An Analogue of Bonami's Lemma for Functions on Spaces of Linear Maps, and 2-2 Games

David Ellis
School of Mathematics, University of Bristol
United Kingdom
david.ellis@bristol.ac.uk

Guy Kindler
Rachel and Selim Benin School of Computer Science and Engineering
Hebrew University of Jerusalem
Israel

Noam Lifshitz
Einstein Institute of Mathematics, Hebrew University of Jerusalem
Israel
noamlifshitz@gmail.com

ABSTRACT

We prove an analogue of Bonami's (hypercontractive) lemma for complex-valued functions on $\mathcal{L}(V, W)$, where V and W are vector spaces over a finite field. This inequality is useful for functions on $\mathcal{L}(V, W)$ whose 'generalised influences' are small, in an appropriate sense. It leads to a significant shortening of the proof of a recent seminal result by Khot, Minzer and Safra that pseudorandom sets in Grassmann graphs have near-perfect expansion, which (in combination with the work of Dinur, Khot, Kindler, Minzer and Safra) implies the 2-2 Games conjecture (the variant, that is, with imperfect completeness).

CCS CONCEPTS

• **Mathematics of computing** → **Continuous functions**; **Combinatoric problems**.

KEYWORDS

datasets, neural networks, gaze detection, text tagging

ACM Reference Format:
David Ellis, Guy Kindler, and Noam Lifshitz. 2023. An Analogue of Bonami's Lemma for Functions on Spaces of Linear Maps, and 2-2 Games. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing (STOC '23), June 20–23, 2023, Orlando, FL, USA.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3564246.3585116

1 INTRODUCTION

Hypercontractive inequalities are of great importance and use in mathematical physics, analysis, geometry, probability theory, combinatorics and theoretical computer science (having first been introduced by Nelson [18], motivated by mathematical physics). In general, for $1 \le p < q \le \infty$, a (p, q)-hypercontractive inequality for a measure space X and an operator $T : L^p(X) \to L^q(X)$ says that $\|T(f)\|_q \le \|f\|_p$ for all $f \in L^p(X)$. One of the most classical, fundamental and useful hypercontractive inequalities is the hypercontractive inequality of Bonami, Beckner and Gross regarding the noise operator on the discrete cube, with the uniform measure. Let us give the statement in full. For $0 \le \rho \le 1$, the noise operator $T_\rho : L^p(\{0,1\}^n) \to L^q(\{0,1\}^n)$ is defined by

$$(T_\rho f)(x) = \mathbb{E}_{y \sim N_\rho(x)}[f(y)] \quad \forall x \in \{0,1\}^n,\ \forall f : \{0,1\}^n \to \mathbb{R},$$

where the distribution $y \sim N_\rho(x)$ is defined as follows: independently for each coordinate $i \in [n]$, we set $y_i = x_i$ with probability ρ, and with probability $1-\rho$ we take $y_i \in \{0,1\}$ uniformly at random (independently of x_i). In other words, we obtain y from x by resampling each coordinate of x independently with probability $1-\rho$, so y is a 'noisy' version of x. Note that $T_1(f) = f$, i.e. T_1 is simply the identity operator; on the other hand, T_0 maps a function f to the constant function with value $\mathbb{E}[f]$. For $0 < \rho < 1$, T_ρ interpolates between these two extremes: the smaller the value of ρ, the greater the degree of 'smoothing'.

The hypercontractive inequality of Bonami [3], Beckner [2] and Gross [8][1] states that

$$\|T_\rho(f)\|_q \le \|f\|_p \quad \forall \rho \le \sqrt{(p-1)/(q-1)},\ \forall f : \{0,1\}^n \to \mathbb{R}.$$

As the spectral norm of T_ρ is 1, this inequality means that it acts as a smoothing operator, smoothing out sharp peaks.

Often, the special case with $q = 4$ and $p = 2$ suffices for applications; this says that

$$\|T_\rho(f)\|_4 \le \|f\|_2 \quad \forall \rho \le 1/\sqrt{3}. \tag{1.1}$$

T_ρ can also be written in terms of the Fourier transform, writing $f : \{0,1\}^n \to \mathbb{R}$ as $f = \sum_{S \subseteq [n]} \hat{f}(S)\chi_S$, where $\chi_S(x) = (-1)^{\sum_{i \in S} x_i}$ for all $x \in \{0,1\}^n$ and $S \subseteq [n]$ (here, $\hat{f}(S) = \langle f, \chi_S \rangle$ for all $S \subseteq [n]$), the noise operator T_ρ is given by

$$T_\rho(f) = \sum_{S \subset [n]} \rho^{|S|} \hat{f}(S)\chi_S.$$

This yields the following corollary of (1.1), known as Bonami's lemma, which is extremely useful.

