2 This volume bound is not necessarily optimal. To take just one simple example, if λ_0^2 is an integer then $D_\ell^{(\ell)}(P(M))$ is an integer. Thus, in the case when $0 < L_1 \dots L_\ell \ll 1$, it is impossible for $D_\ell^{(\ell)}(P(M)) \sim L_1 \dots L_\ell$ to hold and the set \mathcal{N}^{k+1} is empty.

We apply Lemma 5.7 with

$$p = d + 1, \quad M = P(m^{(1)} | \dots | m^{(k)}), \quad \mathcal{M}x = n^{(k+1)}, \quad A = \lambda_0^2, \quad R = L_1 \dots L_{k+1}.$$

We compute that

$$\tau_i \asymp \begin{cases} \max\{L_i, L_{k+1}\} & (k \leq d), \\ L_i & (k > d). \end{cases}$$

We now need to distinguish two cases. If $k \leq d$ then applying Lemma 5.7 gives that

$$\#\mathcal{N}^{k+1} \lesssim 1 + \max_{\substack{\mathcal{I} \subset \{1, \dots, d\} \\ \mathcal{J} \subset \{1, \dots, d\} \\ \#\mathcal{I} = \#\mathcal{J} \\ \sigma: \mathcal{I} \hookrightarrow \mathcal{J}}} \left(\prod_{i \in \mathcal{I}} \mu_i \right) \left(\prod_{\substack{\mu_i^2 > \max(L_{\sigma(i)+1}, L_{k+1}) \\ i \in \mathcal{I}}} \frac{\max(L_{\sigma(i)+1}, L_{k+1})}{\mu_i^2} \right).$$

If instead $k > d$, applying Lemma 5.7 gives that

$$\#\mathcal{N}^{k+1} \lesssim 1 + \max_{\substack{\mathcal{I} \subset \{1, \dots, d\} \\ \mathcal{J} \subset \{1, \dots, d\} \\ \#\mathcal{I} = \#\mathcal{J} \\ \sigma: \mathcal{I} \hookrightarrow \mathcal{J}}} \left(\prod_{i \in \mathcal{I}} \mu_i \right) \left(\prod_{\substack{\mu_i^2 > L_{\sigma(i)+1} \\ i \in \mathcal{I}}} \frac{L_{\sigma(i)+1}}{\mu_i^2} \right).$$

Recall now that this is a bound for the number of choices for $m^{(k+1)}$, given $m^{(1)}, \dots, m^{(k)}$, and that there are $\mu_1 \dots \mu_d$ choices for $m^{(1)}$. Recall also that our object is to estimate that part of the right-hand side of (6.6) for which every μ_i is nonzero. The bound we have proved is

$$\max_{\substack{\lambda_0^2 \sim L_1 \geq \dots \geq L_{\min(b, d+1)} \geq 0 \\ \lambda_0 \geq \mu_1 \geq \dots \geq \mu_d > 0 \\ \mu_i \in 2^{\mathbb{N}}}} \lambda_0 \prod_{\substack{1 \leq i \leq \min(b, d+1) \\ L_i > \delta \lambda_0}} \frac{\delta \lambda_0}{L_i} Z_{d, b}(\vec{\mu}, \vec{L})$$

$$\lesssim \max_{\substack{\lambda_0^2 \sim L_1 \geq \dots \geq L_{\min(b,d+1)} \geq 0 \\ \lambda_0 \geq \mu_1 \geq \dots \geq \mu_d \geq 0}} \max_{\substack{\mathcal{I}_k \subset \{1, \dots, d\} \\ \sigma_k: \mathcal{I}_k \hookrightarrow \{1, \dots, d\}}} \mathcal{F}((L_i), (\mu_i), (\mathcal{I}_k), (\sigma_k)), \tag{6.8}$$