Lemma 1 (Bonami's Lemma). *Let $f : \{0,1\}^n \to \mathbb{R}$ be a function of degree at most d; then*

$$\|f\|_4 \le 3^{d/2}\|f\|_2.$$

(Recall that the *degree* of a function $f : \{0,1\}^n \to \mathbb{R}$ is the maximal size of a set S such that $\hat{f}(S) \neq 0$.) Bonami's lemma bounds the 4-norm of a low-degree function in terms of its 2-norm; roughly speaking, it says that low-degree functions on $\{0,1\}^n$ do not have very large 'peaks' in their modulus (such peaks would lead to their having large 4-norm).

[1] It was discovered independently by these three authors, though Bonami considered only the case $p = 2$, which suffices for most applications.

STOC '23, June 20–23, 2023, Orlando, FL, USA

ACM ISBN 978-1-4503-9913-5/23/06.
https://doi.org/10.1145/3564246.3585116

The Bonami-Beckner-Gross hypercontractive inequality was a crucial ingredient in the proof of the seminal Kahn-Kalai-Linial theorem [9] on the influences of Boolean functions, and of Friedgut's junta theorem [7]; both have been of huge importance in combinatorics and theoretical computer science over the last three decades. (In fact, Bonami's lemma suffices for these two applications.)

The notion of a 'noise' operator (which we defined above for the discrete cube) readily generalises to $L^p(X, \mu)$ for many other measure spaces (X, μ): one just needs to find a (natural) way to resample the input of a function (resampling in a way that is more or less 'extreme', depending on the noise parameter). For example, the noise operator $U_\rho : L^p(\mathbb{R}^n, \gamma^n) \to L^q(\mathbb{R}^n, \gamma^n)$ on n-dimensional Gaussian space (with the standard n-dimensional Gaussian measure γ^n) is defined by

$$U_\rho(f)(x) = \mathbb{E}_{Y\sim\gamma^n}\left[f\left(\rho x + \sqrt{1-\rho^2}Y\right)\right] \quad \forall f \in L^p(\mathbb{R}^n, \gamma^n);$$

this is natural because if X and Y are independent n-dimensional standard Gaussian random variables, then X and $\rho X + \sqrt{1-\rho^2}Y$ are ρ-correlated standard n-dimensional Gaussians. Here, the 'noisy' version of x is the random variable $\mathbb{N}_\rho(x) := \rho x + \sqrt{1-\rho^2}Y$, where $Y \sim \gamma^n$.

Hypercontractive inequalities for natural 'noise' operators on many other spaces have been obtained over the last five decades. A very useful example is the hypercontractive inequality for the noise operator in Gaussian space [2, 8, 19] (an earlier suboptimal version appeared in [18]); this is intimately related to the heat equation. Rothaus proved a sharp hypercontractive inequality [21] for the n-dimensional sphere S^n. Gross [8] proved that a hypercontractive inequality for a space is equivalent to a log-Sobolev inequality for that space, linking two important bodies of work, and proved hypercontractive inequalities over some noncommutative algebras related to quantum field theory. The hypercontractive inequality for the noise operator in Gaussian space was a crucial ingredient in the proof of the seminal Invariance Theorem of Mossel, O'Donnell and Oleszkiewicz [17].

The hypercontractive inequalities we have discussed above hold for *all* functions on the corresponding space. For some important examples of spaces, however, a (strong) hypercontractive inequality does not hold for all functions — even an analogue of Bonami's lemma does not hold, since there are 'badly-behaved' low-degree functions whose 4-norm is large compared to their 2-norm. This is the case for functions on the p-biased cube $(\{0,1\}^n, \mu_p)$, where $p = o(1)$: the 'dictatorship' functions defined by $f(x) = x_i$ for some i have 4-norm $p^{1/4}$, which is much greater than their 2-norm $p^{1/2}$, when $p = o(1)$. (Recall that the p-biased measure on $\{0,1\}^n$ is defined by $\mu_p(x) = p^{\sum_{i=1}^n x_i}(1-p)^{n-\sum_{i=1}^n x_i}$. A 'weak' analogue of Bonami's lemma holds for the p-biased measure, but with $\sqrt{3}$ replaced by function of p which tends to infinity as p tends to zero. This weak analogue is insufficient for many important applications.)

Recently, Keevash, Lifshitz, Long and Minzer [11] proved that 'dictatorships' and similar 'junta-type' constructions are in a sense the only barrier to hypercontractivity: for functions whose norm is not too much affected by restricting the values of a small set of coordinates, a hypercontractive inequality does hold. (Keevash, Lifshitz, Long and Minzer called such functions *global functions*.) A hypercontractive inequality for global functions may be termed a *conditional hypercontractive inequality* (the precise quantitative notion of 'global' may differ according to the context or the application in mind) — the classical Bonami-Beckner-Gross inequality and its Gaussian analogue, on the other hand, are unconditional (the hypercontractive inequality there holds for all functions, not just global ones). In [11], Keevash, Lifshitz, Long and Minzer obtained (conditional) hypercontractive inequalities (for the natural noise operator) for global functions on both the p-biased cube $(\{0,1\}^n, \mu_p)$ and for a general product space (X^n, μ^n); these had several important applications in extremal combinatorics and theoretical computer science (see e.g. [10]). In [6], Filmus, Kindler, Lifshitz and Minzer obtained a (conditional) hypercontractive inequality for global functions on the symmetric group S_n, a non-product space (in the case of the symmetric group, again, a 'strong' hypercontractivity does not hold for all functions, as one can see by considering the indicator function of a point-stabilizer); this in turn was a crucial ingredient in the resolution by Keevash, Lifshitz and Minzer [12] of a well-known open problem of Crane, concerning the largest product-free sets in the alternating groups A_n.

One important family of applications of hypercontractive inequalities (both unconditional and conditional hypercontractive inequalities) is to obtain *small-set expansion* theorems. A small-set expansion theorem for a finite, regular undirected graph G says, roughly speaking, that small sets[2] have very large vertex-boundary in G, much larger than the bound guaranteed by the Cheeger constant[3], the latter bound being sharp only for larger (or non-pseudorandom) sets. More precisely, a small-set expansion theorem for $G = (V, E)$ says that if $S \subset V(G)$ with $|S|$ small (and, possibly, satisfies an additional globalness or psuedorandomness condition), then choosing a uniform random element u of S and a random edge uv of G incident with u, the vertex v (at the other end of the random edge) will lie outside S with probability close to 1. There is a similar notion for weighted graphs, where the edges are weighted with non-negative weights and the weighting is regular (meaning that the sum of the weights of edges incident to each vertex is the same): in this case, the random edge uv is chosen with probability proportional to the weight of the edge uv.

A hypercontractive inequality can often be used to prove a small-set expansion theorem, as we shall now roughly outline. First, given a graph G on a probability space (X, μ), one finds a noise operator T_ρ defined by $T_\rho f(x) = \mathbb{E}_{y\sim N_\rho(x)}[f(y)]$, such that T_ρ satisfies a hypercontractive inequality, and such that the 'noised' version $N_\rho(x)$ of x is concentrated on close neighbours of x in G (i.e., on vertices of G with small graph-distance from x). This means, roughly, that $T_\rho f(x)$ is an average value of $f(y)$ over vertices y that are 'close neighbors' of x. This in turn means that if f is the indicator function of a set S, then the inner product $\langle T_\rho f, f\rangle$ is roughly (or sometimes, exactly) proportional to the probability that if we choose a uniform random vertex u in S and a uniform random edge uv incident with u, traversing the edge from u to v does not take us outside the set S. Partitioning $T_\rho f$ to its low-degree

[2]Possibly, provided they satisfy an additional 'globalness' or 'psuedorandomness' condition, such as having no large density increment on a 'nice' subset.
[3]Or, which is roughly equivalent, the second eigenvalue of the graph.

and high-degree part, the high-degree part contributes little to the inner product because T_ρ shrinks its 2-norm to something very small (this follows from the Fourier transform representation of T_ρ). As for the contribution of the low-degree part, this can be bounded by an expression involving the 4-norm of $T_\rho f$, by using Hölder's inequality. Applying the hypercontractive inequality for T_ρ and rearranging, we obtain an upper bound on the probability of staying inside S, and thus a lower bound on the probability of moving outside it.

In this paper, we obtain an analogue of Bonami's lemma for 'global' functions on the space $\mathcal{L}(V,W)$ of linear maps from V to W, where V and W are finite-dimensional vector spaces over a finite field. This leads to a significant conceptual simplification and streamlining/shortening of the proof of the seminal result of Khot, Minzer and Safra [15] obtaining small-set expansion for pseudorandom sets in the Grassmann graph; the latter was one of the crucial ingredients in the celebrated proof of the 2-2 Games conjecture (with imperfect completeness), along with the earlier results of Khot, Minzer and Safra in [16], of Dinur, Kindler, Khot, Minzer and Safra in [4, 5], and of Barak, Kothari and Steurer in [1].

The Unique Games conjecture of Khot is considered by many to be the second-most important open problem in complexity theory, after the P versus NP problem; yet it is not considered to be out of reach in the same way as the P versus NP problem. The proof of the 2-2 Games conjecture (with imperfect completeness) is one of the greatest breakthroughs in the area, in recent times. We proceed to give a full statement of the problem.