where σ_k is a bijection from \mathcal{I}_k to a subset of $\{1, \dots, d\}$ and

$$\begin{aligned} & \mathcal{F}((L_i), (\mu_i), (\mathcal{I}_k), (\sigma_k)) \\ &= \lambda_0 \left(\prod_{\substack{i=1, \dots, \min(b,d+1) \\ L_i > \delta \lambda_0}} \frac{\delta \lambda_0}{L_i} \right) \left(\prod_{i=1}^d \mu_i \right) \prod_{k=1}^{\min(b-1,d)} \left(\prod_{i \in \mathcal{I}_k} \mu_i \right) \\ & \quad \left(\prod_{\substack{\mu_i^2 > \max\{L_{\sigma_k(i)+1}, L_{k+1}\} \\ i \in \mathcal{I}_k}} \frac{\max(L_{\sigma_k(i)+1}, L_{k+1})}{\mu_i^2} \right) \\ & \quad \prod_{k=\min(b,d+1)}^{b-1} \left(\prod_{i \in \mathcal{I}_k} \mu_i \right) \left(\prod_{\substack{\mu_i^2 > L_{\sigma_k(i)+1} \\ i \in \mathcal{I}_k}} \frac{L_{\sigma_k(i)+1}}{\mu_i^2} \right), \end{aligned}$$

with the understanding that \mathcal{I} might be empty and $\prod_{\mathcal{I}} = 1$ if $\mathcal{I} = \emptyset$.

6.3 The maximization procedure

Our aim is now to find the values of $L_i, \mu_i, \mathcal{I}_k, \sigma_k$ for which the maximum on the right-hand side of (6.8) is attained.

Step 1: Maximizing in (L_i) with the \mathcal{I}'_k s and μ'_i s held fixed. We start with the dependence on (L_i) , and we relax first the condition that they be ordered; we will simply assume that $0 \leq L_i \leq \lambda_0^2$ for each $i = 1, \dots, \min(b, d + 1)$. Next, we claim that the products

$$\left(\prod_{\substack{\mu_i^2 > \max(L_{\sigma_k(i)+1}, L_{k+1}) \\ i \in \mathcal{I}_k}} \frac{\max(L_{\sigma_k(i)+1}, L_{k+1})}{\mu_i^2} \right) \text{ and } \left(\prod_{\substack{\mu_i^2 > L_{\sigma_k(i)+1} \\ i \in \mathcal{I}_k}} \frac{L_{\sigma_k(i)+1}}{\mu_i^2} \right)$$

on the right-hand side of the definition of $\mathcal{F}((L_i), (\mu_i), (\mathcal{I}_k), (\sigma_k))$ can be taken to be empty. Indeed, assume that the maximum of \mathcal{F} is reached at a point such that, for some k in the product, the corresponding product is not empty: it contains $\frac{\max(L_{\sigma_k(i)+1}, L_{k+1})}{\mu_i^2}$

or $\frac{L_{\sigma_k(i)+1}}{\mu_i^2}$ for some i ; we will denote it $\frac{L_{i_0}}{\mu_{i_1}^2}$. Expand \mathcal{F} in powers of the L_i , and consider the exponent of L_{i_0} . It is necessarily ≥ 0 , since only the first factor in the definition of \mathcal{F} contributes a negative power, namely $L_{i_0}^{-1}$. Therefore, \mathcal{F} will be larger, or equal, if we increase the value of $\max(L_{\sigma_k(i)+1}, L_k)$ to $\mu_{i_1}^2$, which has the effect of cancelling the undesirable term.

After this manipulation, the parentheses we mentioned have been cancelled, and the value of some of the L_i 's has been fixed to $\mu_{f(i)}^2$ for some function f . The remaining

L_i contribute $\mathbf{1}_{L_i > \delta\lambda_0} \frac{\delta\lambda_0}{L_i}$, and they might be constrained by inequalities of the type $L_i > \mu_j^2$. Therefore, \mathcal{F} will be maximal if they take the value $\delta\lambda_0$, or $\mu_{f(i)}^2$ for some function f .