One can think of an instance of the Unique Games problem as a system of linear equations over $\mathbb{F}_p$ for some prime p, where every equation (/constraint) is of the form

$$t_{ij}x_i + t'_{ij}x_j = c_{ij},$$

for $i,j \in [n]$, where $x_1, x_2, \ldots x_n$ are variables taking values in $\mathbb{F}_p$, and $c_{ij}, t_{ij}, t'_{ij} \in \mathbb{F}_p$ are constants[4]. The goal is to find an assignment of the variables that satisfies a large fraction of the equations (/constraints). The Unique Games conjecture states that for any $\epsilon > 0$, there exists $p_0(\epsilon) \in \mathbb{N}$ such that for all primes $p \geq p_0(\epsilon)$, given an instance of the Unique Games conjecture for $\mathbb{F}_p$ where we are promised there is an assignment satisfying at least a $(1-\epsilon)$-fraction of the equations, it is an NP-hard problem to find an assignment satisfying (even) at least an ϵ-fraction of the equations.

The 'uniqueness' in the Unique Games problem refers to the fact each equation (/constraint) $\mathcal{E}$ of the form $t_{ij}x_i + t'_{ij}x_j = c_{ij}$ inside an instance actually fixes a one-to-one correspondence between assignments of the variable x_i and assignments of the variable x_j, since if the coefficients c_{ij} and c'_{ij} are non-zero (which indeed we may assume, without loss of generality), then for each assignment of x_i there is a unique assignment of x_j for which $\mathcal{E}$ is satisfied, and vice versa. The 2-2 Games conjecture (the variant, that is, with imperfect completeness) refers to an analogous problem, where each constraint sets a relation between a pair of distinct variables x_i and x_j which, rather than being 'unique' (or 'one-to-one'), is instead 'two-to-two'. (We explain precisely what this means, shortly.) This is a more general set of allowed constraints, and so intuitively one would guess that it would be more difficult to find an assignment that satisfies at least an ϵ-fraction of the constraints, even when one is promised that there exists an assignment satisfying at least a $(1-\epsilon)$-fraction of them. This guess turns out to be correct: it is easy to prove that the 2-2 Games conjecture with imperfect completeness, follows from the Unique Games conjecture, and (as mentioned above) the former has now been proven, whereas the latter remains open.

Now let us explain what a 2-to-2 constraint is. A very simple example is the constraint

$$t_{ij}x_i + t'_{ij}x_j \in \{c_{ij}, c'_{ij}\}$$

on the pair of variables x_i and x_j, where $t_{ij}, t'_{ij} \in \mathbb{F}_p^\times$ and $c_{ij} \neq c'_{ij} \in \mathbb{F}_p$. Now each assignment of x_i that satisfies the constraint has *two* corresponding assignments of x_j that satisfy the constraint, and vice versa. Formally, a constraint on two variables x and y is said to be a *2-to-2 relation on their assignments* if there is a partition of the set of possible assignments of x into a collection of pairs $\mathcal{P}$, and a partition of the possible assignments of y into a collection of pairs $\mathcal{Q}$, along with a perfect matching from $\mathcal{P}$ to $\mathcal{Q}$, such that once two matched pairs are chosen (one pair, p say, in $\mathcal{P}$ and the other pair, q say, in $\mathcal{Q}$), any assignment of x from p and any assignment of y from q will satisfy the constraint; and furthermore, any assignments of x and y that do not come from matched pairs do not satisfy the constraint.

Let us now give a more complicated example of a 2-2 constraint, an example that was crucial in the aforementioned works on the 2-2 Games conjecture. We now index the variables by ℓ-dimensional subspaces of $\mathbb{F}_2^k$, and we impose constraints $C_{L,L'}$ on pairs of variables $x_L, x_{L'}$, where L and L' are ℓ-dimensional subspaces with $\dim(L \cap L') = \ell - 1$. For each ℓ-dimensional subspace L, we seek to assign values (to the variable x_L) which are $\mathbb{F}_2$-linear functionals on L, i.e. the assignments to x_L are elements f_L of the dual space L^*. The constraint $C_{L,L'}$ is defined as follows: an assignment f_L to x_L and an assignment $f_{L'}$ to $x_{L'}$ together satisfy $C_{L,L'}$ if $f_L(x) = f_{L'}(x)$ for all $x \in L \cap L'$, i.e. if the linear functionals f_L and $f_{L'}$ agree on $L \cap L'$. We note that since $L \cap L'$ is of codimension one in L (and also of codimension one in L'), and since we are working over $\mathbb{F}_2$, for any given linear functional g on $L \cap L'$ there are exactly two possible extensions of g to a linear functional on L and exactly two possible extensions of g to a linear functional on L'. It follows that the constraint $C_{L,L'}$ is indeed 2-2 in the above sense.