Step 2: Maximizing in (μ_i) . The result of the maximization in (L_i) is that we can assume that each L_i takes the value either $\mu_{f(i)}^2$, or $\delta\lambda_0$, that $\mu_i^2 \leq \max(L_{\sigma_k(i)+1}, L_{k+1})$ if $i \in \mathcal{I}_k$, with the convention that $L_k = 0$ for $k \geq d + 2$, and that the function to maximize is

$$\lambda_0 \left(\prod_{\substack{i=1, \dots, \min(b, d+1) \\ L_i > \delta\lambda_0}} \frac{\delta\lambda_0}{L_i} \right) \left(\prod_{i=1}^d \mu_i \right) \prod_{k=1}^{b-1} \left(\prod_{i \in \mathcal{I}_k} \mu_i \right). \tag{6.9}$$

We now claim that, at the maximum, the (μ_i) either take the value 1 or λ_0 . To prove this claim, assume that, at the maximum, the μ_i take a number n of distinct values $1 \leq a_1 < \dots < a_n \leq \lambda_0$. Replacing L_i by $\mu_{f(i)}^2$ in the above expression, it takes the form

$$\lambda_0 (\lambda_0 \delta)^{\alpha_0} \prod_{i=1}^n a_i^{\alpha_i}, \quad \text{where } \alpha_i \in \mathbb{Z}.$$

If $\alpha_i > 0$ and $a_i < \lambda_0$, then this expression will increase if the value of a_i is increased until a_{i+1} or λ_0 ; and similarly, if $\alpha_i < 0$ and $a_i > 1$, it will decrease if the value of a_i is decreased until a_{i-1} or 1. This contradicts the maximality of (μ_i) unless a_i only takes the values λ_0 or 1 for $\alpha_i \neq 0$. There remains the case where $\alpha_i = 0$, but then μ_i can be assigned the value λ_0 or 1 indifferently.

Step 3: Maximizing in (\mathcal{I}_k) and (σ_k) . We showed that the maximum of \mathcal{F} is less than the maximum of (6.9), under the constraint that $\mu_i^2 \leq \max(L_{\sigma_k(i)+1}, L_{k+1})$ if $i \in \mathcal{I}_k$ (with the convention that $L_k = 0$ for $k \geq d + 2$); and under the further constraint that L_i can only take the values $\delta\lambda_0, 1, \lambda_0$, and μ_i can only take the values 1, λ_0 .

There are now two cases to consider:

- If $k \leq \min(b - 1, d)$ and $L_{k+1} = \lambda_0^2$, then the optimal choice for \mathcal{I}_k is $\{1, \dots, d\}$.
- Otherwise, \mathcal{I}_k should have the same cardinal as the set of $L_i, i \geq 2$, equal to λ_0^2 , and $\sigma_k + 1$ should map \mathcal{I}_k to this set.

Step 4: Conclusion. As a result of the previous reductions, we find

$$\begin{aligned} & \max_{\substack{\lambda_0^2 \sim L_1 \geq \dots \geq L_{\min(b, d+1)} \geq 0 \\ \lambda_0 \geq \mu_1 \geq \dots \geq \mu_d \geq 0}} \max_{\substack{\mathcal{I}_k \subset \{1, \dots, d\} \\ \sigma_k: \mathcal{I}_k \hookrightarrow \{1, \dots, d\}}} \mathcal{F}((L_i), (\mu_i), (\mathcal{I}_k), (\sigma_k)) \\ & \lesssim \max_{\substack{\lambda_0 = L_1 \geq \dots \geq L_{\min(b, d+1)} \\ \mu_1 \geq \dots \geq \mu_d \\ L_i \in \{\delta\lambda_0, 1, \lambda_0^2\}, \mu_i \in \{1, \lambda_0\}}} \lambda_0 \left(\prod_{\substack{i=1, \dots, \min(b, d+1) \\ L_i > \delta\lambda_0}} \frac{\delta\lambda_0}{L_i} \right) \left(\prod_{i=1}^d \mu_i \right) \prod_{k=1}^{b-1} \left(\prod_{\substack{i \in \{1, \dots, d\} \\ L_{i+1} = \lambda_0^2, \text{ or} \\ (k \leq d \text{ and } L_{k+1} = \lambda_0^2)}} \mu_i \right). \end{aligned}$$

Notice that we are assuming again that the (μ_i) and (L_i) are ordered; a moment of reflection shows that this is possible since the permutation (σ_k) can be freely chosen. This expression is visibly nondecreasing in (μ_i) , so we might as well take all μ_i to be λ_0 .