In [4], Dinur, Khot, Kindler, Minzer and Safra reduced the 2-2 Games conjecture (with imperfect completeness) to a statement called the 'Grassmann Soundness Hypothesis', which concerns constraints of the form $C_{L,L'}$ defined above. To explain further, we need some additional terminology. The *Grassmann graph* $G_{k,\ell}$ denotes the graph whose vertex-set consists of all ℓ-dimensional subspaces of $\mathbb{F}_2^k$, and where two ℓ-dimensional subspaces L and L' are joined by an edge if $\dim(L \cap L') = \ell - 1$. An (ℓ,k)-*Grassmann Test* is a system of constraints where we have a variable x_L for every vertex of the Grassmann graph (i.e. for every ℓ-dimensional subspace L of $\mathbb{F}_2^k$), and a constraint $C_{L,L'}$ as defined above for every edge of the Grassmann graph. The Grassmann Soundness

[4]The original version of the Unique Games Conjecture allowed for more general types of constraints, but it was shown in [13] that one can assume without loss of generality that the constraints are as we describe here.

Hypothesis states (roughly) that if an assignment $(f_L)_{L\in V(G_{k,\ell})}$ satisfies at least an ϵ-fraction of the constraints $(C_{L,L'})_{\{L,L'\}\in E(G_{k,\ell})}$, then there must be a linear functional $f:\mathbb{F}_2^k\to\mathbb{F}_2$ that agrees on L with the assignment $f_L: L\to\mathbb{F}_2$, for many ℓ-dimensional subspaces L. More precisely, there must be a linear functional $f:\mathbb{F}_2^k\to\mathbb{F}_2$, and two subspaces $A\le B\le\mathbb{F}_2^k$ (with A of low dimension and B of low codimension) such that f agrees with a constant fraction of those assignments f_L for which L is sandwiched between A and B. The formal statement is as follows.

Hypothesis 2 (Grassmann Soundness Hypothesis)**.** *For every $\epsilon>0$, there exist $\ell_0\in\mathbb{N}$, $\eta>0$, $d\in\mathbb{N}$ and a function $k_0:\mathbb{N}\to\mathbb{N}$ such that the following holds. If $\ell\ge\ell_0$ and $k\ge k_0(\ell)$, and an assignment is given for the (ℓ,k)-Grassmann Test that satisfies at least an ϵ-fraction of the constraints, then there exists a linear functional $f:\mathbb{F}_2^k\to\mathbb{F}_2$ and subspaces $A\subseteq B\subseteq\mathbb{F}_2^k$ with $\dim(A)+\mathrm{codim}(B)\le d$, such that for at least an η-fraction of the ℓ-dimensional spaces $A\subseteq L\subseteq B$, it holds that f_L (the assignment of x_L) is equal to the restriction of f to L.*

The work of Barak, Kothari and Steurer [1] further reduced the Grassmann Soundness Hypothesis to the 'Grassmann Expansion Hypothesis', a statement about the expansion properties of the Grassmann graph, which we now describe. Given a finite, d-regular graph $G=(V,E)$ and a set of vertices $S\subset V(G)$, we define the *expansion ratio*

$$\Phi_G(S):=\frac{|E_G(S,\overline{S})|}{d|S|},$$

where $E_G(S,\overline{S})$ denotes the set of edges of G with one endpoint in S and the other endpoint in $\overline{S}:=V(G)\setminus S$. (Note that $\Phi_G(S)$ is precisely the probability that, if we pick uniformly at random a vertex u of S and then uniformly at random an edge of G that is incident with u, then the other endpoint of this edge lies outside S.) The Grassmann Expansion Hypothesis states that pseudorandom sets in the Grassmann graph have high expansion ratio, where by 'psuedorandom' we mean that the density of the set on lower-order copies of the Grassmann graph is not too high:

Hypothesis 3 (Grassmann Expansion Hypothesis)**.** *For any $0<\epsilon<1$, there exists $\ell_0=\ell_0(\epsilon)\in\mathbb{N}$, $d\in\mathbb{N}$ and $\eta>0$ such that the following holds. Let $\ell\ge\ell_0$ and let k be sufficiently large depending on ℓ. Let $S\subset V(G_{k,\ell})$ such that for any subspaces A and B of $\mathbb{F}_2^k$ with $A\subseteq B$ and $\dim(A)+\mathrm{codim}(B)\le d$, we have*

$$\frac{|\{L\in S:\ A\subseteq L\subseteq B\}|}{|\{L\in V(G_{k,\ell}):\ A\subseteq L\subseteq B\}|}\le\eta.$$

Then $\Phi_{G_{k,\ell}}(S)\ge 1-\epsilon$.

The proof of the 2-2 Games conjecture (with imperfect completeness) was completed when Khot, Minzer and Safra proved the Grassmann Expansion Hypothesis in the seminal work [15]. The proof in [15], however, is extremely long and technical. In this paper, we find a streamlined proof by first obtaining an (essentially optimal) analogue of Bonami's lemma for complex-valued functions on $\mathcal{L}(V,W)$, where V and W are vector spaces over $\mathbb{F}_q$, and then using the $q=2$ case of this to obtain a small-set expansion theorem for pseudorandom sets in the Shortcode Graph (the graph with vertex-set $\mathcal{L}(V,W)$, where two linear maps A_1 and A_2 are joined by an edge if A_1-A_2 is of rank one); such a small-set expansion theorem was already known to imply the Grassmann Expansion Hypothesis, by the work of Barak, Kothari and Steurer in [1].

We now describe our results in more detail. Our conceptual starting-point is the following (conditional) analogue of Bonami's lemma for global functions on product spaces, obtained by Keevash, Lifshitz, Long and Minzer in [11]. To state it we need some more notation and definitions. If $\Omega=X^n$ is a finite product-space, and $S\subset[n]$, we write $\Omega_S=X^S$. For $x\in\Omega_S$ and a function $f:\Omega\to\mathbb{C}$, we write $f_{S\to x}$ for the 'restricted' function on $\Omega_{[n]\setminus S}$ defined by $f_{S\to x}(y)=f(x,y)$, where (abusing notation slightly) we write (x,y) for the element $z\in\Omega$ with $z_i=x_i$ for all $i\in S$ and $z_i=y_i$ for all $i\in[n]\setminus S$. We equip the product-space Ω with the uniform (product) measure μ on Ω, and similarly we equip the product-space Ω_S with the uniform (product) measure on Ω_S, for any $S\subset[n]$. The *Efron-Stein decomposition* is an orthogonal decomposition of $L^2(\Omega,\mu)$ into spaces V_S (for $S\subset[n]$), where V_S consists of the functions in $L^2(\Omega,\mu)$ that depend only upon the coordinates in S and are orthogonal to any function that depends only upon the coordinates in T, for a proper subsets T of S. For a complex-valued function $f:\Omega\to\mathbb{C}$ and for each $S\subset[n]$, we define $f^{=S}$ to be the orthogonal projection of f onto V_S. We define the *Efron-Stein degree* of f to be $\max\{|S|: f^{=S}\neq 0\}$, and we define the *degree-d truncation* of f to be the function $f^{\le d}$ obtained by orthogonally projecting f onto the linear space of functions of (Efron-Stein) degree at most d (in other words, $f^{\le d}$ is simply the degree-d part of f).

Theorem 4 (Keevash, Lifshitz, Long, Minzer, 2019+)**.** *Let Ω be finite product space. Let $f:\Omega\to\mathbb{C}$ and let $\delta>0$. Suppose that $\|f_{S\to x}\|_2^2\le\delta$ for sets $S\subseteq[n]$ with $|S|\le d$ and all $x\in\Omega_S$. Then $\|f^{\le d}\|_4^4\le 1000^d\delta\|f^{\le d}\|_2^2$.*

We call the functions $f_{S\to x}$ (for $|S|\le d$) the *d-restrictions* of f. The above theorem says that if f is a function whose d-restrictions have small 2-norms, then the 4-norm of the degree-d part of f can be bounded from above in terms of its 2-norm. Theorem 4 was used in [11] to obtain a small-set expansion theorem for noise operators on product spaces; this small-set expansion theorem then played a crucial role in obtaining sharp forbidden intersection theorems for subsets $[m]^n$.

Our first aim in this paper is to obtain an analogue of Theorem 4 for complex-valued functions on $\mathcal{L}(V,W)$, but with Efron-Stein degree replaced by a different notion of degree, namely, the maximum rank of a linear map appearing in the Fourier expansion of f (this turns out to be the same as the 'junta degree', defined below). We note that $\mathcal{L}(V,W)$ could be viewed as a product space by fixing bases of V and W, and it could be equipped with the corresponding Efron-Stein degree, but this notion of degree would not be invariant under changes of basis and would not therefore be useful for applications.

To state our (conditional) Bonami-type lemma for functions on $\mathcal{L}(V,W)$, we need some more definitions. Let q be a prime power, and let V and W be finite-dimensional vector spaces over $\mathbb{F}_q$. We must first define our notion of a *d-restriction* of a function $f:\mathcal{L}(V,W)\to\mathbb{C}$. This is a little notationally cumbersome, if intuitively clear.