In order to evaluate the resulting expression, we need to know the number of L_k equal to, respectively, $\delta\lambda_0$, 1 , λ_0 ; this is also the information needed to determine \mathcal{I}_k ; therefore, we define the numbers

$$\begin{aligned} b_0 &= \#\{1 \leq j \leq \min(b, d + 1) : L_j = \delta\lambda_0\}, \\ b_1 &= \#\{1 \leq j \leq \min(b, d + 1) : L_j = 1\}, \\ b_2 &= \#\{1 \leq j \leq \min(b, d + 1) : L_j = \lambda_0^2\}, \end{aligned}$$

which are such that

$$b_0 + b_1 + b_2 = \min(b, d + 1).$$

Letting $\chi = 1$ if $1 > \delta\lambda_0$ and $\chi = 0$ otherwise, the maximization procedure shows that the function to be optimized is bounded by

$$\begin{aligned} &\max_{b_0+b_1+b_2=\min(b,d+1)} \lambda_0(\delta\lambda_0)^{b_1\chi+b_2}\lambda_0^{-2b_2}\lambda_0^d\lambda_0^{(b_2-1)d}\lambda_0^{(b-b_2)(b_2-1)} \\ &= \max_{b_1+b_2\leq\min(b,d+1)} (\delta\lambda_0)^{b_1\chi}\delta^{b_2}\lambda_0^{-b_2^2+(b+d)b_2+1-b} \end{aligned}$$

We notice first that b_1 can be taken to be zero. Second, there remains to dispose of the assumption that all μ_i are non-zero, which was made at the beginning of Sect. 6.2. By the comments just prior to the start of Sect. 6.2, this equivalent to reducing the dimension d . But some thought shows that the above expression is increasing with d , so that allowing for smaller d is harmless. Overall, the final bound we find is

$$\begin{aligned} &\max_{\substack{\lambda_0^2 \sim L_1 \geq \dots \geq L_{\min(b,d+1)} \geq 0 \\ \lambda_0 \geq \mu_1 \geq \dots \geq \mu_d \geq 0}} \max_{\substack{\mathcal{I}_k \subset \{1, \dots, d\} \\ \sigma_k: \mathcal{I}_k \leftrightarrow \{1, \dots, d\}}} \mathcal{F}((L_i), (\mu_i), (\mathcal{I}_k), (\sigma_k)) \\ &\lesssim \max_{b_2 \leq \min(b,d+1)} \delta^{b_2}\lambda_0^{-b_2^2+(b+d)b_2+1-b}. \end{aligned}$$

6.4 Borel-Cantelli and the end of the argument

The three previous subsections give the estimate

$$\int_{[1,2]^d} \int_{\lambda_0/2}^{\lambda_0} \|P_{\lambda,\delta}\|_{L^1 \rightarrow L^\infty}^b d\lambda d\beta' \lesssim (\log \lambda_0)^{bd+d} \max_{b_2 \leq \min(b,d+1)} \delta^{b_2}\lambda_0^{-b_2^2+(b+d)b_2+1-b},$$

valid uniformly for any off-diagonal coefficients Since $P_{\lambda,\delta}$ varies on a scale $\sim \delta$, this implies that, for $\lambda \in [\frac{\lambda_0}{2}, \lambda_0]$,

$$\int_{[1,2]^d} \sup_{\frac{\lambda_0}{2} < \lambda < \lambda_0} \|P_{\lambda,\delta}\|_{L^1 \rightarrow L^\infty}^b d\beta' \lesssim (\log \lambda_0)^{bd+d} \delta^{-1} \max_{b_2 \leq \min(b,d+1)} \delta^{b_2} \lambda_0^{-b_2^2+(b+d)b_2+1-b}.$$

By Borel-Cantelli as in Appendix A, this implies that, for any $\epsilon > 0$,

$$\|P_{\lambda,\delta}\|_{L^1 \rightarrow L^\infty} \lesssim_{\beta,\epsilon} \delta^{-1/b} \lambda^\epsilon \left[\max_{b_2 \leq \min(b,d+1)} \delta^{b_2} \lambda^{-b_2^2+(b+d)b_2+1-b} \right]^{1/b}, \tag{6.10}$$

both for fixed off-diagonal coefficients and almost all $\beta' \in [1, 2]^d$, and also for almost all matrices (β_{ij}) with $\beta' \in [1, 2]^d$ and small off-diagonal coefficients. This proves (6.1).

We now observe that if b_2 is strictly less than $\min(b, d + 1)$, then

$$\begin{aligned} \frac{\delta^{b_2+1} \lambda^{-(b_2+1)^2+(b+d)(b_2+1)+1-b}}{\delta^{b_2} \lambda^{-b_2^2+(b+d)b_2+1-b}} &= \delta \lambda^{-2b_2-1+b+d} \\ &\geq \delta \lambda^{-2 \min(b-1,d)-1+b+d}. \end{aligned}$$

This will be ≥ 1 provided that $\delta \geq \lambda^{\min(b-d-1,d+1-b)}$, and so for such δ the maximum in (6.10) is reached for $b_2 = \min(b, d + 1)$. We will therefore impose the condition $\delta \geq \lambda^{\min(b-d-1,d+1-b)}$ for convenience rather than because we believe it to be optimal. This yields

$$\|P_{\lambda,\delta}\|_{L^1 \rightarrow L^\infty} \lesssim_{\epsilon,\beta'} \begin{cases} \delta^{1-\frac{1}{b}} \lambda^{d-1+\frac{1}{b}+\epsilon} & (b \leq d, \delta \geq \lambda^{b-d-1}), \\ \delta^{\frac{d}{b}} \lambda^{d-\frac{d}{b}+\epsilon} & (b \geq d + 1, \delta \geq \lambda^{d+1-b}). \end{cases}$$

We use the first of these alternatives when $\delta \geq \lambda^{1-d}$, and otherwise we use the second. In particular by writing $a = d + 1 - b$ if $\delta \geq \lambda^{1-d}$ and $a = b - d - 1$ otherwise, we obtain (6.2). □

Acknowledgements PG was supported by the Simons collaborative grant on weak turbulence. SLRM was supported by a Leverhulme Early Career Fellowship, by the Fields Institute for Research in Mathematical Sciences, and by the National Science Foundation grant DMS 1363013. The authors are thankful to Yu Deng for insightful discussions at an early stage of this project

Declarations

Conflicts of interest, etc. On behalf of all authors, the corresponding author states that there is no conflict of interest. Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix A. Borel-Cantelli lemma

At several points we use versions of the following argument to turn a bound on average over (β_{ij}) into a bound for almost all β .

Let $\beta = (\beta_{ij})$ be as in Definition 1.1. That is, either $\beta_{ij} = \delta_{ij}\beta_i$ with $\beta_i \in [1, 2]$ for each i ; or $\beta_{ij} = \delta_{ij} + h_{ij}$ for each $1 \leq i, j \leq d$ and some $h_{ij} = h_{ji} \in \left[-\frac{1}{10d^2}, \frac{1}{10d^2}\right]$.

We consider a function $\Phi(\beta, \vec{P})$ that depends on (β_{ij}) and also a list of parameters P_1, \dots, P_k , for some $k \in \mathbb{N}$. In applications these parameters will be integer powers of 2, constrained by some inequalities. For example in the proof of Lemma 3.3 we should let \vec{P} take values in the set

$$S = \left\{ (N, T, Q_1, \dots, Q_d) \in (2\mathbb{Z})^{d+2} : 2 \leq Q_j \leq N, \frac{1}{4N} \leq T \leq N^\kappa \right\},$$

and put

$$\Phi(\beta, \vec{P}) = \int_1^2 \dots \int_1^2 \int_0^T \Lambda_{Q_1}(t) \Lambda_{Q_2}(\beta_2 t) \dots \Lambda_{Q_k}(\beta_k t) dt d\beta_2 \dots d\beta_k.$$

Lemma A.1 *Let Φ be as above. Assume that \vec{P} takes values in a subset $S \subset (2\mathbb{Z})^k$ such that for any $\epsilon > 0$ we have*

$$\sum_{\vec{P} \in S} P_1^{-\epsilon} < \infty.$$

If Ψ is some real function of \vec{P} with

$$\int |\Phi(\beta, \vec{P})| d\beta \leq \Psi(\vec{P}),$$

then for every $\vec{P} \in S$, for generic β we have

$$\Phi(\beta, \vec{P}) \lesssim_{\beta, \epsilon} P_1^\epsilon \Psi(\vec{P}).$$

The lemma follows at once from the standard statement of the Borel-Cantelli lemma.