Let V_1 be a subspace of V, let W_1 be a subspace of W, let $T\in\mathcal{L}(V,W)$ be a linear map, and let $f:\mathcal{L}(V,W)\to\mathbb{C}$. The restriction

$f_{(V_1,W_1)\to T}$ is the function from $\mathcal{L}(V/V_1, W_1)$ to $\mathbb{C}$ defined by

$$f_{(V_1,W_1)\to T}(A) = f(A' + T) \quad \forall A \in \mathcal{L}(V/V_1, W_1),$$

where $A' \in \mathcal{L}(V, W)$ is the unique linear map with kernel containing V_1 and satisfying $A' = A \circ Q_{V_1}$, with $Q_{V_1} : V \to V/V_1$ denoting the natural quotient map. If $\dim(V_1) + \mathrm{codim}(W_1) \le d$ then we call such a restriction a *d-restriction*. We note that the linear maps B of the form $A' + T$ in the definition $f_{(V_1,W_1)\to T}$ are precisely the linear maps B such that B agrees with T on V_1 and B^* agrees with T^* on the annihilator of W_1.

Adopting the matrix perspective, the d-restriction of a function f on n by m matrices over $\mathbb{F}_q$ corresponds to restricting f to those matrices where r specific rows and c specific columns take fixed values, where $r + c \le d$ (and possibly translating the domain by a fixed matrix, if the matrix of T has non-zero entries outside the r fixed rows and the c fixed columns).

A function $f : \mathcal{L}(V, W) \to \mathbb{C}$ is said to be a *d-junta* if there exist $v_1, \ldots, v_i \in V, u_{i+1}, \ldots, u_d \in W^*$, such that the value of $f(A)$ is determined once we know the values of $A(v_i)$ and the values of $A^*(u_i)$. The *junta-degree* of a function f is the minimal integer d such that f can be written as a sum of d-juntas. (As mentioned above, we will show that the junta-degree of f is equal to the maximum rank of a linear map that appears in the Fourier expansion of f.) For a function $f : \mathcal{L}(W, V) \to \mathbb{C}$, we let $f^{\le d}$ denote its orthogonal projection onto the (linear) space of all functions with junta-degree at most d (in other words, as before, $f^{\le d}$ is simply the degree-d part of f).

For $d \in \mathbb{N}$ and $\delta > 0$, we say a function $f : \mathcal{L}(V, W) \to \mathbb{C}$ is *(d, δ)-restriction global* if $\|f_{(V_1,W_1)\to T}\|_2^2 \le \delta$ for all $V_1 \le V$ and $W_1 \le W$ with $\dim(V_1) + \mathrm{codim}(W_1) \le d$ and all $T \in \mathcal{L}(V, W)$; in other words, if all the d-restrictions of f have 2-norm at most $\sqrt{\delta}$. This is our notion of 'globalness' for functions on $\mathcal{L}(V, W)$.

We can now state our Bonami-type lemma for global functions on $\mathcal{L}(V, W)$.

Theorem 5. *Let $d \in \mathbb{N}$, let $\delta > 0$, let q be a prime power, let V and W be finite-dimensional vector spaces over $\mathbb{F}_q$, and suppose that $f : \mathcal{L}(V, W) \to \mathbb{C}$ is a (d, δ)-restriction global function. Then*

$$\|f^{\le d}\|_4^4 \le q^{Cd^2} \delta \|f^{\le d}\|_2^2,$$

where $C > 0$ is an absolute constant.

The d^2 in the exponent is sharp, as can be verified by inspecting the function $\sum_{X \in \mathcal{L}(W,V):\, \mathrm{rank}(X)=d} u_X$, where $u_X(A) = \omega^{\tau(\mathrm{Tr}(XA))}$, $\omega = \exp(2\pi i/p)$ and $\tau : \mathbb{F}_q \to \mathbb{F}_p$ is defined by $\tau(x) = x + x^p + \ldots + x^{p^{s-1}}$ for $q = p^s$. (This example shows that one must take $C \ge 1$, for any d and q.)

To motivate our proof of Theorem 5, and to illustrate some of the key ideas in a simpler setting, we will first give a proof of a (slightly weaker) version of Theorem 4 for the product space $\mathbb{F}_p^n$ for p a prime, and with C^d replaced by $(Cd)^d$.

Using Theorem 5, we obtain the following quantitatively sharp small-set expansion theorem for the shortcode graph, which (as mentioned above) implies the Grassmann Expansion Hypothesis.

Theorem 6 (Small-set expansion theorem for the shortcode graph)**.** *There exist absolute constants $C_1, C_2 > 0$ such that the following holds. Let $r \in \mathbb{N}$, and let $S \subseteq \mathcal{L}(V, W)$ be a family of linear maps with 1_S being $\left(C_1 r, q^{-C_2 r^2}\right)$-restriction global. Then*

$$\Pr_{A \sim S,\ B \text{ of rank } 1} [A + B \in S] < q^{-r}.$$

Theorem 6 is sharp up to the values of C_1 and C_2, as can be seen by considering the family $S = \{A \in \mathcal{L}(V, W) : \mathrm{rank}(A) \le n - r\}$, where $\dim(V) = \dim(W) = n$.

2 ARXIV LINK

The rest of this paper contains numerous long mathematical expressions that do not fit nicely in a two-columns format. Hence the reader is referred to the ArXiv version of this paper, available at https://doi.org/10.48550/arXiv.2209.04243, for the remaining details.

REFERENCES

[1] B. Barak, P.K. Kothari and D. Steurer. Small-set Expansion in the Shortcode Graph and the 2-2 Conjecture. *ITCS* 2019, Paper 9.

[2] W. Beckner. Inequalities in Fourier Analysis. Inequalities in Fourier analysis. *Ann. Math.* 102 (1975), 159–182.

[3] A. Bonami. Études des coefficients Fourier des fonctiones de $L^p(G)$. *Ann. Inst. Fourier* 20 (1070), 335–402.

[4] I. Dinur, S. Khot, G. Kindler, D. Minzer and M. Safra. Towards a proof of the 2-to-1 games conjecture? *Electronic Colloquium on Computational Complexity* 23:198, 2016.

[5] I. Dinur, S. Khot, G. Kinder, D. Minzer and S. Safra. On non-optimally expanding sets in Grassmann graphs. *Electronic Colloquium on Computational Complexity* 24:94, 2017.

[6] Y. Filmus, G. Kindler, N. Lifshitz and D. Minzer. Hypercontractivity for the symmetric group. Preprint, 2020. arXiv:2009.05503.

[7] E. Friedgut. Boolean functions with low average sensitivity depend on few coordinates. *Combinatorica* 18 (1998), 27–35.

[8] L. Gross. Logarithmic Sobolev inequalities. *Amer. J. Math.* 97 (1975), 1061–1083.

[9] J. Kahn, G. Kalai and N. Linial. The influence of variables on Boolean functions. *Proceedings of the 29th Annual IEEE Symposium on Foundations of Computer Science* (1988), pp. 68–80.

[10] P. Keevash, N. Lifshitz, E. P. Long and D. Minzer. Global hypercontractivity and its applications. Preprint, 2021. arXiv:2103.04604.

[11] P. Keevash, N. Lifshitz, E. P. Long and D. Minzer. Hypercontractivity for global functions and sharp thresholds. Preprint, 2019. arXiv:1906.05568.

[12] P. Keevash, N. Lifshitz and D. Minzer. On the largest product-free subsets of the alternating groups. Preprint, 2022. arXiv:2205.15191.

[13] S. Khot, G. Kindler, E. Mossel and R. O'Donnell. Optimal inapproximability results for MAX-CUT and other 2-variable CSPs? *SIAM J. Comput.* 37 (2007), 319–357.

[14] S. Khot. On the power of unique 2-prover 1-round games. *Proceedings of the 34th Annual ACM Symposium on the Theory of Computing* (2002), pp. 767–775.

[15] S. Khot, D. Minzer and S. Safra. Pseudorandom sets in Grassmann graphs have near-perfect expansion. *Electronic Colloquium on Communication Complexity* 25:6, 2018.

[16] Subhash Khot, Dor Minzer, and Muli Safra. On independent sets, 2-to-2 games, and Grassmann graphs. In: *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing (STOC)*, 2017, pp. 576–589.

[17] E. Mossel, R. O'Donnell and K. Oleszkiewicz. Noise stability of functions with low influences: Invariance and optimality. *Ann. Math.* 171 (2010), 295–341.

[18] E. Nelson. A quartic interaction in two dimensions. In: *Mathematical Theory of Elementary Particles* (R. Goodman and I. Segal, Eds.), pp. 69–73, MIT Press, Cambridge, MA, 1966.

[19] E. Nelson. Construction of quantum fields from Markoff fields. *J. Funct. Anal.* 12 (1973), 97–112.

[20] R. O'Donnell, *Analysis of Boolean functions.* Cambridge University Press, Cambridge, 2014.

[21] O. S. Rothaus. Hypercontractivity and the Bakry-Emery criterion for compact Lie groups. *J. Funct. Anal.* 65 (1986), 358–367.

Received 2022-11-07; accepted 2023-02-06