References

1. Bourgain, J.: Eigenfunction bounds for the Laplacian on the n -torus. *Int. Math. Res. Not.* (3), 61–66 (1993)
2. Bourgain, J.: Moment inequalities for trigonometric polynomials with spectrum in curved hypersurfaces. *Israel J. Math.* **193**(1), 441–458 (2013)
3. Bourgain, J., Demeter, C.: Improved estimates for the discrete Fourier restriction to the higher dimensional sphere. *Illinois J. Math.* **57**(1), 213–227 (2013)
4. Bourgain, J.: Demeter, Ciprian: New bounds for the discrete Fourier restriction to the sphere in 4D and 5D. *Int. Math. Res. Not. IMRN* (11), 3150–3184 (2015)
5. Bourgain, J., Demeter, C.: The proof of the l^2 decoupling conjecture. *Ann. Math. (2)* **182**(1), 351–389 (2015)
6. Bourgain, J., Shao, P., Sogge, C.D., Yao, X.: On L^p -resolvent estimates and the density of eigenvalues for compact Riemannian manifolds. *Commun. Math. Phys.* **333**(3), 1483–1527 (2015)
7. Bourgain, J., Watt, N.: Mean square of zeta function, circle problem and divisor problem revisited. arXiv e-prints, page [arXiv:1709.04340](https://arxiv.org/abs/1709.04340), (September 2017)
8. Collot, Charles, G. P.: Derivation of the homogeneous kinetic wave equation: longer time scales. arXiv preprint [arXiv:2007.03508](https://arxiv.org/abs/2007.03508), (2020)
9. Colzani, L., Gariboldi, B., Gigante, G.: Variance of lattice point counting in thin annuli. *J. Geom. Anal.* **31**(9), 8903–8924 (2021)
10. Germain, P., Rydin, M., Simon, L.: Bounds for spectral projectors on tori. arXiv preprint, [arXiv:2104.13274](https://arxiv.org/abs/2104.13274)
11. Hickman, J.: Uniform L^p Resolvent Estimates on the Torus. *Math. Res. Rep.* **1**, 31–45 (2020)
12. Hofmann, S., Iosevich, A., Weidinger, D.: Lattice points inside random ellipsoids. *Michigan Math. J.* **52**(1), 13–21 (2004)
13. Jarník, V.: Über Gitterpunkte in mehrdimensionalen Ellipsoiden. *Math. Ann.* **100**(1), 699–721 (1928)
14. Kelmer, D.: On the mean square of the remainder for the euclidean lattice point counting problem. *Israel J. Math.* **221**(2), 637–659 (2017)
15. Kendall, D.G.: On the number of lattice points inside a random oval. *Q. J. Math. Oxford Ser.* **19**, 1–26 (1948)
16. Landau, E.: Über die anzahl der gitterpunkte in gewissen bereichen. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, pages 137–150, (1924)
17. Schmidt, W.M.: A metrical theorem in geometry of numbers. *Trans. Am. Math. Soc.* **95**, 516–529 (1960)
18. Shen, Z.: On absolute continuity of the periodic Schrödinger operators. *Int. Math. Res. Not.* **1**, 1–31 (2001)
19. Sogge, C.D.: Fourier integrals in classical analysis, volume 105 of Cambridge Tracts in Mathematics. Cambridge University Press, Cambridge, (1993)
20. Stein, E.M.: Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals, volume 43 of Princeton Mathematical Series. Princeton University Press, Princeton, NJ, (1993). With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III
21. Tomas, P.A.: A restriction theorem for the Fourier transform. *Bull. Am. Math. Soc.* **81**, 477–478 (1975)
22. Trevisan, J.: Limit laws in the lattice problem. II. The case of ovals. arXiv e-prints, page [arXiv:2109.02378](https://arxiv.org/abs/2109.02378), (September 2021)
23. Zygmund, A.: On Fourier coefficients and transforms of functions of two variables. *Stud. Math.* **50**, 189–201 (1974)